Analysis

Skriptum einer zweisemestrigen Vorlesungsreihe,

gehalten im Studienjahr 2000/2001 an der Universität Essen

Karl–Josef Witsch

Fachbereich Mathematik

der Universität Duisburg–Essen

2. Auflage

Mein Dank gilt Herrn Dr. Sebastian Bauer und Herrn Dipl. Math. Tiemo Pepperl, die mich bei dieser Veranstaltung und der Erstellung des Skriptums kritisch unterstützt haben, Frau cand. math. Stefanie Vanis für ihre sorgfältigen Korrekturen der ersten Auflage und nicht zuletzt Frau Kirsten Wosniack für das geduldige Schreiben des Manuskripts.

Inhaltsverzeichnis

1 Die reellen Zahlen **5**

1.1 Grundlagen aus Logik und Mengenlehre 7

1.2 Geordnete Körper 14

1.3 Die natürlichen Zahlen 24

1.4 Die reellen Zahlen 37

2 Funktionen und Konvergenz **47**

2.1 Funktionen 47

2.2 Über die Mächtigkeit von Mengen 54

2.3 Reelle Funktionen und Konvergenz 58

3 Folgen **69**

3.1 Grenzwerte von Folgen 69

3.2 Grenzwerte monotoner Folgen 73

3.3 Teilfolgen und Häufungspunkte von Folgen 76

4 Stetige Funktionen **94**

4.1 Stetigkeit 94

4.2 Gleichmäßige Stetigkeit 98

4.3 Eigenschaften stetiger Funktionen 99

5 Differenzierbare Funktionen **108**

5.1 Der Ableitungsbegriff . . . 108

5.2 Die Ableitungsregeln . . . 112

5.3 Der Mittelwertsatz der Differentialrechnung . . . 119

5.4 Diskussion der transzendenten Funktionen . . . 127

5.5 Differentiation von Potenzreihen . . . 147

5.6 Eine Liste der Grundableitungen . . . 152

6 Höhere Ableitungen I **153**

6.1 Konvexe Funktionen . . . 154

7 Das Riemannsche Integral **174**

7.1 Zwei Axiome und der Hauptsatz . . . 174

7.2 Integrationstechniken . . . 185

7.3 Riemann–integrierbare Funktionen . . . 198

7.4 Uneigentliche Integrale . . . 206

8 Potenzreihen **212**

8.1 Die Taylorformel . . . 213

8.2 Potenzreihen in $\mathbb{C}$ . . . 227

9 Strukturen der Analysis **244**

9.1 Metrische Räume . . . 244

9.2 Normierte Räume . . . 266

9.3 Differentation in Banachräumen . . . 273

Literaturhinweise **277**

Kapitel 1

Die reellen Zahlen

Die Analysis befasst sich mit reellen Zahlen, Funktionen auf den reellen Zahlen und ganz vielem anderen mehr, was aber letztlich daraus hervorgeht. Ihre Entwicklung ist eng verknüpft mit dem Fortschritt von Physik und Ingenieurwissenschaften. Die grundlegenden Beiträge stammen fast zeitgleich von Gottfried Wilhelm Leibniz (1646 – 1716) und dem vielleicht größten Physiker dieses Jahrtausends Sir Isaac Newton (1642 – 1727). Allerdings war der Zahlbegriff bis ins 18. Jahrhundert hinein noch recht intuitiv und für das Abstraktionsniveau unserer heutigen Mathematik unzureichend. Erst Ende des 19. Jahrhunderts erfolgte die Präzisierung des Zahlbegriffs. Zu nennen sind hier an erster Stelle Karl Weierstraß (1815 – 1897), ferner Bernard Bolzano (1781 – 1848) als Vorläufer und Richard Dedekind (1831 – 1916), Georg Cantor (1845 – 1918), Guiseppe Peano (1858 – 1932).

In dieser Veranstaltung werden zunächst die reellen Zahlen axiomatisch eingeführt. Die Sprache, in der dies geschieht, benötigt einige Grundlagen der Mengenlehre, die Cantor um 1880 herum eingeführt hat. Die Mengenlehre ist ein Zweig der *Grundlagen der Mathematik*, dies auch ein Teilgebiet der Mathematik, auf der Grenze zur Philosophie angesiedelt. Wir interessieren uns aber mehr für die Mathematik, die in Naturwissenschaft und Technik Anwendung findet und benutzen die Mengenlehre i. w. als Sprechweise.

Nach Cantor ist eine Menge *eine Zusammenfassung von wohlbestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unserers Denkens zu einem Ganzen.* Dies ist keine mathematische Definition, aber sie ist für unsere Zwecke ausreichend. Z. B. ist die Menge aller Menschen mit italienischer Staatsangehörigkeit eine Menge, nennen wir sie I. Oder die Menge aller Menschen griechischer Staatsangehörigkeit ist eine Menge GR, ebenso die Mengen D, DK, F, ... der Menschen mit deutscher, dänischer, französischer, ... Staatsangehörigkeit. Ein einzelnes Objekt σ einer Menge M nennt man auch Element von M und schreibt

$$\sigma \in M$$

also z. B.

$$\text{Lionel Jospin} \in F \quad , \quad \text{Giovanni Trapatoni} \in I \quad , \quad \text{Tony Blair} \in GB$$

oder

$$7 \in \mathbb{N},$$

wobei $\mathbb{N}$ die Menge aller natürlichen Zahlen bezeichnet. Genaueres findet man im nächsten Abschnitt.

Es gibt verschiedene Methoden, die reellen Zahlen einzuführen — das sind die Zahlen, die man als abbrechende oder nicht abbrechende Dezimalbrüche schreiben kann. Die vielleicht befriedigendste Methode besteht darin, sie aus den natürlichen Zahlen $\mathbb{N} := \{1, 2, 3, ...\}$ zu konstruieren. Dabei konstruiert man zunächst aus $\mathbb{N}$ die ganzen Zahlen $\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{-n\colon n \in \mathbb{N}\}$ so, dass die Subtraktion unbeschränkt ausgeführt werden kann. Aus $\mathbb{Z}$ konstruiert man die rationalen Zahlen $\mathbb{Q} := \{\frac{p}{q}\colon p \in \mathbb{Z} \wedge q \in \mathbb{N}\}$ so, dass die Division (außer Division durch 0) unbeschränkt durchgeführt werden kann. Diese Konstruktionen sind *algebraischer* Natur.

Schon die alten griechischen Mathematiker haben den Zusammenhang zwischen der Menge der Zahlen und dem geometrischen Objekt *Gerade* gesehen: sie haben die rationalen Zahlen als Punkte auf der *Zahlengeraden* gedeutet. Umgekehrt will man die Menge der Zahlen als die Menge <u>aller</u> Punkte auf der Zahlengeraden deuten. Schon Pythagoras (570 – 497 v. Chr. (?)) wußte, dass $\mathbb{Q}$ nicht die gesamte Zahlengerade ausfüllt. (Einer seiner Schüler hat wohl mit seinem Leben dafür bezahlen müssen, dass er dieses geheime Wissen der Pythagoräer öffentlich gemacht hat.)

Mit einer analytischen Konstruktion kann man schließlich $\mathbb{Q}$ erweitern zur Menge $\mathbb{R}$ der reellen Zahlen, die dann wirklich die gesamte Zahlengerade ausfüllen. Man findet dieses Programm dargestellt in dem Buch [14].

Bei dem beschriebenen Vorgehen war man bei den natürlichen Zahlen gestartet. Man kann auch diese — ausgehend von der Mengenlehre — konstruieren, oder man hält es mit Kronecker (1823 – 1891): *Die natürlichen Zahlen hat der liebe Gott gemacht, alles Übrige ist Menschenwerk.* Dann würde man die natürlichen Zahlen axiomatisch beschreiben (mit Hilfe der Peano-Axiome) und die Konstruktion von $\mathbb{R}$ über $\mathbb{Z}$, $\mathbb{Q}$ durchführen.

Um aber möglichst bald zu dem zu kommen, was uns interessiert, starten wir mit der Angabe eines Axiomensystems für die reellen Zahlen. D. h. wir präzisieren das Grundwissen, das wir von den reellen Zahlen haben, durch ein möglichst einfaches System von Aussagen, aus denen dann alles über die reellen Zahlen hergeleitet wird. Insbesondere werden wir dann auch $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ als Teilmengen von $\mathbb{R}$ definieren.

Wir werden $\mathbb{R}$ einführen durch: *Die Menge* $\mathbb{R}$ *der reellen Zahlen ist ein geordneter, (ordnungs-)vollständiger Körper.*

Dies ist ein Programm für eine Definition: Wir haben zu erklären, was ein Körper ist und was es heißt, dass dieser Körper geordnet und vollständig ist.

Im folgenden Abschnitt sollen einige logische und mengentheoretische Grundlagen bereitgestellt werden. Der Abschnitt kann überschlagen werden. Man sollte ihn zu Rate ziehen, wenn Probleme mit den Schreibweisen auftreten.

1.1 Grundlagen aus Logik und Mengenlehre

Vorbemerkung: Die Sprache – aber nicht der Inhalt – der Mathematik sind Logik und Mengenlehre. So wie man eine Fremdsprache auch nicht einübt, indem man zunächst *alle* Vokabeln und grammatikalischen Regeln lernt und erst dann beginnt, sie zu sprechen, so werden die folgenden Begriffe je nach Bedarf erst nach und nach in der Vorlesung eingeführt. Daher ist dieser Abschnitt auch weniger zum Durcharbeiten und mehr zum Nachschlagen gedacht.

Für die Begriffe "Aussage", "Menge", "Element" einer Menge werden keine mathematischen Definitionen gegeben. Unter einer AUSSAGE verstehen wir ein sprachliches Gebilde, dem ein Wahrheitswert WAHR (w) oder FALSCH (f) eindeutig zugeschrieben werden kann. Unter einer MENGE versteht man "eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen" (G. Cantor 1845 - 1918). Die einzelnen Objekte, die da zusammengefaßt werden, heißen die ELEMENTE (der Menge).

In der mathematischen Logik werden Regeln für die Bildung neuer zulässiger Aussagen aus gegebenen Aussagen festgelegt, indem der Wahrheitswert der neuen in Abhängigkeit von den Wahrheitsswerten der gegebenen Aussagen festgelegt wird. Hierzu verwendet man i.a. "WAHRHEITSTAFELN" (s. u.).

Sind A, B Aussagen, so bedeuten:

$\neg A$	die Aussage	"nicht A"
$A \wedge B$	die Aussage	"A und B"
$A \vee B$	die Aussage	"A oder B"
$A \Rightarrow B$	die Aussage	"A impliziert B"
	alternativ :	"B folgt aus A"
		"wenn A, so B"
		"A ist hinreichend für B"
		"B ist notwendig für A"
$A \Leftrightarrow B$	die Aussage	"A und B sind äquivalent"
	alternativ :	"A genau dann, wenn B"

Die Wahrheitswerte sind dementsprechend durch die folgenden Wahrheitstafeln festgelegt:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Aussagenlogische REGELN ("Tautologien") sind nun solche zusammengesetzte Aussagen, die unabhängig vom Wahrheitsgehalt der Einzelaussagen wahr sind. In den folgenden Beispielen für Tautologien seien A, B, C irgendwelche Aussagen und F eine falsche Aussage. Schließlich legen wir zur Vermeidung allzu vieler Klammern fest, daß wenn nicht durch Klammern anders verlangt, "$\neg$" stets vor und "$\Rightarrow$" sowie "$\Leftrightarrow$" stets nach den Operatoren "$\wedge$" und "$\vee$" durchzuführen sind.

Natürlich kann die folgende Regelliste nicht vollständig sein:

Kommutativgesetze:

$$\left.\begin{array}{l} A \vee B \Leftrightarrow B \vee A \\ A \wedge B \Leftrightarrow B \wedge A \end{array}\right\} \qquad (K)$$

Assoziativgesetze:

$$\left.\begin{array}{l} (A \vee B) \vee C \Leftrightarrow A \vee (B \vee C) \\ (A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C) \end{array}\right\} \qquad (AS)$$

Distributivgesetze:

$$\left.\begin{array}{l} A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C) \\ A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C) \end{array}\right\} \qquad (D)$$

doppelte Verneinung:

$$\neg\,(\neg\, A) \Leftrightarrow A \qquad (DV)$$

Gesetze von de Morgan:

$$\left.\begin{array}{l} \neg\,(A \wedge B) \Leftrightarrow \neg A \vee \neg\, B \\ \neg\,(A \vee B) \Leftrightarrow \neg\, A \wedge \neg\, B \end{array}\right\} \qquad (dM)$$

Umformungen der Implikation,
Regeln für indirekte Beweise:

$$\left.\begin{array}{ccc} (A \Rightarrow B) & \Leftrightarrow & \neg A \vee B \\ \neg(A \Rightarrow B) & \Leftrightarrow & A \wedge \neg\, B \\ (A \Rightarrow B) & \Leftrightarrow & (\neg B \Rightarrow \neg A) \\ (A \Rightarrow B) & \Leftrightarrow & (A \wedge \neg B \Rightarrow \neg A) \\ (A \Rightarrow B) & \Leftrightarrow & (A \wedge \neg B \Rightarrow B) \\ (A \Rightarrow B) & \Leftrightarrow & (A \wedge \neg B \Rightarrow F) \end{array}\right\} \qquad (iB)$$

Regeln für logische Schlüsse:

$$\left.\begin{array}{l} A \wedge B \Rightarrow A \\ A \Rightarrow A \vee B \\ A \wedge (A \Rightarrow B) \Rightarrow B \\ (A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C) \\ (A \Rightarrow B) \wedge (A \Rightarrow C) \Rightarrow (A \Rightarrow B \wedge C) \end{array}\right\} \qquad (S)$$

Widerspruchsfreiheit:

$$\neg(A \wedge \neg A) \qquad (W)$$

Tertium non datur:

$$A \vee \neg A \qquad (T)$$

In der Mathematik, aber auch in der Umgangssprache, treten oft sprachliche Gebilde auf, in denen ein oder mehrere Subjekte durch einen oder mehrere Platzhalter ersetzt sind. Setzt man dann für den oder die Platzhalter geeignete Subjekte ein, so ergibt sich – je nach Einsetzung – eine wahre oder falsche Aussage. Gängiges Beispiel aus der Mathematik ist eine Gleichung, etwa "$x+7=9$": dies ist erst dann eine Aussage, wenn für x eine Zahl, also ein Element aus einer zuvor festgelegten Menge eingetragen wird. Beispiele aus der Umgangssprache sind Sätze mit Pronomina.

Ist M eine Menge und $a(x)$ ein sprachliches Gebilde, welches bei Ersetzung des Platzhalters x durch irgendein Element von M zu einer Aussage wird, so nennt man $a(x)$ eine "AUSSAGEFORM über M". Aus einer solchen Aussageform läßt sich aber nicht nur durch Einsetzung spezieller Elemente aus M eine Aussage machen, man kann auch Aussagen darüber machen, ob und eventuell wie viele Einsetzungen für x es in M gibt, die eine *wahre* Aussage liefern. Die Aussage, daß es überhaupt eine Einsetzung für x in M gibt, durch die $a(x)$ zu einer wahren Aussage wird, nennt man "die EXISTENZAUSSAGE über $a(x)$". Man schreibt

$$\bigvee_{x\in M} a(x)$$

und liest

"Es gibt ein x aus M, für das $a(x)$ gilt" .

Dies ist jetzt wieder eine Aussage, die wahr oder falsch sein kann:

Beispiele:	$\bigvee_{x\in\mathbb{R}}$	$x^2+1=0$	ist falsche Aussage
	$\bigvee_{x\in\mathbb{R}}$	$x^2\geq 0$	ist wahre Aussage
	$\bigvee_{x\in\mathbb{R}}$	$x^2-1=0$	ist ebenfalls wahre Aussage .

(Hier ist $\mathbb{R}$ die Menge der reellen Zahlen.)

Im zweiten Beispiel ist aber noch viel mehr wahr: Es gibt nicht nur *eine* reelle Zahl, deren Quadrat nicht negativ ist, sondern das Quadrat einer *jeden* reellen Zahl ist nicht negativ. Ist $a(x)$ eine Aussageform über M, so nennt man die Aussage, daß *jede* Einsetzung aus M für x $a(x)$ zu einer wahren Aussage macht, "die ALLAUSSAGE über $a(x)$". Man schreibt

$$\bigwedge_{x\in M} a(x) ;$$

und liest

"Für alle $x\in M$ gilt $a(x)$" .

Auch dies ist eine Aussage, die wahr oder falsch sein kann:

Beispiele:	$\bigwedge_{x\in\mathbb{R}}$	$x^2\geq 0$	ist wahr.
	$\bigwedge_{x\in\mathbb{R}}$	$x^2-1=0$	ist falsch.

Sprechweise: Die Variable(n) in einer Aussageform nennt man *freie* Variable. Überführt man eine Aussageform in eine Existenz- oder Allaussage, so sagt man: "Die freie Variable wird durch den Existenz- oder Allquantor gebunden".

Diese Sprechweise ist nützlich, wenn man die Situation beschreiben will, die sich bei Aussageformen mit mehreren Variablen ergibt. Sind M_1, M_2 Mengen und $a(x_1, x_2)$ ein sprachliches Gebilde, welches bei Ersetzung der Platzhalter x_1 und

x_2 durch ein Element aus M_1 und ein Element aus M_2 zu einer Aussage wird, so nennt man $a(x_1, x_2)$ eine (zweistellige) Aussagenform über $M_1 \times M_2$ (zur Bezeichnung "$M_1 \times M_2$" siehe unten). Analog definiert man 3-, 4- ..., n-stellige Aussagenformen über $M_1 \times M_2 \times M_3$, $M_1 \times M_2 \times M_3 \times M_4$ bzw. $M_1 \times \cdots \times M_n$. Bindet man eine der Variablen, etwa x_n, einer n-stelligen Aussageform mit einem Existenz- oder Allquantor, so erhält man eine $(n-1)$-stellige Aussageform:

Beispiele $(n = 2)$:

$x_1 + x_2 = 0$	zweistellige Aussageform über $\mathbb{R} \times \mathbb{R}$
$\left.\begin{array}{l} \bigvee_{x_2 \in \mathbb{R}} \ x_1 + x_2 = 0 \\ \bigwedge_{x_2 \in \mathbb{R}} \ x_1 + x_2 = 0 \end{array}\right\}$	einstellige Aussageformen über $\mathbb{R}$ x_1 ist freie, x_2 gebundene Variable
$\bigwedge_{x_1 \in \mathbb{R}} \ \bigvee_{x_2 \in \mathbb{R}} \ x_1 + x_2 = 0$	(wahre) Aussage: "Zu jedem $x_1 \in \mathbb{R}$ gibt es ein $x_2 \in \mathbb{R}$, so daß $x_1 + x_2 = 0$ ist".
$\bigvee_{x_1 \in \mathbb{R}} \ \bigwedge_{x_2 \in \mathbb{R}} \ x_1 + x_2 = 0$	(falsche) Aussage: "Es gibt ein $x_1 \in \mathbb{R}$, so daß für alle $x_2 \in \mathbb{R}$ gilt: $x_1 + x_2 = 0$".
$x + y = 1 \ \wedge \ x - y = 0$	ist zweistellige Aussageform über $\mathbb{R} \times \mathbb{R}$.

Wie im letzten Beispiel können die logischen Partikel "$\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$, $\neg$" beim Aufbau von Aussageformen verwendet werden.

Wir geben noch eine Auswahl von Regeln für den Umgang mit Quantoren an:

$$\begin{array}{l} \neg\left(\bigwedge_{x \in M} \ a(x)\right) \ \Leftrightarrow \ \bigvee_{x \in M} (\neg\, a(x)) \\ \neg\left(\bigvee_{x \in M} \ a(x)\right) \ \Leftrightarrow \ \bigwedge_{x \in M} (\neg\, a(x)) \end{array} \qquad \text{"De Morgansche Regeln"}$$

Ist $\hat{x}$ ein spezielles Element von M, so gilt stets

$$\bigwedge_{x \in M} a(x) \ \Rightarrow \ a(\hat{x}) \ \Rightarrow \ \bigvee_{x \in M} a(x) \ .$$

Hierbei steht eine *Kette* der Art $A \Rightarrow B \Rightarrow C$ für $(A \Rightarrow B) \wedge (B \Rightarrow C)$. Man beachte in diesem Zusammenhang die vierte Regel für logische Schlüsse. Ist schließlich $a(x, y)$ Aussageform über $X \times Y$, so hat man:

$$\bigwedge_{x \in X} \left(\bigwedge_{y \in Y} \ a(x,y)\right) \ \Leftrightarrow \ \bigwedge_{y \in Y} \left(\bigwedge_{x \in X} \ a(x,y)\right) \ \Leftrightarrow \ \bigwedge_{(x,y) \in X \times Y} \ a(x,y)$$

$$\bigvee_{x \in X} \left(\bigvee_{y \in Y} \ a(x,y)\right) \ \Leftrightarrow \ \bigvee_{y \in Y} \left(\bigvee_{x \in X} \ a(x,y)\right) \ \Leftrightarrow \ \bigvee_{(x,y) \in X \times Y} \ a(x,y) \ ;$$

aber (und hier gilt i. a. *nicht* "$\Leftarrow$"):

$$\bigvee_{x \in X} \left(\bigwedge_{y \in Y} \ a(x,y)\right) \ \Rightarrow \ \bigwedge_{y \in Y} \left(\bigvee_{x \in X} \ a(x,y)\right) \ .$$

Die Klammern werden meist fortgelassen.

Die Quantoren "$\bigwedge$" bzw. "$\bigvee$" können als Verallgemeinerungen von "$\wedge$" bzw. "$\vee$" aufgefaßt werden. Manche Autoren verwenden die Schreibweisen "$\forall$" bzw. "$\exists$".

Auch für Mengen wird keine mathematische Definition gegeben. Eine Menge ist festgelegt, wenn klar ist, welche Elemente sie enthält. Für ihre Beschreibung reservieren wir geschweifte Klammern. So wird eine endliche Menge oft durch Auflisten ihrer Elemente angegeben, z.B. $M := \{1,2,3,4\}$ oder man legt sie durch eine ihren Elementen zukommende Eigenschaft fest, etwa

$$M := \{x\colon x \text{ ist eine natürliche Zahl und kleiner als } 5\} = \{1,2,3,4\}\,.$$

Auch die Schreibweise

$$M := \{\, x \quad | \quad x \text{ ist eine natürliche Zahl und kleiner als } 5\}$$

ist hier gebräuchlich. (Wir haben an dieser Stelle übrigens zum ersten Mal den Definitionsdoppelpunkt ":=" benutzt; die Menge (oder auch Zahl etc.), die auf der Seite des Doppelpunkts steht, wird durch den Ausdruck auf der anderen Seite definiert. Ist eine Aussage zu definieren, so ersetzen wir sinngemäß ":=" durch ":⇔".)

Wir vereinbaren folgende Schreibweisen :

$a \in M$ soll heißen:	"a gehört zur Menge M" oder "a ist Element von M" oder "a liegt in M" .
$a \notin M \;:\Leftrightarrow \neg\,(a \in M)$	"a gehört nicht zur Menge M" etc.
$\emptyset$ bezeichnet	die LEERE MENGE: "$a \in M$" ist stets *falsch* .

Für den Umgang mit $\emptyset$ im Zusammenhang mit einer Aussageform $a(x)$ wird vereinbart:

$$\bigvee_{x\in\emptyset} a(x) \text{ ist stets } \textit{falsch}; \quad \bigwedge_{x\in\emptyset} a(x) \text{ ist stets } \textit{wahr}.$$

Sind L und M Mengen, so sagt man

$$\begin{aligned}
L \subset M \;&:\Leftrightarrow\; L \text{ ist "T}\textsc{eilmenge}\text{" von } M \\
&:\Leftrightarrow\; \bigwedge_{x\in L} x \in M \\
&\;\Leftrightarrow\; \text{Jedes Element von } L \text{ liegt auch in } M\,. \\
L = M \;&:\Leftrightarrow\; L \text{ und } M \text{ sind gleiche Mengen} \\
&:\Leftrightarrow\; L \subset M \;\wedge\; M \subset L\,.
\end{aligned}$$

Aus gegebenen Mengen L und M kann man neue Mengen bilden:

$\mathcal{P}(L) := \{X\colon X \subset L\}$ ist die "Menge aller Teilmengen von L", genannt: "POTENZMENGE von L". Beachte: $\emptyset \in \mathcal{P}(L)$ und $L \in \mathcal{P}(L)$ sind stets wahr.

$X := \{x \in L\colon a(x)\}$ ist die "LÖSUNGSMENGE der Aussageform $a(x)$ über L". Z.B.: $X = \{x \in \mathbb{R}\colon x^2 - 1 = 0\} = \{1, -1\}$.

$L \cap M := \{x \in L\colon x \in M\} = \{x\colon x \in L \wedge x \in M\} = \{x \in M\colon x \in L\}$ ist "der DURCHSCHNITT von M und L". In $L \cap M$ liegen genau die Objekte, die sowohl in L als auch in M liegen.

$L \cup M := \{x\colon x \in L \vee x \in M\}$ ist "die VEREINIGUNG von L und M". Darin liegen die Objekte, die in wenigstens einer der beiden Mengen L und M liegen. Will man zusätzlich ausdrücken, daß $L \cap M$ leer ist, dann schreibt man $L \dot{\cup} M$ ("DISJUNKTE VEREINIGUNG") .

$L \setminus M := \{x \in L\colon x \notin M\}$ ist "die DIFFERENZ von L und M" oder "das KOMPLEMENT von M in L". $L \setminus M$ enthält diejenigen Elemente aus L, die nicht in M liegen. Dabei wird *nicht* vorausgesetzt, daß M Teilmenge von L ist: z.B. ist $\{1, 2, 3\} \setminus \{3, 4, 5\} = \{1, 2\}$.

$L \times M := \{(l, m)\colon l \in L \wedge m \in M\}$ ist "das KARTESISCHE PRODUKT von L und M". Seine Elemente sind "GEORDNETE PAARE aus Elementen von L und M". Ist $L = M$, so schreibt man oft L^2 statt $L \times L$. Daß die Paare (l, m) geordnet sind, bedeutet, daß es auf die Reihenfolge ankommt: ist etwa $L = M = \mathbb{R}$, so ist $(0, 1) \neq (1, 0)$.

Sind die Elemente von L selbst Mengen, so nennt man der sprachlichen Klarheit wegen L auch eine "FAMILIE von Mengen". Die Menge

$$\bigcap_{M \in L} M := \{x\colon \bigwedge_{M \in L} x \in M\}$$

nennt man dann den "Durchschnitt über die Familie L"; er enthält sämtliche Elemente, die allen Mengen aus der Familie L gemeinsam sind. Die Menge

$$\bigcup_{M \in L} M := \{x\colon \bigvee_{M \in L} x \in M\}$$

heißt "Vereinigung über die Familie L"; sie enthält sämtliche Elemente, die in wenigstens einer der Mengen der Familie L vorkommen.

Aussagenlogische Regeln lassen sich auf Mengenverknüpfungen und –relationen übertragen. Hierzu seien L, M, N, W Mengen mit $L \cup M \cup N \subset W$.

Dann gelten:

$\left.\begin{array}{l} L \cup M = M \cup L \\ L \cap M = M \cap L \end{array}\right\}$	(K)
$\left.\begin{array}{l} (L \cup M) \cup N = L \cup (M \cup N) \\ (L \cap M) \cap N = L \cap (M \cap N) \end{array}\right\}$	(AS)
$\left.\begin{array}{l} L \cap (M \cup N) = (L \cap M) \cup (L \cap N) \\ L \cup (M \cap N) = (L \cup M) \cap (L \cup N) \end{array}\right\}$	(D)
$W \backslash (W \backslash L) = L$	(DV)
$\left.\begin{array}{l} W \backslash (L \cap M) = (W \backslash L) \cup (W \backslash M) \\ W \backslash (L \cup M) = (W \backslash L) \cap (W \backslash M) \end{array}\right\}$	(dM)
$L \cap M \subset L\ ;\ L \subset L \cup M$	(S)
$\left.\begin{array}{l} L \subset M \wedge M \subset N \Rightarrow L \subset N \\ L \subset M \wedge L \subset N \Rightarrow L \subset M \cap N \end{array}\right\}$	(S)
$L \setminus L = \emptyset$	(W)
$L \cup (W \backslash L) = W$	(T)

Die Regeln für All- und Existenzquantor übertragen sich auf den Durchschnitt und die Vereinigung einer Mengenfamilie $\mathcal{F}$. Für beliebige Mengen Y gelten nämlich

$$\begin{aligned} Y \setminus \bigcap_{X \in \mathcal{F}} X &= \bigcup_{X \in \mathcal{F}} (Y \setminus X) \\ Y \setminus \bigcup_{X \in \mathcal{F}} X &= \bigcap_{X \in \mathcal{F}} (Y \setminus X) \end{aligned}$$

und, falls $\hat{X}$ eine spezielle Menge der Familie $\mathcal{F}$ ist,

$$\bigcap_{X \in \mathcal{F}} X \subset \hat{X} \subset \bigcup_{X \in \mathcal{F}} X.$$

1.2 Geordnete Körper

Definition 1.1 *Ein Körper $\mathbb{K}$ ist eine nicht–leere Menge, auf der zwei Verknüpfungen "+" (die Addition) und "·" (die Multiplikation) erklärt sind, durch die je zwei Elementen $a, b \in \mathbb{K}$ jeweils ein weiteres Element $a + b$ bzw. $a \cdot b$ aus $\mathbb{K}$ eindeutig zugeordnet ist derart, dass folgende Gesetze gelten:*

(i) *die Kommutativgesetze: für alle $x, y \in \mathbb{K}$ ist*

$$x + y = y + x\ , \quad x \cdot y = y \cdot x \quad ,$$

(ii) *die Assoziativgesetze: für alle* $x, y, z, \in \mathbb{K}$ *ist*

$$(x+y)+z = x+(y+z) =: x+y+z \quad ,$$
$$(x \cdot y) \cdot z = x \cdot (y \cdot z) =: x \cdot y \cdot z \quad ,$$

(iii) *das Distributivgesetz: für alle* $x, y, z \in \mathbb{K}$ *ist*

$$x \cdot (y+z) = (x \cdot y) + (x \cdot z) \quad ,$$

(iv) *die Existenz neutraler Elemente: es existieren spezielle Elemente* $0 \in \mathbb{K}$, $1 \in \mathbb{K}$, *so dass für alle* $x \in \mathbb{K}$ *gelten*

$$x + 0 = x\,, \quad x \cdot 1 = x \quad ,$$

(v) *die Existenz inverser Elemente: zu jedem* $x \in \mathbb{K} \setminus \{0\}$ *existieren ein* $a \in \mathbb{K}$ *und ein* $b \in \mathbb{K}$, *so dass gelten*

$$x + a = 0\,, \quad x \cdot b = 1$$

(dieses a nennen wir “$-x$*”, und das b nennen wir “*$\frac{1}{x}$*” bzw. “*x^{-1}*”.),*

(vi) *schließlich:*

$$1 \neq 0$$

Man kann aus den Körperaxiomen die bekannten Regeln für algebraische Umformungen herleiten. Ich führe hier einige auf, und lege noch übliche Schreibweisen fest:

Für alle Elemente a, b eines Körpers $\mathbb{K}$ gelten

$$\bigvee_{x \in \mathbb{K}}^{1} a + x = b \tag{1.1}$$

(nämlich $x = b + (-a)$),

$$a \neq 0 \Rightarrow \bigvee_{x \in \mathbb{K}}^{1} a \cdot x = b \tag{1.2}$$

(nämlich $x = a^{-1} \cdot b$),

$$a \cdot b = 0 \Leftrightarrow a = 0 \vee b = 0 \tag{1.3}$$

(d.i die Regel vom verschwindenden Produkt),

$$-(a \cdot b) = (-a) \cdot b \tag{1.4}$$

(insbesondere gilt $-b = (-1) \cdot b$ für $a = 1$),

$$(-a) \cdot (-b) = (a \cdot b) \quad . \tag{1.5}$$

Üblicherweise schreibt man

$$b - a := b + (-a) \ , \quad \frac{a}{b} := a \cdot b^{-1} \tag{1.6}$$

(falls $b \neq 0$) und läßt den Punkt "$\cdot$" oft fort, wenn das 2. Argument eine Variable ist oder in einer Klammer steht:

$$ab := a \cdot b \ , \quad 3(5+2) = 21 \quad . \tag{1.7}$$

In (1.6) sind also Subtraktion und Division definiert. Schließlich vereinbart man noch die Vorfahrtsregel *Punktrechnung geht vor Strichrechnung*:

$$ab + cd := (ab) + (cd)$$

Man beachte die Regeln für den Umgang mit Quotienten: Zunächst ist (vgl. (1.3)) $0 \cdot a = 0$ für jedes $a \in \mathbb{K}$. 0 besitzt also kein mulitiplikatives Inverses: man kann durch 0 nicht dividieren. Sind $a, b \in \mathbb{K}$ und $c, d \in \mathbb{K} \setminus \{0\}$, so gelten

$$\frac{a}{c} = \frac{b}{d} \Leftrightarrow ad = bc \tag{1.8}$$

$$\frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd} \tag{1.9}$$

$$\left(\frac{c}{d}\right)^{-1} = \frac{d}{c} \tag{1.10}$$

Insgesamt kann man also mit Klammern und den Grundrechenarten umgehen, wie man es (hoffentlich) gewohnt ist.

Beispiele für Körper sind $\mathbb{F}_3 := \{0; 1; 2\}$ oder $\mathbb{F}_5 := \{0, 1, 2, 3, 4\}$ mit Addition und Multiplikation *modulo 3* bzw. *modulo 5*. Addition und Multiplikation sind durch folgende Tabellen erklärt:

$\mathbb{F}_3$:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

$\mathbb{F}_5$:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Natürlich werden auch $\mathbb{R}$ und $\mathbb{Q}$ Beispiele für Körper sein.

Definition 1.2 *Ein Körper $\mathbb{K}$ heißt geordnet, wenn er eine Teilmenge $\mathbb{K}^+$, die Menge der positiven Elemente, enthält, so dass gelten:*

(i) *Summe und Produkt positiver Elemente sind positiv:*

$$\bigwedge_{a,b\in\mathbb{K}^+} a+b\,,\ ab\in\mathbb{K}^+ \tag{1.11}$$

(ii) 0 *ist nicht positiv:*

$$0\notin\mathbb{K}^+ \tag{1.12}$$

(iii) *Elemente aus $\mathbb{K}$ sind entweder 0, positiv oder negativ:*

$$\bigwedge_{a\in\mathbb{K}\setminus\{0\}} a\in\mathbb{K}^+\vee -a\in\mathbb{K}^+ \tag{1.13}$$

Die Definition impliziert, dass $\mathbb{K}$ disjunkt zerlegt werden kann in

$$\mathbb{K}=\mathbb{K}^+\dot\cup\{0\}\dot\cup\mathbb{K}^- \tag{1.14}$$

mit $\mathbb{K}^-:=\{-a:a\in\mathbb{K}^+\}$. In der Tat ist $\mathbb{K}^+\cap\mathbb{K}^-=\emptyset$; denn läge $a\in\mathbb{K}^+\cap\mathbb{K}^-$, so wären sowohl a als auch $-a$ in $\mathbb{K}^+$ und damit wäre nach (1.11)

$$0=a+(-a)\in\mathbb{K}^+$$

im Widerspruch zu (1.12). Die Elemente aus $\mathbb{K}^+$ nennen wir *positiv* oder *mit Signum* $+$, die aus $\mathbb{K}^-$ *negativ* oder *mit Signum* $-$. In der Menge der reellen Zahlen oder der rationalen Zahlen wird $\mathbb{K}^+$ die Menge der positiven reellen (rationalen) Zahlen sein.

Einige Regeln wollen wir herleiten: Quadrate sind nie negativ; das Produkt zweier negativer Zahlen ist positiv; ein Produkt zweier Zahlen ist negativ, wenn der eine Faktor positiv, der andere negativ ist.

Satz 1.3 *In einem geordneten Körper $\mathbb{K}$ gelten für alle $a,b\in\mathbb{K}$:*

$$ab\in\mathbb{K}^+\Leftrightarrow(a\in\mathbb{K}^+\wedge b\in\mathbb{K}^+)\vee(a\in\mathbb{K}^-\wedge b\in\mathbb{K}^-)\quad, \tag{1.15}$$

$$ab\in\mathbb{K}^-\Leftrightarrow(a\in\mathbb{K}^-\wedge b\in\mathbb{K}^+)\vee(a\in\mathbb{K}^+\wedge b\in\mathbb{K}^-)\quad, \tag{1.16}$$

$$a^2:=a\cdot a\in\mathbb{K}^+\cup\{0\}\,,\quad a^2=0\Leftrightarrow a=0\quad, \tag{1.17}$$

$$\bigwedge_{a\in\mathbb{K}^+}\frac{1}{a}\in\mathbb{K}^+\,,\quad \bigwedge_{a\in\mathbb{K}^-}\frac{1}{a}\in\mathbb{K}^- \tag{1.18}$$

sowie

$$1\in\mathbb{K}^+ \tag{1.19}$$

Beweis:

Zu (1.15) “$\Leftarrow$”

$ab \in \mathbb{K}^+$, *falls* $a \in \mathbb{K}^+ \wedge b \in \mathbb{K}^+$ gilt nach 1.2.i. Sind $a, b \in \mathbb{K}^-$, so liefert 1.2.iii, dass $(-a), (-b) \in \mathbb{K}^+$ sind, und mit 1.2.i und Formel (1.5) erhält man

$$ab = (-a)(-b) \in \mathbb{K}^+ .$$

Zu (1.16) “$\Leftarrow$”

Ist $a \in \mathbb{K}^-$, so ist $(-a) \in \mathbb{K}^+$, und nach 1.2.i, (1.4) folgt für $b \in \mathbb{K}^+$

$$-(ab) = (-a) \cdot b \in \mathbb{K}^+,$$

also $ab \in \mathbb{K}^-$. Analog impliziert $a \in \mathbb{K}^+$, $b \in \mathbb{K}^-$, dass $ab \in \mathbb{K}^-$ ist.

Zu (1.15) “$\Rightarrow$”

Ist $ab \in \mathbb{K}^+$, so sind $a \neq 0$ und $b \neq 0$ nach (1.3). Wäre ein Element positiv und das andere negativ, so hätte man $ab \in \mathbb{K}^-$ wegen “(1.16) $\Leftarrow$”. Also müssen beide Elemente gleiches Signum haben.

(1.16) “$\Rightarrow$” beweist man analog.

Zu (1.17) Für $a \in \mathbb{K} \setminus \{0\}$ ist $a^2 \in \mathbb{K}^+$ wegen (1.15). Wegen $0^2 = 0$ folgt der zweite Teil von (1.17).

(1.19) folgt aus 1.1.vi und (1.17); denn $1 = 1^2$.

(1.18) folgt aus (1.15) bzw. (1.16) und (1.19)(v), denn $\frac{1}{a} \cdot a = 1 \in \mathbb{K}^+$

q.e.d.

Was hat die Festlegung 1.2 nun mit einer *Anordnung* zu tun?

Definition 1.4 *In einem geordneten Körper* $\mathbb{K}$ *schreibt man für* $a, b \in \mathbb{K}$:

$$a > b :\Leftrightarrow b < a :\Leftrightarrow a - b \in \mathbb{K}^+ \tag{1.20}$$

$$a \geq b :\Leftrightarrow b \leq a :\Leftrightarrow a > b \vee a = b \Leftrightarrow a - b \in \mathbb{K}^+ \cup \{0\} \tag{1.21}$$

Man liest “$<$” als kleiner*, “$>$” als* größer*, “$\leq$” als* kleiner oder gleich *und “$\geq$” als* größer oder gleich*.*

Natürlich ist $a > 0$ gleichbedeutend mit $a \in \mathbb{K}^+$ bzw. $a < 0$ mit $a \in \mathbb{K}^-$ sowie $a \geq 0$ mit $a \in \mathbb{K}^+ \cup \{0\}$, $a \leq 0$ mit $a \in \mathbb{K}^- \cup \{0\}$.

Die folgenden Regeln für den Umgang mit den Ordnungsrelationen sollten eigentlich bekannt sein.

Satz 1.5 $\mathbb{K}$ *sei ein geordneter Körper. Dann gilt für alle* $a, b, c, d \in \mathbb{K}$*:*

(i) *Entweder ist* $a = b$ *oder* $a > b$ *oder* $a < b$*; die drei Möglichkeiten schließen einander aus.*

(ii)

$$a < b \wedge b < c \Rightarrow a < c \tag{1.22}$$

Hierbei kann in der Prämisse eines der beiden $<$*–Zeichen durch* $\leq$ *ersetzt werden.*

(iii)

$$a < b \wedge c \leq d \Rightarrow a + c < b + d \tag{1.23}$$

(iv)

$$a < b \wedge c > 0 \Rightarrow ac < bc \quad ; \quad a < b \wedge c < 0 \Rightarrow ac > bc \tag{1.24}$$

(v)

$$a < b \wedge ab > 0 \Rightarrow \frac{1}{a} > \frac{1}{b} \quad ; \quad a < b \wedge ab < 0 \Rightarrow \frac{1}{a} < \frac{1}{b} \tag{1.25}$$

(vi)

$$a < b \Rightarrow a < \frac{a+b}{2} < b \quad \text{, wobei} \quad 2 := 1 + 1 > 1\ . \tag{1.26}$$

In (vi) tritt eine Kette von Ungleichungen auf: so bedeutet natürlich $a < b \leq c$, dass $a < b$ und $b \leq c$ ist, und man kann daraus $a < c$ wegen (ii) folgern. Man beachte insbesondere, dass nach (iv) bei Division und Multiplikation das Ordnungssymbol umgedreht wird, wenn mit einer negativen Zahl mulitpliziert oder dividiert wird. Ist der Multiplikator bzw. der Divisor positiv, so bleibt das Ungleichheitszeichen erhalten.

(ii) ist eine übliche Forderung an eine Ordnungsrelation. Besitzt eine Relation “$<$” zwischen Elementen einer Menge die Eigenschaft (ii), so sagt man, dass auf

M eine *teilweise Ordnung* erklärt ist. Ist etwa $M := \mathcal{P}(S)$ die Potenzmenge einer Menge S, so liefert die Inklusionsbeziehung "$U \subset V$" eine teilweise Ordnung auf M. Eine Menge, auf der eine Relation $<$ mit den Eigenschaften (i) und (ii) erklärt ist, nennt man auch *total geordnet*.

Beweis (von Satz 1.5**):** Wir wollen nur die Punkte (iii), (iv) und (vi) beweisen:

(iii) Zu zeigen ist $(b+d)-(a+c) \in \mathbb{K}^+$. Aber

$$(b+d)-(a+c) = (b-a)+(d-c)$$

Nach Voraussetzung ist $(b-a) \in \mathbb{K}^+$, $(d-c) \in \mathbb{K}^+ \cup \{0\}$ und damit liegt die Summe der beiden Terme in $\mathbb{K}^+$.

(iv) $bc - ac = (b-a)c \in \mathbb{K}^\pm$ wegen $b-a \in \mathbb{K}^+$ und $c \in \mathbb{K}^\pm$.

(vi) Natürlich gelten $\frac{1}{2}+\frac{1}{2}=1$ sowie $\frac{1}{2}>0$. Nun liefert $a<b$:

$$a = \frac{1}{2}a + \frac{1}{2}a < \frac{1}{2}a + \frac{1}{2}b < \frac{1}{2}b + \frac{1}{2}b = b\,.$$

q.e.d.

Aus (vi) kann man erkennen, dass in einem geordneten Körper $\mathbb{K}$ zwischen je zwei verschiedenen Elementen *unendlich viele* weitere Elemente liegen.

Im folgenden Beispiel stelle man sich den geordneten Körper $\mathbb{K}$ als Körper der reellen Zahlen oder der rationalen Zahlen vor. Aber unabhängig davon benutzen wir die Schreibweise

$$2 := 1+1 \quad , \quad 3 := 2+1 \quad , \quad 4 := 3+1 \qquad \text{etc.}$$

Beispiel 1.6 Finde die Menge aller $x \in \mathbb{K}$, für die

$$\frac{x-2}{3x+4} > 1$$

ist.

$$\begin{aligned}
&\frac{x-2}{3x+4} > 1 \\
&\qquad \Leftrightarrow \left(x > -\frac{4}{3} \wedge x-2 > 3x+4\right) \vee \left(x < -\frac{4}{3} \wedge x-2 < 3x+4\right) \\
&\qquad \Leftrightarrow \left(x > -\frac{4}{3} \wedge -6 > 2x\right) \vee \left(x < -\frac{4}{3} \wedge -6 < 2x\right) \\
&\qquad \Leftrightarrow \left(x > -\frac{4}{3} \wedge x < -3\right) \vee \left(x < -\frac{4}{3} \wedge -3 < x\right) \Leftrightarrow -3 < x < -\frac{4}{3}
\end{aligned}$$

Die Lösungsmenge $\mathbb{L} := \{x \in \mathbb{K} : \frac{x-2}{3x+4} > 1\}$ der Ungleichung aus Beispiel 1.6 lässt sich in Intervallschreibweise angeben:

Definition 1.7 *Sei $\mathbb{K}$ ein geordneter Körper und $a, b \in \mathbb{K}$ und $a \leq b$.*

(i) $[\,a\,,\,b\,] := \{x \in \mathbb{K}\colon a \leq x \leq b\}$ *heißt abgeschlossenes Intervall von a bis b.*

(ii) $(\,a\,,\,b\,] := \{x \in \mathbb{K}\colon a < x \leq b\}$ *heißt (links) halboffenes Intervall von a bis b.*

(iii) $[\,a\,,\,b\,) := \{x \in \mathbb{K}\colon a \leq x < b\}$ *heißt (rechts) halboffenes Intervall von a bis b.*

(iv) $(\,a\,,\,b\,) := \{x \in \mathbb{K}\colon a < x < b\}$ *heißt offenes Intervall von a bis b.*

(v) *In diesen vier Fällen nennt man $b - a$ die Länge des Intervalls.*

(vi) $(\,-\infty\,,\,a\,] := \{x \in \mathbb{K}\colon x \leq a\}$ *bzw.* $[\,a\,,\,\infty\,) := \{x \in \mathbb{K}\colon x \geq a\}$ *heißen abgeschlossenes Intervall von $-\infty$ bis a bzw. von a bis ∞.*

(vii) $(\,-\infty\,,\,a\,) := \{x \in \mathbb{K}\colon x < a\}$ *bzw.* $(\,a\,,\,\infty\,) := \{x \in \mathbb{K}\colon x > a\}$ *heißen offenes Intervall von $-\infty$ bis a bzw. von a bis ∞.*

(viii) $(-\infty, \infty) := \mathbb{K}$.

(ix) *Ist $\varepsilon \in \mathbb{K}^+$, so nennt man $U_\varepsilon(a) := (a-\varepsilon, a+\varepsilon)$ bzw. $K_\varepsilon(a) := [a-\varepsilon, a+\varepsilon]$ offene bzw. abgeschlossene ε-Umgebung von a.*

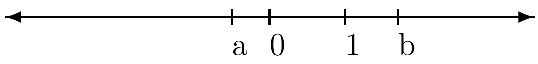
Eine grundlegende Rolle in der Analysis spielt die Betragsfunktion:

Definition 1.8 *Sei $\mathbb{K}$ ein geordneter Körper und $a, b \in \mathbb{K}$. Dann heißen*

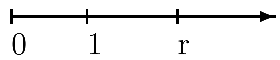
$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases} \tag{1.27}$$

der Betrag von a und $|a - b|$ der Abstand zwischen a und b.

Man stelle sich unter dem geordneten Körper den Körper $\mathbb{R}$ der rellen Zahlen so wie aus der Schule bekannt vor. Man identifiziere ihn mit einer Geraden, der *Zahlengeraden*:



Denkt man sich als Maßband die Halbgerade $\mathbb{R}^+ \cup \{0\}$,



so ist $r :=\mid a - b \mid$ die Zahl die man auf der Halbgerade über der größeren der Zahlen a oder b abliest. Wenn man den Anfangspunkt der Halbgeraden auf die kleinere der beiden Zahlen a oder b legt. Wegen $|a| = |a - 0|$ deutet man $|a|$ als Abstand von a zu 0.

Zum Umgang mit der Betragsfunktion:

Satz 1.9 *In einem geordneten Körper* $\mathbb{K}$ *gelten*

$$\bigwedge_{a \in \mathbb{K} \setminus \{0\}} |a| > 0 \qquad (|\cdot| \text{ ist positiv definit}), \tag{1.28}$$

insbesondere wegen $|0| = 0$

$$|a| = 0 \Leftrightarrow a = 0\,, \tag{1.29}$$

$$\bigwedge_{a,b \in \mathbb{K}} \mid ab \mid = \mid a \mid\mid b \mid\,, \tag{1.30}$$

$$\bigwedge_{a,b \in \mathbb{K}} \mid a + b \mid \leq \mid a \mid + \mid b \mid \qquad \text{(Dreiecksungleichung)} \quad . \tag{1.31}$$

Der Name *Dreiecksungleichung* wird klarer, wenn man (1.31) in der Form

$$\bigwedge_{a,b,c \in \mathbb{K}} \mid a - c \mid \leq \mid a - b \mid + \mid b - c \mid \tag{1.32}$$

aufschreibt: *Der (direkte) Weg von a zu c ist nie größer als der Umweg von a zu c über irgendein b.*

Auch liefert (1.30)

$$\bigwedge_{a,b \in \mathbb{K}} \mid a - b \mid = \mid b - a \mid \tag{1.33}$$

Beweis:

zu (1.30)**:** Es sind $a \in \{| a |, - | a |\}, b \in \{| b |, - | b |\}$, daher ist ab eine der beiden Zahlen $- | a || b |$ oder $| a || b |\}$ und mithin $| ab |=| a || b |$.

zu (1.31)**:**

$$|a+b| = \left\{ \begin{array}{ll} a+b\,, & \text{falls } a+b > 0 \\ -a-b\,, & \text{falls } a+b < 0 \end{array} \right\} \le |a| + |b|\ ,$$

denn für jedes $c \in \mathbb{K}$ ist sowohl $c \le | c |$ als auch $-c \le | c |$.

q.e.d.

Im zweiten Semester wird der Begriff des Abstands und der Länge im allgemeineren Situationen eingeführt. Dabei werden i. w. die in Satz 1.9 beschriebenen Eigenschaften verwendet.

Weitere Eigenschaften der Betragsfunktion:

Satz 1.10 *In einem geordneten Körper* $\mathbb{K}$ *gelten*

$$\bigwedge_{a\in\mathbb{K}} \bigwedge_{b\in\mathbb{K}\setminus\{0\}} \left|\frac{a}{b}\right| = \frac{|a|}{|b|}\ , \tag{1.34}$$

$$\bigwedge_{a,b\in\mathbb{K}} ||a| - |b|| \le |a \pm b|\ , \tag{1.35}$$

$$\bigwedge_{a\in\mathbb{K}} \bigwedge_{\varepsilon\in\mathbb{K}^+} U_\varepsilon(a) = \{x \in \mathbb{K}\colon |x-a| < \varepsilon\} \tag{1.36}$$

$$\bigwedge_{a\in\mathbb{K}} \bigwedge_{\varepsilon\in\mathbb{K}^+} K_\varepsilon(a) = \{x \in \mathbb{K}\colon |x-a| \le \varepsilon\} \tag{1.37}$$

Beweis:

zu (1.34)**:** $| \frac{a}{b} | \cdot | b |=| \frac{a}{b} \cdot b |=| a |\Rightarrow| \frac{a}{b} |= \frac{|a|}{|b|}$.

zu (1.35)**:** Wir nehmen o. B. d. A.[1] an, dass $| a |\ge| b |$ ist. Dann folgt

$$|| a | - | b ||=| a | - | b |=| a{\pm}b{\mp}b | - | b |\le| a{\pm}b | + | b | - | b |=| a{\pm}b |\ .$$

zu (1.36) Im anderen Fall ist $<$ an den entsprechenden Stellen durch $\le$ zu ersetzen.

$$\begin{aligned} | x-a |< \varepsilon &\Leftrightarrow (x \ge a \wedge x-a < \varepsilon) \vee (x < a \wedge a-x < \varepsilon) \\ &\Leftrightarrow x \in [a, a+\varepsilon) \vee x \in (a-\varepsilon, a) \Leftrightarrow x \in U_\varepsilon(a)\ . \end{aligned}$$

[1] ohne Beschränkung der Allgemeinheit

(1.37) beweist man analog zu (1.36).

q.e.d.

1.3 Die natürlichen Zahlen

Zugrunde gelegt sei ein geordneter Körper, den wir ab jetzt $\mathbb{R}$ nennen wollen. Es ist aber in diesem Abschnitt unwichtig, ob das tatsächlich der Körper der reellen Zahlen ist. Dennoch wollen wir seine Elemente *Zahlen* nennen. Jedenfalls enthält er eine Teilmenge, die unseren Vorstellungen von der Menge der natürlichen Zahlen entspricht:

$$1 \quad ; \quad 2 := 1+1 \quad ; \quad 3 := 2+1 \quad ; \quad 4 := 3+1 \quad ; \quad \ldots$$

Die Menge der natürlichen Zahlen soll so definiert werden, dass sie die 1 enthält und alle durch sukzessive Addition von 1 entstehenden Zahlen. Nach deutscher Industrienorm wird auch 0 als natürliche Zahl angegeben. Das wird aber nicht von allen Mathematikern mitgetragen, auch nicht in dieser Vorlesung.

Definition 1.11 *Eine Teilmenge M von $\mathbb{R}$ heißt induktiv, wenn*

$$1 \in M \tag{1.38}$$

und

$$\bigwedge_{m \in M} m+1 \in M \tag{1.39}$$

gelten.

Beispiele für induktive Mengen sind $\mathbb{R}$, $[1,\infty)$, $\mathbb{R}^+ = (0,\infty)$ und natürlich $\mathbb{N}$, die Menge der natürlichen Zahlen. Genauer: Wir werden $\mathbb{N}$ als die *kleinste* induktive Teilmenge von $\mathbb{R}$ definieren.

Satz und Definition 1.12

$$\begin{aligned} \mathbb{N} &:= \bigcap \{M \subset \mathbb{R} \colon M \quad \text{ist induktiv} \quad \} \\ &= \{x \in \mathbb{R} \colon x \quad \text{liegt in jeder induktiven Teilmenge von } \mathbb{R} \quad \} \end{aligned} \tag{1.40}$$

ist (selbst) induktiv und heißt (Modell für die) Menge der natürlichen Zahlen. Ihre Elemente heißen natürliche Zahlen. *Insbesondere ist $\mathbb{N} \neq \emptyset$ und $\mathbb{N} \subset M$ für jede induktive Teilmenge M von $\mathbb{R}$.*

Beweis: Zeigen muss man, dass $\mathbb{N}$ induktiv ist. Dann ist klar, dass $\mathbb{N}$ nicht leer ist ($1 \in \mathbb{N}$), und aus der definierenden Gleichung (1.40) erhält man $\mathbb{N} \subset M$ für jedes induktive M, d. h. $\mathbb{N}$ ist *kleinste* induktive Teilmenge von $\mathbb{R}$. Da $1 \in M$ gilt für jede induktive Teilmenge $M \subset \mathbb{R}$, ist auch $1 \in \mathbb{N}$, und liegt n in $\mathbb{N}$, also in jeder induktiven Teilmenge M von $\mathbb{R}$, so liegt auch $n+1$ in jeder induktiven Teilmenge M von $\mathbb{R}$, d. h . $n+1 \in \mathbb{N}$. $\mathbb{N}$ ist also induktiv.

q.e.d.

Da $[\,1\,,\,\infty\,)$ induktiv ist, können wir festhalten:

Korollar 1.13 $\mathbb{N} \subset [1, \infty)$.

Die Definition von $\mathbb{N}$ als kleinste induktive Teilmenge von $\mathbb{R}$ erlaubt das Beweisprinzip der vollständigen Induktion. Wir stellen noch einmal heraus:

Satz 1.14 *Sei L eine Teilmenge von $\mathbb{N}$. Ist*

$$1 \in L \tag{1.41}$$

und gilt

$$\bigwedge_{\ell \in L} \quad \ell + 1 \in L \quad , \tag{1.42}$$

so ist $L = \mathbb{N}$.

Denn (1.41) und (1.42) bedeuten, dass L induktiv ist, also $\mathbb{N} \subset L$ gilt. Da aber nach Voraussetzung auch $L \subset \mathbb{N}$ ist, muss $\mathbb{N} = L$ sein.

Hat man nun eine Aussageform $A(n)$ über $\mathbb{N}$, zum Beispiel

$$A(n) :\Leftrightarrow \bigwedge_{a \in (-1,\infty)} (1+a)^n \geq 1 + na \ , \tag{1.43}$$

so kann man die Lösungsmenge

$$\mathbb{L} := \{n \in \mathbb{N}\colon\ A(n)\}$$

dieser Aussageform bilden, also die Menge aller natürlichen Zahlen, die $A(n)$ zu einer wahren Aussage machen. Zum Beispiel rechnet man für die Aussageform in (1.43) leicht nach

$$\begin{aligned} A(2) &\Leftrightarrow \bigwedge_{a \in (-1,\infty)} (1+a)^2 \geq 1 + 2a \Leftrightarrow \bigwedge_{a \in (-1,\infty)} 1 + 2a + a^2 \geq 1 + 2a \\ &\Leftrightarrow \bigwedge_{a \in (-1,\infty)} a^2 \geq 0 \quad , \end{aligned}$$

und das stimmt. Also erhält man $2 \in \mathbb{L}$. Will man nun beweisen, dass $A(n)$ für alle $n \in \mathbb{N}$ wahr ist, so muss $\mathbb{L} = \mathbb{N}$ bewiesen werden. Hierzu kann man $A(1)$ zeigen sowie *"Gilt für irgendein $n \in \mathbb{N}$ die Aussage $A(n)$, so gilt auch $A(n+1)$"*. Ein Induktionsbeweis besteht also aus zwei Schritten, der Induktionsverankerung (Nachweis von $A(1)$), und dem Induktionsschluss (Nachweis von $A(n+1)$ unter Annahme der Gültigkeit[1] von $A(n)$). In Formeln:

$$[\bigwedge_{n\in\mathbb{N}} A(n)] \Leftrightarrow [A(1) \wedge \bigwedge_{n\in\mathbb{N}} (A(n) \Rightarrow A(n+1))] \tag{1.44}$$

In unserem Beispiel ist

$$\bigwedge_{n\in[-1,\infty)} (1+a)^1 \geq 1 + 1 \cdot a$$

die Induktionsverankerung und die ist offensichtlich wahr.

Zur Durchführung des Induktionsschlusses machen wir die Induktionsannahme, für irgendein $n \in \mathbb{N}$ gelte $A(n)$. Dann folgt

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

denn $1 + a > 0$ und $(1+a)^n \geq 1 + na$ nach Induktionsannahme.

Damit haben wir die Bernoullische Ungleichung bewiesen:

Satz 1.15

$$\bigwedge_{n\in\mathbb{N}} \quad \bigwedge_{a\in(\,-1\,,\,\infty\,)} \quad (1+a)^n \geq 1 + na \tag{1.45}$$

Es gibt vieles aus der Arithmetik, das für Sie selbstverständlich ist. Hiervon wollen wir einige wenige Beispiele beweisen:

Satz 1.16 *Es gelten:*

$$\bigwedge_{n\in\mathbb{N}} \quad \bigwedge_{m\in\mathbb{N}} \quad m + n \in \mathbb{N}\,, \tag{1.46}$$

$$\bigwedge_{n\in\mathbb{N}} \bigwedge_{m\in\mathbb{N}} m \cdot n \in \mathbb{N} \tag{1.47}$$

[1]Die Annahme der Gültigkeit von $A(n)$ für irgendein $n \in \mathbb{N}$ bezeichnet man als die *Induktionsannahme*

$$\bigwedge_{n\in\mathbb{N}} \quad \bigwedge_{m\in\mathbb{N}} \quad m > n \Rightarrow m - n \in \mathbb{N}\ , \tag{1.48}$$

$$\bigwedge_{n\in\mathbb{N}} \quad (\,n-1\,,\,n\,) \cap \mathbb{N} = \emptyset\ . \tag{1.49}$$

Beweis: Durch Induktionen:

zu (1.46)**:**

$$A(n) :\Leftrightarrow \bigwedge_{m\in\mathbb{N}} \quad m + n \in \mathbb{N}\ .$$

Induktionsverankerung: Da $\mathbb{N}$ induktiv ist, gilt $m + 1 \in \mathbb{N}$ für alle $m \in \mathbb{N}$.
Induktionsschluss: Ist für irgendein $n \in \mathbb{N}$

$$m + n \in \mathbb{N} \quad ,$$

so gilt für $n + 1$:

$$m + (n + 1) = (m + n) + 1 \in \mathbb{N} \quad ;$$

denn $m + n \in \mathbb{N}$ nach Induktionsannahme und $\mathbb{N}$ ist induktiv.

zu (1.47)**:**

$$A(n) :\Leftrightarrow \bigwedge_{m\in\mathbb{N}} m \cdot n \in \mathbb{N}\ .$$

Natürlich gilt $A(1)$, denn $m \cdot 1 = m \in \mathbb{N}$, weil $m \in \mathbb{N}$ ist. Gilt für irgendein $n \in \mathbb{N}$, dass $m \cdot n \in \mathbb{N}$ ist für jedes $m \in \mathbb{N}$, so ist für jedes $m \in \mathbb{N}$ auch

$$m(n + 1) = mn + m \in \mathbb{N}$$

wegen der Induktionsannahme und wegen (1.46).

zu (1.48)**:**

$$A(n) :\Leftrightarrow \bigwedge_{m\in\mathbb{N}} \quad (m > n \Rightarrow m - n \in \mathbb{N})$$

Induktionsverankerung: Wäre $A(1)$ falsch, so gäbe es $m \in \mathbb{N}$ mit $m > 1$ und $m - 1 \notin \mathbb{N}$. Sei dann $M := \mathbb{N} \setminus \{m\}$. Offenbar ist $1 \in M$. Ist ferner $k \in M$, so ist $k + 1$ zunächst einmal aus $\mathbb{N}$ aber insbesondere auch von m verschieden, denn wäre $k + 1 = m$, so hätte man $k = m - 1 \in \mathbb{N}$. M ist also induktiv, und das liefert einen Widerspruch.

