Analysis I

Gunther H. Peichl

Skriptum zur Vorlesung im SS 2011

Institut für Mathematik
Karl–Franzens–Universität Graz

Inhaltsverzeichnis

KAPITEL I

Reelle und komplexe Zahlen

1. Axiomatische Beschreibung der reellen Zahlen

Ein streng axiomatischer Aufbau der Analysis würde erfordern, das System der reellen Zahlen von möglichst einfachen Bausteinen ausgehend, etwa den natürlichen Zahlen, zu entwickeln. Dieser Weg soll hier nicht beschritten werden, hat doch jeder einzelne bereits im Laufe der Zeit mühsam eine intuitive Vorstellung von den reellen (oder zumindest rationalen) Zahlen entwickelt. Wir knüpfen an diese Vorstellung an und beschreiben die für das weitere Rechnen grundlegenden Beziehungen zwischen reellen Zahlen durch eine Reihe sorgfältig gewählter Axiome.

Wir gehen davon aus, daß auf der Menge $\mathbb{R}$ der reellen Zahlen zwei binäre Operationen, Addition und Multiplikation erklärt sind. Somit wird jedem Paar $(x, y) \in \mathbb{R} \times \mathbb{R}$ eindeutig eine reelle Zahl $x + y$ (die Summe von x und y) und ebenso eindeutig eine weitere reelle Zahl xy, manchmal $x \cdot y$ geschrieben (das Produkt von x und y) zugeordnet. Wie diese Summen und Produkte zu bilden sind, spielt dabei keine Rolle. Wesentlich ist nur, daß sie folgenden Axiomen genügen.

($\mathcal{A}$) **Axiome für die Addition**

($\mathcal{A}1$)	$\forall x, y, z \in \mathbb{R} : (x + y) + z = x + (y + z)$	Assoziativgesetz
($\mathcal{A}2$)	$\forall x, y \in \mathbb{R} : x + y = y + x$	Kommutativgesetz
($\mathcal{A}3$)	$\exists 0 \in \mathbb{R}\, \forall x \in \mathbb{R} : x + 0 = x$	additives Neutralelement
($\mathcal{A}4$)	$\forall x \in \mathbb{R}\, \exists \xi \in \mathbb{R} : x + \xi = 0$	additives inverses Element

($\mathcal{M}$) **Axiome für die Multiplikation**

($\mathcal{M}1$)	$\forall x, y, z \in \mathbb{R} : (xy)z = x(yz)$	Assoziativgesetz
($\mathcal{M}2$)	$\forall x, y \in \mathbb{R} : xy = yx$	Kommutativgesetz
($\mathcal{M}3$)	$\exists 1 \in \mathbb{R} : (1 \neq 0 \wedge \forall x \in \mathbb{R} : 1x = x)$	multiplikatives Neutralelement
($\mathcal{M}4$)	$\forall x \neq 0\, \exists \bar{x} \in \mathbb{R} : x\bar{x} = 1$	multiplikatives inverses Element

Die nächsten Axiome verknüpfen Addition und Multiplikation:

($\mathcal{D}$) **Distributivgesetz**
$\forall x, y, z \in \mathbb{R} : x(y + z) = xy + xz$

($\mathcal{O}$) **Ordnungsaxiom:**
Auf $\mathbb{R}$ ist eine Totalordnung $\leq$ erklärt, welche mit der Addition und Multiplikation verträglich ist:

$$\left.\begin{array}{lll}(\mathcal{OA}) & \forall x,y,z\in\mathbb{R}\colon & x\leq y\Rightarrow x+z\leq y+z\\ (\mathcal{OM}) & \forall x,y,z\in\mathbb{R}\colon & x\leq y\wedge 0\leq z\Rightarrow xz\leq yz\end{array}\right\}\text{Monotoniegesetze}$$

Wir erinnern daran, daß jede Ordnung $\leq$ eine strikte Ordnung $<$ in $\mathbb{R}$ induziert. Da $\mathbb{R}$ linear geordnet ist, gilt in $\mathbb{R}$ die Trichotomie. Man kann bei der axiomatischen Beschreibung von $\mathbb{R}$ auch davon ausgehen, daß in $\mathbb{R}$ eine strikte Ordnung $<$ existiert. In diesem Fall muß man allerdings die Monotoniegesetze durch ein weiteres Axiom ergänzen, das die Gültigkeit der Trichotomie in $\mathbb{R}$ verlangt. Die strikte Ordnung $<$ induziert dann eine lineare Ordnung $\leq$ auf $\mathbb{R}$.
Eine vollständige Charakterisierung der reellen Zahlen erfordert ein weiteres Axiom, durch welches Lücken in der Menge der reellen Zahlen ausgeschlossen werden.

($\mathcal{V}$) **Vollständigkeitsaxiom** (R. Dedekind)
Zu jedem Paar von Teilmengen A, B von $\mathbb{R}$ mit
i) $A \neq \emptyset \wedge B \neq \emptyset$
ii) $A \cup B = \mathbb{R}$
iii) $\forall a, b \in \mathbb{R} : a \in A \wedge b \in B \Rightarrow a < b$
gibt es genau ein $\xi \in \mathbb{R}$, sodaß
a) $\forall a \in \mathbb{R} : a < \xi \Rightarrow a \in A$
b) $\forall b \in \mathbb{R} : \xi < b \Rightarrow b \in B$

Man nennt das geordnete Paar (A, B) einen Dedekindschen Schnitt und die durch ihn eindeutig bestimmte Zahl ξ Schnittzahl. Wegen der Eigenschaften eines Schnittes liegt die Schnittzahl ξ entweder in A oder in B. Gilt $\xi \in A$, so sind die Eigenschaften a) und b) von ξ gleichwertig mit $A = \{a \in \mathbb{R}\colon a \leq \xi\}$ und $B = \mathbb{R} \setminus A$, liegt jedoch ξ in B, dann ist $B = \{b \in \mathbb{R}\colon b \geq \xi\}$ und $A = \mathbb{R} \setminus B$. Insbesondere folgt also $a \leq \xi < b$ oder $a < \xi \leq b$ für alle $a \in A$ und für alle $b \in B$.

Bemerkung I-1.1. (1) Die Assoziativgesetze erlauben es, die Ausdrücke $x + y + z$ bzw. xyz sinnvoll zu interpretieren, da jede Klammerung zum selben Ergebnis führt.
(2) Die auffallende Ähnlichkeit der Axiome für die Addition und Multiplikation ist nicht zufällig. Allgemein nennt man eine Menge in der eine binäre Operation erklärt ist, welche den unter $\mathcal{A}$ angegebenen Axiomen genügt, eine **Abelsche Gruppe**. Die Axiome $\mathcal{A}$ bringen somit zum Ausdruck, daß $(\mathbb{R}, +)$ eine additive Abelsche Gruppe ist, $\mathcal{M}$ bedeutet, daß $(\mathbb{R} \setminus \{0\}, \cdot)$ eine multiplikative Abelsche Gruppe darstellt. Jede Menge K, die diese beiden algebraischen Strukturen aufweist und in der außerdem das Distributivgesetz $\mathcal{D}$ gilt, nennt man **Körper**. Ist in K noch zusätzlich eine Totalordnung erklärt, welche den Monotoniegesetzen $\mathcal{OA}$ und $\mathcal{OM}$ genügt, spricht man von einem **geordneten Körper**. Somit kann man die Axiome $\mathcal{A}, \mathcal{M}, \mathcal{D}, \mathcal{O}$ zusammenfassen in der Feststellung: $(\mathbb{R}, +, \cdot, \leq)$ ist ein geordneter Körper. Gilt in einem geordneten Körper überdies das Vollständigkeitsaxiom $\mathcal{V}$, spricht man von einem **vollständigen geordneten Körper**.

(3) $\mathbb{R}^+ := \{x \in \mathbb{R} : x \geq 0\}$ ist die Menge der **nicht negativen** reellen Zahlen,
$\mathbb{R}^- := \{x \in \mathbb{R} : x \leq 0\}$ die Menge der **nicht positiven** reellen Zahlen,
$\mathbb{R}^+ \setminus \{0\}$ bzw. $\mathbb{R}^- \setminus \{0\}$ bezeichnet die Menge der **positiven** bzw. **negativen** reellen Zahlen.
(4) Für die zur Ordnungsrelation $\leq$ inverse Relation ist das Symbol $\geq$ gebräuchlich. Analog ist $>$ erklärt. Anstelle von $x \leq y \wedge y \leq z$ schreibt man oft $x \leq y \leq z$. (analog für $<$.)

Ohne Beweis teilen wir folgenden Satz mit

THEOREM I-1.2. *Es gibt eine Menge* $\mathbb{R}$, *welche den Axiomen* $\mathcal{A}, \mathcal{M}, \mathcal{D}, \mathcal{O}$ *und* $\mathcal{V}$ *genügt.*

Es ist allerdings denkbar, daß es verschiedene vollständige geordnete Körper gibt, etwa $(\mathbb{R}, +, \cdot, \leq)$ und $(\tilde{\mathbb{R}}, \oplus, \odot, \preceq)$. Man kann jedoch zeigen, daß es eine Bijektion $f\colon \mathbb{R} \to \tilde{\mathbb{R}}$ gibt, die mit der algebraischen Struktur und der Ordnungsstruktur in $\mathbb{R}$ verträglich ist, d.h. für alle $x, y \in \mathbb{R}$ gilt

$$\begin{aligned} f(x+y) &= f(x) \oplus f(y) \\ f(x \cdot y) &= f(x) \odot f(y) \\ x \leq y &\Leftrightarrow f(x) \preceq f(y). \end{aligned}$$

Vom mathematischen Standpunkt aus ist es also nicht notwendig, verschiedene Modelle von $\mathbb{R}$ zu unterscheiden.

2. Folgerungen aus den Körperaxiomen

Aus der Bedeutung der Gleichheit ergibt sich unmittelbar folgendes einleuchtende Resultat.

THEOREM I-2.1. *Es seien x, y reelle Zahlen und $x = y$. Dann gilt $x + z = y + z$ und $xz = yz$ für alle $z \in \mathbb{R}$.*

THEOREM I-2.2. *i) Das Neutralelement der Addition ist eindeutig bestimmt.*
ii) Für jede reelle Zahl ist das additive inverse Element ξ eindeutig bestimmt.

BEWEIS. *i)* Es seien 0 und $\tilde{0}$ Neutralelemente der Addition. Setzt man in ($\mathcal{A}3$) für x das Element $\tilde{0}$ ein, folgt $\tilde{0} + 0 = \tilde{0}$. Da $\tilde{0}$ ein weiteres Neutralelement bezüglich der Addition ist, gilt $\forall x \in \mathbb{R} : x + \tilde{0} = x$, somit auch $0 + \tilde{0} = 0$. Mit ($\mathcal{A}2$) schließt man $\tilde{0} + 0 = 0$ und wegen $\tilde{0} + 0 = \tilde{0}$ folgt $0 = \tilde{0}$.
ii) Wir nehmen an, ξ und $\tilde{\xi}$ seien zu x inverse Elemente bezüglich der Addition, d.h. $x + \xi = 0$ und $x + \tilde{\xi} = 0$. Addiert man zu der ersten Gleichung $\tilde{\xi}$, folgt

$$(x + \xi) + \tilde{\xi} = 0 + \tilde{\xi} = \tilde{\xi} + 0 = \tilde{\xi}.$$

Andererseits findet man für $(x + \xi) + \tilde{\xi}$ auch die Darstellung

$$(x + \xi) + \tilde{\xi} = x + (\xi + \tilde{\xi}) = x + (\tilde{\xi} + \xi) = (x + \tilde{\xi}) + \xi = 0 + \xi = \xi.$$

Somit gilt $\xi = \tilde{\xi}$. □

Wir bemerken, daß der Beweis des Satzes nur die Axiome $\mathcal{A}$ benützt. Die Aussage des Satzes trifft somit auf jede Abelsche Gruppe (vgl. Bemerkung I-1.1) zu, insbesondere auch auf $(\mathbb{R} \setminus \{0\}, \cdot)$. Es gilt somit auch folgender Satz:

THEOREM I-2.3. *i) Das Neutralelement der Multiplikation ist eindeutig bestimmt.*
ii) Für jede Zahl $x \neq 0$ ist das multiplikative inverse Element $\bar{x}$ eindeutig bestimmt.

Es ist somit gerechtfertigt, die Neutralelemente mit 0 bzw. 1 zu bezeichnen. Das additive Inverse von x bezeichnet man mit $-x$, das multiplikative Inverse mit x^{-1}. Ferner schreibt man oft $x - y$, $\frac{x}{y}$ anstelle von $x + (-y)$ und xy^{-1}. Es gilt auch eine teilweise Umkehrung von Satz I-2.1.

THEOREM I-2.4. *Für alle reellen Zahlen x, y und z gelten folgende Kürzungsregeln:*

i) *Aus* $x+z=y+z$ *folgt* $x=y$.
ii) *Aus* $xz=yz$ *und* $z\neq 0$ *folgt* $x=y$.

Beweis. Wir zeigen nur die zweite Regel. Es sei $x,y,z\in\mathbb{R}$, $z\neq 0$ und $xz=yz$. Nach $\mathcal{M}4$ existiert das multiplikative inverse Element z^{-1} und es gilt $zz^{-1}=1$. Aus $xz=yz$ folgt mit Satz I-2.1 $(xz)z^{-1}=(yz)z^{-1}$. Dies ergibt wegen $\mathcal{M}1$ $x(zz^{-1})=y(zz^{-1})$, also $x\cdot 1=y\cdot 1$. Wegen $\mathcal{M}2$ ist dies gleichwertig mit $1\cdot x=1\cdot y$ und mit $\mathcal{M}3$ schließen wir auf $x=y$. □

Nun können wir bereits zeigen, daß die Gleichungen $a+x=b$ und $cy=d$, $c\neq 0$, in $\mathbb{R}$ die eindeutigen Lösungen $x=b-a$ bzw. $y=\frac{d}{c}$ besitzen. Man beachte, daß zwei Behauptungen zu beweisen sind: nämlich die Eindeutigkeit einer Lösung und daß die angegebenen Zahlen tatsächlich die Gleichungen lösen. Wir beginnen mit der Eindeutigkeit einer Lösung von $a+x=b$. Wir gehen von der *Annahme* aus, x sei eine Lösung, und zeigen, daß notwendigerweise $x=b-a$ folgt. Addiert man zu $a+x=b$ das additive inverse Element $-a$, erhält man $(a+x)+(-a)=b+(-a)$. Für die linke Seite ergibt sich $(a+x)+(-a)=(x+a)+(-a)=x+(a+(-a))=x+0=x$, wobei wir der Reihe nach $\mathcal{A}2$, $\mathcal{A}1$, $\mathcal{A}4$ und $\mathcal{A}3$ verwendet haben. Es fehlt noch der Nachweis, daß $x=b-a$ Lösung ist. Dazu berechnen wir $a+x$ und erhalten $a+x=a+(b-a)=a+((-a)+b)=(a+(-a))+b=0+b=b$. Dieser Beweis verwendet nur Gruppenaxiome, das Resultat gilt daher in jeder Abelschen Gruppe, insbesondere in $(\mathbb{R}\setminus\{0\},\cdot)$. Eine eigene Diskussion der Gleichung $cy=d$ ist daher nicht notwendig.

Theorem I-2.5. *Für alle* $x,y,z,w\in\mathbb{R}$, $z\neq 0$ *und* $w\neq 0$ *gilt:*

a)	$0x=0$	*f)*	$(-x)+(-y)=-(x+y)$
b)	$-(-x)=x$	*g)*	$(-x)(-y)=xy$
c)	$(w^{-1})^{-1}=w$	*h)*	$\frac{x}{z}\frac{y}{w}=\frac{xy}{zw}$
d)	$(-1)x=-x$	*i)*	$\frac{x}{z}+\frac{y}{w}=\frac{xw+zy}{zw}$
e)	$x(-y)=-(xy)=(-x)y$		

Beweis. a) Aus $0+0=0$ folgt mit Satz I-2.1 $(0+0)x=0x$ für alle $x\in\mathbb{R}$. Das Distributivgesetz $\mathcal{D}$ ergibt $0x+0x=0x$, mit $\mathcal{A}3$ folgt $0x+0x=0x+0=0+0x$ und schließlich mit Satz I-2.4 $0x=0$.
b) Das additive inverse Element zu x bzw. $-x$ erfüllt $x+(-x)=0$ bzw. $(-x)+(-(-x))=0$. Mit Hilfe von $\mathcal{A}2$ folgert man

$$x+(-x)=(-(-x))+(-x).$$

Dies ergibt mit Satz I-2.4 $x=-(-x)$.
c) analog zu b).
d) Wir zeigen: $(-1)x$ ist das additive inverse Element zu x, d.h. $(-1)x+x=0$. Dies folgt aus

$$(-1)x+x\overset{\mathcal{M}3}{=}(-1)x+1x\overset{\mathcal{D}}{=}(-1+1)x\overset{\mathcal{A}4}{=}0x\overset{\text{a)}}{=}0.$$

e) Mit Hilfe des Assoziativgesetzes $\mathcal{M}1$ und der bereits bewiesenen Regel d) schließt man

$$x(-y)\overset{\text{d)}}{=}x((-1)y)\overset{\mathcal{M}1}{=}(x(-1))y\overset{\mathcal{M}2}{=}((-1)x)y\overset{d)}{=}(-x)y$$

und weiter

$$(-x)y\overset{d)}{=}((-1)x)y\overset{\mathcal{M}1}{=}(-1)(xy)\overset{d)}{=}-(xy).$$

f) Wir zeigen, $(-x)+(-y)$ ist das additive inverse Element zu $x+y$. Dies folgt aus

$$\begin{aligned}(x+y)+((-x)+(-y))&\overset{\mathcal{A}2}{=}(x+y)+\big((-y)+(-x)\big)\\&\overset{\mathcal{A}1}{=}\big((x+y)+(-y)\big)+(-x)\overset{\mathcal{A}1}{=}\big(x+(y+(-y))\big)+(-x)\\&\overset{\mathcal{A}4}{=}(x+0)+(-x)\overset{\mathcal{A}3}{=}x+(-x)\overset{\mathcal{A}4}{=}0.\end{aligned}$$

g) Die Behauptung folgt aus

$$(-x)(-y)\overset{\text{e)}}{=}-\big((-x)y\big)\overset{\text{e)}}{=}\big(-(-x)\big)y\overset{\text{b)}}{=}xy.$$

h) Wir zeigen zuerst: $\frac{1}{z}\frac{1}{w}=\frac{1}{zw}$, d.h. $z^{-1}w^{-1}=(zw)^{-1}$ (∗)
Dies folgt aus

$$(zw)\cdot(z^{-1}w^{-1})\overset{\mathcal{M}2}{=}(wz)(z^{-1}w^{-1})\overset{\mathcal{M}1}{=}w(zz^{-1})w^{-1}\overset{\mathcal{M}4}{=}w(1w^{-1})\overset{\mathcal{M}3}{=}ww^{-1}\overset{\mathcal{M}4}{=}1.$$

Somit gilt:

$$\frac{x}{z}\frac{y}{w}=(xz^{-1})(yw^{-1})=x(z^{-1}y)w^{-1}=(xy)(z^{-1}w^{-1})\overset{(*)}{=}xy(zw)^{-1}=\frac{xy}{zw}.$$

i) Mit Hilfe von h) ergibt sich

$$\begin{aligned}\frac{x}{z}+\frac{y}{w} &= \frac{x}{z}\cdot 1+\frac{y}{w}\cdot 1=\frac{x}{z}\frac{w}{w}+\frac{y}{w}\frac{z}{z}\\ &\overset{h)}{=} \frac{xw}{zw}+\frac{yz}{wz}\overset{\mathcal{D}}{=}(xw+yz)(zw)^{-1}=\frac{xw+yz}{zw}\end{aligned}$$

□

THEOREM I-2.6. *Das Produkt von zwei reellen Zahlen ist genau dann ungleich Null, wenn beide Faktoren ungleich Null sind, d.h.* $\forall xy\in\mathbb{R}\colon xy\neq 0\Leftrightarrow x\neq 0\wedge y\neq 0$.

BEWEIS. i) „$\Leftarrow$“ Wir führen einen Widerspruchsbeweis: Angenommen es gäbe $x,y\in\mathbb{R}$ mit $x\neq 0$, $y\neq 0$ und $xy=0$. Dann existiert das multiplikative inverse Element x^{-1} zu x. Aus $xy=0$ folgt einerseits $x^{-1}(xy)=x^{-1}\cdot 0=0$, andererseits $x^{-1}(xy)=(x^{-1}x)y=1\cdot y=y$ und somit $y=0$ im Widerspruch zur Voraussetzung $y\neq 0$.
ii) „$\Rightarrow$“ Wir führen den Beweis indirekt und zeigen:
$\forall x,y\in\mathbb{R}: x=0\vee y=0\Rightarrow xy=0$. Dies folgt jedoch unmittelbar aus I-2.5 (a). □

BEMERKUNG I-2.7. 1.) Satz I-2.6 ist äquivalent zu der Feststellung, daß das Produkt zweier reeller Zahlen genau dann Null ist, wenn mindestens einer der beiden Faktoren Null ist. In den Anwendungen führt diese Situation häufig zu einer Fallunterscheidung.
2.) Die Beweise dieses Abschnittes verwenden nur die Axiome $\mathcal{A},\mathcal{M},\mathcal{D}$. Die Behauptungen treffen daher in jedem Körper zu.
3.) Wir vereinbaren vorerst nur als Schreibweise:

$$2:=1+1, \tag{1}$$

$$\forall a\in\mathbb{R}\colon a^2:=a\cdot a. \tag{2}$$

3. Folgerungen aus den Ordnungsaxiomen

Wir erinnern daran, daß eine Ordnung $\leq$ eine strikte Ordnung $<$ induziert, welche durch $x<y\Leftrightarrow x\leq y\wedge x\neq y$ definiert ist. Geht man umgekehrt von einer strikten Ordnung aus, wird durch $x\leq y\Leftrightarrow x<y\vee x=y$ eine Ordnung erklärt. Diese Ordnung ist total, falls zusätzlich gefordert wird, daß für alle x, $y\in\mathbb{R}$ genau eine der drei Alternativen $x<y$, $x=y$, $y<x$ zutrifft. Das Ordnungsaxiom $\mathcal{O}$ kann daher durch folgendes gleichwertige Axiom ersetzt werden:

($\mathcal{O}^*$) **Ordnungsaxiom:**

In $\mathbb{R}$ ist eine strikte Ordnung $<$ erklärt, welche mit der Addition und Multiplikation verträglich ist:

$$\left.\begin{array}{lll}(\mathcal{OA}^*) & \forall x,y,z\in\mathbb{R}\colon & x<y\Rightarrow x+z<y+z\\ (\mathcal{OM}^*) & \forall x,y,z\in\mathbb{R}\colon & x<y\wedge 0<z\Rightarrow xz<yz\end{array}\right\}\text{Monotoniegesetze}$$

Es gilt die **Trichotomie**: für alle x, $y\in\mathbb{R}$ trifft genau eine der drei Alternativen $x<y$, $x=y$, $y<x$ zu.

Wir überlassen es dem Leser, die folgenden Resultate für die strikte Ordnung $<$ auf die Ordnung $\leq$ zu übertragen.

THEOREM I-3.1. *Für alle* $x,y\in\mathbb{R}$ *gilt*

a) $x<0\Leftrightarrow -x>0$ *c)* $x<y\Leftrightarrow y-x>0$
b) $x>0\Leftrightarrow -x<0$ *d)* $x<y\Leftrightarrow -y<-x.$

Beweis. c) „$\Rightarrow$: “ Aus $x < y$ folgt mit $\mathcal{OA}^*$ $x + (-x) < y + (-x)$, also $0 < y - x$.
„$\Leftarrow$: “ Es sei $y - x > 0$. Mit $\mathcal{OA}^*$ folgt $y + (-x) + x > 0 + x$, somit $y + 0 > x$. Dies hat $x < y$ zur Folge.
a) Setze $y = 0$ in c).
b) Ersetze x durch $-x$ in a) und verwende $-(-x) = x$ (Satz I-2.5-b).
d) $x < y \overset{c)}{\Leftrightarrow} y - x > 0 \Leftrightarrow -x - (-y) > 0 \overset{c)}{\Leftrightarrow} -y < -x$. □

Theorem I-3.2. *Für alle $x, y, z, w \in \mathbb{R}$ gilt*

a) $x < y \wedge z < w \Rightarrow x + z < y + w$,
b) $0 < x < y \wedge 0 < z < w \Rightarrow xz < yw$.

Beweis. a) Es seien $x, y, z, w \in \mathbb{R}$, $x < y$ und $z < w$. Wegen $\mathcal{OA}^*$ gilt dann auch $x + z < y + z$ und $y + z < y + w$. Da die strikte Ordnung $<$ transitiv ist, folgt $x + z < y + w$.
b) analog. □

Satz I-3.2 stellt also sicher, daß gleichsinnige Ungleichungen addiert werden dürfen; gleichsinnige Ungleichungen, in denen sämtliche Glieder nicht negativ sind, dürfen multipliziert werden.

Theorem I-3.3. *Das Produkt zweier reeller Zahlen ist genau dann positiv, wenn beide Faktoren ungleich Null sind und dasselbe Vorzeichen haben.*

Beweis. „$\Leftarrow$: “ Für $x > 0$ und $y > 0$ folgt die Behauptung aus $\mathcal{OM}^*$ und Satz I-2.5. Falls $x < 0$ und $y < 0$ folgt $-x > 0$ und $-y > 0$. Nach dem vorhin Bewiesenen gilt daher $(-x)(-y) > 0$. Nach Satz I-2.5 ist aber $(-x)(-y) = xy$, somit gilt $xy > 0$.
„$\Rightarrow$: “ Wir führen einen Widerspruchsbeweis und nehmen ohne Beschränkung der Allgemeinheit an, es gäbe relle Zahlen x, y mit $xy > 0$, $x \leq 0$ und $y > 0$. Ist $x = 0$ ergibt sich ein Widerspruch zu Satz I-2.5–(a). Es sei $x < 0$, also nach Satz I-3.1 $-x > 0$, und nach dem ersten Teil des Beweises ergibt sich

$$(-x)y = -(xy) > 0.$$

Mit Satz I-3.1 folgert man $xy < 0$. □

Korollar I-3.4. Für jede reelle Zahl x ungleich Null gilt $x^2 > 0$. Insbesondere folgt $1 > 0$.

Theorem I-3.5. *Für alle $x, y, z \in \mathbb{R}$ gilt*

i) $x < y \wedge z < 0 \Rightarrow yz < xz$,
ii) $0 < x \Rightarrow 0 < x^{-1}$,
iii) $0 < x < y \Rightarrow 0 < y^{-1} < x^{-1}$.

Beweis. i) Es sei $z < 0$, also $-z > 0$. Multipliziert man die Ungleichung $x < y$ mit $(-z)$, erhält man $x(-z) < y(-z)$, bzw. $-(xz) < -(yz)$. Nach Satz I-3.1 ist dies gleichwertig mit $yz < xz$.
ii) Aus dem Korollar I-3.4 ergibt sich $0 < 1 = xx^{-1}$, wegen Satz I-3.3 und $0 < x$ muß auch $0 < x^{-1}$ gelten.
iii) Man zeige $x^{-1}y^{-1} > 0$ und multipliziere $0 < x < y$ mit $x^{-1}y^{-1}$. □

Theorem I-3.6. *Zwischen zwei verschiedenen reellen Zahlen liegt stets eine weitere reelle Zahl, genauer $\forall a, b \in \mathbb{R}\colon a < b \Rightarrow a < \frac{a+b}{2} < b$.*
Die reelle Zahl $\frac{a+b}{2}$ heißt ***arithmetische Mittel*** *von a und b.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $a < b$, also $b - a > 0$. Es folgt

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

Analog zeigt man $\frac{a+b}{2} < b$. □

Insbesondere folgt aus Satz I-3.6, daß es keine kleinste positive reelle Zahl gibt.

KOROLLAR I-3.7. Es seien $a, b \in \mathbb{R}$ und $a < b+\varepsilon$ für beliebiges $\varepsilon > 0$. Dann ist $a \leq b$.

BEWEIS. Wir führen einen Beweis durch Widerspruch und nehmen an es gäbe reelle Zahlen a, b mit $a > b$ und $a < b+\varepsilon$ für alle $\varepsilon > 0$. Dann ist $a-b > 0$ und daher auch $b < b + \frac{a-b}{2} = \frac{a+b}{2}$. Wegen Satz I-3.6 muß $\frac{a+b}{2} < a$ gelten, dies widerspricht der Voraussetzung $a < b + \varepsilon$ für $\varepsilon = \frac{a-b}{2} > 0$. □

Wir notieren den nützlichen Spezialfall $a \geq 0$ und $b = 0$:

KOROLLAR I-3.8. Es sei $x \geq 0$ und für alle $\varepsilon > 0$ gelte $x \leq \varepsilon$. Dann ist $x = 0$.

DEFINITION I-3.9. *Für jede reelle Zahl x heißt*

$$|x| := \begin{cases} x, & x \geq 0, \\ -x, & x < 0, \end{cases}$$

absoluter Betrag *(Betrag) von x und*

$$\mathit{sign}\, x := \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0 \end{cases}$$

*heißt **Signum (Vorzeichen)** von x.*

Das Vorzeichen und der Betrag einer reellen Zahl x sind daher verknüpft durch die Beziehung

$$x = |x| \mathrm{sign}\, x.$$

Folgende Eigenschaften des Betrages sind eine unmittelbare Konsequenz der Definition.

LEMMA I-3.10. *Es seien x, $b \in \mathbb{R}$ und $b \geq 0$. Dann gilt*

a) $|x| \geq 0$ *b)* $x \leq |x|$

c) $|x| = |-x|$, *d)* $|x| \leq b \Leftrightarrow -b \leq x \leq b$.

BEWEIS. Wir zeigen nur die Eigenschaft d). Es seien x, $b \in \mathbb{R}$, $b \geq 0$ und $|x| \leq b$. Wir machen eine Fallunterscheidung nach dem Vorzeichen von x. Ist $x \geq 0$, dann gilt $|x| = x$, also auch $x \leq b$. Die zweite Ungleichung $x \geq -b$ folgt aus der Transitivität der Ordnung. Ist $x \leq 0$, also $|x| = -x$, ist die Ungleichung $|x| \leq b$ gleichwertig mit $-x \leq b$. Aus Satz I-3.1 folgt $x \geq -b$. Die Ungleichung $x \leq b$ ist wieder trivialerweise erfüllt. In jedem Falle gilt somit $-b \leq x \leq b$. Gehen wir nun umgekehrt von den Ungleichungen $x \leq b$ und $x \geq -b$ aus. Für $x \geq 0$ folgt direkt $|x| \leq b$. Ist $x \leq 0$ schließen wir von $x \geq -b$ auf $-x \leq b$. Dies ist gleichwertig mit $|x| \leq b$ □

THEOREM I-3.11. *Für alle $x, y \in \mathbb{R}$ gilt*

a) $|x| = 0 \Leftrightarrow x = 0$

b) $|xy| = |x||y|$

c) $|x+y| \leq |x| + |y|$ ***Dreiecksungleichung***

d) $\big||x| - |y|\big| \leq |x-y|$

e) $2|x||y| \leq |x|^2 + |y|^2$

Beweis. a) Für $x = 0$ folgt aus der Definition $|x| = 0$. Ist $x \neq 0$, dann ist auch $-x \neq 0$, somit ergibt sich ebenfalls $|x| \neq 0$.
b) Wir machen eine Fallunterscheidung und betrachten zuerst den Fall $x \geq 0$ und $y \geq 0$. Dann gilt $xy \geq 0$, $|x| = x$ und $|y| = y$. Es folgt $|xy| = xy = |x||y|$.
Als zweiten Fall untersuchen wir $x \geq 0$ und $y \leq 0$. Nun ist $xy \leq 0$, $|x| = x$ und $|y| = -y$. Es folgt $|xy| = -(xy) = x(-y) = |x||y|$.
Durch Vertauschen von x und y führt man den Fall $x \leq 0$ und $y \geq 0$ auf die eben betrachteten Situationen zurück. Es bleibt der Fall $x \leq 0$ und $y \leq 0$. Dann ist nach Satz I-3.3 $xy \geq 0$ und $|x| = -x$, $|y| = -y$. Aus Satz I-2.5-g folgt somit $|x| \cdot |y| = (-x)(-y) = xy = |xy|$.
c) Aus Lemma I-3.10 folgt $-|x| \leq x \leq |x|$ und $-|y| \leq y \leq |y|$. Addiert man die Ungleichungen ergibt sich

$$-(|x| + |y|) = -|x| - |y| \leq x + y \leq |x| + |y|,$$

also $|x + y| \leq |x| + |y|$ nach I-3.10-d).
d) Mit Hilfe der Dreiecksungleichung ergibt sich $|x| = |(x-y)+y| \overset{c}{\leq} |x-y|+|y|$ und somit $|x|-|y| \leq |x-y|$. Vertauscht man x und y erhält man $|y| - |x| \leq |y - x| = |x - y|$ und daraus mit Satz I-3.1 $-|x - y| \leq -(|y| - |x|) = |x| - |y|$. Insgesamt finden wir also $-|x - y| \leq |x| - |y| \leq |x - y|$, und daher

$$\big||x| - |y|\big| \leq |x - y|.$$

e) Mit Hilfe von Korollar I-3.4 schließen wir $0 \leq (|x|-|y|)(|x|-|y|) = |x|^2 - |x||y| - |y||x| + |y|^2 = |x|^2 - (1+1)|x||y| + |y|^2 = |x|^2 - 2|x||y| + |y|^2$. Wegen $\mathcal{OA}$ zeigt dies die behauptete Ungleichung. □

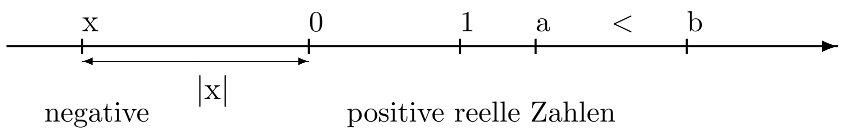
Bemerkung I-3.12. Die Beweise dieses Abschnittes stützen sich nur auf Eigenschaften, die für jeden Körper zutreffen, und auf das Axiom $\mathcal{O}$. Die Behauptungen gelten demnach in jedem geordneten Körper.

Definition I-3.13. *Es seien $a, b \in \mathbb{R}$ und $a \leq b$.*

i) $[a, b] := \{x \in \mathbb{R}\colon a \leq x \leq b\}$ *heißt* ***abgeschlossenes Intervall***,
ii) $(a, b) := \{x \in \mathbb{R}\colon a < x < b\}$ *heißt* ***offenes Intervall***,
iii) $(a, b] := \{x \in \mathbb{R}\colon a < x \leq b\}$, $[a, b) := \{x \in \mathbb{R}\colon a \leq x < b\}$ *heißen* ***halboffene Intervalle***.

*a ist **linker**, b **rechter Endpunkt** des Intervalls.*

Ist $a = b$, so ist $[a, b] = \{a\}$, die anderen Intervalle sind leer.
Da $\mathbb{R}$ linear geordnet ist, läßt sich $\mathbb{R}$ (wie jede linear geordnete Menge) mit Hilfe einer Geraden folgendermaßen veranschaulichen: Auf einer beliebigen Geraden in der Ebene werden zwei Punkte ausgezeichnet, die wir mit den reellen Zahlen 0 und 1 identifizieren. Es liege etwa 0 links von 1. Jeder reellen Zahl entspricht ein eindeutig bestimmter Punkt auf dieser Zahlengeraden. Es gilt $a < b$ genau dann, wenn der Punkt a links von b liegt. Den positiven reellen Zahlen entsprechen also Punkte jenes Halbstrahles, in dem der mit 1 bezeichnete Punkt liegt. In dieser Phase der Diskussion von $\mathbb{R}$ können wir aber noch nicht sicherstellen, daß umgekehrt jedem Punkt der Geraden tatsächlich genau eine reelle Zahl entspricht.



Reelle Zahlengerade

4. Die natürlichen Zahlen

Wir sind von einer axiomatischen Beschreibung der reellen Zahlen ausgegangen. Obwohl die natürlichen Zahlen vertraute Objekte sind, müssen wir nun zeigen, wie sie mit Hilfe unseres Axiomensystems als Teilmenge von $\mathbb{R}$ charakterisiert werden können.

DEFINITION I-4.1. *1) Für jedes $x \in \mathbb{R}$ heißt $x+1$ Nachfolger von x.*
2) Eine Teilmenge $I \subset \mathbb{R}$ heißt ***induktiv*** $\underset{Def}{\Leftrightarrow}$

i) $1 \in I$
ii) $\forall x \in \mathbb{R}\colon x \in I \Rightarrow x+1 \in I$.

Beispielsweise sind die Mengen $\mathbb{R}$ und $\{x \in \mathbb{R} : x \geq 1\}$ induktiv.

LEMMA I-4.2. *Es sei $\mathcal{A} \neq \emptyset$ eine Familie induktiver Mengen. Dann ist auch $\cap\mathcal{A}$ induktiv.*

BEWEIS. Da $1 \in A$ für jedes $A \in \mathcal{A}$ gilt, folgt $1 \in \cap\mathcal{A}$. Es sei nun $x \in \cap\mathcal{A}$, d.h. $x \in A$ für alle $A \in \mathcal{A}$. Da jede Menge A in $\mathcal{A}$ induktiv ist, folgt $x+1 \in A$ für jedes $A \in \mathcal{A}$, d.h. $x+1 \in \cap\mathcal{A}$. □

Es bezeichne $\mathfrak{J}$ das System aller induktiven Mengen. Wir haben uns bereits von $\mathfrak{J} \neq \emptyset$ überzeugt. Nach Lemma I-4.2 ist $\cap\mathfrak{J}$ selbst induktiv. Offenbar ist $\cap\mathfrak{J}$ die kleinste induktive Menge.

DEFINITION I-4.3. *i) $\mathbb{N} := \cap\mathfrak{J}$ heißt die Menge der* ***natürlichen Zahlen****. Ferner setzen wir $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.*
ii) $n \in \mathbb{N}$ heißt gerade $\underset{Def}{\Leftrightarrow} \exists m \in \mathbb{N}\colon n = 2m$
$n \in \mathbb{N}$ heißt ungerade $\underset{Def}{\Leftrightarrow} \exists m \in \mathbb{N}_0\colon n = 2m+1$

Es ist üblich und zweckmäßig, für folgende Nachfolger von 1 spezielle Symbole zu verwenden: $2 := 1+1$, $3 := 2+1, \dots, 10 := 9+1$. Wir werden später zeigen, daß wir mit den Symbolen $(0, 1, ..., 9, +, -,$ „,“) sämtliche reellen Zahlen darstellen können.

Aus der Definition I-4.3 ergibt sich für jede induktive Menge A, $\mathbb{N} \subset A$. Dieser Umstand ist die Grundlage für das überaus nützliche Induktionsprinzip.

THEOREM I-4.4 (Prinzip der vollständigen Induktion). *Es sei $P(n)$ eine Aussageform über $\mathbb{N}$ und es gelte*

i) $P(1)$,
ii) $\forall n \in \mathbb{N}\colon P(n) \Rightarrow P(n+1)$.

Dann gilt $P(n)$ für alle $n \in \mathbb{N}$, d.h. $\forall n \in \mathbb{N}\colon P(n)$.

BEWEIS. Es sei $E = \{n \in \mathbb{N}\colon P(n)\} \subset \mathbb{N}$ die Erfüllungsmenge von P. Wegen der Voraussetzungen ist E induktiv, daher gilt auch $\mathbb{N} \subset E$. Somit folgt $\mathbb{N} = E$. □

BEMERKUNG I-4.5. Oft trifft eine Aussageform $P(n)$ erst für $n_0 > 1$ zu. Man überzeuge sich davon, daß das Prinzip der vollständigen Induktion auch für beliebigen Induktionsanfang $n_0 \in \mathbb{N}$ gilt. Mit Hilfe der Transformation $n = n_0 + k - 1$ und $\tilde{P}(k) = P(n_0 + k - 1)$ führt man den Fall $n_0 > 1$ auf Satz I-4.4 zurück.

THEOREM I-4.6. *Es seien $n, m \in \mathbb{N}$. Dann gilt*

$$\begin{array}{llll} a) & n \geq 1, & c) & n+m \in \mathbb{N}, \\ b) & n-1 \in \mathbb{N}_0, & d) & nm \in \mathbb{N}. \end{array}$$

BEWEIS. a) Übung.
b) Es sei $G = \{n \in \mathbb{N} \colon n-1 \in \mathbb{N}_0\} \subseteq \mathbb{N}$. Wir zeigen: G ist induktiv. Offensichtlich gilt $1 \in G$. Für $n \in G$ schließt man

$$(n+1)-1 = n+(1-1) = n \in \mathbb{N}_0, \text{ d.h. } n+1 \in G. \text{ Es folgt } G = \mathbb{N}.$$

c) Es sei $P(n)$ die Aussageform $\forall m \in \mathbb{N} \colon n+m \in \mathbb{N}$. Wegen der Induktivität von $\mathbb{N}$ gilt $\forall m \in \mathbb{N} \colon 1+m \in \mathbb{N}$, d.h. $P(1)$. Es gelte $P(n)$, d.h. $n+m \in \mathbb{N}$ für jedes $m \in \mathbb{N}$. Wir verwenden noch einmal die Induktivität von $\mathbb{N}$, und folgern $(n+m)+1 = (n+1)+m \in \mathbb{N}$ für alle $m \in \mathbb{N}$, also $P(n+1)$. Wegen Satz I-4.4 trifft $P(n)$ für alle $n \in \mathbb{N}$ zu.
d) Übung. □

$\mathbb{N}$ ist somit abgeschlossen bezüglich der Addition und Multiplikation.

THEOREM I-4.7. *Für natürliche Zahlen n und m gilt $n > m$ genau dann, wenn $n - m \in \mathbb{N}$.*

BEWEIS. "$\Rightarrow$" Es sei $P(m)$ die Aussageform $\forall n \in \mathbb{N} \colon m < n \Rightarrow n - m \in \mathbb{N}$. Aus $1 < n$ folgt mit $(\mathcal{OA}^*)$ $0 < n-1$. Nach Satz I-4.6 trifft $n-1 \in \mathbb{N}_0$ zu, also $n-1 = 0$ oder $n-1 \in \mathbb{N}$. Wegen $n-1 > 0$ gilt zwangsläufig $n-1 \in \mathbb{N}$, d.h. $P(1)$. Es gelte nun $P(m)$ und $n \in \mathbb{N}$ erfülle $m+1 < n$, (n sonst beliebig), d.h. $m < n-1$. Aus $1 \leq m < n-1$ schließen wir auf $n-1 \in \mathbb{N}$. Somit folgt aus $m < n-1$ wegen $P(m)$ die Gültigkeit von $(n-1)-m = n-(m+1) \in \mathbb{N}$, also von $P(m+1)$.
"$\Leftarrow$" Es sei $n - m = l$ für ein $l \in \mathbb{N}$. Die Behauptung folgt nun aus $n - m = l \geq 1 > 0$ durch Addition von m. □

KOROLLAR I-4.8. Für natürliche Zahlen n und m gilt $n > m$ genau dann, wenn $n \geq m+1$.

KOROLLAR I-4.9. Zwischen zwei unmittelbar aufeinander folgenden natürlichen Zahlen und zwischen Null und 1 liegt keine weitere natürliche Zahl.

BEWEIS. Formal geschrieben lautet die Behauptung $\forall m \in \mathbb{N}_0 \colon (m, m+1) \cap \mathbb{N} = \emptyset$. Wir führen den Beweis durch Widerspruch und nehmen vorerst an, es gäbe n, $m \in \mathbb{N}$ mit $m < n < m+1$. Wegen Satz I-4.7 gilt $n - m \in \mathbb{N}$, also nach Satz I-4.6 $n - m \geq 1$. Aus $n < m+1$ ergibt sich der Widerspruch $n - m < 1$. Für $m = 0$ erhalten wir einen Widerspruch zu I-4.6–(a). □

DEFINITION I-4.10. *$M \subset \mathbb{R}$ sei nicht leer.*

i) *$\mu \in \mathbb{R}$ heißt* ***Minimum*** *von M, $\mu = \min M$,* $\underset{Def}{\Leftrightarrow} \mu \in M \wedge \forall x \in M \colon \mu \leq x.$

ii) *$\nu \in \mathbb{R}$ heißt* ***Maximum*** *von M, $\nu = \max M$,* $\underset{Def}{\Leftrightarrow} \nu \in M \wedge \forall x \in M \colon x \leq \nu.$

Nicht jede Teilmenge reeller Zahlen besitzt Minimum oder Maximum, z.B. $\mathbb{R}$ oder $(0,1)$. Existiert ein Maximum (Minimum) einer Teilmenge reeller Zahlen, so ist diese Zahl wegen der linearen Ordnung in $\mathbb{R}$ natürlich eindeutig bestimmt.

THEOREM I-4.11 (Wohlordnungssatz). *Jede nicht leere Menge von natürlichen Zahlen besitzt ein Minimum.*

BEWEIS. Nehmen wir an, es gäbe $A \in \mathcal{P}(\mathbb{N}) \setminus \{\emptyset\}$ und A besitze kein Minimum. Es sei $S = \{n \in \mathbb{N} \colon (\forall a \in A \colon n < a)\}$. Da A nicht leer ist, folgt $S \subsetneq \mathbb{N}$. Wegen Satz I-4.6 a) ist $1 \in S$, sonst wäre $1 = \min A$. Es sei $n \in S$ und somit nach Korollar I-4.8

$n+1 \leq a$ für alle $a \in A$. Wäre $n+1$ nicht Element von S, dann gäbe es $\mu \in A$ mit $n < \mu \leq n+1$. Wegen Korollar I-4.9 müßte $\mu = n+1$ gelten. Dies hätte $\mu = \min A$ zur Folge, im Widerspruch zur Wahl von A. Also gilt $n+1 \in S$ und somit $S = \mathbb{N}$. Dies ist ein Widerspruch zu $S \subsetneq \mathbb{N}$. □

Manchmal kommt es vor, daß für den Induktionsschritt nicht nur die unmittelbar vorangehende Aussage gebraucht wird, sondern alle vorausgehenden. Als Anwendung des Wohlordnungssatzes zeigen wir folgende Variante des Induktionsprinzipes.

Theorem I-4.12. *$P(n)$ sei eine Aussageform über $\mathbb{N}$, $n_0 \in \mathbb{N}$ mit den Eigenschaften:*

i) *$P(n_0)$,*
ii) *$\forall n \geq n_0 \forall k \in \mathbb{N}\colon (n_0 \leq k \leq n \wedge P(k)) \Rightarrow P(n+1)$.*

Dann gilt $P(n)$ für alle $n \geq n_0$.

Beweis. Wir führen einen Widerspruchsbeweis und nehmen an, es gelte i) und ii), aber die Aussageform $P(n)$ trifft nicht für alle $n \geq n_0$ zu. Dann ist die Menge $\{n \in \mathbb{N}\colon n \geq n_0 \wedge \neg P(n)\}$ nicht leer und besitzt wegen des Wohlordnungssatzes ein Minimum $m_0 \in \mathbb{N}$. Dann gilt $P(k)$ für alle $n_0 \leq k \leq m_0 - 1$ und somit wegen ii) auch für $k = m_0$. Dies ist wegen der Bedeutung des Index m_0 nicht möglich. □

Definition I-4.13. *$x \in \mathbb{R}$ heißt **ganze Zahl** $\underset{Def}{\Leftrightarrow} x \in \mathbb{N}_0 \vee -x \in \mathbb{N}_0$.*
Die Menge der ganzen Zahlen wird mit $\mathbb{Z}$ bezeichnet.

Es sei dem Leser überlassen, die Sätze I-4.6 c), d), I-4.7 und Korollar I-4.8, I-4.9 auf $\mathbb{Z}$ zu übertragen.

Das Prinzip der vollständigen Induktion ist auch die Grundlage für sogenannte *rekursive Definitionen*, bei denen für alle $n \in \mathbb{N}$ Begriffe $B(n)$ definiert werden, indem man auf die bereits erklärten Begriffe $B(1), \ldots, B(n-1)$ zurückgreift. Betrachten wir zum Beispiel die Folge 1, 2, 4, 8, 16,... (der Begriff einer Folge wird später genauer definiert). Die Ellipsis ... suggeriert das Bildungsgesetz $a_n = 2^{n-1}$, $n \in \mathbb{N}$. Dieselbe Folge kann jedoch auch rekursiv beschrieben werden. Dazu setzt man $a_1 = 1$ und $a_{n+1} = 2a_n$ für alle $n \in \mathbb{N}$. Aus a_1 ergibt sich a_2, aus a_2 wiederum a_3 usw. Durch die Rekursionsvorschrift wird zumindest in diesem Beispiel eine Folge eindeutig bestimmt. Das rekursive Vorgehen kann durch folgenden Satz gerechtfertigt werden.

Theorem I-4.14 (Rekursionssatz). *Es sei B eine nichtleere Menge und $b \in B$. Für jedes $n \in \mathbb{N}$ sei eine Abbildung $F_n\colon B^n \to B$ gegeben. Dann gibt es genau eine Abbildung $\varphi\colon \mathbb{N}_0 \to B$ mit den Eigenschaften*

(1) *$\varphi(0) = b$.*
(2) *$\varphi(n+1) = F_{n+1}(\varphi(0), \ldots, \varphi(n))$ für alle $n \in \mathbb{N}_0$.*

Natürlich kann die Rekursion auch bei jedem $n_0 \in \mathbb{N}_0$ beginnen.

Beweis. Wir zeigen zunächst durch vollständige Induktion, daß es höchstens eine derartige Abbildung φ geben kann. Angenommen es gäbe zwei Abbildungen φ und $\psi\colon \mathbb{N}_0 \to B$ mit $\varphi(0) = \psi(0) = b$ und

$$\begin{aligned}\varphi(n+1) &= F_{n+1}(\varphi(0), \dots, \varphi(n))\\ \psi(n+1) &= F_{n+1}(\psi(0), \dots, \psi(n))\end{aligned}$$

für $n \in \mathbb{N}_0$. Dann gilt $\varphi(0) = \psi(0)$. Aus der Induktionsannahme $\varphi(k) = \psi(k)$ für $0 \le k \le n$ folgt $\varphi(n+1) = \psi(n+1)$ und daher $\varphi = \psi$ nach dem Induktionsprinzip in der Form von Satz I-4.12. Für die Konstruktion von φ benötigen wir Hilfsfunktionen $\varphi_n\colon \{0, \dots, n\} \to B$, $n \in \mathbb{N}_0$, mit folgenden Eigenschaften

$$\begin{aligned}&\varphi_n(0) = b,\\ (*)\qquad &\varphi_n(k) = \varphi_k(k),\\ &\varphi_n(k+1) = F_{k+1}(\varphi_n(0), \dots, \varphi_n(k)), \qquad 0 \le k \le n-1.\end{aligned}$$

Definiert man nun $\varphi\colon \mathbb{N}_0 \to B$ durch

$$\varphi(n) = \begin{cases} b, & n = 0,\\ \varphi_n(n), & n \in \mathbb{N},\end{cases}$$

folgt aus den Eigenschaften von φ_n

$$\begin{aligned}\varphi(n+1) = \varphi_{n+1}(n+1) &= F_{n+1}(\varphi_{n+1}(0), \dots, \varphi_{n+1}(n))\\ &= F_{n+1}(\varphi_0(0), \dots, \varphi_n(n)) = F_{n+1}(\varphi(0), \dots, \varphi(n)),\end{aligned}$$

d.h. φ ist die gesucht Funktion.
Wir zeigen nun induktiv die Existenz der Hilfsfunktionen φ_n. Für $n = 0$ ist φ_0 allein durch $\varphi_0(0) = b$ bestimmt. Die beiden zusätzlichen Eigenschaften in $(*)$ sind trivial erfüllt, da es keine natürliche Zahl $0 \le k < 0$ gibt, für welche sie nicht zutreffen könnten. Für den Induktionsschritt nehmen wir an, es seien bereits Funktionen $\varphi_0, \dots, \varphi_n$ mit den in $(*)$ geforderten Eigenschaften konstruiert und setzen

$$\varphi_{n+1}(k) = \begin{cases} \varphi_n(k) & 0 \le k \le n\\ F_{n+1}(\varphi_n(0), \dots, \varphi_n(n)) & k = n+1.\end{cases}$$

Verwendet man für $0 \le k < n$ die Induktionsvoraussetzung und für $k = n$ die Definition von φ_{n+1}, erhält man

$$\varphi_{n+1}(k) = \varphi_n(k) = \varphi_k(k), \qquad 0 \le k \le n.$$

Für $0 \le k < n+1$ schließt man wieder mit Hilfe der Induktionsvoraussetzung auf

$$\begin{aligned}\varphi_{n+1}(k+1) = \varphi_n(k+1) &= F_{k+1}(\varphi_n(0), \dots, \varphi_n(k))\\ &= F_{k+1}(\varphi_{n+1}(0), \dots, \varphi_{n+1}(k)), \qquad 0 \le k \le n-1.\end{aligned}$$

Für $k = n$ ergibt sich

$$\begin{aligned}\varphi_{n+1}(n+1) &= F_{n+1}(\varphi_n(0), \dots, \varphi_n(n))\\ &= F_{n+1}(\varphi_{n+1}(0), \dots, \varphi_{n+1}(n)).\end{aligned}$$

Somit ist der Induktionsschritt bewiesen und die Existenz der Funktionen φ_k gesichert. □

Definition I-4.15. *Es sei $x \in \mathbb{R}$. Wir definieren*

$$x^1 := x, \quad x^{n+1} := x^n x, \quad n \in \mathbb{N},$$

Man nennt x^n, $n \in \mathbb{N}_0$, n-te ***Potenz*** *von x, x heißt* **Basis** *und n* **Exponent**.
Für $x \in \mathbb{R} \setminus \{0\}$setzen wir

$$x^0 := 1, \quad x^{-n} := (x^n)^{-1}, n \in \mathbb{N}.$$

Diese Definition kann folgendermaßen auf den Rekursionssatz zurückgeführt werden: setze $B = \mathbb{R}$, $b = x$ und $F_n\colon \mathbb{R}^{n-1} \to \mathbb{R}$, $F_n(x_1 \dots, x_{n-1}) = x_{n-1}x$. Die Funktion φ erfüllt dann $\varphi(1) = x$ und $\varphi(n+1) = \varphi(n)x$, $n \in \mathbb{N}$. Anstelle von $\varphi(n)$ schreibt man x^n.

Lemma I-4.16. *Es sei $n \in \mathbb{N}$, $x \in \mathbb{R}$ und $x \ne 0$. Dann gilt $\frac{1}{x^n} = (\frac{1}{x})^n$ bzw. $(x^n)^{-1} = (x^{-1})^n$.*

Beweis. Wir führen einen Induktionsbeweis. Für $n = 1$ stimmt die Behauptung mit der Definition des multiplikativen inversen Elementes von x überein. Es gelte $\frac{1}{x^n} = (\frac{1}{x})^n$. Der Induktionsschritt ergibt sich aus

$$(\frac{1}{x})^{n+1} = (\frac{1}{x})^n \frac{1}{x} = \frac{1}{x^n}\frac{1}{x} = \frac{1}{x^n x} = \frac{1}{x^{n+1}} = (x^{n+1})^{-1}.$$

□

Theorem I-4.17 (Potenzgesetze). *1.) Für $x, y \in \mathbb{R} \setminus \{0\}$, $n, m \in \mathbb{Z}$ gilt*
 a) $x^m x^n = x^{m+n}$,
 b) $x^n y^n = (xy)^n$,
 c) $(x^m)^n = x^{mn}$,

Für $n, m \in \mathbb{N}$ gelten a)–c) für $x, y \in \mathbb{R}$.
2.) Für $n \in \mathbb{N}$ und $x \geq 0, y > 0$ gilt

$$x < y \Leftrightarrow x^n < y^n.$$

Beweis. Wir zeigen nur a), die übrigen Aussagen möge der Leser als Übung beweisen. Für $m = 0$ oder $n = 0$ ist die Behauptung klar. Es sei $m \in \mathbb{Z}$ beliebig gewählt und $P(n)$ das Prädikat $x^m x^n = x^{m+n}$ für $n \in \mathbb{N}$. Für $m \in \mathbb{N}$ ist $P(1)$ äquivalent zur rekursiven Definition der Potenz. Für $m < 0$ ergibt sich

$$x^m x = \frac{x}{x^{|m|}} = \frac{x}{x^{|m|-1}x} = \frac{1}{x^{|m|-1}} = \frac{1}{x^{-(m+1)}} = x^{m+1},$$

d.h. es gilt $P(1)$. (Die zweite Gleichheit begründet man mit der Definition I-4.15). Es gelte nun $P(n)$. Verwendet man die Definition I-4.15 und $P(n)$, ergibt sich

$$x^m x^{n+1} = x^m(x^n x) = (x^m x^n)x = x^{m+n}x = x^{m+n+1},$$

also $P(n+1)$. Die letzte Gleichheit wird durch $P(1)$ mit $m+n$ anstelle von m gerechtfertigt. Schließlich seien n, $m \in \mathbb{Z}$, $n, m < 0$. Dann gilt

$$x^n x^m = \frac{1}{x^{|n|}}\frac{1}{x^{|m|}} = \frac{1}{x^{|n|}x^{|m|}} = \frac{1}{x^{|n|+|m|}} = x^{-(|n|+|m|)} = x^{n+m}.$$

□

Wir beenden diesen Abschnitt mit weiteren Beispielen zur vollständigen Induktion.

Lemma I-4.18 (Bernoullische Ungleichung). *Es sei $x \in \mathbb{R}$ und $n \in \mathbb{N}$. Für $x > -1$, $x \neq 0$ und $n > 1$ gilt dann*

$$(1+x)^n > 1 + nx.$$

Beweis. i) Induktionsanfang: Für $n = 2$ gilt $(1+x)^2 > 1 + 2x$.
ii) Induktionsschritt: Es sei $n \in \mathbb{N}$ und es gelte $(1+x)^n > 1+nx$. Nach Voraussetzung ist $1+x > 0$. Somit folgt aus der Induktionsvoraussetzung $(1+x)^{n+1} > (1+nx)(1+x)$. Die Abschätzung

$$(1+nx)(1+x) = 1 + (n+1)x + nx^2 > 1 + (n+1)x$$

ergibt wegen der Transitivität von $>$ die gewünschte Ungleichung

$$(1+x)^{n+1} > 1 + (n+1)x.$$

□

Ein weiteres Beispiel einer rekursiven Definition ist folgende kompakte Schreibweise für endliche Summen $a_0 + a_2 + \cdots + a_n$ und endliche Produkte $a_0 \cdots a_n$. Die rekursive Definition präzisiert die Bedeutung der Ellipsis. Dazu setzen wir im Rekursionssatz I-4.14 $B = \mathbb{R}$, $F_n\colon B^n \to B$, $F_n(x_0, \ldots, x_{n-1}) = x_{n-1} * a_n$ und erhalten so eine Funktion $\varphi\colon \mathbb{N}_0 \to B$ mit $\varphi(0) = a_0$ und $\varphi(n+1) = \varphi(n) * a_{n+1}$. Bezeichnet $*$ die Addition in $\mathbb{R}$, schreibt man $\sum_{k=0}^n a_k$ für $\varphi(n)$, steht $*$ für die Multiplikation, schreibt man $\prod_{k=0}^n a_k$:

Definition I-4.19. *Es seien* $a_0, a_1, \ldots, \in \mathbb{R}$.

i) $\sum_{j=0}^{0} a_j := a_0$ *und* $\sum_{j=0}^{n} a_j := \sum_{j=0}^{n-1} a_j + a_n$, $n \in \mathbb{N}$.
ii) $\prod_{j=0}^{0} a_j = a_0$ *und* $\prod_{j=0}^{n} a_j = \prod_{j=0}^{n-1} a_j \cdot a_n$, $n \in \mathbb{N}$.

Für die leere Summe vereinbaren wir den Wert 0, das leere Produkt erhält den Wert 1.

$\sum$ heißt Summenzeichen, der Buchstabe j ist ein (gebundener) Summationsindex. Er kann beliebig umbenannt werden. Analoges gilt für das Produkt.

Definition I-4.20. *Die Zahlen*

i) $0! := 1, \forall n \in \mathbb{N}_0 \colon (n+1)! = n!(n+1)$ *heißen* ***Fakultäten****.*
ii) *Für* $n \in \mathbb{N}_0$ *und* $0 \leq k \leq n$ *heißt*
$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$
Binomialkoeffizient *(*$\binom{n}{k}$ *wird „n über k“ gelesen).*

Es gilt also $n! = \prod_{i=1}^{n} i$. Die Fakultäten wachsen sehr rasch an: $1! = 1, 2! = 2, 3! = 6, 4! = 24, 10! = 3\ 628\ 800$. Diese Zahlen werden bei Abzählproblemen verwendet. Beispielsweise ist $n!$ die Anzahl der Möglichkeiten, n unterscheidbare Objekte in verschiedener Reihenfolge anzuordnen. Der Binomialkoeffizient $\binom{n}{k}$ gibt die Anzahl der Teilmengen mit k Elementen in einer Grundmenge von n Elementen.

Lemma I-4.21. *Für alle* $n, k \in \mathbb{N}_0$, $0 \leq k \leq n-1$ *gilt*

i) $\binom{n}{k} \in \mathbb{N}$, $\binom{n}{0} = \binom{n}{n} = 1$.
ii) $\binom{n}{k+1} + \binom{n}{k} = \binom{n+1}{k+1}$,

Beweis. Übung. □

Ordnet man die Binomialkoeffizenten im sogenannten **Pascalschen Dreieck** an,

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \end{array}$$

dann sagt Lemma I-4.21-ii), daß man die Binomialkoeffizienten der $n+1$-ten Reihe als Summe der beiden unmittelbar darüber stehenden Binomialkoeffizienten berechnen kann

Theorem I-4.22 (Binomischer Satz). *Es seien* x, y *reelle Zahlen. Dann gilt für alle* $n \in \mathbb{N}$*:*
$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

Beweis. 1) Induktionsanfang: Für $n = 1$ ergibt sich

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

2) Induktionsschritt: Sei $n \in \mathbb{N}$ und es gelte $(x+y)^n = \sum\limits_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$. Durch Multiplikation mit $x + y$ folgt

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

Ersetzt man in der zweiten Summe den Summationsindex k durch $j - 1$, ergibt sich weiter

$$\begin{aligned}
(x+y)^{n+1} &= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k} x^{n+1-k} y^k + \sum_{j=1}^{n} \binom{n}{j-1} x^{n+1-j} y^j + y^{n+1} \\
&= x^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] x^{n+1-k} y^k + y^{n+1} \\
&\overset{I-4.21}{=} \binom{n+1}{0} x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} x^{n+1-k} y^k + \binom{n+1}{n+1} y^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k} y^k
\end{aligned}$$

□

Ohne Beweis notieren wir die Identitäten:

Lemma I-4.23. i) $\forall n \in \mathbb{N}\, \forall q \in \mathbb{R} \setminus \{1\} \colon \sum\limits_{k=0}^{n} q^k = \frac{q^{n+1}-1}{q-1}$,

ii) $\forall n \in \mathbb{N}\, \forall x, y \in \mathbb{R} \colon x^n - y^n = (x - y) \sum\limits_{i=0}^{n-1} x^{n-1-i} y^i$ *(Hornersche Regeln).*

5. Die rationalen Zahlen

Definition I-5.1.

$$\mathbb{Q} := \{x \in \mathbb{R} : x = \frac{p}{q}, p \in \mathbb{Z}, q \in \mathbb{N}\}$$

heißt die Menge der ***rationalen Zahlen****. $\mathbb{R} \setminus \mathbb{Q}$ ist die Menge der* ***irrationalen Zahlen****.*

Es kann leicht gezeigt werden, daß für $x, y \in \mathbb{Q}$ auch $-x, x+y, xy$ und, falls $x \neq 0$, auch x^{-1} zu $\mathbb{Q}$ gehören. $\mathbb{Q}$ erfüllt die Axiome $\mathcal{A}, \mathcal{M}, \mathcal{D}, \mathcal{O}$. $\mathbb{Q}$ ist somit ein geordneter Unterkörper von $\mathbb{R}$, in dem alle bisher abgeleiteten Sätze gelten. Jedoch besitzen bereits sehr einfache Probleme keine Lösung in $\mathbb{Q}$:

Theorem I-5.2. *Die Gleichung $x^2 = 2$ besitzt keine Lösung in $\mathbb{Q}$.*

Beweis. Angenommen es gäbe ein $x \in \mathbb{Q}$ mit $x^2 = 2$. Es ist $x = \frac{p}{q}$, $p \in \mathbb{Z}, q \in \mathbb{N}$. O.B.d.A. seien p und q nicht beide zugleich gerade. Aus $p^2 = 2q^2$ folgt p^2 und somit auch p sind gerade, d.h. es existiert eine ganze Zahl n mit $p = 2n$. Dies hat $4n^2 = 2q^2$, also $q^2 = 2n^2$ zur Folge. Somit ist auch q^2 und damit auch q gerade. Dies ist ein Widerspruch zur Wahl von p und q. In diesem Beweis haben wir folgende Eigenschaft ganzer Zahlen verwendet, deren einfachen Nachweis wir dem Leser überlassen: $\forall p \in \mathbb{Z}\colon p^2$ gerade $\Rightarrow p$ gerade. $\square$

Analysieren wir diesen Sachverhalt genauer: Es sei

$$A = \{p \in \mathbb{Q} : p < 0 \vee p^2 < 2\} \text{ und } B = \{p \in \mathbb{Q} : p \geq 0, p^2 > 2\}.$$

Wir zeigen: A besitzt kein Maximum und B kein Minimum in $\mathbb{Q}$. Zu diesem Zweck definieren wir für jedes rationale $p > 0$ die Zahl $q \in \mathbb{Q}$ durch

$$q = p - \frac{p^2 - 2}{p + 2} \begin{cases} > p & \text{für } p \in A \text{ und } p \geq 0, \\ < p & \text{für } p \in B. \end{cases}$$

Eine einfache Rechnung zeigt

$$q^2 - 2 = 2\frac{p^2 - 2}{(p+2)^2} \begin{cases} < 0 & \text{für } p \in A, \\ > 0 & \text{für } p \in B. \end{cases}$$

Aus diesen Beziehungen folgt für jedes nicht negative $p \in A$ einerseits $p < q$, andererseits $q^2 - 2 < 0$, d.h. $q \in A$. Es gibt somit kein Maximum in A. Für jedes $p \in B$ ergibt sich $q < p$ und $q^2 - 2 > 0$, d.h. $q \in B$. B hat demnach kein Minimum. Andererseits ist $A \cup B = \mathbb{Q}$. Das Paar (A, B) definiert einen Dedekindschen Schnitt in $\mathbb{Q}$, aber es existiert keine Schnittzahl in $\mathbb{Q}$. Das System der rationalen Zahlen hat also Lücken, welche in $\mathbb{R}$ durch das Vollständigkeitsaxiom geschlossen werden.

6. Folgerungen aus dem Vollständigkeitsaxiom

Definition I-6.1. *Es sei $M \subset \mathbb{R}$.*

(1) *Eine reelle Zahl o heißt **obere Schranke** für M $\underset{Def}{\Leftrightarrow} \forall x \in M : x \leq o$.*
*M heißt nach **oben beschränkt**, wenn M eine obere Schranke besitzt.*
(2) *Eine reelle Zahl u heißt **untere Schranke** für M $\underset{Def}{\Leftrightarrow} \forall x \in M : x \geq u$.*
*M heißt nach **unten beschränkt**, wenn M eine untere Schranke besitzt.*
(3) *$M \subset \mathbb{R}$ ist **beschränkt** $\underset{Def}{\Leftrightarrow}$ M ist nach oben und unten beschränkt.*

DEFINITION I-6.2. *a) Es sei $M \subset \mathbb{R}$ nach oben beschränkt.*
$\alpha \in \mathbb{R}$ heißt ***Supremum*** *von M, $\alpha = \sup M \underset{Def}{\Leftrightarrow} \alpha$ ist die kleinste obere Schranke.*
b) Es sei $M \subset \mathbb{R}$ nach unten beschränkt.
$\beta \in \mathbb{R}$ heißt ***Infimum*** *von M, $\beta = \inf M \underset{Def}{\Leftrightarrow} \beta$ ist die größte untere Schranke.*

Wegen der Trichotomie in $\mathbb{R}$ kann eine nichtleere Teilmenge von $\mathbb{R}$ höchstens ein Supremum bzw. Infimum besitzen. Ist das Supremum einer Menge M selbst Element von M, gilt $\sup M = \max M$. Umgekehrt, besitzt eine Menge M ein Maximum, so ist natürlich auch $\max M = \sup M$.

BEISPIEL I-6.3. a) $M = \{x \in \mathbb{R} : x \geq 1\}$ ist nicht nach oben beschränkt. M besitzt somit kein Supremum, M ist nach unten beschränkt ($x \geq 1 \Rightarrow 1$ ist untere Schranke) und $1 \in M$. Somit gilt $\inf M = \min M = 1$.
2) Für $N = (0, 1]$ gilt $\sup N = \max N = 1 \in N$, $\inf N = 0 \notin N$.
3) Die leere Menge besitzt in $\mathbb{R}$ weder Supremum noch Infimum.

Es sei $\alpha = \sup M$ und $x < \alpha$. Da α die kleinste obere Schranke von M ist, kann α keine obere Schranke für M sein. Somit existiert $y \in M$ mit $x < y$. Dies führt auf folgende Charakterisieung des Supremums:

THEOREM I-6.4. *Es sei $\emptyset \neq M \subset \mathbb{R}$. Dann ist $\alpha = \sup M$ charakterisiert durch*
i) $\forall x \in M \colon x \leq \alpha$,
ii) $\forall \varepsilon > 0 \exists y \in M \colon \alpha - \varepsilon < y$.

BEWEIS. Die erste Eigenschaft besagt, daß α eine obere Schranke für M dastellt. Die Eigenschaft ii) stellt fest, daß jede reelle Zahl $z < \alpha$ keine obere Schranke für M sein kann. $\square$

Dieser Satz hat ein merkwürdiges Häufungsphänomen zur Folge: Besitzt eine nicht leere Menge M ein Supremum, jedoch kein Maximum, so liegen für jedes $\varepsilon > 0$ „unendlich" viele Elemente von M zwischen $\sup M - \varepsilon$ und $\sup M$.

Wir formulieren nun eine wichtige Konsequenz aus dem Vollständigkeitsaxiom:

THEOREM I-6.5 (Supremum Prinzip). *Jede nicht leere, nach oben beschränkte Menge reeller Zahlen besitzt ein Supremum in $\mathbb{R}$.*

BEWEIS. Es sei $M \subset \mathbb{R}$, $M \neq \emptyset$, nach oben beschränkt. Ferner sei B die Menge aller oberen Schranken von M und $A = \mathbb{R} \setminus B$. Da M nicht leer ist, gibt es $x \in M$. Es folgt $x - 1 \in A$. Somit gilt $A \neq \emptyset$ und $B \neq \emptyset$, $A \cup B = \mathbb{R}$. Es seien $a \in A$ und $b \in B$ beliebig gewählt. Dann ist a keine obere Schranke für M, also gibt es $x \in M$ mit $a < x \leq b$, d.h. es ist $a < b$. Die Mengen A und B definieren also einen Dedekindschen Schnitt. Das Vollständigkeitsaxioms $\mathcal{V}$ sichert die Existenz einer eindeutig bestimmten Schnittzahl ξ, welche entweder in A oder in B liegen muß. Wäre $\xi \in A$, also

$$A = \{x \in \mathbb{R} \colon x \leq \xi\} \tag{$*$}$$

gäbe es (wie vorher) $x \in M$, sodaß $\xi < x$. Wegen Satz I-3.6 würde auch $\xi < \frac{\xi+x}{2} < x$ gelten, d.h. $\frac{\xi+x}{2} \in A$. Diese Ungleichung ist ein Widerspruch zu $(*)$. Es folgt $\xi = B$ und daher

$$(**) \qquad B = \{x \in \mathbb{R}\colon \xi \leq x\}$$

Somit ist ξ eine obere Schranke für M und wegen $(**)$ sogar die kleinste, d.h. $\xi = \sup M$. □

Korollar I-6.6. Jede nicht leere, nach unten beschränkte Menge reeller Zahlen besitzt ein Infimum in $\mathbb{R}$.

Beweis. Es sei $N \subset \mathbb{R}$, $N \neq \emptyset$, nach unten beschränkt und U die Menge der unteren Schranken. Somit ist $U \neq \emptyset$ und jedes Element aus N ist eine obere Schranke für U. Wegen Satz I-6.5 existiert $s = \sup U$. Wir behaupten $s = \inf N$. Dies folgt aus $s \in U$, da dann $s = \max U$ gilt. Angenommen, es wäre $s \notin U$, dann gäbe es $x \in N$, sodaß $x < s = \sup U$. Da s die kleinste obere Schranke von U ist, kann x nicht obere Schranke von U sein. Somit existiert $y \in U$ mit $x < y$. Wegen $x \in N$ kann y keine untere Schranke von N sein. Dies führt auf den Widerspruch $y \notin U$. □

Ein einfacherer Beweis kann auf die Beobachtung aufgebaut werden, daß $N \subset \mathbb{R}$ genau dann nach unten beschränkt ist, wenn $-N := \{t \in \mathbb{R}\colon -t \in N\}$ nach oben beschränkt ist. Die angegebene Beweisvariante verwendet keine speziellen Eigenschaften von $\mathbb{R}$. In jeder linear geordneten Menge folgt daher aus dem Supremumprinzip das Infimumprinzip und umgekehrt.

Theorem I-6.7. $\mathbb{N}$ *ist nicht nach oben beschränkt.*

Beweis. Angenommen, $\mathbb{N}$ ist nach oben beschränkt. Nach Satz I-6.5 existiert dann $\alpha = \sup \mathbb{N}$. Da $\alpha - 1$ keine obere Schranke für $\mathbb{N}$ sein kann, existiert $n \in \mathbb{N}$, sodaß $\alpha - 1 < n$. Es folgt $\alpha < n + 1$, ein Widerspruch zu $\forall m \in \mathbb{N} : m \leq \alpha$. □

Korollar I-6.8. $\mathbb{R}$ ist **archimedisch** geordnet, d.h. für alle $a, b \in \mathbb{R}$, $a > 0$, existiert $n \in \mathbb{N}$, sodaß $na > b$.

Beweis. Da $\mathbb{N}$ nicht nach oben beschränkt ist, gibt es $n \in \mathbb{N}$, sodaß $\frac{b}{a} < n$. □

Korollar I-6.9. $\forall \varepsilon > 0\ \exists n \in \mathbb{N}\colon \frac{1}{n} < \varepsilon$.

Beweis. Wähle $b = 1, a = \varepsilon$ in Korollar I-6.8. □

Theorem I-6.10. *Es sei $x \in \mathbb{R}$. Dann gibt es eine eindeutig bestimmte ganze Zahl g, sodaß*

$$(*) \qquad g \leq x < g + 1, \qquad (\Leftrightarrow x - 1 < g \leq x).$$

*Man nennt g **Größtes Ganzes** von x und bezeichnet diese Zahl mit $\lfloor x \rfloor$.*

Beweis. Wir zeigen zunächst die Eindeutigkeit und nehmen an, es gäbe zwei verschiedene Zahlen $g, \tilde{g} \in \mathbb{Z}$, o.B.d.A. $g < \tilde{g}$, sodaß $(*)$ gilt. Es folgt $0 < \tilde{g} - g \in \mathbb{N}$ und $\tilde{g} \leq x < g + 1$, d.h. $\tilde{g} - g < 1$. Insgesamt gilt $0 < \tilde{g} - g < 1$, dies widerspricht

Korollar I-4.9. Die Menge $\{n \in \mathbb{N}\colon |x| < n\}$ ist nicht leer und besitzt daher ein kleinstes Element n_0 (Satz I-4.11). Somit gilt $n_0 - 1 \le x < n_0$, falls $x \ge 0$ und $-n_0 < x \le -(n_0 - 1)$ falls $x < 0$. Wenn $x \ge 0$, wähle $g = n_0 - 1$. Ist $x < 0$, setze $g = -n_0$ falls $x \notin \mathbb{Z}$ und $g = -n_0 + 1$ falls $x \in \mathbb{Z}$. $\square$

Wir zeigen nun, daß die rationalen Zahlen „dicht" in $\mathbb{R}$ liegen.

Theorem I-6.11. *Zwischen zwei verschiedenen reellen Zahlen liegt stets eine rationale Zahl.*

Beweis. Es seien x, y reelle Zahlen und $x \ne y$. O.B.d.A. können wir $x < y$, also $y - x > 0$ annehmen. Nach Korollar I-6.9 existiert $n_0 \in \mathbb{N}$, sodaß $0 < \frac{1}{n_0} < y - x$, d.h. $1 + n_0 x < n_0 y$. Nach Satz I-6.10 (mit $1 + n_0 x$ anstelle von x) folgt die Existenz von $k_0 \in \mathbb{Z}$, mit $n_0 x < k_0 \le n_0 x + 1 < n_0 y$. Für $p_0 = \frac{k_0}{n_0} \in \mathbb{Q}$ gilt dann $x < p_0 < y$. $\square$

Theorem I-6.12 (Intervallschachtelung). *Es sei $\{I_n\}_{n=1}^\infty$ eine Familie abgeschlossener Intervalle $I_n := [a_n, b_n]$, $a_n \le b_n$, in $\mathbb{R}$ mit folgenden Eigenschaften:*

a) $\forall n \in \mathbb{N}\colon I_{n+1} \subset I_n$,
b) $\forall \varepsilon > 0\ \exists n \in \mathbb{N}\colon b_n - a_n < \varepsilon$.

Dann gibt es genau ein $z \in \mathbb{R}$, sodaß $\bigcap_{n=1}^\infty I_n = \{z\}$.

Beweis. Wegen a) gilt: $a_n \le a_{n+1} \le b_{n+1} \le b_n$ für alle $n \in \mathbb{N}$. Dies hat

$$\forall n, k \in \mathbb{N}\colon a_n \le b_k. \tag{1}$$

zur Folge. Angenommen, es gäbe Indizes $n_0, k_0 \in \mathbb{N}$ mit $b_{k_0} < a_{n_0}$. Ist $n_0 \le k_0$, müßte auch $b_{k_0} < a_{n_0} \le a_{k_0}$ gelten, im Widerspruch zu $a_{k_0} \le b_{k_0}$. Ist $k_0 < n_0$, dann folgt $b_{n_0} \le b_{k_0} < a_{n_0}$ im Gegensatz zu $a_{n_0} \le b_{n_0}$. Setzt man $A = \{a_n\colon n \in \mathbb{N}\}$ und $B = \{b_n\colon n \in \mathbb{N}\}$, ergibt sich aus (1), daß jedes b_n eine obere Schranke für A, und jedes a_n eine untere Schranke für B darstellt. Somit existieren $\alpha = \sup A$ und $\beta := \inf B$ und es gilt $\alpha \le b_n$ für alle $n \in \mathbb{N}$. Es ist also α eine untere Schranke für B, was $\alpha \le \beta$ zur Folge hat. Ferner gilt $a_n \le \alpha \le \beta \le b_n$ für alle $n \in \mathbb{N}$, d.h. $[\alpha, \beta] \subset \bigcap_{n=1}^\infty I_n$. Es folgt $\bigcap_{n=1}^\infty I_n \ne \emptyset$. Es gilt aber auch $\bigcap_{n=1}^\infty I_n \subset [\alpha, \beta]$ und daher auch $\bigcap_{n=1}^\infty I_n = [\alpha, \beta]$. Denn $x \in \bigcap_{n=1}^\infty I_n$ ist gleichwertig mit $\forall n \in \mathbb{N}\colon a_n \le x \le b_n$. Somit ist x eine obere Schranke für A und gleichzeitig eine untere Schranke für B. Es folgt $\alpha \le x \le \beta$, also $x \in [\alpha, \beta]$. Wäre $\alpha < \beta$, dann müßte auch $b_n - a_n \ge \beta - \alpha$ für alle $n \in \mathbb{N}$ gelten. Dies widerspricht jedoch der Voraussetzung b). Es gilt somit $\bigcap_{n=1}^\infty I_n = [\alpha, \alpha] = \sup A = \inf B$. $\square$

Bemerkung I-6.13. 1.) Setzt man den Begriff einer Nullfolge voraus, verlangt die Voraussetzung b), daß die Folge der Intervalllängen eine Nullfolge darstellt.
2.) Man kann zeigen, daß das Vollständigkeitsaxiom, das Supremumprinzip und der Satz über die Intervallschachtelung äquivalent sind.

Definition I-6.14. *Es seien* $A, B \subset \mathbb{R}$ *nicht leer und* $r \in \mathbb{R}$. *Wir definieren*

$$\begin{aligned} A + B &:= \{a + b : a \in A, b \in B\}, \\ AB &:= \{ab : a \in A, b \in B\}, \\ rA &:= \{ra : a \in A\}. \end{aligned}$$

Theorem I-6.15. $A, B \subset \mathbb{R}$ *seien nicht leer und nach oben beschränkt. Dann gilt*

i) *Falls* $A \subset B$, *dann gilt* $\sup A \leq \sup B$,
ii) $\sup(A + B) = \sup A + \sup B$,
iii) $\forall r \geq 0 : \sup(rA) = r \sup A$,
iv) $\forall r \leq 0 : \inf(rA) = r \sup A$,
v) $A \subset \mathbb{R}^+ \wedge B \subset \mathbb{R}^+ \Rightarrow \sup AB = \sup A \cdot \sup B$.

Sind A und B nach unten beschränkt, so gilt ein entsprechender Satz für das Infimum.

Der Beweis sei dem Leser als Übung überlassen.

7. Wurzeln

Mit Hilfe des Supremumprinzips kann man die Existenz n-ter Wurzeln aus nichtnegativen Zahlen zeigen. Wir benötigen folgende Hilfsresultate.

Lemma I-7.1. *Für alle* $0 < x < y \leq 1$ *und alle natürlichen Zahlen* $n \geq 2$ *gilt*

$$y^n - x^n < n(y - x).$$

Beweis. Der Beweis ergibt sich aus den Hornerschen Regeln Lemma I-4.23. □

Lemma I-7.2. *Für alle reellen Zahlen* $0 < q < 1$ *und* $\varepsilon > 0$ *existiert* $n \in \mathbb{N}$ *mit* $0 < q^n < \varepsilon$.

Beweis. Wir zeigen für ein geeignetes $n \in \mathbb{N}$ die gleichwertige Ungleichung

$$0 < \frac{1}{\varepsilon} < (\frac{1}{q})^n.$$

Dazu beachte man $\frac{1}{q} > 1$, also $\frac{1}{q} = 1 + x$ für ein $x > 0$. Aus der Bernoulli Ungleichung I-4.18

$$(\frac{1}{q})^n = (1 + x)^n > 1 + nx.$$

Die Behauptung folgt nun aus der archimedischen Ordnung von $\mathbb{R}$ (Korollar I-6.8). □

Theorem I-7.3. *Für alle natürlichen Zahlen* $n \geq 2$ *und alle reellen Zahlen* $a \geq 0$ *hat die Gleichung* $\xi^n = a$ *genau eine nichtnegative Lösung.*

Beweis. Die Eindeutigkeit folgt aus Satz I-4.17- 2. Für $a = 0$ setzen wir $\xi = 0$, für $a = 1$ ist die Lösung $\xi = 1$. Wir betrachten zuerst den Fall $a < 1$. Mit vollständiger Induktion zeigt man, daß $a^n < a$ für alle natürlichen Zahlen $n \geq 2$ gilt. Für $0 < x \leq a$ folgt daher $x^n \leq a^n < a$. Somit muß die Lösung der Gleichung

$\xi^n = a$ im Intervall $(a, 1)$ liegen. Wir konstruieren nun mit vollständiger Induktion eine Intervallschachtelung $I_k = [a_k, b_k]$ mit folgenden Eigenschaften

$$(*) \qquad \begin{aligned} &a_k^n \le a \le b_k^n, \\ &b_k - a_k = 2^{-k}(b_0 - a_0), \quad k \in \mathbb{N}_0. \end{aligned}$$

Für $k = 0$ setzen wir $I_0 = [a, 1]$. Dann ist $(*)$ erfüllt. Es sei I_k mit der Eigenschaft $(*)$ bereits konstruiert. Für $m = \frac{1}{2}(a_k + b_k)$ setzen wir

$$I_{k+1} = \begin{cases} [a_k, m], & \text{falls } m^n \ge a, \\ [m, b_k], & \text{falls } m^n < a. \end{cases}$$

Da wegen der Induktionsvoraussetzung $a_k \le a \le b_k$ gilt, folgt entweder $a_k^n \le a \le m^n$ oder $m^n \le a \le b_k^n$, also $a \in [a_{k+1}^n, b_{k+1}^n]$. Es gilt auch $b_{k+1} - a_{k+1} = 2^{-k-1}(b_0 - a_0)$, da I_{k+1} durch Halbierung des Intervalles I_k entsteht. Somit gilt $I_{k+1} \subset I_k$ und I_{k+1} besitzt die Eigenschaft $(*)$. Wegen Lemma I-7.2 sind die Voraussetzungen von Satz I-6.12 erfüllt.
Wir zeigen nun, daß auch die Intervalle $J_k = [a_k^n, b_k^n]$, $k \in \mathbb{N}_0$, eine Intervallschachtelung bilden. Aus $I_{k+1} \subset I_k$ folgt mit Satz I-4.17- 2 $J_{k+1} \subset J_k$. Wegen Lemma I-7.1 gilt

$$b_k^n - a_k^n < n(b_k - a_k).$$

Da die Intervalle I_k eine Intervallschachtelung bilden, gibt es für jedes $\varepsilon > 0$ einen Index ℓ mit $b_\ell - a_\ell < \frac{\varepsilon}{n}$. Somit folgt

$$b_\ell^n - a_\ell^n < \varepsilon.$$

Nach Satz I-6.12 gibt es daher eindeutig bestimmte Zahlen α, $\xi \in \mathbb{R}$ mit $\{\alpha\} = \cap_{k\in\mathbb{N}_0} J_k$ und $\{\xi\} = \cap_{k\in\mathbb{N}_0} I_k$. Dann gilt aber auch $\xi^n \in \cap_{k\in\mathbb{N}_0} J_k$ und somit $\xi^n = \alpha$. Wegen $(*)$ muß auch $a \in \cap_{k\in\mathbb{N}_0} J_k$ und somit $\xi^n = a$ gelten.
Den Fall $a > 1$ führen wir mittels $b = \frac{1}{a}$ auf den Fall $a < 1$ zurück. Es sei $\beta > 0$ die eindeutige Lösung der Gleichung $\beta^n = b$ und $\xi = \frac{1}{\beta}$. Dann gilt

$$\xi^n = (\frac{1}{\beta})^n = \frac{1}{\beta^n} = \frac{1}{b} = a.$$

□

Korollar I-7.4. Ist $n \in \mathbb{N}$ ungerade, dann besitzt die Gleichung $x^n = a$ für jedes $a \in \mathbb{R}$ genau eine Lösung $\xi \in \mathbb{R}$.

Beweis. Es genügt, den Fall $a < 0$ zu betrachten. Dann ist $-a > 0$ und die Gleichung $x^n = -a$ hat nach Satz I-7.3 genau eine positive Lösung $\tilde{\xi}$. Setzt man $\xi = -\tilde{\xi}$, folgt

$$\xi^n = (-\tilde{\xi})^n = (-1)^n \tilde{\xi}^n = (-1)^n(-a) = -(-a) = a,$$

da n ungerade ist. □

Definition I-7.5 (Wurzel). *Es seien entweder $n \in \mathbb{N}$ gerade und $a \in \mathbb{R}^+$ oder $n \in \mathbb{N}$ ungerade und $a \in \mathbb{R}$. Dann bezeichnet man mit $\sqrt[n]{a}$ die eindeutig bestimmte Lösung in $\mathbb{R}^+$ (falls $n \in \mathbb{N}$ gerade), bzw. in $\mathbb{R}$ (falls $n \in \mathbb{N}$ ungerade) der Gleichung $x^n = a$. Man nennt $\sqrt[n]{a}$ die n-te Wurzel aus a. Anstelle von $\sqrt[n]{a}$ schreibt man auch $a^{\frac{1}{n}}$. Im Fall $n = 2$ ist die Notation $\sqrt{a}$ für $\sqrt[2]{a}$ gebräuchlich.*

Bemerkung I-7.6. 1) Wegen Satz I-7.3 gilt mit dieser Definition $(a^{\frac{1}{n}})^n = a$ für alle $a \in \mathbb{R}$ im Fall n ungerade und für alle $a \in \mathbb{R}^+$ im Fall n gerade.
2) Ist $n \in \mathbb{N}$ gerade und $a > 0$, dann liefert $a^{\frac{1}{n}}$ *eine* positive Lösung der Gleichung $x^n = a$. Ist n gerade und $a > 0$, dann ist $-\sqrt[n]{a}$ eine weitere Lösung der Gleichung $x^n = a$. Dies wird durch die Kurzschreibweise $x_{1,2} = \pm\sqrt[n]{a}$ zum Ausdruck gebracht. Dies darf keinesfalls dahingehend missverstanden werden, $\sqrt[n]{a}$ habe zweierlei Vorzeichen. Es gilt stets $\sqrt[n]{x} \geq 0$ für alle $x \geq 0$ und $n \in \mathbb{N}$.
3)Ein häufig auftretender Fehler ist es, $\sqrt{x^2} = x$ zu setzen. Richtig ist $\sqrt{x^2} = |x|$.
4) Die Definition der n-ten Wurzel ist nicht einheitlich: viele Autoren definieren $\sqrt[n]{a}$ nur für $a \geq 0$ für *alle* $n \in \mathbb{N}$.

Korollar I-7.7. 1.) Für alle positiven reellen Zahlen a, b und $n, m \in \mathbb{N}$ gilt

i) $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$, $(\frac{a}{b})^{\frac{1}{n}} = \frac{a^{\frac{1}{n}}}{b^{\frac{1}{n}}}$,
ii) $(a^{\frac{1}{n}})^m = (a^m)^{\frac{1}{n}}$,
iii) $(a^{\frac{1}{n}})^{\frac{1}{m}} = (a^{\frac{1}{m}})^{\frac{1}{n}} = a^{\frac{1}{nm}}$.

2.) Ist a, $b > 0$, dann gilt

$$a < b \Leftrightarrow a^{\frac{1}{n}} < b^{\frac{1}{n}}$$

für alle $n \in \mathbb{N}$. Für n ungerade gilt diese Beziehung für alle a, $b \in \mathbb{R}$.

Beweis. i) $(a^{\frac{1}{n}} b^{\frac{1}{n}})^n = (a^{\frac{1}{n}})^n (b^{\frac{1}{n}})^n = ab$. Wegen der Eindeutigkeit der Wurzel folgt daher $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$.
ii) $[(a^{\frac{1}{n}})^m]^n = (a^{\frac{1}{n}})^{mn} = [(a^{\frac{1}{n}})^n]^m = a^m$.
iii) $[(a^{\frac{1}{n}})^{\frac{1}{m}}]^{nm} = [[(a^{\frac{1}{n}})^{\frac{1}{m}}]^m]^n = (a^{\frac{1}{n}})^n = a$.
Dies zeigt $(a^{\frac{1}{n}})^{\frac{1}{m}} = a^{\frac{1}{nm}}$. Wegen der Symmetrie bezüglich n und m in $a^{\frac{1}{nm}}$ folgt $(a^{\frac{1}{n}})^{\frac{1}{m}} = (a^{\frac{1}{m}})^{\frac{1}{n}}$.
2.) Für $0 < a < b$ folgt die Behauptung aus

$$a^{\frac{1}{n}} < b^{\frac{1}{n}} \overset{I-4.17}{\Leftrightarrow} (a^{\frac{1}{n}})^n < (b^{\frac{1}{n}})^n \overset{I-7.3}{\Leftrightarrow} a < b.$$

Es sei n ungerade und $a < b < 0$. Dies ist gleichwertig mit $0 < -b < -a$. Aus dem eben Bewiesenen folgt

$$0 < -b < -a \Leftrightarrow (-b)^{\frac{1}{n}} < (-a)^{\frac{1}{n}} \Leftrightarrow 0 < -b^{\frac{1}{n}} < -a^{\frac{1}{n}} \Leftrightarrow a^{\frac{1}{n}} < b^{\frac{1}{n}} < 0.$$

Die übrigen Fälle sind trivial. □

Die Darstellung rationaler Zahlen durch Brüche ist nicht eindeutig. Gilt etwa $p = \frac{r}{s} = \frac{u}{v} > 0$ d.h. $rv = us$ für $r, s, u, v \in \mathbb{N}$ und setzt man für $a > 0$ $x = (a^r)^{\frac{1}{s}}$, $y = (a^u)^{\frac{1}{v}}$, folgt

$$x^{sv} = (x^s)^v = (a^r)^v = a^{rv} = a^{us} = (a^u)^s = (y^v)^s = y^{vs},$$

wegen der Eindeutigkeit der Wurzel gilt dann $x = y$. Beachtet man noch Korollar I-7.7-1 ergibt sich $(a^{\frac{1}{s}})^r = (a^r)^{\frac{1}{s}} = (a^u)^{\frac{1}{v}} = (a^{\frac{1}{v}})^u$. Es kommt somit auf die Reihenfolge des Potenzierens und Radizierens nicht an.
Somit ist folgende Definition sinnvoll:

Definition I-7.8 (Potenz mit rationalen Exponenten). *Es sei $a > 0$ und $r = \frac{p}{q} > 0$, $p, q \in \mathbb{N}$.*

$$0^r := 0, \quad a^r := (a^p)^{\frac{1}{q}}, \quad a^{-r} := \frac{1}{a^r}.$$

Natürlich hätte man wegen Korollar I-7.7. 1-(ii) auch $(a^{\frac{1}{q}})^p$ zur Definition von a^r verwenden können. Man beachte, daß diese Definition für $a < 0$ nicht immer sinnvoll ist. Als Beispiel betrachte man

$$\sqrt[6]{(-27)^2} = 3, \quad \text{aber } \sqrt[3]{-27} = -3.$$

Aus diesem Grunde wird häufig die n-te Wurzel nur für nicht negative reelle Zahlen definiert.

Theorem I-7.9 (Potenzgesetze). *Es seien $a, b \in \mathbb{R}^+$ und $r, s \in \mathbb{Q}$. Folgende Regeln gelten, falls alle beteiligten Potenzen definiert sind.*

i) $a^r a^s = a^{r+s}$,
ii) $(a^r)^s = a^{rs}$,
iii) $a^r b^r = (ab)^r$,
iv) $a > 0 \Rightarrow \frac{a^r}{a^s} = a^{r-s}$,
v) $b > 0 \Rightarrow \frac{a^r}{b^r} = (\frac{a}{b})^r$.

Beweis. Wir beweisen nur die erste Regel. Wir betrachten vorerst den Fall r, $s > 0$, also $r = \frac{p}{q}$, $s = \frac{m}{n}$ mit geeigneten natürlichen Zahlen p, q, n und m. Wegen $r = \frac{np}{nq}$ und $s = \frac{mq}{qn}$ kann man diesen Fall auf Satz I-4.17 zurückführen:

$$a^r a^s = (a^{\frac{1}{nq}})^{np}(a^{\frac{1}{nq}})^{mq} = (a^{\frac{1}{nq}})^{np+mq} = a^{\frac{np+mq}{nq}} = a^{\frac{p}{q}+\frac{m}{n}} = a^{r+s}.$$

Es sei nun $rs < 0$ und ohne Beschränkung der Allgemeinheit $r < 0$, $s > 0$, also $r = -\frac{p}{q}$. Es folgt

$$\begin{aligned} a^r a^s &= \frac{a^s}{a^{|r|}} = \frac{(a^{\frac{1}{nq}})^{mq}}{(a^{\frac{1}{nq}})^{np}} = (a^{\frac{1}{nq}})^{mq-np} \\ &= \begin{cases} a^{\frac{mq-pn}{nq}} = a^{\frac{m}{n}-\frac{p}{q}} = a^{s+r} & mq - pn \geq 0, \\ \frac{1}{(a^{\frac{1}{nq}})^{pn-mq}} = \frac{1}{a^{\frac{p}{q}-\frac{m}{n}}} = \frac{1}{a^{-(r+s)}} = a^{r+s} & mq - pn < 0. \end{cases} \end{aligned}$$

Schließlich betrachten wir noch den Fall $rs > 0$ und $r < 0$, und $s < 0$. Man erhält

$$a^r a^s = \frac{1}{a^{|r|}} \frac{1}{a^{|s|}} = \frac{1}{a^{|r|+|s|}} = a^{-|r|-|s|} = a^{r+s}.$$

□

Theorem I-7.10. *Es sei $a > 0, b > 0$ und $r \in \mathbb{Q}$. Dann gilt*

i) $a < b \Leftrightarrow a^r < b^r$, *falls $r > 0$,*
ii) $a < b \Leftrightarrow b^r < a^r$, *falls $r < 0$.*

Beweis. i) folgt aus I-7.7 iv) und Satz I-4.17 2.). Für den Nachweis von ii) beachte man $a^r = \frac{1}{a^{|r|}}$. □

Theorem I-7.11. *Es sei $a \in \mathbb{R}, a > 0$, $r, s \in \mathbb{Q}$ und $r < s$. Dann gilt*

i) $a^r < a^s \Leftrightarrow a > 1$,
ii) $a^r > a^s \Leftrightarrow a < 1$.

Beweis. i) Wegen $s - r > 0$ und $1^{s-r} = 1$ folgt

$$a > 1 \overset{I-7.10}{\Leftrightarrow} a^{s-r} > 1 \Leftrightarrow \frac{a^s}{a^r} > 1 \Leftrightarrow a^s > a^r.$$

ii) $a < 1 \Leftrightarrow \frac{1}{a} > 1$. □

8. Komplexe Zahlen

Aus Korollar I-3.4 folgt, daß die Gleichung $x^2 + 1 = 0$ keine reelle Lösung besitzt. Will man erreichen, daß diese und viele andere in $\mathbb{R}$ nicht lösbare Gleichungen lösbar werden, ist eine Erweiterung des reellen Zahlensystems erforderlich.

Definition I-8.1.

i) *z heißt **komplexe Zahl*** $\underset{Def}{\Leftrightarrow}$ $z = (a, b) \in \mathbb{R} \times \mathbb{R}$.
Die Menge der komplexen Zahlen bezeichnet man mit $\mathbb{C}$.

ii) *Auf* $\mathbb{C} \times \mathbb{C}$ *erklärt man Addition* $+$ *und Multiplikation* $\cdot$ *: Es sei* $z = (a, b)$, $w = (c, d)$*:*

$$\begin{aligned} z + w &:= (a + c, b + d) \\ z \cdot w = zw &:= (ac - bd, ad + bc). \end{aligned}$$

$\mathbb{C}$ und $\mathbb{R} \times \mathbb{R}$ sind also als Mengen identisch, charakteristisch für $\mathbb{C}$ ist die durch Addition und Multiplikation aufgeprägte algebraische Struktur.

Theorem I-8.2. *Ersetzt man* $\mathbb{R}$ *durch* $\mathbb{C}$*,* 0 *durch* $(0, 0)$*,* 1 *durch* $(1, 0)$*, dann sind die Axiome* $\mathcal{A}, \mathcal{M}$ *und* $\mathcal{D}$ *erfüllt. Das Tripel* $(\mathbb{C}, +, \cdot)$ *ist somit ein Körper.*

Beweis. Für $z = (a, b) \in \mathbb{C}$ ist die additive Inverse

$$-z = (-a, -b),$$

für $z = (a, b) \neq (0, 0)$ ist die multiplikative Inverse

$$z^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

Exemplarisch zeigen wir die Assoziativität der Multiplikation: Es sei $x = (a, b)$, $y = (c, d)$ und $z = (e, f)$. Dann ist

$$\begin{aligned} x(yz) &= (a, b)(ce - df, cf + de) \\ &= (a(ce - df) - b(cf + de), b(ce - df) + a(cf + de)) \\ &= (ace - adf - bcf - bde, bce - bdf + acf + ade) \\ &= (ac - bd, ad + bc)(e, f) \\ &= (xy)z. \end{aligned}$$

□

Die Sätze aus Abschnitt 2 sowie jene Resultate, welche nicht auf die Ordnung in $\mathbb{R}$ Bezug nehmen, gelten daher auch in $\mathbb{C}$. Für die Paare $(a, 0), (b, 0) \in \mathbb{C}$ gilt

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

(für die letzte Beziehung setzen wir natürlich $a \neq 0$ voraus). Die Menge $\{(a,0)\colon a \in \mathbb{R}\} \subset \mathbb{C}$ bildet also einen Unterkörper von $\mathbb{C}$, der dieselben arithmetischen Eigenschaften wie $\mathbb{R}$ besitzt. Man kann daher $\mathbb{R}$ und $\{(a,0)\colon a \in \mathbb{R}\}$ identifizieren, indem man jeder reellen Zahl $a \in \mathbb{R}$ das geordnete Paar $(a,0)$ zuordnet und umgekehrt. Es ist somit sinnvoll, die komplexen Zahlen $(a,0)$ reell zu nennen und man schreibt anstelle von $(a,0)$ kurz a.

Definition I-8.3.

$$i := (0,1) \textit{ heißt } \boldsymbol{imaginäre\ Einheit}.$$

Wegen

$$(0,1)(0,1) = (-1,0) = -(1,0)$$

folgt

$$i^2 = -1,$$

somit ist die komplexe Zahl i Lösung von $x^2 + 1 = 0$. Für $b \in \mathbb{R}$ folgt

$$ib = (0,1)(b,0) = (0,b).$$

Komplexe Zahlen dieser Form heißen **imaginär**. Wegen a, $b \in \mathbb{R}$ und

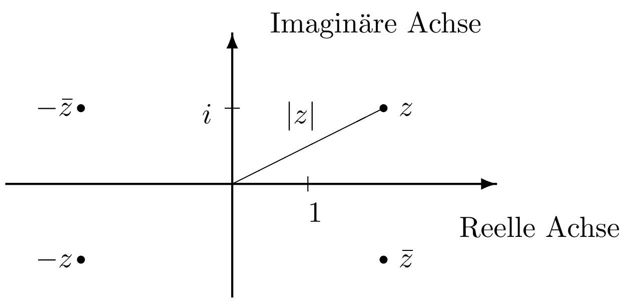
$$(a,b) = (a,0) + (0,b) = a + ib.$$

kann jede komplexe Zahl (a,b) in der bequemeren Form $a + ib$ geschrieben werden.