Induktionsschluss: Gilt nun für irgendein $n \in \mathbb{N}$ und jedes $m \in \mathbb{N}$ mit $m > n$, dass $m - n \in \mathbb{N}$ ist, so gilt für jedes $m > n + 1$:

$$m - (n + 1) = (m - 1) - n \in \mathbb{N} \quad ;$$

denn $m - 1 \in \mathbb{N}$ nach Induktionsverankerung, $m - 1 > n$, und damit ist nach Induktionsannahme $(m - 1) - n \in \mathbb{N}$.

zu (1.49)**:**

$$A(n) :\Leftrightarrow \mathbb{N} \cap (\, n - 1 \,,\, n \,) = \emptyset \,.$$

Das ist richtig für $n = 1$, denn $\mathbb{N} \subset [1, \infty)$ und damit ist $\mathbb{N} \cap (0, 1) \subset [1, \infty) \cap (0, 1) = \emptyset$. Wenn es für irgendein $n \in \mathbb{N}$ richtig ist, so gilt $(n, n+1) \cap \mathbb{N} = \emptyset$; wäre nämlich $p \in (n, n + 1) \cap \mathbb{N}$, so wäre $p > 1$ und $p - 1 \in \mathbb{N}$ nach (1.48). Also wäre $p - 1 \in (n - 1, n) \cap \mathbb{N}$, ein Widerspruch zur Induktionsannahme.

q.e.d.

Eng verwandt mit dem Beweisprinzip der vollständigen Induktion ist das Prinzip von der rekursiven Definition. Hiermit kann man Folgen definieren. Dabei ist eine Folge eine Funktion[1] von $\mathbb{N}$ in irgendeine nicht–leere Menge S. Jeder natürlichen Zahl wird also ein Element von S in eindeutiger Weise zugeordnet. Man spricht genauer von einer Folge in S.

Beispiele für Folgen sind etwa

$$f : \mathbb{N} \to \mathbb{R} \quad , \quad n \mapsto \frac{n^2 + 7n + 8}{n^5 + 6} \quad , \tag{1.50}$$

meist in der Form

$$\left(\frac{n^2 + 7n + 8}{n^5 + 6} \right)_{n \in \mathbb{N}} \tag{1.51}$$

geschrieben. Hier müssen Sie präzise sein: Mit (1.50) oder (1.51) wird das gesamte mathematische Objekt *Folge* beschrieben, während $\frac{n^2+7n+8}{n^5+6}$ ein Term ist, der bei Einsetzung einer Zahl n eine andere Zahl als Ergebnis “ausspuckt”. Ist k fixiert, so ist $\frac{k^2+7k+8}{k^5+6}$ der Funktionswert von f bzw. von $(\frac{n^2+7n+8}{n^5+6})_{n \in \mathbb{N}}$ an der Stelle k.

Nun tritt hier so etwas wie k^5 auf. Was ist das? Sei $a \in \mathbb{R}$; dann definieren wir

$$a^1 := a \,.$$

[1] In diesem Kapitel wird bereits an einigen Stellen der Begriff einer Funktion (synonym: Abbildung) verwendet. Dieser Begriff wurde in der *Linearen Algebra* eingeführt. Auch findet man eine Einführung im folgenden Kapitel

Wenn für irgendein $n \in \mathbb{N}$ festgelegt ist, was a^n bedeutet, so definiert man

$$a^{n+1} := a \cdot a^n \ .$$

Jetzt weiß man, was a^5 ist:

$$\begin{aligned}
a^1 &= a \ , \\
a^2 &= a \cdot a^1 = a \cdot a \ , \\
a^3 &= a \cdot a^2 = a \cdot a \cdot a \ , \\
a^4 &= a \cdot a^3 = a \cdot a \cdot a \cdot a \ , \\
a^5 &= a \cdot a^4 = a \cdot a \cdot a \cdot a \cdot a \ .
\end{aligned}$$

Ohne Beweis soll von Ihnen das Prinzip der rekursiven Definition geglaubt werden. Konkret: Mit obigem Verfahren wird eine Folge $(a^n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ erklärt. Die abstrakte Formulierung lassen wir teilweise unbewiesen:

Satz und Definition 1.17 *Sei S eine nicht–leere Menge und $\hat{s} \in S$ vorgegeben. Zu jedem $n \in \mathbb{N}$ sei $\varphi_n : S \longrightarrow S$ eine Funktion. $(\varphi_n)_{n\in\mathbb{N}}$ ist also eine Folge in $\mathcal{F}(S, S)$, der Menge der Funktionen von S in sich. Dann existiert genau eine Folge*

$$\Phi : \mathbb{N} \longrightarrow S$$

in S mit den Eigenschaften

$$\Phi(1) = \hat{s} \tag{1.52}$$

$$\bigwedge_{n\in\mathbb{N}} \Phi(n+1) = \varphi_n(\Phi(n)) \tag{1.53}$$

Beweis (der Eindeutigkeit): Ist sowohl Φ als auch Ψ eine solche Folge, gilt also (1.52), (1.53) auch mit Ψ statt Φ, so ist

$$\Phi(1) = \Psi(1) = \hat{s} \ ,$$

und ist für irgendeinein $n \in \mathbb{N}$

$$\Phi(n) = \Psi(n) \quad ,$$

so gilt

$$\Phi(n+1) = \varphi_n(\Phi(n)) = \varphi_n(\Psi(n)) = \Psi(n+1) \ .$$

Damit ist $\Phi(n) = \Psi(n)$ für alle $n \in \mathbb{N}$, und Φ und Ψ sind gleiche Folgen.

q.e.d.

Beispiel 1.18 *Seien* $S = \mathbb{R}$, $\hat{s} = 1$ *und*

$$\varphi_n \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ x \longmapsto x + \frac{1}{n+1}\ .$$

Definiert man Φ *gemäß* 1.17, *so erhält man*

$$\begin{aligned}
&\Phi(1) = 1;\\
&\Phi(2) = \varphi_1(\Phi(1)) = 1 + \frac{1}{2};\\
&\Phi(3) = \varphi_2(\Phi(2)) = 1 + \frac{1}{2} + \frac{1}{3};\\
&\Phi(4) = \varphi_3(\Phi(3)) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4};\\
&\Phi(5) = \varphi_4(\Phi(4)) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}\ .
\end{aligned}$$

Die folgenden Definitionen liefern weitere Beispiele.

Definition 1.19 *Für* $a \in \mathbb{R}$ *seien*

$$a^1 := a \quad \text{und} \quad \bigwedge_{n\in\mathbb{N}} \quad a^{n+1} := a \cdot a^n\ , \tag{1.54}$$

$$a^0 := 1\ , \tag{1.55}$$

$$a^{-n} := \frac{1}{a^n}\ , \quad \text{falls} \quad n \in \mathbb{N}\,,\ a \neq 0\ . \tag{1.56}$$

In Definition 1.18 ist $S = \mathbb{R}$, $\hat{s} = a$ und $\varphi_n \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ t \longmapsto at$. Natürlich hätte man auch die Definition von a^n für $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ in der Form

$$a^0 := 1 \quad \text{und} \quad \bigwedge_{n\in\mathbb{N}} \quad a^n = a \cdot a^{n-1}$$

aussprechen können.

Machen Sie sich an dieser Stelle klar, dass die üblichen Potenzgesetze gelten:

$$\bigwedge_{p\in\mathbb{Z}} \quad \bigwedge_{a,b\in\mathbb{R}^+} \quad (ab)^p = a^p b^p\ , \tag{1.57}$$

$$\bigwedge_{p,q\in\mathbb{Z}} \quad \bigwedge_{a\in\mathbb{R}^+} a^p a^q = a^{p+q}\ , \tag{1.58}$$

$$\bigwedge_{p,q\in\mathbb{Z}} \quad \bigwedge_{a\in\mathbb{R}^+} (a^p)^q = a^{pq}\ . \tag{1.59}$$

Hat man es nur mit nicht negativen ganzzahligen Exponenten zu tun, kann jeweils unter dem zweiten Quantor $\mathbb{R}^+$ durch $\mathbb{R}$ ersetzt werden.

Definition 1.20 *$(a_n)_{j\in\mathbb{N}}$ sei eine Folge in $\mathbb{R}$. Die Folgen*

$$\left(\sum_{j=1}^{n} a_j\right)_{n\in\mathbb{N}} \quad \text{bzw.} \quad \left(\prod_{j=1}^{n} a_j\right)_{n\in\mathbb{N}}$$

heißen die Reihe bzw. das Produkt über (die Folge) $(a_j)_{j\in\mathbb{N}}$, wobei

$$\sum_{j=1}^{1} a_j := a_1 \;\wedge\; \bigwedge_{n\in\mathbb{N}} \sum_{j=1}^{n+1} a_j := \left(\sum_{j=1}^{n} a_j\right) + a_{n+1}\,, \tag{1.60}$$

$$\prod_{j=1}^{1} a_j := a_1 \;\wedge\; \bigwedge_{n\in\mathbb{N}} \prod_{j=1}^{n+1} a_j = \left(\prod_{j=1}^{n} a_j\right) \cdot a_{n+1}\,. \tag{1.61}$$

Ist eine weitere Zahl a_0 festgelegt, so schreibt man mit $n \in \mathbb{N}_0$

$$\sum_{j=0}^{n} a_j := \sum_{j=1}^{n+1} a_{j-1}\,, \quad \prod_{j=0}^{n} a_j := \prod_{j=1}^{n+1} a_{j-1} \tag{1.62}$$

sowie für $n\,,\, m \in \mathbb{N}_0$

$$\sum_{j=m}^{n} a_j := \begin{cases} \sum_{i=0}^{n-m} a_{m+i} & \textit{falls } n \geq m \\ 0 & \textit{falls } n < m \end{cases}\,, \tag{1.63}$$

$$\prod_{j=m}^{n} a_j := \begin{cases} \prod_{i=0}^{n-m} a_{m+i} & \textit{falls } n \geq m \\ 1 & \textit{falls } n < m \end{cases}\,. \tag{1.64}$$

Den Index j nennt man *den Laufindex*. Sie können ihn durch (fast) jedes andere Symbol ersetzen. Ist $(a_j)_{j\in\mathbb{N}}$ gegeben, so ist $\sum_{j=m}^{n} a_j$ eine Funktion von m und n, und statt j dürfen Sie alles außer n und m schreiben.

Die in Beispiel 1.18 beschriebene Folge Φ läßt sich also in der Gestalt

$$\bigwedge_{n\in\mathbb{N}} \quad \Phi(n) = \sum_{j=1}^{n} \frac{1}{j} = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$$

schreiben, wobei der letzte Term nur verdeutlichen soll, was $\sum_{j=1}^{n} \frac{1}{j}$ bedeutet.

Einige Regeln für den Umgang mit Summen und Produkten sollten Sie beherrschen:

Satz 1.21 *Seien* $(a_j)_{j\in\mathbb{N}}, (b_j)_{j\in\mathbb{N}}$. *Folgen in* $\mathbb{R}$ *sowie* $\alpha \in \mathbb{R}$. *Dann gelten für alle* $n\,,\, m \in \mathbb{N}$*:*

$$\sum_{j=1}^{n}(\alpha a_j) = \alpha \cdot \sum_{j=1}^{n} a_j \quad ; \quad \prod_{j=1}^{n}(\alpha a_j) = \alpha^n \prod_{j=1}^{n} a_j \tag{1.65}$$

$$\sum_{j=1}^{n} a_j + \sum_{j=1}^{n} b_j = \sum_{j=1}^{n}(a_j + b_j) \quad ; \quad \left(\prod_{j=1}^{n} a_j\right) \cdot \left(\prod_{j=1}^{n} b_j\right) = \prod_{j=1}^{n}(a_j b_j) \tag{1.66}$$

$$\sum_{j=1}^{m+n} a_j = \sum_{j=1}^{m} a_j + \sum_{j=m+1}^{n+m} a_j \quad ; \quad \prod_{j=1}^{m+n} a_j = \left(\prod_{j=1}^{m} a_j\right) \cdot \left(\prod_{j=m+1}^{n+m} a_j\right) \tag{1.67}$$

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} a_{\pi(j)} \quad ; \quad \prod_{j=1}^{n} a_j = \prod_{j=1}^{n} a_{\pi(j)} \tag{1.68}$$

sofern π *eine bijektive Abbildung von* $\{1, 2, \ldots, n\}$ *auf sich bezeichnet (also eine Permutation der Zahlen* $1, \ldots, n$*)*

Beweis: (1.65) – (1.67) beweist man leicht mit vollständiger Induktion bezüglich n. Der Beweis von(1.68) soll für die Summe durchgeführt werden.

Ist $n = 1$, so ist nichts zu sagen. Wir nehmen an, dass (1.68) für irgendein $n \in \mathbb{N}$ richtig ist. Sei π eine Permutation von $\{1, \ldots, n+1\}$ mit $n + 1 = \pi(k)$.

Dann ist

$$\tau \;:\; \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}\,,\; j \longmapsto \begin{cases} \pi(j) & \text{falls } j < k \\ \pi(j+1) & \text{falls } j \geq k \end{cases}$$

eine Permutation von $\{1, \ldots, n\}$. Es gilt (für Begründungen siehe die Fußnoten)

$$\begin{aligned} \sum_{j=1}^{n+1} a_{\pi(j)} &\overset{1}{=} \sum_{j=1}^{k} a_{\pi(j)} + \sum_{j=k+1}^{n+1} a_{\pi(j)} \overset{2}{=} \left(\sum_{j=1}^{k-1} a_{\pi(j)} + a_{n+1}\right) + \sum_{j=k}^{n} a_{\pi(j+1)} \\ &\overset{3}{=} \left(\sum_{j=1}^{k-1} a_{\tau(j)} + \sum_{j=k}^{n} a_{\tau(j)}\right) + a_{n+1} \overset{1}{=} \sum_{j=1}^{n} a_{\tau(j)} + a_{n+1} \overset{4}{=} \sum_{j=1}^{n} a_j + a_{n+1} \overset{5}{=} \sum_{j=1}^{n+1} a_j \quad . \end{aligned}$$

q.e.d.

[1]nach (1.67)
[2]nach (1.60) oder im Fall $k = 1$ nach (1.63)
[3]nach Assoziativ- und Kommutativgesetz (Def. 1.1) sowie der Definition von τ
[4]nach Induktionsannahme
[5]nach (1.60)

Die Regeln werden analog zu Def. 1.1 ebenfalls Assoziativ– bzw. Kommutativgesetz genannt.

Mit den oben eingeführten Begriffen läßt sich mit Induktion die allgemeine binomische Formel beweisen. Hierzu definieren wir mit Hilfe der Fakultäten

$$k! := \prod_{j=1}^{k} j \quad \text{für } k \in \mathbb{N}_0 \quad \text{(lies: “k Fakultät”)} \tag{1.69}$$

die Binomialkoeffizienten:

$$\begin{pmatrix} n \\ j \end{pmatrix} := \frac{\prod\limits_{i=0}^{j-1}(n-i)}{\prod\limits_{i=1}^{j} i} = \frac{n!}{j!(n-j)!} \quad \text{für } j, n \in \mathbb{N}_0,\ j \leq n. \tag{1.70}$$

Insbesondere gelten also

$$0! = 1\ , \tag{1.71}$$

$$\bigwedge_{n\in\mathbb{N}} \begin{pmatrix} n \\ n \end{pmatrix} = \begin{pmatrix} n \\ 0 \end{pmatrix} = 1\ , \tag{1.72}$$

$$\bigwedge_{n\in\mathbb{N}} \quad \bigwedge_{j\in\mathbb{N}_0,\, j\leq n} \begin{pmatrix} n \\ j \end{pmatrix} = \begin{pmatrix} n \\ n-j \end{pmatrix}\ . \tag{1.73}$$

Schließlich ist

$$\bigwedge_{n\in\mathbb{N}} \quad \bigwedge_{j\in\mathbb{N},\, j\leq n} \begin{pmatrix} n \\ j \end{pmatrix} + \begin{pmatrix} n \\ j-1 \end{pmatrix} = \begin{pmatrix} n+1 \\ j \end{pmatrix}\ ; \tag{1.74}$$

denn

$$\begin{aligned} \begin{pmatrix} n \\ j \end{pmatrix} + \begin{pmatrix} n \\ j-1 \end{pmatrix} &= \frac{n!}{j!(n-j)!} + \frac{n!}{(j-1)!(n-j+1)!} \\ &= \frac{n![(n-j+1)+j]}{j!(n-j+1)!} = \frac{(n+1)!}{j!((n+1)-j)!} = \begin{pmatrix} n+1 \\ j \end{pmatrix} \end{aligned}$$

Satz 1.22 *Für alle $n \in \mathbb{N}$ gilt*

$$\bigwedge_{x,y\in\mathbb{R}} \quad (x+y)^n = \sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} x^j y^{n-j} \tag{1.75}$$

Beweis (durch Induktion über n): Wegen (1.72) gilt (1.75) für $n = 1$:

$$\bigwedge_{x,y\in\mathbb{R}} \quad (x+y)^1 = y + x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} x^0 y^1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} x^1 y^0 = \sum_{j=0}^{1} \begin{pmatrix} 1 \\ j \end{pmatrix} x^j y^{1-j} \ .$$

Gilt (1.75) für irgendein $n \in \mathbb{N}$, so erhält man für alle x , $y \in \mathbb{R}$ (Begründungen findet man in den Fußnoten):

$$\begin{aligned}
&(x+y)^{n+1} \\
&\quad = (x+y)(x+y)^n = (x+y) \cdot \sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} x^j y^{n-j} \\
&\quad = \sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} x^{j+1} y^{n-j} + \sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} x^j y^{n+1-j} \\
&\quad \overset{1}{=} \quad \sum_{j=0}^{n-1} \begin{pmatrix} n \\ j \end{pmatrix} x^{j+1} y^{(n+1)-(j+1)} + \begin{pmatrix} n \\ n \end{pmatrix} x^{n+1} y^0 \\
&\qquad\qquad\qquad + \sum_{i=1}^{n} \begin{pmatrix} n \\ i \end{pmatrix} x^i y^{(n+1)-i} + \begin{pmatrix} n \\ 0 \end{pmatrix} x^0 y^{n+1} \\
&\quad \overset{2}{=} \quad x^0 y^{n+1} + \sum_{i=1}^{n} \left[\begin{pmatrix} n \\ i-1 \end{pmatrix} + \begin{pmatrix} n \\ i \end{pmatrix} \right] x^i y^{(n+1)-i} + x^{n+1} y^0 \\
&\quad \overset{3}{=} \quad \sum_{i=0}^{n+1} \begin{pmatrix} n+1 \\ i \end{pmatrix} x^i y^{(n+1)-i}
\end{aligned}$$

und das ist (1.75) für $n + 1$.

q.e.d.

In Zukunft benötigen wir auch die geometrische Summenformel

Satz 1.23 *Für alle $n \in \mathbb{N}$ gelten*

$$\bigwedge_{q\in\mathbb{R}} \quad (1-q) \sum_{j=0}^{n} q^j = 1 - q^{n+1}, \tag{1.76}$$

$$\bigwedge_{q\in\mathbb{R}\setminus\{1\}} \quad \sum_{j=0}^{n} q^j = \frac{1-q^{n+1}}{1-q} \tag{1.77}$$

---

[1] nach (1.67)
[2] nach (1.72) und (1.66)
[3] nach (1.72), (1.67) und (1.74)

Beweis: Division beider Seiten von (1.76) durch $1 - q \neq 0$ liefert (1.77).

Man erhält (1.76) aus

$$\begin{aligned}(1-q)\sum_{j=0}^{n} q^j &= \sum_{j=0}^{n} q^j - q\sum_{j=0}^{n} q^j = \sum_{j=0}^{n} q^j - \sum_{j=0}^{n} q^{j+1} \\ &= \sum_{j=0}^{n} q^j - \sum_{j=1}^{n+1} q^j = q^0 + \sum_{j=1}^{n} q^j - \sum_{j=1}^{n} q^j - q^{n+1} = 1 - q^{n+1}\end{aligned}$$

q.e.d.

Eine andere Variante des vollständigen Induktionsprinzips ist der Wohlordnungssatz. Hierzu benötigen wird den Begriff des Minimums.

Definition 1.24 *Ist $T \subset \mathbb{R}$ eine nicht–leere Teilmenge und erfüllt $\tau \in T$ die Ungleichung*

$$\bigwedge_{t\in T} \quad \tau \leq t \quad (bzw. \bigwedge_{t\in T} \quad \tau \geq t),$$

so nennt man τ das Minimum (bzw. das Maximum) von T. Wenn in T ein Minimum (bzw. Maximum) existiert, sagt man: T besitzt ein Minimum (bzw. Maximum).

Zum Beispiel besitzt $(\,0\,,\,1\,]$ das Maximum 1, aber kein Minimum. Man beachte, dass T höchstens ein Minimum und höchstens ein Maximim besitzt. Man schreibt dafür $\min T$, bzw. $\max T$.

Satz 1.25 (Wohlordnungssatz)
Jede nicht–leere Teilmenge T von $\mathbb{N}$ besitzt ein Minimum.

Beweis: (indirekt) Annahme: T besitzt kein Minimum. Sei

$$U := \{n \in \mathbb{N}\colon \bigwedge_{t\in T} n \leq t\} \quad .$$

Dann ist $U \cap T = \emptyset$, denn ein Element von $U \cap T$ wäre gerade das Minimum von T.

Offenbar ist $1 \in U$, denn $1 \leq t$ für jede natürliche Zahl t. Liegt nun irgendein $n \in \mathbb{N}$ in U, so gilt $n < t$ für alle $t \in T$, da $U \cap T = \emptyset$ ist. Dann ist auch $n+1 \in U$:

anderenfalls gäbe es $t \in T$ mit $n < t < n+1$, ein Widerspruch zu (1.49). Nach dem Induktionsprinzip ist also $U = \mathbb{N}$ und damit $T = \mathbb{N} \cap T = U \cap T = \emptyset$ im Widerspruch zur Voraussetzung.

q.e.d.

Definition 1.26 *Für $n \in \mathbb{N}$ sei $\mathbb{N}(n) := \{k \in \mathbb{N}\colon k \leq n\} = \mathbb{N} \cap [1\,,\,n]$. Eine Menge M heißt endlich, wenn $M = \emptyset$ ist oder wenn es ein $n \in \mathbb{N}$ und eine bijektive Abbildung von M auf $\mathbb{N}(n)$ gibt. Die Zahl n ist dann eindeutig bestimmt und heißt Mächtigkeit oder Kardinalzahl von M. Wir schreiben dafür $\#M$ und legen noch $\#\emptyset := 0$ fest.*

Dass die Kardinalzahl einer endlichen Menge tatsächlich eindeutig definiert ist, soll hier nicht bewiesen werden. Man benötigt dazu einen Satz des Inhalts, dass für $n, m \in \mathbb{N}$ dann und nur dann eine bijektive Abbildung von $\mathbb{N}(n)$ auf $\mathbb{N}(m)$ existiert, wenn $n = m$ ist.

Ebenfalls ohne Beweis halten wir fest:

Satz 1.27 *Ist $T \subset \mathbb{R}$ eine nicht–leere endliche Menge, so besitzt T Maximum und Minimum.*

Versuchen Sie selbst einen Beweis, z. B. durch Induktion über $\#T$. Hierzu formuliere man Satz 1.27 in der Form

$$\bigwedge_{n \in \mathbb{N}} \quad (\#T = n \Rightarrow T \text{ besitzt Maximum und Minimum}).$$

Man kann nun die Menge $\mathbb{Z}$ der ganzen Zahlen und die Menge $\mathbb{Q}$ der rationalen Zahlen definieren:

Definition 1.28

$$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{p \in \mathbb{R} : -p \in \mathbb{N}\} = \mathbb{N} \cup \{0\} \cup \{-n : n \in \mathbb{N}\}$$

heißt Menge der ganzen Zahlen.

$$\mathbb{Q} := \{\alpha \in \mathbb{R}\colon \bigvee_{p \in \mathbb{Z}} \bigvee_{q \in \mathbb{N}} \alpha = \frac{p}{q}\}$$

heißt Menge der rationalen Zahlen.

Wir haben zu Beginn dieses Unterabschnitts betont, dass $\mathbb{R}$ irgendein geordneter Körper sein kann. So lassen sich ganze und rationale Zahlen in jedem geordneten Körper definieren und sind charakterisiert durch die Eigenschaften: $\mathbb{Z}$ ist der kleinste Ring in $\mathbb{R}$, der $\mathbb{N}$ enthält; $\mathbb{Q}$ ist der kleinste Körper in $\mathbb{R}$, der $\mathbb{N}$ enthält. Dabei ist die mathematische Struktur eines Rings (genauer: eines kommutativen Rings mit 1) definiert wie ein Körper (siehe Definition 1.1), jedoch ohne die Forderung nach der Existenz inverser Elemente bezüglich der Multiplikation. Spricht man übrigens von einem Ring bzw. Körper in $\mathbb{R}$, so meint man, dass Addition und Multiplikation durch die entsprechenden Operationen in $\mathbb{R}$ erklärt sind. $\mathbb{Q}$ ist übrigens selbst ein geordneter Körper.

1.4 Die reellen Zahlen

Wie eben gesagt, bildet die Menge der rationalen Zahlen einen geordneten Körper. Aber wie man schon seit Pythagoras weiß, entsprechen die rationalen Zahlen nicht allen Punkten auf einer Geraden. Zum Beispiel lässt sich die Länge der Diagonalen eines Quadrats der Seitenlänge 1 nicht als rationale Zahl deuten. Der klassische Beweis gehört zur Allgemeinbildung nicht nur eines Mathematikers:

Satz 1.29 *Es gibt keine rationale Zahl, deren Quadrat* 2 *ist:*

$$\bigwedge_{\alpha\in\mathbb{Q}} \alpha^2 \neq 2\ .$$

Dem Beweis schicken wir einen Hilfsatz voraus

Lemma 1.30 *Mit* $\mathbb{G} := \{2n\colon n \in \mathbb{N}\}, \mathbb{U} := \{2n-1\colon n \in \mathbb{N}\}$ *gelten*

(i) $\mathbb{N} = \mathbb{U} \cup \mathbb{G}$

(ii) $\mathbb{U} \cap \mathbb{G} = \emptyset$

(iii) $\bigwedge_{n\in\mathbb{U}} n^2 \in \mathbb{U}$

(iv) $\bigwedge_{n\in\mathbb{G}} n^2 \in \mathbb{G}$

(i) und (ii) sagen aus, dass jede natürliche Zahl entweder gerade oder ungerade ist. Versuchen Sie selbst, das zu beweisen. Ist $n \in \mathbb{U}$ also ungerade, so ist mit einem $m \in \mathbb{N}$

$$n = 2m - 1,$$

also

$$n^2 = 4m^2 - 4m + 1 = 2(2m^2 - 2m + 1) - 1 \ .$$

Da $2m^2 - 2m + 1$ eine natürliche Zahl ist, ist also n^2 ebenfalls ungerade. Analog zeigt man (iv).

Beweis (von Satz 1.29): Wir führen die Annahme, es gäbe eine rationale Zahl, etwa α, für die $\alpha^2 = 2$ ist, zum Widerspruch. Dann können wir α als positiv annehmen und in der Form $\alpha = \frac{m}{n}$ mit $m, n \in \mathbb{N}$ schreiben.

Sei

$$p = \min\{n \in \mathbb{N}\colon \bigvee_{m \in \mathbb{N}} \alpha = \frac{m}{n}\} = \min\{n \in \mathbb{N}\colon \alpha n \in \mathbb{N}\},$$

und setze

$$q := \alpha p \in \mathbb{N} \ .$$

Dann ist

$$q^2 = \alpha^2 p^2 = 2p^2 \in \mathbb{G}$$

Nach Lemma 1.30 ist auch $q \in \mathbb{G}$. Es gibt also ein $s \in \mathbb{N}$ mit $q = 2s$. Setzt man dies in obige Gleichung ein, erhält man

$$4s^2 = 2p^2,$$

also

$$p^2 = 2s^2 \ .$$

Dann ist aber auch p^2 und damit p gerade, $p = 2t$ mit $t \in \mathbb{N}$, und es folgt

$$\alpha = \frac{q}{p} = \frac{2s}{2t} = \frac{s}{t} \ .$$

Da $t < p$ ist, widerspricht dies der Festlegung von p als Minimum der möglichen Nenner von α.

q.e.d.

Zur Definition der (Ordnungs)vollständigkeit benötigt man den Begriff der oberen bzw. unteren Schranke und der kleinsten oberen bzw. größten unteren Schranke (des Infimums und des Supremums).

Definition 1.31 *Ist T Teilmenge eines geordneten Körpers K, so heißt*

$$\begin{aligned} T \text{ nach oben beschränkt} &:\Leftrightarrow \bigvee_{s\in\mathbb{R}} \bigwedge_{t\in T} t \leq s \,, \\ T \text{ nach unten beschränkt} &:\Leftrightarrow \bigvee_{s\in\mathbb{R}} \bigwedge_{t\in T} t \geq s \,, \\ T \text{ beschränkt} &:\Leftrightarrow T \text{ ist nach oben und nach unten beschränkt} \\ &\Leftrightarrow \bigvee_{s\in\mathbb{R}} \bigwedge_{t\in T} |\, t \,| \leq s \,. \end{aligned}$$

Jede Zahl $s \in \mathbb{R}$, für die gilt

$$\bigwedge_{t\in T} t \geq s \quad (\text{bzw. } t \leq s),$$

heißt untere bzw. obere Schranke von T. Das Infimum bzw. Supremum von T ist dann die größte untere bzw. kleinste obere Schranke von T:

$$\inf\ T := \max\{s \in \mathbb{R}\colon s \text{ ist untere Schranke von } T\} \,,$$

$$\sup\ T := \min\{s \in \mathbb{R}\colon s \text{ ist obere Schranke von } T\} \,.$$

Infimum oder Supremum einer Menge sind somit nur definiert, wenn die Menge der unteren bzw. oberen Schranken tatsächlich ein Maximum bzw. Minimum besitzt. Wenn T selbst ein Minimum bzw. Maximum besitzt, so ist dieses auch das Infimum bzw. Supremum. Infimum und Supremum einer Menge T sind charakterisiert durch

$$a = \inf T \iff \left(\bigwedge_{t\in T} a \leq t\right) \wedge \left(\bigwedge_{\varepsilon>0} \bigvee_{t\in T} t < a + \varepsilon\right) \,, \tag{1.78}$$

$$b = \sup T \iff \left(\bigwedge_{t\in T} b \geq t\right) \wedge \left(\bigwedge_{\varepsilon>0} \bigvee_{t\in T} t > b - \varepsilon\right) \,. \tag{1.79}$$

Ist $\mathbb{K} = \mathbb{Q}$, der geordnete Körper der rationalen Zahlen, so gelten:

$\inf\ (0\,,\,1] = 0$,
$\min\ (0\,,\,1]$ existiert nicht,
$\sup\ (0\,,\,1] = 1$,
$\max\ (0\,,\,1] = 1$,
$\inf\ \{\alpha \in \mathbb{Q}\colon \alpha > 0, \alpha^2 \geq 2\}$ existiert nicht.

Der Körper $\mathbb{R}$ der reellen Zahlen ist durch die Eigenschaft, (ordnungs)-vollständig zu sein, ausgezeichnet.

Definition 1.32 *Ein geordneter Körper $\mathbb{K}$ heißt* (ordnungs)-vollständig*, wenn jede nach oben beschränkte Teilmenge $T \subset \mathbb{K}$ ein Supremum besitzt.*

Dann besitzt auch jede nach unten beschränkte Menge ein Infimum.

Damit sind die Begriffe in unserem Definitionsprogramm für die reellen Zahlen geklärt.

Definition 1.33 *Einen geordneten, (ordnungs)-vollständigen Körper nennt man ein Modell für den Körper $\mathbb{R}$ der rellen Zahlen.*

Man kann beweisen, dass je zwei vollständige geordnete Körper isomorph sind: Sind $\mathbb{K}_1$ und $\mathbb{K}_2$ ordnungsvollständige geordnete Körper, so gibt es (sogar genau eine) eine bijektive Abbildung

$$\varphi : \mathbb{K}_1 \to \mathbb{K}_2$$

mit den Eigenschaften

$$\bigwedge_{r,s\in\mathbb{K}_1} \begin{cases} \varphi(r+s) = \varphi(r) + \varphi(s) \wedge \varphi(r \cdot s) = \varphi(r) \cdot \varphi(s) & , \\ r < s \Rightarrow \varphi(r) < \varphi(s) & . \end{cases}$$

φ erhält also die algebraische und die Ordnungsstruktur. Will man diesen Beweis wirklich durchführen, benötigt man die folgende *Archimedische Eigenschaft*:

Satz und Definition 1.34

(i) *$\mathbb{N}$ ist unbeschränkt. (Archimedische Eigenschaft von $\mathbb{N}$)*

(ii) *Jedes Intervall $I \subset \mathbb{R}$ positiver Länge enthält rationale Zahlen*

(iii) *$\mathbb{Q}$ ist dicht in $\mathbb{R}$, d. h. zu jedem $x \in \mathbb{R}$ und $\varepsilon > 0$ gibt es ein $q \in \mathbb{Q} \cap U_\varepsilon(x)$.*

(iv) *Jedes $r \in \mathbb{R}$ ist Supremum einer Menge rationaler Zahlen, genauer:*

$$\bigwedge_{r\in\mathbb{R}} \quad r = \sup\left((-\infty\, ,\, r\,) \cap \mathbb{Q}\right)$$

Beweis: (i) Da $\mathbb{N}$ nach unten durch 1 beschränkt ist, muss gezeigt werden, dass $\mathbb{N}$ nicht nach oben beschränkt ist. Wäre $\mathbb{N}$ nach oben beschränkt, dann gäbe es wegen der Vollständigkeit eine reelle Zahl r mit

$$r = \sup \mathbb{N}\ .$$

Die Zahl $s := r - 1/2$ kann dann nicht obere Schranke von $\mathbb{N}$ sein, d. h. es gibt ein $n \in \mathbb{N}$ mit

$$n > r - 1/2 \, .$$

Aber dann ist $n+1 > r-1/2+1 = r+1/2 > r$, d. h. r war keine obere Schranke von $\mathbb{N}$.

(ii) Seien $a, b \in \mathbb{R}$ mit $a < b$ und $a = \inf I, b = \sup I$. Bestimme $k \in \mathbb{N}$ mit $k \geq -a + 1$ und definiere

$$\hat{a} := a + k \quad , \quad \hat{b} = b + k \, .$$

Dann ist $\hat{b} > \hat{a} \geq 1$.

Wir zeigen, dass $(\, \hat{a} \, , \, \hat{b} \,)$ eine rationale Zahl α enthält. Dann ist $\beta := \alpha - k \in \mathbb{Q} \cap I$.

O. B. d. A. kann angenommen werden, dass $(\, \hat{a} \, , \, \hat{b} \,)$ keine natürliche Zahl enthält. Sonst wäre diese das gesuchte α. Man bestimme nun $n \in \mathbb{N}$ mit $(\, \hat{a} \, , \, \hat{b} \,) \subset (\, n \, , \, n+1 \,)$ sowie ein $q \in \mathbb{N}$ mit

$$\frac{1}{q} < \hat{b} - \hat{a} \, .$$

gemäß (i). Bestimme ferner $j \in \mathbb{N}$ als die kleinste Zahl, für die

$$n + \frac{j}{q} > \hat{a}$$

ist: $j = \min\{i \in \mathbb{N} \colon i > q(\hat{a} - n)\}$. Dann ist

$$\alpha := n + \frac{j}{q} \in (\, \hat{a} \, , \, \hat{b} \,) \, ,$$

denn

$$\hat{a} < \alpha = n + \frac{j-1}{q} + \frac{1}{q} \leq \hat{a} + \frac{1}{q} < \hat{a} + (\hat{b} - \hat{a}) = \hat{b} \, .$$

Natürlich ist $\alpha \in \mathbb{Q}$.

(iii) folgt direkt aus (ii).

(iv) r ist sicher obere Schranke von $M := \mathbb{Q} \cap (-\infty, r)$. Ist $s < r$, so existiert $q \in \mathbb{Q}$ mit $q \in (s, r)$. Dann ist aber $q \in M$ und daher s nicht obere Schranke von M. Es gibt also keine kleinere obere Schranke von M als r.

q.e.d.

In der Menge der reellen Zahlen existieren Wurzeln aus positiven Zahlen. Zunächst formulieren wir aber als Lemma:

Lemma 1.35 *Für alle $n \in \mathbb{N}$ und $x, y \in \mathbb{R}$ mit $0 \leq x < y$ gilt*

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

Das zeigt man leicht mit vollständiger Induktion oder mit der Formel (vgl. Satz 1.23)

$$y^n - x^n = (y-x)\left(\sum_{j=0}^{n-1} x^j y^{n-j-1} > 0\right) \quad .$$

Satz 1.36 *Zu jedem $k \in \mathbb{N}$ und $a \in \mathbb{R}^+$ existiert genau ein $b \in \mathbb{R}^+$ mit*

$$b^k = a \; .$$

Schreibweisen: $\sqrt[k]{a} := a^{1/k} := b$, $\sqrt{a} := \sqrt[2]{a}$.
Man erklärt auch $\sqrt[k]{0} := 0^{1/k} := 0$

Beweis: Wir definieren

$$M(a,k) := \{x \in \mathbb{R}^+ \colon x^k \leq a\} \; .$$

$M(a,k)$ ist nicht leer: ist $a \leq 1$, so ist $a \in M(a,k)$, ist $a > 1$, so ist $1 \in M(a,k)$. Auch ist $M(a,k)$ nach oben beschränkt, etwa durch 1 im Fall $a \leq 1$ oder durch a im Fall $a > 1$. Es existiert mithin

$$b := \sup M(a,k) > 0 \; .$$

Sei nun

$$0 < \varepsilon \leq 1 \; .$$

Dann ist nach der binomischen Formel

$$\begin{aligned}(b+\varepsilon)^k &= \sum_{j=0}^{k} \binom{k}{j} \varepsilon^j b^{k-j} \\ &= b^k + \varepsilon \sum_{j=1}^{k} \binom{k}{j} \varepsilon^{j-1} b^{k-j} \leq b^k + c \cdot \varepsilon\end{aligned}$$

mit

$$c := \sum_{j=1}^{k} \binom{k}{j} b^{k-j} \; ,$$

denn $0 < \varepsilon \leq 1$. Wäre $b^k < a$, so könnte man $\varepsilon \leq \frac{a-b^k}{2c}$ wählen und erhielte

$$(b+\varepsilon)^k \leq b^k + c\left(\frac{a-b^k}{2c}\right) = \frac{1}{2}b^k + \frac{1}{2}a < a\ .$$

b wäre also nicht obere Schranke von $M(a,b)$.

Andererseits liefert die Bernoullische Ungleichung (Satz 1.15) für $\varepsilon \in (\,0\,,\,b\,)$

$$(b-\varepsilon)^k = b^k(1-\frac{\varepsilon}{b})^k \geq b^k(1-\frac{k\varepsilon}{b}) = b^k - kb^{k-1}\varepsilon$$

Wäre $b^k > a$, so hätte man mit $\varepsilon := \frac{1}{kb^{k-1}}(b^k-a)$:

$$(b-\varepsilon)^k \geq b^k - kb^{k-1}\cdot\frac{b^k-a}{kb^{k-1}} = a \quad .$$

Wegen des Lemmas 1.35 wäre für $x \in M(a,k)$

$$x \leq b - \varepsilon\ .$$

Damit wäre b nicht *kleinste* untere Schranke von $M(a,k)$.

Insgesamt muss also

$$b^k = a$$

sein.

Schließlich folgt wieder aus Lemma 1.35, dass keine weitere positive Zahl y der Gleichung $y^k = a$ genügen kann.

q.e.d.

Wichtige Bemerkungen:

(i) Beachten Sie, dass für $a \geq 0$ stets

$$\sqrt{a} \geq 0$$

ist. Daher erhält man

$$\bigwedge_{x\in\mathbb{R}} \quad \sqrt{x^2} = |\ x\ |,$$

und die Betragsstriche darf man auf keinen Fall vergessen.

(ii) Der Ausdruck $a^{1/k}$ bleibt für negatives k undefiniert, auch dann, wenn k ungerade ist. Eine Begründung hierfür geben wir später.

Nun kann man Potenzen mit rationalen Exponenten definieren. Beachten Sie bitte, dass man dies nur für positive Basen tut (allenfalls noch für die Basis 0), da sonst Widersprüche entstehen.

Satz und Definition 1.37 *Seien $a \in \mathbb{R}^+$, $p \in \mathbb{Z}$, $q, r \in \mathbb{Z} \setminus \{0\}$. Dann ist*

$$(a^p)^{1/q} = (a^{rp})^{1/(rq)} . \tag{1.80}$$

Für $\alpha := \frac{p}{q} \in \mathbb{Q}$ sei

$$a^\alpha := (a^p)^{1/q} . \tag{1.81}$$

Formel (1.80) zeigt, dass a^α durch 1.81 wohldefiniert ist. Man mache sich klar, dass 1.80 unsinnig wird, wenn man $a < 0$ zuläßt und etwa $p = 1, q = 3, r = 2$ setzt.

Man definiert auch

$$0^\alpha := 0 \text{ für } \alpha \in \mathbb{Q}, \alpha > 0 . \tag{1.82}$$

Beweis: Mit $x := (a^p)^{1/q}, y := (a^{rp})^{1/(rq)}$ genügt es $x^{rq} = y^{rq}$ zu zeigen wegen des Lemmas 1.35. Nach Satz 1.36 ist

$$y^{rq} = a^{rp} .$$

Ferner liefern Satz 1.36 und (1.59)

$$x^{rq} = (x^q)^r = (a^p)^r = a^{rp} = y^{rq} .$$

q.e.d.

Für die Potenzen und rationalen Exponenten gelten wieder die Potenzgesetze (vgl. (1.57)–(1.59)):

$$\bigwedge_{\alpha \in \mathbb{Q}} \bigwedge_{a,b \in \mathbb{R}^+} (ab)^\alpha = a^\alpha b^\alpha \tag{1.83}$$

$$\bigwedge_{\alpha,\beta \in \mathbb{Q}} \bigwedge_{a \in \mathbb{R}^+} a^{\alpha+\beta} = a^\alpha a^\beta \tag{1.84}$$

$$\bigwedge_{\alpha,\beta \in \mathbb{Q}} \bigwedge_{a \in \mathbb{R}^+} (a^\alpha)^\beta = a^{\alpha\beta} \tag{1.85}$$

Wir wollen noch die erweiterte reelle Achse einführen:

$$\hat{\mathbb{R}} := \mathbb{R} \cup \{-\infty; \infty\} \tag{1.86}$$

Damit werden zum Körper $\mathbb{R}$ noch zwei Elemente hinzugefügt. Rechenregeln hierfür setzen wir später fest. Bezüglich der Ordnung wird

$$\bigwedge_{r \in \mathbb{R}} \quad -\infty < r < \infty \tag{1.87}$$

vereinbart. Es ist üblich Supremum und Infimum einer beliebigen Teilmenge von $\mathbb{R}$ als Element in $\hat{\mathbb{R}}$ festzulegen, indem man für $M \subset \mathbb{R}$ setzt

$$\sup M := \infty \text{ , falls } M \text{ nicht nach oben beschränkt ist.} \tag{1.88}$$
$$\inf M := -\infty \text{ , falls } M \text{ nicht nach unten beschränkt ist.} \tag{1.89}$$
$$\sup \emptyset := -\infty \tag{1.90}$$
$$\inf \emptyset := \infty \tag{1.91}$$

Damit hat jede Teilmenge von $\mathbb{R}$ ein Supremum und ein Infimum in $\hat{\mathbb{R}}$. Ist M nicht leer, so ist

$$\inf M \leq \sup M \ .$$

Ferner gilt für Teilmengen $S\,, T \subset \mathbb{R}$ mit $S \subset T$

$$\inf T \leq \inf S \ , \qquad \sup S \leq \sup T \quad , \tag{1.92}$$

denn jede untere bzw. obere Schranke von T ist auch untere bzw. obere Schranke von S.

Schließlich halten wir noch fest:

Satz 1.38 *Für jedes $n \in \mathbb{N}$ sei $I_n \subset \mathbb{R}$ ein abgeschlossenes, beschränktes und nicht–leeres Intervall. Es gelte*

$$\bigwedge_{n \in \mathbb{N}} \quad I_{n+1} \subset I_n \ . \tag{1.93}$$

Dann ist

$$\bigcap_{n \in \mathbb{N}} \quad I_n \neq \emptyset;$$

es gibt also eine reelle Zahl x, die in allen Intervallen $I_n, n \in \mathbb{N}$, liegt.

Beweis: Sei $I_n = [\, a_n \, , \, b_n \,]$. Dann gilt für alle $n \in \mathbb{N}$

$$a_1 \le a_n \le b_n \le b_1 \ . \tag{1.94}$$

Also sind

$$\underline{M} := \{a_n \colon n \in \mathbb{N}\} \quad , \qquad \overline{M} := \{b_n \colon n \in \mathbb{N}\}$$

beschränkte nicht–leere Mengen, und

$$a := \sup \underline{M} \quad , \qquad b := \inf \overline{M}$$

existieren. Wegen $a_n \in I_m$ für $m \le n$ und $b_m \in I_n$ für $m \ge n$ läßt sich (1.94) verschärfen zu

$$\bigwedge_{n,m \in \mathbb{N}} \quad a_1 \le a_n \le b_m \le b_1 \ . \tag{1.95}$$

Hieraus erhält man

$$\bigwedge_{n \in \mathbb{N}} \quad a \le b_n \ \wedge \ b \ge a_n \ . \tag{1.96}$$

Wegen $b = \inf \overline{M}$ folgt $a \le b$.

Dies fassen wir mit (1.96) zu

$$\bigwedge_{n \in \mathbb{N}} \quad a_n \le a \le b \le b_n$$

zusammen. Folglich ist $a, b \in \bigcap\limits_{n \in \mathbb{N}} I_n$, genauer sogar $[\, a \, , \, b \,] \subset \bigcap\limits_{n \in \mathbb{N}} I_n$, und $[\, a \, , \, b \,]$ ist nicht leer.

q.e.d.

Kapitel 2

Reelle Funktionen und der Konvergenzbegriff

2.1 Funktionen

Eines der wichtigsten Konzepte in der Mathematik ist das Konzept der Funktion. Wie beim Begriff der Menge kann man vielleicht ohne eine formale Definiton auskommen: Sind S und T zwei nicht–leere Mengen, so ist eine Funktion (oder Abbildung) von S nach (in) T eine Zuordnung, vermöge der jedem Element $s \in S$ genau ein Element $t \in T$ zugeordnet ist. Die Funktion f ist also definiert durch Angabe ihres Definitionsbereichs S, ihres Zielbereichs T und einer Vorschrift, wie einem vorgegebenem $s \in S$ ein Element $t \in T$ zuzuordnen ist. Zum Beispiel kann S die Menge aller Telefonnummern sein, die im Ortsnetz von Essen vergeben werden, und T ist die Menge aller Essener Telefonkunden. Dann kann man etwa f als die Funktion definieren, die jeder Telefonnummer den Kunden zuordnet, der diese Telefonnummer besitzt. Da einzelne Kunden auch verschiedene Nummern haben können (ISDN), habe ich die Funktion f als Funktion von S nach T definiert. Umgekehrt hat man keine Funktion. Um das Wirken einer Funktion zu veranschaulichen, notieren wir sie meist in der Form

$$f \ : \ S \longrightarrow T \, , \ s \longmapsto f(s) \, .$$

Dabei ist S der Definitionsbereich von f, T der Zielbereich und $f(s)$ die Vorschrift, wie zu gegebenem $s \in S$ (statt s darf da auch ein anderes Symbol stehen) der *Funktionswert* gebildet wird. Trägt man für s ein spezielles Element der Menge S ein, so bezeichnet $f(s)$ den Funktionswert von f an der Stelle s, das ist dasjenige Element aus der Menge T, welches s zugeordnet ist.

In diesem Semester werden S und T meist Teilmengen von $\mathbb{R}$ sein. In diesem Fall sprechen wir von *reellen Funktionen*.

Beispiele 2.1

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; s \longmapsto s^2 + 8s + 17 \tag{2.1}$$

Dann ist etwa $f(2) = 2^2 + 8 \cdot 2 + 17 = 37$.

$$g \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; p \longmapsto \begin{cases} 1 & , \textit{ falls } \; p \;\; \textit{rational ist} \\ 0 & , \textit{ falls } \; p \;\; \textit{irrational ist} \end{cases} . \tag{2.2}$$

Eine irrationale Zahl ist eine Zahl aus $\mathbb{R}\setminus\mathbb{Q}$. *Wir haben* $g(\frac{17}{29}) = 1$, $g(11+\sqrt{2}) = 0$.

$$h \;:\; \mathbb{R}\setminus\{2\} \longrightarrow \mathbb{R}\,,\;\; t \longmapsto \frac{t^2}{t-2}\,. \tag{2.3}$$

Dann ist $h(3) = 9$.

$$k \;:\; [\,-1\,,\,1\,] \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \sqrt{1-x^2}\,. \tag{2.4}$$

Dann ist $k(\frac{1}{2}) = \frac{1}{2}\sqrt{3}$. *Berechnen Sie* $g(k(1/2))$ *sowie* $k(g(1/2))$.

$$z \;:\; (\,-1\,,\,1\,) \longrightarrow \mathbb{R}\,,\;\; q \longmapsto \frac{1}{1-q^2} \tag{2.5}$$

Mit einer Funktion $f : S \to T$ ist der Graph von f assoziiert:

$$\Gamma_f := \{(s, f(s))\colon s \in S\} \subset S \times T\,.$$

Für reelle Funktionen kann man sich den Graphen von f als Punktmenge einer Ebene veranschaulichen. Man zeichne z. B. zwei senkrecht aufeinander stehende Geraden, die man als reelle Zahlengeraden interpretiert und die sich im jeweiligen Nullpunkt schneiden. Die beiden Geraden nennen wir erste und zweite Achse oder x– und y–Achse. Jetzt entspricht jedem Punkt der Ebene ein Zahlenpaar $(x_1\,,\,x_2)$, das man geometrisch erhält, indem man jeweils das Lot auf die erste und die zweite Achse fällt. Umgekehrt erhält man zu jedem Paar (x_1, x_2) den zugehörigen Punkt der Ebene als Schnittpunkt der in x_1 bzw. x_2 auf die erste bzw. zweite Achse errichteten Lote. Γ_f ist interpretierbar als Teilmenge der Ebene mit der Maßgabe, dass auf jeder Parallelen zur zweiten Achse höchstens ein Punkt von Γ_f liegt.

Der Graph Γ_f einer Funktion $f : S \to T$ ist aber eigentlich nur die andere Seite ein und derselben Medaille. Eine mathematische Definition einer Funktion läßt sich wie folgt aussprechen:

Definition 2.2 *Seien S, T zwei nicht–leere Mengen. Eine Funktion f von S in T ist eine Teilmenge von $S \times T$ mit der Eigenschaft*

$$\bigwedge_{s \in S} \ \overset{1}{\bigvee_{t \in T}} \ (s,t) \in f \ . \tag{2.6}$$

Für das Element t in (2.6) schreiben wir $f(s)$.

S heißt Definitionsbereich von f, den wir manchmal auch mit $D(f)$ abkürzen. Der Wertebereich $W(f)$ von f ist die Menge aller Elemente aus T, zu denen ein $s \in S$ mit $(s,t) \in f$ existiert

$$W(f) = \{t \in T\colon \bigvee_{s \in S} \ (s,t) \in f\}$$

Sind $S' \subset S$, $T' \subset T$, so nennen wir

$$\begin{aligned} f(S') &= \{t \in T\colon \bigvee_{s' \in S'} (s',t) \in f\} \ , \\ f^{-1}(T') &= \{s \in S\colon \bigvee_{t' \in T'} (s,t') \in f\} \end{aligned}$$

das Bild von S' bzw. das Urbild von T'. Insbesondere ist also $W(f) = f(S)$, $D(f) = f^{-1}(T)$. Mit $\mathcal{F}(S,T)$ bezeichnen wir die Menge aller Funktionen von S nach T.

Eigentlich haben wir in Definition 2.2 den Graphen von f definiert. Wir wollen aber bei der suggestiveren Schreibweise

$$f \ : \ S \longrightarrow T \, , \ s \longmapsto f(s) \tag{2.7}$$

bleiben und für den Graphen von f weiterhin Γ_f schreiben. Mit der Schreibweise (2.7) hat man dann

$$\begin{aligned} f(S') &= \{f(s)\colon s \in S'\} \, , \\ f^{-1}(T') &= \{s \in S\colon f(s) \in T'\} \\ f = \Gamma_f &= \{(s \, , \, f(s))\colon s \in S\} \subset S \times T \end{aligned}$$

Definition 2.3 *Eine Funktion $f : S \to T$ heißt*

$$\begin{aligned} \text{injektiv} \quad &:\Leftrightarrow \quad \bigwedge_{t \in T} \ \#f^{-1}(\{t\}) \le 1 \Leftrightarrow \bigwedge_{s',s'' \in S} \ (f(s') = f(s'') \Rightarrow s' = s'') \\ \text{surjektiv} \quad &:\Leftrightarrow \quad W(f) = T \\ \text{bijektiv} \quad &:\Leftrightarrow \quad f \quad \text{ist injektiv und surjektiv.} \end{aligned}$$

Bei einer surjektiven Abbildung $f : S \to T$ spricht man auch von einer Abbildung auf T.

Man kann diese drei Begriffe deuten durch das Lösungsverhalten bei Gleichungen des Typs

$$f(x) = y \tag{2.8}$$

mit vorgegebenen $y \in T$:

f ist injektiv, wenn für jedes vorgelegte $y \in T$ die Gleichung (2.8) höchstens eine Lösung besitzt.

f ist surjektiv, wenn man für jedes vorgelegte $y \in T$ die Gleichung (2.8) überhaupt lösen kann.

f ist bijektiv, wenn Gleichung (2.8) für jedes vorgelegte $y \in T$ genau eine Lösung x besitzt.

Definition 2.4 *Sind R, S, T nicht–leere Mengen und $f \in \mathcal{F}(R,S)$, $g \in \mathcal{F}(S,T)$, so heißt*

$$g \circ f \ : \ R \longrightarrow T \ , \ r \longmapsto g(f(r))$$

die Hintereinanderschaltung von g nach f (lies: g "nach" f.)

Beispiel 2.5 *Seien: $f \ : \ \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto 3x+4$ und $g \ : \ \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto 4x+5$. Dann erhält man*

$$\begin{aligned} g \circ f \ &: \ \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto 4(3x+4)+5 = 12x+21 \\ f \circ g \ &: \ \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto 3(4x+5)+4 = 12x+19 \end{aligned}$$

Offenbar ist $g \circ f$ verschieden von $f \circ g$. Zwar haben $g \circ f$ und $f \circ g$ beide den gleichen Definitionsbereich und die gleiche Zielmenge, aber es gibt wenigstens eine Zahl, z. B. 61, der durch $f \circ g$ und durch $g \circ f$ verschiedene Funktionswerte zugeordnet werden:

$$f \circ g(61) : 751 \quad , \quad g \circ f(61) = 753$$

Beachten Sie hierbei: Zwei Funktionen F, G sind gleich, wenn sie gleichen Definitionsbereich und gleichen Zielbereich haben und wenn für jedes Element des gemeinsamen Definitionsbereichs $F(z) = G(z)$ ist.

Beispiel 2.6 *Seien*

$$\begin{aligned} g \ &: \ [\,0\,,\infty\,) \longrightarrow \mathbb{R} \ , \ x \longmapsto \sqrt{x} \ , \\ f \ &: \ \mathbb{R} \longrightarrow [\,0\,,\infty\,) \ , \ x \longmapsto x^2 \ . \end{aligned}$$

Dann ist

$$g \circ f \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ x \longmapsto \sqrt{f(x)} = \sqrt{x^2} = \mid x \mid ,$$
$$f \circ g \ : \ [\,0\,,\,\infty\,) \longrightarrow [\,0\,,\,\infty\,)\,, \ x \longmapsto g(x)^2 = (\sqrt{x})^2 = x\,.$$

Ist aber

$$G \ : \ [\,0\,,\,\infty\,) \longrightarrow [\,0\,,\,\infty\,)\,, \ x \longmapsto \sqrt{x}\,,$$
$$F \ : \ [\,0\,,\,\infty\,) \longrightarrow [\,0\,,\,\infty\,)\,, \ x \longmapsto x^2\,,$$

so gilt

$$F \circ G = G \circ F \ : \ [\,0\,,\,\infty\,) \longrightarrow [\,0\,,\,\infty\,)\,, \ x \longmapsto x$$

Definition 2.7 *Ist S eine nicht–leere Menge, so heißt*

$$id := id_S \ : \ S \longrightarrow S\,, \ s \longmapsto s$$

die Identität *auf S. Ist $R \subset S$ eine nicht–leere Teilmenge von S, so heißt*

$$\iota := \iota_{R,S} \ : \ R \longrightarrow S\,, \ s \longmapsto s$$

die Einbettung von R in S. Ist T eine weitere nicht–leere Menge und $f \in \mathcal{F}(S,T)$, so heißt $f_{|R} := f \circ \iota_{R,S}$ die Einschränkung von f auf R. Schließlich heißt $f \in \mathcal{F}(S,T)$ Fortsetzung einer Funktion $h \in \mathcal{F}(R,T)$, wenn $f_{|R} = h$ ist.

Im obigem Beispiel ist $F = f_{|[\,0\,,\,\infty\,)}$.

Satz und Definition 2.8 *Seien S, T nicht–leere Mengen und $f \in \mathcal{F}(S,T)$.*

(i) *f ist genau dann surjektiv, wenn es eine Funktion $f^r \in \mathcal{F}(T,S)$ gibt mit $f \circ f^r = id_T$. f^r heißt eine Rechtsinverse von f.*

(ii) *f ist genau dann injektiv, wenn es eine Funktion $f^l \in \mathcal{F}(T,S)$ gibt mit $f^l \circ f = id_S$. f^l heißt eine Linksinverse von f.*

(iii) *Ist f bijektiv, so existieren genau eine Rechts– (f^r) und genau eine Links-inverse (f^l) von f. Sie stimmen überein und werden die inverse Funktion (Inverse) von f genannt.*
Bezeichnung: $f^{-1} := f^r = f^l$.

Beweis: zu (i): Wir nehmen an, es gibt eine Rechtsinverse $f^r \in \mathcal{F}(T, S)$ zu f. Ist $t \in T$, so gilt mit $s = f^r(t)$ die Beziehung $f(s) = f(f^r(t)) = t$. D. h. jedes $t \in T$ ist Bild eines Elementes $s \in S$ unter f.

Ist umgekehrt f surjektiv, so erhält man die Existenz einer Rechtsinversen aus dem Auswahlaxiom der Mengenlehre, das wir als eine grundlegende Tatsache aus der Mengenlehre akzeptieren wollen:

Axiom 2.9 (Auswahlaxiom) *Sei S eine nicht–leere Menge und $\mathcal{P}(S)$ die Potenzmenge von S. Es gibt eine Funktion*

$$\phi : \mathcal{P}(S) \setminus \{\emptyset\} \to S$$

mit der Eigenschaft

$$\phi(M) \in M$$

für jede nicht–leere Teilmenge M von S.

Ist nun f surjektiv, so ist $f^{-1}(\{t\}) \neq \emptyset$ für jedes $t \in T$. Dann ist

$$f^r \; : \; T \longrightarrow S \, , \; t \longmapsto \phi \circ f^{-1}(\{t\})$$

eine Rechtsinverse.

zu (ii): Besitzt f eine Linksinverse und gilt für $s_1, s_2 \in S$

$$f(s_1) = f(s_2),$$

so folgt

$$s_1 = f^l(f(s_1)) = f^l(f(s_2)) = s_2,$$

f ist also injektiv.

Ist andererseits f injektiv, so existiert zu jedem $t \in W(f)$ genau ein $s \in S$ mit $f(s) = t$. Ist ferner $\hat{s}$ ein beliebiges Element aus S, so ist

$$f^l \; : \; T \longrightarrow S \, , \; t \longmapsto \begin{cases} s & , \quad \text{falls} \quad f(s) = t \\ \hat{s} & , \quad \text{falls} \quad t \notin W(f) \end{cases}$$

eine Linksinverse zu f.

zu (iii): Im Beweis zu (ii) erkennt man:
f^l muss auf $W(f)$ notwendig durch "$f^l(t) = s$ mit $t = f(s)$" definiert sein. Da im Fall bijektiven f's aber $W(f) = T$ ist, gibt es genau eine Linksinverse, und

diese ist auch Rechtsinverse. Im Beweis von (i) sieht man: $f^r(t)$ muss notwendig ein Element aus der Menge $f^{-1}(\{t\})$ sein. Ist aber f bijektiv, so enthält $f^{-1}(\{t\})$ genau ein Element, und dieses muss dann $f^r(t)$ sein. Damit gibt es genau eine Rechtsinverse.

q.e.d.

In der modernen Analysis werden Funktionen oft als Punkte eines Raumes aufgefasst. Entsprechend betrachtet man später auch Funktionen die anderen Funktionen, also Punkten eines Raumes, wieder Punkte eines Raumes zuordnen. Zum Beispiel im Raum $\mathbb{P}$ der Polynome, das sind Funktionen vom Typ

$$p \ : \ \mathbb{R} \longrightarrow \mathbb{R}\ , \ t \longmapsto \sum_{j=0}^{n} a_j t^j \qquad \text{mit gegebenen } a_1, \ldots, a_n \in \mathbb{R}\ ,$$

ist etwa

$$q \ : \ \mathbb{R} \longrightarrow \mathbb{R}\ , \ t \longmapsto 4t^3 + 17t^2 + 13t + 2$$

ein spezieller Punkt. Die Funktion

$$D \ : \ \mathbb{P} \longrightarrow \mathbb{P}\ , \ p \longmapsto p'$$

ordnet jedem Polynom seine Ableitung zu — Sie wissen aus der Schule, wie man Polynome differenziert. Dann ist

$$D(q) = q' \text{ mit } q'(t) = 12t^2 + 34t + 13\ .$$

Die eben beschriebene Sichtweise untermauert folgender Satz, den wir ohne Beweis aufschreiben wollen:

Satz 2.10 *Sei S eine nicht–leere Menge und $\mathcal{J}$ die Menge aller bijektiven Funktionen von S in sich:*

$$\mathcal{J}(S) := \{f \in \mathcal{F}(S,S) \colon f \ \textit{ ist bijektiv}\}\ .$$

Dann ist $(\mathcal{J}(S), \circ)$ eine Gruppe: Für jedes $f, g \in \mathcal{J}(S)$ ist $f \circ g \in \mathcal{J}(S)$ und es gelten:

(i) $\bigwedge_{f,g,h \in \mathcal{J}(S)} \quad f \circ (g \circ h) = (f \circ g) \circ h$

(ii) $\bigwedge_{f \in \mathcal{J}(S)} \quad f \circ id = id \circ f = f$

(iii) $\bigwedge_{f \in \mathcal{J}(S)} \quad f \circ f^{-1} = f^{-1} \circ f = id$

Man beachte, dass (i) und (ii) in allgemeinerem Kontext gelten:

$$\bigwedge_{f\in\mathcal{F}(S,T)} \quad \bigwedge_{g\in\mathcal{F}(R,S)} \quad \bigwedge_{h\in\mathcal{F}(Q,R)} f\circ(g\circ h) = (f\circ g)\circ h =: f\circ g\circ h\,,$$

$$\bigwedge_{f\in\mathcal{F}(S,T)} \quad f\circ id_S = id_T\circ f = f\ .$$

(Hier sind Q, R, S, T nicht–leere Mengen).

2.2 Über die Mächtigkeit von Mengen

In Definition 1.26 haben wir festgelegt, was eine endliche Menge ist. Unter einer unendlichen Menge (oder einer Menge mit unendlich vielen Elementen) versteht man eine Menge, die nicht endlich ist. Wir wollen aber den Grad der Unendlichkeit einer Menge noch genauer spezifizieren:

Definition 2.11 *Eine Menge S heißt abzählbar unendlich, wenn es eine bijektive Abbildung $\phi \in \mathcal{F}(\mathbb{N}, S)$ gibt. S heißt überabzählbar unendlich, wenn S unendlich, aber nicht abzählbar unendlich ist. Ist S abzählbar unendlich, so schreiben wir*

$$\#S = \aleph_0\,, \qquad\qquad \aleph\ :\ \textit{Aleph, hebräischer Buchstabe}$$

S heißt abzählbar, wenn S endlich oder abzählbar unendlich ist.

Ohne Beweis bemerken wir:

Satz 2.12 *Sei S eine nicht–leere Menge. S ist genau dann abzählbar, wenn es eine surjektive Abbildung $\phi \in \mathcal{F}(\mathbb{N}, S)$ gibt.*

Es folgt:

Satz 2.13

(i) *Ist S abzählbar, so ist jede Teilmenge von S abzählbar.*

(ii) *Sind S, T abzählbar, so auch $S \times T$.*

(iii) *Ist $\mathcal{F}$ eine abzählbare Familie abzählbarer Mengen, so ist*

$$X = \bigcup \mathcal{F} = \{x\colon \bigvee_{F \in \mathcal{F}} \; x \in F\}$$

abzählbar.

Beweis:
zu (i): Sei T nicht–leere Teilmenge der abzählbaren Menge S und $\varphi : \mathbb{N} \to S$ surjektiv. Ferner sei $\hat{t} \in T$. Dann ist

$$\psi \ : \ \mathbb{N} \longrightarrow T \ , \ n \longmapsto \begin{cases} \varphi(n) & , \ \text{falls} \quad \varphi(n) \in T \\ \hat{t} & , \ \text{falls} \quad \varphi(n) \notin T \end{cases}$$

surjektive Abbildung von $\mathbb{N}$ auf T.

zu (ii): (Beweisskizze) Zunächst soll die Idee erläutert werden. Wir schreiben die surjektiven Funktionen $\varphi \in \mathcal{F}(\mathbb{N}, S)$ und $\psi \in \mathcal{F}(\mathbb{N}, T)$ *als Folgen*: $\varphi = (s_n)_{n \in \mathbb{N}}$, $\psi = (t_n)_{n \in \mathbb{N}}$ und denken uns $S \times T$ notiert in einem Schema in der Form eines Quadranten:

$$\begin{array}{ccccccccc}
(s_1,t_1) & , & (s_1,t_2) & , & (s_1,t_3) & , & (s_1,t_4) & , & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & \dots \\
(s_2,t_1) & , & (s_2,t_2) & , & (s_2,t_3) & , & (s_2,t_4) & , & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & \dots \\
(s_3,t_1) & , & (s_3,t_2) & , & (s_3,t_3) & , & (s_3,t_4) & , & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & \dots \\
(s_4,t_1) & , & (s_4,t_2) & , & (s_4,t_3) & , & (s_4,t_4) & , & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & \dots \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots
\end{array}$$

In der i-ten Zeile und j-ten Spalte steht $(s_i, t_j) = (\varphi(i), \psi(j))$. Nun wird $S \times T$ *abgezählt*, indem man nacheinander die Elemente (s_i, t_j) abzählt, für die die Indexsumme fest ist: $i + j = k$.

Ausgeschrieben erhält man

$$\Theta : \mathbb{N} \to S \times T$$

wie folgt: man mache sich zunächst klar, dass

$$M_k := \{(i,j) \in \mathbb{N}\colon i + j = k\} \quad , \quad k \in \mathbb{N} \quad \text{mit} \quad k \geq 2$$

$k - 1$ Elemente enthält:

$$\#M_k = k - 1,$$

und dass

$$\bigwedge_{k\in\mathbb{N}} \quad \sum_{\nu=1}^{k} \nu = \frac{k}{2}(k+1) \; .$$

Wir definieren

$$\zeta \; : \; \mathbb{N} \longrightarrow \mathbb{N} \, , \; n \longmapsto \min\{k \in \mathbb{N}\colon n \leq \frac{k}{2}(k+1)\} \; .$$

Dann ist

$$\Theta \; : \; \mathbb{N} \longrightarrow S \times T \, , \; n \longmapsto (s_{n-\frac{\zeta(n)}{2}(\zeta(n)-1)}, t_{\frac{\zeta(n)}{2}(\zeta(n)+1)-n+1})$$

eine surjektive Abbilung.

(iii): kann man ähnlich beweisen.

q.e.d.

Für unser Zahlensystem erhalten wir

Satz 2.14 **(i)** $\mathbb{N}, \mathbb{Z}$ *und* $\mathbb{Q}$ *sind abzählbar unendlich.*

(ii) $\mathbb{R}$ *ist überabzählbar unendlich.*

Ist eine Menge M überabzählbar unendlich und gibt es eine bijetive Abbildung von $\mathbb{R}$ auf M, so schreibt man

$$\# M = c$$

Man nennt c die Mächtigkeit des Kontinuums. Die Kontinuumshypothese, von Cantor 1878 aufgestellt, besagt, dass jede überabzählbare Menge mindestens die Mächtigkeit des Kontinuums besitzt, also surjektiv auf $\mathbb{R}$ abgebildet werden kann. Es ist nicht gelungen, diese Hypothese zu beweisen — im Gegenteil: die Kontinuumshypothese ist auf der Grundlage der heute verwendeten mengentheoretischen Axiomensysteme weder beweisbar (Cohen 1968) noch widerlegbar (Gödel 1938).

Beweis (von Satz 2.14**):**
zu (i): Natürlich ist $\mathbb{N}$ abzählbar unendlich, denn $id_{\mathbb{N}} \in \mathcal{F}(\mathbb{N}, \mathbb{N})$ ist bijektiv. Wegen $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-n\colon n \in \mathbb{N}\}$ ist nach 2.13 (iii) auch $\mathbb{Z}$ abzählbar. Schließlich ist

$$\psi \; : \; \mathbb{Z} \times \mathbb{N} \longrightarrow \mathbb{Q} \, , \; (p,q) \longmapsto \frac{p}{q}$$

surjektiv. Nach 2.13 (ii) existiert eine surjektive Abbildung $\varphi \in \mathcal{F}(\mathbb{N}, \mathbb{Z} \times \mathbb{N})$. Dann ist $\psi \circ \varphi \in \mathcal{F}(\mathbb{N}, \mathbb{Q})$ surjektiv.

zu (ii): Wir nehmen an, $\mathbb{R}$ wäre abzählbar und $\varphi \in \mathcal{F}(\mathbb{N}, \mathbb{R})$ wäre surjektiv.

Sei

$$I_1 := [\, a_1 \, , \, b_1 \,] \quad , \quad a_1 < b_1$$

ein abgeschlossenes beschränktes Intervall, in dem $\varphi(1)$ *nicht* liegt:

$$\varphi(1) \notin I_1 \; ,$$

etwa $I_1 := [\, \varphi(1) + 1 \, , \, \varphi(1) + 2 \,]$. Wir definieren eine Folge $(I_n)_{n \in \mathbb{N}}$ abgeschlossener nicht–leerer Intervalle mit den Eigenschaften

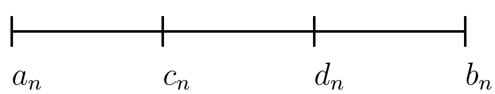
$$\bigwedge_{n \in \mathbb{N}} \quad I_n = [\, a_n \, , \, b_n \,] \quad \text{mit} \quad a_n < b_n \; . \tag{2.9}$$

$$\bigwedge_{n \in \mathbb{N}} \quad I_{n+1} \subset I_n \tag{2.10}$$

$$\bigwedge_{n \in \mathbb{N}} \quad \varphi(n) \notin I_n \; . \tag{2.11}$$

I_1 haben wir soeben definiert. Ist I_n für irgendein $n \in \mathbb{N}$ definiert, $I_n = [\, a_n \, , \, b_n \,]$, so sei mit $c_n := \frac{2}{3} a_n + \frac{1}{3} b_n$ und $d_n := \frac{1}{3} a_n + \frac{2}{3} b_n$

$$I_{n+1} := \begin{cases} [\, a_n \, , \, c_n \,] \, , & \text{falls} \;\; \varphi(n+1) \in [\, d_n \, , \, b_n \,] \\ [\, d_n \, , \, b_n \,] & \text{sonst} \end{cases} \; .$$



Die so definierte Folge von Intervallen besteht aus abgeschlossenen Intervallen, und eine leichte Induktion zeigt, dass (2.9), (2.10), (2.11) gelten. Nach Satz 1.38 existiert ein $x \in \mathbb{R}$ mit

$$\bigwedge_{n \in \mathbb{N}} \quad x \in I_n \; .$$

Aber $x = \varphi(m)$ für ein $m \in \mathbb{N}$. Daher ist $x \notin I_m$, und das ist ein Widerspruch.

q.e.d.

In der Menge $\mathbb{I} := \mathbb{R} \setminus \mathbb{Q}$ sind die irrationalen Zahlen zusammengefaßt. Wegen des Satzes 2.13 muss es davon überabzählbar viele geben. Eine $\mathbb{Q}$ umfassende

Teilmenge von $\mathbb{R}$ ist die Menge $\mathbb{A}$ der algebraischen Zahlen. Das sind alle Zahlen, die als Nullstellen irgendeines nicht konstanten Polynoms mit rationalen Koeffizienten auftreten[1]. Zum Beispiel ist $\sqrt{2}$ als Nullstelle des Polynoms $x^2 - 2$ eine algebraische Zahl. Da es nur abzählbar viele Polynome mit rationalen Koeffizienten gibt und jedes nicht konstante Polynom höchstens endlich viele Nullstellen besitzt, ist auch die Menge der algebraischen Zahlen abzählbar. Das Komplement $\mathbb{R} \setminus \mathbb{A}$, die Menge der *transzendenten Zahlen*, ist folglich überabzählbar. *Ferdinand von Lindemann (1852–1939)* hat bewiesen, dass e (siehe Definition 3.38) und π (siehe Definition und Satz 5.40(vi)) transzendente Zahlen sind. Einen Nachweis für die Irrationalität von e werden wir am Ende des Skriptums angeben.

2.3 Reelle Funktionen und Konvergenz

Definition 2.15 *Eine reelle Funktion ist eine Funktion, deren Definitionsbereich eine Teilmenge von $\mathbb{R}$ und deren Zielbereich $\mathbb{R}$ ist.*

Zum Beispiel:

$$\left(\frac{5+n}{n^2+1}\right)_{n\in\mathbb{N}}$$

ist eine reelle Funktion. Ihr Definitionsbereich ist $\mathbb{N} \subset \mathbb{R}$, ihr Zielbereich $\mathbb{R}$.

Ferner

$$f \;:\; [\,0\,,\,1\,) \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \frac{1}{(x-1)(x+1)}$$

ist eine reelle Funktion mit Definitionsbereich $D(f) = [\,0\,,\,1\,)$

Oft schreibt man: *Die Funktion f ist durch (z. B.) $f(x) := x^{1/2} + (x-2)^{-1}$ definiert.* Dann wollen wir darunter die reelle Funktion f verstehen, deren Zielbereich $\mathbb{R}$ und deren Definitionsbereich aus all den $x \in \mathbb{R}$ besteht, für die die angegebene Abbildungsvorschrift durchführbar ist. In obigem Beispiel erhält man$D(f) = (\,0\,,\,2\,) \cup (\,2\,,\,\infty\,) = \mathbb{R}^+ \setminus \{2\}$.

Die Funktion g sei durch $g(x) = x^x$ definiert. Da wir bisher Potenzen nur für rationale Exponenten definiert haben, müssten wir jetzt $D(g) = (\mathbb{Q} \cap [\,0\,,\,\infty\,)) \cup \mathbb{Z}$ als Definitionsbereich aufschreiben. Wir werden später für positives $a \in \mathbb{R}$ und jedes $r \in \mathbb{R}$ auch die Potenz a^r definieren. Mit diesem Wissen erhält man dann $D(g) = \mathbb{R}^+ \cup \mathbb{Z}$.

[1]Erkenntnisse über die Nullstellen solcher Polynome zu gewinnen, ist ein grundlegendes Problem der Algebra; daher der Name *algebraische* Zahlen.

Die Operation der Hintereinanderschaltung wollen wir etwas modifizieren:

Seien

$$f \,:\, [\,0\,,\,\infty\,) \longrightarrow \mathbb{R}\,,\; x \longmapsto x^{1/2} \quad , \quad g \,:\, \mathbb{R} \longrightarrow \mathbb{R}\,,\; x \longmapsto x^2\,.$$

Da der Zielbereich von g ganz $\mathbb{R}$ ist, kann man $f \circ g$ gemäß Definition 2.4 nicht bilden.

Aber in diesem Fall kann man argumentieren, das der Wertebereich von g in $[\,0\,,\,\infty\,)$ enthalten ist, und man könnte g als Funktion mit Zielbereich $[\,0\,,\,\infty\,)$ auffassen und wieder $f \circ g$ schreiben. Ersetzt man nun g durch

$$g \,:\, \mathbb{R} \longrightarrow \mathbb{R}\,,\; x \longmapsto x^3 - 1$$

so liegt der Wertebereich von g nicht in $[\,0\,,\,\infty\,)$. Man kann aber $f \circ g\,|_{[\,1\,,\,\infty\,)}$ bilden.

Für reelle Funktionen legen wir fest:

Definition 2.16 *Sind f, g reelle Funktionen, so verstehen wir unter $f \circ g$ die reelle Funktion*

$$f \circ g \,:\, D(f \circ g) \longrightarrow \mathbb{R}\,,\; x \longmapsto f(g(x))$$

mit dem Definitionsbereich

$$D(f \circ g) := \{x \in D(g)\colon\, g(x) \in D(f)\} = g^{-1}(D(f)),$$

sofern diese Menge nicht leer ist.

Sind also f und g durch $f(x) = x^{1/2}, g(x) = x^3 - 1$ gegeben, so erhält man

$$\begin{array}{rcl} f \circ g &:& [\,1\,,\,\infty\,) \longrightarrow \mathbb{R}\,,\; x \longmapsto (x^3 - 1)^{1/2} \\ g \circ f &:& [\,0\,,\,\infty\,) \longrightarrow \mathbb{R}\,,\; x \longmapsto x^{3/2} - 1 \end{array}$$

Da man reelle Zahlen addieren und multiplizieren kann, kann man auch reelle Funktionen addieren und multiplizieren:

Definition 2.17 *Seien f, g reelle Funktionen. Dann heißen*

$$\begin{array}{l} f + g \,:\, D(f+g) := D(f) \cap D(g) \longrightarrow \mathbb{R}\,,\; x \longmapsto f(x) + g(x) \\ f - g \,:\, D(f-g) := D(f) \cap D(g) \longrightarrow \mathbb{R}\,,\; x \longmapsto f(x) - g(x) \\ f \cdot g \,:\, D(f \cdot g) := D(f) \cap D(g) \longrightarrow \mathbb{R}\,,\; x \longmapsto f(x) \cdot g(x) \\ f/g \,:\, D(f/g) := D(f) \cap D(g) \setminus N(g) \longrightarrow \mathbb{R}\,,\; x \longmapsto f(x)/g(x) \end{array}$$

(mit $N(g) := \{x \in D(g)\colon g(x) = 0\}$) Summe, Produkt, Differenz bzw. Quotient von f und g, sofern $D(f) \cap D(g)$ bzw. $D(f) \cap D(g) \setminus N(g)$ nicht leer sind.

Mit $\alpha \in \mathbb{R}$ sei schließlich

$$\alpha f \;:\; D(f) \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \alpha f(x)\,.$$

Man beachte, dass bei der Festlegung von Summe, Produkt etc. von Funktionen es nicht darauf ankommt, dass der Definitionsbereich eine Teilmenge von $\mathbb{R}$ ist, sondern der Zielbereich muss $\mathbb{R}$ oder allgemeiner ein Körper sein. Genauer gilt:

Satz 2.18 *Ist S eine nicht–leere Menge und $\mathbb{K}$ ein Körper, so ist $\mathcal{F}(S,\mathbb{K})$ ein Vektorraum über $\mathbb{K}$, wenn man für $f, g \in \mathcal{F}(S,\mathbb{K})$ und $\alpha \in \mathbb{K}$ definiert*

$$f+g \;:\; S \longrightarrow \mathbb{K}\,,\;\; x \longmapsto f(x)+g(x)\,,\quad \alpha f \;:\; S \longrightarrow \mathbb{K}\,,\;\; x \longmapsto \alpha f(x)\,.$$

Auf den Beweis wollen wir verzichten.

Der grundlegende Begriff für die Analysis ist der Begriff des Grenzwerts. Hat man eine reelle Funktion, etwa gegeben durch

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

wüsste man gern, ob sich die Werte von f einem gewissen Wert annähern, wenn man sich mit der freien Variablen x einem der Werte 2 oder -2 oder aber sogar ∞ oder $-\infty$ annähert:

Zunächst benötigen wir den Begriff eines Häufungspunkts. Hierzu erinnern wir an den Begriff einer offenen ε-Umgebung eines Punktes $a \in \mathbb{R}$ in Definiton 1.7. Um $-\infty$ und ∞ mitbehandeln zu können definieren wir für $\varepsilon > 0$

$$U_\varepsilon(\infty) := (\,\tfrac{1}{\varepsilon}\,,\,\infty\,) \quad,\quad U_\varepsilon(-\infty) := (\,-\infty\,,\,-1/\varepsilon\,)$$

und nennen diese Mengen ε-Umgebungen von ∞ bzw. von $-\infty$.

Definition 2.19 **(i)** *Sei $a \in \hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. Eine Umgebung von a ist eine Teilmenge $U \subset \mathbb{R}$, die mit einem $\varepsilon > 0$ eine ε-Umgebung von a enthält. $\mathcal{U}(a)$ bezeichnet die Familie aller Umgebungen von a, also*

$$\mathcal{U}(a) := \{U \subset \mathbb{R}\colon \bigvee_{\varepsilon>0} U_\varepsilon(a) \subset U\}$$

(ii) *Sei M eine Teilmenge von $\mathbb{R}$. Ein Punkt $a \in \mathbb{R}$ heißt Häufungspunkt von M, wenn jede Umgebung von a wenigstens einen von a verschiedenen Punkt aus M enthält:*

$$\bigwedge_{U \in \mathcal{U}(a)} U \cap (M \setminus \{a\}) \neq \emptyset \tag{2.12}$$

$b = \infty$ bzw. $b = -\infty$ heißt uneigentlicher Häufungspunkt von M, wenn

$$\bigwedge_{U \in \mathcal{U}(b)} U \cap M \neq \emptyset \, . \tag{2.13}$$

Beachten Sie: (2.12) ist äquivalent mit

$$\bigwedge_{\varepsilon > 0} \bigvee_{x \in M} 0 <\mid x - a \mid< \varepsilon \tag{2.14}$$

und (2.13) mit

$$\bigwedge_{\varepsilon > 0} \bigvee_{x \in M} x > 1/\varepsilon \qquad (\text{im Fall} \quad b = \infty) \tag{2.15}$$

$$\bigwedge_{\varepsilon > 0} \bigvee_{x \in M} x < -1/\varepsilon \quad (\text{im Fall} \quad b = -\infty) \tag{2.16}$$

Wir kommen zur Grenzwertdefinition:

Definition 2.20 *Sei f reelle Funktion und $x_0 \in \hat{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ ein (uneigentlicher) Häufungspunkt von $D(f)$. $a \in \hat{\mathbb{R}}$ heißt Grenzwert in $\hat{\mathbb{R}}$ von f bei x_0, wenn*

$$\bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x \in U_\delta(x_0) \cap D(f) \setminus \{x_0\}} f(x) \in U_\varepsilon(a) \, . \tag{2.17}$$

Wir schreiben

$$a = \hat{\mathbb{R}}\text{-}\lim_{x \to x_0} f(x) \quad \text{oder} \quad f(x) \to a \quad \text{für} \quad x \to x_0 \, , \tag{2.18}$$

und sagen in diesem Fall: f besitzt bei x_0 einen Grenzwert in $\hat{\mathbb{R}}$. Liegt a in $\mathbb{R}$, also $a \notin \{-\infty, \infty\}$, so sagen wir: f konvergiert (gegen a) und schreiben auch

$$a = \lim_{x \to x_0} f(x) \, . \tag{2.19}$$

Gibt es kein $a \in \mathbb{R}$, gegen das f für x gegen x_0 konvergiert, so sagt man: f divergiert für x gegen x_0. Ist ∞ (oder $-\infty$) Grenzwert in $\hat{\mathbb{R}}$ von f bei x_0, so sagt man auch: f divergiert gegen ∞ (oder gegen $-\infty$) bei x_0.

Die Aussage (2.17) kann man wie folgt interpretieren. Ganz gleich, wie groß eine Fehlerschranke (ε) angenommen wird, liegt die unabhängige Variable x nur hinreichend nah (δ) bei x_0, so liegt der Funktionswert innerhalb des vorgeschriebenen Fehlerintervalls.

Zu (2.17) ist äquivalent:

$$\bigwedge_{U \in U(a)} \quad \bigvee_{V \in U(x_0)} \quad f(V \cap D(f) \setminus \{x_0\}) \subset U \; . \tag{2.20}$$

Eine Funktion kann bei $x_0 \in \hat{\mathbb{R}}$ höchstens einen Grenzwert $a \in \hat{\mathbb{R}}$ besitzen:

Satz 2.21 *Besitzt f bei x_0 einen Grenzwert in $\hat{\mathbb{R}}$, so ist dieser eindeutig bestimmt.*

Beweis: Sind a, b verschiedene Punkte in $\hat{\mathbb{R}}$, so gibt es ε-Umgebungen $U_\varepsilon(a)$ von a und $U_\varepsilon(b)$ von b mit

$$U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset \; . \tag{2.21}$$

Sind $a, b \in \mathbb{R}$, so wähle man $\varepsilon =\mid a - b \mid /2$. Ist $a = \infty$, so gibt es eine Zahl $R > 0$ mit $b < R$,und (2.21) gilt mit $\varepsilon = \frac{1}{2R}$. Ist $a = -\infty$, so existiert eine Zahl $R > 0$ mit $b > -R$ und 2.21 gilt wieder mit $\varepsilon = \frac{1}{2R}$. Gäbe es nun zwei verschiedene Grenzwerte a, b in $\hat{\mathbb{R}}$ für f bei x_0, so könnte man $\varepsilon > 0$ so bestimmen, dass (2.21) gültig wäre. Gemäß (2.21) gäbe es aber ein $\delta_a > 0$ und ein $\delta_b > 0$, so dass

$$\bigwedge_{x \in U_{\delta_a}(x_0) \cap D(f) \setminus \{x_0\}} f(x) \in U_\varepsilon(a) \tag{2.22}$$

$$\bigwedge_{x \in U_{\delta_b}(x_0) \cap D(f) \setminus \{x_0\}} f(x) \in U_\varepsilon(b) \tag{2.23}$$

Dann gilt mit $\delta := min\{\delta_a, \delta_b\}$:

$$\bigwedge_{x \in U\delta(x_0) \cap D(f) \setminus \{x_0\}} f(x) \in (U_\varepsilon(b) \cap U_\varepsilon(a)) = \emptyset$$

Dies ist ein Widerspruch, denn $U_\delta(x_0) \cap D(f) \setminus \{x_0\}$ ist nicht leer.

q.e.d.

Es ist also vernünftig, von *dem* Grenzwert bei x_0 zu sprechen.

Natürlich ist (für $f(x) = x$, $x_0 \in \hat{\mathbb{R}}$).

$$\hat{\mathbb{R}}\text{–} \lim_{x \to x_0} x = x_0 \; .$$

Wenn Sie nämlich eine Fehlerschranke $\varepsilon > 0$ vorschreiben, so kann ich Ihnen eine Zahl δ nennen, so dass

$$x \in U_\varepsilon(x_0) \ ,$$

sofern $x \in U_\delta(x_0) \setminus \{x_0\}$ ist, nämlich $\delta := \varepsilon$.

Ebenfalls ist (für $f(x) = \alpha$ mit einer Zahl $\alpha \in \mathbb{R}$, $x_0 \in \hat{\mathbb{R}}$)

$$\lim_{x \to x_0} \alpha = \alpha \ .$$

Zu vorgegebenen ε wähle man $\delta > 0$ irgendwie, z. B. $\delta = 7$. Dann ist

$$f(x) = \alpha \in U_\varepsilon(\alpha),$$

sofern $x \in U_\delta(x_0) \setminus \{x_0\}$ ist.

Mit Grenzwerten kann man rechnen:

Satz 2.22 *Seien f, g reelle Funktionen und x_0 ein (eventuell uneigentlicher) Häufungspunkt von $D(f) \cap D(g)$. Es mögen $\lim\limits_{x \to x_0} f(x)$ und $\lim\limits_{x \to x_0} g(x)$ existieren. Dann gelten*

$$\lim_{x \to x_0} [f(x) \pm g(x)] = \lim_{x \to x_0} f(x) \pm \lim_{x \to x_0} g(x) \tag{2.24}$$

$$\lim_{x \to x_0} (f(x)g(x)) = \left(\lim_{x \to x_0} f(x) \right) \cdot \left(\lim_{x \to x_0} g(x) \right) \ , \tag{2.25}$$

sofern die Grenzwerte auf der rechten Seite existieren.

Ist x_0 (uneigentlicher) Häufungspunkt von $D(f/g)$, so ist

$$\lim_{x \to x_0} \left(\frac{f(x)}{g(x)} \right) = \frac{\lim\limits_{x \to x_0} f(x)}{\lim\limits_{x \to x_0} g(x)} \ , \tag{2.26}$$

sofern $\lim\limits_{x \to x_0} g(x) \neq 0$ ist.

Beispiel 2.23

$$\begin{aligned}
\lim_{x \to 1} \frac{x^3 - 1}{x^2 - 1} &= \lim_{x \to 1} \frac{(x-1)(x^2+x+1)}{(x-1)(x+1)} \\
&= \lim_{x \to 1} \frac{x^2+x+1}{x+1} = \frac{\lim\limits_{x \to 1} x \cdot \lim\limits_{x \to 1} x + \lim\limits_{x \to 1} x + \lim\limits_{x \to 1} 1}{\lim\limits_{x \to 1} x + \lim\limits_{x \to 1} 1} \\
&= \frac{1 \cdot 1 + 1 + 1}{1 + 1} = \frac{3}{2} \ .
\end{aligned}$$

Beweis: Sei $a := \lim\limits_{x\to x_0} f(x)$, $b := \lim\limits_{x\to x_0} g(x)$.
zu (2.24)**:** Zunächst gilt

$$| f(x) \pm g(x) - (a \pm b) | \leq | f(x) - a | + | g(x) - b | \ . \tag{2.27}$$

Ist $\varepsilon > 0$ vorgegeben, so existieren $\hat{\delta} > 0$ und $\tilde{\delta} > 0$ derart, dass

$$| f(x) - a | < \varepsilon/2 \quad , \quad | g(x) - b | < \varepsilon/2,$$

falls $x \in (D(f) \setminus \{x_0\}) \cap U_{\hat{\delta}}(x_0)$ und $x \in (D(g) \setminus \{x_0\}) \cap U_{\hat{\delta}}(x_0)$.
Man beachte, dass stets

$$U_{\hat{\delta}}(s_\circ) \cap U_{\tilde{\delta}}(x_0) = U_\delta(x_0) \quad \text{mit} \quad \delta := min(\hat{\delta}, \tilde{\delta}) \ .$$

gilt. Daher und wegen (2.25) ist mit $\delta := min\{\hat{\delta}, \tilde{\delta}\}$ für jedes $x \in [D(f) \cap D(g)] \setminus \{x_0\} \cap U_\delta(x_0)$:

$$| f(x) \pm g(x) - (a \pm b) | < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

d. h. $f(x) \pm g(x) \in U_\varepsilon(a \pm b)$.

zu (2.25)**:** Zunächst gilt

$$\begin{aligned} | f(x)g(x) - ab | &= | f(x)(g(x) - b) + (f(x) - a)b | \\ &\leq | f(x) || g(x) - b | + | b || f(x) - a | \qquad (2.28) \\ &\leq | f(x) - a || g(x) - b | + | a || g(x) - b | + | b || f(x) - a | \qquad (2.29) \end{aligned}$$

Sei nun $\varepsilon > 0$ vorgegeben. Dann gibt es $\delta > 0$, so, dass für alle $x \in U_\delta(x_0)$ gelten:

$$| g(x) - b | < \frac{1}{3} \cdot \frac{\varepsilon}{\max\{| a |, 1\}} \quad , \quad | f(x) - a | < \min \left\{ 1 \ ; \ \frac{\varepsilon}{\max\{| b |, 1\}} \right\} \ .$$

Dann folgt aus (2.26) für diese $x \in U_\delta(x_0)$:

$$| f(x)g(x) - ab | < 1 \cdot \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \ \text{, also} \quad f(x)g(x) \in U_\varepsilon(ab) \ .$$

zu (2.26)**:** Wegen (2.25) genügt es, den Fall der durch $f(x) = 1$ gegebenen Funktion zu behandeln. Zunächst ist für alle $x \in D(1/g) = D(g) \setminus N(g)$:

$$\left| \frac{1}{g(x)} - \frac{1}{b} \right| = \left| \frac{b - g(x)}{g(x)b} \right| = \frac{1}{| g(x) || b |} \, | g(x) - b | \ . \tag{2.30}$$

Wähle nun δ so, dass für $x \in U_\delta(x_0)$ gelten

$$| g(x) - b | < \frac{| b |}{2} \tag{2.31}$$

$$| g(x) - b | < \frac{| b |^2}{2} \varepsilon \, . \tag{2.32}$$

Für diese x ist zunächst wegen (2.31)

$$| g(x) | \geq | b | - | g(x) - b | > \frac{| b |}{2}, \tag{2.33}$$

und (2.32) liefert zusammen mit (2.33):

$$\left(\frac{1}{g(x)} - \frac{1}{b} \right) = \frac{1}{| g(x) || b |} | g(x) - b | \leq \frac{2}{| b |^2} | g(x) - b | < \varepsilon$$

q.e.d.

Wir wollen noch die *uneigentlichen* Grenzwerte ∞ und $-\infty$ mit einbeziehen und definieren in gewissen Situationen Summen, Produkte, Differenzen und Quotienten für Elemente $a \in \hat{\mathbb{R}}$ und $b \in \{-\infty, \infty\}$:

Definition 2.24 *Im folgenden sei $a \in \mathbb{R}$, $b \in \{-\infty, \infty\}$ und*

$$-b := \begin{cases} \infty & , \ \textit{falls} \quad b = -\infty \\ -\infty & , \ \textit{falls} \quad b = \infty \end{cases} .$$

Dann werden festgelegt:

$$\begin{aligned} a + b &:= b + a = b \\ a - b &:= -b \\ b + b &:= b \\ -b - b &:= -b \\ a/b &:= 0 \\ a \cdot b &= b \cdot a = b \quad \text{falls} \quad a > 0 \\ a \cdot b &= b \cdot a = -b \quad \text{falls} \quad a < 0 \\ b \cdot b &:= \infty \\ (-b) \cdot b &:= b \cdot (-b) := -\infty \end{aligned}$$

Beachten Sie, dass "$\frac{a}{0}$", "$0 \cdot \infty$", "$\frac{\infty}{\infty}$" undefiniert bleiben!

Nun kann Satz 2.22 erweitert werden zu

Satz 2.25 *Seien f, g reelle Funktionen und x_0 ein (eventuell uneigentlicher) Häufungspunkt von $D(f) \cap D(g)$. Die $\hat{\mathbb{R}}$-Limiten von f und g bei x_0 mögen existieren. Dann gelten*

$$\hat{\mathbb{R}}\text{-}\lim_{x \to x_0} (f(x) \pm g(x)) = \hat{\mathbb{R}}\text{-}\lim_{x \to x_0} f(x) \pm \hat{\mathbb{R}}\text{-}\lim_{x \to x_0} g(x) \tag{2.34}$$

$$\hat{\mathbb{R}}\text{–}\lim_{x\to x_0}\left(f(x)\cdot g(x)\right) = \left(\hat{\mathbb{R}}\text{–}\lim_{x\to x_0} f(x)\right)\cdot\left(\hat{\mathbb{R}}\text{–}\lim_{x\to x_0} g(x)\right)\ , \tag{2.35}$$

$$\hat{\mathbb{R}}\text{–}\lim_{x\to x_0}\left(\frac{f(x)}{g(x)}\right) = \frac{\hat{\mathbb{R}}\text{–}\lim\limits_{x\to x_0} f(x)}{\hat{\mathbb{R}}\text{–}\lim\limits_{x\to x_0} g(x)}\ , \tag{2.36}$$

sofern jeweils der Ausdruck auf der rechten Seite definiert ist.

Beachten Sie, dass in (2.36) *die rechte Seite nur dann definiert sein kann, wenn* $\hat{\mathbb{R}}\text{–}\lim\limits_{x\to x_0} g(x) \neq 0$ *ist. Dann ist aber* $N(g)\cap U_\delta(x_0)\setminus\{x_0\}\neq\emptyset$ *für alle* $\delta>0$, *und damit ist* x_0 *(eventuell uneigentlicher) Häufungspunkt von* $D(f/g)$.

Beweis: Es soll nur der Beweis von (2.34) mit $\pm := +$ durchgeführt werden. Wir schreiben $\hat{\mathbb{R}}\text{–}\lim\limits_{x\to x_0} f(x) =: a$, $\hat{\mathbb{R}}\text{–}\lim\limits_{x\to x_0} g(x) =: b$ und nehmen wegen des Satzes 2.22 an, dass $b=\infty$ ist. (Den Fall $b=-\infty$ überlassen wir dem Leser). Dann ist $a\in\hat{\mathbb{R}}$ und $a\neq-\infty$. Zunächst gibt es ein $\hat{\delta}>0$, so dass für $x\in(U_{\hat{\delta}}(x_0)\setminus\{x_0\})\cap D(f)$ gilt

$$f(x)\in U_1(a)\ .$$

Daher folgt mit

$$c := \begin{cases} 1+\mid a\mid & ,\ \text{falls}\ \ a\neq\infty \\ 1 & ,\ \text{falls}\ \ a=\infty \end{cases}$$

für $x\in U_\delta(x_0)\cap D(f)\setminus\{x_0\}$:

$$f(x) > \left\{\begin{array}{lll} a-1 & , & \text{falls}\ \ a\neq\infty \\ 1 & , & \text{falls}\ \ a=\infty \end{array}\right\} \geq -c\ .$$

Zu vorgegebenen $\varepsilon>0$ wähle nun $\delta>0$ so, dass $\delta\leq\hat{\delta}$ und dass für alle $x\in U_\delta(x_0)\cap D(f)\setminus\{x_0\}$

$$g(x) > c+\frac{1}{\varepsilon}\ ,$$

also

$$g(x)\in U_{\frac{\varepsilon}{1+c\varepsilon}}(\infty)$$

ist. Dann folgt für $x\in U_\delta(x_0)\cap D(f)\cap D(g)\setminus\{x_0\}$

$$f(x)+g(x) > -c+c+1/\varepsilon = 1/\varepsilon\ ,\quad \text{also}\quad f(x)+g(x)\in U_\varepsilon(\infty)\ .$$

q.e.d.

Beispiele 2.26

$$\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{x^3-1}{x^2-1}=\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{x-\frac{1}{x^2}}{1-\frac{1}{x^2}}=\frac{\infty-0}{1-0}=\infty$$

$$\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{x^2+4x+15}{2x^2+3x+1}=\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{1+\frac{4}{x}+\frac{15}{x^2}}{2+\frac{3}{x}+\frac{1}{x^2}}=\frac{1+0+0}{2+0+0}=\frac{1}{2}$$

Im zweiten Fall liegt Konvergenz gegen $1/2$ *vor.*

$$\begin{aligned}\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{x^4+3x^3+2x+1}{x^7+3x}&=\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{\frac{1}{x^3}+\frac{3}{x^4}+\frac{2}{x^6}+\frac{1}{x^7}}{1+\frac{3}{x^6}}\\&=\frac{0+0+0+0}{1+0}=0\end{aligned}$$

Beachten Sie: sind P, Q Polynome, also Funktionen vom Typ

$$f\ :\ \mathbb{R}\longrightarrow\mathbb{R}\,,\ x\longmapsto\sum_{j=0}^{N}a_jx^{N-j}$$

mit Konstanten $a_0,\ldots,a_N\in\mathbb{R}$, $a_0\neq 0$, so bestimmt man

$$\hat{\mathbb{R}}-\lim_{x\to\infty}\frac{P(x)}{Q(x)}\ ,$$

indem man den Bruch $\frac{P(x)}{Q(x)}$ mit der höchsten im Nenner auftretenden Potenz kürzt.

Für die Grenzwertbildung bei geschachtelten Funktionen formulieren wir

Satz 2.27 *f, g seien reelle Funktionen und a ein (eventuell uneigentlicher) Häufungspunkt von $D(f\circ g)$. In $\hat{\mathbb{R}}$ existiere*

$$b:=\hat{\mathbb{R}}-\lim_{x\to a}g(x)$$

und sei ein (eventuell uneigentlicher) Häufungspunkt von $D(f)$. Mit $c\in\hat{\mathbb{R}}$ gelte

$$\hat{\mathbb{R}}-\lim_{y\to b}f(y)=c, \tag{2.37}$$

$$f(b)=c\quad \textit{im Fall}\ b\in D(f)\ . \tag{2.38}$$

Dann ist

$$\hat{\mathbb{R}}-\lim_{x\to a}f\circ g(x)=\hat{\mathbb{R}}-\lim_{y\to b}f(y)=c$$

Beweis: Die Grenzwertdefinition (2.17) zusammen mit der Bedingung (2.38) liefert

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{y\in U_\delta(b)\cap D(f)} f(y) \in U_\varepsilon(c) \ . \tag{2.39}$$

Unter dem letzten Allquantor kann also die Einschränkung $y \neq b$ entfallen. Ist nun $\varepsilon > 0$ vorgelegt, so findet man $\hat{\delta} > 0$ mit $f(y) \in U_\varepsilon(c)$ für $y \in U_{\hat{\delta}}(b) \cap D(f)$. Andererseits gibt es zu diesem $\hat{\delta}$ eine Zahl $\delta > 0$ mit $g(x) \in U_\delta(b)$, falls

$$x \in (U_\delta(a) \setminus \{a\}) \cap D(g) \ .$$

Für alle $x \in (U_\delta(a) \setminus \{a\}) \cap D(f \circ g) \subset (U_\delta(a) \setminus \{a\}) \cap D(g)$ gilt dann $f(g(x)) \in U_\varepsilon(c)$, und das war zu zeigen.

q.e.d.

Grundlegende Begriffsbildungen der Analysis, etwa Stetigkeit, Differenzierbarkeit verwenden den Grenzwertbegriff. Ist f reelle Funktion mit $b \in D(f)$, so ist (2.39) gerade die Definition der Stetigkeit von f im Punkte b. Wir werden darauf noch genauer eingehen, geben aber für die Stetigkeit einer reellen Funktion bereits die Definition — allerdings ohne genauere Erläuterung des Begriffs:

Definition 2.28 *Sei f eine reelle Funktion und $x_0 \in D(f)$. f heißt stetig in $x_0 \in D(f)$, wenn* (2.39) *gilt:*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{x\in U_\delta(x_0)\cap D(f)} f(x) \in U_\varepsilon(f(x_0)) \tag{2.40}$$

(2.40) *ist äquivalent zu*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{x\in D(f)} (\mid x - x_0 \mid < \delta \Rightarrow \mid f(x) - f(x_0) \mid < \varepsilon)$$

f ist also stetig in x_0, wenn

$$\lim_{x\to x_0} f(x) = f(x_0)$$

gilt, sofern x_0 Häufungspunkt von $D(f)$ ist, oder wenn x_0 kein Häufungspunkt von $D(f)$ ist. Insbesondere beinhaltet Satz 2.27 also, dass

$$\lim_{x\to a} f \circ g(x) = \lim_{y\to b} f(y)$$

ist, sofern $a \in \hat{\mathbb{R}}$ Häufungspunkt von $D(f \circ g)$ ist,

$$b := \lim_{x\to a} g(x)$$

existiert, in $D(f)$ liegt und f stetig in b ist.

Kapitel 3

Folgen

3.1 Grenzwerte von Folgen

Am leichtesten ist der Grenzwertbegriff zu verstehen, wenn man Folgen untersucht. Eine Folge $(a_n)_{n\in\mathbb{N}}$ (in $\mathbb{R}$) ist bekanntlich eine auf $\mathbb{N}$ definierte reelle Funktion, und nur ∞ ist Häufungspunkt in $\hat{\mathbb{R}}$ von $\mathbb{N}$. Für eine Folge läßt sich die Grenzwertdefinition auch in folgender Form aussprechen:

Satz 3.1 *Eine Folge $(a_n)_{n\in\mathbb{N}}$ kovergiert genau dann gegen einen Grenzwert $a \in \mathbb{R}$, wenn zu jedem $\varepsilon > 0$ ein Index $N \in \mathbb{N}$ existiert, so daß für alle höheren Indizes $n \in \mathbb{N}$, $n \geq N$, gilt:*

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

Mit Quantoren:

$$\bigwedge_{\varepsilon>0} \ \bigvee_{N\in\mathbb{N}} \ \bigwedge_{n\geq N, n\in\mathbb{N}} \ | a_n - a | < \varepsilon \ .$$

Dies folgt aus der Archimedischen Eigenschaft der reellen Zahlen, Satz 1.34, wonach zu jedem $\delta \in \mathbb{R}^+$ eine natürliche Zahl $N \in \mathbb{N}$ mit $U_{1/N}(\infty) \subset U_\delta(\infty)$ existiert.

Natürlich ist für die Folge $(n)_{n\in\mathbb{N}}$:

$$\hat{\mathbb{R}}\text{–} \lim_{n\to\infty} n = \infty$$

und daher liefert Satz 2.25:

$$\lim_{n\to\infty} \frac{1}{n} = 0 \ .$$

$(\frac{1}{n})_{n\in\mathbb{N}}$ ist also eine Nullfolge:

Definition 3.2 *Eine Nullfolge ist eine Folge $(b_n)_{n\in\mathbb{N}}$ mit*

$$\lim_{n\to\infty} b_n = 0 \; .$$

Wir erinnern an die Definition 1.31 des Supremums und des Infimums und die Erweiterung dieser Begriffe auf $\hat{\mathbb{R}}$ (Formeln (1.88) – (1.91)).

Definition 3.3 *Für eine Funktion $f : T \to \mathbb{R}$ ist*

$$\sup f := \sup_{x\in T} f(x) := \sup W(f) \quad , \quad \inf f := \inf_{x\in T} f(x) := \inf W(f) \; ,$$

verstanden als Elemente der erweiterten reellen Achse $\hat{\mathbb{R}}$. f heißt nach oben bzw. nach unten beschränkt, falls

$$\sup f < \infty \quad \textit{bzw.} \quad \inf f > -\infty$$

ist. f heißt beschränkt, wenn f sowohl nach unten als auch nach oben beschränkt ist.

Satz 3.4 *Ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge und $(b_n)_{n\in\mathbb{N}}$ eine beschränkte Folge, so ist $(a_n b_n)_{n\in\mathbb{N}}$ ebenfalls eine Nullfolge.*

Beweis: Da (b_n) beschränkt ist, gibt es eine Konstante $c > 0$, so daß

$$\bigwedge_{n\in\mathbb{N}} \quad | \; b_n \; | \leq c \; .$$

Für c kann man irgendeine positive Zahl wählen, welche

$$c \geq | \sup_{n\in\mathbb{N}} b_n \; | \quad \text{und} \quad c \geq | \inf_{n\in\mathbb{N}} b_n \; |$$

erfüllt. Die Abschätzung

$$\bigwedge_{n\in\mathbb{N}} \quad | \; a_n b_n \; | \leq c \; | \; a_n \; |$$

zeigt, daß man zu gegebenem $\varepsilon > 0$ den Index $N \in \mathbb{N}$ so wählen sollte, daß

$$| \; a_n \; | < \frac{\varepsilon}{c}$$

ist für Indizes $n \geq N$.

q.e.d.

Aus Satz 3.4 folgt auch, daß$(a_n)_{n\in\mathbb{N}}$ genau dann Nullfolge ist, wenn $(| \; a_n \; |)_{n\in\mathbb{N}}$ Nullfolge ist.

Offenbar gelten:

Lemma 3.5 *Für eine Folge $(a_n)_{n\in\mathbb{N}}$ und eine Zahl $a \in \mathbb{R}$ sind äquivalent*

(i) $\lim\limits_{n\to\infty} a_n = a$

(ii) *Es gibt eine Nullfolge $(b_n)_{n\in\mathbb{N}}$ mit*

$$\bigwedge_{n\in\mathbb{N}} \mid a_n - a \mid \leq b_n$$

(iii) *Es gibt Folgen $(\check{a}_n)_{n\in\mathbb{N}}$ und $(\hat{a}_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} \hat{a}_n = \lim\limits_{n\to\infty} \check{a}_n = a$ und*

$$\bigwedge_{n\in\mathbb{N}} \check{a}_n \leq a_n \leq \hat{a}_n$$

Lemma 3.6 *Eine Folge $(a_n)_{n\in\mathbb{N}}$ divergiert genau dann gegen ∞ (bzw. $-\infty$), wenn es eine gegen ∞ divergierende Folge (b_n) gibt mit*

$$\bigwedge_{n\in\mathbb{N}} a_n \geq b_n \quad (\textit{bzw.} \quad a_n \leq -b_n) \; .$$

Beweis (nur von Lemma 3.5**):** Wählt man $b_n :=\mid a_n - a \mid$, so erkennt man, daß (ii) aus (i) folgt.

Umgekehrt folgt auch (i) aus (ii). Denn für vorlegtes $\varepsilon > 0$ gibt es $N \in \mathbb{N}$ mit $\mid b_n \mid < \varepsilon$ für $n \geq N$. Dann folgt auch $\mid a_n - a \mid \leq b_n < \varepsilon$ für $n \geq N$.

Natürlich folgt wieder (iii) aus (i), indem man $\check{a}_n \equiv \hat{a}_n = a_n$ setzt für jedes $n \in \mathbb{N}$.

Umgekehrt liefert (iii) für jedes $n \in \mathbb{N}$

$$\mid a_n - a \mid \leq \mid \hat{a}_n - a \mid + \mid \check{a}_n - a \mid =: b_n \to 0 \; .$$

q.e.d.

Beispiel 3.7 *Mit Lemma* 3.5 *erkennt man, daß bei gegebenen $q \in (-1, 1)$ die Folge $(q^n)_{n\in\mathbb{N}}$ eine Nullfolge ist:*

Schreibe

$$\mid q \mid := (1 + h)^{-1}$$

mit

$$h := \frac{1}{\mid q \mid} - 1 > 0 \ .$$

Dann folgt unter Verwendung der Benoullischen Ungleichung (Satz 1.15*):*

$$\mid q^n \mid = \frac{1}{(1+h)^n} \leq \frac{1}{1+hn} \ ,$$

und $\left(\frac{1}{1+hn}\right)_{n\in\mathbb{N}}$ *ist Nullfolge, weil*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty}(1+hn) = 1 + \infty \cdot h = \infty$$

ist.

Beispiel 3.8 *Ist* $(a_n)_{n\in\mathbb{N}}$ *eine Folge, so kann man die Folge* $\left(\sum\limits_{j=1}^{n} a_j\right)_{n\in\mathbb{N}}$ *bilden. Man nennt diese die "Reihe der* (a_j)*". Zum Beispiel ist das vierte Glied von* $\left(\sum\limits_{j=1}^{n} a_j\right)_{n\in\mathbb{N}}$ *durch*

$$\sum_{j=1}^{4} a_j = a_1 + a_2 + a_3 + a_4$$

gegeben. Für $q \in \mathbb{R}$ *heißt* $\left(\sum\limits_{j=0}^{n} q^j\right)_{n\in\mathbb{N}}$ *die geometrische Reihe (zur Basis* q*).*

Aus der Formel (1.77) *erkennt man für* $q \neq 1$*:*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty} = \hat{\mathbb{R}} - \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \begin{cases} \infty & \textit{falls } q > 1 \\ \frac{1}{1-q} \ , & \textit{falls } q \in (\,-1\,,\,1\,) \end{cases} \tag{3.1}$$

Ist $q \leq -1$*, so existiert* $\hat{\mathbb{R}}\text{–}\lim\limits_{n\to\infty} \sum_{j=0}^{n} q^j$ *nicht, und ist* $q = 1$*, so folgt*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty} \sum_{j=0}^{n} 1^j = \hat{\mathbb{R}}\text{–}\lim_{n\to\infty}(n+1) = \infty \ .$$

Es ist üblich, mit $\sum\limits_{j=1}^{n} a_j$ sowohl die Folge $\left(\sum\limits_{j=1}^{n} a_j\right)_{n\in\mathbb{N}}$ als auch den eventuell existierenden $\hat{\mathbb{R}}$-Grenzwert zu bezeichnen, je nachdem in welchem Kontext das

Symbol verwendet wird. Oben haben wir gezeigt, dass $\sum\limits_{j=0}^{\infty} q^j$ keinen Grenzwert in $\hat{\mathbb{R}}$ besitzt, falls $q \leq -1$ ist. Hier steht $\sum\limits_{j=0}^{\infty} q^j$ für die Folge $\left(\sum\limits_{j=0}^{n} q^j\right)_{n\in\mathbb{N}}$. In der Formel

$$\sum_{j=0}^{\infty} q^j = \begin{cases} \infty & \text{falls } q \geq 1 \\ \frac{1}{1-q} & \text{falls } q \in (-1\,,\,1) \end{cases}$$

steht $\sum\limits_{j=0}^{\infty} q^j$ für den $\hat{\mathbb{R}}$–Grenzwert.

3.2 Grenzwerte monotoner Folgen

Bei den bisherigen Konvergenznachweisen war es nötig, den Grenzwert bereits zu ahnen. Wir benötigen aber auch Kriterien, für die Konvergenz von Folgen, die ohne eine Kenntnis über den Grenzwert greifen. Am einfachsten ist das für monotone Folgen.

Definition 3.9 *Eine reelle Funktion f heißt*

- *monoton wachsend, wenn $f(s) \geq f(t)$ ist für alle $s\,,\,t \in D(f)$ mit $s > t$.*
- *streng monoton wachsend, wenn $f(s) > f(t)$ ist für alle $s\,,\,t \in D(f)$ mit $s > t$.*
- *monoton fallend bzw. streng monoton fallend, wenn $f(s) \leq f(t)$ bzw. $f(s) < f(t)$ ist für alle $s\,,\,t \in D(f)$ mit $s > t$.*

Eine Folge ist monoton wachsend (streng monoton wachsend, monoton fallend bzw. streng monoton fallend), wenn

$$\bigwedge_{n\in\mathbb{N}} a_n \leq a_{n+1} \qquad (a_n < a_{n+1}\,, \quad a_n \geq a_{n+1} \text{ bzw. } a_n > a_{n+1})$$

ist, wie man leicht mit vollständiger Induktion beweist.

Jede monotone Folge hat einen Grenzwert in $\hat{\mathbb{R}}$:

Satz 3.10 *Für eine monoton wachsende bzw. fallende Folge $(a_n)_{n\in\mathbb{N}}$ ist*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \qquad (\textit{bzw.} \quad \hat{\mathbb{R}}\text{–}\lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n)$$

Insbesondere sind beschränkte monotone Folgen konvergent.

Beweis: Wir beschränken uns auf den Fall einer monoton wachsenden, beschränkten Folge. Die übrigen Fälle beweise der Leser analog. Setze $a := \sup\limits_{n\in\mathbb{N}} a_n$. Ist $\varepsilon > 0$ vorgelegt, so existiert ein Index $N \in \mathbb{N}$ mit

$$a - \varepsilon < a_N \leq a\;;.$$

Da (a_n) monoton wächst und a obere Schranke ist, erhält man für alle Indizes $n \geq N$

$$a - \varepsilon < a_N \leq a_n \leq a\;.$$

Also konvergiert (a_n) gegen a.

q.e.d.

Beispiel 3.11 *Die Folge* $(a_n)_{n\in\mathbb{N}} := \left(\frac{1}{\sqrt{n}}\right)_{n\in\mathbb{N}}$ *ist streng monoton fallend, denn für jedes* $n \in \mathbb{N}$ *ist*

$$\sqrt{n+1} - \sqrt{n} = \frac{\sqrt{n+1}^2 - \sqrt{n}^2}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} > 0$$

und daher

$$a_n = \frac{1}{\sqrt{n}} > \frac{1}{\sqrt{n+1}} = a_{n+1}\;.$$

Da $a_n \geq 0$ *ist für alle* $n \in \mathbb{N}$, *ist* $(a_n)_{n\in\mathbb{N}}$ *konvergent, etwa gegen* $a \in \mathbb{R}$.

Es folgt aus Satz 2.20*:*

$$a^2 = (\lim_{n\to\infty} a_n) \cdot (\lim_{n\to\infty} a_n) = \lim_{n\to\infty} a_n^2 = \lim_{n\to\infty} \frac{1}{n} = 0\;.$$

Daher ist $a = 0$*; die Folge* $\left(\frac{1}{\sqrt{n}}\right)_{n\in\mathbb{N}}$ *ist eine Nullfolge.*

Beispiel 3.12 *Ist* $(a_n)_{n\in\mathbb{N}}$ *eine Folge nicht negativer Zahlen, so ist die Reihe der* (a_n) *eine monton wachsende Folge. Unter der harmonischen Reihe versteht man die Folge* $\left(\sum\limits_{j=1}^{n} \frac{1}{j}\right)$ *(oder in der in Beispiel* 3.8 *eingeführten Sprechweise:* $\sum\limits_{j=1}^{\infty} \frac{1}{j}$ *ist die harmonische Reihe). Für diese Reihe gilt:*

$$\sum_{j=1}^{\infty} \frac{1}{j} = \hat{\mathbb{R}}\text{-}\lim_{n\to\infty} \sum_{j=1}^{n} \frac{1}{j} = \infty\;. \tag{3.2}$$

Es muss nachgewiesen werden, dass die Reihe $\sum_{j=1}^{\infty} \frac{1}{j}$ *unbeschränkt ist. Die Idee erkennt man, wenn man sich* $\sum_{j=1}^{n} \frac{1}{j}$ *für ein großes* n *ausgeschrieben denkt:*

$$\sum_{j=1}^{n} \frac{1}{j} = \underbrace{1}_{=1} + \underbrace{\frac{1}{2}}_{=1/2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 2\cdot\frac{1}{4}=1/2} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq 4\cdot\frac{1}{8}=1/2}$$
$$+ \underbrace{\frac{1}{9} + \ldots + \frac{1}{16}}_{\geq 8\cdot\frac{1}{16}=1/2} + \underbrace{\frac{1}{17} + \ldots \frac{1}{32}}_{\geq 16\cdot\frac{1}{32}=1/2} + \ldots$$

Wir fassen die unterklammerten Terme zusammen und schätzen jeden der Summanden über einer Klammer durch den letzten der Summanden nach unten ab wie oben beschrieben. Genau formuliert: Für alle $m \in \mathbb{N}$ *ist*

$$\begin{aligned} \sum_{j=1}^{2^m} \frac{1}{j} &= 1 + \sum_{k=1}^{m} \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{j} \geq 1 + \sum_{k=1}^{m} \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{2^k} \\ &\overset{1}{=} 1 + \sum_{k=1}^{m} (2^{k-1} \cdot \tfrac{1}{2^k}) = 1 + \tfrac{m}{2} \end{aligned}$$

Hieraus erkennt man (3.2).

Beispiel 3.13 *Mit einer vorgegebenen Zahl* $\delta > 0$ *gilt:*

$$\sum_{j=1}^{\infty} \left(\frac{1}{j}\right)^{1+\delta} < \infty$$

Die Reihe $\sum\limits_{j=1}^{\infty} \left(\frac{1}{j}\right)^{1+\delta}$ *konvergiert also.*

Zum Nachweis ordnen wir die Summanden etwas anders als in Beispiel 3.12 *und schätzen die unterklammerten Summanden jeweils durch ihren ersten nach oben ab:*

$$\sum_{j=1}^{n} \left(\frac{1}{j}\right)^{1+\delta} = \underbrace{1}_{\leq 1} + \underbrace{\left(\frac{1}{2}\right)^{1+\delta} + \left(\frac{1}{3}\right)^{1+\delta}}_{\leq 2\cdot\left(\frac{1}{2}\right)^{1+\delta}} + \underbrace{\left(\frac{1}{4}\right)^{1+\delta} + \ldots + \left(\frac{1}{7}\right)^{1+\delta}}_{\leq 4\cdot\left(\frac{1}{4}\right)^{1+\delta}}$$
$$+ \underbrace{\left(\frac{1}{8}\right)^{1+\delta} + \ldots + \left(\frac{1}{15}\right)^{1+\delta}}_{\leq 8\cdot\left(\frac{1}{8}\right)^{1+\delta}} + \left(\frac{1}{16}\right)^{1+\delta} \ldots$$

[1]Denn die innere Summe hat 2^{k-1} Summanden.

Genauer: Für jedes $n \in \mathbb{N}$ gilt

$$\begin{aligned}
\sum_{j=1}^{n} \left(\frac{1}{j}\right)^{1+\delta} &\leq \sum_{j=1}^{2^n-1} \left(\frac{1}{j}\right)^{1+\delta} = \sum_{k=0}^{n-1} \sum_{j=2^k}^{2^{k+1}-1} \left(\frac{1}{j}\right)^{1+\delta} \\
&\leq \sum_{k=0}^{n-1} \sum_{j=2^k}^{2^{k+1}-1} \left(\tfrac{1}{2^k}\right)^{1+\delta} = \sum_{k=0}^{n-1} 2^k \cdot \left(\tfrac{1}{2^k}\right)^{1+\delta} \\
&= \sum_{k=0}^{n-1} \left(\tfrac{1}{2^\delta}\right)^k \overset{2}{\leq} \frac{1}{1-\frac{1}{2^\delta}} = \frac{2^\delta}{2^\delta - 1}
\end{aligned}$$

3.3 Teilfolgen und Häufungspunkte von Folgen

Eine wesentliche Rolle spielt der Begriff der Teilfolge einer Folge. Man stelle sich die Folgenglieder einer Folge $(a_n)_{n\in\mathbb{N}}$ in der durch die Abbildungsvorschrift gegebenen Reihenfolge aufgeschrieben vor:

$$a_1; a_2; a_3; a_4; \ldots; a_n, \ldots$$

Durchläuft man diese Zahlenschnur von links nach rechts und markiert hin und wieder ein Folgenglied, so bilden die markierten Glieder eine neue Folge, sofern man nicht nach endlich vielen Markierungen seinen Durchlauf beendet. Die so erhaltene Folge nennt man eine Teilfolge. Präzise:

Definition 3.14 *Eine Teilfolge einer Folge $(a_n)_{n\in\mathbb{N}}$ ist die Hintereinanderschaltung von a_n mit einer streng monoton wachsenden Abbildung von $\mathbb{N}$ in sich.*

Ist also

$$\pi : \mathbb{N} \to \mathbb{N}$$

streng monoton wachsend – die "Auswählfunktion" –, so ist

$$\left(a_{\pi(n)}\right)_{n\in\mathbb{N}}$$

eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$. Man beachte, dass für jede Auswählfunktion π gilt:

$$\bigwedge_{n\in\mathbb{N}} \pi(n) \geq n \; . \tag{3.3}$$

Das erhält man mit einer leichten Induktion aus der strengen Monotonie von π.

[2] wegen (3.1)

Beispiel 3.15

(i) $\pi : \mathbb{N} \longrightarrow \mathbb{N}$, $n \longmapsto 2n - 1$ *wählt aus einer Folge alle Glieder mit ungeraden Indizes aus. Das dritte Glied der Teilfolge* $\left(a_{\pi(n)}\right)_{n\in\mathbb{N}}$ *ist also* $a_{\pi(3)} = a_5$.