Definition I-8.4. *Es sei $z = a + ib \in \mathbb{C}$.*

i) $\operatorname{Re} z := a$ *heißt **Realteil** von z,*
$\operatorname{Im} z := b$ *heißt **Imaginärteil** von z.*
ii) $\bar{z} := a - ib$ *heißt die zu z **konjungiert komplexe** Zahl,*
iii) $|z| := \sqrt{a^2 + b^2}$ *heißt **Absolutbetrag** von z.*
iv) $A \subset \mathbb{C}$ *heißt beschränkt* $\underset{Def}{\Leftrightarrow}$ $\{|z|\colon z \in A\}$ *ist beschränkt in $\mathbb{R}$.*

Wegen $i^2 = -1 < 0$ ist es nicht möglich, auf $\mathbb{C}$ eine Ordnung zu definieren, die den Axiomen $\mathcal{O}$ genügt. Es müßte dann nämlich $x^2 \geq 0$ für alle $x \in \mathbb{C}$ gelten (vgl. Korollar I-3.4). Die Ergebnisse aus Abschnitt 3 lassen sich daher *nicht* auf $\mathbb{C}$ übertragen: *Ungleichungen zwischen komplexen Zahlen sind sinnlos!*
Die komplexen Zahlen lassen sich als Menge geordneter Paare in der **Gaußschen Zahlenebene** veranschaulichen: Nach Wahl eines cartesischen Koordinatensystems wird die komplexe Zahl $z = x + iy$ durch den Punkt (x,y) dargestellt. Die reellen Zahlen werden mit den Punkten der reellen Achse identifiziert. $|z|$ ist der Abstand des Punktes (x,y) von 0. Die Definition des Absolutbetrages für komplexe Zahlen ist die natürliche Verallgemeinerung des Absolutbetrages reeller Zahlen als deren Abstand von dem mit Null bezeichneten Punkt auf der reellen Zahlengeraden.



Gaußsche Zahlenebene

Theorem I-8.5. *Für $z, w \in \mathbb{C}$ gilt*

a)	$\overline{z+w} = \bar{z} + \bar{w}$,	*d)*	$\operatorname{Re} z = \frac{1}{2}(z+\bar{z})$, $\operatorname{Im} z = \frac{1}{2i}(z-\bar{z})$,
b)	$\overline{zw} = \bar{z} \cdot \bar{w}$,	*e)*	$z = \bar{z} \Leftrightarrow z \in \mathbb{R}$,
c)	$z\bar{z} = \lvert z\rvert^2$,	*f)*	$\overline{(\frac{1}{z})} = \frac{1}{\bar{z}}$ *falls* $z \neq 0$.

Theorem I-8.6. *Für $z, w \in \mathbb{C}$ gilt*

a)	$\lvert z\rvert = 0 \Leftrightarrow z = 0$,	*d)*	$\lvert z\rvert = \lvert\bar{z}\rvert$,
b)	$\lvert zw\rvert = \lvert z\rvert\lvert w\rvert$,	*e)*	$\lvert\operatorname{Re} z\rvert \leq \lvert z\rvert$, $\lvert\operatorname{Im} z\rvert \leq \lvert z\rvert$,
c)	$\lvert z+w\rvert \leq \lvert z\rvert + \lvert w\rvert$,	*f)*	$\big\lvert\lvert z\rvert - \lvert w\rvert\big\rvert \leq \lvert z-w\rvert$

Die Ungleichung (c) nennt man Dreiecksungleichung.

Beweis. Für den Beweis von c) beachte man $\overline{z\bar{w}} = \bar{z}w$. Somit folgt

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

□

Theorem I-8.7 (Ungleichung von Cauchy-Schwarz-Bunjakowski).
Es seien $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ komplexe Zahlen. Dann gilt

$$\Big|\sum_{j=1}^{n} a_j \bar{b_j}\Big|^2 \leq \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

Gleichheit gilt genau dann, wenn es ein $z \in \mathbb{C}$ gibt, sodaß $a_j = b_j z$, $j = 1, \ldots, n$, oder $b_j = a_j z$, $j = 1, \ldots, n$, gilt.

Beweis. Abkürzend schreiben wir $\alpha = \sum_{j=1}^{n} |a_j|^2$, $\beta = \sum_{j=1}^{n} |b_j|^2$ und $\gamma = \sum_{j=1}^{n} a_j \bar{b_j}$. O.B.d.A. können wir $\beta > 0$ annehmen (die Ungleichung ist für $\beta = 0$ trivial). Es folgt

$$\begin{aligned}
0 \quad &\leq \quad \sum_{j=1}^{n} |\beta a_j - \gamma b_j|^2 = \sum_{j=1}^{n} (\beta a_j - \gamma b_j)(\beta \bar{a}_j - \bar{\gamma}\bar{b_j}) \\
&= \quad \sum_{j=1}^{n} (\beta^2 |a_j|^2 - \beta\bar{\gamma} a_j \bar{b_j} - \beta\gamma b_j \bar{a}_j + |\gamma|^2 |b_j|^2) \\
&= \quad \beta^2 \sum_{j=1}^{n} |a_j|^2 - \beta\bar{\gamma} \sum_{j=1}^{n} a_j \bar{b_j} - \beta\gamma \sum_{j=1}^{n} b_j \bar{a}_j + |\gamma|^2 \sum_{j=1}^{n} |b_j|^2 \\
&= \quad \beta^2 \alpha - 2\beta|\gamma|^2 + \beta|\gamma|^2 = \beta^2\alpha - \beta|\gamma|^2 = \beta(\alpha\beta - |\gamma|^2).
\end{aligned}$$

Wegen Satz I-3.3 und $\beta > 0$ folgt $|\gamma|^2 \leq \alpha\beta$. Dies ist die gewünschte Ungleichung. Gleichheit tritt genau dann ein, wenn $\beta a_j = \gamma b_j$ für alle $j = 1, \ldots, n$ gilt. Wegen $\beta > 0$ ist dies gleichbedeutend mit $a_j = \frac{\gamma}{\beta} b_j$, $j = 1, \ldots, n$. □

Mit Hilfe der Ungleichung von Cauchy läßt sich die Dreicksungleichung I-8.6-c) erheblich verallgemeinern.

Theorem I-8.8 (Minkowski Ungleichung). *Es seien $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ komplexe Zahlen. Dann gilt*

$$(*) \qquad \left(\sum_{j=1}^{n} |a_j + b_j|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{j=1}^{n} |a_j|^2 \right)^{\frac{1}{2}} + \left(\sum_{j=1}^{n} |b_j|^2 \right)^{\frac{1}{2}}.$$

Gleichheit gilt genau dann, wenn es ein $z \in \mathbb{R}^+$ gibt, sodaß $a_j = b_j z$, $j = 1, \ldots, n$ oder $b_j = a_j z$, $j = 1, \ldots, n$ gilt.

Beweis. Es seien α, β und γ definiert wie im Beweis von I-8.7 und es sei $\beta > 0$. Es folgt

$$\begin{aligned}
\sum_{j=1}^{n} |a_j + b_j|^2 &= \alpha + \gamma + \bar{\gamma} + \beta = \alpha + 2\operatorname{Re}\gamma + \beta \\
&\leq \alpha + 2|\gamma| + \beta \\
&\leq \alpha + 2|\alpha|^{\frac{1}{2}}|\beta|^{\frac{1}{2}} + \beta = (\alpha^{\frac{1}{2}} + \beta^{\frac{1}{2}})^2.
\end{aligned}$$

Angenommen es gilt in $(*)$ Gleichheit. Dann muß auch

$$|\gamma|^2 = \alpha\beta \quad \text{und} \quad |\gamma| = \operatorname{Re}\gamma$$

gelten. Aus $|\gamma|^2 = \alpha\beta$ folgt wegen $\beta > 0$ und I-8.7 die Existenz von $z \in \mathbb{C}$ mit $a_j = b_j z$, $j = 1, \ldots, n$. Somit gilt $\gamma = z\beta$, wegen $|\gamma| = \operatorname{Re}\gamma$ ergibt sich schließlich $\operatorname{Im}\gamma = 0$, somit $z \in \mathbb{R}$ und $z \geq 0$. Umgekehrt zeigt man, daß die angegebenen Bedingungen hinreichend sind für die Gleichheit in $(*)$. □

KAPITEL II

Folgen und Reihen

In diesem Kapitel bezeichnet $\mathbb{K}$ den Körper der reellen oder komplexen Zahlen.

1. Normierte Räume

DEFINITION II-1.1. *Es sei X ein Vektorraum über $\mathbb{K}$ und $\|\cdot\|: X \to \mathbb{R}^+$. Das Paar $(X, \|\cdot\|)$ heißt* ***normierter Raum*** *und $\|\cdot\|$* ***Norm****, wenn die Abbildung $\|\cdot\|$ folgende Eigenschaften besitzt:*

($\mathcal{N}$1) $\forall x \in X: \|x\| \geq 0$
$\forall x \in X: \|x\| = 0 \Leftrightarrow x = 0$ — DEFINITHEIT
($\mathcal{N}$2) $\forall x \in X\, \forall \lambda \in \mathbb{K}: \|\lambda x\| = |\lambda| \|x\|$ — HOMOGENITÄT
($\mathcal{N}$3) $\forall x, y \in X: \|x + y\| \leq \|x\| + \|y\|$ — DREIECKSUNGLEICHUNG

Häufig schreibt man X anstelle von $(X, \|\cdot\|)$, beispielsweise wenn die spezielle Norm ohne Bedeutung ist oder aus dem Zusammenhang klar ersichtlich ist.

LEMMA II-1.2. *Es sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gilt für alle $x, y \in X$*

$$\big| ||x|| - ||y|| \big| \leq ||x - y||. \tag{3}$$

BEISPIEL II-1.3. 1) Es sei $X = \mathbb{K}^n$. Für $x \in \mathbb{K}^n$, $x = (\xi_1, \dots, \xi_n)$, definieren wir

$$\|x\|_1 = \sum_{i=1}^n |\xi_i|,$$
$$\|x\|_2 = (\sum_{i=1}^n |\xi_i|^2)^{1/2},$$
$$\|x\|_\infty = \max\{|\xi_i| : i = 1, \dots, n\}.$$

Die Räume $(\mathbb{R}^n, \|\cdot\|_p)$, $p \in \{1, 2, \infty\}$, sind normierte Räume. Die Dreiecksungleichung folgt für $p = 2$ aus der Minkowski-Ungleichung I-8.8, für $p = 1$ und $p = \infty$ ergibt sie sich leicht aus der Gültigkeit der Dreiecksungleichung in $\mathbb{K}$.

DEFINITION II-1.4. (1) *Eine Funktion $f: D \to Y$ heißt* ***beschränkt*** $\underset{Def}{\Leftrightarrow}$ $\exists M > 0\, \forall x \in D: \|f(x)\| \leq M$.

(2) *Wir setzen*

$$\mathcal{B}(D, Y) = \{f : D \to Y,\ f \textit{ ist beschränkt}\}.$$

Beispiel II-1.5. Man überzeuge sich davon, daß $\mathcal{B}(D,Y)$ ein Vektorraum ist. Durch

$$\|f\|_\infty = \sup_{x\in D} \|f(x)\|$$

wird eine Norm auf $\mathcal{B}(D,Y)$ erklärt. Für den Nachweis der Dreiecksungleichung beachte man für $f,\ g \in \mathcal{B}(D,Y)$ und $x \in D$

$$\|(f+g)(x)\| = \|f(x)+g(x)\| \le \|f(x)\| + \|g(x)\| \le \|f\|_\infty + \|g\|_\infty.$$

Somit ist $\|f\|_\infty + \|g\|_\infty$ eine obere Schranke für $\{\|(f+g)(x)\| \colon x \in D\}$. Daraus folgt $f+g \in \mathcal{B}(D,Y)$ und $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$. $(\mathcal{B}(D,Y), \|\cdot\|_\infty)$ ist somit ein normierter Raum.

Man beachte, daß $\|x-y\|_2$ für $x,\ y \in \mathbb{R}^3$ mit dem euklidischen Abstand des Punktes x vom Punkt y übereinstimmt. Es ist auch für andere Normen sinnvoll, $\|x-y\|$ als Abstand von x zu y zu interpretieren. Der Begriff des Kreises in $\mathbb{R}^2$ bzw. einer Kugel in $\mathbb{R}^3$ läßt sich daher in naheliegender Weise verallgemeinern:

Definition II-1.6. *Es sei $(X, ||\cdot||)$ ein normierter Raum, $x_0 \in X$ und $r > 0$.*

i) $K(x_0,r) := \{x \in X \colon ||x-x_0|| < r\}$,
*heißt **offene Kugel** mit Radius r und Mittelpunkt x_0.*
$\bar{K}(x_0,r) := \{x \in X \colon ||x-x_0|| \le r\}$,
*heißt **abgeschlossene Kugel** mit Radius r und Mittelpunkt x_0.*

ii) *$O \subset X$ heißt **offen** in $(X, ||\cdot||)$* $\underset{Def}{\Leftrightarrow} \forall x \in O \exists r > 0 \colon K(x,r) \subset O$.

iii) *$A \subset X$ heißt **abgeschlossen** in $(X, ||\cdot||)$* $\underset{Def}{\Leftrightarrow}$ $\complement A$ *ist offen.*

iv) *$U \subset X$ heißt **Umgebung** von x_0* $\underset{Def}{\Leftrightarrow} \exists r > 0 \colon K(x_0,r) \subset U$.

$\mathcal{U}(x) = \{U \subset X \colon U$ *ist eine Umgebung von* $x\}$.
*Man nennt die Kugeln $K(x,\varepsilon)$ auch **ε-Umgebung** von x.*

Die spezielle Notation für abgeschlossene Kugeln wird später gerechtfertigt.
Eine Teilmenge $O \subset X$ ist demnach offen, wenn es zu jedem Punkt von O eine Umgebung gibt, die in O enthalten ist. Der Begriff offen hängt von dem X zugrunde gelegten Abstandsbegriff (Norm) ab. Wir zeigen nun, daß die Bezeichnung offene Kugel für die Menge $K(x,r)$ konsistent ist mit der Definition einer offenen Menge:

Theorem II-1.7. *Jede offene (abgeschlossene) Kugel eines normierten Raumes ist offen (abgeschlossen).*

Beweis. Es sei $(X, ||\cdot||)$ ein normierter Raum, $x_0 \in X$ und $K(x_0,r) \subset X$. Für jedes beliebige $\xi \in K(x_0,r)$ setzen wir $s = r - ||x_0 - \xi|| > 0$ und betrachten die Kugel $K(\xi,s)$. Für $\eta \in K(\xi,s)$ folgt mit Hilfe der Dreiecksungleichung

$$||\eta - x_0|| \le ||\xi - x_0|| + ||\xi - \eta|| < ||\xi - x_0|| + s = r,$$

also

$$K(\xi,s) \subset K(x_0,r),$$

d.h. $K(x_0,r)$ ist offen.

Abbildung 1. Einheitskugeln in $\mathbb{R}^2$

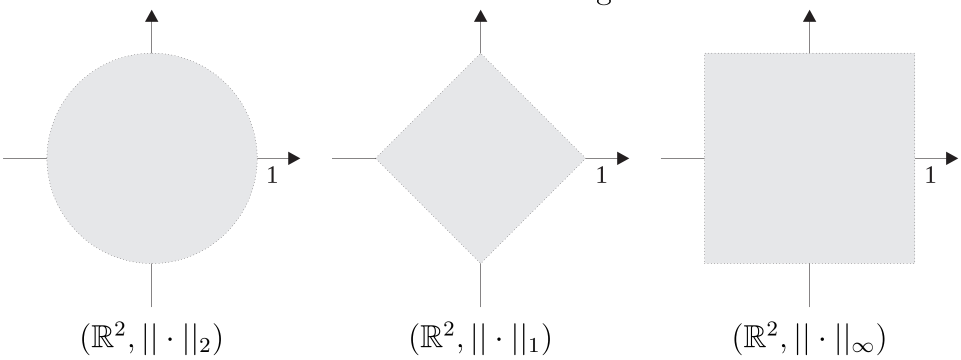
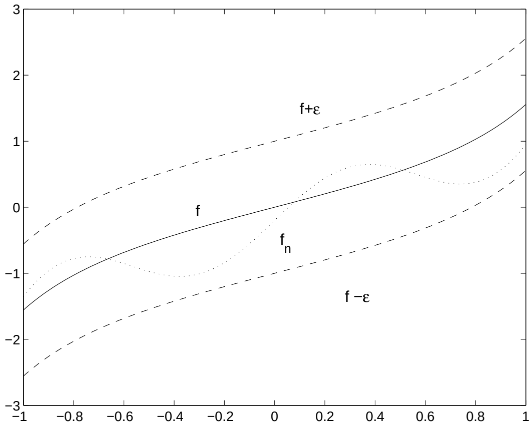


Abbildung 2. ε-Schlauch $K(f, \varepsilon)$



Der Nachweis für abgeschlossene Kugeln sei dem Leser als Übung überlassen. $\square$

Beispiel II-1.8. (1) In $\mathbb{R}$, versehen mit dem üblichen, euklidischen Abstand, sind die offenen Kugeln gerade die offenen Intervalle $(x_0 - r, x_0 + r)$.

(2) $(\mathbb{R}^2, ||\cdot||_2)\colon K(0,1) = \{(\xi,\eta) \in \mathbb{R}^2\colon \sqrt{|\xi|^2 + |\eta|^2} < 1\}$
$(\mathbb{R}^2, ||\cdot||_1)\colon K(0,1) = \{(\xi,\eta) \in \mathbb{R}^2\colon |\xi| + |\eta| < 1\}$
$(\mathbb{R}^2, ||\cdot||_\infty)\colon K(0,1) = \{(\xi,\eta) \in \mathbb{R}^2\colon \max\{|\xi|, |\eta|\} < 1\}$

(3) Es sei I ein nichtleeres Intervall.
$(\mathcal{B}(I,\mathbb{R}), ||\cdot||_\infty)\colon K(f,\varepsilon) = \{g \in \mathcal{B}(I,\mathbb{R})\colon \sup_{x\in I} |f(x) - g(x)| < \varepsilon\}$
$K(f,\varepsilon)$ beschreibt einen "ε-Schlauch" von beschränkten Funktionen um f mit Radius ε.

Offene Mengen in einem normierten Raum besitzen folgende Eigenschaften.

Theorem II-1.9. *Es sei* $(X, ||\cdot||)$ *ein normierter Raum und* $\mathcal{T}$ *die Menge aller offenen Teilmengen von* X*. Dann gilt*

($\mathcal{T}1$) *$\emptyset \in \mathcal{T}$, $X \in \mathcal{T}$.*

($\mathcal{T}2$) *Es sei I eine beliebige endliche Indexmenge:*

$$(\forall i \in I \colon O_i \in \mathcal{T}) \Rightarrow \bigcap_{i \in I} O_i \in \mathcal{T}.$$

($\mathcal{T}3$) *Es sei I eine beliebige Indexmenge:*

$$(\forall i \in I \colon O_i \in \mathcal{T}) \Rightarrow \bigcup_{i \in I} O_i \in \mathcal{T}.$$

Beweis. 1) Die Bedingung aus Definition II-1.6-ii) kann für $O = \emptyset$ nie verletzt werden, für $O = X$ ist sie trivialerweise erfüllt.
2) Es sei $\{O_1, \dots, O_n\} \subset \mathcal{T}$. Im Falle $\bigcap_{i=1}^n O_i = \emptyset$ ist nichts zu beweisen. Sei $x_0 \in \bigcap_{i=1}^n O_i$. Da jede Menge O_i, $i = 1, \dots, n$, offen ist, gibt es Kugeln $K(x_0, r_i) \subset O_i$, $r_i > 0$, $i = 1, \dots, n$. Setzt man $r = \min\{r_1, \dots, r_n\}$, ergibt sich

$$K(x_0, r) \subset K(x_0, r_i) \subset O_i, \quad i = 1, \dots, n,$$

also

$$K(x_0, r) \subset \bigcap_{i=1}^n O_i,$$

d.h. $\bigcap_{i=1}^n O_i$ ist offen.
3) Es sei I eine beliebige nichtleere Indexmenge, $O_i \in \mathcal{T}$ für alle $i \in I$ und $x_0 \in \bigcup_{i \in I} O_i$ ($\neq \emptyset$). Es gibt also einen Index $i_0 \in I$ mit $x_0 \in O_{i_0}$. Folglich existiert eine Kugel $K(x_0, r_0)$ mit

$$K(x_0, r_0) \subset O_{i_0} \subset \bigcup_{i \in I} O_i.$$

□

Theorem II-1.10 (Eigenschaften der abgeschlossenen Mengen). *Es sei $(X, ||\cdot||)$ ein normierter Raum. Dann gilt*

i) *$\emptyset$, X sind abgeschlossen.*
ii) *Für jede endliche Indexmenge I:*
($\forall i \in I \colon A_i$ ist abgeschlossen) $\Rightarrow \bigcup_{i \in I} A_i$ ist abgeschlossen.
iii) *Für jede beliebige Indexmenge I:*
($\forall i \in I \colon A_i$ ist abgeschlossen) $\Rightarrow \bigcap_{i \in I} A_i$ ist abgeschlossen.

Beweis. Satz II-1.9 und de Morgansche Regeln. □

Man nennt jede Familie $\mathcal{T}$ von Teilmengen einer nichtleeren Grundmenge X, welche die Eigenschaften ($\mathcal{T}1$)-($\mathcal{T}3$) erfüllt, eine **Topologie**, auf X. Die Elemente von $\mathcal{T}$ heißen **offene Mengen**, das Paar $(X, \mathcal{T})$ bildet einen **topologischen Raum**. Satz II-1.9 zeigt, daß das System der offenen Mengen eines normierten Raumes $(X, ||\cdot||)$ eine Topologie $\mathcal{T}_{||\cdot||}$ auf X, die metrische Topologie auf X, bildet. Die Definition II-1.6 der offenen Menge in einem normierten Raum ist somit konsistent mit der topologischen Definition. Sinngemäß läßt sich der Umgebungsbegriff in einem topologischen

Raum einführen. Topologische Räume stellen das theoretische Werkzeug dar, um die unterschiedlichen Konvergenz- und Stetigkeitsbegriffe der Analysis unter einem einheitlichen Gesichtspunkt zu betrachten. In dieser Einführung beschränken wir uns auf die Darstellung der topologischen Grundlagen im Rahmen der normierten Räume.

2. Konvergenz von Folgen

DEFINITION II-2.1. *Es sei X ein normierter Raum.*
i) Eine Abbildung $x\colon \mathbb{N} \to X$, heißt ***Folge****. $x_n := x(n)$ wird* **n-tes Glied** *der Folge genannt.*
ii) Eine Abbildung $\varphi\colon \mathbb{N} \to \mathbb{N}$ ist **streng monoton wachsend** $\underset{Def}{\Leftrightarrow}$ $\forall n, m \in \mathbb{N}\colon n < m \Rightarrow \varphi(n) < \varphi(m)$.
iii) Es sei $x\colon \mathbb{N} \to X$ eine Folge und $\varphi\colon \mathbb{N} \to \mathbb{N}$ streng monoton wachsend. Dann heißt $x \circ \varphi$ ***Teilfolge*** *von x.*

Es ist üblich, Folgen in der Form $(x_n)_{n\geq 1}$, $(x_n)_{n=1}^{\infty}$, (x_n) oder $x_1, x_2, \dots,$ zu schreiben. Teilfolgen $(x_{\varphi(k)})_{k\geq 1}$ bezeichnet man oft auch mit $(x_{n_k})_{k\geq 1}$ (es wird dabei $\varphi(k)$ durch n_k ersetzt). Es ist also $n_1 < n_2 < \cdots$. Als Übung beweise man, daß $\varphi(n) \geq n$ für alle $n \in \mathbb{N}$ gilt, falls $\varphi\colon \mathbb{N} \to \mathbb{N}$ streng monoton wächst. Die Schreibweise $(x_n)_{n\geq 1} \subset X$ bedeutet $x_n \in X$ für alle $n \in \mathbb{N}$. Die Indizierung einer Folge kann natürlich bei einem beliebigen $n_0 \in \mathbb{N}_0$ beginnen. Ist X ein Funktionenraum, heißt $(f_n)_{n\geq 1}$ **Funktionenfolge**.

BEISPIEL II-2.2.

(1)	$x_n = \frac{1}{n}$	:	$1, \frac{1}{2}, \frac{1}{3}, \dots$
(2)	$y_n = i^n$	:	$i, -1, -i, 1, i, -1, \dots$
(3)	$z_n = \frac{n}{n+1}$	:	$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \dots$
(4)	$w_n = n^2$	:	$1, 4, 9, 16, \dots,$
(5)	$f_n(x) = \frac{1}{n}x^n$	:	$x, \frac{1}{2}x^2, \frac{1}{3}x^3, \dots,$ für $x \in [0,1]$.

Diese Folgen weisen ein recht unterschiedliches Verhalten auf: Die Glieder x_n werden immer kleiner, z_n nähert sich immer mehr der Zahl 1, y_n nimmt abwechselnd die Werte $i, -1, -i, 1$ an und w_n wächst über alle Grenzen, f_n kommt der Nullfunktion auf $[0,1]$ immer näher. Wir präzisieren nun die vage Formulierung, eine Folge nähere sich immer mehr einem bestimmten Element $x \in X$.

DEFINITION II-2.3. *Es sei $(x_n)_{n\geq 1} \subset X$ und $x \in X$.*
1) $(x_n)_{n\geq 1}$ ***konvergiert*** *gegen x* $\underset{Def}{\Leftrightarrow}$
$\forall \varepsilon > 0\, \exists N(\varepsilon) \in \mathbb{N}\, \forall n \in \mathbb{N}\colon n \geq N(\varepsilon) \Rightarrow \|x_n - x\| < \varepsilon$.
In diesem Falle heißt x ***Grenzwert*** *von $(x_n)_{n\geq 1}$ und man schreibt*

$$\lim_{n\to\infty} x_n = x \qquad \textit{oder auch} \qquad x_n \to x \qquad \textit{für } n \to \infty.$$

2) $(x_n)_{n\geq 1}$ heißt ***konvergent*** $\underset{Def}{\Leftrightarrow}$ $\exists x \in X\colon \lim_{n\to\infty} x_n = x$.
3) $(x_n)_{n\geq 1}$ heißt ***divergent*** $\underset{Def}{\Leftrightarrow}$ *$(x_n)_{n\geq 1}$ ist nicht konvergent.*

4) $x \in X$ heißt ***Häufungswert*** *der Folge $(x_n)_{n\geq 1}$, wenn es eine Teilfolge $(x_{\varphi(n)})_{n\geq 1} \subset (x_n)_{n\geq 1}$ gibt, welche gegen x konvergiert.*

$\lim_{n\to\infty} x_n = x$ bedeutet demnach, daß in jeder beliebigen ε–Umgebung von x, *sei sie auch noch so klein*, alle Folgenglieder mit Ausnahme von höchstens endlich vielen liegen müssen. Außerhalb einer beliebigen ε–Umgebung können daher höchstens endlich viele Glieder der Folge liegen. Die Konvergenz einer Folge und ihr Grenzwert werden also nicht beeinflußt, wenn man endlich viele Glieder der Folge abändert. Durch formale Negation erhält man: „$(x_n)_{n\geq 1}$ konvergiert nicht gegen x" ist gleichwertig mit:

$$\exists \varepsilon_0 > 0\, \forall N \in \mathbb{N}\, \exists n \in \mathbb{N}\colon n \geq N \wedge \|x_n - x\| \geq \varepsilon_0.$$

Bemerkung II-2.4. $(x_n)_{n\geq 1}$ konvergiert nicht gegen x $\Leftrightarrow$ es gibt eine Teilfolge $(x_{\varphi(k)})_{k\geq 1}$ und $\varepsilon_0 > 0$ so, daß $\forall k \in \mathbb{N}\colon \|x_{\varphi(k)} - x\| \geq \varepsilon_0$.

Bemerkung II-2.5. Die Konvergenz einer Funktionenfolge (f_n) in $\mathcal{B}(D, X)$ gegen $f\colon D \to X$ ist charakterisiert durch

$$\forall \varepsilon > 0 \exists N(\varepsilon) \in \mathbb{N} \forall n \in \mathbb{N}\colon n \geq N(\varepsilon) \Rightarrow \|f_n - f\|_\infty = \sup\{\|f_n(s) - f(s)\|\colon s \in D\} < \varepsilon.$$

Eine einfache Überlegung zeigt, daß dies gleichwertig ist mit

$$(*) \qquad \forall \varepsilon > 0 \exists N(\varepsilon) \in \mathbb{N} \forall n \in \mathbb{N} \forall s \in D\colon n \geq N(\varepsilon) \Rightarrow \|f_n(s) - f(s)\| < \varepsilon.$$

Es wird also gefordert, daß zu jedem $\varepsilon > 0$ ein Index $N(\varepsilon)$ gefunden werden kann mit $\|f_n(s) - f(s)\| < \varepsilon$, $s \in D$ und $n \geq N(\varepsilon)$ Wesentlich ist dabei, daß der Index $N(\varepsilon)$ uniform für alle $s \in D$ gewählt werden kann. Man nennt daher diese Art der Konvergenz einer Funktionenfolge **gleichmäßige Konvergenz**.

Beispiel II-2.6. 1.) Wir zeigen: $\lim_{n\to\infty} \frac{1}{n} = 0$. Zu beliebigen $\varepsilon > 0$ gibt es nach Korollar I-6.9 eine natürliche Zahl $N(\varepsilon)$ mit $0 < \frac{1}{N(\varepsilon)} < \varepsilon$. Für $n \geq N(\varepsilon)$ gilt dann erst recht $0 < \frac{1}{n} \leq \frac{1}{N(\varepsilon)} < \varepsilon$, also $\lim_{n\to\infty} \frac{1}{n} = 0$.
2.) Für die Folge $y_n = i^n$ gilt $y_{4n} = 1$, $y_{4n+1} = i$, $y_{4n+2} = -1$ und $y_{4n+3} = -i$, $n \in \mathbb{N}_0$. Die Folge (y_n) besitzt somit die Häufungswerte 1, i, -1 und $-i$.
3.) Wir zeigen $\lim_{n\to\infty} f_n = 0$ für $f_n(x) = \frac{1}{n} x^n$, $x \in [0, 1]$. (Man achte auf die unterschiedliche Bedeutung des Symbols 0). Der Grenzübergang bezieht sich auf den normierten Raum $X = \mathcal{B}([0, 1], \mathbb{R}^n)$. Wegen $(*)$ müssen also Folgendes zeigen:

$$\forall \varepsilon > 0 \exists N(\varepsilon) \in \mathbb{N} \forall n \in \mathbb{N} \forall x \in [0, 1]\colon n \geq N(\varepsilon) \Rightarrow |f_n(x) - 0| < \varepsilon.$$

Für $x \in [0, 1]$ gilt $0 \leq x^n \leq 1$. Wählen wir daher zu $\varepsilon > 0$ den Index $N(\varepsilon)$ wie im vorangehenden Beispiel, folgt

$$|f_n(x)| = \frac{1}{n} x^n \leq \frac{1}{n} \leq \frac{1}{N(\varepsilon)} < \varepsilon.$$

Da diese Abschätzung für alle $x \in [0, 1]$ gilt und der Index $N(\varepsilon)$ NICHT von x abhängt, ist $\lim_{n\to\infty} f_n = 0$ gezeigt.

Theorem II-2.7 (Eindeutigkeit des Grenzwertes). *Eine Folge hat höchstens einen Grenzwert.*

Beweis. Angenommen eine Folge $(x_n) \subset X$ konvergiere gegen x und $y \in X$. Nach Definition II-2.3 gibt es zu jedem $\varepsilon > 0$ Indices N_1 und N_2, sodaß

$$n \geq N_1 \Rightarrow \|x_n - x\| < \varepsilon \quad \text{sowie} \quad n \geq N_2 \Rightarrow \|x_n - y\| < \varepsilon$$

gilt. Setzt man $N_0 = \max(N_1, N_2)$ gilt für $n \geq N_0$

$$\|x - y\| \leq \|x - x_n\| + \|x_n - y\| < 2\varepsilon,$$

wegen Korollar II-I-3.8 folgt $x = y$. □

Wir zeigen nun zwei nützliche notwendige Konvergenzbedingungen.

Theorem II-2.8. *Es sei* $(x_n)_{n\geq 1} \subset X$ *konvergent mit* $\lim_{n\to\infty} x_n = x$. *Dann gilt:*

(1) $(x_n)_{n\geq 1}$ *ist beschränkt.*
(2) *Jede Teilfolge* $(x_{\varphi(n)})_{n\geq 1}$ *ist konvergent mit* $\lim_{n\to\infty} x_{\varphi(n)} = x$.

Beweis. 1) Wir wählen $\varepsilon = 1$ in Definition II-2.3. Wegen $\lim_{n\to\infty} x_n = x$ gibt es $N_0 \in \mathbb{N}$, sodaß $\|x_n - x\| < 1$ für alle $n \geq N_0$ zutrifft. Für $n \geq N_0$ gilt dann auch

$$\|x_n\| \leq \|x_n - x\| + \|x\| < 1 + \|x\|.$$

Setzt man $\alpha = \max\{\|x_n\| \colon 1 \leq n < N_0\}$ ergibt sich die Behauptung

$$\forall n \in \mathbb{N} \colon \|x_n\| \leq \max\{1 + \|x\|, \alpha\}.$$

2) Es sei $(x_{\varphi(n)})$ eine Teilfolge von (x_n). Für ein beliebig gewähltes $\varepsilon > 0$ gibt es einen Index $N(\varepsilon)$, sodaß $\|x_n - x\| < \varepsilon$ für alle $n \geq N(\varepsilon)$ zutrifft. Für $n \geq N(\varepsilon)$ gilt daher auch $\varphi(n) \geq \varphi(N(\varepsilon)) \geq N(\varepsilon)$ und somit $\|x_{\varphi(n)} - x\| < \varepsilon$, d.h. $\lim_{n\to\infty} x_{\varphi(n)} = x$. □

Satz II-2.8 zeigt, daß die Folgen (y_n) und (w_n) in Beispiel II-2.2 nicht konvergieren.

Theorem II-2.9. *i) Es sei* $(x_n)_{n\geq 1} \subset X$ *eine* **Nullfolge**, *d.h.* $\lim_{n\to\infty} x_n = 0$, *und* $(\lambda_n)_{n\geq 1} \subset \mathbb{K}$ *beschränkt. Dann gilt* $\lim_{n\to\infty} \lambda_n x_n = 0$.
ii) Es sei $(x_n)_{n\geq 1} \subset X$ *konvergent und* $\lim_{n\to\infty} x_n = x$. *Dann gilt* $\lim_{n\to\infty} \|x_n\| = \|x\|$.

Beweis. Es sei $\varepsilon > 0$ und $M > 0$ so, daß $|\lambda_n| \leq M$ für alle $n \in \mathbb{N}$ gilt. Da (x_n) eine Nullfolge ist, gibt es einen Index $N(\varepsilon)$ so, daß $\|x_n\| < \frac{\varepsilon}{M}$ für alle $n \geq N(\varepsilon)$ gilt. Für beliebiges $n \geq N(\varepsilon)$ folgt dann

$$\|\lambda_n x_n\| = |\lambda_n|\,\|x_n\| \leq M\|x_n\| < \varepsilon.$$

ii) Dies folgt aus Lemma II-1.2. □

Theorem II-2.10. *Es seien* (x_n) *und* (y_n) *konvergente Folgen in einem normierten Raum und* $\lambda \in \mathbb{K}$. *Dann konvergieren auch die Folgen* (λx_n) *und* $(x_n + y_n)$ *und es gilt*

$$\lim_{n\to\infty} \lambda x_n = \lambda \lim_{n\to\infty} x_n,$$
$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n.$$

Beweis. Wir zeigen nur die Konvergenz der Summenfolge. Es sei $\lim x_n = x$ und $\lim y_n = y$. Zu beliebig gewähltem $\varepsilon > 0$ gibt es Indices $N_1(\varepsilon)$ und $N_2(\varepsilon)$ derart, daß

$$\forall n \in \mathbb{N}\colon n \geq N_1(\varepsilon) \Rightarrow \|x_n - x\| < \frac{\varepsilon}{2},$$
$$\forall n \in \mathbb{N}\colon n \geq N_2(\varepsilon) \Rightarrow \|y_n - y\| < \frac{\varepsilon}{2}.$$

Für $n \geq \max\{N_1(\varepsilon), N_2(\varepsilon)\}$ folgt dann

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Auf die Konvergenz der Summenfolge kann nur geschlossen werden, wenn die Konvergenz der einzelnen Summanden gesichert ist. Wählt man beispielsweise $x_n = (-1)^n$, $y_n = -x_n$ und $z_n = x_n$, dann ist die Folge $(x_n + y_n)$ konvergent, die Folge $(x_n + z_n)$ hingegen divergent.
Die Konvergenz einer Folge hängt natürlich auch von der jeweiligen Norm ab. Das Beispiel II-1.3 zeigt, daß es viele Möglichkeiten gibt, einen Vektorraum zu normieren. Welchen Einfluß hat dies auf die Konvergenz einer Folge?

Definition II-2.11. *Es sei X ein Vektorraum. Zwei Normen $\|\cdot\|$ und $|||\cdot|||$ auf X heißen* ***äquivalent****, wenn es Konstante $0 < m \leq M$ gibt, so daß*

$$m\|x\| \leq |||x||| \leq M\|x\|$$

für alle $x \in X$ gilt.

Die Äquivalenz von Normen ist eine Äquivalenzrelation in der Menge aller Normen auf X.

Beispiel II-2.12. Die Normen $\|\cdot\|_p$, $p \in \{1, 2, \infty\}$ in $\mathbb{C}^n$ sind äquivalent. Wir zeigen zuerst die Äquivalenz von $\|\cdot\|_1$ und $\|\cdot\|_2$. Aus der Ungleichung von Cauchy-Schwarz-Bunjakowski II-I-8.7 folgt

$$\|x\|_1 \leq \sqrt{n}\|x\|_2.$$

Eine einfache Induktion zeigt die Gültigkeit der Ungleichung

$$(\sum_{i=1}^{n} |x_i|)^2 \geq \sum_{i=1}^{n} |x_i|^2$$

für alle $n \in \mathbb{N}$. Dies ist gleichwertig mit

$$\|x\|_1 \geq \|x\|_2.$$

Als Übung zeige der Leser die Ungleichungen

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

Somit sind die Normen $\|x\|_1$ und $\|x\|_\infty$ und daher auch $\|x\|_2$ und $\|x\|_\infty$ äquivalent.

Dieses Beispiel ist ein Spezialfall eines wesentlich stärkeren Resultates:

Theorem II-2.13. *Es sei X ein endlich dimensionaler Vektorraum. Dann sind alle Normen auf X äquivalent.*

Beweis. Der Beweis dieses Satzes wird nachgetragen, sobald die erforderlichen Hilfsmittel zur Verfügung stehen. $\square$

Die Äquivalenz von Normen ist aus folgendem Grund nützlich:

Theorem II-2.14. *Es seien $\|\cdot\|$ und $|||\cdot|||$ äquivalente Normen auf einem Vektorraum X. Eine Folge $(x_n) \subset X$ konvergiert genau bezüglich $\|\cdot\|$, wenn sie bezüglich $|||\cdot|||$ konvergiert.*

Beweis. Angenommen es gilt $\lim_{n\to\infty} x_n = x$ bezüglich der Norm $\|\cdot\|$ und es sei $\varepsilon > 0$. Dann gibt es einen Index $N(\varepsilon)$, sodaß für alle $n \geq N(\varepsilon)$

$$\|x_n - x\| \leq \frac{\varepsilon}{M}.$$

($M > 0$ bezeichnet die Konstante aus der Normenäquivalenz). Wegen der Äquivalenz der beiden Normen folgt dann für $n \geq N(\varepsilon)$

$$|||x_n - x||| \leq M\|x_n - x\| \leq M\frac{\varepsilon}{M} = \varepsilon.$$

Somit konvergiert die Folge (x_n) auch bezüglich der Norm $|||\cdot|||$ (gegen denselben Grenzwert). Für die Umkehrung vertausche man die Rollen der beiden Normen. $\square$

Für Konvergenzuntersuchungen kann man daher eine möglichst bequeme äquivalente Norm wählen. Wir machen davon im Beweis des folgenden Satzes Gebrauch:

Theorem II-2.15. *Eine Folge in $\mathbb{K}^n$ ist genau dann konvergent, wenn jede Komponentenfolge konvergiert:*

$$\lim_{k\to\infty} x_k = x \qquad \Leftrightarrow \qquad \lim_{k\to\infty} x_i^k = x_i, \quad i = 1, \ldots, n,$$

($x_k = (x_1^k, \ldots, x_n^k)$, $x = (x_1, \ldots, x_n)$).

Beweis. Da in $\mathbb{K}^n$ sämtliche Normen äquivalent sind, können wir $\|x\| = \|\cdot\|_\infty$ wählen. Gilt $\lim_{k\to\infty} x_k = x$, gibt es für jedes $\varepsilon > 0$ einen Index $N(\varepsilon)$ mit

$$\|x_k - x\|_\infty = \max\{|x_i^k - x_i| \colon 1 \leq i \leq n\} \leq \varepsilon$$

für $n \geq N(\varepsilon)$. Dann ist aber auch

$$|x_i^k - x_i| \leq \varepsilon, \qquad 1 \leq i \leq n$$

für $n \geq N(\varepsilon)$. Dies bedeutet

$$(*) \qquad \lim_{k\to\infty} x_i^k = x_i, \quad i = 1, \ldots, n$$

Gilt umgekehrt $(*)$, gibt es für jedes $\varepsilon > 0$ Indices $N_i(\varepsilon)$, $1 \leq i \leq n$, mit

$$|x_i^k - x_i| \leq \varepsilon, \qquad 1 \leq i \leq n$$

für $n \geq N_i(\varepsilon)$. Setzt man $N(\varepsilon) = \max\{N_i(\varepsilon)\colon 1 \leq i \leq n\}$ folgt für $n \geq N(\varepsilon)$

$$\max\{|x_i^k - x_i|\colon 1 \leq i \leq n\} = \|x_k - x\|_\infty \leq \varepsilon,$$

also $\lim_{k\to\infty} x_k = x$. □

Wir beschließen diesen Abschnitt mit der Charakterisierung der Konvergenz einer Folge komplexer Zahlen.

THEOREM II-2.16. *Es sei $(z_n)_{n\geq 1} \subset \mathbb{C}$, $z_n = x_n + iy_n$, $x_n, y_n \in \mathbb{R}$.*
(z_n) ist genau dann konvergent, wenn die reellen Folgen (x_n) und (y_n) konvergent sind. In diesem Falle gilt

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} x_n + i \lim_{n\to\infty} y_n.$$

BEWEIS. Die Behauptung folgt aus den Abschätzungen

$$\begin{aligned}\max(|x - x_n|, |y - y_n|) &\leq |z - z_n| \\ &= |(x - x_n) + i(y - y_n)| \leq |x - x_n| + |y - y_n|\end{aligned}$$

□

3. Nützliche Grenzwerte

In diesem Abschnitt berechnen wir einige Grenzwerte, auf die im Folgenden immer wieder zurückgegriffen wird.

BEISPIEL II-3.1.

i) $\lim_{n\to\infty} \frac{1}{n} = 0$,
ii) $\forall p \in \mathbb{Q}\colon p > 0 \Rightarrow lim_{n\to\infty} n^{-p} = 0$,
iii) $\forall a \in \mathbb{R}^+ \setminus \{0\}\colon lim_{n\to\infty} \sqrt[n]{a} = 1$,
iv) $lim_{n\to\infty} \sqrt[n]{n} = 1$,
v) $\forall k \in \mathbb{N}_0 \forall z \in \mathbb{C}\colon |z| > 1 \Rightarrow lim_{n\to\infty} \frac{n^k}{z^n} = 0$

BEWEIS. i) Siehe Korollar II-I-6.9
ii) Zu $\varepsilon > 0$ wählen wir $N(\varepsilon) = \lfloor \varepsilon^{-\frac{1}{p}} \rfloor + 1$ ($\lfloor x \rfloor$ wurde in Satz II-I-6.10 definiert). Für $n > N(\varepsilon)$ folgt dann mit Satz II-I-7.10 $\frac{1}{n^p} < \frac{1}{N(\varepsilon)^p} < \varepsilon$.
iii) Die Behauptung ist trivial für $a = 1$. Es sei daher zunächst $a > 1$ und somit auch $\sqrt[n]{a} > 1$ (Korollar II-I-7.7). Setzt man $x_n = \sqrt[n]{a} - 1 > 0$, d.h. $a = (1 + x_n)^n$, folgt mit Hilfe der Bernoullischen Ungleichnung II.I-4.18 $a = (1 + x_n)^n > 1 + nx_n$, somit $0 < x_n < \frac{a-1}{n}$. Für beliebiges $\varepsilon > 0$ setzen wir $N(\varepsilon) = \lfloor \frac{a-1}{\varepsilon} \rfloor + 1$. Somit ergibt sich für $n \geq N(\varepsilon)$

$$|\sqrt[n]{a} - 1| = x_n \leq \frac{a-1}{N(\varepsilon)} < \varepsilon.$$

Der Beweis für $a \in (0, 1)$ wird vorerst zurückgestellt.
iv) Wir setzen wieder $x_n = \sqrt[n]{n} - 1 \geq 0$, d.h. $n = (1 + x_n)^n$. Mit Hilfe des binomischen

Lehrsatzes II-I-4.22 erhält man für $n \geq 2$

$$n = (1 + x_n)^n \geq 1 + \binom{n}{2} x_n^2 = 1 + \frac{1}{2} n(n-1) x_n^2$$

und folglich

$$x_n \leq \sqrt{\frac{2}{n}}.$$

Für $\varepsilon > 0$ wählen wir $N(\varepsilon) > \frac{2}{\varepsilon^2}$. Dann folgt für alle $n \geq N(\varepsilon)$

$$|\sqrt[n]{n} - 1| = x_n \leq \sqrt{\frac{2}{n}} \leq \sqrt{\frac{2}{N(\varepsilon)}} < \varepsilon.$$

v) Es sei $|z| = 1 + x$, $x > 0$, und $p \in \mathbb{N}$, $p > k$. Für jedes $n > 2p$ folgt aus der Binomialentwicklung

$$(1 + x)^n > \binom{n}{p} x^p = \frac{n(n-1) \cdots (n-p+1)}{p!} x^p > \frac{n^p x^p}{2^p p!}$$

(wir benutzten $n - j > n - p > \frac{n}{2}, 0 \leq j \leq p - 1$, und $p < \frac{n}{2}$). Somit folgt für $n > 2p$ (beachte $p - k > 0$)

$$\left| \frac{n^k}{z^n} \right| = \frac{n^k}{(1+x)^n} < \frac{2^p p!}{x^p} \frac{1}{n^{p-k}} < \frac{2^p p!}{x^p} \frac{1}{n}.$$

Folglich ist $|\frac{n^k}{z^n}| < \varepsilon$, soferne $n \geq N(\varepsilon) > \max\{2p, \frac{2^p p!}{x^p \varepsilon}\}$. □

Bemerkung II-3.2. Das letzte Beispiel bringt zum Ausdruck, daß für $|z| > 1, z \in \mathbb{C}$, die Folge $(|z|^n)_{n \geq 1}$ schneller anwächst, als jede noch so große Potenz von n.

Theorem II-3.3. *Es sei $z \in \mathbb{C}$. Dann gilt*

(1) $\lim_{n\to\infty} z^n = 0$ *für* $|z| < 1$.
(2) $(z^n)_{n\geq 1}$ *ist divergent, falls* $|z| \geq 1$ *und* $z \neq 1$.

Beweis. (1) Die Behauptung ist trivial für $z = 0$. Es sei $0 < |z| < 1$, also $\frac{1}{|z|} > 1$. Ersetzt man z in Beispiel Beispiel II-3.1–(v) durch $\frac{1}{z}$ und wählt $k = 0$, ergibt sich die Behauptung.

(2) Es sei nun $|z| \geq 1$ und $z \neq 1$. Dann ist auch $|z|^n \geq 1$ für alle $n \in \mathbb{N}$. Wir nehmen an, $(z^n)_{n\geq 1}$ sei konvergent, d.h. es existiert $\alpha \in \mathbb{C}$ mit $\lim_{n\to\infty} z^n = \alpha$. Somit gibt es zu $\varepsilon = \frac{1}{3}|z - 1| > 0$ einen Index $N(\varepsilon)$, sodaß $|z^n - \alpha| < \varepsilon$ für alle $n \geq N(\varepsilon)$. Dies hat

$$\begin{aligned} 3\varepsilon = |z - 1| &\leq |z^{N(\varepsilon)}||z - 1| = |z^{N(\varepsilon)+1} - z^{N(\varepsilon)}| \\ &\leq |z^{N(\varepsilon)+1} - \alpha| + |\alpha - z^{N(\varepsilon)}| < \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

zur Folge. Dieser Widerspruch zeigt die Divergenz von (z^n). □

4. Rechenregeln für konvergente Folgen

Die Konvergenzuntersuchung komplizierter Folgen läßt sich oft wesentlich vereinfachen, indem man sie auf die Untersuchung der Konvergenz einfacherer Folgen zurückführt.

Theorem II-4.1. *Es seien $(z_n)_{n\geq 1}$, $(w_n)_{n\geq 1}$ konvergente Folgen in $\mathbb{C}$. Dann gilt*

i) $\lim(z_n + w_n) = \lim z_n + \lim w_n$,
ii) $\forall a \in \mathbb{C}\colon \lim a z_n = a \lim z_n$,
iii) $\lim z_n w_n = \lim z_n \lim w_n$,
iv) *Gilt* $\lim w_n \neq 0$ *und* $w_n \neq 0$ *für alle* $n \in \mathbb{N}$, *dann ist* $\lim \frac{z_n}{w_n} = \frac{\lim z_n}{\lim w_n}$.

Beweis. i) Vergleiche Satz II-2.10.
iii) Es sei $\lim z_n = z$ und $\lim w_n = w$. Nach Satz II-2.8 gibt es $b > 0$, sodaß $|z_n| \leq b$ für alle $n \in \mathbb{N}$ zutrifft. Zu $\varepsilon > 0$ seien $N_1(\varepsilon)$ und $N_2(\varepsilon)$ so bestimmt, daß

$$\forall n \in \mathbb{N}\colon n \geq N_1(\varepsilon) \Rightarrow |z_n - z| < \frac{\varepsilon}{2(|w|+1)},$$
$$\forall n \in \mathbb{N}\colon n \geq N_2(\varepsilon) \Rightarrow |w_n - w| < \frac{\varepsilon}{2b}.$$

Für $n \geq \max\{N_1(\varepsilon), N_2(\varepsilon)\}$ folgt

$$\begin{aligned} |z_n w_n - zw| &\leq |z_n w_n - z_n w| + |z_n w - zw| \\ &\leq |z_n||w_n - w| + |w||z_n - z| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

ii) Folgt aus (iii) mit $w_n = a$ für alle $n \in \mathbb{N}$.
iv) Wegen iii) genügt es $\lim \frac{1}{w_n} = \frac{1}{w}$ zu zeigen. Vorerst bestimmen wir einen Index $\tilde{N}$, sodaß $|w_n - w| < \frac{1}{2}|w|$ für alle $n \geq \tilde{N}$ gilt. Für $n \geq \tilde{N}$ folgt somit

$$|w_n| \geq |w| - |w_n - w| > \frac{1}{2}|w|.$$

Es sei nun $\varepsilon > 0$ beliebig gewählt und $N_1(\varepsilon)$ so bestimmt, daß $|w_n - w| < \frac{\varepsilon}{2}|w|^2$ für alle $n \geq N_1(\varepsilon)$ gilt. Für $n \geq \max\{N_1(\varepsilon), \tilde{N}\}$ ergibt sich

$$|\frac{1}{w_n} - \frac{1}{w}| = \frac{|w - w_n|}{|w||w_n|} < 2\frac{|w - w_n|}{|w|^2} < \varepsilon.$$

$\square$

Es soll noch einmal betont werden, daß diese Regeln nur auf konvergente Folgen angewendet werden dürfen. Der Schluß von der Konvergenz der Summenfolge oder Produktfolge auf die Konvergenz der Summanden bzw. Faktoren ist im Allgemeinen falsch. Es sei etwa $z_n = (-1)^n$ und $w_n = (-1)^{n+1}$. Die Folgen (z_n) und (w_n) sind divergent, trotzdem existieren die Grenzwerte

$$\lim_{n\to\infty}(z_n + w_n) = \lim_{n\to\infty}((-1)^n + (-1)^{n+1}) = 0,$$
$$\lim_{n\to\infty} z_n w_n = \lim_{n\to\infty} \frac{z_n}{w_n} = \lim_{n\to\infty}(-1)^n(-1)^{n+1} = -1.$$

Wir demonstrieren die Anwendung von Satz II-4.1 an folgendem Beispiel:

Beispiel II-4.2.

$$\lim_{n\to\infty} \frac{2n^2+n+3}{n^2+1} = \lim \frac{2+\frac{1}{n}+\frac{3}{n^2}}{1+\frac{1}{n^2}} \overset{(iv)}{=} \frac{\lim(2+\frac{1}{n}+\frac{3}{n^2})}{\lim(1+\frac{1}{n^2})}$$
$$\overset{(i)}{=} \frac{\lim 2 + \lim\frac{1}{n} + \lim\frac{3}{n^2}}{\lim 1 + \lim\frac{1}{n^2}} \overset{(ii)}{=} \frac{2+\lim\frac{1}{n}+3\lim\frac{1}{n^2}}{1+\lim\frac{1}{n^2}} = \frac{2+0+0}{1+0} = 2$$

Die vorletzte Gleichheit stützt sich auf Beispiel II-3.1–(i), –(ii). Die zu Beginn getroffene Annahme, die Folge sei konvergent, wird durch die angedeuteten Argumente gerechtfertigt.

Wir können nun auch den Beweis von Beispiel II-3.1–(iii) zu Ende führen.

Beispiel II-4.3. $\forall a \in \mathbb{R}^+ \setminus \{0\}\colon\ \lim_{n\to\infty} \sqrt[n]{a} = 1$

Beweis. In Beispiel II-3.1 wurde bereits der Fall $a \geq 1$ erledigt. Es sei nun also $0 < a < 1$. Dann ist $\frac{1}{a} > 1$ und wegen Beispiel II-3.1 $\lim_{n\to\infty} \sqrt[n]{\frac{1}{a}} = 1$. Die Behauptung folgt nun aus der Identität $\sqrt[n]{a} = \frac{1}{\sqrt[n]{\frac{1}{a}}}$ und Satz II-4.1. □

Das nächste Resultat stellt sicher, daß sich Ungleichungen zwischen den Gliedern konvergenter Folgen auf deren Grenzwerte übertragen.

Theorem II-4.4. *Es seien $(x_n)_{n\geq 1}$, $(y_n)_{n\geq 1}$ konvergente Folgen reeller Zahlen und $N_0 \in \mathbb{N}$. Gilt $x_n \leq y_n$ für alle $n \geq N_0$, dann folgt $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.*

Beweis. Es sei $x = \lim_{n\to\infty} x_n$, $y = \lim_{n\to\infty} y_n$ und $y < x$. Zu $\varepsilon = \frac{1}{2}(x-y) > 0$ gibt es einen Index $N_1 > N_0$, sodaß $|x_n - x| < \varepsilon$ und $|y_n - y| < \varepsilon$ für alle $n \geq N_1$ gilt. Für jedes $n \geq N_1$ folgt somit

$$y_n < y + \varepsilon = \frac{1}{2}(x+y) = x - \varepsilon < x_n$$

im Widerspruch zu $x_n \leq y_n$ für $n \geq N_0$. □

An Hand der Folge $(\frac{1}{n})_{n\geq 1}$ macht man sich klar, daß strikte Ungleichungen beim Übergang zum Grenzwert nicht immer erhalten bleiben.

Korollar II-4.5. Die Folge $(x_n)_{n\geq 1} \subset X$ konvergiere gegen x. Ferner seien $\alpha \in \mathbb{R}^+$, $z \in X$ und $N \in \mathbb{N}$. Aus $\|x_n - z\| \leq \alpha$ für alle $n \geq N$ folgt dann die Ungleichung $\|x - z\| \leq \alpha$.

Beweis. Aus $x = \lim_{n\to\infty} x_n$ ergibt sich $\lim_{n\to\infty}(x_n - z) = x - z$ und wegen Satz II-2.9 auch $\|x - z\| = \lim_{n\to\infty} \|x_n - z\|$. Die Behauptung folgt nun unmittelbar aus Satz II-4.4. □

Der Beweis des folgenden Einschachtelungssatzes sei dem Leser als Übung überlassen.

Theorem II-4.6. *Es seien $(x_n)_{n\geq1}$, $(y_n)_{n\geq1}$ und $(z_n)_{n\geq1}$ Folgen reeller Zahlen. Sind die Folgen (x_n) und (y_n) konvergent mit $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$ und gilt $x_n \leq z_n \leq y_n$ für alle $n \in \mathbb{N}$, dann ist auch (z_n) konvergent mit $\lim_{n\to\infty} z_n = \lim_{n\to\infty} x_n$.*

5. Konvergenzkriterien

Bisher können wir Konvergenz einer Folge nur untersuchen, wenn wir den Grenzwert der Folge bereits kennen. Meistens ist aber der Grenzwert nicht bekannt und es ergibt sich die Frage, wie man aus Eigenschaften der Folgenglieder auf Konvergenz der Folge schließen kann.

Definition II-5.1. *Eine reelle Folge $(x_n)_{n\geq1}$ heißt*

(1) **monoton wachsend (fallend)** $\underset{Def}{\Leftrightarrow} \forall n \in \mathbb{N}\colon x_n \leq x_{n+1} \quad (x_n \geq x_{n+1})$

(2) **streng monoton wachsend (fallend)** $\underset{Def}{\Leftrightarrow} \forall n \in \mathbb{N}\colon x_n < x_{n+1} \quad (x_n > x_{n+1})$

(3) **monoton** $\underset{Def}{\Leftrightarrow}$ (x_n) *ist monoton wachsend oder monoton fallend*

Theorem II-5.2 (Monotoniekriterium). *Eine beschränkte, monoton wachsende (fallende) Folge (x_n) reeller Zahlen ist konvergent und es gilt*

$$\lim_{n\to\infty} x_n = \sup\{x_n\colon n \in \mathbb{N}\} \qquad (\lim_{n\to\infty} x_n = \inf\{x_n\colon n \in \mathbb{N}\})$$

Beweis. Es sei $(x_n)_{n\geq1}$ eine monoton wachsende, beschränkte Folge. Das Supremum Prinzip II-I-6.5 sichert die Existenz von $\xi = \sup\{x_n\colon n \geq 1\}$. Wir behaupten: $\xi = \lim_{n\to\infty} x_n$. Für beliebiges $\varepsilon > 0$ gibt es nach Satz II-I-6.4 ein Folgenglied x_N mit $\xi - \varepsilon < x_N \leq \xi$. Wegen der Monotonie folgt für alle $n > N$: $\xi - \varepsilon < x_N \leq x_n \leq \xi < \xi + \varepsilon$, d.h. $|\xi - x_n| < \varepsilon$. Der Beweis für monoton fallende Folgen ist analog. □

Wegen Satz II-2.8 konvergiert eine monotone Folge genau dann, wenn sie beschränkt ist. Ist eine Folge monoton wachsend (fallend), genügt eine obere (untere) Schranke für den Nachweis der Beschränktheit. Da die Konvergenz einer Folge und deren Grenzwert nicht durch eine endliches Anfangsstück der Folge beeinflusst werden, ist das Monotoniekriterium auch dann anwendbar, wenn die Folge erst ab einem Index $n_0 > 1$ monoton ist. Zur Illustration des Monotoniekriteriums betrachten wir folgende Beispiele:

Beispiel II-5.3. Die rekursiv definierte Folge $(x_n)_{n\geq1}$

$$x_1 = 1, \qquad x_{n+1} = \frac{x_n^2 + 2}{2x_n}$$

ist konvergent mit $\lim_{n\to\infty} x_n = \sqrt{2}$.

1) $\forall n \geq 2\colon x_n > \sqrt{2}$: Da $x_1 > 0$ ist, folgt induktiv $x_n > 0$, $n \in \mathbb{N}$. Somit ist die Division durch x_n gerechtfertigt. Mit Satz II-I-3.11–(e) ergibt sich die Abschätzung

$$x_n^2 + 2 = 2x_{n+1}x_n \leq x_{n+1}^2 + x_n^2.$$

Dies impliziert $x_{n+1} \geq \sqrt{2}$, $n \in \mathbb{N}$. Wäre $x_{n_0} = \sqrt{2}$ für ein $n_0 \in \mathbb{N}$, dann müsste $x_n = \sqrt{2}$ für alle $n \in \mathbb{N}$ gelten. Wegen $x_1 \neq \sqrt{2}$ folgt $x_n > \sqrt{2}$ für $n \in \mathbb{N}$.
2) $\forall n \geq 2\colon x_{n+1} < x_n$, d.h. die Folge $(x_n)_{n\geq 2}$ ist strikt monoton fallend.
Es ist $x_3 = \frac{17}{12} < \frac{3}{2} = x_2$. Es gelte nun $x_n < x_{n-1}$. Die Ungleichung

$$x_{n+1} = \frac{x_n^2 + 2}{2x_n} < \frac{x_{n-1}^2 + 2}{2x_{n-1}} = x_n$$

ist äquivalent zu

$$\begin{gathered} 2x_n^2 x_{n-1} + 4x_{n-1} < 2x_{n-1}^2 x_n + 4x_n \Leftrightarrow \\ 2x_n x_{n-1}(x_n - x_{n-1}) < 4(x_n - x_{n-1}) \Leftrightarrow x_n x_{n-1} > 2 \end{gathered}$$

Die Gültigkeit der letzten Ungleichung ist eine Folge von 1). Die Behauptung folgt nun mit Hilfe des Induktionsprinzips.
3) Wegen $\sqrt{2} < x_n < x_2$, $n \geq 3$, ist (x_n) beschränkt, somit nach dem Monotoniekriterum konvergent. Es sei $\alpha = \lim x_n$. Wegen 1) und Satz II-4.4 folgt $\alpha \geq \sqrt{2}$. Mit Hilfe der Regeln in Satz II-4.1 berechnet man

$$\alpha = \frac{\alpha^2 + 2}{2\alpha} \text{ d.h. } 2\alpha^2 = \alpha^2 + 2 \text{ und daher } \alpha^2 = 2.$$

Somit folgt $\alpha = \lim x_n = \sqrt{2}$.

Dieses Beispiel zeigt, daß man bei rekursiv definierten Folgen die Rekursionsvorschrift verwenden kann, um einen Kandidaten für den Grenzwert zu bestimmen. Der Nachweis der Konvergenz ist meist der schwierigere Teil der Konvergenzuntersuchung.

Beispiel II-5.4 (Eulersche Zahl).
Für $n \in \mathbb{N}$ sei $a_n = (1 + \frac{1}{n})^n$ und $b_n = (1 + \frac{1}{n})^{n+1}$. Es gilt:
(1) (a_n) ist streng monoton wachsend, d.h. $\forall n \in \mathbb{N}\colon a_n < a_{n+1}$
(2) (b_n) ist streng monoton fallend, d.h. $\forall n \in \mathbb{N}\colon b_{n+1} < b_n$
(3) $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Der gemeinsame Grenzwert der beiden Folgen heißt **Eulersche Zahl** e. Es gilt $2 < e < 4$.

Beweis. Für $n > 1$ schließen wir mit Hilfe der Bernoulli Ungleichung II-I-4.18

$$(*) \qquad \frac{a_n}{b_{n-1}} = (\frac{n+1}{n})^n (\frac{n}{n-1})^{-n} = (\frac{n^2-1}{n^2})^n = (1 - \frac{1}{n^2})^n > 1 - \frac{1}{n} = \frac{n-1}{n}$$

und

$$(\dagger) \qquad \frac{b_{n-1}}{a_n} = (\frac{n^2}{n^2-1})^n = (1 + \frac{1}{n^2-1})^n > (1 + \frac{1}{n^2})^n \overset{II-I-4.18}{>} 1 + \frac{1}{n} = \frac{n+1}{n}.$$

Aus $(*)$ folgt

$$a_n > \frac{n-1}{n}(\frac{n}{n-1})^n = (\frac{n}{n-1})^{n-1} = (1+\frac{1}{n-1})^{n-1} = a_{n-1}$$

und aus $(\dagger)$

$$b_{n-1} > \frac{n+1}{n}(\frac{n+1}{n})^n = (1+\frac{1}{n})^{n+1} = b_n.$$

Für $n > 2$ folgt demnach

$$2 = a_1 < a_{n-1} < a_n < b_n < b_{n-1} < b_1 = 4.$$

Die Konvergenz der Folgen ergibt sich aus Satz II-5.2. Aus Satz II-4.4 folgt

$$2 < \lim a_n \leq \lim b_n < 4,$$

die beiden äußeren Ungleichungen sind strikt wegen der strikten Monotonie von (a_n) und (b_n). Wegen $b_n = (1+\frac{1}{n})a_n$ erhält man schließlich $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$. □

Theorem II-5.5 (Bolzano–Weierstrass). *Jede beschränkte Folge in $\mathbb{C}$ besitzt eine konvergente Teilfolge.*

Beweis. Fall 1: $(x_n)_{n\geq 1}$ ist eine beschränkte reelle Folge.
Wegen der Beschränktheit gibt es $a_0, b_0 \in \mathbb{R}$ mit $a_0 \leq x_n \leq b_0$ für alle $n \in \mathbb{N}$. Halbiert man das Intervall $J_0 = [a_0, b_0]$, so liegen in mindestens einem der beiden Teilintervalle $[a_0, \frac{1}{2}(a_0+b_0)]$, $[\frac{1}{2}(a_0+b_0), b_0]$ unendlich viele Glieder der Folge. Wir bezeichnen dieses Intervall mit $J_1 = [a_1, b_1]$ und wenden das eben beschriebene Verfahren auf J_1 an. Man erhält somit eine Folge von abgeschlossenen Intervallen $J_k = [a_k, b_k]$, für welche $J_{k+1} \subset J_k$ gilt und deren Länge $b_k - a_k = 2^{-k}(b_0 - a_0)$ beträgt. Aus Beispiel II-3.1-(v) folgert man $\lim_{k\to\infty}(b_k - a_k) = 0$. Die Folge von Intervallen $(J_k)_{k\geq 0}$ bildet daher eine Intervallschachtelung, welche eine reelle Zahl $\alpha \in \bigcap J_k$ eindeutig bestimmt. Wegen $\alpha \in [a_k, b_k]$, $k \in N$, folgt $0 \leq |a_k - \alpha| \leq b_k - a_k$ und mit Satz II-4.6 weiter $\lim_{k\to\infty} a_k = \alpha$. Wegen $b_k = a_k + (b_k - a_k)$ folgt dann $\lim_{k\to\infty} b_k = \lim_{k\to\infty} a_k = \alpha$. Wir wählen nun einen Index $\varphi(1)$, sodaß $x_{\varphi(1)}$ in J_1 liegt. Da J_2 unendlich viele Glieder der Folge enthält, können wir $\varphi(2) > \varphi(1)$ bestimmen, sodaß $x_{\varphi(2)}$ in J_2 liegt, u.s.w. Im n–ten Schritt könnten wir zum Beispiel $\varphi(n) = \min\{k \in \mathbb{N} \colon x_k \in J_n, k > \varphi(n-1)\}$ wählen. Wir erhalten auf diese Weise eine streng monoton wachsende Folge $(\varphi(n))_{n\geq 1}$, die zugehörige Teilfolge $(x_{\varphi(n)})_{n\geq 1}$ erfüllt dann

$$a_n \leq x_{\varphi(n)} \leq b_n, \qquad n \in \mathbb{N}.$$

Die Konvergenz von $(x_{\varphi(n)})$ folgt nun aus Satz II-4.6.
Fall 2: Es sei nun $(z_n)_{n\geq 1} \subset \mathbb{C}$ und beschränkt, d.h. es gibt $b > 0$ mit $|z_n| \leq b$, $n \in \mathbb{N}$. Ferner setzen wir $z_n = x_n + iy_n$. Wegen $\max\{|x_n|, |y_n|\} \leq |z_n|$ sind auch die reellen Folgen (x_n) und (y_n) beschränkt. Somit gibt es eine konvergente Teilfolge $(x_{\varphi(n)})_{n\geq 1}$

von (x_n), ihr Grenzwert sei x. Weiters enthält $(y_{\varphi(n)})_{n\geq 1}$ eine konvergente Teilfolge $(y_{(\psi\circ\varphi)(n)})_{n\geq 1}$ mit Grenzwert y. Mit Hilfe der Sätze II-2.16 und II-2.8 folgern wir

$$\lim_{n\to\infty} z_{(\psi\circ\varphi)(n)} = \lim_{n\to\infty}\big(x_{(\psi\circ\varphi)(n)} + iy_{(\psi\circ\varphi)(n)}\big) = x + iy.$$

□

KOROLLAR II-5.6. Jede beschränkte Folge in $\mathbb{C}^n$ besitzt eine konvergente Teilfolge.

BEWEIS. BEWEIS. Es sei also $(x_k) \subset \mathbb{C}^n$ eine beschränkte Folge. Dann gibt es eine reelle Zahl $r > 0$ mit $(x_k) \subset \bar{K}_\infty(0,r)$. Schreibt man $x_k = (\xi_1^k, \dots, \xi_n^k)$, gilt $-r \leq \xi_j^k \leq r$ für $1 \leq j \leq n$. Somit kann man aus der Folge (ξ_1^k) eine konvergente Teilfolge $(\xi_1^{\varphi_1(k)}) \subset (\xi_1^k)$ auswählen (vgl. Satz II-5.5). Im nächsten Schritt wählt man aus $(\xi_2^{\varphi_1(k)}) \subset (\xi_2^k)$ eine konvergente Teilfolge $(\xi_2^{\varphi_2\circ\varphi_1(k)}) \subset (\xi_2^{\varphi_1(k)})$ aus. Man beachte, daß dann auch $(\xi_1^{\varphi_2\circ\varphi_1(k)})$ konvergiert. Setzt man $\sigma_1 = \varphi_1$ und $\sigma_j = \varphi_j \circ \sigma_{j-1}$, $2 \leq j \leq n$, erhält man induktiv fortfahrend konvergente Teilfolgen $(\xi_j^{\sigma_j(k)}) \subset ((\xi_j^{\sigma_{j-1}(k)}))$ so, daß auch alle Teilfolgen $(\xi_l^{\sigma_j(k)})$ für $1 \leq l \leq j-1$ konvergieren. Für $j = n$ ergibt sich schließlich eine Teilfolge $(x_{\sigma_n(k)}) \subset (x_k)$, für welche alle Komponentenfolgen $(\xi_j^{\sigma_n(k)})$, $1 \leq j \leq n$, konvergieren. Nach Satz II-2.15 ist die Teilfolge $(x_{\sigma_n(k)})$ konvergent. □

□

Das Monotoniekriterium ist zwar, wie wir gesehen haben, recht bequem einzusetzen. Sein Anwendungsbereich ist allerdings auf eine doch recht spezielle Klasse *reeller* Folgen eingeschränkt. Das folgende Kriterium, das auf A. CAUCHY (1789 —1857) zurückgeht, erlaubt es, die Konvergenz beliebiger Folgen in $\mathbb{C}$ zu untersuchen. Dazu benötigen wir einen neuen, fundamentalen Begriff.

DEFINITION II-5.7. *Eine Folge $(x_n)_{n\geq 1} \subset X$ heißt* ***Cauchy Folge*** $\underset{Def}{\Leftrightarrow}$

$$(*) \qquad \forall\varepsilon > 0\, \exists N(\varepsilon) \in \mathbb{N}\, \forall n, m \in \mathbb{N}\colon (n, m \geq N(\varepsilon)) \Rightarrow \|x_n - x_m\| < \varepsilon.$$

Anschaulich bedeutet diese Definition, daß die Folgenglieder immer näher zusammenrücken müssen.

THEOREM II-5.8. *Jede konvergente Folge $(x_n)_{n\geq 1} \subset X$ ist eine Cauchy Folge.*

BEWEIS. Es sei $\lim_{n\to\infty} x_n = x$ und $\varepsilon > 0$. Dann gibt es einen Index $N(\varepsilon)$, sodaß $\|x - x_n\| < \frac{1}{2}\varepsilon$ für alle $n \geq N(\varepsilon)$ gilt. Ist $n \geq N(\varepsilon)$ und $m \geq N(\varepsilon)$, folgt

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x_m - x\| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

□

Die Umkehrung ist in normierten Räumen nicht immer richtig!

LEMMA II-5.9. *In einem normierten Raum ist jede Cauchy Folge beschränkt.*

BEWEIS. Es sei (x_n) eine Cauchy Folge in X. Für $\varepsilon = 1$ gibt es dann einen Index $N \in \mathbb{N}$ derart, daß $\|x_n - x_m\| < 1$ für alle $n, m \geq N$ zutrifft. Für $n \geq N$ folgt daher

$$\|x_n\| \leq \|x_n - x_N\| + \|x_N\| < 1 + \|x_N\|.$$

Somit ist $(x_n)_{n\geq 1}$ beschränkt durch $\max\{\|x_1|, \dots, \|x_{N-1}\|, 1 + \|x_N\|\}$. □

THEOREM II-5.10 (Cauchy Kriterium). *Eine Folge $(z_n)_{n\geq 1} \subset \mathbb{C}^k$ ist konvergent genau dann, wenn sie eine Cauchy Folge ist.*

BEWEIS. Die Notwendigkeit der Cauchy Bedingung $(*)$ für Konvergenz wurde bereits in Satz II-5.8 nachgewiesen. Wir zeigen nun, daß sie auch hinreichend ist. Es sei also $(z_n)_{n\geq 1}$ eine Cauchy Folge. Wegen Lemma II-5.9 ist sie beschränkt und besitzt daher nach dem Satz von Bolzano Weierstrass **??** eine konvergente Teilfolge $(z_{\varphi(n)})_{n\geq 1}$. Es sei z deren Grenzwert. Zu $\varepsilon > 0$ gibt es dann natürliche Zahlen N_1, N_2 mit der Eigenschaft

$$\forall n, m \in \mathbb{N}\colon (n \geq N_1 \wedge m \geq N_1) \Rightarrow \|z_n - z_m\| < \frac{1}{2}\varepsilon$$
$$\forall n \in \mathbb{N}\colon n \geq N_2 \Rightarrow \|z_{\varphi(n)} - z\| < \frac{1}{2}\varepsilon.$$

($\|\cdot\|$ bezeichnet eine beliebige Norm in $\mathbb{C}^n$). Für $N = \max\{N_1, N_2\}$ ergibt sich nun für alle $n \geq N$ (man beachte, daß $\varphi(n) \geq n$, $n \in \mathbb{N}$, als Folge der strengen Monotonie von φ gelten muß)

$$\|z_n - z\| \leq \|z_n - z_{\varphi(n)}\| + \|z_{\varphi(n)} - z\| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

□

BEMERKUNG II-5.11. 1.) Die Stärke des Cauchy Kriteriums ist seine Universalität: es ist auf jede Folge (vorerst in $\mathbb{C}$) anwendbar. Allerdings gibt es keinen Hinweis auf den Grenzwert.
2.) Durch Negieren erhalten wir folgende Divergenzbedingung: $(z_n)_{n\geq 1} \subset \mathbb{C}^k$ ist divergent genau dann, wenn

$$\exists \varepsilon_0 > 0\, \forall N \in \mathbb{N}\, \exists n, m \in \mathbb{N}\colon n \geq N \wedge m \geq N \wedge \|z_n - z_m\| \geq \varepsilon_0.$$

Eine Folge (z_n) ist also genau dann divergent, wenn es nach jedem Glied der Folge weitere Folgenglieder gibt, deren Abstand voneinander größer als eine feste Schranke ist.

DEFINITION II-5.12. *Ein normierter Raum, in welchem jede Cauchy Folge konvergiert heißt **vollständig** oder **Banach Raum**.*

Satz II-5.10 zeigt, daß $\mathbb{C}^n$ und somit auch $\mathbb{R}^n$ vollständige Räume sind. Nicht jeder normierte Raum ist jedoch vollständig!

6. Limes inferior und Limes superior

Beschränkte Zahlenfolgen sind nicht notwendig konvergent, sie besitzen aber nach dem Satz von Bolzano Weierstrass II-5.5 stets konvergente Teilfolgen. Es wird nun ein Begriff vorgestellt, der als Ersatz für den fehlenden Grenzwert einer Folge dienen kann. Da die Ordnung in $\mathbb{R}$ dabei eine wesentliche Rolle spielt, ist dieser Abschnitt auf *reelle* Folgen beschränkt. Ausgangspunkt ist folgende Beobachtung:

LEMMA II-6.1. *Die Menge der Häufungswerte einer beschränkten Folge reeller Zahlen ist nicht leer und beschränkt.*

BEWEIS. Es sei $(x_n) \subset \mathbb{R}$ eine beschränkte Folge und $\mathcal{H}$ die Menge ihrer Häufungswerte. Nach dem Satz von Bolzano-Weierstraß II-5.5 ist $\mathcal{H}$ nicht leer. Da (x_n) beschränkt ist, existiert $M \geq 0$ mit

$$|x_n| \leq M, \qquad n \in \mathbb{N}.$$

Diese Ungleichung gilt auch für alle Teilfolgen und somit für alle Häufungswerte der Folge (x_n). □

Somit besitzt die Menge der Häufungswerte einer beschränkten Folge das Supremum und Infimum. Diese Beobachtung kann verschärft werden.

THEOREM II-6.2. *Die Menge der Häufungswerte einer beschränkten Folge reeller Zahlen hat ein größtes und ein kleinstes Element.*

BEWEIS. Wir bezeichnen wieder mit $\mathcal{H}$ die Menge der Häufungswerte einer beschränkten Folge (x_n) in $\mathbb{R}$. Es sei $\alpha = \sup \mathcal{H}$. Wir zeigen $\alpha \in \mathcal{H}$. Aus den Eigenschaften des Supremums folgt für jedes $n \in \mathbb{N}$ die Existenz eines Häufungswertes α_n der Folge (x_n) mit

$$\alpha - \frac{1}{2n} < \alpha_n \leq \alpha.$$

Da α_n ein Häufungswert von (x_n) ist, gibt es für jedes $n \in \mathbb{N}$ auch ein Folgenglied $x_{\varphi(n)}$ mit

$$|\alpha_n - x_{\varphi(n)}| \leq \frac{1}{2n}$$

($\varphi\colon \mathbb{N} \to \mathbb{N}$ kann streng monoton steigend gewählt werden). Somit folgt

$$(*) \qquad |\alpha - x_{\varphi(n)}| \leq |\alpha - \alpha_n| + |\alpha_n - x_{\varphi(n)}| < \frac{1}{n},$$

d.h. die Teilfolge $(x_{\varphi(n)})$ von (x_n) konvergiert gegen α, also $\alpha \in \mathcal{H}$. Der Beweis für die Existenz des Minimums von $\mathcal{H}$ verläuft analog. □

Dieser Satz motiviert folgende Verallgemeinerung des Grenzwertbegriffes.

DEFINITION II-6.3. *Es sei (x_n) eine beschränkte Folge reeller Zahlen und $\mathcal{H}$ die Menge ihrer Häufungswerte. Man nennt* $\max \mathcal{H}$ ***Limes superior*** *von (x_n) und* $\min \mathcal{H}$ ***Limes inferior*** *von (x_n). Folgende Notation ist üblich*

$$\limsup_{n\to\infty} x_n \equiv \overline{\lim}_{n\to\infty} x_n = \max \mathcal{H}$$
$$\liminf_{n\to\infty} x_n \equiv \underline{\lim}_{n\to\infty} x_n = \min \mathcal{H}.$$

BEISPIEL II-6.4. Es sei $(x_n)_{n\geq 1}$ definiert durch $x_n = (-1)^n(1 + \frac{1}{n})$. Eine einfache Überlegung zeigt

$$\limsup_{n\to\infty} x_n = 1 \quad \text{und} \quad \liminf_{n\to\infty} x_n = -1$$

Wir notieren einfache Folgerungen aus dieser Definition:

LEMMA II-6.5. *1.) Eine beschränkte reelle Folge* $(x_n)_{n\geq 1}$ *ist konvergent genau dann, wenn* $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$. *Im Falle der Konvergenz von* (x_n) *gilt* $\lim_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.
2.) $(x_n)_{n\geq 1} \subset \mathbb{R}$ *sei beschränkt. Es gilt stets*

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

BEMERKUNG II-6.6. Eine genauere Analyse der vorangehenden Betrachtungen zeigt, daß die Existenz des Limes superior (inferior) gesichert werden kann, wenn (x_n) nur nach oben (unten) beschränkt ist und mindestens einen Häufungswert besitzt. In diesem Fall kann $\mathcal{H} \neq \emptyset$ nicht durch den Satz von Bolzano-Weierstrass sichergestellt werden, sondern muß in die Voraussetzungen aufgenommen werden.

Wir stellen nun eine weitere Charakterisierung des Limes superior (Limes inferior) bereit.

THEOREM II-6.7. *Es sei* (x_n) *eine beschränkte Folge reeller Zahlen.*
1.) Es ist $\alpha = \limsup_{n\to\infty} x_n$ *genau dann, wenn für alle* $\varepsilon > 0$ *Folgendes gilt*

$$\begin{array}{ll} i) & \exists N(\varepsilon) \in \mathbb{N} \forall n \geq N(\varepsilon) \colon x_n < \alpha + \varepsilon \\ ii) & \exists (x_{\varphi(n)}) \subset (x_n) \forall n \in \mathbb{N} \colon x_{\varphi(n)} > \alpha - \varepsilon \end{array}$$

2.) Es ist $\beta = \liminf_{n\to\infty} x_n$ *genau dann, wenn für alle* $\varepsilon > 0$ *Folgendes gilt*

$$\begin{array}{ll} i) & \exists N(\varepsilon) \in \mathbb{N} \forall n \geq N(\varepsilon) \colon x_n > \beta - \varepsilon \\ ii) & \exists (x_{\varphi(n)}) \subset (x_n) \forall n \in \mathbb{N} \colon x_{\varphi(n)} < \beta + \varepsilon \end{array}$$

BEWEIS. Übung. □

Der Limes Superior α ist also dadurch charaktrisiert, daß für alle $\varepsilon > 0$ fast alle Folgenglieder kleiner sind als $\alpha + \varepsilon$ und unendlich viele Folgenglieder größer sind als $\alpha - \varepsilon$. Die erste Bedingung bedeutet, daß jede Zahl $\xi > \alpha$ nicht Häufungswert der Folge sein kann, die zweite Bedingung (gemeinsam mit der ersten) impliziert, daß α ein Häufungswert ist.

Wir zeigen nun, daß für den Limes superior (Limes inferior) ähnliche Rechenregeln gelten, wie für den gewöhnlichen Limes.

THEOREM II-6.8. *Es seien* $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ *beschränkte, reelle Folgen. Es gelte* $a_n \leq b_n$ *für alle* $n \in \mathbb{N}$. *Dann folgt*

$$\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n \quad \textit{und} \quad \liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n.$$

BEWEIS. Übung. □

THEOREM II-6.9. *Es seien $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ beschränkte, reelle Folgen. Dann gilt*

i) $\limsup_{n\to\infty}(a_n+b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$
ii) $\liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n \leq \liminf_{n\to\infty}(a_n+b_n)$
iii) $\liminf_{n\to\infty}(a_n+b_n) \leq \liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty}(a_n+b_n)$
iv) $\forall\alpha\geq 0\colon \limsup_{n\to\infty} \alpha a_n = \alpha \limsup_{n\to\infty} a_n$
v) $\forall\alpha\leq 0\colon \limsup_{n\to\infty} \alpha a_n = \alpha \liminf_{n\to\infty} a_n$

BEWEIS. ii) Da die Folge (a_n+b_n) beschränkt ist, gibt es nach Satz II-6.2 eine konvergente Teilfolge $(a_{\varphi(n)} + b_{\varphi(n)})_{n\geq 1}$ mit

$$\lim_{n\to\infty}(a_{\varphi(n)}+b_{\varphi(n)}) = \liminf_{n\to\infty}(a_n+b_n).$$

Man beachte, daß die Teilfolgen $(a_{\varphi(n)})$ und $(b_{\varphi(n)})$ selbst nicht konvergent sein müssen. Wir wählen daher aus $(a_{\varphi(n)})$ eine konvergente Teilfolge $(a_{(\psi\circ\varphi)(n)})$ und aus $(b_{(\psi\circ\varphi)(n)})$ eine konvergente Teilfolge $(b_{(\chi\circ\psi\circ\varphi)(n)})$ aus. Es sei $\nu=\chi\circ\psi\circ\varphi$. Dann sind die Folgen $(a_{\nu(n)})$, $(b_{\nu(n)})$ und $(a_{\nu(n)}+b_{\nu(n)})$ konvergent und es folgt

$$\liminf_{n\to\infty}(a_n+b_n) = \lim_{n\to\infty}(a_{\nu(n)}+b_{\nu(n)}) = \lim_{n\to\infty} a_{\nu(n)} + \lim_{n\to\infty} b_{\nu(n)} \geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n.$$

i) Analog.
iii) Wir beweisen nur die Ungleichung

$$\liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty}(a_n+b_n).$$

Wegen Satz II-6.7 gibt es für jedes beliebig gewählte $\varepsilon>0$ einen Index $N_1(\varepsilon)$ und eine Teilfolge $(b_{\varphi(n)})_{n\geq 1}$ von (b_n) so, daß

$$(*)\qquad \begin{cases}\forall n\in\mathbb{N}\colon n\geq N_1(\varepsilon)\Rightarrow a_n > \liminf_{k\to\infty} a_k - \varepsilon\\ \forall n\in\mathbb{N}\colon b_{\varphi(n)} > \limsup_{k\to\infty} b_k - \varepsilon\end{cases}$$

Da $(\varphi(n))_{n\geq 1}$ strikt monoton ist, gilt für alle $n\geq N_1(\varepsilon)$ auch $\varphi(n)\geq n\geq N_1(\varepsilon)$. Durch Addition der beiden Ungleichungen in $(*)$ folgt daher für alle $n\geq N_1(\varepsilon)$

$$a_{\varphi(n)}+b_{\varphi(n)} > \liminf_{k\to\infty} a_k + \limsup_{k\to\infty} b_k - 2\varepsilon.$$

Nun ist $(a_{\varphi(n)}+b_{\varphi(n)})_{n\geq 1}$ eine Teilfolge von (a_n+b_n), somit folgt mit Satz II-6.8

$$\limsup_{n\to\infty}(a_n+b_n) \geq \limsup_{n\to\infty}(a_{\varphi(n)}+b_{\varphi(n)}) \overset{II-6.8}{\geq} \liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n - 2\varepsilon.$$

Wegen Korollar II-I-3.7 ist dies gleichwertig mit

$$\liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty}(a_n+b_n).$$

Der Beweis von $\liminf_{n\to\infty}(a_n+b_n) \leq \liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$ und der beiden restlichen Behauptungen sei dem Leser als Übung überlassen. □

Bei manchen Konvergenzuntersuchungen ist folgendes Ergebnis nützlich.

THEOREM II-6.10. *Es sei $(a_n)_{n\geq 1}\subset\mathbb{R}^+\setminus\{0\}$ und beschränkt. Dann gilt*

$$(*)\qquad \liminf_{n\to\infty}\frac{a_{n+1}}{a_n} \leq \liminf_{n\to\infty}\sqrt[n]{a_n} \leq \limsup_{n\to\infty}\sqrt[n]{a_n} \leq \limsup_{n\to\infty}\frac{a_{n+1}}{a_n}.$$

Existiert insbesondere $\lim_{n\to\infty}\frac{a_{n+1}}{a_n}$, dann ist auch $(\sqrt[n]{a_n})_{n\geq 1}$ konvergent mit $\lim_{n\to\infty}\sqrt[n]{a_n} = \lim_{n\to\infty}\frac{a_{n+1}}{a_n}$.

Beweis. Wir zeigen nur die Ungleichung $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n\to\infty} \sqrt[n]{a_n}$. Der Beweis der letzten Ungleichung in (∗) verläuft analog, die mittlere wurde bereits in Lemma II-6.5 bewiesen. Es sei $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \alpha \geq 0$. Ist $\alpha = 0$, gibt es nichts zu beweisen. Es sei also $\alpha > 0$. Zu $\varepsilon > 0$ gibt es nach Satz II-6.7 einen Index $N(\varepsilon) \in \mathbb{N}$, sodaß $\frac{a_{n+1}}{a_n} > \alpha - \varepsilon > 0$ für alle $n \geq N(\varepsilon)$ zutrifft. Eine einfache Induktion zeigt dann die Gültigkeit von

$$\frac{a_n}{a_{N(\varepsilon)}} = \prod_{i=N(\varepsilon)}^{n-1} \frac{a_{i+1}}{a_i} > (\alpha - \varepsilon)^{n-N(\varepsilon)}$$

d.h. von

$$a_n > (\alpha - \varepsilon)^{n-N(\varepsilon)} a_{N(\varepsilon)}$$

für alle $n > N(\varepsilon)$. Es folgt für $\alpha - \varepsilon > 0$

$$\sqrt[n]{a_n} > (\alpha - \varepsilon) \sqrt[n]{a_{N(\varepsilon)}(\alpha - \varepsilon)^{-N(\varepsilon)}}, \qquad n > N(\varepsilon).$$

Mit Hilfe der Sätze II-6.8, Lemma II-6.5 und Beispiel II-3.1-iii) ergibt sich

$$\begin{aligned} \liminf_{n\to\infty} \sqrt[n]{a_n} &\geq (\alpha - \varepsilon) \liminf_{n\to\infty} \sqrt[n]{a_{N(\varepsilon)}(\alpha - \varepsilon)^{-N(\varepsilon)}} \\ &= (\alpha - \varepsilon) \lim_{n\to\infty} \sqrt[n]{a_{N(\varepsilon)}(\alpha - \varepsilon)^{-N(\varepsilon)}} = \alpha - \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig wählbar war, folgt $\liminf_{n\to\infty} \sqrt[n]{a_n} \geq \alpha = \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$. Existiert $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ so gilt in (∗) nach Lemma II-6.5 überall die Gleichheit, daraus folgt die Behauptung. □

Beispiel II-6.11. 1.) Als Anwendung zeigen wir den Grenzwert

$$\lim_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = 0.$$

Dies folgt aus $\lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0$ und Satz II-6.10.
2.) Die Folge $(a_n)_{n\geq 1}$ sei definiert durch $a_n = \frac{n^n}{n!}$. Es folgt $\frac{a_{n+1}}{a_n} = (1 + \frac{1}{n})^n$ und mit Beispiel II-5.4 ergibt sich $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = e$. Mit Satz II-6.10 schließen wir dann $\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \frac{n}{\sqrt[n]{n!}} = e$.

7. Die erweiterten reellen Zahlen und uneigentliche Grenzwerte

Die Konvergenz von Folgen wurde bisher nur für beschränkte Folgen erklärt. Es ist aber manchmal sinnvoll, auch unbeschränkten reellen Folgen einen Grenzwert zuzuweisen. Zu diesem Zweck ergänzen wir das System der reellen Zahlen mit zwei verschiedenen festen Objekten, die wir mit „∞“ (gelesen: unendlich) und „$-\infty$“ bezeichnen. Diese Objekte sind *keine* reellen Zahlen. Wir bilden $\bar{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$ und nennen $\bar{\mathbb{R}}$ **erweiterte reelle Zahlen**. Behält man innerhalb von $\mathbb{R}$ die ursprüngliche (strikte) Ordnung bei und setzt

i) $-\infty < \infty$,

ii) $\forall x \in \mathbb{R}\colon -\infty < x < \infty$,

so wird damit eine strikte Ordnung auf $\bar{\mathbb{R}}$ erklärt. Für die neuen Symbole gelten folgende Rechenregeln: Es sei $x \in \mathbb{R}$, dann gilt

i) $x + \infty = \infty + x = x - (-\infty) = \infty$,
$x + (-\infty) = -\infty + x = x - \infty = -\infty$.

ii) $x > 0\colon \quad \infty \cdot x = x \cdot \infty = \infty$
$\qquad (-\infty) \cdot x = x \cdot (-\infty) = -\infty$

$x < 0\colon \quad \infty \cdot x = x \cdot \infty = -\infty$
$\qquad (-\infty) \cdot x = x \cdot (-\infty) = \infty$

iii) $\infty + \infty = \infty, \quad (-\infty) + (-\infty) = -\infty, \quad \infty \cdot \infty = \infty$

Die Ausdrücke $\infty + (-\infty)$, $0 \cdot \infty$, $\frac{\infty}{\infty}$, ..., sind nicht definiert.

Definition II-7.1. i) *Ist eine Teilmenge* $E \subset \bar{\mathbb{R}}$ *in* $\mathbb{R}$ *nicht nach oben beschränkt, setzt man* $\sup E = \infty$.

ii) *Ist eine Teilmenge* $E \subset \bar{\mathbb{R}}$ *in* $\mathbb{R}$ *nicht nach unten beschränkt, setzt man* $\inf E = -\infty$.

Diese Definition stellt eine natürliche Verallgemeinerung des Supremums (Infimums) auf beliebige nicht leeren Mengen reeller Zahlen dar. Jede Teilmenge von $\bar{\mathbb{R}}$ besitzt also in $\bar{\mathbb{R}}$ sowohl Supremum als auch Infimum. Da jede reelle Zahl zugleich obere und untere Schranke für die leere Menge ist, erhält man hat die merkwürdige Konsequenz:

Bemerkung II-7.2. $\sup \emptyset = -\infty$, $\inf \emptyset = \infty$.

Die Definition eines Intervalls I-3.13 kann in offensichtlicher Weise auf $\bar{\mathbb{R}}$ übertragen werden. Die Intervalle in $\bar{\mathbb{R}}$

$$(\xi, \infty], \qquad [-\infty, \xi)$$

sind ξ-Umgebungen von ∞ bzw. $-\infty$.

Definition II-7.3. *Es sei* $(x_n)_{n\geq 1} \subset \bar{\mathbb{R}}$.
Die Folge (x_n) *hat den* ***uneigentlichen Grenzwert*** ∞ *(*$-\infty$*) genau dann, wenn*

$$\forall \xi \in \mathbb{R}\, \exists N(\xi)\, \forall n \in \mathbb{N}\colon n \geq N(\xi) \Rightarrow x_n > \xi \quad (x_n < \xi).$$

Man sagt auch, die Folge (x_n) **divergiert (konvergiert uneigentlich)** *nach* ∞ *(*$-\infty$*).*
Schreibweise: $\lim_{n\to\infty} x_n = \infty$ *(* $\lim_{n\to\infty} x_n = -\infty$*)*

Eine unmittelbare Folge dieser Definition ist folgende Erweiterung des Monotoniekriteriums Satz II-5.2.

Theorem II-7.4. *Jede monotone Folge in* $\bar{\mathbb{R}}$ *hat einen Grenzwert in* $\bar{\mathbb{R}}$.

Wegen Definition II-7.1 ist folgende Vereinbarung sinnvoll:

DEFINITION II-7.5. *Es sei* $(x_n)_{n\geq 1} \subset \bar{\mathbb{R}}$.

(1) $\limsup_{n\to\infty} x_n = \infty \underset{Def}{\Leftrightarrow} (x_n)$ *ist nicht nach oben beschränkt*
(2) $\liminf_{n\to\infty} x_n = -\infty \underset{Def}{\Leftrightarrow} (x_n)$ *ist nicht nach unten beschränkt*

Man kann sich leicht davon überzeugen, daß den Regeln für das Rechnen mit den Symbolen ∞, $-\infty$ Regeln für die Grenzwerte von Folgen in $\bar{\mathbb{R}}$ entsprechen, welche nach ∞ bzw. $-\infty$ divergieren. Es gelte beispielsweise $\lim_{n\to\infty} x_n = x \in \mathbb{R}$ und $\lim_{n\to\infty} y_n = \infty$. Zu $\xi \in \mathbb{R}$ gibt es Indices N_1 und $N_2(\xi)$ so,daß $|x_n - x| < 1$ für $n \geq N_1$ und $y_n > \xi + 1 - x$ für $n \geq N_2(\xi)$ gilt. Für $n \geq \max\{N_1, N_2(\xi)\}$ folgt daher $x_n + y_n > x - 1 + \xi + 1 - x = \xi$. Somit konvergiert die Folge $(x_n + y_n)$ uneigentlich gegen ∞. Dies entspricht der Regel $x + \infty = \infty$. Auf analoge Weise können auch die übrigen Regeln interpretiert werden. Divergieren die Folgen (x_n) und (y_n) gegen ∞, so können die Folgen $(x_n - y_n)$, $(\frac{x_n}{y_n})$ ein recht unterschiedliches Verhalten aufweisen:

BEISPIEL II-7.6.

i)	$x_n = \sqrt{n+1} - \sqrt{n}$,	$\lim_{n\to\infty} x_n = 0$
ii)	$y_n = n - n^2$,	$\lim_{n\to\infty} y_n = -\infty$
iii)	$z_n = \frac{n}{n^2}$,	$\lim_{n\to\infty} z_n = 0$
iv)	$w_n = \frac{1+2+\cdots+n}{n^2}$,	$\lim_{n\to\infty} w_n = \frac{1}{2}$

Entsprechendes gilt auch für $(x_n y_n)$, falls $\lim_{n\to\infty} x_n = 0$ und $\lim_{n\to\infty} y_n = \infty$. Den Ausdrücken der Form $\infty - \infty$, $\frac{\pm\infty}{\pm\infty}$ und $0 \cdot (\pm\infty)$ kann somit a priori kein sinnvoller Wert zugewiesen werden. Mann nennt sie deshalb auch **unbestimmte Ausdrücke**.

8. Doppelfolgen

DEFINITION II-8.1. *Eine Abbildung* $x\colon \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ *heißt* ***Doppelfolge*** *komplexer Zahlen. Anstelle von* $x(n,m)$ *schreibt man* x_{nm}*, entsprechend bezeichnen wir eine Doppelfolge mit* $(x_{nm})_{n,m\geq 1}$.

Ein großer Teil der elementaren Theorie der Konvergenz einfacher Folgen läßt sich ohne große Änderungen auf Doppelfolgen übertragen. Wir beschränken uns daher auf die Darstellung einiger charakteristischer Unterschiede. Vorerst legen wir jedoch fest, was wir unter dem Grenzwert einer Doppelfolge verstehen:

DEFINITION II-8.2. *Es sei* $(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ *eine Doppelfolge und* $\alpha \in \mathbb{C}$.

$$(x_{nm})_{n,m\geq 1}\textbf{konvergiert}\ \textit{gegen}\ \alpha \underset{Def}{\Leftrightarrow} \forall \varepsilon > 0\ \exists N(\varepsilon) \in \mathbb{N}\ \forall n, m \in \mathbb{N} :$$
$$n \geq N(\varepsilon) \wedge m \geq N(\varepsilon) \Rightarrow |x_{nm} - \alpha| < \varepsilon.$$

Man schreibt $\lim_{n,m\to\infty} x_{nm} = \alpha$.

Die Eindeutigkeit des Grenzwertes, falls er existiert, läßt sich genauso wie in Satz II-2.7 zeigen. Ebenso lassen sich die Rechenregeln für konvergente Folgen aus Abschnitt 2 auf Doppelfolgen übertragen. Es gilt auch ein Cauchy-Kriterium, welches wir ohne Beweis notieren:

THEOREM II-8.3. *Eine Doppelfolge* $(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ *ist genau dann konvergent, wenn sie eine Cauchy-Folge ist, d.h. wenn*

$$\forall \varepsilon > 0\ \exists N(\varepsilon)\ \forall n,m,r,s \in \mathbb{N}\colon n,m,r,s \geq N(\varepsilon) \Rightarrow |x_{nm} - x_{rs}| < \varepsilon.$$

Oft ordnet man die Glieder einer Doppelfolge (x_{nm}) in einem Rechteckschema an,

$$\begin{array}{cccccc} x_{11} & x_{12} & \dots & x_{1m}\dots & \longrightarrow & \xi_1 \\ x_{21} & x_{22} & \dots & x_{2m}\dots & \longrightarrow & \xi_2 \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm}\dots & \longrightarrow & \xi_n \\ \downarrow & \downarrow & & \downarrow & & \vdots \\ \eta_1 & \eta_2 & \dots & \eta_m\dots & & \end{array}$$

und nennt den ersten Index von x_{nm} entsprechend Zeilenindex, den zweiten Index Spaltenindex. Es wird deutlich, daß man (x_{nm}) zumindest auf zwei Weisen als Folge von Folgen betrachten kann: Einerseits kann man die Folgen in den Zeilen des obigen Schemas betrachten. Falls diese Folgen konvergieren, für alle $n \in \mathbb{N}$ sei etwa $\lim_{m\to\infty} x_{nm} = \xi_n$, dann kann man auch die Folge $(\xi_n)_{n\geq 1}$ auf Konvergenz untersuchen. Es existiere $\xi = \lim_{n\to\infty} \xi_n$, ξ ist dann gegeben durch

$$\xi = \lim_{n\to\infty}(\lim_{m\to\infty} x_{nm})$$

und man nennt ξ den **zeileniterierten Grenzwert** der Doppelfolge (x_{nm}). Andererseits kann man die Folgen in den Spalten des Rechteckschemas betrachten. Bezeichnet man mit $\eta_m = \lim_{n\to\infty} x_{nm}$ und mit $\eta = \lim_{m\to\infty} \eta_m$ (soferne diese Grenzwerte existieren), gelangt man zum **spalteniterierten Grenzwert** η der Doppelfolge (x_{nm}),

$$\eta = \lim_{m\to\infty}(\lim_{n\to\infty} x_{nm}).$$

Einer Doppelfolge (x_{nm}) kann man also auf natürliche Weise drei charakteristische Größen zuordnen: Ihren Grenzwert $x = \lim_{n,m\to\infty} x_{nm}$, den man der Deutlichkeit halber manchmal auch als **Doppellimes** bezeichnet, und die beiden iterierten Grenzwerte. Natürlich stellt sich die Frage, wie diese Größen zusammenhängen. Wir demonstrieren an einem Beispiel, daß ohne Zusatzbedingung kein Zusammenhang zu erwarten ist.

BEISPIEL II-8.4. 1) $x_{nm} = (-1)^{m+n}(\frac{1}{n} + \frac{1}{m})$. Es existiert $\lim_{n,m\to\infty} x_{nm} = 0$, aber es existiert keiner der beiden iterierten Grenzwerte, da etwa für jedes $n \in \mathbb{N}$ die Folge $(x_{nm})_{m\geq 1}$ divergiert. Die Existenz des Doppellimes impliziert also nicht die Existenz der iterierten Grenzwerte.
2) $x_{nm} = (-1)^m(\frac{1}{n} + \frac{1}{m})$. Es gilt wieder $\lim_{n,m\to\infty} x_{nm} = 0$. Für jedes $m \in \mathbb{N}$ ist nun aber $\lim_{n\to\infty} x_{nm} = (-1)^m \lim_{n\to\infty}(\frac{1}{n} + \frac{1}{m}) = (-1)^m\frac{1}{m}$ und somit $\lim_{m\to\infty}(\lim_{n\to\infty} x_{nm}) = 0$. Der zeileniterierte Limes existiert nicht.

3) $x_{nm} = \frac{nm}{n^2+m^2}$. Halten wir $n \in \mathbb{N}$ fest, dann gilt $\lim_{m\to\infty} x_{nm} = 0$ und somit auch $\lim_{n\to\infty}(\lim_{m\to\infty} x_{nm}) = 0$. Aus Symmetriegründen gilt $\lim_{m\to\infty}(\lim_{n\to\infty} x_{nm}) = 0$. Die beiden iterierten Grenzwerte existieren und sind gleich, trotzdem existiert der Doppellimes nicht: Für alle $n \in \mathbb{N}$ ist nämlich $x_{nn} = \frac{1}{2}$ und $x_{n,2n} = \frac{2}{5}$, das Cauchy-Kriterium II-8.3 kann nicht erfüllt werden.
4) $x_{nm} = \frac{1}{n} + \frac{1}{m}$. In diesem Falle existieren alle drei Limiten und sind gleich Null.

Theorem II-8.5. *Es sei $(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ eine konvergente Doppelfolge. Existiert für jedes $n \in \mathbb{N}$ der Zeilenlimes $\xi_n = \lim_{m\to\infty} x_{nm}$, dann existiert der zeileniterierte Limes $\lim_{n\to\infty}(\lim_{m\to\infty} x_{nm})$ und es gilt*

$$\lim_{n,m\to\infty} x_{nm} = \lim_{n\to\infty}(\lim_{m\to\infty} x_{nm}).$$

Beweis. Aus der Existenz des Doppellimes x schließen wir für jedes $\varepsilon > 0$ auf einen Index $N(\varepsilon)$, sodaß $|x_{nm} - x| < \varepsilon$ für alle $n, m \geq N(\varepsilon)$ gilt. Nach Korollar II-4.5 kann man in dieser Ungleichung den Grenzübergang $m \to \infty$ durchführen und erhält für alle $n \geq N(\varepsilon)$

$$|\xi_n - x| \leq \varepsilon,$$

d.h.

$$\lim_{n\to\infty} \xi_n = \lim_{n\to\infty}(\lim_{m\to\infty} x_{nm}) = x.$$

□

Dieser Satz zeigt, daß bei einer konvergenten Doppelfolge die iterierten Limiten nur dann nicht existieren können, wenn die entsprechenden „inneren“ Limiten nicht existieren. Genauer gilt

Korollar II-8.6. Es sei $(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ eine konvergente Doppelfolge. Existieren für alle $n, m \in \mathbb{N}$ die einfachen Grenzwerte

$$\xi_n = \lim_{m\to\infty} x_{nm} \quad \text{und} \quad \eta_m = \lim_{n\to\infty} x_{nm},$$

dann existieren auch die beiden iterierten Grenzwerte und es gilt

$$\lim_{n,m\to\infty} x_{nm} = \lim_{n\to\infty}(\lim_{m\to\infty} x_{nm}) = \lim_{m\to\infty}(\lim_{n\to\infty} x_{nm}).$$

Wir untersuchen nun das letzte Beispiel in II-8.4 etwas genauer: Für jedes n betrachten wir die Zeilenfolge $m \mapsto x_{nm}$. Es gilt $\lim_{m\to\infty} x_{nm} = \frac{1}{n}$, denn $|x_{nm} - \frac{1}{n}| = \frac{1}{m}$. Zu einem beliebigen $\varepsilon > 0$ wählen wir $N_0 \in \mathbb{N}$ mit $N_0 > \frac{1}{\varepsilon}$. *Für alle* $n \in \mathbb{N}$ folgt dann $|x_{nm} - \frac{1}{n}| = \frac{1}{m} < \frac{1}{N_0} < \varepsilon$ soferne $m \geq N_0$. Wir erkennen, daß für jede Zeilenfolge derselbe Index N_0 gewählt werden kann, um den Approximationsfehler unter eine vorgegebene Toleranzgrenze zu drücken. Die Möglichkeit einer gleichmäßigen Wahl von N_0 erlaubt es auch, aus der Existenz eines der iterierten Limiten auf den im allgemeinen komplizierter zu berechnenden Doppellimes zu schließen.

Definition II-8.7. *Es seien $\Lambda \neq \emptyset$ eine Indexmenge und $\{(x_{n\lambda})_{n\geq 1} \colon \lambda \in \Lambda\}$ eine Familie konvergenter Folgen komplexer Zahlen, d.h. für alle $\lambda \in \Lambda$ existiert $x_\lambda \in \mathbb{C}$ mit $x_\lambda = \lim_{n\to\infty} x_{n\lambda}$.*

Die Folgen $(x_{n\lambda})$ konvergieren ***gleichmäßig*** *bezüglich $\lambda \in \Lambda$ gegen $x_\lambda \underset{Def}{\Leftrightarrow}$*

$$\forall \varepsilon > 0\ \exists N(\varepsilon) \in \mathbb{N}\ \forall \lambda \in \Lambda\ \forall n \in \mathbb{N}\colon n \geq N(\varepsilon) \Rightarrow |x_{n\lambda} - x_\lambda| < \varepsilon.$$

Lemma II-8.8. *$(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ sei eine konvergente Doppelfolge. Falls für alle $n \in \mathbb{N}$ der Zeilenlimes $\xi_n = \lim_{m\to\infty} x_{nm}$ existiert, dann sind die Zeilenfolgen $(x_{nm})_{m\geq 1}$, $n \in \mathbb{N}$, gleichmäßig konvergent bezüglich n.*

Beweis. Wegen der Existenz des Doppellimes $x = \lim_{n,m\to\infty} x_{nm}$ existiert für alle $\varepsilon > 0$ ein Index $N_0(\varepsilon)$, sodaß für alle $n, m \geq N_0(\varepsilon)$

$$(*) \qquad |x_{nm} - x| < \varepsilon$$

zutrifft. Da aber auch für jedes $n \in \mathbb{N}$ der Zeilenlimes $\xi_n = \lim_{m\to\infty} x_{nm}$ existiert, kann man in $(*)$ den Grenzübergang $m \to \infty$ durchführen und erhält nach Korollar II-4.5 für alle $n \geq N_0(\varepsilon)$

$$|\xi_n - x| \leq \varepsilon.$$

Somit folgt für $n, m \geq N_0(\varepsilon)$

$$(**) \qquad |x_{nm} - \xi_n| \leq |x_{nm} - x| + |x - \xi_n| < 2\varepsilon.$$

Es konvergieren auch die Zeilenfolgen mit $n = 1, \ldots, N_0(\varepsilon) - 1$. Da es sich nur um endlich viele Folgen handelt, gibt es einen Index $N_1(\varepsilon)$, sodaß

$$|x_{nm} - \xi_n| < \varepsilon$$

für $m \geq N_1(\varepsilon)$ und $n = 1, \ldots, N_0(\varepsilon) - 1$ zutrifft. Setzt man nun $N(\varepsilon) = \max\{N_0(\varepsilon), N_1(\varepsilon)\}$, so gilt $(**)$ für alle $n \in \mathbb{N}$ und für alle $m \geq N(\varepsilon)$. $\square$

Dieses Lemma zeigt, daß unter der Voraussetzung der Existenz der Zeilenlimiten, die Gleichmäßigkeit ihrer Konvergenz eine notwendige Bedingung für die Existenz des Doppellimes ist. Im folgenden Satz wird gezeigt, daß umgekehrt diese Bedingung die Vertauschbarkeit von Zeilen– und Spaltenlimes sicherstellt.

Theorem II-8.9 (Iterierte Grenzwerte). *Existieren für alle $n, m \in \mathbb{N}$ die einfachen Grenzwerte*

$$\xi_n = \lim_{m\to\infty} x_{nm} \quad \textit{und} \quad \eta_m = \lim_{n\to\infty} x_{nm}$$

einer komplexen Doppelfolge $(x_{nm})_{n,m\geq 1}$ und ist die Konvergenz mindestens einer der beiden Familien von Folgen gleichmäßig, so existieren sowohl beide iterierte Grenzwerte, als auch der Doppellimes und es gilt

$$\lim_{n,m\to\infty} x_{nm} = \lim_{n\to\infty}\big(\lim_{m\to\infty} x_{nm}\big) = \lim_{m\to\infty}\big(\lim_{n\to\infty} x_{nm}\big).$$

Beweis. Wir nehmen an, die Konvergenz der Zeilenfolgen $(x_{nm})_{m\geq 1}, n \in \mathbb{N}$, gegen ξ_n sei gleichmäßig, d.h. für jedes $\varepsilon > 0$ gibt es einen Index $N(\varepsilon)$, sodaß für alle $m \geq N(\varepsilon)$

$$|x_{nm} - \xi_n| < \varepsilon$$

für alle $n \in \mathbb{N}$ gilt. Wir zeigen, daß die Folge $(\xi_n)_{n\geq 1}$ eine Cauchy-Folge ist. Wir wählen einen festen Index $q \geq N(\varepsilon)$. Da nach Voraussetzung auch die Spaltenfolge

$(x_{nq})_{n\geq 1}$ konvergiert, also eine Cauchy-Folge ist, gibt es zu dem bereits gewählten ε einen weiteren Index $K = K(\varepsilon, q) \geq N(\varepsilon)$, sodaß $|x_{rq} - x_{sq}| < \varepsilon$ für alle $r, s \geq K$ gilt. Somit folgt für $r, s \geq K(\varepsilon, q)$

$$|\xi_r - \xi_s| \leq |\xi_r - x_{rq}| + |x_{rq} - x_{sq}| + |x_{sq} - \xi_s| < 3\varepsilon.$$

Die Folge $(\xi_n)_{n\geq 1}$ ist also eine Cauchy-Folge und somit konvergent:

$$\lim_{n\to\infty} (\lim_{m\to\infty} x_{nm}) = \lim_{n\to\infty} \xi_n = x.$$

Es existiert also der zeileniterierte Limes. Wir zeigen nun die Existenz des Doppellimes. Wegen der Konvergenz von (ξ_n) gibt es zu jedem $\varepsilon > 0$ einen Index $M(\varepsilon)$, sodaß $|x - \xi_n| < \varepsilon$ für alle $n \geq M(\varepsilon)$ zutrifft. Setzt man $\tilde{N}(\varepsilon) = \max\{M(\varepsilon), N(\varepsilon)\}$ ergibt sich für $n, m \geq \tilde{N}(\varepsilon)$

$$|x_{nm} - x| \leq |x_{nm} - \xi_n| + |\xi_n - x| < 2\varepsilon,$$

d.h.

$$\lim_{n,m\to\infty} x_{nm} = x.$$

Die Existenz des spalteniterierten Grenzwertes und dessen Gleichheit mit dem Doppellimes folgt aus Satz II-8.5. □

9. Reihen

Das Paradoxon des Zenon von Elea *(495 – 435 v.Chr.): Ein Läufer, der eine konstante Geschwindigkeit einhält, kann niemals das Ende der Rennbahn erreichen.* Zenon argumentiert, der Läufer müsse ja zuerst die Hälfte der Strecke, dann die Hälfte der zweiten Hälfte, dann die Hälfte des verbleibenden Viertels, u.s.w, zurücklegen. Der Läufer muß also unendlich viele Teilstrecken durchlaufen, für welche er jeweils eine bestimmte Zeit benötigt. Somit kann der Läufer sein Ziel nie erreichen.
Nehmen wir an, der Läufer benötige für die erste Hälfte der Rennstrecke gerade eine Zeiteinheit. Das Durchlaufen des nächsten Viertels erfordert eine weitere halbe Zeiteinheit, das anschließende Achtel der Strecke wird in der nächsten viertel Zeiteinheit zurückgelegt. Die Bewältigung der ersten n Teilstrecken erfordert also

$$t_n = 1 + \frac{1}{2} + (\frac{1}{2})^2 + \cdots + (\frac{1}{2})^{n-1}$$

Zeiteinheiten. Zenon, der den Begriff des Grenzwertes noch nicht kannte, war also offensichtlich der Ansicht, die „Addition“ unendlich vieler Zeitintervalle, auch wenn diese immer kleiner werden, müsse eine unendlich große Zeitspanne ergeben. Wir können das Paradoxon klären, da wegen Lemma II-I-4.23 die Laufzeit für n Teilstrecken durch

$$t_n = 2 - (\frac{1}{2})^{n-1}$$

gegeben ist und somit die Gesamtlaufzeit $T = \lim_{n\to\infty} t_n = 2$ Zeiteinheiten beträgt. Dieses Beispiel zeigt auch, wie das Aufsummieren von „unendlich“ vielen Summanden präzisiert werden kann.

DEFINITION II-9.1. *Es sei X ein normierter Raum, $(a_k)_{k\geq 1}$ eine Folge in X.*

i) *$S_n = \sum_{k=1}^{n} a_k$, $n \in \mathbb{N}$, heißt* ***n-te Partialsumme*** *der Folge (a_k).*
ii) *Die Folge der n-ten Partialsummen $(S_n)_{n\geq 1}$ heißt* ***unendliche Reihe****. Anstelle von $(S_n)_{n\geq 1}$ schreibt man $\sum_{k=1}^{\infty} a_k$. Man nennt a_k das* **k-te Glied** *der unendlichen Reihe.*
iii) *$\sum_{k=1}^{\infty} a_k$ heißt* ***konvergent (divergent)*** *, wenn die Folge der Partialsummen $(S_n)_{n\geq 1}$ konvergiert (divergiert). Man nennt den Grenzwert der Partialsummenfolge* ***Summe*** *(Grenzwert) der Reihe $\sum_{k=1}^{\infty} a_k$.*

Wir stellen fest, daß das Symbol $\sum_{n=1}^{\infty} a_n$ eine zweifache Bedeutung besitzt: Es bezeichnet einerseits die unendliche Reihe und andererseits, im Falle der Konvergenz, auch deren Grenzwert. Wir betonen nachdrücklich, daß die Summe einer konvergenten Reihe *nicht* als Ergebnis simultanen Aufsummierens unendlich vieler Summanden zu verstehen ist, sondern durch den Grenzwert der Folge der Partialsummen bestimmt ist. Die Reihenfolge der Reihenglieder geht also ganz wesentlich in die Definition der Summe einer Reihe ein (der Wert jeder einzelnen Partialsumme ist natürlich unabhängig von der Reihenfolge ihrer Summanden). Wir können somit viele Ergebnisse für Folgen unmittelbar auf Reihen übertragen. Beispielsweise kann man Satz II-2.10 in Rechenregeln für *konvergente* Reihen übersetzen.

THEOREM II-9.2. *Es seien $\sum_{n=1}^{\infty} a_n$, $\sum_{n=1}^{\infty} b_n$ konvergente Reihen in X und $\lambda \in \mathbb{K}$. Dann konvergieren auch die Reihen $\sum_{n=1}^{\infty} \lambda a_n$ und $\sum_{n=1}^{\infty}(a_n + b_n)$. Ferner gilt*

i) $\sum_{n=1}^{\infty} \lambda a_n = \lambda \sum_{n=1}^{\infty} a_n$,
ii) $\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n$.

Aus der Definition folgt unmittelbar eine nützliche *notwendige* Konvergenzbedingung:

THEOREM II-9.3. *Es sei $(a_n)_{n\geq 1}$ eine Folge in X. Wenn die Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert, dann gilt $\lim_{n\to\infty} a_n = 0$.*

BEWEIS. Wegen der Konvergenz der Reihe existiert $\lim_{n\to\infty} S_n$. Die Behauptung folgt nun aus $\lim_{n\to\infty} a_n = \lim_{n\to\infty}(S_{n+1} - S_n) = 0$. ☐

THEOREM II-9.4 (Cauchykriterium). *Es sei X ein* vollständiger*, normierter Raum und $(a_n)_{n\geq 1} \subset X$. Die unendliche Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert genau dann, wenn folgendes gilt:*

$$(*) \qquad \forall \varepsilon > 0\ \exists N(\varepsilon) \in \mathbb{N}\ \forall n, m \in \mathbb{N}\colon n \geq m \geq N(\varepsilon) \Rightarrow \Big\| \sum_{k=m+1}^{n} a_k \Big\| < \varepsilon.$$

Für $n = m$ gilt wegen unserer Konvention über leere Summen $\sum_{k=m+1}^{n} a_k = 0$.

BEWEIS. Da X vollständig ist, konvergiert die unendliche Reihe $\sum_{n=1}^{\infty} a_n$ genau dann, wenn (S_n) eine Cauchy Folge ist. Dies ist gleichwertig zu: für jedes $\varepsilon > 0$ existiert ein Index $N(\varepsilon)$, sodaß für alle $n, m \geq N(\varepsilon)$ und $n \geq m$

$$\|S_n - S_m\| = \Big\|\sum_{k=1}^{n} a_k - \sum_{k=1}^{m} a_k\Big\| = \Big\| \sum_{k=m+1}^{n} a_k\Big\| < \varepsilon$$

zutrifft. □

Als Folgerung zeigen wir nun, daß aus der Konvergenz von $\sum_{n=1}^{\infty} \|a_n\|$ die Konvergenz von $\sum_{n=1}^{\infty} a_n$ folgt.

KOROLLAR II-9.5. Es sei $(a_n)_{n\geq 1}$ eine Folge in einem vollständigen, normierten Raum X.
Konvergiert die Reihe $\sum_{n=1}^{\infty} \|a_n\|$, dann ist auch $\sum_{n=1}^{\infty} a_n$ konvergent.

BEWEIS. Bezeichnen wir mit $S_n = \sum_{k=1}^{n} a_k$ und $T_n = \sum_{k=1}^{n} \|a_k\|$, folgt die Behauptung mit Hilfe der Dreiecksungleichung (es sei $n > m$)

$$\|S_n - S_m\| = \Big\| \sum_{k=m+1}^{n} a_k \Big\| \leq \sum_{k=m+1}^{n} \|a_k\| = \|T_n - T_m\|,$$

und dem Cauchy-Kriterium. □

BEMERKUNG II-9.6. Führt man in $(*)$ den Grenzübergang $n \to \infty$ durch, ergibt sich eine weitere notwendige Konvergenzbedingung:

$$\sum_{n=1}^{\infty} a_n \text{ ist konvergent } \Rightarrow \forall \varepsilon > 0\, \exists N(\varepsilon) \in \mathbb{N}\, \forall m \geq N(\varepsilon) \colon \Big\| \sum_{k=m+1}^{\infty} a_k \Big\| < \varepsilon.$$

Wir betrachten nun den Prototyp einer divergenten und einer konvergenten Reihe.

THEOREM II-9.7 (Harmonische Reihe). $\sum_{n=1}^{\infty} \frac{1}{n}$ *ist divergent.*

BEWEIS. Es sei $S_n = \sum_{k=1}^{n} \frac{1}{k}$, $n \in \mathbb{N}$. Wir zeigen

$$(*) \qquad \forall n \in \mathbb{N} \colon S_{2^n} \geq \frac{1}{2}(n+2),$$

somit ist $(S_n)_{n\geq 1}$ unbeschränkt und daher divergent. Den Nachweis von $(*)$ führen wir mit Hilfe vollständiger Induktion. Es ist $S_{2^1} = \frac{3}{2}$. Mit Hilfe von $S_{2^n} \geq \frac{1}{2}(n+2)$ schätzen wir $S_{2^{n+1}}$ folgendermaßen ab

$$\begin{aligned} S_{2^{n+1}} &= S_{2^n} + \sum_{k=2^n+1}^{2^{n+1}} \frac{1}{k} \geq \frac{1}{2}(n+2) + \sum_{k=2^n+1}^{2^{n+1}} 2^{-n-1} \\ &= \frac{1}{2}(n+2) + 2^{-n-1}(2^{n+1} - (2^n+1) + 1) \\ &= \frac{1}{2}(n+2) + \frac{1}{2} = \frac{1}{2}(n+3). \end{aligned}$$

□

Da (S_n) monoton wächst, können wir der Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ den uneigentlichen Grenzwert ∞ zuweisen. Das Beispiel der harmonischen Reihe zeigt, daß die notwendige Konvergenzbedingung aus Satz II-9.3 nicht hinreichend ist.

THEOREM II-9.8 (Geometrische Reihe). *Es sei $z \in \mathbb{C}$. Die Reihe $\sum_{n=0}^{\infty} z^n$ konvergiert gegen die Summe $\frac{1}{1-z}$ genau dann, wenn $|z| < 1$.*

Beweis. Nach Lemma II-I-4.23 ist die n-te Partialsumme gegeben durch

$$S_n = \sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z}.$$

Ist $|z| < 1$, folgt weiter aus den Sätzen II-3.3 und II-4.1

$$\lim S_n = \frac{1 - \lim z^n}{1 - z} = \frac{1}{1 - z}.$$

Für $|z| \geq 1$ ist $|z^n| \geq 1$. Somit ist nach Satz II-9.3 $\sum_{k=0}^{\infty} z^n$ divergent. □

10. Reelle Reihen mit nicht negativen Gliedern

Die Ergebnisse dieses Abschnittes sind auch bei Konvergenzuntersuchungen beliebiger komplexer Reihen nützlich. Wir bemerken, daß sich die Diskussion der bloßen Konvergenz von Reihen mit den Mitteln dieses Kapitels meist relativ leicht bewerkstelligen läßt, wesentlich schwieriger, oft sogar unmöglich, ist es, die Summe einer Reihe zu ermitteln.

Theorem II-10.1. $(a_n)_{n\geq 1} \subset \mathbb{R}^+$. *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *konvergiert genau dann, wenn die Folge* $(S_n)_{n\geq 1}$ *der Partialsummen beschränkt ist.*

Beweis. Wegen $(a_n) \subset \mathbb{R}^+$ ist die Folge (S_n) monoton wachsend, d.h. für alle $n \in \mathbb{N}$ gilt $S_n \leq S_{n+1}$. Die Behauptung folgt nun aus dem Monotoniekriterium Satz II-5.2. □

Theorem II-10.2 (Vergleichskriterium). *Es seien* $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$, *reelle Folgen und es gelte* $0 \leq a_n \leq b_n$ *für alle* $n \in \mathbb{N}$. *Dann folgt*

i) $\sum_{n=1}^{\infty} b_n$ *ist konvergent* $\Rightarrow$ $\sum_{n=1}^{\infty} a_n$ *ist konvergent.*
Man nennt die Reihe $\sum_{n=1}^{\infty} b_n$ *eine* ***konvergente Majorante*** *für* $\sum_{n=1}^{\infty} a_n$.

ii) $\sum_{n=1}^{\infty} a_n$ *ist divergent* $\Rightarrow$ $\sum_{n=1}^{\infty} b_n$ *ist divergent.*
Man nennt die Reihe $\sum_{n=1}^{\infty} a_n$ *eine* ***divergente Minorante*** $\sum_{n=1}^{\infty} b_n$.

Beweis. Wir bezeichnen die n-ten Partialsummen von $\sum_{k=1}^{\infty} a_k$ mit A_n und von $\sum_{k=1}^{\infty} b_k$ mit B_n. Es folgt

$$(*) \qquad A_n = \sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} b_k = B_n.$$

Beide Folgen (A_n) und (B_n) sind monoton wachsend. Konvergiert die Reihe $\sum_{k=1}^{\infty} b_k$ so ist die Folge (B_n) und folglich auch (A_n) beschränkt. Daraus folgt die Konvergenz von $\sum_{k=1}^{\infty} a_k$ mit Satz II-10.1. Ist die Reihe $\sum_{k=1}^{\infty} a_k$ divergent, so ist die Folge (A_n) und wegen $(*)$ auch (B_n) unbeschränkt. Die Reihe $\sum_{k=1}^{\infty} b_k$ ist somit divergent. □

Bemerkung II-10.3. 1.) Da die Konvergenzeigenschaften einer Reihe nicht vom Verhalten endlich vieler Glieder abhängen, genügt es, wenn die Voraussetzung $0 \leq a_n \leq b_n$ erst ab einem bestimmten Index $N \geq 1$ zutrifft.
2.) Einer divergenten Reihe $\sum_{n=1}^{\infty} a_n$ mit nicht negativen Gliedern kann man auf

Grund der Ausführungen in Abschnitt 7 den uneigentlichen Grenzwert ∞ zuweisen. Man sagt, die Reihe divergiere (aber nicht: konvergiere) nach ∞ und schreibt $\sum_{n=1}^{\infty} a_n = \infty$.
3.) Es sei $(a_n)_{n\geq 1} \subset \mathbb{C}$ und $\sum_{n=1}^{\infty} b_n$ eine konvergente Majorante für die Reihe $\sum_{n=1}^{\infty} |a_n|$. Dann ist auch $\sum_{n=1}^{\infty} a_n$ konvergent.

Theorem II-10.4 (Verdichtungskriterium nach Cauchy). *Es sei $(a_n)_{n\geq 0}$ eine monoton fallende Folge nicht negativer Zahlen. Die Reihe $\sum_{n=0}^{\infty} a_n$ konvergiert genau dann, wenn die „verdichtete" Reihe $\sum_{n=0}^{\infty} 2^n a_{2^n}$ konvergiert.*

(Wir beginnen die Indizierung der Reihenglieder mit 0, um die Darstellung des Beweises etwas zu vereinfachen.)

Beweis. Wir nehmen zuerst an, es konvergiere die Reihe $\sum_{n=0}^{\infty} a_n$. Wegen $a_{n+1} \leq a_n$, $n \in \mathbb{N}$, ergibt sich für $m \in \mathbb{N}$

$$\begin{aligned}\frac{1}{2}\sum_{n=0}^{m} 2^n a_{2^n} &= \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_8 + \dots + 2^{m-1}a_{2^m}\\ &\leq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \dots +\\ &\quad + (a_{2^{m-1}+1} + a_{2^{m-1}+2} + \dots + a_{2^m}) < \sum_{n=0}^{\infty} a_n.\end{aligned}$$

Mit Hilfe des Monotoniekriteriums ergibt sich die Konvergenz der verdichteten Reihe $\sum_{n=0}^{\infty} 2^n a_{2^n}$. Es konvergiere nun die Reihe $\sum_{n=0}^{\infty} 2^n a_{2^n}$. Für $m \in \mathbb{N}$ sei $k \in \mathbb{N}$ so gewählt, daß $2^k > m$ gilt. Es folgt für die m-te Partialsumme von $\sum_{n=0}^{\infty} a_n$

$$\begin{aligned}\sum_{n=0}^{m} a_n &\leq a_0 + a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \dots\\ &\quad + (a_{2^k} + a_{2^k+1} + \dots + a_{2^{k+1}-1})\\ &\leq a_0 + a_1 + 2a_2 + 4a_4 + \dots + 2^k a_{2^k} \leq a_0 + \sum_{k=0}^{\infty} 2^k a_{2^k}.\end{aligned}$$

Das Monotoniekriteriums sichert nun die Konvergenz von $\sum_{n=0}^{\infty} a_n$. Beide Teile des Beweises stützen sich auf die Tatsache, daß in $[2^k+1, 2^{k+1}] \cap \mathbb{N}$ genau 2^k natürliche Zahlen liegen. □

Theorem II-10.5. *Es sei $s \in \mathbb{Q}$. $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergiert genau dann, wenn $s > 1$.*

Beweis. Für $s < 0$ ist $(\frac{1}{n^s})$ nicht konvergent und somit die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$ nach Korollar II-9.3 divergent. Es sei also $s \geq 0$. Nach Satz II-10.4 ist $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergent genau dann, wenn

$$\sum_{n=1}^{\infty} 2^n \frac{1}{(2^n)^s} = \sum_{n=1}^{\infty} 2^{(1-s)n}$$

konvergiert. Die geometrische Reihe $\sum_{n=1}^{\infty} 2^{(1-s)n}$ konvergiert nach Satz II-9.8 genau dann, wenn $2^{1-s} < 1$. Dies ist wegen Satz II-I-7.11 gleichbedeutend mit $s > 1$. □

Die Gültigkeit dieses Satzes wird später auf alle $s \in \mathbb{R}$ ausgedehnt.
In vielen Fällen ist es schwierig, eine konvergente Majorante bzw. eine divergente Minorante für eine Reihe zu finden oder es ist nicht klar, ob nach einer konvergenten Majorante oder einer divergenten Minorante gesucht werden soll. In solchen Fällen kann das Grenzwertkriterium hilfreich sein.

Theorem II-10.6 (Grenzwertkriterium). *Es seien $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ Folgen positiver reeller Zahlen.*

i) *Ist* $\overline{\lim}\,\frac{a_n}{b_n} < \infty$ *und* $\sum_{n=1}^{\infty} b_n$ *konvergent, dann konvergiert auch* $\sum_{n=1}^{\infty} a_n$.
ii) *Ist* $\underline{\lim}\,\frac{a_n}{b_n} > 0$ *und* $\sum_{n=1}^{\infty} b_n$ *divergent, dann divergiert auch* $\sum_{n=1}^{\infty} a_n$.

Beweis. i) Es sei $\overline{\lim}\,\frac{a_n}{b_n} < \beta < \infty$. Nach Satz II-6.7 gibt es $N_0 \in \mathbb{N}$, sodaß $\frac{a_n}{b_n} < \beta$ gilt für alle $n \geq N_0$. Somit folgt $a_n < \beta b_n$ für $n \geq N_0$. Die Konvergenz von $\sum_{n=1}^{\infty} b_n$ zieht nun die Konvergenz von $\sum_{n=1}^{\infty} a_n$ unter Berufung auf das Vergleichskriterium Satz II-10.2 nach sich.
ii) Es sei $\underline{\lim}\,\frac{a_n}{b_n} > \alpha > 0$ und $\sum_{n=1}^{\infty} b_n$ divergent. Nach Satz II-6.7 gibt es $N_1 \in \mathbb{N}$, sodaß $\frac{a_n}{b_n} > \alpha$ zutrifft für alle $n \geq N_1$. Es folgt $a_n > \alpha b_n > 0$ für $n \geq N_1$ und wieder mit Hilfe des Vergleichskriteriums die Divergenz von $\sum_{n=1}^{\infty} a_n$. $\square$

Wir demonstrieren die Anwendung des Grenzwertkriteriums an einem Beispiel:

Beispiel II-10.7. $\sum_{n=1}^{\infty} \frac{\sqrt{n+1}-\sqrt{n}}{n^{2/3}}$ ist konvergent.

Beweis. Wegen $\sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1}+\sqrt{n}}$ verhält sich das n-te Glied der Reihe für große n wie (man sagt dafür auch n ist „**asymptotisch gleich**“) $\frac{1}{2}n^{-\frac{7}{6}}$. Dies legt nahe, als einfachere Vergleichsreihe $\sum_{n=1}^{\infty} n^{-\frac{7}{6}}$ zu wählen. Wegen Satz II-10.5 ist diese Reihe konvergent und aus

$$\frac{a_n}{b_n} = \frac{\sqrt{n+1}-\sqrt{n}}{n^{2/3}} \cdot n^{\frac{7}{6}} = (\sqrt{n+1}-\sqrt{n})\sqrt{n} = \frac{\sqrt{n}}{\sqrt{n+1}+\sqrt{n}}$$

folgt $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{1}{2}$ und mit Hilfe von Satz II-10.6 auch die Konvergenz der ursprünglichen Reihe. $\square$

11. Alternierende Reihen

Wir wenden uns nun einer wichtigen Klasse von Reihen zu, nämlich jenen, deren Glieder reell und abwechselnd positiv und negativ sind. Solche Reihen nennt man **alternierend**.

Theorem II-11.1 (Leibniz Kriterium). *Es sei $(a_n)_{n\geq 1}$, eine Folge reeller, nicht negativer Zahlen. Gilt darüber hinaus*

i) $\forall n \in \mathbb{N}\colon a_{n+1} \leq a_n$,
ii) $\lim_{n\to\infty} a_n = 0$,

dann ist die alternierende Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ *konvergent.*

Beweis. Es sei $S_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$. Wir zeigen, daß die Folgen der geraden und ungeraden Partialsummen konvergieren und ihre Grenzwerte gleich sind, d.h. $\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty} S_{2n+1} = S$. Daraus folgt die Konvergenz der Reihe und

$\sum_{n=1}^{\infty}(-1)^{n+1}a_n = S$ (Übung). Für $n \geq 1$ gilt

$$\begin{aligned} S_{2n+2} - S_{2n} &= -a_{2n+2} + a_{2n+1} \geq 0 \\ S_{2n+2} - S_{2n+1} &= -a_{2n+2} \leq 0 \\ S_{2n+1} - S_{2n-1} &= a_{2n+1} - a_{2n} \leq 0, \end{aligned}$$

somit folgt

$$(*) \qquad S_2 \leq \cdots \leq S_{2n} \leq S_{2n+2} \leq S_{2n+1} \leq S_{2n-1} \leq \cdots \leq S_1.$$

Die Folge $(S_{2n})_{n\geq 1}$ ist daher monoton wachsend und nach oben beschränkt (durch S_1), $(S_{2n+1})_{n\geq 0}$ ist monoton fallend und nach unten beschränkt (durch S_2). Nach dem Monotoniekriterium Satz II-5.2 existieren die Grenzwerte $\lim_{n\to\infty} S_{2n}$ und $\lim_{n\to\infty} S_{2n-1}$. Aus $S_{2n} = S_{2n-1} - a_{2n}$ folgt die Gleichheit der Grenzwerte. □

Aus den Ungleichungen $(*)$ lassen sich auch Schranken für den Fehler ableiten, der sich ergibt, wenn man eine alternierende Reihe durch die n-te Partialsumme approximiert: Es stellt sich heraus, daß der Betrag des Fehlers nicht größer sein kann, als der Betrag des ersten vernachlässigten Gliedes der Reihe.

Korollar II-11.2. $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ erfülle die Voraussetzungen des Leibniz Kriteriums und es sei $S_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$, $S = \sum_{k=1}^{\infty}(-1)^{k+1}a_k$. Es gilt dann für alle $n \in \mathbb{N}$:

$$0 \leq S - S_{2n} \leq a_{2n+1} \text{ und } 0 \leq S_{2n+1} - S \leq a_{2n+2}.$$

Beweis. Aus $(*)$ folgt für alle $n \in \mathbb{N}$:

$$S_{2n} \leq S \leq S_{2n+1}.$$

Subtrahiert man in dieser Ungleichungn S_{2n}, erhält man

$$0 \leq S - S_{2n} \leq S_{2n+1} - S_{2n} = a_{2n+1}.$$

Die zweite Fehlerabschätzung folgt durch Subtraktion von S_{2n+1} von $S_{2n+2} \leq S \leq S_{2n+1}$. □

Beispiel II-11.3. (Alternierende harmonische Reihe, Leibniz Reihe) $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ ist konvergent. Der Nachweis der Konvergenz der alternierenden harmonischen Reihe ist eine einfache Anwendung des Leibniz Kriteriums. Wir können allerdings noch nicht ihre Summe bestimmen. Wir werden später sehen, daß $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = \ln 2 = 0.693147\ldots$. Diese Reihe konvergiert sehr langsam gegen ihre Summe. Korollar II-11.2 zeigt, daß wir bei Vernachlässigung von Rundungsfehlern 10^6 Glieder der Reihe berücksichtigen müßten, um ihre Summe bis auf einen Fehler von 10^{-6} berechnen zu können. Dieses Beispiel zeigt auch, daß die Umkehrung von Satz II-9.5 nicht gilt.

12. Absolute und bedingte Konvergenz von Reihen

Definition II-12.1. *Es sei* $(a_n)_{n\geq 1} \subset \mathbb{C}$.

i) *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *heißt* ***absolut konvergent*** $\underset{Def}{\Leftrightarrow}$ $\sum_{n=1}^{\infty} |a_n|$ *ist konvergent.*

ii) *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *heißt* ***bedingt konvergent*** $\underset{Def}{\Leftrightarrow}$ $\sum_{n=1}^{\infty} a_n$ *ist konvergent aber nicht absolut konvergent.*

Nach Satz Satz II-9.5 ist jede absolut konvergente Reihe auch konvergent. Bei reellen Reihen mit nicht negativen Gliedern sind die beiden Begriffe absolute Konvergenz und Konvergenz natürlich identisch. Die alternierende harmonische Reihe ist ein Beispiel einer bedingt konvergenten Reihe. Jeder Test aus Abschnitt 10 kann zur Untersuchung der absoluten Konvergenz herangezogen werden. Wir ergänzen nun diese Liste mit zwei weiteren Konvergenzkriterien. Allerdings ist man mit beiden Tests nur dann in der Lage, die absolute Konvergenz einer Reihe festzustellen, wenn sich diese asymptotisch wie die geometrische Reihe verhält. Auf Divergenz kann man mit diesen Tests nur bei Reihen schließen, die ohnehin die notwendige Konvergenzbedingung Korollar II-9.3 verletzen.

Theorem II-12.2 (Wurzelkriterium). *Es sei* $(a_n)_{n\geq 1} \subset \mathbb{C}$ *und*

$$\rho = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}.$$

Dann gilt

i) *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *ist absolut konvergent, falls* $\rho < 1$.

ii) *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *divergiert, falls* $\rho > 1$.

Für $\rho = 1$ *ist keine Aussage möglich.*

Beweis. i) Es sei $\rho < 1$ und $\beta \in (\rho, 1)$. Nach Satz II-6.7 gibt es einen Index $N(\beta) \in \mathbb{N}$, sodaß $\sqrt[n]{|a_n|} < \beta$, d.h. $|a_n| < \beta^n$ für alle $n \geq N(\beta)$ zutrifft. Die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert daher nach Satz II-9.8 und dem Vergleichskriterium.
ii) Es sei nun $\rho > 1$. Wegen Satz II-6.7 gilt dann $\sqrt[n]{|a_n|} > 1$ und somit $|a_n| > 1$ für unendlich viele $n \in \mathbb{N}$. Die Folge (a_n) kann daher nicht nach 0 konvergieren. Somit ist die Reihe $\sum_{n=1}^{\infty} a_n$ nach Korollar II-9.3 divergent. □

Das Wurzelkriterium ermöglicht es aber, nicht nur die Konvergenz, sondern auch die **Konvergenzgeschwindigkeit** festzustellen:

Korollar II-12.3. Für die Folge $(a_n)_{n\geq 1} \subset \mathbb{C}$ seien $\beta \in (0, 1)$ und $N(\beta) \in \mathbb{N}$ so gewählt, daß $\sqrt[n]{|a_n|} < \beta$ für alle $n \geq N(\beta)$ gilt. Dann läßt sich der Fehler zwischen der n-ten Partialsumme $S_n = \sum_{k=1}^{n} a_k$ und der Summe $S = \sum_{k=1}^{\infty} a_k$ abschätzen durch

$$|S - S_n| \leq \frac{\beta^{n+1}}{1-\beta}, \quad n \geq N(\beta),$$

Beweis. Für $m > n \geq N(\beta)$ findet man

$$
\begin{aligned}
|S_n - S_m| = |a_{n+1} + \cdots + a_m| &\leq \sum_{k=n+1}^{m} |a_k| \\
&\leq \sum_{k=n+1}^{m} \beta^k < \beta^{n+1} \sum_{k=0}^{\infty} \beta^k = \frac{\beta^{n+1}}{1-\beta}.
\end{aligned}
$$

Die behauptete Abschätzung folgt nun aus $\lim_{m\to\infty} S_m = S$ und Korollar II-4.5. □

Theorem II-12.4 (Quotientenkriterium). *Es sei $(a_n)_{n\geq 1} \subset \mathbb{C} \setminus \{0\}$. Dann gilt*

i) *Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, falls $\overline{\lim}_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$.*

ii) *Die Reihe $\sum_{n=1}^{\infty} a_n$ divergiert für $\underline{\lim}_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$.*

Falls $\overline{\lim}_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \geq 1$ und $\underline{\lim}_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \leq 1$ ist keine Aussage möglich.

Beweis. Satz II-6.10 und Wurzelkriterium. □

Korollar II-12.5. Es sei $(a_n)_{n\geq 1} \subset \mathbb{C}$, $\beta \in (0,1)$ und $N(\beta) \in \mathbb{N}$ so, daß $\frac{|a_{n+1}|}{|a_n|} \leq \beta < 1$ für alle $n \geq N(\beta)$ gilt. Dann läßt sich der Fehler zwischen der n-ten Partialsumme $S_n = \sum_{k=1}^{n} a_n$ und der Summe $S = \sum_{k=1}^{\infty} a_k$ abschätzen durch

$$|S - S_n| \leq \frac{\beta}{1-\beta} |a_n|, \quad n \geq N(\beta).$$

Beweis. Für alle $k \in \mathbb{N}$ und $n \geq N(\beta)$ gilt $|a_{n+k}| \leq \beta^k |a_n|$. Somit erhält man für alle $m > n \geq N(\beta)$

$$|S_m - S_n| \leq \sum_{k=1}^{m-n} |a_{n+k}| \leq \sum_{k=1}^{m-n} \beta^k |a_n| \leq \frac{\beta}{1-\beta} |a_n|.$$

Die Behauptung folgt nun wie im Beweis von Korollar II-12.3. □

Bemerkung II-12.6. (1) Die strikten Ungleichungen im Wurzel- und Quotientenkriterium können nicht zu $\leq$ abgeschwächt werden. Als Beispiel betrachten wir die divergente harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ und die konvergente Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$, für welche jeweils $\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$ gilt.

(2) Der Beweis des Quotientenkriteriums zeigt, daß jede Reihe, auf die das Quotientenkriterium anwendbar ist, auch mit dem Wurzelkriterium untersucht werden kann, obgleich die Rechnung dann erheblich komplizierter sein kann. Das Wurzelkriterium ist tatsächlich umfassender als das Quotientenkriterium, wie die folgenden beiden Beispiele zeigen:

(3) Es sei $\sum_{n=1}^{\infty} a_n$ mit $a_n = 2^{-n}$, falls n gerade ist, und $a_n = 3^{-n}$, falls n ungerade ist. Aus $\sum a_n < \sum 2^{-n}$ folgt die Konvergenz der Reihe. Es gilt $\overline{\lim} \sqrt[n]{a_n} = \frac{1}{2} < 1$ und $\overline{\lim} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{a_{2n}}{a_{2n-1}} = \lim_{n\to\infty} \frac{3^{2n-1}}{2^{2n}} = \infty$. Das Wurzelkriterium deckt also im Gegensatz zum Quotientenkriterium die Konvergenz der Reihe auf. Dies zeigt auch, daß $\underline{\lim}$ in II-12.4-ii) nicht durch $\overline{\lim}$ ersetzt werden kann. Man beachte ferner

$$\underline{\lim}_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{a_{2n+1}}{a_{2n}} = \lim_{n\to\infty} \frac{2^{2n}}{3^{2n+1}} = \frac{1}{3} \lim_{n\to\infty} (\frac{2}{3})^{2n} = 0.$$

(4) Es sei nun $\sum_{n=1}^{\infty} a_n$ mit $a_n = 2^n$ falls n gerade ist, und $a_n = 2^{-n}$, falls n ungerade ist. Die Reihe ist natürlich divergent. Ferner gilt $\overline{\lim} \sqrt[n]{a_n} = 2$ und $\underline{\lim} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{a_{2n+1}}{a_{2n}} = 0$. Wiederum ist das Wurzelkriterium erfolgreicher als das Quotientenkriterium. Ferner gilt $\underline{\lim} \sqrt[n]{a_n} = \frac{1}{2} < 1$. Gemeinsam mit

der letzten Beobachtung in 3) belegt dies, daß $\overline{\lim}$ in II-12.2 und II-12.4-i nicht durch $\underline{\lim}$ ersetzt werden kann.

Wir führen noch zwei weitere Konvergenzkriterien an. Dazu benötigen wir folgendes Hilfsmittel:

Lemma II-12.7 (Abelsche partielle Summation). *Es seien $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ komplexe Folgen und $S_n = \sum_{k=1}^n a_k$, $n \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$ gilt dann*

$$\sum_{k=1}^n a_k b_k = \sum_{k=1}^n S_k(b_k - b_{k+1}) + S_n b_n.$$

Beweis. Setzen wir $S_0 = 0$, so gilt $a_k = S_k - S_{k-1}$ für alle $k \in \mathbb{N}$ und somit

$$\begin{aligned}\sum_{k=1}^n a_k b_k &= \sum_{k=1}^n (S_k - S_{k-1}) b_k = \sum_{k=1}^n S_k b_k - \sum_{k=1}^n S_{k-1} b_k \\ &= \sum_{k=1}^n S_k b_k - \sum_{k=0}^{n-1} S_k b_{k+1} = \sum_{k=1}^{n-1} S_k(b_k - b_{k+1}) + S_n b_n\end{aligned}$$

□

Eine unmittelbare Folgerung ist

Lemma II-12.8. *Es seien $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$ komplexe Folgen und $S_n = \sum_{k=1}^n a_k$ für $n \in \mathbb{N}$. Konvergieren die Folge $(S_n b_{n+1})_{n\geq 1}$ und die Reihe $\sum_{n=1}^\infty S_n(b_n - b_{n+1})$, so ist auch die Reihe $\sum_{n=1}^\infty a_n b_n$ konvergent.*

Theorem II-12.9 (Abelsches Kriterium). *Ist die komplexe Reihe $\sum_{n=1}^\infty a_n$ konvergent und die reelle Folge $(b_n)_{n\geq 1}$ monoton und beschränkt, so konvergiert $\sum_{k=1}^\infty a_k b_k$.*

Beweis. Es gilt $\sum_{k=1}^n (b_k - b_{k+1}) = b_1 - b_{n+1}$. Somit ist die Reihe $\sum_{k=1}^\infty (b_k - b_{k+1})$ konvergent und zwar sogar absolut konvergent, da für alle $k \in \mathbb{N}$ wegen der Monotonie von (b_n) stets $b_k - b_{k+1} \geq 0$ oder stets $b_k - b_{k+1} \leq 0$ gilt. Wegen der Konvergenz von $\sum_{n=1}^\infty a_n$ ist die Folge $(S_n)_{n\geq 1}$, $S_n = \sum_{k=1}^n a_k$ beschränkt, d.h. es gibt $M > 0$ sodaß $|S_n| \leq M$ für alle $n \in \mathbb{N}$ gilt. Daraus folgt mit

$$\left|\sum_{k=1}^\infty S_k(b_k - b_{k+1})\right| \leq \sum_{k=1}^\infty |S_k||b_k - b_{k+1}| \leq M \sum_{k=1}^\infty |b_k - b_{k+1}|$$

die absolute Konvergenz von $\sum_{k=1}^\infty S_k(b_k - b_{k+1})$. Da auch $(S_n b_{n+1})_{n\geq 1}$ konvergiert, folgt die Konvergenz mit Lemma II-12.8. □

Theorem II-12.10 (Dirichlet Kriterium). *Sind die Partialsummen der komplexen Reihe $\sum_{n=1}^\infty a_n$ beschränkt und strebt die reelle Folge $(b_n)_{n\geq 1}$ monoton nach Null, so konvergiert die Reihe $\sum_{n=1}^\infty a_n b_n$.*

Beweis. Analog zu Beweis von Satz II-12.9. □

Ehe wir den Unterschied zwischen absolut konvergenten Reihen und bedingt konvergenten Reihen darlegen, bemerken wir, daß bei *konvergenten* Reihen die Glieder beliebig durch Klammern zusammengefaßt werden können, da dies den Übergang auf eine Teilfolge von Partialsummen bedeutet. Schon vorhandene Klammerungen in einer konvergenten Reihe dürfen jedoch nur dann weggelassen werden, wenn die entstehende Reihe wieder konvergiert. Das Weglassen der Klammern in der konvergenten Reihe $(1-1)+(1-1)+(1-1)+\ldots$ führt zum Beispiel auf die divergente Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}$. Durch das Setzen von Klammern wird die Reihenfolge der Glieder der Reihe nicht verändert. Wir untersuchen nun den Einfluß des Vertauschens von unendlich vielen Gliedern auf die Summe der Reihe.

Definition II-12.11. *Es sei* $(a_n)_{n\geq 1}\subset\mathbb{C}$ *und* $\sigma\colon\mathbb{N}\to\mathbb{N}$ *eine Bijektion. Die Reihe* $\sum_{k=1}^{\infty}a_{\sigma(k)}$ *heißt* ***Umordnung*** *der Reihe* $\sum_{k=1}^{\infty}a_k$.

Beispiel II-12.12. Wir betrachten folgende Umordnung der bedingt konvergenten alternierenden harmonischen Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}=\alpha$.

$$\sum_{n=0}^{\infty}a_n=1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\frac{1}{5}-\frac{1}{10}-\frac{1}{12}+\frac{1}{7}-\ldots.$$

Das allgemeine Glied der Reihe ist gegeben durch

$$\begin{aligned} a_{3k}&=\frac{1}{2k+1},\\ a_{3k+1}&=-\frac{1}{4k+2}, \qquad k=0,1,2,\ldots\\ a_{3k+2}&=-\frac{1}{4(k+1)}. \end{aligned}$$

(Der Einfachheit halber wurde mit der Indizierung der Reihe bei 0 begonnen.) Wir betrachten die Partialsummen $(k\geq 1)$

$$\begin{aligned} S_{3k-1}&=\sum_{j=0}^{k-1}(a_{3j}+a_{3j+1}+a_{3j+2})\\ &=\sum_{j=0}^{k-1}(\frac{1}{2j+1}-\frac{1}{4j+2}-\frac{1}{4(j+1)})=\frac{1}{2}\sum_{j=0}^{k-1}(\frac{1}{2j+1}-\frac{1}{2(j+1)})\\ &=\frac{1}{2}(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots-\frac{1}{2k}). \end{aligned}$$

Daraus folgt $\lim_{k\to\infty}S_{3k-1}=\frac{1}{2}\alpha$. Ferner gilt für alle $k\geq 1$

$$\begin{aligned} S_{3k}&=S_{3k-1}+a_{3k}=S_{3k-1}+\frac{1}{2k+1},\\ S_{3k+1}&=S_{3k-1}+a_{3k}+a_{3k+1}=S_{3k-1}+\frac{1}{2k+1}-\frac{1}{4k+2}, \end{aligned}$$

und somit auch $\lim_{k\to\infty}S_{3k}=\lim_{k\to\infty}S_{3k+1}=\lim_{k\to\infty}S_{3k-1}=\frac{1}{2}\alpha$. Wir erhalten also, daß die umgeordnete Reihe nicht gegen α sondern gegen $\frac{1}{2}\alpha$ konvergiert.

Wir merken an, daß die Summe einer absolut konvergenten Reihe unabhängig ist von der Reihenfolge ihrer Glieder. Absolut konvergente Reihen verhalten sich also sehr ähnlich den endlichen Summen.

Theorem II-12.13. *Es sei $(a_k)_{k\geq 1} \subset \mathbb{C}$ und die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent. Dann ist auch jede Umordnung $\sum_{k=1}^{\infty} a_{\sigma(k)}$ absolut konvergent und es gilt*

$$\sum_{k=1}^{\infty} a_{\sigma(k)} = \sum_{k=1}^{\infty} a_k.$$

Beweis. Wir verwenden folgende Bezeichnungen: $S_n = \sum_{k=1}^{n} a_k$, $U_n = \sum_{k=1}^{n} a_{\sigma(k)}$. Wegen der absoluten Konvergenz von $\sum_{k=1}^{\infty} a_k$ existiert nach Satz II-9.4 zu jedem $\varepsilon > 0$ ein Index N_0, sodaß

$$\sum_{k=N_0+1}^{\infty} |a_k| < \varepsilon$$

zutrifft. Zu N_0 bestimmen wir einen weiteren Index $K > N_0$ so, daß

$$(*) \qquad \{1, \dots, N_0\} \subset \{\sigma(1), \sigma(2), \dots, \sigma(K)\}.$$

Wegen $(*)$ treten für alle $n \geq K$ die Zahlen $a_1, \dots, a_{N_0}$ sowohl in den Partialsummen S_n als auch in U_n auf und fallen aus der Differenz $S_n - U_n$ heraus. Somit hat $S_n - U_n$ die Form

$$S_n - U_n = \delta_{N_0+1} a_{N_0+1} + \dots + \delta_{N_0+p} a_{N_0+p}, \quad \delta_j \in \{-1, 0, 1\}$$

für ein geeignet gewähltes $p \in \mathbb{N}$. Dies hat für alle $n \in \mathbb{N}$, $n \geq K$,

$$|S_n - U_n| \leq \sum_{k=N_0+1}^{\infty} |a_k| < \varepsilon$$

zur Folge. Somit gilt $\lim(S_n - U_n) = 0$ und die Folge (U_n) ist wegen $U_n = U_n - S_n + S_n$ konvergent mit

$$\lim U_n = \lim(S_n - U_n) + \lim S_n = \lim S_n.$$

□

Wir haben in Beispiel II-12.12 bereits demonstriert, daß bedingt konvergente Reihen sehr sensibel sein können in Bezug auf die Anordnung ihrer Glieder. Wir werden zeigen, daß durch Umordnung die Konvergenz sogar zerstört werden kann:

Theorem II-12.14 (Umordnungssatz von Riemann). *Es sei $\sum_{n=1}^{\infty} a_n$ eine bedingt konvergente Reihe reeller Zahlen und $\alpha, \beta \in \bar{\mathbb{R}}$: $-\infty \leq \alpha \leq \beta \leq \infty$. Dann gibt es eine Umordnung $\sum a_{\sigma(k)}$ derart, daß deren Partialsummen $S_n = \sum_{k=1}^{n} a_{\sigma(k)}$*

$$\underline{\lim}_{n\to\infty} S_n = \alpha \quad \textit{und} \quad \overline{\lim}_{n\to\infty} S_n = \beta$$

erfüllen.

Bemerkung II-12.15. Insbesondere folgt, daß durch Umordnen einer bedingt konvergenten Reihe deren Summe beliebig verändert werden kann. Wählt man $\alpha = \beta = \infty$, gibt es sogar eine Umordnung der Reihe, welche nach ∞ divergiert.

Beweisidee des Umordnungssatzes. Wir demonstrieren die Beweisidee an einer etwas einfacheren Situation und zeigen: Für jedes ξ in $\mathbb{R}$ existiert eine Umordnung der Reihe, welche gegen ξ konvergiert. Zunächst überlegen wir, daß beide Teilreihen, die entstehen, wenn man nur die positiven oder nur die negativen Glieder berücksichtigt, divergieren. Für $n \in \mathbb{N}$ setzen wir

$$a_n^+ = \frac{1}{2}(|a_n| + a_n) = \begin{cases} a_n & \text{falls } a_n > 0 \\ 0 & \text{falls } a_n \leq 0 \end{cases}$$

$$a_n^- = \frac{1}{2}(|a_n| - a_n) = \begin{cases} 0 & \text{falls } a_n > 0 \\ -a_n & \text{falls } a_n \leq 0. \end{cases}$$

Es gilt dann $a_n = a_n^+ - a_n^-$ und $|a_n| = a_n^+ + a_n^-$, $a_n^+ \geq 0$, $a_n^- \geq 0$. Wegen der Konvergenz von $\sum_{n=1}^\infty a_n$ und $a_n = a_n^+ - a_n^-$ können die beiden Reihen $\sum_{n=1}^\infty a_n^+$ und $\sum_{n=1}^\infty a_n^-$ nur beide konvergieren oder beide divergieren. Wären beide Reihen konvergent, dann müßte wegen $|a_n| = a_n^+ + a_n^-$ die Reihe $\sum_{n=1}^\infty a_n$ absolut konvergieren. Beide Reihen $\sum_{n=1}^\infty a_n^+$ und $\sum_{n=1}^\infty a_n^-$ divergieren also nach ∞ und es gilt überdies $\lim_{n\to\infty} a_n^+ = \lim_{n\to\infty} a_n^- = 0$. Man summiert nun solange nicht negative Reihenglieder a_n^+ auf, bis deren Summe zum ersten Male ξ übertrifft, anschließend subtrahiert man solange Reihenglieder a_n^-, bis die resultierende Summe zum ersten Male kleiner als ξ ausfällt, usw. $\square$

13. Doppelreihen

Definition II-13.1. *$(x_{nm})_{n,m\geq 1} \subset \mathbb{C}$ sei eine Doppelfolge.*

i) *$S_{nm} = \sum_{i=1}^n \sum_{j=1}^m x_{ij}$, $n, m \in \mathbb{N}$ heißt* ***(n,m)-te Partialsumme****.*
Die Doppelfolge $(S_{nm})_{n,m\geq 1} \subset \mathbb{C}$ heißt ***Doppelreihe****, $\sum_{n,m=1}^\infty x_{nm}$.*

ii) *Die Doppelreihe $\sum_{n,m=1}^\infty x_{nm}$ ist* **konvergent** *gegen $s \underset{Def}{\Leftrightarrow} s = \lim_{n,m\to\infty} S_{nm}$.*
Man schreibt auch $s = \sum_{n,m=1}^\infty x_{nm}$.

iii) *Die Doppelreihe $\sum_{n,m=1}^\infty x_{nm}$ ist* **absolut konvergent** *$\underset{Def}{\Leftrightarrow} \sum_{n,m=1}^\infty |x_{nm}|$ ist konvergent.*

Die Doppelreihe $\sum_{n,m=1}^\infty x_{nm}$ konvergiert somit nach s genau dann, wenn für jedes $\varepsilon > 0$ ein Index $N(\varepsilon)$ gefunden werden kann, daß $|S_{nm} - s| < \varepsilon$ für alle $n, m \geq N(\varepsilon)$ gilt. Die Resultate über Doppelfolgen übertragen sich somit sinngemäß auf Doppelreihen. In diesem Abschnitt wollen wir uns allerdings auf absolut konvergente Doppelreihen beschränken.

Lemma II-13.2. *Eine Doppelreihe $\sum_{i,j=1}^\infty z_{ij}$ ist absolut konvergent genau dann, wenn $\{\sum_{i=1}^n \sum_{j=1}^m |z_{ij}| \colon n, m \in \mathbb{N}\}$ beschränkt ist.*

Beweis. „$\Leftarrow$" Es sei $x = \sup\{\sum_{i=1}^{n}\sum_{j=1}^{m}|z_{ij}|\colon n, m \in \mathbb{N}\} < \infty$. Zu jedem $\varepsilon > 0$ gibt es $(n_\varepsilon, m_\varepsilon) \in \mathbb{N} \times \mathbb{N}$ mit

$$x - \varepsilon < \sum_{i=1}^{n_\varepsilon}\sum_{j=1}^{m_\varepsilon}|z_{ij}|.$$

Setzt man $N(\varepsilon) = \max\{n_\varepsilon, m_\varepsilon\}$ folgt für alle $n, m \geq N(\varepsilon)$

$$|x - \sum_{i=1}^{n}\sum_{j=1}^{m}|z_{ij}|| = x - \sum_{i=1}^{n}\sum_{j=1}^{m}|z_{ij}| \leq x - \sum_{i=1}^{n_\varepsilon}\sum_{j=1}^{m_\varepsilon}|z_{ij}| < \varepsilon,$$

d.h. $\sum_{i,j=1}^{\infty} z_{ij}$ ist absolut konvergent. Die Umkehrung ist offensichtlich. □

Ähnlich wie bei Doppelfolgen, kann man zu einer Doppelreihe $\sum_{n,m=1}^{\infty} z_{nm}$ auch die iterierten Reihen $\sum_{n=1}^{\infty}(\sum_{m=1}^{\infty} z_{nm}) = \lim_{n\to\infty}(\lim_{m\to\infty} S_{nm})$ (Reihe der Zeilensummen) bzw. $\sum_{m=1}^{\infty}(\sum_{n=1}^{\infty} z_{nm})$ (Reihe der Spaltensummen) bilden. Da letztere oft wesentlich einfacher zu berechnen sind, ergibt sich die Frage nach ihrer Beziehung zur Doppelreihe.

Theorem II-13.3. *Es sei $(a_{nm})_{n,m\geq 1} \subset \mathbb{R}^+$. Konvergiert eine der beiden iterierten Reihen $\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_{ij})$ bzw. $\sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} a_{ij})$, dann sind sowohl die andere iterierte Reihe, als auch die Doppelreihe $\sum_{i,j=1}^{\infty} a_{ij}$ konvergent und es gilt*

$$\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_{ij}) = \sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} a_{ij}) = \sum_{i,j=1}^{\infty} a_{ij}.$$

Beweis. Es sei etwa die iterierte Reihe $\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_{ij})$ konvergent, d.h. es existiert $S = \lim_{n\to\infty}(\lim_{m\to\infty} S_{nm})$ mit $S_{nm} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}$. Für alle $n, m \in \mathbb{N}$ gilt $S_{nm} \leq \lim_{k\to\infty} S_{nk} \leq S$. S ist daher eine obere Schranke für $A = \{S_{nm}\colon n, m \in \mathbb{N}\}$. Nach Lemma II-13.2 konvergiert daher die Doppelreihe $\sum_{i,j=1}^{\infty} a_{ij}$ gegen $\sup A$. Aus der Ungleichung $S_{nm} \leq \sup A$ folgt aber auch die Existenz von $\lim_{n\to\infty} S_{nm}$ (Monotoniekriterium) für alle $m \in \mathbb{N}$. Die Behauptung folgt nun aus Korollar II-8.6 □

Theorem II-13.4 (Doppelreihensatz). *Es sei $(a_{nm})_{n,m\geq 1} \subset \mathbb{C}$. Konvergiert eine der beiden iterierten Reihen $\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty}|a_{ij}|)$ bzw. $\sum_{j=1}^{\infty}(\sum_{i=1}^{\infty}|a_{ij}|)$, dann sind sowohl die andere iterierte Reihe, als auch die Doppelreihe $\sum_{i,j=1}^{\infty} a_{ij}$ absolut konvergent und es gilt*

$$\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty} a_{ij}) = \sum_{j=1}^{\infty}(\sum_{i=1}^{\infty} a_{ij}) = \sum_{i,j=1}^{\infty} a_{ij}$$

Beweis. Es sei z.B. die Reihe der Zeilensummen $\sum_{i=1}^{\infty}(\sum_{j=1}^{\infty}|a_{ij}|)$ konvergent. Nach Satz II-13.3 ist die Doppelreihe $\sum_{i,j=1}^{\infty} a_{ij}$ absolut konvergent und es existiert auch $\sum_{j=1}^{\infty}(\sum_{i=1}^{\infty}|a_{ij}|)$. Somit sind auch für alle $j \in \mathbb{N}$ die Reihen $\sum_{i=1}^{\infty} a_{ij}$ absolut konvergent. Bezeichnet man mit S_{nm} die (n,m)-te Partialsumme der Doppelreihe, $S_{nm} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}$, $n,m \in \mathbb{N}$, existieren der Doppellimes $\lim_{n,m\to\infty} S_{nm}$ und für alle $n,m \in \mathbb{N}$ auch die einfachen Limiten $\lim_{j\to\infty} S_{nj}$ und $\lim_{i\to\infty} S_{im}$. Die Behauptung folgt nun aus Korollar II-8.6. □

THEOREM II-13.5. *Es sei $(a_{nm})_{n,m\geq 1} \subset \mathbb{C}$ und die Doppelreihe $\sum_{n,m=1}^{\infty} a_{nm}$ sei absolut konvergent. Dann konvergieren auch die beiden iterierten Reihen absolut und es gilt*

$$\sum_{n,m=1}^{\infty} a_{nm} = \sum_{n=1}^{\infty}\Big(\sum_{m=1}^{\infty} a_{nm}\Big) = \sum_{m=1}^{\infty}\Big(\sum_{n=1}^{\infty} a_{nm}\Big).$$

BEWEIS. Nach Lemma II-13.2 gibt es $k > 0$, sodaß

$$\forall n,m \in \mathbb{N}\colon \sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}| \leq k$$

gilt. Daraus folgt für jedes $n \in \mathbb{N}$ auch $\sum_{i=1}^{n}\sum_{j=1}^{\infty} |a_{ij}| \leq k$ und schließlich $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}| \leq k$. Die Behauptung ergibt sich nun aus Satz II-13.4. □

Abschließend zeigen wir, daß die Summe einer absolut konvergenten Doppelreihe unabhängig ist von der Reihenfolge ihrer Glieder.

DEFINITION II-13.6. *Es sei $\sum_{n,m=1}^{\infty} a_{nm}$ eine komplexe Doppelreihe und $\varphi\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ eine Bijektion. Dann heißt die Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$* ***Anordnung*** *der Doppelreihe (in eine einfache Reihe).*

LEMMA II-13.7. *Es sei $\sum_{n,m=1}^{\infty} a_{nm}$ eine konvergente reelle Doppelreihe mit nicht-negativen Gliedern und $\varphi\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ eine Bijektion. Dann gilt*

$$\sum_{n,m=1}^{\infty} a_{nm} = \sum_{k=1}^{\infty} a_{\varphi(k)}.$$

BEWEIS. Es sei $\alpha = \sum_{n,m=1}^{\infty} a_{nm} = \sup\{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\colon n,m \in \mathbb{N}\}$ (vgl. Lemma II-13.2). Zu $\varepsilon > 0$ gibt es somit $N_1(\varepsilon), N_2(\varepsilon) \in \mathbb{N}$ mit

$$\alpha - \varepsilon < \sum_{i=1}^{N_1(\varepsilon)}\sum_{j=1}^{N_2(\varepsilon)} a_{ij} \leq \alpha.$$

Da φ surjektiv ist, existiert $K(\varepsilon) \in \mathbb{N}$ mit

$$\mathcal{I}(N_1(\varepsilon)) \times \mathcal{I}(N_2(\varepsilon)) \subset \varphi\big(\mathcal{I}(K(\varepsilon))\big).$$

(Zur Erinnerung: $\mathcal{I}(n) = \{K \in \mathbb{N}\colon 1 \leq K \leq n\}$). Wegen $a_{nm} \geq 0$, $n,m \in \mathbb{N}$, ergibt sich

$$\sum_{k=1}^{K(\varepsilon)} a_{\varphi(k)} \geq \sum_{i=1}^{N_1(\varepsilon)}\sum_{j=1}^{N_2(\varepsilon)} a_{ij}.$$

Für alle $k \geq K(\varepsilon)$ gilt daher auch

$$\alpha - \varepsilon < \sum_{i=1}^{N_1(\varepsilon)} \sum_{j=1}^{N_2(\varepsilon)} a_{ij} \leq \sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)} \leq \sum_{\ell=1}^{k} a_{\varphi(\ell)} \leq \alpha.$$

□

THEOREM II-13.8. *Es sei* $\sum_{n,m=1}^{\infty} a_{nm}$ *eine komplexe, absolut konvergente Doppelreihe und* $\varphi\colon \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ *eine Bijektion. Dann gilt*

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{n,m=1}^{\infty} a_{nm}.$$

BEWEIS. Nach II-13.7 ist die Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ absolut konvergent. Es ist somit nur die Gleichheit der Limiten zu zeigen. Es sei $\alpha = \sum_{n,m=1}^{\infty} a_{nm}$ und $s = \sum_{k=1}^{\infty} a_{\varphi(k)}$. Zu $\varepsilon > 0$ existieren somit Indizes $N(\varepsilon)$ und $K(\varepsilon)$ mit

$$\forall n, m \in \mathbb{N}\colon n \geq N(\varepsilon) \wedge m \geq N(\varepsilon) \Rightarrow \Big|\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} - \alpha\Big| < \frac{\varepsilon}{3}$$

und

$$\Big|\sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)} - s\Big| < \frac{\varepsilon}{3}, \qquad \sum_{\ell=K(\varepsilon)+1}^{\infty} |a_{\varphi(\ell)}| < \frac{\varepsilon}{3}.$$

Wir wählen nun $N_1(\varepsilon) \in \mathbb{N}$ so, daß $N_1(\varepsilon) \geq N(\varepsilon)$ und

$$\varphi(\mathfrak{I}(K(\varepsilon)) \subset \mathfrak{I}(N_1(\varepsilon)) \times \mathfrak{I}(N_1(\varepsilon)).$$

Jeder Term der Summe $\sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)}$ kommt dann auch als Summand in $\sum_{i=1}^{N_1(\varepsilon)} \sum_{j=1}^{N_1(\varepsilon)} a_{ij}$ vor. Somit gilt

$$\Big|\sum_{i=1}^{N_1(\varepsilon)} \sum_{j=1}^{N_1(\varepsilon)} a_{ij} - \sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)}\Big| \leq \sum_{\ell=K(\varepsilon)+1}^{\infty} |a_{\varphi(\ell)}| < \frac{\varepsilon}{3}.$$

Daraus folgt schließlich

$$\begin{aligned} |s - \alpha| \leq & \Big|s - \sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)}\Big| + \Big|\sum_{\ell=1}^{K(\varepsilon)} a_{\varphi(\ell)} - \sum_{i=1}^{N_1(\varepsilon)} \sum_{j=1}^{N_1(\varepsilon)} a_{ij}\Big| \\ & + \Big|\sum_{i=1}^{N_1(\varepsilon)} \sum_{j=1}^{N_1(\varepsilon)} a_{ij} - \alpha\Big| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

d.h. $s = \alpha$. □

In der Sprechweise der Definition II-13.6 bedeutet der vorige Satz, daß man eine absolut konvergente Doppelreihe auf beliebige Weise in eine einfache Reihe anordnen kann und daß diese Reihe gegen die Summe der Doppelreihe konvergiert. Ohne Beweis teilen wir noch folgenden Satz mit:

THEOREM II-13.9. *Es sei* $\sum_{k=1}^{\infty} a_{\varphi(k)}$ *eine beliebige Anordnung (in eine einfache Reihe) einer Doppelreihe* $\sum_{i,j=1}^{\infty} a_{ij}$. *Ist die Reihe* $\sum_{k=1}^{\infty} a_{\varphi(k)}$ *absolut konvergent, dann ist auch die Doppelreihe* $\sum_{i,j=1}^{\infty} a_{ij}$ *absolut konvergent und es gilt*

$$\sum_{i,j=1}^{\infty} a_{ij} = \sum_{k=1}^{\infty} a_{\varphi(k)}.$$

KAPITEL III

Stetige Funktionen

1. Der Grenzwert einer Funktion

Der Grenzwert ist einer der zentralen Begriffe der Analysis. Dazu benötigen wir folgendes Konzept. In diesem Kapitel bezeichnen X und Y normierte Räume über demselben Skalarenkörper $\mathbb{K}$.

Definition III-1.1. *$D \subset X$ sei nicht leer.*
i) $x_0 \in X$ heißt ***Häufungspunkt*** *von D*

$$\forall U \in \mathcal{U}(x_0)\colon (U \setminus \{x_0\}) \cap D \neq \emptyset.$$

ii) $x_0 \in D$ heißt ***isolierter Punkt*** *von D*

$$\exists U \in \mathcal{U}(x_0)\colon (U \setminus \{x_0\}) \cap D = \emptyset.$$

iii) $\bar{D} = D \cup \{x \in X\colon x \text{ ist Häufungspunkt von } D\}$ heißt ***abgeschlossene Hülle*** *von D.*

Da jede Umgebung von x_0 eine hinreichend kleine Kugel $K(x_0, r)$ enthält, kann man in der Definition des Häufungspunktes $U \in \mathcal{U}(x_0)$ durch Kugeln $K(x_0, r_n)$ mit $\lim_{n\to\infty} r_n = 0$ ersetzen. Jede dieser Kugeln enthält ein Element $x_n \in D$ mit $x_n \neq x_0$ und $\|x_n - x_0\| < r_n$, $n \in \mathbb{N}$. Somit folgt $\lim_{n\to\infty} x_n = x_0$. Gibt es umgekehrt eine Folge $(x_n) \subset D\setminus\{x_0\}$, welche gegen x_0 konvergiert, dann ist natürlich x_0 ein Häufungspunkt von D. Es gilt also

Lemma III-1.2. *$D \subset X$ sei nicht leer. Äquivalent sind folgende Aussagen:*
i) $x_0 \in X$ ist ein Häufungspunkt von D.
ii) Es gibt eine Folge $(x_n) \subset D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$.

Bemerkung III-1.3. Häufungspunkte von D sind Stellen, in deren Nähe sich die Elemente von D anschaulich häufen. Eine endliche Menge kann daher keine Häufungspunkte besitzen. Sie besteht nur aus isolierten Punkten. Man beachte, daß ein Häufungspunkt von D nicht Element der Menge D sein muß.

Beispiel III-1.4. Es sei $-\infty < a < b < \infty$ und D eines der Intervalle (a, b), $(a, b]$, $[a, b)$ oder $[a, b]$. Dann ist $[a, b]$ die Menge der Häufungspunkte von D, $\bar{D} = [a, b]$.

Beispiel III-1.5. $x_0 \in X$ is ein Häufungspunkt von $D = K(y_0, r)$ genau dann, wenn $x_0 \in \bar{K}(y_0, r)$. Die Notation für die abgeschlossenen Kugel in Definition II-1.6 ist somit konsistent mit der Bezeichnung der abgeschlossenen Hülle der offenen Kugel.

Es sei $x_0 \in \bar{K}(y_0, r)$. Wir betrachten vorerst den Fall $x_0 \neq y_0$ und setzen $x_n = r_n y_0 + (1 - r_n) x_0$, $r_n \in (0,1)$ und $\lim_{n\to\infty} r_n = 0$. Dann folgt

$$x_n - x_0 = r_n(y_0 - x_0), \qquad x_n - y_0 = (1 - r_n)(x_0 - y_0),$$

also $x_n \neq x_0$, $n \in \mathbb{N}$, und

$$\begin{aligned} \|x_n - y_0\| &= (1 - r_n)\|x_0 - y_0\| < \|x_0 - y_0\| \leq r, \\ \|x_n - x_0\| &= r_n\|y_0 - x_0\| \leq r_n r < r. \end{aligned}$$

Dies zeigt $x_n \in K(y_0, r)$, $n \in \mathbb{N}$ und $\lim_{n\to\infty} x_n = x_0$. Im Fall $x_0 = y_0$ wähle man $\xi_0 \in K(y_0, r)$ und wende die vorausgehende Argumentation auf

$$x_n = y_0 + r_n(\xi_0 - y_0) = r_n\xi_0 + (1 - r_n)y_0$$

an. Jedes Element der abgeschlossenen Kugel $\bar{K}(y_0, r)$ ist somit ein Häufungspunkt von $K(y_0, r)$.
Umgekehrt sei nun $x_0 \notin \bar{K}(y_0, r)$, also $\eta = \|x_0 - y_0\| > r$. Wir behaupten:

$$K(x_0, \eta - r) \cap \bar{K}(y_0, r) = \emptyset.$$

Somit kann x_0 kein Häufungspunkt von D sein. Wäre dieser Durchschnitt nicht leer, gäbe es ein Element $\xi \in X$ mit

$$\|\xi - x_0\| < \eta - r \quad \text{und } \|\xi - y_0\| \leq r.$$

Dies hätte den Widerspruch

$$\|x_0 - y_0\| \leq \|x_0 - \xi\| + \|\xi - y_0\| < \eta - r + r = \eta$$

zur Folge.

Beispiel III-1.6. Es sei $X = \mathbb{R}$ und $D = \mathbb{Q}$. Wir zeigen: $\bar{\mathbb{Q}} = \mathbb{R}$.
Jede rationale Zahl ist natürlich Häufungspunkt von $\mathbb{Q}$. Es sei nun $x_0 \in \mathbb{R} \setminus \mathbb{Q}$. Wir setzen $x_1 = \lfloor x_0 \rfloor \in \mathbb{Q}$. Wegen $x_0 \notin \mathbb{Q}$ ist $x_1 < x_0$ (Satz I-6.10). Mit Satz I-6.11 schließen wir auf die Existenz von $x_2 \in \mathbb{Q}$ mit $\frac{1}{2}(x_0 + x_1) < x_2 < x_0$ und $|x_2 - x_0| < \frac{1}{2}|x_1 - x_0|$. Induktiv fortfahrend konstruiert man eine Folge rationaler Zahlen (x_n) mit $\frac{1}{2}(x_0 + x_{n-1}) < x_n < x_0$ und $|x_n - x_0| < (\frac{1}{2})^{n-1}|x_1 - x_0|$, also $\lim_{n\to\infty} x_n = x_0$.

Theorem III-1.7. *Eine Menge $D \subset X$ ist abgeschlossen genau dann, wenn $D = \bar{D}$.*

Beweis. Die Behauptung trifft für die leere Menge zu: wegen $\complement\emptyset = X$ ist $\emptyset$ abgeschlossen und natürlich gilt auch $\bar{\emptyset} = \emptyset$.
Es sei also $D \subset X$ nicht leer und abgeschlossen. Wir zeigen: $\forall x_0 \in \complement D\colon x_0 \notin \bar{D}$. Dann muß $\bar{D} \subset D$ gelten, woraus $\bar{D} = D$ folgt. Da $\complement D$ offen ist, existiert $r > 0$, so daß $K(x_0, r) \subset \complement D$. Dann kann es aber keine Folge $(x_n) \subset D \setminus \{x_0\}$ geben, welche gegen x_0 konvergiert. Folglich kann x_0 kein Häufungspunkt von D sein.
Umgekehrt sei nun $\bar{D} = D$. Wir zeigen, daß dann $\complement D$ offen ist. Dies ist äquivalent zu $\forall x \in \complement D\, \exists r_x > 0\colon K(x, r_x) \subset \complement D$. Angenommen diese Aussage wäre falsch, d.h. $\exists x_0 \in \complement D\, \forall r > 0\colon K(x_0, r) \cap D \neq \emptyset$. Dann gäbe es eine Folge $(x_n) \subset D$ mit $\lim_{n\to\infty} x_n = x_0$.

Somit wäre x_0 ein Häufungspunkt von D. Es müßte also $x_0 \in D = \bar{D}$ gelten, im Widerspruch zu $x_0 \in \complement D$. □

DEFINITION III-1.8. *Es sei $f\colon D \to Y$, $D \subset X$ nicht leer, und $x_0 \in X$ ein Häufungspunkt von D.*
$F \in Y$ heißt ***Grenzwert*** *von f an der Stelle $x_0 \in X$, $F = \lim_{x \to x_0} f(x)$, genau dann, wenn*

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x \in D \setminus \{x_0\}\colon \|x - x_0\|_X < \delta \Rightarrow \|F - f(x)\|_Y < \varepsilon.$$

Eine gleichwertige Formulierung verwendet den Umgebungsbegriff:

$$F = \lim_{x \to x_0} f(x) \Leftrightarrow \forall K(F, \varepsilon) \, \exists K(x_0, \delta)\colon f((K(x_0, \delta) \setminus \{x_0\}) \cap D) \subset K(F, \varepsilon).$$

Mit dem Konzept des Grenzwertes will man das lokale Verhalten einer Funktion in der Nähe einer Stelle $x_0 \in X$ erfassen. Dabei ist es unerheblich, ob x_0 im Definitionsbereich von f liegt. Falls $x_0 \in D$ hat der Funktionswert $f(x_0)$ keinen Einfluß auf den Grenzwert, die Funktion wird ja nicht in x_0 ausgewertet. Von x_0 wird nur gefordert, daß man x_0 durch Elemente aus D "beliebig"genau annähern kann, d.h. x_0 muß ein Häufungspunkt von D sein.

THEOREM III-1.9. *Eine Funktion $f\colon D \to Y$, $D \subset X$, besitzt in einem Häufungspunkt von D höchstens einen Grenzwert.*

BEWEIS. Vergleiche den Beweis von Satz II-2.7. □

Es besteht ein enger Zusammenhang zwischen dem Grenzwert für Folgen und jenem für Funktionen.

THEOREM III-1.10. *Es sei $f\colon D \to Y$. $D \subset X$ und $x_0 \in \bar{D}$. Äquivalent sind folgende Aussagen.*
i) $F = \lim_{x \to x_0} f(x)$.
ii)Für jede Folge $(x_n) \subset D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$ ist die Folge der Bilder konvergent und es gilt $\lim_{n\to\infty} f(x_n) = F$.

BEWEIS. *i) ⇒ ii)* Zu $\varepsilon > 0$ wählen wir $\delta > 0$ gemäß Definition III-1.8. Für dieses δ existiert wegen $\lim_{n\to\infty} x_n = x_0$ ein Index $N(\delta) \in \mathbb{N}$ mit $||x_n - x_0||_X < \delta$ für $n \geq N(\delta)$. Aus $F = \lim_{x \to x_0} f(x)$ folgt wegen der Wahl von δ daher $||f(x_n) - F||Y < \varepsilon$ für $n \geq N(\delta)$, also $\lim_{n\to\infty} f(x_n) = F$ (man beachte $\delta = \delta(\varepsilon)$).
ii) ⇒ i) Wir zeigen $\neg i) \Rightarrow \neg ii)$ Aus der Negation von i) folgt

$$\exists \varepsilon_0 \, \exists (x_n) \subset D \setminus \{x_0\}\colon \lim_{n\to\infty} x_n = x_0 \wedge \|f(x_n) - F\| \geq \varepsilon_0.$$

Die Folge $(f(x_n)$ kann also nicht gegen F konvergieren. □

Diese Charakterisierung eignet sich besonders gut, um nachzuweisen, daß der Grenzwert nicht existiert: es genügt, zwei Folgen (x_n) und $(\tilde{x}_n) \subset D \setminus \{x_0\}$ zu finden mit $\lim_{n\to\infty} x_n = \lim_{n\to\infty} \tilde{x}_n = x_0$, aber $\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(\tilde{x}_n)$. Es reicht auch bereits eine einzige Folge $(x_n) \subset D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$, für welche die Folge der Bilder $(f(x_n)) \subset Y$ nicht konvergiert.

Beispiel III-1.11. Zu untersuchen ist die Existenz des Grenzwertes für $f\colon \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$, $f(x,y) = \frac{xy^2}{2x^2+3y^4}$. Wegen $f(x,0) = 0$, $x \neq 0$, folgt $\lim_{x\to 0} f(x,0) = 0$. Betrachtet man aber die Folge $(x_n, y_n) = (t_n, \sqrt{t_n})$, $t_n > 0$, mit $\lim_{n\to\infty} t_n = 0$, gilt $\lim_{n\to\infty}(x_n, y_n) = (0,0)$ und

$$f(x_n, y_n) = \frac{t_n^2}{2t_n^2 + 3t_n^2} = \frac{1}{5}.$$

Somit gilt auch $\lim_{n\to\infty} f(x_n, y_n) = \frac{1}{5}$. Die Funktion f besitzt daher keinen Grenzwert in $(0,0)$.

Wie bei Folgen kann man auf die bloße Existenz eines Grenzwertes schließen, ohne bereits dessen Wert vermuten zu müssen. Dies erfordert jedoch die Vollständigkeit des Bildraumes Y.

Theorem III-1.12. *Es sei $f\colon D \to Y$, $D \subset X$, $x_0 \in \bar{D}$ und Y vollständig. Äquivalent sind folgende Aussagen:*
i) f besitzt einen Grenzwert in x_0.
ii) Für jede Folge $(x_n) \subset D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$ ist die Folge der Bilder $(f(x_n))$ eine Cauchy Folge in Y.

Beweis. *i)* $\Rightarrow$ *ii)* ist klar, da eine konvergente Folge immer eine Cauchy Folge ist.
ii) $\Rightarrow$ *i)* Es seien (x_n) und $(\tilde{x}_n) \subset D \setminus \{x_0\}$ Folgen, welche nach x_0 konvergieren. Dann sind die Folgen der Bilder $f(x_n)$ und $f(\tilde{x}_n)$ Cauchy Folgen und wegen der Vollständigkeit von Y konvergent. Es seien $y = \lim_{n\to\infty} f(x_n)$ und $\tilde{y} = \lim_{n\to\infty} f(\tilde{x}_n)$. Wir zeigen $y = \tilde{y}$ und betrachten dazu die Mischfolge $z_{2n-1} = x_n$, $z_{2n} = \tilde{x}_n$, $n \in \mathbb{N}$. Es folgt $\lim_{n\to\infty} z_n = x_0$. Wäre $y \neq \tilde{y}$ könnte die Folge der Bilder $(f(z_n))$ nicht konvergieren und somit auch keine Cauchy Folge sein. Es muß daher $y = \tilde{y}$ gelten. Der Grenzwert F der Bildfolge $(f(x_n))$ ist daher unabhängig von der Folge (x_n), welche zur Approximation von x_0 verwendet wird. Für jede gegen x_0 konvergente Folge $(x_n) \subset D \setminus \{x_0\}$ gilt also $\lim_{n\to\infty} f(x_n) = F$. Somit ist F der Grenzwert von f in x_0. □

Die Berechnung von Grenzwerten kann durch folgende Regeln vereinfacht werden.

Theorem III-1.13. *Es seien f, $g\colon D \to Y$, $D \subset X$ und $x_0 \in \bar{D}$. Existieren die Grenzwerte $\lim_{x\to x_0} f(x)$ und $\lim_{x\to x_0} g(x)$, dann besitzen auch die Funktionen λf für $\lambda \in \mathbb{K}$ und $f + g$ einen Grenzwert in x_0 und es gilt*

$$\lim_{x\to x_0} (\lambda f)(x) = \lambda \lim_{x\to x_0} f(x),$$
$$\lim_{x\to x_0} (f+g)(x) = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x).$$

Beweis. Übung. □

Theorem III-1.14. *Es seien $f\colon D \to V$, $g\colon V \to Z$, $D \subset X$, $V \subset Y$, $x_0 \in \bar{D}$, $y_0 \in \bar{V}$ und $f(x) \neq y_0$ für $x \in D$. Existieren $\lim_{x\to x_0} f(x) = y_0$ und $\lim_{y\to y_0} g(y) = z_0$, dann besitzt $g \circ f$ einen Grenzwert in x_0 und es gilt $\lim_{x\to x_0}(g \circ f)(x) = z_0$.*

Beweis. Übung. □

Abschließend gehen wir näher auf den Fall $X = \mathbb{K}$ und / oder $Y = \mathbb{K}$ ein. Ist $Y = \mathbb{K}$ kann man wegen der Körperstruktur die Regeln aus Satz III-1.13 ergänzen.

Theorem III-1.15. *Es seien f, $g\colon D \to \mathbb{K}$, $D \subset X$ und $x_0 \in \bar{D}$. Existieren die Grenzwerte $\lim_{x\to x_0} f(x)$ und $\lim_{x\to x_0} g(x)$, dann besitzen auch die Funktionen fg und falls $\lim_{x\to x_0} g(x) \neq 0$ auch $\frac{f}{g}$ Grenzwerte in x_0 und es gilt*

$$\lim_{x\to x_0}(fg)(x) = \lim_{x\to x_0} f(x) \lim_{x\to x_0} g(x),$$

$$\lim_{x\to x_0}(\frac{f}{g})(x) = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}.$$

Als Anwendung zeigen wir, daß Polynome überall einen Grenzwert besitzen. Dazu betrachten wir die Funktionen $p_i\colon \mathbb{K} \to \mathbb{K}$, $i \in \mathbb{N}_0$, $p_i(x) = x^i$ (wir vereinbaren für diese Überlegung $0^0 = 1$). Die Existenz der Grenzwerte für $x_0 \in \mathbb{K}$ und $i \in \mathbb{N}_0$ ist eine einfache Konsequenz aus Satz III-1.15 und der Beobachtung $p_i = p_{i-1}p_1$. Es gilt $\lim_{x\to x_0} p_i(x) = x_0^i = p_i(x_0)$. Aus Satz III-1.13 folgt daher für $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$ die Existenz des Grenzwertes für die Polynomfunktion $p = \sum_{i=0}^n a_i p_i$ und

$$\lim_{x\to x_0} p(x) = \sum_{i=0}^n a_i \lim_{x\to x_0} p_i(x) = \sum_{i=0}^n a_i x_0^i = p(x_0).$$

Zum Vergleich führen wir nun einen Beweis, der auf die Definition des Grenzwertes zurückgreift. Es sei also $x_0 \in \mathbb{K}$ und $\varepsilon > 0$ gewählt. Mit Hilfe der Hornerschen Regeln erhält man

$$p(x) - p(x_0) = \sum_{i=0}^n a_i(x^i - x_0^i) = (x - x_0)\sum_{i=0}^n a_i \sum_{k=0}^{i-1} x^{i-1-k}x_0^k.$$

Für $x \in K(x_0, 1)$ gilt die Abschätzung

$$|x| < 1 + |x_0|.$$

Somit folgt für $x \in K(x_0, 1)$

$$\begin{aligned}|p(x) - p(x_0)| &\leq |x - x_0| \sum_{i=0}^n |a_i| \sum_{k=0}^{i-1} |x|^{i-1-k}|x_0|^k \\ &\leq |x - x_0| \sum_{i=0}^n |a_i| \sum_{k=0}^{i-1} (1 + |x_0|)^{i-1-k}|x_0|^k = M\,|x - x_0|\end{aligned}$$

mit $M = \sum_{i=0}^n |a_i| \sum_{k=0}^{i-1}(1 + |x_0|)^{i-1-k}|x_0|^k > 0$. Falls $|x - x_0| < \delta$ erhält man

$$(*) \qquad |p(x) - p(x_0)| < M\delta.$$

Die Abschätzung

$$|p(x) - p(x_0)| < \varepsilon$$

folgt dann für alle $x \in K(x_0, \delta)$, falls $\delta < \frac{\varepsilon}{M}$ gewählt wird. Da bei der Herleitung von $(*)$ die a priori Annahme $x \in K(x_0, 1)$ verwendet wurde, muß zusätzlich $\delta < 1$ sein, d.h

$$\delta < \min(\frac{\varepsilon}{M}, 1).$$

Es sei nun r eine **rationale Funktion**, d.h.

$$r = \frac{p}{q}, \qquad p, q \text{ Polynome.}$$

Der Definitionsbereich von r ist $D(r) = \{x \in \mathbb{K}\colon q(x) \neq 0\}$. Mit Hilfe von Satz III-1.15 schließt man für $x_0 \in D(r)$ auf die Existenz des Grenzwertes von r und

$$\lim_{x \to x_0} r(x) = \lim_{x \to x_0} \frac{p}{q}(x) = \frac{\lim_{x \to x_0} p(x)}{\lim_{x \to x_0} q(x)} = \frac{p(x_0)}{q(x_0)} = r(x_0).$$

Wir fassen zusammen:

Theorem III-1.16. *Polynome und rationale Funktionen besitzen an jeder Stelle des Definitionsbereiches einen Grenzwert. Dieser stimmt mit dem Funktionswert überein.*

Es sei nun $X = \mathbb{R}$. Wegen der zusätzlichen Ordnungsstruktur kann man sinnvollerweise von einer Annäherung von links, bzw. rechts an x_0 sprechen. Dies motiviert die Einführung einseitiger Grenzwertbegriffe:

Definition III-1.17. *Es sei $f\colon D \to Y$, $D \subset \mathbb{R}$ nicht leer, und $x_0 \in \bar{D}$.*
i) $L \in Y$ heißt ***linksseitiger Grenzwert*** *von f an der Stelle $x_0 \in \bar{D}$, $L = \lim_{x \to x_0^-} f(x)$, genau dann, wenn*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D\colon 0 < x_0 - x < \delta \Rightarrow \|f(x) - L\|_Y < \varepsilon,$$

ii) $R \in Y$ heißt ***rechtsseitiger Grenzwert*** *von f an der Stelle $x_0 \in \bar{D}$, $R = \lim_{x \to x_0^+} f(x)$ genau dann, wenn*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D\colon 0 < x - x_0 < \delta \Rightarrow \|f(x) - R\|_Y < \varepsilon.$$

Zwischen den einseitigen Grenzwerten und dem Grenzwert besteht folgender naheliegende Zusammenhang:

Theorem III-1.18. *Es sei $f\colon D \to Y$, $D \subset \mathbb{R}$ nicht leer, und $x_0 \in \bar{D}$. Äquivalent sind folgende Aussagen:*
i) Es existiert $\lim_{x \to x_0} f(x)$.
ii) Es existieren der rechts- und linksseitige Grenzwert in x_0 und beide sind gleich.

Beweis. Übung. □

Für die Untersuchung des asymptotischen Verhaltens einer Funktion ist folgende Variante des Grenzwertbegriffes nützlich:

Definition III-1.19. *Es sei $f\colon D \to Y$, $D \subset \mathbb{R}$, D nach oben unbeschränkt.*

*$L \in Y$ heißt **Grenzwert** von f bei ∞,* $\lim_{x\to\infty} f(x)$, *genau dann, wenn*
$\forall \varepsilon > 0\ \exists \xi \in \mathbb{R}\ \forall x \in D\colon x > \xi \Rightarrow \|f(x) - L\|_Y < \varepsilon.$
Entsprechend definiert man den Grenzwert bei $-\infty$.

Beispiel III-1.20. $\lim_{x\to\infty}(\sqrt{x+1} - \sqrt{x}) = 0$. Zum Beweis benützen wir die für $x > 0$ gültige Abschätzung

$$\sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1} + \sqrt{x}} < \frac{1}{2\sqrt{x}}.$$

Für $x > \xi = \frac{1}{4\varepsilon^2}$ folgt damit $|\sqrt{x+1} - \sqrt{x}| < \varepsilon$.

Die Untersuchung der Grenzwerte von $f\colon D \to \mathbb{K}$ bei Unendlich kann durch die Substitution $x \mapsto \xi = \frac{1}{x}$ auf die Untersuchung einseitiger Grenzwerte bei 0 zurückgeführt werden: Setzt man $\varphi(x) = f(\frac{1}{x})$, falls $\frac{1}{x} \in D(f)$, dann gilt $\lim_{x\to\infty} f(x) = \lim_{\xi\to 0^+} \varphi(\xi)$.
Bisher wurde immer angenommen, daß f beschränkt ist. Auch für unbeschränkten Funktionen kann man einen (uneigentlichen) Grenzwert definieren:

Definition III-1.21. *Es sei $f\colon D \to \mathbb{R}$, $D \subset X$ und $x_0 \in \bar{D}$. f hat an der Stelle x_0 den uneigentlichen Grenzwert ∞,*

$$\lim_{x\to x_0} f(x) = \infty \Leftrightarrow \forall \xi > 0 \exists \delta > 0 \forall x \in D\colon 0 < \|x - x_0\|_< \delta \Rightarrow f(x) > \xi.$$

Es ist für das Verständnis sehr hilfreich, sich klar zu machen, daß alle bisherigen Varianten des Grenzwertbegriffes folgendem allgemeinen Prinzip genügen:

$$\forall V \in \mathfrak{U}(L)\ \exists U \in \mathfrak{U}(x_0)\colon f((U \setminus \{x_0\}) \cap D) \subset V.$$

2. Stetigkeit

Wir erinnern daran, daß bei der Untersuchung des Grenzwertes einer Funktion die betreffende Stelle x_0 nur ein Häufungspunkt des Definitionsbereiches sein mußte. Es war völlig unerheblich, ob die Funktion an der Stelle x_0 definiert war. Falls der Funktionswert $f(x_0)$ exisitierte, wurde er bei der Berechnung von $\lim_{x\to x_0} f(x)$ nicht benutzt. Die Untersuchung der Stetigkeit einer Funktion an der Stelle x_0 erfordert nun einen radikalen Wechsel des Standpunktes: nun *muß* x_0 im Definitionsbereich von f liegen und der Funktionswert $f(x_0)$ *muß* zum lokalen Verhalten der Funktion passen:

Definition III-2.1. *Es sei $f\colon D \to Y$, $D \subset X$ nicht leer.*
*i) f heißt **stetig** in $x_0 \in D$ genau dann,*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D\colon \|x - x_0\|_X < \delta \Rightarrow \|f(x) - f(x_0)\|_Y < \varepsilon.$$

ii) f heißt stetig auf $S \subset D$ genau dann, wenn f in jeder Stelle von S stetig ist.
iii) $C(D, Y) = \{f\colon D \to Y,\ f$ ist stetig$\}$

Vergleicht man diese Definition mit der Definition des Grenzwertes, erhält man folgende gleichwertige Charakterisierung:

> f ist stetig in einem Häufungspunkt $x_0 \in D$ genau dann, wenn der Grenzwert von f in x_0 existiert und wenn $\lim_{x \to x_0} f(x) = f(x_0)$ gilt.

In einem isolierten Punkt des Definitionsbereiches ist jede Funktion trivialerweise stetig.

Beispiel III-2.2. Polynome und rationale Funktionen sind stetig.

Beispiel III-2.3. Die Norm in einem normierten linearen Raum ist stetig. Dies ist eine unmittelbare Folge der Abschätzung

$$|\|x\| - \|y\|| \leq \|x - y\|$$

Beispiel III-2.4. Es sei $f\colon D \to Y$, $D \subset X$ und es existiere eine Konstante $L \geq 0$, so daß

$$(*) \qquad \|f(x) - f(y)\|_Y \leq L\|x - y\|.$$

für alle x, $y \in D$ gilt. Dann ist f stetig.
Im Fall $L = 0$ ist f konstant und daher trivialerweise stetig. Für $L > 0$ wähle man $\delta < \frac{\varepsilon}{L}$ für ein gegebenes $\varepsilon > 0$. Man nennt die Abschätzung $(*)$ eine **Lipschitz Bedingung** und Funktionen mit dieser Eigenschaft **Lipschitz stetig**.

Beispiel III-2.5. Wir zeigen, daß die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \to f(x, y) = \frac{x^3 y}{2x^2+3y^2}$, $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$ stetig in $(0, 0)$ ist. Zur Abstandsmessung in $\mathbb{R}^2$ verwenden wir die euklidische Norm. Zu zeigen ist also:

$$(*) \qquad \forall \varepsilon > 0\ \exists \delta > 0\ \forall (x, y) \in \mathbb{R}^2\colon \sqrt{x^2 + y^2} < \delta \Rightarrow |\frac{x^3 y}{2x^2 + 3y^2}| < \varepsilon.$$

Für $(x, y) \neq (0, 0)$ und $\|(x, y)\| = \sqrt{x^2 + y^2} < \delta$ (für ein noch zu bestimmendes $\delta > 0$) erhalten wir die Abschätzung

$$|f(x, y)| \leq \frac{1}{2} \frac{x^2}{x^2 + y^2} |xy| \leq \frac{1}{2}|xy| \leq \frac{1}{4}(x^2 + y^2) < \frac{1}{4}\delta^2 \overset{!}{<} \varepsilon$$

Die letzte Ungleichung ist als Bedingung an δ zu verstehen. Sie gilt falls

$$\delta < 2\sqrt{\varepsilon}.$$

Somit ist $(*)$ gezeigt.

Wegen der engen Verwandtschaft zwischen dem Grenzwertbegriff und der Stetigkeit kann man die grundlegenden Eigenschaften von Grenzwerten mühelos auf stetige Funktionen übertragen.

Theorem III-2.6. *Die Funktionen f, $g\colon D \to Y$, $D \subset X$ nicht leer, seien stetig in $x_0 \in D$.*
i) Dann sind auch die Funktionen λf, $\lambda \in \mathbb{K}$, und $f + g$ stetig in x_0.
ii) Ist $Y = \mathbb{K}$, dann sind die Funktionen fg, und falls $g(x_0) \neq 0$ auch $\frac{f}{g}$ stetig in x_0.

Als Konsequenz dieser Regeln halten wir fest, daß $C(D, Y)$ ein Vektorraum ist.

Theorem III-2.7. *Es sei $f\colon D \to Y$, $D \subset X$ und $x_0 \in D$. Äquivalent sind folgende Aussagen.*
i) f ist stetig in x_0.
ii)Für jede Folge $(x_n) \subset D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$ ist die Folge der Bilder konvergent und es gilt $\lim_{n\to\infty} f(x_n) = f(x_0)$.

Der Satz bleibt richtig, wenn man in ii) beliebige Folgen in D zuläßt.
Die Stetigkeit bedeutet also die Vertauschbarkeit eines Grenzwertes und Abbildungsvorschrift:

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n).$$

Da sich die Menge der konvergenten Folgen eines normierten Raumes nicht ändert, wenn man zu einer äquivalenten Norm übergeht, erhält man aus Satz III-2.7 folgendes nützliche Ergebnis.

THEOREM III-2.8. *Es sei $f\colon D \to Y$, $D \subset X$, stetig in $x_0 \in D$. Ersetzt man in X und/oder Y die jeweiligen Normen durch äquivalente Normen, bleibt f stetig in x_0.*

Der Grenzwert einer Funktion kann verwendet werden, um Lücken im Definitionsbereich sinnvoll zu schließen:

THEOREM III-2.9. *Es sei $f\colon D \to Y$, $D \subset X$ und $x_0 \in \bar{D} \setminus D$. Ferner existiere $\lim_{x\to x_0} f(x)$. Definiert man $F\colon D \cup \{x_0\} \to Y$ durch*

$$F(x) = \begin{cases} f(x), & x \in D \\ \lim_{x\to x_0} f(x), & x = x_0 \end{cases}$$

*dann ist F stetig in x_0. Man nennt F **stetige Fortsetzung** von f nach x_0.*

BEWEIS. Übung. □

DEFINITION III-2.10. *Eine Teilmenge M von $D \subset X$ heißt **dicht** in D, wenn $D \subset \bar{M}$.*

Die Menge der rationalen Zahlen beispielsweise ist dicht in $\mathbb{R}$.

THEOREM III-2.11. *Stimmen zwei stetige Funktionen f, $g\colon D \to Y$, $D \subset X$, auf einer dichten Teilmenge von D überein, dann sind sie gleich. Insbesondere gibt es höchstens eine stetige Fortsetzung von D auf $\bar{D}$.*

BEWEIS. Übung. □

Die Unstetigkeit einer Funktion an einer Stelle x_0 kann somit zwei Ursachen haben:

- Es existiert zwar der Grenzwert von f in x_0, aber $\lim_{x\to x_0} f(x) \neq f(x_0)$. In diesem Fall liegt eine **hebbare Unstetigkeit** vor.
- Der Grenzwert von f in x_0 existiert nicht. Man nennt x_0 eine **Unstetigkeit 2. Art**. Im Fall $D \subset \mathbb{R}$ ist dies beispielsweise der Fall, wenn die beiden einseitigen Grenzwerte existieren, aber nicht gleich sind. Es liegt dann eine **Sprungstelle** vor.