(ii) *Die Folge* $\left(\frac{1}{n^2}\right)_{n\in\mathbb{N}}$ *ist die Teilfolge der Folge* $(\frac{1}{n})_{n\in\mathbb{N}}$, *die durch die durch* $\pi(n) := n^2$ *gegebene Auswählfunktion entsteht.*

(iii) *Die mittels* $\pi(n) = n + 28$ *ausgewählte Teilfolge von* (a_n) *entsteht durch Streichung der ersten* 28 *Folgeglieder von* $(a_n)_{n\in\mathbb{N}}$.

Aus Satz 3.1 folgt direkt — für den ersten Teil des Satzes beachte man (3.3) — :

Satz 3.16 *Sei* $(a_n)_{n\in\mathbb{N}}$ *eine Folge.*

(i) *Existiert* $a := \hat{\mathbb{R}}\text{–}\lim_{n\to\infty} a_n$, *so ist auch*

$$\hat{\mathbb{R}}\text{–}\lim a_{\pi(n)} = a$$

für jede Teilfolge $\left(a_{\pi(n)}\right)_{n\in\mathbb{N}}$ *von* $(a_n)_{n\in\mathbb{N}}$.

(ii) *Ist mit einer Zahl* $k \in \mathbb{N}$ *die Auswählfunktion* π *durch* $\pi(n) = n + k$ *gegeben und existiert* $a := \hat{\mathbb{R}} - \lim_{n\to\infty} a_{\pi(n)}$, *so gilt*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty} a_n = a$$

auch für die ursprüngliche Folge (a_n).

Beispiel 3.17 *Mit* $p \in \mathbb{R}^+$ *sei* $(a_n)_{n\in\mathbb{N}}$ *durch*

$$a_n := \frac{p^n}{n!}$$

definiert. Dann ist $a_n > 0$ *für jedes* $n \in \mathbb{N}$ *und*

$$\frac{a_{n+1}}{a_n} = \frac{p}{n+1} .$$

Ist

$$k := \max\{j \in \mathbb{N}\colon\ j \le p\} ,$$

so folgt für $n > k$

$$\frac{a_{n+1}}{a_n} = \frac{p}{n+1} \leq \frac{k+1}{n+1} < 1 \; .$$

Die Teilfolge $(a_{n+k})_{n\in\mathbb{N}}$ ist also monton fallend und beschränkt und konvergiert nach Satz 3.10. Satz 3.16(ii) liefert, dass (a_n) konvergiert. Bezeichnen wir den Grenzwert mit a, so konvergiert auch $(a_{n+1})_{n\in\mathbb{N}}$ gegen a. Andererseits ist $a_{n+1} = \frac{p}{n+1}a_n$ für jedes $n \in \mathbb{N}$. Mit Satz 2.22 folgt

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \left(\frac{p}{n+1}a_n\right) = 0 \cdot a = 0 \; .$$

Insgesamt ist also

$$\lim_{n\to\infty} \frac{p^n}{n!} = 0 \; .$$

Etwas anders als für Teilmengen von $\mathbb{R}$ (Punktmengen) definiert man Häufungspunkte von Folgen:

Definition 3.18 *Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge, so heißt $a \in \hat{\mathbb{R}}$ ein Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$, wenn gilt*

$$\bigwedge_{\varepsilon\in\mathbb{R}^+} \; \bigwedge_{m\in\mathbb{N}} \; \bigvee_{n\in\mathbb{N}} \; n \geq m \wedge a_n \in U_\varepsilon(a) \; . \tag{3.4}$$

Liegt a in $\mathbb{R}$, so nennt man a Häufungspunkt; ist $a \in \{-\infty, \infty\}$ so nennt man a uneigentlichen Häufungspunkt.

Definition 3.18 besagt, dass in jeder ε–Umgebung von a unendlich viele Folgenglieder liegen (d. h. Folgenglieder zu unendlich vielen Indizes).

Man beachte: ein $a \in \mathbb{R}$ kann Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$ sein, ohne gleichzeitig Häufungspunkt des Wertebereichs

$$W := \{a_n\colon n \in \mathbb{N}\}$$

der Folge zu sein: z. B. $((-1)^n)_{n\in\mathbb{N}}$ hat die Zahlen -1 und 1 als Häufungspunkte, aber der Wertebereich ist $W = \{-1, 1\}$ und hat keinen Häufungspunkt.

Lemma 3.19 *$a \in \hat{\mathbb{R}}$ ist genau dann Häufungspunkt in $\hat{\mathbb{R}}$ der Folge $(a_n)_{n\in\mathbb{N}}$, wenn es eine Teilfolge $\left(a_{\pi(n)}\right)_{n\in\mathbb{N}}$ gibt mit*

$$\hat{\mathbb{R}}\text{–}\lim_{n\to\infty} a_{\pi(n)} = a \; .$$

Beweis: Wenn es eine in $\hat{\mathbb{R}}$ gegen a strebende Teilfolge $a_{\pi(n)}$ gibt, so zeigen wir, dass a Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$ ist:

Sind $\varepsilon > 0$ und $m \in \mathbb{N}$ vorgelegt, so gibt es zunächst einmal eine Zahl $N \in \mathbb{N}$ mit

$$\bigwedge_{k\geq N} a_{\pi(k)} \in U_\varepsilon(a)\ .$$

Wähle nun $k \in \mathbb{N}$ so, dass

$$k \geq \max\{N, m\};$$

dann ist $a_n \in U_\varepsilon(a)$ für

$$n := \pi(k) \geq k \geq m\ .$$

(beachte (3.3)).

Ist nun umgekehrt a ein Häufungspunkt in $\hat{\mathbb{R}}$ von a, so wollen wir eine Teilfolge $(a_{\pi(n)})_{n\in\mathbb{N}}$ konstruieren, die gegen a strebt. Konstruktionsidee: Wähle für ε sukzessive die Glieder einer Nullfolge, etwa $(1/n)_{n\in\mathbb{N}}$ und finde zu jedem n eine natürliche Zahl $\pi(n)$ mit $a_{\pi(n)} \in U_{1/n}(a)$. Man muss dabei nur darauf achten, dass π streng monoton wächst! Die Funktion π wird mit dem Rekursionsprinzip definiert:

Zu $\varepsilon = 1$ gibt es eine Zahl $\pi(1) \in \mathbb{N}$ mit $a_{\pi(1)} \in U_1(a)$, etwa

$$\pi(1) := \min\{k \in \mathbb{N}\colon a_k \in U_1(a)\}\ .$$

Dies wird durch (3.4) garantiert.

Wir nehmen an, wir hätten zu einem $n \in \mathbb{N}$ ein $\pi(n)$ mit

$$a_{\pi(n)} \in U_{1/n}(a)$$

gefunden. Dann gibt es nach (3.4) eine Zahl $\pi(n+1) \in \mathbb{N}$ mit

$$\pi(n+1) \geq \pi(n) + 1 \quad \text{und} \quad a_{\pi(n+1)} \in U_{1/(n+1)}(a)\ ,$$

etwa

$$\pi(n+1) := \min\{k \in \mathbb{N}\colon k \geq \pi(n) + 1 \ \wedge\ a_k \in U_{1/(n+1)}(a)\}\ .$$

Die so definierte Funktion $\pi : \mathbb{N} \to \mathbb{N}$ ist streng monoton wachsend und

$$\bigwedge_{n\in\mathbb{N}} a_{\pi(n)} \in U_{1/n}(a)\ . \tag{3.5}$$

Ist nun $a \in \mathbb{R}$, so ist $\lim\limits_{n\to\infty} a_{\pi(n)} = a$, denn

$$\bigwedge_{n\in\mathbb{N}} \quad | a_{\pi(n)} - a |< 1/n \to 0 \; .$$

Ist $a = \infty$ bzw. $a = -\infty$, so ist $a_{\pi(n)} > n$ bzw. $a_{\pi(n)} < -n$ für alle $n \in \mathbb{N}$ und daher $\hat{\mathbb{R}}\text{--}\lim\limits_{n\to\infty} a_{\pi(n)} = a$.

q.e.d.

Satz 3.20 *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und*

$$H := \{a \in \mathbb{R} \colon a \text{ ist Häufungspunkt von } (a_n)\} \; . \tag{3.6}$$

Ist $H \neq \emptyset$, so sind $\sup H$ *und* $\inf H$ *Häufungspunkte in $\hat{\mathbb{R}}$ von (a_n).*

Beweis: Sei $H \neq \emptyset$. Dann ist $b := \inf H < \infty$. Zu gegebenem $\varepsilon > 0$ existiert ein $h \in H$ welcher in $U_\varepsilon(b)$ liegt. Fixiere nun $\delta > 0$ so, dass $U_\delta(h) \subset U_\varepsilon(b)$ ist. Ist nun $m \in \mathbb{N}$ vorgegeben, so existiert ein Index $n > m$ mit $a_n \in U_\delta(h) \subset U_\varepsilon(b)$, denn h ist ein Häufungspunkt der Folge (a_n).

q.e.d.

Nun definieren wir $\liminf\limits_{n\to\infty} a_n$ bzw. $\limsup\limits_{n\to\infty} a_n$ als kleinsten bzw. größten Häufungspunkt in $\hat{\mathbb{R}}$ von (a_n):

Definition 3.21 *Seien $(a_n)_{n\in\mathbb{N}}$ eine Folge, $H \subset \mathbb{R}$ die Menge der Häufungspunkte von (a_n) (wie in (3.6)) und $\hat{H} \subset \hat{\mathbb{R}}$ Die Menge der Häufungspunkte in $\hat{\mathbb{R}}$ von (a_n). Dann heißen*

$$\liminf_{n\to\infty} a_n := \begin{cases} -\infty \, , & \text{falls } -\infty \in \hat{H} \\ \inf H \, , & \text{falls } H \neq \emptyset \text{ und } \inf H > -\infty \\ \infty & \text{falls } \hat{H} = \{\infty\} \end{cases} \tag{3.7}$$

$$\limsup_{n\to\infty} a_n := \begin{cases} \infty \, , & \text{falls } \infty \in \hat{H} \\ \sup H \, , & \text{falls } H \neq \emptyset \text{ und } \sup H < \infty \\ -\infty & \text{falls } \hat{H} = \{-\infty\} \end{cases} \tag{3.8}$$

Limes inferior bzw. Limes superior der Folge (a_n).

Nach einem fundamentalen Resultat der Analysis, dem Satz von Bolzano und Weierstraß, besitzt jede Folge einen Häufungspunkt in $\hat{\mathbb{R}}$. Dies ist äquivalent mit der Aussage, dass Limes inferior und Limes superior einer jeden Folge $(a_n)_{n\in\mathbb{N}}$ existieren. Um den Beweis dieses Satzes vorzubereiten, wird eine alternative Möglichkeit zur Bestimmung von Limes superior und inferior angegeben.

Lemma 3.22 *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und für jedes $n \in \mathbb{N}$*

$$b_n := \inf\{a_j\colon j \geq n\} \quad , \qquad c_n := \sup\{a_j\colon j \geq n\} \ . \tag{3.9}$$

Dann gelten (i) oder (ii) bzw. (iii) oder (iv):

(i) $\bigwedge_{n\in\mathbb{N}} \quad b_n \in \mathbb{R}.$
In diesem Fall existiert $\hat{\mathbb{R}}-\lim_{n\to\infty} b_n$ und

$$\liminf_{n\to\infty} a_n = \hat{\mathbb{R}}-\lim_{n\to\infty} b_n$$

(ii) $\bigwedge_{n\in\mathbb{N}} \quad b_n = -\infty$
In diesem Fall gilt

$$\liminf_{n\to\infty} a_n = -\infty \ .$$

(iii) $\bigwedge_{n\in\mathbb{N}} \quad c_n \in \mathbb{R}$
In diesem Fall existiert $\hat{\mathbb{R}} - \lim_{n\to\infty} c_n$ und

$$\limsup_{n\to\infty} a_n = \hat{\mathbb{R}}-\lim_{n\to\infty} c_n$$

(iv) $\bigwedge_{n\in\mathbb{N}} \quad c_n = \infty$
In diesem Fall gilt

$$\limsup_{n\to\infty} a_n = \infty \ .$$

Beweis: Wir führen den Beweis nur für die Alternative "(i)$\vee$(ii)", da die andere völlig analog bewiesen wird.

Mit

$$A_n := \{a_j\colon j \geq n\}$$

folgt für jedes $n \in \mathbb{N}$

$$\emptyset \neq A_{n+1} \subset A_n \ .$$

Dann liefert (1.92)

$$b_n = \inf A_n \leq \inf A_{n+1} = b_{n+1} < \infty \tag{3.10}$$

Gilt $b_k = -\infty$ für einen Index $k \in \mathbb{N}$, dann folgt aus (3.10) für Indizes $n < k$, dass $b_n = -\infty$ ist. Für $n > k$ ist $-\infty = b_k = \min\{a_k, a_{k+1}, \ldots, a_{n-1}, b_n\}$. Da $a_k, \ldots, a_{n-1} \in \mathbb{R}$ sind, muss $b_n = -\infty$ sein. Ist also $b_k = -\infty$ für ein $k \in \mathbb{N}$, so auch für alle weiteren $k \in \mathbb{N}$, und wir sind im Fall (ii). Sind in diesem Fall $\varepsilon > 0$ und $m \in \mathbb{N}$ vorgelegt, so gibt es wegen $b_m = \inf A_m$ einen Index $n \geq m$ mit

$$a_n \in U_\varepsilon(-\infty),$$

d. h. $-\infty$ ist uneigentlicher Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ und gemäß (3.7) gilt

$$\liminf_{n\to\infty} a_n = -\infty \ . \tag{3.11}$$

Andererseits folgt aus (3.10), dass

$$\inf_{n\in\mathbb{N}} a_n \leq \liminf_{n\to\infty} a_n = -\infty$$

ist, also $b_1 = -\infty$. Wie oben folgt, dass wir dann im Fall (ii) sind.

Gilt $b_n \in \mathbb{R}$ für alle $n \in \mathbb{N}$, so ist (b_n) eine monton wachsende Folge und

$$b := \hat{\mathbb{R}} - \lim_{n\to\infty} b_n = \sup_{n\in\mathbb{N}} b_n > -\infty$$

existiert wegen Satz 3.10.

Wir zeigen, dass b Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$ ist. Sind $\varepsilon > 0$ und $m \in \mathbb{N}$ vorgelegt, so gibt es eine Zahl $k \in \mathbb{N}$ mit $k > m$ und

$$b_k \in U_\varepsilon(b) \ .$$

Insbesondere ist $b_k \leq b$. Wegen $b_k = \inf A_k$ existiert ein Index $n \geq k$ mit

$$b_k \leq a_n < b_k + \varepsilon \ , \text{ also } a_n \in U_\varepsilon(b)$$

Mithin ist b Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$.

b ist kleinster Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$. Gäbe es einen $\hat{\mathbb{R}}$–Häufungspunkt $\hat{b}$ von $(a_n)_{n\in\mathbb{N}}$ mit $\hat{b} < b$, so wäre – wir sind in Fall (i) – nach dem ersten Teil des Beweises $\hat{b} \neq -\infty$. Wegen $\hat{b} < b$ folgt $\hat{b} \in \mathbb{R}$. Es gibt ein $\varepsilon \in \mathbb{R}^+$ mit

$$\hat{b} + \varepsilon < b \ .$$

Dann gibt es $\delta > 0$, so dass $U_\varepsilon(\hat{b}) \cap U_\delta(b) = \emptyset$ ist. Zu diesem δ findet man $m \in \mathbb{N}$ mit

$$b_m \in U_\delta(b) \quad , \qquad b_m \leq b \ .$$

Für alle $n \geq m$ ist dann

$$\hat{b} + \varepsilon < b_m \leq a_n \ .$$

Wir haben also gesehen

$$\bigvee_{\varepsilon \in \mathbb{R}^+} \ \bigvee_{m \in \mathbb{N}} \ \bigwedge_{n \geq m} a_n \notin U_\varepsilon(\hat{b}) \ .$$

Dies heißt, dass $\hat{b}$ nicht Häufungspunkt von (a_n) war.

Insgesamt haben wir also gezeigt: Entweder tritt Fall (i) oder Fall (ii) ein. Im Fall (ii) ist $-\infty$ uneigentlicher Häufungspunkt von (a_n). Im Fall (i) ist (mit H wie in Definition 3.21)

$$b = \inf H > -\infty,$$

und dieses Infimum ist ∞, falls $H = \emptyset$ ist. In jedem Fall folgt

$$b = \liminf_{n \to \infty} a_n \ .$$

q.e.d.

Wir ziehen aus Lemma 3.22 einige Schlüsse:

Korollar 3.23 *Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge, so gilt für jeden Häufungspunkt a von $(a_n)_{n\in\mathbb{N}}$*

$$\inf_{n\in\mathbb{N}} a_n \leq \liminf_{n\to\infty} a_n \leq a \leq \limsup_{n\to\infty} a_n \leq \sup_{n\in\mathbb{N}} a_n$$

Denn mit den Folgen (b_n) und (c_n) aus dem Beweis von Lemma 3.22 ist

$$\begin{aligned} \inf_{n\in\mathbb{N}} a_n = b_1 \leq \sup_{n\in\mathbb{N}} b_n = \liminf_{n\in\mathbb{N}} a_n \\ \limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} c_n \leq c_1 = \sup_{n\in\mathbb{N}} a_n \ , \end{aligned}$$

und $\liminf\limits_{n\to\infty} a_n$ bzw. $\limsup\limits_{n\to\infty} a_n$ sind kleinster bzw. größter Häufungspunkt in $\hat{\mathbb{R}}$.

Satz 3.24 *(von Bolzano-Weierstraß)*
Jede Folge $(a_n)_{n\in\mathbb{N}}$ besitzt wenigstens einen Häufungspunkt in $\hat{\mathbb{R}}$. Insbesondere enthält jede beschränkte Folge eine konvergente Teilfolge.

Beweis: $\liminf_{n\to\infty} a_n$ und $\limsup_{n\to\infty} a_n$ sind Häufungspunkte in $\hat{\mathbb{R}}$ von (a_n). Nach Korollar 3.23 besitzt eine beschränkte Folge keinen uneigentlichen Häufungspunkt; eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ aber, die gegen einen Häufungspunkt $a \in \mathbb{R}$ strebt, ist konvergent.

q.e.d.

Eine in der Analysis wesentliche Begriffsbildung steht im Zusammenhang mit dem Satz von Bolzano – Weierstraß:

Definition 3.25 *Eine Teilmenge $M \subset \mathbb{R}$ heißt kompakt, wenn jede Folge in M eine konvergente Teilfolge mit Grenzwert in M besitzt.*

Eigentlich spricht man in diesem Fall von "Folgenkompaktheit" im Gegensatz zu "Überdeckungskompaktheit". Letzteren Begriff werden wir später noch kennenlernen. Beide Begriffe werden auch in allgemeineren Situationen definiert. Für Teilmengen von $\mathbb{R}$ sind sie aber äquivalent.

Definition 3.26 *Eine Teilmenge $M \subset \mathbb{R}$ heißt abgeschlossen, wenn jede Häufungspunkt von M selbst zu M gehört.*

Beachten Sie: Da Kompaktheit und Abgeschlossenheit von M durch eine Allaussage über M definiert sind, ist insbesondere die leere Menge sowohl kompakt als auch abgeschlossen.

Der Satz von Bolzano – Weierstraß läßt sich nun auch in folgender Formulierung enthalten:

Satz 3.27 *$M \subset \mathbb{R}$ ist genau dann kompakt, wenn M beschränkt und abgeschlossen ist.*

Als weiteres Konvergenzkriterium für Folgen halten wir fest:

Satz 3.28 *Eine Folge (a_n) konvergiert genau dann in $\hat{\mathbb{R}}$, wenn*

$$\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n \ .$$

Dieser gemeinsame Wert ist dann der Grenzwert von (a_n) in $\hat{\mathbb{R}}$.

Beweis: Wir setzen $a := \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$ und setzen wie im Beweis von Lemma 3.22:

$$b_n := \inf\{a_k\colon k \geq n\} \quad , \quad c_n := \sup\{a_k\colon k \geq n\} \ .$$

Dann gilt für alle $n \in \mathbb{N}$

$$b_n \leq a_n \leq c_n \ . \tag{3.12}$$

Ist a endlich, also $a \in \mathbb{R}$, so folgt mit Lemma 3.5

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

denn $b_n, c_n \to a$.

Ist $a = \infty$ bzw. $a = -\infty$, so divergiert $(b_n)_{n\in\mathbb{N}}$ gegen ∞ bzw. $(c_n)_{n\in\mathbb{N}}$ gegen $-\infty$, und die Behauptung folgt wegen (3.12) unter Verwendung von Lemma 3.6.

q.e.d.

Man kann auf folgende Weise aus der Konvergenz von Teilfolgen auf die Konvergenz der ursprünglichen Folge schließen:

Satz 3.29 *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $a \in \hat{\mathbb{R}}$. Wenn jede Teilfolge $(a_{\pi(n)})_{n\in\mathbb{N}}$ eine Teilfolge $(a_{\pi(\mu(n))})_{n\in\mathbb{N}}$ enthält, die in $\hat{\mathbb{R}}$ gegen a strebt, so gilt für die ursprüngliche Folge*

$$\hat{\mathbb{R}} - \lim_{n\to\infty} a_n = a \ .$$

Beweis: Die Folge $(a_n)_{n\in\mathbb{N}}$ hat a als einzigen $\hat{\mathbb{R}}$–Häufungspunkt. Ist nämlich $\tilde{a}$ ein Häufungspunkt in $\hat{\mathbb{R}}$ von $(a_n)_{n\in\mathbb{N}}$, so strebt eine Teilfolge $(a_{\pi(n)})_{n\in\mathbb{N}}$ gegen $\tilde{a}$. Damit strebt aber jede Teilfolge von $(a_{\pi(n)})_{n\in\mathbb{N}}$ gegen $\tilde{a}$. Eine dieser Teilfolgen aber strebt gegen a. Damit muss $\tilde{a} = a$ sein.

Andererseits ist a auch Häufungspunkt in $\hat{\mathbb{R}}$ von (a_n): es gibt ja eine gegen a strebende Teilfolge von (a_n).

Wir folgern

$$a = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n,$$

und mit Satz 3.28

$$\hat{\mathbb{R}} - \lim_{n\to\infty} a_n = a \ .$$

q.e.d.

Eine weitere Methode, die Konvergenz von Folgen nachzuweisen, liefert das Konzept der Cauchy-Folge (Augustin Louis Cauchy, 1789-1857)

Definition 3.30 *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *heißt Cauchyfolge, wenn*

$$\bigwedge_{\varepsilon\in\mathbb{R}^+} \bigvee_{N\in\mathbb{N}} \bigwedge_{\substack{n,m\in\mathbb{N}\\ n,m\geq N}} \mid a_n - a_m \mid < \varepsilon \tag{3.13}$$

Bei vorgelegtem ε unterscheiden sich die Glieder zu genügend großen Indizes also um weniger als ε.

Äquivalente Formulierungen liefert folgender Satz

Satz 3.31 *Für eine Folge* $(a_n)_{n\in\mathbb{N}}$ *sind äquivalent*

(i) $(a_n)_{n\in\mathbb{N}}$ *ist Cauchyfolge.*

(ii) *Es gilt*

$$\bigwedge_{\varepsilon\in\mathbb{R}^+} \bigvee_{N\in\mathbb{N}} \bigwedge_{\substack{n\geq N\\ n\in\mathbb{N}}} \bigwedge_{k\in\mathbb{N}} \mid a_{n+k} - a_n \mid < \varepsilon\ . \tag{3.14}$$

(iii) *Es gilt*

$$\bigwedge_{\varepsilon\in\mathbb{R}^+} \bigvee_{N\in\mathbb{N}} \bigwedge_{l\in\mathbb{N}} \mid a_{N+l} - a_N \mid < \varepsilon\ . \tag{3.15}$$

Beweis: Die Äquivalenz von (i) und (ii) sowie die Implikation (ii) $\Rightarrow$ (iii) sind unmittelbar klar. Wir beweisen die Implikation (iii) $\Rightarrow$ (ii):

Ist $\varepsilon > 0$ vorgelegt, so existiert $N \in \mathbb{N}$ derart, dass

$$\bigwedge_{l\in\mathbb{N}} \mid a_{N+l} - a_N \mid < \varepsilon/2\ .$$

Dann folgt für alle $n \in \mathbb{N}$ mit $n \geq N$ und $k \in \mathbb{N}$:

$$\begin{aligned} &\mid a_{n+k} - a_n \mid \leq \mid a_{n+k} - a_N \mid + \mid a_n - a_N \mid \\ &= \mid a_{N+(n+k-N)} - a_N \mid + \mid a_{N+(n-N)} - a_N \mid < \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

denn $n+k-N \in \mathbb{N}$, $n-N \in \mathbb{N} \cup \{0\}$, (und für $n = N$ ist die Abschätzung $\mid a_{N+(n-N)} - a_N \mid = 0 < \varepsilon/2$ natürlich richtig).

q.e.d.

Beachtet man

Satz 3.32 *Cauchyfolgen sind beschränkt,*

so erhält man eine weitere nützliche Äquivalenz:

Satz 3.33 $(a_n)_{n\in\mathbb{N}}$ *ist genau dann Cauchyfolge, wenn eine Nullfolge* $(b_n)_{n\in\mathbb{N}}$ *existiert, mit der*

$$\bigwedge_{k,n\in\mathbb{N}} \mid a_{n+k} - a_n \mid \leq b_n \tag{3.16}$$

abgeschätzt werden kann.

Beweis: von 3.32
Zu $\varepsilon := 1$ bestimme man N so, dass für alle $\ell \in \mathbb{N}$ gemäß (3.15) gilt:

$$\mid a_{N+\ell} - a_N \mid < 1 \ .$$

Dann folgt für alle $n > N$:

$$\mid a_n \mid \leq \mid a_n - a_N \mid + \mid a_N \mid < \mid a_N \mid + 1$$

Also ist

$$\sup\{\mid a_k \mid : k \in \mathbb{N}\} \leq \max\{\mid a_1 \mid; \mid a_2 \mid; \ldots; \mid a_N \mid, \mid a_N \mid + 1\} < \infty \ .$$

q.e.d.

Beweis: von 3.33
Ist $(a_n)_{n\in\mathbb{N}}$ Cauchyfolge, so ist wegen Satz 3.32

$$b_n := \sup_{k\in\mathbb{N}} \mid a_{n+k} - a_n \mid < \infty$$

für alle $n \in \mathbb{N}$. Die Folge (b_n) ist wegen (3.15) eine Nullfolge, und

$$\bigwedge_{k,n\in\mathbb{N}} \mid a_{n+k} - a_n \mid \leq b_n \ .$$

Ist umgekehrt $(b_n)_{n\in\mathbb{N}}$ eine Nullfolge, mit der Abschätzung (3.16) gilt, so existiert für vorgelegtes $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für $n \in \mathbb{N}$ mit $n \geq N$ gilt

$$b_n < \varepsilon,$$

und damit

$$| a_{n+k} - a_n | \leq b_n < \varepsilon$$

für alle $k \in \mathbb{N}$ und $n \geq N$.

Konvergiert eine Folge $(a_n)_{n\in\mathbb{N}}$ und ist etwa a ihr Grenzwert, so gilt für alle $n, k \in \mathbb{N}$

$$| a_{n+k} - a_n | \leq | a_{n+k} - a | + | a_n - a | \ .$$

Ist $\varepsilon > 0$ vorgelegt, so existiert $N \in \mathbb{N}$ mit

$$| a_m - a | < \varepsilon/2$$

für alle $m \geq N$. Daher folgt für $n \geq N$ und $k \in \mathbb{N}$

$$| a_{n+k} - a_n | < \varepsilon/2 + \varepsilon/2 = \varepsilon \ .$$

q.e.d.

Konvergente Folgen sind also Cauchyfolgen. Aber auch das Umgekehrte ist richtig:

Satz 3.34 *(Vollständigkeitssatz) Eine Folge ist genau dann konvergent, wenn sie Cauchyfolge ist.*

Ist nämlich $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, so ist sie beschränkt und enthält nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(a_{\pi(n)})_{n\in\mathbb{N}}$. Deren Grenzwert sei a. Beachtet man $\pi(n) \geq n$ für die Auswählfunktion π und alle $n \in \mathbb{N}$, so folgt für $n \in \mathbb{N}$

$$| a_n - a | \leq | a_n - a_{\pi(n)} | + | a_{\pi(n)} - a | \leq b_n + | a_{\pi(n)} - a | \ ,$$

wobei (b_n) eine (3.16) erfüllende Nullfolge ist. Insgesamt ist $(b_n + | a_{\pi(n)} - a |)$ Nullfolge, und

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

folgt mit Lemma 3.5.

Eine erste Anwendung ist das Majorantenkriterium für die Konvergenz von Reihen:

Satz 3.35 *Gegeben sei eine Reihe* $\sum_{j=1}^{\infty} a_j$, *und es gebe eine konvergente Reihe* $\sum_{j=1}^{\infty} b_j$ *mit*

$$\bigvee_{k\in\mathbb{N}} \quad \bigwedge_{j\in\mathbb{N},j\geq k} \quad | a_j |\leq b_j \ .$$

Dann ist auch $\sum_{j=1}^{\infty} a_j$ *konvergent.*

Man nennt $\sum_{n=1}^{\infty} b_n$ eine Majorante der Reihe $\sum_{n=1}^{\infty} a_n$. Beachten Sie, dass dieses Majorantenkriterium auch die Konvergenz der Reihe $\sum_{n=1}^{\infty} | a_n |$ impliziert. Indem man $b_n :=| a_n |$ setzt, erkennt man auch, dass die Konvergenz von $\sum_{n=1}^{\infty} | a_n |$ die Konvergenz von $\sum_{n=1}^{\infty} a_n$ impliziert.

Definition 3.36 *Die Reihe* $\sum_{j=1}^{\infty} a_j$ *heißt absolut konvergent, wenn* $\sum_{j=1}^{\infty} | a_j |$ *konvergent ist.*

Das Majorantenkriterium liefert also sogar die absolute Konvergenz der vorgelegten Reihe.

Beispiel 3.37 *Die Exponentialreihe* $\sum_{j=0}^{\infty} \frac{x^j}{j!}$ *konvergiert für jedes vorgelegte* $x \in \mathbb{R}$. *Denn für jedes* $n \in \mathbb{N}$ *ist*

$$\left|\frac{x^n}{n!}\right| \leq \frac{|x|^n}{n!} \leq \left(\frac{(2|x|)^n}{n!}\right)\left(\frac{1}{2}\right)^n \ ;$$

nach Beispiel 3.17 *ist* $(\frac{(2|x|)^n}{n!})$ *eine Nullfolge, also beschränkt durch eine Zahl* $c > 0$, *und die Reihe über* $c \cdot \left(\frac{1}{2}\right)^n$ *konvergiert.*

Definition 3.38

$$e := \sum_{j=0}^{\infty} \frac{1}{j!}$$

heißt Eulersche Zahl *(Leonhard Euler 1707-1783).*

Satz 3.39 *Die Eulersche Zahl e ist irrational.*

Der Beweis wird am Ende von Kapitel 5 ausgeführt.

Beweis (von Satz 3.35**):** Für $n, k \in \mathbb{N}$ ist

$$| \sum_{j=1}^{n+k} a_j - \sum_{j=1}^{n} a_j |=| \sum_{j=n+1}^{n+k} a_j |\leq \sum_{j=n+1}^{n+k} | a_n |\leq \sum_{j=n+1}^{n+k} b_j \leq \sum_{j=n+1}^{\infty} b_j;$$

und durch

$$d_n := \sum_{j=n+1}^{\infty} b_j = \sum_{j=1}^{\infty} b_j - \sum_{j=1}^{n} b_j \quad ,n \in \mathbb{N},$$

ist eine Nullfolge definiert. Satz 3.33 liefert die Cauchykonvergenz von $\left(\sum\limits_{j=1}^{n} a_j\right)_{n\in\mathbb{N}}$, und die Behauptung folgt aus dem Vollständigkeitssatz 3.34.

q.e.d.

Beispiele 3.40 *Die "Cosinusreihe"*

$$\sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!}$$

und die "Sinusreihe"

$$\sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!}$$

konvergieren für jedes $x \in \mathbb{R}$. Denn für jedes $j \in \mathbb{N}$ gelten

$$\left|(-1)^j \frac{x^{2j}}{(2j)!}\right| \leq \frac{(| x |^2)^j}{j!} \; , \quad \left|(-1)^j \frac{x^{2j+1}}{(2j+1)!}\right| \leq | x | \frac{|x|^{2j}}{j!} \quad ;$$

also ist mit $p :=| x |^2$ die Exponentialreihe $\sum\limits_{j=0}^{\infty} \frac{p^j}{j!}$ bzw. $\sum\limits_{j=0}^{\infty} \frac{|x|p^j}{j!}$ eine konvergente Majorante.

Sind N Zahlen $a_1, \ldots, a_N$ gegeben ($N \in \mathbb{N}$) und ist

$$\alpha : \{1, \ldots, N\} \to \{1, \ldots, N\}$$

eine Bijektion (also α eine Permutation der Zahlen $1, \ldots, N$), so ist

$$\sum_{n=1}^{N} a_n = \sum_{n=1}^{N} a_{\alpha(n)},$$

d. h. man kann die Reihenfolge, in der summiert wird, beliebig umordnen. Dies stimmt aber nicht mehr allgemein, wenn man es mit Reihen zu tun hat. In diesem Fall gilt der "große Umordnungssatz":

Satz 3.41 *Sei* $\sum\limits_{j=1}^{\infty} a_j$ *eine konvergente Reihe.*

(i) *Ist* $\sum\limits_{j=1}^{\infty} a_j$ *absolut konvergent, so gilt für alle Bijektionen* $\alpha : \mathbb{N} \to \mathbb{N}$:

$$\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} a_{\alpha(j)}$$

(ii) *Ist* $\sum\limits_{j=1}^{\infty} \mid a_j \mid = \infty$, *so gibt es zu jedem* $x \in \hat{\mathbb{R}}$ *eine Bijektion* $\alpha : \mathbb{N} \to \mathbb{N}$, *so dass* $\sum\limits_{j=1}^{\infty} a_{\alpha(j)} = x$ *ist.*

Beweis (zu (i)): Seien $\alpha : \mathbb{N} \to \mathbb{N}$ eine Bijektion, $a := \sum\limits_{n=1}^{\infty} a_j$ und $\varepsilon > 0$ vorgegeben. Es gibt eine Zahl $N_0 \in \mathbb{N}$, so dass für alle $m \geq n \geq N_0$ gelten

$$\left| a - \sum_{j=1}^{n} a_j \right| < \frac{\varepsilon}{2} \,, \qquad \sum_{j=n}^{m} \mid a_j \mid < \frac{\varepsilon}{2} \,.$$

Wähle nun $N \in \mathbb{N}$ so groß, dass

$$\{1, \ldots, N_0\} \subset \{\alpha(1), \ldots, \alpha(N)\} \,.$$

Dann ist $N \geq N_0$ und es gibt eine Permutation $\pi : \{1, \ldots, N\} \to \{1, \ldots, N\}$, die die Zahlen von 1 bis N_0 so ordnet, dass $\alpha(\pi(j)) = j$ ist.

Für $j > N$ und für $i \in \{N_0 + 1, \ldots, N\}$ folgt dann

$$\alpha(j) > N_0 \,, \qquad \alpha(\pi(i)) > N_0 \,.$$

Für $n > N$ gilt dann mit $m := \max\{\pi(N_0+1);\ldots,\pi(N); N+1;\ldots;n\}$:

$$\begin{aligned}
\left| a - \sum_{j=1}^{n} a_{\alpha(j)} \right| &= \left| a - \sum_{j=1}^{N} a_{\alpha(j)} - \sum_{j=N+1}^{n} a_{\alpha(j)} \right| \\
&= \left| a - \sum_{j=1}^{N} a_{\alpha(\pi(j))} - \sum_{j=N+1}^{n} a_{\alpha(j)} \right| \\
&\leq \left| a - \sum_{j=1}^{N_0} a_{\alpha(\pi(j))} \right| + \sum_{j=N_0+1}^{N} \mid a_{\alpha(\pi(j))} \mid + \sum_{j=N+1}^{n} \mid a_{\alpha(j)} \mid \\
&\leq \left| a - \sum_{j=1}^{N_0} a_j \right| + \sum_{j=N_0+1}^{m} \mid a_j \mid \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \, .
\end{aligned}$$

q.e.d.

Beweis (– Skizze zu (ii)): Wir skizzieren nur den Fall $x \in \mathbb{R}^+$. Da $\sum\limits_{j=1}^{\infty} a_j$ konvergiert, nicht aber $\sum\limits_{j=1}^{\infty} \mid a_j \mid$, enthält $(a_j)_{j\in\mathbb{N}}$ sowohl unendlich viele positive als auch unendlich viele negative Folgenglieder. $(a_{\pi(j)})_{j\in\mathbb{N}}$ bzw. $(a_{\gamma(j)})_{j\in\mathbb{N}}$ seien die Teilfolgen der positiven bzw. der nicht positiven Folgenglieder. Es gelten dann

$$\bigwedge_{j\in\mathbb{N}} \quad a_{\pi(j)} > 0 \wedge a_{\gamma(j)} \leq 0,$$

$$\mathbb{N} = \mathcal{W}(\pi) \stackrel{\bullet}{\cup} \mathcal{W}(\gamma), \tag{3.17}$$

$$\sum_{j=1}^{\infty} a_{\pi(j)} = -\sum_{j=1}^{\infty} a_{\gamma(j)} = \infty \, . \tag{3.18}$$

Die Bijektion α erhält man, indem man sukzessive Indizes aus $\mathcal{W}(\pi)$, dem Wertebereich von π, auswählt bis $\sum\limits_{j=1}^{k_1} a_{\alpha(j)} > x$ ist; dann wählt man Indizes aus dem Wertebereich $\mathcal{W}(\gamma)$ von γ aus, bis $\sum\limits_{j=1}^{k} a_{\alpha(j)} \leq x$ ist; dann wählt man wieder aus $\mathcal{W}(\pi)$ aus u.s.w.

Genau definiert man α rekursiv wie folgt:

$$\alpha(1) := \pi(1)$$

und für $k \in \mathbb{N}$

$$\alpha(k+1) := \begin{cases} \pi(j), \quad j := \min\{i : \pi(i) \notin \{\alpha(1), \ldots, \alpha(k)\}\} \text{ , falls } \sum\limits_{i=1}^{k} a_{\alpha(i)} \leq x \\ \gamma(j), \quad j := \min\{i : \gamma(i) \notin \{\alpha(1), \ldots, \alpha(k)\}\} \text{ , falls } \sum\limits_{i=1}^{k} a_{\alpha(i)} > x \end{cases}$$

Dass dieses Verfahren eine Bijektion liefert, folgt aus (3.17) und (3.18). Dass dann tatsächlich

$$\sum_{j=1}^{\infty} a_{\alpha(j)} = x$$

ist, folgt aus der Tatsache, dass (a_j) Nullfolge ist.

q.e.d.

Kapitel 4

Stetige Funktionen

4.1 Stetigkeit

Wir wollen nun reelle Funktionen untersuchen, deren Definitionsbereich nicht auf $\mathbb{N}$ beschränkt ist. Insbesondere interessieren uns Funktionen, deren Definitionsbereich z. B. ein Intervall ist. Die riesige Menge solcher Funktionen möchten wir aber durch für Anwendungen vernünftige Bedingungen einschränken. Eine erste solche Bedingung wird mit dem Begriff der "Stetigkeit" formuliert.

Man stelle sich eine Funktion

$$f : [\,a\,,\,b\,] \to \mathbb{R}$$

vor, die eine physikalische Messung beschreibt: Ist etwa $h \in [\,a\,,\,b\,]$ die Höhe einer Quecksilbersäule (in mm) so sei $f(h)$ die Temperatur (in °C) die um die Säule herum (oder an ihrem Fußpunkt) herrscht. Nehmen wir an, es herrsche die Temperatur $\theta_0 = f(h_0)$. Da physikalische Messungen immer ungenau sind, wird man eine Höhe h der Quecksilbersäule ablesen, der nahe bei h_0 liegt und auf die Temperatur θ schließen läßt, und man erwartet, dass dann auch θ nahe bei θ_0 liegt, dass also ein geringer Ablesefehler zu einem gerinen Interpretationsfehler führt.

Genauer: Ist eine Fehlertoleranz ε für die zu bestimmende Temperatur vorgegeben, so fragt es sich, wie groß die Ablesetoleranz δ sein muss, damit beim Ablesen eines Werts h im Ablesetoleranzbereich $(h_0 - \delta, h_0 + \delta)$ die daraus interpretierte Temperatur $f(h)$ im Toleranzbereich $(\theta_0 - \varepsilon, \theta_0 + \varepsilon)$ liegt.

Handelt es sich bei dem Thermometer etwa um ein Fieberthermometer, so interessiert besonders der Wert $\theta_0 = 37$ und man wird den Patienten als weder unterkühlt noch fiebrig auffassen, wenn z. B. $\varepsilon = 0,4$ ist. Die dann zugelassene Ablesegenauigkeit δ wird von θ_0 abhängen.

Hat man aber keine Temperatur θ_0 vorgegeben, so ist man daran interessiert, dass die Ablesetoleranz δ nicht von θ_0 abhängt. Diese beiden Standpunkte führen auf zwei Stetigkeitsbegriffe: der erste führt auf "Stetigkeit", der zweite auf den schärferen Begriff der "gleichmäßigen Stetigkeit".

Definition 4.1 *Eine reelle Funktion f heißt genau dann stetig in einem Punkt $\hat{x} \in D(f)$, wenn*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} f(D(f) \cap U_\delta(\hat{x})) \subset U_\varepsilon \left(f(\hat{x})\right) . \tag{4.1}$$

f heißt stetig in einer Teilmenge $M \subset D(f)$, wenn f stetig in jedem Punkt aus M ist. f heißt stetig, wenn f stetig in $D(f)$ ist.

Bevor wir Beispiele für stetige Funktionen angeben, sollen einige äquivalente Bedingungen zu (4.1) formuliert werden:

Satz 4.2 *Sei f eine reelle Funktion und $\hat{x} \in D(f)$. Dann sind äquivalent:*

(i) *f ist stetig in $\hat{x}$*

(ii) *die Bedingung*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{x\in D(f)} (\mid x - \hat{x} \mid < \delta \Rightarrow \mid f(x) - f(\hat{x}) \mid < \varepsilon) \; ; \tag{4.2}$$

(iii) *die Bedingung*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} D(f) \cap U_\delta(\hat{x}) \subset f^{-1}(U_\varepsilon(f(\hat{x}))) \; ; \tag{4.3}$$

(iv) $\lim\limits_{x\to\hat{x}} f(x) = f(\hat{x})$*, sofern $\hat{x}$ Häufungspunkt von $D(f)$ ist;*

(v) $\lim\limits_{n\to\infty} f(x_n) = f(\hat{x})$ *für jede gegen $\hat{x}$ konvergente monotone Folge $(x_n)_{n\in\mathbb{N}}$ in $D(f)$;*

(vi) $\lim\limits_{n\to\infty} f(x_n) = f(\hat{x})$ *für jede gegen $\hat{x}$ konvergente Folge $(x_n)_{n\in\mathbb{N}}$ in $D(f)$.*

(iv) schreibt man auch oft in der Form $\lim\limits_{h\to 0} f(\hat{x} + h) = f(\hat{x})$.

Beweis: Die Äquivalenz von (i), (ii) und (iii) ist unmittelbar einsichtig.

Auch erkennt man direkt, dass (4.1) die Bedingung (iv) impliziert. Sei nun andererseits (iv) erfüllt. Für $\hat{x}$ müssen zwei Möglichkeiten diskutiert werden: Entweder ist $\hat{x}$ kein Häufungspunkt von $D(f)$; dann ist für ein $\delta > 0$ der Punkt $\hat{x}$ einziger Punkt in $D(f) \cap U_\delta(\hat{x})$ und mithin $f(D(f) \cap U_\delta(\hat{x})) = \{f(\hat{x})\} \subset U_\varepsilon(f(\hat{x}))$ für jedes $\varepsilon > 0$, d. h. f ist stetig in $\hat{x}$. Oder $\hat{x}$ ist Häufungspunkt von $D(f)$; dann liefert die Bedingung $\lim\limits_{x\to\hat{x}} f(x) = f(\hat{x})$, dass zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $f(U_\delta(\hat{x}) \setminus \{\hat{x}\}) \subset U_\varepsilon(f(\hat{x}))$. Da natürlich $f(\hat{x}) \in U_\varepsilon(f(\hat{x}))$ ist, folgt sogar $f(U_\delta(\hat{x})) \subset U_\varepsilon(f(\hat{x}))$; f ist also auch in diesem Fall stetig in $\hat{x}$.

Ferner folgt (vi) aus (iv): Ist $\hat{x}$ kein Häufungspunkt von $D(f)$, so müssen gegen $\hat{x}$ konvergente Folgen $(x_n)_{n\in\mathbb{N}}$ in $D(f)$ von einem Index $\tilde{n}$ an konstant $\hat{x}$ sein: $x_n = \hat{x}$ für $n \geq \tilde{n}$. In diesem Fall folgt also (vi) aus (iv). Ist aber $\hat{x}$ Häufungspunkt von $D(f)$ und ist $(x_n) \subset D(f)$ eine monotone Folge mit $\lim\limits_{n\to\infty} x_n = \hat{x}$, so folgt (v) mit Satz 2.27.

Das folgende Lemma liefert, dass (iv) aus (v) folgt, und da (vi) trivalerweise (v) impliziert, ist dann der Satz bewiesen.

q.e.d.

Lemma 4.3 *Seien f eine reelle Funktion, z ein $\hat{\mathbb{R}}$-Häufungspunkt von $D(f)$ und $a \in \hat{\mathbb{R}}$. Dann sind (i) und (ii) äquivalent:*

(i) $\hat{\mathbb{R}} - \lim\limits_{x\to z} f(x) = a$

(ii) *Für jede streng monotone Folge $(x_n)_{n\in\mathbb{N}}$ in $D(f)$ ist $\hat{\mathbb{R}} - \lim\limits_{n\to\infty} f(x_n) = a$.*

Beweis: Wir beschränken uns auf den Nachweis "ii $\Rightarrow$ i" — das ist der Schluss, der zur Vervollständigung des Beweis von Satz 4.2 noch benötigt wird. Gilt (ii) und wäre (i) falsch, so würde gelten

$$\bigvee_{\varepsilon>0} \ \bigwedge_{\delta>0} \ \bigvee_{x\in U_\delta(z)\cap D(f)} x \neq z \wedge f(x) \notin U_\varepsilon(a) \ .$$

Wir wählen also ein solches ε und finden (setze $\delta = 1/n$) zu jedem $n \in \mathbb{N}$ eine Zahl $x_n \in U_{1/n}(z) \cap D(f)$ mit $f(x_n) \notin U_\varepsilon(a)$ und $x_n \neq z$. Dann ist $\hat{\mathbb{R}} - \lim\limits_{n\to\infty} x_n = z$, und (x_n) ist Folge in $D(f) \setminus \{z\}$.

Nun ist $x_n > z$ für unendlich viele Indizes n oder $x_n < z$ für unendlich viele Indizes. Ingesamt gibt es also eine Teilfolge $(x_{\pi(n)})_{n\in\mathbb{N}}$ mit $x_{\pi(n)} > z$ für alle $n \in \mathbb{N}$ oder $x_{\pi(n)} < z$ für alle $n \in \mathbb{N}$. Wegen $\hat{\mathbb{R}} - \lim x_{\pi(n)} = z$ kann man aus

$(x_{\pi(n)})_{n\in\mathbb{N}}$ wieder eine Teilfolge $(x_{\pi(\alpha(n))})_{n\in\mathbb{N}}$ auswählen, die streng monoton ist. Wegen (ii) folgt dann

$$\hat{\mathbb{R}} - \lim f(x_{\pi\circ\alpha(n)}) = a$$

im Widerspruch zu $f(x_{\pi(\alpha(n))}) \notin U_\varepsilon(a)$.

q.e.d.

Beachten Sie: Wegen des Lemmas 4.3 läßt sich 4.2(v) auch in der Form

$$\left.\begin{array}{ll} \text{wenn } \hat{x} \text{ Häufungspunkt von } D(f) \cap (-\infty\,,\,\hat{x}\,) \text{ ist:} & \lim\limits_{x \nearrow \hat{x}} f(x) \\ \text{wenn } \hat{x} \text{ Häufungspunkt von } D(f) \cap (\,\hat{x}\,,\,\infty\,) \text{ ist:} & \lim\limits_{x \searrow \hat{x}} f(x) \end{array}\right\} = f(\hat{x})\ . \quad (4.4)$$

Hierbei sind

$$\lim_{x\nearrow\hat{x}} := \lim_{x\to\hat{x}} f\,|_{(\,-\infty\,,\,\hat{x}\,)\cap D(f)}\ , \quad \lim_{x\searrow\hat{x}} := \lim_{x\to\hat{x}} f\,|_{(\,\hat{x}\,,\,\infty\,)\cap D(f)} \qquad (4.5)$$

die *Grenzwerte von unten und von oben* von f in $\hat{x}$. Statt *unten* bzw. *oben* sagt man auch *links* bzw. *rechts*.

Satz 4.4 *Sind die reellen Funktionen f, g in $\hat{x} \in D(f) \cap D(g)$ stetig, so auch $f+g$, $f-g$, $f \cdot g$, f/g — letzteres, sofern $g(\hat{x}) \neq 0$ ist.*

Das erhält man aus Satz 2.25. Satz 2.27 liefert

Satz 4.5 *Sind f, g reelle Funktionen und sind g in $\hat{x} \in D(g)$, f in $g(\hat{x}) \in D(f)$ stetig, so ist $f \circ g$ in $\hat{x}$ stetig.*

So erhält man

Beispiel 4.6 *Gebrochen rationale Funktionen (das sind Quotienten von Polynomen) sind stetig.*

Ferner:

Beispiel 4.7 *Die durch $r(x) := |\,x\,|$ definierte Funktion ist stetig.*

Denn für alle $x\,,\,h \in \mathbb{R}$ ist

$$|\,r(x+h) - r(x)\,| = ||\,x+h\,| - |\,x\,|| \leq |\,x+h-x\,| = |\,h\,| \to 0 \quad \textit{für} \quad h \to 0\ .$$

4.2 Gleichmäßige Stetigkeit

Nach Definition 4.1 und nach 4.2(ii) ist eine Funktion f genau dann stetig, wenn

$$\bigwedge_{s\in D(f)} \bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{t\in D(f)} \mid s-t\mid<\delta \Rightarrow \mid f(s)-f(t)\mid<\varepsilon\ . \tag{4.6}$$

Die zu wählende "Abszissentoleranz" δ hängt also nicht nur von der vorgegebenen "Ordinatentoleranz" ε sondern auch vom jeweiligen Punkt s ab, um den herum die Toleranz δ anzugeben ist. Beim Begriff der gleichmäßigen Stetigkeit soll die Toleranz δ vom Punkt s unabhängig angegeben werden können: der sich auf s beziehende Allquantor muss hinter den Existenzquantor rücken.

Definition 4.8 *Die reelle Funktion f heißt gleichmäßig stetig, wenn*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{s,t\in D(f)} (\mid s-t\mid<\delta \Rightarrow \mid f(s)-f(t)\mid<\varepsilon)$$

Beispiel 4.9 *Durch $f(x) := \frac{x}{1+x^2}$ ist eine gleichmäßig stetige Funktion gegeben. Denn für alle $s,t\in\mathbb{R}$ ist*

$$\begin{aligned}\mid f(s)-f(t)\mid &= \left|\frac{s}{1+s^2}-\frac{t}{1+t^2}\right| = \frac{\mid 1-st\mid}{(1+s^2)(1+t^2)}\mid s-t\mid \\ &\leq \frac{1+\mid st\mid}{1+s^2+t^2+s^2t^2}\mid s-t\mid \leq \mid s-t\mid\end{aligned}$$

(wegen $\mid st\mid \leq 1/2(s^2+t^2)$). Ist also $\varepsilon>0$ vorgegeben, so gilt mit $\delta:=\varepsilon$, dass $\mid f(s)-f(t)\mid<\varepsilon$ ist, sofern $\mid s-t\mid<\delta$ ist.

Beispiel 4.10 *Die durch $f(x)=x^2$ gegebene Funktion ist nicht gleichmäßig stetig.*

Denn für alle $s,t\in\mathbb{R}$ ist

$$\mid f(s)-f(t)\mid = \mid s+t\mid\mid s-t\mid\ .$$

Hieraus ersieht man: Ist z. B. $\varepsilon=1$, so existieren zu jedem $\delta>0$ Zahlen $s,t\in\mathbb{R}$ mit $\mid s-t\mid<\delta$ und $\mid f(s)-f(t)\mid\geq 1$. Setze zum Beispiel $s:=\frac{2}{\delta}, t:=s+\delta/2$.

Aber für Funktionen auf kompakten Intervallen fallen "Stetigkeit" und "gleichmäßige Stetigkeit" zusammen:

Satz 4.11 *Jede stetige Funktion f mit kompaktem Definitionsbereich $D(f)$ ist gleichmäßig stetig.*

Beweis: Wäre eine stetige Funktion f mit kompaktem Definitionsbereich nicht gleichmäßig stetig, dann gäbe es ein $\varepsilon_0 > 0$ so, dass

$$\bigwedge_{\delta>0} \quad \bigvee_{t,s\in D(f)} \quad (\mid t-s \mid < \delta \wedge \mid f(t)-f(s) \mid \geq \varepsilon_0) \ .$$

Dann gibt es insbesondere Folgen $(t_n)_{n\in\mathbb{N}}$ und $(s_n)_{n\in\mathbb{N}}$ in $D(f)$, für die gelten

$$\bigwedge_{n\in\mathbb{N}} \left(\mid t_n - s_n \mid < \frac{1}{n} \wedge \mid f(t_n)-f(s_n) \mid \geq \varepsilon_0 \right) . \tag{4.7}$$

Da $D(f)$ kompakt ist, enthält $(t_n)_{n\in\mathbb{N}}$ eine gegen ein $\hat{t} \in D(f)$ konvergierende Teilfolge $(t_{\pi(n)})_{n\in\mathbb{N}}$, und wegen

$$\mid s_{\pi(n)} - \hat{t} \mid \leq \mid s_{\pi(n)} - t_{\pi(n)} \mid + \mid t_{\pi(n)} - \hat{t} \mid \leq \frac{1}{\pi(n)} + \mid t_{\pi(n)} - \hat{t} \mid \to 0$$

konvergiert auch $(s_{\pi(n)})_{n\in\mathbb{N}}$ gegen $\hat{t}$. Aber f ist stetig in $\hat{t}$; also gibt es zu $\frac{\varepsilon_0}{2}$ eine Zahl $\delta > 0$ mit

$$\bigwedge_{t\in D(f)} \quad \mid f(t) - f(\hat{t}) \mid < \frac{\varepsilon_0}{2} \ ,$$

sofern $\mid t - \hat{t} \mid < \delta$ ist.

Für hinreichend großes n liegen aber $t_{\pi(n)}$ und $s_{\pi(n)}$ in $U_\delta(\hat{t})$, und daher ist

$$\mid f(t_{\pi(n)}) - f(s_{\pi(n)}) \mid \leq \mid f(t_{\pi(n)}) - f(\hat{t}) \mid + \mid f(s_{\pi(n)}) - f(\hat{t}) \mid < \varepsilon_0$$

im Widerspruch zu (4.7).

q.e.d.

4.3 Eigenschaften stetiger Funktionen

Wir wollen nun drei wesentliche Eigenschaften stetiger Funktionen nachweisen: Stetige Funktionen bilden kompakte Mengen in kompakte Mengen ab; stetige Funktionen bilden zusammenhängende Mengen in zusammenhängende Mengen ab; auf einem Intervall definierte stetige, injektive Funktionen besitzen eine stetige Umkehrfunktion.

Satz 4.12 *Jede stetige Funktion f mit kompaktem Definitionsbereich hat einen kompakten Wertebereich.*

Beweis: Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge im Wertebereich von f. Dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in $D(f)$ mit $y_n = f(x_n)$, und wegen der Kompaktheit von $D(f)$ enthält (x_n) eine gegen ein $x \in D(f)$ konvergente Teilfolge $(x_{\pi(n)})$. Nach Satz 4.2(vi) folgt

$$\lim_{n\to\infty} y_{\pi(n)} = \lim_{n\to\infty} f(x_{\pi(n)}) = f(x) \in W(f),$$

also ist $W(f)$ kompakt.

q.e.d.

Man beachte, dass kompakte Mengen Minimum und Maximum besitzen. Kompakte Mengen sind nämlich beschränkt, und Infimum bzw. Supremum einer beschränkten Menge M sind zumindest Häufungspunkte von M, wenn sie nicht bereits Elemente von M sind. Da aber eine kompakte Menge abgeschlossen ist, müssen Infimum und Supremum in ihr enthalten, also Minimum bzw. Maximum sein.

Als Korollar von Satz 4.12 erhält man also:

Korollar 4.13 *Ist f eine stetige Funktion mit kompaktem Definitionsbereich, so ist f beschränkt und nimmt ihr Minimum und ihr Maximum an.*

Korollar 4.13 wird häufig in dem Fall angewendet, dass $D(f)$ ein kompaktes, also ein beschränktes und abgeschlossenes Intervall ist.

Bevor wir die zweite Eigenschaft formulieren, wollen wir noch gewisse topologische Grundbegriffe einführen.

Definition 4.14 *Eine Teilmenge $M \subset \mathbb{R}$ heißt offen, wenn mit jedem Punkt aus M eine ganze Umgebung des Punktes in M enthalten ist:*

$$\bigwedge_{x\in M} \ \bigvee_{\delta>0} \ U_\delta(x) \subset M\ .$$

Beispiele für offene Mengen sind die *offenen Intervalle* $(\,a\,,\,b\,)$ mit $a < b$, $a, b \in \hat{\mathbb{R}}$. Insbesondere sind also alle offenen δ–Umgebungen $U_\delta(a)$ von Punkten $a \in \hat{\mathbb{R}}$ offen. Man mache sich klar, daß eine Menge sowohl *nicht* offen als auch *nicht* abgeschlossen sein kann. Z. B. $M := (\,0\,,\,1\,]$ ist weder offen noch abgeschlossen. Aber das Komplement einer offenen Menge ist stets abgeschlossen (und umgekehrt):

Satz 4.15 *$M \subset \mathbb{R}$ ist genau dann offen, wenn das Komplement $\mathbb{R} \setminus M$ abgeschlossen ist.*

Beachten Sie:
Das Intervall $(0\,,\,1]$ ist weder offen noch abgeschlossen. Ist aber $-\infty \leq a < b \leq \infty$, so ist $(a\,,\,b)$ stets offen.

Beweis (von Satz 4.15)**:** Ist M offen und ist z ein Häufungspunkt von $\mathbb{R} \setminus M$, so liegen in jeder Umgebung von z Punkte aus $\mathbb{R} \setminus M$; d. h. keine Umgebung von z liegt ganz in M. Mithin ist $z \notin M$, also $z \in \mathbb{R} \setminus M$. $\mathbb{R} \setminus M$ ist also abgeschlossen.

Ist umgekehrt $\mathbb{R} \setminus M$ abgeschlossen, so ist jeder Punkt $x \in M$ sicher nicht Häufungspunkt von $\mathbb{R} \setminus M$, d. h. jeder Punkt $x \in M$ besitzt eine Umgebung, in der kein Punkt aus $\mathbb{R} \setminus M$ liegt, die also ganz in M liegt. M ist also offen.

q.e.d.

Satz 4.16 *A_1, A_2 seien abgeschlossene, V_1, V_2 offene Teilmengen von $\mathbb{R}$. $\mathbb{A}$ sei eine Familie abgeschlossener, $\mathbb{O}$ eine Familie offener Teilmengen von $\mathbb{R}$. Es gelten*

(i) *$\emptyset$ ist offen*
(ii) *$\mathbb{R}$ ist offen*
(iii) *$V_1 \cap V_2$ ist offen*
(iv) *$\bigcup_{U\in\mathbb{O}} U$ ist offen*

(i') *$\mathbb{R}$ ist abgeschlossen*
(ii') *$\emptyset$ ist abgeschlossen*
(iii') *$A_1 \cup A_2$ ist abgeschlossen*
(iv') *$\bigcap_{A\in\mathbb{A}} A$ ist abgeschlossen*

Beweis: Es genügt wegen Satz 4.15, sich auf die Aussagen (i)–(iv) zu beschränken, die die offenen Mengen betreffen.

(i) und (ii) sind trivial.

Ist $x \in V_1 \cap V_2$, so gibt es $\delta_1 > 0$ und $\delta_2 > 0$ mit $U_{\delta_1}(x) \subset V_1$, $U_{\delta_2}(x) \subset V_2$. Mithin ist $U_\delta(x) \subset V_1 \cap V_2$ mit $\delta := \min\{\delta_1, \delta_2\} > 0$. Also gilt (iii).

Ist $x \in \bigcup\limits_{U\in\mathbb{O}} U$, so gibt es eine offene Menge $V \in \mathbb{O}$, in der x und somit auch $U_\delta(x)$ mit geeignetem $\delta > 0$ liegt. Es folgt

$$U_\delta(x) \subset V \subset \bigcup_{U\in\mathbb{O}} U\ .$$

Damit ist (iv) bewiesen.

q.e.d.

Man nennt eine Menge X einen *topologischen Raum*, wenn eine Familie $\mathcal{S}$, bestehend aus Teilmengen von X, ausgezeichnet ist mit (i) - (iv) entsprechenden Eigenschaften:

$$\emptyset \in \mathcal{S}\ ,$$
$$X \in \mathcal{S}\ ,$$
$$\bigwedge_{V_1,V_2\in\mathcal{S}} \quad V_1 \cap V_2 \in \mathcal{S}\ ,$$
$$\bigwedge_{\mathbb{O}\subset\mathcal{S}} \quad \bigcup_{V\in\mathbb{O}} V \in \mathcal{S}\ .$$

Eine Menge aus der Familie $\mathcal{S}$ nennt man eine *offene* Teilmenge von X. In topologischen Räumen kann man die Begriffe der Konvergenz und der Stetigkeit formulieren, und eine Reihe von anschaulichen Eigenschaften lassen sich beschreiben, indem man nur aus “offenen Mengen” abgeleitete Begriffe benutzt. Solche Eigenschafen nennt man dann “topologische Eigenschaften”. Zum Beispiel ist die Überdeckungskompaktheit (vgl. die Ausführung hinter Definition 3.25) ein topologischer Begriff, und wir werden später sehen, dass Satz 4.12 in weit größerer Allgemeinheit gilt.