BEISPIEL III-2.12. (Dirichlet Funktion)

$$d\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \begin{cases} 0 & x \in \mathbb{Q} \\ 1 & x \notin \mathbb{Q} \end{cases} \end{cases}$$

d ist für alle $x \in \mathbb{R}$ unstetig, denn jede δ-Umgebung von x enthält sowohl rationale als auch irrationale Zahlen.

Im folgenden Beispiel ist die Anschauung wenig hilfreich bei der Diskussion der Stetigkeit und analytische Methoden sind unumgänglich:

Beispiel III-2.13. Wir betrachten die Abbildung

$$f\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto 0 & x\{0\} \cup \mathbb{R} \setminus \mathbb{Q} \\ x \mapsto \frac{1}{q} & x = \frac{p}{q}, \quad p \in \mathbb{Z}, q \in \mathbb{N} \text{ teilerfremd} \end{cases}$$

Wir betrachten zuerst eine rationale Stelle $x_0 = \frac{p_0}{q_0} \neq 0$ (p_0, q_0 teilerfremd) und die Folge $(x_n) \subset \mathbb{R}\backslash\mathbb{Q}$, $x_n = x_0 + \frac{1}{n}\sqrt{2}$, $n \in \mathbb{N}$. Es ist $\lim_{n\to\infty} x_n = x_0$, aber $\lim_{n\to\infty} f(x_n) = 0 \neq \frac{1}{q_0} = f(x_0)$, d.h. f ist an jeder rationalen Stelle ungleich Null unstetig.
Es sei nun x_0 irrational, also $f(x_0) = 0$. Zu $\varepsilon > 0$ wählen wir N so, daß $\frac{1}{N} < \varepsilon$. Im Intervall $(-\frac{1}{N} + x_0, \frac{1}{N} + x_0)$ gibt es nur endlich viele rationale Zahlen $\frac{p}{q}$, deren Nenner q kleiner ist als N. Wir wählen nun

$$\delta = \min\{|x_0 - \frac{p}{q}|\colon p \in \mathbb{Z}, q \in \mathbb{N}, q < N, |x_0 - \frac{p}{q}| < \frac{1}{N}\}.$$

Der Nenner q jeder rationalen Zahl $\frac{p}{q}$ mit $|x_0 - \frac{p}{q}| < \delta$ muß also zwangsläufig größer (mindestens gleich) N sein. Somit gilt auch

$$|f(x) - f(x_0)| = |f(x)| = \begin{cases} \frac{1}{q} < \frac{1}{N} < \varepsilon & \text{für } x = \frac{p}{q}, \\ 0 < \varepsilon & \text{für } x \notin \mathbb{Q}. \end{cases}$$

Dieses Argument ist auch für $x = 0$ mit $\delta < \frac{1}{N}$ anwendbar. Diese Funktion ist also an jeder irrationalen Stelle und in $x = 0$ stetig. Die Abbildung dieses Beispiels weist somit ein sehr komplexes Verhalten auf: sowohl die Menge der Stetigkeitsstellen, als auch die Menge der Unstetigkeitsstellen liegen dicht in $\mathbb{R}$!

Ein anderer Beweis der Stetigkeitseigenschaften dieses Beispiels kann auch auf folgender Beobachtung aufgebaut werden:

Lemma III-2.14. *Es sei $\xi \in \mathbb{R} \setminus \mathbb{Q}$ und $(r_n) \subset \mathbb{Q}$ mit $\frac{p_n}{q_n}$, $p_n \in \mathbb{Z}$, $q_n \in \mathbb{N}$ teilerfremd. Dann gilt:*

$$\lim_{n\to\infty} r_n = \xi \Rightarrow \lim_{n\to\infty} q_n = \infty$$

Besonders nützlich ist das Resultat über die Verkettung stetiger Funktionen.

Theorem III-2.15. *Es sei $f\colon D \to Y$, $g\colon V \to Z$, $D \subset X$, $V \subset Y$ und $f(D) \subset V$. Ist f stetig in $x_0 \in D$ und g stetig in $f(x_0) \in V$, dann ist $g \circ f\colon D \to Z$ stetig in x_0.*

Beweis. Da g stetig in $f(x_0)$ ist, existiert zu $\varepsilon > 0$ eine σ-Umgebung $K(\sigma, f(x_0))$so, daß $\|g(y) - g(f(x_0))\|_Z < \varepsilon$ für alle $y \in K(\sigma, f(x_0)) \cap V$ zutrifft. Wir wählen nun zu σ entsprechend der Stetigkeit von f in x_0 eine δ-Umgebung von x_0 so, daß $\|f(x) - f(x_0)\|_Y < \sigma$ für alle $x \in K(\delta, x_0) \cap D$ zutrifft. Für diese Argumente folgt dann $\|g(f(x)) - g(f(x_0))\|_Z < \varepsilon$. □

Besonders einfach gestaltet sich der Nachweis der Stetigkeit für *lineare Abbildungen.*

Theorem III-2.16. *Es seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ normierte Räume und $f\colon X \to X$ linear. Folgende Aussagen sind äquivalent.*

i) *f ist stetig,*
ii) *f ist stetig in 0,*
iii) *$\exists M > 0\, \forall x \in X\colon \|f(x)\|_Y \leq M\|x\|_X$.*

Beweis. i) $\Rightarrow$ ii) trivial.
ii) $\Rightarrow$ iii) Es sei f stetig in 0. Wegen $f(0) = 0$ gibt es zu $K_Y(0, 1)$ ein $\delta > 0$ mit $f(K_X(0, \delta)) \subset K_Y(0, 1)$. Es sei $x \in X$, $x \neq 0$. Wir setzen $z_x = \frac{x}{\|x\|_X}\frac{\delta}{2}$. Wegen $z_x \in K_X(0, \delta)$ folgt $\|f(z_x)\|_Y < 1$. Wegen der Linearität von f folgt

$$\|f(z_x)\|_Y = \left\|f\left(\frac{x}{\|x\|_X}\frac{\delta}{2}\right)\right\|_Y = \left\|\frac{\delta}{2\|x\|_X} f(x)\right\|_Y = \frac{\delta}{2\|x\|_X}\|f(x)\|_Y < 1,$$

d.h. es gilt iii) mit $M = \frac{2}{\delta}$.
iii) $\Rightarrow$ i) Dies folgt aus

$$\|f(x) - f(y)\|_Y = \|f(x-y)\|_Y \leq M\|x-y\|_X.$$

$\square$

Eine lineare Abbildung mit der Eigenschaft iii) aus Satz III-2.16 nennt man **beschränkt**. Dies ist nach unserer bisherigen Definition gleichbedeutend mit der Beschränktheit von $f|_{\bar{K}_X(0,1)}$. Es ist üblich, die Beschränktheit einer *linearen* Abbildung durch die Bedingung

$$\|f\| := \sup\{\|f(x)\|_Y \colon \|x\|_X \leq 1\} < \infty$$

zu charakterisieren. Nach Definition II-1.4 wäre nur die lineare Abbildung $f \equiv 0$ beschränkt. Satz III-2.16 zeigt, daß für lineare Abbildungen die Stetigkeit äquivalent ist zur Beschränktheit der Abbildung.

Bemerkung III-2.17. Als Übung überlege man sich, daß die Ungleichung

$$\|x\|_X \leq m\|f(x)\|_Y$$

für eine lineare Abbildung $f\colon X \to Y$ die Stetigkeit der Umkehrfunktion $f^{-1}\colon f(X) \to X$ bedeutet.

Insbesondere sind lineare Abbildungen zwischen endlich dimensionalen normierten Räumen stetig.

Theorem III-2.18. *Es sei $A \in \mathbb{K}^{n\times m}$. Dann ist die durch die Matrix $A = (a_{ik})$ definierte Abbildung $x \to Ax$, $\mathbb{K}^m \to \mathbb{K}^n$ linear und stetig.*

Beweis. Wegen Satz III-2.8 können wir in $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ jeweils $\|\cdot\|_\infty$ als Norm verwenden. Die Linearität der Abbildung $x \to Ax$ wird in der Linearen Algebra gezeigt. Wegen Satz III-2.16 genügt es daher, die Beschränktheit der Abbildung $x \to Ax$ nachzuweisen. Wir erinnern daran, daß die i-te Koordinate des Bildes gegeben ist durch

$$(Ax)_i = \sum_{k=1}^m a_{ik}x_k, \qquad i = 1, \dots, n,$$

$(x = (x_1, \dots, x_m))$. Es folgt

$$\begin{aligned}
|(Ax)_i| &= |\sum_{k=1}^m a_{ik}x_k| \leq \sum_{k=1}^m |a_{ik}|\,|x_k| \\
&\leq \max\{|x_k|, 1 \leq k \leq m\} \sum_{k=1}^m |a_{ik}| = \|x\| \sum_{k=1}^m |a_{ik}| \\
&\leq \max\{\sum_{k=1}^m |a_{ik}|, 1 \leq i \leq n\}\|x\|.
\end{aligned}$$

Diese Ungleichung gilt für alle $1 \le i \le n$. Folglich gilt auch

$$\|Ax\| \le M\|x\|,$$

mit $M = \max\{\sum_{k=1}^m |a_{ik}|, 1 \le i \le n\}$. Somit ist die Abbildung $x \to Ax$ beschränkt. □

Abschließend skizzieren wir eine Technik, welche es erlaubt, die Stetigkeit einer Funktion in mehreren Veränderlichen aus der (bekannten) Stetigkeit von Funktionen in einer Veränderlichen abzuleiten.

LEMMA III-2.19. *Es sei $I \subset \mathbb{K}$, $\varphi\colon I \to \mathbb{K}$ stetig in $x_0 \in I$ und $J \subset \mathbb{K}$. Dann ist die Funktion $\Phi\colon I \times J \to \mathbb{K}$, $\Phi(x,y) = \varphi(x)$, stetig in (x_0, y) für alle $y \in J$.*

BEWEIS. Übung. □

Die Reihenfolge der Argumente von Φ kann ohne Einfluß auf die Stetigkeit vertauscht werden.

LEMMA III-2.20. *Es seien I, $J \subset \mathbb{K}$ und $\varphi\colon I \to \mathbb{K}$ stetig in $x_0 \in I$ und $\psi\colon J \to \mathbb{K}$ stetig in $y_0 \in J$. Dann sind die Funktionen f und $g\colon I \times J \to \mathbb{K}$, $f(x,y) = \varphi(x) + \psi(y)$, bzw. $g(x,y) = \varphi(x)\psi(y)$ stetig in (x_0, y_0).*

BEWEIS. Wir zeigen die Stetigkeit von f: nach Lemma III-2.19 sind die Funktionen Φ, $\Psi\colon I \times J \to \mathbb{K}$, $\Phi(x,y) = \varphi(x)$ bzw. $\Psi(x,y) = \psi(y)$ stetig in (x_0, y_0). Die Behauptung folgt nun aus $f = \Phi + \Psi$. □

BEISPIEL III-2.21. Als Anwendung zeigen wir, daß die Funktion aus Beispiel III-2.5 für alle $(x,y) \ne (0,0)$ stetig ist. Dazu beachte man, daß die Funktionen $x \to x^3$, $x \to y^2$ stetig sind als Funktionen in einer Veränderlichen, Beispiel III-2.2. Nach Lemma III-2.20 sind daher auch die Funktionen $(x,y) \to x^3 y$ und $(x,y) \to 2x^2 + 3y^2$ stetig, somit auch ihr Quotien, falls $(x,y) \ne (0,0)$, Satz III-2.6.

3. Kompakte Mengen

DEFINITION III-3.1. *Es sei $(X, ||\cdot||)$ ein normierter Raum und $K \subset X$.*

i) *Eine Familie $\mathcal{C}$ offener Mengen heißt (offene)* ***Überdeckung*** *von $K \underset{Def}{\Leftrightarrow}$ $K \subset \bigcup \mathcal{C}$.*
Eine Teilfamilie $\mathcal{S} \subset \mathcal{C}$ heißt ***Teilüberdeckung*** *von $K \underset{Def}{\Leftrightarrow}$ $\mathcal{S}$ ist Überdeckung von K.*

ii) *$K \subset X$ heißt* ***kompakt*** *$\underset{Def}{\Leftrightarrow}$ jede offene Überdeckung $\mathcal{C}$ von K besitzt eine endliche Teilüberdeckung (d.h. eine Überdeckung mit einer endlichen Anzahl von Elementen aus $\mathcal{C}$).*

BEMERKUNG III-3.2. Die Kompaktheit ist eine **topologische** Eigenschaft, d.h. sie bleibt erhalten, wenn man die zugrundeliegende Norm durch eine äquivalente Norm ersetzt.

Beispiel III-3.3. 1) Jede endliche Teilmenge eines normierten Raumes ist kompakt. Man braucht aus einer beliebigen Überdeckung von K für jeden Punkt $x \in K$ nur ein Element der Überdeckung auswählen, welches x enthält.
2) Es sei $K = (0, 1] \subset \mathbb{R}$. K ist nicht kompakt. Um dies einzusehen, betrachten wir die Überdeckung

$$\mathcal{C} = \{(\frac{1}{n}, 2) \colon n \in \mathbb{N}\}$$

von K. Würden bereits endlich viele Elemente von $\mathcal{C}$, etwa die Mengen $(\frac{1}{n_i}, 2)$, $i = 1, \ldots, N$, $n_1 > n_2 > \cdots > n_N$, die Menge K überdecken, dann müßte $(0, 1] \subset \bigcup_{i=1}^{N} (\frac{1}{n_i}, 2) = (\frac{1}{n_1}, 2)$ gelten. Dies ist nicht möglich.
3) $K = [0, \infty)$ ist nicht kompakt, man betrachte etwa die Überdeckung

$$\mathcal{C} = \{(-1, n) \colon n \in \mathbb{N}\}$$

von K. $\mathcal{C}$ enthält keine endliche Teilüberdeckung von K.

Die beiden letzten Beispiele spiegeln notwendige Bedingungen für die Kompaktheit einer Menge wider.

Lemma III-3.4. *Jede kompakte Teilmenge eines normierten Raumes ist beschränkt.*

Beweis. Es sei $K \subset X$ kompakt und $x_0 \in X$. Die Familie

$$\mathcal{C} = \{K(x_0, n) \colon n \in \mathbb{N}\}$$

ist eine offene Überdeckung von X und daher auch von K. Da K kompakt ist, wird K bereits durch endlich viele Kugeln $K(x_0, n)$ überdeckt. Wegen $K(x_0, n) \subset K(x_0, n+1)$, $n \in \mathbb{N}$, gibt es also ein $N \in \mathbb{N}$ mit $K \subset K(x_0, N)$. Somit ist K beschränkt. $\square$

Theorem III-3.5. *Jede kompakte Teilmenge eines normierten Raumes* $(X, ||\cdot||)$ *ist abgeschlossen und beschränkt.*

Beweis. Es sei K eine kompakte Teilmenge eines normierten Raumes $(X, ||\cdot||)$. Nach Lemma III-3.4 ist K beschränkt und somit $\complement K \neq \emptyset$. Sei also $x \in \complement K$. Für jedes $y \in K$ wählen wir disjunkte Umgebungen $U_y \in \mathcal{U}(x)$ und $V_y \in \mathcal{U}(y)$. Die Familie $\{V_y \colon y \in K\}$ ist eine offene Überdeckung von K. Somit gibt es endlich viele Punkte $y_1, \ldots, y_n \in K$ mit $K \subset \bigcup_{i=1}^{n} V_{y_i}$. Setzt man $U = \bigcap_{i=1}^{n} U_{y_i}$, folgt U ist offen und $U \in \mathcal{U}(x)$. Ferner gilt $U \subset \complement K$. Gäbe es nämlich ein $z \in K \cap U$, dann müßte z auch in einer Umgebung $V_{y_{i_z}}$ und damit auch in $U_{y_{i_z}}$ liegen, im Widerspruch zu $U_{y_{i_z}} \cap V_{y_{i_z}} = \emptyset$. Somit ist x ist innerer Punkt von $\complement K$ und $\complement K$ ist offen. $\square$

Die Umkehrung dieses Satzes ist jedoch in allgemeinen normierten Räumen falsch. Es jedoch folgende Charakterisierung:

Theorem III-3.6. *Eine Teilmenge K eines normierten Raumes ist kompakt genau dann, wenn K abgeschlossen ist und jede Folge $(x_n)_{n\geq 1} \subset K$ eine konvergente Teilfolge enthält.*

Beweis. Für den umfangreichen Beweis verweisen wir auf die weiterführende Literatur. □

Als Folgerung halten wir fest:

Korollar III-3.7. Abgeschlossene Teilmengen kompakter Mengen sind kompakt.

Beweis. Es sei $A \subset K \subset X$, A abgeschlossen und K kompakt. Dann hat jede Folge $(x_n) \subset A$ einen Häufungswert in K. Da A abgeschlossen ist, muß jeder Häufungswert sogar in A liegen, d.h. A ist kompakt. □

Wir können nun den Beweis der Äquivalenz aller Normen in einem endlichdimensionalen Vektorraum nachholen. Dazu benötigen wir folgenden Spezialfall eines wesentlich allgmeineren Resultates.

Lemma III-3.8. *Die Kugeln $\bar{K}_n(0,r)$ und deren Rand $S_n(0,r) = \{x \in \mathbb{K}^n \colon \|x\|_\infty = r\}$, $r > 0$, sind kompakt (in $(\mathbb{K}^n, \|\cdot\|_\infty)$).*

Beweis. Da $S_n(0,r)$ abgeschlossen ist (Übung), folgt die Behauptung aus Satz III-3.6 und Korollar II-5.6. □

Bemerkung III-3.9. Die Kompaktheit von $\bar{K}_n(0,r)$ und $S_n(0,r)$ wird vorerst nur für die Maximumnorm gezeigt. Wir werden später sehen, daß diese Behauptung für jede Norm in $\mathbb{K}^n$ gilt.

Theorem III-3.10. *Es sei $(X,\mathbb{K})$ ein endlichdimensionaler Vektorraum,* $\dim X = n$, *und $\mathcal{A} = (a_1,\dots,a_n)$ eine geordnete Basis von X. Bezeichnet man mit $[x]_{\mathcal{A}} = (\xi_1,\dots,\xi_n) \in \mathbb{K}^n$ den Koordinatenvektor von $x \in X$ bezüglich der Basis $\mathcal{A}$, wird durch $\varphi(x) = [x]_{\mathcal{A}}$ ein Isomorphismus $\varphi\colon X \to \mathbb{K}^n$ festgelegt. Statten man $\mathbb{K}^n$ mit der Norm $\|\cdot\|_\infty$ aus, sind φ und φ^{-1} stetig. Es gibt also Konstante* $0 < m \le M$, *so daß*

$$m\|x\| \le \|\varphi(x)\|_\infty \le M\|x\|$$

für alle $x \in X$ gilt.

Beweis. Für $x = \sum_{i=1}^n \xi_i a_i$ folgt $[x]_{\mathcal{A}} = (\xi_1,\dots,\xi_n)$ und somit

$$(*) \qquad \|x\| \le \sum_{i=1}^n |\xi_i|\ \|a_i\| \le (\sum_{i=1}^n \|a_i\|) \max_{1\le i\le n} |\xi_i| = (\sum_{i=1}^n \|a_i\|)\|\varphi(x)\|_\infty.$$

Es gilt also

$$m\|x\| \le \|\varphi(x)\|_\infty$$

mit $m = (\sum_{i=1}^n \|a_i\|)^{-1}$. Somit ist φ^{-1} stetig. Für den Nachweis der Stetigkeit von φ, also von

$$\exists M \ge 0 \forall x \in X\colon \|\varphi(x)\|_\infty \le M\|x\|,$$

nehmen wir an, diese Ungleichung wäre für jedes $M \ge 0$ falsch. Dann gibt es eine Folge (x_k) mit

$$\|\varphi(x_k)\|_\infty > k\|x_k\|, \qquad k \in \mathbb{N}.$$

Setzt man $z_k = \frac{x_k}{\|\varphi(x_k)\|_\infty}$ erhält man

$$\|z_k\| < \frac{1}{k},$$

also

$$\lim_{k\to\infty} z_k = 0.$$

Andererseits folgt auch $\|\varphi(z_k)\|_\infty = 1$, d.h. $\varphi(z_k) \in S_n(0,1)$. Da $S_n(0,1)$ in $(\mathbb{K}^n, \|\cdot\|_\infty)$ kompakt ist, gibt es eine Teilfolge $(\varphi(z_{\psi(k)})) \subset (\varphi(z_k))$ und $\xi \in S_n(0,1)$ mit

$$\lim_{k\to\infty} \varphi(z_{\psi(k)}) = \xi.$$

Setzt man $z = \varphi^{-1}(\xi)$, dann müsste $z \neq 0$ sein. Einerseits folgt dann wegen der Stetigkeit von φ^{-1}

$$\lim_{k\to\infty} \varphi^{-1}(\varphi(z_{\psi(k)})) = \varphi^{-1}(\xi) = z,$$

andererseits muß auch

$$\lim_{k\to\infty} \varphi^{-1}(\varphi(z_{\psi(k)})) = \lim_{k\to\infty} z_{\psi(k)} = 0$$

gelten,d.h. es müsste $z = 0$ sein. □

Wir sind nun in der Lage den Normenäquivalenzsatz III-3.11 zu beweisen.

THEOREM III-3.11. *Es sei X ein endlich dimensionaler Vektorraum. Dann sind alle Normen auf X äquivalent.*

BEWEIS. Es sei $\varphi\colon (X, \|\cdot\|) \to (\mathbb{K}^n, \|\cdot\|_\infty)$ der Isomorphismus aus Satz III-3.10. Dann definiert $\|\cdot\|_\varphi\colon X \to \mathbb{R}$

$$\|x\|_\varphi = \|\varphi(x)\|_\infty$$

eine Norm in X (Übung). Wegen Satz III-3.10 sind die Normen $\|\cdot\|$ und $\|\cdot\|_\varphi$ äquivalent. Dies gilt für jede Norm in X. Sind nun $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei beliebige Normen in X, existieren Konstante $0 < m_i < M_i$, $i = 1, 2$, mit

$$m_i\|x\|_i \le \|x\|_\varphi \le M_i\|x\|_i, \qquad i = 1, 2.$$

Daraus folgt

$$\frac{m_1}{M_2}\|x\|_1 \le \|x\|_2 \le \frac{M_1}{m_2}\|x\|_1.$$

□

Eine einfache Folgerung ist nun die Charakterisierung kompakter Mengen in $\mathbb{K}^n$.

THEOREM III-3.12. *Eine Teilmenge $K \subset \mathbb{K}^n$ ist kompakt genau dann, wenn K abgeschlossen und beschränkt ist.*

BEWEIS. Ohne Beschränkung der Allgemeinheit können wir $\mathbb{K}^n$ mit der Norm $\|\cdot\|_\infty$ ausstatten (warum?). Da K beschränkt ist, gilt für $r > 0$ hinreichend groß $K \subset \bar{K}(0, r)$. Die Behauptung folgt nun aus Lemma III-3.8 und Korollar III-3.7. □

Theorem III-3.13. *Jede lineare Abbildung L eines endlich dimensionalen Raumes X in einen endlich dimensionalen Raum Y ist stetig.*

Beweis. Es sei $\dim X = m$ und $\dim Y = n$, $\mathcal{A} = (a_1, \dots, a_m)$ eine geordnete Basis in X, $\mathcal{B} = (b_1, \dots, b_n)$ eine geordnete Basis in Y und $A \in \mathbb{K}^{n\times m}$ die Matrixdarstellung von L bezüglich der gewählten Basen in X bzw. Y. Dann gilt

$$Lx = y \qquad \Leftrightarrow \qquad A[x]_{\mathcal{A}} = [y]_{\mathcal{B}}.$$

Verwendet man die Isomorphismen $\varphi_X \colon X \to \mathbb{K}^m$, $\varphi_Y \colon Y \to \mathbb{K}^n$ aus Satz III-3.10, ist $Lx = y$ daher gleichwertig mit $(\varphi_Y^{-1} A \varphi_X)x = y$, d.h

$$L = \varphi_Y^{-1} A \varphi_X.$$

Somit ist L als Verkettung stetiger Abbildungen (Satz III-3.10 und Satz III-2.18) stetig. □

4. Stetigkeit und Kompaktheit

Wir beleuchten nun das Zusammenspiel von Stetigkeit und Kompaktheit.

Theorem III-4.1. *Das stetige Bild einer kompakten Menge ist kompakt.*

Beweis. Es sei $f \colon D \to Y$, $D \subset X$, stetig und $K \subset D$ kompakt. Um die Kompaktheit von $f(K)$ zu zeigen, gehen wir von einer Folge von Bildern (y_n), also $y_n = f(x_n)$ mit $(x_n) \subset K$ aus. Da K kompakt ist, existieren eine Teilfolge $(x_{\varphi(n)}) \subset (x_n)$ und $x \in K$, so daß $\lim_{n\to\infty} x_{\varphi(n)} = x$. Wegen der Stetigkeit von f gilt dann $\lim_{n\to\infty} f(x_{\varphi(n)}) = f(x)$, d.h. die Teilfolge $(y_{\varphi(n)}) \subset (y_n)$ konvergiert gegen $f(x) \in f(K)$. Somit ist $f(K)$ kompakt. □

Korollar III-4.2 (Satz von Weierstraß). Eine stetige, reellwertige Funktion nimmt auf kompakten Teilmengen Maximum und Minimum an.

Beweis. Es sei $f \colon D \to \mathbb{R}$, $D \subset X$, stetig und $K \subset D$ kompakt. Dann ist $f(K)$ kompakt, also abgeschlossen und beschränkt. Somit existiert $\alpha = \sup\{f(x) \colon x \in K\} < \infty$. Es sei $(x_n) \subset K$ eine Folge mit $\lim_{n\to\infty} f(x_n) = \alpha$. Wegen der Kompaktheit von K existiert eine Teilfolge $(x_{\varphi(n)}) \subset (x_n) \subset K$, welche gegen ein Element $x \in K$ konvergiert. Da f stetig ist, folgert man $\alpha = \lim_{n\to\infty} f(x_{\varphi(n)}) = f(x)$, also $\alpha \in f(K)$. □

Korollar III-4.3. Eine stetige Funktion $f \colon [a, b] \to \mathbb{R}$, $-\infty < a \le b < \infty$, nimmt auf $[a, b]$ Maximum und Minimum an.

Der Satz von Weierstraß ist falsch, wenn auf eine der drei Voraussetzungen verzichtet wird. Er bildet die theoretische Grundlage beispielsweise für die Existenz von Lösungen von zahlreichen Optimierungsproblemen.
Als weitere Anwendung von Satz III-4.1 können wir nun die Charakterisierung kompakter Mengen in endlichdimensionalen Räumen abschließen.

Theorem III-4.4. *Eine Teilmenge K eines endlichdimensionalen Vektorraumes X, ist kompakt genau dann, wenn K abgeschlossen und beschränkt ist.*

Beweis. Es sei $\dim X = n$, $\varphi\colon X \to \mathbb{K}^n$ der Isomorphismus aus Satz III-3.10 (bezüglich irgendeiner Norm in X) und $K \subset X$ abgeschlossen und beschränkt. Da φ eine beschränkte, lineare Abbildung ist, existiert eine Konstante $m > 0$ mit

$$\|\varphi(x)\|_\infty \leq m\|x\|, \qquad x \in X.$$

Dies hat die Beschränktheit von $\varphi(K) \subset \mathbb{K}^n$ zur Folge. $\varphi(K)$ ist auch abgeschlossen: es sei $(\xi_k) \subset \varphi(K)$ eine konvergente Folge mit Grenzwert $\xi \in \mathbb{K}^n$. Wegen der Stetigkeit von φ^{-1} konvergieren dann auch die Urbilder $x_k = \varphi^{-1}(\xi_k) \in K$ gegen $x = \varphi^{-1}(\xi) \in X$. Da K abgeschlossen ist, folgt $x \in K$ und somit $\xi = \varphi(x) \in \varphi(K)$. Also ist $\varphi(K) \subset \mathbb{K}^n$ abgeschlossen und beschränkt, und somit kompakt. Die Behauptung folgt nun aus $K = \varphi^{-1}(\varphi(K))$ und Satz III-4.1. □

Diese einfache Charakterisierung kompakter Mengen ist in Räumen unendlicher Dimension falsch, vgl. Beispiel III-4.6.
Ist der Definitionsbereich eine kompakte Menge K, definiert

$$\|f\|_\infty = \max\{\|f(x)\|_Y : x \in K\}$$

eine Norm in $C(K,Y)$ (Übung). Den normierten Raum $(C(K,Y), \|\cdot\|_\infty)$ bezeichnen wir wieder mit $C(K,Y)$.

Theorem III-4.5. *Es sei $K \subset X$ nicht leer und kompakt und Y vollständig. Dann ist $C(K,Y)$ vollständig.*

Beweis. Es sei (f_n) eine Cauchy Folge in $C(K,Y)$. Für $\varepsilon > 0$ existiert also ein Index $N(\varepsilon)$, so daß für alle n, $m \geq N(\varepsilon)$

$$\|f_n - f_m\|_\infty = \max\{\|f_n(x) - f_m(x)\|_Y : x \in K\} < \varepsilon,$$

also

$$\|f_n(x) - f_m(x)\|_Y \leq \varepsilon, \qquad \text{für alle } x \in K$$

gilt. Insbesondere ist also für jedes $x \in K$ die Folge $f_n(x)$ eine Cauchy Folge in Y. Wegen der Vollständigkeit von Y existiert der Grenzwert $\lim_{n\to\infty} f_n(x)$, den wir mit $f(x)$ bezeichnen. Somit existiert eine Grenzfunktion $f\colon K \to Y$. Führt man für ein festes $x \in K$ den Grenzübergang $n \to \infty$ durch, erhält man

$$(*) \qquad \|f(x) - f_m(x)\|_Y \leq \varepsilon$$

für alle $x \in K$ und $m \geq N(\varepsilon)$. Wir zeigen, daß f stetig ist. Es seien also $\varepsilon > 0$ und $x_0 \in K$ beliebig gewählt. Da $f_{N(\varepsilon)}$ stetig ist, gibt es ein $\delta > 0$, so daß $\|f_{N(\varepsilon)}(x) - f_{N(\varepsilon)}(x_0)\|_Y \leq \varepsilon$ für alle $x \in K(x_0,\delta) \cap D$ gilt. Wegen der Wahl von $N(\varepsilon)$ folgt daher aus $(*)$

$$\begin{aligned}\|f(x) - f(x_0)\|_Y &\leq \|f(x) - f_{N(\varepsilon)}(x)\|_Y + \|f_{N(\varepsilon)}(x) - f_{N(\varepsilon)}(x_0)\|_Y + \|f_{N(\varepsilon)}(x_0) - f(x_0)\|_Y \\ &\leq 2\|f - f_{N(\varepsilon)}\|_\infty + \|f_{N(\varepsilon)}(x) - f_{N(\varepsilon)}(x_0)\|_Y < 3\varepsilon.\end{aligned}$$

Diese Abschätzung gilt für alle $x \in K(x_0, \delta) \cap D$, somit ist f stetig. Die Funktion $x \to \|f(x) - f_m(x)\|_Y$ nimmt daher auf K das Maximum an und es gilt wegen $(*)$

$$\|f - f_m\|_\infty \leq \varepsilon$$

für $m \geq N(\varepsilon)$. Somit konvergiert die Funktionenfolge (f_m) im Sinne der Norm $\|\cdot\|_\infty$ gegen die Grenzfunktion f. $\square$

Beispiel III-4.6. Die abgeschlossene Einheitskugel in $(C([0,1], \mathbb{R}), ||\cdot||_\infty)$ ist nicht kompakt. Dazu betrachten wir die Folge stetiger Funktionen

$$f_n(x) = \begin{cases} -nx + 1, & 0 \leq x \leq \frac{1}{n}, \\ 0, & \text{sonst} \end{cases}, \quad n \in \mathbb{N}.$$

Natürlich ist $\|f_n\|_\infty = \max\{|f_n(x)| : x \in [0,1]\} = 1$. Für jedes $x \in [0,1]$ existiert der Grenzwert $\lim_{n\to\infty} f_n(x)$, welchen wir $f(x)$ nennen,

$$f(x) = \begin{cases} 1, & x = 0, \\ 0, & \text{sonst} \end{cases}$$

Als Grenzwert der Folge (f_n) oder einer beliebigen Teilfolge kommt daher nur die Funktion f in Frage Da f nicht stetig ist, kann die abgeschlossene Einheitskugel in $(C([0,1], \mathbb{R}), ||\cdot||_\infty)$ nicht kompakt sein. Als Übung überlege man, daß (f_n) keine Cauchyfolge sein kann.

5. Gleichmäßige Stetigkeit

Wir wenden uns nun einer wichtigen Variante des Stetigkeitsbegriffes zu. Dazu untersuchen wir vorerst ein einfaches Beispiel etwas genauer:

Beispiel III-5.1. Es sei $f\colon (0,1] \to \mathbb{R}$ die Abbildung $x \mapsto \frac{1}{x}$, $x_0 \in (0,1]$ und $\varepsilon > 0$. Eine einfache Rechnung ergibt

$$|f(x) - f(x_0)| = \frac{|x - x_0|}{x x_0},$$

für $0 < \delta < x_0$ und $|x - x_0| \leq \delta$ erhält man weiter

$$|f(x) - f(x_0)| \leq \frac{\delta}{x_0(x_0 - \delta)}.$$

Die Abschätzung kann nicht verbessert werden, da für $x = x_0 - \delta$ Gleichheit eintritt. Somit gilt $|f(x) - f(x_0)| < \varepsilon$ für $|x - x_0| < \delta$ soferne $\frac{\delta}{x_0(x_0-\delta)} < \varepsilon$. Das Supremum aller δ, die mit dieser Ungleichung verträglich sind, ist gegeben durch $\Delta(\varepsilon, x_0) = \frac{\varepsilon x_0^2}{1+\varepsilon x_0}$. Aus $\inf_{x_0 \in (0,1]} \Delta(\varepsilon, x_0) = 0$ folgt, daß es nicht möglich ist, zu $\varepsilon > 0$ ein $\delta = \delta(\varepsilon)$ anzugeben, welches für alle $x \in (0,1]$ die $\varepsilon\delta$-Charakterisierung der Stetigkeit verifiziert.

Definition III-5.2. *Es sei $f\colon D \to Y$, $D \subset X$ und $K \subset D$ nicht leer.*
f heißt ***gleichmäßig stetig*** *auf K, wenn*

$$\forall \varepsilon > 0\, \exists \delta = \delta(\varepsilon)\, \forall x, y \in K\colon ||x - y||_X < \delta \Rightarrow ||f(x) - f(y)||_Y < \varepsilon.$$

Insbesondere folgt, daß eine auf K gleichmäßig stetige Funktion an jeder Stelle $y \in K$ stetig ist. Man beachte jedoch, daß eine für alle $y \in K$ gültige passende δ Umgebung von y postuliert wird, welche die Erfordernisse der $\varepsilon\delta$-Charakterisierung der Stetigkeit von f in $y \in K$ erfüllt. Das vorhergehende Beispiel zeigt, daß diese zusätzliche Bedingung allgemein nicht erfüllbar ist. Gleichmäßige Stetigkeit einer Funktion auf K verlangt also echt mehr als die bloße Stetigkeit auf K.

Beispiel III-5.3. Lipschitz stetige Funktionen sind gleichmäßig stetig.

Durch Negation erhält man leicht folgende Charakterisierung:

$$f \text{ ist nicht gleichmäßig stetig auf } K \Leftrightarrow$$
$$\exists \varepsilon_0 > 0\, \exists (x_n)_{n\geq 1}, (y_n)_{n\geq 1} \subset K\, \forall n \in \mathbb{N}\colon ||x_n - y_n||_X < \tfrac{1}{n} \wedge ||f(x_n) - f(y_n)||_Y \geq \varepsilon_0.$$

Man verifiziere diese Bedingung in Beispiel III-5.1.

Theorem III-5.4. *Es sei $f\colon D \to Y$, $D \subset X$, stetig. Dann ist f auf jeder kompakten Teilmenge $K \subset D$ gleichmäßig stetig.*

Beweis. Es sei $K \subset D$ kompakt und f stetig, aber nicht gleichmäßig stetig auf K. Somit gibt es $\varepsilon_0 > 0$ und Folgen $(x_n)_{n\geq 1}$, $(y_n)_{n\geq 1} \subset K$ mit

$$||x_n - y_n||_X < \frac{1}{n} \text{ und } ||f(x_n) - f(y_n)||_Y \geq \varepsilon_0.$$

Da K kompakt ist, enthält (x_n) eine konvergente Teilfolge $(x_{\varphi(n)})$: Es gibt also ein $x \in K$ mit $\lim_{n\to\infty} x_{\varphi(n)} = x$. Aus

$$||y_{\varphi(n)} - x||_X \leq ||y_{\varphi(n)} - x_{\varphi(n)}||_X + ||x_{\varphi(n)} - x||_X \leq \frac{1}{\varphi(n)} + ||x_{\varphi(n)} - x||_X$$

folgt

$$\lim_{n\to\infty} y_{\varphi(n)} = x.$$

Aus der Abschätzung

$$\varepsilon_0 \leq ||f(x_{\varphi(n)}) - f(y_{\varphi(n)})||_Y \leq ||f(x_{\varphi(n)}) - f(x)||_Y + ||f(x) - f(y_{\varphi(n)})||_Y$$

folgt dann, daß f an der Stelle x nicht stetig sein kann. ☐

Eine Anwendung der gleichmäßigen Stetigkeit ist die Möglichkeit der stetigen Fortsetzung einer Funktion:

Theorem III-5.5. *Es sei $f\colon D \to Y$, $D \subset X$, gleichmäßig stetig und Y vollständig. Dann gibt es genau eine auf der abgeschlossenen Hülle $\bar{D}$ gleichmäßig stetige Funktion $F\colon \bar{D} \to Y$, welche auf D mit f übereinstimmt.*

Beweis. Es sei $x_0 \in \bar{D} \setminus D$, die Folge $(x_n) \subset D$ konvergiere gegen x_0. Wir zeigen, daß $(f(x_n))$ eine Cauchy Folge ist. Wir wählen also $\varepsilon > 0$ und bestimmen dazu $\delta > 0$ gemäß Definition III-5.2. Ferner bestimmen wir einen Index $N(\delta)$, so daß $\|x_n - x_0\| < \frac{\delta}{2}$ für alle $n \geq N(\delta)$ zutrifft. Für $n > m \geq N(\delta)$ gilt dann $\|x_n - x_m\| < \delta$ und somit wegen der gleichmäßigen Stetigkeit von f auch $\|f(x_n) - f(x_m)\| < \varepsilon$. Somit ist $(f(x_n))$ eine Cauchy Folge und es existiert $\lim_{x\to x_0} f(x)$ (Satz III-1.12). Dann kann f stetig nach x_0 fortgesetzt werden (Satz III-2.9). Die Funktion $F\colon \bar{D} \to Y$, $F(x) = f(x)$ für $x \in D$ und $F(x) = \lim_{y\to x} f(y)$, $x \in \bar{D} \setminus D$, ist also stetig. Ist $\tilde{F}$ eine weitere stetige Fortsetzung, dann stimmen F und $\tilde{F}$ auf D überein. Da D dicht in $\bar{D}$ ist, folgt $\tilde{F} = F$, Satz III-2.11. Es bleibt noch zu zeigen, daß F gleichmäßig stetig auf $\bar{D}$ ist. Zu $\varepsilon > 0$ wählen wir $\delta > 0$ entsprechend der gleichmäßigen Stetigkeit von f und $z \in \bar{D}$. Für u, $v \in K(z, \frac{\delta}{2}) \cap D$ folgt $\|u - v\| < \delta$ und somit

$$\|f(u) - f(v)\| < \varepsilon.$$

Führt man in dieser Ungleichung den Grenzübergang $v \to z$ durch, findet man

$$\|f(u) - F(z)\| \leq \varepsilon, \qquad \text{für alle } \|u - z\| < \frac{\delta}{2}.$$

Es seien nun x, und y beliebige Elemente aus $\bar{D}$ mit $\|x - y\| < \frac{\delta}{3}$ und u_x und u_y Elemente aus D mit $\|u_x - x\| < \frac{\delta}{3}$ bzw. $\|u_y - y\| < \frac{\delta}{3}$. Dann gilt

$$\|u_x - u_y\| \leq \|u_x - x\| + \|x - y\| + \|y - u_y\| < \delta$$

und somit

$$\begin{aligned}\|F(x) - F(y)\| &\leq \|F(x) - F(u_x)\| + \|F(u_x) - F(u_y)\| + \|F(u_y) - F(y)\| \\ &\leq \|F(x) - f(u_x)\| + \|f(u_x) - f(u_y)\| + \|f(u_y) - F(y)\| < 3\varepsilon,\end{aligned}$$

somit ist F gleichmäßig stetig. □

6. Globale Stetigkeit

Wir formulieren zuerst die Stetigkeit einer Abbildung mit Hilfe von Umgebungen.

Theorem III-6.1. *Es seien $(X, ||\cdot||_X)$ und $(Y, ||\cdot||_Y)$ normierte Räume und $f\colon D \to Y$, $D \subset X$. Äquivalent sind folgende Aussagen:*

(1) *f ist stetig in $x_0 \in D$.*
(2) *Zu jeder Umgebung V von $f(x_0)$ gibt es eine Umgebung U von x_0 mit $f(U \cap D) \subset V$.*

Beweis. Übung. □

Die globale Stetigkeit einer Funktion kann sehr elegant mit dem Begriff der offenen (abgeschlossenen) Mengen charakterisiert werden. Wir benötigen dazu folgende Verallgemeinerung dieser Konzepte.

Definition III-6.2. *Es sei $(X, ||\cdot||)$ ein normierter Raum und $D \subset X$. Eine Teilmenge $U \subset D$ heißt* ***relativ offen (abgeschlossen) in D****, wenn es eine offene (abgeschlossene) Menge $M \subset X$ gibt mit*

$$U = D \cap M.$$

Man überzeuge sich davon, daß eine Menge U genau dann offen in D ist, wenn $D \setminus U$ abgeschlossen in D ist

Beispiel III-6.3. Es sei $X = \mathbb{R}$ und $D = [0,1]$, $E = (0,1)$. Dann ist $[0, \frac{1}{2})$ relativ offen in D und $(0, \frac{1}{2}]$ relativ abgeschlossen in E. Natürlich ist auch D selber relativ offen in D.

Theorem III-6.4. *Es seien $(X, ||\cdot||_X)$ und $(Y, ||\cdot||_Y)$ normierte Räume und $f\colon D \to Y$, $D \subset X$. Folgende Aussagen sind äquivalent*

i) *f ist stetig auf D.*
ii) *Das Urbild offener Teilmengen in Y ist relativ offen in D.*
iii) *Das Urbild abgeschlossener Teilmengen in Y ist relativ abgeschlossen in D.*

Beweis. i) $\Rightarrow$ ii) Es sei f stetig auf X und V offen in Y. Zu zeigen ist: $f^{-1}(V) = \{x \in D\colon f(x) \in V\}$ ist offen in X. Dies ist trivial für $f^{-1}(V) = \emptyset$. Es sei also $x \in f^{-1}(V)$. Da f in x stetig ist, existiert eine offene Umgebung $U_x \in \mathfrak{U}_X(x)$ mit $f(U_x \cap D) \subset V$, d.h. $U_x \cap D \subset f^{-1}(V)$. Die Behauptung folgt nun aus der Darstellung

$$f^{-1}(V) = \bigcup \{U_x \cap D\colon x \in f^{-1}(V)\} = \bigcup \{U_x\colon x \in f^{-1}(V)\} \cap D,$$

und Satz II-1.9.
ii) $\Rightarrow$ iii) Es sei $A \subset Y$ abgeschlossen, d.h. $\complement A$ ist offen in Y. Somit ist $f^{-1}(\complement A)$ relativ offen in D, d.h es gibt eine offene Menge $O \subset X$ mit $f^{-1}(\complement A) = O \cap D$. Wegen $f^{-1}(\complement A) = D \setminus f^{-1}(A)$ folgt

$$f^{-1}(A) = D \setminus (D \setminus f^{-1}(A)) = D \setminus (O \cap D) = D \cap \complement O,$$

d.h. $f^{-1}(A)$ ist relativ abgeschlossen.
iii) $\Rightarrow$ i) Wir wählen $x_0 \in X$ und $V \in \mathfrak{U}_Y(f(x_0))$. Dann ist V° eine offene Umgebung von $f(x_0)$ und somit $f^{-1}(\complement V^\circ)$ relativ abgeschlossen in D, d.h es gibt eine abgeschlossene Menge $A \subset X$ mit $f^{-1}(\complement V^\circ) = A \cap D$. Wie vorhin folgt nun

$$f^{-1}(V^\circ) = D \setminus (D \setminus f^{-1}(V^\circ)) = D \setminus (A \cap D) = D \cap \complement A.$$

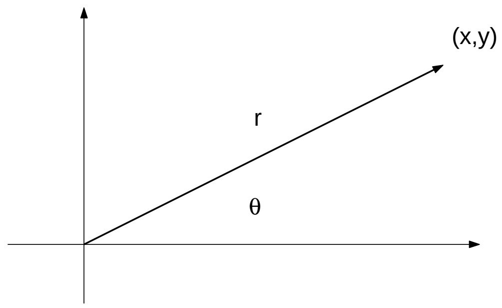
Setzt man $U = \complement A$, dann ist offenbar U eine offene Umgebung von x_0 mit $f(U \cap D) \subset V^\circ \subset V$, d.h. f ist stetig in x_0. $\square$

Definition III-6.5. *Eine bijektive Abbildung f einer Teilmenge U eines normierten Raumes $(X, ||\cdot||_X)$ auf eine Teilmenge V eines normierten Raumes $(Y, ||\cdot||_Y)$ heißt **Homöomorphismus**, wenn $f\colon U \to V$ und $f^{-1}\colon V \to U$ stetig sind. Man sagt auch, f bildet U homöomorph auf V ab, bzw. U ist homöomorph zu V.*

Beispiel III-6.6 (Ebene Polarkoordinaten). Wir betrachten die Abbildung

$$\phi\colon \begin{cases} [0,\infty) \times [-\pi,\pi] \to \mathbb{R}^2, \\ (r,\varphi) \mapsto (r\cos\varphi, r\sin\varphi). \end{cases}$$

Man nennt (r,φ) Polarkoordinaten des Punktes in $\mathbb{R}^2$ mit den kartesischen Koordinaten $(x,y) = (r\cos\varphi, r\sin\varphi)$.



Die Abbildung ϕ ist stetig und surjektiv. Sie ist aber nicht injektiv, denn es gilt $\phi(0,\varphi) = (0,0)$ für alle $\varphi \in [-\pi,\pi]$ und $\phi(r,\pi) = \phi(r,-\pi) = (-r,0)$ für alle $r > 0$. Offenbar bildet ϕ jedoch den halboffenen Halbstreifen $(0,\infty) \times (-\pi,\pi]$ bijektiv auf die punktierte Ebene $\mathbb{R}^2 \setminus \{(0,0)\}$ ab. Definiert man $h\colon \mathbb{R} \to \mathbb{R}$ durch

$$h(y) = \begin{cases} 1, & y \geq 0, \\ -1, & y < 0, \end{cases}$$

kann man die Umkehrabbildung $\phi^{-1}\colon \mathbb{R}^2 \setminus \{(0,0)\} \to (0,\infty) \times (-\pi,\pi]$ kompakt schreiben in der Form

$$\phi^{-1}(x,y) = (\sqrt{x^2+y^2}, h(y) \arccos \frac{x}{\sqrt{x^2+y^2}}).$$

Als Übung beweise man die Identitäten $\phi \circ \phi^{-1} = id_{\mathbb{R}^2 \setminus \{(0,0)\}}$ und $\phi^{-1} \circ \phi = id_{(0,\infty)\times(-\pi,\pi]}$. Allerdings besitzt ϕ^{-1} in keinem Punkt des Halbstrahls $\{(-r,0)\colon r > 0\}$ einen Grenzwert: es gilt nämlich

$$\lim_{n\to\infty} \phi^{-1}(-r, \frac{1}{n}) = (r,\pi),$$
$$\lim_{n\to\infty} \phi^{-1}(-r, -\frac{1}{n}) = (r,-\pi).$$

Man überzeuge sich davon, daß ϕ^{-1} auf $\mathbb{R}^2 \setminus \{(-r,0)\colon r > 0\}$ stetig ist. ϕ bildet also $(0,\infty) \times (-\pi,\pi)$ homöomorph auf die geschlitzte Ebene $\mathbb{R}^2 \setminus \{(-r,0)\colon r > 0\}$ ab.

KAPITEL IV

Reelle Funktionen

1. Zwischenwertsatz

Im folgenden untersuchen wir die speziellen Eigenschaften von reellen Funktionen, deren Definitionsbereich ein Intervall ist. Wir beginnen mit einer Charakterisierung von Intervallen.

THEOREM IV-1.1. *Es sei $A \subset \mathbb{R}$ und $A \neq \emptyset$. Äquivalent sind folgende Aussagen:*

a) *A ist ein Intervall.*
b) *Mit je zwei Punkten von A liegt auch deren Verbindungsstrecke in A, d.h. $\forall x, y \in A\colon x \leq y \Rightarrow [x, y] \subset A$.*

BEWEIS. $a \Rightarrow b$ trivial
$b \Rightarrow a$: Wir unterscheiden vier Fälle, je nachdem ob A nach unten oder nach oben beschränkt ist oder ob dies nicht der Fall ist.
Fall 1: A ist beschränkt nach unten, unbeschränkt nach oben. Wir zeigen, daß in diesem Falle A ein Intervall des Typs (α, ∞) oder $[\alpha, \infty)$ ist, $\alpha = \inf A$. Da α eine untere Schranke von A ist, folgt $A \subset [\alpha, \infty)$. Die Behauptung folgt dann aus der Inklusion $(\alpha, \infty) \subset A$. Es sei also $r \in (\alpha, \infty)$ beliebig gewählt. Nach Satz II-I-6.4 (modifiziert für inf) gibt es ein $a \in A$ mit $\alpha \leq a < r$. Andererseits ist A nach oben unbeschränkt, somit gibt es $b \in A$ mit $r < b$, also $a < r < b$. Aus der Bedingung b) folgt $[a, b] \subset A$ und somit a fortiori auch $r \in A$.
Fall 2: A ist beschränkt nach oben, unbeschränkt nach unten. Dies läßt sich durch Übergang auf die Menge -A auf Fall 1 zurückführen.
Fall 3: A ist beschränkt. Analog zum Fall 1 zeigt man, daß mit $\alpha := \inf A$ und $\beta := \sup A$ die Inklusionen

$$(\alpha, \beta) \subset A \subset [\alpha, \beta]$$

gelten. Somit ist A eines der Intervalle $(\alpha, \beta), (\alpha, \beta], [\alpha, \beta)$ oder $[\alpha, \beta]$.
Fall 4: A ist nach oben und unten unbeschränkt. Für alle $r \in \mathbb{R}$ gibt es dann $a, b \in A$, $a \leq b$, mit $r \in [a, b] \subset A$. Somit ist $A = \mathbb{R}$. □

Wir zeigen nun, daß eine stetige, reelle Funktion, die in den Endpunkten eines *abgeschlossenen* Intervalles Werte verschiedenen Vorzeichens annimmt, in diesem Intervall mindestens eine Nullstelle besitzt.

THEOREM IV-1.2. *Es sei $I = [a, b]$, $-\infty < a < b < \infty$, $f\colon I \to \mathbb{R}$ stetig und $f(a)f(b) < 0$. Dann gibt es ein $\alpha \in (a, b)$ mit $f(\alpha) = 0$.*

BEWEIS. Wir nehmen an, es sei $f(a) > 0$ und $f(b) < 0$ (anderenfalls geht man von f auf $-f$ über). Die Menge $A = \{x \in [a, b] : f(x) \geq 0\}$ ist nicht leer ($a \in A$) und beschränkt, somit existiert $\alpha = \sup A$. Es sei $(x_n)_{n \geq 1} \subset A$ eine Folge mit $\lim_{n\to\infty} x_n = \alpha$, dann gilt $\alpha \in I$. Das Supremum α liegt aber sogar in A. Aus $f(x_n) \geq 0$, $n \in \mathbb{N}$, folgt nämlich mit Satz II-4.4 und Satz III-2.7

$$f(\alpha) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) \geq 0.$$

Wegen $f(b) < 0$ ist insbesondere $\alpha < b$ und $f(y) < 0$ für alle $y \in (\alpha, b]$. Ist daher $(y_n)_{n\geq 1} \subset (\alpha, b]$ eine Folge mit $\lim_{n\to\infty} y_n = \alpha$, muß notwendigerweise auch $f(\alpha) = \lim f(y_n) \leq 0$, insgesamt also $f(\alpha) = 0$ gelten. $\square$

Dieses Resultat bildet die theoretische Grundlage für das **Bisektionsverfahren** zur Berechnung der Nullstelle einer stetigen Funktion.
Wir erweitern nun die Anwendbarkeit dieses Satzes auf beliebige Intervalle.

THEOREM IV-1.3 (Zwischenwertsatz). *Es sei $I \subset \mathbb{R}$ ein Intervall und $f \in C(I, \mathbb{R})$. Dann ist $f(I)$ ein Intervall.*

BEWEIS. Es sei $x, y \in f(I)$, $x < y$. Nach Satz IV-1.1 genügt es, $[x, y] \subset f(I)$ nachzuweisen. Zu diesem Zweck wählen wir $k \in \mathbb{R}$ mit $x < k < y$. Wegen $x, y \in f(I)$ gibt es $a, b \in I$ mit $x = f(a)$ und $y = f(b)$. Da $x \neq y$ und f eine Funktion ist, gilt $a \neq b$. Wir bezeichnen nun mit J das abgeschlossene Intervall mit den Endpunkten a und b. Da I ein Intervall ist, folgt $J \subset I$. Wir definieren nun $g\colon J \to \mathbb{R}$ durch

$$g : \begin{cases} J \to \mathbb{R} \\ x \mapsto f(x) - k. \end{cases}$$

Da $f|_J$ stetig ist, ist auch g stetig. Darüberhinaus erfüllt g die Bedingungen

$$\begin{aligned} g(a) &= f(a) - k = x - k < 0, \\ g(b) &= f(b) - k = y - k > 0. \end{aligned}$$

Satz IV-1.2 sichert die Existenz von $c \in J$ mit $g(c) = 0$. Das bedeutet aber $f(c) = k$, dh. $k \in f(I)$. $\square$

BEMERKUNG IV-1.4. Aus dem Beweis wird klar, warum dieses Ergebnis Zwischenwertsatz genannt wird: Jede reelle Zahl zwischen zwei Bildern ist selbst ein Bildelement.

Es ist leicht, Beispiele zu finden, die belegen, daß man beim Zwischenwertsatz weder auf die Stetigkeit von f, noch auf die Voraussetzung, daß I ein Intervall ist, verzichten kann.

KOROLLAR IV-1.5. Es sei $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$ stetig. Für alle $x \in I$ sei $f(x) \neq 0$. Dann ist entweder $f(I) \subset (0, \infty)$ oder $f(I) \subset (-\infty, 0)$.

BEWEIS. Wir führen einen Widerspruchsbeweis: angenommen es gäbe $a, b \in I$ mit $f(a) < 0$ und $f(b) > 0$. Nach dem Zwischenwertsatz wäre $0 \in f(I)$ im Widerspruch zu $f(x) \neq 0$ für $x \in I$. $\square$

2. Monotonie und Stetigkeit

DEFINITION IV-2.1. *Es sei $f\colon D \to \mathbb{R}$ und $D \subset \mathbb{R}$.*

i) *f heißt **monoton wachsend (fallend)*** $\underset{Def}{\Leftrightarrow}$

$$\forall x, y \in D\colon x < y \Rightarrow f(x) \leq f(y) \quad (f(x) \geq f(y)).$$

ii) *f heißt* ***streng (strikt) monoton wachsend (fallend)*** $\underset{Def}{\Leftrightarrow}$

$$\forall x, y \in D\colon x < y \Rightarrow f(x) < f(y) \quad (f(x) > f(y)).$$

iii) *f heißt* ***(streng) monoton*** $\underset{Def}{\Leftrightarrow}$ *f ist (streng) monoton wachsend oder fallend.*

Strenge Monotonie einer Funktion ist deswegen von Bedeutung, da sie die Existenz der Umkehrfunktion sicherstellt.

THEOREM IV-2.2. *Es sei $f\colon D \to \mathbb{R}$, $D \subset \mathbb{R}$, streng monoton. Dann ist f injektiv, die Umkehrfunktion existiert und ist ebenfalls streng monoton im selben Sinn wie f.*

BEWEIS. Eine streng monotone Funktion ist injektiv, somit existiert die Umkehrfunktion. O.B.d.A. sei f streng monoton wachsend (gegebenenfalls ersetze man f durch $-f$). Angenommen f^{-1} wäre nicht streng monoton wachsend, d.h.

$$\exists y_1, y_2 \in f(D)\colon y_1 < y_2 \wedge f^{-1}(y_1) \geq f^{-1}(y_2).$$

Da $f^{-1}(y_i) \in D$, $i = 1, 2$, folgt aus der strengen Monotonie von f die Beziehung $y_2 = f(f^{-1}(y_2)) \leq f(f^{-1}(y_1)) = y_1$ im Widerspruch zu $y_1 < y_2$. □

Allerdings ist die strenge Monotonie nur eine hinreichende Bedingung für Injektivität:

BEISPIEL IV-2.3. Wir definieren $f\colon [0, 1] \to [0, 1]$ durch

$$f(x) = \begin{cases} x & x \in \mathbb{Q} \cap [0, 1] \\ 1 - x & x \in [0, 1] \setminus \mathbb{Q}. \end{cases}$$

Diese Funktion ist sogar bijektiv (zum Beweis betrachte man die Restriktionen $f|_{\mathbb{Q}\cap[0,1]}$ und $f|_{[0,1]\setminus\mathbb{Q}}$), aber nicht monoton: Für $\frac{1}{2} < \frac{\sqrt{2}}{2} < \frac{3}{4}$ folgt $f(\frac{1}{2}) = \frac{1}{2} > f(\frac{\sqrt{2}}{2})$, $\frac{3}{4} = f(\frac{3}{4}) > f(\frac{\sqrt{2}}{2})$.

Wir merken an, daß die Menge der monotonen Funktionen keinen Vektorraum bilden. Die Charakterisierung des Bildbereiches einer stetigen monotonen Funktion auf einem Intervall ist besonders einfach.

THEOREM IV-2.4. *Es sei $I \subset \mathbb{R}$ ein Intervall mit den Randpunkten $a = \inf I$ und $b = \sup I$ und $f \in C(I, \mathbb{R})$ monoton. Dann ist $f(I)$ ein Intervall mit den Randpunkten $\alpha = \inf f$ und $\beta = \sup f$. Ist f sogar streng monoton wachsend, dann sind I und $f(I)$ Intervalle vom selben Typ, also*

$$\alpha \in f(I) \Leftrightarrow a \in I \quad \textit{und} \quad \beta \in f(I) \Leftrightarrow b \in I.$$

Eine analoge Aussage gilt für streng monoton fallende Funktionen.

BEWEIS. Da f stetig ist, ist $f(I)$ ein Intervall (Zwischenwertsatz). Der weitere Beweis verläuft vollkommen analog zum Beweis von Satz IV-1.1: Man weist die Inklusion $(\alpha, \beta) \subset f(I) \subset [\alpha, \beta]$ nach (mit geeigneten Modifikationen falls $\alpha, \beta \in \bar{\mathbb{R}}$). Es sei nun f streng monoton wachsend und $f(I)$ links abgeschlossen, dh. $\alpha = \inf f \in f(I)$. Somit gibt es ein $a \in I$ mit $f(a) = \alpha$. Wegen der Injektivität von f folgt

$$\forall y \in I\colon y \neq a \Rightarrow f(y) > f(a),$$

und da f^{-1} ebenfalls streng monoton wächst ergibt sich

$$\forall y \in I\colon y \neq a \Rightarrow y > a, \text{ dh. } a = \min I.$$

I ist somit links abgeschlossen. Umgekehrt setzen wir nun diese Eigenschaft für I voraus, dh. $a = \min I$. Somit gilt $f(x) \geq f(a)$ für alle $x \in I$, dh. $f(I)$ ist links abgeschlossen. Die Behauptung, daß die rechten Randpunkte von $f(I)$ und I beide entweder zu $f(I)$ bzw. I gehören oder nicht, kann analog bewiesen werden. □

Satz IV-2.2 zeigt, daß die strenge Monotonie einer Funktion hinreichend (wegen Beispiel IV-2.3 aber nicht notwendig) ist für deren Injektivität. Im Raum der stetigen Funktionen allerdings ist diese Bedingung sogar notwendig.

Lemma IV-2.5. *Es sei $I = [a, b]$, $-\infty < a < b < \infty$. Jede stetige, injektive Abbildung $f : I \to \mathbb{R}$ ist streng monoton.*

Beweis. Wegen der Injektivität von f ist $f(a) \neq f(b)$. Wir nehmen an, es sei $f(a) < f(b)$ (anderenfalls ersetzt man f durch $-f$).
1. Schritt: $f(a) < f(x) < f(b)$ für alle $x \in (a, b)$. Angenommen, es gäbe ein $\xi \in (a, b)$ mit $f(\xi) \leq f(a)$ oder $f(\xi) \geq f(b)$. Wir betrachten zuerst den Fall $f(\xi) \leq f(a)$. Da f injektiv ist und $\xi > a$, muß $f(\xi) < f(a)$ gelten. Somit ist $f(a)$ ein Zwischenwert der stetigen Funktion $f|_{[\xi,b]}$. Der Zwischenwertsatz für $f|_{[\xi,b]}$ sichert die Existenz von $\eta \in (\xi, b)$ mit $f(\eta) = f(a)$. Wegen $a < \xi < \eta$ ergibt sich ein Widerspruch zur Injektivität von f, es muß also $f(a) < f(x)$ für alle $x \in (a, b)$ gelten. Analog schließt man die Möglichkeit $f(\xi) \geq f(b)$ aus.
2. Schritt: $x < y \Rightarrow f(x) < f(y)$ für alle $x, y \in [a, b]$. Die Behauptung ist wahr für $x = a$, $y = b$. Sie ergibt sich in den Fällen $x = a < y < b$, $a < x < y = b$ aus Schritt 1. Schließlich sei $a < x < y < b$. Für das Tripel $a < y < b$ folgt mit Schritt 1: $f(a) < f(y) < f(b)$. Ersetzt man somit in Schritt 1 f durch die stetige Funktion $\tilde{f} := f|_{[a,y]}$ folgt aus $a < x < y$ die Ungleichung $\tilde{f}(a) < \tilde{f}(x) < \tilde{f}(y)$, dh. $f(a) < f(x) < f(y)$. □

Wir erweitern nun den Gültigkeitsbereich dieses Ergebnisses auf beliebige Intervalle:

Theorem IV-2.6. *Es sei $I \subset \mathbb{R}$ ein Intervall. Jede stetige, injektive Abbildung $f : I \to \mathbb{R}$ ist streng monoton.*

Beweis. Es sei $r, s \in I : r < s$ und o.B.d.A $f(r) < f(s)$ und somit f nach Lemma IV-2.5 streng monoton steigend auf $[r, s]$. Es sei nun $x, y \in I$ und $x < y$ Dann ist f auf $[\min\{r, x\}, \max\{s, y\}]$ streng monoton steigend und somit $f(x) < f(y)$. □

Im Raum der stetigen Funktionen, deren Definitionsbereich ein Intervall ist, sind die injektiven Funktionen also genau die streng monotonen. Es ist naheliegend, die Stetigkeit der Umkehrfunktion zu untersuchen. Es ist ein überraschendes Resultat, daß in diesem Falle die Stetigkeit von f^{-1} automatisch gesichert ist (unabhängig von der Stetigkeit von f). Die Grundlage ist die anschauliche Eigenschaft, daß eine monotone Funktion stetig sein muß, wenn ihr Bild ein Intervall ist.

Theorem IV-2.7. *Es sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}$, monoton. Ist $f(D)$ ein Intervall, dann ist f stetig.*

Beweis. O.B.d.A sei f monoton wachsend. Wir führen einen Widerspruchsbeweis und nehmen daher an, daß zwar $f(D)$ ein Intervall ist, es aber eine Stelle $x_0 \in D$ gibt, an der f nicht stetig ist, dh. es gibt ein $\varepsilon_0 > 0$ und eine Folge $(x_n) \subset D$, welche zwar gegen x_0 konvergiert, für deren Bilder jedoch $|f(x_n) - f(x_0)| > \varepsilon_0$ für alle $n \in \mathbb{N}$ gilt (warum darf man „$>$" behaupten?). Es gibt also eine Teilfolge $(x_{\varphi(n)}) \subset (x_n)$, die einer der beiden Bedingungen

$$f(x_{\varphi(n)}) < f(x_0) - \varepsilon_0 < f(x_0) \text{ oder } f(x_0) < f(x_0) + \varepsilon_0 < f(x_{\varphi(n)}), \quad n \in \mathbb{N},$$

genügt. Es sei z.B. die zweite erfüllt. Nach Voraussetzung ist $f(D)$ ein Intervall. Daher muß $[f(x_0), f(x_{\varphi(1)})] \subset f(D)$ und daher erst recht $f(x_0) + \varepsilon_0 \in f(D)$ gelten. Es gibt also $\xi \in D$ mit $f(x_0) + \varepsilon_0 = f(\xi)$. Aus $f(x_0) < f(\xi) < f(x_{\varphi(n)})$, $n \in \mathbb{N}$, folgert man wegen der Monotonie von f, daß $x_0 < \xi < x_{\varphi(n)}$, $n \in \mathbb{N}$, zutreffen muß, also gilt auch $x_0 < \xi \leq \lim x_{\varphi(n)} = x_0$, ein Widerspruch. □

Theorem IV-2.8 (Satz von der Umkehrfunktion). *Es sei $I \subset \mathbb{R}$ ein Intervall und $f \colon I \to \mathbb{R}$ streng monoton. Dann existiert die Umkehrfunktion $f^{-1} \colon f(I) \to I$, f^{-1} ist streng monoton (im selben Sinne wie f) und stetig.*

Beweis. Satz IV-2.2 und Satz IV-2.7. □

Oft findet man eine schwächere Formulierung des Umkehrsatzes.

Theorem IV-2.9. *Es sei $I \subset \mathbb{R}$ ein Intervall und $f \in C(I, \mathbb{R})$ injektiv. Dann ist $f^{-1} \colon f(I) \to I$ stetig.*

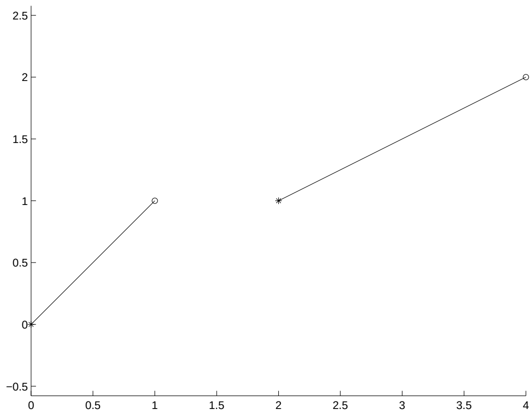
Beweis. Satz IV-2.6, IV-2.8. □

Der Satz von der Umkehrfunktion zeigt, daß einzig der Umstand, daß $D(f)$ kein Intervall ist, verhindern kann, daß die Umkehrfunktion einer streng monotonen Funktion stetig ist.

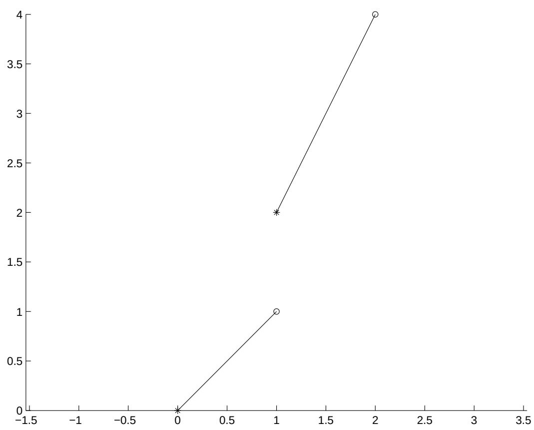
Beispiel IV-2.10.

$$f(x) = \begin{cases} x, & x \in [0,1), \\ \frac{1}{2}x, & x \in [2,4], \end{cases} \qquad \operatorname{bild} f = [0,1) \cup [1,2]$$

$$f^{-1}(x) = \begin{cases} x, & x \in [0,1), \\ 2x, & x \in [1,2]. \end{cases}$$



Graph von $f(x)$

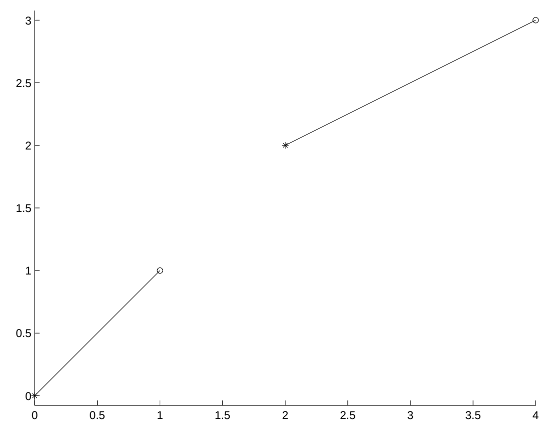


Graph von $f^{-1}(x)$

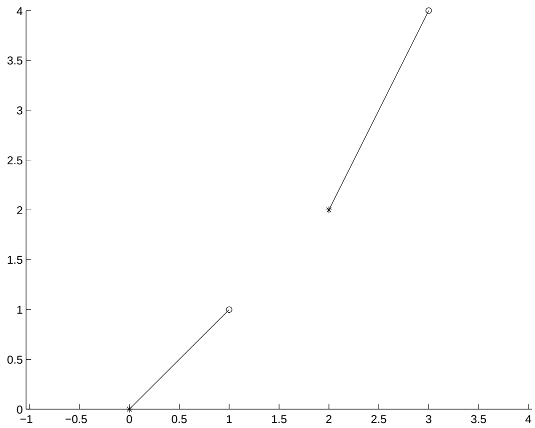
Die Bedingung „def f ist ein Intervall“ ist jedoch keineswegs notwendig für die Stetigkeit von f^{-1}: Wir ändern Beispiel IV-2.10 geringfügig ab:

$$f(x) = \begin{cases} x, & x \in [0,1), \\ 1 + \frac{1}{2}x, & x \in [2,4], \end{cases} \quad f^{-1}(x) = \begin{cases} x, & x \in [0,1), \\ 2x-2, & x \in [2,3]. \end{cases}$$

Man beachte, daß $f^{-1}(x)$ stetig ist und streng monoton wächst.



Graph von $f(x)$



Graph von $f^{-1}(x)$

3. Funktionenfolgen und Funktionenreihen

Wir wenden uns nun einer weiteren, überaus flexiblen Methode zu, aus bekannten Funktionen neue Funktionen aufzubauen. Es sei (f_n) eine Folge von *Funktionen* auf D. Existiert für alle $x \in D$ der Grenzwert $\lim_{n\to\infty} f_n(x)$, wird durch $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in D$, eine Abbildung auf D erklärt. Von Interesse sind die Eigenschaften der Grenzfunktion, insbesondere wollen wir Bedingungen finden, die gewährleisten, daß sich die Stetigkeit von f_n auf die Grenzfunktion überträgt. Zuerst untersuchen wir den angedeuteten Grenzprozeß:

Definition IV-3.1. *Es sei $(f_n)_{n\geq 1}$ eine Folge von Funktionen $f_n\colon D \to Y$, $D \subset X$.*

(f_n) heißt ***punktweise konvergent*** *gegen die Grenzfunktion $f\colon D \to Y$, $\lim_{n\to\infty} f_n = f$, pw. $\underset{Def}{\Leftrightarrow}$*

i) $\forall x \in D\colon \exists \lim_{n\to\infty} f_n(x)$

ii) $\forall x \in D\colon f(x) := \lim_{n\to\infty} f_n(x)$

Wegen der Eindeutigkeit des Grenzwertes der Zahlenfolgen $(f_n(x))$ ist auch die Grenzfunktion eindeutig bestimmt. Ist Y vollständig, folgt aus der Definition, daß die Funktionenfolge (f_n) genau dann punktweise konvergiert, wenn für alle $x \in D$ die Folge der Bilder $(f_n(x))$ eine Cauchy-Folge ist. Die üblichen Rechenregeln für Grenzwerte übertragen sich auf punktweise konvergente Funktionenfolgen.

Beispiel IV-3.2. (1) $D = [0,1]$, $f_n(x) = x^n$. Die Grenzfunktion ist

$$f(x) = \begin{cases} 0, & x \in [0,1), \\ 1, & x = 1. \end{cases}$$

Zum Beweis greift man auf Satz II-3.3 zurück. Man beachte, daß f unstetig, jedes f_n aber stetig ist. Wir merken an, daß $0 \le x^n < \varepsilon < 1$ für alle $x \in [0,1)$, und für alle $n > \left[\frac{\log \varepsilon}{\log x}\right] + 1 =: N(\varepsilon, x)$ gilt. $N(\varepsilon, x)$ ist optimal.

(2) $D = [0,2]$.

$$f_n(x) = \begin{cases} nx, & x \in [0, \frac{1}{n}], \\ 2 - nx, & x \in (\frac{1}{n}, \frac{2}{n}], \\ 0, & x \in (\frac{2}{n}, 2]. \end{cases}$$

Es gilt: $\lim_{n\to\infty} f_n = 0$. Zum Beweis fixieren wir ein beliebiges $x \in (0,2]$ und bestimmen $N_x \ge \frac{2}{x}$. Für alle $n \ge N_x$ ist dann $f_n(x) = 0$. Je kleiner x gewählt wurde, desto größer ist also N_x. Man beachte, daß in diesem Beispiel sowohl f, als auch jede approximierende Funktion f_n stetig ist.

(3) $D = \mathbb{R}$, $f_n(x) = \frac{1}{n}\lfloor nx \rfloor$, $n \in \mathbb{N}$ (vgl. Satz I-6.10). Wir behaupten: $\lim_{n\to\infty} f_n = id$. Für jedes $n \in \mathbb{N}$, $x \in \mathbb{R}$ gibt es $k \in \mathbb{N}$ mit

$$\lfloor nx \rfloor = k \Leftrightarrow k \le nx < k+1 \Leftrightarrow \frac{k}{n} \le x < \frac{k+1}{n}.$$

Somit folgt für $x \in [\frac{k}{n}, \frac{k+1}{n})$, $\varepsilon > 0$,

$$|x - f_n(x)| = |x - \frac{1}{n}\lfloor nx \rfloor| = |x - \frac{k}{n}| < \frac{1}{n} < \varepsilon.$$

Die letzte Ungleichung gilt für alle $n \ge N(\varepsilon) > \frac{1}{\varepsilon}$. In diesem Beispiel sind die approximierenden Funktionen f_n unstetig, trotzdem ist die Grenzfunktion stetig.

(4) $D = \mathbb{R}^+$, $f_n(x) = \frac{x}{1+nx}$. Die Abschätzung für alle $x \in \mathbb{R}^+$, $\varepsilon > 0$

$$0 \le f_n(x) \le \frac{1}{n}\frac{nx}{1+nx} \le \frac{1}{n} < \varepsilon,$$

gültig für $n \ge N(\varepsilon) > \frac{1}{\varepsilon}$, zeigt $\lim_{n\to\infty} f_n = 0$. In diesem Beispiel sind alle f_n und die Grenzfunktion f stetig.

Diese Beispiele illustrieren, daß der punktweise Grenzwert stetiger Funktionen nicht notwendig stetig sein muß. Umgekehrt kann eine Folge unstetiger Funktionen einen stetigen Grenzwert besitzen. Worauf läuft die Stetigkeit der Grenzfunktion an einer Stelle $x_0 \in D$ eigentlich hinaus? Es sei $(f_n) \subset C(D, \mathbb{R})$, $f = \lim f_n$, $(x_n) \subset D$: $\lim_{n\to\infty} x_n = x_0$. Nach Satz III-2.7 impliziert die Stetigkeit von f in x_0: $f(x_0) = \lim_{k\to\infty} f(x_k)$, dh.

$$f(x_0) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty}\left[\lim_{n\to\infty} f_n(x_k)\right] = \lim_{n\to\infty}\left[\lim_{k\to\infty} f_n(x_k)\right] = \lim_{n\to\infty} f_n(x_0).$$

Die Stetigkeit der Grenzfunktion wäre also gewährleistet, wenn die Vertauschung der beiden Grenzwerte gerechtfertigt wäre. Dies führt zu folgendem, stärkeren Konvergenzbegriff:

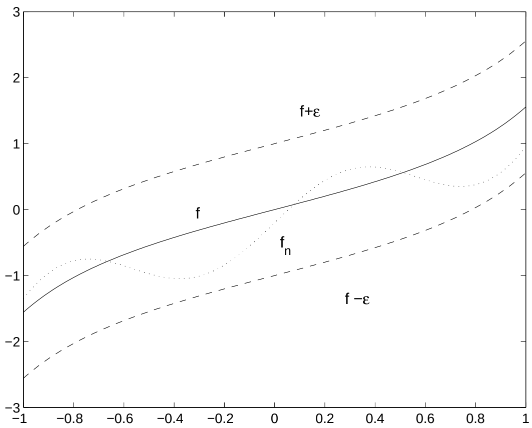
DEFINITION IV-3.3. $(f_n)_{n\geq 1}$ *sei eine Folge von Funktionen* $f_n\colon D \to Y$ *und* $f\colon D \to Y$, $D \subset X$.
(f_n) *heißt* ***gleichmäßig konvergent*** *gegen die Grenzfunktion* f, $\lim_{n\to\infty} f_n = f$, *glm., genau dann, wenn*

$$\forall \varepsilon > 0\ \exists N(\varepsilon) \in \mathbb{N}\ \forall x \in D\ \forall n \geq N(\varepsilon)\colon \|f_n(x) - f(x)\| < \varepsilon.$$

Wir erinnern: Punktweise Konvergenz von (f_n) gegen f bedeutet

$$\forall \varepsilon > 0\ \forall x \in D\ \exists N(\varepsilon, \boldsymbol{x})\ \forall n \geq N(\varepsilon, x)\colon \|f_n(x) - f(x)\| < \varepsilon.$$

Man beachte die Reihenfolge der Quantoren! Die Folge (f_n) konvergiert also gleichmäßig genau dann, wenn ein Index $N(\varepsilon)$ **unabhängig** von $x \in D$ gefunden werden kann, sodaß sämtliche Funktionswerte von f_n, $n \geq N(\varepsilon)$, in einem „ε-Schlauch“ um f liegen.



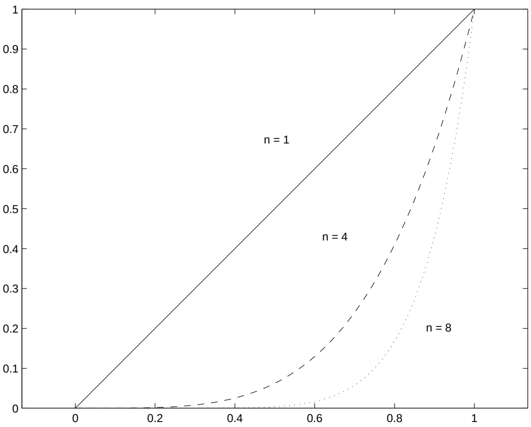
Die Folge (f_n) konvergiert demnach *nicht* gleichmäßig gegen f genau dann, wenn

$$\exists \varepsilon_0 > 0\ \forall n \in \mathbb{N}\ \exists x_n \in D\ \exists k \geq n\colon \|f_k(x_n) - f(x_n)\| \geq \varepsilon_0$$
$$\Leftrightarrow$$
$$(*) \qquad \exists \varepsilon_0 > 0 \exists (f_{\varphi(n)}) \subset (f_n) \exists (x_n)\colon \|f_{\varphi(n)}(x_n) - f(x_n)\| \geq \varepsilon_0.$$

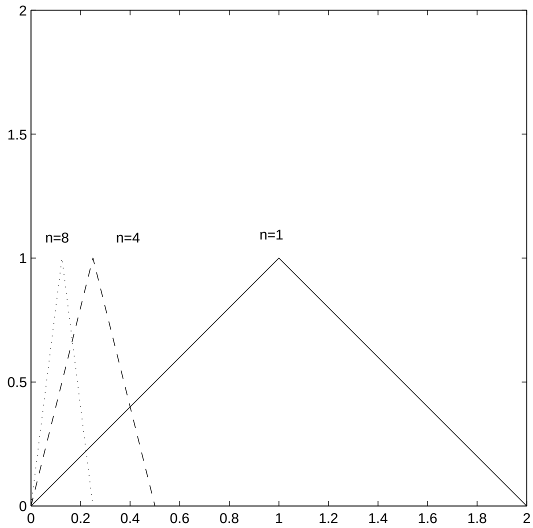
Betrachten wir die Funktionenfolgen in Beispiel IV-3.2: Die Folgen in (3) und (4) sind gleichmäßig konvergent, jene in (1) und (2) konvergieren nicht gleichmäßig: Man verifiziere die Bedingung $(*)$ mit folgender Wahl:
IV-3.2-(1): $\varepsilon_0 = \frac{1}{4}$, $\varphi(n) = n$, $x_n = 1 - \frac{1}{n}$, $n \geq 2$:

$$f_n(x_n) = (1 - \frac{1}{n})^n = \frac{1}{(1 + \frac{1}{n-1})^n} \overset{\text{Beispiel II-5.4}}{\geq} \frac{1}{4}$$

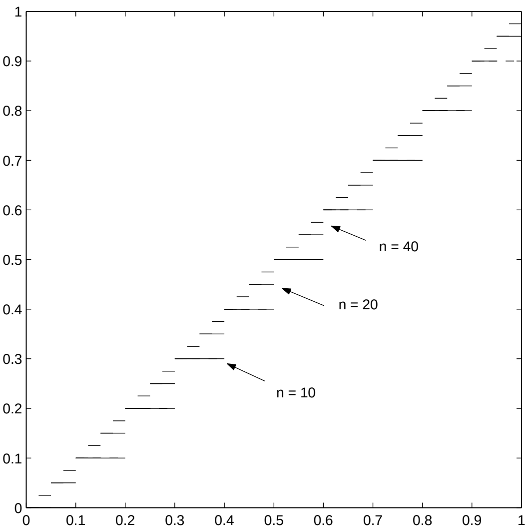
IV-3.2-(2): $\varepsilon_0 = \frac{1}{2}$, $\varphi(n) = n$, $x_n = \frac{1}{n}$, $n \geq 1$: $f_n(x_n) = 1$.



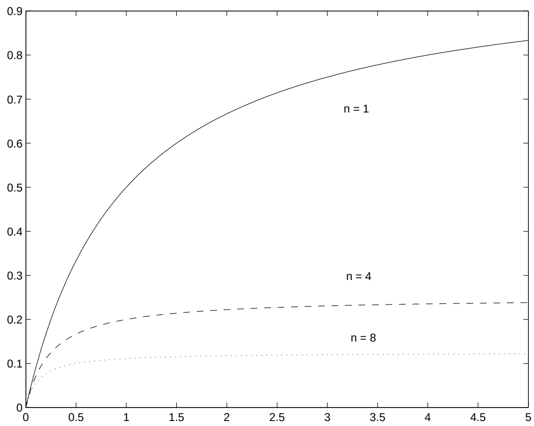
Beispiel 1



Beispiel 2



Beispiel 3



Beispiel 4

In den folgenden Plots ist die nicht gleichmäßige Konvergenz in den Beispielen (1) und (2), bzw. die gleichmäßige Konvergenz in den Beispielen (3) und (4) deutlich erkennbar:

Bemerkung IV-3.4. Wir erinnern daran, daß im Vektorraum der beschränkten Funktionen $\mathcal{B}(D, Y)$ durch

$$\|f\|_\infty = \sup\{\|f(x)\|_Y : x \in D\}$$

eine Norm erklärt wird (Beispiel II-1.5). Man überzeuge sich davon, daß eine Folge $(f_n) \subset \mathcal{B}(D,Y)$ genau dann gegen $f \in \mathcal{B}(D,Y)$ bezüglich der Norm $\|\cdot\|_\infty$ konvergiert, wenn f_n gleichmäßig gegen f konvergiert. Auch in $C(K,Y)$, $K \subset X$ kompakt, bedeutet Konvergenz einer Folge deren gleichmäßige Konvergenz.

Es ist klar, daß als Grenzfunktion bei gleichmäßiger Konvergenz nur jene Funktion in Frage kommt, die durch die punktweise Konvergenz bestimmt wird. Wie bei gewöhnlichen Folgen kann auch die gleichmäßige Konvergenz einer Funktionenfolge durch eine Cauchy Bedingung charakterisiert werden. Diese hat wieder den Vorteil, daß ihr Nachweis die Kenntnis der Grenzfunktion nicht voraussetzt.

Theorem IV-3.5. *Es sei $(f_n)_{n\ge 1}$ eine Folge von Funktionen $f_n\colon D \to Y$, $D \subset X$ und Y vollständig. Die Folge (f_n) ist genau dann gleichmäßig konvergent, wenn (f_n) eine* **gleichmäßige Cauchy Folge** *bildet, dh.*

$$(*) \qquad \forall \varepsilon > 0\ \exists N(\varepsilon)\ \forall x \in D\ \forall n,m \ge N(\varepsilon)\colon \|f_n(x) - f_m(x)\| < \varepsilon.$$

Beweis. Übung. □

Theorem IV-3.6. *Eine Folge von Funktionen $f_n\colon D \to Y$, $D \subset X$, $n \in \mathbb{N}$, konvergiere gleichmäßig gegen die Funktion $f\colon D \to Y$. Ist jede Abbildung f_n in $x_0 \in D$ stetig, dann ist auch f in x_0 stetig. Gilt insbesondere $(f_n) \subset C(D,Y)$, dann ist f stetig.*

Beweis. Wir schreiben $f(x) - f(x_0)$ in der Form

$$f(x) - f(x_0) = (f(x) - f_n(x)) + (f_n(x) - f_n(x_0)) + (f_n(x_0) - f(x_0)).$$

und argumentieren dann wie im Beweis von Satz III-4.5. □

Dieser Satz ist auch von Nutzen beim Nachweis, daß bestimmte Grenzwerte nicht gleichmäßig sind. Z.B. kann die Konvergenz in Beispiel IV-3.2-(1) nicht gleichmäßig sein, da die Grenzfunktion f unstetig, die approximierenden Funktionen f_n jedoch stetig sind.

Eine teilweise Umkehrung dieses Satzes bringt folgendes Resultat, welches auf U. Dini zurückgeht. Der Beweis erfordert allerdings Hilfsmittel, die uns noch nicht zur Verfügung stehen.

Theorem IV-3.7 (Dini). *Es sei $I = [a,b]$ und $(f_n) \subset C(I,\mathbb{R})$ und (f_n)* ***konvergiere monoton*** *gegen f, dh. $f(x) = \lim_{n\to\infty} f_n(x)$, und $f_n(x) \le f_{n+1}(x)$ (bzw. $f_n(x) \ge f_{n+1}(x)$) für alle $x \in I$ und $n \in \mathbb{N}$. Ist die Grenzfunktion stetig, dann ist die Konvergenz gleichmäßig.*

Der Satz von Dini ist falsch, wenn man auch nur auf eine der Voraussetzungen verzichtet. In den folgenden Beispielen sind jeweils alle Voraussetzungen bis auf eine einzige Ausnahme erfüllt, die Konvergenz gegen die Grenzfunktion ist jeweils *nicht* gleichmäßig. Wir überlassen diesen Nachweis dem Leser und notieren nur die Voraussetzung, die verletzt wird:

Beispiel IV-3.8.
(1) $I = [0,1]$, $f_n(x) = 1 - x^n$. Die Grenzfunktion $f(x) = 1$ für $x \in [0,1)$, $f(1) = 0$, ist unstetig.
(2) $I = \mathbb{R}$, $f_n(x) = -\frac{x^2}{n}$. Der Definitionsbereich ist kein abgeschlossenes und **beschränktes** Intervall.
(3) $I = [0,1]$, $f_n(x) = 0$, $0 < x < \frac{1}{n}$, $f_n(x) = 1$, $x \in \{0\} \cup [\frac{1}{n}, 1]$. Die Funktionen f_n sind unstetig.
(4) $I = [0,2]$, $f_n(x) = 0$, $x \in (\frac{2}{n}, 2]$, $f_n(x) = n^2x^2 - 2nx$, $x \in [0, \frac{2}{n}]$. Die Konvergenz von (f_n) gegen 0 ist nicht monoton.

Wir betrachten nun Funktionenreihen $\sum_{n=1}^{\infty} f_n$, die natürlich aufzufassen sind als Folgen von Partialsummen $(s_n)_{n\geq 1}$, $s_n = \sum_{i=1}^{n} f_i$. Somit gibt es auch für Funktionenreihen zwei natürliche Konvergenzbegriffe:

Definition IV-3.9. *Es sei $(f_n)_{n\geq 1}$ eine Folge von Funktionen $f_n\colon D \to Y$, $D \subset X$ und $(s_n)_{n\geq 1}$ die Folge der n-ten Partialsummen $s_n = \sum_{i=1}^{n} f_i$.*

i) *Die Reihe $\sum_{n=1}^{\infty} f_n$ konvergiert **punktweise** gegen $f\colon D \to Y$, $\sum_{n=1}^{\infty} f_n = f$, pw. $\underset{Def}{\Leftrightarrow} \lim_{n\to\infty} s_n = f$ punktweise.*
ii) *Die Reihe $\sum_{n=1}^{\infty} f_n$ konvergiert **gleichmäßig** gegen $f\colon D \to Y$, $\sum_{n=1}^{\infty} f_n = f$ glm. $\underset{Def}{\Leftrightarrow} \lim_{n\to\infty} s_n = f$ gleichmäßig.*
iii) *Die Reihe $\sum_{n=1}^{\infty} f_n$ konvergiert punktweise (gleichmäßig)*
$\underset{Def}{\Leftrightarrow} \exists f\colon D \to Y, \sum_{n=1}^{\infty} f_n = f$ *punktweise (gleichmäßig).*

Satz IV-3.6 läßt sich auch für Funktionenreihen formulieren:

Theorem IV-3.10. *Es seien $f_n\colon D \to Y$, $D \subset X$. Die Folge der Partialsummen (s_n), $s_n = \sum_{i=1}^{n} f_i$, konvergiere* gleichmäßig *gegen $f\colon D \to Y$. Sind alle Abbildungen f_n in $x_0 \in D$ stetig, dann ist auch $f = \sum_{n=1}^{\infty} f_n$ in x_0 stetig. Gilt insbesondere $(f_n) \subset C(D, \mathbb{R})$, dann ist auch f stetig.*

Wir überlassen es dem Leser, das Cauchy Kriterium IV-3.5 für Funktionenreihen anzugeben. Eine überaus nützliche *hinreichende* Bedingung für (absolute) gleichmäßige Konvergenz ist folgende Adaptation des Vergleichskriteriums:

Theorem IV-3.11 (Weierstrass m-Test). *Es sei $(f_n)_{n\geq 1}$ eine Folge von Funktionen $f_n\colon D \to Y$, $D \subset X$, und Y vollständig. Gibt es eine Folge $(M_n) \subset \mathbb{R}^+$, sodaß*

i) $\sum_{n=1}^{\infty} M_n < \infty$,
ii) $\|f_n(x)\| \leq M_n$ *für alle $x \in D$ und $n \in \mathbb{N}$,*

dann ist die Reihe $\sum_{n=1}^{\infty} f_n$ (absolut) gleichmäßig konvergent.

Beweis. Die Bedingungen des Satzes implizieren, daß die Reihe $\sum_{n=1}^{\infty} M_n$ für alle $x \in D$ eine konvergente Majorante für $\sum_{n=1}^{\infty} \|f_n(x)\|$ darstellt. Die Behauptung folgt nun aus Satz II-9.5. □

Der Weierstraß m-Test führt notwendig auf *absolute* gleichmäßige Konvergenz. Wir weisen jedoch darauf hin, daß die beiden Begriffe gleichmäßige Konvergenz und absolute gleichmäßige Konvergenz voneinander unabhängig sind. Die hinreichenden Bedingungen für bedingte Konvergenz aus Kapitel II lassen sich jedoch leicht auf Funktionenreihen übertragen:

Theorem IV-3.12 (Leibniz). *Die Funktionen $(f_n)\colon D \to [0, \infty)$, $D \subset \mathbb{R}$, $n \in \mathbb{N}$, erfüllen die Bedingungen*

(1) *Die Folge $(f_n(x))$ ist monoton fallend für alle $x \in D$.*
(2) *Die Funktionenfolge (f_n) konvergiere gleichmäßig gegen* 0.

Dann ist die Reihe $\sum_{n=1}^{\infty}(-1)^n f_n$ *gleichmäßig konvergent.*

Beweis. Die Folge $(f_n(x))$ ist für alle $x \in D$ konvergent. Wir bezeichnen den punktweisen Grenzwert der Reihe mit $f(x)$ (Leibniz-Kriterium II-11.1). Nach Korollar II-11.2 gilt die Abschätzung

$$|f(x) - \sum_{k=1}^{n} f_k(x)| \leq |f_{n+1}(x)|.$$

Die Behauptung ergibt sich nun aus Bedingung (2). □

Theorem IV-3.13 (Abel). *Es seien* $(u_n), (v_n)$ *Folgen von Funktionen* $u_n,\ v_n\colon D \to \mathbb{R}$, $D \subset \mathbb{R}$. *Es gelte:*

(1) $\sum_{n=1}^{\infty} u_n$ *ist gleichmäßig konvergent,*
(2) (v_n) *ist* **gleichmäßig beschränkt**, *dh.,*

$$\exists M > 0\ \forall x \in D\ \forall n \in \mathbb{N}\colon |v_n(x)| \leq M$$

(3) *Die Folge* $(v_n(x))_{n\geq 1}$ *ist monoton für alle* $\forall x \in D$.

Dann ist die Reihe $\sum_{n=1}^{\infty} u_n v_n$ *gleichmäßig konvergent.*

Theorem IV-3.14 (Dirichlet). *Es seien* $(u_n), (v_n)$ *Folgen von Funktionen* $u_n,\ v_n\colon D \to \mathbb{R}$, $D \subset \mathbb{R}$. *Es gelte:*

(1) *Die Folge der Partialsummen* $(s_n)_{n\geq 1}$, $s_n = \sum_{k=1}^{n} u_k$, *ist gleichmäßig beschränkt.*
(2) *Die Funktionenfolge* (v_n) *konvergiert gleichmäßig und monoton nach* 0.

Dann ist die Reihe $\sum_{n=1}^{\infty} u_n v_n$ *gleichmäßig konvergent.*

Beispiel IV-3.15. Wir untersuchen die Konvergenz der Reihe $\sum_{n=1}^{\infty}(-1)^{n+1} f_n(x)$ mit $D = \mathbb{R}$, $f_n(x) = \frac{x^2}{(1+x^2)^n}$.
i) $|\frac{f_{n+1}(x)}{f_n(x)}| = \frac{1}{1+x^2} < 1$, $x \neq 0$, $f_n(0) = 0$, $n \in \mathbb{N}$. Das Quotientenkriterium garantiert somit die absolute, punktweise Konvergenz der Reihe $\sum_{n=1}^{\infty}(-1)^{n+1} f_n(x)$. Für festes $x \in \mathbb{R}$ liegt eine geometrische Reihe vor. Eine leichte Rechnung liefert nun die Grenzfunktion

$$f(x) = -x^2\big[\frac{1}{1 + \frac{1}{1+x^2}} - 1\big] = \frac{x^2}{2 + x^2}.$$

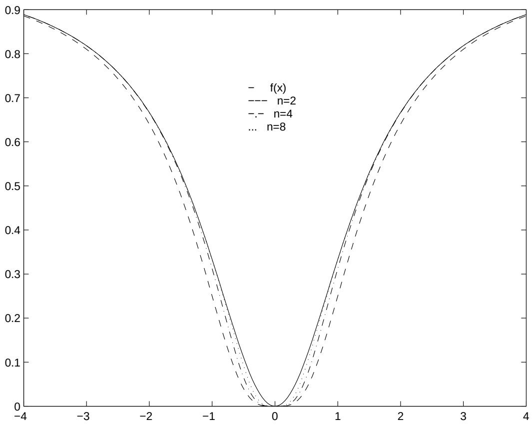
ii) Setzen wir $u_n(x) = (-1)^{n+1}$, $v_n(x) = f_n(x)$, $x \in \mathbb{R}$. Die Folge der Partialsummen $(s_n) = (1, 0, 1, 0, \dots)$ ist beschränkt, es gilt $f_{n+1}(x) \leq f_n(x)$, $x \in \mathbb{R}$. Wir zeigen, daß (f_n) gleichmäßig gegen die Nullfunktion konvergiert:

$$f_n(x) \overset{\text{I-4.21}}{\leq} \frac{x^2}{1 + nx^2} \leq \frac{1}{n} \quad \text{für alle } x \in \mathbb{R},$$

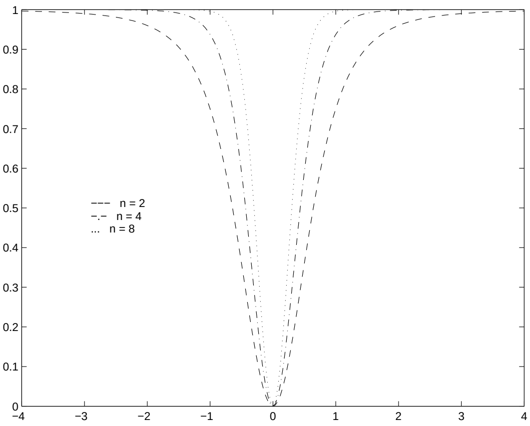
somit konvergiert die Reihe gleichmäßig (vgl. die Sätze IV-3.14 bzw. IV-3.12).
iii) Die Reihe ist jedoch nicht absolut gleichmäßig konvergent, dh. die Reihe $\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$ konvergiert nicht gleichmäßig auf $\mathbb{R}$. Wegen i) konvergiert diese Reihe absolut und punktweise, und zwar gegen

$$\tilde{f}(x) = \begin{cases} 0 & \text{für } x = 0 \\ 1 & \text{für } x \neq 0 \end{cases}$$

Da die Grenzfunktion in $x = 0$ unstetig ist, alle Funktionen f_n jedoch stetig sind, kann nach Satz IV-3.10 die Reihe $\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$ nicht gleichmäßig auf $\mathbb{R}$ konvergieren.



gleichmäßige Konvergenz



absolute, nicht gleichmäßige Konvergenz

Bemerkung IV-3.16. Abschließend bemerken wir, daß die Kriterien von Abel und Dirichlet auch dann gelten, wenn die Funktionen u_n Werte in $\mathbb{C}$ annehmen (wegen der Monotonievoraussetzung für (v_n) müssen die Funktionen v_n natürlich nach wie vor reelle Funktionen sein).

4. Potenzreihen

Definition IV-4.1. *Es sei $(a_n)_{n\geq 0} \subset \mathbb{C}$ und $z_0 \in \mathbb{C}$. Die komplexe Reihe*

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n, \qquad z \in \mathbb{C},$$

*heißt **Potenzreihe** mit Entwicklungszentrum z_0. Man nennt a_n, $n \in \mathbb{N}_0$, den n-ten **Koeffizienten** der Reihe.*

Je nach Wahl von $z \in \mathbb{C}$ kann eine Potenzreihe entweder konvergieren oder divergieren. Trivialerweise konvergiert sie immer für $z = z_0$. Allerdings sind die Punkte, für welche eine Potenzreihe konvergiert, nicht regellos in der Gaußschen Zahlenebene verteilt. Der folgende Satz zeigt nämlich, daß es zu jeder Potenzreihe einen Kreis mit Mittelpunkt z_0 gibt, sodaß im Inneren dieses Kreises die Potenzreihe konvergiert und im äußeren divergiert:

Theorem IV-4.2. *Es sei $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ eine Potenzreihe und $\rho = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}$. Wir definieren $R \in \bar{\mathbb{R}}^+$ durch*

$$R = \begin{cases} \frac{1}{\rho} & \text{für } 0 < \rho < \infty, \\ \infty & \text{für } \rho = 0. \end{cases}$$

Die Potenzreihe konvergiert absolut für alle $z \in \mathbb{C}$ mit $|z - z_0| < R$ und divergiert für alle $z \in \mathbb{C}$ mit $|z - z_0| > R$. Ist $\rho = \infty$, konvergiert die Potenzreihe nur im Entwicklungszentrum z_0.

Beweis. Nach dem Wurzelkriterium II-12.2 konvergiert die Reihe $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ absolut, falls (vgl. Satz II-6.9–(iv))

$$\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n (z - z_0)^n|} = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} \cdot |z - z_0| = \rho \cdot |z - z_0| < 1$$

und divergiert, falls $\rho|z - z_0| > 1$. Ist die Folge $(\sqrt[n]{|a_n|})$ unbeschränkt, d.h. $\rho = \infty$, kann somit die Potenzreihe für $z \in \mathbb{C}$ mit $|z - z_0| > 0$ nicht konvergieren, ist andererseits $\rho = 0$, so ist die Bedingung $\rho|z - z_0| < 1$ für alle $z \in \mathbb{C}$ erfüllt. Gilt $\rho \in (0, \infty)$, dann konvergiert die Reihe absolut für alle $z \in \mathbb{C}$ mit $|z - z_0| < \frac{1}{\rho}$. $\square$

Definition IV-4.3. *Es gelten die Bezeichnungen von Satz IV-4.2. Für $R \in (0, \infty)$ nennt man $K(z_0, R) = \{z \in \mathbb{C} \colon |z - z_0| < R\}$* ***Konvergenzkreis*** *der Potenzreihe und R ihren* ***Konvergenzradius****. Sind sämtliche Koeffizienten der Potenzreihe reell und ist $z_0 \in \mathbb{R}$, so heißt $K(z_0, R) \cap \mathbb{R} = (z_0 - R, z_0 + R)$* ***Konvergenzintervall*** *der (reellen) Potenzreihe.*

Jede Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ mit Konvergenzkreis $K(z_0, R)$ definiert durch $z \mapsto \sum_{n=0}^{\infty} a_n(z - z_0)^n$ eine Abbildung $f : K(z_0, R) \to \mathbb{C}$, die eine Reihe außergewöhnlicher Eigenschaften aufweist. Die Darstellung einer Funktion f durch eine Potenzreihe bietet eine Möglichkeit, Funktionswerte von f innerhalb des Konvergenzkreises mit beliebiger Genauigkeit zu berechnen.