Auch folgende Charakterisierung der Stetigkeit gilt in weitaus größerer Allgemeinheit:

Satz 4.17 *Eine reelle Funktion f ist genau dann stetig, wenn zu jeder offenen Menge $W \subset \mathbb{R}$ eine offene Menge $V \subset \mathbb{R}$ so existiert, dass*

$$f^{-1}(W) = D(f) \cap V\ . \tag{4.8}$$

Ist insbesondere $D(f)$ offen, so heißt dies: das Urbild offener Mengen ist offen.

Beweis: Sei f stetig und $z \in f^{-1}(W)$. Wähle $\varepsilon_z > 0$ so, dass $U_{\varepsilon_z}(f(z)) \subset W$. Dann gibt es $\delta_z > 0$ mit $D(f) \cap U_{\delta_z}(z) \subset f^{-1}(U_{\varepsilon_z}(f(x))) \subset f^{-1}(W)$. Nach Satz 4.16(iv) ist $V := \bigcup\{U_{\delta_z}(z)\colon z \in f^{-1}(W)\}$ offen und

$$f^{-1}(W) = D(f) \cap V\ .$$

Existiert umgekehrt zu jeder offenen Menge $W \subset \mathbb{R}$ eine offene Menge $V \subset \mathbb{R}$ mit (4.8), so ist zu jedem $\varepsilon > 0$

$$f^{-1}(U_\varepsilon(f(z))) = D(f) \cap V$$

mit einer offenen Menge V, welche z enthält. V enthält dann auch eine geeignete δ-Umgebung von z:

$$\bigvee_{\delta>o} \quad D(f) \cap U_\delta(z) \subset f^{-1}(U_\varepsilon(f(z))),$$

d. h. f ist stetig in z und damit insgesamt stetig, denn z war beliebig.

q.e.d.

Wir wollen nun *zusammenhängende* Mengen definieren. Das ist für Teilmengen von $\mathbb{R}$ einfach: Zusammenhängende Teilmengen von $\mathbb{R}$ sind genau die Intervalle. Aber tatsächlich ist *Zusammenhang* eine topologische Eigenschaft:

Satz 4.18 *Für eine Teilmenge X von $\mathbb{R}$ sind äquivalent:*

(i) *X ist ein Intervall*

(ii) *Sind U und V offene Teilmengen von $\mathbb{R}$ mit den Eigenschaften*

$$X \cap U \cap V = \emptyset \quad , \quad (X \cap U) \cup (X \cap V) = X \quad , \tag{4.9}$$

so ist $X \cap U = \emptyset$ oder $X \cap V = \emptyset$.

Teilmengen von $X \subset \mathbb{R}$, die sich als Durchschnitt von X mit einer offenen Menge O schreiben lassen, nennt man auch *relativ offen* (in X).

Satz 4.18 besagt, dass die Intervalle genau die Teilmengen von $\mathbb{R}$ sind, die sich nicht als punktfremde Vereinigung zweier nicht–leerer relativ offener Teilmengen darstellen lassen.

Beweis (von Satz 4.18**):** Der Leser mache sich klar, dass ein Intervall $I \subset \mathbb{R}$ durch die Bedingung

$$\bigwedge_{x,y \in I} \quad x < y \Rightarrow (x, y) \subset I \tag{4.10}$$

charakterisiert ist: Gilt (4.10) für eine Menge I, so ist I ein Intervall mit den Grenzen $\inf I$ und $\sup I$ oder $I = \emptyset$; aber die leere Menge ist auch ein Intervall.

Zu “(i) $\Rightarrow$ (ii)”: Wenn (ii) nicht gilt, so gibt es offene Mengen $U, V \subset \mathbb{R}$, die (4.9) sowie $X \cap U \neq \emptyset$, $X \cap V \neq \emptyset$ erfüllen. Seien dann $u \in X \cap U$, $v \in X \cap V$ und o. B. d. A. $u < v$. Wir definieren

$$t := \sup\{s \in [\,u\,,\, v\,] \colon s \in U\} = \sup\,[\,u\,,\, v\,] \cap U \ .$$

Da U und V offen sowie $(X \cap U \cap V) = \emptyset$ sind, muss $u < t < v$ sein. t kann nicht in U liegen, sonst wäre t keine obere Schranke für $[\,u\,,\, v\,] \cap U$. t kann aber auch nicht in V liegen, sonst gäbe es eine kleinere obere Schranke für $[\,u\,,\, v\,] \cap U$ als t. Dies aber widerspricht der zweiten Gleichung in (4.9).

Zu “(ii) $\Rightarrow$ (i)”: Ist X kein Intervall, so gibt es $x, y \in X$ mit $x < y$ und ein $z \in (x, y)$ mit $z \notin X$. Mit $U := (\,-\infty\,,\,z\,)$ und $V := (\,z\,,\,\infty\,)$ ist dann (4.9) erfüllt, aber $x \in X \cap U$ und $y \in X \cap V$ im Widerspruch zu (ii).

q.e.d.

Definition 4.19 *Eine Teilmenge $X \subset \mathbb{R}$ heißt zusammenhängend, wenn die Bedingung* 4.18*(ii) gilt: Sind U, V offene Teilmengen von $\mathbb{R}$ mit den Eigenschaften $X \cap U \cap V = \emptyset$ und $(X \cap U) \cup (X \cap V) = X$, so ist $X \cap U = \emptyset$ oder $X \cap V = \emptyset$.*

Wegen Satz 4.18 liefert Definition 4.19 eine topologische Charakterisierung von Intervallen.

Wir kommen nun zur zweiten wesentlichen Eigenschaft stetiger Funktionen

Satz 4.20 *Ist f eine stetige Funktion mit zusammenhängendem Definitionsbereich, so ist der Wertebereich von f auch zusammenhängend.*

Etwas lässig gesprochen, beinhaltet Satz 4.20 die Eigenschaft stetiger Funktionen, dass man ”sie durchzeichnen kann”. (sofern ihr Definitionsbereich ein Intervall ist). Dies wird im folgenden Korollar aufgeschrieben, welches man den *Zwischenwertsatz* nennt.

Korollar 4.21 *Sei f eine stetige Funktion, für die $[\,s\,,\,t\,] \subset D(f)$ ist. Es sei $a := \min\{f(s)\,;\,f(t)\}$, $b := \max\{f(s)\,;\,f(t)\}$. Dann gibt es zu jedem $c \in (\,a\,,\,b\,)$ einen Punkt $\xi \in (\,s\,,\,t\,)$, für den $c = f(\xi)$ ist.*

Siehe Abbildung auf der folgenden Seite.

Beweis (von Satz 4.20**):** Ist $W(f)$ nicht zusammenhängend, so gibt es zwei offene Mengen $\hat{U}$, $\hat{V}$ mit den Eigenschaften

$$W(f) \cap \hat{U} \cap \hat{V} = \emptyset \tag{4.11}$$

$$(W(f) \cap \hat{U}) \cup (W(f) \cap \hat{V}) = W(f) \tag{4.12}$$

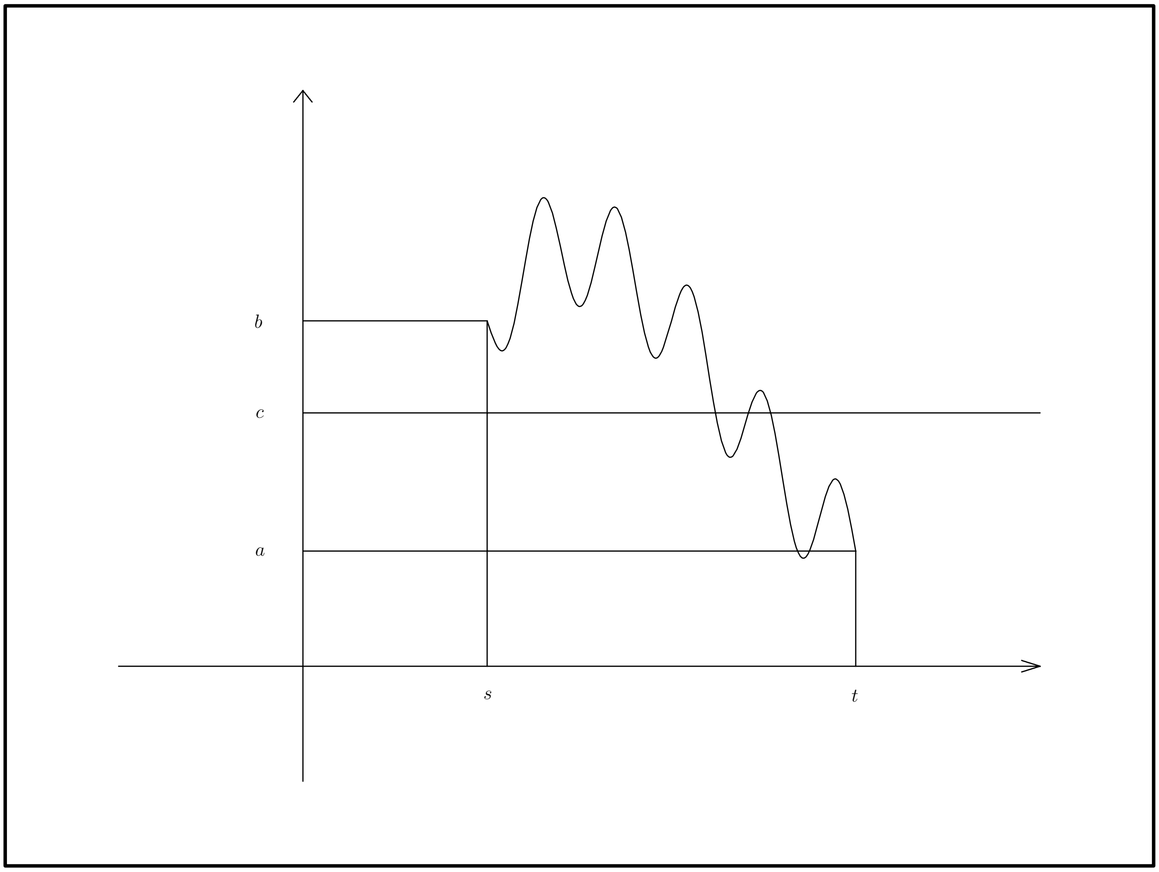
$$\hat{U} \cap W(f) \neq \emptyset \quad , \quad \hat{V} \cap W(f) \neq \emptyset\,. \tag{4.13}$$

Nun gibt es nach Satz 4.17 offene Mengen U, V mit

$$f^{-1}(\hat{U}) = U \cap D(f) \quad , \quad f^{-1}(\hat{V}) = V \cap D(f)\;.$$

Wegen (4.11) ist $U \cap V \cap D(f) = \emptyset$, wegen (4.12) ist $(U \cap D(f)) \cup (V \cap D(f)) = D(f)$ und wegen (4.13) ist $U \cap D(f) \neq \emptyset$ sowie $V \cap D(f) \neq \emptyset$. Dann aber ist $D(f)$ nicht zusammenhängend.

q.e.d.



Irgendwo schneidet der Graph von f die Parallele zur Abszisse durch die Ordinate c.

Kombiniert man die Sätze 4.12 und 4.20, so erhält man

Korollar 4.22 *Ist der Definitionsbereich der stetigen reellen Funktion f ein kompaktes Intervall, so ist auch ihr Wertebereich ein kompaktes Intervall.*

Die dritte wesentliche Eigenschaft stetiger Funktionen ist die Tatsache, dass jede injektive stetige Funktion, deren Definitionsbereich ein Intervall ist, eine Umkehrfunktion besitzt. Dabei definiert man natürlich

Definition 4.23 *Ist f eine reelle injektive Funktion, so heißt*

$$f^{-1} \ : \ W(f) \longrightarrow \mathbb{R}\,, \ y := f(x) \longmapsto x$$

Umkehrfunktion von f.

Der Definitionsbereich von f^{-1} ist also der Wertebereich von f, und f^{-1} ordnet jedem $y \in W(f)$ die eindeutige Lösung der Gleichung $f(x) = y$ zu.

Satz 4.24 *Ist f stetige und injektive reelle Funktion, und ist der Definitionsbereich von f ein Intervall, so ist f streng monoton und die Umkehrfunktion f^{-1} ist stetig.*

Damit erhält man die Stetigkeit der durch $f(x) = x^\alpha$ mit rationalem α gegebenen Funktion oder — allgemeiner — von algebraischen Funktionen. Das sind Funktionen, die man mittels algebraischen Operationen aus $id : \mathbb{R} \longrightarrow \mathbb{R}\,,\ x \longmapsto x$ und den konstanten Funktionen erhält. Algebraische Operationen sind: Addition, Multiplikation, Division und Wurzelziehen (zum Exponenten $n \in \mathbb{N}$).

Beweis (von Satz 4.24): Wir beweisen zunächst die strenge Monotonie. Hierzu seien $p, q, u, v \in D(f)$ mit $p < q$ und $u < v$. Wir haben zu zeigen, dass

$$[f(q) - f(p)] \cdot [f(v) - f(u)] > 0 \tag{4.14}$$

ist. Setze $\gamma := f(v) - f(u)$. Da f injektiv ist, ist $\gamma \neq 0$.

Da $D(f)$ ein Intervall ist, ist

$$g \;:\; [\,0\,,\,1\,] \longrightarrow \mathbb{R}\,,\;\; t \longmapsto \gamma\left(f\left((1-t)\,v + tq\right) - f\left((1-t)\,u + tp\right)\right)$$

eine wohldefinierte und stetige Funktion, und $g(0) = \gamma^2 > 0$. g besitzt keine Nullstelle: Ist nämlich $g(t) = 0$ für ein $t \in [\,0\,,\,1\,]$, so ist wegen der Injektivität von f

$$(1-t)v + tq = (1-t)u + tp \Leftrightarrow (1-t)(v-u) + t(q-p) = 0\ .$$

Aber $(1-t)(v-u) + t(p-q) > 0$. Nach Korollar 4.22 ist daher $g(1) > 0$, und das ist (4.14). f ist also streng monoton.

Wir nehmen nun o. B. d. A. f als streng monoton wachsend an — anderenfalls gehen wir zu $-f$ über. Ist $y \in W(f)$ und ist $(y_n)_{n\in\mathbb{N}}$ eine gegen y konvergente monotone Folge in $W(f)$, so ist auch die durch $x_n := f^{-1}(y_n)$ definierte Folge im gleichen Sinne wie (y_n) monoton. Ist (y_n) monoton wachsend, so folgt

$$\bigwedge_{n\in\mathbb{N}} \quad x_1 \leq x_n \leq f^{-1}(y)\ ,$$

und die Ungleichheitszeichen kehren sich um, wenn (x_n) fallend ist. Jedenfalls existiert $x := \lim_{n\to\infty} x_n$ und liegt in $D(f)$, da $D(f)$ Intervall ist. Satz 4.2(iv) liefert

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} y_n = y,$$

folglich $f^{-1}(y) = x$. Also gilt

$$\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(y)$$

für jede gegen y konvergente monoton wachsende Folge (y_n) in $D(f^{-1}) = W(f)$. Analog argumentiert man für eine monoton fallende Folge.

q.e.d.

Schließlich führen wir noch als Schreibweise ein

Definition 4.25 *Sei* $\emptyset \neq D \subset \mathbb{R}$*.*

$$C(D) := C^0(D) := \{f \in \mathcal{F}(D, \mathbb{R})\colon f \;\textit{ ist stetig}\}$$

heißt Raum der stetigen reellwertigen Funktionen auf D*.*

Man überlege sich, dass $C(D)$ ein Vektorraum ist, wenn man mit $\alpha, \beta \in \mathbb{R}$ und $f, g \in C(D)$ erklärt

$$\alpha f + \beta g \;:\; D \longrightarrow \mathbb{R}\,,\; x \longmapsto \alpha f(x) + \beta g(x)\;.$$

Kapitel 5

Differenzierbare Funktionen

5.1 Der Ableitungsbegriff

Es gibt verschiedene Vorstellungen, die zur Ableitung einer Funktion f in einem Punkt ξ ihres Definitionsbereichs führen. Newton stellt sich die Funktion f vor, als eine Vorschrift, die jedem Zeitpunkt t eine Weglänge $f(t)$ zuordnet, die etwa ein Massenpunkt bis zu diesem Zeitpunkt zurückgelegt hat. Die Ableitung dient dann der Geschwindigkeitsdefinition.

Bei Leibniz steht die Frage nach der Definition einer Tangente im Vordergrund. Wir wollen von der Vorstellung der Approximation von f durch eine affin lineare Funktion (ein Polynom ersten Grades) ausgehen (Brook Taylor (1685-1731)).

Zunächst:

Definition 5.1 *Ein Punkt x heißt innerer Punkt einer Teilmenge $M \subset \mathbb{R}$, wenn eine Umgebung $U_\varepsilon(x)$ von x mit $U_\varepsilon(x) \subset M$ existiert.*

Eine offene Menge besteht also nur aus inneren Punkten.

Definition 5.2 *Sei f eine reelle Funktion und x ein innerer Punkt von $D(f)$. f heißt differenzierbar in x mit Ableitung $\alpha \in \mathbb{R}$ in x, wenn es eine in 0 stetige Funktion R_x mit $R_x(0) = 0$ gibt, so dass für alle $h \in \mathbb{R}$ mit $x + h \in D(f)$ gilt*

$$f(x+h) = f(x) + \alpha h + R_x(h) \cdot h \, . \tag{5.1}$$

Satz und Definition 5.3 *Sei f eine reelle Funktion und ξ ein innerer Punkt von $D(f)$.*

(i) *Ist f differenzierbar in ξ, so ist die Ableitung in ξ eindeutig bestimmt und wird mit $f'(\xi)$ bezeichnet.*

(ii) *f ist genau dann differenzierbar in ξ wenn*

$$\lim_{h\to 0} \frac{f(\xi+h)-f(\xi)}{h}$$

existiert. Dieser Grenzwert ist gleich $f'(\xi)$.

(iii) *f ist genau dann differenzierbar in ξ, wenn es eine in ξ stetige Funktion $F : D(f) \to \mathbb{R}$ gibt, so dass für alle $x \in D(f)$ gilt*

$$f(x) = f(\xi) + F(x)(x-\xi) \ . \tag{5.2}$$

Dann ist $f'(\xi) = F(\xi)$.

Beweis:
Zu i): Sei $\delta > 0$ so, dass $U_\delta(\xi) \subset D(f)$. Sind sowohl α als auch $\hat{\alpha}$ Ableitungen von f in ξ, so gilt gemäß (5.1) mit in 0 stetigen und verschwindenden Funktionen R_x und $\tilde{R}_x$:

$$\bigwedge_{|h|<\delta} \quad (\alpha - \tilde{\alpha})h = (R_x(h) - \tilde{R}_x(h))h$$

Folglich ist für $h \in U_\delta(0) \setminus \{0\}$

$$\alpha - \tilde{\alpha} = R_x(h) - \tilde{R}_x(h) \to 0 \quad (\text{für } \ h \to 0);$$

d. h. $\alpha = \tilde{\alpha}$, und das war zu zeigen.

Zu ii): Ist f differenzierbar in ξ, so gilt nach (5.1) mit $\alpha := f'(\xi)$

$$\lim_{h\to 0} \frac{f(\xi+h)-f(\xi)}{h} = \lim_{h\to 0} \frac{\alpha h + R_\xi(h)\cdot h}{h} = \lim_{h\to 0}(\alpha + R_\xi(h)) = \alpha$$

Ist umgekehrt

$$\lim_{h\to 0} \frac{f(\xi+h)-f(\xi)}{h} = \alpha \ ,$$

so ist durch

$$R_\xi(h) := \begin{cases} \frac{f(\xi+h)-f(\xi)}{h} - \alpha \, , & \text{falls } h \neq 0 \text{ und } \xi + h \in D(f) \\ 0 \, , & \text{falls } h = 0 \end{cases}$$

eine in 0 stetige Funktion R_ξ definiert, für die (5.1) (mit ξ statt x) gilt:

Zu iii): Dies ist klar mit den Beziehungen

$$F(x) = R_\xi(x-\xi) + f'(\xi)\ , \quad R_\xi(h) = F(\xi+h) - f'(\xi)\ .$$

q.e.d.

Aus 5.3(iii) erhält man sofort:

Korollar 5.4 *Ist f differenzierbar in ξ, so ist f stetig in ξ.*

Um einen Anschluss an die Leibnizsche Vorstellung zu gewinnen, definieren wir

Definition 5.5 *Ist f differenzierbar in ξ, so heißt die durch*

$$T_{f,\xi}\ :\ \mathbb{R} \longrightarrow \mathbb{R}\,,\ x \longmapsto f(\xi) + f'(\xi)(x-\xi)$$

gegebene affin lineare Funktion die Tangente an f in ξ.

Die Definition 5.2 der Ableitung wird Ihnen ungewohnt erscheinen. Sie erlaubt im Gegensatz zur Formulierung in Satz 5.3(ii) eine Verallgemeinerung auf Funktionen, deren Definitionsbereiche in $\mathbb{R}^N$ oder noch allgemeineren Vektorräumen liegen. Dennoch erlaubt Satz 5.3, Grundableitungen so zu berechnen, wie Sie es vielleicht von der Schule her kennen.

Definition 5.6 *Ist f eine reelle Funktion, so heißt*

$$D(f') := \{x \in D(f)\colon f\, \textit{ist differenzierbar in}\, x\}$$

der Definitionsbereich der Ableitung von f, *und*

$$f'\ :\ D(f') \longrightarrow \mathbb{R}\,,\ x \longmapsto f'(x)$$

heißt die Ableitung von f, *sofern $D(f')$ nicht leer ist.*

Ist $D(f') = D(f)$, so heißt f differenzierbar.

Man beachte, dass $D(f')$ stets eine Teilmenge von $D(f)$ ist.

Wir wollen nun die Ableitungen möglichst vieler Funktionen gewinnen. Beispiele, die wir direkt mit der Definition gewinnen wollen, sind

Beispiel 5.7 *Sei id die Identität, also gegeben durch* $id(x) = x$. *Dann ist*

$$D(id') = D(id) = \mathbb{R}\ , \quad id'(x) = 1 \quad \text{für alle } x \in \mathbb{R}\ .$$

Denn für $x, h \in \mathbb{R}$ *ist*

$$id(x+h) - id(x) = h = 1 \cdot h + 0 \cdot h\ .$$

Die Bedingung (5.1) *aus Definition* 5.2 *ist also mit* $\alpha = 1$ *und* $R_x(h) = 0$ *(für alle* h*) erfüllt.*

Beispiel 5.8 *Mit einer gegebenen Zahl* $\beta \in \mathbb{R}$ *sei* k *durch* $k(x) = \beta$ *gegeben.* k *ist also die konstante Funktion mit Wert* β. *Dann ist* $D(k') = D(k) = \mathbb{R}$ *und* $k'(x) = 0$ *für alle* $x \in \mathbb{R}$.

Denn

$$k(x+h) - k(x) = 0 = 0 \cdot h + 0 \cdot h\ ,$$

d. i. (5.1) *mit* $\alpha = 0$ *und* $R_x(h) = 0$ *(für alle* h*).*

Beispiel 5.9 *Sei* $K = \frac{1}{id}$, *also* $K(x) = \frac{1}{x}$ *für alle* $x \in \mathbb{R} \setminus \{0\}$. *Dann ist* $D(K) = \mathbb{R} \setminus \{0\}$ *und für* $x \in \mathbb{R} \setminus \{0\}$, $h \in \mathbb{R} \setminus \{-x\}$ *gilt*

$$K(x+h) - K(x) = -\frac{h}{x(x+h)} = -\frac{1}{x^2}h + \left(\frac{1}{x^2} - \frac{1}{x(x+h)}\right) h\ .$$

Da die durch $R_x(h) = \frac{1}{x^2} - \frac{1}{x(x+h)}$ *gegebene Funktion in* $h = 0$ *stetig ist mit Wert* $R_x(0) = 0$, *ist*

$$D(K') = \mathbb{R} \setminus \{0\} = D(K)\ , \quad K'(x) = -\frac{1}{x^2}\ .$$

Beispiel 5.10 *Sei* r *die Betragsfunktion:* $r(x) := |x|$ *für* $x \in \mathbb{R}$. *Diese ist für jedes* $x \in \mathbb{R} \setminus \{0\}$ *differenzierbar: Mit* $\delta := |x|$ *gilt nämlich für alle* $h \in U_\delta(0)$

$$r(x+h) = r(x) + \begin{cases} 1 \cdot h + 0 \cdot h & \text{falls } x > 0 \text{ ist} \\ -1 \cdot h + 0 \cdot h & \text{falls } x < 0 \text{ ist} \end{cases} \quad ;$$

also ist $r'(x) = \frac{x}{|x|}$ *für* $x \neq 0$.

In $x = 0$ *aber ist* r *nicht differenzierbar, denn für* $h \neq 0$ *ist*

$$\frac{r(0+h) - r(0)}{h} = \frac{|\,h\,|}{h} = \begin{cases} 1 & \text{für } h > 0 \\ -1 & \text{für } h < 0 \end{cases} \quad ;$$

der Grenzwert für h *gegen* 0 *existiert also nicht. Wohl aber ist* r *überall stetig nach Beispiel* 4.7. *Wir haben also eine Funktion aus* $C(\mathbb{R})$ *gefunden, die* nicht *überall differenzierbar ist.*

5.2 Die Ableitungsregeln

Nun sollen Sätze angegeben werden, wie man aus den Ableitungen zweier Funktionen auf die Ableitung einer Funktion schließen kann, die aus diesen beiden Funktionen aufgebaut ist.

Satz 5.11 *Seien f, g zwei in $x \in \mathbb{R}$ differenzierbare Funktionen. Dann sind die Funktionen $f+g$, $f-g$, $f \cdot g$ und, sofern $g(x) \neq 0$ ist, f/g in x differenzierbar mit den Ableitungen*

$$(f \pm g)'(x) = f'(x) \pm g'(x)\ , \tag{5.3}$$

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x) \qquad \textit{(Produktregel)}, \tag{5.4}$$

$$(f/g)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} \qquad \textit{(Quotientenregel)}. \tag{5.5}$$

Insbesondere ist für $\alpha, \beta \in \mathbb{R}$

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x) \qquad \textit{(Linearität)}.$$

Definition 5.12 *Ist D eine offene Teilmenge von $\mathbb{R}$ so ist*

$$\begin{aligned} C^1(D) &:= \{f \in C(D) \colon f \quad \textit{ist differenzierbar und} \quad f' \in C(D)\}\ , \\ C^0(D) &:= C(D)\ . \end{aligned}$$

$C^1(D)$ heißt Raum der einmal stetig differenzierbaren Funktionen auf D.

Gemäß Satz 5.11 ist durch die Ableitung also ein linearer Operator von $C^1(D)$ nach $C^0(D)$ definiert:

$$d\ :\ C^1(D) \longrightarrow C^0(D)\ ,\ f \longmapsto f'$$

ist linear.
Beweis (von Satz 5.11): Wir verwenden Satz (5.3)(iii) und schreiben

$$\begin{aligned} f(t) &= f(x) + F(t)(t-x) \\ g(t) &= g(x) + G(t)(t-x) \end{aligned}$$

mit Funktionen F, G, die in einer Umgebung U von x definiert und in x stetig sind mit den Werten $F(x) = f'(x)$, $G(x) = g'(x)$.

Dann erhält man für $t \in U$

$$\begin{aligned} (f \pm g)(t) &= (f(x) \pm g(x)) + (F(t) \pm G(t))(t-x) \\ (f \cdot g)(t) &= (f(x) + F(t)(t-x))(g(x) + G(t)(t-x)) \\ &= f(x) \cdot g(x) + H(t)(t-x) \end{aligned}$$

mit

$$H(t) = F(t)g(x) + f(x)G(t) + F(t)G(t)(t-x)$$

Offenbar sind $F \pm G$ bzw. H in einer Umgebung U von x definiert und in x stetig mit den Werten $(F \pm G)(x) = f'(x) \pm g'(x)$ bzw. $H(x) = f'(x)g(x) + f(x)g'(x)$. Damit sind (5.3) und (5.4) bewiesen.

Mittels der Kettenregel, die im folgenden Satz formuliert wird, wird gezeigt werden, dass

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)} := \frac{g'(x)}{(g(x))^2} \tag{5.6}$$

ist. Die Quotientenregel (5.5) erhält man dann direkt aus (5.6) und der Produktregel (5.4). Bis auf Formel (5.6) ist Satz 5.11 jetzt bewiesen.

q.e.d.

Die Kettenregel beschreibt, wie man die Hintereinanderschaltung zweier Funktionen differenziert:

Satz 5.13 *(Kettenregel) Ist g differenzierbar in x und ist f differenzierbar in $g(x)$, so ist $f \circ g$ differenzierbar in x und*

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

Die Funktion $1/g$, für die (5.6) noch nachzuweisen ist, läßt sich mit K aus Beispiel 5.9 schreiben in der Form

$$1/g = K \circ g \ .$$

Die Kettenregel liefert dann

$$(1/g)'(x) = K'(g(x))g'(x) = -\frac{1}{(g(x))^2}g'(x) \ ,$$

also (5.6).

Beweis (der Kettenregel): Da eine Umgebung $U_\varepsilon(z)$ von $z := g(x)$ existiert, welche in $D(f)$ liegt, und da g stetig ist in x, gibt es eine Umgebung $U_\delta(x)$, für welche $g(U_\delta(x) \cap D(g)) \subset U_\varepsilon(z)$ ist. Da zudem x innerer Punkt von $D(g)$ ist, kann δ so gewählt werden, dass $U_\delta(x) \subset D(g)$. Mithin ist $U_\delta(x) \subset D(f \circ g)$. Also ist x innerer Punkt von $D(f \circ g)$.

Wir verwenden zum Nachweis der Kettenregel Satz 5.3(iii). Danach gibt es Funktionen F,G, die in $D(f)$ bzw. $D(g)$ definiert sind, in z bzw. x stetig sind und mit denen gelten

$$\bigwedge_{s \in D(f)} f(s) = f(z) + F(s)(s-z) \quad , \quad \bigwedge_{t \in D(g)} g(t) = g(x) + G(t)(t-x)$$

sowie

$$F(z) = f'(z) = f'(g(x)) \quad , \quad G(x) = g'(x) \; .$$

Dann ist für $t \in D(f \circ g)$

$$\begin{aligned} f \circ g(t) &= f(g(t)) = f(g(x)) + F(g(t)) \cdot (g(t) - g(x)) \\ &= f \circ g(x) + F \circ g(t) \cdot G(t) \cdot (t - x) \; , \end{aligned}$$

und $F \circ g \cdot G$ ist stetig in x mit Wert $F \circ g \cdot G(x) = f'(g(x)) \cdot g'(x)$ Damit ist die Kettenregel bewiesen.

q.e.d.

Wir sind jetzt in der Lage, Potenzfunktionen, Polynome und gebrochen rationale Funktionen zu differenzieren:

Beispiel 5.14 *Für $n \in \mathbb{N}$ sei*

$$f_n(x) = x^n \; .$$

Nach Beispiel 5.7 *ist f_1 differenzierbar, $f_1'(x) = 1$.*

Nach der Produktregel gilt: Ist f_n differenzierbar, so auch $f_{n+1} = f_n \cdot f_1$ und

$$f_{n+1}'(x) = f_n'(x) \cdot f_1(x) + f_n(x) \cdot f_1'(x) = x f_n'(x) + f_n(x) \; .$$

Mit Induktion folgt nun: f_n ist für alle $n \in \mathbb{N}$ differenzierbar und

$$\bigwedge_{n \in \mathbb{N}} \bigwedge_{x \in \mathbb{R}} f_n'(x) = n x^{n-1} \; .$$

Beispiel 5.15 *Für $n \in \mathbb{N}$ sei*

$$f_{-n}(x) = x^{-n} = \frac{1}{f_n(x)} \; .$$

Nach der Quotientenregel ist f_{-n} differenzierbar und zwar ist

$$\bigwedge_{n \in \mathbb{N}} \bigwedge_{x \in \mathbb{R} \setminus \{0\}} f_{-n}'(x) = -n x^{-n-1} \; .$$

Beispiel 5.16 *Sei P ein Polynom:*

$$\bigwedge_{x \in \mathbb{R}} P(x) = \sum_{j=0}^{n} a_j x^j \; .$$

Nach Beispielen 5.14, 5.8 *und nach Satz* 5.11 *ist P differenzierbar und*

$$\bigwedge_{x \in \mathbb{R}} P'(x) = \sum_{j=1}^{n} a_j j x^{j-1} = \sum_{i=0}^{n-1} (i+1) a_{i+1} x^i$$

Beispiel 5.17 *Mit zwei Polynomen P und Q, $Q \not\equiv 0$, sei f die durch*

$$f(x) = \frac{P(x)}{Q(x)}$$

gegebene gebrochen rationale Funktion. Nach Beispiel 5.16 *und der Quotientenregel ist f differenzierbar und*

$$\bigwedge_{x \in D(f)} \quad f'(x) = \frac{P'(x)Q(x) - P(x)Q'(x)}{(Q(x))^2} \ .$$

f' ist also wieder eine gebrochen rationale Funktion.

Die Sätze 5.11 und 5.13 wollen wir etwas formaler aufschreiben. Hierzu führen wir für zwei reelle Funktionen u, v die Relation

$$u \prec v \ :\Leftrightarrow \ u \text{ ist Einschränkung von } v \tag{5.7}$$

ein. Dann lassen sich die Sätze 5.11 und 5.13 wie folgt formulieren:

Satz 5.18 *Für reelle Funktionen f, g gelten*

$$f' \pm g' \prec (f \pm g)' \ , \tag{5.8}$$

$$f'g + fg' \prec (fg)' \ , \tag{5.9}$$

$$\frac{f'g - fg'}{g^2} \prec \left(\frac{f}{g}\right)' \ , \tag{5.10}$$

$$(f' \circ g)g' \prec (f \circ g)' \ , \tag{5.11}$$

sofern der Definitionsbereich des Ausdrucks auf der linken Seite jeweils nicht leer ist.

Die Schreibweise in diesem Satz macht deutlich, dass eine Funktion u, die aus f und g mittels einer der Operationen $+$, $-$, ... , $\circ$ gebildet ist, in einem Punkt differenzierbar sein kann, obwohl die entsprechende Ableitungsregel nicht anwendbar ist.

Beispiel 5.19 *Nach der Kettenregel ist die durch*

$$f(x) = \mid x \mid^3$$

gegebene Funktion für $x \neq 0$ differenzierbar:

$$f'(x) = 3 \mid x \mid^2 \cdot \frac{x}{\mid x \mid} = 3x \cdot \mid x \mid \ .$$

Die Stelle $x = 0$ muss gesondert untersucht werden: für alle $t \in \mathbb{R}$ ist

$$f(t) = \mid t \mid^3 = (\mid t \mid t)t = f(0) + F(t)t$$

mit $F(t) := \mid t \mid t$, F ist stetig in 0 und $F(0) = 0$. Daher ist f auch in 0 differenzierbar mit Ableitung 0. Insgesamt ist also f differenzierbar und für alle $x \in \mathbb{R}$ gilt

$$f'(x) = 3 \mid x \mid \cdot x \ .$$

Wir führen nun *transzendente Funktionen* ein, ohne zunächst Angaben machen zu wollen, wie man Funktionswerte konkret berechnet. Die beiden folgenden Sätze werden erst nach und nach im Laufe dieses Kapitels bewiesen.

Satz und Definition 5.20 *Es gibt genau eine reelle Funktion f mit den Eigenschaften*

$$D(f) = \mathbb{R} \ , \tag{5.12}$$

$$f(0) = 1 \ , \tag{5.13}$$

$$f' = f \ . \tag{5.14}$$

Diese Funktion nennen wir die Exponentialfunktion *und schreiben in Zukunft* $\exp := f$.

Satz und Definition 5.21 *Es gibt genau ein Paar von reellen Funktionen $(C \, , \, S)$ mit den Eigenschaften*

$$D(C) = D(S) = \mathbb{R} \ , \tag{5.15}$$

$$C(0) = 1 \quad , \qquad S(0) = 0 \ , \tag{5.16}$$

$$C' = -S \ \wedge \ S' = C \ . \tag{5.17}$$

Diese beiden Funktionen C, S nennen wir Cosinus- und Sinusfunktion und schreiben in Zukunft $\cos := C$ *bzw.* $\sin := S$.

Es ist üblich, $\cos x := \cos(x)$ und $\sin x := \sin(x)$ sowie $\cos^n x := (\cos x)^n$ und $\sin^n x := (\sin x)^n$ zu schreiben.

Die Funktion exp bzw. das Funktionenpaar (cos , sin) sind als Lösungen eines sogenannten Anfangswertproblems zu einer gewöhnlichen Differentialgleichung bzw. zu einem gewöhnlichen Differentialgleichungssystem definiert. Es gibt recht allgemeine Sätze darüber, wann solche Probleme eindeutige Lösungen besitzen, und diese Sätze lassen sich auch in unserem Zusammenhang anwenden. Wir werden solche Aussagen im kommenden Semester kennenlernen.

Beispiel 5.22 *Auch wenn über Exponential–, Sinus– und Cosinusfunkton noch kaum etwas bekannt ist — insbesondere nichts, was letztere mit den aus der Schule bekannten Winkelfunktionen verbindet —, können wir doch Funktionen ableiten, in denen diese Funktionen als Bausteine auftreten. Ist etwa f gegeben durch*

$$f(x) := x^4 \exp(\cos^2 x) ,$$

so erhält man aus Produkt– und Kettenregel, dass f differenzierbar ist mit Ableitung

$$f'(x) = 4x^3 \exp(\cos^2 x) - 2x^4(\sin x)(\cos x) \exp(\cos^2 x) .$$

Ist f eine injektive differenzierbare Funktion und bezeichnen wir die Umkehrfunktion mit g, so kann man aus der Kettenregel folgern: Ist g in x differenzierbar, so ist $f'(g(x)) \neq 0$ und

$$g'(x) = \frac{1}{f'(g(x))} .$$

Denn für alle $x \in D(g)$ ist $f \circ g(x) = x$ und Differentiation liefert

$$f' \circ g(x) \cdot g'(x) = 1 .$$

In der Tat gilt

Satz 5.23 *Die Funktion f sei auf dem Intervall $D(f)$ stetig und injektiv. Es sei $g := f^{-1}$ mit Definitionsbereich $D(g) := W(f)$.*

(i) *Ist x innerer Punkt von $D(g)$ und ist f in $z := g(x)$ differenzierbar mit Ableitung $f'(z) \neq 0$, so ist g in x differenzierbar und*

$$g'(x) = \frac{1}{f'(z)} = \frac{1}{f' \circ g(x)} .$$

(ii) *Ist $f'(z) = 0$, so ist g nicht differenzierbar in x.*

Wie in Satz 5.18 kann man schreiben:

$$\frac{1}{f' \circ f^{-1}} \prec (f^{-1})' ,$$

wobei f^{-1} hier die Umkehrfunktion von f bezeichnet.

Beweis: Für alle $s \in D(f)$ gilt

$$f(s) = f(z) + F(s)(s - z)$$

mit einer in z stetigen Funktion $F : D(f) \to \mathbb{R}$, und $F(z) = f'(z)$. Beachtet man $z = g(x)$, so folgt für alle $t \in D(g)$

$$f(g(t)) = f(g(x)) + F(g(t))(g(t) - g(x))$$

und hieraus

$$t - x = (F \circ g)(t)(g(t) - g(x)) \ .$$

Da f injektiv und $F(z) \neq 0$ ist, besitzt $F \circ g$ keine Nullstelle, und es folgt für alle $t \in D(g)$

$$g(t) = g(x) + \frac{1}{F \circ g(t)}(t - x) \ .$$

Da mit f auch g stetig ist, ist $\frac{1}{F \circ g}$ eine in $D(g)$ stetige Funktion und

$$\frac{1}{F \circ g(x)} = \frac{1}{F(z)} = \frac{1}{f'(z)} \ .$$

Damit ist (i) bewiesen. Behauptung (ii) ist bereits mit der Vorbemerkung zu Satz 5.23 bewiesen.

q.e.d.

Beispiel 5.24 *Die Funktion*

$$g \ : \ [\,0\,,\,\infty\,) \ \longrightarrow \ \mathbb{R}\,, \ \ t \ \longmapsto \ t^{\frac{1}{n}}$$

mit $n \in \mathbb{N}$ ist die Umkehrfunktion von

$$f \ : \ [\,0\,,\,\infty\,) \ \longrightarrow \ \mathbb{R}\,, \ \ z \ \longmapsto \ z^n \ .$$

Da f differenzierbar ist in $(\,0\,,\,\infty\,)$, ist auch g in $(\,0\,,\,\infty\,)$ differenzierbar, und für $x > 0$ gilt

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{n[g(x)]^{n-1}} = \frac{1}{n} x^{\frac{1}{n}-1}$$

Beispiel 5.25 *Mit $\alpha \in \mathbb{Q}$ sei f gegeben durch*

$$f(x) := x^{\alpha} \ .$$

Wir nehmen $\alpha \notin \mathbb{Z}$ an und stellen es dar in der Form

$$\alpha = \frac{p}{q}\,, \quad p \in \mathbb{Z} \quad , \quad q \in \mathbb{N} \ .$$

Dann ist f in $(0, \infty)$ differenzierbar und Beispiele 5.14, 5.15 *und* 5.24 *sowie die Kettenregel liefern wegen $f = F \circ G$ mit $G(x) := x^{\frac{1}{q}}$ und $F(z) := z^p$:*

$$f'(x) = F'(G(x))G'(x) = p(x^{\frac{1}{q}})^{p-1} \cdot \frac{1}{q} x^{\frac{1}{q}-1} = \alpha x^{\alpha-1} \ .$$

Damit sind wir in der Lage, algebraische Funktionen zu differenzieren; das sind Funktionen, die aus der Identität und konstanten Funktionen durch addieren, multiplizieren, dividieren, potentzieren mit rationalen Exponenten und hintereinanderschalten gebildet werden können. Man beachte in diesem Zuammenhang die Komplikation, die aus der Tatsache entsteht, dass für positives, nicht ganzzahliges rationales α die Funktion f nicht in 0 differenzierbar ist.

5.3 Der Mittelwertsatz der Differentialrechnung

Wir wollen uns nun überlegen, welche qualitativen Auswirkungen die Ableitung auf den Verlauf des Graphen einer Funktion hat. Ausgehend von Satz 5.3(iii) denken wir uns eine in $\xi \in D(f)$ differenzierbare Funktion f dargestellt in der Form

$$\bigwedge_{t \in D(f)} \quad f(t) = f(\xi) + F(t)(t - \xi)$$

mit einer in ξ stetigen Funktion $F : D(f) \to \mathbb{R}$, mit Wert $f'(\xi)$ in ξ. Ist nun $f'(\xi) > 0$ (bzw. $f'(\xi) < 0$), so ist wegen der Stetigkeit $F > 0$ (bzw. $F < 0$) in einer Umgebung $U_\delta(\xi)$ von ξ. Folglich gilt für alle $t \in U_\delta(\xi) \setminus \{\xi\}$

$$(f(t) - f(\xi))(t - \xi) = F(t)(t - \xi)^2 > 0 \quad (\text{bzw.} < 0) \ .$$

Wir erhalten also

Satz 5.26 *Ist f differenzierbar in $\xi \in D(f)$ und ist $f'(\xi) > 0$ (bzw. $f'(\xi) < 0$), so gibt es eine Zahl $\delta > 0$ so, dass gelten*

(i) $\displaystyle\bigwedge_{x \in (\xi, \xi+\delta)} \quad f(x) > f(\xi) \quad$ *(bzw. $f(x) < f(\xi)$)*

(ii) $\displaystyle\bigwedge_{x \in (\xi-\delta, \xi)} \quad f(x) < f(\xi) \quad$ *(bzw. $f(x) > f(\xi)$)*

Ist also $f'(\xi) > 0$, so sind in einer Umgebung von ξ die Werte an Stellen "links von ξ" kleiner als der Wert in ξ, während die Werte an Stellen "rechts von ξ" größer als der Wert in ξ sind.

Hieraus folgt aber: Nimmt f in ξ sein Maximum oder Minimum an, so ist $f'(\xi) = 0$. Dies ist bereits richtig für lokale Extrema:

Definition 5.27 *Die reelle Funktion f besitzt in $x_0 \in D(f)$ ein lokales (oder relatives) Minimum (bzw. Maximum), wenn es eine Umgebung $U_\delta(x_0)$ von x_0 so gibt, dass*

$$\bigwedge_{x \in D(f) \cap U_\delta(x_0)} f(x) \geq f(x_0) \quad (f(x) \leq f(x_0)) \, .$$

Den Wert $f(x_0)$ *nennt man das lokale (oder auch relative) Minimum bzw. Maximum. Besitzt f in x_0 ein relatives Minimum oder ein relatives Maximum, so sagt man auch: f besitzt in x_0 ein relatives (lokales) Extremum. x_0 nennt man Extremalstelle bzw. Minimal- oder Maximalstelle.*

Beachten Sie, dass das Minimum bzw. Maximum einer Funktion f auch ein lokales Minimum bzw. Maximum ist, nicht aber umgekehrt.

In folgender Skizze ist f auf $I_1 \cup I_2 \cup \{w\}$ erklärt. Die Punkte p, q, r, s, t, u, v, w sowie alle Punkte aus I_2 sind Extremalstellen von f. In p, r, t, v, w sowie in jedem Punkt des Intervalls I_2 liegen lokale Maxima; in q, s, u, w sowie in jedem Punkt von I_2 liegen lokale Minima. Das Maximum von f liegt in v, das Minimum in s und u.

Aus Satz 5.26 gewinnt man ein notwendiges Kriterium für das Vorliegen eines Extremums:

Satz 5.28 *Die reelle Funktion f besitze in $x_0 \in D(f)$ ein lokales Extremum. Dann gilt (i) oder (ii):*

(i) *f ist in x_0 nicht differenzierbar*

(ii) *f ist in x_0 differenzierbar und $f'(x_0) = 0$.*

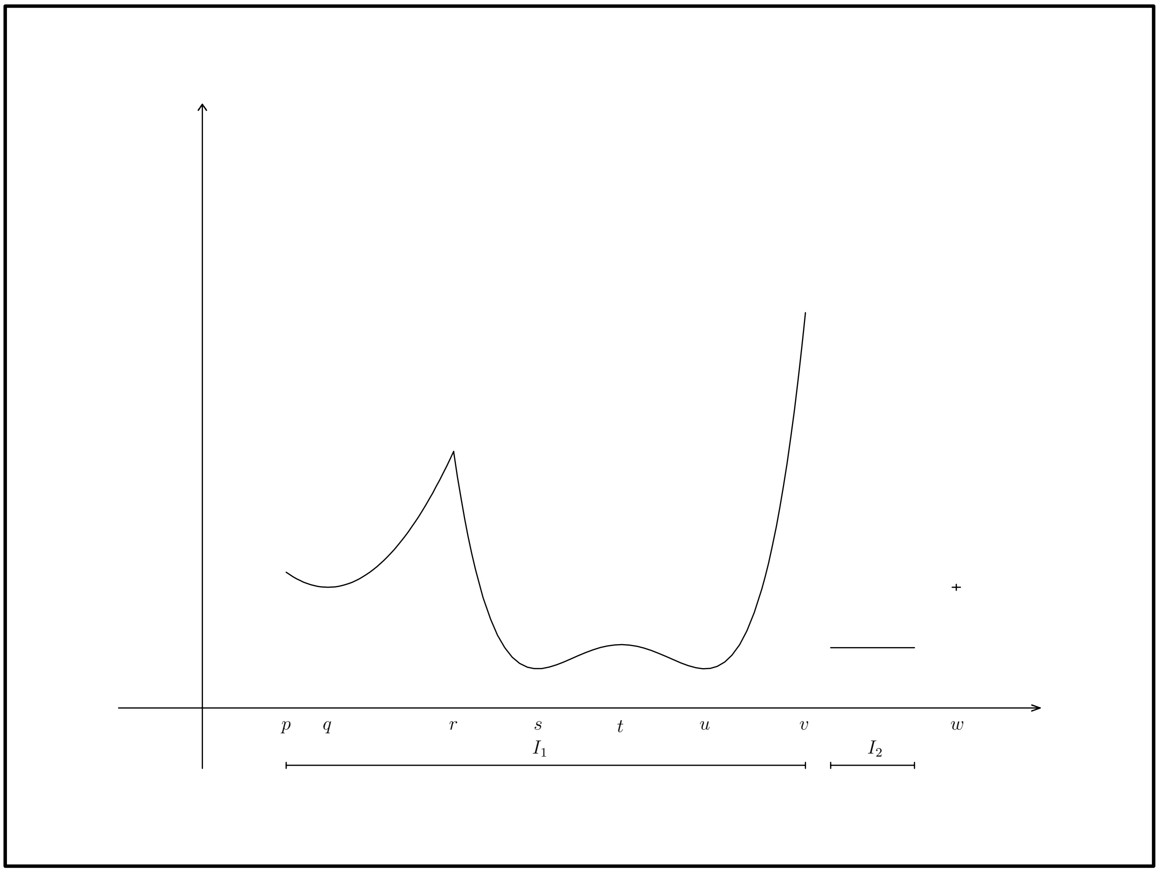


Abbildung zu Definition 5.27

Zunächst erhalten wir hieraus ein Verfahren zur Bestimmung des Maximums und des Minimums einer Funktion auf einem kompakten Intervall.

Beispiel 5.29 *Bestimme Maximum und Minimum von*

$$f \ : \ [\,-1\,,\,2\,] \longrightarrow \mathbb{R}\,,\ x \longmapsto |\,x\,| + x^3\ .$$

Lösung: *Da f stetig ist und $D(f) = [\,-1\,,\,2\,]$ kompakt ist, müssen* $\min f$ *und* $\max f$ *angenommen werden. Nach Satz* 5.28 *kann dies nur in Punkten geschehen, in denen f nicht differenzierbar ist oder in denen f die Ableitung* 0 *besitzt. Das liefert eine Kandidatenliste: f ist nicht differenzierbar in -1, in 0 und in 2,*

$$f'(x) = \frac{x}{|\,x\,|} + 3x^2 = \left\{ \begin{array}{rcr} -1+3x^2 & in & (\,-1\,,\,0\,) \\ 1+3x^2 & in & (\,0\,,\,2\,) \end{array} \right\} \overset{!}{=} 0 \Leftrightarrow x = -\frac{1}{3}\sqrt{3}\ .$$

Die Kandidatenliste ist also

$$K := \{-1, -\frac{1}{3}\sqrt{3}, 0, 2\}\ ,$$

und in einem der vier Punkte aus dieser Liste muss f jeweils Minimum und Maximum annehmen. Wo das ist, erkennt man durch einen Wertevergleich:

$$f(-1) = 0\,; \quad f(-\frac{1}{3}\sqrt{3}) = \frac{2}{9}\sqrt{3}\,; \quad f(0) = 0\,; \quad f(2) = 10\,.$$

Also ist 0 *das Minimum von* f*, und dieses wird bei* -1 *und bei* 0 *angenommen; das Maximum von* f *wird bei* 2 *angenommen und hat den Wert* 10*.*

Man erkennt übrigens auch, dass f *in* $-\frac{1}{3}\sqrt{3}$ *ein lokales Maximum annimmt, denn mit gleichem Argument wie oben sieht man, dass* $f_{|\,[-1\,,\,0]}$ *sein Maximum in* $-\frac{1}{3}\sqrt{3}$ *annimmt.*

Beispiel 5.30 *Man bestimme Supremum und Infimum von*

$$f\;:\;\mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \frac{|\,x\,|}{1+x^2}\,.$$

Handelt es sich um ein Maximum bzw. ein Minimum?

Lösung: *Die Funktion liegt in* $C(\mathbb{R})$ *und ist differenzierbar in* $\mathbb{R} \setminus \{0\}$ *und für* $x \neq 0$ *ist*

$$f'(x) = \frac{\frac{x}{|x|}(1+x^2) - 2x\,|\,x\,|}{(1+x^2)^2} = \frac{(1-x^2)x}{|\,x\,|\,(1+x^2)^2}\,.$$

Kandidaten für Extrema sind daher die Nullstellen 1 *und* -1 *der Ableitung sowie* 0 *als Punkt, in dem* f *nicht differenzierbar ist:*

$$K = \{-1\,;\,0\,;\,1\}$$

(In der Tat ist f *in* 0 *nicht differenzierbar*[1]*. Anderenfalls wäre* r*, darstellbar durch*

$$r(x) := |\,x\,| = (1+x^2)f(x) \quad ,$$

in 0 *differenzierbar.)*

Es gelten $f(-1) = f(1) = 1/2\;, \quad f(0) = 0$*. Wegen* $f \geq 0$ *nimmt* f *in* 0 *sein Minimum* 0 *an. Ferner ist*

$$\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0\,.$$

Daher existiert eine Zahl $R > 1$ *mit*

$$f(x) < 1/2 \quad \textit{für} \quad |\,x\,| > R\,.$$

[1] Ob in einem Punkt f tatsächlich nicht differenzierbar ist, ist für die Aufnahme in die Kandidatenliste unerheblich. Wichtig ist nur, dass jeder Punkt, in dem f nicht differenzierbar ist, in der Kandidatenliste auftaucht. Für die Aufnahme in die Liste genügt also der Verdacht.

Es folgt

$$\max f|_{[\,-R\,,\,R\,]} = 1/2 \quad , \quad f(-1) = f(1) = 1/2$$

sowie

$$f|_{\mathbb{R}\setminus[\,-R\,,\,R\,]} < 1/2\ .$$

Daher ist $1/2$ *das Maximum von* f *und wird in* 1 *und in* -1 *angenommen.*

Man führe den Beweis des folgenden Satzes selbst aus.

Satz 5.31 *Sei* f *eine auf dem offenen Intervall* $I = (\,a\,,\,b\,)$ *definierte stetige Funktion,* $\infty \leq a < b \leq \infty$. *In den Punkten* $x_1, \ldots, x_K \in I$ *sei* f *nicht differenzierbar oder habe Ableitung* 0. *Es mögen*

$$A := \hat{\mathbb{R}} - \lim_{x\to a} f(x)\ , \quad B := \hat{\mathbb{R}} - \lim_{x\to b} f(x)$$

existieren. Dann ist das größte der Elemente $A, f(x_1), \ldots, f(x_K), B$ *das Supremum von* f, *und das kleinste von* $A, f(x_1), \ldots, f(x_K), B$ *ist das Infimum von* f. *Ist für ein* $k \in \{1, \ldots, K\}$ *die Zahl* $f(x_k)$ *das Supremum bzw. das Infimum, dann nimmt* f *in* x_k *sein Maximum bzw. Minimum an.*

Ein wesentliches Resultat für den Zusammenhang zwischen Funktion und Ableitung ist der Mittelwertsatz der Differentialrechnung, der auf Joseph Louis Lagrange (1736 - 1819) zurückgeht. Man leitet diesen Satz aus einem Spezialfall, dem Satz von Rolle her (Michel Rolle 1652 - 1719). Diesen stellen wir als Lemma voran.

Lemma 5.32 *(Rolle) Sei* $I := [\,a\,,\,b\,]$ *ein kompaktes Intervall und* $f \in C(I)$ *sei im Inneren von* I *differenzierbar, und es gelte* $f(a) = f(b)$. *Dann gibt es ein* $\xi \in (\,a\,,\,b\,)$ *mit*

$$f'(\xi) = 0\ .$$

Beweis:
Unter den angegebenen Voraussetzungen nimmt f im Inneren von I, etwa in ξ, sein Maximum oder sein Minimum an. Nach Satz 5.28 gilt an dieser Stelle

$$f'(\xi) = 0\ .$$

q.e.d.

Wir beseitigen nun die Voraussetzung $f(a) = f(b)$ und erhalten den Mittelwertsatz:

Satz 5.33 *Sei $-\infty < a < b < \infty$, und $f \in C([\,a\,,\,b\,])$ sei in $(\,a\,,\,b\,)$ differenzierbar. Dann gibt es ein $\xi \in (\,a\,,\,b\,)$, so dass*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \ .$$

Zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ gibt es also eine Tangente an den Graphen von f in einem Punkt zwischen a und b, welche parallel zur Sekante ist. Siehe nachfolgende Skizze

Beweis (von Satz 5.33**):** Die Sekante durch $(a, f(a)), (b, f(b))$ ist der Graph der durch $\ell(x) := f(a) + \frac{f(b)-f(a)}{b-a} \cdot (x - a)$ gegebenen affin linearen Funktion. Die Differenz $f - \ell =: \tilde{f}$ genügt den Voraussetzungen des Satzes von Rolle 5.32: $\tilde{f} \in C(I)$ ist in $\overset{\circ}{I}$ differenzierbar und

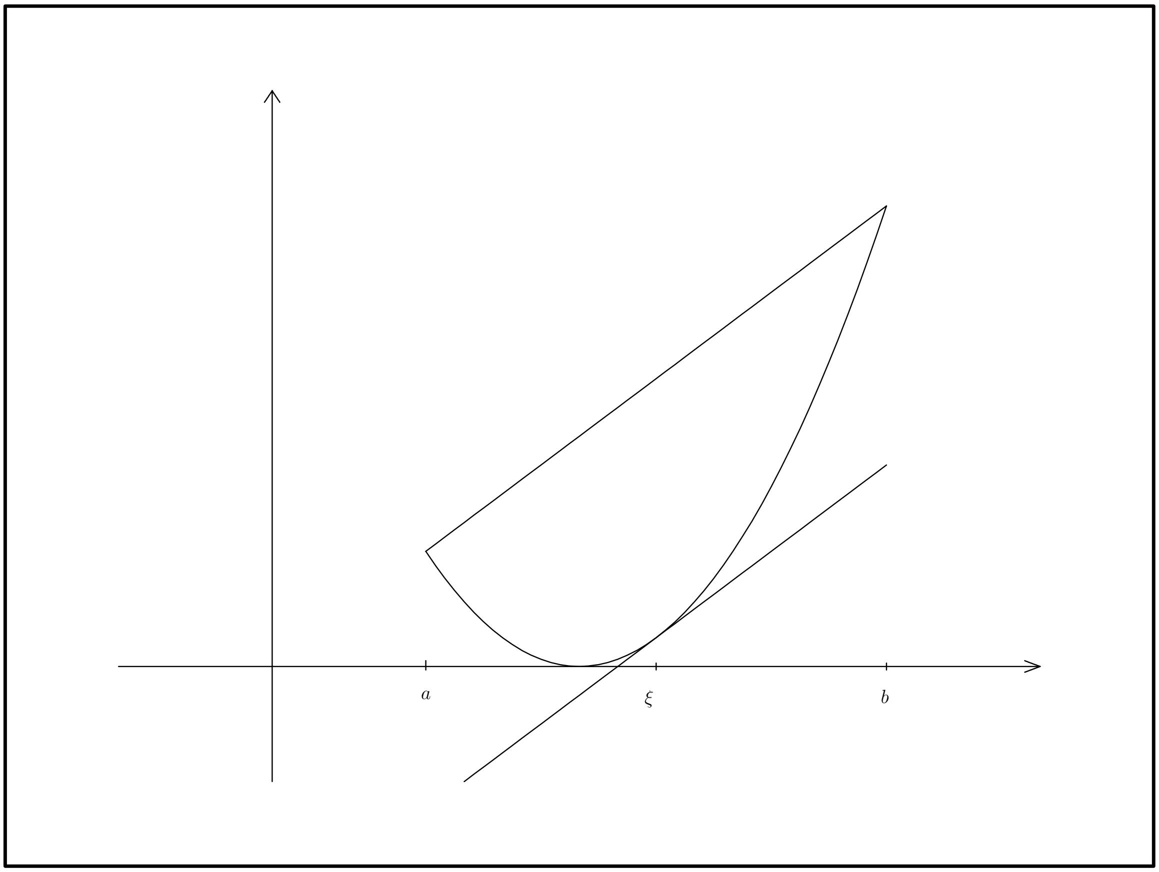
$$\tilde{f}(a) = \tilde{f}(b) = 0 \ .$$

Mithin existiert ein $\xi \in \overset{\circ}{I}$ mit

$$\tilde{f}'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0 \quad ,$$

und das war zu zeigen.

q.e.d.



Skizze zum Mittelwertsatz 5.33

Folgendes Korollar enthält zwei weitere Formulierungen des Mittelwertsatzes sowie die Mittelwertabschätzung:

Korollar 5.34 *Sei I ein Intervall, und $f \in C(I)$ sei differenzierbar in $\mathring{I}$, dem Inneren von I. Dann gelten*

(i) *Zu jedem x_1,$x_2 \in I$ mit $x_1 < x_2$ existiert ein $\xi \in (x_1, x_2)$ mit*

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)\ .$$

(ii) *Zu jedem $x \in I$ und $h \in \mathbb{R} \setminus \{0\}$ mit $x + h \in I$ gibt es ein $\theta \in (\,0\,,\,1\,)$ mit*

$$f(x + h) = f(x) + hf'(x + \theta h)\ .$$

(iii) *Ist f' eine beschränkte Funktion, so gilt mit*

$$L := \| f' \|_\infty := \sup_{\overset{\circ}{I}} | f' |$$

die Abschätzung

$$\bigwedge_{x_1, x_2 \in I} | f(x_2) - f(x_1) | \leq L \mid x_2 - x_1 \mid \ .$$

Ist in (ii) übrigens x innerer Punkt von (i), so kann man trivialerweise auch $h = 0$ zulassen.

Definition 5.35 *Gilt für eine reelle Funktion f mit einer Konstanten $L \geq 0$ die Abschätzung*

$$\bigwedge_{x_1, x_2 \in D(f)} | f(x_1) - f(x_2) | \leq L \mid x_2 - x_1 \mid \quad , \tag{5.18}$$

so nennt man f Lipschitzstetig *mit der* Lipschitzkonstanten L.
(Rudolf Lipschitz 1832–1903)

Beweis (von Satz 5.34**):**
(i) ist der Mittelwertsatz 5.33 angewandt auf das Intervall $[\, x_1 \, , \, x_2 \,]$.
(ii) erhält man, indem man je nach dem Vorzeichen von h in (i) setzt

$$x_1 := x \ , \quad x_2 := x + h \qquad \text{bzw. } x_1 := x + h, x_2 := x$$

und beachtet, dass dann jeder Punkt ξ aus (x_1, x_2) in der Form $\xi = x + \theta h$ mit einem $\theta \in (\, 0 \, , \, 1 \,)$ geschrieben werden kann.

(iii) folgt direkt aus (i).

q.e.d.

Aus Satz 5.34 erkennt man den Einfluss der Ableitung auf den Verlauf einer Funktion.

Satz 5.36 *Ist I ein Intervall und $f \in C(I)$ im Inneren von I differenzierbar, so gelten:*

(i) *Ist $f'(t) > 0$ bzw $(f'(t) < 0)$ für alle $t \in \overset{\circ}{I}$, so ist f streng monoton wachsend (bzw. streng monoton fallend).*

(ii) *f ist genau dann monoton wachsend (bzw. fallend), wenn $f'(t) \geq 0$ (bzw. $f'(t) \leq 0$) ist für jedes $t \in \overset{\circ}{I}$.*

(iii) *f ist genau dann eine konstante Funktion (d. h. $\bigvee_{c\in\mathbb{R}} \bigwedge_{x\in I} f(x)=c$), wenn $f'(t)=0$ ist für alle $t \in \overset{\circ}{I}$.*

Man beachte, dass in 5.34(i) *keine* Äquivalenz, sondern nur eine Implikation steht.

Beweis: Ist $f'>0$ in $\overset{\circ}{I}$, so liefert Korollar 5.34(i) für alle $x_1, x_2 \in I$ mit $x_1 < x_2$ und mit einem $\xi \in (\, x_1 \, , \, x_2 \,)$

$$f(x_2) = f(x_1) + f'(\xi)(x_2 - x_1) > f(x_1)\ . \tag{5.19}$$

Also ist f streng monoton wachsend.

Ist $f' \geq 0$ in $\overset{\circ}{I}$, so erhält man (5.19) mit $\geq$ statt $>$. f ist also monoton wachsend. Ist umgekehrt f monoton wachsend und wäre $f'(t)<0$ für ein $t \in \overset{\circ}{I}$, so erhielte man einen Widerspruch zu Satz 5.26. Die Behauptungen für fallende Monotonie erhält man analog.

Insbesondere ist f sowohl monoton wachsend als auch monoton fallend, falls $f'=0$ ist in $\overset{\circ}{I}$. Das aber ist genau (iii).

q.e.d.

5.4 Diskussion der transzendenten Funktionen

Der Mittelwertsatz, insbesondere der hieraus folgende Satz 5.36, hilft uns bei der Aufklärung des Verlaufs der Exponential– sowie der Sinus– und Cosinusfunktion.

Satz 5.37 **(i)** $\bigwedge_{x\in\mathbb{R}} \exp(x) > 0$

(ii) $\exp$ *ist streng monoton wachsend*

(iii) *Ist I ein Intervall und $f \in C(I)$ differenzierbar in $\overset{\circ}{I}$, so gilt*

$$\bigwedge_{x\in\overset{\circ}{I}} f'(x) = f(x) \Leftrightarrow \bigvee_{\alpha\in\mathbb{R}} \bigwedge_{x\in I} f(x) = \alpha\exp(x)$$

Insbesondere ist $\exp$ *durch* (5.12), (5.13) *und* (5.14) *eindeutig festgelegt.*

(iv) *Für alle $s,t\in\mathbb{R}$ gilt*

$$\exp(s+t) = \exp(s)\cdot\exp(t)\ .$$

Insbesondere ist

$$\exp(-t) = \frac{1}{\exp(t)}$$

für alle $t \in \mathbb{R}$.

(v) *Für alle $n \in \mathbb{N}_0$ und $x \in [\,0\,,\,\infty\,)$ gelten die Abschätzungen*

$$\sum_{j=0}^{n} \frac{x^j}{j!} \leq \exp(x) \leq \sum_{j=0}^{n} \frac{x^j}{j!} + \frac{x^{n+1}}{(n+1)!} \exp(x)$$

(vi) *Für jedes $\alpha \in \mathbb{Q}$[1] ist*

$$\lim_{x\to\infty} [\exp(-x)] x^\alpha = 0 \ .$$

Insbesondere ist

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

(vii) *Für jedes $\alpha \in \mathbb{Q}$ und $x \in \mathbb{R}$ ist*

$$(\exp x)^\alpha = \exp(\alpha x) \ .$$

Beweis: Wir zeigen zunächst

$$\bigwedge_{t\in\mathbb{R}} \quad \exp(t)\cdot\exp(-t) = 1 \quad :$$

Setze

$$g \ : \ \mathbb{R} \longrightarrow \mathbb{R} \, , \ t \longmapsto \exp(t)\cdot\exp(-t) \ .$$

Dann ist $g(0) = \exp(0)\cdot\exp(0) = 1$ und für alle $t \in \mathbb{R}$

$$g'(t) = \exp(t)\exp(-t) - \exp(t)\exp(-t) = 0 \ .$$

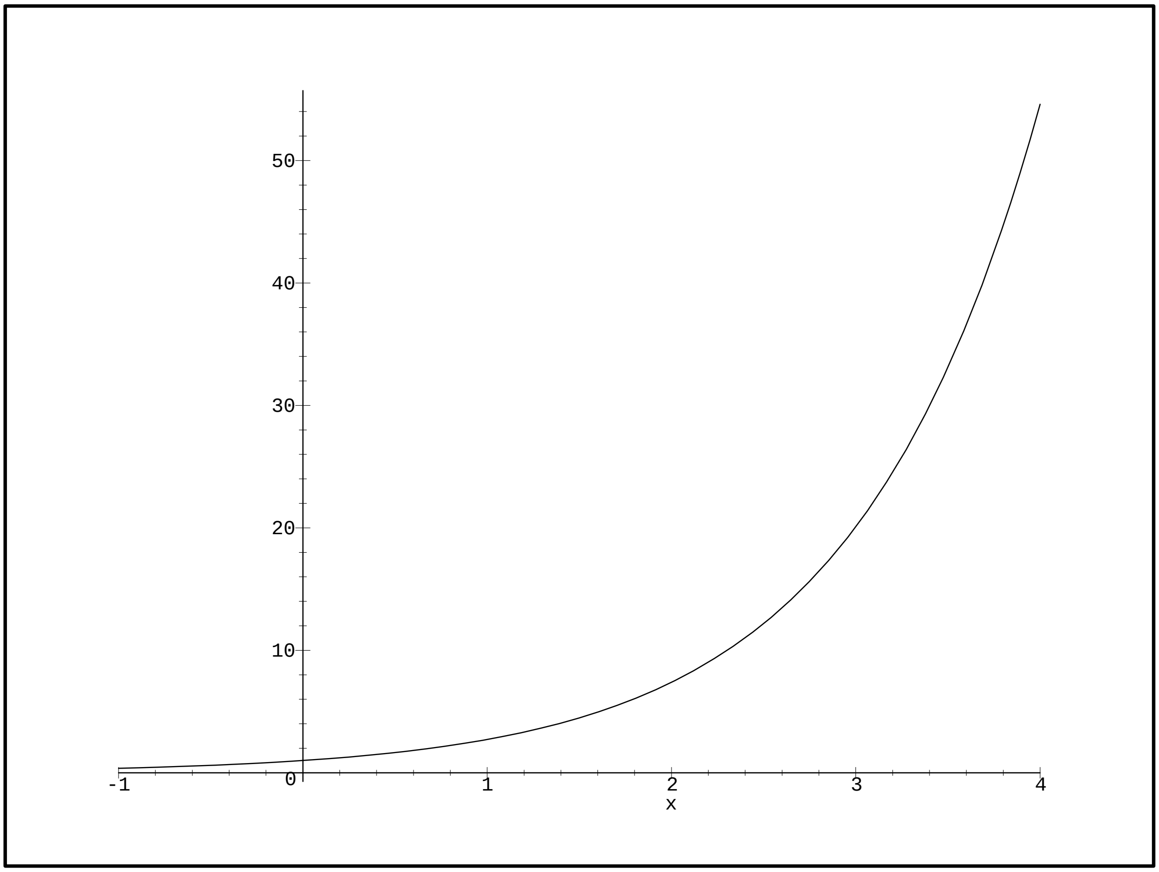
Also ist g konstante Funktion gemäß 5.36(iii) und zwar

$$\bigwedge_{t\in\mathbb{R}} \quad g(t) = g(0) = 1 \ .$$

Insbesondere kann exp keine Nullstelle besitzen. Da exp stetig ist und im Nullpunkt positiv, ist $\exp(x) > 0$ für alle $x \in \mathbb{R}$. Damit ist (i) gezeigt.

[1] Das gilt auch für $\alpha \in \mathbb{R}$, was wir bald definieren werden.

Der Graph der Exponentialfunktion:



Man beachte, dass die beiden Achsen verschieden skaliert sind. Die Ableitung der Exponentialfunktion in 0 ist 1, auch wenn das hier nicht so aussieht

(ii) folgt mit 5.36(i) aus $\exp' = \exp$ und (i).

Zum Beweis von (iii) "$\Rightarrow$" definieren wir

$$g \;:\; I \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \frac{f(x)}{\exp(x)}$$

und erhalten wegen $f' = f$ in $\overset{\circ}{I}$:

$$\bigwedge_{x \in \overset{\circ}{I}} \quad g'(x) = \frac{f'(x)\exp(x) - f(x)\exp(x)}{(\exp(x))^2} = 0\ .$$

Nach 5.36(iii) existiert $\alpha \in \mathbb{R}$, so dass $g(x) = \alpha$ ist für alle $x \in I$. Aber genau das war zu zeigen. Die umgekehrte Richtung in (iii) ist trivial.

Zum Beweis von (iv) fixieren wir $s \in \mathbb{R}$ und definieren

$$f \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ t \longmapsto \exp(s+t)\ .$$

Dann ist für alle $t \in \mathbb{R}$

$$f'(t) = \exp(s+t) = f(t) \quad ,$$

nach (iii) also

$$\bigwedge_{t\in\mathbb{R}} \quad f(t) = \alpha \exp(t)$$

mit einem $\alpha \in \mathbb{R}$. Man erhält α aus

$$\alpha = f(0) = \exp(s+0) = \exp(s)\ .$$

Damit ist (iv) bewiesen.

(v) beweist man mit vollständiger Induktion: Setze hierzu für $n \in \mathbb{N}_0$ und $x \in \mathbb{R}$

$$E_n(x) := \sum_{j=0}^{n} \frac{x^j}{j!}\ .$$

Für $n = 0$ erhält man wegen (ii)

$$\bigwedge_{x\geq 0} \quad \exp(x) \geq \exp(0) = 1 = E_0(x)\ .$$

Wir nehmen an, dass für irgendein $n \in \mathbb{N}_0$ gilt

$$\bigwedge_{x\geq 0} \quad \exp(x) - E_n(x) \geq 0\ . \tag{5.20}$$

Man rechnet leicht nach, dass $E'_{n+1}(x) = E_n(x)$ ist für alle $x \in \mathbb{R}$. Nun ist

$$\exp(0) - E_{n+1}(0) = 0$$

und für $x > 0$ nach 5.36(i) mit einem $\xi \in (\,0\,,\,x\,)$

$$(\exp(x) - E_{n+1}(x)) = x \cdot (\exp(\xi) - E_n(\xi)) \geq 0$$

wegen (5.20).

Analog weist man die Abschätzung nach oben nach: Definiere mit $n \in \mathbb{N}_0$ die Funktion R_n durch

$$R_n(x) := E_n(x) + \frac{x^{n+1}}{(n+1)!}\exp(x) - \exp(x)\ .$$

Dann ist $R_n(0) = 0$ und für alle $x \in \mathbb{R}^+$

$$R_0'(x) = x\exp(x) > 0 \ .$$

Folglich ist $R_0 > 0$ für $x > 0$.

Ist $R_{n-1} \geq 0$ für alle $x \geq 0$ und irgendein $n \in \mathbb{N}$, so folgt für $x > 0$

$$\begin{aligned} R_n'(x) &= E_{n-1}(x) + \frac{x^n}{n!}\exp(x) - \exp(x) + \frac{x^{n+1}}{(n+1)!}\exp(x) \\ &= R_{n-1}(x) + \frac{x^{n+1}}{(n+1)!}\exp(x) > 0 \ . \end{aligned}$$

Es folgt $R_n(x) > 0$ für $x > 0$.

Zum Beweis von (vi) sei $\alpha \in \mathbb{Q}$ und $n \in \mathbb{N}$ mit $n > \alpha$. Dann gilt für alle $x > 0$ mit E_n wie oben

$$0 \leq x^\alpha \exp(-x) = \frac{x^\alpha}{\exp(x)} \leq \frac{x^\alpha}{E_n(x)} \to 0 \quad \text{für} \quad x \to \infty \ .$$

Folglich ist $\lim\limits_{x\to\infty} x^\alpha \exp(-x) = 0$. Mit $\alpha = 0$ und wegen $\exp(-x) = \frac{1}{\exp(x)}$ folgen

$$\hat{\mathbb{R}} - \lim_{x\to\infty} \exp(x) = \infty \quad , \quad \lim_{x\to-\infty} \exp(x) = 0 \quad ,$$

und wegen $W(\exp) \subset (\,0\,,\,\infty\,)$ sogar

$$W(\exp) = (\,0\,,\,\infty\,)$$

nach dem Zwischenwertsatz 4.21.

Schließlich erhält man (vii) für $\alpha \neq 0$, indem man

$$g \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ x \longmapsto \frac{(\exp x)^\alpha}{\exp(\alpha x)}$$

differenziert,

$$\bigwedge_{x\in\mathbb{R}} \quad g'(x) = \frac{\alpha(\exp(x))^{\alpha-1}\exp(x)\exp(\alpha x) - \alpha\exp(\alpha x)(\exp x)^\alpha}{(\exp(\alpha x))^2} = 0$$

und $g(0) = 1$ beachtet und Satz 5.36(iii) verwendet. Für $\alpha = 0$ ist die Behauptung trivial.

q.e.d.

Da exp streng monoton wachsend ist, besitzt exp eine (ebenfalls streng monoton wachsende) Umkehrfunktion, die man den (natürlichen) Logarithmus nennt:

Satz und Definition 5.38 *Die Umkehrfunktion von* exp

$$\ln \ : \ \mathbb{R}^+ \longrightarrow \mathbb{R}\,, \ \exp(t) \longmapsto t$$

heißt (natürlicher) Logarithmus. Hierfür gelten:

(i) $W(\ln) = \mathbb{R}$ *sowie*

$$\hat{\mathbb{R}} - \lim_{x\to 0} \ln x = -\infty \ ,$$
$$\hat{\mathbb{R}} - \lim_{x\to \infty} \ln x = \infty \ .$$

(ii) *Für alle* $x > 0$ *ist*

$$\ln'(x) = \frac{1}{x} \ .$$

(iii)

$$\ln(1) = 0 \ .$$

(iv) *Für alle* $s, t \in \mathbb{R}^+$ *ist*

$$\ln(st) = \ln s + \ln t \ .$$

Insbesondere ist für jedes $t > 0$

$$\ln \frac{1}{t} = -\ln t \ .$$

(v) *Für alle* $a > 0$ *und* $\tau \in \mathbb{Q}$ *ist*

$$a^\tau = \exp(\tau \ln a)$$

Dies führt auf die Definition von Potenzen mit positiver Basis und beliebigem reellen Exponenten: Für $a \in \mathbb{R}^+$ *und* $\tau \in \mathbb{R}$ *ist*

$$a^\tau := \exp(\tau \ln a) \ .$$

Insbesondere ist für jedes $x \in \mathbb{R}$

$$\exp(x) = e^x \ .$$

Ist $\tau > 0$*, so setzt man auch* $0^\tau := 0$*.*

(vi) *Für alle* $s > 0$ *und* $\tau \in \mathbb{R}$ *ist*

$$\ln(s^\tau) = \tau \ln s \ .$$

(vii) *Für alle $\tau > 0$ gelten*

$$\lim_{x\to\infty} \frac{\ln x}{x^\tau} = \lim_{x\to 0} x^\tau \ln x = 0 \; .$$

Der Graph des Logarithmus ist umseitig abgebildet.

Korollar 5.39 **(i)** *Für Potenzen mit reellen Exponenten gelten die Potenzgesetze: Für alle $a, b \in \mathbb{R}^+$ und $\sigma, \tau \in \mathbb{R}$ gelten*

$$a^{\tau+\sigma} = a^\tau \cdot a^\sigma \; , \quad (a^\tau)^\sigma = a^{\tau\sigma} \; , \quad (ab)^\tau = a^\tau b^\tau$$

(ii) *Mit $\tau \in \mathbb{R}$ sei f definiert durch $f(x) = x^\tau$. f ist differenzierbar in $(\,0\,,\,\infty\,)$ und für alle $x \in (\,0\,,\,\infty\,)$ gilt*

$$f'(x) = \tau x^{\tau-1} \; .$$

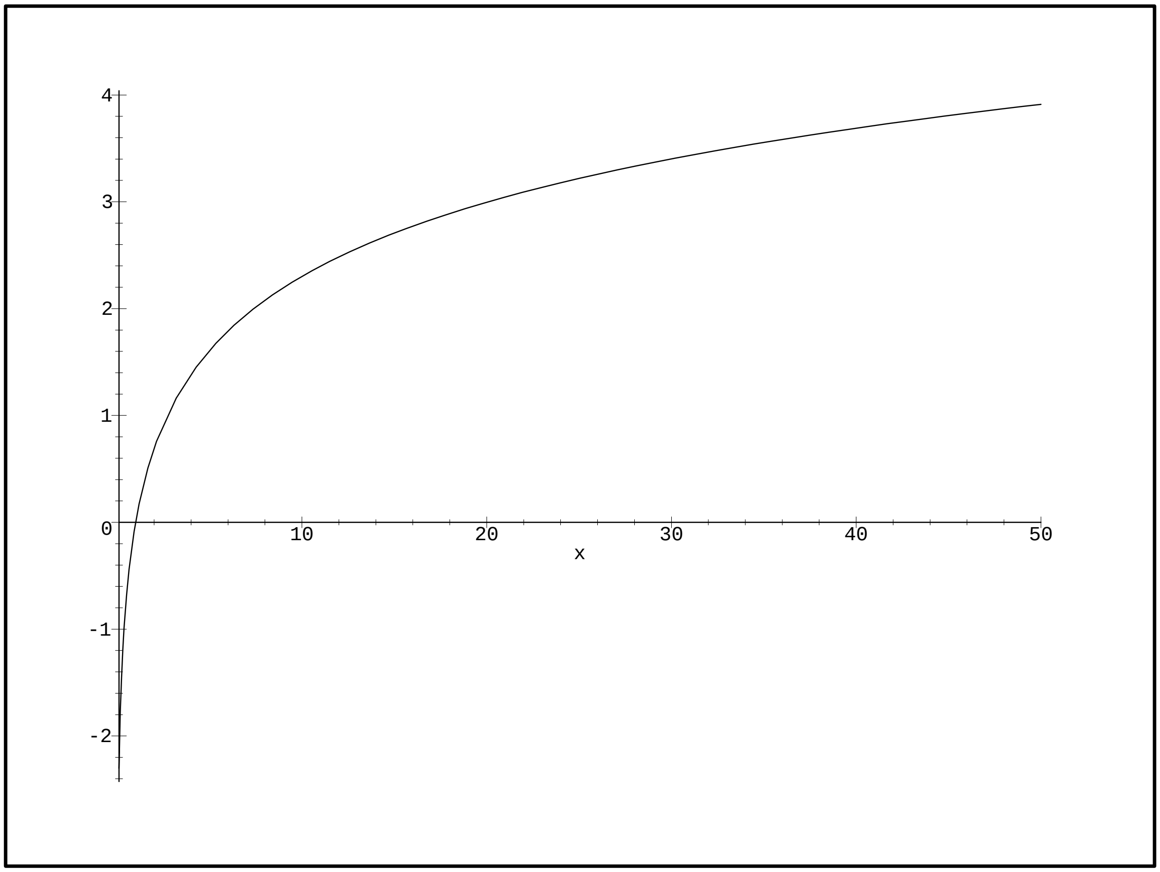
(iii) *Für alle $\tau > 0$ gilt $\lim\limits_{x\to 0} x^\tau = 0$.*

Man beachte: Ist $a > 0$ und $a \neq 1$ so ist $f_a \; : \; \mathbb{R} \longrightarrow \mathbb{R}\,,\; x \longmapsto a^x$ injektiv. Man nennt f_a eine Exponentialfunktion zur Basis a und ihre Umkehrfunktion den Logarithmus zur Basis a. Man schreibt dafür $\log_a$ und im Fall $a = 10$ auch $\lg := \log_{10}$. Zwischen $\log_a$ und $\ln$ besteht die Beziehung

$$\log_a x = \frac{\ln x}{\ln a}$$

für jedes $x \in \mathbb{R}^+$.

Der Graph der Logarithmusfunktion:



Man beachte, dass die beiden Achsen verschieden skaliert sind. Die Ableitung der Logarithmusfunktion in 1 ist 1, auch wenn das hier nicht so aussieht

Beweis (von Satz 5.38**):** (i) folgt aus $W(\ln) = D(\exp)$ sowie der strengen Monotonie.

Nach der Differentiationsregel für die Umkehrfunktion, Satz 5.23, erhält man (ii):

$$\bigwedge_{x \in \mathbb{R}^+} \quad \ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln x)} = \frac{1}{x} \, .$$

(iii) erhält man aus $\exp(0) = 1$, also

$$\ln(1) = \ln(\exp(0)) = 0 \, .$$

Zum Nachweis von (iv) setze man

$$s = \exp(x) \, , \quad t = \exp(y) \quad \text{, also} \quad x = \ln s \, , \quad y = \ln t \, .$$

Dann erhält man mit 5.37(iv):

$$\ln(st) = \ln(\exp(x) \cdot \exp(y)) = \ln(\exp(x+y)) = x + y = \ln s + \ln t\ .$$

(v) erhält man aus 5.37(vii) und aus $a = \exp(\ln a)$ für $a \in \mathbb{R}^+$; denn ln ist Umkehrfunktion von exp. Ersetzt man in (v) a durch s und wendet auf beide Seiten der Gleichheit ln an, so erhält man (vi).

Zum Nachweis von (vii) schreibt man

$$\lim_{x\to\infty} \frac{\ln x}{x^\tau} = \lim_{x\to\infty} \frac{1}{\tau} \frac{\tau \ln x}{\exp(\tau \ln x)} = \lim_{t\to\infty} \frac{1}{\tau} \frac{t}{\exp(t)} = 0$$

nach 5.37(vi), sowie

$$\lim_{x\to 0} x^\tau \ln x = \lim_{x\to 0} \frac{1}{\tau}[\exp(\tau \ln x)](\tau \ln x) = \lim_{t\to -\infty} \frac{t}{\tau} \exp(t) = 0$$

nach 5.37(vi).

q.e.d.

Beweis (von Korollar 5.39**):**
Die Potenzgesetze erhält man sofort aus der Definition und den Eigenschaften von exp.

Beim Nachweis von (ii) hat man die durch

$$f(x) = \exp(\tau \ln x)$$

definierte Funktion zu differenzieren. Mit der Kettenregel erhält man:

$$\begin{aligned} f'(x) &= \tau \cdot \frac{1}{x} \exp(\tau \ln x) = \tau \exp(-\ln x) \cdot \exp(\tau \ln x) \\ &= \tau \exp((\tau - 1) \ln x) = \tau x^{\tau - 1}\ . \end{aligned}$$

q.e.d.

Ähnlich diskutieren wir nun die trigonometrischen Funktionen Sinus und Cosinus.

Satz und Definition 5.40

(i) *Für alle* $t \in \mathbb{R}$ *ist* $\cos^2 t + \sin^2 t := [\cos(t)]^2 + [\sin(t)]^2 = 1$.

(ii) *Ist* I *Intervall und sind* $f, g \in C(I)$ *in* $\mathring{I}$ *differenzierbar mit*

$$f' = g\ , \quad g' = -f\ ,$$

so gibt es $\alpha, \beta \in \mathbb{R}$ derart, dass für alle $t \in I$ gelten

$$f(t) = \alpha \cos t + \beta \sin t \ , \quad g(t) = -\alpha \sin t + \beta \cos t \ .$$

Insbesondere sind sin *und* cos *durch die Bedingungen*

$$D(\sin) = D(\cos) = \mathbb{R} \ , \quad \sin(0) = 0 \ , \quad \cos(0) = 1 \ , \quad \cos' = -\sin$$

eindeutig bestimmt.

(iii) cos *ist gerade und* sin *ist ungerade* [1]

(iv) *Für alle $t, s \in \mathbb{R}$ gelten*

$$\sin(t+s) = \sin t \cos s + \cos t \sin s \tag{5.21}$$

$$\cos(t+s) = \cos t \cos s - \sin t \sin s \tag{5.22}$$

(v) *Für alle $n \in \mathbb{N}_0$ und $t \in [\,0\,,\,\infty\,)$ gelten:*

$$\sum_{k=0}^{2n+1} (-1)^k \frac{t^{2k}}{(2k)!} \le \cos t \le \sum_{k=0}^{2n} (-1)^k \frac{t^{2k}}{(2k)!} \tag{5.23}$$

$$\sum_{k=0}^{2n+1} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \le \sin t \le \sum_{k=0}^{2n} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \tag{5.24}$$

Insbesondere konvergieren die Cosinus– bzw. die Sinusreihe für jedes feste $t \in \mathbb{R}$ gegen $\cos t$ bzw. $\sin t$:

$$\cos t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} \ , \quad \sin t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \ . \tag{5.25}$$

(vi) *Die Funktion* cos *besitzt im Intervall $[\,0\,,\,\sqrt{6}\,]$ genau eine Nullstelle p und es gilt $p \in [\,\sqrt{2}\,,\,\sqrt{6-\sqrt{12}}\,]$.*
Wir definieren

$$\pi := 2p \ .$$

(vii) *Einige Werte:*

$$\begin{aligned}
\cos\frac{\pi}{2} &= 0 \quad ; \quad \sin\frac{\pi}{2} = 1 \ ; \\
\cos\pi &= -1 \quad ; \quad \sin\pi = 0 \ ; \\
\cos\frac{3\pi}{2} &= 0 \quad ; \quad \sin\frac{3\pi}{2} = -1 \ ; \\
\cos 2\pi &= 1 \quad ; \quad \sin 2\pi = 0 \ .
\end{aligned}$$

[1] Eine Funktion f heißt gerade (bzw. ungerade), wenn $D(f) = -D(f) := \{-t\colon t \in D(f)\}$ ist und für jedes $t \in D(f)$ gilt $f(t) = f(-t)$ (bzw. $f(t) = -f(-t)$)

(viii) *Für alle $t \in \mathbb{R}$ gelten*

$$\cos(t+2\pi) = \cos t \quad ; \quad \sin(t+2\pi) = \sin t \ ,$$

(Man sagt: cos *und* sin *sind 2π-periodisch.)*

$$\cos t = \sin(t+\pi/2) \quad ; \quad \cos t = -\cos(\pi - t) \ .$$

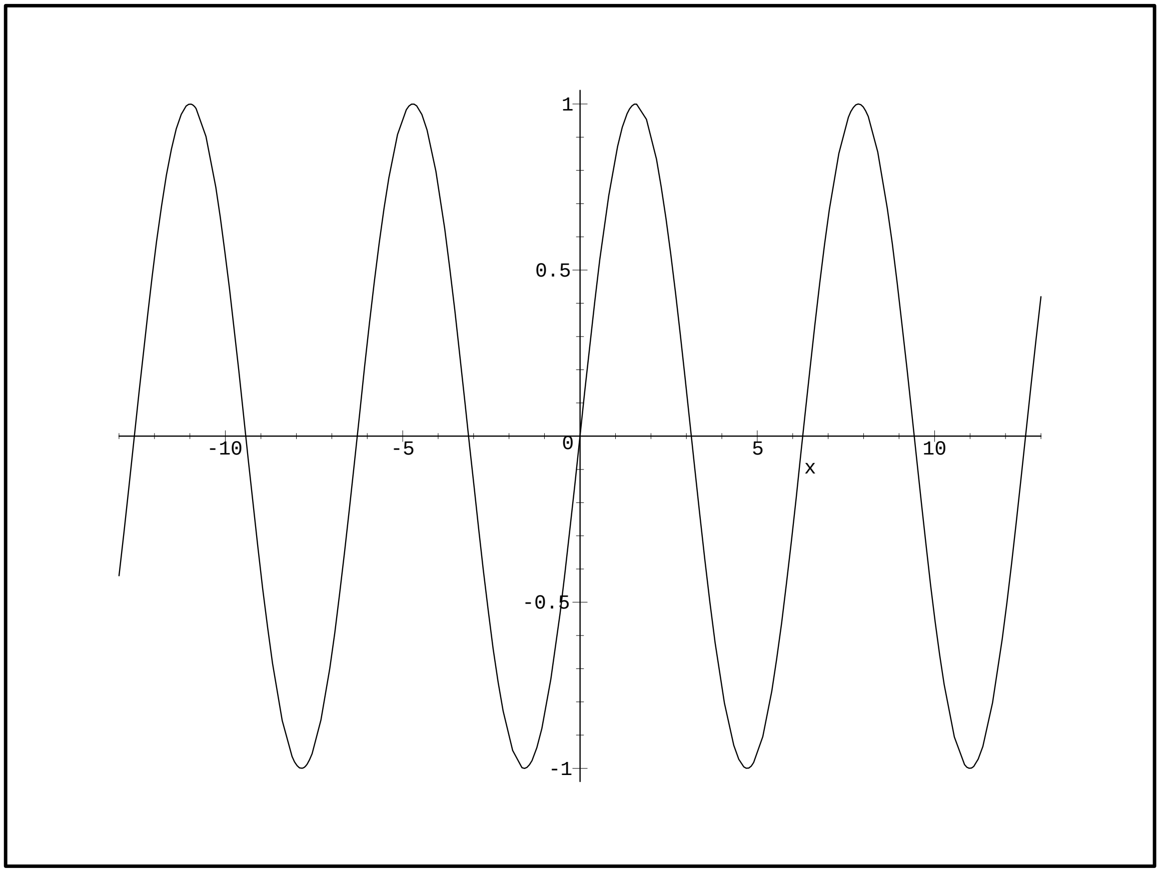
(ix) *Die Nullstellen und Extremstellen von* cos *und* sin *sind:*

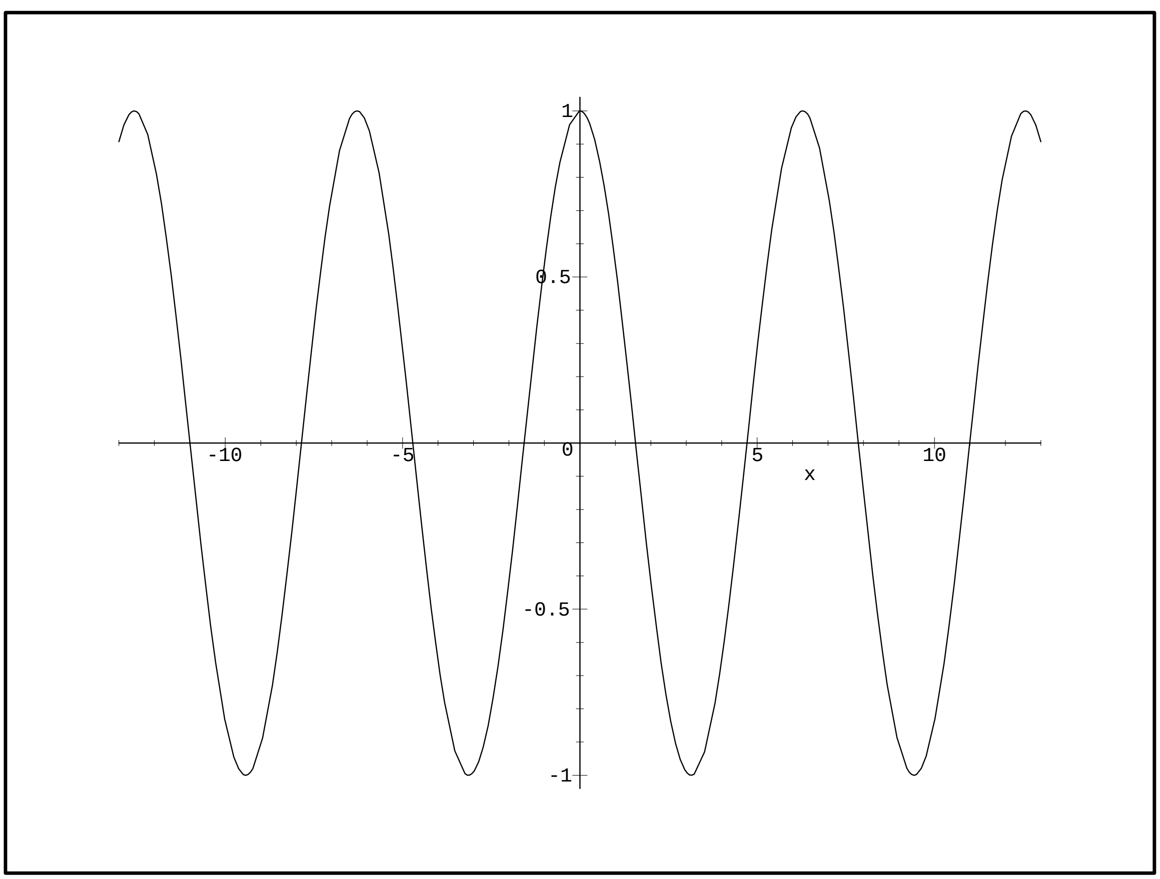
$$N(\cos) = \{\frac{\pi}{2} + k\pi \colon k \in \mathbb{Z}\} \quad , \quad N(\sin) = \{k\pi \colon k \in \mathbb{Z}\} \ ,$$
$$N(\cos -1) = \{2k\pi \colon k \in \mathbb{Z}\} \quad , \quad N(\sin -1) = \{\frac{\pi}{2} + 2k\pi \colon k \in \mathbb{Z}\} \ ,$$
$$N(\cos +1) = \{(2k+1)\pi \colon k \in \mathbb{Z}\} \quad , \quad N(\sin +1) = \{\frac{3\pi}{2} + 2k\pi \colon k \in \mathbb{Z}\} \ .$$

(x) cos *ist streng monoton fallend in* $[\,0\,,\,\pi\,]$ *und streng monoton wachsend in* $[\,\pi\,,\,2\pi\,]$. sin *ist streng monoton wachsend in* $[\,-\frac{\pi}{2}\,,\,\frac{\pi}{2}\,]$ *und streng monoton fallend in* $[\,\frac{\pi}{2}\,,\,\frac{3\pi}{2}\,]$*.*

In (vii) sind einige Werte angegeben. Mittels der Additionstheoreme (5.21), (5.22) und mittels 5.40(i) lassen sich weitere Werte berechnen. Auch kann man weitere Formeln wie in 5.40(viii) finden.

Umseitig sind die Graphen des Sinus (1) und des Cosinus (2) abgebildet. Man beachte, dass die beiden Achsen jeweils verschieden skaliert sind, die Skalierungen aber in beiden Abbildungen gleich sind. Die Nullstellen des Sinus sind alle ganzzahligen Vielfachen von π, die Nullstellen des Cosinus sind alle ungeraden Vielfachen von $\frac{\pi}{2}$ und $-\frac{\pi}{2}$.


1
0.5
-0.5
-1
-10
-5
0
5
10
x


1
0.5
-0.5
-1
-10
-5
0
5
10
x

Beweis:
Zu (i): Man differenziere die durch $f(t) = \cos^2 t + \sin^2 t$ gegebene Funktion und erhält $f' = 0$. Wegen $f(0) = \cos^2(0) + \sin^2(0) = 1$ folgt die Behauptung.

Zu (ii): Indem man das lineare Gleichungssystem

$$\begin{bmatrix} f(t) \\ g(t) \end{bmatrix} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} F(t) \\ G(t) \end{bmatrix}$$

nach $F(t)$, $G(t)$ auflöst, erkennt man: Zu f, g existieren Funktionen $F, G \in C(I)$, die in $\mathring{I}$ differenzierbar sind und mit denen gilt

$$f(t) = F(t)\cos t + G(t)\sin t\ , \quad g(t) = -F(t)\sin t + G(t)\cos t\ . \tag{5.26}$$

F, G errechnen sich nämlich zu

$$F(t) = f(t)\cos t - g(t)\sin t\ , \quad G(t) = f(t)\sin t + g(t)\cos t\ . \tag{5.27}$$

Wir differenzieren F und G, und erhalten wegen $f' = g$, $g' = -f$ für alle $t \in \mathring{I}$:

$$\begin{aligned} F'(t) &= -f(t)\sin t + f'(t)\cos t - g(t)\cos t - g'(t)\sin t = 0\ , \\ G'(t) &= f(t)\cos t + f'(t)\sin t - g(t)\sin t + g'(t)\cos t = 0\ . \end{aligned}$$

Nach Satz 5.36(iii) gibt es $\alpha, \beta \in \mathbb{R}$ so, dass für alle $t \in I$ gelten

$$F(t) = \alpha\ , \quad G(t) = \beta\ .$$

Setzt man dies in (5.26) ein, so erhält man die Behauptung.

Zu (iii): Definiert man f, g durch

$$f(t) = -\sin(-t)\ , \quad g(t) = \cos(-t)$$

so ist $D(f) = D(g) = \mathbb{R}$, $f(0) = 0$, $g(0) = 1$ sowie $f' = g$, $g' = -f$. Also ist $f = \sin$, $g = \cos$, und das war zu zeigen.

Zu (iv): Fixiere $s \in \mathbb{R}$ und definiere für $t \in \mathbb{R}$

$$f(t) = \sin(t + s)\ , \quad g(t) = \cos(t + s)\ . \tag{5.28}$$

Dann ist $f' = g$, $g' = -f$, $D(f) = D(g) = \mathbb{R}$; also gilt mit $\alpha, \beta \in \mathbb{R}$ für alle $t \in \mathbb{R}$

$$\bigwedge_{t \in \mathbb{R}} \quad f(t) = \alpha\sin t + \beta\cos t\ , \quad g(t) = \alpha\cos t - \beta\sin t\ . \tag{5.29}$$

α und β erhält man aus (5.28) und (5.29), indem man diese Gleichungen für $t = 0$ auswertet:

$$\beta = f(0) = \sin s\ , \quad \alpha = g(0) = \cos s\ .$$

Trägt man diese Werte für α und β in (5.29) ein, erhält man die Behauptung.

Zu (v): Wir definieren für gerade nicht–negative ganze Zahlen $2m, m \in \mathbb{N}_0$:

$$\bigwedge_{t\in\mathbb{R}} \quad P_{2m}(t) = \sum_{k=0}^{m} (-1)^k \frac{t^{2k}}{(2k)!}\ .$$

Für ungerade natürliche Zahlen $2m+1$, $m \in \mathbb{N}_0$ schreiben wir

$$\bigwedge_{t\in\mathbb{R}} P_{2m+1}(t) = \sum_{k=0}^{m} (-1)^k \frac{t^{2k+1}}{(2k+1)!}\ .$$

Für vorgegebenen $n \in \mathbb{N}_0$ ist dann P_n ein Polynom n-ter Ordnung und es gilt

$$\bigwedge_{n\in\mathbb{N}} \quad P_n'(t) = (-1)^{n-1} P_{n-1}(t) \tag{5.30}$$

Nun schreiben sich (5.23) und (5.24) in der Form

$$\bigwedge_{m\in\mathbb{N}_0} \quad \bigwedge_{n\in\mathbb{N}} \bigwedge_{t\in\mathbb{R}\setminus\{0\}} \quad \left\{ \begin{array}{ccccc} P_{4m+2}(t) & \leq & \cos t & \leq & P_{4m}(t) \\ P_{4m+3}(t) & \leq & \sin t & \leq & P_{4m+1}(t) \end{array} \right. \tag{5.31}$$

Wir zeigen zunächst:
Gilt für ein $m \in \mathbb{N}$ und alle $t \in [\,0\,,\,\infty\,)$

$$\cos t \leq P_{4m}(t)\ , \tag{5.32}$$

so gelten auch für alle $t \in [\,0\,,\,\infty\,)$

$$\sin t \leq P_{4m+1}(t)\ , \tag{5.33}$$

$$P_{4m+2}(t) \leq \cos t\ , \tag{5.34}$$

$$P_{4m+3}(t) \leq \sin t\ . \tag{5.35}$$

Zum Beweis definieren wir für $t \in [\,0\,,\,\infty\,)$

$$u_0(t) := P_{4m}(t) - \cos t\ , \quad u_1(t) = P_{4m+1}(t) - \sin t$$

$$u_2(t) := \cos t - P_{4m+2}(t)\ , \quad u_3(t) := \sin t - P_{4m+3}(t)$$

und berechnen

$$u_0(0) = u_1(0) = u_2(0) = u_3(0) = 0$$

sowie für alle $t \in (0\,,\,\infty)$

$$\begin{aligned}
u_1'(t) &= u_0(t)\ ,\\
u_2'(t) &= u_1(t)\ ,\\
u_3'(t) &= u_2(t)\ .
\end{aligned}$$

Satz 5.36 liefert, dass u_1 in $[0\,,\,\infty)$ monoton wächst, und daher $u_1(t) \geq 0$ für $t \geq 0$ erfüllt. Analog folgen dann, dass u_2 und schließlich u_3 für $t \geq 0$ ebenfalls größer oder gleich 0 sind, und das ist (5.33–5.35).

Nun beweist man (5.31) durch vollständige Induktion. Wegen Teil (i) des Satzes ist $\cos t \leq 1$ für alle t, und damit gilt $\cos t \leq P_0(t)$ für alle $t \geq 0$. Der obige Schluss von (5.32) auf (5.33–5.35) zeigt nun, dass die übrigen Ungleichungen im Fall $m = 0$ erfüllt sind.

Nimmt man nun (5.31) für irgendein $m \in \mathbb{N}_0$ als wahr an, so ist nach dieser Annahme

$$u_3 := \sin - P_{4m+3}$$

in $[0\,,\,\infty)$ nirgends negativ. Für

$$u_4 := P_{4m+4} - \cos$$

erhält man $u_4(0) = 0$ sowie für $t > 0$:

$$u_4'(t) = u_3(t) \geq 0\ .$$

Daher ist u_4 monoton wachsend in $[0\,,\,\infty)$ und insbesondere nirgends negativ. Der obige Schluss von (5.32) auf (5.33–5.35) liefert dann die übrigen Ungleichungen aus (5.31) für $m + 1$.

Zu (vi): Nach (5.24) mit $n = 0$ ist für $t \in [0\,,\,\infty)$

$$t\left(1 - \frac{t^2}{6}\right) \leq \sin t\ .$$

Für $t \in (0\,,\,\sqrt{6})$ erhält man hieraus

$$0 < \sin t\ .$$

Damit ist cos streng monoton fallend in $[0\,,\,\sqrt{6}]$ und kann in diesem Intervall höchstens eine Nullstelle besitzen.

Nun liefert (5.23) mit $n = 0$ bzw. $n = 1$ für $t \in [0\,,\,\infty)$

$$P_2(t) = 1 - \frac{t^2}{2} \leq \cos t \leq 1 - \frac{t^2}{2} + \frac{t^4}{24} = P_4(t)\ .$$

Mit $t = \sqrt{2}$ liefert die erste Ungleichung

$$0 \leq \cos\sqrt{2}$$

und mit $t = \sqrt{6 - \sqrt{12}}$ — das ist die kleinste positive Nullstelle von P_4 — liefert die zweite Ungleichung

$$\cos\sqrt{6 - \sqrt{12}} \leq 0 \ .$$

Als stetige Funktion muss cos in $[\,\sqrt{2}\,,\ \sqrt{6 - \sqrt{12}}\,]$ tatsächlich eine Nullstelle besitzen, die wir auf den Namen $\frac{\pi}{2}$ taufen.

Zu (vii): Da $\cos = \sin'$ positiv ist in $(\,0\,,\ \frac{\pi}{2}\,)$, und nach Teil (i) ist $\sin\frac{\pi}{2} = 1$. Die übrigen Werte erhält man dann mit den Additionstheoremen (5.21), (5.22).

Zu (viii): Erhält man aus den Additionstheoremen (5.21), (5.22) und aus (vii).

Bei den Beweisen von (ix) und (x) beachte man, dass man sich wegen der Perodizität auf das Intervall $[\,0\,,\ 2\pi\,]$ beschränken kann. Man kann dann alles mit (5.21), (5.22), (vii) und (viii) folgern.

q.e.d.

Unter Verwendung von sin und cos definiert man den Tangens und den Cotangens.

Satz und Definition 5.41 *Die Funktionen Tangens* (tan) *und Cotangens* (cot) *sind durch*

$$\tan x := \frac{\sin x}{\cos x}\ , \qquad \cot x := \frac{\cos x}{\sin x} = \frac{1}{\tan x}$$

definiert. Es gelten:

(i) $D(\tan) = \mathbb{R} \setminus N(\cos) = \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi\colon\ k \in \mathbb{Z}\}$
$D(\cot) = \mathbb{R} \setminus N(\sin) = \mathbb{R} \setminus \{k\pi\colon\ k \in \mathbb{Z}\}$

(ii) *Tangens und Cotangens sind π-periodisch, d. h.: für alle $t \in D(\tan)$ (bzw. für alle $t \in D(\cot)$) ist $t + \pi \in D(\tan)$ (bzw. $t + \pi \in D(\cot)$) und es gelten*

$$\tan(t + \pi) = \tan t \quad \textit{bzw.} \ \cot(t + \pi) = \cot t \ .$$

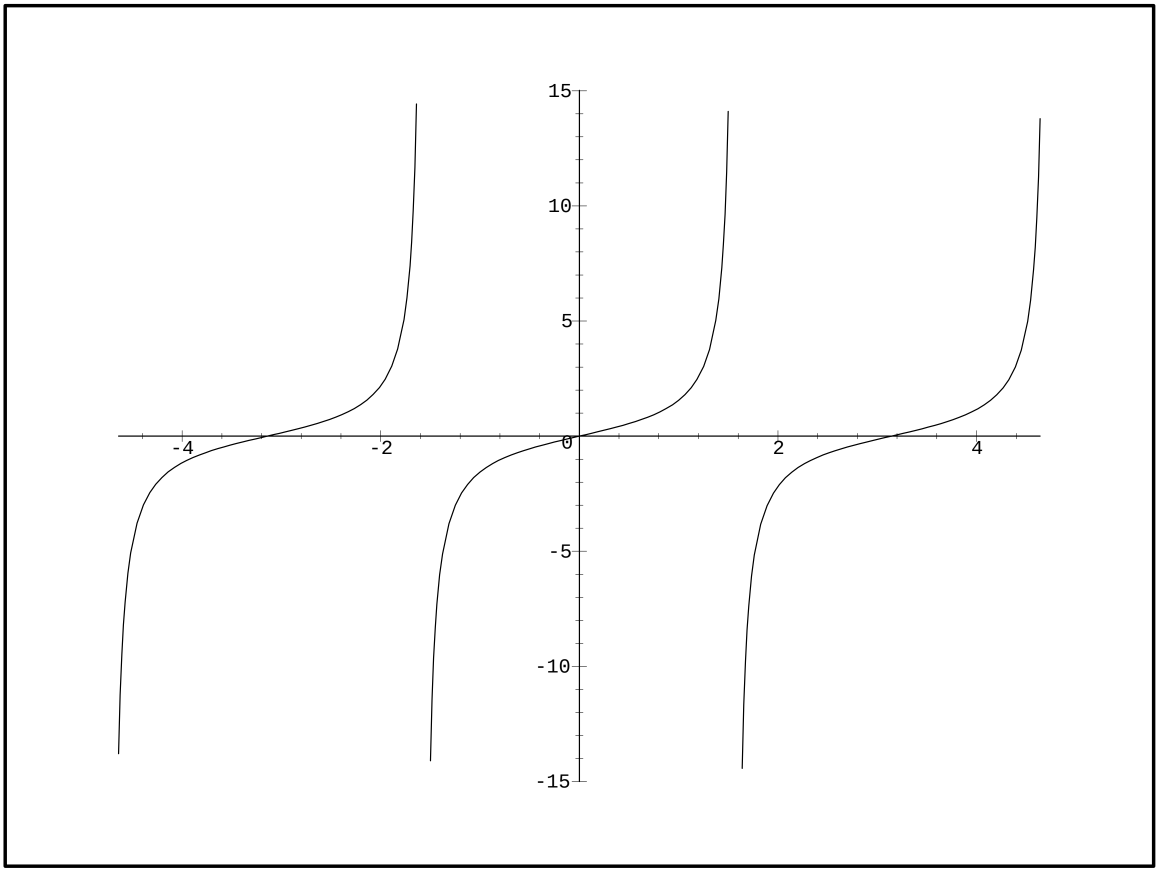
(iii)

$$\tan' = \frac{1}{\cos^2} = 1 + \tan^2\ , \qquad \cot' = -\frac{1}{\sin^2} = -(1 + \cot^2)\ .$$

(iv)

$$\cot t = -\tan\left(t + \frac{\pi}{2}\right) \ .$$

Der Graph des Tangens:



Man beachte, dass die beiden Achsen verschieden skaliert sind.

Beweis: (i) ist klar.

Zu (ii): Man beachte, dass nach den Additionstheoremen (5.21), (5.22) und wegen Satz 5.40(vii) für alle $t \in \mathbb{R}$ gelten

$$\begin{aligned} \cos(t+\pi) &= \cos t \cos\pi - \sin t \sin\pi = -\cos t \ , \\ \sin(t+\pi) &= \sin t \cos\pi + \cos t \sin\pi = -\sin t \ , \end{aligned}$$

mithin für alle $t \in D(\tan)$

$$\tan(t+\pi) = \frac{-\sin t}{-\cos t} = \tan t$$

(analog für den Cotangens).

(iii) folgt direkt aus Quotientenregel und Satz 5.40(i).

(iv) erhält man schließlich unter Verwendung der letzten Gleichung in 5.40(viii).
q.e.d.

Da die trigonomen Funktionen sin, cos, tan, cot periodisch sind, können sich nicht injektiv sein. Schränkt man aber die Funktionen auf geeignete Intervalle ein, so sind diese Einschränkungen injektiv.

Wir nennen

$$\mathrm{Sin} := \sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}, \qquad \mathrm{Cos} := \cos|_{[0, \pi]}, \tag{5.36}$$

$$\mathrm{Tan} := \tan|_{(-\frac{\pi}{2}, \frac{\pi}{2})}, \qquad \mathrm{Cot} := \cot|_{(0, \pi)} \tag{5.37}$$

die *Hauptzweige* von sin, cos, tan bzw. cot. Die Intervalle, auf denen die Hauptzweige definiert sind, sind so gewählt, dass die Hauptzweige injektiv sind: Nach 5.40(x) ist sin bzw. cos streng monoton wachsend bzw. fallend, und tan bzw. cot sind streng monoton wachsend bzw. fallend gemäß 5.41(iii).

Satz und Definition 5.42 *Die Hauptzweige der trigonometrischen Funktionen besitzen Umkehrfunktionen*

$$\begin{aligned} \operatorname{arc\,sin} &:= \mathit{Sin}^{-1} && \textit{(Arcussinus)}\,, \\ \operatorname{arc\,cos} &:= \mathit{Cos}^{-1} && \textit{(Arcuscosinus)}\,, \\ \operatorname{arc\,tan} &:= \mathit{Tan}^{-1} && \textit{(Arcustangens)}\,, \\ \operatorname{arc\,cot} &:= \mathit{Cot}^{-1} && \textit{(Arcuscotangens)}\,. \end{aligned}$$

Es gelten:

$$\begin{aligned} &D(\operatorname{arc\,sin}) = D(\operatorname{arc\,cos}) = [-1, 1]\,, \\ &W(\operatorname{arc\,sin}) = [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]\,, \\ &W(\operatorname{arc\,cos}) = [0, \pi]\,; \end{aligned}$$

Für alle $t \in (0, 1)$ *ist*

$$\operatorname{arc\,sin}'(t) = \frac{1}{\sqrt{1-t^2}}\,, \quad \operatorname{arc\,cos}'(t) = -\frac{1}{\sqrt{1-t^2}}\,. \tag{5.38}$$

$$D(\text{arc tan }) = D(\text{arc cot }) = \mathbb{R}\ , \tag{5.39}$$
$$W(\text{arc tan }) = (\,-\tfrac{\pi}{2}\,,\,\tfrac{\pi}{2}\,)\ , \tag{5.40}$$
$$W(\text{arc cot }) = (\,0\,,\,\pi\,)\ . \tag{5.41}$$

Für alle $t \in \mathbb{R}$ ist

$$\text{arc tan }'(t) = \frac{1}{1+t^2}\ , \quad \text{arc cot }'(t) = -\frac{1}{1+t^2}\ . \tag{5.42}$$

Für alle $t \in [\,-1\,,\,1\,]$ gilt

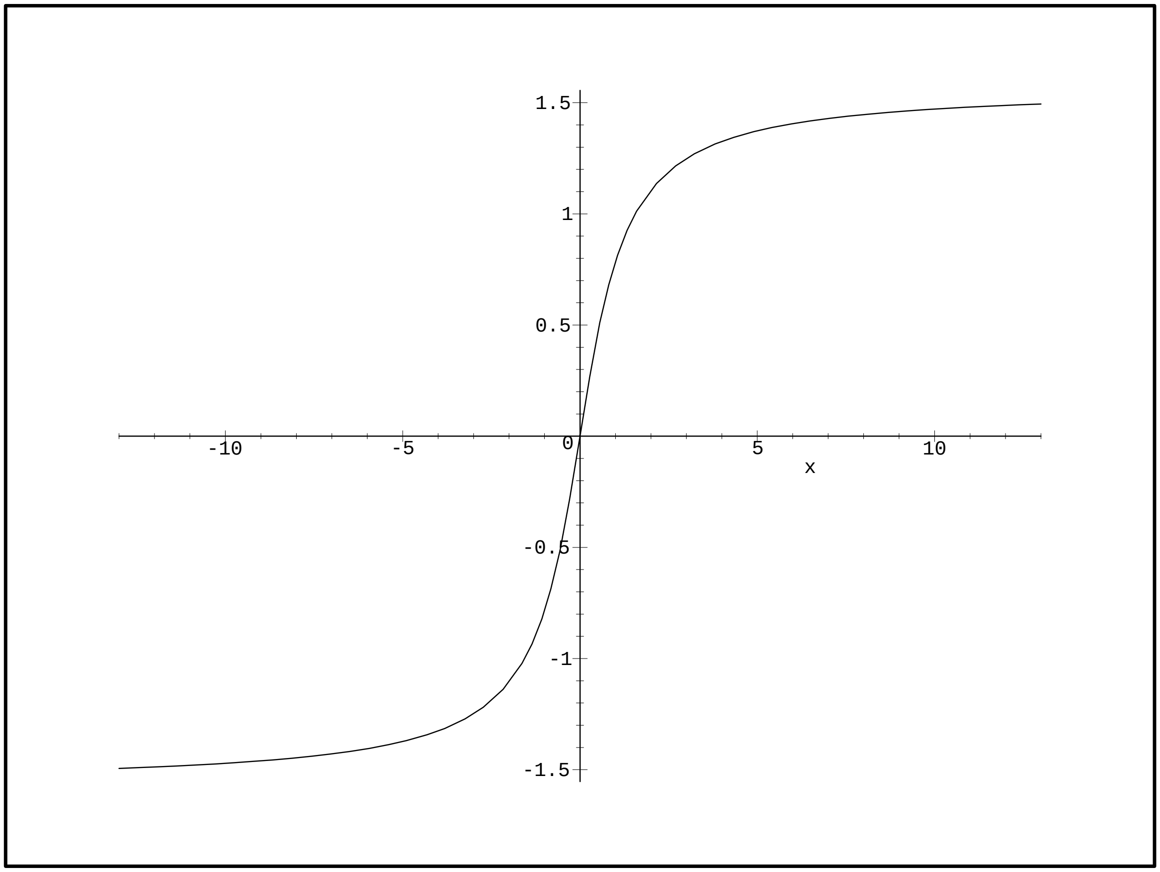
$$\text{arc cos } t = \frac{\pi}{2} - \text{arc sin } t\ . \tag{5.43}$$

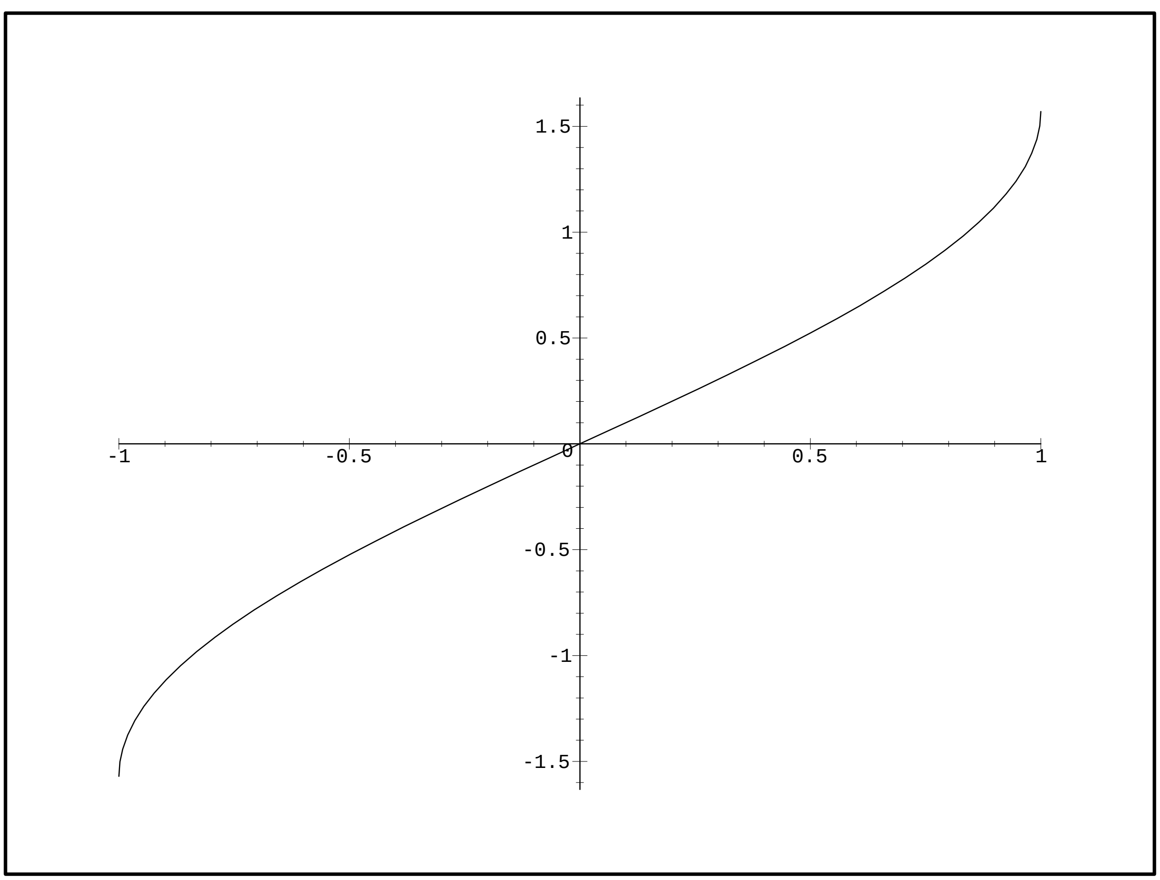
Für alle $t \in \mathbb{R}$ gilt

$$\text{arc cot } t = \frac{\pi}{2} - \text{arc tan } t\ . \tag{5.44}$$

$$\lim_{t\to\infty} \text{arc tan } t = \frac{\pi}{2}\ , \ \lim_{t\to-\infty} \text{arc tan } t = -\frac{\pi}{2}\ . \tag{5.45}$$

Auf der folgenden Seite sind die Graphen des Arcustangens (1)und des Arcussinus (2) abgebildet. Man beachte, dass alle Achsen verschieden skaliert sind.


1.5
1
0.5
0
-10
-5
5
10
x
-0.5
-1
-1.5


1.5
1
0.5
0
-1
-0.5
0.5
1
-0.5
-1
-1.5

Beweis: Wir beschränken uns auf die Nachweise von (5.38) - (5.43). Nach der Regel für die Ableitung der Umkehrfunktion, Satz 5.23, ist für $t \in (-1, 1)$

$$\text{arc sin}'(t) = \frac{1}{\cos(\text{arc sin } t)} .$$

Nun ist arc sin $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$, und für $s \in (-\pi/2, \pi/2)$ ist

$$\cos s = \sqrt{1 - \sin^2(s)}$$

denn cos ist positiv in $(-\pi/2, \pi/2)$ und $\sin^2 + \cos^2 = 1$.

Folglich ist

$$\text{arc sin}'t = \frac{1}{\sqrt{1 - [\sin(\text{arc sin } t)]^2}} = \frac{1}{\sqrt{1-t^2}} .$$

Analog erhält man

$$\text{arc cos}'t = -\frac{1}{\sin(\text{arc cos } t)} = -\frac{1}{\sqrt{1 - (\cos(\text{arc cos } t))^2}} = -\frac{1}{\sqrt{1-t^2}}$$

Nun liefert (5.39) zusammen mit Satz 5.36(iii), dass mit einer Konstanten $\alpha \in \mathbb{R}$

$$\text{arc cos } t + \text{arc sin } t = \alpha$$

für alle $t \in [-1, 1]$. Aus arc cos $0 = \frac{\pi}{2}$, arc sin $0 = 0$ erhält man $\alpha = \frac{\pi}{2}$

q.e.d.