Es ist nicht immer notwendig, auf Satz IV-4.2 zur Berechnung des Konvergenzradius zurückzugreifen. Oft ist es einfacher, das Quotientenkriterium unmittelbar auf die Potenzreihe anzuwenden (man vergleiche in diesem Zusammenhang auch Satz II-6.10).

Beispiel IV-4.4. $\sum_{n=1}^{\infty} \frac{n^n z^n}{n!} = \sum_{n=1}^{\infty} c_n$. Die Anwendung des Quotientenkriteriums ergibt

$$\frac{c_{n+1}}{c_n} = \frac{(n+1)^{n+1} z^{n+1} n!}{n^n z^n (n+1)!} = (1 + \frac{1}{n})^n z,$$

und mit Hilfe von Beispiel II-5.4 folgt

$$\lim_{n\to\infty} |\frac{c_{n+1}}{c_n}| = e|z|.$$

Die Reihe konvergiert somit absolut für $e|z| < 1$, d.h. für $|z| < \frac{1}{e}$, und divergiert für $|z| > \frac{1}{e}$. Der Konvergenzradius ist also $R = \frac{1}{e}$.

Die folgenden Beispiele zeigen, daß das Konvergenzverhalten von Potenzreihen auf dem Rand des Konvergenzkreises sehr unterschiedlich sein kann.

Beispiel IV-4.5. 1) $\sum_{n=0}^{\infty} z^n$, $R = 1$. Die Reihe divergiert in jedem Punkt mit $|z| = 1$, da die notwendige Konvergenzbedingung in Korollar II-9.3 verletzt ist.
2) $\sum_{n=1}^{\infty} \frac{z^n}{n(n+1)}$, $R = 1$. Die Reihe konvergiert (sogar absolut) in jedem Randpunkt von $K(0, 1)$, denn für $|z| = 1$ ist $\sum_{n=1}^{\infty} \frac{1}{n^2}$ eine konvergente Majorante.
3) $\sum_{n=1}^{\infty} \frac{z^n}{n}$, $R = 1$. Von dieser Reihe kann man zeigen, daß sie in jedem Punkt mit $|z| = 1$ und $z \neq 1$ bedingt konvergiert. Für $z = -1$ erhält man die Leibniz Reihe.
4) $\sum_{n=1}^{\infty} n^n z^n$, $R = 0$.
5) $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$, $R = \infty$.

Theorem IV-4.6 (Eulersche Zahl e und Exponentialreihe). i) $e = \sum_{k=0}^{\infty} \frac{1}{k!}$

ii) *e ist irrational.*

iii) *Der Fehler der rationalen Approximation $S_n = \sum_{k=0}^{n} \frac{1}{k!}$, $n \in \mathbb{N}$, ist beschränkt durch*

$$0 < e - S_n < \frac{1}{n!n}.$$

Beweis. i) Wir erinnern daran, daß e in Beispiel II-5.4 durch

$$e = \lim_{n\to\infty} (1 + \frac{1}{n})^n = \lim_{n\to\infty} a_n$$

definiert wurde. Nach Beispiel IV-4.5-5 existiert $S = \sum_{k=0}^{\infty} \frac{1}{k!}$. Mit Hilfe des binomischen Lehrsatzes folgt für alle $n \geq 1$

$$\begin{aligned} a_n &= 1 + \sum_{k=1}^{n} \frac{1}{k!} \big(\prod_{i=0}^{k-1} (n-i) \big) \frac{1}{n^k} = 1 + \sum_{k=1}^{n} \big(\prod_{i=0}^{k-1} (1 - \frac{i}{n}) \big) \frac{1}{k!} \\ &< 1 + \sum_{k=1}^{n} \frac{1}{k!} < S \end{aligned}$$

und daher auch

$$e = \lim_{n\to\infty} a_n \leq S. \tag{1}$$

Zu beliebigem $\varepsilon > 0$ wählt man $p \in \mathbb{N}$ so, daß

$$S - \varepsilon < S_p. \tag{2}$$

Für jedes $n > p$ folgt aus der Monotonie der Folge (a_n)

$$e > a_n = 1 + \sum_{k=1}^{n} \big(\prod_{i=0}^{k-1} (1 - \frac{i}{n}) \big) \frac{1}{k!} > 1 + \sum_{k=1}^{p} \big(\prod_{i=0}^{k-1} (1 - \frac{i}{n}) \big) \frac{1}{k!}.$$

Man beachte, daß auf der rechten Seite dieser Abschätzung eine feste Anzahl von Summanden steht. Man kann daher den Grenzübergang $n \to \infty$ durchführen und erhält

$$e \geq 1 + \sum_{k=1}^{p} \frac{1}{k!} = S_p.$$

Zusammen mit (1) und (2) ergibt sich daher für jedes $\varepsilon > 0$

$$S - \varepsilon < S_p \leq e \leq S,$$

d.h. $S = e$.

iii) Der Approximationsfehler $\sum_{k=n+1}^{\infty} \frac{1}{k!}$ kann abgeschätzt werden durch

$$\sum_{k=n+1}^{\infty} \frac{1}{k!} = \frac{1}{n!}(\frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \dots)$$
$$< \frac{1}{n!}(\frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \dots) = \frac{1}{n!}(\frac{1}{1-\frac{1}{n+1}} - 1) = \frac{1}{n!}\frac{1}{n}.$$

Dies zeigt

$$0 < e - S_n < \frac{1}{n!n}$$

ii) Wäre e rational, d.h. $e = \frac{p}{q}$, mit $p, q \in \mathbb{N}$, dann müßte auch

$$0 < \frac{p}{q} - S_q < \frac{1}{q!q}$$

bzw.

$$0 < p(q-1)! - S_q q! < \frac{1}{q} \le 1$$

gelten. Dann wäre $p(q-1)! - S_q q! = p(q-1)! - \sum_{k=0}^{q} \frac{q!}{k!}$ eine natürliche Zahl, welche in $(0,1)$ liegen müßte. Dies ist ein Widerspruch zu Satz II-I-4.6. $\square$

Approximation der Eulerschen Zahl $e = 2,718281828459...$

n	$e_n = \sum_{k=0}^{n} \frac{1}{k!}$	$e - e_n$
2	2.50000000	0.21828182
4	2.70833333	0.00994849
6	2.71805555	0.00022627
8	2.71827876	0.00000305
10	2.71828180	0.00000002

n	$e_n = (1+\frac{1}{n})^n$	$e - e_n$
20	2.653	0.064
40	2.685	0.033
60	2.695	0.022
80	2.701	0.016
100	2.704	0.013

Diese Ergebnisse veranschaulichen die rasche Konvergenz der Potenzreihe und die außerordentlich langsame Konvergenz der Folgenglieder $a_n = (1+\frac{1}{n})^n$.

Wir betrachten die Potenzreihe $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ mit Konvergenzradius R und bezeichnen die Summenfunktion mit f. Es ist also $f\colon K(z_0, R) \to \mathbb{C}$. Man sagt auch, f wird durch die Potenzreihe dargestellt, bzw. f besitzt die Reihenentwicklung $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ um z_0. Im folgenden wollen wir stets $R > 0$ voraussetzen.

Es sei $g(z) = \sum_{k=0}^{\infty} b_k(z-z_0)^k$ eine weitere Potenzreihe mit Konvergenzradius $\tilde{R}$ und $r = \min\{R, \tilde{R}\}$. Wegen der absoluten Konvergenz der Reihen f und g konvergieren auch die Reihen $f \pm g$ und $f \cdot g$. Insbesondere ist die Summe der Produktreihe unabhängig von der speziellen Reihenfolge in der die Produkte $a_k b_j (z-z_0)^{k+j}$ aufsummiert werden. Berechnen wir fg über das Cauchy Produkt (Korollar **??**), dann ist auch die Produktreihe eine Potenzreihe. Zusammenfassend gilt also für $|z - z_0| < r = \min\{R, \tilde{R}\}$:

$$f(z) \pm g(z) = \sum_{k=0}^{\infty} (a_k \pm b_k)(z - z_0)^k,$$

$$f(z)g(z) = \sum_{k=0}^{\infty} c_k(z - z_0)^k, \quad c_k = \sum_{i=0}^{k} a_i b_{k-i},\ k \in \mathbb{N}_0.$$

Der Einfachheit halber wollen wir in diesem Abschnitt o.B.d.A $z_0 = 0$ voraussetzen.

THEOREM IV-4.7. *Wir betrachten die durch die Potenzreihe* $\sum_{k=0}^{\infty} a_k z^k$ *im Konvergenzkreis* $K(0, R)$ *dargestellte Funktion* f. *Es bezeichne* $R_n(z)$ *den Reihenrest* $\sum_{k=n}^{\infty} a_k z^k$. *Dann gilt*

i) f *ist stetig auf* $K(0, R)$,
ii) f *ist beschränkt auf* $\bar{K}(0, \delta)$, $0 < \delta < R$,
iii) $\forall n \in \mathbb{N}\ \exists M_n > 0\ \forall z \in \bar{K}(0, \delta)\colon |R_n(z)| \leq M_n |z|^n$

BEWEIS. Für alle z aus der abgeschlossenen Kreisscheibe $\bar{K}(0, \delta)$, $\delta < R$, gilt die Abschätzung:

$$|f(z)| = |\sum_{k=0}^{\infty} a_k z^k| \leq \sum_{k=0}^{\infty} |a_n||z|^k \leq \sum_{k=0}^{\infty} |a_k|\delta^k := M_\delta < \infty$$

(wegen $\delta < R$ konvergiert die Reihe $\sum_{k=0}^{\infty} |a_k|\delta^k$). Dies zeigt ii), aber auch i). Nach dem m-Test von Weierstraß ist die Reihe $\sum_{k=0}^{\infty} a_k z^k$ auf $\bar{K}(0, \delta)$ nämlich gleichmäßig konvergent und die Grenzfunktion $f|_{\bar{K}(0,\delta)}$ somit nach Satz IV-3.10 stetig. Also ist f auch in jedem $z \in K(0, R)$ stetig (man beachte $z \in \bar{K}(0, \delta_0)$, etwa für $\delta_0 = \frac{1}{2}(R + |z|)$). Die Abschätzung des Reihenrestes folgt aus

$$|R_n(z)| \leq \sum_{k=n}^{\infty} |a_k||z|^k \leq |z|^n \sum_{k=0}^{\infty} |a_{k+n}| \cdot \delta^k$$

und der Beobachtung, daß die Potenzreihen $\sum_{k=0}^{\infty} a_k z^k$ und $\sum_{k=0}^{\infty} a_{n+k} z^k$, $n \in \mathbb{N}$, denselben Konvergenzradius besitzen. □

Dieser Satz garantiert zwar die Stetigkeit einer durch eine Potenzreihe dargestellten Funktion auf $K(0, R)$, es wird aber nichts ausgesagt über das Verhalten dieser Funktion in den Randpunkten von $K(0, R)$. Eine teilweise Lösung dieses Problems bringt der folgende Satz:

THEOREM IV-4.8 (Grenzwertsatz von Abel). *Es sei* $(a_n)_{n \geq 0} \subset \mathbb{C}$ *und es konvergiere die Reihe* $\sum_{n=0}^{\infty} a_n$. *Dann konvergiert die Potenzreihe*

$$\sum_{n=0}^{\infty} a_n x^n, \quad x \in \mathbb{R},$$

gleichmäßig auf $[0,1]$ *und stellt auf* $[0,1]$ *eine stetige Funktion* f *dar, dh. es gilt*

$$\lim_{x\to 1^-}\sum_{n=0}^{\infty} a_n x^n = \lim_{x\to 1^-} f(x) = f(1) = \sum_{n=0}^{\infty} a_n.$$

Beweis. Wir verwenden das Kriterium von Abel mit $u_n = a_n$, $v_n(x) = x^n$, $x \in D = [0,1]$ und zeigen damit die gleichmäßige Konvergenz von $\sum_{n=0}^{\infty} a_n x^n$ auf $[0,1]$. Nach Satz IV-3.10 ist somit die Summenfunktion f stetig auf $[0,1]$. □

Die Stetigkeit von f in $x = 1$ ist natürlich als linksseitige Stetigkeit aufzufassen. Konvergiert die Reihe $\sum_{n=0}^{\infty}(-1)^n a_n$, folgt die gleichmäßige Konvergenz der Potenzreihe auf $[-1,0]$ und insbesondere die Stetigkeit von f auf $[-1,0]$.

Korollar IV-4.9. Die Potenzreihe $\sum_{n=0}^{\infty} a_n z^n$ habe den Konvergenzradius $0 < R < \infty$. Konvergiert die Potenzreihe auch in $x = R$, dann konvergiert die Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad x \in \mathbb{R},$$

für alle $0 \le \delta < R$ gleichmäßig auf $[-\delta, R]$ und f ist stetig auf $(-R, R]$. Ein analoges Resultat gilt auch wenn die Potenzreihe in $x = -R$ konvergiert.

Beweis. Wir bezeichnen mit $\xi = \frac{x}{R}$ und betrachten die Potenzreihe $\tilde{f}(\xi) = f(R\xi) = \sum_{n=0}^{\infty} a_n R^n \xi^n$. Aus

$$\overline{\lim}\sqrt[n]{|a_n| R^n} = R\,\overline{\lim}\sqrt[n]{|a_n|} = 1$$

folgt, daß die Potenzreihe $\tilde{f}$ den Konvergenzradius $\tilde{R} = 1$ besitzt. Die Voraussetzung des Satzes bedeutet demnach, daß $\sum_{k=0}^{\infty} a_n R^n$ gegen $\tilde{f}(1)$ konvergiert. Die Behauptung folgt nun aus Satz IV-4.8. □

Beispiel IV-4.10. Vorerst ohne Beweis erwähnen wir die Entwicklung

$$\ln(1+x) = \sum_{i=1}^{\infty}(-1)^{i-1}\frac{x^i}{i}, \quad |x| < 1.$$

Die Potenzreihe konvergiert auch in $x = 1$ (vgl.Satz II-11.1). Aus dem Abel'schen Grenzwertsatz folgt daher

$$\sum_{i=1}^{\infty}(-1)^{i-1}\frac{1}{i} = \ln 2.$$

Wir kennen somit die Summe der alternierenden harmonischen Reihe.

Bemerkung IV-4.11. Die Grenzwerte $\lim_{x\to 1^-}\sum_{n=0}^{\infty} a_n x^n = \lim_{x\to 1^-} f(x)$ können auch existieren, ohne daß die Reihe $\sum_{n=0}^{\infty} a_n$ konvergent ist. Es sei zum Beispiel $a_n = (-1)^n$ und

$$f(x) = \sum_{n=0}^{\infty}(-1)^n x^n = \frac{1}{1+x}, \quad |x| < 1.$$

Es existiert zwar

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} \sum_{n=0}^{\infty} (-1)^n x^n = \lim_{x\to 1^-} \frac{1}{1+x} = \frac{1}{2},$$

aber die Reihe $\sum_{n=0}^{\infty}(-1)^n$ ist divergent.

Das nächste Resultat deckt eine sehr starke Bindung zwischen den Funktionswerten einer durch eine Potenzreihe dargestellten Funktion auf:

Theorem IV-4.12 (Identitätssatz). *Die Potenzreihen $f(z) = \sum_{k=0}^{\infty} a_k z^k$ und $g(z) = \sum_{k=0}^{\infty} b_k z^k$ mögen zumindest auf $K(0,r)$ für ein $r > 0$ konvergieren. Gibt es auch nur eine einzige Folge $(z_n)_{n\ge 1} \subset K(0,r) \setminus \{0\}$, sodaß*

i) $\lim_{n\to\infty} z_n = 0$,
ii) $f(z_n) = g(z_n)$ *für alle* $n \in \mathbb{N}$

gilt, dann ist $a_n = b_n$ für alle $n \in \mathbb{N}_0$, dh. die Potenzreihen sind identisch.

Beweis. Angenommen die Behauptung wäre falsch, dh. es gibt ein $k_0 \in \mathbb{N}$ mit $a_{k_0} \neq b_{k_0}$. Wir nehmen an, k_0 ist der kleinste derartige Index, dh. $a_k = b_k$, $k = 0, \dots, k_0 - 1$ falls $k_0 \ge 1$, bzw. $a_0 \neq b_0$ falls $k_0 = 0$. Somit gilt

$$(f-g)(z) = \sum_{i=k_0}^{\infty} (a_i - b_i) z^i \text{ und } (f-g)(z_n) = 0,\ n \in \mathbb{N}.$$

Wir wenden nun die Abschätzung in Satz IV-4.7-iii) folgendermaßen an: Zu $0 < \delta < r$ gibt es $M_{k_0+1} > 0$, sodaß für alle $z \in \bar{K}(0,\delta)$

$$|(f-g)(z) - (a_{k_0} - b_{k_0}) z^{k_0}| = |R_{k_0+1}(z)| \le M_{k_0+1} |z|^{k_0+1}$$

zutrifft. Für $z = z_n$ (n so groß, daß $|z_n| < \delta$) folgt also

$$|a_{k_0} - b_{k_0}| \cdot |z_n|^{k_0} \le M_{k_0+1} |z_n|^{k_0+1}$$

und daraus

$$|a_{k_0} - b_{k_0}| \le M_{k_0+1} |z_n|.$$

$\lim_{n\to\infty} z_n = 0$ führt auf $a_{k_0} = b_{k_0}$ im Widerspruch zur Wahl von k_0. □

Eine überraschende Konsequenz des Identitätssatzes ist, daß eine durch eine Potenzreihe darstellbare Funktion durch die Werte auf einer gegen das Entwicklungszentrum konvergenten Folge (z_n) bereits im gesamten Inneren des Konvergenzkreises eindeutig bestimmt ist. Weiters sind die Koeffizienten der Entwicklung einer Funktion in eine Potenzreihe um z_0 eindeutig bestimmt. Als Übung beweise man weiters, daß in der Potenzreihe einer geraden (ungeraden) Funktion nur gerade (ungerade) Potenzen auftreten.

Der Gültigkeitsbereich der Reihenentwicklung $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ einer *gegebenen* Funktion f ist auf den Konvergenzkreis $K(z_0, R)$ der Reihe beschränkt. Die Funktion f selbst kann ohne weiteres außerhalb von $K(z_0, R)$ definiert sein. Natürlich kann sie dort nicht durch die Reihe $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ dargestellt werden. Oft kann man durch eine geeignete Wahl des Entwicklungszentrums eine Reihenentwicklung $\sum_{k=0}^{\infty} b_k(z-z_1)^k$ mit Konvergenzkreis $K(z_1, R_1)$ finden, sodaß $K(z_1, R_1) \cap \complement K(z_0, R) \neq \emptyset$. Die Grundlage hiefür ist der folgende Satz:

Theorem IV-4.13 (Transformationssatz). *Die Potenzreihe $f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k$ konvergiere in $K(z_0, R)$. Es sei $z_1 \in K(z_0, R)$ und $\rho := R - |z_1 - z_0| > 0$. Für alle $z \in K(z_1, \rho)$ gilt dann die Identität*

$$f(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k = \sum_{k=0}^{\infty} b_k(z - z_1)^k$$

mit

$$b_k = \sum_{n=k}^{\infty} \binom{n}{k} a_n(z_1 - z_0)^{n-k}, \quad k \in \mathbb{N}_0.$$

Beweis. Setzt man die Identität

$$(z - z_0)^n = [(z - z_1) + (z_1 - z_0)]^n = \sum_{k=0}^{n} \binom{n}{k} (z - z_1)^k (z_1 - z_0)^{n-k}$$

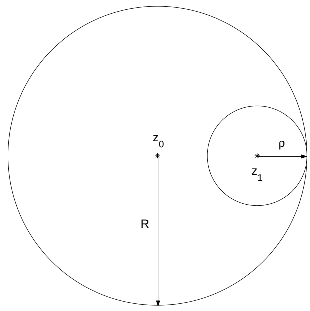
in die Potenzreihe um z_0 ein, ergibt eine formale Rechnung (wir setzen dabei $\binom{n}{k} = 0$ für $k > n$)

$$\begin{aligned}
\sum_{n=0}^{\infty} a_n(z - z_0)^n &= \sum_{n=0}^{\infty} a_n \left[\sum_{k=0}^{n} \binom{n}{k} (z - z_1)^k (z_1 - z_0)^{n-k}\right] \\
&= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{\infty} \binom{n}{k} a_n(z_1 - z_0)^{n-k} (z - z_1)^k\right] \\
&=^{*)} \sum_{k=0}^{\infty} \left[\sum_{n=0}^{\infty} \binom{n}{k} a_n(z_1 - z_0)^{n-k}\right] (z - z_1)^k \\
&= \sum_{k=0}^{\infty} \left[\sum_{n=k}^{\infty} \binom{n}{k} a_n(z_1 - z_0)^{n-k}\right] (z - z_1)^k = \sum_{k=0}^{\infty} b_k(z - z_1)^k
\end{aligned}$$

Einzig der mit *) markierte Schritt, in dem die Summationsreihenfolge vertauscht wird, ist noch zu rechtfertigen. Die Vertauschung ist nach dem Doppelreihensatz **??** zulässig, falls die absolute Konvergenz einer der iterierten Reihen nachgewiesen werden kann, da diese die Summierbarkeit der Doppelreihe zur Folge hat:

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} |a_n| \cdot |z_1 - z_0|^{n-k} |z - z_1|^k =$$

$$\sum_{n=0}^{\infty} |a_n| \sum_{k=0}^{n} \binom{n}{k} |z_1 - z_0|^{n-k} |z - z_1|^k \overset{\text{II-I-4.22}}{=} \sum_{n=0}^{\infty} |a_n| \cdot [|z_1 - z_0| + |z - z_1|]^n < \infty.$$



zum Transformationssatz

Die Konvergenz der letzten Reihe ergibt sich aus der Wahl von ρ ($z \in K(z_1, \rho) \Rightarrow |z - z_1| + |z_1 - z_0| < R$), und der absoluten Konvergenz der Reihe $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ auf $K(z_0, R)$. $\square$

Hat man also eine *gegebene* Funktion in eine Potenzreihe um z_0 mit Konvergenzradius R entwickelt, dann kann man jeden Punkt z_1 von $K(z_0, R)$ als neues Entwicklungszentrum wählen, und der Transformationssatz garantiert die Konvergenz der neuen Entwicklung um z_1 zumindest auf dem größten Kreis mit Mittelpunkt z_1, der noch vollständig in $K(z_0, R)$ enthalten ist. Oft ist der Konvergenzradius der neuen Reihe jedoch bedeutend größer.

Beispiel IV-4.14. Wir betrachten die Entwicklung

$$f(z) = \frac{1}{1-z} = \sum_{i=0}^{\infty} z^i, \qquad |z| < 1.$$

Nach dem Transformationssatz läßt sich f auch um $z_1 = -\frac{1}{2}$ in eine Potenzreihe entwickeln, die zumindest für $|z + \frac{1}{2}| < \frac{1}{2}$ konvergiert. Wir erhalten die Entwicklung

$$(*) \qquad \frac{1}{1-z} = \sum_{k=0}^{\infty} b_k(z+\frac{1}{2})^k, \quad b_k = \sum_{i=k}^{\infty} \binom{i}{k}(-\frac{1}{2})^{i-k}, \quad k \in \mathbb{N}_0.$$

Die Schwierigkeit bei der Anwendung des Transformationssatzes ist die Berechnung der neuen Koeffizienten b_k. Oft ist es zielführender, auf bereits bekannte Entwicklungen zurückzugreifen: Da wir f um $z_1 = -\frac{1}{2}$ entwickeln wollen, liegt folgende Vorgangsweise nahe:

$$f(z) = \frac{1}{1-z} = \frac{1}{1-(z+\frac{1}{2})+\frac{1}{2}} = \frac{2}{3}\frac{1}{1-\frac{2}{3}(z+\frac{1}{2})} = \frac{2}{3}\sum_{i=0}^{\infty}(\frac{2}{3})^i(z+\frac{1}{2})^i.$$

Diese Reihe konvergiert für $|z + \frac{1}{2}| < \frac{3}{2}$ und stimmt auf $|z + \frac{1}{2}| < \frac{1}{2}$ mit (*) überein. Nach dem Identitätssatz muß also $b_k = \sum\limits_{i=k}^{\infty} \binom{i}{k}(-\frac{1}{2})^{i-k} = (\frac{2}{3})^{k+1}$, $k \in \mathbb{N}_0$, gelten. Die Koeffizienten b_k können übrigens auch unabhängig von dieser Überlegung über die

Entwicklung

$$\frac{1}{(1-z)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{k} z^n, \quad |z| < 1,$$

berechnet werden (Übung).

THEOREM IV-4.15 (Divisionssatz). *Die Potenzreihe* $f(z) = \sum\limits_{n=0}^{\infty} a_n z^n$ *konvergiere für* $|z| < R$ *und es sei* $f(0) \neq 0$. *Dann läßt sich* $\frac{1}{f}$ *in einer hinreichend kleinen* ρ*–Umgebung von* 0 *ebenfalls in eine Potenzreihe entwickeln:*

$$\frac{1}{f(z)} = \sum_{k=0}^{\infty} b_k z^k, \quad |z| < \rho.$$

BEWEIS. O.B.d.A. können wir $a_0 = f(0) = 1$ annehmen. Wegen der Stetigkeit von $\tilde{f}(z) = \sum_{k=1}^{\infty} a_k z^k$ und $\tilde{f}(0) = 0$ gibt es ein $\delta \in (0, R)$, sodaß für $|z| < \delta$ stets $|\tilde{f}(z)| < 1$ bleibt. Für $|z| < \delta$ gilt also

$$\frac{1}{f(z)} = \frac{1}{1-(-\tilde{f}(z))} = \sum_{j=0}^{\infty} (-\tilde{f}(z))^j = \sum_{j=0}^{\infty} \left[-\sum_{i=1}^{\infty} a_i z^i \right]^j.$$

Wegen der absoluten Konvergenz von $\tilde{f}$ kann man $\tilde{f}^j$ z.B. mit der Methode von Cauchy berechnen und erhält mit geeigneten Koeffizienten a_{jk}, $j, k \in \mathbb{N}_0$,

$$(-\tilde{f}(z))^j = \sum_{k=0}^{\infty} a_{jk} z^k, \quad j \in \mathbb{N}_0.$$

Dies ergibt

$$\frac{1}{f(z)} = \sum_{j=0}^{\infty} \left[\sum_{k=0}^{\infty} a_{jk} z^k \right] =^{*)} \sum_{k=0}^{\infty} \left[\sum_{j=0}^{\infty} a_{jk} \right] z^k = \sum_{k=0}^{\infty} b_k z^k.$$

Der Beweis ist vollständig, wenn die in *) durchgeführte Vertauschung der Summationsreihenfolge gerechtfertigt werden kann. Nach dem Doppelreihensatz genügt es, die Konvergenz von $\sum_{j=0}^{\infty} \left[\sum_{k=0}^{\infty} |a_{jk}||z|^k \right]$ nachzuweisen: Es ist

$$|\tilde{f}(z)| \leq \sum_{i=1}^{\infty} |a_i||z|^i \equiv g(z).$$

Wie für $\tilde{f}$ gibt es $0 < \rho < \delta$, sodaß $g(z) < 1$ für alle $z \in K(0, \rho)$ gilt. Auch g^j läßt sich mit geeigneten Koeffizienten d_{jk}, $j, k \in \mathbb{N}_0$, als Potenzreihe darstellen,

$$g(z)^j = \sum_{k=0}^{\infty} d_{jk} |z|^k, \quad j \in \mathbb{N}_0.$$

Es konvergiert also für $|z| < \rho$ auch die Reihe

$$\sum_{j=0}^{\infty} g(z)^j = \sum_{j=0}^{\infty} \left[\sum_{k=0}^{\infty} d_{jk} |z|^k \right].$$

Aus der Definition von g ergibt sich

$$|a_{jk}| \leq d_{jk}, \quad j, k \in \mathbb{N}_0, \tag{1}$$

und daraus mit dem Majorantenkriterium die Konvergenz von $\sum_{j=0}^{\infty} \left[\sum_{k=0}^{\infty} |a_{jk}||z|^k \right]$. Die Gültigkeit von (1) für $j = 2$ ergibt sich aus der Definition des Cauchy-Produktes: demnach ist $a_{2k} = \sum_{l=0}^{k} a_{k-l} a_l$ und $d_{2k} = \sum_{l=0}^{k} |a_{k-l}||a_l|$ und somit folgt $|a_{2k}| \leq \sum_{l=0}^{k} |a_{k-l}||a_l| = d_{2k}$. Mit vollständiger Induktion kann man nun die Gültigkeit von (1) für $j > 2$ nachweisen. □

Die praktische Berechnung der (oder einiger) Koeffizienten von $\frac{1}{f}$ beruht meist auf der Methode des Koeffizientenvergleichs, die wir an einem einfachen Beispiel demonstrieren.

Beispiel IV-4.16. Es sei $f(z) = 1 - z$. Wegen $f(0) = 1$ läßt sich nach Satz IV-4.15 $\frac{1}{f}$ in eine Reihe $\sum_{k=0}^{\infty} b_k z^k$ entwickeln, welche auf $K(0,\rho)$, $\rho > 0$ hinreichend klein, konvergiert. Für $|z| < \rho$ ergibt sich die Identität

$$\begin{aligned} 1 &= \left(\sum_{k=0}^{\infty} b_k z^k\right)(1-z) = \sum_{k=0}^{\infty} b_k z^k - \sum_{k=1}^{\infty} b_{k-1} z^k \\ &= b_0 + \sum_{k=1}^{\infty} (b_k - b_{k-1}) z^k. \end{aligned}$$

Wegen des Identitätssatzes IV-4.12 muß demnach

$$b_0 = 1 \qquad \text{und} \qquad b_k - b_{k-1} = 0, \quad k \in \mathbb{N},$$

gelten. Daraus folgt $b_k = 1$, $k \in \mathbb{N}_0$.

Der Divisionssatz ist ein Spezialfall eines weitaus allgemeineren Resultates, das analog bewiesen werden kann.

Theorem IV-4.17 (Einsetzungssatz). *Die Konvergenzradien der Potenzreihen $f(z) = \sum_{k=0}^{\infty} f_k z^k$ und $g(z) = \sum_{k=0}^{\infty} g_k z^k$ seien R_f bzw. R_g, ferner gelte $|f_0| < R_g$. Es gibt dann ein $\delta < R_f$, sodaß $g \circ f$ auf $K(0,\delta)$ definiert und in eine Potenzreihe entwickelbar ist:*

$$(g \circ f)(z) = \sum_{k=0}^{\infty} c_k z^k, \quad |z| < \delta.$$

5. Elementare Funktionen

Die Untersuchung der Eigenschaften von Funktionen kann oft erheblich vereinfacht werden, wenn die Funktionen gewisse Symmetrien aufweisen.

Definition IV-5.1. *D sei eine* **symmetrische** *Teilmenge von $\mathbb{K}$, d.h.*

$$\forall x \in \mathbb{K} \colon x \in D \Leftrightarrow -x \in D,$$

und $f \in \colon D \to \mathbb{K}$.

i) *f heißt* ***gerade*** $\underset{Def}{\Leftrightarrow} \forall x \in D \colon f(-x) = f(x)$,

ii) *f heißt* ***ungerade*** $\underset{Def}{\Leftrightarrow} \forall x \in D \colon f(-x) = -f(x)$.

Eine ungerade Funktion nimmt an der Stelle $x = 0$ den Wert 0 an. Die Umkehrfunktion einer ungeraden Funktion (falls sie existiert) ist ebenfalls ungerade. Der Graph einer geraden Funktion ist symmetrisch zur y-Achse, der Graph einer ungeraden Funktion ist symmetrisch zum Ursprung.

5.1. Potenzfunktion, Wurzelfunktion.

Wir haben bereits Potenzen mit rationalem Exponenten erklärt und betrachten nun diese Operation unter einem abbildungstheoretischen Blickwinkel. Wir beschränken uns vorerst auf Exponenten aus $\mathbb{N}$.

Theorem IV-5.2 (Potenzfunktion). *Für $n \in \mathbb{N}$ hat die* ***Potenzfunktion*** *p_n, erklärt durch*

$$p_n : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^n \end{cases}$$

folgende Eigenschaften.

Fall 1. *n ist gerade:*

i) *p_n ist eine gerade Funktion,*
ii) *$p_n|_{\mathbb{R}^+}$ ist streng monoton wachsend, $p_n|_{\mathbb{R}^-}$ ist streng monoton fallend,*
iii) $p_n(\mathbb{R}) = \mathbb{R}^+$.

Fall 2. *n ist ungerade:*
i) *p_n ist eine ungerade Funktion,*
ii) *p_n ist streng monoton wachsend,*
iii) $p_n(\mathbb{R}) = \mathbb{R}$.

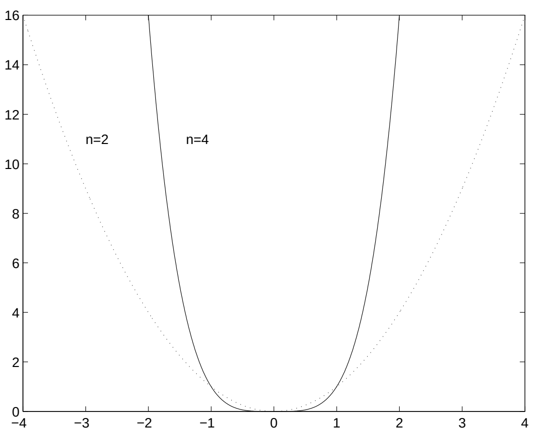
Weiters gilt: Die Potenzfunktionen p_n, $n \in \mathbb{N}$ sind stetig.

Beweis. Die Beziehung $p_n(-x) = (-1)^n x^n = (-1)^n p_n(x)$ (vgl. Satz I-4.17) zeigt, daß p_n eine gerade Funktion ist für n gerade und eine ungerade Funktion für n ungerade. Wegen Satz I-4.17-d ist $p_n|_{\mathbb{R}^+}$ für jedes $n \in \mathbb{N}$ streng monton wachsend. Es sei nun $x < y < 0$, d.h. $0 < -y < -x$. Somit folgt $p_n(-y) < p_n(-x)$, also gilt

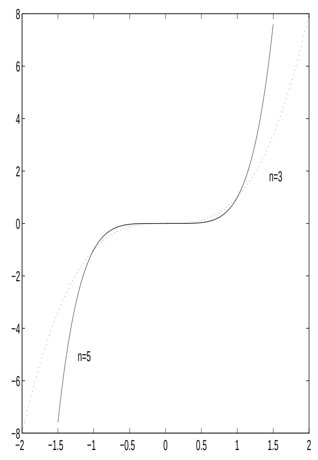
$$(4) \qquad \begin{aligned} p_n(y) &< p_n(x) && \text{falls } n \text{ gerade ist und} \\ -p_n(y) &< -p_n(x), && \text{d.h. } p_n(x) < p_n(y), \text{ falls } n \text{ ungerade ist.} \end{aligned}$$

Die Charakterisierung des Bildes von p_n folgt aus Satz IV-2.4, die Stetigkeit aus Satz III-1.16. $\square$

Wir veranschaulichen die Potenzfunktion, indem wir die Graphen von p_n, $n = 2, 3, 4, 5$ skizzieren.



gerade Potenzen



ungerade Potenzen

Es sei n gerade und $g_n = p_n|_{\mathbb{R}^+}$. Aus dem eben Bewiesenen folgt, daß g_n eine Bijektion ist. Somit existiert die Umkehrfunktion $g_n^{-1}\colon \mathbb{R}^+ \to \mathbb{R}^+$. Für alle $y \in \mathbb{R}^+$ ist die Beziehung $x = g_n^{-1}(y)$ gleichwertig mit $y = g_n(x) = x^n$, mit Satz I-7.3 folgt also $g_n^{-1}(y) = \sqrt[n]{y}$. Es gilt $g_n^{-1}(\mathbb{R}^+) = D(g_n) = \mathbb{R}^+$ und da g_n streng monoton wächst, ist auch g_n^{-1} streng monoton wachsend (Satz IV-2.2). Eine analoge Betrachtung gilt für n ungerade und $\mathbb{R}$ anstelle von $\mathbb{R}^+$. Wir fassen diese Diskussion zusammen in

THEOREM IV-5.3 (Wurzelfunktion). *Für $n \in \mathbb{N}$, $n \geq 2$ wird die* ***Wurzelfunktion*** *w_n, erklärt durch*

$$w_n : \begin{cases} \mathbb{R}^+ \to \mathbb{R}^+, & n \textit{ gerade}, \\ \mathbb{R} \to \mathbb{R}, & n \textit{ ungerade}, \\ x \mapsto \sqrt[n]{x}. \end{cases}$$

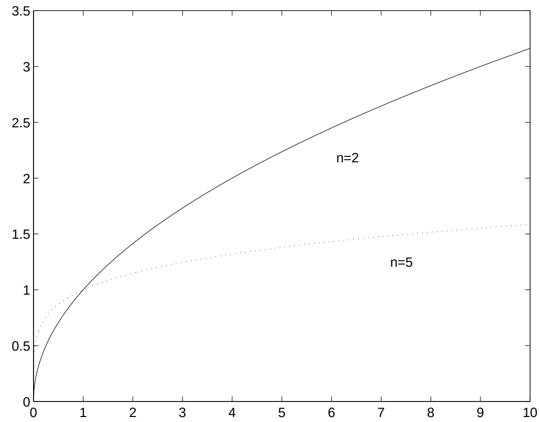
Sie hat folgende Eigenschaften:

i) *w_n ist streng monoton wachsend,*
ii) *$w_n(\mathbb{R}^+) = \mathbb{R}^+$ für n gerade und $w_n(\mathbb{R}) = \mathbb{R}$ für n ungerade.*
iii) *Die Wurzelfunktionen w_n sind stetig.*

BEWEIS. iii) Satz IV-2.8.

□

Wir skizzieren den Graph der Wurzelfunktionen w_2, w_5 auf $\mathbb{R}^+$.



Wurzelfunktion

5.2. Die Exponentialfunktion. Wir erinnern daran, daß die Reihe $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ für alle $z \in \mathbb{C}$ konvergiert (Beispiel-IV-4.5-5)). Dies rechtfertigt folgende Definition.

DEFINITION IV-5.4. *Die* ***Exponentialfunktion*** exp *ist definiert durch*

$$\exp\colon \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto \sum\limits_{k=0}^{\infty} \frac{z^k}{k!} \end{cases}$$

(Wir vereinbaren für diese Definition $0^0 = 1$.)

THEOREM IV-5.5. *Eigenschaften der Exponentialfunktion in $\mathbb{C}$.*

i) $\forall z, w \in \mathbb{C}\colon\ \exp(z)\exp(w) = \exp(z + w)$,
ii) $\exp(0) = 1, \exp(1) = e$,
iii) $\forall z \in \mathbb{C}\colon\ \exp(z) \neq 0$, $\exp(-z) = \frac{1}{\exp(z)}$,
iv) $\exp \in C(\mathbb{C}, \mathbb{C})$.

Beweis. i) Die Koeffizienten c_k des Cauchy Produktes der beiden Reihen $\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ und $\exp(w) = \sum\limits_{k=0}^{\infty} \frac{w^k}{k!}$ sind

$$c_k = \sum_{i=0}^{k} \frac{z^i}{i!} \frac{w^{k-i}}{(k-i)!} = \sum_{i=0}^{k} \binom{k}{i} z^i w^{k-i} \frac{1}{k!} = \frac{1}{k!}(z+w)^k.$$

Daraus folgt die Behauptung.
ii) Satz IV-4.6.
iii) folgt aus i) und ii): $\exp(0) = 1 = \exp(z)\exp(-z)$.
iv) Satz Satz IV-4.7. □

Wir werden die Exponentialfunktion auf $\mathbb{C}$ hier nicht weiter untersuchen und beschränken uns im folgenden auf reelle Argumente. Wir bezeichnen auch die Einschränkung der Exponentialfunktion auf $\mathbb{R}$ mit exp.

Theorem IV-5.6. *Weitere Eigenschaften der Exponentialfunktion in* $\mathbb{R}$*:*

i) $\forall x \in \mathbb{R}\colon \exp(x) > 0$,
ii) exp *ist streng monoton wachsend auf* $\mathbb{R}$,
iii) $\lim_{x\to\infty} \exp(x) = \infty$, $\lim_{x\to-\infty} \exp(x) = 0$,
iv) $\exp(\mathbb{R}) = (0, \infty)$,
v) *Es gilt* $\lim_{x\to\infty} x^{-n} \exp x = \infty$ *und* $\lim_{x\to\infty} x^n \exp(-x) = 0$ *für alle* $n \in \mathbb{N}$.

Beweis. i) Aus der Reihendarstellung folgt $\exp(\mathbb{R}) \subset \mathbb{R}$ und

$$(*) \qquad \forall x > 0\colon \exp(x) > 1 + x > 1.$$

Somit gilt für alle $x \in \mathbb{R}^+\colon \exp(x) \geq 1 > 0$. i) folgt nun aus

$$\exp(-x) = (\exp(x))^{-1} > 0.$$

ii) Es sei $x < y$, also $y - x > 0$. Aus (*) folgt $1 < \exp(y-x) = \exp(y)\exp(-x) = \exp(y)(\exp(x))^{-1}$.
iii) $\lim_{x\to\infty} \exp(x) = \infty$ folgt aus (*), dies zusammen mit Satz IV-5.5-iii) impliziert $\lim_{x\to-\infty} \exp(x) = 0$.
iv) Ergibt sich aus ii), iii) und Satz IV-2.4.
v) Für festes $n \in \mathbb{N}$ und $x > 0$ erhält man aus der Reihendarstellung die Abschätzung

$$\frac{\exp(x)}{x^n} > \frac{x^{n+1}}{(n+1)!} \frac{1}{x^n} = \frac{x}{(n+1)!}$$

und somit $\lim_{x\to\infty} \frac{\exp(x)}{x^n} \geq \lim_{x\to\infty} \frac{x}{(n+1)!} = \infty$. □

Der letzte Punkt des Satzes zeigt, daß die Exponentialfunktion stärker wächst als jede noch so hohe Potenz von x.

Rationale Potenzen einer positiven reellen Zahl wurden bereits in Definition I-7.8 eingeführt. Dies führt zu keinem Konflikt mit den entsprechenden Werten der Exponentialfunktion: Wegen Satz IV-5.5-ii) gilt nämlich $\exp(1) = e$, also auch $\exp(n) = e^n$,

$n \in \mathbb{N}$ (Satz IV-5.5-i). Auf analoge Weise findet man $e = \exp(\frac{1}{n} \cdot n) = (\exp(\frac{1}{n}))^n$, dh. $\exp(\frac{1}{n}) = e^{\frac{1}{n}}$. Somit ergibt sich für $r = \frac{p}{q} \in \mathbb{Q}^+$, $p \in \mathbb{N}_0$, $q \in \mathbb{N}$

$$\exp(r) = \exp(\frac{p}{q}) = (\exp(\frac{1}{q}))^p = (e^{\frac{1}{q}})^p = e^r.$$

Für $r \le 0$ verwenden wir die Identität

$$\exp(r) = \frac{1}{\exp(-r)} = \frac{1}{e^{-r}} = e^r.$$

Somit stimmen die Funktionen exp und $r \mapsto e^r$ auf $\mathbb{Q}$ überein. Wegen der Stetigkeit der Exponentialfunktion kann exp als stetige Fortsetzung der Funktion $x \to e^x$ von von $\mathbb{Q}$ auf $\mathbb{R}$ aufgefaßt werden. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, ist diese stetige Fortsetzung eindeutig bestimmt. Dies rechtfertigt auch die Bezeichnung Exponentialfunktion für exp.

Definition IV-5.7. *Für alle $x \in \mathbb{R}$ setzen wir $e^x = \exp(x)$.*

Im Folgenden verzichten wir auf die Unterscheidung zwischen $x \to \exp(x)$ und $x \to e^x$. Nach Satz IV-5.6 besitzt die Exponentialfunktion eine Umkehrfunktion, den natürlichen Logarithmus.

Definition IV-5.8 (Logarithmusfunktion). *Die Umkehrfunktion der Exponentialfunktion $x \to e^x$, heißt **natürliche Logarithmusfunktion**,* ln*:*

$$\ln\colon \begin{cases} (0,\infty) \to \mathbb{R} \\ y \mapsto \ln(y). \end{cases}$$

Den Wert $\ln(y)$ *der Logarithmusfunktion an der Stelle* y *nennt man **natürlicher Logarithmus** von* y. *Er ist bestimmt durch die Beziehung*

$$x = \ln y \Leftrightarrow y = e^x.$$

Theorem IV-5.9. *Die Logarithmusfunktion ist stetig und hat folgende Eigenschaften:*

i) $\forall x, y \in (0,\infty)\colon\ \ln(xy) = \ln x + \ln y$,
ii) $\forall x \in (0,\infty)\, \forall \rho \in \mathbb{Q}\colon\ \ln x^\rho = \rho \ln x$,
iii) $\ln 1 = 0$, $\ln e = 1$,
iv) $\mathrm{bild}(\ln) = \mathbb{R}$,
v) ln *ist streng monoton steigend.*
vi) $\lim_{x\to\infty} \ln x = \infty$, $\lim_{x\to 0^+} \ln x = -\infty$.

Beweis. Die Stetigkeit ergibt sich aus Satz IV-2.8. Ausgehend von der Identität $\forall x \in (0,\infty)\colon x = e^{\ln x}$ schließt man

$$e^{\ln xy} = xy = e^{\ln x} e^{\ln y} = e^{\ln x + \ln y},$$
$$e^{\ln x^\rho} = x^\rho = (e^{\ln x})^\rho = e^{\rho \ln x}.$$

Wegen der Injektivität der Exponentialfunktion folgen die Behauptungen i) und ii), iii) ergibt sich aus Satz IV-5.5-ii), iv) gilt, weil jede Exponentialfunktion auf $\mathbb{R}$ definiert ist und v) ist eine Konsequenz aus Satz IV-2.2. $\square$

Die Gültigkeit der Potenzregel wird später auf beliebige reelle Exponenten ausgedehnt.

5.3. Potenzen mit Reellen Exponenten.

Im vorigen Abschnitt wurden reelle Potenzen der Basis e eingeführt. Es sei nun $a > 0$. Ausgehend von der Identität $a = e^{\ln a}$ findet man für $r \in \mathbb{Q}$

$$a^r = e^{\ln(a^r)} = e^{r \ln a}.$$

Approximiert man $x \in \mathbb{R}$ durch eine Folge rationaler Zahlen $(r_n) \subset \mathbb{Q}$, $\lim_{n\to\infty} r_n = x$, folgt aus der Stetigkeit der Exponentialfunktion die Existenz des Grenzwertes

$$\lim_{n\to\infty} a^{r_n} = \lim_{n\to\infty} e^{r_n \ln a} = e^{\lim_{n\to\infty} r_n \ln a} = e^{x \ln a}$$

Dieser Grenzwert hängt nur von a und x ab. Ist x selbst rational, ergibt sich

$$\lim_{n\to\infty} a^{r_n} = e^{x \ln a} = e^{\ln(a^x)} = a^x.$$

Diese Überlegungen motivieren folgende Definition.

Definition IV-5.10. *Für $a > 0$ und $x \in \mathbb{R}$ definieren wir die* ***allgemeine Potenz***

$$a^x = e^{x \ln a},$$

für $x > 0$ setzen wir ferner

$$0^x = 0.$$

Aus der Definition des natürlichen Logarithmus folgt dann die Identität

$$\ln(a^x) = x \ln a$$

(man beachte: $a^x > 0$ für $x \in \mathbb{R}$).
Potenzen mit reellen Exponenten erfüllen die üblichen Eigenschaften von Potenzen:

Theorem IV-5.11 (Reelle Potenzen). *Es seien a, $b > 0$ und $x, y \in \mathbb{R}$. Dann gilt:*

(1) $a^{x+y} = a^x a^y$,
(2) $(ab)^x = a^x b^x$,
(3) $(a^x)^y = a^{xy}$,
(4) $a^x > 0$ *und* $a^{-x} = \frac{1}{a^x} = (\frac{1}{a})^x$,
(5) $0 < a^x < b^x$ *für* $0 < a < b$ *und* $x > 0$,
$0 < b^x < a^x$ *für* $0 < a < b$ *und* $x < 0$,
(6) $a^x < a^y$ *für* $x < y$ *und* $a > 1$,
(7) $a^y < a^x$ *für* $x < y$ *und* $0 < a < 1$.

Beweis. Wir zeigen $(a^x)^y = a^{xy}$. Dies folgt aus

$$(a^x)^y = e^{y \ln(a^x)} = e^{xy \ln a} = a^{xy}.$$

Der Nachweis der anderen Regeln sei dem Leser als Übung überlassen. □

THEOREM IV-5.12 (Potenzfunktion). *Für $\rho \in \mathbb{R}$ hat die Potenzfunktion p_ρ, definiert durch*

$$p_\rho\colon \begin{cases} (0,\infty) \to \mathbb{R} \\ x \mapsto x^\rho \end{cases}$$

folgende Eigenschaften:

(1) *p_ρ ist stetig.*
(2) $\operatorname{bild} p_\rho = (0,\infty)$ *für $\rho \neq 0$.*
(3) *p_ρ ist streng monoton wachsend für $\rho > 0$ und streng monoton fallend für $\rho < 0$.*
(4) $\lim_{x\to 0^+} p_\rho(x) = 0$, $\lim_{x\to\infty} p_\rho(x) = \infty$ *für $\rho > 1$.*
(5) *Für $\rho = 0$ ist $p_\rho \equiv 1$.*

BEWEIS. Die Gleichung $x^\rho = y$ besitzt für jedes $y > 0$ die eindeutige Lösung $x = y^{\frac{1}{\rho}}$. Somit ist $(0,\infty) \subset \operatorname{bild} p_\rho \subset (0,\infty)$. Die Monotonieeigenschaft für $\rho > 0$ wurde bereits in Satz IV-5.11-(5) festgestellt. Für den Fall $\rho < 0$ beachte man $x^\rho = (\frac{1}{x})^{|\rho|}$. Die Grenzwerte in iii) folgen aus Satz IV-5.6. □

BEMERKUNG IV-5.13. Für nicht negative Exponenten ρ wird die Potenzfunktion stetig nach 0 fortgesetzt, wenn man $p_\rho(0) = 0$ vereinbart.

5.4. Exponentialfunktion, Logarithmusfunktion.

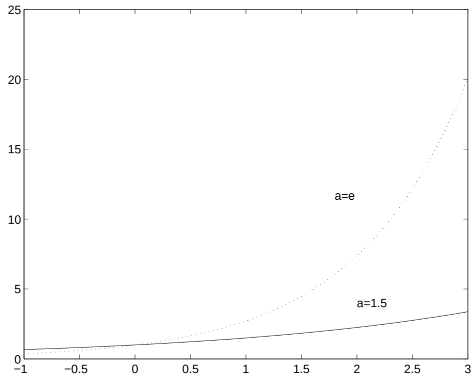
THEOREM IV-5.14 (Exponentialfunktion). *Für $a > 0$ ist die* ***Exponentialfunktion*** *zur* ***Basis*** **a**, *definiert durch*

$$\begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto a^x \end{cases}$$

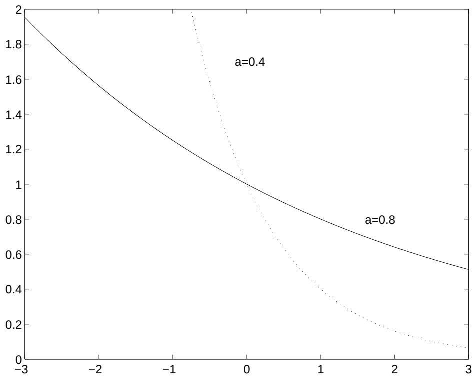
stetig und besitzt folgende Eigenschaften:

i) $\forall x, y \in \mathbb{R}\colon a^{x+y} = a^x a^y$,
ii) $a^0 = 1$, $a^1 = a$,
iii) $\forall x \in \mathbb{R}\colon a^x > 0$ *und* $a^x = \frac{1}{a^{-x}}$,
iv) *$x \to a^x$ ist streng monoton wachsend für $a > 1$ und streng monoton fallend für $a < 1$,*
v) *der Wertebereich der Exponentialfunktion ist $(0,\infty)$ für $a \neq 1$.*

BEWEIS. Die Stetigkeit der Exponentialfunktion folgt aus Satz IV-5.5 und Satz III-2.15. Die Aussagen i) – iv) folgen aus Satz IV-5.11. Die Behauptung v) für $a > 1$ ergibt sich aus Satz IV-2.4 und $\sup_{x\in\mathbb{R}} a^x = \infty$, $\inf_{x\in\mathbb{R}} a^x = 0$. □



$x \mapsto a^x$, $a > 1$



$x \mapsto a^x$, $a < 1$

Jede Exponentialfunktion mit $a \neq 1$ ist also injektiv und besitzt daher eine Umkehrfunktion.

THEOREM IV-5.15 (Logarithmusfunktion). *Die Umkehrfunktion der Exponentialfunktion*
$x \to a^x$, $a > 0$, $a \neq 1$, *heißt **Logarithmusfunktion zur Basis** **a**,* $\log_a$*:*

$$\log_a \colon \begin{cases} (0,\infty) \to \mathbb{R} \\ y \mapsto \log_a(y). \end{cases}$$

Den Wert der Logarithmusfunktion an der Stelle y*,* $\log_a(y)$*, nennt man **Logarithmus*** **von** y **zur Basis** a*. Er ist bestimmt durch die Beziehung*

$$x = \log_a y \Leftrightarrow y = a^x.$$

Die Logarithmusfunktionen sind stetig und haben folgende Eigenschaften:

i) $\forall x, y \in (0,\infty)\colon\ \log_a(xy) = \log_a x + \log_a y$,
ii) $\forall x \in (0,\infty)\, \forall \rho \in \mathbb{R}\colon\ \log_a x^\rho = \rho \log_a x$,
iii) $\forall a \in (0,\infty) \setminus \{1\}\colon\ \log_a 1 = 0$, $\log_a a = 1$,
iv) $\mathrm{bild}(\log_a) = \mathbb{R}$,
v) $\log_a$ *ist streng monoton steigend für* $a > 1$ *und streng monoton fallend für* $0 < a < 1$.

BEWEIS. Die Behauptungen i) und ii) beweist man wie in Satz IV-5.9, iii) ergibt sich aus Satz IV-5.14-ii), iv) gilt, weil jede Exponentialfunktion auf $\mathbb{R}$ definiert ist und v) ist eine Konsequenz aus Satz IV-2.2. □

Die eindeutige Lösung dr Gleichung $y = a^x$ ist also $x = \log_a y$. Aus der Identität

$$y = a^x = e^{x \ln a}$$

folgt $\ln y = x \ln a = \log_a y \ln a$, also

$$\log_a y = \frac{\ln y}{\ln a}.$$

Das rasche Wachstum der Exponentialfunktion (mit einer Basis $a > 1$), spiegelt sich im extrem langsamen Wachsen der Logarithmen wieder. Analog zu Satz IV-5.6 gilt nämlich:

THEOREM IV-5.16. *Für jedes $\alpha > 0$ gilt*

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = 0, \qquad \lim_{x\to 0^+} x^\alpha \ln x = 0.$$

BEWEIS. Es genügt (warum) den ersten Grenzwert nachzuweisen: dieser folgt aus der Abschätzung

$$\frac{\ln x}{x^\alpha} = \frac{\ln x}{e^{\alpha \ln x}} \le \frac{2\ln x}{\alpha^2(\ln x)^2} = \frac{2}{\alpha^2 \ln x},$$

für $x \ge 1$ und Satz IV-5.9. □

Der (natürliche) Logarithmus wächst somit schwächer, als jede noch so kleine positive Potenz von x. Satz IV-5.16 läßt sich leicht auf Logarithmen mit beliebiger Basis $a > 1$ übertragen.

5.5. Trigonometrische Funktionen. Wir definieren die trigonometrischen Funktionen durch Potenzreihen in $\mathbb{C}$ und zeigen, daß diese Funktionen eingeschränkt auf $\mathbb{R}$ tatsächlich die bekannten Eigenschaften besitzen. Wir schreiben e^z anstelle von $\exp(z)$.

DEFINITION IV-5.17. *Die **Sinus**-Funktion* sin *und die **Cosinus**-Funktion* cos *sind definiert durch* $(0^0 := 1)$

$$\sin z\colon \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto \sum\limits_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \end{cases}, \qquad \cos z\colon \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto \sum\limits_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} \end{cases}.$$

Man kann leicht verifizieren, daß die Potenzreihen für $\sin z$ und $\cos z$ für alle $z \in \mathbb{C}$ konvergieren.

THEOREM IV-5.18.

(1) sin *und* cos *sind stetig auf* $\mathbb{C}$.

(2) EULER*sche Formeln. Für alle $z \in \mathbb{C}$ gilt:*

i) $e^{iz} = \cos z + i \sin z$, *iii)* $\sin z = \frac{1}{2i}[e^{iz} - e^{-iz}]$,

ii) $e^{-iz} = \cos z - i \sin z$, *iv)* $\cos z = \frac{1}{2}[e^{iz} + e^{-iz}]$.

BEWEIS. Es genügt, i) und ii) von (2) zu beweisen. Wegen $(-1)^n z^{2n} = (iz)^{2n}$ und $i(-1)^n z^{2n+1} = (iz)^{2n+1}$, $n \in \mathbb{N}_0$, folgt i) aus

$$\begin{aligned} \cos z + i \sin z &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(iz)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(iz)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} = e^{iz}. \end{aligned}$$

ii) folgt aus i) indem man z durch $-z$ ersetzt. □

THEOREM IV-5.19 (Trigonometrische Identitäten). *Für alle $z, w \in \mathbb{C}$ gilt*
i) $\sin(-z) = -\sin z$, *dh.* $\sin$ *ist ungerade,*
ii) $\cos(-z) = \cos z$, *dh.* $\cos$ *ist gerade,*
iii) $\sin^2 z + \cos^2 z = 1$,
iv) $\sin(z + w) = \sin z \cos w + \cos z \sin w$,
v) $\cos(z + w) = \cos z \cos w - \sin z \sin w$,
vi) $\sin z - \sin w = 2 \cos \frac{z+w}{2} \sin \frac{z-w}{2}$,
vii) $\cos z - \cos w = -2 \sin \frac{z+w}{2} \sin \frac{z-w}{2}$.

BEWEIS. i), ii) sind klar.
iii) folgt aus $1 = e^{iz}e^{-iz} = (\cos z + i \sin z)(\cos z - i \sin z)$ (vgl.Satz IV-5.18-i,ii)).
iv) Aus Satz IV-5.18-iii) ergibt sich

$$\begin{aligned} 4i \sin(z + w) &= 2(e^{i(z+w)} - e^{-i(z+w)}) \\ &= e^{i(z+w)} + e^{i(z-w)} - e^{-i(z-w)} - e^{-i(z+w)} \\ &\quad + e^{i(z+w)} - e^{i(z-w)} + e^{-i(z-w)} - e^{-i(z+w)} \\ &= (e^{iz} - e^{-iz})(e^{iw} + e^{-iw}) + (e^{iz} + e^{-iz})(e^{iw} - e^{-iw}) \\ &= 4i(\sin z \cos w + \cos z \sin w). \end{aligned}$$

v) analog.
vi) und vii) folgen aus iv) und v). □

KOROLLAR IV-5.20. Für alle $x \in \mathbb{R}$ gilt $|e^{ix}| = 1$.

BEWEIS. Die Behauptung folgt aus Satz IV-5.19-iii) und der Identität

$$e^{ix} = \cos x + i \sin x.$$

□

Wir wenden uns nun den Eigenschaften von $\sin|_{\mathbb{R}}$ und $\cos|_{\mathbb{R}}$ zu: Zuerst demonstrieren wir die Approximation des Sinus auf $[0, 6]$ durch die ersten 5 Partialsummen seiner Potenzreihendarstellung (Abbildung 5.5)

THEOREM IV-5.21 (Die Zahl π). *Es gibt eine eindeutig bestimmte Zahl π mit folgenden Eigenschaften*

i) $0 < \pi < 4$,
ii) $\cos \frac{\pi}{2} = 0$,
iii) *es ist* $\cos x > 0$ *für alle* $x \in [0, \frac{\pi}{2})$.

Die Zahl $\frac{\pi}{2}$ ist also die kleinste positive Nullstelle des Cosinus.

BEWEIS. Angenommen es gäbe zwei Zahlen π und π', etwa $\pi < \pi'$, mit den Eigenschaften i) – iii). Wegen i) folgt $0 < \frac{\pi}{2} < \frac{\pi'}{2}$ und aus iii) $\cos x > 0$ für $x \in [0, \frac{\pi'}{2})$, dh. $\cos \frac{\pi}{2} > 0$. Es kann somit höchstens eine derartige Zahl existieren.

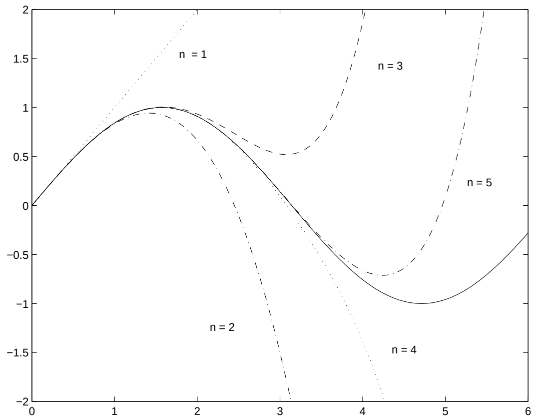


ABBILDUNG 1. Approximation des Sinus

Aus Korollar IV-II-11.2 folgt

$$\cos 2 \leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = -\frac{1}{3} < 0.$$

Da $\cos 0 = 1 > 0$ und cos stetig ist, muß nach dem Zwischenwertsatz IV-1.1 in $(0,2)$ mindestens eine Nullstelle des Cosinus liegen, dh.

$$\mathcal{N} = \{x \in [0,2] \colon \cos x = 0\} \neq \emptyset.$$

Wir setzen $\frac{\pi}{2} := \inf \mathcal{N}$. Wegen der Stetigkeit von cos gilt $\frac{\pi}{2} \in \mathcal{N}$. Aus $\cos 0 = 1$ folgt dann $\frac{\pi}{2} > 0$, dh. π erfüllt i) und ii). Wegen der Minimaleigenschaft von $\frac{\pi}{2}$ und $\cos 0 = 1$ muß dann aber auch iii) erfüllt sein: Wäre $\cos \xi < 0$ für ein $\xi \in (0, \frac{\pi}{2})$, müßte es in $(0, \xi)$ nach dem Zwischenwertsatz eine weitere Nullstelle $\eta < \frac{\pi}{2}$ geben. □

Es ist derzeit natürlich nicht einsichtig, daß die eben definierte Zahl π in irgendeinem Zusammenhang mit dem Umfang oder Flächeninhalt eines Kreises steht. Wir können diese Frage erst dann befriedigend klären, wenn wir uns mit dem Problem der Längen– und Flächenmessung auseinandergesetzt haben.

LEMMA IV-5.22.

$$\forall x \in [0,2] \colon \sin x \geq \frac{x}{3}.$$

Insbesondere ist $\sin x > 0$ *für* $x \in (0, 2]$.

BEWEIS. Aus der Reihe für sin und Korollar IV-II-11.2 folgt

$$\sin x \geq x - \frac{x^3}{3!} = x(1 - \frac{x^2}{6}) \geq \frac{x}{3}, \quad x \in [0,2].$$

□

THEOREM IV-5.23. sin *ist streng monoton wachsend und* cos *streng monoton fallend auf* $[0, \frac{\pi}{2}]$.

Beweis. Es sei $0 \leq x < y \leq \frac{\pi}{2}$, also $0 < \frac{x+y}{2} < \frac{\pi}{2}$ und $0 < \frac{y-x}{2} < \frac{\pi}{2}$. Aus Satz IV-5.19-vi,vii, i, ii, Lemma IV-5.22, Satz IV-5.21, folgt nun

$$\sin x - \sin y = 2\cos\frac{x+y}{2}\sin\frac{x-y}{2} = -2\cos\frac{x+y}{2}\sin\frac{y-x}{2} < 0$$
$$\cos x - \cos y = -2\sin\frac{x+y}{2}\sin\frac{x-y}{2} = 2\sin\frac{x+y}{2}\sin\frac{y-x}{2} > 0.$$

□

Wir komplettieren nun das Bild der Graphen von Sinus und Cosinus:

Theorem IV-5.24. i) sin *ist streng monoton wachsend auf* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *und streng monoton fallend auf* $[-\pi, -\frac{\pi}{2}]$ *und auf* $[\frac{\pi}{2}, \pi]$.

ii) cos *ist streng monoton fallend auf* $[0, \pi]$ *und streng monoton wachsend auf* $[-\pi, 0]$.

iii) $\sin(-\pi) = \sin 0 = \sin\pi = 0$, $\sin(-\frac{\pi}{2}) = -1$, $\sin\frac{\pi}{2} = 1$.

iv) $\cos(-\frac{\pi}{2}) = \cos\frac{\pi}{2} = 0$, $\cos(-\pi) = \cos\pi = -1$, $\cos 0 = 1$.

v) $\forall x \in \mathbb{R}\colon \sin(x + \frac{\pi}{2}) = \cos x$, $\cos(x + \frac{\pi}{2}) = -\sin x$.

vi) $\forall x \in \mathbb{R}\colon \sin(x + 2\pi) = \sin x$, $\cos(x + 2\pi) = \cos x$.

vii) $\forall x \in \mathbb{R}\colon |\sin x| \leq 1$, $|\cos x| \leq 1$.

Beweis. Wegen $\cos\frac{\pi}{2} = 0$ muß nach Satz IV-5.19-iii) $\sin\frac{\pi}{2} = \pm 1$ sein, wegen Lemma IV-5.22 folgt $\sin\frac{\pi}{2} = 1$ und da sin ungerade ist, auch $\sin(-\frac{\pi}{2}) = -1$. Für $-\frac{\pi}{2} \leq x < y \leq 0$, also $0 \leq -y < -x \leq \frac{\pi}{2}$, finden wir

$$-\sin y = \sin(-y) < \sin(-x) = -\sin x \quad \text{d.h.} \quad \sin x < \sin y.$$

Also ist sin streng monoton wachsend auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Aus dem Additionstheorem Satz IV-5.19-iv) folgt nun $\sin(x + \frac{\pi}{2}) = \cos x$ und $\cos(x + \frac{\pi}{2}) = -\sin x$, $x \in \mathbb{R}$. Nach Satz IV-5.23 ist also sin streng monoton fallend auf $[\frac{\pi}{2}, \pi]$ und damit aus Symmetriegründen auch auf $[-\pi, -\frac{\pi}{2}]$. Darüber hinaus gilt $\sin\pi = \cos\frac{\pi}{2} = 0 = \sin(-\pi)$. Somit sind i), iii) und v) bewiesen. Aus v) folgt für $x \in \mathbb{R}$

$$\sin(x + 2\pi) = \cos(x + \frac{3\pi}{2}) = -\sin(x + \pi) = -\cos(x + \frac{\pi}{2}) = \sin x$$

und analog für cos. Dies zeigt vi). Die Eigenschaften des Cosinus können nun aus denen des Sinus abgelesen werden: $\cos(-\pi) = \sin(-\pi + \frac{\pi}{2}) = \sin(-\frac{\pi}{2}) = -1$. Für $x \in [-\pi, 0]$ ist $x + \frac{\pi}{2} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ und daher $\cos x = \sin(x + \frac{\pi}{2})$ streng monoton wachsend auf $[-\pi, 0]$ und aus Symmetriegründen streng monoton fallend auf $[0, \pi]$. vii) ist eine einfache Konsequenz aus Satz IV-5.19-iii). □

Bemerkung IV-5.25. 1) Zusammen mit den Eulerschen Formeln erhalten wir nun die überraschende Beziehung $e^{i\pi} = -1$.
2) Die Eigenschaft vi) bedeutet, daß sin und cos **periodische Funktionen** mit **Periode** 2π sind. Man nennt eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ periodisch mit Periode p, wenn $f(x + p) = f(x)$ für alle $x \in \mathbb{R}$ gilt.

LEMMA IV-5.26. *Die Menge der Nullstellen des Cosinus ist* $\{(k+\frac{1}{2}\pi\colon k\in\mathbb{Z}\}$, *die Menge der Nullstellen des Sinus ist* $\{k\pi\colon k\in\mathbb{Z}\}$.

BEWEIS. Die einzige Nullstelle des Cosinus in $(-\frac{\pi}{2},\frac{\pi}{2}]$ ist $\frac{\pi}{2}$. Wegen $\cos(x+\pi)=-\cos x$ sind also $\frac{\pi}{2}$ und $\frac{3\pi}{2}$ die einzigen Nullstellen in $(-\frac{\pi}{2},\frac{3\pi}{2}]$. Dieses Intervall hat die Länge 2π. Wegen der Periodizität des Cosinus ergeben sich alle anderen Nullstellen durch Addition von $2\pi k$, $k\in\mathbb{Z}$. Die Nullstellen des Sinus erhält man wegen $\cos(x+\frac{\pi}{2})=-\sin x$ aus den Nullstellen des Cosinus durch eine Verschiebung um $\frac{\pi}{2}$. □

KOROLLAR IV-5.27. 2π ist die kleinste positive Periode von Cosinus und Sinus.

BEWEIS. Angenommen p mit $0<p<2\pi$ wäre eine weitere Periode des Sinus. Mit Lemma IV-5.26 folgt aus $\sin 0=\sin(0+p)=0$ notwendigerweise $p=\pi$. Wegen $\sin(-\frac{\pi}{2})=-1$ und $\sin(-\frac{\pi}{2}+\pi)=1$ kann π keine Periode sein. □

THEOREM IV-5.28. *1) Die Gleichung* $e^z=1$ *hat in* $\mathbb{C}$ *nur die Lösungen* $z=2k\pi i$, $k\in\mathbb{Z}$.
2) Die Gleichung $e^z=-1$ *hat in* $\mathbb{C}$ *nur die Lösungen* $z=(2k+1)\pi i$, $k\in\mathbb{Z}$.

BEWEIS. 1) Wegen $e^{2k\pi i}=\cos 2k\pi+i\sin 2k\pi=1$ sind die Zahlen $z=2k\pi i$ Lösungen. Wir zeigen daß es keine weiteren Lösungen geben kann. Es sei $z=x+iy$. Aus Korollar IV-5.20 ergibt sich

$$|e^z|=e^x|e^{iy}|=e^x=1.$$

Daraus folgt $x=0$. Es ist also $e^{iy}=\cos y+i\sin y=1$, dh. $\cos y=1$ und $\sin y=0$. Dies hat $y=k\pi$, $k\in\mathbb{Z}$ zur Folge, ungerade $k=2\ell+1$, $\ell\in\mathbb{Z}$, sind wegen $\cos(2\ell+1)\pi=-1$ jedoch nicht zulässig. Der Beweis von 2) wird analog geführt. □

KOROLLAR IV-5.29. Cosinus und Sinus haben auch in $\mathbb{C}$ nur die Nullstellen $(k+\frac{1}{2})\pi$ bzw. $k\pi$, $k\in\mathbb{Z}$.

Wir können nun die Tangens- bzw. Cotangens-Funktion dort definieren, wo Cosinus bzw. Sinus keine Nullstellen haben.

DEFINITION IV-5.30. *Der* ***Tangens*** *bzw.* ***Cotangens*** *sind folgendermaßen definiert:*

$$\tan\colon\begin{cases}\mathbb{R}\setminus\{(k+\frac{1}{2})\pi\colon k\in\mathbb{Z}\}\to\mathbb{R}\\ x\mapsto\frac{\sin x}{\cos x}\end{cases},\quad \cot\colon\begin{cases}\mathbb{R}\setminus\{(k\pi\colon k\in\mathbb{Z}\}\to\mathbb{R}\\ x\mapsto\frac{\cos x}{\sin x}\end{cases}.$$

Wir führen einige Eigenschaften dieser beiden Funktionen an:

THEOREM IV-5.31. i) $\tan$ *und* $\cot$ *sind* π*-periodisch.*
ii) $\tan$ *ist streng monoton wachsend auf* $(-\frac{\pi}{2},\frac{\pi}{2})$,
iii) $\cot$ *ist streng monoton fallend auf* $(0,\pi)$.
iv) $\tan(x+\frac{\pi}{2})=-\cot x$ *für alle* $x\in\mathbb{R}\setminus\{(k\pi\colon k\in\mathbb{Z}\}$,
v) *es gilt* $\lim_{x\to\frac{\pi}{2}^-}\tan(x)=\infty$ *und* $\lim_{x\to-\frac{\pi}{2}^+}\tan(x)=-\infty$.

Beweis. Wir zeigen nur ii). Es sei $-\frac{\pi}{2} < x < y < \frac{\pi}{2}$, also $0 < y - x < \pi$. Die Behauptung folgt aus

$$0 < \sin(y - x) = \sin y \cos x - \cos y \sin x \Leftrightarrow \tan x < \tan y$$

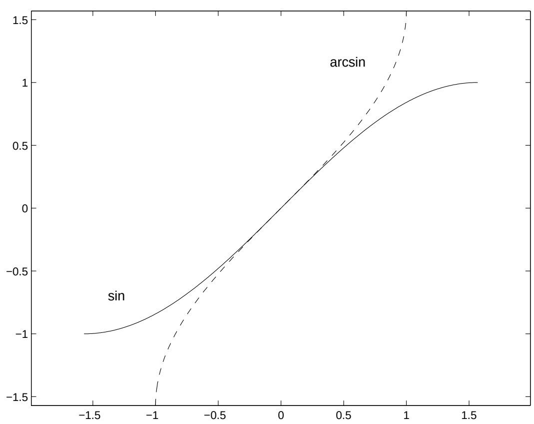
(man beachte: $\cos x > 0$, $\cos y > 0$). $\square$

Die Einschränkungen von sin auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$, cos auf $[0, \pi]$, tan auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ und cot auf $(0, \pi)$ sind streng monoton. Somit besitzen sie Umkehrfunktionen, die sogenannten **Arcus-Funktionen**:

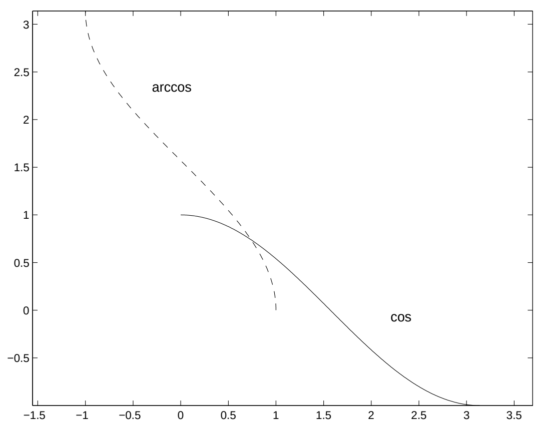
Definition IV-5.32 (Arcusfunktionen).

Funktion	*Umkehrfunktion*
$\cos\colon [0, \pi] \to [-1, 1]$	$\arccos\colon [-1, 1] \to [0, \pi]$
$\sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$	$\arcsin\colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$
$\tan\colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$	$\arctan\colon \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$
$\cot\colon (0, \pi) \to \mathbb{R}$	$\mathrm{arccot}\ \colon \mathbb{R} \to (0, \pi)$

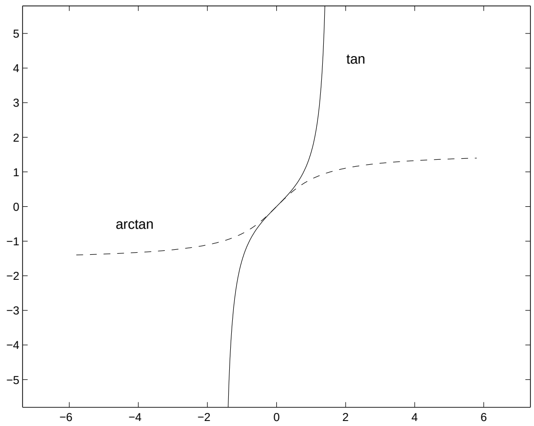
(für arccos *lies „arcus-cosinus“, etc.)*



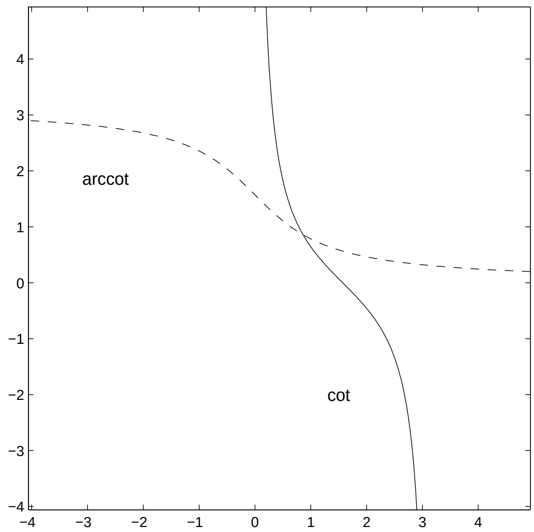
Sinus und Arcussinus



Cosinus und Arcuscosinus



Tangens und Arcustangens



Cotangens und Arcuscotangens

5.6. Hyperbelfunktionen. Bestimmte Linearkombinationen von Exponentialfunktionen treten in Anwendungen so häufig auf, daß man ihnen eigene Namen gegeben hat.

DEFINITION IV-5.33. *Die **Hyperbelfunktionen** sind folgendermaßen definiert:*

i) ***sinus hyperbolicus***, sinh

$$\sinh\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \frac{1}{2}(e^x - e^{-x}) \end{cases}$$

ii) ***cosinus hyperbolicus***, cosh

$$\cosh\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \frac{1}{2}(e^x + e^{-x}) \end{cases}$$

iii) ***tangens (cotangens) hyperbolicus***

$$\tanh\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \frac{\sinh x}{\cosh x} \end{cases} \qquad \coth\colon \begin{cases} \mathbb{R}\setminus\{0\} \to \mathbb{R} \\ x \mapsto \frac{\cosh x}{\sinh x} \end{cases}$$

Der hyperbolische Cosinus beschreibt die Form eines Seiles, welches zwischen zwei Befestigungspunkten frei durchhängt. Zeichnet man die Punktmenge $\{(\cosh t, \sinh t)\colon t \in \mathbb{R}\} \subset \mathbb{R}^2$, erhält man den rechten Zweig der Hyperbel $x^2 - y^2 = 1$. Dies erklärt den Zusatz „hyperbolicus“ in der Namensgebung. Es besteht aber auch ein enger Zusammenhang mit den trigonometrischen Funktionen:

$$\forall x \in \mathbb{R}\colon \sinh x = -i \sin ix, \quad \cosh x = \cos ix.$$

Damit lassen sich nun leicht trigonometrische Identitäten auf Hyperbelfunktionen übertragen:

THEOREM IV-5.34. *Für alle $x, y \in \mathbb{R}$ gilt*

i) $\cosh^2 x - \sinh^2 x = 1$,
ii) $\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y$,
iii) $\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$,
iv) $\sinh x - \sinh y = 2 \cosh \frac{x+y}{2} \sinh \frac{x-y}{2}$,
v) $\cosh x - \cosh y = 2 \sinh \frac{x+y}{2} \sinh \frac{x-y}{2}$.

BEWEIS. übung. $\square$

THEOREM IV-5.35. (1) *Eigenschaften des* $\sinh$*:*
i) $\sinh$ *ist ungerade,*
ii) $\sinh$ *ist streng monoton wachsend,*
iii) $\sinh 0 = 0$, $\lim_{x\to\infty} \sinh x = \infty$,
iv) $\operatorname{bild} \sinh = \mathbb{R}$.

(2) *Eigenschaften des* $\cosh$*:*
i) $\cosh$ *ist gerade,*
ii) $\cosh|_{\mathbb{R}^+}$ *ist streng monoton wachsend,*
iii) $\cosh 0 = 1$, $\lim_{x\to\infty} \cosh x = \infty$,
iv) $\operatorname{bild} \cosh = [1, \infty)$.

(3) *Eigenschaften des* $\tanh$*:*
i) $\tanh$ *ist ungerade,*
ii) $\tanh$ *ist streng monoton wachsend,*
iii) $\tanh 0 = 0$, $\lim_{x\to\infty} \tanh x = 1$, $\lim_{x\to-\infty} \tanh = -1$,
iv) $\operatorname{bild} \tanh = (-1, 1)$.

(4) *Eigenschaften des* $\coth$*:*
i) $\coth$ *ist ungerade,*
ii) $\coth|_{(-\infty,0)}$ *und* $\coth|_{(0,\infty)}$ *sind streng monoton fallend,*
iii) $\lim_{x\to\infty} \coth x = 1$, $\lim_{x\to-\infty} \coth x = -1$, $\lim_{x\to 0^+} \coth x = \infty$, $\lim_{x\to 0^-} \coth x = -\infty$,
iv) $\operatorname{bild} \coth = \mathbb{R} \setminus [-1, 1]$.

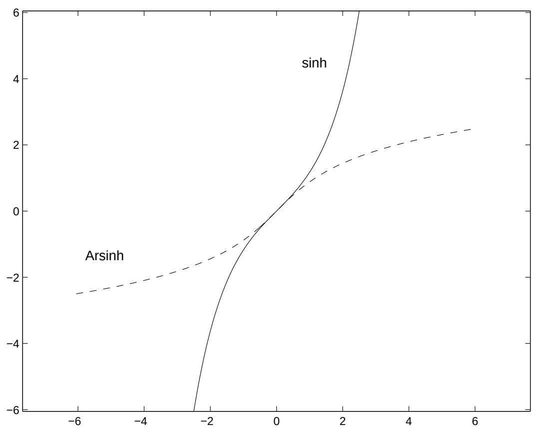
BEWEIS. übung. $\square$

Der hyperbolische Sinus, Tangens und Cotangens und die Restriktion von cosh auf $\mathbb{R}^+$ sind injektiv. Ihre Umkehrfunktionen sind die **Areafunktionen**, die wir im folgenden tabellarisch zusammenfassen. Wegen des engen Zusammenhanges der Hyperbelfunktionen mit der Exponentialfunktion, ist es nicht überraschend, daß die Areafunktionen sich durch den natürlichen Logarithmus darstellen lassen.

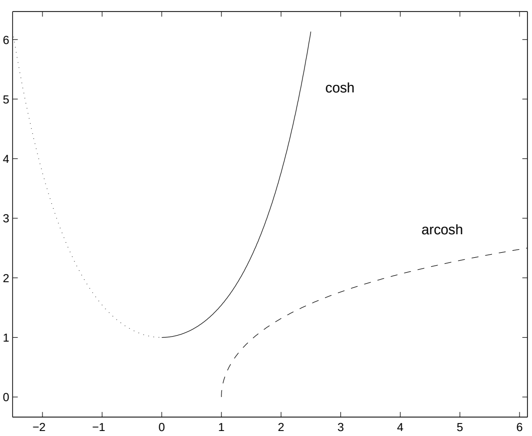
Definition IV-5.36.

Funktion	*Umkehrfunktion*
$\cosh\colon [0,\infty) \to [1,\infty)$	$\operatorname{Arcosh}\colon \begin{cases} [1,\infty) \to [0,\infty) \\ y \mapsto \ln(y+\sqrt{y^2-1}) \end{cases}$
$\sinh\colon \mathbb{R} \to \mathbb{R}$	$\operatorname{Arsinh}\colon \begin{cases} \mathbb{R} \to \mathbb{R} \\ y \mapsto \ln(y+\sqrt{y^2+1}) \end{cases}$
$\tanh\colon \mathbb{R} \to (-1,1)$	$\operatorname{Artanh}\colon \begin{cases} (-1,1) \to \mathbb{R} \\ y \mapsto \frac{1}{2}\ln\frac{1+y}{1-y} \end{cases}$
$\coth\colon \mathbb{R}\setminus\{0\} \Rightarrow \mathbb{R}\setminus[-1,1]$	$\operatorname{Arcoth}\colon \begin{cases} \mathbb{R}\setminus[-1,1] \mapsto \mathbb{R}\setminus\{0\} \\ y \mapsto \frac{1}{2}\ln\frac{1+y}{y-1} \end{cases}$

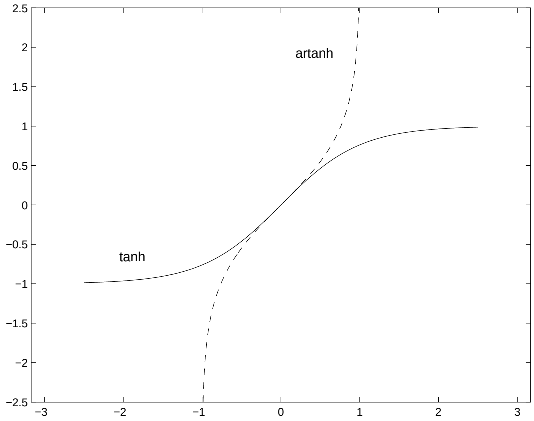
In den folgenden Plots stellen die durchgezogenen Kurven den invertierbaren Zweig der hyperbolischen Funktionen dar. Die Areafunktionen werden strichliert dargestellt:



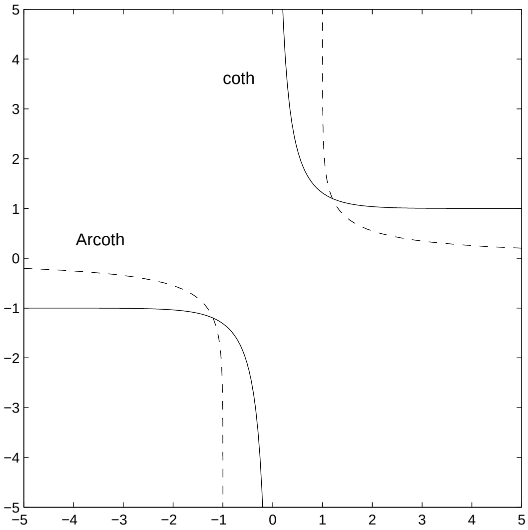
sinh und Arsinh



cosh und Arcosh



tanh und Artanh



coth und Arcoth

KAPITEL V

Differenzierbare Funktionen

1. Differenzierbarkeit

Die Stetigkeit einer Abbildung f an einer Stelle x_0 bedeutet, daß man lokal, dh. in einer Umgebung V von x_0, $f(x)$ durch $f(x_0)$ approximieren kann. Natürlich bringt dies i.a. nur eine sehr grobe Näherung an den tatsächlichen Funktionsverlauf. Bedeutend besser kann man den Graph von f durch eine affine Funktion $t_1(x) = f(x_0)+\alpha(x-x_0)$ approximieren. Die Forderung, der Approximationsfehler $r(x) = f(x) - t_1(x)$ möge in einer Umgebung von x_0 klein sein, liefert jedoch allein keine Bedingung für α. Es ist also notwendig, die Anforderungen an die Approximationsgüte zu erhöhen:

Definition V-1.1. *Es sei $f\colon D \to \mathbb{K}$, $D \subset \mathbb{K}$ und $x_0 \in D$ ein Häufungspunkt von D.*

(1) *f heißt **differenzierbar** an der Stelle $x_0 \in D \underset{Def}{\Leftrightarrow}$*
Es gibt $\alpha \in \mathbb{K}$ und $r\colon D - \{x_0\} \to \mathbb{K}$, sodaß
i) $\forall x \in D\colon f(x) = f(x_0) + \alpha(x - x_0) + r(x - x_0)$,
ii) $\lim_{x\to x_0} \frac{r(x-x_0)}{x-x_0} = 0$.
(2) *Die Zahl α heißt **Ableitung** von **f an der Stelle $\mathbf{x_0}$**.*
Schreibweise: $\alpha = f'(x_0) = \frac{df}{dx}(x_0)$.
(3) *f heißt **differenzierbar** $\underset{Def}{\Leftrightarrow}$ f ist an jeder Stelle $x_0 \in D$ differenzierbar.*
*Die Abbildung $f'\colon D \to \mathbb{K}$, $x \mapsto f'(x)$ heißt **Ableitung** von f.*

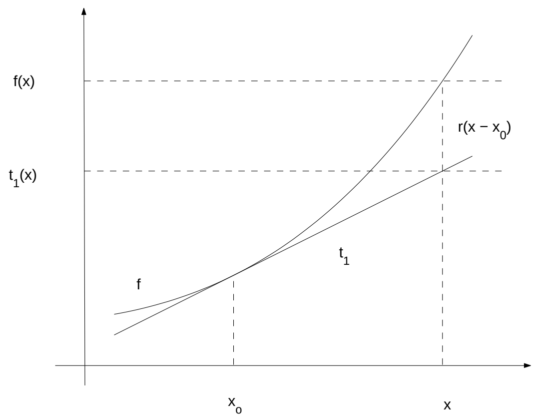
Wir zeigen zuerst, daß diese Definition sinnvoll ist, dh. daß durch die beiden Bedingungen i) und ii) α und r eindeutig festgelegt sind: Angenommen es gäbe ein weiteres Paar $(\tilde{\alpha}, \tilde{r})$, $\tilde{\alpha} \in \mathbb{K}$,$\tilde{r}\colon D - \{x_0\} \to \mathbb{K}$ mit den geforderten Eigenschaften. Für alle $x \in D$ müßte dann

$$\begin{aligned} f(x) &= f(x_0) + \alpha(x - x_0) + r(x - x_0) \\ &= f(x_0) + \tilde{\alpha}(x - x_0) + \tilde{r}(x - x_0) \end{aligned}$$

also insbesondere für $x \neq x_0$ auch

$$\alpha + \frac{r(x - x_0)}{x - x_0} = \tilde{\alpha} + \frac{\tilde{r}(x - x_0)}{x - x_0}$$

gelten. Aus der Bedingung ii) folgt dann $\alpha = \tilde{\alpha}$ und daraus $r = \tilde{r}$. Somit ist es gerechtfertigt von *der* Ableitung von f an der Stelle x_0 zu sprechen.



Durch die Bedingungen i) und ii) wird also eine eindeutige affine Approximation $t_1(x) = f(x_0) + f'(x_0)(x - x_0)$ an f bestimmt. Für $\mathbb{K} = \mathbb{R}$ ist der Graph von t_1 eine Gerade durch den Punkt $(x_0, f(x_0))$ mit Anstieg $f'(x_0)$. Man nennt sie **Tangente** an (den Graph von) f in x_0.

THEOREM V-1.2. *Es sei $f\colon D \to \mathbb{K}, D \subset \mathbb{K}$ und $x_0 \in D$ ein Häufungspunkt von D. Äquivalent sind*

i) *f ist differenzierbar in x_0,*
ii) *es existiert $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$.*

BEWEIS. „i) $\Rightarrow$ ii)“ Für $x \neq x_0$, $x \in D$, folgt mit $\alpha = f'(x_0)$

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) + \frac{r(x - x_0)}{x - x_0}.$$

Führt man den Grenzübergang $x \to x_0$ durch, ergibt sich wegen 1-ii) in Definition V-1.1

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

„ii) $\Leftarrow$ i)“ Es sei $\alpha = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$. Definieren wir

$$R(x) = \begin{cases} 0 & x = x_0, \\ \frac{f(x)-f(x_0)-\alpha(x-x_0)}{x-x_0} & x \in D \setminus \{x_0\}, \end{cases}$$

dann ist die Voraussetzung äquivalent zu

$$(*) \qquad \lim_{x\to x_0} R(x) = 0 = R(x_0),$$

dh. R ist stetig in x_0. Somit gilt auch die Gleichung

$$f(x) = f(x_0) + \alpha(x - x_0) + R(x)(x - x_0).$$

Setzt man $r(x - x_0) := R(x)(x - x_0)$, so erfüllt das Paar (α, r) die Bedingungen von Definition V-1.1, dh. f ist in x_0 differenzierbar und es gilt $\alpha = f'(x_0)$. $\square$

Der Beweis enthält eine weitere Charakterisierung der Differenzierbarkeit:

THEOREM V-1.3. *Äquivalent sind*

i) *f ist differenzierbar in x_0,*
ii) *es gibt $F\colon D \to \mathbb{K}$ mit folgenden Eigenschaften:*
  a) $\forall x \in D\colon f(x) = f(x_0) + F(x)(x - x_0)$,
  b) *F ist stetig in x_0.*

Es gilt $F(x_0) = f'(x_0)$.

BEWEIS. Man setzt $F(x) = \alpha + R(x)$, R wie im Beweis von Satz V-1.2. □

Es ist nun unmittelbar ersichtlich, daß die Differenzierbarkeit von f an einer Stelle x_0 eine stärkere Eigenschaft ist als die Stetigkeit von f an der Stelle x_0.

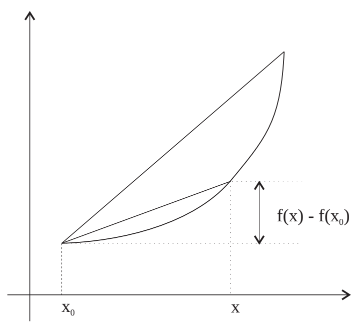
KOROLLAR V-1.4. Es sei $f\colon D \to \mathbb{K}$) und $x_0 \in D$ ein Häufungspunkt von D. Ist f differenzierbar an der Stelle $x_0 \in D$, dann ist f auch stetig in x_0.

BEWEIS. Wegen der Darstellung in Satz V-1.3,

$$f(x) = f(x_0) + F(x)(x - x_0),$$

ist f als Summe stetiger Funktionen selbst stetig in x_0. □

BEMERKUNG V-1.5. 1) Anschaulich bedeutet die Definition der Differenzierbarkeit einer Funktion f an einer Stelle x_0, daß es möglich ist, f in einer Umgebung von x_0 so durch eine affine Funktion zu approximieren, daß der Approximationsfehler $r(x - x_0)$ rascher als $x - x_0$ nach Null strebt. Der Vorteil dieses Zugangs ist, daß sich Definition V-1.1 leicht auf allgemeinere Situationen übertragen läßt (etwa auf Funktionen in mehreren Veränderlichen).
2) Den Quotient $\frac{f(x)-f(x_0)}{x-x_0}$ bezeichnet man als **Differenzenquotient**. Geometrisch gesehen stellt er den Anstieg der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$ dar. Satz V-1.2 bedeutet somit, daß Differenzierbarkeit von f in x_0 auch dadurch charakterisiert werden kann, daß die Anstiege der Sekanten einen eindeutig bestimmten Grenzwert besitzen, wenn man x gegen x_0 streben läßt. Es ist somit gerechtfertigt, diesen Grenzwert als Anstieg von f (bzw. des Graph von f) zu betrachten.



3) Ist D ein Intervall und $x_0 \in D$ ein Randpunkt des Intervalls (dies ist einer der Gründe, in Definition V-1.1 einen *Häufungspunkt* x_0 des Definitionsbereiches zu betrachten), dann sind in Definition V-1.1 und Satz V-1.2 natürlich nur *einseitige* Annäherungen an x_0 möglich. Dies führt nun zwangsläufig auf das Konzept der einseitigen Ableitungen. Dabei können die Funktionswerte von f ohne weiteres komplex seinj.

Definition V-1.6. *Es sei $f\colon D \to \mathbb{K}$, $D \subset \mathbb{R}$ und $x_0 \in D$ ein Häufungspunkt von D.*

i) *f heißt in $\mathbf{x_0} \in \mathbf{D}$* ***linksseitig (rechtsseitig) differenzierbar*** *$\underset{Def}{\Leftrightarrow} f|_{D\cap(-\infty,x_0]}$ $(f|_{D\cap[x_0,\infty)})$ ist in x_0 differenzierbar.*
ii) *Die Ableitung von $f|_{D\cap(-\infty,x_0]}$ $(f|_{D\cap[x_0,\infty)})$ heißt linksseitige Ableitung, $f'_-(x_0)$, (rechtsseitige Ableitung, $f'_+(x_0)$) von f an der Stelle x_0.*

Eine Abbildung $f\colon [a,b] \to \mathbb{R}$ nennt man differenzierbar, wenn f für alle $x \in (a,b)$ differenzierbar ist und in den Randpunkten von $[a,b]$ die entsprechenden einseitigen Ableitungen existieren.
Aus Lemma III-1.18, Korollar V-1.4 und Satz V-1.2 folgt unmittelbar

Theorem V-1.7. *Es sei $f\colon D \to \mathbb{R}$, $D \subset \mathbb{R}$, $x_0 \in D$ ein Häufungspunkt von D. Äquivalent sind*

(1) *f ist differenzierbar an der Stelle x_0,*
(2) *f ist stetig in x_0, es existieren die links- und rechtsseitige Ableitung von f an der Stelle x_0 und sind gleich: $f'_-(x_0) = f'_+(x_0)$.*

Rechnet man die Differenzierbarkeit von f an einer Stelle x_0 gemäß Definition V-1.1 oder Satz V-1.2 nach, ist es zweckmäßig, die Nachbarstelle x in der Form x_0+h zu schreiben. Somit ist entweder für die Funktion $r(h) = f(x_0+h) - f(x_0) - \alpha h$ die Existenz des Grenzwertes $\lim_{h\to 0} \frac{r(h)}{h} = 0$ nachzuweisen oder der Grenzwert des Differenzenquotienten

$$\frac{1}{h}(f(x_0+h) - f(x_0))$$

zu untersuchen:

Beispiel V-1.8. 1) $D = \mathbb{C}$, $f(x) = x^n$. Es gilt $f'(x_0) = nx_0^{n-1}$ für alle $x_0 \in D$.
a) über Definition V-1.1:
Wir versuchen, in der Differenz $f(x_0+h) - f(x_0)$ den *linearen* Anteil αh zu isolieren:

$$\begin{aligned} f(x_0+h) - f(x_0) = (x_0+h)^n - x_0^n &= \sum_{k=0}^{n} h^k x_0^{n-k}\binom{n}{k} - x_0^n \\ &= nx_0^{n-1}h + \sum_{k=2}^{n}\binom{n}{k}h^k x_0^{n-k}. \end{aligned}$$

Der kleinste Exponent von h in $r(h) := \sum\limits_{k=2}^{n}\binom{n}{k}h^k x_0^{n-k}$ ist 2, somit folgt $\lim_{h\to 0}\frac{1}{h}r(h) = 0$, dh. $f'(x_0) = nx_0^{n-1}$.
b) über Satz V-1.2: Aus

$$\frac{1}{h}(f(x_0+h) - f(x_0)) = \frac{1}{h}((x_0+h)^n - x_0^n) = \sum_{i=0}^{n-1}(x_0+h)^i x_0^{n-1-i}$$

folgt

$$\lim_{h\to 0}\frac{1}{h}(f(x_0+h)-f(x_0))=\sum_{i=0}^{n-1}\lim_{h\to 0}(x_0+h)^i x_0^{n-1-i}=\sum_{i=0}^{n-1}x_0^i x_0^{n-1-i}=nx_0^{n-1}.$$

2) $D=\mathbb{R}$, $f(x)=|x|$.
i) $x_0<0$, $f(x_0)=-x_0$. Für alle $h\in\mathbb{R}$ mit $h<|x_0|$ folgt $x_0+h<0$ und somit

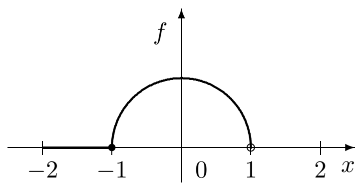
$$|x_0+h|=-x_0-h=f(x_0)-h.$$

Dies legt nahe, $r(h):=|x_0+h|+x_0+h$ zu setzen. Für $h<|x_0|$ gilt dann $r(h)=0$ und daher auch $\lim_{h\to 0}\frac{r(h)}{h}=0$. Daraus folgt $f'(x_0)=-1$.
ii) Analog zeigt man $f'(x_0)=1$ für $x_0>0$. Einfacher läßt sich dies über den Differenzenquotienten berechnen.
iii) $x_0=0$. Wir zeigen $f'_+(0)=1$ und $f'_-(0)=-1$. Somit ist f in $x_0=0$ nicht differenzierbar:

$$f'_+(0)=\lim_{h\to 0+}\frac{1}{h}(f(h)-f(0))=\lim_{h\to 0+}\frac{1}{h}|h|=1,$$
$$f'_-(0)=\lim_{h\to 0-}\frac{1}{h}(f(h)-f(0))=\lim_{h\to 0-}\frac{1}{h}|h|=-1.$$

3) $D=[-2,1]$ und

$$f(x)=\begin{cases}0 & -2\le x\le -1\\ \sqrt{1-x^2} & -1<x\le 1.\end{cases}$$



i) $x_0\in(-1,1)$, $f'(x_0)=-\frac{x_0}{\sqrt{1-x_0^2}}$. Wir wählen $h_0>0$ mit $|x_0+h_0|<1$. Für $|h|<h_0$ folgt dann

$$\begin{aligned}\frac{1}{h}(f(x_0+h)-f(x_0))&=\frac{1}{h}\left(\sqrt{1-(x_0+h)^2}-\sqrt{1-x_0^2}\right)\\&=\frac{1}{h}\frac{-2x_0h-h^2}{\sqrt{1-(x_0+h)^2}+\sqrt{1-x_0^2}}\\&=-\frac{2x_0+h}{\sqrt{1-(x_0+h)^2}+\sqrt{1-x_0^2}}.\end{aligned}$$

Daraus ergibt sich die Behauptung wegen der Stetigkeit der Wurzel.
ii) $x_0\in[-2,-1)$, $f'(x_0)=0$,
iii) $x_0=-1$, $f'_-(x_0)=0$. Die rechtsseitige Ableitung existiert nicht (in $\mathbb{R}$), denn es ist für $0<h<h_0$

$$\frac{1}{h}(f(-1+h)-f(-1))=\frac{1}{h}\sqrt{2h-h^2}=\frac{1}{\sqrt{h}}\sqrt{2-h}\underset{h\to 0+}{\longrightarrow}\infty.$$

iv) $x_0=1$, $f'_-(x_0)$ existiert nicht.

2. Rechenregeln für differenzierbare Funktionen

In diesem Abschnitt entwickeln wir einige Regeln, die es erlauben, die Ableitung komplizierterer Funktionen mit Hilfe der bekannten Ableitung einfacherer Funktionen zu berechnen.

THEOREM V-2.1. *Es seien $f,g\colon D\to\mathbb{K}$, $D\subset\mathbb{K}$, differenzierbar in einem Häufungspunkt $x_0\in D$ und $\lambda\in\mathbb{K}$. Dann sind auch $f+g$, λf, fg und, falls $f(x_0)\neq 0$*

ist, auch $\frac{1}{f}$ *an der Stelle* x_0 *differenzierbar und es gilt*

$$
\begin{array}{ll}
(f+g)'(x_0) = f'(x_0) + g'(x_0) & \\
(\lambda f)'(x_0) = \lambda f'(x_0) & \\
(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0) & \text{Produktregel} \\
(\frac{1}{f})'(x_0) = -\frac{f'(x_0)}{f^2(x_0)} & \text{Reziprokregel} \\
(\frac{g}{f})'(x_0) = \frac{g'(x_0)f(x_0) - g(x_0)f'(x_0)}{f^2(x_0)} & \text{Quotientenregel}
\end{array}
$$

Beweis. Die Differenzierbarkeit von f und g an der Stelle x_0 bedeutet nach Satz V-1.3 die Existenz in x_0 stetiger Funktionen F und G mit $F(x_0) = f'(x_0)$, $G(x_0) = g'(x_0)$ und

$$
\begin{array}{l} f(x) = f(x_0) + F(x)(x - x_0), \\ g(x) = g(x_0) + G(x)(x - x_0), \end{array} \qquad x \in D.
$$

Die Differenzierbarkeit von $f + g$ und λf in x_0 ist nun unmittelbar ersichtlich. Die Produktregel ergibt sich durch Ausmultiplizieren

$$
\begin{aligned}
(fg)(x) = f(x)g(x) = f(x_0)g(x_0) \\
&+ (F(x)g(x_0) + f(x_0)G(x) + F(x)G(x)(x - x_0))(x - x_0),
\end{aligned}
$$

und der Stetigkeit von $H(x) = F(x)g(x_0) + f(x_0)G(x) + F(x)G(x)(x - x_0)$. Nach Satz V-1.3 existiert $(fg)'(x_0)$ und ist bestimmt durch

$$
H(x_0) = F(x_0)g(x_0) + f(x_0)G(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).
$$

Wir betrachten nun

$$
\begin{aligned}
(\frac{1}{f})(x) - (\frac{1}{f})(x_0) &= \frac{1}{f(x)} - \frac{1}{f(x_0)} = -\frac{f(x) - f(x_0)}{f(x)f(x_0)} \\
&= -\frac{F(x)}{f(x)f(x_0)}(x - x_0) \equiv K(x)(x - x_0).
\end{aligned}
$$

Man beachte nun die Stetigkeit von K in x_0 und $K(x_0) = -\frac{f'(x_0)}{f(x_0)^2}$. Die Quotientenregel ergibt sich nun als Kombination der Produkt- und Reziprokregel:

$$
(\frac{g}{f})'(x_0) = g'(x_0)(\frac{1}{f})(x_0) + g(x_0)(\frac{1}{f})'(x_0).
$$

□

Insbesondere folgt, daß die Menge aller in x_0 differenzierbaren Abbildungen einen Vektorraum bildet. Besonders nützlich für die praktische Rechnung ist:

Theorem V-2.2 (Kettenregel). *Es seien mit* $f \colon D \to \mathbb{K}$, $g \colon V \to \mathbb{K}$ *und* $f(D) \subset V$. f *sei in* $x_0 \in D$ *und* g *in* $f(x_0) \in V$ *differenzierbar. Dann ist* $g \circ f \colon D \to \mathbb{K}$ *in* x_0 *differenzierbar und es gilt*

$$
(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).
$$

Beweis. Es seien F und G wie im Beweis von Satz V-2.1, dh.

$$\begin{aligned} f(x) &= f(x_0) + F(x)(x - x_0), && x \in D \\ g(y) &= g(f(x_0)) + G(y)(y - f(x_0)), && y \in V \end{aligned}$$

und $F(x_0) = f'(x_0)$, $G(f(x_0)) = g'(f(x_0))$. Somit folgt

$$\begin{aligned} (g \circ f)(x) &= g(f(x)) = g(f(x_0) + F(x)(x - x_0)) \\ &= g(f(x_0)) + G(f(x_0) + F(x)(x - x_0))F(x)(x - x_0) \\ &= g(f(x_0)) + H(x)(x - x_0). \end{aligned}$$

Wiederum ist $H(x) = G(f(x_0) + F(x)(x - x_0))F(x)$ stetig in x_0 und es ist $H(x_0) = G(f(x_0))F(x_0) = g'(f(x_0))f'(x_0)$. Die Behauptung folgt nun mit Satz V-1.3 □

Wir demonstrieren die Anwendung der Kettenregel an folgendem Beispiel, das man sich als nützliche Heuristik einprägen sollte:

Beispiel V-2.3. $f\colon D \to \mathbb{K}$, $D \subset \mathbb{K}$ sei injektiv. Ferner seien f in $x_0 \in D$ und f^{-1} in $f(x_0)$ differenzierbar. Falls $f'(x_0) \neq 0$, gilt

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Dies folgt unmittelbar durch Anwenden der Kettenregel auf die Identität $f^{-1} \circ f = \mathrm{id}_D$:

$$(f^{-1})'(f(x_0))f'(x_0) = 1.$$

Die Schwachstelle bei dieser Argumentation ist natürlich, daß die Differenzierbarkeit der Umkehrfunktion bereits gesichert sein muß. Oft weiß man zu wenig über die Umkehrfunktion, um einen direkten Nachweis der Differenzierbarkeit zu führen (man denke etwa an die Arcus Funktionen). Für die Umkehrfunktion folgt allerdings bereits aus der Stetigkeit von f^{-1} deren Differenzierbarkeit!

Theorem V-2.4. *Es sei I ein Intervall und $f\colon I \to \mathbb{R}$ streng monoton. Ist f in $x_0 \in I$ differenzierbar mit $f'(x_0) \neq 0$, dann ist f^{-1} in $y_0 = f(x_0)$ differenzierbar und es gilt*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)} \quad \textit{bzw.} \quad (f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Beweis. Nach Satz V-1.3 gibt es eine in x_0 stetige Abbildung $F\colon I \to \mathbb{R}$ mit

$$f(x) - f(x_0) = F(x)(x - x_0).$$

Wegen $y = f(x)$, also $x = f^{-1}(y)$, läßt sich dies schreiben in der Form

$$y - y_0 = F(f^{-1}(y))(f^{-1}(y) - f^{-1}(y_0)).$$

Da I ein Intervall ist, ist f^{-1} stetig, Satz IV-2.8. Auf Grund der Voraussetzung ist auch $F \circ f^{-1}$ stetig in y_0 und wegen $(F \circ f^{-1})(y_0) = F(x_0) = f'(x_0) \neq 0$ ist auch $\frac{1}{F \circ f^{-1}}$ stetig in y_0. Aus

$$f^{-1}(y) = f^{-1}(y_0) + \frac{1}{(F \circ f^{-1})(y)}(y - y_0)$$

folgt nun mit Satz V-1.3 die Differenzierbarkeit von f^{-1} in y_0 und

$$(f^{-1})'(y_0) = \frac{1}{(F \circ f^{-1})(y_0)} = \frac{1}{f'(x_0)}.$$

□

Bemerkung V-2.5. An Stellen $y_0 = f(x_0)$ mit $f'(x_0)$ ist die Umkehrfunktion f^{-1} nicht differenzierbar. Wäre f^{-1} in $f(x_0)$ differenzierbar, müßte nach Beispiel V-2.3

$$0 = (f^{-1})'(f(x_0))f'(x_0) = 1$$

gelten.

Wir demonstrieren die Anwendung dieser Regeln bei der Berechnung der Ableitung der elementaren Funktionen:

3. Ableitung der elementaren Funktionen

Potenzfunktion: $p_n(x) = x^n$, $n \in \mathbb{N}$. Wir haben die Differenzierbarkeit von p_n zwar bereits in Beispiel V-1.8 untersucht, können nun aber einen einfacheren Beweis führen, der nur die Differenzierbarkeit von p_1 verwendet. Es sei $x_0 \in \mathbb{R}$ beliebig. Wir behaupten: $p_n'(x_0) = nx_0^{n-1}$, $n \in \mathbb{N}$, (wir vereinbaren hier, $0^0 = 1$ zu setzen) und führen einen Induktionsbeweis. Es gilt $p_1'(x_0) = 1$. Als Induktionsvoraussetzung verwenden wir: p_n ist differenzierbar in x_0 und $p_n'(x_0) = nx_0^{n-1}$. Wegen $p_{n+1} = p_n p_1$ ist dann nach Satz V-2.1 auch p_{n+1} an der Stelle x_0 differenzierbar und es gilt:

$$\begin{aligned} p_{n+1}'(x_0) &= p_n'(x_0)p_1(x_0) + p_n(x_0)p_1'(x_0) \\ &= nx_0^{n-1}x_0 + x_0^n = (n+1)x_0^n. \end{aligned}$$

Aus der Reziprokregel folgt nun für $x_0 \neq 0$:

$$\frac{d}{dx}(\frac{1}{p_n})(x_0) = -\frac{p_n'(x_0)}{(p_n(x_0))^2} = -\frac{nx_0^{n-1}}{x_0^{2n}} = -nx_0^{-n-1}.$$

Wurzelfunktion: $w_n(x) = \sqrt[n]{x}$, $x \geq 0$, $n \in \mathbb{N}$. Wir können die Regel für die Differentiation der Umkehrfunktion verwenden: diese garantiert einerseits die Differenzierbarkeit von w_n für alle $x_0 > 0$, andererseits gilt

$$\begin{aligned} w_n'(x_0) &= \frac{1}{p_n'(w_n(x_0))} = \frac{1}{n(w_n(x_0))^{n-1}} \\ &= \frac{1}{n}\frac{1}{x_0^{\frac{n-1}{n}}} = \frac{1}{n}x_0^{\frac{1}{n}-1}. \end{aligned}$$

Exponentialfunktion: $\exp(z)$, $z \in \mathbb{C}$. Dies bedarf etwas mehr Aufwand. Wir zeigen zunächst einen nützlichen Grenzwert:

Lemma V-3.1.

$$\lim_{z \to 0} \frac{1}{z}(e^z - 1) = 1.$$

Beweis. Aus der Potenzreihe für e^z folgt

$$e^z - 1 = z\sum_{k=1}^{\infty}\frac{z^{k-1}}{k!} =: zf(z).$$

Der Konvergenzradius von f ist ebenfalls ∞. Somit ist f stetig auf $\mathbb{C}$ und es gilt

$$\lim_{z\to 0}\frac{1}{z}(e^z - 1) = \lim_{z\to 0} f(z) = f(0) = 1.$$

□

Eine überraschende Folgerung ergibt sich, wenn wir $z = ix$ setzen. Wir erinnern daran, daß eine Folge $(z_n) \subset \mathbb{C}$ genau dann konvergent ist, wenn $(\mathrm{Re}\ z_n) \subset \mathbb{R}$, $(\mathrm{Im}\ z_n) \subset \mathbb{R}$ konvergieren: Wegen $e^{ix} = \cos x + i\sin x$ ergeben sich durch Vergleich der Real- und Imaginärteile die nützlichen reellen Grenzwerte:

Korollar V-3.2.

$$\lim_{x\to 0}\frac{\sin x}{x} = 1, \qquad \lim_{x\to 0}\frac{\cos x - 1}{x} = 0.$$

Es sei nun $z_0 \in \mathbb{C}$ fest und $h \in \mathbb{C}$. Aus

$$\frac{1}{h}(e^{z_0+h} - e^{z_0}) = \frac{1}{h}e^{z_0}(e^h - 1)$$

ergibt sich mit Lemma V-3.1

$$\frac{de^z}{dz}\Big|_{z=z_0} = e^{z_0}.$$

Wir benützen die Schreibweise: $f'(z)|_{z=z_0} := f'(z_0)$. Dies hilft Mehrdeutigkeiten zu vermeiden bei Funktionen, bei denen man traditionell Funktion und Funktionswert an einer Stelle x gleich anschreibt (etwa die Funktion e^x, $\sin x$, ...).
Die Exponentialfunktion hat also die erstaunliche Eigenschaft

$$\forall z \in \mathbb{C}\colon \frac{d}{dz}e^z = e^z.$$

Wir werden später nachweisen, daß die Abbildungen $z \to \lambda e^z$, $\lambda \in \mathbb{C}$, die einzigen Funktionen mit dieser Eigenschaft sind.
Logarithmus: $x \mapsto \ln x$, $x > 0$. Die Differenzierbarkeit des Logarithmus folgt wieder aus Satz V-2.4. Weiters gilt

$$\frac{d}{dx}\ln x\Big|_{x=x_0} = \frac{1}{\left(\frac{d}{dx}e^x\right)\Big|_{x=\ln x_0}} = \frac{1}{e^{\ln x_0}} = \frac{1}{x_0}$$

Trigonometrische Funktionen: Die Ableitung der trigonometrischen Funktionen an einer Stelle $x_0 \in \mathbb{R}$ ergeben sich aus der Ableitung der Exponentialfunktion: Wegen Satz IV-5.18 gilt die Darstellung

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$$

und somit folgt mit Hilfe der Kettenregel

$$\begin{aligned}\frac{d}{dx}\sin x\Big|_{x=x_0} &= \frac{1}{2i}\left(\frac{d}{dz}e^z\Big|_{z=ix_0}\frac{d(ix)}{dx}\Big|_{x=x_0} - \frac{d}{dz}e^z\Big|_{z=-ix_0}\frac{d(-ix)}{dx}\Big|_{x=x_0}\right)\\ &= \frac{1}{2i}(e^{ix_0}i - e^{-ix_0}(-i)) = \frac{1}{2}(e^{ix_0} + e^{-ix_0}) = \cos x_0.\end{aligned}$$

Die Rechnung für den Kosinus verläuft analog. Man kann aber auch von der Identität

$$\cos x = \sin(x + \frac{\pi}{2}), \quad x \in \mathbb{R},$$

ausgehen und erhält wieder mit Hilfe der Kettenregel

$$\begin{aligned}\frac{d}{dx}\cos x\Big|_{x=x_0} &= \frac{d}{dx}\sin(x+\frac{\pi}{2})\Big|_{x=x_0} = \frac{d}{dx}\sin x\Big|_{x=x_0+\frac{\pi}{2}}\frac{d(x+\frac{\pi}{2})}{dx}\Big|_{x=x_0}\\ &= \cos(x_0 + \frac{\pi}{2}) = -\sin x_0.\end{aligned}$$

Ein alternativer Beweis verwendet die Grenzwerte in Korollar V-3.2 Die Ableitung des Tangens und Kotangens ergibt sich aus der Quotientenregel

$$\begin{aligned}\frac{d}{dx}\tan x\Big|_{x=x_0} &= \frac{d}{dx}\frac{\sin x}{\cos x}\Big|_{x=x_0} = \frac{\frac{d}{dx}\sin x\Big|_{x=x_0}\cos x_0 - \sin x_0\frac{d}{dx}\cos x\Big|_{x=x_0}}{\cos^2 x_0}\\ &= \frac{\sin^2 x_0 + \cos^2 x_0}{\cos^2 x_0} = \frac{1}{\cos^2 x_0}, \quad x_0 \in (-\frac{\pi}{2}, \frac{\pi}{2}),\end{aligned}$$

analog

$$\frac{d}{dx}\cot x|_{x=x_0} = -\frac{1}{\sin^2 x_0}, \quad x_0 \in (0, \pi).$$

Arcus Funktionen: Die Differenzierbarkeit der Arcusfunktionen als Umkehrfunktionen der (Einschränkungen von) trigonometrischen Funktionen folgt aus Satz V-2.4. Weiters ergibt sich

$$\begin{aligned}\frac{d}{dx}\arcsin x\Big|_{x=x_0} &= \frac{1}{\frac{d}{dx}\sin x\Big|_{x=\arcsin x_0}} = \frac{1}{\cos(\arcsin x_0)}\\ &= \frac{1}{\sqrt{1-\sin^2(\arcsin x_0)}} = \frac{1}{\sqrt{1-x_0^2}}, \quad x_0 \in (-1, 1).\end{aligned}$$

Im vorletzten Schritt ist das positive Vorzeichen bei der Wurzel zu wählen, da $\arcsin x_0 \in (-\frac{\pi}{2}, \frac{\pi}{2})$ und somit $\cos(\arcsin x_0) > 0$ ist. Auf gleiche Weise folgt

$$\frac{d}{dx}\arccos x\Big|_{x=x_0} = -\frac{1}{\sqrt{1-x_0^2}}, \quad x_0 \in (-1, 1).$$

arcsin und arccos sind in $x = \pm 1$ wegen Bemerkung V-2.5 nicht differenzierbar. Ähnlich findet man für $x \in \mathbb{R}$

$$\begin{aligned}\frac{d}{dx} \arctan x\Big|_{x=x_0} &= \frac{1}{\frac{d}{dx} \tan x\big|_{x=\arctan x_0}} = \cos^2(\arctan x_0) \\ &= \frac{1}{1+\tan^2(\arctan x_0)} = \frac{1}{1+x_0^2},\end{aligned}$$

und

$$\frac{d}{dx} \operatorname{arccot} x\Big|_{x=x_0} = -\frac{1}{1+x_0^2}.$$

Allgemeine Potenzfunktion: $p_\rho(x) = x^\rho$, $x > 0$, $\rho \neq 0$. Ausgehend von der Identität

$$p_\rho(x) = \exp(\rho \ln x),$$

findet man mit Hilfe der Kettenregel ($x_0 > 0$)

$$\begin{aligned}p'_\rho(x_0) &= \frac{d}{dx} e^x\Big|_{x=\rho \ln x_0} \frac{d}{dx}(\rho \ln x)|_{x=x_0} \\ &= e^{\rho \ln x_0} \rho \frac{1}{x_0} = \rho x_0^{\rho-1}.\end{aligned}$$

Exponentialfunktion zur Basis a: $x \mapsto a^x$. Ausgehend von der Identität

$$a^x = \exp(x \ln a)$$

findet man für $x_0 \in \mathbb{R}$ mit Hilfe der Kettenregel

$$\begin{aligned}\frac{d}{dx} a^x\Big|_{x=x_0} &= \frac{d}{dx} e^x\Big|_{x=x_0 \ln a} \frac{d}{dx}(x \ln a)\Big|_{x=x_0} \\ &= (\ln a) e^{x_0 \ln a} = (\ln a) a^{x_0}.\end{aligned}$$

Wir fassen die Ergebnisse tabellarisch zusammen:

Funktion $f(x)$		Ableitung $f'(x)$
x^n,	$x \in \mathbb{R}, n \in \mathbb{N}$	nx^{n-1}
x^n,	$x \neq 0, n \in \mathbb{Z}$	nx^{n-1}
x^ρ,	$x > 0, \rho \in \mathbb{R}$	$\rho x^{\rho-1}$
e^x,	$x \in \mathbb{R}$	e^x
a^x,	$x \in \mathbb{R}, a > 0$	$(\ln a)a^x$
$\sin x$,	$x \in \mathbb{R}$,	$\cos x$
$\cos x$,	$x \in \mathbb{R}$,	$-\sin x$
$\tan x$,	$x \in \mathbb{R}$,	$\frac{1}{\cos^2 x} = 1 + \tan^2 x$
$\cot x$,	$x \in \mathbb{R}$,	$-\frac{1}{\sin^2 x} = -1 - \cot^2 x$
$\arcsin x$,	$x \in (-1, 1)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos x$,	$x \in (-1, 1)$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan x$,	$x \in \mathbb{R}$	$\frac{1}{1+x^2}$
$\operatorname{arccot} x$,	$x \in \mathbb{R}$	$-\frac{1}{1+x^2}$

Wir betrachten nun eine Abbildung $f\colon D \to \mathbb{K}$, welche in jedem Punkt $x \in D$ differenzierbar ist. Dann existiert die Abbildung $f'\colon x \mapsto f'(x)$. Ist f' selbst wieder an einer Stelle $x_0 \in D$ differenzierbar, nennt man die Ableitung von f' die **2. Ableitung** von f an der Stelle x_0 und bezeichnet sie mit $f''(x_0)$ oder auch $\frac{d^2 f}{dx^2}(x_0)$. Es gilt also $f''(x_0) = (f')'(x_0)$. Es sind auch die Bezeichnungen $f^{(0)} = f$, $f^{(1)} = f'$ und $f^{(2)} = f'', \ldots$ usw. üblich. Allgemein definiert man nun höhere Ableitungen rekursiv. Wenn für $n \in \mathbb{N}$ die $(n-1)$-te Ableitung $f^{(n-1)}$ existiert und in $x_0 \in D$ differenzierbar ist, setzt man

$$f^{(n)}(x_0) = (f^{(n-1)})'(x_0).$$

Man nennt $f^{(n)}(x_0)$ auch Ableitung n-ter Ordnung an der Stelle x_0. Besitzt f an jeder Stelle $x \in D$ Ableitungen n-ter Ordnung, heißt f n-mal differenzierbar. Existieren an einer Stelle x_0 sämtliche Ableitungen $f^{(n)}(x_0)$ für alle $n \in \mathbb{N}$, sagt man f ist **unendlich oft differenzierbar in $\mathbf{x_0}$**. Ist f an jeder Stelle $x \in D$ unendlich oft differenzierbar, nennt man f **unendlich oft differenzierbar**, kurz, f ist eine **$\mathbf{C^\infty}$-Funktion**. Wir können nun folgende Unterräume von $C(D, \mathbb{K})$ betrachten:

Definition V-3.3.

$$\begin{aligned}
C^0(D, \mathbb{K}) &= C(D, \mathbb{K}) \\
C^n(D, \mathbb{K}) &= \{f \in C(D, \mathbb{K})\colon f \textit{ ist n-mal differenzierbar, } f^{(n)} \in C(D, \mathbb{K})\ \} \\
C^\infty(D, \mathbb{K}) &= \{f \in C(D, \mathbb{K})\colon f \textit{ ist unendlich oft differenzierbar}\}
\end{aligned}$$

Die folgende Kette von Inklusionen ergibt sich unmittelbar aus der Definition:

$$C^\infty(D,\mathbb{K}) \subsetneq \cdots \subsetneq C^{n+1}(D,\mathbb{K}) \subsetneq C^n(D,\mathbb{K}) \subsetneq C(D,\mathbb{K}).$$

Polynome, trigonometrische Funktionen, Exponentialfunktionen sind Beispiele für C^∞-Funktionen. Als weiteres Beispiel betrachten wir $f\colon x \mapsto |x|^3$, $x \in \mathbb{R}$. Wir überlassen es dem Leser, folgende Fakten zu verifizieren:

$$f'(x) = \begin{cases} 3x^2 & x \geq 0 \\ -3x^2 & x < 0 \end{cases} \qquad\qquad f''(x) = \begin{cases} 6x & x \geq 0 \\ -6x & x < 0 \end{cases}$$

$$f'''(x) = \begin{cases} 6 & x > 0 \\ -6 & x < 0 \end{cases} \qquad\qquad f'''(0) \text{ existiert nicht.}$$

Es gilt also $f \in C^2(\mathbb{R}) \setminus C^3(\mathbb{R})$.

Beispiel V-3.4. Wir betrachten die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Mit Hilfe der Regeln von Satz V-2.1, V-2.2 berechnet man leicht

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x} \qquad \text{für } x \neq 0.$$

Man beachte: $\nexists \lim_{x\to 0} f'(x)$. Existierte $\lim_{x\to 0} f'(x)$, wäre f automatisch in $x_0 = 0$ differenzierbar (Beweis mit Mittelwertsatz, Übung) und $f \in C^1(\mathbb{R},\mathbb{R})$. Für den Nachweis der Differenzierbarkeit von f in $x_0 = 0$ gehen wir daher auf die Definition zurück und betrachten

$$\frac{1}{h}(f(h) - f(0)) = \frac{1}{h} h^2 \sin \frac{1}{h} = h \sin \frac{1}{h}.$$

Wegen $|\sin x| \leq 1$, $x \in \mathbb{R}$, folgt daraus $f'(0) = 0$.

Dieses Beispiel zeigt, daß es Funktionen gibt, die zwar überall differenzierbar, aber nicht stetig differenzierbar sind.
Auch für höhere Ableitungen gelten die üblichen Rechenregeln:

Lemma V-3.5. *Es seien $f, g\colon D \to \mathbb{K}$, $D \subset \mathbb{K}$, n-mal differenzierbar in $x_0 \in D$ und $\lambda \in \mathbb{K}$. Dann sind auch λf, $f + g$ und fg n-mal differenzierbar und es gilt*

i) $(\lambda f)^{(n)}(x_0) = \lambda f^{(n)}(x_0)$
ii) $(f+g)^{(n)}(x_0) = f^{(n)}(x_0) + g^{(n)}(x_0)$
iii) $(fg)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x_0) g^{(k)}(x_0)$ (Leibniz Regel) .

Beweis. Den Nachweis von i) und ii) überlassen wir dem Leser als Übung und führen einen Induktionsbeweis zu iii). Für $n = 1$ ergibt sich die Produktregel aus

Satz V-2.1. Wir nehmen nun an, die Behauptung iii) sei für ein $n-1 \in \mathbb{N}$ erfüllt. Aus der rekursiven Definition der n-ten Ableitung an der Stelle x_0 ergibt sich (Übung)

$$(*) \qquad f^{(n)}(x_0) = (f')^{(n-1)}(x_0) \text{ und } g^{(n)}(x_0) = (g')^{(n-1)}(x_0),$$

d.h. die Ableitungen f' und g', die zumindest in einer Umgebung von x_0 existieren müssen, sind an der Stelle x_0 $(n-1)$-mal differenzierbar. Wegen der Induktionsvoraussetzung besitzen dann auch $f'g$ und $g'f$ Ableitungen $(n-1)$-ter Ordnung an der Stelle x_0 und es gilt nach ii)

$$\begin{aligned}(f'g)^{(n-1)}(x_0) + (g'f)^{(n-1)}(x_0) &= (f'g + g'f)^{(n-1)}(x_0) \\ &= ((fg)')^{(n-1)}(x_0) \overset{(*)}{=} (fg)^{(n)}(x_0),\end{aligned}$$

d.h. fg ist n-mal differenzierbar an der Stelle x_0. Wegen der Induktionsvoraussetzung gilt weiter

$$\begin{aligned}(fg)^{(n)}(x_0) &= (f'g)^{(n-1)}(x_0) + (g'f)^{(n-1)}(x_0) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} (f')^{(n-1-k)}(x_0) g^{(k)}(x_0) + \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(n-1-k)}(x_0)(g')^{(k)}(x_0) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(n-k)}(x_0) g^{(k)}(x_0) + \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(n-1-k)}(x_0) g^{(k+1)}(x_0) \\ &= \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x_0) g^{(k)}(x_0).\end{aligned}$$

Das letzte Gleichheitszeichen kann genauso wie im Beweis des binomischen Lehrsatzes I-4.22 gerechtfertigt werden. □

Lemma V-3.6. *Es seien $D, V \subset \mathbb{K}$, $f\colon D \to \mathbb{K}$, $g\colon V \to \mathbb{K}$ und $f(D) \subset V$. f sei in $x_0 \in D$ und g sei in $f(x_0) \in V$ n-mal differenzierbar. Dann ist auch $g \circ f$ an der Stelle x_0 n-mal differenzierbar.*

Beweis. Wir führen wieder einen Induktionsbeweis: Für $n = 1$ folgt die Behauptung aus der Kettenregel. Wir nehmen nun an, die Behauptung sei richtig für ein $n-1 \in \mathbb{N}$. Es gilt (in einer Umgebung von x_0):

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

g' und f sind in x_0 $n-1$ mal differenzierbar, wegen der Induktionsvoraussetzung besitzt auch $g' \circ f$ Ableitungen $(n-1)$-ter Ordnung in x_0. Nach Lemma V-3.5 ist auch $(g' \circ f)f'$ und somit auch $(g \circ f)'$ $(n-1)$-mal differenzierbar an der Stelle x_0. □

Die bisherigen Beispiele könnten den Eindruck vermitteln, daß stetige Funktionen an allen Stellen differenzierbar sind, ausgenommen an vereinzelten Stellen, in denen „Ecken" oder „Spitzen" auftreten. Dies ist nicht der Fall. Wir beenden diesen Abschnitt mit einem Beispiel einer stetigen, aber *nirgends* differenzierbaren Funktion. Dies zeigt auch $C^1(I, \mathbb{R}) \subsetneq C(I, \mathbb{R})$.

Beispiel V-3.7. Wir definieren die Abbildung φ auf $[-1,1]$ durch $x \mapsto |x|$ und setzen φ periodisch auf $\mathbb{R}$ fort, d.h. wir setzen

$$\forall x \in \mathbb{R}\colon \varphi(x+2) = \varphi(x).$$

Insbesondere gilt also

$$\forall x \in \mathbb{R}\, \forall \ell \in \mathbb{Z}\colon \varphi(x+2\ell) = \varphi(x).$$

φ ist Lipschitz stetig: Ist $|x-y| \geq 2$, so gilt $|\varphi(x)-\varphi(y)| \leq |\varphi(x)| + |\varphi(y)| \leq 2 \leq |x-y|$. Ist $|x-y| < 2$, existiert eine eindeutig bestimmte Zahl $l \in \mathbb{Z}$, sodaß entweder $x-2l$, $y-2l \in [-1,1]$ oder $x-2l$, $y-2l \in [0,2]$. Im ersten Fall gilt

$$\begin{aligned} |\varphi(x)-\varphi(y)| &= |\varphi(x-2l) - \varphi(y-2l)| = \big||x-2l| - |y-2l|\big| \\ &\leq \big|(x-l2) - (y-2l)\big| = |x-y|, \end{aligned}$$

während im zweiten Fall die Abschätzung

$$\begin{aligned} |\varphi(x)-\varphi(y)| &= |\varphi(x-2l) - \varphi(y-2l)| = \big|1 - |x-2l-1| - (1 - |y-2l-1|)\big| \\ &= \big||x-2l-1| - |y-2l-1|\big| \leq |x-y| \end{aligned}$$

die Lipschitz-Stetigkeit von φ zeigt.
Für $n \in \mathbb{N}_0$ definieren wir $f_n\colon \mathbb{R} \to \mathbb{R}$ durch $f_n(x) = (\frac{3}{4})^n \varphi(4^n x)$ und setzen

$$f = \sum_{n=0}^{\infty} f_n.$$

Wegen $0 \leq \varphi(x) \leq 1$, $x \in \mathbb{R}$, ist die Reihe gleichmäßig konvergent auf $\mathbb{R}$ und da jedes f_n auf $\mathbb{R}$ stetig ist, folgt mit Satz V-IV-3.10 die Stetigkeit von f auf $\mathbb{R}$. Wir zeigen nun, daß f in keinem Punkt eine endliche Ableitung besitzt. Es sei also $x_0 \in \mathbb{R}$ beliebig gewählt. Für jedes $m \in \mathbb{N}$ setzen wir

$$h_m = \begin{cases} \frac{1}{2}4^{-m} & \text{falls } 0 \leq 4^m x_0 - [4^m x_0] < \frac{1}{2} \\ -\frac{1}{2}4^{-m} & \text{falls } \frac{1}{2} \leq 4^m x_0 - [4^m x_0] < 1. \end{cases}$$

Für $n > m$ ist $4^n h_m = \pm\frac{1}{2}4^{n-m}$ eine gerade ganze Zahl; somit folgt

$$f_n(x_0+h_m) - f_n(x_0) = (\frac{3}{4})^n(\varphi(4^n x_0 + 4^n h_m) - \varphi(4^n x_0)) = 0, \quad n > m.$$

Wegen der Wahl von h_m liegt keine ganze Zahl zwischen $4^m x_0$ und $4^m(x_0+h_m)$ und somit gilt $\varphi(4^m x_0 + 4^m h_m) - \varphi(4^m x_0) = 4^m h_m$. Man erhält daher für $n = m$

$$|f_m(x_0+h_m) - f_m(x_0)| = (\frac{3}{4})^m |\varphi(4^m x_0 + 4^m h_m) - \varphi(4^m x_0)| = 3^m |h_m|,$$

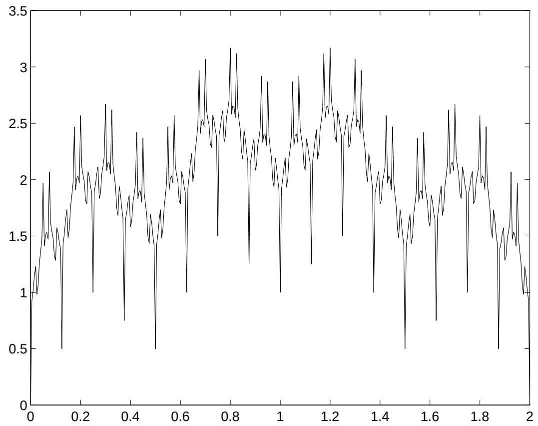
für $0 \leq n \leq m-1$ ergibt sich mit $(*)$ die Abschätzung

$$|f_n(x_0+h_m) - f_n(x_0)| \leq 3^n \cdot |h_m|.$$

Der Differenzenquotient von f kann nun folgendermaßen abgeschätzt werden

$$\begin{aligned} &|\frac{1}{h_m}(f(x_0+h_m) - f(x_0)| = |\frac{1}{h_m}\sum_{n=0}^{m}(f_n(x_0+h_m) - f_n(x_0))| \\ &\geq |\frac{1}{h_m}(f_m(x_0+h_m) - f_m(x_0))| - |\frac{1}{h_m}\sum_{n=0}^{m-1}(f_n(x_0+h_m) - f_n(x_0))| \\ &\geq 3^m - \sum_{n=0}^{m-1} 3^n = \frac{1}{2}(3^m+1). \end{aligned}$$

Da (h_m) eine Nullfolge ist, kann f an der Stelle x_0 nicht differenzierbar sein.



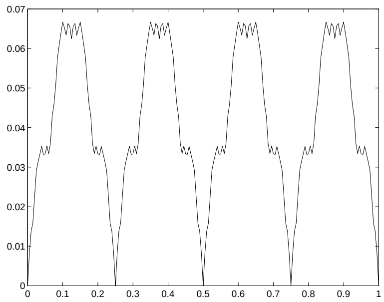
Es ist verblüffend, daß eine bereits geringfügige Modifikation dieses Beispieles dazu führt, daß die Oszillationen von f makroskopisch nicht mehr sichtbar sind. Dazu ersetzt man φ durch

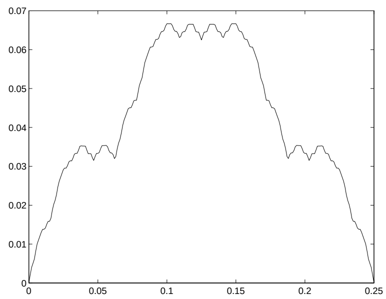
$$\tilde{\varphi}(x) = |x|, \quad x \in [-\frac{1}{2}, \frac{1}{2}],$$

setzt $\tilde{\varphi}$ periodisch auf $\mathbb{R}$ fort und definiert $\tilde{f}_n$ durch

$$\tilde{f}_n(x) = \tilde{\varphi}(4^n x), \quad x \in \mathbb{R}, n \in \mathbb{N}_0.$$

Ähnlich wie vorher kann man dann zeigen, daß $\tilde{f} = \sum_{n=0}^{\infty} \tilde{f}_n$ ebenfalls nirgends differenzierbar ist. Allerdings läßt sich dies nicht mehr aus dem nun glatt erscheinenden Graph von $\tilde{f}$ ablesen (Tagaki Funktion).





4. Lokale Extrema, Mittelwertsatz

Nach dem Satz von Weierstraß (Korollar III-4.2) nimmt eine stetige Funktion f auf einem abgeschlossenen und beschränkten Intervall I Maximum und Minimum an. Im Fall des Maximums bedeutet dies: $\exists x_0 \in I \, \forall x \in I \colon f(x) \leq f(x_0)$. Dieser Satz bringt eine *globale* Eigenschaft stetiger Funktionen zum Ausdruck, da sämtliche Funktionswerte zum Vergleich zugelassen sind. Er gibt aber keinerlei Hinweis darauf, *wo* ein globales Extremum liegt. Bei differenzierbaren Funktionen ist es jedoch möglich, aus dem *lokalen* Verhalten der Funktion, d.h. dem Verhalten in einer *Umgebung* einer Stelle x_0, auf das Vorliegen eines Extremums (relativ zur Umgebung) in x_0 zu schließen.

Definition V-4.1. *Es sei $f \colon D \to \mathbb{R}$, $D \subset \mathbb{K}$ und $x_0 \in D$.*

i) *f besitzt in x_0 ein **lokales Maximum** (**Minimum**) $\underset{Def}{\Leftrightarrow}$ es gibt eine Umgebung U von x_0 mit der Eigenschaft*

$$\forall x \in U \cap D\colon f(x) \leq f(x_0) \quad (f(x) \geq f(x_0)).$$

ii) *f besitzt in $x_0 \in D$ ein **lokales Extremum** $\underset{Def}{\Leftrightarrow}$ f besitzt in x_0 ein lokales Maximum oder lokales Minimum.*

Theorem V-4.2. *Es sei I ein* offenes *Intervall, $f\colon I \to \mathbb{R}$ und $x_0 \in I$. Ist f differenzierbar in x_0 und besitzt f in x_0 ein lokales Extremum, dann gilt $f'(x_0) = 0$.*

Beweis. O.B.d.A. besitze f in x_0 ein lokales Maximum (anderenfalls betrachte man $-f$). Nach Definition V-4.1 gibt es ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subset I$ und

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geq 0 & \text{für } x_0 - \delta < x < x_0 \\ \leq 0 & \text{für } x_0 < x < x_0 + \delta. \end{cases}$$

Führt man den Grenzübergang $x \to x_0$ durch, ergibt sich $f'(x_0) = 0$. □

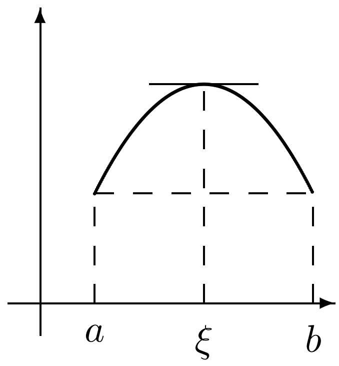
Bemerkung V-4.3. (1) Die notwendige Bedingung für das Auftreten eines lokalen Extremums gilt nur in *inneren* Punkten des Definitionsbereiches von f (wir haben dies erzwungen, indem wir als Definitionsbereich ein offenes Intervall wählten). Der Satz gilt nicht in den Randpunkten von I: Als Beispiel betrachte man $f = \mathrm{id}\,|_{[0,1]}$. f besitzt ein lokales Minimum in $x = 0$ und ein lokales Maximum in $x = 1$, aber es ist $f'(0) = f'(1) \neq 0$.
(2) Die Bedingung $f'(x_0) = 0$ ist nicht hinreichend: Für $f(x) = x^3$, $x \in [-1, 1]$ gilt $f'(0) = 0$, aber f besitzt in $x = 0$ kein lokales Extremum.
(3) Es gibt auch innere lokale Extrema, die durch Satz V-4.2 nicht erfaßt werden: $f(x) = |x|$, $x \in [-1, 1]$, besitzt in $x = 0$ ein lokales Minimum und ist an dieser Stelle nicht differenzierbar.
(4) Als Kandidaten für lokale Extremstellen einer Abbildung $f\colon [a, b] \to \mathbb{R}$ kommen also in Frage

a) die Randpunkte a und b,
b) die Nullstellen von f',
c) jene Stellen, in denen f nicht differenzierbar ist.

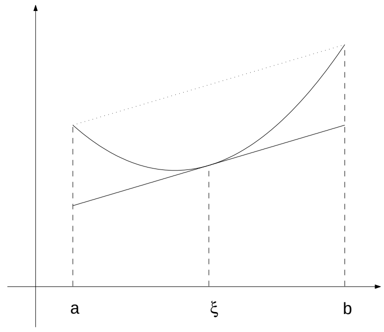
Die nächste Satzgruppe erscheint trivial, trotzdem gehören diese Sätze zu den nützlichsten Hilfsmitteln der Analysis:

Theorem V-4.4 (Satz von Rolle). *Es sei $f\colon [a, b] \to \mathbb{R}$ stetig, $f(a) = f(b)$, und f sei differenzierbar auf (a, b). Dann gibt es eine Zwischenstelle $\xi \in (a, b)$ mit $f'(\xi) = 0$.*

Beweis. Ist f konstant, dann gilt $f'(\xi) = 0$ für alle $\xi \in (a, b)$. Andernfalls nimmt f als stetige Funktion auf $[a, b]$ nach Korollar III-4.2 ein Maximum und Minimum an. Mindestens einer der Extremwerte ist von $f(a) = f(b)$ verschieden. Somit wird ein Extremum an einer Stelle $\xi \in (a, b)$ angenommen. Nach Satz V-4.2 ist $f'(\xi) = 0$. □



Satz von Rolle



Mittelwertsatz

Theorem V-4.5 (Verallgemeinerter Mittelwertsatz, Cauchy). *Die Funktionen f, $g\colon [a,b] \to \mathbb{R}$ seien stetig auf $[a,b]$ und differenzierbar auf (a,b). Dann gibt es eine Zwischenstelle $\xi \in (a,b)$ mit*

$$(f(b)-f(a))g'(\xi) = (g(b)-g(a))f'(\xi).$$

Beweis. Wir definieren die Abbildung $\varphi \in C([a,b],\mathbb{R})$ durch

$$\varphi(x) = (f(b)-f(a))g(x) - (g(b)-g(a))f(x).$$

Man überprüfe, daß φ die Voraussetzungen des Satzes von Rolle erfüllt. Somit gibt es ein $\xi \in (a,b)$ mit $\varphi'(\xi) = (f(b)-f(a))g'(\xi) - (g(b)-g(a))f'(\xi) = 0$. □

Korollar V-4.6 (Mittelwertsatz, Lagrange). Es sei $f\colon [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar auf (a,b). Dann gibt es eine Zwischenstelle $\xi \in (a,b)$ mit

$$f(b)-f(a) = f'(\xi)(b-a).$$

Beweis. Setze $g = \mathrm{id}$ in Satz V-4.5 □

Oft formuliert man den Mittelwertsatz folgendermaßen: Es gibt eine Zahl $\vartheta \in (0,1)$ mit

$$f(b) = f(a) + f'(a+\vartheta(b-a))(b-a).$$

Als unmittelbare Anwendung des Mittelwertsatzes ergeben sich folgende Fakten:

Korollar V-4.7. Es sei $f \in C([a,b],\mathbb{R})$ differenzierbar auf (a,b). Es gilt:

i) f konstant auf $[a,b]$ genau dann wenn $f'(x) = 0$ für alle $x \in (a,b)$.
ii) f monoton wachsend (fallend) auf $[a,b]$ genau dann, wenn $f'(x) \geq 0$ ($f'(x) \leq 0$) für alle $x \in (a,b)$.
iii) Gilt $f'(x) > 0$ ($f'(x) < 0$) für alle $x \in (a,b)$, dann ist f streng monoton wachsend (fallend) auf $[a,b]$.

Beweis. Es sei $x, y \in [a,b]$ und $x < y$. Wir wenden den MWS auf $f|_{[x,y]}$ an und erhalten:

$$f(y) - f(x) = f'(\xi)(y-x)$$

für ein $\xi \in (x, y)$. Aus dieser Darstellung liest man alle Behauptungen ab. $\square$

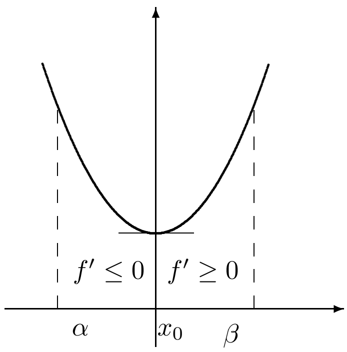
Bemerkung V-4.8. Auf die Voraussetzung, daß der Definitionsbereich ein Intervall ist, kann nicht verzichtet werden: Es sei etwa $f\colon D \to \mathbb{R}$, $D = [0,1] \cup [2,3]$ und $f(x) = 1$, $x \in [0,1]$ und $f(x) = 2$, $x \in [2,3]$. Es ist $f' = 0$, aber f ist nicht konstant. Das Beispiel $f(x) = |x|$, $x \in [-1,1]$, zeigt, daß die Behauptung des Satzes von Rolle nicht zutreffen muß, wenn die Funktion auch nur an einer einzigen Stelle $x \in (a,b)$ nicht differenzierbar ist.

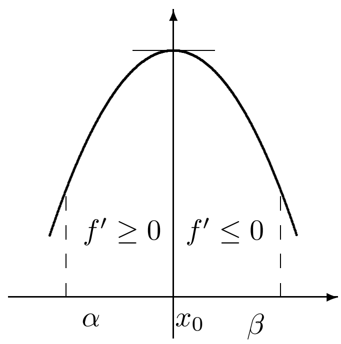
Wir sind nun auch in der Lage, eine hinreichende Bedingung für das Auftreten lokaler Extrema zu formulieren.

Theorem V-4.9. *Es sei $f\colon (a,b) \to \mathbb{R}$ differenzierbar auf einer Umgebung $(x_0 - \delta, x_0 + \delta)$ von $x_0 \in (a,b)$. Ist $f'(x_0) = 0$, so hat f in x_0*

a) *ein lokales Maximum, falls $f'(x) \geq 0$ für alle $x \in (x_0 - \delta, x_0)$ und $f'(x) \leq 0$ für alle $x \in (x_0, x_0 + \delta)$,*
b) *ein lokales Minimum, falls $f'(x) \leq 0$ für alle $x \in (x_0 - \delta, x_0)$ und $f'(x) \geq 0$ für alle $x \in (x_0, x_0 + \delta)$.*

Beweis. zu a): Nach Korollar V-4.7 ist f monoton wachsend auf $(x_0 - \delta, x_0]$ und monoton fallend auf $[x_0, x_0 + \delta)$, d.h. für alle $x \in (x_0 - \delta, x_0 + \delta)$ gilt $f(x) \leq f(x_0)$. $\square$





Korollar V-4.10. Es sei $f\colon (a,b) \to \mathbb{R}$ differenzierbar auf einer δ-Umgebung von $x_0 \in (a,b)$. Ist $f'(x_0) = 0$ und existiert $f''(x_0)$, so hat f in x_0

a) ein lokales Maximum, falls $f''(x_0) < 0$,
b) ein lokales Minimum, falls $f''(x_0) > 0$.

Beweis. zu a) Es sei $f''(x_0) < 0$. Aus

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

schließen wir auf die Existenz von $\sigma > 0$, sodaß

$$\forall x \in \mathbb{R}\colon 0 < |x - x_0| < \sigma \Rightarrow \frac{f'(x)}{x - x_0} < 0$$

zutrifft. In einer σ-Umgebung von x_0 ist also

$$\begin{aligned} f'(x) > 0 \quad &\text{für} \quad x < x_0 \\ f'(x) < 0 \quad &\text{für} \quad x > x_0, \end{aligned}$$

nach Satz V-4.9 liegt somit in x_0 ein lokales Maximum vor. □

Der nächste Satz zeigt, daß nicht jede Funktion eine Ableitung sein kann:

Theorem V-4.11 (Zwischenwertsatz). *Es sei $f \in C([a,b],\mathbb{R})$ differenzierbar auf $[a,b]$ und $f'(a) \neq f'(b)$. Dann nimmt f' in (a,b) jeden Wert zwischen $f'(a)$ und $f'(b)$ an.*

Beweis. Es sei o.B.d.A. $f'(a) < \lambda < f'(b)$. Wir setzen $g(x) = f(x) - \lambda x$. Wegen $g'(a) < 0$ und $g'(b) > 0$ gibt es Konstante δ_0, $\delta_1 > 0$, sodaß $g(x) < g(a)$ für alle $x \in (a, a+\delta_0)$ und $g(x) < g(b)$ für alle $x \in (b-\delta_1, b)$ gilt. Da g stetig ist, nimmt g auf $[a,b]$ das Minimum an. Die vorausgehende Überlegung zeigt, daß das Minimum nicht in den Randpunkten a und b angenommen werden kann. Nach Satz V-4.2 gibt es daher eine Stelle $x_0 \in (a,b)$ mit $g'(x_0) = f'(x_0) - \lambda = 0$. □

Korollar V-4.12. Es sei $f \in C([a,b],\mathbb{R})$ differenzierbar auf $[a,b]$. Dann kann f' keine Unstetigkeiten 1. Art auf $[a,b]$ haben.

Beweis. Übung (Man beachte, daß die Ableitung einer auf $[a,b]$ differenzierbaren Funktion keine hebbaren Unstetigkeiten aufweisen kann). □

5. Regel von de L'Hospital

Wir wenden uns nun Grenzwerten zu, die auf sogenannte **unbestimmte Formen** führen: $\frac{0}{0}, \frac{\infty}{\infty}, \infty - \infty, \infty^0, 0^0, 0 \cdot \infty$. Es genügt, die unbestimmten Formen $\frac{0}{0}$ bzw. $\frac{\infty}{\infty}$ zu betrachten, da sich die anderen unbestimmten Formen durch geeignete Transformationen auf diese beiden Möglichkeiten zurückführen lassen.

Wir gehen von folgender Beobachtung aus: Zu berechnen ist $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ und es sei $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$. Dieser Grenzwert läßt sich mühelos bestimmen, falls f und g in x_0 differenzierbar sind und $g'(x_0) \neq 0$ ist: da in diesem Falle f und g in x_0 stetig sind, gilt $f(x_0) = g(x_0) = 0$ und es folgt

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

Diese Vorgangsweise läßt sich erheblich verallgemeinern. Insbesondere kann auf die Differenzierbarkeit von f und g in x_0 verzichtet werden, mehr noch, f und g müssen in x_0 nicht einmal definiert sein. Im folgenden sind die auftretenden Grenzwerte entweder im eigentlichen oder uneigentlichen Sinne zu verstehen:

Theorem V-5.1 (Regel von de L'Hospital). *Es seien $f, g\colon (a,b) \to \mathbb{R}$, $-\infty \le a < b \le \infty$, f, g differenzierbar auf (a,b) und es sei $g'(x) \neq 0$ für alle $x \in (a,b)$. Es existiere $\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ in $\bar{\mathbb{R}}$. In jeder der beiden Situationen*

a) $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$,
b) $\lim_{x\to a^+} g(x) = \infty$

existiert dann auch $\lim_{x\to a^+} \frac{f(x)}{g(x)}$ *und es ist*

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

(Eine analoge Aussage gilt auch für $x \to b^-$ bzw. $g(x) \to -\infty$.)

Beweis. Fall 1: $L = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} \in [-\infty, \infty)$. Wir wählen $q \in \mathbb{R}$ mit $L < q$ und anschließend $r \in \mathbb{R}$ mit $L < r < q$. Zu r gibt es ein $\xi_r \in \mathbb{R}$, sodaß

$$\forall x \in (a, \xi_r)\colon \frac{f'(x)}{g'(x)} < r$$

zutrifft. Es sei nun $a < x < y < \xi_r$. Wegen $g'(x) \neq 0$ auf (a, b) gilt nach Satz V-4.11 entweder $g'(x) > 0$ oder $g'(x) < 0$ auf (a, b). Daher ist g strikt monoton, insbesondere gilt $g(x) - g(y) \neq 0$. Nach Satz V-4.5 gibt es eine Zwischenstelle $\tau \in (x, y)$, sodaß

$$(*) \qquad \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(\tau)}{g'(\tau)} < r.$$

Es sei a) erfüllt. Läßt man x gegen a rücken, folgt daher

$$\frac{f(y)}{g(y)} \leq r < q \text{ für alle } y \in (a, \xi_r).$$

(Man beachte, daß die Zwischenstelle τ von x und y abhängig ist.) Es sei nun b) erfüllt. Wir halten wieder y fest und bestimmen $\xi_y \in (a, y)$ so, daß

$$g(x) > \max\{0, g(y)\} \qquad \text{für alle } x \in (a, \xi_y)$$

zutrifft. Es ist also $(g(x) - g(y))/g(x)$ positiv. Multipliziert man $(*)$ mit diesem Bruch ergibt sich für alle $x \in (a, \xi_y)$

$$\frac{f(x) - f(y)}{g(x)} < r - \frac{g(y)}{g(x)} r$$

bzw.

$$\frac{f(x)}{g(x)} < r + \frac{f(y)}{g(x)} - r\frac{g(y)}{g(x)}.$$

Die rechte Seite dieser Ungleichung konvergiert für $x \to a$ nach r. Daraus läßt sich jedoch nicht ableiten, daß auch $\frac{f(x)}{g(x)} \leq r$ für alle $x \in (a, \xi_y)$ zutrifft. Wohl aber kann man auf die Existenz von $\sigma \in (a, \xi_y)$ schließen, sodaß

$$\frac{f(x)}{g(x)} < q \qquad \text{für alle } x \in (a, \sigma)$$

zutrifft (erst an dieser Stelle des Beweises machen wir von q Gebrauch). Für beide Situationen a) und b) wurde also bisher folgendes gezeigt:

$$(1) \qquad \forall q > L\, \exists x_1 \in (a, b)\, \forall x \in (a, x_1)\colon \frac{f(x)}{g(x)} < q.$$

Fall 2: $L = \lim_{x\to a^+} \frac{f'(x)}{g'(x)} \in (-\infty, \infty]$. Analog wie im ersten Fall kann gezeigt werden:

$$(2) \qquad \forall p < L\, \exists x_2 \in (a, b)\, \forall x \in (a, x_2)\colon \frac{f(x)}{g(x)} > p.$$

Die Behauptung folgt nun leicht aus (1) und (2): Gilt $L = -\infty$ bzw. $L = \infty$, ist die Behauptung durch (1) bzw. (2) bereits bewiesen. Es sei $L \in \mathbb{R}$. Zu $\varepsilon > 0$ wählen wir $q = L + \varepsilon$ in (1), bzw. $p = L - \varepsilon$ in (2) und setzen $\bar{x} = \min\{x_1, x_2\}$. Für $x \in (a, \bar{x})$ gilt dann $|\frac{f(x)}{g(x)} - L| < \varepsilon$. □

Wir haben die Regel von de L'Hospital für Randpunkte des Definitionsbereiches formuliert. Dies bedeutet jedoch keine Einschränkung, da man immer auf einseitige Limiten geeigneter Einschränkungen zurückgehen kann.

Beispiel V-5.2. Man verifiziere in den folgenden Beispielen die Voraussetzungen von Satz V-5.1:

(1) $\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \cos x = 1.$ (Form: $\frac{0}{0}$)

(2) $\forall \alpha > 0\colon\ \lim_{x\to\infty} \frac{e^{\alpha x}}{x} = \infty.$ (Form: $\frac{\infty}{\infty}$)

(3) $\lim_{x\downarrow 0} x \ln x = \lim_{x\downarrow 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x\downarrow 0} \frac{\frac{1}{x}}{-(\frac{1}{x})^2} = -\lim_{x\downarrow 0} x = 0.$ (Form: $0\cdot\infty$)

(4) $\lim_{x\to 0^+} x^x = 1$ (Form: 0^0)

Wir verwenden die Identität $x^x = e^{x\ln x}$. Wegen der Stetigkeit der Exponentialfunktion genügt es vorerst $\lim_{x\to 0^+} x\ln x$ zu untersuchen. Nach (3) gilt $\lim_{x\to 0^+} x \ln x = 0$ und daher $\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} e^{x\ln x} = \exp(\lim_{x\to 0^+} x\ln x) = 1.$

(5) $\lim_{x\to\infty} (e^x + x)^{\frac{1}{x}} = e.$ (Form: ∞^0)

Es gilt $(e^x + x)^{\frac{1}{x}} = \exp(\frac{1}{x}\ln(e^x + x))$. Aus

$$\lim_{x\to\infty} \frac{\ln(e^x + x)}{x} = \lim_{x\to\infty} \frac{e^x + 1}{e^x + x} = 1$$

folgt wie in Beispiel 4 die Behauptung.

(6) $\lim_{x\to 1} \left(\frac{x}{x-1} - \frac{1}{\ln x}\right) = \frac{1}{2}.$ (Form: $\infty - \infty$)

$$\begin{aligned}
\lim_{x\to 1} \left(\frac{x}{x-1} - \frac{1}{\ln x}\right) &= \lim_{x\to 1} \frac{x\ln x - x + 1}{(x-1)\ln x} \qquad (\text{Form: } \frac{0}{0})\\
&= \lim_{x\to 1} \frac{\ln x + 1 - 1}{\ln x + 1 - \frac{1}{x}} = \lim_{x\to 1} \frac{\ln x}{\ln x + 1 - \frac{1}{x}} \qquad (\text{Form:} \frac{0}{0})\\
&= \lim_{x\to 1} \frac{\frac{1}{x}}{\frac{1}{x} + \frac{1}{x^2}} = \frac{1}{2}.
\end{aligned}$$

Dieses Beispiel zeigt, daß es in manchen Fällen notwendig ist, die Regel von de L'Hospital mehrfach anzuwenden. Da der Rechenaufwand rasch anwächst, sei vor einer reflexartigen Anwendung dieser Regel gewarnt. Viele Grenzwerte lassen sich durch einen Potenzreihenansatz bedeutend einfacher berechnen.

(7) $\lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} = 1.$ (Form: $\frac{\infty}{\infty}$)

Eine blinde Anwendung von Satz V-5.1 ergibt

$$\lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} = \lim_{x\to\infty} \frac{e^x + e^{-x}}{e^x - e^{-x}} = \lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} = \dots,$$

also eine endlose Schleife, während eine einfache Umformung sofort auf den gewünschten Grenzwert führt:

$$\lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} = \lim_{x\to\infty} \frac{1 - e^{-2x}}{1 + e^{-2x}} = 1.$$

Zuletzt zeigen wir noch, daß Satz V-5.1 sich nicht (ohne Zusatzvoraussetzungen) auf komplexwertige Funktionen übertragen läßt:

Beispiel V-5.3. $g, f\colon (0,1) \to \mathbb{C}$ seien definiert durch $g(x) = x + x^2 e^{\frac{i}{x^2}}$ und $f(x) = x$. Wegen $|e^{\frac{i}{x^2}}| = 1$ für $x \in (0,1)$ folgt

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = 1.$$

Satz V-5.1 führt auf

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{1}{(1+(2x-\frac{2i}{x})e^{\frac{i}{x^2}})}.$$

Wegen

$$|g'(x)| \geq |2x - \frac{2i}{x}| - 1 \geq \frac{2}{x} - 1$$

und somit

$$|\frac{f'(x)}{g'(x)}| \leq \frac{x}{2-x},$$

erhält man

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = 0.$$

Da $g'(x) \neq 0$ würde Satz V-5.1 zu dem falschen Ergebnis

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = 0$$

führen.

6. Differenzierbarkeit und gleichmäßige Konvergenz

Wir haben bereits in Abschnitt IV-5 gesehen, daß Funktionenfolgen bzw. Funktionenreihen ein sehr nützliches Instrument zur Darstellung stetiger Funktionen sind. Wir wenden uns nun der Differenzierbarkeit der Grenzfunktion zu.

Beispiel V-6.1. Wir betrachten die Funktionenfolge $(f_n)_{n\geq 1}$, $f_n(x) = \frac{1}{\sqrt{n}} \sin nx$, $x \in \mathbb{R}$. Die Abschätzung $\|f_n\|_\infty \leq \frac{1}{\sqrt{n}}$ zeigt, daß f_n *gleichmäßig* auf $\mathbb{R}$ gegen $f = 0$ konvergiert. Es gilt $f_n'(x) = \sqrt{n}\cos nx$ und $f'(x) = 0$. Wir behaupten, daß für jedes $x \in \mathbb{R}$ die Folge $(f_n'(x))_{n\geq 1}$ unbeschränkt ist und daher insbesondere nicht gegen $f'(x)$ konvergieren kann. Wir wählen also ein beliebiges $x \in \mathbb{R}$. Wenn $n \in \mathbb{N}$ und $|\cos nx| < \frac{1}{2}$ ist, dann folgt

$$|\cos 2nx| = |2\cos^2(nx) - 1| > \frac{1}{2}.$$

Für jedes $n \in \mathbb{N}$ gilt daher entweder

$$|\cos nx| \geq \frac{1}{2} \quad \text{oder} \quad |\cos 2nx| \geq \frac{1}{2}$$

und daher auch

$$\sup\{|f_n'(x)| \colon n \in \mathbb{N}\} = \infty.$$

Dieses Beispiel zeigt, daß es möglich ist, daß eine Folge (f_n) (sogar beliebig oft) differenzierbarer Funktionen gleichmäßig konvergiert, aber die Ableitungsfolge (f_n') sogar nirgends konvergiert. Das Beispiel zeigt auch auf, daß man den Hebel bei der Folge der Ableitungen ansetzen muß:

Theorem V-6.2. *Es sei $(f_n)_{n\geq 1} \subset C([a,b],\mathbb{R})$, $a < b$, eine Folge differenzierbarer Funktionen. Es gelte*

(1) *(f_n') ist gleichmäßig konvergent, d.h. $\exists g \colon [a,b] \to \mathbb{R}$, $\lim_{n\to\infty} f_n' = g$, glm.,*
(2) *$\exists x_0 \in [a,b] \colon (f_n(x_0))_{n\geq 1}$ ist konvergent.*

Dann konvergiert die Folge (f_n) *gleichmäßig gegen eine differenzierbare Funktion* $f \in C([a,b],\mathbb{R})$ *und es gilt*

$$f' = g = \lim_{n\to\infty} f_n'.$$

Beweis. 1. Schritt: Wir zeigen zuerst die gleichmäßige Konvergenz von (f_n) und schätzen dazu ab:

$$|f_n(x) - f_m(x)| \le |f_n(x) - f_m(x) - (f_n(x_0) - f_m(x_0))| + |f_n(x_0) - f_m(x_0)|.$$

Auf die Abbildung $f_n - f_m$ wenden wir den Mittelwertsatz an und erhalten mit ξ zwischen x und x_0:

$$(*) \qquad \begin{aligned} |f_n(x) - f_m(x)| &= |f_n'(\xi) - f_m'(\xi)||x - x_0| + |f_n(x_0) - f_m(x_0)| \\ &\le \|f_n' - f_m'\|_\infty (b-a) + |f_n(x_0) - f_m(x_0)|. \end{aligned}$$

Wegen der gleichmäßigen Konvergenz von (f_n') existiert nach Satz IV-3.5 zu $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ so, daß für alle $n, m \ge N(\varepsilon)$

$$\|f_n' - f_m'\|_\infty < \frac{\varepsilon}{2(b-a)},$$

wegen der Konvergenz von $(f_n(x_0))_{n\ge 1}$, kann man darüber hinaus $N(\varepsilon)$ so groß wählen, daß für alle $n, m \ge N(\varepsilon)$ auch

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$$

zutrifft. Somit gilt für alle $n, m \ge N(\varepsilon)$ wegen $(*)$

$$\|f_n - f_m\|_\infty < \varepsilon,$$

d.h. (f_n) ist eine gleichmäßige Cauchyfolge und daher nach Satz IV-3.5 gleichmäßig konvergent. Die Grenzfunktion f ist wegen der Stetigkeit der f_n nach Satz IV-3.6 selbst stetig.

2. Schritt: f ist an jeder Stelle $\xi \in [a,b]$ differenzierbar mit $f'(\xi) = \lim_{n\to\infty} f_n'(\xi)$. Jede Abbildung f_n ist an der Stelle ξ differenzierbar, d.h. es gibt eine in ξ stetige Funktion $R_n\colon [a,b] \to \mathbb{R}$, sodaß

$$\begin{aligned} \forall x \in [a,b]\colon f_n(x) &= f_n(\xi) + f_n'(\xi)(x-\xi) + R_n(x)(x-\xi) \\ R_n(\xi) &= 0. \end{aligned}$$

Wenn wir nun zeigen können, daß die Funktionenfolge (R_n) gleichmäßig konvergiert, erhalten wir mit $R(x) = \lim_{n\to\infty} R_n(x)$ durch Grenzübergang

$$\begin{aligned} &f(x) = f(\xi) + g(\xi)(x-\xi) + R(x)(x-\xi) \\ &R(\xi) = 0, \quad R \text{ stetig in } \xi, \end{aligned}$$

d.h. f ist an der Stelle ξ differenzierbar und es ist $f'(\xi) = \lim_{n\to\infty} f_n'(\xi) = g(\xi)$. Für $x \ne \xi$ erhalten wir

$$R_m(x) - R_n(x) = \frac{f_m(x) - f_m(\xi) - (f_n(x) - f_n(\xi))}{x - \xi} - (f_m'(\xi) - f_n'(\xi)).$$

Wie im ersten Schritt ergibt eine Anwendung des Mittelwertsatzes

$$|R_m(x) - R_n(x)| \leq 2\|f_n' - f_m'\|_\infty, \quad x \neq \xi,$$

somit wegen $R_m(\xi) - R_n(\xi) = 0$ auch

$$\|R_m - R_n\| \leq 2\|f_n' - f_m'\|_\infty.$$

Die Folge (R_n) ist also eine gleichmäßige Cauchy Folge. Nach Satz IV-3.5 ist die Folge (R_n) gleichmäßig konvergent und da R_n, $n \in \mathbb{N}$, an der Stelle 0 stetig ist, muß auch R an der Stelle 0 stetig sein. □

Wenden wir diesen Satz auf gleichmäßig konvergente Reihen differenzierbarer Funktionen an, erhalten wir

Korollar V-6.3. Es sei $(f_n) \subset C([a,b],\mathbb{R})$ eine Folge differenzierbarer Funktionen. Es gelte

i) $\sum\limits_{n=1}^{\infty} f_n'$ ist gleichmäßig konvergent,

ii) $\exists x_0 \in [a,b]\colon \sum\limits_{n=1}^{\infty} f_n(x_0)$ ist konvergent.

Dann konvergiert die Reihe $\sum\limits_{n=1}^{\infty} f_n$ gleichmäßig gegen eine differenzierbare Funktion $f \in C([a,b],\mathbb{R})$ und es gilt

$$\forall x \in [a,b]\colon f'(x) = \left(\sum_{n=1}^{\infty} f_n(x)\right)' = \sum_{n=1}^{\infty} f_n'(x).$$

Satz V-6.2 bzw. Korollar V-6.3 geben hinreichende Bedingungen an, die die Vertauschbarkeit von Limes und Ableitung gewährleisten. Unter diesen Voraussetzungen darf man also gliedweise differenzieren.

Korollar V-6.4. Eine Potenzreihe $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$, $(a_n)_{n\geq 1} \subset \mathbb{R}$, stellt im Inneren des Konvergenzintervalles $(x_0 - R, x_0 + R)$ eine beliebig oft differenzierbare Funktion f dar, d.h. $f \in C^\infty((x_0-R, x_0+R),\mathbb{R})$. Es gilt

$$\forall x \in (x_0 - R, x_0 + R)\colon f^{(k)}(x) = \sum_{n=k}^{\infty} \prod_{j=0}^{k-1} (n-j) a_n (x-x_0)^{n-k}.$$

Beweis. Die gliedweise differenzierte Potenzreihe

$$h(x) = \sum_{n=1}^{\infty} a_n n (x-x_0)^{n-1}$$

besitzt ebenfalls den Konvergenzradius R (Übung). Die Potenzreihe h konvergiert daher gleichmäßig auf jedem abgeschlossenen und beschränkten Teilintervall von

$(x_0 - R, x_0 + R)$. Mit Korollar V-6.3 folgt daher $h(x) = f'(x)$, $x \in (x_0 - R, x_0 + R)$. Die Behauptung folgt nun mit vollständiger Induktion. □

7. Taylorpolynome und Taylorreihen

Der Differenzierbarkeitsbegriff wurde in Abschnitt 6.1 mit dem Bestreben motiviert, eine Funktion f lokal durch eine möglichst einfache Funktion, z.B. ein Polynom ersten Grades $T_1(x) = a_0 + (x - x_0)a_1$, zu approximieren. Die Forderung $T_1(x_0) = f(x_0)$ ergab $a_0 = f(x_0)$. Weiters sollte der Approximationsfehler $r(x - x_0) = f(x) - T_1(x)$ hinreichend rasch abklingen: $\lim_{x \to x_0} \frac{r(x-x_0)}{x-x_0} = 0$. Diese Bedingung legt nicht nur den Wert von a_1 fest, sie sichert auch die Eindeutigkeit der Approximation. Es ist intuitiv klar, daß eine Verbesserung der Approximationsgüte zu erwarten ist, wenn Polynome n-ten Grades zur Approximation verwendet werden.

Theorem VI-7.1. *Es sei $f\colon I \to \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall, n-mal an der Stelle $x_0 \in I$ differenzierbar. Das Polynom*

$$T_n(x) = \sum_{i=0}^{n} \frac{1}{i!} f^{(i)}(x_0)(x - x_0)^i$$

ist das eindeutig bestimmte Polynom n-ten Grades mit

$$T_n^{(k)}(x_0) = f^{(k)}(x_0), \quad k = 0, \dots, n.$$

*T_n heißt **n-tes Taylorpolynom** von f um die Entwicklungsstelle x_0.*

Beweis. Wir setzen T_n in der Form

$$T_n(x) = \sum_{i=0}^{n} a_i (x - x_0)^i, \quad a_i \in \mathbb{R}, \quad i = 0, \dots, n,$$

an. Aus Korollar V-6.4 folgt

$$T_n^{(k)}(x) = \sum_{i=k}^{n} \prod_{j=0}^{k-1} (i - j) a_i (x - x_0)^{i-k}$$

und insbesondere

$$T_n^{(k)}(x_0) = \prod_{j=0}^{k-1} (k - j) a_k = k! a_k.$$

Die Bedingung $T_n^{(k)}(x_0) = f^{(k)}(x_0)$ ergibt dann

$$a_k = \frac{1}{k!} f^{(k)}(x_0), \quad k = 0, \dots, n.$$

Da Polynome n-ten Grades durch Angabe der $n+1$ Koeffizienten eindeutig bestimmt sind, folgt, daß T_n das einzige Polynom mit den geforderten Eigenschaften ist. □

Lemma VI-7.2. *Es sei $f\colon I \to \mathbb{R}$, $I \subset \mathbb{R}$ ein Intervall, n-mal an der Stelle $x_0 \in I$ differenzierbar. T_n bezeichne das n-te Taylorpolynom von f an der Stelle x_0. Dann gilt*

$$\lim_{x\to x_0} \frac{f(x) - T_n(x)}{(x-x_0)^n} = 0.$$

Beweis. Die Abbildung $R_n(x) := f(x) - T_n(x)$, $x \in I$, ist n mal differenzierbar an der Stelle x_0 und erfüllt

$$R_n^{(k)}(x_0) = 0, \quad k = 0, \ldots, n.$$

Wenden wir die Regel von de L'Hospital $n-1$ mal an erhalten wir

$$\begin{aligned}
\lim_{x\to x_0} \frac{R_n(x)}{(x-x_0)^n} &= \frac{1}{n!} \lim_{x\to x_0} \frac{R_n^{(n-1)}(x)}{x-x_0} \\
&= \frac{1}{n!} \lim_{x\to x_0} \frac{f^{(n-1)}(x) - T_n^{(n-1)}(x)}{x-x_0} \\
&= \frac{1}{n!} \lim_{x\to x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0) - f^{(n)}(x_0)(x-x_0)}{x-x_0} = 0.
\end{aligned}$$

Die letzte Gleichung folgt aus der Existenz der n-ten Ableitung von f an der Stelle x_0. Wir bemerken, daß eine unmittelbare n-fache Anwendung der Regel von de L'Hospital nicht möglich ist, da $f^{(n)}(x)$ für $x \neq x_0$ nicht zu existieren braucht. □

Dieses Lemma rechtfertigt zwar, ein Taylorpolynom als lokale Approximation von f zu betrachten, gibt aber keine Information über die Größe des Approximationsfehlers. Dies erfordert zusätzliche Glattheit von f (d.h. bessere Differenzierbarkeitseigenschaften):

Theorem VI-7.3 (Satz von Taylor). *Es sei $f \in C^n([x_0, x], \mathbb{R})$, $x_0 < x$, und es existiere $f^{(n+1)}$ zumindest auf (x_0, x). Dann gibt es für alle $s > 0$ eine Zwischenstelle $\xi \in (x_0, x)$ mit*

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{sn!} \left(\frac{x-\xi}{x-x_0}\right)^{n-s+1} (x-x_0)^{n+1}.$$

*Der letzte Summand heißt **Restglied nach Schlömilch**. Eine analoge Aussage gilt auch für $x < x_0$.*

Beweis. Wir gehen aus von der Darstellung

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k + R_n(x),$$

mit einem noch zu bestimmenden Restglied $R_n(x)$. Wir betrachten die Abbildung $F \in C([x_0, x], \mathbb{R})$ definiert durch

$$F(t) = f(x) - \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(t)(x-t)^k$$

(die Summe ist nichts anderes, als das n-te Taylorpolynom von f um die Entwicklungsstelle t). Es gilt

$$(*) \qquad F(x) = 0, \quad F(x_0) = R_n(x).$$

F ist auf (x_0, x) differenzierbar und es ist

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Dies sieht man folgendermaßen ein: Die Produktregel ergibt

$$\begin{aligned} F'(t) &= -f'(t) - \sum_{k=1}^{n} \frac{1}{k!}[f^{(k+1)}(t)(x-t)^k - kf^{(k)}(t)(x-t)^{k-1}] \\ &= -\sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(t)(x-t)^k - \frac{f^{(n+1)}(t)}{n!}(x-t)^n + \sum_{j=0}^{n-1} \frac{1}{j!} f^{(j+1)}(t)(x-t)^j. \end{aligned}$$

Es sei nun $\psi \in C([x_0, x], \mathbb{R})$ eine beliebige, streng monotone Abbildung, differenzierbar auf (x_0, x) mit $\psi'(\xi) \neq 0$ für $\xi \in (x_0, x)$. Wir wenden nun den verallgemeinerten Mittelwertsatz V-4.5 auf die Funktionen F und ψ an und erhalten

$$\frac{F(x) - F(x_0)}{\psi(x) - \psi(x_0)} = \frac{F'(\xi)}{\psi'(\xi)}, \quad \text{für ein } \xi \in (x_0, x).$$

Aus $(*)$ folgt somit

$$\begin{aligned} R_n(x) = F(x_0) &= -\frac{F(x) - F(x_0)}{\psi(x) - \psi(x_0)}(\psi(x) - \psi(x_0)) \\ &= -(\psi(x) - \psi(x_0))\frac{F'(\xi)}{\psi'(\xi)} = \frac{\psi(x) - \psi(x_0)}{\psi'(\xi)} \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n. \end{aligned}$$

Wählt man speziell $\psi(t) = (x-t)^s$, $s > 0$, erhält man die Schlömilchsche Form des Restgliedes

$$R_n(x) = (x-x_0)^s \frac{f^{(n+1)}(\xi)}{sn!}(x-\xi)^{n-s+1} = \frac{f^{(n+1)}(\xi)}{sn!} \cdot \left(\frac{x-\xi}{x-x_0}\right)^{n-s+1} (x-x_0)^{n+1}.$$

□

Ohne Schwierigkeit ergeben sich nun leichter handhabbare Formen des Restgliedes:

Korollar VI-7.4. f sei wie in Satz VI-7.3 und $f(x) = T_n(x) + R_n(x)$. Es gibt eine Zwischenstelle $\xi \in (x_0, x)$ $(\xi \in (x, x_0))$ so, daß R_n folgende Darstellungen besitzt:

i) $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$. (Restglied nach **Lagrange**)

ii) $R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}\left(\frac{x-\xi}{x-x_0}\right)^n (x-x_0)^{n+1}$. (Restglied nach **Cauchy**)

Beweis. Setze $s = n+1$ bzw. $s = 1$ in VI-7.3. □

Wir weisen darauf hin, daß die Zwischenstelle ξ von x_0, x, n und s abhängt. Da der Satz von Taylor keine Information liefert, wo die Zwischenstelle liegt, ist man auf Abschätzungen von $|R_n(x)|$ angewiesen, um beurteilen zu können, wie gut

$T_n(x)$ den Funktionswert $f(x)$ approximiert. Oft ist es bei derartigen Abschätzungen zweckmäßig, die Zwischenstelle ξ in der Form

$$\xi = x_0 + \vartheta(x - x_0), \quad \vartheta \in (0,1),$$

zu schreiben. Setzt man noch $h = x - x_0$ und $R_n^X(h) = R_n^X(x) = R_n^X(x_0 + h)$, nimmt der Satz von Taylor folgende Form an:

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0) h^k + R_n^X(h), \quad X = S, L, C,$$

mit

$$\begin{aligned}
R_n^S(h) &= \frac{f^{(n+1)}(x_0 + \vartheta h)}{sn!}(1-\vartheta)^{n-s+1} h^{n+1} && \text{Schlömilch},\\
R_n^L(h) &= \frac{f^{(n+1)}(x_0 + \vartheta h)}{(n+1)!} h^{n+1} && \text{Lagrange},\\
R_n^C(h) &= \frac{f^{(n+1)}(x_0 + \vartheta h)}{n!}(1-\vartheta)^{n} h^{n+1} && \text{Cauchy}.
\end{aligned}$$

Beispiel VI-7.5. Als Beispiel betrachten wir die Logarithmusfunktion $f(x) = \ln(1+x)$ in einer Umgebung von $x_0 = 0$. Man verifiziert induktiv

$$\begin{aligned}
&\forall k \in \mathbb{N}\colon f^{(k)}(x) = (-1)^{k-1}(k-1)!(1+x)^{-k} \quad \text{und somit}\\
&\forall k \in \mathbb{N}\colon f^{(k)}(0) = (-1)^{k-1}(k-1)!.
\end{aligned}$$

Es gilt daher für alle $h > -1$ die Darstellung

$$\ln(1+h) = \sum_{i=1}^{n} (-1)^{i-1} \frac{h^i}{i} + R_n^L(h).$$

Der Approximationsfehler ist nach Korollar VI-7.4 gegeben durch

$$R_n^L(h) = (-1)^n \frac{1}{(1+\vartheta h)^{n+1}} \frac{h^{n+1}}{n+1}.$$

Für $h > 0$ folgt

$$|R_n^L(h)| < \frac{h^{n+1}}{n+1},$$

da $1 + \vartheta h > 1$ gilt (eine bessere Abschätzung ist nicht möglich, da ϑ beliebig nahe bei 0 liegen kann). Für $h \in (-1, 0)$ ergibt sich

$$|R_n^L(h)| < \frac{1}{n+1} \frac{|h|^{n+1}}{(1-|h|)^{n+1}}.$$

Diese Abschätzung für $|R_n^L|$ kann sehr schlecht sein, z.B. ergibt sich für $h = -\frac{2}{3}$

$$|R_n^L(-\frac{2}{3})| \le \frac{2^{n+1}}{n+1},$$

also etwa $|R_9^L(-\frac{2}{3})| \le \frac{1024}{10}$. Verwenden wir jedoch die Cauchysche Form des Restgliedes, erhalten wir

$$R_n^C(h) = (-1)^n \left(\frac{1-\vartheta}{1+\vartheta h}\right)^n (1+\vartheta h)^{-1} h^{n+1}, \quad \vartheta \in (0,1).$$

Für $h \in (-1, 0)$ folgt nun wegen

$$1 + \vartheta h > 1 - \vartheta \quad \text{also} \quad 0 < \frac{1-\vartheta}{1+\vartheta h} < 1$$

die Abschätzung

$$(*) \qquad |R_n^C(h)| < \frac{1}{1-|h|} |h|^{n+1}.$$

Für $h = -\frac{2}{3}$ ergibt sich nun die bedeutend bessere Fehlerabschätzung

$$|R_n^C(-\frac{2}{3})| < 3(\frac{2}{3})^{n+1},$$

z.B. $|R_9^C(-\frac{2}{3})| < 0{,}053$.

Beispiel VI-7.6. Wir entwickeln nun die Potenzfunktion $f(x) = x^\alpha$, $\alpha \in \mathbb{R}$, in einer Umgebung von $x_0 = 1$. Für alle $h > -1$ gilt mit dem Restglied nach Lagrange die Darstellung

$$(*) \qquad (1+h)^\alpha = \sum_{k=0}^{n} \binom{\alpha}{k} h^k + \binom{\alpha}{n+1}(1+\vartheta h)^{\alpha-n-1}h^{n+1},$$

in der wir die Definition der Binomialkoeffizienten in naheliegender Weise erweitert haben:

$$\binom{\alpha}{0} := 1, \qquad \forall n \in \mathbb{N}\colon \binom{\alpha}{n} := \frac{1}{n!}\prod_{j=0}^{n-1}(\alpha - j).$$

Induktiv bestätigt man

$$\forall k \in \mathbb{N}\colon f^{(k)}(x) = \prod_{j=0}^{k-1}(\alpha - j)x^{\alpha-k} = k!\binom{\alpha}{k}x^{\alpha-k}, \qquad \text{also}$$

$$\forall k \in \mathbb{N}\colon f^{(k)}(1) = \prod_{j=0}^{k-1}(\alpha - j) = k!\binom{\alpha}{k},$$

dies impliziert $(*)$. Als Näherung für $(1+h)^\alpha$ verwendet man oft das 1. Taylorpolynom

$$(1+h)^\alpha \approx 1 + h\alpha, \qquad R_1^L(h) = \binom{\alpha}{2}(1+\vartheta h)^{\alpha-2}h^2$$

(das Symbol „$\approx$“ bedeutet ungefähr gleich), insbesondere ergibt sich

$$\sqrt{1+h} \approx 1 + \frac{h}{2}, \qquad \frac{1}{\sqrt{1+h}} \approx 1 - \frac{h}{2} \quad (|h| \text{ klein }).$$

Damit der Fehler klein ist, sollte also $|h|$ klein sein. Will man etwa $\sqrt[3]{2}$ berechnen, kann man z.B. von $2 = \frac{2}{x^3}x^3$, $x > 0$, ausgehen. Wir erhalten dann

$$\sqrt[3]{2} = x\sqrt[3]{1 + (\frac{2}{x^3} - 1)}.$$

Damit h klein ist wählen wir x so, daß $\frac{2}{x^3} - 1$ klein ist, z.B. $x = \frac{5}{4}$. Wir erhalten dann

$$\sqrt[3]{2} = \frac{5}{4}\sqrt[3]{1 + \frac{3}{125}} = \frac{5}{4}(1 + \frac{1}{125} + R_1^L(\frac{3}{125})).$$

Da $R_1^L(h)$ für $|h| < 1$ negativ ist $\left(\binom{\frac{1}{3}}{2} = -\frac{1}{9}\right)$, ergibt die Abschätzung

$$|R_1^L(\frac{3}{125})| < \frac{1}{9}(\frac{3}{125})^2$$

die Ungleichungen

$$\frac{5}{4} + \frac{1}{100} - \frac{5}{4}\frac{1}{(125)^2} < \sqrt[3]{2} < \frac{5}{4} + \frac{1}{100},$$

d.h.

$$1{,}25992 < \sqrt[3]{2} < 1{,}26.$$

Die Dezimalentwicklung von $\sqrt[3]{2}$ beginnt also mit 1,2599...

Als weitere Anwendung des Satzes von Taylor können wir nun die hinreichende Bedingung aus Satz V-4.9 für das Auftreten innerer, lokaler Extrema erweitern.

Theorem VI-7.7. *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^n(I, \mathbb{R})$. An der Stelle $x_0 \in I$ gelte*

$$f^{(k)}(x_0) = 0, \quad k = 1, \ldots, n-1 \quad \textit{und} \quad f^{(n)}(x_0) \neq 0.$$

Dann gilt:

(1) *Ist n ungerade, dann liegt in x_0 kein lokales Extremum vor.*
(2) *Ist n gerade, dann liegt in x_0 ein lokales Minimum vor falls $f^{(n)}(x_0) > 0$, und ein lokales Maximum, falls $f^{(n)}(x_0) < 0$.*

Beweis. Unter den Voraussetzungen des Satzes liefert die Taylorsche Formel mit dem Restglied nach Lagrange die Entwicklung

$$f(x_0 + h) = f(x_0) + \frac{f^{(n)}(x_0 + \vartheta h)}{n!} h^n, \quad \vartheta \in (0, 1),$$

aus der man die Behauptung ablesen kann: Wegen der Stetigkeit von $f^{(n)}$ stimmt das Vorzeichen von $f^{(n)}(x_0 + \vartheta h)$ für $|h|$ hinreichend klein mit dem Vorzeichen von $f^{(n)}(x_0) \neq 0$ überein. Ist n ungerade, ändert das Restglied sein Vorzeichen, wenn man h durch $-h$ ersetzt. Es kann an der Stelle x_0 daher kein lokales Extremum vorliegen. Ist hingegen n gerade, dann wird das Vorzeichen des Restgliedes in einer hinreichend kleinen Umgebung von x_0 durch das Vorzeichen von $f^{(n)}(x_0)$ bestimmt. Auf dieser Umgebung gilt dann

$$f(x_0 + h) \geq f(x_0) \qquad \text{falls } f^{(n)}(x_0) > 0$$

bzw.

$$f(x_0 + h) \leq f(x_0) \qquad \text{falls } f^{(n)}(x_0) < 0,$$

was zu beweisen war. □

Die Restgliedabschätzung in Beispiel VI-7.5 läßt vermuten, daß der Approximationsfehler mit wachsendem Grad n des Taylorpolynoms auf einer immer größer werdenden Umgebung von x_0 immer kleiner wird. Ist $f \in C^\infty$, ist es naheliegend, die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n$$

zu betrachten.

Definition VI-7.8. *Es sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f \in C^\infty(I, \mathbb{R})$. Die Potenzreihe*

$$\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n$$

heißt ***Taylorreihe*** *von f um x_0. (Im Falle $x_0 = 0$ ist auch die Bezeichnung* ***Maclaurin Reihe*** *üblich.)*

Die Taylorreihe kann zwar für jedes $f \in C^\infty$ angeschrieben werden, es ergeben sich aber sofort zwei Fragen:

- Für welche x konvergiert die Taylorreihe?
- Stimmt die Summenfunktion der Taylorreihe mit f überein?

Da die Taylorreihe eine Potenzreihe ist, kann die Konvergenzfrage mit den Methoden aus Abschnitt IV-11 vollständig beantwortet werden. Die zweite Frage wird durch folgenden Satz geklärt.

Theorem VI-7.9. *Es sei $I \subset \mathbb{R}$ ein Intervall, $J \subset I$ ein Teilintervall, $x_0 \in J$ und $f \in C^\infty(I,\mathbb{R})$. f wird genau dann auf J durch seine Taylorreihe um x_0 dargestellt, d.h. es gilt*

$$\forall x \in J\colon f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n,$$

wenn für das Restglied $R_n(x)$ in der Taylorschen Formel für alle $x \in J$

$$\lim_{n\to\infty} R_n(x) = 0$$

gilt. Insbesondere ist dies der Fall, wenn es positive Konstanten α und M gibt mit

$$\forall x \in J\, \forall n \in \mathbb{N}_0\colon |f^{(n)}(x)| \leq \alpha M^n.$$

Beweis. Bezeichnen wir mit $T_n(x)$ das n-te Taylorpolynom von f an der Stelle x_0, dann folgt aus Korollar VI-7.4

$$R_n(x) = f(x) - T_n(x).$$

Gilt also $\lim_{n\to\infty} R_n(x) = 0$, folgt

$$0 = f(x) - \lim_{n\to\infty} T_n(x) = f(x) - \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n,$$

d.h. f wird durch seine Taylorreihe dargestellt und umgekehrt bedeutet dieses Faktum $0 = f(x) - \lim_{n\to\infty} T_n(x) = \lim_{n\to\infty}(f(x) - T_n(x)) = \lim_{n\to\infty} R_n(x)$. Genügt $f^{(n)}$ der angegebenen Abschätzung, ergibt sich

$$|R_n^L(x)| \leq \alpha \frac{(M|x-x_0|)^{n+1}}{(n+1)!}.$$

Die rechte Seite ist das $(n+1)$-te Glied der Exponentialreihe für $\alpha e^{M|x-x_0|}$ und strebt daher für $n \to \infty$ nach 0. □

Dieser Satz ist nicht trivial: Es gibt C^∞-Funktionen, die nicht durch ihre Taylorreihe dargestellt werden:

Beispiel VI-7.10. Wir betrachten die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} 0 & x \leq 0 \\ e^{-\frac{1}{x}} & x > 0. \end{cases}$$

Da $x \mapsto \frac{1}{x}$ und $x \mapsto e^x$ beide C^∞-Funktionen auf $\mathbb{R}_+ \setminus \{0\}$ sind, folgt mit Satz V-2.2 $f \in C^\infty(\mathbb{R} \setminus \{0\}, \mathbb{R})$. f besitzt aber auch an der Stelle $x = 0$ Ableitungen beliebig hoher Ordnung: Für $x > 0$ gilt nämlich

$$f'(x) = \frac{1}{x^2} e^{-\frac{1}{x}}.$$

Mit Hilfe der Abschätzung

$$(*) \qquad e^{-\frac{1}{x}} \leq k!x^k \quad \text{für } x > 0 \text{ und beliebiges } k \in \mathbb{N}$$

folgt

$$\lim_{x\to 0^+} f'(x) = 0 \text{ und } \lim_{x\to 0^+} f(x) = 0 = f(0),$$

daher existiert $f'_+(0)$ und es gilt $f'_+(0) = f'_-(0) = 0$. f ist also in $x_0 = 0$ differenzierbar und es gilt $f'(0) = 0$. Induktiv zeigen wir nun: Es gibt Polynome p_n, $\operatorname{grad} p_n = 2n$, sodaß

$$f^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Die Behauptung stimmt für $n = 1$. $f^{(n)}$ besitze nun die angegebene Form. Für $x > 0$ ist $f^{(n)}$ differenzierbar mit

$$f^{(n+1)}(x) = \left(\frac{d}{dx}p_n(\frac{1}{x}) + p_n(\frac{1}{x})\frac{1}{x^2}\right) e^{-\frac{1}{x}}.$$

$\frac{d}{dx}p_n(\frac{1}{x})$ ist ein Polynom in $\frac{1}{x}$ vom Grade $2n+1$. Aus $\operatorname{grad}(p_n(\frac{1}{x})\frac{1}{x^2}) = 2n+2$ folgt $\operatorname{grad}\left(\frac{d}{dx}p_n(\frac{1}{x}) + p_n(\frac{1}{x})\frac{1}{x^2}\right) = 2(n+1)$. Dies beweist die Struktur von $f^{(n)}(x)$ für $x > 0$. Mit $(*)$ für $k > 2n$ folgt aber

$$\lim_{x\to 0^+} f^{(n)}(x) = 0,$$

und daraus wie bei f' auch $f^{(n)}(0) = 0$. Alle Koeffizienten der Taylorreihe von f um 0 sind daher Null. Trivialerweise konvergiert diese Reihe auf ganz $\mathbb{R}$. Da $f(x)$ für $x > 0$ positiv ist, kann die Taylorreihe f an keiner Stelle $x > 0$ darstellen.

Beispiel VI-7.11. Es sei f die Summenfunktion einer Potenzreihe $\sum\limits_{k=0}^{\infty} a_k(x - x_0)^k$ mit Konvergenzradius $R > 0$. Nach Korollar V-6.4 ist $f \in C^\infty((x_0 - R, x_0 + R), \mathbb{R})$ und

$$\forall k \in \mathbb{N}_0 \colon f^{(k)}(x_0) = k!a_k.$$

Die Taylorreihe einer durch eine Potenzreihe mit Entwicklungszentrum x_0 dargestellen Funktion um x_0 stimmt demnach mit der Potenzreihe überein. Dies ist natürlich nicht mehr der Fall, wenn man die Taylorreihe um eine andere Stelle als x_0 betrachtet.

Beispiel VI-7.12. Wir führen nun die Diskussion der Taylorreihe der Logarithmusfunktion $\ln(1+x)$, $x > -1$, (vgl. Beispiel VI-7.5)

$$\sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k},$$

zu Ende. Für $x > 0$ ergibt sich für das Restglied nach Lagrange

$$R_n^L(x) = (-1)^n \frac{1}{(1+\vartheta x)^{n+1}} \cdot \frac{x^{n+1}}{n+1},\ \vartheta \in (0,1),\ n \in \mathbb{N},$$

wegen $1 + \vartheta x > 1$ die Abschätzung

$$|R_n^L(x)| < \frac{x^{n+1}}{n+1}.$$

Für $x \in [0,1]$ gilt daher

$$\lim_{n\to\infty} R_n^L(x) = 0.$$

Für $x < 0$ könnte $1 + \vartheta x$ beliebig klein sein. Daher ist es vorteilhaft, in diesem Falle die Cauchysche Form des Restgliedes zu verwenden. Aus der Abschätzung $(*)$ in Beispiel VI-7.5 erhält man unmittelbar für $|x| < 1$

$$\lim_{n\to\infty} R_n^C(x) = 0.$$

Nach Satz VI-7.9 ist die Darstellung

$$(*) \qquad \ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}$$

für alle $x \in (-1,1]$ gültig. Für $x = -1$ ist $\ln(1+x)$ nicht definiert, entsprechend divergiert die Taylorreihe. Für $|x| > 1$ ist die Taylorreihe divergent.

Beispiel VI-7.13 (Binomialreihe). Als weiteres Beispiel betrachten wir die Taylorreihe von $f(x) = (1+x)^\alpha$, $\alpha \in \mathbb{R}$, def $f = (-1,\infty)$, in $x_0 = 0$,

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,$$

vgl. Beispiel VI-7.6. Das Restglied nach Cauchy ist für alle $x > -1$ gegeben durch

$$\begin{aligned} R_n^C(x) &= \frac{1}{n!} \prod_{j=0}^{n} (\alpha - j) \left(\frac{1-\vartheta}{1+\vartheta x}\right)^n (1+\vartheta x)^{\alpha-1} x^{n+1} \\ &= \alpha x \left(\frac{1-\vartheta}{1+\vartheta x}\right)^n (1+\vartheta x)^{\alpha-1} \prod_{j=1}^{n} [(\frac{\alpha}{j} - 1)x]. \end{aligned}$$

mit $\vartheta = \vartheta(n,x) \in (0,1)$. Wir fixieren nun x, β so, daß

$$0 < |x| < \beta < 1$$

gilt und wählen $N(\beta) \in \mathbb{N}$ derart, daß für alle $j > N(\beta)$

$$|(\frac{\alpha}{j} - 1)x| < \beta$$

zutrifft. Wir können nun das mögliche Anwachsen von $|\prod_{j=0}^{n}(\alpha - j)|$ folgendermaßen ausgleichen: für $n > N(\beta)$ gilt nämlich

$$\begin{aligned}|\prod_{j=1}^{n}(\frac{\alpha}{j} - 1)x| &= |\prod_{j=1}^{N(\beta)}(\frac{\alpha}{j} - 1)x| \cdot |\prod_{j=N(\beta)+1}^{n}(\frac{\alpha}{j} - 1)x| \\ &\leq (|\alpha| + 1)^{N(\beta)}\beta^{N(\beta)} \cdot \beta^{n-N(\beta)} = (|\alpha| + 1)^{N(\beta)}\beta^{n}.\end{aligned}$$

Zusammen mit den (bezüglich n gleichmäßigen) Abschätzungen

$$|\frac{1-\vartheta}{1+\vartheta x}| < 1$$

und

$$0 < (1 + \vartheta x)^{\alpha-1} \leq M := \begin{cases} 2^{\alpha-1} & \text{falls } \alpha \geq 1 \\ (1-\beta)^{\alpha-1} & \text{falls } \alpha < 1, \end{cases}$$

erhalten wir endlich

$$|R_n^C(x)| \leq M|\alpha|(|\alpha| + 1)^{N(\beta)} \cdot \beta^{n+1}$$

also $\lim_{n\to\infty} R_n^C(x) = 0$ für $|x| < 1$. Nach Satz VI-7.9 gilt daher die Darstellung

$$(1 + x)^{\alpha} = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad (\textsc{Binomialreihe})$$

für alle $x \in (-1, 1)$. Als Spezialfall erhält man für $\alpha \in \mathbb{N}$ den binomischen Lehrsatz. Die Untersuchung des Konvergenzverhaltens der Binomialreihe in den Randpunkten des Konvergenzintervalles erfordert Hilfsmittel, die wir nicht entwickelt haben.

Wir betonen noch einmal: Die Bestimmung des Konvergenzradius einer Taylorreihe gibt allgemein keinen Aufschluß darüber, ob eine Funktion durch ihre Taylorreihe tatsächlich dargestellt wird. Die meist erheblich kompliziertere Diskussion des Restgliedes erledigt beide Fragen gleichzeitig. Oft kann diese Diskussion nur auf einer echten Teilmenge J des Konvergenzintervalles K der Taylorreihe durchgeführt werden und es bleibt die Frage offen, ob die Taylorreihe auch auf $K \setminus J$ die Funktion f darstellt. Die Entscheidung, ob dies tatsächlich der Fall ist, erfordert tiefer reichende Methoden der komplexen Analysis.
Die komplizierte Restglieddiskussion kann man manchmal durch eine andere Methode ersetzen:

Beispiel VI-7.14. Wir betrachten wieder die Taylorreihe von $f(x) = \ln(1 + x)$ um $x_0 = 0$, also die Reihe

$$g(x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}$$

und weisen die Gültigkeit von $f(x) = g(x)$ für $|x| < 1$ nach. Die Reihe konvergiert für $|x| < 1$, somit ist $g \in C^\infty((-1,1),\mathbb{R})$ und es ist (vgl. Korollar V-6.4)

$$g'(x) = \sum_{n=1}^{\infty}(-1)^{n-1}x^{n-1} = \sum_{k=0}^{\infty}(-1)^k x^k = \frac{1}{1+x},\ |x| < 1.$$

Für $|x| < 1$ gilt also

$$f'(x) - g'(x) = 0,$$

nach Korollar V-4.7 gibt es also eine Konstante α mit

$$\forall x \in (-1,1)\colon f(x) - g(x) = \alpha.$$

Aus $f(0) = g(0) = 0$ folgt $\alpha = 0$ und damit auch die Darstellung

$$\ln(1+x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n} \quad \text{für } x \in (-1,1).$$

Wir demonstrieren diese Methode noch einmal am Beispiel der Binomialreihe:

Beispiel VI-7.15. Die Taylorreihe von $f(x) = (x+1)^\alpha$, $\alpha \in \mathbb{R}$, $x > -1$,

$$g(x) = \sum_{n=0}^{\infty}\binom{\alpha}{n}x^n,$$

konvergiert für $|x| < 1$ und kann dort gliedweise differenziert werden:

$$g'(x) = \sum_{n=1}^{\infty}\binom{\alpha}{n}nx^{n-1} = \alpha\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n$$

(wir verwenden $n\binom{\alpha}{n} = \alpha\binom{\alpha-1}{n-1}$). Multipliziert man mit $(1+x)$, erhält man

$$\begin{aligned}(1+x)g'(x) &= \alpha\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^n + \alpha\sum_{n=0}^{\infty}\binom{\alpha-1}{n}x^{n+1}\\ &= \alpha + \alpha\sum_{n=1}^{\infty}\left[\binom{\alpha-1}{n} + \binom{\alpha-1}{n-1}\right]x^n = \alpha\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n\\ &= \alpha g(x).\end{aligned}$$

(vgl. II-I-4.21). Wir betrachten nun auf $(-1,1)$ die differenzierbare Abbildung $h = \frac{g}{f}$ (man beachte $f(x) \neq 0$ für $x > -1$). Es folgt

$$h'(x) = \frac{g'f - gf'}{f^2}(x) = \frac{(x+1)^\alpha\alpha(1+x)^{-1}g - \alpha(x+1)^{\alpha-1}g}{f(x)^2} = 0,$$

also ist h konstant. Aus $h(0) = 1$ folgt dann $f = g$, d.h.

$$(1+x)^\alpha = \sum_{n=0}^{\infty}\binom{\alpha}{n}x^n \quad \text{für } |x| < 1.$$

Wir beenden diesen Abschnitt mit dem Beispiel einer C^∞-Funktion, deren Taylorreihe nur an der Entwicklungsstelle konvergiert:

Beispiel VI-7.16. Die Funktion f sei definiert durch

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cos n^2 x, \quad x \in \mathbb{R}.$$

Diese Reihe konvergiert gleichmäßig auf $\mathbb{R}$, somit ist f stetig auf $\mathbb{R}$. Die Reihe der Ableitungen

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

ist ebenfalls gleichmäßig konvergent auf $\mathbb{R}$, da die Majorante $\sum \frac{n^2}{2^n}$ konvergiert. Nach Korollar V-6.3 ist f differenzierbar und es gilt

$$f'(x) = -\sum_{n=1}^{\infty} \frac{n^2}{2^n} \sin n^2 x.$$

Ähnlich zeigt man, daß auch sämtliche höheren Ableitungen existieren,

$$f^{(2k)}(x) = (-1)^k \sum_{n=1}^{\infty} \frac{n^{4k}}{2^n} \cos n^2 x,$$
$$f^{(2k+1)}(x) = (-1)^{k+1} \sum_{n=1}^{\infty} \frac{n^{4k+2}}{2^n} \sin n^2 x, \quad k \in \mathbb{N}_0.$$

Wegen $f^{(2k+1)}(0) = 0$, $k \in \mathbb{N}_0$, lautet die Taylorreihe von f um $x_0 = 0$

$$\sum_{k=0}^{\infty} \frac{1}{(2k)!} f^{(2k)}(0) x^{2k}.$$

Für alle $n, k \in \mathbb{N}$ gilt

$$|f^{(2k)}(0)| = \sum_{j=1}^{\infty} \frac{j^{4k}}{2^j} > \frac{n^{4k}}{2^n}$$

und somit

$$\left|\frac{1}{(2k)!} f^{(2k)}(0) x^{2k}\right| > \frac{(n^2|x|)^{2k}}{2^n (2k)!}.$$

Speziell für $n = 2k$ ergibt sich die Abschätzung

$$\left|\frac{1}{(2k)!} f^{(2k)}(0) x^{2k}\right| > \frac{1}{2^{2k}} (2k|x|)^{2k} \frac{(2k)^{2k}}{(2k)!} \geq (k|x|)^{2k},$$

sodaß die Taylorreihe von f für $x \neq 0$ nicht konvergieren kann.