5.5 Differentiation von Potenzreihen

Die Aussagen über Exponential- und trigonometrische Funktionen beruhen auf der gewagten Annahme, dass tatsächlich auf ganz $\mathbb{R}$ definierte Funktionen exp, cos, sin mit den Eigenschaften

$$\begin{aligned} &\exp' = \exp , \quad \exp(0) = 1 , \\ &\cos' = -\sin , \quad \sin' = \cos , \quad \cos(0) = 1 , \quad \sin(0) = 0 \end{aligned}$$

existieren. Wir haben in den Sätzen 5.37 und 5.40 gesehen, dass exp, cos und sin, sofern es sie denn gibt, durch Exponential–, Cosinus– und Sinusreihe gegeben sein müssen. Wir definieren nun die Funktionen

$$E \; : \; \mathbb{R} \longrightarrow \mathbb{R} \, , \; t \longmapsto \sum_{j=0}^{\infty} \frac{t^j}{j!} \, ,$$

$$C \; : \; \mathbb{R} \longrightarrow \mathbb{R} \, , \; t \longmapsto \sum_{j=0}^{\infty} (-1)^j \frac{t^{2j}}{(2j)!} \, ,$$

$$S \; : \; \mathbb{R} \longrightarrow \mathbb{R} \, , \; t \longmapsto \sum_{j=0}^{\infty} (-1)^j \frac{t^{2j+1}}{(2j+1)!}$$

Die Funktionen E, C, S sind tatsächlich auf ganz $\mathbb{R}$ definiert, und natürlich ist

$$E(0) = C(0) = 1 \, , \quad S(0) = 0 \, .$$

Der folgende Satz zeigt, dass auch

$$E' = E \, , \quad C' = -S \, , \quad S' = C$$

gelten. Damit ist dann die Existenz von Exponential–, Cosinus– und Sinusfunktion bewiesen.

Satz 5.43 *Sei $\delta > 0$ und $(a_n)_{n \in \mathbb{N}}$ eine Folge mit der Eigenschaft, dass für jedes $x \in (\, -\delta \, , \, \delta \,)$ die Reihe*

$$\sum_{n=0}^{\infty} a_n x^n \tag{5.46}$$

absolut konvergiert. Dann konvergiert für jedes $x \in (\, -\delta \, , \, \delta \,)$ auch die Reihe

$$\sum_{n=0}^{\infty} a_{n+1} \cdot (n+1) x^n \, . \tag{5.47}$$

Die Funktion

$$f \; : \; (\, -\delta \, , \, \delta \,) \longrightarrow \mathbb{R} \, , \; x \longmapsto \sum_{n=0}^{\infty} a_n x^n$$

ist differenzierbar, und für alle $x \in (\, -\delta \, , \, \delta \,)$ ist

$$f'(x) = \sum_{n=0}^{\infty} a_{n+1}(n+1) x^n \, .$$

Beweis:
Die Reihe in (5.47) konvergiert absolut: Wähle zu $x \in (\, -\delta \,,\, \delta \,)$ eine Zahl r mit

$$\mid x \mid < r < \delta \,.$$

Man schätzt jetzt

$$\mid a_{n+1}(n+1)x^n \mid \leq \frac{1}{r}(n+1)\left(\frac{\mid x \mid}{r}\right)^n \mid a_{n+1} r^{n+1} \mid \leq c \mid a_{n+1} \mid r^{n+1}$$

ab mit

$$c := \frac{1}{r}\sup_{n\in\mathbb{N}}(n+1)\left(\frac{\mid x \mid}{r}\right)^n$$

(wegen $\frac{|x|}{r} < 1$ gilt ja $\lim\limits_{n\to\infty}\left((n+1)\left(\frac{|x|}{r}\right)^n\right) = 0$).

Mit dem Majorantenkriterium (Satz 3.35) folgt die absolute Konvergenz der Reihe (5.47), denn $\sum\limits_{n=0}^{\infty} \mid a_n \mid r^n$ ist ja konvergent nach Voraussetzung. Mit einem ähnlichen Argument erhält man die Konvergenz von $\sum\limits_{n=2}^{\infty} n(n-1) \mid a_n \mid\mid t \mid^{n-2}$ für $t \in (\, -\delta \,,\, \delta \,)$:

$$\sum_{n=2}^{\infty} n(n-1) \mid a_n \mid\mid t \mid^{n-2} < \infty \,. \tag{5.48}$$

Wir benötigen nun eine Abschätzung: Für alle $n \in \mathbb{N}$ mit $n \geq 2$ und alle $x, h \in \mathbb{R}$ gilt

$$\mid (x+h)^n - x^n - nx^{n-1}h \mid \leq n(n-1)(\mid x \mid + \mid h \mid)^{n-2}h^2 \,. \tag{5.49}$$

Zu ihrem Beweis definieren wir

$$g \;:\; [\,0\,,\,1\,] \longrightarrow \mathbb{R}\,,\;\; t \longmapsto (x+th)^n + n(x+th)^{n-1}(1-t)h \,.$$

Dann ist

$$g(1) - g(0) = (x+h)^n - x^n - nx^{n-1}h$$

und für $t \in (\,0\,,\,1\,)$:

$$g'(t) = n(n-1)(x+th)^{n-2}(1-t)h^2 \,.$$

Eine etwas grobe Abschätzung liefert für $t \in (\,0\,,\,1\,)$

$$\mid g'(t) \mid \leq n(n-1)(\mid x \mid + \mid h \mid)^{n-2}h^2 \quad ,$$

und (5.49) folgt mit der Mittelwertabschätzung (Korollar 5.34(iii)).

Nun schreibt man

$$\tilde{f}(x) := \sum_{n=0}^{\infty} a_{n+1}(n+1)x^n = \sum_{n=1}^{\infty} a_n n x^{n-1}$$

und erhält für $x \in (-\delta, \delta)$ und h, $0 <| h |< (\delta- | x |)/2$

$$\begin{aligned}
\left|\frac{f(x+h)-f(x)}{h} - \tilde{f}(x)\right| &= \left|\sum_{n=1}^{\infty} a_n \left(\frac{(x+h)^n - x^n - hnx^{n-1}}{h}\right)\right| \\
&\leq | h | \sum_{n=2}^{\infty} | a_n | n(n-1)(| x | + | h |)^{n-2} \\
&\leq A(x) \, | h |
\end{aligned}$$

mit (beachte $|x| + |h| \leq \frac{1}{2}(|x| + \delta) < \delta$)

$$A(x) := \sum_{n=2}^{\infty} | a_n | n(n-1) \left(\frac{| x |}{2} + \frac{\delta}{2}\right)^{n-2} < \infty$$

nach (5.48). Wegen $\lim\limits_{h\to 0} A(x) \, | h |= 0$ folgt, dass $\tilde{f}$ tatsächlich die Ableitung von f ist.

q.e.d.

Wir geben noch einen Beweis für die Irrationalität von e:
Beweis (von Satz 3.39**):** Wäre e rational, so könnte man es in der Form

$$e = \frac{p}{q}$$

mit $p, q \in \mathbb{N}$ und $p > q$ (wegen $e > 1$) schreiben. Satz 5.37(v) liefert für $e = \exp(1)$ die für alle $n \in \mathbb{N}$ gültige Abschätzung

$$\sum_{j=0}^{n} \frac{1}{j!} < \frac{p}{q} \leq \sum_{j=0}^{n} \frac{1}{j!} + \frac{p}{(n+1)!q} \; .$$

Die erste Ungleichung ist in der Tat strikt, da $(\sum_{j=0}^{n} \frac{1}{j!})_{n\in\mathbb{N}}$ eine streng monoton wachsende Folge ist. Nach Subtraktion des Terms auf der äußersten linken Seite dieser Ungleichung und Multiplikation mit $n!$ erhält man mit

$$a_n := \frac{n!p}{q} - \sum_{j=0}^{n} \frac{n!}{j!}$$

für jedes $n \in \mathbb{N}$ die Abschätzung

$$0 < a_n \leq \frac{p}{(n+1)q} \ .$$

Wertet man diese Abschätzung für $n := p$ aus, so erhält man

$$0 < a_p \leq \frac{p}{(p+1)q} < 1 \ .$$

Wegen $p > q$ ist aber a_p eine ganze Zahl. Dies liefert einen Widerspruch.

q.e.d.

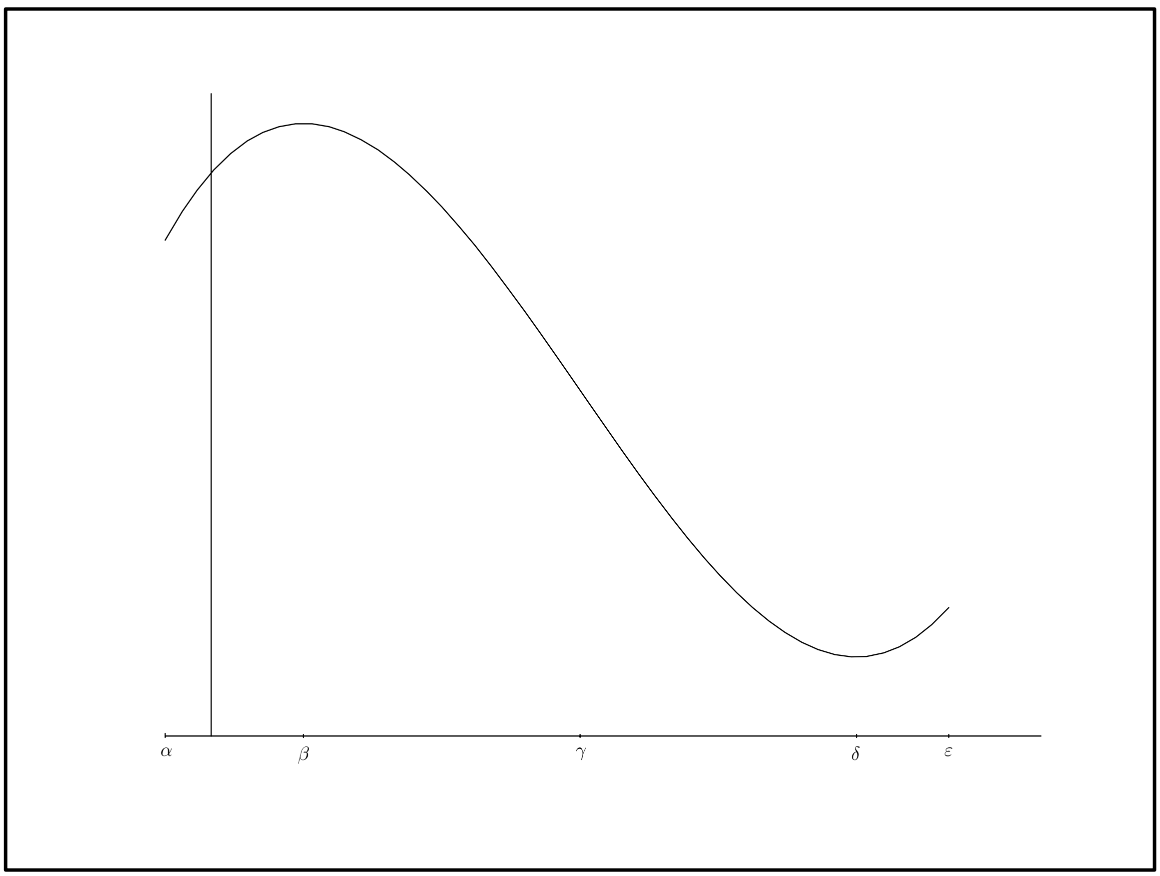
5.6 Eine Liste der Grundableitungen

$f(x) =$	mit	$D(f) =$	$f'(x) =$	$D(f') =$
x^α	$\alpha \in \mathbb{N}$	$\mathbb{R}$	$\alpha x^{\alpha-1}$	$D(f)$
	$\alpha \in \mathbb{Z}\,,\ \alpha < 0$	$\mathbb{R} \setminus \{0\}$		$D(f)$
	$\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$	$[\,0\,,\,\infty\,)$		$D(f) \setminus \{0\}$
	$\alpha \in \mathbb{R}^- \setminus \mathbb{Z}$	$(\,0\,,\,\infty\,)$		$D(f)$
$\lvert x\rvert$		$\mathbb{R}$	$\frac{x}{\lvert x\rvert}$	$D(f) \setminus \{0\}$
$\exp(x) = e^x$		$\mathbb{R}$	$\exp(x)$	$D(f)$
a^x	$a \in \mathbb{R}^+$	$\mathbb{R}$	$(\ln a)a^x$	$D(f)$
$\ln x$		$\mathbb{R}^+$	$\frac{1}{x}$	$D(f)$
$\ln \lvert x\rvert$		$\mathbb{R} \setminus \{0\}$	$\frac{1}{x}$	$D(f)$
$\sin x$		$\mathbb{R}$	$\cos x$	$D(f)$
$\cos x$		$\mathbb{R}$	$-\sin x$	$D(f)$
$\tan x$		$\mathbb{R} \setminus \{\frac{(2k-1)\pi}{2}\colon\ k \in \mathbb{Z}\}$	$1 + \tan^2 x$	$D(f)$
$\cot x$		$\mathbb{R} \setminus \{k\pi\colon\ k \in \mathbb{Z}\}$	$-1 - \cot^2 x$	$D(f)$
$\text{arc sin } x$		$[\,-1\,,\,1\,]$	$(1-x^2)^{-\frac{1}{2}}$	$D(f) \setminus \{-1\,,\,1\}$
$\text{arc cos } x$		$[\,-1\,,\,1\,]$	$-(1-x^2)^{-\frac{1}{2}}$	$D(f) \setminus \{-1\,,\,1\}$
$\text{arc tan } x$		$\mathbb{R}$	$(1+x^2)^{-1}$	$D(f)$
$\text{arc cot } x$		$\mathbb{R}$	$-(1+x^2)^{-1}$	$D(f)$

Kapitel 6

Höhere Ableitungen I

6.1 Konvexe Funktionen



Man stelle sich den Graphen Γ_f einer Funktion f vor, etwa den oben abgebildeten.

Außer den Monotoniebereichen $[\alpha\,,\,\beta]$, $[\delta\,,\,\varepsilon]$ bzw. $[\beta\,,\,\delta]$. (darin ist f monoton wachsend [1] bzw. fallend) bemerkt man zwei Bereiche mit je einer weiteren geometrischen Besonderheit: in $[\alpha\,,\,\gamma]$ ist Γ_f rechts- und in $[\gamma\,,\,\varepsilon]$ linksgekrümmt.

Anschaulich: Fährt man mit einem Fahrrad in Richtung wachsender Abszissenwerte den Graphen entlang, so muss man in $[\alpha\,,\,\gamma]$ den Lenker rechts und in $[\gamma\,,\,\varepsilon]$ links einschlagen. Beachten Sie, dass beim Übergang von f zu $-f$ rechts und links vertauscht werden müssen. Das heißt aber, dass "rechts" oder "links" von der Orientierung der Koordinatenachsen abhängt. Wir werden deshalb lieber von "konkav" [2] und "konvex" [3] sprechen.

Definition 6.1 *Sei $I \subset D(f)$ ein nicht entartetes Intervall (d. h. $\mathring{I} \neq \emptyset$) innerhalb des Definitionsbereichs $D(f)$ einer reellen Funktion f. f heißt konvex in I, wenn*

$$\bigwedge_{s,t\in I} \quad \bigwedge_{\lambda\in(0,1)} \quad f((1-\lambda)s+\lambda t) \leq (1-\lambda)f(s)+\lambda f(t)\ . \tag{6.1}$$

f heißt streng konvex in I, wenn (6.1) *mit der strikten Ungleichung "$<$" (statt "$\leq$") wahr ist. f heißt konkav bzw. streng konkav in I, wenn* (6.1) *mit "$\geq$" bzw. "$>$" statt "$\leq$" wahr ist.*

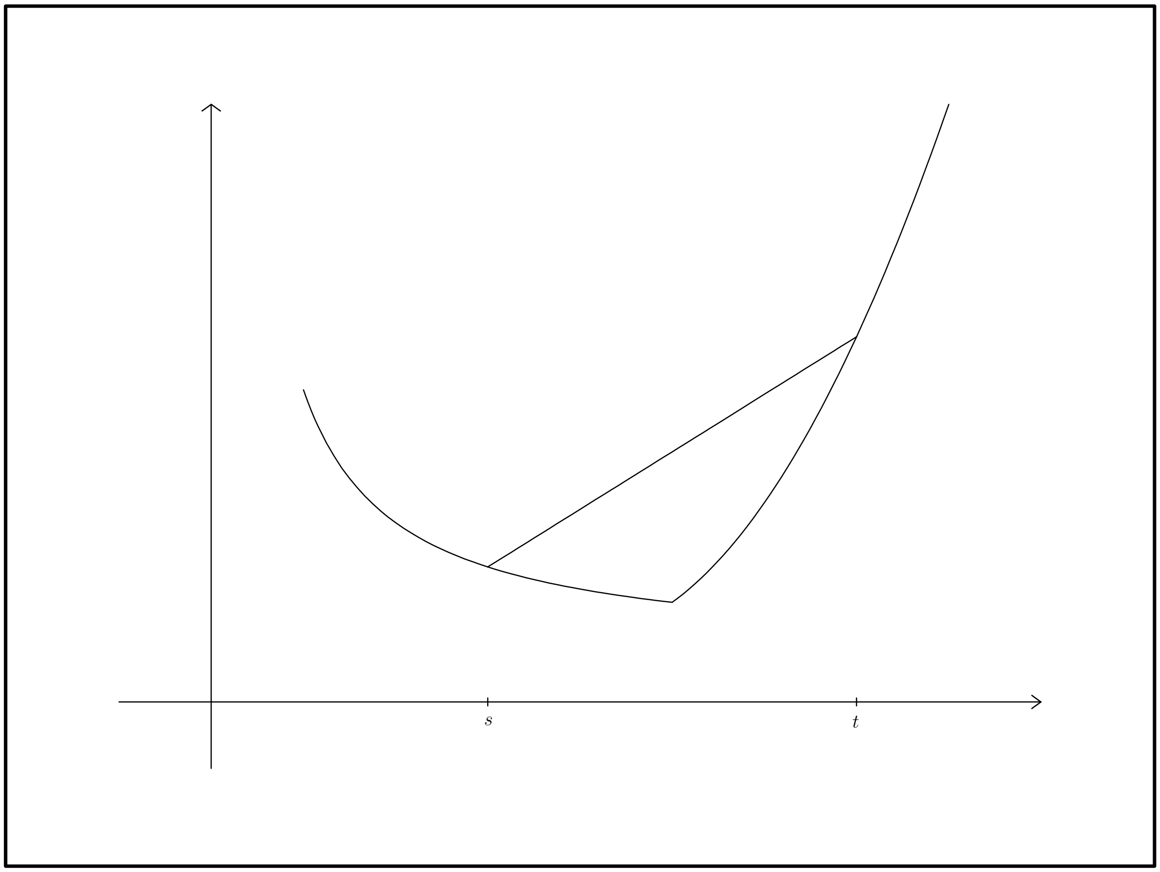
Ist $D(f)$ ein nicht entartetes Intervall und ist f (streng) konvex bzw. (streng) konkav in $D(f)$, so nennt man f (streng) konvex bzw. (streng) konkav.

Geometrisch bedeutet (strenge) Konvexität von f in I, dass der Graph von f zwischen zwei verschiedenen beliebigen Punkten nirgends überhalb (stets unterhalb) der Sekante durch diese beiden Punkte liegt:

[1] Achtung: f ist monoton wachsend in $[\alpha\,,\,\beta]$ und in $[\delta\,,\,\varepsilon]$, <u>nicht</u> aber in $[\alpha\,,\,\beta]\cup[\delta\,,\,\varepsilon]$

[2] statt rechtsgekrümmt

[3] statt linksgekrümmt



(6.1) ist nämlich äquivalent mit

$$\bigwedge_{s,t\in I, s<t} \quad \bigwedge_{\xi\in(s,t)} \quad f(\xi) \leq f(s)\frac{t-\xi}{t-s} + f(t)\frac{\xi-s}{t-s} \quad , \tag{6.2}$$

und die Funktion

$$S \; : \; (\, s \, , \, t \,) \longrightarrow \mathbb{R} \, , \; \xi \longmapsto f(s)\frac{t-\xi}{t-s} + f(t)\frac{\xi-s}{t-s}$$

hat als Graphen die Sekante zwischen den Punkten $(s, f(s))$ und $(t, f(t))$.

Beispiel 6.2 *Folgende Funktionen sind streng konvex:*

$$f \; : \; \mathbb{R} \longrightarrow \mathbb{R} \, , \; x \longmapsto x^2 \tag{6.3}$$

$$f \; : \; \mathbb{R} \longrightarrow \mathbb{R} \, , \; x \longmapsto \exp(x) \tag{6.4}$$

Folgende Funktionen sind konvex

$$r \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto |x| \tag{6.5}$$

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \alpha x + \beta \quad (\textit{mit } \alpha, \beta \in \mathbb{R}) \tag{6.6}$$

Die Funktion aus (6.6) *ist sowohl konvex als auch konkav.*

Die Konvexität in obigen Beispielen "sieht man", wenn man die Graphen skizziert. Sie ist leicht nachzuweisen, sobald Kriterien für Konvexität bekannt sind.

Wir werden im Folgenden meist Resultate für konvexe Funktionen herleiten. Für konkave Funktionen erhält man dann entsprechendes, wenn man beachtet

Satz 6.3 *Die reelle Funktion f ist im Intervall I genau dann (streng) konkav, wenn $-f$ in I (streng) konvex ist.*

Dies folgt sofort aus der Definition 6.1.

Der folgende Satz bestätigt die Anschauung:

Satz 6.4 *Ist f konvexe Funktion auf dem kompakten Intervall $[\,a\,,\,b\,] = D(f)$, so nimmt f ihr Maximum an einer der Intervallgrenzen an:*

$$\max_{[\,a\,,\,b\,]} f = \max\{f(a), f(b)\}\;. \tag{6.7}$$

Besitzt f im Inneren $(\,a\,,\,b\,)$ von $D(f)$ ein lokales Extremum, so ist dieses ein lokales Minimum.

Beispiel 6.5

$$f \;:\; [\,0\,,\,1\,] \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \begin{cases} 1 & , \textit{falls} \quad x = 1 \\ 0 & \textit{sonst} \end{cases}$$

ist konvex und in jedem inneren Punkt von $D(f)$, also in jedem $x \in (\,0\,,\,1\,)$, besitzt f ein lokales Extremum. Dieses ist sowohl lokales Maximum als auch lokales Minimum.

Beweis (von Satz 6.4**):** Formel (6.7) folgt direkt aus der Definition der Konvexität: Ist $x \in [\,a\,,\,b\,]$, so liegt der Punkt $(x, f(x))$ nicht oberhalb der Sekante durch $(a, f(a))$ und $(b, f(b))$:

$$f(x) \leq f(b)\frac{x-a}{b-a} + f(a)\frac{b-x}{b-a} \leq \max\{f(a); f(b)\}\;.$$

Besitzt nun f in $\xi \in (\,a\,,\,b\,)$ ein lokales Maximum, so existiert ein $\delta > 0$ mit

$$f(\xi) = \max_{x \in U_\delta(\xi)} f(x)\ .$$

Daher und nach (6.1) ist für jedes $x \in U_\delta(\xi)$ mit $\theta := x - \xi$:

$$f(\xi) \leq \frac{1}{2} f(\xi + \theta) + \frac{1}{2} f(\xi - \theta) \leq \frac{1}{2} f(\xi) + \frac{1}{2} f(x) \quad , \tag{6.8}$$

also

$$\frac{1}{2} f(\xi) \leq \frac{1}{2} f(x)\ . \tag{6.9}$$

Mithin liegt in ξ zugleich ein lokales Minimum von f.

q.e.d.

Man beachte: Ist f streng konvex in I, so besitzt f im Inneren von I kein lokales Maximum. In (6.8) könnte nämlich die erste Ungleichung zu einer strikten Ungleichung verschärft werden und (6.9) (mit "$<$" statt '$\leq$") lieferte einen Widerspruch.

Wir formen Formel (6.2) äquivalent um, indem wir von beiden Seiten $f(s)$ subtrahieren:

$$f(\xi) - f(s) \leq f(s) \left(\frac{t - \xi}{t - s} - 1 \right) + f(t) \frac{\xi - s}{t - s} = \frac{f(t) - f(s)}{t - s} \cdot (\xi - s) \quad ,$$

und erhalten

Satz 6.6 *$f : I \longrightarrow \mathbb{R}$ ist im Intervall I genau dann konvex, wenn für alle $x_1, x_2, x_3 \in I$ mit $x_1 < x_2 < x_3$ gilt*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1}\ . \tag{6.10}$$

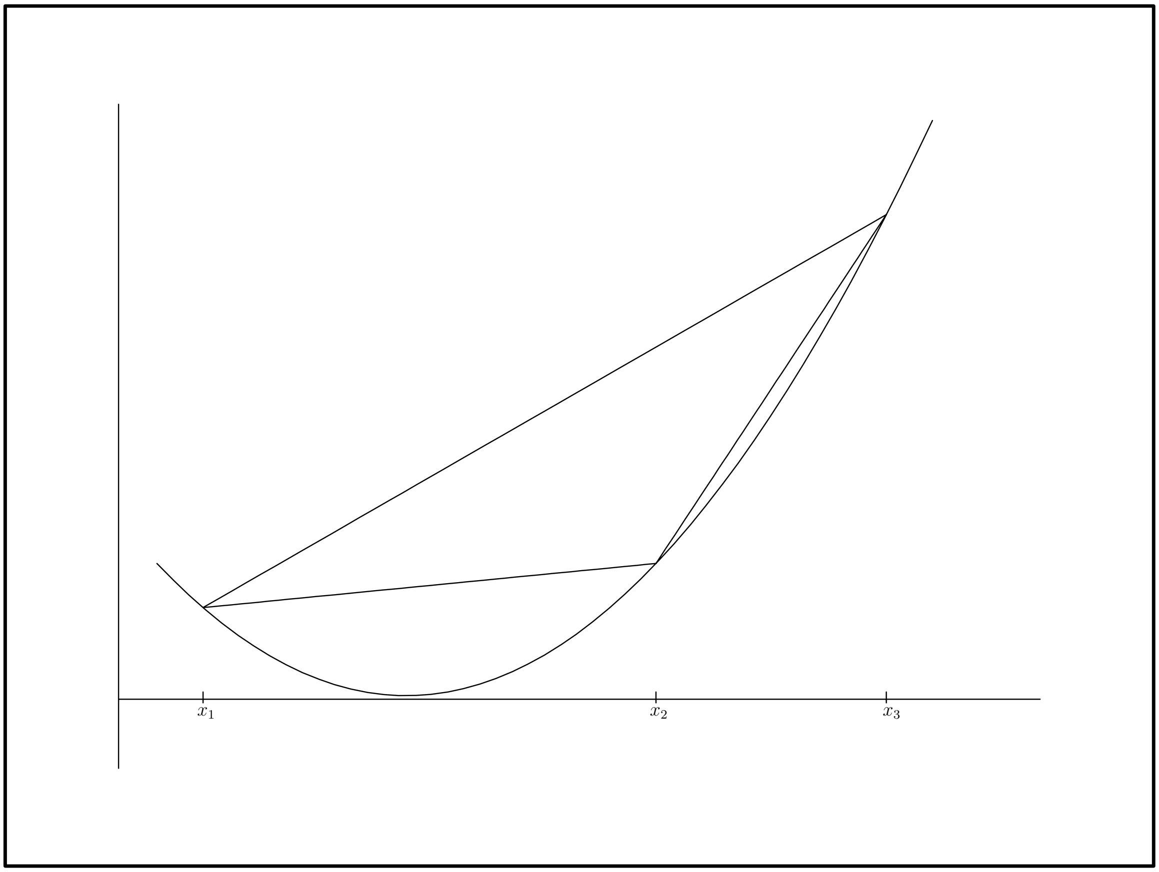
f ist genau dann streng konvex wenn (6.10) mit "$<$" statt "$\leq$" gilt.

Ist f konvex in I, so gilt ferner für x_1, x_2, x_3 wie oben:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} \tag{6.11}$$

(mit scharfen Ungleichungen im Fall strenger Konvexität).

Man mache sich die Ungleichungskette (6.11) anhand folgender Abbildung klar:



Die in Differenzenquotienten in (6.11) sind die jeweiligen Sekantensteigerungen.

Beweis: Es bleibt, die zweite Ungleichung in (6.11) zu beweisen. Hierzu setzt man $s := x_1, \xi := x_2, t := x_3$, und subtrahiert $f(t)$ von beiden Seiten der Ungleichung (6.2) und erhält

$$f(\xi) - f(t) \leq \frac{f(t) - f(s)}{t - s}(\xi - t) \; .$$

Dividiere beide Seiten dieser Ungleichung durch $\xi - t$ und beachte $\xi - t < 0$.

q.e.d.

Korollar 6.7 *Die Funktion* $f \;:\; I \;\longrightarrow\; \mathbb{R}$ *ist im Intervall* I *genau dann* *(streng) konvex, wenn für jedes $p \in I$ die Funktion*

$$\Phi_p \ : \ I \setminus \{p\} \longrightarrow \mathbb{R}\ , \ x \longmapsto \frac{f(x) - f(p)}{x - p}$$

(streng) monoton wachsend ist.

Beweis: Seien $p \in I$ und $s, t \in I \setminus \{p\}$ mit $s < t$. Dann sind folgende Konfigurationen möglich.

- $s < t < p$: Wende Satz 6.6 an mit $x_1 := s$, $x_2 := t$, $x_3 := p$.
- $s < p < t$: Wende Satz 6.6 an mit $x_1 := s$, $x_2 := p$, $x_3 := t$.
- $p < s < t$: Wende Satz 6.6 an mit $x_1 := p$, $x_2 := s$, $x_3 := t$.

In jedem Fall erhält man $\Phi_p(s) \leq \Phi_p(t)$ (bzw. $\Phi_p(s) < \Phi_p(t)$).

q.e.d.

Aus Korollar 6.7 erhält man, dass f im Inneren eines Konvexitätsintervalls stetig und in jedem Punkt im Intervallinneren rechts- und linksseitig differenzierbar ist.

Zunächst soll einseitige Differenzierbarkeit definiert werden.

Definition 6.8 *Sei f eine reelle Funktion und $x \in D(f)$ ein Häufungspunkt von $D_x^+ := D(f) \cap (\, x\, ,\, \infty\,)$ (bzw. von $D_x^- := D(f) \cap (\, -\infty\, ,\, x\,)$). f heißt rechtsseitig (bzw. linksseitig) differenzierbar in x, wenn es eine in x stetige Funktion F^+ (bzw. F^-), $F^\pm : D_x^\pm \longrightarrow \mathbb{R}$, gibt mit*

$$\bigwedge_{t \in D_x^\pm} f(t) = f(x) + F^\pm(t)(t - x)\ .$$

Die Zahl

$$d^+ f(x) := F^+(x) \quad bzw. \quad d^- f(x) := F^-(x)\ .$$

heißt dann rechts- (bzw. linksseitige) Ableitung von f in x.

Wie die Ableitung, so sind auch rechts- und linksseitige Ableitung eindeutig.

Äquivalent sind die Bedingungen

$$\lim_{n \to \infty} \left[f(x_n) - f(x) \right] / (x_n - x) = d^\pm f(x) \tag{6.12}$$

für jede streng monoton fallende (bzw. wachsende) gegen x konvergente Folge (x_n) in $D(f)$ mit $\lim_{n\to\infty} x_n = x$. Oder (vgl. Formel (4.6)):

$$\lim_{t\searrow x} \frac{f(t)-f(x)}{t-x} = d^+f(x) \quad (\text{bzw.} \lim_{t\nearrow x} \frac{f(t)-f(x)}{t-x} = d^-f(x)) \ . \tag{6.13}$$

Natürlich existiert in einem inneren Punkt x von $D(f)$ die Ableitung von f genau dann, wenn rechts- und linksseitige Ableitung in x existieren und übereinstimmen. Dann ist

$$f'(x) = d^+f(x) = d^-f(x) \ .$$

Da für konvexe Funktionen laut Korollar 6.7 für jedes p der Differenzenquotient bei p monoton wachsend ist, erhält man:

Korollar 6.9 *Ist f im <u>offenen</u> Intervall I konvex, so existiert in jedem $p \in I$ sowohl rechts- als auch linksseitige Ableitung, und es gilt für alle $p \in I$*

$$d^-f(p) \le d^+f(p)$$

Um etwa die Existenz von $d^+f(p)$ in p nachzuweisen, beachte man, dass für eine streng monoton fallende, gegen p konvergente Folge $(x_n)_{n\in\mathbb{N}}$ in I die durch

$$\Phi_p(x_n) := \frac{f(x_n)-f(p)}{x_n-p}$$

gegebene Folge ebenfalls monoton fällt und (mit irgend einem $s \in I$, $s < p$) durch $\Phi_p(s)$ beschränkt ist. Daher ist $(\Phi_p(x_n))$ konvergent gegen $\inf_{n\in\mathbb{N}} \Phi_p(x_n)$ (vgl. Satz 3.10).

Man überlegt sich aber leicht, dass

$$\inf_{n\in\mathbb{N}} \Phi_p(x_n) = \inf_{x>p} \Phi_p(x) \quad ,$$

und damit existiert

$$d^+f(p) = \inf_{x>p} \Phi_p(x) \ .$$

Analog zeigt man die Existenz von $d^-f(p)$ als

$$d^-f(p) = \sup_{x<p} \Phi_p(x) \ .$$

Da für jedes $s, t \in I$ mit $s < p < t$ gilt

$$\Phi_p(s) \le \Phi_p(t) \quad ,$$

folgt

$$d^-f(p) \le d^+f(p) \ .$$

Korollar 6.10 *Ist f auf dem offenen Intervall I konvex, so ist f stetig auf I.*

Denn für jede streng monotone Folge $(x_n)_{n\in\mathbb{N}}$ in I, die gegen $x \in I$ konvergiert, gilt

$$\begin{aligned}\lim_{n\to\infty} f(x_n) &= \lim_{n\to\infty}\left[f(x) + \frac{f(x_n)-f(x)}{x_n - x}\cdot(x_n - x)\right]\\ &= f(x) + d^{\pm}f(x)\cdot 0 = f(x).\end{aligned}$$

Schaut man sich den Graphen einer konvexen Funktion an, so erwartet man, dass die Ableitung – so existent – monoton wächst:

Satz 6.11 *Sei I ein Intervall, und $f \in C(I)$ sei differenzierbar im Inneren von I. Dann gilt: f ist genau dann (streng) konvex, wenn f' (streng) monoton wachsend ist.*

Beweis: Wir nehmen zunächst f als konvex an und zeigen die Monotonie der Ableitung. Hierzu seien p, q Punkte im Inneren von I mit $p < q$, und $(p_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$ seien streng monoton fallende Folgen in I mit Grenzwerten p bzw. q sowie $p_n < q$ für alle $n \in \mathbb{N}$.

Dann liefert (6.11):

$$\frac{f(p_n)-f(p)}{p_n - p} \le \frac{f(q_n)-f(p)}{q_n - p} \le \frac{f(q_n)-f(q)}{q_n - q}$$

und folglich

$$f'(p) = \lim_{n\to\infty}\frac{f(p_n)-f(p)}{p_n - p} \le \lim_{n\to\infty}\frac{f(q_n)-f(q)}{q_n - q} = f'(q)\ ;$$

f' ist also monoton wachsend.

Ist f sogar streng konvex, so ist zunächst f' monoton wachsend. Wäre nun für zwei Zahlen p, q aus dem Inneren von I

$$\alpha := f'(p) = f'(q) \quad \text{und} \quad p < q \quad ,$$

so wäre für alle $x \in [\,p\,,\,q\,]$

$$f'(x) = \alpha$$

und nach Satz 5.36

$$\bigwedge_{x\in[p,q]} \quad f(x) = \alpha(x-p) + f(p)\ .$$

Denn

$$g \ : \ [\,p\,,\,q\,] \longrightarrow \mathbb{R}\,, \ x \longmapsto f(x) - [\alpha(x-p) + f(p)]$$

verschwindet in p und hat Ableitung 0 in $(\,p\,,\,q\,)$.

Dann aber ist für alle $x \in (\,p\,,\,q\,]$

$$\frac{f(x)-f(p)}{x-p} = \alpha$$

im Widerspruch zu Satz 6.6.

Ist nun umgekehrt f' (streng) monoton wachsend, so erhält man (strenge) Konvexität mit Hilfe von Satz 6.6:

Sind $x_1, x_2, x_3 \in I$ mit $x_1 < x_2 < x_3$, so gibt es $s \in (\,x_1\,,\,x_2\,)$ und $t \in (\,x_2\,,\,x_3\,)$ (nach dem Mittelwertsatz) mit

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} = f'(s) \quad ; \quad \frac{f(x_3)-f(x_2)}{x_3-x_2} = f'(t) \overset{>}{\geq} f'(s)\ .$$

Also ist

$$\begin{aligned}
\frac{f(x_3)-f(x_1)}{x_3-x_1} &= \frac{[f(x_3)-f(x_2)] + [f(x_2)-f(x_1)]}{x_3-x_1} \\
&= f'(t)\frac{x_3-x_2}{x_3-x_1} + f'(s)\frac{x_2-x_1}{x_3-x_1} \\
&\overset{>}{\geq} f'(s)\frac{x_3-x_2}{x_3-x_1} + f'(s)\frac{x_2-x_1}{x_3-x_1} \\
&= f'(s) \\
&= \frac{f(x_2)-f(x_1)}{x_2-x_1},
\end{aligned}$$

und das ist (6.10).

q.e.d.

Ein Kriterium für die Monotonie der Ableitung f' erhält man, indem man auf f' den Satz 5.38 anwendet. Hierzu muss man die Ableitung f'' von f' bilden, welche man die zweite Ableitung von f nennt.

Rekursiv definiert man höhere Ableitungen einer Funktion:

Definition 6.12 *Sei f eine reelle Funktion (mit Definitionsbereich $D(f)$).*

Wir setzen

$$f^{(0)} := f \quad ,$$

und nennen $f^{(0)}$ auch "die 0–te Ableitung von f". Für $n \in \mathbb{N}_0$ existiere die n-te Ableitung $f^{(n)}$, eine reelle Funktion mit Definitionsbereich $D(f^{(n)})$.

Ist

$$D_{n+1} := \{x \in D(f^{(n)}) \colon f^{(n)} \text{ ist differenzierbar in } x\}$$

nicht leer, so heißt

$$f^{(n+1)} \ : \ D(f^{(n+1)}) := D_{n+1} \longrightarrow \mathbb{R}\,, \ x \longmapsto (f^{(n)})'(x)$$

die $(n+1)$-te Ableitung von f. Ist D_{n+1} leer, so existiert weder die $(n+1)$-te Ableitung noch irgend eine höhere Ableitung.

f heißt n-mal differenzierbar in $x \in D(f)$, wenn $x \in D(f^{(n)})$ ist. f heißt n-mal differenzierbar in $M \subset D(f)$, wenn $M \subset D(f^{(n)})$, f heißt n-mal differenzierbar, wenn $D(f) = D(f^{(n)})$ ist.

Ist n "genügend klein", so ersetzt man in der Schreibung $f^{(n)}$ den Exponenten (n) durch n Striche:

$$f^{(5)} =: f''''' \,.$$

Nun liefert Satz 5.36 in Zusammenhang mit Satz 6.11:

Satz 6.13 *Ist f stetig im Intervall I und zweimal differenzierbar im Inneren $\overset{\circ}{I}$ von I, so gelten*

f ist genau dann konvex in I, wenn $f''(x) \geq 0$ ist für alle $x \in \overset{\circ}{I}$.

Ist $f''(x) > 0$ für alle $x \in \overset{\circ}{I}$, so ist f streng konvex in I.

Beispiel 6.14 *Man beachte:*

$$f \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ x \longmapsto x^4$$

ist streng konvex, aber die zweite Ableitung

$$f'' \ : \ \mathbb{R} \longrightarrow \mathbb{R}\,, \ x \longmapsto 12x^2$$

besitzt in 0 eine Nullstelle.

Beispiel 6.15 *Die Funktion f genüge im offenen Intervall I der Differentialgleichung*

$$f''(x) + f(x) = 0$$

für alle $x \in I$. Dann ist f in jedem Teilintervall, wo es positiv ist, streng konkav, und wo es negativ ist, streng konvex. Das gilt also insbesondere für die Cosinus- und die Sinusfunktion.

Definition 6.16 *Sei f reelle Funktion, $w \in D(f)$. Ein Der Punkt $(w, f(w)) \in \mathbb{R}^2$ heißt "Wendepunkt" von f (oder: f besitzt in w einen Wendepunkt, w ist Wendestelle von f), wenn w innerer Punkt von $D(f)$ ist und es ein $\delta > 0$ so gibt, dass f konvex (bzw. konkav) im Intervall $[\, w - \delta \,,\, w \,]$ und konkav (bzw. konvex) im Intervall $[\, w \,,\, w + \delta \,]$ ist.*

Beispiel 6.17

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \sin x$$

besitzt in 0 einen Wendepunkt. Denn

$$f''(x) = -\sin x$$

ist negativ für $x \in (\,0\,,\,\pi\,)$ und positiv für $x \in (\,-\pi\,,\,0\,)$. Entsprechend liegen in allen Nullstellen der Sinusfunktion Wendepunkte des Sinus.

Merke:
Ein Kandidat für die Wendestelle einer Funktion ist eine Stelle w, wo $f''(w) = 0$ ist.

Beispiel 6.18

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto x \cdot |x|$$

besitzt in 0 eine Wendestelle. Denn f' ist gegeben durch $f'(x) = 2|x|$ und f'' durch $f''(x) = 2\frac{x}{|x|}$.

Offenbar ist f'' negativ in $\mathbb{R}^-$ und positiv in $\mathbb{R}^+$. Aber f ist in 0 nicht zweimal differenzierbar.

Merke:
Stellen, in denen f nicht zweimal differenzierbar ist, sind Kandidaten für Wendestellen.

Beispiel 6.19 *Mit*

$$sgn \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \begin{cases} \frac{x}{|x|} & , \;\; \textit{falls} \;\; x \neq 0 \\ 0 & , \;\; \textit{falls} \;\; x = 0 \end{cases}$$

sei

$$f \;:\; \mathbb{R} \longrightarrow \mathbb{R}\,,\;\; x \longmapsto (sgn\, x) \cdot |x|^{1/3}\,.$$

f ist die Umkehrfunktion von $g : \mathbb{R} \longrightarrow \mathbb{R}\,,\ x \longmapsto x^3$, also insbesondere stetig. f besitzt in 0 eine Wendestelle, denn

$$f'(x) = \frac{1}{3}(x^2)^{-1/3} \quad , \quad f''(x) = -\frac{2}{9}x(x^8)^{-1/3}$$

für $x \in \mathbb{R} \setminus \{0\}$. Offenbar ist f'' positiv in $\mathbb{R}^-$ und negativ in $\mathbb{R}^+$, existiert aber nicht in 0.

Beispiel 6.20 *Die Wendepunktdefinition impliziert auch, dass jeder Punkt Wendestelle von $r : \mathbb{R} \longrightarrow \mathbb{R}\,,\ x \longmapsto |x|$ ist.*

Beispiel 6.21 *Wir wollen eine Kurvendiskussion durchführen. Gegeben sei eine Funktion f durch eine "Abbildungsvorschrift" $f(x)$. Ein vernünftiges Schema scheint die Bearbeitung der folgenden Punkte zu sein:*

- *Bestimmung von $D(f), D(f'), D(f'')$, den Abbildungsvorschriften für f' und f'' sowie die Angabe aller Punkte, wo f nicht stetig ist.*
- *Bestimmung der Nullstellen von f.*
- *Aufklärung des Verhaltens von f an den Grenzen von $D(f)$ und an eventuellen Unstetigkeitsstellen. Mit "Grenzen des Definitionsbereichs" sind hier alle Häufungspunkte in $\hat{\mathbb{R}}$ von $D(f)$ gemeint.*
- *Bestimmung aller Intervalle, in denen f monoton wächst bzw. fällt und Angabe lokaler und globaler Extrema.*
- *Bestimmung aller Intervalle, wo f konvex bzw. konkav ist und Angabe der Wendepunkte.*
- *Skizze auf Grund der gefundenen Resultate.*

Untersucht werden die durch

$$f(x) := (x^3 + x^2)^{1/2} \quad , \quad g(x) := (x^3 + x^2)^{1/4}$$

gegebenen Funktionen:

- $D(f) = D(g) = \{x \in \mathbb{R}\colon x^3 + x^2 \geq 0\} = [\,-1\,,\,\infty\,)$
  f, g sind stetig (als Hintereinanderschaltungen stetiger Funktionen)
  $f'(x) = \frac{1}{2}x(3x+2)(x^3+x^2)^{-1/2} \quad , \quad g'(x) = \frac{1}{4}x(3x+2)(x^3+x^2)^{-3/4}$
  $D(f') = D(g') = (\,-1\,,\,0\,) \cup (\,0\,,\,\infty\,)$
  Dass f und g tatsächlich nicht in 0 differenzierbar sind, erkennt man aus

  $$d^+f(0) = 1\,, \quad d^-f(0) = -1$$

$$\left|\frac{g(0+h)-g(0)}{h}\right| = \frac{|h|^{1/2}|1+h|^{1/4}}{|h|} \to \infty \quad \textit{für} \quad h \to 0 \; .$$

sowie mit $D(f'') = D(f') = D(g'') = D(g') = (\,-1\,,\,0\,) \cup (\,0\,,\,\infty\,)$*:*

$$f''(x) = \frac{1}{4}x^3(x^3+x^2)^{-3/2}(3x+4)\;,$$
$$g''(x) = -\frac{3}{16}x^2(x^3+x^2)^{-7/4}\left(\left(x+\frac{2}{3}\right)^2+\frac{8}{9}\right)\;.$$

- $f(x)=0 \Leftrightarrow g(x)=0 \Leftrightarrow x^3+x^2=0 \Leftrightarrow x=0 \vee x=-1$.
  Also ist $N(g) = N(f) = \{-1\,;\,0\}$

- $\hat{\mathbb{R}} - \lim\limits_{x\to\infty} f(x) = \hat{\mathbb{R}} - \lim\limits_{x\to\infty} g(x) = \infty$

- $f'(x) > 0 \Leftrightarrow g'(x) > 0 \Leftrightarrow x \in (\,-1\,,\,-\frac{2}{3}\,) \cup (\,0\,,\,\infty\,)$
  $f'(x) < 0 \Leftrightarrow g'(x) < 0 \Leftrightarrow x \in (\,-\frac{2}{3}\,,\,0\,)$
  $f'(x) = 0 \Leftrightarrow g'(x) = 0 \Leftrightarrow x = -\frac{2}{3}$.
  Hieraus erkennt man
  f *und* g *sind streng monoton wachsend in* $[\,-1\,,\,-\frac{2}{3}\,]$ *und in* $[\,0\,,\,\infty\,)$*; sie sind streng monoton fallend in* $[\,-\frac{2}{3}\,,\,0\,]$*. Sowohl* f *als auch* g *besitzen in* $\frac{2}{3}$ *jeweils ein lokales Maximum,*

  $$b := f(-\frac{2}{3}) = \frac{2}{9}\sqrt{3}\;, \quad a := g(-\frac{2}{3}) = \frac{1}{3}\sqrt{2}\cdot\sqrt[4]{3}\;.$$

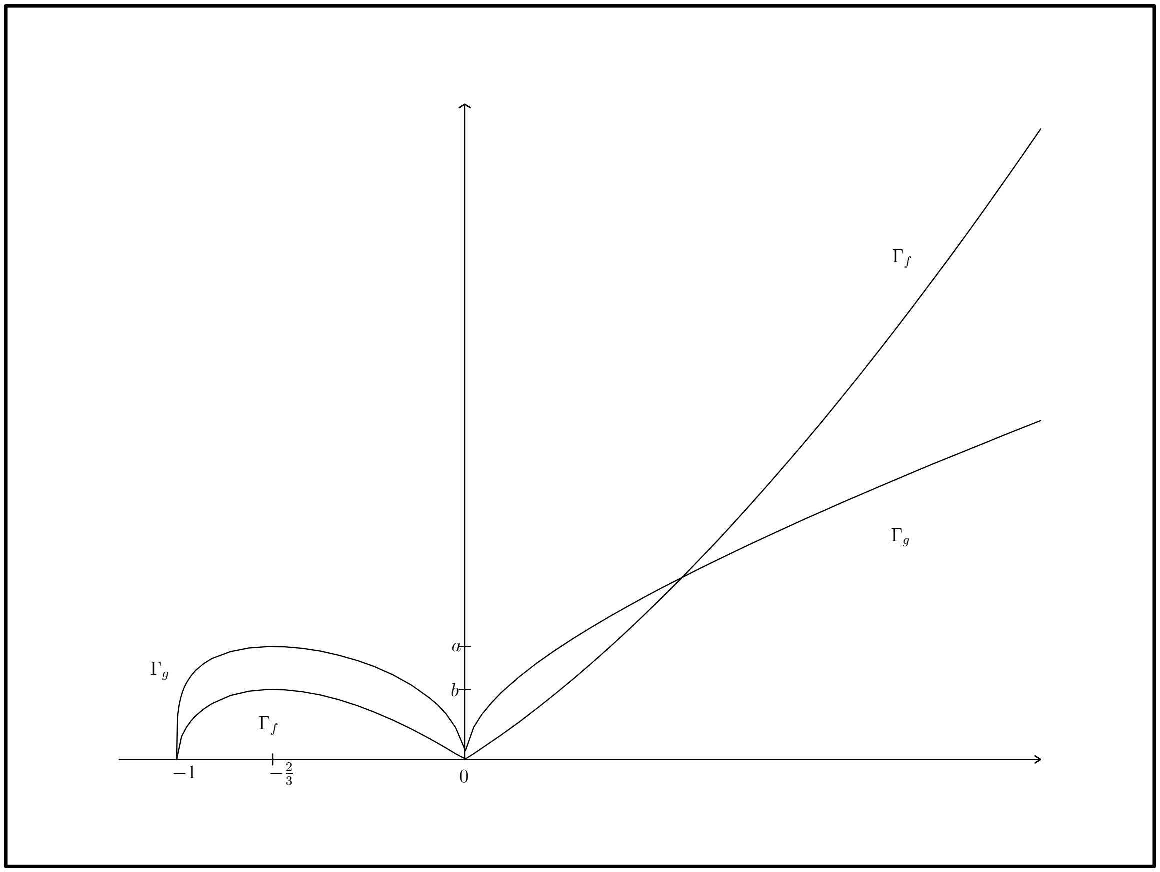
  f *und* g *nehmen in* -1 *und in* 0 *ihr (globales) Minimum an:*

  $$f(0) = g(0) = f(-1) = g(-1) = 0\;.$$

  Weder f *noch* g *besitzt ein globales Maximum wegen des Verhaltens bei* ∞*.*

- $f''(x) < 0$ *in* $(\,-1\,,\,0\,)$
  $f''(x) > 0$ *in* $(\,0\,,\,\infty\,)$
  $g''(x) < 0$ *in* $(\,-1\,,\,0\,) \cup (\,0\,,\,\infty\,)$
  Damit ist f *konkav in* $[\,-1\,,\,0\,]$ *und konvex in* $[\,0\,,\,\infty\,]$*;* g *ist konkav in* $[\,-1\,,\,0\,]$ *und in* $[\,0\,,\,\infty\,]$*.* f *besitzt in* 0 *einen Wendepunkt, nicht aber* g*.*
  Man kann sich nun fragen, ob g *konkav ist; aber das ist es nicht: nach Satz* 6.3 *wäre das Extremum in* 0 *ein lokales Maximum, wenn* g *konkav wäre.*

- *Skizze:*



Wir verwenden nun das Konzept der Konvexität zum Nachweis einiger wichtiger Ungleichungen.

Satz 6.22 *(Jensensche Ungleichung, J. L. Jensen 1895 - 1925)*
Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ eine konvexe Funktion und $N \in \mathbb{N}$. Sind $\lambda_1, \ldots, \lambda_N \in (0, 1)$ mit

$$\sum_{i=1}^{N} \lambda_i = 1, \tag{6.14}$$

so gilt für alle $x_1, \ldots, x_N \in I$

$$f(\sum_{i=1}^{N} \lambda_i x_i) \leq \sum_{i=1}^{N} \lambda_i f(x_i) \, . \tag{6.15}$$

Ist f sogar streng konvex, so gilt in (6.13) *Gleichheit genau dann, wenn*

$$x_1 = x_2 = \ldots = x_N \tag{6.16}$$

ist.

Sind $\lambda_1, \ldots, \lambda_N \geq 0$ mit $\sum_{i=1}^{N} \lambda_i = 1$, so nennt man $\sum_{i=1}^{N} \lambda_i x_i$ eine konvexe Kombination der Punkte $x_1, \ldots, x_N$. Der Begriff der konvexen Kombination läßt sich auch definieren, wenn $x_1, \ldots, x_N$ Punkte eines Vektorraums sind. Zeichnen Sie einmal einige Punkte des $\mathbb{R}^2$ und versuchen Sie, sich die Menge aller konvexen Kombinationen dieser Punkte vorzustellen.

Man zeigt übrigens leicht, dass alle konvexen Kombinationen von Punkten aus einem Intervall I ebenfalls in I liegen.

Beweis (von Satz 6.22**):** Ist (6.16) erfüllt, so steht auf beiden Seiten von (6.15) die gleiche Zahl, ebenso im Fall $N = 1$. Wir haben also (6.15) (mit scharfer Ungleichung im streng konvexen Fall) zu beweisen für den Fall, dass $x_1, \ldots, x_N$ nicht alle gleich sind. Den Beweis führt man mittels Induktion über N. Für $N = 2$ entspricht die Behauptung gerade der Definition 6.1 der Konvexität bzw. der strengen Konvexität.

Wir nehmen nun an, die Behauptung sei für irgendein $N \geq 2$ richtig. Seien nun $\lambda_1, \ldots, \lambda_{N+1} \in (\,0\,,\,1\,)$ mit

$$\sum_{i=1}^{N+1} \lambda_i = 1$$

und $x_1, \ldots, x_{N+1} \in I$ mögen nicht alle übereinstimmen.
O. B. d. A. nehmen wir

$$x_1 = \min\{x_1, \ldots, x_{N+1}\} < \max\{x_1, \ldots, x_{N+1}\} = x_{N+1}$$

an.

Setze

$$\mu := \lambda_{N+1} \quad , \quad t := x_{N+1} \quad , \quad s := \sum_{i=1}^{N} \frac{\lambda_i}{1-\mu} x_i \; .$$

und beachte

$$\sum_{i=1}^{N} \frac{\lambda_i}{1-\mu} = 1 \; .$$

Dann ist $t > s$ und nach 6.1 gilt (mit "$<$" statt des ersten "$\leq$" im Fall der strengen Konvexität)

$$\begin{aligned}\sum_{i=1}^{N+1} f(\lambda_i x_i) &= f((1-\mu)s + \mu t) \\ &\leq (1-\mu)f(s) + \mu f(t) \\ &\leq (1-\mu)\sum_{i=1}^{N} \frac{\lambda_i}{1-\mu} f(x_i) + \lambda_{N+1} f(x_{N+1}) \\ &= \sum_{i=1}^{N+1} \lambda_i f(x_i) \ .\end{aligned}$$

In der zweiten Abschätzung wurde die Induktionsannahme benutzt.

q.e.d.

Verwendet man die Exponentialfunktion als streng konvexe Funktion in der Jensenschen Ungleichung, so erhält man für alle $\lambda_1, \ldots, \lambda_N \in (\,0\,,\,1\,)$ mit

$$\sum_{i=1}^{N} \lambda_i = 1$$

und alle $x_1, \ldots, x_N \in \mathbb{R}$ die Abschätzung

$$\prod_{i=1}^{N} [\exp(x_i)]^{\lambda_i} = \exp\left(\sum_{i=1}^{N} \lambda_i x_i\right) \leq \sum_{i=1}^{N} \lambda_i \exp(x_i) \quad ;$$

Gleichheit gilt genau dann, wenn alle x_i übereinstimmen. Beachtet man, dass jede positive Zahl als Wert der Exponentialfunktion auftritt, erhält man die wichtige Ungleichung

Korollar 6.23 *Sind $y_1, \ldots, y_N > 0$ und $\lambda_1, \ldots, \lambda_N \in (\,0\,,\,1\,)$ mit*

$$\sum_{i=1}^{N} \lambda_i = 1 \quad ,$$

so gilt

$$\prod_{i=1}^{N} y_i^{\lambda_i} \leq \sum_{i=1}^{N} \lambda_i y_i \ . \tag{6.17}$$

Im Fall $\lambda_i := \frac{1}{N}$ ist dies die Ungleichung zwischen geometrischem und arithmetischem Mittel:

$$\left(\prod_{i=1}^{N} y_i\right)^{1/N} \leq \frac{1}{N}\sum_{i=1}^{N} y_i \tag{6.18}$$

Eine ganz wichtige Folgerung ist die Höldersche Ungleichung, welche die Cauchy-Schwarzsche Ungleichung verallgemeinert, die Sie aus der linearen Algebra kennen:

Korollar 6.24 *(Höldersche Ungleichung, Otto Hölder 1859 - 1937)*
Seien $p, q \in (\,1\,,\,\infty\,)$ mit

$$\frac{1}{p} + \frac{1}{q} = 1\ . \tag{6.19}$$

Dann gilt für alle $x, y \in \mathbb{R}^N$

$$\left|\sum_{i=1}^{N} x_i y_i\right| \leq \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{N} |y_i|^q\right)^{1/q}\ . \tag{6.20}$$

Beweis: Wegen

$$\left|\sum_{i=1}^{N} x_i y_i\right| \leq \sum_{i=1}^{N} |x_i||y_i|$$

kann o. B. d. A. $x_i, y_i \geq 0$ für $i = 1, \ldots, N$ angenommen werden.

Wir betrachten zunächst den Fall, dass $x_i, y_i \in \mathbb{R}^+$ sind für alle $i = 1, \ldots, N$ und dass gilt

$$\sum_{i=1}^{N} {y_i}^q = 1\ . \tag{6.21}$$

Wendet man die Jensensche Ungleichung mit der streng konvexen Funktion

$$f\ :\ [\,0\,,\,\infty\,) \longrightarrow \mathbb{R}\,,\ t \longmapsto t^p$$

und mit $\lambda_i := {y_i}^q$ auf die Zahlen $x_1 {y_1}^{1-q}, \ldots, x_N {y_N}^{1-q}$ an, so erhält man

$$\begin{aligned} \left(\sum_{i=1}^{N} x_i y_i\right)^p &= \left(\sum_{i=1}^{N} (x_i {y_i}^{1-q}){y_i}^q\right)^p \leq \sum_{i=1}^{N} {y_i}^q (x_i {y_i}^{1-q})^p \\ &= \sum_{i=1}^{N} {x_i}^p {y_i}^{q+p-pq} = \sum_{i=1}^{N} {x_i}^p\ ; \end{aligned} \tag{6.22}$$

denn $\frac{1}{p} + \frac{1}{q} = 1$ impliziert $q + p - pq = 0$.

Schwächt man nun (6.21) zu

$$\eta := \left(\sum_{i=1}^{N} {y_i}^q \right)^{1/q} > 0$$

ab, so ist

$$\sum_{i=1}^{N} \left(\frac{y_i}{\eta} \right)^q = 1 \quad ,$$

und man erhält aus (6.22)

$$\left(\sum x_i y_i \right)^p = \eta^p \left(\sum_{i=1}^{N} x_i \frac{y_i}{\eta} \right)^p \leq \eta^p \sum_{i=1}^{N} {x_i}^p \ . \tag{6.23}$$

Indem man beide Seiten mit $1/p$ potenziert, erhält man (6.20).

Schließlich sei P die Menge aller Indizes i, für die $x_i y_i > 0$ ist:

$$P := \{ i \in \mathbb{N} \colon 1 \leq i \leq N \wedge x_i y_i > 0 \}$$

Ist $P = \emptyset$, so ist (6.20) trivial.

Ist $L := \#P > 0$ und ist

$$\pi \ : \ \{1, \ldots, L\} \ \longrightarrow \ P$$

eine bijektive Funktion (also eine Abzählung von P), so liefert das bisher Bewiesene

$$\begin{aligned} \sum_{i=1}^{N} x_i y_i = \sum_{j=1}^{L} x_{\pi(j)} y_{\pi(j)} &\leq \left(\sum_{j=1}^{L} {x_{\pi(j)}}^p \right)^{1/p} \left(\sum_{j=1}^{L} {y_{\pi(j)}}^q \right)^{1/q} \\ &\leq \left(\sum_{i=1}^{N} {x_i}^p \right)^{1/p} \left(\sum_{i=1}^{N} {y_i}^q \right)^{1/q} \quad . \end{aligned}$$

q.e.d.

Wir werden später sehen, dass man die Funktion

$$\rho_p \ : \ \mathbb{R}^N \ \longrightarrow \ \mathbb{R} \, , \ \ x \ \longmapsto \ \left(\sum_{j=1}^{N} |x_j|^p \right)^{1/p}$$

zur Längenmessung (als ”Norm”) verwenden kann. In diesem Zusammenhang benötigt man die Minkowski-Ungleichung:

Korollar 6.25 *(Minkowski-Ungleichung, Hermann Minkowski 1865 - 1909)*
Sei $p \in (1, \infty)$. Für alle $x, y \in \mathbb{R}^N$ gilt

$$\left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{N} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |y_n|^p\right)^{1/p} .$$

Beweis: Sei $q \in (1, \infty)$ so, dass

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{6.24}$$

ist, also $q := p/(p-1)$. Für $x, y \in \mathbb{R}^N$ gilt dann zunächst unter Verwendung der Dreiecksungleichung und dann der Hölderschen Ungleichung:

$$\begin{aligned} \sum_{n=1}^{N} |x_n + y_n|^p &\leq \sum_{n=1}^{N} |x_n + y_n|^{p-1} |x_n| + \sum_{n=1}^{N} |x_n + y_n|^{p-1} |y_n| \\ &\leq \left[\sum_{n=1}^{N} |x_n + y_n|^{q(p-1)}\right]^{1/q} \left(\sum_{n=1}^{N} |x_n|^p\right)^{1/p} \\ &+ \left[\sum_{n=1}^{N} |x_n + y_n|^{q(p-1)}\right]^{1/q} \left(\sum_{n=1}^{N} |y_n|^p\right)^{1/p} \\ &= \left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{1/q} \left[\left(\sum_{n=1}^{N} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |y_n|^p\right)^{1/p}\right] . \end{aligned} \tag{6.25}$$

Für $x = -y$ ist die Minkowskiungleichung trivialerweise erfüllt. Ist $x \neq -y$, so erhält man die Minkowskiungleichung, indem man in (6.25) durch $\left(\sum_{n=1}^{N} |x_n + y_n|^p\right)^{1/q}$ dividiert und (6.24) beachtet.

q.e.d.