8. Konvexe Funktionen

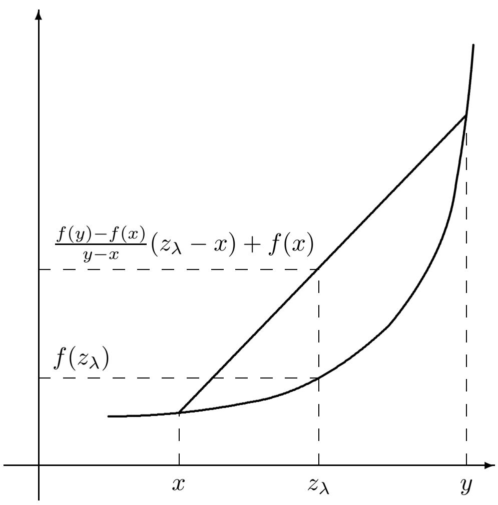
Konvexe Funktionen sind u.a. in der Optimierung von großer Bedeutung. Außerdem verknüpfen sie die zweite Ableitung mindestens zweimal differenzierbarer Funktionen mit einer globalen Eigenschaft dieser Funktion, ähnlich, wie die erste Ableitung einer Funktion Monotonieeigenschaften zum Ausdruck bringt.

Definition VI-8.1. *Es sei $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$.*

i) *f heißt **konvex** (auf I) $\underset{Def}{\Leftrightarrow} \forall x, y \in I$, $\forall \lambda \in [0,1]$:*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

ii) *f heißt **konkav** (auf I) $\underset{Def}{\Leftrightarrow} -f$ ist konvex auf I.*



Bemerkung VI-8.2. Die Konvexitätsbedingung besitzt eine einfache geometrische Veranschaulichung: Es sei $z_\lambda = \lambda x + (1-\lambda)y$, x, y, λ wie in Definition VI-8.1 und $x < y$. Man verifiziert leicht $z_\lambda \in [x, y]$. Umgekehrt läßt sich jeder Zwischenpunkt $z \in [x, y]$ in der Form

$$z = \underbrace{\frac{y-z}{y-x}}_{\lambda} x + \underbrace{\frac{z-x}{y-x}}_{1-\lambda} y = \lambda x + (1-\lambda) y$$

darstellen. Die Konvexitätsbedingung ist daher gleichwertig mit

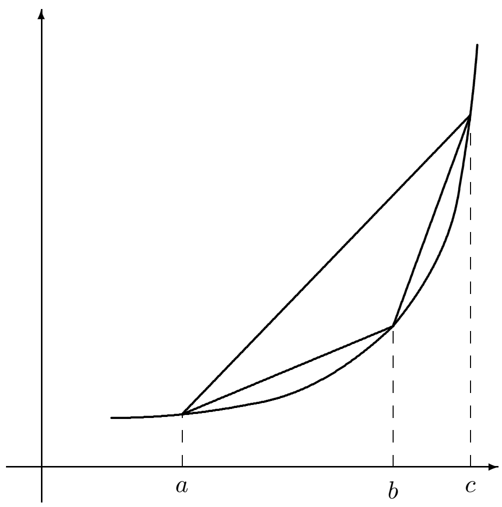
$$\begin{aligned} f(z_\lambda) &\leq \frac{y - z_\lambda}{y-x} f(x) + \frac{z_\lambda - x}{y-x} f(y) \\ &= \frac{f(y) - f(x)}{y-x}(z_\lambda - x) + f(x), \end{aligned}$$

d.h. f ist konvex genau dann, wenn der Graph von f stets unterhalb der Verbindungsstrecke zweier beliebiger Punkte $(x, f(x))$ und $(y, f(y))$ liegt.

Konvexität einer Funktion hat bereits überraschende Differenzierbarkeitseigenschaften zur Folge:

LEMMA VI-8.3. *Es sei $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$ konvex. Für alle $a, b, c \in I$ mit $a < b < c$ gilt*

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(a)}{c-a} \leq \frac{f(c)-f(b)}{c-b}.$$



BEWEIS. Mit $\lambda = \frac{b-a}{c-a} \in (0,1)$ gilt

$$(*) \qquad b = \frac{c-b}{c-a}a + \frac{b-a}{c-a}c = (1-\lambda)a + \lambda c.$$

Da f konvex ist, folgt

$$\begin{aligned} f(b) &\leq (1-\lambda)f(a) + \lambda f(c) \\ &= \lambda(f(c) - f(a)) + f(a). \end{aligned}$$

Dies impliziert

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(a)}{c-a}.$$

Ersetzt man λ in $(*)$ durch $\mu = 1 - \lambda = \frac{c-b}{c-a}$, erhält man

$$f(b) \leq \mu f(a) + (1-\mu) f(c)$$

und daraus folgt wie vorhin die zweite Ungleichung. □

Im Folgenden benötigen wir eine wichtige Eigenschaft monotoner Funktionen:

THEOREM VI-8.4. *Es sei $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$ monoton. Dann ist f stetig auf I mit Ausnahme von höchstens abzählbar vielen Sprungstellen.*

BEWEIS. Nach Satz V-**??** existieren an jeder Stelle x des Definitionsbereiches die einseitigen Grenzwerte $f(x^+)$ und $f(x^-)$. Wegen der Monotonie kann f keine hebbare Unstetigkeit aufweisen. Eine monotone Funktion ist also in $x \in I$ genau dann unstetig, wenn in x eine Sprungstelle vorliegt. Wir nehmen nun an, f sei monoton wachsend und betrachten in jeder Sprungstelle x das Intervall $S_x := [f(x^-), f(x^+)]$. Da f monoton wächst, folgt $S_x \cap S_y = \emptyset$ falls $x < y$. In einer Sprungstelle gilt $f(x^-) < f(x^+)$, also gibt es in S_x eine rationale Zahl r_x. Ordnet man jeder Sprungstelle x genau eine derartige rationale Zahl r_x zu, ergibt sich die Abzählbarkeit der Menge der Unstetigkeitstellen von f. □

THEOREM VI-8.5. *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ konvex. Dann gilt*

i) *f ist rechts- und linksseitig differenzierbar.*
ii) *f'_+ und f'_- sind monoton wachsend auf I.*
iii) $\forall x \in I : f'_-(x) \le f'_+(x)$.
iv) *f ist differenzierbar ausgenommen in höchstens abzählbar vielen Punkten von I.*
v) *$f|_{[\alpha,\beta]}$ ist Lipschitz stetig, d.h.*

$$\forall x, y \in [\alpha, \beta]\colon |f(x) - f(y)| \le M|x - y|,$$

wobei $[\alpha, \beta] \subset I$ und $M := \max\{|f'_+(\alpha)|, |f'_-(\beta)|\}$ gesetzt wurde.

vi) *f ist stetig.*

BEWEIS. Wir wählen $x_0 \in I$ beliebig und bestimmen $h_0 > 0$ so, daß $[x_0 - h_0, x_0 + h_0] \subset I$ (dies ist möglich, da I ein offenes Intervall ist). Die Abbildung $\varphi\colon [-h_0, h_0] \setminus \{0\} \to \mathbb{R}$, definiert durch

$$h \to \varphi(h) = \frac{1}{h}(f(x_0 + h) - f(x_0)),$$

ist nach Lemma VI-8.3 monoton wachsend und beschränkt: Für $0 < k < h < h_0$ gilt nämlich

$$\begin{aligned} \frac{f(x_0 - h_0) - f(x_0)}{-h_0} &\le \frac{f(x_0 - h) - f(x_0)}{-h} \le \frac{f(x_0 - k) - f(x_0)}{-k} \\ &\le \frac{f(x_0 + k) - f(x_0)}{k} \le \frac{f(x_0 + h) - f(x_0)}{h} \le \frac{f(x_0 + h_0) - f(x_0)}{h_0}, \end{aligned}$$

d.h.

$$\varphi(-h_0) \le \varphi(-h) \le \varphi(-k) \le \varphi(k) \le \varphi(h) \le \varphi(h_0).$$

Aus Satz V-**??** folgt daher die Existenz der einseitigen Grenzwerte $\varphi(0^+) = \inf\{\varphi(h)\colon 0 < h \le h_0\}$ und $\varphi(0^-) = \sup\{\varphi(h)\colon -h_0 \le h < 0\}$. Insbesondere gilt

$$(*)\qquad f'_-(x_0) = \varphi(0^-) = \lim_{h\to 0^-} \varphi(h) = \sup\{\varphi(h)\colon h<0\}$$
$$\le \inf\{\varphi(h)\colon h>0\} = \lim_{h\to 0^+}\varphi(h) = \varphi(0^+) = f'_+(x_0).$$

Dies beweist i) und iii).
Wir zeigen nun die Monotonie von f'_+: Es sei $a, b \in I$ und $a < b$. Für alle $c > b$, $c \in I$, und $x_0 = a$ folgt

$$(\dagger)\qquad f'_+(a) = \inf\{\varphi(h)\colon 0 < h \le h_0\} \le \frac{f(b)-f(a)}{b-a} \overset{\text{VI-8.3}}{\le} \frac{f(c)-f(b)}{c-b}.$$

Durch Grenzübergang $c \downarrow b$ ergibt sich $f'_+(a) \le f'_+(b)$. Analog zeigt man die Monotonie von f'_-.
Es bezeichne $\mathcal{S}$ die Menge der Sprungstellen von f'_+ und f'_-. Nach Satz VI-8.4 sind f'_+ und f'_- stetig auf $I\setminus\mathcal{S}$. Es sei nun $a \in I\setminus\mathcal{S}$, $b, c \in I$ und $a < b < c$. Führt man in Lemma VI-8.3 einerseits den Grenzübergang $b \downarrow a$ und andererseits den Grenzübergang $b \uparrow c$ durch, erhält man für alle $c > a$

$$f'_+(a) \le \frac{f(c)-f(a)}{c-a} \le f'_-(c),$$

und da f'_- in a stetig ist, ergibt sich

$$f'_-(a) \le f'_+(a) \le \lim_{c\downarrow a} f'_-(c) = f'_-(a),$$

d.h.

$$f'_-(a) = f'_+(a).$$

Somit ist iv) bewiesen.
Es sei nun $[\alpha,\beta] \subset I$, $M = \max\{|f'_+(\alpha)|, |f'_-(\beta)|\}$ und $\alpha \le x < y \le \beta$. Es folgt

$$-M \le f'_+(\alpha) \overset{(ii)}{\le} f'_+(x) \overset{(\dagger)}{\le} \frac{f(y)-f(x)}{y-x} \le f'_-(y) \le f'_-(\beta) \le M.$$

Diese Ungleichung zeigt v) und daraus folgt die Stetigkeit von f auf dem offenen Intervall I. □

Das Beispiel $f\colon [0,1) \to \mathbb{R}^-$, $x \mapsto -\sqrt{x}$ zeigt, daß einige Behauptungen des Satzes nicht zutreffen, wenn als Definitionsbereich von f ein halboffenes (oder abgeschlossenes) Intervall zugelassen wird.
Aus Satz VI-8.5 ergibt sich eine nützliche Charakterisierung konvexer differenzierbarer Funktionen:

Theorem VI-8.6. *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ differenzierbar auf I. Gleichwertig sind*

(1) *f ist konvex,*
(2) *f' ist monoton wachsend.*

Beweis. (1) $\Rightarrow$ (2) Dies folgt aus Satz VI-8.5-ii) und $f'(x) = f'_+(x) = f'_-(x)$, $x \in I$.
(2) $\Rightarrow$ (1) Wir zeigen diese Implikation durch Kontraposition und nehmen an, f sei nicht konvex. Nach Bemerkung VI-8.2 ist dies gleichwertig mit der Existenz von $a, b, x \in I$ und $a < x < b$ und

$$f(x) > \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b).$$

Daraus folgt mit $1 = \frac{b-x}{b-a} + \frac{x-a}{b-a}$

$$(f(x) - f(a)) \frac{b-x}{b-a} > (f(b) - f(x)) \frac{x-a}{b-a}$$

bzw.

$$\frac{f(x) - f(a)}{x-a} > \frac{f(b) - f(x)}{b-x}.$$

Wir wenden nun den Mittelwertsatz auf jedes der Intervalle $[a, x]$ und $[x, b]$ an und schließen auf die Existenz von Zwischenstellen $\xi \in (a, x)$ und $\eta \in (x, b)$, also $\xi < \eta$, mit

$$f'(\xi) > f'(\eta),$$

d.h. f' ist nicht monoton wachsend. □

Korollar VI-8.7. Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ zweimal differenzierbar auf I. Gleichwertig sind

(1) f ist konvex,
(2) $\forall x \in I\colon f''(x) \geq 0$.

Mit Hilfe dieses Kriteriums ergibt sich nun leicht, daß z.B. die Abbildungen

$$\begin{aligned} &x \mapsto e^x, \quad x \mapsto e^{-x}, \quad x \in \mathbb{R}, \\ &x \mapsto x^\alpha, \quad \alpha \geq 1, \qquad x \in \mathbb{R}^+, \end{aligned}$$

konvex sind, und die Abbildung

$$x \mapsto \ln x, \qquad x > 0,$$

konkav ist.

Wir können nun auch die geometrische Anschauung untermauern, daß der Graph einer konvexen Funktion immer oberhalb der Tangente liegt:

Theorem VI-8.8. *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ konvex. Ferner sei $x_0 \in I$ und $m \in [f'_-(x_0), f'_+(x_0)]$. Dann gilt*

$$f(x) \geq m(x - x_0) + f(x_0)$$

für alle $x \in I$.

Den Graph der affinlinearen Funktion auf der rechten Seite in der letzten Ungleichung nennt man **Stützgerade** von f. Ist f an der Stelle x_0 differenzierbar, ist die Stützgerade eindeutig bestimmt und fällt mit der Tangente in x_0 an den Graph von f zusammen.

BEWEIS. Aus Lemma VI-8.3 folgt für $x > x_0$

$$\frac{f(x) - f(x_0)}{x - x_0} \geq f'_+(x_0) \geq m,$$

und für $x < x_0$

$$\frac{f(x) - f(x_0)}{x - x_0} \leq f'_-(x_0) \leq m.$$

□

THEOREM VI-8.9 (Jensensche Ungleichung). *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ konvex. Ferner sei $\{x_j : j = 1, \ldots, n\} \subset I$, $\{\lambda_j : j = 1, \ldots, n\} \subset [0, 1]$ und $\sum\limits_{j=1}^{n} \lambda_j = 1$. Dann gilt*

$$f(\sum_{j=1}^{n} \lambda_j x_j) \leq \sum_{j=1}^{n} \lambda_j f(x_j).$$

BEWEIS. Wir nehmen an, es gelte

$$x_1 \leq x_2 \leq \cdots \leq x_n$$

und setzen $x_0 = \sum\limits_{i=1}^{n} \lambda_i x_i$. Es gilt $x_1 \leq x_0 \leq x_n$. Es sei m wie in Satz VI-8.8; mit $x = x_0$, folgt aus Satz VI-8.8

$$(*) \qquad f(x_j) \geq m(x_j - x_0) + f(x_0), \quad j = 1, \ldots, n.$$

Multipliziert man diese Ungleichungen mit λ_j und addiert, erhält man die gewünschte Ungleichung

$$\sum_{j=1}^{n} \lambda_j f(x_j) \geq m \underbrace{\sum_{j=1}^{n} \lambda_j x_j}_{=x_0} - m x_0 + f(x_0) = f(x_0).$$

□

Wählt man z.B. $f(x) = e^x$, erhält man die Ungleichung

$$\exp(\sum_{i=1}^{n} \lambda_i x_i) \leq \sum_{i=1}^{n} \lambda_i e^{x_i}.$$

Setzt man $y_i = e^{x_i}$, ergibt sich

$$\prod_{i=1}^{n} y_i^{\lambda_i} \leq \sum_{i=1}^{n} \lambda_i y_i$$

Für $\lambda_i = \frac{1}{n}$, $i = 1, \dots, n$, erhält man daraus die Ungleichung zwischen arithmetischem und geometrischem Mittel:

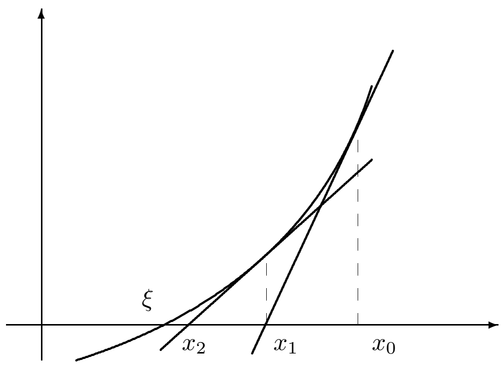
$$\sqrt[n]{y_1 \cdots y_n} \leq \frac{1}{n} \sum_{i=1}^{n} y_i.$$

9. Das Newton Verfahren

Sehr viele Problemstellungen führen letztendlich auf die Aufgabe, eine Nullstelle einer nichtlinearen Funktion f zu bestimmen, also die Gleichung

$$f(x) = 0$$

nach x aufzulösen. Wir wenden uns dem einfachsten Fall, $I \subset \mathbb{R}$ und $f \in C^1(I, \mathbb{R})$, zu. Die Idee des Newton Verfahrens ist denkbar einfach: Gesucht ist der Schnittpunkt ξ des Graphen $G(f)$ mit der x-Achse. Kennt man bereits irgendeinen Näherungswert x_0 für ξ, ersetzt man f lokal durch die Tangente im Punkt $(x_0, f(x_0))$ und bestimmt den Schnittpunkt x_1 der Tangente mit der x-Achse. Unter bestimmten Voraussetzungen liegt x_1 näher bei ξ als x_0 und Iteration dieses Verfahrens liefert eine Folge (x_n), die gegen die gesuchte Nullstelle ξ konvergiert.



Newton Verfahren

Die Tangente an den Graph von f im Punkt $(x_0, f(x_0))$ besitzt die Gleichung $T_1(x) = f(x_0) + f'(x_0)(x - x_0)$ und schneidet daher die x-Achse an der Stelle

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Eine analoge Formel gilt auch für den allgemeinen Schritt $x_n \to x_{n+1}$. Das Newton Verfahren wird also durch folgende Iteration beschrieben:

Algorithmus:
Schritt 0: Wahl des Startwertes x_0, hinreichend nahe bei ξ,

Schritt n: Korrektur von x_n:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Man beendet die Iteration, wenn eine bestimmte Abbruchbedingung erfüllt ist, etwa wenn für eine vorgegebene Toleranz tol

$$|x_{n+1} - x_n| < \text{tol}$$

erreicht wird.

Dieses Verfahren bzw. Varianten davon funktioniert so gut und ist so effizient, daß es zum Standardwerkzeug der numerischen Mathematik zählt. Wir geben nun Bedingungen an, die sicherstellen, daß der Startwert x_0 hinreichend nahe bei ξ liegt, sodaß die Iterationsfolge (x_n) tatsächlich gegen die gesuchte Nullstelle ξ konvergiert. In der Praxis wendet man das Newton Verfahren jedoch oft an, ohne die Voraussetzungen zu verifizieren.

Theorem VI-9.1. *Es sei $I = [a, b]$ und $f \in C^2([a, b], \mathbb{R})$. Es gelte*

i) $f(a)f(b) < 0$,
ii) $\forall x \in I\colon f'(x) \neq 0$,
iii) *f ist konvex oder konkav,*
iv) *die Iterationswerte x_1 zu $x_0 = a$ und zu $x_0 = b$ liegen in $[a, b]$.*

Dann gilt

a) *f besitzt in (a, b) genau eine Nullstelle ξ.*
b) *Für jeden Startwert $x_0 \in [a, b]$ ist die Folge der Iterationen $(x_n)_{n \geq 1}$ eine Teilmenge von I und konvergiert monoton gegen ξ.*
c) *Die Konvergenz ist quadratisch, d.h. es gibt eine Konstante $M > 0$ mit*

$$\forall n \in \mathbb{N}\colon |x_n - \xi| \leq M|x_n - x_{n-1}|^2.$$

Beweis. Die Behauptung a) ist eine Folge des Zwischenwertsatzes und der strengen Monotonie von f. Wir nehmen nun an, f sei streng monoton wachsend und konvex, d.h.

$$\forall x \in I\colon f'(x) > 0 \quad \text{und} \quad f''(x) \geq 0$$

(der Beweis für die anderen drei Möglichkeiten verläuft analog). Wir sammeln die Eigenschaften der Iterationsvorschrift

$$\mathfrak{I}(x) := x - \frac{f(x)}{f'(x)}, \quad x \in I. \tag{1}$$

Es folgt

$$\mathfrak{I}'(x) = \frac{f(x)f''(x)}{f'^2(x)} \begin{cases} \leq 0, & x \in [a, \xi] \\ \geq 0, & x \in (\xi, b] \end{cases}.$$

Nach Korollar V-4.7 nimmt $\mathfrak{I}$ an der Stelle ξ das (globale) Minimum an, d.h.

$$\forall x \in I\colon \mathfrak{I}(x) \geq \mathfrak{I}(\xi) = \xi.$$

Wegen iv) gilt daher auch

$$\mathfrak{I}(I) \subset [\xi, b]. \tag{2}$$

Für $x \in [\xi, b]$ ist $f(x) \geq 0$, aus (1) folgt daher

$$\mathfrak{I}(x) \leq x \quad \text{für } x \in [\xi, b]. \tag{3}$$

Wir untersuchen nun die Iterationsfolge $(x_n)_{n\geq 0}$ mit

$$\forall n \in \mathbb{N}_0 \colon x_{n+1} = \mathfrak{I}(x_n)$$

zu einem beliebigen Startwert $x_0 \in [a, b]$. Wegen (2) gilt $x_1 \in [\xi, b]$ und wegen (3) auch $x_2 = \mathfrak{I}(x_1) \leq x_1$. Eine einfache Induktion zeigt nun

$$\forall n \in \mathbb{N} \colon \xi \leq x_{n+1} \leq x_n.$$

Die Iterationsfolge ist also ab dem Index 1 monoton fallend und nach unten beschränkt, somit konvergent nach Satz II-I-3.2. Es sei $\tau = \lim_{n\to\infty} x_n$. Führt man in der Iterationsvorschrift

$$x_{n+1} = \mathfrak{I}(x_n) \quad \text{d.h.} \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

den Grenzübergang $n \to \infty$ durch, erhält man

$$\tau = \tau - \frac{f(\tau)}{f'(\tau)},$$

also $f(\tau) = 0$. Wegen der Eindeutigkeit der Nullstelle von f muß $\tau = \xi$ gelten, womit b) bewiesen ist.
Nach dem MWS V-4.6 gilt für alle $x_n \neq \xi$

$$\left|\frac{f(x_n) - f(\xi)}{x_n - \xi}\right| \geq \min_{\tau \in I} |f'(\tau)| =: m > 0,$$

also

$$|x_n - \xi| \leq \frac{1}{m}|f(x_n)|. \tag{4}$$

Um $|f(x_n)|$ abzuschätzen, verwenden wir die Taylorsche Formel für f um x_{n-1}:

$$f(x_n) = f(x_{n-1}) + f'(x_{n-1})(x_n - x_{n-1}) + \frac{1}{2}f''(x_{n-1} + \theta(x_n - x_{n-1}))(x_n - x_{n-1})^2$$

für ein $\theta \in (0, 1)$. Berücksichtigt man die Rekursionsvorschrift in der Form

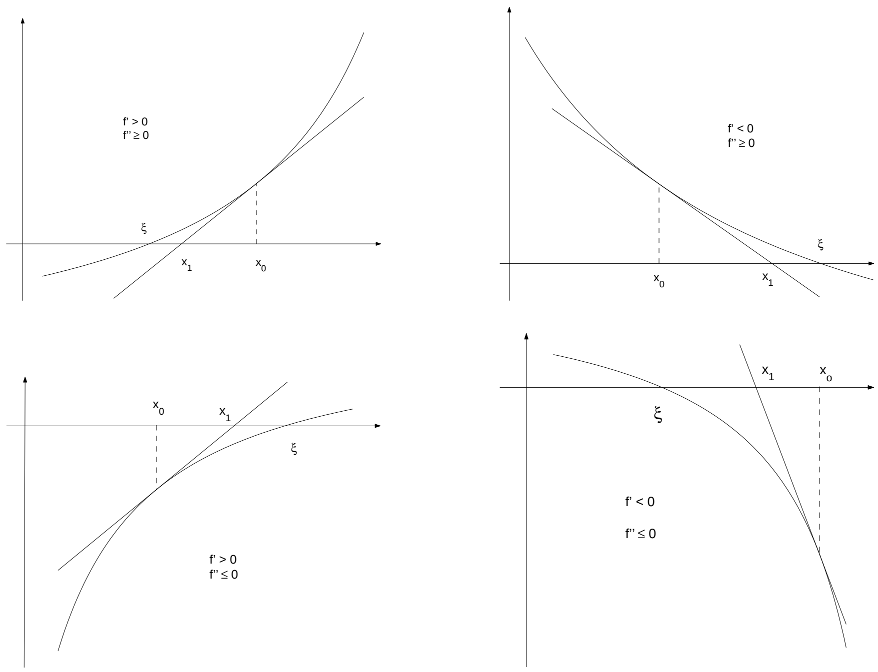
$$(x_n - x_{n-1})f'(x_{n-1}) = -f(x_{n-1})$$

erhält man

$$|f(x_n)| \leq \frac{1}{2} \max_{x \in I} |f''(x)|(x_n - x_{n-1})^2. \tag{5}$$

Kombiniert man (4) und (5), ergibt sich die gesuchte Fehlerabschätzung mit $M = \frac{1}{2m} \max_{x\in I} |f''(x)|$. $\square$

Wir erwähnen, daß die Voraussetzung iv) in der im Beweis vorliegenden Situation $f'(x) > 0$, $f''(x) \geq 0$, $x \in I$, für den Startwert $x_0 = b$ automatisch erfüllt ist (vgl. (1)). $x_0 = b$ ist daher eine gute Wahl für den Startwert falls weitere Informationen fehlen. In den folgenden Abbildungen deuten wir eine Heuristik für die Wahl des Startwertes an.



Beispiel VI-9.2. Gesucht ist die Lösung der transzendenten Gleichung

$$f(x) = e^{-x} - x = 0.$$

Wegen $f'(x) = -e^{-x} - 1 < 0$, $x \in \mathbb{R}$, ist f streng monoton fallend auf $\mathbb{R}$. Somit kann f höchstens eine Nullstelle besitzen. Aus $f(0) = 1$ und $f(1) = e^{-1} - 1 < 0$ ergibt sich, daß in $[0, 1]$ eine Nullstelle von f liegt. Ferner gilt $f''(x) = e^{-x} > 0$, also ist f konvex. In der folgenden Tabelle sind die Ergebnisse für die ersten 5 Iterationen zum Startwert $x_0 = 1$ zusammengestellt:

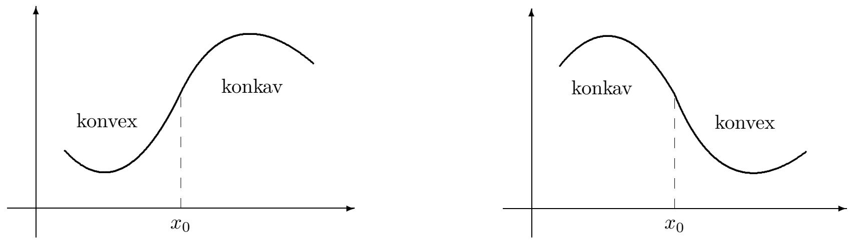
x_k	$f(x_k)$	$x_k - x_{k-1}$
1.00000000000000	-0.63212055882856	—
0.53788284273999	0.04610048629169	-0,46211715726001
0.56698699140541	0.00024494986384	0.02910414866542
0.56714328598912	0.00000000692781	0.00015629458371
0.56714329040978	0	0.00000000442066
0.56714329040978	0	0

10. Kurvendiskussion

Unsere bisher erzielten Ergebnisse ermöglichen es, weitgehende qualitative Aussagen über den Verlauf einer Funktion f und die Gestalt ihres Graphen zu treffen. Wir ergänzen die Diskussion mit einem weiteren Begriff:

Definition VI-10.1. *Es sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$.*

$$x_0 \in I \text{ heißt } \textbf{\textit{Wendepunkt}} \text{ von } f \underset{Def}{\Leftrightarrow}$$
$$\exists \alpha, \beta \in I, \alpha < x_0 < \beta \colon f|_{(\alpha,x_0]} \text{ ist konvex und } f|_{[x_0,\beta)} \text{ ist konkav oder umgekehrt.}$$



Für $f \in C^2(I, \mathbb{R})$ ist ein Wendepunkt x_0 dadurch charakterisiert, daß f'' in x_0 sein Vorzeichen wechselt. Es gilt also

$$\begin{aligned} f''(x) &\geq 0 \quad \text{für } x \in (\alpha, x_0), \\ f''(x) &\leq 0 \quad \text{für } x \in (x_0, \beta), \end{aligned}$$

oder umgekehrt. In einem Wendepunkt x_0 gilt also

$$f''(x_0) = 0.$$

Eine systematische Kurvendiskussion sollte folgende Aspekte behandeln:

- Bestimmung des maximalen Definitionsbereiches
- Symmetrieeigenschaften (gerade, ungerade)
- Nullstellen von f, f', f'': Dies ist in den meisten praktischen Fällen nur näherungsweise möglich, etwa mit dem Newton-Verfahren
- Monotonieeigenschaften und lokale Extrema
- Teilmengen von $\operatorname{def} f$, auf welchen f konvex bzw. konkav ist, Wendepunkte
- Verhalten von f in Randpunkten des Definitionsbereiches
- $\lim_{x\to\pm\infty} f(x)$ soferne die Limiten existieren, Asymptoten
- Skizze

Eine **Asymptote** ist eine Gerade, beschrieben durch die Gleichung $a(x) = kx + d$, $k, d \in \mathbb{R}$, mit der Eigenschaft

$$\lim_{x\to\infty} (f(x) - a(x)) = 0, \qquad \text{oder} \qquad \lim_{x\to-\infty} (f(x) - a(x)) = 0.$$

Wir bemerken, daß im allgemeinen das asymptotische Verhalten einer Funktion für $x \to \infty$ und $x \to -\infty$ getrennt zu untersuchen ist. Die Unbekannten k, d erhält man durch folgende Überlegung: Angenommen wir hätten bereits eine Asymptote bestimmt. Dann gilt auch

$$0 = \lim_{x\to\infty} \frac{f(x) - a(x)}{x} = \lim_{x\to\infty} (\frac{f(x)}{x} - k - \frac{d}{x}) = \lim_{x\to\infty} (\frac{f(x)}{x} - k).$$

Der Anstieg der Asymptote ist demnach festgelegt durch

$$k = \lim_{x\to\infty} \frac{f(x)}{x},$$

der Wert von $d = a(0)$ ergibt sich *anschließend* aus dem Grenzwert

$$d = \lim_{x\to\infty} (f(x) - kx).$$

Analoge Formeln gelten für das asymptotische Verhalten $x \to -\infty$.

Beispiel VI-10.2. Man diskutiere die Abbildung $x \mapsto f(x) = \sqrt{\frac{x^3}{x-1}}$.
Die Forderung $\frac{x^3}{x-1} \geq 0$ ergibt den maximalen Definitionsbereich $\operatorname{def} f = (-\infty, 0] \cup (1, \infty)$. Die Abbildung besitzt genau eine Nullstelle und zwar in $x = 0$, sie ist auf dem gesamten Definitionsbereich stetig und auf $(-\infty, 0) \cup (1, \infty)$ beliebig oft differenzierbar. Eine etwas längere Rechnung ergibt

$$\begin{aligned} f'(x) &= \frac{1}{2} \frac{1}{f(x)} \frac{2x-3}{(x-1)^2} x^2, \\ f''(x) &= \frac{3}{4} \frac{1}{f(x)} \frac{x}{(x-1)^3}, \end{aligned}$$

für $x \in (-\infty, 0) \cup (1, \infty)$. Es gilt $\lim_{x\to 1^+} f(x) = \infty$, die Umformung $f(x) = \frac{|x|}{\sqrt{1-\frac{1}{x}}}$ zeigt $\lim_{x\to\infty} f(x) = \infty$ und $\lim_{x\to-\infty} f(x) = \infty$. Wegen $\lim_{x\to 0^-} \frac{x^2}{f(x)} = 0$ folgt $\lim_{x\to 0^-} f'(x) = 0$, somit ist f auch in $x = 0$ (linksseitig) differenzierbar und es gilt $f'(0^-) = 0$. Die 1. Ableitung besitzt daher in $x = 0$ und in $x = \frac{3}{2}$ eine Nullstelle. Aus den Ungleichungen

$$\begin{aligned} f'(x) &< 0, \quad x \in (-\infty, 0) \cup (1, \frac{3}{2}), \\ f'(x) &> 0, \quad x \in (\frac{3}{2}, \infty), \end{aligned}$$

liest man ab, daß f auf $(-\infty, 0] \cup (1, \frac{3}{2}]$ streng monoton fällt, auf $[\frac{3}{2}, \infty)$ streng monoton steigt. Insbesondere nimmt f in $x = \frac{3}{2}$ ein lokales Minimum an. Wegen $f \geq 0$ wird das globale Minimum in $x = 0$ angenommen. Die Ungleichung $f'' > 0$ auf $(-\infty, 0) \cup (1, \infty)$ (man beachte, daß$f''(0^-)$ nicht existiert) zeigt, daß die Einschränkungen von f auf die Intervalle $(-\infty, 0]$ und $(1, \infty)$ strikt konvex sind. Schließlich diskutieren wir noch das asymptotische Verhalten von f und untersuchen die Existenz von Asymptoten.

Dazu betrachten wir die Grenzwerte

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \sqrt{\frac{x}{x-1}} = 1,$$

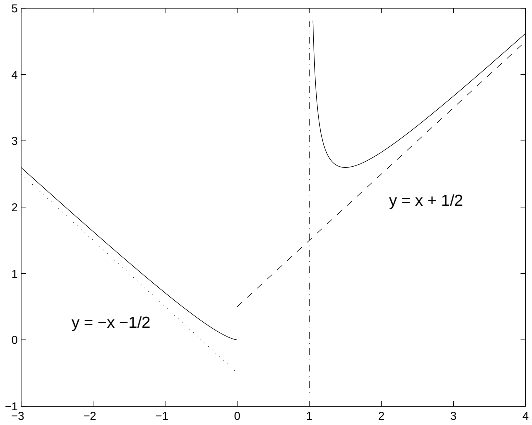
$$\lim_{x\to\infty} (f(x) - x) = \lim_{x\to\infty} \left(\sqrt{\frac{x^3}{x-1}} - x\right) = \lim_{x\to\infty} \frac{\frac{x}{x-1}}{\sqrt{\frac{x}{x-1}} + 1} = \frac{1}{2}.$$

Die Gerade $y = x + \frac{1}{2}$ beschreibt also das asymptotische Verhalten von f für $x \to \infty$. Eine ähnliche Rechnung ergibt

$$\lim_{x\to-\infty} \frac{f(x)}{x} = -1$$

$$\lim_{x\to-\infty} (f(x) - x) = -\frac{1}{2},$$

die Asymptote ist gegeben durch $y = -x - \frac{1}{2}$. Das qualitative Verhalten von f wird in der folgenden Graphik zusammengefaßt.



$$f(x) = \sqrt{\frac{x^3}{x-1}}$$

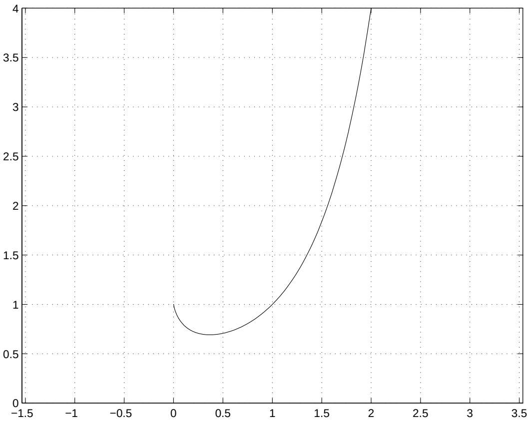
Beispiel VI-10.3. $f(x) = x^x = e^{x \ln x}$.
Es ist $\operatorname{def} f = \mathbb{R}^+ \setminus \{0\}$ und daher $\operatorname{bild} f \subset \mathbb{R}^+ \setminus \{0\}$. Ferner gilt

$$f'(x) = x^x(1 + \ln x),$$
$$f''(x) = x^{x-1} + x^x(1 + \ln x)^2 > 0.$$

f'' zeigt, daß f konvex auf $\operatorname{def} f$ ist. f' besitzt die einzige Nullstelle $x_1 = \frac{1}{e}$ und ist negativ auf $(0, \frac{1}{e})$ und positiv auf $(\frac{1}{e}, \infty)$. Es liegt also in x_1 ein Minimum vor mit $f(\frac{1}{e}) \approx 0.69$. Ferner gilt

$$\lim_{x\to0^+} f(x) = 1, \qquad \lim_{x\to0^+} f'(x) = -\infty,$$
$$\lim_{x\to\infty} f(x) = \infty, \qquad \lim_{x\to\infty} f'(x) = \infty.$$

Es gibt keine Asymptote.



$f(x) = x^x$

KAPITEL VII

Integralrechnung

Historisch wurzelt der Integralbegriff in der Ermittelung von Flächeninhalten. Aber auch physikalische Problemstellungen führen auf Probleme der Integralrechnung: Etwa die Berechnung des zurückgelegten Weges während einer Zeitspanne $[a, b]$, wenn man die Geschwindigkeit v als Funktion der Zeit kennt. Besonders einfach ist die Lösung dieses Problems, wenn die Geschwindigkeit auf jedem Intervall $[t_{i-1}, t_i]$, $t_0 = a < t_1 < \cdots < t_{i-1} < t_i < \cdots < t_n = b$, einen konstanten Wert v_i, $i = 1, \ldots, n$ annimmt. Da der bei konstanter Geschwindigkeit v_i während der Zeitspanne $t_i - t_{i-1}$ zurückgelegte Weg durch $v_i(t_i - t_{i-1})$ gegeben ist, ergibt sich für den gesamten Weg der Ausdruck $\sum_{i=1}^{n} v_i(t_i - t_{i-1})$. Dies ist ein Beispiel eines besonders einfachen Integrals. Selbstverständlich genügt es nicht, sich auf einen derart engen Integralbegriff zu beschränken. Wir zeigen im folgenden, wie man das Integral von stückweise konstanten Funktionen wie oben auf eine für die Praxis ausreichend reichhaltige Klasse von Funktionen systematisch ausdehnen kann.

Notation: In diesem Kapitel bezeichnet I stets das abgeschlossene, beschränkte Intervall $[a, b] \subset \mathbb{R}$. Diese Voraussetzung wird im Folgenden daher nicht mehr gesondert ausgewiesen. Wir erinnern an die Norm von $f \in \mathcal{B}(I, \mathbb{K})$

$$||f||\colon = \sup\{|f(x)|\colon x \in I\},$$

das ist die **Norm** von f. Die gleichmäßige Konvergenz einer Folge von Funktionen (f_n) gegen f werden wir durch $f_n \rightrightarrows f$ andeuten. Mit $\mathbb{K}$ bezeichnen wir den Körper der reellen oder komplexen Zahlen.

1. Treppenfunktionen, Regelfunktionen

Wir präzisieren vorerst die Klasse der stückweise konstanten Funktionen, welchen auf anschauliche Weise ein Integralwert zugeordnet werden kann.

Definition VII-1.1. *Es sei $f\colon I \to \mathbb{K}$ und $n \in \mathbb{N}$.*

i) *$\mathcal{Z} = \{x_0, x_1, \ldots, x_n\}$ heißt **Zerlegung** von I $\underset{Def}{\Leftrightarrow}$ $a = x_0 < x_1 < \cdots < x_n = b$.*
   *$|\mathcal{Z}| := \max\{x_i - x_{i-1} : x_i \in \mathcal{Z}, i = 1, \ldots, n\}$ heißt **Feinheitsmaß** der Zerlegung $\mathcal{Z}$.*

ii) *f heißt **Treppenfunktion** $\underset{Def}{\Leftrightarrow}$ es gibt eine Zerlegung $\mathcal{Z} = \{x_0, x_1, \ldots, x_n\}$ von I und Zahlen c_i, $i = 1, \ldots, n$, derart, daß $f|_{(x_{i-1}, x_i)} = c_i$,*

iii) *$\mathcal{T}(I, \mathbb{K}) := \{f\colon I \to \mathbb{K}$ ist eine Treppenfunktion$\}$.*

Eine Treppenfunktion nimmt nur endlich viele Werte an, somit gilt trivialerweise

$$\mathcal{T}(I,\mathbb{K}) \subset \mathcal{B}(I,\mathbb{K}).$$

Da f auf I definiert ist, sind natürlich auch die Funktionswerte von $f(x_i)$, $i=0,\ldots,n$, festgelegt. Sie sind aber im Folgenden ohne Bedeutung. Verschiedene Zerlegungen $\mathcal{Z}$ von I können zur gleichen Treppenfunktion f führen, z.B. wenn man noch weitere Teilpunkte in eine Zerlegung einfügt. Dies ist zweckmäßig etwa beim Nachweis, daß für $f,g \in \mathcal{T}(I,\mathbb{K})$ auch $f+g \in \mathcal{T}(I,\mathbb{K})$ gilt. Bezeichnet man mit $\mathcal{Z}_f$ bzw. $\mathcal{Z}_g$ den Treppenfunktionen f bzw. g zugrundeliegenden Zerlegungen von I, und mit $\mathcal{Z}$ jene Zerlegung, in der sämtliche Teilpunkte von $\mathcal{Z}_f$ und $\mathcal{Z}_g$ vorkommen, d.h. $\mathcal{Z}$ ist eine **Verfeinerung** von $\mathcal{Z}_f$ und $\mathcal{Z}_g$, dann sind sowohl f als auch g und damit auch $f+g$ (aber auch fg) auf den Teilintervallen von $\mathcal{Z}$ konstant. Da natürlich $\lambda f \in \mathcal{T}(I,\mathbb{K})$ für $f \in \mathcal{T}(I,\mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt, ergibt sich, daß $\mathcal{T}(I,\mathbb{K})$ ein (komplexer) Vektorraum ist. Im nächsten Abschnitt werden wir zeigen, daß man auf $\mathcal{T}(I,\mathbb{K})$ in naheliegender Weise einen Integralbegriff erklären kann. Um zu verstehen, auf welche Funktionen sich dieser Integralbegriff erweitern läßt, ist es notwendig zu untersuchen, welche Funktionen als *gleichmäßige* Grenzwerte von Treppenfunktionen darstellbar sind. Dies trifft beispielsweise für stetige Funktionen zu.

THEOREM VII-1.2. *Es sei $f \in C(I,\mathbb{K})$. Dann gibt es eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$, die gleichmäßig gegen f konvergiert:* $\lim_{n\to\infty} \|f-t_n\| = 0$.

BEWEIS. Wegen der Kompaktheit von I ist f gleichmäßig stetig, d.h.

$$(*) \qquad \forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) \forall x,y \in I\colon |x-y| < \delta \Rightarrow |f(x)-f(y)| < \varepsilon.$$

Es sei $\varepsilon > 0$ und δ durch $(*)$ gegeben. Wir wählen eine Zerlegung $\mathcal{Z} = \{x_0,\ldots,x_n\}$ mit Feinheitsmaß $|\mathcal{Z}| \le \delta$ und definieren eine Treppenfunktion t_ε, indem wir $\xi_i \in [x_{i-1},x_i)$ beliebig wählen und

$$t_\varepsilon(x) = \begin{cases} f(\xi_i), & x \in [x_{i-1},x_i), \quad i=1,\ldots,n, \\ f(b), & x = b \end{cases}$$

festsetzen. Da jedes $x \in I$ in genau einem der Intervalle $[x_{i-1},x_i)$, $i=1,\ldots,n$, liegt (oder $x=b$ gilt), folgt mit $|\xi_i - x| < \delta$ aus $(*)$ für $x \ne b$

$$|f(x)-t_\varepsilon(x)| = |f(x)-f(\xi_i)| < \varepsilon,$$

(für $x=b$ erhält man $|f(b)-t_\varepsilon(b)| = 0$) also

$$\|f-t_\varepsilon\| \le \varepsilon.$$

Läßt man nun ε eine Nullfolge durchlaufen, etwa $(\frac{1}{n})$, erhält man eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ mit

$$\|f-t_n\| \le \frac{1}{n},$$

d.h. (t_n) konvergiert gleichmäßig gegen f. $\square$

Auf die Kompaktheit von I kann nicht verzichtet werden: es ist nicht möglich, die stetige Funktion $x \mapsto \frac{1}{x}$ auf $(0,1)$ gleichmäßig durch eine Folge von Treppenfunktionen zu approximieren.
Der Beweis zeigt, wie man vorgehen muß, wenn man zu einer gegebenen *stetigen* Funktion f eine approximierende Folge von Treppenfunktionen (t_n) konstruieren soll.
Aus der Definition VII-1.1 geht hervor, daß Treppenfunktionen alle sinnvollen links- und rechtsseitigen Grenzwerte besitzen. Diese Eigenschaft überträgt sich auch auf die gleichmäßigen Grenzwerte von Treppenfunktionen.

Theorem VII-1.3. *Die Abbildung $f\colon I \to \mathbb{K}$ werde gleichmäßig durch eine Folge $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ approximiert. Dann besitzt f an jeder Stelle – wo dies möglich ist – sowohl den rechtsseitigen als auch den linksseitigen Grenzwert.*

Beweis. Zu zeigen ist
$$\forall x \in [a,b) \exists f(x^+) \text{ und } \forall x \in (a,b] \exists f(x^-).$$
Wir zeigen nur die Existenz von $f(x^+)$ und überlassen den entsprechenden Nachweis für $f(x^-)$ dem interessierten Leser. Zu $\varepsilon > 0$ wählen wir einen festen Index N_ε so , daß
$$|f(x) - t_{N_\varepsilon}(x)| \leq ||f - t_{N_\varepsilon}|| < \varepsilon$$
für alle $x \in [a,b]$ gilt. Es sei $\xi \in I$. Die Treppenfunktion t_{N_ε} besitzt den rechtsseitigen Grenzwert in ξ, d.h. zu $\varepsilon > 0$ existiert $\delta = \delta(\xi,\varepsilon)$, sodaß für alle $x, y \in (\xi, \xi+\delta)$
$$|x-y| < \delta \Rightarrow |t_{N_\varepsilon}(x) - t_{N_\varepsilon}(y)| < \varepsilon.$$
zutrifft. Insgesamt ergibt sich für solche x, y
$$(*) \qquad |f(x) - f(y)| \leq |f(x) - t_{N_\varepsilon}(x)| + |t_{N_\varepsilon}(x) - t_{N_\varepsilon}(y)| + |t_{N_\varepsilon}(y) - f(y)| < 3\varepsilon.$$
Die Behauptung folgt nun mit Hilfe des Cauchy Kriteriums: es sei (x_n) eine beliebige Folge mit $x_n \downarrow \xi$. Es sei nun $\tilde{N}(\varepsilon)$ so gewählt, daß x_n für $n \geq \tilde{N}(\varepsilon)$ im Intervall $(\xi, \xi+\delta)$ liegt. Für n, $m \geq \tilde{N}(\varepsilon)$ gilt dann wegen $(*)$
$$|f(x_n) - f(x_m)| \leq 3\varepsilon.$$
Somit ist $(f(x_n))$ eine Cauchyfolge und besitzt wegen der Vollständigkeit von $\mathbb{K}$ einen Grenzwert L. Man überzeuge sich, daß L unabhängig von der Wahl der approximierenden Folge (x_n) ist. Es folgt $L = f(\xi^+)$. □

Definition VII-1.4. *Es sei $f\colon I \to \mathbb{K}$.*

i) *f heißt* ***Regelfunktion*** $\underset{Def}{\Leftrightarrow} \exists (t_n) \subset \mathcal{T}(I,\mathbb{K})\colon \lim_{n\to\infty} t_n = f$, *glm.*

ii) $\mathcal{R}(I,\mathbb{K}) := \{f\colon I \to \mathbb{K}, f$ *ist Regelfunktion* $\}$.

Satz VII-1.3 zeigt, daß eine Regelfunktion nur Unstetigkeiten 1. Art aufweisen kann. Diese für Regelfunktionen notwendige Bedingung ist auch hinreichend.

Theorem VII-1.5. *Besitzt $f\colon I \to \mathbb{K}$ überall den links- und rechtsseitigen Grenzwert (wo dies möglich ist), dann ist f eine Regelfunktion.*

Beweis. Es sei $\varepsilon > 0$ gegeben. Für alle $\xi \in [a,b]$ gibt es nach Voraussetzung eine Umgebung $K(\xi, \delta_\xi)$, $\delta_\xi > 0$, mit folgender Eigenschaft:

$$(*) \qquad \forall x, y \in K(\xi, \delta_\xi) \cap I \colon (x-\xi)(y-\xi) > 0 \Rightarrow |f(x) - f(y)| < \varepsilon.$$

(Die Bedingung $(x-\xi)(y-\xi) > 0$ bedeutet, daß entweder $x, y < \xi$ oder $x, y > \xi$ zutrifft). Die Familie $\{K(\xi, \delta_\xi) \colon \xi \in I\}$ bildet eine offene Überdeckung von I. Da I kompakt ist, genügen bereits endlich viele dieser Umgebungen, etwa $K(\xi_i, \delta_i)$, $i = 1, \ldots, k$, um I zu überdecken. Die Punkte ξ_i und die Endpunkte der Intervalle $K(\xi_i, \delta_i)$ denken wir uns der Größe nach geordnet und erhalten dadurch eine Zerlegung $\mathcal{Z}$:

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Aus jedem offenen Intervall (x_{i-1}, x_i) wählen wir einen Punkt z_i und definieren $t_\varepsilon \in \mathfrak{T}(I, \mathbb{K})$ durch

$$t_\varepsilon(x) = \begin{cases} f(z_i), & x \in (x_{i-1}, x_i), \qquad i = 1, \ldots, n, \\ f(x_i), & x = x_i \qquad\qquad\quad\; i = 0, \ldots, n. \end{cases}$$

Für alle $x \in I$ gilt dann

$$|f(x) - t_\varepsilon(x)| < \varepsilon.$$

Dies ist trivialerweise richtig für $x = x_j$. Zu jedem anderen x gibt es genau einen Index j und mindestens einen Index $1 \le \ell \le k$ mit $x \in (x_{j-1}, x_j) \subset K(\xi_\ell, \delta_\ell)$. Dies ist trivial, falls x_{i-1} oder x_i mit einem der Mittelpunkte ξ_ℓ übereinstimmt. Im Fall $x_{i-1} = \xi_\ell - \delta_\ell$ oder $x_i = \xi_\ell + \delta_\ell$ folgt zwangsläufig $\delta_\ell \ge x_i$ bzw. $\delta_\ell \le x_{i-1}$. Andere Fälle kann man analog behandeln, problematisch ist lediglich die Situation $x_{i-1} = \xi_r + \delta_r$ und $x_i = \xi_s - \delta_s$ mit $1 \le r, s \le k$. Dann muß es ein weiteres Intervall $K(\xi_\ell, \delta_\ell)$ mit $(x_{i-1}, x_i) \subset K(\xi_\ell, \delta_\ell)$ geben, da andernfalls $I \subset \bigcup_{\nu=1}^k K(\xi_\nu, \delta_\nu)$ verletzt wäre.
Da somit (x_{j-1}, x_j) entweder links oder rechts von ξ_ℓ liegt, folgt aus $(*)$

$$|f(x) - t_\varepsilon(x)| = |f(x) - f(z_j)| < \varepsilon.$$

Läßt man ε eine Nullfolge durchlaufen, erhält man eine Folge (t_n) mit $\lim_{n\to\infty} t_n = f$, glm. □

Kombiniert man die Sätze VII-1.5 und VII-1.3, ergibt sich folgende Charakterisierung von Regelfunktionen.

Theorem VII-1.6. *Es sei $f \colon I \to \mathbb{K}$. Äquivalent sind*

(1) *f ist eine Regelfunktion,*
(2) *f besitzt überall den links- und rechtsseitigen Grenzwert (wo dies möglich ist).*

Ist $f \colon [a,b] \to \mathbb{R}$ monoton, dann existiert für alle $x \in (a,b]$ der rechtsseitige Grenzwert, für alle $x \in [a,b)$ der linksseitige Grenzwert und es gilt

$$f(x^+) = \inf_{s>x} f(s), \qquad \text{bzw. } f(x^-) = \sup_{s<x} f(s).$$

Monotone Funktionen sind also Regelfunktionen.

Sind f, g Regelfunktionen, folgt aus Satz VII-1.6, daß auch $f + g$ bzw. λf für $\lambda \in \mathbb{K}$ Regelfunktionen sind. $\mathcal{R}(I, \mathbb{K})$ ist somit ein Vektorraum. Es gilt

$$\mathcal{R}(I, \mathbb{K}) \subset \mathcal{B}(I, \mathbb{K}).$$

Abschließend zeigen wir, daß $\mathcal{R}(I, \mathbb{K})$ gegenüber der Bildung gleichmäßiger Grenzwerte abgeschlossen ist:

THEOREM VII-1.7. *Es sei $(f_n)_{n \geq 1} \subset \mathcal{R}(I, \mathbb{K})$ und $f_n \rightrightarrows f$. Dann ist f eine Regelfunktion.*

BEWEIS. Es sei $\varepsilon > 0$ gegeben. Zu jedem f_n existiert $t_n \in \mathcal{T}(I, \mathbb{K})$ mit

$$\|f_n - t_n\| < \frac{1}{n}.$$

Zu $\varepsilon > 0$ gibt es nach Voraussetzung einen Index $N(\varepsilon)$ mit $\|f - f_n\| < \varepsilon$ und $\frac{1}{n} < \varepsilon$ für alle $n \geq N(\varepsilon)$. Somit folgt für alle $n \geq N(\varepsilon)$

$$\|f - t_n\| \leq \|f - f_n\| + \|f_n - t_n\| < 2\varepsilon.$$

Somit kann auch f durch eine Folge von Treppenfunktionen gleichmäßig approximiert werden und ist daher selbst eine Regelfunktion. $\square$

2. Das Cauchy Integral

Es sei $t \in \mathcal{T}(I, \mathbb{K})$ und $\mathcal{Z} = \{x_0, x_1, \ldots, x_n\}$ eine zu t passende Zerlegung von I. Ferner bezeichnen wir mit c_i den Wert von t auf (x_{i-1}, x_i), $i = 1, \ldots, n$. Wir definieren die Abbildung $\mathcal{I}\colon \mathcal{T}(I, \mathbb{K}) \to \mathbb{C}$ durch

$$\mathcal{I}(t) = \sum_{k=1}^{n} c_k (x_k - x_{k-1}).$$

Der Wert von $\mathcal{I}(t)$ könnte möglicherweise nicht nur von t, sondern auch von der jeweiligen Zerlegung $\mathcal{Z}$ abhängen. Tatsächlich ist dies aber nicht der Fall: Man überzeugt sich davon, indem man sich vorerst überlegt, daß sich $\mathcal{I}(t)$ nicht ändert, wenn man einen Teilpunkt in $\mathcal{Z}$ einfügt, oder einen (redundanten) Teilpunkt von $\mathcal{Z}$, in dem t stetig ist, wegläßt. Es seien nun t eine Treppenfunktion und $\mathcal{Z}_i$, $i = 1, 2$, zwei nach Definition VII-1.1 zugeordnete Zerlegungen von I. Wir bilden die Zerlegung $\mathcal{Z}_{12}$, in welcher sämtliche Teilpunkte von $\mathcal{Z}_1$ und $\mathcal{Z}_2$ nach der Größe geordnet auftreten. Ferner sei $\mathcal{I}_i(t)$ der Wert der Summe, wenn man der Berechnung die Zerlegung $\mathcal{Z}_i$ zugrundelegt, $i = 1, 2, 12$. Man kann nun der Zerlegung $\mathcal{Z}_1$ bzw. $\mathcal{Z}_2$ schrittweise einen Teilpunkt hinzufügen, ohne daß sich der Wert von $\mathcal{I}_i(t)$, $i = 1, 2$ verändert. Nach endlich vielen Schritten ist man schließlich bei der Zerlegung $\mathcal{Z}_{12}$ angelangt und erhält

$$\mathcal{I}_1(t) = \mathcal{I}_{12}(t) = \mathcal{I}_2(t).$$

Somit ist folgende Definition sinnvoll:

Definition VII-2.1. *Es sei $t \in \mathcal{T}(I, \mathbb{K})$ und $\mathcal{Z} = \{x_0, \ldots, x_n\}$ eine Zerlegung von I. Ferner seien $c_i \in \mathbb{K}$, $i = 1, \ldots, n$, die Werte von t auf (x_{i-1}, x_i).*

$$\mathcal{I}(t) := \sum_{k=1}^{n} c_k (x_k - x_{k-1})$$

*heißt **(bestimmtes) Integral** von t. Man schreibt auch*

$$\mathcal{I}(t) = \int_a^b t = \int_a^b t(x)dx$$

*und nennt t **Integrand**, a die* **untere** *und b die* **obere** ***Integrationsgrenze**.*

Beispiel VII-2.2. Wir betrachten die Abbildung (für festes $n \in \mathbb{N}$)

$$t(x) = \frac{1}{n}[nx], \quad x \in [0, 1].$$

(vgl. Beispiel IV-3.2). t nimmt auf $[\frac{k}{n}, \frac{k+1}{n})$ den Wert $c_k = \frac{k}{n}$ an, $k = 0, \ldots, n-1$. Eine zu t passende Zerlegung von $[0, 1]$ ist die äquidistante Einteilung

$$0 < \frac{1}{n} < \cdots < \frac{k}{n} < \cdots < \frac{n-1}{n} < 1,$$

d.h. $x_i = \frac{i}{n}$, $i = 0, \ldots, n$, und $c_i = \frac{i-1}{n}$, $i = 1, \ldots, n$. Somit folgt

$$\mathcal{I}(t) = \sum_{i=1}^{n} \frac{i-1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \sum_{i=1}^{n} (i-1) = \frac{1}{n^2} \sum_{i=0}^{n-1} i = \frac{n(n-1)}{n^2 \cdot 2} = \frac{1}{2} \frac{n-1}{n}.$$

Wir zeigen nun, daß $\mathcal{I} \colon \mathcal{T}(I, \mathbb{K}) \to \mathbb{K}$ eine lineare Abbildung ist:

Theorem VII-2.3. *Für alle $f, g \in \mathcal{T}(I, \mathbb{K})$, $I = [a, b]$, und für alle $\lambda \in \mathbb{K}$ gilt*

(1) $\mathcal{I}(f + g) = \mathcal{I}(f) + \mathcal{I}(g)$,
(2) $\mathcal{I}(\lambda f) = \lambda \mathcal{I}(f)$.

Beweis. Für den Beweis von (1) bildet man die Zerlegung $\mathcal{Z}$, in der alle Teilpunkte der zu f und zu g gehörenden Zerlegungen auftreten. Dann sind sowohl f als auch g konstant auf den Teilintervallen von $\mathcal{Z}$. Die Behauptung (1) ist nun evident. Die Behauptung (2) folgt unmittelbar aus der Definition von $\mathcal{I}$. $\square$

Theorem VII-2.4. *1. Es sei $f \in \mathcal{T}(I, \mathbb{R})$ und $f(x) \geq 0$ für alle $x \in I$. Dann ist $\mathcal{I}(f) \geq 0$ (d.h. $\mathcal{I}$ ist eine **positive** lineare Abbildung).*
*2. Es seien $f, g \in \mathcal{T}(I, \mathbb{R})$ und $f(x) \geq g(x)$ für alle $x \in I$. Dann ist $\mathcal{I}(f) \geq \mathcal{I}(g)$ (d.h. $\mathcal{I}$ ist eine **monotone** lineare Abbildung).*
3. Für alle $f \in \mathcal{T}(I, \mathbb{K})$ gilt

$$|\mathcal{I}(f)| \leq \mathcal{I}(|f|) \qquad \text{und} \qquad |\mathcal{I}(f)| \leq \|f\|(b-a).$$

Somit folgt $\mathcal{I} \in \mathcal{L}(\mathcal{T}(I, \mathbb{K}), \mathbb{K})$.

Beweis. Die Aussage 1. liest man unmittelbar aus der Definition ab und 2. folgt aus 1.
ad 3. Es sei $\mathcal{Z} = \{x_0, \ldots, x_n\}$ eine zu f gehörige Zerlegung von I und c_i der Wert von f auf (x_{i-1}, x_i). Aus

$$|c_i| \leq \sup\{|f(x)| \colon x \in I\} = \|f\|, \quad i = 1, \ldots, n,$$

folgt

$$|\mathfrak{I}(f)| \leq \sum_{i=1}^{n} |c_i|(x_i - x_{i-1}) \leq \|f\|(b-a).$$

□

Neben der Additivität bezüglich des Integranden hat das Integral auch eine Additivitätseigenschaft bezüglich des Integrationsintervalls: Dazu bemerken wir, daß die Einschränkung einer Treppenfunktion $f \in \mathfrak{T}(I, \mathbb{K})$ auf ein beliebiges Teilintervall $J \subset I$ wieder eine Treppenfunktion ist, die wir der Einfachheit halber wieder mit f bezeichnen:

Theorem VII-2.5. *Es sei $f \in \mathfrak{T}(I, \mathbb{K})$ und $c \in I$. Dann gilt*

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx.$$

Beweis. Die Behauptung ergibt sich unmittelbar indem man den Punkt c in die Zerlegung aufnimmt, welche für die Berechnung von $\int_a^b f(x)dx$ verwendet wird. □

Die Beschränktheit von $\mathfrak{I}$ ermöglicht es, die Definition des Integrals von $\mathfrak{T}(I, \mathbb{K})$ auf $\mathfrak{R}(I, \mathbb{K})$ zu erweitern: Dazu betrachten wir eine Regelfunktion f und eine Folge von Treppenfunktionen (t_n), welche gleichmäßig gegen f konvergiert. Es folgt

$$|\mathfrak{I}(t_n) - \mathfrak{I}(t_m)| \leq (b-a)\|t_n - t_m\|.$$

Wegen der Gleichmäßigkeit der Konvergenz von (t_n) ist für alle $x \in I$ die Folge $(t_n(x))$ eine (gleichmäßige) Cauchy Folge, d.h. es existiert ein $N(\varepsilon) \in \mathbb{N}$ derart, daß für alle $n, m \geq N(\varepsilon)$

$$\|t_n - t_m\| < \frac{\varepsilon}{b-a}$$

zutrifft. Dies zeigt für $n, m \geq N(\varepsilon)$

$$|\mathfrak{I}(t_n) - \mathfrak{I}(t_m)| < \varepsilon,$$

d.h. $(\mathfrak{I}(t_n))_{n \geq 1}$ ist eine Cauchy Folge in $\mathbb{K}$ und somit konvergent. Es liegt nahe

$$\mathfrak{I}(f) := \lim_{n \to \infty} \mathfrak{I}(t_n)$$

zu setzen. Diese Definition ist sinnvoll, soferne gezeigt werden kann, daß $\lim_{n\to\infty} \mathfrak{I}(t_n)$ unabhängig ist von der jeweiligen Folge (t_n), welche gegen f konvergiert: Es gelte also für zwei Folgen von Treppenfunktionen $t_n \rightrightarrows f$ und $s_n \rightrightarrows f$. Zu zeigen ist:

$\lim_{n\to\infty} \mathfrak{I}(t_n) = \lim_{n\to\infty} \mathfrak{I}(s_n)$ bzw. $\lim_{n\to\infty} \mathfrak{I}(t_n - s_n) = 0$. Letzteres ist aber eine unmittelbare Folge der Beschränktheit von $\mathfrak{I}$

$$|\mathfrak{I}(t_n - s_n)| \leq (b-a)\|t_n - s_n\|$$

und des Faktums $\lim_{n\to\infty} \|t_n - s_n\| \leq \lim_{n\to\infty} \|t_n - f\| + \lim_{n\to\infty} \|f - s_n\| = 0$.
Es ist daher sinnvoll, folgende Definition zu vereinbaren:

Definition VII-2.6. *Es sei $f \in \mathcal{R}(I,\mathbb{K})$ und $(t_n) \subset \mathcal{T}(I,\mathbb{K})$ mit $t_n \rightrightarrows h$.*

$$\mathfrak{I}(f) := \lim_{n\to\infty} \mathfrak{I}(t_n) \text{ heißt } \textbf{Integral der Regelfunktion } f.$$

Für das Integral $\mathfrak{I}(f)$ schreibt man auch

$$\mathfrak{I}(f) = \int_a^b f = \int_a^b f(x)dx.$$

Wegen $\mathcal{T}(I,\mathbb{K}) \subset \mathcal{R}(I,\mathbb{K})$ kann man für $f \in \mathcal{T}(I,\mathbb{K})$ die konstante Folge $t_n = f$, $n \in \mathbb{N}$, zur Approximation verwenden. Da der Grenzwert $\mathfrak{I}(f)$ unabhängig von der approximierenden Folge von Treppenfunktionen ist, ergibt sich, daß das Integral gemäß Definition VII-2.6 auf $\mathcal{T}(I,\mathbb{K})$ mit dem Integral gemäß Definition VII-2.1 übereinstimmt. Durch Definition VII-2.6 wird also der Integralbegriff aus Definition VII-2.1 von $\mathcal{T}(I,\mathbb{K})$ auf $\mathcal{R}(I,\mathbb{K})$ fortgesetzt. Satz VII-1.6 kann nun als notwendige *und* hinreichende Integrabilitätsbedingung betrachtet werden: Die Klasse der **integrierbaren Funktionen** ist der Vektorraum der Funktionen, für welche überall (soferne sinnvoll) rechts- und linksseitige Grenzwerte existieren.
Wir veranschaulichen die Handhabung der Definition an einigen Beispielen:

Beispiel VII-2.7. 1) Wir betrachten $f(x) = x$, $x \in [0,1]$. Da f stetig ist, ist f eine Regelfunktion. Als approximierende Folge von Treppenfunktionen können wir $t_n(x) = \frac{1}{n}[nx]$, $n \in \mathbb{N}$, verwenden (vgl. IV-3.2). Nach Beispiel VII-2.2 gilt

$$\mathfrak{I}(t_n) = \frac{1}{2}(1 - \frac{1}{n}),$$

und daher

$$\mathfrak{I}(f) = \int_0^1 f(x)dx = \lim_{n\to\infty} \mathfrak{I}(t_n) = \frac{1}{2}.$$

2) Für $f(x) = e^x$, $x \in [a,b]$, zeigen wir nun

$$\mathfrak{I}(f) = \int_a^b f(x)dx = e^b - e^a.$$

Auch in diesem Falle können wir eine äquidistante Zerlegung $\mathcal{Z}_n$ von I verwenden:

$$a = x_0 < x_1 < \cdots < x_n = b, \quad x_i = \frac{b-a}{n} i + a, \quad i = 0, \ldots, n.$$

Wir setzen
$$t_n(x) = \begin{cases} e^{x_{i-1}}, & x \in [x_{i-1}, x_i), \quad i = 1, \dots, n, \\ e^b, & x = b. \end{cases}$$
Auf Grund des Beweises von Satz VII-1.2 wissen wir zwar bereits, daß wegen $\lim_{n\to\infty} |\mathcal{Z}_n| = 0$ auch $t_n \rightrightarrows f$ gelten muß. Wir wollen uns von der gleichmäßigen Konvergenz der Folge (t_n) jedoch direkt überzeugen: Jedes $x \neq b$ liegt in einem eindeutig bestimmten Intervall $[x_{i-1}, x_i)$. Es folgt also mit Hilfe des Mittelwertsatzes die Existenz von $\xi_i \in (x_{i-1}, x_i)$ mit
$$|f(x) - t_n(x)| = |e^x - e^{x_{i-1}}| \le e^{\xi_i}|x - x_{i-1}| \le e^b|x - x_{i-1}| \le e^b|\mathcal{Z}_n|.$$
Für $x = b$ ist diese Abschätzung trivialerweise efüllt. Daraus folgt
$$\|f - t_n\| \le e^b|\mathcal{Z}_n|.$$
Als nächstes berechnen wir $\mathfrak{I}(t_n)$:
$$\begin{aligned}\mathfrak{I}(t_n) &= \sum_{k=1}^{n} \exp(a + \frac{k-1}{n}(b-a))\frac{b-a}{n} \\ &= \frac{b-a}{n}e^a \sum_{k=1}^{n} \exp(\frac{k-1}{n}(b-a)) = \frac{b-a}{n}e^a \sum_{k=1}^{n} (\exp(\frac{b-a}{n}))^{k-1} \\ &= \frac{b-a}{n}e^a \frac{\exp(b-a) - 1}{\exp\frac{b-a}{n} - 1} = \frac{\frac{b-a}{n}}{\exp\frac{b-a}{n} - 1}(e^b - e^a) \underset{n\to\infty}{\longrightarrow} e^b - e^a,\end{aligned}$$
wobei $\lim_{z\to 0} \frac{z}{e^z - 1} = 1$ verwendet wurde.

Theorem VII-2.8. *Für alle $f, g \in \mathcal{R}(I, \mathbb{K})$ und alle $\lambda \in \mathbb{K}$ gilt*

(1) $\mathfrak{I}(f + g) = \mathfrak{I}(f) + \mathfrak{I}(g)$,
(2) $\mathfrak{I}(\lambda f) = \lambda \mathfrak{I}(f)$.

Beweis. Es seien $(t_n), (s_n) \subset \mathfrak{T}(I, \mathbb{K})$ mit
$$t_n \rightrightarrows f, \qquad s_n \rightrightarrows g.$$
Dann gilt $t_n + s_n \in \mathfrak{T}(I, \mathbb{K})$ und
$$t_n + s_n \rightrightarrows f + g.$$
Auf $\mathfrak{T}(I, \mathbb{K})$ ist $\mathfrak{I}$ linear, also gilt
$$\mathfrak{I}(t_n + s_n) = \mathfrak{I}(t_n) + \mathfrak{I}(s_n).$$
Durch Grenzübergang erhält man
$$\mathfrak{I}(f + g) = \lim_{n\to\infty} \mathfrak{I}(t_n + s_n) = \lim_{n\to\infty} \mathfrak{I}(t_n) + \lim_{n\to\infty} \mathfrak{I}(s_n) = \mathfrak{I}(f) + \mathfrak{I}(g).$$
Analog ergibt sich (2). □

Theorem VII-2.9. *Für alle $f \in \mathcal{R}(I, \mathbb{K})$ gilt*
$$|\mathfrak{I}(f)| \le \mathfrak{I}(|f|) \qquad \text{und} \qquad |\mathfrak{I}(f)| \le (b-a)\|f\|.$$

Beweis. Bezeichnet man mit (t_n) eine Folge von Treppenfunktionen, welche gleichmäßig gegen f konvergiert, folgt die Behauptung aus der Abschätzung

$$|\mathcal{I}(t_n)| \leq (b-a)||t_n|| \leq (b-a)(||t_n - f|| + ||f||).$$

$\square$

Die Sätze VII-2.8, VII-2.9 zeigen, daß die Erweiterung von $\mathcal{I}$ eine *lineare* und *stetige* Abbildung $\mathcal{R}(I,\mathbb{K}) \to \mathbb{K}$ darstellt. Wir weisen nun nach, daß auch die Positivität erhalten bleibt.

Theorem VII-2.10. *1) Es sei $f \in \mathcal{R}(I,\mathbb{R})$ und $f(x) \geq 0$ für alle $x \in I$. Dann gilt $\mathcal{I}(f) \geq 0$.*
2) Es sei $f \in C(I,\mathbb{R})$ und $f(x) \geq 0$ für alle $x \in I$. Ist $f(x_0) > 0$ für ein $x_0 \in I$, dann ist $\mathcal{I}(f) > 0$.

Beweis. 1) Es sei $(t_n) \in \mathcal{T}(I,\mathbb{R})$ eine f approximierende Folge von Treppenfunktionen, d.h. $\lim_{n\to\infty} \|f - t_n\| = 0$, und $\varepsilon > 0$. Es gibt also ein $N(\varepsilon) \in \mathbb{N}$, sodaß für alle $n \geq N(\varepsilon)$ und für alle $x \in I$

$$t_n(x) \geq f(x) - \varepsilon \geq -\varepsilon$$

zutrifft. Nach Satz VII-2.4-2) gilt

$$\mathcal{I}(t_n) \geq -\varepsilon(b-a)$$

und daher auch $\mathcal{I}(f) = \lim_{n\to\infty} \mathcal{I}(t_n) \geq -\varepsilon(b-a)$. Da $\varepsilon > 0$ beliebig gewählt war, gilt die Behauptung.
2) O.B.d.A können wir $x_0 \in (a,b)$ annehmen. Es sei $\delta > 0$ so gewählt, daß für alle $x \in (x_0-\delta, x_0+\delta) \subset I$,

$$f(x) > \frac{1}{2} f(x_0) > 0.$$

Wir definieren nun die Treppenfunktion

$$t(x) = \begin{cases} \frac{1}{2} f(x_0), & x \in (x_0-\delta, x_0+\delta), \\ 0, & \text{sonst}, \end{cases}$$

sodaß $f(x) \geq t(x)$ für alle $x \in I$ gilt. Wegen 1) folgt

$$\mathcal{I}(f) \geq \mathcal{I}(t) = \frac{1}{2} f(x_0) 2\delta > 0.$$

$\square$

Eine unmittelbar Folge der Linearität und Positivität ist wieder die Monotonie:

Korollar VII-2.11. Es seien $f, g \in \mathcal{R}(I,\mathbb{R})$ und $f(x) \geq g(x)$ für alle $x \in I$. Dann gilt

$$\mathcal{I}(f) \geq \mathcal{I}(g).$$

Sind darüber hinaus f und g stetig und gilt an mindestens einer Stelle $x_0 \in I$ $f(x_0) > g(x_0)$, dann gilt

$$\mathcal{I}(f) > \mathcal{I}(g).$$

Bemerkung VII-2.12. Geht man den beschriebenen Fortsetzungsprozeß von $\mathcal{I}$ und den Beweis der Sätze VII-2.8 und VII-2.10 noch einmal durch, erkennt man, daß weder die spezielle Bedeutung von $\mathcal{I}$ – das Integral von Treppenfunktionen – noch spezielle Eigenschaften von $\mathbb{K}$ verwendet wurden. Die Beweise beruhen lediglich auf den *strukturellen* Eigenschaften Linearität und Beschränkheit (also Stetigkeit) und der Vollständigkeit $\mathbb{K}$. Die Konstruktion des Cauchy Integrals kann daher auf Funktionen $f\colon I \to X$, X ein Banachraum übertragen werden.

Theorem VII-2.13. *Es sei $f \in \mathcal{R}(I, \mathbb{K})$ und $c \in I$. Dann gilt*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Beweis. Zunächst überlegen wir, daß die Restriktion einer
Regelfunktion auf ein Teilintervall von I wieder eine Regelfunktion ist. Dies folgt unmittelbar aus der Charakterisierung in Satz VII-1.6. Es sei nun $(t_n) \subset \mathcal{T}(I, \mathbb{K})$ eine approximierende Folge von Treppenfunktionen. Für jedes $n \in \mathbb{N}$ gilt nach Satz VII-2.5

$$\int_a^b t_n(x)dx = \int_a^c t_n(x)dx + \int_c^b t_n(x)dx.$$

Da natürlich $(t_n|_{[a,c]})$ gleichmäßig gegen $f|_{[a,c]}$ konvergiert (und analog für $[c, b]$), ergibt sich die Behauptung aus der Definition des Integrals. □

Es ist zweckmäßig, auf die Voraussetzung $a \le b$ zu verzichten und für $a > b$ zu definieren

$$\int_a^b f(x)dx := -\int_b^a f(x)dx.$$

Es seien nun a, b und c beliebige reelle Zahlen und $f\colon [\min(a, b, c), \max(a, b, c)] \to \mathbb{K}$ eine Regelfunktion. Mit Hilfe dieser erweiterten Integraldefinition gilt kann man sich von der Gültigkeit folgender Gleichheit überzeugen

$$\int_u^v f(x)\,dx + \int_v^w f(x)\,dx = \int_u^w f(x)\,dx.$$

Theorem VII-2.14. *[2. Mittelwertsatz der Integralrechnung] Es sei $f \in C(I, \mathbb{R})$, $g \in \mathcal{R}(I, \mathbb{R})$ und $g(x) \ge 0$ für alle $x \in I$. Dann gibt es eine Zwischenstelle $\xi \in I$ mit*

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx.$$

Beweis. Wegen der Kompaktheit von I nimmt f nach Korollar III-4.3 Maximum und Minimum an, d.h. es existieren

$$m = \min\{f(x)\colon x \in I\} \text{ und } M = \max\{f(x)\colon x \in I\}.$$

Somit gilt für alle $x \in I$

$$mg(x) \le f(x)g(x) \le Mg(x),$$

und wegen der Monotonie und Linearität des Integrals auch

$$m\int_a^b g(x)\,dx \le \int_a^b f(x)g(x)\,dx \le M\int_a^b g(x)\,dx.$$

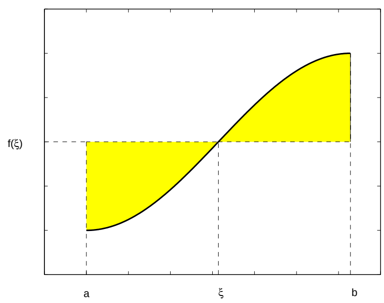
Falls $\int_a^b g(x)\,dx = 0$, muß auch $\int_a^b f(x)g(x)\,dx = 0$ gelten und die Behauptung des Satzes ist richtig. Es sei nun $\int_a^b g(x)\,dx \ne 0$. Dann folgt $\kappa = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \in [m, M]$. Aus dem Zwischenwertsatz IV-1.3 folgt nun die Existenz einer Zwischenstelle $\xi \in I$ mit $\kappa = f(\xi)$. □

Eine eingehende Analyse würde zeigen, daß ξ sogar in (a, b) gewählt werden kann. Für $g \equiv 1$ ergibt sich als Spezialfall:

Korollar VII-2.15 (1. Mittelwertsatz der Integralrechnung). Es sei $f \in C(I, \mathbb{R})$. Dann gibt es ein $\xi \in I$ mit

$$\int_a^b f(x)dx = f(\xi)(b-a).$$

Deutet man $\int_a^b f(x)dx$ (mit $f \ge 0$ auf I) als die Fläche zwischen dem Graph von f und der x-Achse, dann ist nach dem 1. Mittelwertsatz diese gleich dem Flächeninhalt eines Rechteckes mit den Seitenlängen $b - a$ und $f(\xi)$:



3. Stammfunktion

Viele Probleme in Naturwissenschaft und Technik führen auf die Aufgabe, den Differentiationsprozeß umzukehren, d.h. zu einer gegebenen Funktion $f\colon I \to \mathbb{K}$ eine Funktion $F\colon I \to \mathbb{K}$ so zu bestimmen, daß $F' = f$ gilt.

Definition VII-3.1. *Es sei $f\colon I \to \mathbb{K}$. $F\colon I \to \mathbb{K}$ heißt* ***Stammfunktion*** *von f* $\underset{Def}{\Leftrightarrow}$

(1) *F ist stetig.*
(2) *Es gibt eine höchstens abzählbare Teilmenge $A \subset I$ so, daß F auf $I \setminus A$ differenzierbar ist mit*

$$F'(x) = f(x), \quad x \in I \setminus A.$$

Beispiel VII-3.2. Zu jeder Ableitung gibt es eine Stammfunktion (wir lassen in der folgenden Tabelle die jeweiligen Definitionsbereiche weg):

Funktion f	Stammfunktion F	Funktion f	Stammfunktion F
$f(x)$	$F(x)$	$f(x)$	$F(x)$
$x^c \quad (c \neq -1)$	$\dfrac{1}{c+1}x^{c+1}$	$\dfrac{1}{\cos^2 x}$	$\tan x$
$\dfrac{1}{x}$	$\ln x$	$\dfrac{1}{\sin^2 x}$	$-\cot x$
$a^x \quad (a \neq 1)$	$\dfrac{1}{\ln a}a^x$	$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\exp ax \quad (a \neq 0)$	$\dfrac{1}{a} \cdot \exp ax$	$\dfrac{1}{1+x^2}$	$\arctan x$
$\exp ix$	$-i \cdot \exp ix$	$\dfrac{f'(x)}{f(x)}$	$\ln \lvert f(x) \rvert$
$\cos x$	$\sin x$	$\dfrac{f'(x)}{1+(f(x))^2}$	$\arctan f(x)$
$\sin x$	$-\cos x$	$\sum\limits_{k=0}^{\infty} a_k x^k$	$\sum\limits_{k=0}^{\infty} \dfrac{a_k}{k+1} x^{k+1}$

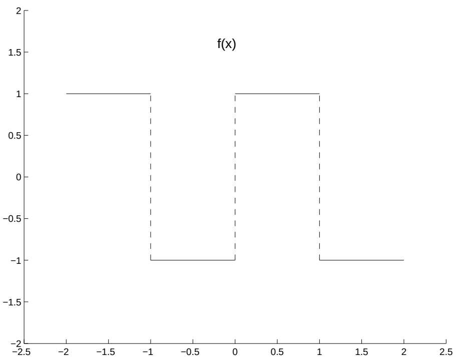
Die meisten Lehrbücher verlangen für eine Stammfunktion F die Differenzierbarkeit und Identität $F' = f$ auf *ganz* I. Bereits einfache Anwendungen in der Technik erfordern aber einen allgemeineren und flexibleren Begriff der Stammfunktion.

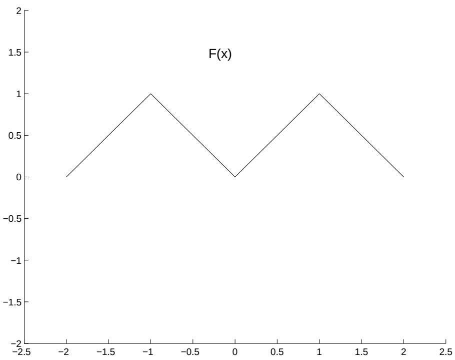
Beispiel VII-3.3. Es sei $f\colon \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} 1, & 2n \leq x < 2n+1, \quad n \in \mathbb{Z}, \\ -1 & \text{sonst.} \end{cases}$$

Eine Stammfunktion von f ist gegeben durch

$$F(x) = \begin{cases} x - 2n, & 2n \leq x < 2n+1, \\ -x + 2n + 2 & 2n+1 \leq x < 2n+2 \end{cases} n \in \mathbb{Z}.$$





Wegen der Linearität der Differentiation ist klar, daß mit zwei Stammfunktionen F und G für f bzw.g auch $\alpha F+\beta G$, $\alpha, \beta \in \mathbb{K}$, Stammfunktionen von $\alpha f+\beta g$ sind (dabei wird auch verwendet, daß die Vereinigung von zwei abzählbaren Mengen abzählbar ist. Trivialerweise ist mit F natürlich auch $F + \alpha$, $\alpha \in \mathbb{K}$, eine Stammfunktion von f. Wir zeigen nun, daß sich umgekehrt je zwei Stammfunktionen von f höchstens um eine additive Konstante unterscheiden können. Dieses Resultat ist eine unmittelbare Konsequenz von Korallar V-4.7 falls $A = \emptyset$. Der allgemeinere Fall folgt aus

Theorem VII-3.4. *Es sei $f \in C(I, \mathbb{K})$. Ferner gebe es eine abzählbare Teilmenge $A \subset I$ und eine Konstante $L > 0$ derart, daß gilt*

(1) *f ist rechtsseitig differenzierbar auf $I \setminus A$,*
(2) *$\forall x \in I \setminus A$: $|f'_+(x)| \leq L$.*

Für alle $x, y \in I$ gilt dann

$$|f(x) - f(y)| \leq L|x - y|,$$

d.h. f ist Lipschitz-stetig.

Beweis. Es sei $x < y$. Für beliebiges $\varepsilon > 0$ definieren wir $F_\varepsilon \colon [x, y] \to \mathbb{R}$,

$$F_\varepsilon(\xi) := |f(\xi) - f(x)| - (L + \varepsilon)(\xi - x), \quad \xi \in [x, y],$$

und zeigen $F_\varepsilon(y) \leq 0$. Daraus folgt für $\varepsilon \to 0$ die Behauptung. Angenommen es wäre $F_\varepsilon(y) > 0$. Da $F_\varepsilon(A)$ abzählbar, $[0, F_\varepsilon(y)]$ aber überabzählbar ist, muß es ein $\gamma \in \mathbb{R}$ geben mit

$$0 = F_\varepsilon(x) < \gamma < F_\varepsilon(y) \text{ und } \gamma \notin F_\varepsilon(A).$$

Nach dem Zwischenwertsatz IV-1.3 gibt es eine Zwischenstelle $\xi^* \in (x, y)$ mit $F_\varepsilon(\xi^*) = \gamma$. Dies läßt sich so einrichten (vgl. den Beweis zu IV-1.2), daß für alle $\xi \in (\xi^*, y)$

$$F_\varepsilon(\xi) > \gamma$$

zutrifft. Dann gilt einerseits

$$(*) \qquad \varphi(\xi) := \frac{F_\varepsilon(\xi) - F_\varepsilon(\xi^*)}{\xi - \xi^*} > 0 \text{ für alle } \xi \in (\xi^*, y],$$

andererseits wegen der Definition von F_ε

$$\begin{aligned}\varphi(\xi) &= (\xi - \xi^*)^{-1}\big[|f(\xi) - f(x)| - |f(\xi^*) - f(x)| - (L + \varepsilon)(\xi - x - (\xi^* - x))\big] \\ &= (\xi - \xi^*)^{-1}(|f(\xi) - f(x)| - |f(\xi^*) - f(x)|) - (L + \varepsilon) \\ &\leq \frac{|f(\xi) - f(\xi^*)|}{\xi - \xi^*} - L - \varepsilon.\end{aligned}$$

Nun ist $\xi^* \notin A$, da $\gamma \notin F_\varepsilon(A)$. f besitzt also in ξ^* eine rechtsseitige Ableitung. Wegen $|f'_+(\xi^*)| \leq L$ gibt es also eine rechtsseitige Umgebung $(\xi^*, \xi^* + \delta)$, $\delta > 0$, so, daß

$$\frac{|f(\xi) - f(\xi^*)|}{\xi - \xi^*} \leq L + \varepsilon/2 \quad \text{ für alle } \xi \in (\xi^*, \xi^* + \delta)$$

zutrifft. Dies hat jedoch

$$\varphi(\xi) \leq -\frac{\varepsilon}{2} < 0$$

für alle $\xi \in (\xi^*, \xi^* + \delta)$ zur Folge, ein Widerspruch zu $(*)$. □

Korollar VII-3.5. Es sei $f \in C(I, \mathbb{K})$, und $A \subset I$ abzählbar. Auf $I \setminus A$ sei f rechtsseitig differenzierbar mit $f'_+(x) = 0$, $x \in I \setminus A$. Dann ist f konstant.

Beweis. Setze $L = 0$ in Satz VII-3.4. □

Für Stammfunktionen bedeutet Korollar VII-3.5:

Theorem VII-3.6. *Es seien $f\colon I \to \mathbb{K}$ und $F, G \in C(I, \mathbb{K})$ Stammfunktionen von f. Dann ist $F - G$ konstant.*

4. Der Hauptsatz der Differential- und Integralrechnung

In diesem Abschnitt zeigen wir, daß die Integration in gewisser Weise die Umkehroperation zur Differentiation ist. Wir beginnen mit einer Eigenschaft von Regelfunktionen:

Theorem VII-4.1. *Jede Regelfunktion $f\colon I = [a, b] \to \mathbb{K}$ besitzt höchstens abzählbar viele Unstetigkeiten.*

Beweis. Es sei $(t_n) \subset \mathcal{T}(I, \mathbb{K})$ eine approximierende Folge von Treppenfunktionen, d.h. es gilt $f = \lim_{n\to\infty} t_n$ glm. auf I. Jede Treppenfunktion t_n ist stetig, ausgenommen auf einer höchstens endlichen Menge $A_n \subset I$. Somit ist ist $A := \bigcup_{n=1}^\infty A_n$ abzählbar. Da jede Treppenfunktion t_n auf $I \setminus A$ stetig ist, folgt aus Satz IV-3.6, daß auch f auf $I \setminus A$ stetig ist. □

THEOREM VII-4.2. *Es sei $f \in \mathcal{R}(I, \mathbb{K})$ und $F\colon I \to \mathbb{K}$ die Abbildung*

$$x \to F(x) = \int_a^x f(t)dt, \quad x \in I.$$

Dann folgt:

(1) *F ist Lipschitz stetig auf I,*
(2) *F ist auf $[a,b)$ rechtsseitig differenzierbar und auf $(a,b]$ linksseitig differenzierbar und es gilt*

$$\forall x \in [a,b)\colon F'_+(x) = f(x^+),$$
$$\forall x \in (a,b]\colon F'_-(x) = f(x^-).$$

BEWEIS. 1) Da $f|_{[a,x]}$ eine Regelfunktion ist, ist F sinnvoll definiert. Aus Satz VII-2.13 ergibt sich für $x, y \in I$, $x < y$, die Abschätzung

$$\begin{aligned}|F(x) - F(y)| &= |\int_a^x f(t)dt - \int_a^y f(t)dt| = |\int_a^x f(t)dt - \int_a^x f(t)dt - \int_x^y f(t)dt| \\ &= |\int_x^y f(t)dt| \leq \|f\| \cdot |y - x|.\end{aligned}$$

2) Wir führen den Beweis nur für die rechtsseitige Ableitung an $x \in [a,b)$. Für $h > 0$ (h so klein, daß $x + h \in I$) ergibt sich für den rechtsseitigen Differenzenquotienten von F

$$\frac{1}{h}\big(F(x+h) - F(x)\big) = \frac{1}{h}\Big(\int_a^{x+h} f(t)dt - \int_a^x f(t)\,dt\Big) = \frac{1}{h}\int_x^{x+h} f(t)\,dt.$$

Zu $\varepsilon > 0$ wählen wir nun $\delta > 0$ so, daß $|f(t) - f(x^+)| < \varepsilon$ für alle $t \in (x, x+\delta)$ zutrifft. Damit folgt für $0 < h < \delta$ mit Satz VII-2.9

$$\begin{aligned}\Big|\frac{1}{h}\big(F(x+h) - F(x)\big) - f(x^+)\Big| &= \Big|\frac{1}{h}\int_x^{x+h} f(t)\,dt - f(x^+)\Big| = \frac{1}{h}\Big|\int_x^{x+h} \big(f(t) - f(x^+)\big)\,dt\Big| \\ &\leq \frac{1}{h}\int_x^{x+h} |f(t) - f(x^+)|dt \leq \frac{1}{h}\int_x^{x+h} \sup_{t \in (x,x+h)} |f(t) - f(x^+)|dt < \varepsilon.\end{aligned}$$

$\square$

THEOREM VII-4.3 (Hauptsatz der Differential- und Integralrechnung).

1) Für jedes $f \in \mathcal{R}(I, \mathbb{K})$, $I = [a, b]$, ist die Abbildung $f \colon I \to \mathbb{K}$,

$$F(x) = \int_a^x f(t)\, dt, \quad x \in I,$$

eine Stammfunktion von f. Insbesondere ist F in den Stetigkeitsstellen von f differenzierbar und es gilt dort

$$F'(x) = f(x).$$

2) Mit einer beliebigen Stammfunktion Φ von f gilt

$$\int_a^b f(t)\, dt = \Phi(b) - \Phi(a).$$

Für die Differenz $\Phi(b) - \Phi(a)$ schreibt man auch $\Phi|_a^b$.

Beweis. 1) Nach Satz VII-4.2 ist F auf I stetig und in allen Stetigkeitsstellen von f gilt

$$F'_+(x) = f(x^+) = f(x) = f(x^-) = F'_-(x),$$

also

$$F'(x) = f(x),$$

d.h. F ist dort differenzierbar (in den Randpunkten von I natürlich nur rechts- bzw. linksseitig). Da nach Satz VII-4.1 f nur abzählbar viele Unstetigkeitsstellen besitzt, ist F eine Stammfunktion von f im Sinne von Definition VII-3.1.
2) Die Behauptung ist trivial für F. Für jede weitere Stammfunktion gilt nach Satz VII-3.6 $F(x) - \Phi(x) = c$, $x \in I$, für ein geeignetes $c \in \mathbb{K}$. Es folgt somit

$$\int_a^b f(t)\, dt = F(b) - F(a) = (\Phi(b) + c) - (\Phi(a) + c) = \Phi(b) - \Phi(a).$$

□

Korollar VII-4.4. Für jedes $F \in C^1(I, \mathbb{K})$, gilt für alle $x \in I$

$$\int_a^x F'(t)\, dt = F(x) - F(a).$$

Bemerkung VII-4.5. 1) Der Hauptsatz beinhaltet die theoretisch sehr zufriedenstellende Erkenntnis, daß zumindest jede integrierbare Funktion (Regelfunktion) eine Stammfunktion besitzt. Diese ist in der Form eines Integrals mit fester unterer und variabler oberer Grenze gegeben. Dieser Zusammenhang zwischen Integration und Differentiation ist umso bemerkenswerter, wenn man sich die höchst unterschiedlichen Ausgangspositionen für den Begriff der Stammfunktion und des Integrals vor Augen hält.

2) Für die Menge aller Stammfunktionen einer Regelfunktion f,

$$\{x \mapsto \Phi(x) = \int_a^x f(t)\,dt + c \colon c \in \mathbb{K}\}$$

ist auch das Symbol

$$\int f(x)\,dx \text{ oder } \int f\,dx \text{ oder } \int f$$

gebräuchlich, welches man **unbestimmtes Integral** von f nennt. Der Einfachheit halber wird dasselbe Symbol auch zur Bezeichnung irgendeiner Stammfunktion benützt, wie etwa in

$$\int \sin x\,dx = -\cos x, \qquad \int e^{\alpha x}\,dx = \frac{1}{\alpha}e^{\alpha x}, \quad \alpha \neq 0.$$

Diese Bezeichnung führt zu keinen Mißverständnissen, wenn man berücksichtigt, daß mit $\int f(x)\,dx = F$ auch $\int f(x)\,dx = F + c$, $c \in \mathbb{K}$, gilt. Hat man auf einem Intervall I die Beziehungen $\int f\,dx = F$ und $\int f\,dx = G$ gefunden, so ist es nicht zulässig, auf $F = G$ zu schließen. Vielmehr gilt dann $F - G = c$ für ein $c \in \mathbb{K}$.
3) Die Ableitung einer differenzierbaren Funktion $F \colon I \to \mathbb{K}$ muß keine Regelfunktion sein. Als Beispiel betrachte man $F \colon [-1, 1] \to \mathbb{R}$, gegeben durch

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

F ist überall differenzierbar, aber in $x = 0$ besitzt F' weder einen rechts-, noch eine linksseitigen Grenzwert. F' ist also keine Regelfunktion auf I. Insbesondere kann daher eine Stammfunktion von F' *nicht* durch eine Integration (nach Cauchy) gefunden werden.
4) Mit dem Hauptsatz läßt sich die oft mühsame Berechnung eines unbestimmten Integrals über die Definition zurückführen auf das Auffinden einer Stammfunktion. Der Hauptsatz ermöglicht es auch, die Produktregel und die Kettenregel aus der Differentialrechnung in häufig benutzte Integrationsregeln umzusetzen. Um die Beziehung $\int u'\,dx = u$ zu haben, formulieren wir diese Regeln für *stetig* differenzierbare Funktionen:

4.1. Partielle Integration.

Theorem VII-4.6. *Es sei $f \in C(I, \mathbb{K})$, $g \in C^1(I, \mathbb{K})$ und F eine Stammfunktion von f. Dann gilt*

$$\int_a^b f(x)g(x)\,dx = F(x)g(x)\Big|_a^b - \int_a^b F(x)g'(x)\,dx \qquad \textit{(partielle Integration).}$$

Beweis. Es gilt $F \in C^1(I, \mathbb{K})$ und somit $Fg \in C^1(I, \mathbb{K})$. Nach der Produktregel gilt

$$(Fg)' = F'g + Fg' = fg + Fg'.$$

Somit ist Fg eine Stammfunktion von $fg + Fg'$ und es folgt

$$(Fg)(b) - (Fg)(a) = \int\limits_a^b f(x)g(x)\,dx + \int\limits_a^b F(x)g'(x)\,dx.$$

Dies ist äquivalent zur Behauptung. □

Häufig wird die Regel für die partielle Integration in der einprägsameren Form

$$\int\limits_a^b u'v\,dx = uv\big|_a^b - \int\limits_a^b uv'\,dx$$

bzw. als unbestimmtes Integral

$$\int u'v\,dx = uv - \int uv'\,dx$$

formuliert.

Beispiel VII-4.7. 1) $\int x^\alpha \ln x\,dx = \frac{x^{\alpha+1}}{(\alpha+1)^2}((\alpha+1)\ln x - 1)$, $\alpha \neq -1$.
Wir definieren

$$f(x) = x^\alpha \qquad g(x) = \ln x$$

(bzw. $u'(x) = x^\alpha$, $v = \ln x$, für $x > 0$). Eine Stammfunktion von f ist gegeben durch $F(x) = \frac{x^{\alpha+1}}{\alpha+1}$. Wir erhalten somit

$$\begin{aligned}\int x^\alpha \ln x\,dx &= \frac{x^{\alpha+1}}{\alpha+1}\ln x - \int \frac{x^{\alpha+1}}{\alpha+1}\cdot\frac{1}{x}\,dx\\ &= \frac{x^{\alpha+1}}{\alpha+1}\ln x - \frac{x^{\alpha+1}}{(\alpha+1)^2}.\end{aligned}$$

2) $\int \sqrt{1-x^2}\,dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$, $x \in [-1, 1]$.
Für $|x| \leq 1 - \varepsilon$, $0 < \varepsilon < 1$ setzen wir

$$f(x) = 1, \qquad g(x) = \sqrt{1-x^2}$$

(bzw. $u'(x) = 1$, $v(x) = \sqrt{1-x^2}$), und erhalten

$$\begin{aligned}\int 1\cdot\sqrt{1-x^2}\,dx &= x\sqrt{1-x^2} - \int x\cdot\frac{(-x)}{\sqrt{1-x^2}}\,dx\\ &= x\sqrt{1-x^2} - \int \frac{1-x^2-1}{\sqrt{1-x^2}}\,dx\\ &= x\sqrt{1-x^2} + \int \frac{dx}{\sqrt{1-x^2}} - \int \sqrt{1-x^2}\,dx,\end{aligned}$$

d.h.

$$2\int \sqrt{1-x^2}\,dx = x\sqrt{1-x^2} + \int \frac{dx}{\sqrt{1-x^2}}.$$

Nach Beispiel VII-3.2 ist $\arcsin x$ eine Stammfunktion von $\frac{1}{\sqrt{1-x^2}}$ auf $(-1,1)$. Daraus folgt die behauptete Formel zunächst auf $(-1,1)$. Da $f\colon x \mapsto \sqrt{1-x^2}$ stetig auf $[-1,1]$ ist, besitzt f eine Stammfunktion auf I. Die Abbildung $x \mapsto \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)$ ist auf $[-1,1]$ stetig und stellt daher auch auf $[-1,1]$ eine Stammfunktion dar.

Als Anwendung der partiellen Integration beweisen wir folgende nützliche Variante des Satzes von Taylor:

Theorem VII-4.8. *Es sei $f \in C^{n+1}(I,\mathbb{K})$ und $x_0 \in I$. Dann gilt für alle $x \in I$*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt.$$

Beweis. Wir führen den Beweis mit vollständiger Induktion. Für $f \in C^1(I,\mathbb{K})$ ergibt der Hauptsatz VII-4.3 die Identität

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,dt, \quad x \in I.$$

Dies ist gerade die Behauptung für $n = 0$. Die Behauptung gelte nun für alle C^n-Funktionen und es sei $f \in C^{n+1}(I,\mathbb{K})$. Insbesondere gilt also $f \in C^n(I,\mathbb{K})$ und daher auch

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\,dt, \quad x \in I.$$

$f^{(n)} \in C^1(I,\mathbb{K})$, somit können wir partiell integrieren und erhalten

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k - \frac{(x-t)^n}{n!} f^{(n)}(t)\Big|_{x_0}^{x} + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt. \end{aligned}$$

$\square$

4.2. Integration durch Substitution. Die Kettenregel der Differentialrechnung ermöglicht die Integration durch Substitution. Diese Technik kann in zwei unterschiedlichen Situationen eingesetzt werden:

Fall 1: Der Integrand besitzt die Form $f(g(t))g'(t)$: Als Beispiel betrachte man den Integranden

$$h(t) = (\sin^3 t + e^{\sin t}) \cos t,$$

welcher offensichtlich die Ableitung von

$$H(t) = \frac{1}{4} \sin^4 t + e^{\sin t}$$

ist. H ist also Stammfunktion von h. Der folgende Satz beschreibt die allgemeine Situation.

Theorem VII-4.9 (1. Substitutionsregel). *Es sei $f \in C(I, \mathbb{K})$, $g \in C^1(J, \mathbb{R})$, $I = [a, b]$, $J = [\alpha, \beta]$ und $g(J) \subset I$. Dann gilt für alle $x, y \in J$*

$$\int\limits_x^y f(g(t))g'(t)\,dt = \int\limits_{g(x)}^{g(y)} f(u)du$$

und

$$\int f(g(t))g'(t)\,dt = \int f(u)du\Big|_{u=g(t)}.$$

Dabei bedeutet $\int f(u)du\big|_{u=g(t)}$ die Auswertung einer Stammfunktion von f an der Stelle $g(t)$.

Beweis. Da f stetig ist, besitzt f eine Stammfunktion $F \in C^1(I, \mathbb{K})$. Wegen $g(J) \subset \operatorname{def} F$ ist die Komposition $F \circ g$ möglich. Wir zeigen, daß $F \circ g$ eine Stammfunktion von $(f \circ g)g'$ ist: Nach der Kettenregel V-2.2 gilt für alle $t \in J$

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)$$

(die letzte Gleichheit gilt überall, weil F eine Stammfunktion von f ist und f stetig ist). Damit folgt

$$\int\limits_\alpha^\beta f(g(t))g'(t)\,dt = (F \circ g)(\beta) - (F \circ g)(\alpha) \ = F(g(\beta)) - F(g(\alpha)) \overset{\text{S.VII-4.3}}{=} \int\limits_{g(\alpha)}^{g(\beta)} f(u)du.$$

$\square$

Beispiel VII-4.10. 1) $\mathcal{I} = \int (\cos t + \cos^3 t)\,dt$.
Um den Satz anwenden zu können, schreiben wir das Integral in der Form

$$\mathcal{I}_t = \int (1 + \cos^2 t) \cos t\,dt = \int (2 - \sin^2 t) \cos t\,dt.$$

1. Schritt: *formale* Substitution: $g(t) = u$, $g'(t)\,dt = du$,
hier: $g(t) = \sin t = u$, $\cos t\,dt = du$,

$$\mathcal{I}_u = \int (2 - u^2)du.$$

2. Schritt: unbestimmte Integration nach u:

$$\mathcal{I}_u = 2u - \frac{1}{3}u^3.$$

3. Schritt: Rücksubstitution $u = g(t)$:

$$\mathcal{I}_t = 2\sin t - \frac{1}{3}\sin^3 t.$$

Diese formale Vorgangsweise spiegelt die Anwendung der Substitutionsregel folgendermaßen wieder: Es ist $f(g(t))g'(t) = (2-\sin^2 t)\cos t$ mit $g(t) = \sin t$ und $f(u) = 2-u^2$. Gemäß der Substitutionsregel hat man die Stammfunktion von f an der Stelle $u = \sin t$ auszuwerten.

2)

$$\mathcal{I} = \int_0^2 e^{t^2} t\,dt = \frac{1}{2}\int_0^2 e^{t^2} 2t\,dt.$$

1. Schritt: *formale* Substitution $t^2 = u$, $2t\,dt = du$.
2. Schritt: Transformation der Grenzen: $t = 0 \Rightarrow u = 0$, $t = 2 \Rightarrow u = 4$.
3. Schritt: Integration: $\mathcal{I} = \int_0^4 e^u du = e^4 - e^0 = e^4 - 1$.

Fall 2: Einsetzen einer beliebigen Substitutionsfunktion g. Dies ist der eigentliche Anwendungsbereich der Substitutionsregel. Man führt ein Integral $\int f(x)\,dx$ durch geschickte Wahl einer Substitutionsfunktion g in die Form

$$\int f(g(t))g'(t)\,dt,$$

über, in der Hoffnung, daß letzteres Integral einfacher als das Ausgangsintegral ist.

Theorem VII-4.11 (2. Substitutionsregel). *Es sei $f \in C(I,\mathbb{K})$, $g \in C^1(J,\mathbb{R})$, $J = [\alpha,\beta]$ und $g(J) \subset I$. Ferner sei g injektiv. Dann gilt für alle $u, v \in g(J)$*

$$\int_u^v f(x)\,dx = \int_{g^{-1}(u)}^{g^{-1}(v)} f(g(t))g'(t)\,dt$$

bzw.

$$\int f(x)\,dx = \int f(g(t))g'(t)\,dt\Big|_{t=g^{-1}(x)}.$$

Beweis. Wie im Beweis von Satz VII-4.9 sieht man, daß $F \circ g$ eine Stammfunktion von $(f \circ g)g'$ ist, falls F eine Stammfunktion von f ist. Für die linke Seite erhält man daher

$$\int_u^v f(x)\,dx = F(v) - F(u),$$

für die rechte

$$\int\limits_{g^{-1}(u)}^{g^{-1}(v)} f(g(t))g'(t)\,dt = (F\circ g)(g^{-1}(v)) - (F\circ g)(g^{-1}(u)) = F(v) - F(u).$$

□

Beispiel VII-4.12. 1) $\int \frac{dx}{\sin x} = \ln|\tan\frac{x}{2}|$.
Ist der Integrand eine rationale Funktion in sin und cos, $R(\sin x, \cos x)$, kann man immer – soferne sich nicht andere, einfachere Methoden anbieten – folgende Substitution ansetzen:

$$x = g(u) = 2\arctan u \quad \text{bzw} \quad u = \tan\frac{x}{2}.$$

Ersetzt man in den Identitäten

$$\cos 2z = \frac{1-\tan^2 z}{1+\tan^2 z}, \quad \sin 2z = \frac{2\tan z}{1+\tan^2 z}$$

z durch $\arctan u$, erhält man

$$\cos(g(u)) = \frac{1-u^2}{1+u^2}, \quad \sin(g(u)) = \frac{2u}{1+u^2}.$$

Wegen Satz VII-4.11 gilt die Identität

$$\int R(\sin x, \cos x)\,dx = \int R(\sin(g(u)), \cos g(u))g'(u)du\Big|_{u=g^{-1}(x)}$$
$$= \int R(\frac{2u}{1+u^2}, \frac{1-u^2}{1+u^2})\frac{2}{1+u^2}du\Big|_{u=\tan\frac{x}{2}}.$$

Umgelegt auf das konkrete Beispiel bedeutet dies

$$\int \frac{dx}{\sin x} = \int \frac{1+u^2}{2u}\frac{2}{1+u^2}du = \int \frac{du}{u} = \ln|u|\Big|_{u=\tan\frac{x}{2}} = \ln|\tan\frac{x}{2}|.$$

2) $\mathcal{I} = \int\limits_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{\pi}{2}$.
Dieses Integral stellt somit eine erste Verbindung zwischen der Zahl π und dem Einheitskreis her: Interpretiert man das Integral als Flächeninhalt, ergibt sich, daß $\frac{\pi}{2}$ – die kleinste positive Nullstelle des Kosinus – den Flächeninhalt eines Halbkreises mit Radius 1 angibt.
1. Schritt: Wahl der Substitutionsfunktion. Zu $f(x) = \sqrt{1-x^2}$ kann man, um die Wurzel zu eliminieren, als Substitutionsfunktion g wählen:

$$g(t) = \cos t, \quad t \in [0, \pi].$$

Man beachte, daß g auf $[0,\pi]$ streng monoton ist und $g([0,\pi]) = [-1, 1]$ gilt. Natürlich ist g stetig differenzierbar.
2. Schritt: formale Substitution: Auf $[0,\pi]$ gilt

$$f(g(t)) = \sqrt{1-\cos^2 t} = \sin t, \quad g'(t) = -\sin t.$$

Wir erhalten also das unbestimmte Integral

$$\int \sqrt{1-x^2}\,dx = -\int \sin^2 t\,dt.$$

3. Schritt: Transformation der Grenzen: $x=-1 \Rightarrow t = \arccos(-1) = \pi$, $x = 1 \Rightarrow t = \arccos 1 = 0$
4. Schritt: Integration durch Auswertung der Stammfunktion von $f(g(t))g'(t)$ in $t=0$ und $t=\pi$.
Zur Bestimmung der Stammfunktion $F \circ g$ von $t \mapsto -\sin^2 t$ verwenden wir die Umformung

$$\sin^2 t = \frac{1}{2} - \frac{1}{2}\cos 2t,$$

und erhalten somit unmittelbar

$$(F \circ g)(t) = -\frac{1}{2}t + \frac{1}{4}\sin 2t.$$

Nach Satz VII-4.12 erhalten wir

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = (F \circ g)(t)\Big|_{t=\pi}^{t=0} = \frac{1}{2}(-t\Big|_{\pi}^{0} + \frac{1}{2}\sin 2t\Big|_{\pi}^{0}) = \frac{\pi}{2}.$$

4.3. Partialbruchzerlegung. Ausgangspunkt ist folgende **kanonische Produktdarstellung** für Polynome, welche wir ohne Beweis mitteilen.

THEOREM VII-4.13. *Jedes Polynom p mit* $\operatorname{grad} p = n \geq 1$ *läßt sich mit Hilfe seiner paarweise verschiedenen Nullstellen $z_i \in \mathbb{C}$, $i = 1, \ldots, m$, in der Form*

$$p(z) = a_n \prod_{i=1}^{m} (z - z_i)^{\nu_i}$$

darstellen, wobei $a_n \in \mathbb{C}$, $a_n \neq 0$. Die Zahlen $\nu_i \in \mathbb{N}$, $i = 1, \ldots, m$, sind eindeutig bestimmt und erfüllen

$$\sum_{i=1}^{m} \nu_i = n.$$

*Man nennt ν_i die **Vielfachheit (Multiplizität)** der Nullstelle z_i.*

BEWEIS. Übung. □

Sind alle Koeffizienten des Polynoms $p(z) = \sum_{k=0}^{n} a_k z^k$ reell und ist $p(\xi) = 0$, so gilt wegen $a_k = \bar{a}_k$, $k = 1, \ldots, n$,

$$p(\bar{\xi}) = \sum_{k=0}^{n} a_k (\bar{\xi})^k = \sum_{k=0}^{n} a_k \overline{\xi^k} = \sum_{k=0}^{n} \overline{a_k \xi^k} = \overline{\sum_{k=0}^{n} a_k \xi^k} = \overline{p(\xi)} = 0.$$

Mit $\xi \in \mathbb{C}$ ist also auch $\bar{\xi} \in \mathbb{C}$ eine Nullstelle von p. Besitzt ξ die Vielfachheit μ, d.h. gilt

$$p(z) = (z - \xi)^\mu q(z), \quad q(\xi) \neq 0, \quad z \in \mathbb{C},$$

dann folgt mit einer ähnlichen Rechnung wie vorhin

$$p(\bar{z}) = \overline{p(z)} = (\bar{z} - \bar{\xi})^\mu q(\bar{z}), \quad \bar{z} \in \mathbb{C}.$$

Schreibt man für $\bar{z}$ wieder z, erhält man

$$p(z) = (z - \bar{\xi})^\mu q(z),$$

also ist $\bar{\xi}$ eine Nullstelle von p (dies wissen wir bereits) mit der Vielfachheit $\mu' \geq \mu$. Die erste Überlegung zeigt auch $q(\bar{\xi}) = \overline{q(\xi)} \neq 0$. Somit gilt $\mu = \mu'$ und die Vielfachheiten der Nullstellen ξ und $\bar{\xi}$ stimmen also überein.

Wir betrachten nun den Fall $\xi = \alpha + i\beta$, $\beta \neq 0$. Für jedes *reelle* x gilt dann

$$\begin{aligned}(x - \xi)(x - \bar{\xi}) &= (x - (\alpha + i\beta))(x - (\alpha - i\beta)) = (x - \alpha)^2 + \beta^2 \\ &= x^2 + Ax + B,\end{aligned}$$

mit

$$A = -2\alpha, \quad B = \alpha^2 + \beta^2.$$

Das quadratische Polynom $x^2 + Ax + B$ besitzt also keine *reellen* Nullstellen. Zusammen mit Satz VII-4.13 erhalten wir die *reelle kanonische Produktdarstellung* von p:

Theorem VII-4.14. *Es sei* $p(z) = \sum\limits_{k=0}^{n} a_k z^k$ *ein Polynom mit* reellen *Koeffizienten* a_k, $k = 0, \ldots, n$. *Dann gilt*

(1) *Komplexe Nullstellen* ξ *mit* $\operatorname{Im} \xi \neq 0$ *treten in konjugierten Paaren auf: Mit* ξ *ist auch* $\bar{\xi}$ *eine Nullstelle gleicher Multiplizität wie* ξ.

(2) *Sind* $x_1, \ldots, x_r$ *alle verschiedenen reellen Nullstellen von* p *mit der Multiplizität* ρ_i, $i = 1, \ldots, r$, *gilt die Darstellung*

$$p(x) = a_n \prod_{i=1}^{r} (x - x_i)^{\rho_i} \prod_{j=1}^{s} (x^2 + A_j x + B_j)^{\sigma_j}, \quad x \in \mathbb{R}.$$

Die Polynome $x^2 + A_j x + B_j$, $A_j, B_j \in \mathbb{R}$, $j = 1, \ldots, s$, *sind paarweise verschieden und besitzen keine reellen Nullstellen. Die natürlichen Zahlen* ρ_i, $i = 1, \ldots, r$, *und* σ_j, $j = 1, \ldots, s$, *erfüllen*

$$\sum_{i=1}^{r} \rho_i + 2 \sum_{j=1}^{s} \sigma_j = n.$$

Theorem VII-4.15 (Partialbruchzerlegung). *Es sei* $r = \frac{p}{q}$ *eine rationale Funktion mit* $\operatorname{grad} p < \operatorname{grad} q = n$. *Das Nennerpolynom habe die kanonische Produktdarstellung*

$$q(z) = a_n \prod_{j=1}^{m} (z - z_j)^{\nu_j}, \quad z \in \mathbb{C},$$

*mit $z_j \neq z_k$ für $j \neq k$. Dann besitzt r eine eindeutig bestimmte Summendarstellung (**Partialbruchzerlegung**) der Form*

$$r(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z - z_i)^j}, \quad z \in \mathbb{C} \setminus \{z_1, \dots, z_m\},$$

mit komplexen Zahlen a_{ij}.

Beweis. 1. Existenz: Wir führen den Beweis durch Induktion nach $\operatorname{grad} q = n$. Im Fall $n = 1$ gilt $\operatorname{grad} p < 1$, d.h. p ist eine Konstante $c \in \mathbb{C}$. Es gilt also

$$r(z) = \frac{c}{a_1(z - z_1)} = \frac{a_{11}}{z - z_1} \quad \text{mit} \quad a_{11} := \frac{c}{a_1}.$$

Induktionsschritt: Es sei $n > 1$. Wir nehmen an, der Satz sei für alle rationalen Funktionen $\frac{P}{Q}$ mit $\operatorname{grad} P < \operatorname{grad} Q \leq n - 1$ bewiesen. Wir schreiben q in der Form

$$q(z) = (z - z_1)^{\nu_1} s(z) \quad \text{mit} \quad s(z) = a_n \prod_{j=2}^{m} (z - z_j)^{\nu_j}.$$

Wegen $z_1 \neq z_i$, $i = 2, \dots, n$, ist $s(z_1) \neq 0$. Für $a \in \mathbb{C}$ berechnen wir

$$\frac{p(z)}{q(z)} - \frac{a}{(z - z_1)^{\nu_1}} = \frac{p(z) - as(z)}{(z - z_1)^{\nu_1} s(z)}.$$

Wählt man speziell $a = \frac{p(z_1)}{s(z_1)}$, gilt $p(z_1) - as(z_1) = 0$. Gilt sogar

$$p(z) - as(z) = 0 \quad \text{für } \textit{alle } z \in \mathbb{C},$$

folgt

$$r(z) = \frac{a}{(z - z_1)^{\nu_1}}$$

und wir sind fertig. Ist jedoch $p - as$ nicht das Nullpolynom, kann man nach Satz VII-4.13 den Linearfaktor $z - z_1$ abspalten, d.h. es gibt ein Polynom P mit

$$\begin{aligned} p(z) - as(z) &= (z - z_1) P(z) \\ \operatorname{grad} P &\leq \max\{\operatorname{grad} p, \operatorname{grad} s\} - 1 \leq n - 2. \end{aligned}$$

Somit folgt

$$r(z) = \frac{a}{(z - z_1)^{\nu_1}} + \frac{P(z)}{Q(z)},$$

mit

$$Q(z) = (z - z_1)^{\nu_1 - 1} s(z) = (z - z_1)^{\nu_1 - 1} a_n \prod_{j=2}^{m} (z - z_j)^{\nu_j},$$

also

$$\operatorname{grad} Q = n - 1.$$

Nach Induktionsvoraussetzung besitzt $\frac{P}{Q}$ eine Partialbruchzerlegung

$$\frac{P(z)}{Q(z)} = \sum_{j=1}^{\nu_1 - 1} \frac{a_{1j}}{(z-z_1)^j} + \sum_{i=2}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j},$$

und daher mit $a_{1\nu_1} = a$

$$r(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j}.$$

2. Eindeutigkeit. Es genügt zu zeigen, daß sämtliche Koeffizienten der Zerlegung von $r_0(z) \equiv 0$ verschwinden. Angenommen es ist

$$(*) \qquad r_0(z) = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} = 0, \quad z \in \mathbb{C} \setminus \{z_1, \ldots, z_m\}.$$

Wir betrachten eine feste Nullstelle z_{i_0} und wählen $1 \le j_0 \le \nu_{i_0}$ so, daß $a_{i_0 j} = 0$ für $j > j_0$ zutrifft. Multipliziert man $(*)$ mit $(z - z_{i_0})^{j_0}$, erhält man

$$0 = \Bigg(\sum_{\substack{i=1 \\ i \ne i_0}}^{\nu_i} \sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} \Bigg) (z - z_{i_0})^{j_0} + \sum_{j=1}^{j_0} (z - z_{i_0})^{j_0 - j} a_{i_0 j}.$$

Durch Grenzübergang $z \to z_{i_0}$ erhält man $a_{i_0 j_0} = 0$. Schrittweise ergibt sich auf diese Weise für sämtliche Koeffizienten $a_{ij} = 0$. $\square$

Für die praktische Berechnung der Partialbruchzerlegung gibt es mehrere Möglichkeiten:

1. **Koeffizientenvergleich:** Multipliziert man die Darstellung von $r = \frac{p}{q}$ mit q, erhält man

$$\begin{aligned} p(z) &= \sum_{i=1}^{m} \Bigg(\sum_{j=1}^{\nu_i} \frac{a_{ij}}{(z-z_i)^j} \Bigg) a_n \prod_{k=1}^{m} (z - z_k)^{\nu_k} \\ &= a_n \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} a_{ij} (z - z_i)^{\nu_i - j} \prod_{\substack{k=1 \\ k \ne i}}^{m} (z - z_k)^{\nu_k}, \quad z \in \mathbb{C}. \end{aligned}$$

Der Identitätssatz IV-4.12 ist nun die Grundlage für den Koeffizientenvergleich, bei dem man die Koeffizienten gleicher Potenzen von z^k in p und im Polynom auf der rechten Seite gleichsetzt. Dies führt auf ein lineares Gleichungssystem für die Unbekannten a_{ij}, welches wegen der Existenz- und Eindeutigkeitsaussage in Satz VII-4.15 genau eine Lösung besitzen *muß*.

2. **Substitutionsmethode:** Ein anderes lineares Gleichungssystem für a_{ij} läßt sich gewinnen, wenn man für z nacheinander n verschiedene und möglichst zweckmäßig gewählte Werte $\xi_1, \ldots, \xi_n$ einsetzt.

3. **Grenzwertmethode:** Die „höchsten“ Koeffizienten $a_{i\nu_i}$ erhält man besonders einfach, indem man die Darstellung für $r(z)$ mit $(z-z_i)^{\nu_i}$ multipliziert und den Grenzwert $z \to z_i$ betrachtet (in der Praxis bedeutet dies natürlich, daß man nach der Multiplikation den gekürzten Ausdruck in $z = z_i$ auswertet). Hat man $a_{i\nu_i}$ berechnet, kann man den Term $\frac{a_{i\nu_i}}{(z-z_i)^{\nu_i}}$ auf die linke Seite bringen und das Verfahren mit dem nächstniedrigeren Koeffizienten $a_{i\nu_i-1}$ wiederholen.

Beispiel VII-4.16. 1) $r(z) = \frac{z+1}{z(z-1)(z-i)(z+i)}$.
Da sämtliche Nullstellen des Nennerpolynoms Vielfachheit 1 haben, führt folgender Ansatz zum Ziel ($a_{11} = A$, $a_{21} = B$, $a_{31} = C$, $a_{41} = D$):

$$(*) \qquad r(z) = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{z-i} + \frac{D}{z+i}.$$

a) Koeffizientenvergleich:
Wir multiplizieren $(*)$ mit $z(z-1)(z-i)(z+i)$ und erhalten die Identität

$$\begin{aligned} z+1 &= A(z-1)(z^2+1) + Bz(z^2+1) + Cz(z-1)(z+i) + Dz(z-1)(z-i) \\ &= A(z^3-z^2+z-1) + B(z^3+z) + C(z^3+(i-1)z^2-iz) \\ &\quad + D(z^3-(1+i)z^2+iz) \\ &= (A+B+C+D)z^3 + (-A+(i-1)C-(1+i)D)z^2 \\ &\quad + (A+B-iC+iD)z - A. \end{aligned}$$

Vergleicht man die Koeffizienten gleicher Potenzen, erhält man das Gleichungssystem

$$\begin{array}{rrrrcl} A & +B & +C & +D & = & 0 \\ -A & & +(i-1)C & -(1+i)D & = & 0 \\ A & +B & -iC & +iD & = & 1 \\ -A & & & & = & 1\ . \end{array}$$

b) Substitutionsmethode:
Setzt man nacheinander in $(*)$ etwa $z = -1, -2, 2, 3$ ein, erhält man das Gleichungssystem

$$\begin{array}{rrcrrrr} z = & -1: & 0 & = & -A & -\frac{1}{2}B & -\frac{1}{1+i}C & +\frac{1}{-1+i}D \\ z = & -2: & \frac{-1}{30} & = & -\frac{1}{2}A & -\frac{1}{3}B & -\frac{1}{2+i}C & +\frac{1}{-2+i}D \\ z = & 2: & \frac{3}{10} & = & \frac{1}{2}A & +B & +\frac{1}{2-i}C & +\frac{1}{2+i}D \\ z = & 3: & \frac{4}{60} & = & \frac{1}{3}A & +\frac{1}{2}B & +\frac{1}{3-i}C & +\frac{1}{3+i}D\ . \end{array}$$

c) Grenzwertmethode:
Wir multiplizieren $(*)$ mit z. Dies ergibt

$$\frac{z+1}{(z-1)(z-i)(z+i)} = A + z\left(\frac{B}{z-1} + \frac{C}{z-i} + \frac{D}{z+i}\right).$$

Für $z = 0$ erhält man $A = -1$. Multipliziert man mit $z - 1$, folgt

$$\frac{z+1}{z(z-i)(z+i)} = B + (z-1)\left(-\frac{1}{z} + \frac{C}{z-i} + \frac{D}{z+i}\right),$$

für $z = 1$ ergibt sich $B = 1$. Ganz entsprechend findet man $C = \frac{i}{2}$, $D = -\frac{i}{2}$ und damit die Partialbruchzerlegung

$$\frac{z+1}{z(z-1)(z-i)(z+i)} = -\frac{1}{z} + \frac{1}{z-1} + \frac{1}{2}\frac{i}{z-i} - \frac{1}{2}\frac{i}{z+i}.$$

2) $r(z) = \frac{z^2+1}{z(z-1)^2}$.
1 ist eine zweifache Nullstelle des Nennerpolynoms, somit ist ($a_{11} = A$, $a_{21} = B$, $a_{22} = C$)

$$r(z) = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{(z-1)^2}$$

anzusetzen. Multipliziert man mit z und wertet in $z = 0$ aus, erhält man $A = 1$. Multiplikation mit $(z-1)^2$ liefert

$$\frac{z^2+1}{z} = \frac{(z-1)^2}{z} + B(z-1) + C.$$

Für $z = 1$ erhält man also $C = 2$. Setzt man nun in

$$r(z) = \frac{z^2+1}{z(z-1)^2} = \frac{1}{z} + \frac{B}{z-1} + \frac{2}{(z-1)^2}$$

$z = 2$ ein, erhält man

$$\frac{5}{2} = \frac{1}{2} + B + 2,$$

also $B = 0$. Die gesuchte Partialbruchzerlegung lautet somit

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

Das Beispiel zeigt, daß es vorteilhaft sein kann, verschiedene Methoden zur Berechnung der Partialbruchzerlegungen zu kombinieren.

Ist die rationale Funktion reell, d.h. sind alle Koeffizienten der Polynome p und q reell, ist es z.B. für Anwendungen in der Integralrechnung oft zweckmäßig, eine *reelle* Partialbruchzerlegung zur Verfügung zu haben.

Theorem VII-4.17. *Es sei $r = \frac{p}{q}$ eine rationale Funktion mit* $\operatorname{grad} p < \operatorname{grad} q = n$. *Sämtliche Koeffizienten der Polynome p und q seien reell. Das Nennerpolynom besitze die reelle kanonische Produktdarstellung*

$$q(x) = a_n \prod_{i=1}^{r} (x - x_i)^{\rho_i} \prod_{j=1}^{s} (x^2 + A_j x + B_j)^{\sigma_j}, \quad x \in \mathbb{R}$$

(vgl. Satz VII-4.14). Dann besitzt r eine eindeutig bestimmte Partialbruchzerlegung der Form

$$r(x) = \sum_{i=1}^{r}\sum_{j=1}^{\rho_i} \frac{a_{ij}}{(x-x_i)^j} + \sum_{i=1}^{s}\sum_{j=1}^{\sigma_i} \frac{\alpha_{ij}x+\beta_{ij}}{(x^2+A_ix+B_i)^j}, \quad x \in \mathbb{R},$$

wobei die Koeffizienten a_{ij}, α_{ij}, β_{ij} reelle Zahlen sind.

Beispiel VII-4.18. $r(x) = \frac{x+1}{x(x-1)(x^2+x+1)}, x \in \mathbb{R} \setminus \{0,1\}$.
Die einzigen reellen Nullstellen von $q(x) = x(x-1)(x^2+x+1)$, $x = 0$ und $x = 1$, besitzen Multiplizität 1. Somit lautet der Ansatz für die reelle Partialbruchzerlegung

$$r(x) = \frac{A}{x} + \frac{B}{x-1} + \frac{Cx+D}{x^2+x+1}.$$

Die Koeffizienten A und B erhält man leicht mit Hilfe der Grenzwertmethode. Man findet

$$A = -1, \quad B = \frac{2}{3}.$$

Es ist also

$$\frac{x+1}{x(x-1)(x^2+x+1)} = -\frac{1}{x} + \frac{2}{3}\frac{1}{x-1} + \frac{Cx+D}{x^2+x+1}.$$

Setzt man die Werte $x = -1$ und $x = 2$ ein, erhält man das Gleichungssystem

$$0 = 1 - \frac{1}{3} - C + D$$
$$\frac{3}{14} = -\frac{1}{2} + \frac{2}{3} + \frac{2}{7}C + \frac{1}{7}D,$$

woraus $C = \frac{1}{3}$ und $D = -\frac{1}{3}$ folgt. Insgesamt gilt also die Darstellung

$$f(x) = -\frac{1}{x} + \frac{2}{3}\frac{1}{x-1} + \frac{1}{3}\frac{x-1}{x^2+x+1}.$$

Die reelle bzw. komplexe Partialbruchzerlegung ermöglicht die systematische, geschlossene Integration rationaler Funktionen $r(x) = \frac{p(x)}{q(x)}$. Im Falle $\operatorname{grad} p \geq \operatorname{grad} q$ ist vorerst der polynomiale Anteil in r durch Division von p durch q abzuspalten. Wir demonstrieren die Methode an einigen einfachen Beispielen:

Beispiel VII-4.19. 1) Mit Hilfe von Beispiel VII-4.16-1) finden wir

$$\begin{aligned}\int \frac{(x+1)\,dx}{x(x-1)(x^2+1)} &= -\int \frac{dx}{x} + \int \frac{dx}{x-1} + \frac{i}{2}\int \left(\frac{1}{x-i} - \frac{1}{x+i}\right) dx \\ &= -\ln|x| + \ln|x-1| + \frac{i}{2}\int \frac{2i}{x^2+1}\,dx \\ &= -\ln|x| + \ln|x-1| - \arctan x.\end{aligned}$$

(Wir haben $\frac{1}{x-i} - \frac{1}{x+i}$ durch $\frac{2i}{x^2+1}$ ersetzt, da wir den natürlichen Logarithmus vorerst nur für *reelle* Argumente definiert haben).

2) Aus Beispiel VII-4.16-2) folgt

$$\int \frac{x^2+1}{x(x-1)^2}\,dx = \int \frac{dx}{x} + 2\int \frac{dx}{(x-1)^2} = \ln|x| - \frac{2}{x-1}.$$

3) Wir benützen Beispiel VII-4.18 in

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

Lediglich das letzte Integral erfordert etwas Aufwand:

$$\begin{aligned}\int \frac{x-1}{x^2+x+1}\,dx &= \frac{1}{2}\int \frac{2x+1}{x^2+x+1}\,dx - \frac{3}{2}\int \frac{dx}{x^2+x+1}\\ &= \frac{1}{2}\ln(x^2+x+1) - \frac{3}{2}\int \frac{dx}{x^2+x+1}\end{aligned}$$

(man beachte $x^2+x+1>0$ für $x\in\mathbb{R}$). Den Integranden im letzten Integral formt man um zu

$$x^2+x+1 = (x+\frac{1}{2})^2 + \frac{3}{4} = \frac{3}{4}\big((\frac{2}{\sqrt{3}}(x+\frac{1}{2}))^2+1\big).$$

Substituiert man $g(x) = \frac{2}{\sqrt{3}}(x+\frac{1}{2})$, erhält man

$$\int \frac{dx}{x^2+x+1} = \frac{2}{\sqrt{3}}\arctan(\frac{2}{\sqrt{3}}(x+\frac{1}{2})).$$

Insgesamt ergibt sich also

$$\begin{aligned}\int \frac{x+1}{x(x-1)(x^2+x+1)}\,dx = &-\ln|x| + \frac{2}{3}\ln|x-1| + \frac{1}{6}\ln(x^2+x+1)\\ &-\sqrt{3}\arctan(\frac{2}{\sqrt{3}}(x+\frac{1}{2})).\end{aligned}$$

Bemerkung VII-4.20. Im Prinzip ist es nicht notwendig, für die Integration einer *reellen* rationalen Funktion, die *reelle* Partialbruchzerlegung aus Satz VII-4.17 zu verwenden. Man kann natürlich auch die komplexe Partialbruchzerlegung zur Integration heranziehen und erhält

$$\begin{aligned}\int r(x)\,dx &= \sum_{k=1}^{m}\sum_{j=1}^{\nu_k}\int \frac{a_{kj}}{(x-z_k)^j}\,dx\\ &= \sum_{k=1}^{m}\Big[\int \frac{a_{k1}}{x-z_k}\,dx - \sum_{j=2}^{\nu_k}\frac{a_{kj}}{j-1}(x-z_k)^{-j+1}\Big].\end{aligned}$$

Problematisch ist nur die Berechnung der Stammfunktion von $(x-z_k)^{-1}$ für $z_k\in\mathbb{C}\backslash\mathbb{R}$, da die Logarithmen nur für reelle Argumente definiert wurden. Es sei $z_k=\alpha_k+i\beta_k$, $\beta_k\neq 0$. Mit der Umformung

$$\frac{1}{x-z_k} = \frac{x-\alpha_k+i\beta_k}{(x-\alpha_k)^2+\beta_k^2}$$

erhalten wir

$$\begin{aligned}\int \frac{dx}{x-z_k} &= \int \frac{x-\alpha_k}{(x-\alpha_k)^2+\beta_k^2}\,dx + i\beta_k \int \frac{dx}{(x-\alpha_k)^2+\beta_k^2}\\ &= \frac{1}{2}\int \frac{2(x-\alpha_k)}{(x-\alpha_k)^2+\beta_k^2}\,dx + i\beta_k \int \frac{dx}{(x-\alpha_k)^2+\beta_k^2}\\ &= \frac{1}{2}\ln((x-\alpha_k)^2+\beta_k^2) + i\arctan\frac{x-\alpha_k}{\beta_k}.\end{aligned}$$

Der gesamte Ausdruck für $\int r(x)\,dx$ ist natürlich wieder reell. Allgemein sind jedoch die *reellen* Gleichungssysteme, auf welche die reelle Partialbruchzerlegung führt, händisch leichter zu lösen, als Gleichungssysteme mit komplexen Koeffizienten.

5. Integration und Grenzübergang

In diesem Abschnitt untersuchen wir die Vertauschbarkeit von Limes und Integral. Diese ist im allgemeinen *nicht* gegeben:

Beispiel VII-5.1. $f_n(x) = \begin{cases} n & x \in [\frac{1}{2n}, \frac{1}{n}] \\ 0 & x \in [0,1] \setminus [\frac{1}{2n}, \frac{1}{n}] \end{cases}.$

Es gilt $\lim f_n(x) = f(x) = 0$ für alle $x \in [0,1]$ und $\int_0^1 f_n(x)\,dx = \frac{1}{2}$, $n \in \mathbb{N}$, und $\int_0^1 f(x)\,dx = 0$. Man beachte, daß in diesem Beispiel die Folge (f_n) nur punktweise gegen f konvergiert und darüber hinaus $\{\|f_n\| : n \in \mathbb{N}\}$ unbeschränkt ist.

In Hinblick auf die Definition des Cauchy Integrals ist es jedoch nicht weiter überraschend, daß die Integrale einer gleichmäßig konvergenten Folge von Regelfunktionen gegen das Integral der Grenzfunktion konvergieren (diese Eigenschaft ist gewissermaßen von Haus aus in das Integral eingebaut):

Theorem VII-5.2. *Es sei $(f_n)_{n\geq 1} \subset \mathcal{R}(I,\mathbb{K})$, $f\colon I \to \mathbb{K}$, und $f_n \rightrightarrows f$. Dann ist f eine Regelfunktion, die Folge der Integrale $(\mathcal{I}(f_n))$ ist konvergent und es gilt*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Beweis. Wegen Satz VII-1.7 gilt $f \in \mathcal{R}(I,\mathbb{K})$. Aus der Beschränktheit des Integrals Satz VII-2.9 folgt

$$|\mathcal{I}(f_n) - \mathcal{I}(f)| = |\mathcal{I}(f_n - f)| \leq (b-a)\|f_n - f\|.$$

□

Korollar VII-5.3. Es sei $(f_n) \subset \mathcal{R}(I,\mathbb{K})$ und die Reihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig konvergent. Dann gilt

$$\sum_{n=0}^{\infty} \int_a^b f_n(x)\,dx = \int_a^b \sum_{n=0}^{\infty} f_n(x)\,dx.$$

Dieses Ergebnis eröffnet eine weitere Möglichkeit, Potenzreihenentwicklungen zu berechnen:

Beispiel VII-5.4. Aus der Binomialreihe (Beispiel VI-7.13) ergibt sich für $|t| < 1$ die Reihenentwicklung

$$(1-t^2)^{-\frac{1}{2}} = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} t^{2k}.$$

Auf kompakten Teilintervallen von $(-1,1)$ ist die Potenzreihe gleichmäßig konvergent. Auf $[0,x]$ bzw. $[x,0]$, $|x| < 1$, kann nach Korollar VII-5.3 die Reihe gliedweise integriert werden: Es gilt somit

$$\int\limits_0^x \frac{dt}{\sqrt{1-t^2}} = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \int\limits_0^x t^{2k}\,dt = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \frac{x^{2k+1}}{2k+1}.$$

Die Stammfunktion auf der linken Seite ist $\arcsin x$. Die rechte Seite ergibt mit $(k \geq 1)$

$$\binom{-\frac{1}{2}}{k} = \frac{1}{k!} \prod_{j=0}^{k-1} \left(-\frac{1}{2} - j\right) = \frac{1}{k!} (-1)^k \frac{1}{2^k} \prod_{j=0}^{k-1} (1+2j)$$

die Reihenentwicklung

$$\arcsin x = x + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{1}{2^k} \frac{1}{2k+1} \prod_{j=0}^{k-1} (1+2j) x^{2k+1},$$

die zumindest für $|x| < 1$ konvergiert. Tatsächlich läßt sich zeigen, daß diese Darstellung sogar für $|x| \leq 1$ gültig ist.

Unser nächstes Ziel ist es, Satz VII-5.2 erheblich zu verallgemeinern. Dazu benötigen wir zwei technische Resultate. Im folgenden wollen wir unter einer **Elementarmenge** eine endliche Vereinigung von disjunkten, beschränkten Intervallen verstehen. Ist $E \subset \mathbb{R}$ eine Elementarmenge, d.h. ist

$$E = \bigcup_{i=1}^{n} J_i,$$

$J_i \subset \mathbb{R}$ beschränktes Intervall, $i = 1, \ldots, n$, und bezeichnet $\lambda(J)$ die Länge eines Intervalls J, dann nennen wir

$$\lambda(E) = \sum_{i=1}^{n} \lambda(J_i)$$

Länge von E. Wir übergehen hier den Nachweis, daß $\lambda(E)$ unabhängig ist von der speziellen Darstellung $\bigcup_{i=1}^{n} J_i$ von E. Ohne Beweis zitieren wir folgendes Resultat:

LEMMA VII-5.5. *Es sei* $(E_n)_{n\geq 1}$ *eine Folge von Elementarmengen mit* $E_n \subset I = [a, b]$, $n \in \mathbb{N}$, *und es existiere* $c > 0$ *so, daß*

$$\forall n \in \mathbb{N} \colon \lambda(E_n) \geq c > 0.$$

Dann gibt es ein Element $x_0 \in I$, *das zu unendlich vielen Elementarmengen gehört.*

LEMMA VII-5.6. *Es sei* $t \in \mathcal{T}(I, \mathbb{K})$, $I = [a, b]$. *Gilt für ein* $\varepsilon > 0$

$$\left|\int_a^b t(x)\,dx\right| \geq \varepsilon,$$

dann ist die Menge

$$E = \{x \in I \colon |t(x)| \geq \frac{\varepsilon}{2(b-a)}\}$$

eine Elementarmenge, für deren Länge die Abschätzung

$$\lambda(E) \geq \frac{\varepsilon}{2\|t\|}$$

zutrifft.

BEWEIS. Es sei $\mathcal{Z} \colon a = x_0 < \cdots < x_i < \cdots < x_n = b$ eine zu t gehörende Zerlegung von I und es sei c_i der Wert von t auf $J_i = (x_{i-1}, x_i)$. Dann ist klar, daß E eine endliche Vereinigung von disjunkten Intervallen ist (es können nur Intervalle des Typs J_i oder $[x_i, x_i]$ auftreten). Es gilt

$$\int_a^b t(x)\,dx = \sum_{i=1}^n c_i(x_i - x_{i-1}) = \sum_{i=1}^n c_i \lambda(J_i).$$

Wir spalten nun die Summe folgendermaßen auf:

$$\sum_{i=1}^n c_i \lambda(J_i) = \sum_{|c_i| \geq \frac{\varepsilon}{2(b-a)}} c_i \lambda(J_i) + \sum_{|c_i| < \frac{\varepsilon}{2(b-a)}} c_i \lambda(J_i).$$

Dies ergibt die Abschätzung

$$\begin{aligned}\varepsilon \leq \left|\int_a^b t(x)\,dx\right| &\leq \sum_{|c_i| \geq \frac{\varepsilon}{2(b-a)}} |c_i| \lambda(J_i) + \sum_{|c_i| < \frac{\varepsilon}{2(b-a)}} |c_i| \lambda(J_i) \\ &\leq \|t\| \cdot \lambda(E) + \frac{\varepsilon}{2(b-a)} \cdot (b-a)\end{aligned}$$

und somit (man beachte $t \neq 0$ und daher auch $\|t\| > 0$)

$$\lambda(E) \geq \frac{\varepsilon}{2\|t\|}.$$

□

THEOREM VII-5.7. *Es sei* $(t_n)_{n\geq 1} \subset \mathcal{T}(I, \mathbb{K})$ *mit*

i) $\lim_{n\to\infty} t_n(x) = 0$ *für alle* $x \in I$ *(*t_n *konvergiert punktweise gegen* 0*),*
ii) $\exists c > 0\, \forall n \in \mathbb{N}\colon \|t_n\| \le c$ *((*t_n*) ist beschränkt).*

Dann gilt

$$\lim_{n\to\infty} \int_a^b t_n(x)\, dx = 0.$$

Beweis. Angenommen, die Folge $(\mathcal{I}(t_n))$ wäre keine Nullfolge, dann gäbe es ein $\varepsilon_0 > 0$ und eine Teilfolge $(\mathcal{I}(t_{\varphi(n)}))$ mit $|\mathcal{I}(t_{\varphi(n)})| \ge \varepsilon_0$, $n \in \mathbb{N}$. Wir bezeichnen der Einfachheit halber nun diese Teilfolge wieder mit $\mathcal{I}(t_n)$ und gehen daher von der Annahme aus:

$$\forall n \in \mathbb{N}\colon |\mathcal{I}(t_n)| \ge \varepsilon_0.$$

Nach Lemma VII-5.6 besitzt für jedes t_n die Elementarmenge

$$E_n = \{x \in I\colon |t_n(x)| \ge \frac{\varepsilon_0}{2(b-a)}\},$$

die Länge

$$\lambda(E_n) \ge \frac{\varepsilon_0}{2\|t_n\|}.$$

Wegen $\|t_n\| \le c$, $n \in \mathbb{N}$, gilt daher

$$\lambda(E_n) \ge \frac{\varepsilon_0}{2c}, \quad n \in \mathbb{N}.$$

Nach Lemma VII-5.5 gibt es daher eine Stelle $x_0 \in I$, welche in unendlich vielen Elementarmengen E_n liegt. Für unendlich viele $n \in \mathbb{N}$ gilt also

$$|t_n(x_0)| \ge \frac{\varepsilon_0}{2(b-a)},$$

im Widerspruch zur Voraussetzung $\lim_{n\to\infty} t_n(x_0) = 0$. □

Theorem VII-5.8 (Arzela-Osgood). *Die Folge* $(f_n) \subset \mathcal{R}(I, \mathbb{K})$ *konvergiere punktweise gegen die Regelfunktion* f*. Ferner sei* $\{\|f_n\|\colon n \in \mathbb{N}\}$ *beschränkt. Dann gilt*

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx = \int_a^b \lim_{n\to\infty} f_n(x)\, dx.$$

Wir bemerken, daß aus der *punktweisen* Konvergenz von f_n gegen eine Funktion f nicht gefolgert werden kann, daß f eine Regelfunktion ist. Es ist daher notwendig, $f \in \mathcal{R}(I, \mathbb{K})$ vorauszusetzen.

Beweis. Da f nach Voraussetzung eine Regelfunktion ist, gilt auch $f_n - f \in \mathcal{R}(I, \mathbb{K})$. Da $(f - f_n)$ eine Nullfolge und $(\|f - f_n\|)$ beschränkt ist, genügt es wegen der Linearität des Integrals den Satz für den Spezialfall

$$(*) \qquad \lim_{n\to\infty} f_n(x) = 0, \quad x \in I,$$

zu beweisen. Zu jeder Regelfunktion f_n wählen wir – wie im Beweis von Satz VII-1.7 – eine Treppenfunktion $t_n \in \mathcal{T}(I, \mathbb{K})$ mit

$$\|t_n - f_n\| \leq \frac{1}{n}, \quad n \in \mathbb{N}.$$

Daraus folgt die Abschätzung

$$\|t_n\| \leq \|f_n\| + \|f_n - t_n\| < \|f_n\| + \frac{1}{n},$$

aus welcher wir die Beschränktheit von $(\|t_n\|)$ ablesen. Insbesondere folgt aber auch für jedes $x \in I$

$$|t_n(x)| \leq |f_n(x) - t_n(x)| + |f_n(x)| \leq \frac{1}{n} + |f_n(x)|.$$

Somit konvergiert auch die Folge von Treppenfunktionen (t_n) punktweise gegen die Nullfunktion. Aus Satz VII-5.7 folgt daher

$$\lim_{n\to\infty} \mathcal{I}(t_n) = 0.$$

Zusammen mit der Beschränktheit des Integrals ergibt sich daraus die Behauptung

$$\begin{aligned} |\mathcal{I}(f_n)| &\leq |\mathcal{I}(t_n)| + |\mathcal{I}(t_n - f_n)| \leq |\mathcal{I}(t_n)| + (b-a)\|t_n - f_n\| \\ &< |\mathcal{I}(t_n)| + \frac{1}{n}(b-a). \end{aligned}$$

□

Wir formulieren den Satz von Arzela-Osgood auch für Funktionenreihen:

Korollar VII-5.9. Es sei $(f_n) \subset \mathcal{R}(I, \mathbb{K})$. Die Reihe $\sum_{k=1}^\infty f_k$ konvergiere punktweise gegen eine Regelfunktion. Ferner sei die Folge der Partialsummen gleichmäßig beschränkt, d.h.

$$\exists c \geq 0 \, \forall x \in I \, \forall n \in \mathbb{N} \colon |\sum_{k=1}^n f_k(x)| \leq c.$$

Dann gilt

$$\sum_{k=0}^\infty \left(\int_a^b f_k(x)\, dx \right) = \int_a^b \left(\sum_{k=0}^\infty f_k(x) \right) dx.$$

6. Parameterabhängige Integrale

Wir gehen von einer Funktion $f \colon I \times D \to \mathbb{K}$, $D \subset \mathbb{K}$, mit folgender Eigenschaft aus:

$$\forall t \in D \colon x \mapsto f(x,t) \text{ ist eine Regelfunktion.}$$

(Die Abbildung $x \mapsto f(x,t)$ bezeichnen wir mit $f(\cdot,t)$). Man bezeichnet t in diesem Zusammenhang oft als **Parameter**. Definiert man

$$F : \begin{cases} D \to \mathbb{K} \\ t \mapsto \int\limits_a^b f(x,t)\,dx, \end{cases}$$

sagt man, F sei durch ein **parameterabhängiges Integral** gegeben. Wir untersuchen nun die Eigenschaften von F.

THEOREM VII-6.1. $f \in \mathcal{B}(I \times D, \mathbb{K})$, $D \subset \mathbb{K}$, *habe folgende Eigenschaften:*

i) $\forall t \in D\colon f(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
ii) $\forall x \in I\colon f(x,\cdot) \in C(D,\mathbb{K})$.

Dann ist die Abbildung $F\colon D \to \mathbb{K}$,

$$F(t) = \int\limits_a^b f(x,t)\,dx$$

stetig auf D.

BEWEIS. Es sei $t_0 \in D$ ein Häufungspunkt von D und $(t_n) \subset D$ konvergiere gegen t_0. Wir betrachten nun die Folge von Regelfunktionen (g_n),

$$g_n(x) = f(x,t_n), \quad x \in I, n \in \mathbb{N}.$$

Aus ii) folgt für alle $x \in I$

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} f(x,t_n) = f(x,t_0).$$

Es gilt $f(\cdot,t_0) \in \mathcal{R}(I,\mathbb{K})$ und wegen $f \in \mathcal{B}(I \times D,\mathbb{K})$ ist die Folge $(\|g_n\|)$, d.h. $(\|f(\cdot,t_n)\|)$ beschränkt. Nach dem Satz von Arzela-Osgood VII-5.8 folgt daher

$$\begin{aligned}\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty} \int\limits_a^b f(x,t_n)\,dx &= \lim_{n\to\infty} \int\limits_a^b g_n(x)\,dx = \int\limits_a^b \lim_{n\to\infty} g_n(x)\,dx \\ &= \int\limits_a^b f(x,t_0)\,dx = F(t_0).\end{aligned}$$

□

Der Beweis zeigt, daß die Forderung der globalen Beschränktheit von f abgeschwächt werden kann zu

$$\forall t \in D\, \exists U \in \mathcal{U}(t)\colon f \text{ ist beschränkt auf } I \times (U \cap D).$$

Beispiel VII-6.2. 1) Es sei $I = [a,b]$, $0 < a < b$, $D = \mathbb{R}$, und $f(x,t) = x^t$. Aus dem Satz folgt, daß

$$F(t) = \begin{cases} \frac{b^{t+1}-a^{t+1}}{t+1}, & t \neq -1, \\ \ln \frac{b}{a}, & t = -1, \end{cases}$$

auf beliebigen kompakten Intervallen von $\mathbb{R}$ und damit auf $\mathbb{R}$ selbst stetig ist. Insbesondere folgt

$$\lim_{t\to -1} \frac{b^{t+1} - a^{t+1}}{t+1} = F(-1) = \ln \frac{b}{a}.$$

2) Es sei $I = [a,b] \subset \mathbb{R}$, $D = (0,\infty)$ und $f(x,t) = t^x$. Mit Hilfe von Satz VII-6.1 schließen wir auf die Stetigkeit von

$$F(t) = \begin{cases} \frac{1}{\ln t}(t^b - t^a), & t \in (0,\infty) \setminus \{1\}, \\ b - a, & t = 1, \end{cases}$$

und insbesondere $F(1) = \int\limits_a^b 1\, dx = b - a = \lim_{t\to 1} \frac{1}{\ln t}(t^b - t^a)$.

Das folgende Resultat benötigt den Begriff der partiellen Ableitungen, der im Abschnitt über die Differenzierbarkeit von Funktionen in mehreren Veränderlichen diskutiert wird.

Theorem VII-6.3 (Vertauschung von Integration und Differentiation). *Es sei $f\colon I \times D \to \mathbb{K}$, und $D \subset \mathbb{K}$. f besitze folgende Eigenschaften:*

i) $\forall t \in D\colon f(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
ii) $\forall x \in I\colon$ *existiere die partielle Ableitung* $\frac{\partial f}{\partial t}\colon I \times D \to \mathbb{K}$,
iii) $\forall t \in D\colon \frac{\partial f}{\partial t}(\cdot,t) \in \mathcal{R}(I,\mathbb{K})$,
iv) $\frac{\partial f}{\partial t} \in \mathcal{B}(I \times D,\mathbb{K})$, *d.h.* $\exists c \geq 0\, \forall (x,t) \in I \times D\colon \left|\frac{\partial f}{\partial t}(x,t)\right| \leq c$.

Dann ist die Abbildung $F\colon D \to \mathbb{K}$, definiert durch

$$F(t) = \int\limits_a^b f(x,t)\, dx,$$

auf D differenzierbar und es gilt

$$F'(t) = \int\limits_a^b \frac{\partial f}{\partial t}(x,t)\, dx.$$

Beweis. Es sei $t \in D$ und $(t_n) \subset D$ mit $\lim_{n\to\infty} t_n = t$ und $t_n \neq t$, $n \in \mathbb{N}$. Wir betrachten den Differenzenquotient von F:

$$\frac{F(t_n) - F(t)}{t_n - t} = \int\limits_a^b \frac{f(x,t_n) - f(x,t)}{t_n - t}\, dx =: \int\limits_a^b g_n(x)\, dx, \quad n \in \mathbb{N}.$$

Wegen i) gilt $(g_n) \subset \mathcal{R}(I, \mathbb{K})$ und aus ii) folgt für alle $x \in I$

$$\lim_{n\to\infty} g_n(x) = \frac{\partial f}{\partial t}(x,t).$$

Nach iii) ist $\frac{\partial f}{\partial t}$, also der punktweise Grenzwert der Folge (g_n), eine Regelfunktion. Wir zeigen nun die Beschränktheit von $(\|g_n\|)$: Der Mittelwertsatz V-4.6 zeigt

$$g_n(x) = \frac{f(x,t_n) - f(x,t)}{t_n - t} = \frac{\partial f}{\partial t}(x, t^*(x)), \quad t^*(x) \in (\min\{t, t_n\}, \max\{t, t_n\})$$

und iv) ergibt

$$\forall x \in I\, \forall n \in \mathbb{N} \colon |g_n(x)| \le c.$$

Aus dem Satz von Arzela-Osgood VII-5.8 folgt somit

$$\begin{aligned} F'(t) &= \lim_{n\to\infty} \frac{F(t_n) - F(t)}{t_n - t} = \lim_{n\to\infty} \int_a^b g_n(x)\, dx \\ &= \int_a^b \lim_{n\to\infty} g_n(x)\, dx = \int_a^b \frac{\partial f}{\partial t}(x,t)\, dx. \end{aligned}$$

□

Man beachte, daß in Satz VII-6.3 nicht die Beschränktheit von f, sondern jene von $\frac{\partial f}{\partial t}$ gefordert wird.

Beispiel VII-6.4 (Gaußsches Fehlerintegral). Wir betrachten das parameterabhängige Integral

$$F(t) = \int_0^1 \frac{e^{-(1+x^2)t^2}}{1+x^2}\, dx$$

für $t \in \mathbb{R}$. Wir setzen also $I = [0,1]$ und

$$f := \begin{cases} [0,1] \times \mathbb{R} \to \mathbb{R}, \\ (x,t) \mapsto \frac{e^{-(1+x^2)t^2}}{1+x^2}. \end{cases}$$

Wir verifizieren die Voraussetzungen von Satz VII-6.3: Für jedes $t \in \mathbb{R}$ ist die Abbildung $x \mapsto f(x,t)$ stetig, also eine Regelfunktion und für jedes $x \in [0,1]$ ist $t \mapsto f(x,t)$ differenzierbar. Es gilt

$$\frac{\partial f}{\partial t}(x,t) = -2te^{-(1+x^2)t^2},$$

somit ist $\frac{\partial f}{\partial t}(\cdot, t) \in \mathcal{R}(I, \mathbb{R})$ für alle $t \in \mathbb{R}$. Ferner gilt für alle $(x,t) \in I \times \mathbb{R}$

$$\Big|\frac{\partial f}{\partial t}(x,t)\Big| = 2|t|e^{-(1+x^2)t^2} \le \begin{cases} 2|t| \le 2 & |t| \le 1, x \in I, \\ \frac{2|t|}{(1+x^2)t^2} \le \frac{2}{|t|} \le 2 & |t| \ge 1, x \in I, \end{cases}$$

also ist $\frac{\partial f}{\partial t} \in \mathcal{B}(I \times \mathbb{R}, \mathbb{R})$. Nach Satz VII-6.3 ist F auf $\mathbb{R}$ differenzierbar und es gilt

$$F'(t) = -2e^{-t^2} \int_0^1 e^{-x^2t^2} t\, dx.$$

Substituiert man für festes t im Integral $g(x) = xt$, also $g(0) = 0$ und $g(1) = t$, erhält man mit Hilfe von Satz VII-4.9

$$F'(t) = -2e^{-t^2} \int_0^t e^{-u^2} du.$$

Beschränken wir uns vorübergehend auf ein kompaktes t-Intervall, folgt wegen der Stetigkeit von $u \mapsto e^{-u^2}$ aus dem Hauptsatz VII-4.3 die Beziehung

$$\frac{d}{dt}\Big(\int_0^t e^{-u^2} du\Big)^2 = 2 \int_0^t e^{-u^2} du \cdot e^{-t^2}.$$

Wir erhalten also

$$F'(t) = -\frac{d}{dt}\Big(\int_0^t e^{-u^2} du\Big)^2,$$

d.h.

$$F(t) = -\Big(\int_0^t e^{-u^2} du\Big)^2 + c, \quad t \in \mathbb{R}.$$

Die Konstante c ist festgelegt durch

$$c = F(0) = \int_0^1 \frac{dx}{1+x^2} = \arctan x\Big|_0^1 = \frac{\pi}{4}.$$

Wir erhalten daraus für alle $t \in \mathbb{R}$

$$\left(\int_0^t e^{-u^2} du\right)^2 = \frac{\pi}{4} - F(t).$$

Wir betrachten nun eine beliebige Folge (t_n) mit $\lim_{n\to\infty} t_n = \infty$. Für alle $n \in \mathbb{N}$ und $x \in I$ gilt dann

$$\lim_{n\to\infty} f(x, t_n) = 0$$

und

$$|f(x, t_n)| \leq 1.$$

Aus dem Satz von Arzela-Osgood folgt daher

$$\lim_{n\to\infty} F(t_n) = \lim_{n\to\infty} \int_0^1 f(x,t_n)\,dx = \int_0^1 \lim_{n\to\infty} f(x,t_n)\,dx = 0,$$

und folglich

$$\lim_{t\to\infty} \int_0^t e^{-u^2} du = \frac{\sqrt{\pi}}{2}.$$

Vorausgreifend bemerken wir, daß man den Grenzwert auf der linken Seite mit dem Symbol $\int_0^\infty e^{-u^2} du$ bezeichnet. Somit wurde gezeigt

$$\int_0^\infty e^{-u^2} du = \frac{\sqrt{\pi}}{2}.$$

Schließlich betrachten wir noch die Integration einer durch ein parameterabhängiges Integral definierten Funktion F. Unter den Voraussetzungen von Satz VII-6.1 ist F stetig und somit auf kompakten Teilintervallen von D integrierbar. Es sei etwa $[\alpha,\beta] \subset D$, dann gilt

$$\int_\alpha^\beta F(t)\,dt = \int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt.$$

Die Frage liegt nahe, ob die Gleichung

$$\int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx$$

gilt, d.h. ob es zulässig ist, die Integrationsreihenfolge zu vertauschen. Da die Variablen x und t bei dieser Problemstellung gleichberechtigt auftreten, fordern wir, daß $f(\cdot,t) \in C(I,\mathbb{K})$ für alle $t \in D$ ist.

Theorem VII-6.5 (Vertauschung der Integrationsreihenfolge). *Es sei $f \in \mathcal{B}([a,b]\times[\alpha,\beta],\mathbb{K})$ und es gelte*

i) $\forall x \in [a,b]\colon f(x,\cdot) \in C([\alpha,\beta],\mathbb{K})$,
ii) $\forall t \in [\alpha,\beta]\colon f(\cdot,t) \in C([a,b],\mathbb{K})$.

Dann gilt

$$\int_\alpha^\beta \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx.$$

Es ist klar, daß $f \in C([a,b]\times[\alpha,\beta],\mathbb{K})$ eine hinreichende Bedingung für i) und ii) darstellt.

Beweis. Nach Satz VII-6.1 ist die Abbildung $F\colon [\alpha, \beta] \to \mathbb{K}$

$$F(t) = \int_a^b f(x,t)\, dx$$

stetig auf $[\alpha, \beta]$ und somit integrierbar. Wir definieren die Abbildung $H\colon [\alpha, \beta] \to \mathbb{K}$ durch

$$H(y) = \int_\alpha^y F(t)\, dt, \quad y \in [\alpha, \beta].$$

Nach dem Hauptsatz VII-4.3 gilt für alle $y \in [\alpha, \beta]$

$$H'(y) = F(y) = \int_a^b f(x,y)\, dx.$$

Wir untersuchen nun die Abbildung $g\colon [a,b] \times [\alpha, \beta] \to \mathbb{K}$, welche gegeben ist durch

$$g(x,y) = \int_\alpha^y f(x,t)\, dt,$$

und verifizieren die Voraussetzungen von Satz VII-6.3. Für jedes $y \in [\alpha, \beta]$ ist $g(\cdot, y) \in C([a,b], \mathbb{K})$ nach Satz VII-6.1 (mit y als Parameter). Für jedes $x \in [a,b]$ ist $g(x, \cdot)$ nach dem Hauptsatz VII-4.3 wegen der Stetigkeit von $f(x, \cdot)$ stetig differenzierbar auf $[\alpha, \beta]$ und es gilt für alle $y \in [\alpha, \beta]$, $x \in [a,b]$

$$\frac{\partial g}{\partial y}(x,y) = f(x,y).$$

Es ist also auch $\frac{\partial g}{\partial y}(\cdot, y) \in C([a,b], \mathbb{K})$ und $\frac{\partial g}{\partial y} \in \mathcal{B}([a,b] \times [\alpha, \beta], \mathbb{K})$. Nach Satz VII-6.3 ist die Abbildung $G\colon [\alpha, \beta] \to \mathbb{K}$, definiert durch

$$G(y) = \int_a^b g(x,y)\, dx,$$

auf $[\alpha, \beta]$ differenzierbar mit

$$G'(y) = \int_a^b \frac{\partial g}{\partial y}(x,y)\, dx = \int_a^b f(x,y)\, dx.$$

Somit folgt für alle $y \in [\alpha, \beta]$

$$H'(y) = G'(y).$$

Es gibt daher eine Konstante $c \in \mathbb{K}$ mit

$$H(y) = G(y) + c, \quad y \in [\alpha, \beta].$$

Wegen $H(\alpha) = G(\alpha) = 0$ ($g(x,\alpha) = 0$ für alle $x \in [a,b]$) folgt sogar $H = G$, d.h. es ist

$$\int_\alpha^y \Big[\int_a^b f(x,t)\,dx\Big]\,dt = \int_a^b \Big[\int_\alpha^y f(x,t)\,dt\Big]\,dx, \quad y \in [\alpha,\beta].$$

Für $y = \beta$ ergibt sich die Behauptung. □

7. Uneigentliche Integrale

Die bisher entwickelte Integrationstheorie bezog sich auf Funktionen, deren Definitionsbereich ein *kompaktes* Intervall in $\mathbb{R}$ ist. Dies hat zwei starke Einschränkungen der Anwendbarkeit der Theorie zur Folge: Einerseits muß der Integrationsbereich beschränkt sein, andererseits sind integrierbare Funktionen notwendigerweise beschränkt. Wir zeigen nun, wie man Funktionen, die diesen Anforderungen nicht genügen, unter bestimmten Umständen einen sinnvollen Wert des Integrals zuordnen kann.

Definition VII-7.1. *Es sei $I = [a,b] \subset \bar{\mathbb{R}}$, $-\infty \le a < b \le \infty$ und $f\colon I \cap \mathbb{R} \to \mathbb{K}$. Wir definieren folgende* ***uneigentliche Integrale*** $\int_a^b f(x)\,dx$*:*

(1) *Es sei $a \in \mathbb{R}$ und $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{K})$ für alle $\beta \in [a,b)$. Existiert $\lim\limits_{\beta\uparrow b} \int_a^\beta f(x)\,dx$, so definiert man*

$$\int_a^b f(x)\,dx := \lim_{\beta\uparrow b} \int_a^\beta f(x)\,dx.$$

(2) *Eine analoge Definition gilt, falls $b \in \mathbb{R}$ für $\alpha \downarrow a$.*

(3) *Es sei $-\infty \le a < b \le \infty$. Existieren für ein $c \in (a,b)$ die uneigentlichen Integrale $\int_a^c f(x)\,dx$ und $\int_c^b f(x)\,dx$, definieren wir*

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

(4) *Das uneigentliche Integral $\int_a^b f(x)\,dx$ heißt* ***divergent****, wenn der entsprechende Grenzwert nicht existiert.*

Die Unabhängigkeit des uneigentlichen Integrals in (3) von der Wahl der Zwischenstelle wird später behandelt. Ist $I = [a,b] \subset \mathbb{R}$ und $f \in \mathcal{R}(I,\mathbb{K})$, so stimmt das uneigentliche Integral mit dem Integral gemäß Definition VII-2.1 überein. Wegen der stetigen Abhängigkeit etwa von der oberen Grenze gilt nämlich

$$\int_a^b f(x)\,dx = \lim_{\beta\uparrow b} \int_a^\beta f(x)\,dx.$$

(Beweis: Übung).
Diese Definition erfaßt die für die Praxis relevanten Typen von uneigentlichen Integralen:

Beispiel VII-7.2. 1) $\int\limits_1^\infty x^{-s}\,dx$.
(Typ 1: $a \in \mathbb{R}$, $b = \infty$, unbeschränkter Integrationsbereich). Der Integrand ist stetig auf $[1, \beta]$ für $\beta \in [1, \infty)$. Man erhält

$$\int_1^\beta x^{-s}\,dx = \begin{cases} \frac{1}{s-1}(1-\beta^{1-s}), & s \neq 1, \\ \ln\beta, & s = 1. \end{cases}$$

Somit existiert $\lim_{\beta\to\infty} \int\limits_1^\beta x^{-s}\,dx$ genau dann, wenn $s > 1$ gilt. Sein Wert ist

$$\int_1^\infty x^{-s}dx = \lim_{\beta\to\infty} \int_1^\beta x^{-s}ds = \frac{1}{s-1}.$$

2) $\int\limits_0^1 x^{-s}dx$.
(Typ 2: a, $b \in \mathbb{R}$, Integrand unbeschränkt am Rand des Integrationsbereiches). Der Integrand ist stetig auf $[\alpha, 1]$ für alle $\alpha \in (0, 1]$. Man erhält

$$\int_\alpha^1 x^{-s}dx = \begin{cases} \frac{1}{1-s}(1-\alpha^{1-s}), & s \neq 1, \\ -\ln\alpha, & s = 1. \end{cases}$$

Das uneigentliche Integral $\int_0^1 x^{-s}dx$ existiert also genau für $s < 1$. Es hat den Wert

$$\int_0^1 x^{-s}dx = \lim_{\alpha\downarrow 0} \int_\alpha^1 x^{-s}dx = \frac{1}{1-s}.$$

3) $\int\limits_{-\infty}^\infty \frac{dx}{1+x^2} = \pi$.

(Typ 3: $a = -\infty, b = \infty, c = 0$). Wie vorhin untersuchen wir die beiden uneigentlichen Integrale

$$\int\limits_0^\infty \frac{dx}{1+x^2} = \lim_{\beta\to\infty} \int\limits_0^\beta \frac{dx}{1+x^2} = \lim_{\beta\to\infty} \arctan\beta = \frac{\pi}{2}$$

$$\int\limits_{-\infty}^0 \frac{dx}{1+x^2} = \lim_{\alpha\to-\infty} \int\limits_\alpha^0 \frac{dx}{1+x^2} = -\lim_{\alpha\to-\infty} \arctan\alpha = \frac{\pi}{2}.$$

Beide zusammen ergeben die Behauptung. Aus Symmetriegründen könnte man sich in diesem Beispiel die Berechnung von $\int\limits_{-\infty}^0 \frac{dx}{1+x^2}$ ersparen.

4) $\int\limits_{-1}^1 \frac{dx}{\sqrt{|x|}} = 4.$

(Typ 4: a, $b \in \mathbb{R}$, $a < b$, $c = 0$, Integrand unbeschränkt im Inneren des Integrationsbereiches). Nach Beispiel 2) existieren die beiden uneigentlichen Integrale

$$\int\limits_0^1 \frac{dx}{\sqrt{x}} = \int\limits_{-1}^0 \frac{dx}{\sqrt{|x|}} = 2.$$

Daraus folgt die Behauptung.

Wir weisen darauf hin, daß bei uneigentlichen Integralen vom Typ 3) und 4) die Grenzwerte in den beiden uneigentlichen Integralen $\int\limits_a^c f(x)\,dx$ und $\int\limits_c^b f(x)\,dx$ voneinander *unabhängig* durchzuführen sind. Durch eine geeignete Koppelung der beiden Grenzwerte kann man manchmal die Existenz des Grenzwertes erzwingen: Als Beispiel betrachte man das uneigentliche Integral

$$\int\limits_a^b \frac{1}{x}\,dx, \quad a < 0 < b.$$

Dieses existiert nicht, da die beiden uneigentlichen Integrale $\int_a^0 \frac{1}{x}\,dx$ und $\int_0^b \frac{1}{x}\,dx$ nicht existieren. Koppelt man jedoch die beiden Grenzwerte $\lim_{\beta\uparrow 0} \int\limits_a^\beta \frac{1}{x}\,dx$ und $\lim_{\alpha\downarrow 0} \int\limits_\alpha^b \frac{1}{x}\,dx$ in der Form

$$\lim_{\varepsilon\downarrow 0} \left(\int\limits_a^{-\varepsilon} \frac{1}{x}\,dx + \int\limits_\varepsilon^b \frac{1}{x}\,dx \right) = \lim_{\varepsilon\downarrow 0} (\ln|-\varepsilon| - \ln|a| + \ln b - \ln\varepsilon) = \ln\frac{b}{|a|},$$

existiert der Grenzwert. Man nennt diesen *speziellen* Grenzwert **Cauchyscher Hauptwert** des divergenten uneigentlichen Integrals und schreibt

$$(C)\int_a^b \frac{1}{x}\,dx = \ln\frac{b}{|a|}.$$

Existiert das uneigentliche Integral $\int\limits_a^\infty f(x)\,dx$ für ein $a \in \mathbb{R}$, dann existiert auch das uneigentliche Integral $\int_\alpha^\infty f(x)\,dx$ für alle $\alpha \geq a$ und sein Wert ist

$$\int_\alpha^\infty f(x)\,dx = \int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx.$$

So selbstverständlich dies erscheinen mag, bedarf es trotzdem einer Rechtfertigung. Die Existenz von $\int_a^\infty f(x)\,dx$ bedeutet

$$(*) \qquad \forall\varepsilon > 0\,\exists\xi \geq a\,\forall\beta \geq \xi\colon \Big|\int_a^\infty f(x)\,dx - \int_a^\beta f(x)\,dx\Big| < \varepsilon.$$

Es sei nun $\alpha \geq a$ und $\varepsilon > 0$ beliebig und ξ entsprechend $(*)$ gewählt (O.B.d.A. können wir $\xi \geq \alpha$ annehmen). Für $\beta \geq \alpha$ gilt dann

$$\Big|\Big(\int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx\Big) - \int_\alpha^\beta f(x)\,dx\Big| = \Big|\int_a^\infty f(x)\,dx - \int_a^\beta f(x)\,dx\Big| < \varepsilon,$$

d.h.

$$\lim_{\beta\to\infty}\int_\alpha^\beta f(x)\,dx = \int_a^\infty f(x)\,dx - \int_a^\alpha f(x)\,dx.$$

Auch uneigentliche Integrale sind also additiv in Bezug auf das Integrationsintervall und verhalten sich in dieser Hinsicht wie das Cauchy Integral. Insbesondere gilt auch für alle $\alpha \leq \beta$

$$\int_\alpha^\beta f(x)\,dx = \int_\alpha^\infty f(x)\,dx - \int_\beta^\infty f(x)\,dx$$

soferne die uneigentlichen Integrale existieren. Tatsächlich gilt diese Formel für alle $\alpha, \beta \geq a$.

Ähnlich der Situation bei Reihen ist es meist nicht möglich, uneigentliche Integrale zu berechnen, sodaß man sich mit der bloßen Existenz (man sagt auch Konvergenz) der uneigentlichen Integrale bescheiden muß. Hiefür gibt es verschiedene Kriterien, welche wir nur für den Typ $\int_a^\infty f(x)\,dx$ formulieren. Die Modifikation für die anderen Möglichkeiten sind evident. Universell einsetzbar ist das Cauchy-Kriterium:

THEOREM VII-7.3 (Cauchy-Kriterium). *Es sei* $f\colon [a,\infty) \to \mathbb{K}$ *und* $f|_{[a,\beta]} \in \mathfrak{R}([a,\beta],\mathbb{K})$ *für alle* $\beta \geq a$. *Das uneigentliche Integral* $\int_a^\infty f(x)\,dx$ *existiert genau dann, wenn gilt:*

$$\forall \varepsilon > 0\, \exists \xi \geq a\, \forall \alpha, \beta \geq \xi\colon \Big|\int\limits_\alpha^\beta f(x)\,dx\Big| < \varepsilon.$$

BEWEIS. Man betrachte $\lim_{\beta\to\infty} F(\beta)$ für $F(\beta) = \int\limits_a^\beta f(x)\,dx$, $\beta \geq a$. □

BEISPIEL VII-7.4. $\int_0^\infty \frac{\sin x}{x}\,dx$.
Wir definieren die Funktion $f\colon [0,\infty) \to \mathbb{R}$ durch

$$f(x) = \begin{cases} 1, & x = 0, \\ \frac{\sin x}{x}, & x > 0, \end{cases}$$

und bemerken, daß f auf $[0,\infty)$ stetig ist. Insbesondere ist f integrierbar auf $[0,a]$ und auf $[a,\beta]$ für beliebige $0 < a < \beta < \infty$. Spalten wir daher $\int_0^\beta f(x)\,dx$ auf in

$$\int\limits_0^\beta f(x)\,dx = \int\limits_0^a f(x)\,dx + \int\limits_a^\beta f(x)\,dx,$$

genügt es, die Existenz von $\lim_{\beta\to\infty} \int\limits_a^\beta f(x)\,dx$ nachzuweisen. Dazu formen wir vorerst das Integral mit partieller Integration um:

$$\int\limits_a^\beta \frac{1}{x}\sin x\,dx = -\frac{1}{x}\cos x\Big|_a^\beta - \int\limits_a^\beta \frac{1}{x^2}\cos x\,dx.$$

Der erste Term auf der rechten Seite besitzt einen Grenzwert für $\beta \to \infty$. Auf das Integral wenden wir das Cauchy-Kriterium an und erhalten für $a \leq u < v$

$$\Big|\int\limits_u^v \frac{\cos x}{x^2}\,dx\Big| \leq \int\limits_u^v \frac{dx}{x^2} = \frac{1}{u} - \frac{1}{v} < \frac{1}{u}.$$

Wählt man $\varepsilon > 0$ und setzt $\xi(\varepsilon) = \frac{1}{\varepsilon}$, erhält man für $\xi(\varepsilon) < u < v$

$$\Big|\int\limits_u^v \frac{\cos x}{x^2}\,dx\Big| < \varepsilon,$$

und damit die Konvergenz von $\int_a^\infty \frac{\cos x}{x^2}\,dx$. Somit existiert auch $\int_0^\infty \frac{\sin x}{x}\,dx$.

Definition VII-7.5. *Es sei* $f\colon [a,\infty) \to \mathbb{K}$ *und* $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{K})$ *für alle* $\beta \in [a,\infty)$*. Das uneigentliche Integral* $\int\limits_a^\infty f(x)\,dx$ *heißt* ***absolut konvergent****, wenn* $\int_a^\infty |f(x)|\,dx$ *existiert.*

Aus der absoluten Konvergenz eines uneigentlichen Integrals folgt dessen Konvergenz (Existenz). Dies ergibt sich mit Hilfe des Cauchy-Kriteriums aus der Ungleichung ($u < v$)

$$\Big|\int\limits_u^v f(x)\,dx\Big| \le \int\limits_u^v |f(x)|\,dx.$$

Die Umkehrung ist – wie bei Reihen – nicht richtig: Aus der bloßen Existenz eines uneigentlichen Integrals folgt nicht dessen absolute Konvergenz. Als Beispiel betrachten wir das konvergente Integral $\int_0^\infty \frac{\sin x}{x}\,dx$. Die Abbildung $F\colon [0,\infty) \to \mathbb{R}$ sei gegeben durch

$$F(\beta) = \int\limits_0^\beta \Big|\frac{\sin x}{x}\Big|\,dx.$$

Für $\beta = n\pi$ erhalten wir

$$F(n\pi) = \int\limits_0^{n\pi} \Big|\frac{\sin x}{x}\Big|\,dx = \sum_{i=1}^n \int\limits_{(i-1)\pi}^{i\pi} |\sin x|\frac{1}{x}\,dx \ge \sum_{i=1}^n \frac{1}{i\pi} \int\limits_{(i-1)\pi}^{i\pi} |\sin x|\,dx = \frac{2}{\pi}\sum_{i=1}^n \frac{1}{i}.$$

Wegen der Divergenz der harmonischen Reihe kann also $\lim_{\beta\to\infty} F(\beta)$ nicht existieren.

Theorem VII-7.6 (Vergleichskriterium). *Es sei* $f\colon [a,\infty) \to \mathbb{R}$*,* $g\colon ([a,\infty) \to \mathbb{R}_+)$ *und* $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R})$*,* $g|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R}_+)$ *für alle* $\beta \ge a$*. Gilt* $|f(t)| \le g(t)$ *für alle* $t \in [a,\infty)$ *und existiert* $\int_a^\infty g(t)\,dt$*, dann ist* $\int_a^\infty f(t)\,dt$ *absolut konvergent.*
Umgekehrt: Gilt $f(t) \ge g(t) \ge 0$ *für alle* t *(notwendigerweise ist* $\operatorname{bild} f \subset \mathbb{R}$*) und ist* $\int_a^\infty g(t)\,dt$ *divergent, dann divergiert auch das uneigentliche Integral* $\int_a^\infty f(t)\,dt$*.*

Beweis. Die erste Behauptung ergibt sich aus der Abschätzung

$$\Big|\int\limits_u^v f(x)\,dx\Big| \le \int\limits_u^v |f(x)|\,dx \le \int\limits_u^v g(x)\,dx$$

für $a \le u < v$ und dem Cauchy-Kriterium VII-7.3. Die zweite Behauptung folgt aus

$$\int\limits_a^\beta g(x)\,dx \le \int\limits_a^\beta f(x)\,dx$$

für $\beta \in [a,\infty)$ und der Unbeschränktheit von $\{\int\limits_a^\beta g(x)\,dx \colon \beta \in [a,\infty)\}$. $\square$

KOROLLAR VII-7.7. Es sei $f\colon [a,\infty) \to \mathbb{R}$, $a > 0$, und $f|_{[a,\beta]} \in \mathcal{R}([a,\beta],\mathbb{R})$ für alle $\beta \in [a,\infty)$. Gilt

$$|f(x)| \leq \frac{1}{x^s}, \quad \text{für ein } s > 1$$

für alle $x \geq x_0 \geq a$, dann ist $\int_a^\infty f(x)\,dx$ absolut konvergent. Ist $\operatorname{bild} f \subset \mathbb{R}$ und gilt

$$0 < \frac{1}{x^s} \leq f(x), \quad \text{für ein } s \leq 1$$

für alle $x \geq x_0 \geq a$, dann ist $\int_a^\infty f(x)\,dx$ divergent.

Die manchmal recht mühsamen Abschätzungen bei der Anwendung des Vergleichskriteriums kann man u.U. auch umgehen:

THEOREM VII-7.8 (Grenzwertkriterium). *Es seien $f, g\colon [a,\infty),\mathbb{R}$ und $f|_{[a,\beta]}, g|_{[a,\beta]} \in \mathcal{R}([a,\beta]),\mathbb{R})$ für alle $\beta \geq a$. Darüber hinaus gelte $f(x) \geq 0$ und $g(x) > 0$ für alle $x \in [a,\infty)$. Es existiere*

$$\rho := \lim_{x\to\infty} \frac{f(x)}{g(x)} \in \bar{\mathbb{R}}.$$

i) *Gilt $\rho \in (0,\infty)$, dann sind $\int_a^\infty f(x)\,dx$ und $\int_a^\infty g(x)\,dx$ beide konvergent oder beide divergent.*
ii) *Gilt $\rho = 0$ und existiert $\int_a^\infty g(x)\,dx$, dann existiert auch $\int_a^\infty f(x)\,dx$.*
ii) *Gilt $\rho = \infty$ und divergiert $\int_a^\infty g(x)\,dx$, dann divergiert auch $\int_a^\infty f(x)\,dx$.*

BEWEIS. i) Es gibt ein $\xi \geq a$ so, daß für alle $x \geq \xi$ die Abschätzung

$$\frac{\rho}{2} g(x) \leq f(x) \leq \frac{3\rho}{2} g(x)$$

gilt. Die Behauptung folgt nun aus dem Vergleichskriterium.
ii) Es gibt ein $\xi \geq a$ so, daß für alle $x \geq \xi$ die Abschätzung

$$0 \leq f(x) \leq g(x)$$

zutrifft.
iii) In diesem Falle gilt für alle hinreichend großen x die Abschätzung

$$1 \leq g(x) \leq f(x).$$

□

Ersetzt man im Grenzwertkriterium f durch $|f|$, erhält man natürlich ein Kriterium für die absolute Konvergenz des uneigentlichen Integrals. Wir betonen noch einmal, daß analoge Kriterien auch für die übrigen Typen von uneigentlichen Integralen gelten.

BEISPIEL VII-7.9 (Eulersches Γ-Integral).

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\,dx, \quad s > 0.$$

Da beide Grenzen kritisch sind, spalten wir das Integral auf in

$$\Gamma(s) = \int_0^1 x^{s-1}e^{-x}\,dx + \int_1^\infty x^{s-1}e^{-x}\,dx.$$

Für das erste Integral verwenden wir die Vergleichsfunktion $g(x) = x^{s-1}$. Nach Beispiel VII-7.2 existiert $\int_0^1 x^{s-1}\,dx$ genau für $s > 0$. Mit $f(x) = x^{s-1}e^{-x}$ ergibt sich

$$\rho = \lim_{x\downarrow 0} \frac{f(x)}{g(x)} = \lim_{x\downarrow 0} e^{-x} = 1.$$

Nach dem Grenzwertkriterium existiert $\int_0^1 x^{s-1}e^{-x}\,dx$ genau für $s > 0$. Für das zweite Integral verwenden wir die Vergleichsfunktion $g(x) = e^{-\frac{x}{2}}$ (um das mögliche Anwachsen von $x \mapsto x^{s-1}$ zu kompensieren). Wir erhalten nun

$$\rho = \lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x^{s-1}e^{-x}}{e^{-x/2}} = \lim_{x\to\infty} x^{s-1}e^{-\frac{x}{2}} = 0.$$

Aus der Konvergenz von $\int_1^\infty e^{-\frac{x}{2}}\,dx$ folgt daher mit Satz VII-7.8 die Existenz von $\int_1^\infty x^{s-1}e^{-x}\,dx$ (sogar für alle $s \in \mathbb{R}$). Somit existiert $\Gamma(s)$ genau für $s > 0$. Als Übung beweise man

$$\forall n \in \mathbb{N}\colon \Gamma(n) = \int_0^\infty x^{n-1}e^{-x}\,dx = (n-1)!.$$

Mit Hilfe uneigentlicher Integrale läßt sich in manchen Fällen bequem die Konvergenz unendlicher Reihen nachweisen:

Theorem VII-7.10 (Integralkriterium). *Es sei $f\colon [1,\infty) \to \mathbb{R}$ monoton fallend und $f(x) \geq 0$ für alle $x \in [1,\infty)$. Die unendliche Reihe*

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

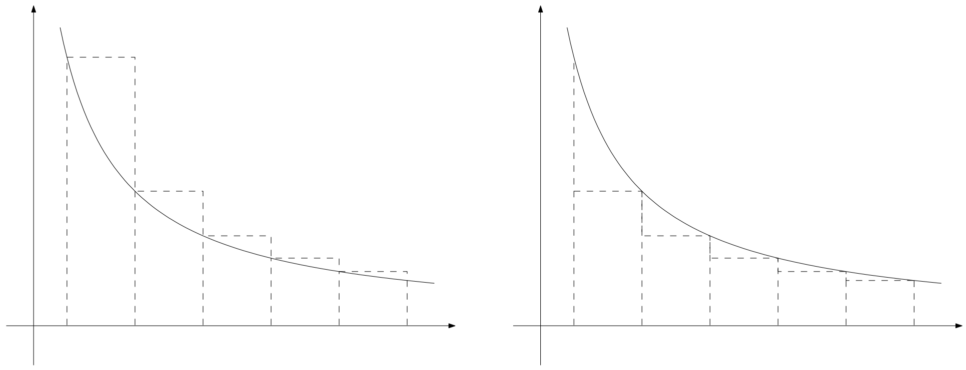
konvergiert genau dann, wenn das uneigentliche Integral

$$\int_1^\infty f(x)\,dx$$

existiert.

Beweis. Da f monoton fällt, ist $f|_{[1,\beta]} \in \mathcal{R}([1,\beta],\mathbb{R})$ für alle $\beta \geq 1$. Ferner gilt für alle $k \in \mathbb{N}$ die Abschätzung

$$f(k+1) \leq \int_k^{k+1} f(x)\,dx \leq f(k).$$



Durch Addition folgt

$$\sum_{k=2}^{n} f(k) \leq \int_{1}^{n} f(x)\,dx \leq \sum_{k=1}^{n-1} f(k), \quad n \in \mathbb{N}.$$

Bezeichnet S_n die n-te Partialsumme der Reihe $\sum_{k=1}^{\infty} f(k)$ erhält man

$$\begin{aligned} &S_n - f(1) \leq \int_{1}^{n} f(x)\,dx \leq S_{n-1}, \\ &S_n \leq S_{n+1}, \quad n \in \mathbb{N}. \end{aligned}$$

Existiert $\int_{1}^{\infty} f(x)\,dx$, ergibt sich aus

$$S_n \leq f(1) + \int_{1}^{n} f(x)\,dx \leq f(1) + \int_{1}^{\infty} f(x)\,dx$$

die Konvergenz der Folge (S_n) aus dem Monotoniekriterium. Divergiert hingegen $\int_1^\infty f(x)\,dx$, ist wegen

$$\int_{1}^{n} f(x)\,dx \leq S_{n-1}$$

auch die Reihe $\sum_{k=1}^{\infty} f(k)$ divergent. □

Beispiel VII-7.11. $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^s}$ konvergiert für $s > 1$ und divergiert für $s \leq 1$. Die Abbildung $x \to f(x) = \frac{1}{x}(\ln x)^{-s}$ erfüllt für $x \geq 2$ die Voraussetzungen des Integralkriteriums. Substitutiert man $g(x) = \ln x$, erhält man

$$\int_{2}^{n} \frac{dx}{x(\ln x)^s} = \int_{\ln 2}^{\ln n} \frac{dt}{t^s} = \frac{1}{1-s}[(\ln n)^{1-s} - (\ln 2)^{1-s}], \quad s \neq 1.$$

Für $s > 1$ folgt die Existenz des uneigentlichen Integrals

$$\int_2^\infty \frac{dx}{x(lnx)^s},$$

für $s < 1$ dessen Divergenz. Für $s = 1$ ergibt sich wegen

$$\int_3^n \frac{dx}{x\ln x} = \int_{\ln 3}^{\ln n} \frac{dt}{t} = \ln(\ln n) - \ln(\ln 3)$$

ebenfalls die Divergenz des uneigentlichen Integrals. Dies zeigt die Behauptung.

Da uneigentliche Integrale durch einen Grenzübergang definiert sind, bleiben Linearität, Positivität und Monotonie erhalten. Der Betrag des Integrals kann aber nicht mehr gegen die Intervallänge und $\|f\|$ abgeschätzt werden. Dies wurde jedoch wesentlich im Beweis des Satzes von Arzela-Osgood verwendet. Wir zeigen nun, daß dieser Satz für uneigentliche Integrale nicht mehr gilt:

Beispiel VII-7.12. Es sei $(f_n) \subset C(\mathbb{R})$ gegeben durch

$$f_n(x) = \frac{n}{n^2 + x^2}, \quad x \in \mathbb{R}, n \in \mathbb{N}.$$

Es gilt $\lim_{n\to\infty} f_n = 0$ gleichmäßig auf $\mathbb{R}$. Für jedes f_n existiert das uneigentliche Integral

$$\int_0^\infty \frac{n}{n^2 + x^2}\, dx = \int_0^\infty \frac{dt}{1 + t^2} = \frac{\pi}{2}.$$

Es gilt also *nicht* $\lim_{n\to\infty} \int_0^\infty f_n(x)\, dx = \int_0^\infty \lim_{n\to\infty} f_n(x)\, dx$.

8. Parameterabhängige uneigentliche Integrale

Wie bei Funktionenreihen spielt bei parameterabhängigen uneigentlichen Integralen die Gleichmäßigkeit der Konvergenz eine wesentliche Rolle:

Definition VII-8.1. *Es sei $f\colon I \times D \to \mathbb{K}$, $I = [a, \infty)$, $D \subset \mathbb{R}$. Für alle $\beta \in [a, \infty)$ und für alle $t \in D$ sei $f|_{[a,\beta]}(\cdot, t) \in \mathcal{R}([a, \beta], \mathbb{K})$.*

$\int_a^\infty f(x, t)\, dx$ *heißt* ***gleichmäßig konvergent*** *(auf D)* $\underset{Def}{\Leftrightarrow}$

$$\forall \varepsilon > 0\, \exists \xi \in [a, \infty)\, \forall \alpha, \beta \geq \xi\, \forall t \in D\colon |\int_\alpha^\beta f(x, t)\, dx| < \varepsilon.$$

Eine offensichtlich hinreichende Bedingung für gleichmäßige Konvergenz des uneigentlichen Integrals ist die Existenz einer bezüglich t gleichmäßigen Majorante $g\colon [a, \infty) \to \mathbb{R}$. Gilt also für alle $x \in [a, \infty)$ und für alle $t \in D$

$$|f(x, t)| \leq g(x),$$

und existiert $\int_a^\infty g(x)\,dx$, dann ist $\int_a^\infty f(x,t)\,dx$ gleichmäßig konvergent. Dies folgt unmittelbar aus der Abschätzung ($\alpha \leq \beta$)

$$|\int_\alpha^\beta f(x,t)\,dx| \leq \int_\alpha^\beta g(x)\,dx.$$

Beispiel VII-8.2. $\int_a^\infty \frac{1}{x}\exp((-t+i)x)\,dx$, $a>0$, $t\geq 0$. Für $t \geq c > 0$ ergibt sich aus $|\frac{1}{x}\exp((-t+i)x)| \leq \frac{1}{c}\frac{1}{x^2}$ die gleichmäßige Konvergenz auf $[c,\infty)$. Wir dehnen nun die Gleichmäßigkeit der Konvergenz auf $t \geq 0$ aus. Durch partielle Integration erhält man für $a \leq u \leq v$

$$\int_u^v \frac{1}{x}e^{(-t+i)x}\,dx = \frac{1}{x}\cdot\frac{1}{-t+i}e^{(-t+i)x}\Big|_u^v + \int_u^v \frac{1}{x^2}\frac{1}{-t+i}e^{(-t+i)x}\,dx.$$

Wegen $|\frac{1}{-t+i}| \leq 1$ und $|e^{(-t+i)x}| = e^{-tx} \leq 1$ für $t \geq 0$, ergibt sich die Abschätzung

$$\Big|\int_u^v \frac{1}{x}e^{(-t+i)x}\,dx\Big| \leq \frac{1}{u}+\frac{1}{v}+\int_u^v \frac{1}{x^2}\,dx = \frac{2}{u}$$

und damit die gleichmäßige Konvergenz für $t \geq 0$. Durch Zerlegung in Real- und Imaginärteil folgt dann die gleichmäßige Konvergenz der reellen Integrale

$$\int_a^\infty \frac{\cos x}{x}e^{-tx}\,dx \quad \text{und} \quad \int_a^\infty \frac{\sin x}{x}e^{-tx}\,dx, \quad a>0,$$

auf $t \geq 0$. Da $x \mapsto \frac{\sin x}{x}$ stetig nach 0 fortgesetzt werden kann, ist sogar

$$\int_0^\infty \frac{\sin x}{x}e^{-tx}\,dx$$

gleichmäßig konvergent auf $[0,\infty)$.

Theorem VII-8.3. *Für $f \in \mathcal{B}(I\times D,\mathbb{K})$, $I=[a,\infty)$, $D=[c,d]\subset\mathbb{R}$, gelte*

i) *$\int_a^\infty f(x,t)\,dx$ sei gleichmäßig konvergent auf D,*
ii) *$\forall x \in I\colon f(x,\cdot) \in C(D,\mathbb{K})$.*

Dann ist die Abbildung $F\colon D \to \mathbb{K}$, definiert durch

$$F(t) = \int_a^\infty f(x,t)\,dx,$$

stetig auf D.

Beweis. Es sei $t_0 \in D$. Für $\alpha \in [a, \infty)$ betrachten wir

$$\begin{aligned}|F(t) - F(t_0)| &= |\int_a^\infty f(x,t)\,dx - \int_a^\infty f(x,t_0)\,dx| \\ &\leq |\int_a^\alpha f(x,t)\,dx - \int_a^\alpha f(x,t_0)\,dx| + |\int_\alpha^\infty f(x,t)\,dx| + |\int_\alpha^\infty f(x,t_0)\,dx|.\end{aligned}$$

Wegen der gleichmäßigen Konvergenz von $\int_a^\infty f(x,t)\,dx$ für $t \in D$ gibt es zu $\varepsilon > 0$ ein $\xi \geq a$ derart, daß für alle $\xi \leq \alpha < \infty$ und für alle $t \in D$

$$|\int_\alpha^\infty f(x,t)\,dx| < \varepsilon.$$

Wir fixieren nun ein derartiges α. Auf $[a, \alpha]$ erfüllt f die Voraussetzungen von Satz VII-6.1. Also gibt es ein $\delta > 0$ so, daß $|t - t_0| < \delta$ und $t \in D$

$$|\int_a^\alpha f(x,t)\,dx - \int_a^\alpha f(x,t_0)\,dx| < \varepsilon$$

nach sich zieht. Insgesamt erhält man daher für $|t - t_0| < \delta$ und $t \in D$

$$|F(t) - F(t_0)| < 3\varepsilon.$$

□

Theorem VII-8.4. *Für $f \in \mathcal{B}(I \times D, \mathbb{K})$, $I = [a, \infty)$, $D = [\alpha, \beta] \subset \mathbb{R}$, gelte*

i) *$\int_a^\infty f(x,t)\,dx$ sei gleichmäßig konvergent auf D,*
ii) *$\forall t \in D\colon f(\cdot, t) \in C(I, \mathbb{K})$,*
iii) *$\forall x \in I\colon f(x, \cdot) \in C(D, \mathbb{K})$.*

Dann konvergiert das uneigentliche Integral $\int_a^\infty [\int_\alpha^\beta f(x,t)\,dt]\,dx$ und es gilt

$$\int_\alpha^\beta [\int_a^\infty f(x,t)\,dx]\,dt = \int_a^\infty [\int_\alpha^\beta f(x,t)\,dt]\,dx.$$

Beweis. Nach Satz VII-8.3 ist $t \mapsto \int_a^\infty f(x,t)\,dx$ stetig auf D. Somit existiert das iterierte Integral $\int_\alpha^\beta [\int_a^\infty f(x,t)\,dx]\,dt$. Es sei $\varepsilon > 0$ beliebig gewählt. Wegen der gleichmäßigen Konvergenz von $\int_a^\infty f(x,t)\,dx$ gibt es ein $\xi \geq a$, so daß für alle $u \geq \xi$

und $t \in D$ die Abschätzung

$$\Big|\int_a^\infty f(x,t)\,dx - \int_a^u f(x,t)\,dx\Big| < \varepsilon$$

zutrifft. Damit ergibt sich die Ungleichung

$$\begin{aligned}&\Big|\int_\alpha^\beta \Big[\int_a^\infty f(x,t)\,dx\Big]\,dt - \int_\alpha^\beta \Big[\int_a^u f(x,t)\,dx\Big]\,dt\Big| \\ &\leq \int_\alpha^\beta \Big|\int_a^\infty f(x,t)\,dx - \int_a^u f(x,t)\,dx\Big|\,dt \leq \varepsilon(\beta-\alpha).\end{aligned}$$

Auf Grund der Voraussetzungen an f kann man in $\int_\alpha^\beta[\int_a^u f(x,t)\,dx]\,dt$ nach Satz VII-6.5 die Integrationsreihenfolge vertauschen und erhält für alle $u \geq \xi$

$$\Big|\int_\alpha^\beta \Big[\int_a^\infty f(x,t)\,dx\Big]\,dt - \int_a^u \Big[\int_\alpha^\beta f(x,t)\,dt\Big]\,dx\Big| \leq \varepsilon(\beta-\alpha).$$

Daraus folgt die Behauptung. □

THEOREM VII-8.5. *Für $f\colon I \times D \to \mathbb{K}$, $I = [a,\infty)$, $D = [\alpha,\beta] \subset \mathbb{R}$, gelte:*

i) *Es existiere $\int_a^\infty f(x,\alpha)\,dx$,*
ii) *Es existiere $\frac{\partial f}{\partial t}\colon I \times D \to \mathbb{K}$,*
iii) $\frac{\partial f}{\partial t} \in \mathcal{B}(I \times D, \mathbb{K})$,
iv) $\forall t \in D\colon \frac{\partial f}{\partial t}(\cdot,t) \in C(I,\mathbb{K})$,
v) $\forall x \in I\colon \frac{\partial f}{\partial t}(x,\cdot) \in C(D,\mathbb{K})$,
vi) $\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx$ *sei gleichmäßig konvergent auf D.*

Dann konvergiert

$$F(t) := \int_a^\infty f(x,t)\,dx$$

für alle $t \in D$, F ist differenzierbar auf D und es gilt

$$F'(t) = \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx.$$

BEWEIS. $\frac{\partial f}{\partial t}$ erfüllt die Voraussetzungen von Satz VII-8.4 auf $I \times [\alpha, y]$ für alle $y \in D$. Somit konvergiert für alle $y \in D$ das uneigentliche Integral $\int_a^\infty[\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt]\,dx$

und es gilt

$$(*) \qquad \int_\alpha^y \Big[\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx\Big]\,dt = \int_a^\infty \Big[\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt\Big]\,dx.$$

Für festes $x \in I$ ist $\frac{\partial f}{\partial t}(x,\cdot) \in C(D,\mathbb{K})$. Nach dem Hauptsatz VII-4.3 folgt daher für alle $y \in D$

$$\int_\alpha^y \frac{\partial f}{\partial t}(x,t)\,dt = f(x,y) - f(x,\alpha).$$

Wegen der Konvergenz von $\int_a^\infty f(x,\alpha)\,dx$ ergibt sich damit aus $(*)$ die Konvergenz von $\int_a^\infty f(x,y)\,dx$ und

$$F(y) := \int_a^\infty f(x,y)\,dx = \int_a^\infty f(x,\alpha)\,dx + \int_\alpha^y \Big[\int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx\Big]\,dt, \quad y \in D.$$

$\frac{\partial f}{\partial t}$ erfüllt auch die Voraussetzungen von Satz VII-8.3. Also ist $t \mapsto \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx$ stetig auf D. Der Hauptsatz VII-4.3 sichert daher die Differenzierbarkeit von F auf D und

$$F'(t) = \int_a^\infty \frac{\partial f}{\partial t}(x,t)\,dx.$$

$\square$

Beispiel VII-8.6. 1) Wir betrachten das komplexe uneigentliche Integral

$$\int_0^\infty e^{(-1+it)x}\,dx = -\frac{1}{-1+it}, \quad t \in \mathbb{R}.$$

Dies folgt aus

$$\int_0^\beta e^{(-1+it)x}\,dx = \frac{e^{-\beta}e^{it\beta}}{-1+it} - \frac{1}{-1+it}.$$

Durch Aufspalten in Real- und Imaginärteil ergibt sich

$$\int_0^\infty e^{(-1+it)x}\,dx = \int_0^\infty e^{-x}\cos tx\ dx + i\int_0^\infty e^{-x}\sin tx\ dx = \frac{1+it}{1+t^2},$$

und damit die reellen uneigentlichen Integrale

$$(*) \qquad \left.\begin{aligned} \int_0^\infty e^{-x}\cos tx \; dx &= \frac{1}{1+t^2} \\ \int_0^\infty e^{-x}\sin tx \; dx &= \frac{t}{1+t^2} \end{aligned}\right\}, \quad t \in \mathbb{R}.$$

Aus $|e^{(-1+it)x}| \leq e^{-x}$, $(t,x) \in [0,\infty) \times \mathbb{R}$, folgt die gleichmäßige Konvergenz der uneigentlichen Integrale. Wir schränken nun t auf $D = [0,1]$ ein. Die Abbildung $f\colon [0,\infty) \times D \to \mathbb{R}$, $f(x,t) = e^{-x}\cos tx$, erfüllt die Voraussetzungen von Satz VII-8.4. Somit existiert $\int_0^\infty [\int_0^1 e^{-x}\cos tx \, dt] \, dx$ und es gilt

$$\int_0^1 \Big[\int_0^\infty e^{-x}\cos tx \; dx\Big] \, dt = \int_0^\infty \Big[\int_0^1 e^{-x}\cos tx \; dt\Big] \, dx.$$

Auf der linken Seite erhält man mit $(*)$

$$\int_0^1 \frac{dt}{1+t^2} = \arctan 1 - \arctan 0 = \frac{\pi}{4},$$

auf der rechten Seite kann man das innere Integral ausführen:

$$\int_0^\infty \Big[\int_0^1 e^{-x}\cos tx \, dt\Big] \, dx = \int_0^\infty e^{-x}\frac{\sin x}{x} \, dx.$$

(Man beachte, daß der Integrand stetig nach $x = 0$ fortgesetzt werden kann). Insgesamt wurde also gezeigt

$$\int_0^\infty e^{-x}\frac{\sin x}{x} \, dx = \frac{\pi}{4}.$$

2) Für $t \geq 0$ definieren wir die Abbildung F durch

$$F(t) = \int_0^\infty e^{-tx}\frac{\sin x}{x} \, dx.$$

Die Existenz von $F(0)$ wurde in Beispiel VII-7.4 nachgewiesen, für $t > 0$ kann man wie in Beispiel 1) argumentieren. Dort wurde auch

$$F(1) = \frac{\pi}{4}$$

gezeigt. Wir setzen $f\colon [0,\infty) \times [0,1] \to \mathbb{R}$ fest durch

$$f(x,t) = e^{-tx}\frac{\sin x}{x},$$

und erhalten

$$\frac{\partial f}{\partial t}(x,t) = -e^{-tx}\sin x.$$

Für jedes $t > 0$ ergibt eine einfache Rechnung wie in 1)(∗) das uneigentliche Integral

$$\int_0^\infty \frac{\partial f}{\partial t}(x,t)\,dx = -\int_0^\infty e^{-tx}\sin x\,dx = \frac{-1}{1+t^2},$$

welches gleichmäßig für $t \in [\alpha, 1]$, $\alpha \in (0,1]$ konvergiert. Somit erfüllt f sämtliche Voraussetzungen von Satz VII-8.5 auf $[0,\infty) \times [\alpha, \beta]$ für alle $\alpha \in (0,1]$. Daher ist F auf $(0,1]$ differenzierbar und es gilt

$$F'(t) = -\int_0^\infty e^{-tx}\sin x\,dx = -\frac{1}{1+t^2}.$$

Folglich ist mit einer Konstanten $c \in \mathbb{R}$

$$F(t) = -\arctan t + c.$$

Der Wert von c ergibt sich aus

$$F(1) = \frac{\pi}{4} = -\arctan 1 + c$$

zu

$$c = \frac{\pi}{2}.$$

Für alle $t \in (0,1]$ gilt daher

$$F(t) = \frac{\pi}{2} - \arctan t.$$

Nach Satz VII-8.3 ist F stetig auf $[0,1]$, soferne $\int_0^\infty f(x,t)\,dx$ gleichmäßig auf $[0,1]$ konvergiert. Nehmen wir dies vorerst an, kann man auf $F(0) = \frac{\pi}{2}$ schließen. Also gilt

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Für den Nachweis der gleichmäßigen Konvergenz von $\int_0^\infty f(x,t)\,dx$ für $t \in [0,1]$ bemerken wir, daß es genügt, die uneigentlichen Integrale $\int_1^\infty f(x,t)\,dx$ zu betrachten. Integrieren wir partiell und beachten

$$\int e^{-tx}\sin x\,dx = -\frac{1}{1+t^2}[t\sin x + \cos x]e^{-tx},$$

erhalten wir

$$\begin{aligned}\int_1^\beta e^{-tx}\frac{\sin x}{x}\,dx &= -\frac{1}{x}\frac{e^{-tx}}{1+t^2}[t\sin x+\cos x]\Big|_1^\beta - \frac{1}{1+t^2}\int_1^\beta \frac{1}{x^2}e^{-tx}[t\sin x+\cos x]\,dx \\ &= -\frac{1}{\beta}\frac{e^{-t\beta}}{1+t^2}[t\sin\beta+\cos\beta] + \frac{e^{-t}}{1+t^2}[t\sin 1+\cos 1] \\ &\quad - \frac{1}{1+t^2}\int_1^\beta \frac{1}{x^2}e^{-tx}[t\sin x+\cos x]\,dx.\end{aligned}$$

Der erste Term auf der rechten Seite kann gleichmäßig in $t \in [0,1]$ durch $\frac{2}{\beta}$ abgeschätzt werden. Ebenso kann das Integral durch $2\int_1^\beta \frac{1}{x^2}\,dx$ beschränkt werden. Daraus folgt die gleichmäßige Konvergenz von $\int_1^\infty e^{-tx}\frac{\sin x}{x}\,dx$ für $t \in [0,1]$.