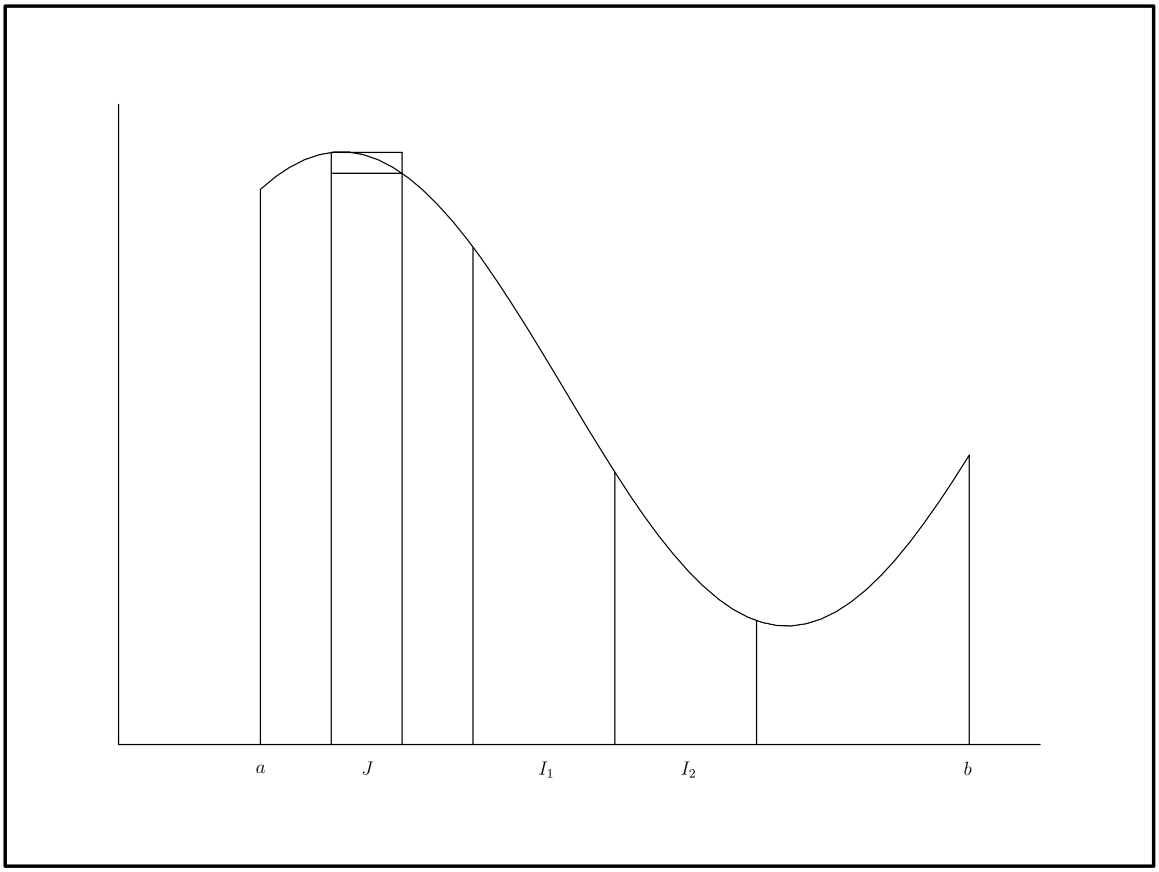
Kapitel 7

Das Riemannsche Integral

Die Berechnung von Flächen (und Volumina), die von krummlinigen Kurven (bzw. Flächen) begrenzt sind, führt auf den Begriff des Integrals. Methoden, solche Flächen zu berechnen, gehen bereits auf Archimedes zurück (287? – 212 v. Chr.). Aber erst von Leibniz und Newton wurde der Zusammenhang zwischen Inhaltsmessung und Differentiation entdeckt. Eine präzise Formulierung geht dann zunächst auf Augustin C. Cauchy (1789 – 1857) und allgemeiner auf Bernhard Riemann (1826 – 1866) zurück. Henri Lebesgue (1875 – 1941) hat zu Beginn des inzwischen letzten Jahrhunderts einen weitaus flexibleren Integralbegriff eingeführt, wie er in der modernen Analysis nötig gebraucht wird. Das Lebesgueintegral wird im kommenden Semester eingeführt. Wir wollen bei der Einführung des Riemannintegrals einer Idee von Serge Lang [8] folgen, und ein leicht zu verstehendes Axiomensystem für das Integral angeben. Daraus werden wir sehr schnell den Hauptsatz der Differential- und Integralrechnung gewinnen können, welcher Differentiation und Integration verknüpft.

7.1 Eine Axiomatische Begründung und der Hauptsatz der Differential– und Integralrechnung

Auf einem kompakten Intervall $I := [\,a\,,\,b\,]$, $a < b$, welches für den Rest dieses Abschnitts vorgegeben ist, sei eine reelle Funktion f gegeben. Bei der axiomatischen Begründung wollen wir uns diese Funktion zunächst positiv vorstellen, obwohl das nicht vorausgesetzt werden soll.



Der Inhalt der Fläche über $I_1 \cup I_2$ ist die Summe der Inhalte der Flächen über I_1 und I_2. Der Inhalt Fläche über J liegt zwischen den Fächeninhalten des ein– und des umbeschriebenen Rechtecks

Zunächst einige Bezeichnungen: Die Menge aller beschränkten Funktionen auf dem Intervall I bezeichnen wir mit

$$\mathcal{B}(I) := \{f \in \mathcal{F}(I, \mathbb{R})\colon\, f \text{ ist beschränkt}\}\ . \tag{7.1}$$

Die Familie aller kompakten Teilintervalle von I nennen wir

$$\mathcal{I} := \mathcal{I}(I) := \{J \subset I\colon\, J \text{ ist kompaktes Intervall}\}\ . \tag{7.2}$$

Im Fall $I = [\,0\,,\,1\,]$ liegen zum Beispiel $[\,\frac{1}{3}\,,\,\frac{2}{3}\,]$, $[\,0\,,\,1/2\,]$, $[\,0\,,\,1\,]$ in $\mathcal{I}$.

Eine Intervallfunktion über I ist eine Funktion

$$\mu\ :\ \mathcal{I}(I)\ \longrightarrow\ \mathbb{R}\quad ,$$

die jedem kompakten Teilintervall von I eine reelle Zahl zuordnet. Zum Beispiel ist

$$|\cdot| \,:\, \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\ [\,c\,,\,d\,] \longmapsto d-c$$

eine Intervallfunktion, und $|J|$ nennt man die Länge von J.

Definition 7.1 *Sei $f \in \mathcal{B}(I)$. Eine Flächeninhaltsfunktion (Fif) für f ist eine Intervallfunktion μ_f über I mit den beiden Eigenschaften:*

$$\bigwedge_{J\in\mathcal{I}(I)} \quad (\inf_J f)\cdot|J| \le \mu_f(J) \le (\sup_J f)|J| \tag{7.3}$$

$$\bigwedge_{I_1,I_2\in\mathcal{I}(I)} \quad \#(I_1\cap I_2) = 1 \Rightarrow \mu_f(I_1\cup I_2) = \mu_f(I_1)+\mu_f(I_2)\ . \tag{7.4}$$

f heißt "Riemann-integrierbar" auf I, wenn es für f eine und nur eine Flächeninhaltsfunktion über I gibt. Die Menge der auf I Riemann-integrierbaren Funktionen bezeichnen wir mit $\mathcal{R}(I)$. Für $[\,c\,,\,d\,] \in \mathcal{I}(I)$ und $f \in \mathcal{R}(I)$ schreiben wir

$$\int_c^d f(t)\,dt := \mu_f([\,c\,,\,d\,])\ . \tag{7.5}$$

(Hier kann t durch ein anderes Symbol ersetzt werden.)

Interpretiert man für *positives* $f \in \mathcal{B}(I)$ und $[\,c\,,\,d\,] \in \mathcal{I}(I)$ die Zahl $\mu_f([\,c\,,\,d\,])$ als den Flächeninhalt — was immer das sein mag —, der von dem Intervall $[\,c\,,\,d\,]$ auf der Abszisse, den Liniensegmenten von $\begin{bmatrix} c\\ 0\end{bmatrix}$ bis $\begin{bmatrix} c\\ f(c)\end{bmatrix}$ sowie von $\begin{bmatrix} d\\ 0\end{bmatrix}$ bis $\begin{bmatrix} d\\ f(d)\end{bmatrix}$ und dem Graphen von T_f umschlossenen Fläche, so sagt Bedingung (7.3), dass dieser Flächeninhalt zwischen dem des kleinsten einbeschriebenen und dem des größten umbeschriebenen Rechtecks mit achsenparallelen Kanten liegt. Die oben beschriebene Fläche wollen wir kurz "die Fläche über $[\,c\,,\,d\,]$" nennen. Dann sagt die zweite Bedingung, dass sich die Flächeninhalte der Flächen über $[\,c\,,\,d\,]$ und $[\,d\,,\,e\,]$ (mit $c \le d \le e$) zum Inhalt der Fläche über $[\,c\,,\,e\,]$ aufaddieren — was erwartet man anderes von einer Flächeninhaltsfunktion?

Auch wenn diese Interpretation für positives f gegeben wurde, so ist die Definition 7.1 doch für *jedes* $f \in \mathcal{B}(I)$ formuliert!

Die Schreibweise (7.5) ist nur sinnvoll, wenn $\mu_f([\,c\,,\,d\,])$ tatsächlich nur von dem Intervall $[\,c\,,\,d\,]$ und der Einschränkung von f auf $[\,c\,,\,d\,]$ abhängt. Dies wird sich aber als richtig herausstellen. Da wir zur Zeit mit einem festen Intervall I arbeiten, auf dem f definiert ist, steht es uns frei, die Schreibweise (7.5) bereits jetzt zu benutzen.

Wir werden gleich sehen, dass für jede Funktion $f \in \mathcal{B}(I)$ eine Fif über I existiert. Der wesentliche Punkt bei der Definition wird also die Eindeutigkeit der Fif sein.

Beispiel 7.2 *Sei*

$$f \;:\; I \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \begin{cases} 0\,, & \textit{falls}\; x \in \mathbb{R} \setminus \mathbb{Q} \\ 1\,, & \textit{falls}\; x \in \mathbb{Q} \end{cases} \quad .$$

Man sieht leicht, dass sowohl

$$\mu_f \;:\; \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\;\; J \longmapsto 0$$

als auch

$$\nu_f \;:\; \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\;\; J \longmapsto |J|$$

Flächeninhaltsfunktionen über I für f sind. Insbesondere ist $f \notin \mathcal{R}(I)$.

Beispiel 7.3 *Sei $\alpha \in \mathbb{R}$ eine vorgegebene Zahl. Für die konstante Funktion*

$$f \;:\; I \longrightarrow \mathbb{R}\,,\;\; t \longmapsto \alpha$$

existiert offenbar genau eine Flächeninhaltsfunktion auf I, nämlich

$$\mu_f \;:\; \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\;\; J \longmapsto \alpha|J|\;.$$

Konstante Funktionen sind also Riemann–integrierbar.

Ist μ_f eine Fif für f über I und ist $J \in \mathcal{I}(I)$, so gilt nach (7.4)

$$\mu_f(J) = \sum_{n=1}^{N} \mu_f(J_n) \quad , \tag{7.6}$$

sofern $J_1, \ldots, J_N$ kompakte Teilintervalle von J derart sind, dass

$$\#(J_n \cap J_{n+1}) = 1 \qquad \text{für alle} \quad n = 1, \ldots, n-1 \tag{7.7}$$

und

$$\bigcup_{n=1}^{N} J_n = J \tag{7.8}$$

gelten, sofern also $J_1, \ldots, J_N$ das Intervall J so in kompakte Teilntervalle zerlegen, von denen jeweils zwei der Teilintervalle höchstens einen Punkt gemeinsam haben. Mit (7.3) folgt dann

$$\sum_{n=1}^{N} (\inf_{J_n} f)|J_n| \le \mu_f(J) \le \sum_{n=1}^{N} (\sup_{J_n} f)|J_n| \tag{7.9}$$

Definition 7.4 *Sei* $[\,c\,,\,d\,] \in \mathcal{I}(I)$. *Eine endliche Teilmenge* $z \subset [\,c\,,\,d\,]$ *mit* $c, d \in z$ *heißt "Zerlegung des Intervalls"* $[\,c\,,\,d\,]$. $\mathcal{Z}([\,c\,,\,d\,])$ *ist die Menge aller Zerlegungen von* $[\,c\,,\,d\,]$*:*

$$\mathcal{Z}([\,c\,,\,d\,]) = \{z \subset [\,c\,,\,d\,] \colon z \text{ ist endlich und } c, d \in z\}$$

Enthält eine Zerlegung z *genau* $(n+1)$ *Punkte, so schreiben wir* $z = \{z_0, z_1, \ldots, z_n\}$ *mit* $z_{i-1} < z_i$ *für* $i = 1, \ldots, n$, *sowie* $n(z) := n$. *Für* $f \in \mathcal{B}(I)$, $J \in \mathcal{I}(I)$ *und* $z \in \mathcal{Z}(J)$ *heißen*

$$\overline{S}(z, J, f) := \sum_{i=1}^{n(z)} \left(\sup_{[\,z_{i-1}\,,\,z_i\,]} f \right) (z_i - z_{i-1}) \tag{7.10}$$

bzw.

$$\underline{S}(z, J, f) := \sum_{i=1}^{n(z)} \left(\inf_{[\,z_{i-1}\,,\,z_i\,]} f \right) (z_i - z_{i-1}) \tag{7.11}$$

Ober– bzw. Untersumme für f *über* J *zur Zerlegung* z.

Laut (7.9) gilt für jede Obersumme $\overline{S}(z, J, f)$ und jede Untersumme $\underline{S}(z, J, f)$

$$\underline{S}(z, J, f) \leq \mu_f(J) \leq \overline{S}(z, J, f) \quad , \tag{7.12}$$

sofern μ_f eine Fif für f über I ist.

Lemma 7.5 *Mit* $f \in \mathcal{B}(I)$,$J \in \mathcal{I}(I)$ *und* $z, \hat{z} \in \mathcal{Z}(J)$ *gelten:*

(i) $\overline{S}(\hat{z}, J, f) \leq \overline{S}(z, J, f)$ *falls* $z \subset \hat{z}$*;*

(ii) $\underline{S}(\hat{z}, J, f) \geq \underline{S}(z, J, f)$ *falls* $z \subset \hat{z}$ *;*

(iii) $\underline{S}(\hat{z}, J, f) \leq \overline{S}(z, J, f)$

Beweis: Zum Beweis von (i) beachte man, dass es Zahlen

$$i_0 = 0 < i_1 < i_2 \ldots < i_k = n(\hat{z})$$

gibt mit

$$\hat{z}_{i_j} = z_j \quad \text{für} \quad j = 0, \ldots, k := n(z) \ .$$

Mit dieser Bezeichnung errechnet man

$$\begin{aligned}
\overline{S}(\hat{z}, J, f) &= \sum_{l=1}^{n(\hat{z})} \left(\sup_{[\,\hat{z}_{l-1}\,,\,\hat{z}_l\,]} f \right) (\hat{z}_l - \hat{z}_{l-1}) \\
&= \sum_{j=0}^{k-1} \sum_{l=i_j+1}^{i_{j+1}} \left(\sup_{[\,\hat{z}_{l-1}\,,\,\hat{z}_l\,]} f \right) (\hat{z}_l - \hat{z}_{l-1}) \\
&\leq \sum_{j=0}^{k-1} \sum_{l=i_j+1}^{i_{j+1}} \left(\sup_{[\,\hat{z}_{i_j}\,,\,\hat{z}_{i_{j+1}}\,]} f \right) (\hat{z}_l - \hat{z}_{l-1}) \\
&= \sum_{j=0}^{k-1} \left(\sup_{[\,\hat{z}_{i_j}\,,\,\hat{z}_{i_{j+1}}\,]} f \right) (\hat{z}_{i_{j+1}} - \hat{z}_{i_j}) \\
&= \sum_{j=0}^{k-1} \left(\sup_{[\,z_j\,,\,z_{j+1}\,]} f \right) (z_{j+1} - z_j) = \overline{S}(z, J, f) \; .
\end{aligned}$$

Man erhält (ii) mit analoger Rechnung oder unter Beachtung von

$$\underline{S}(z, J, f) = -\overline{S}(z, J, -f)$$

(und entsprechend für $\hat{z}$).

(iii) erhält man schließlich aus (i) und (ii):

$$\underline{S}(\hat{z}, J, f) \overset{\text{(ii)}}{\leq} \underline{S}(z \cup \hat{z}, J, f) \leq \overline{S}(z \cup \hat{z}, J, f) \overset{\text{(i)}}{\leq} \overline{S}(z, J, f) \; .$$

Die mittlere Ungleichung ist trivial.

q.e.d.

Aus Lemma 7.5 folgt, dass für $f \in \mathcal{B}(I)$ und $J \in \mathcal{I}(I)$ die Menge aller Obersummen zu f über J

$$\overline{\mathcal{S}}(J, f) := \{\overline{S}(z, J, f)\colon\, z \in \mathcal{Z}(J)\}$$

nach unten und die Menge aller Untersummen

$$\underline{\mathcal{S}}(J, f) := \{\underline{S}(z, J, f)\colon\, z \in \mathcal{Z}(J)\}$$

nach oben beschränkt ist. Daher können wir definieren

Definition 7.6 *Für $f \in \mathcal{B}(I)$, $J \in \mathcal{I}(I)$ heißen*

$$\overline{\int_J} f(t)\,dt := \inf \overline{\mathcal{S}}(J, f)$$

$$\underline{\int_J} f(t)\,dt := \sup \underline{\mathcal{S}}(J, f)$$

das Ober– bzw. das Unterintegral von f über J.

Anders als die Schreibweise $\int_J f(t)\,dt$ bereiten Schreibweisen $\overline{\int}_J f(t)\,dt$ bzw. $\underline{\int}_J f(t)\,dt$ keine Kopfschmerzen. Die definierten Werte hängen offenbar nur vom Intervall J und von der Einschränkung von f auf J ab.

Ober– und Unterintegral liefern Flächeninhaltsfunktionen für f auf I: Wie bereits angekündigt, gibt es also zu jedem $f \in \mathcal{B}(I)$, eine Flächeninhaltsfunktion. Außerdem sind sie die "kleinste" und die "größte" Fif für f:

Satz 7.7 *Sei $f \in \mathcal{B}(I)$. Es gelten:*

(i) $$\bigwedge_{J\in\mathcal{I}(I)} \quad \underline{\int}_J f(t)\,dt \le \overline{\int}_J f(t)\,dt = -\underline{\int}_J (-f(t))\,dt$$

(ii) $\overline{\mu}_f \;:\; \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\;\; J \longmapsto \overline{\int}_J f(t)\,dt$

und

$\underline{\mu}_f \;:\; \mathcal{I}(I) \longrightarrow \mathbb{R}\,,\;\; J \longmapsto \underline{\int}_J f(t)\,dt$

sind Fif für f über I

(iii) *Für jede Flächeninhaltsfunktion μ_f für f über I gilt*

$$\bigwedge_{J\in\mathcal{I}(I)} \quad \underline{\mu}_f(J) \le \mu_f(J) \le \overline{\mu}_f(J)\ .$$

Beweis:
(i) folgt aus Lemma 7.5(iii) bzw. aus $\overline{S}(z,J,f) = -\underline{S}(z,J,-f)$ für jede Zerlegung $z \in \mathcal{Z}(J)$.
Zu (ii): Zunächst beweisen wir (7.3). Ist z irgendeine Zerlegung eines Intervalles $J \in \mathcal{I}(I)$, so ist

$$\begin{aligned}\overline{S}(z,J,f) &= \sum \sup_{[\,z_{j-1}\,,\,z_j\,]} f\cdot(z_j - z_{j-1})\\ &\le \sum \sup_J f\cdot(z_j-z_{j-1}) = \sup_J f\cdot|J|\ .\end{aligned}$$

Analog ist für jede Zerlegung $z \in \mathcal{Z}(J)$:

$$(\inf_J f)|J| \le \underline{S}(z,J,f)\ .$$

Folglich ist

$$(\inf_J f)\cdot|J| \le \underline{\int}_J f(t)\,dt \le \overline{\int}_J f(t)\,dt \le (\sup_J f)\cdot|J|\ .$$

Zum Nachweis von (7.4) beachte man, dass es genügt, $\overline{\mu}_f$ zu betrachten; denn für jedes Intervall $J \in \mathcal{I}(I)$ ist

$$\overline{\int}_J f(t)\,dt = -\underline{\int}_J (-f)(t)\,dt\ .$$

Seien nun

$$J_1 := [\,c\,,\,d\,], J_2 := [\,d\,,\,e\,] \in \mathcal{I}(I), J := J_1 \cup J_2 = [\,c\,,\,e\,]\ .$$

Sind $z^{(1)} \in \mathcal{Z}(J_1), z^{(2)} \in \mathcal{Z}(J_2)$, so ist $z^{(1)} \cup z^{(2)} \in \mathcal{Z}(J)$ und

$$\overline{S}(z^{(1)} \cup z^{(2)}, J, f) = \overline{S}(z^{(1)}, J_1, f) + \overline{S}(z^{(2)}, J_2, f) \quad ,$$

folglich

$$\begin{aligned}\overline{\int}_J f(t)\,dt &= \inf \overline{\mathcal{S}}(J,f) \qquad (7.13)\\ &\leq \inf \overline{\mathcal{S}}(J_1,f) + \inf \overline{\mathcal{S}}(J_2,f) = \overline{\int}_{J_1} f(t)\,dt + \overline{\int}_{J_2} f(t)\,dt\ .\end{aligned}$$

Andererseits gibt es zu jedem $\varepsilon > 0$ eine Zerlegung $z \in \mathcal{Z}(J)$ von J mit

$$\overline{\int}_J f(t)\,dt + \varepsilon \geq \overline{S}(z, J, f) \geq \overline{S}(z \cup \{d\}, J, f) \qquad (7.14)$$

Die letzte Ungleichung folgt mit Lemma 7.5(i). Nun sind

$$z^{(1)} := (z \cup \{d\}) \cap J_1 \in \mathcal{Z}(J_1)\ ,\ z^{(2)} := (z \cup \{d\}) \cap J_2 \in \mathcal{Z}(J_2)$$

und

$$\begin{aligned}\overline{S}(z \cup \{d\}, J, f) &= \overline{S}(z^{(1)}, J_1, f) + \overline{S}(z^{(2)}, J_2, f) \qquad (7.15)\\ &\geq \overline{\int}_{J_1} f(t)\,dt + \overline{\int}_{J_2} f(t)\,dt\ .\end{aligned}$$

Die Formeln (7.13) , (7.14) und (7.15) liefern

$$\overline{\int}_{J_1} f(t)\,dt + \overline{\int}_{J_2} f(t)\,dt = \overline{\int}_J f(t)\,dt\ ,$$

also (7.4) für $\overline{\mu}_f$.

(iii) erhält man schließlich aus Formel (7.12).

q.e.d.

Wir sind nun in der Lage, den Hauptsatz der Differential– und Integralrechnung zu beweisen:

Satz 7.8 *(Hauptsatz) Sei $I := [\,a\,,\,b\,]$ ein kompaktes Intervall. Dann gelten*

(i) $C(I) \subset \mathcal{R}(I)$

(ii) *Für jedes $f \in C(I)$ ist*

$$F \;:\; I \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \int_a^x f(t)\,dt$$

stetig und in $\overset{\circ}{I}$ differenzierbar. Es gelten

$$F(a) = 0$$

sowie für alle $x \in \overset{\circ}{I}$

$$F'(x) = f(x)\;.$$

(iii) *Sind $f, G \in C(I)$, und ist G differenzierbar in $\overset{\circ}{I}$ mit Ableitung $G' = f$ in $\overset{\circ}{I}$, so ist für alle Intervalle $[\,c\,,\,d\,] \in \mathcal{I}(I)$*

$$\int_c^d f(t)\,dt = G(d) - G(c)\;.$$

Schreibweise:

$$G(t)|_c^d := G(d) - G(c)$$

Beweis:
Wegen Satz 7.7 gibt es für $f \in C(I)$ eine Flächeninhaltsfunktion μ_f über I. Mit so einer Flächeninhaltsfunktion setzen wir

$$F \;:\; I \longrightarrow \mathbb{R}\,,\;\; x \longmapsto \mu_f([\,a\,,\,x\,])\;.$$

Dann ist $F(a) = 0$ und wegen (7.3) und (7.4) gilt für jedes $x \in \overset{\circ}{I}$ und positives $h \leq b - x$:

$$\min_{[\,x\,,\,x+h\,]} f \leq \frac{1}{h}(F(x+h) - F(x)) = \frac{1}{h}\mu_f([\,x\,,\,x+h\,]) \leq \max_{[\,x\,,\,x+h\,]} f \quad ,$$

und für $h \searrow 0$ konvergieren

$$\min_{[\,x\,,\,x+h\,]} f \text{ und } \max_{[\,x\,,\,x+h\,]} f$$

gegen $f(x)$.

Ist $h < 0$ und $a \leq x + h$, so erhält man

$$\min_{[\,x+h\,,\,x\,]} f \leq \frac{1}{h}(F(x+h) - F(x)) = \frac{1}{|h|}\mu_f([\,x+h\,,\,x\,]) \leq \max_{[\,x+h\,,\,x\,]} f\ .$$

Auch hier erhält man wieder Konvergenz gegen $f(x)$ für $h \nearrow 0$.

Die beiden Argumente liefern sogar, dass F in a rechts- und in b linksseitig differenzierbar ist:

$$d^+F(a) = f(a) \quad , \quad d^-F(b) = f(b)\ .$$

Der Schluss aus Korollar 6.10 liefert, dass F stetig ist.

Ist nun $\tilde{\mu}_f$ eine weitere Fif für f über I, so gilt auch für die auf I durch $\tilde{F}(x) := \tilde{\mu}_f([\,a\,,\,x\,])$ definierte Funktion:

$$\tilde{F}(a) = 0 \quad \text{und} \quad \tilde{F}'(x) = f(x) \quad \text{für alle} \quad x \in \overset{\circ}{I}\ .$$

Folglich ist

$$(\tilde{F} - F)(a) = 0 \quad , \quad (\tilde{F} - F)'(x) = 0 \quad \text{in} \quad \overset{\circ}{I} \quad ,$$

also $\tilde{F} = F$ nach Satz 5.38(iii). Ist $J := [\,c\,,\,d\,] \in \mathcal{I}(I)$ so folgt nach (7.4)

$$\mu_f(J) = \mu_f([\,a\,,\,d\,]) - \mu_f([\,a\,,\,c\,]) = F(d) - F(c) = \tilde{F}(d) - \tilde{F}(c) = \tilde{\mu}_f(J)\ .$$

Damit sind (i) und (ii) bewiesen.

Ist schließlich $G \in C(I)$ und $G' = f$ in $\overset{\circ}{I}$, so liefert Satz 5.36(iii)

$$G(x) = F(x) + \alpha$$

mit einer Zahl $\alpha \in \mathbb{R}$ für alle $x \in I$. Folglich ist für $[\,c\,,\,d\,] \in \mathcal{I}(I)$:

$$\int_c^d f(t)\,dt = \mu_f([\,c\,,\,d\,]) = F(d) - F(c) = G(d) - G(c)\ .$$

q.e.d.

Nach dem Hauptsatz der Differential- und Integralrechnung sind alle auf einem kompakten Intervall I stetigen Funktionen Riemann-integrierbar. Zugleich besagt der Satz, dass jede auf I stetige Funktion f eine Stammfunktion besitzt; das ist eine auf I stetige Funktion, deren Ableitung im Inneren von I mit f übereinstimmt. Außerdem liefert der Hauptsatz eine Methode, wie man Integrale zu gewissen Funktionen berechnen kann.

Definition 7.9 *Sei I ein Intervall und $f \in \mathcal{F}(I, \mathbb{R})$. Unter einer Stammfunktion zu f versteht man eine Funktion $F \in C(I)$, die für alle $x \in \overset{\circ}{I}$*

$$F'(x) = f(x)$$

erfüllt.

Beispiel 7.10

(i) $F : (0, \infty) \longrightarrow \mathbb{R}, \ x \longmapsto \ln x$ *ist Stammfunktion von*

$$f : (0, \infty) \longrightarrow \mathbb{R}, \ x \longmapsto \frac{1}{x} .$$

(ii) $F : [-1, 1] \longrightarrow \mathbb{R}, \ x \longmapsto \arcsin x$ *ist Stammfunktion von*

$$f : [-1, 1] \longrightarrow \mathbb{R}, \ x \longmapsto \begin{cases} (1-x^2)^{-1/2} & \textit{für} \quad |x| < 1 \\ 27 & \textit{für} \quad |x| = 1 . \end{cases}$$

(iii) $F : [-1, 1] \longrightarrow \mathbb{R}, \ x \longmapsto \frac{1}{3}x^3$ *ist Stammfunktion von*

$$f : [-1, 1] \longrightarrow \mathbb{R}, \ x \longmapsto x^2 .$$

Korollar 7.11 *Zu jeder auf einem kompakten Intervall I stetigen Funktion f gibt es eine Stammfunktion*

$$F : I \longrightarrow \mathbb{R} .$$

Ist F eine Stammfunktion zu f, so ist

$$\{F + \alpha \colon \alpha \in \mathbb{R}\}$$

die Menge aller Stammfunktionen von f.

Natürlich:

$$F : I \longrightarrow \mathbb{R}, \ x \longmapsto \int_a^x f(t)\,dt$$

ist Stammfunktion zu f. Zwei Stammfunktionen zu f unterscheiden sich wegen Satz 5.38(iii) nur um eine additive Konstante.

Bevor zu den Integrationstechniken übergegangen wird, soll der oberen Integrationsgrenze gestattet werden, kleiner als die untere zu sein:

Satz und Definition 7.12 *Mit* $-\infty < a < b < \infty$ *sei* $I := [\,a\,,\,b\,]$. *Wir schreiben für* $f \in \mathcal{R}(I)$:

$$\int_b^a f(t)\,dt := -\int_a^b f(t)\,dt\ .$$

Dann gilt für alle $c, d, e \in I$

$$\int_c^d f(t)\,dt + \int_d^e f(t)\,dt = \int_c^e f(t)\,dt\ .$$

Dies zeigt man, indem man die möglichen Reihenfolgen der Punkte c, d, e diskutiert.

7.2 Integrationstechniken

Der Hauptsatz der Differential- und Integralrechnung erlaubt zunächst die Integration all der Funktionen, von denen eine Stammfunktion bekannt ist. Hierzu lese man die Tabelle am Schluss des Abschnitts 5.6 von rechts nach links. Auch kann man die Linearität des Integrals ausnutzen, die später noch allgemein bewiesen wird. Für den Augenblick sei bemerkt, dass zumindest für stetige Funktionen $f, g \in C(I)$ und $\alpha, \beta \in \mathbb{R}$ die Formel

$$\int_a^b (\alpha f(t) + \beta g(t))\;dt = \alpha \int_a^b f(t)\,dt + \beta \int_a^b g(t)\,dt \tag{7.16}$$

gültig ist. Sind nämlich F und G Stammfunktionen von f bzw. g, so ist $\alpha F + \beta G$ Stammfunktion von $\alpha f + \beta g$. So erhält man aus

$$\bigwedge_{\alpha \in \mathbb{R}} \quad \int_a^b (\alpha + 1) t^\alpha\,dt = t^{\alpha+1}|_a^b\ , \tag{7.17}$$

(falls t^α für jedes $t \in [\,a\,,\,b\,]$ definiert ist) die Integrale

$$\bigwedge_{\alpha \in \mathbb{R}\setminus\{-1\}} \quad \int_a^b t^\alpha\,dt = \frac{1}{\alpha + 1} t^{\alpha+1}|_a^b\ , \tag{7.18}$$

$$\int_a^x (t^5 + 12t^4 - 13t)\,dt = \left.\left(\frac{t^6}{6} + \frac{12}{5}t^5 - \frac{13}{2}t^2\right)\right|_a^x\ . \tag{7.19}$$

Liegt 0 nicht im Integrationsintervall, so ist

$$\int_a^x \frac{1}{t}\,dt = \ln|t||_a^x = \ln\frac{x}{a}\ . \tag{7.20}$$

Die Integration algebraischer oder gebrochen rationaler Funktionen kann auf Funktionen führen, die weder algebraisch noch transzendent sind:

$$\int_a^x \frac{1}{\sqrt{1-t^2}}\,dt = \arcsin t|_a^x \quad \text{, falls} \quad a, x \in (\,-1\,,\,1\,) \tag{7.21}$$

$$\int_a^x \frac{1}{1+t^2}\,dt = \arctan t|_a^x\,. \tag{7.22}$$

Man kann aus den Differentiationsregeln weitere Integrationstechniken gewinnen. Zum Beispiel aus der Produktregel

$$(uv)' = u'v + uv' \tag{7.23}$$

für zwei Funktionen u, v. Kennt man eine Stammfunktion einer der beiden Summanden auf der rechten Seite von (7.23), so kennt man auch eine Stammfunktion des anderen Summanden. Dies führt auf die Methode der partiellen Integration. Wir geben hier aber nicht die schärfsten Voraussetzungen an u und v an.

Zunächst jedoch eine Schreibweise:

Definition 7.13 *Sei $n \in \mathbb{N}$ und I ein Intervall. Unter $C^n(I)$ verstehen wir den Vektorraum all der Funktionen f aus $C(I)$, die im Inneren $\overset{\circ}{I}$ von I n–mal differenzierbar sind und deren Ableitungen $f^{(1)}, \ldots, f^{(n)}$ stetig auf ganz I fortsetzbar sind. Mit $f^{(k)}$ soll die stetige Fortsetzung der k–ten Ableitung von f auf I bezeichnet werden.*

Satz 7.14 (Partielle Integration)
Sei $I := [\,a\,,\,b\,]$ ein kompaktes Intervall und $u, v \in C^1(I)$.
Dann gilt

$$\int_a^b u'(t)v(t)\,dt = (u(t)v(t))|_a^b - \int_a^b u(t)v'(t)\,dt\,. \tag{7.24}$$

Denn

$$\int_a^b u'(t)v(t)\,dt + \int_a^b u(t)v'(t)\,dt = \int_a^b (u\cdot v)'(t)\,dt = (uv)(t)|_a^b\,.$$

Bei der Darstellung der Beispiele markieren wir den Funktionsterm für u' mit einem Pfeil nach oben und den für v mit einem Pfeil nach unten:

$$\int_a^b \underset{\uparrow}{u'}\,(t)\,\underset{\downarrow}{v}\,(t)\,dt = u(t)v(t)|_a^b - \int_a^b u(t)v'(t)\,dt$$

u' wird integriert, v wird differenziert.

Beispiel 7.15 *Wir geben einige Beispiele:*

$$\int_a^x \underset{\uparrow}{t^\alpha} \quad \underset{\downarrow}{\ln t} \quad dt = \left(\frac{1}{\alpha+1} t^{\alpha+1} \ln t\right)\Big|_a^b - \int_a^b \frac{1}{\alpha+1} t^{\alpha+1} \cdot \frac{1}{t}\, dt \tag{7.25}$$
$$= \left(\frac{1}{\alpha+1} t^{\alpha+1} \ln t - \frac{1}{(\alpha+1)^2} t^{\alpha+1}\right)\Big|_a^b$$

(für $\alpha \in \mathbb{R} \setminus \{-1\}$, $a, x > 0$).

$$\int_a^x \underset{\uparrow}{\tfrac{1}{t}} \quad \underset{\downarrow}{\ln t} \quad dt = (\ln t)^2\big|_a^x - \int_a^x (\ln t) \cdot \frac{1}{t}\, dt \tag{7.26}$$
$$\Rightarrow \int_a^x \frac{1}{t} \ln t \, dt = \frac{1}{2} (\ln t)^2\big|_a^x \;.$$

Mit $n \in \mathbb{N}$:

$$\int_a^x \underset{\downarrow}{t^n} \quad \underset{\uparrow}{e^t} \quad dt = t^n e^t\big|_a^x - n \int_a^x t^{n-1} e^t \, dt = \ldots = e^t \sum_{j=0}^n (-1)^j \frac{n!}{(n-j)!} t^{n-j} \Bigg|_a^x \;. \tag{7.27}$$

Ferner:

$$\int_a^x \sin^2 t \; dt = \int_a^x \underset{\uparrow}{\sin t} \quad \underset{\downarrow}{\sin t} \quad dt = -\cos t \sin t\big|_a^x + \int_a^x \cos^2 t \, dt$$
$$= -\cos t \sin t\big|_a^x + \int_a^x (1 - \sin^2 t)\, dt$$
$$\Rightarrow \int_a^x \sin^2 t \, dt = \frac{1}{2}\left(\int_a^x 1\, dt - \cos t \sin t \Bigg|_a^x\right) \tag{7.28}$$
$$= \frac{1}{2}\,(t - \cos t \sin t)\big|_a^x \;.$$

Mit $n, k \in \mathbb{N}$, $n \neq k$:

$$\int_0^\pi \underset{\uparrow}{\sin nx} \quad \underset{\downarrow}{\sin kx} \quad dx = -\frac{1}{n} \cos nx \sin kx\big|_0^\pi + \frac{k}{n} \int_0^\pi \underset{\uparrow}{\cos nx} \quad \underset{\downarrow}{\cos kx} \quad dx$$
$$= 0 + \frac{k}{n}\left(\frac{1}{n} \sin nx \cos kx \Bigg|_0^\pi + \frac{k}{n} \int_0^\pi \sin nx \sin kx \, dx\right)$$
$$= \frac{k^2}{n^2} \int_0^\pi \sin nx \sin kx \, dx$$
$$\Rightarrow \int_0^\pi \sin nx \sin kx \, dx = 0 \;. \tag{7.29}$$

Als weitere Differentiationsregel haben wir die Kettenregel kennengelernt: Sei $g \in C^1([\,a\,,\,b\,])$, $f \in C(J)$, wobei J ein Interall ist, welches $g(I)$ umfasst: $J \supset g(I)$. Dann ist die durch

$$G(x) := \int_{g(a)}^{g(x)} f(s)\,ds$$

für $x \in [\,a\,,\,b\,]$ definierte Funktion G stetig und für $x \in (\,a\,,\,b\,)$ gilt

$$G'(x) = g'(x)f(g(x))\ ;$$

genauer:

$$G\ :\ [\,a\,,\,b\,] \longrightarrow \mathbb{R}\,,\ \ x \longmapsto \int_{g(a)}^{g(x)} f(s)\,ds \tag{7.30}$$

ist eine Stammfunktion von

$$h\ :\ [\,a\,,\,b\,] \longrightarrow \mathbb{R}\,,\ \ x \longmapsto f \circ g \cdot g'\ .$$

Da zudem $G(a) = 0$ ist, gilt die Regel

Satz 7.16 (Substitutionsmethode, Transformationssatz)
Seien $[\,a\,,\,b\,]$ ein kompaktes Intervall, J ein Intervall, sowie $f \in C(J)$ und $g \in C^1([\,a\,,\,b\,])$ mit $g([\,a\,,\,b\,]) \subset J$.

Dann ist

$$\int_a^b (f \circ g)(t)g'(t)\,dt = \int_{g(a)}^{g(b)} f(s)\,ds\ . \tag{7.31}$$

Die oben gegebene Begründung für den Transformationssatz enthält noch eine Lücke: Natürlich ist die Funktion G aus (7.30) stetig auf $[\,a\,,\,b\,]$, und nach der Kettenregel ist auch

$$G'(x) = (f \circ g)(x)g'(x) \quad , \tag{7.32}$$

zunächst aber nur für diejenigen $x \in (\,a\,,\,b\,)$, für die $g(x)$ kein Randpunkt von J ist. aber indem man sich f konstant über den Randpunkt hinaus fortgesetzt denkt, kann o. B. d. A. J als offen angenommen werden.

Beispiel 7.17 *Im folgenden sind a, x so zu wählen, dass der Integrand in $[\,a\,,\,x\,]$ definiert ist. Zum Beispiel müssen im letzten Integral* (7.36) *beide Grenzen a und x entweder in $(\,0\,,\,1\,)$ oder in $(\,1\,,\,\infty\,)$ liegen.*

$$\begin{aligned}\int_a^x \frac{1}{t}\ln t\,dt &= \int_{\ln a}^{\ln x} s\,ds = \frac{1}{2}s^2\Big|_{\ln a}^{\ln x} \\ &= \frac{1}{2}(\ln x)^2 - \frac{1}{2}(\ln a)^2 = \frac{1}{2}(\ln t)^2\Big|_a^x\ .\end{aligned} \tag{7.33}$$

$$\begin{aligned}\int_a^x t^3\sqrt{1-t^2}\,dt &= -\int_{1-a^2}^{1-x^2}\frac{1}{2}(1-s)\sqrt{s}\,ds = -\frac{1}{2}\int_{1-a^2}^{1-x^2}(s^{1/2}-s^{3/2})\,ds \\ &= -\left(\frac{1}{3}s^{3/2}-\frac{1}{5}s^{5/2}\right)\Bigg|_{1-a^2}^{1-x^2} \\ &= \left(\frac{1}{3}\left(1-t^2\right)^{3/2}-\frac{1}{5}\left(1-t^2\right)^{5/2}\right)\Bigg|_x^a \;. \qquad (7.34)\end{aligned}$$

$$\begin{aligned}\int_a^x \tan t\,dt &= -\int_a^x \frac{1}{\cos t}(-\sin t)\,dt \qquad (7.35)\\ &= -\int_{\cos a}^{\cos x}\frac{1}{s}\,ds = \ln|s|\big|_{\cos x}^{\cos a} = \ln\left(\frac{\cos a}{\cos x}\right)\end{aligned}$$

$$\int_a^x \frac{1}{t\ln t}\,dt = \int_{\ln a}^{\ln x}\frac{1}{s}\,ds = \ln|s|\big|_{\ln a}^{\ln x} = \ln\left(\frac{\ln x}{\ln a}\right) \qquad (7.36)$$

Bisher haben wir (7.31) von links nach rechts gelesen: ein Integral der Form, wie es auf der linken Seite steht, wird berechnet, indem man es in die Form der rechten Seite bringt und dann ausrechnet. Manchmal geht es aber auch in anderer Richtung:

Beispiel 7.18

$$\int_a^x (1-t^2)^{3/2}\,dt = \int_\alpha^\xi (1-\sin^2 s)^{3/2}\cos s\,ds \quad , \quad a,x\in[-1\,,\,1] \qquad (7.37)$$

Hierbei sind α *und* ξ *so zu wählen, dass* $\sin\alpha = a$ *und* $\sin\xi = x$ *gilt, also etwa*

$$\alpha := \arcsin a \quad , \quad \xi = \arcsin x \quad ; \qquad (7.38)$$

aber zum Beispiel auch

$$\alpha := (\arcsin a) + 28\pi \quad , \quad \xi = \arcsin x - 12\pi$$

wäre möglich.

Mit der Wahl aus (7.38) *sind* $\alpha,\xi\in[-\frac{\pi}{2}\,,\,\frac{\pi}{2}]$, *also ist* $\cos s\geq 0$ *zwischen* α *und* ξ. *Daher:*

$$\begin{aligned}\int_\alpha^\xi (1-\sin^2 s)^{3/2}\cos s\,ds &= \int_\alpha^\xi \cos^4 s\,ds = \int_\alpha^\xi \underset{\uparrow}{\cos s}\cdot\underset{\downarrow}{\cos^3 s}\;ds \\ &= \sin s\cdot\cos^3 s\big|_\alpha^\xi + 3\int_\alpha^\xi \sin^2 s\cos^2 s\,ds \\ &= \sin s\cos^3 s\big|_\alpha^\xi + 3\int_\alpha^\xi \cos^2 s\,ds - 3\int_\alpha^\xi \cos^4 s\,ds\;.\end{aligned}$$

Folglich ist

$$\int_\alpha^\xi (1-\sin^2 s)^{3/2}\cos s\, ds = \frac{1}{4}\sin s\cos^3 s\Big|_\alpha^\xi + \frac{3}{4}\int_\alpha^\xi \cos^2 s\, ds\ . \tag{7.39}$$

Dass wir in (7.28) *bereits* $\sin^2$ *integriert haben, machen wir uns zu Nutze und erhalten*

$$\int_\alpha^\xi \cos^2 s\, ds = \int_\alpha^\xi (1-\sin^2 s)\, ds = \frac{1}{2}(s+\cos s\sin s)\Big|_\alpha^\xi\ . \tag{7.40}$$

Kombiniert man (7.37), (7.39) *und* (7.40), *und setzt man für* α *und* ξ *gemäß* (7.38) *ein, so erhält man*

$$\begin{aligned}\int_a^x (1-t^2)^{3/2}\, dt &= \left(\frac{1}{4}\sin s\cos^3 s + \frac{3}{8}\sin s\cos s + \frac{3}{8}s\right)\Big|_{\arcsin a}^{\arcsin x}\\ &= \left(\frac{1}{4}t(1-t^2)^{3/2} + \frac{3}{8}t(1-t^2)^{1/2} + \frac{3}{8}\arcsin t\right)\Big|_a^x\end{aligned}$$

Die folgenden Beispiele bereiten die Integration gebrochen rationaler Funktionen vor.

Beispiel 7.19 *Mit* $n\in\mathbb{N}$ *ist*

$$\begin{aligned}\int_a^x (t^2+1)^{-n}\, dt &= \int_a^x \underset{\uparrow}{1}\ \cdot\ \underset{\downarrow}{(t^2+1)^{-n}}\, dt \qquad (7.41)\\ &= t(1+t^2)^{-n}\big|_a^x + 2n\int_a^x t^2(1+t^2)^{-(n+1)}\, dt\\ &= t(1+t^2)^{-n}\big|_a^x + 2n\int_a^x (1+t^2)^{-n}\, dt - 2n\int_a^x (1+t^2)^{-(n+1)}\, dt\ .\end{aligned}$$

Mithin ist

$$2n\int_a^x \frac{1}{(1+t^2)^{n+1}}\, dt = \frac{t}{(1+t^2)^n}\Big|_a^x + (2n-1)\int_a^x \frac{1}{(1+t^2)^n}\, dt\ .$$

Dies liefert eine Rekursion zur Bestimmung von $\int_a^x \frac{1}{(1+t^2)^n}\, dt$.

Für $n=1$ *ist nämlich*

$$\int_a^x \frac{1}{1+t^2}\, dt = \arctan t\big|_a^x \tag{7.42}$$

Ist $\int_a^x \frac{1}{(1+t^2)^n}\, dt$ *für ein* $n\in\mathbb{N}$ *bekannt, so ist*

$$\int_a^x \frac{1}{(1+t^2)^{n+1}}\, dt = \frac{1}{2n}\frac{t}{(1+t^2)^n}\Big|_a^x + \frac{2n-1}{2n}\int_a^x \frac{1}{(1+t^2)^n}\, dt\ . \tag{7.43}$$

Mit $p \in \mathbb{R}$ und $n \in \mathbb{N}$ sind

$$\int_a^x \frac{1}{(t-p)^n}\,dt = \begin{cases} \ln|t-p|\big|_a^x & \text{, falls} \quad n=1 \\ -\frac{1}{n-1}\frac{1}{(t-p)^{n-1}}\big|_a^x & \text{, falls} \quad n>1 \end{cases} \tag{7.44}$$

$$\int_a^x \frac{t}{(1+t^2)^n}\,dt = \frac{1}{2}\int_{1+a^2}^{1+x^2} s^{-n}\,ds = \begin{cases} \frac{1}{2}\ln s\Big|_{1+a^2}^{1+x^2} & \text{, falls} \quad n=1 \\ -\frac{1}{2(n-1)}s^{1-n}\Big|_{1+a^2}^{1+x^2} & \text{, falls} \quad n>1 \end{cases} \tag{7.45}$$

Sind $p, q \in \mathbb{R}$ mit $q - p^2/4 > 0$, so liefert der Transformationssatz 7.16 *mit $n \in \mathbb{N}$ und $d := \sqrt{q-p^2/4}$:*

$$\begin{aligned} \int_a^x \frac{1}{(t^2+pt+q)^n}\,dt &= \int_a^x \frac{1}{[(t+\frac{p}{2})^2+d^2]^n}\,dt = d^{-2n}\int_a^x \frac{1}{[(\frac{2t+p}{2d})^2+1]^n}\,dt \\ &= d^{1-2n}\int_{\frac{2a+p}{2d}}^{\frac{2x+p}{2d}} \frac{1}{(s^2+1)^n}\,ds \end{aligned} \tag{7.46}$$

sowie

$$\begin{aligned} \int_a^x \frac{t}{(t^2+pt+q)^n}\,dt &= \frac{1}{2}\int_a^x \frac{2t+p}{(t^2+pt+q)^n}\,dt - \frac{p}{2}\int_a^x \frac{1}{(t^2+pt+q)^n} \\ &= \frac{1}{2}\int_{a^2+pa+q}^{x^2+px+q} s^{-n}\,ds - \frac{p}{2}\int_a^x \frac{1}{(t^2+pt+q)^n}\,dt\ . \end{aligned} \tag{7.47}$$

Ist also g gegeben durch eine der Gleichungen

$$g(t) = \frac{1}{(t-p)^n} \quad , \quad g(t) = \frac{1}{(t^2+pt+q)^n} \quad , \quad g(t) = \frac{t}{(t^2+pt+q)^n} \tag{7.48}$$

mit $n \in \mathbb{N}$, $p, q \in \mathbb{R}$ und $q - p^2/4 > 0$, so liefern (7.42), (7.43), (7.44), (7.45), (7.46) *und* (7.47) *Formeln für die Berechnung von*

$$\int_a^x g(t)\,dt\ . \tag{7.49}$$

Das kann man ausnutzen, um gebrochen rationale Funktionen zu integrieren. Vorgelegt sei f in der Gestalt

$$f(t) = \frac{P(t)}{Q(t)} \tag{7.50}$$

mit zwei Polynomen P, Q. Wir setzen voraus, dass Q den Grad $N \in \mathbb{N}$ hat mit 1 als führendem Koeffizienten:

$$Q(t) = t^N + \sum_{n=0}^{N-1} \gamma_n t^n\ .$$

Der Grad von P sei *kleiner* als der von Q. Ist etwa $\tilde{P}$ ein Polynom, dessen Grad größer oder gleich dem von Q ist, so kann man wegen des Euklidischen Divisionsalgorithmus

$$\frac{\tilde{P}(t)}{Q(t)} = R(t) + \frac{P(t)}{Q(t)}$$

schreiben, wobei R, P Polynome sind und der Grad von P kleiner als der von Q ist.

Der Fundamentalsatz der Algebra, welcher in einem der folgenden Kapitel noch bewiesen wird, erlaubt es, das Polynom Q zu faktorisieren in der Form

$$Q(t) = \prod_{j=1}^{J} (t - \xi_j)^{\alpha(j)} \prod_{k=1}^{K} \left((t - \zeta_k)^{\beta(k)} \left(t - \overline{\zeta}_k\right)^{\beta(k)} \right) . \tag{7.51}$$

Dabei sind $\xi_1, \ldots, \xi_J$ die (paarweise verschiedenen) reellen Nullstellen von Q; $\zeta_1, \ldots, \zeta_K$ sind die (paarweise verschiedenen) komplexen Nullstellen mit positivem Imaginärteil

$$\zeta_k = \sigma_k + i\tau_k \quad , \quad \sigma_k \in \mathbb{R}, \tau_k \in \mathbb{R}^+ ,$$

und $\overline{\zeta}_1, \ldots, \overline{\zeta}_K$ sind die konjugiert komplexen Zahlen zu $\zeta_1, \ldots, \zeta_K$:

$$\overline{\zeta}_k = \sigma_k - i\tau_k .$$

Es gilt

$$\sum_{j=1}^{J} \alpha(j) + 2 \sum_{k=1}^{K} \beta(k) = N = \operatorname{Grad} Q .$$

Da — und das setzen wir voraus — unsere Polynome reelle Koeffizienten haben, sind $\overline{\zeta}_1, \ldots, \overline{\zeta}_K$ in der Tat die Nullstellen mit negativem Imaginärteil von Q. $\alpha(j), j = 1, \ldots, J$ bzw. $\beta(k), k = 1, \ldots, K$, sind die Vielfachheiten der Nullstellen von Q.

Der erste Schritt zur Integration von f besteht darin, das Nennerpolynom Q in der Gestalt

$$Q(t) = \prod_{j=1}^{J} (t - \xi_j)^{\alpha(j)} \prod_{k=1}^{K} (t^2 + p_k t + q_k)^{\beta(k)} \tag{7.52}$$

zu schreiben. Dabei ist

$$p_k = -2\sigma_k \quad , \quad q_k = |\zeta_k|^2 = \sigma_k^2 + \tau_k^2 ,$$

also

$$(t-\zeta_k)(t-\overline{\zeta}_k) = t^2 + p_k t + q_k \ .$$

Zum Beispiel ist

$$Q(t) := t^4 + 4t^3 + 7t^2 + 6t + 2 = (t+1)^2(t^2+2t+2) \quad , \tag{7.53}$$

und wir haben in diesem Fall $J = 1$, $\alpha(1) = 2$, $\xi_1 = -1$ sowie $K = 1$, $\beta(1) = 1$ und $\zeta_1 = -1 + i$, $\overline{\zeta}_1 = -1 - i$.

Liegt Q in der Gestalt (7.52) vor, so kann man folgenden Satz über die Partialbruchzerlegung benutzen:

Satz 7.20 *Gegeben seien zwei Polynome P, Q, $0 \leq \mathrm{Grad}\, P < \mathrm{Grad}\, Q =: N$, wobei Q in der Gestalt* (7.52) *dargestellt ist mit ($j = 1, \ldots, J$, $k = 1 \ldots K$) $\xi_j, p_k, q_k \in \mathbb{R}$, $q_k - p_k^2/4 > 0$ sowie $\alpha(j), \beta(k) \in \mathbb{N}$ und*

$$\sum_{j=1}^{J} \alpha(j) + 2 \sum_{k=1}^{K} \beta(k) = N \ . \tag{7.54}$$

Dann existieren eindeutig bestimmte reelle Zahlen A_{ij}, b_{lk}, C_{lk}, $j = 1, \ldots, J$, $i = 1, \ldots, \alpha(j)$, $k = 1, \ldots, K$, $l = 1, \ldots, \beta(k)$), so dass für alle $t \in \mathbb{R} \setminus \{\xi_1, \ldots, \xi_J\}$ gilt

$$\frac{P(t)}{Q(t)} = \sum_{j=1}^{J} \sum_{i=1}^{\alpha(j)} \frac{A_{ij}}{(t-\xi_j)^i} + \sum_{k=1}^{K} \sum_{l=1}^{\beta(k)} \frac{B_{lk}(t + \frac{p_k}{2}) + C_{lk}}{(t^2 + p_k t + q_k)^l} \ . \tag{7.55}$$

Der Satz soll nicht vollständig bewiesen werden. Wir begnügen uns mit dem Beweis für den Fall $K = 0$.
Beweis ((im Fall $K = 0$)):
Die Behauptung ist äquivalent zu

$$\bigwedge_{t \in \mathbb{R}} \quad P(t) = \sum_{j=1}^{J} \sum_{i=1}^{\alpha(j)} A_{ij} Q_{ij}(t) \quad , \tag{7.56}$$

wobei

$$Q_{ij}(t) = \frac{Q(t)}{(t-\xi_j)^i} \quad \text{für } t \neq \xi_j \tag{7.57}$$

bzw. für alle t:

$$Q_{ij}(t) = \left(\prod_{\substack{m=1 \\ m \neq j}}^{J} (t-\xi_m)^{\alpha(m)} \right) \cdot (t-\xi_j)^{\alpha(j)-i} \ . \tag{7.58}$$

Bei den Polynomen Q_{ij} handelt es sich wegen (7.54) um N Polynome jeweils höchstens $(N-1)$-ten Grades. Sie sind, wie wir gleich sehen, linear unabhängig, und bilden daher eine Basis des N-dimensionalen Vektorraums der Polynome maximal $(N-1)$-ten Grades. Daher gilt (7.56) mit eindeutig bestimmten Zahlen A_{ij}.

Die Menge $\{Q_{ij} \colon j = 1, \ldots, J\,,\ i = 1, \ldots, \alpha(j)\}$ ist in der Tat linear unabhängig: Ist $\mathbb{O}$ das Nullpolynom und gilt mit reellen Koeffizienten (A_{ij})

$$\sum_{j=1}^{J} \sum_{i=1}^{\alpha(j)} A_{ij} Q_{ij} = \mathbb{O} \quad , \tag{7.59}$$

so gilt auch für die Ableitungen

$$\sum_{j=1}^{J} \sum_{i=1}^{\alpha(j)} A_{ij} Q_{ij}^{(n)} = \mathbb{O}\ . \tag{7.60}$$

Wählt man einen Index $m \in \{1, \ldots, J\}$ und wertet man (7.59) an der Stelle ξ_m aus, so erhält man

$$A_{\alpha(m),1} = 0 \quad ,$$

denn

$$\begin{aligned}
&Q_{\alpha(m),1}(\xi_m) \neq 0 \quad , \\
&Q_{i,m}(\xi_m) = 0 \quad , \text{falls} \quad 1 \leq i < \alpha(m) \quad , \\
&Q_{i,j}(\xi_m) = 0 \quad , \text{falls} \quad j \neq m \quad \text{und} \quad 1 \leq i \leq \alpha(j)\ .
\end{aligned}$$

Ist $\alpha(m) > 1$, so erhält man sukzessive

$$A_{\alpha(m)-n,m} = 0 \quad , \quad n = 1, \ldots, \alpha(m) - 1$$

indem man (7.60) jeweils für $n = 1, \ldots, \alpha(m) - 1$ an der Stelle ξ_m auswertet, denn

$$\begin{aligned}
&Q^{(n)}_{\alpha(m)-n,m}(\xi_m) \neq 0\ , \\
&Q^{(n)}_{i,m}(\xi_m) = 0 \quad \text{für} \quad 1 \leq i < \alpha(m) - n\ , \\
&Q^{(n)}_{i,j}(\xi_m) = 0 \quad \text{für} \quad j \neq m\,,\ 1 \leq i \leq \alpha(j)\ .
\end{aligned}$$

Insgesamt erhält man also $A_{im} = 0$ für $i = 1, \ldots, \alpha(m)$.

q.e.d.

Beim Beweis im allgemeinen Fall ist es vernünftig, in $\mathbb{C}$ zur arbeiten. Das Vorgehen ist dann ähnlich.

Beispiel 7.21 *In allen folgenden Beispielen sind a und x so zu wählen, dass zwischen a und x keine Nullstelle des Nennerpolynoms liegt.*

$$\int_a^x \frac{7t^2 - 9t + 1}{t^3 - 7t + 6}\, dt = ?\ . \tag{7.61}$$

Beachte

$$\bigwedge_{t\in\mathbb{R}} t^3 - 7t + 6 = (t-1)(t-2)(t+3)$$

und mache den Ansatz:

$$\bigwedge_{t\in\mathbb{R}\setminus\{1,2-3\}} \quad \frac{7t^2 - 9t + 1}{(t-1)(t-2)(t+3)} = \frac{A_1}{t-1} + \frac{A_2}{t-2} + \frac{A_3}{t+3}\ .$$

Indem man auf beiden Seiten mit dem Nennerpolynom multipliziert erhält man die Gleichung

$$\bigwedge_{t\in\mathbb{R}} 7t^2 - 9t + 1 \stackrel{!}{=} A_1(t-2)(t+3) + A_2(t-1)(t+3) + A_3(t-1)(t-2)\ .$$

Wertet man diese Gleichung an den Stellen $1, 2$ *und* -3 *aus, erhält man*

$$\begin{aligned}
&7\cdot 1^2 - 9\cdot 1 + 1 = -1 \stackrel{!}{=} A_1(1-2)(1+3) \quad \textit{also} \quad A_1 = 1/4\\
&7\cdot 2^2 - 9\cdot 2 + 1 = 11 \stackrel{!}{=} A_2(2-1)(2+3) \quad \textit{also} \quad A_2 = 11/5\\
&7\cdot (-3)^2 - 9\cdot (-3) + 1 = 91 \stackrel{!}{=} A_3(-3-1)(-3-2) \quad \textit{also} \quad A_3 = 91/20\ .
\end{aligned}$$

Damit folgt

$$\begin{aligned}
\int_a^x \frac{7t^2 - 9t + 1}{t^3 - 7t + 6}\, dt &= \frac{1}{4}\int_a^x \frac{dt}{t-1} + \frac{11}{5}\int_a^x \frac{dt}{t-2} + \frac{91}{20}\int_a^x \frac{dt}{t+3}\\
&= \left(\frac{1}{4}\ln|t-1| + \frac{11}{5}\ln|t-2| + \frac{91}{20}\ln|t+3|\right)\Bigg|_a^x\ .
\end{aligned}$$

$$\int_a^x \frac{t^2 + 1}{(t-2)^2(t+3)}\, dt = ? \tag{7.62}$$

Der Ansatz

$$\bigwedge_{t\in\mathbb{R}\setminus\{2,-3\}} \quad \frac{t^2+1}{(t-2)^2(t+3)} = \frac{A_{11}}{(t-2)} + \frac{A_{21}}{(t-2)^2} + \frac{A_{12}}{t+3}$$

zusammen mit der Umbenennung $A := A_{11}, B := A_{21}, C := A_{12}$ *liefert*

$$\bigwedge_{t\in\mathbb{R}} \quad t^2+1 \stackrel{!}{=} A(t-2)(t+3) + B(t+3) + C(t-2)^2$$

Auswertung an den Stellen 2 *und* -3 *liefert*

$$\begin{aligned} 2^2+1 = 5 = B(2+3) \quad &, \textit{ also} \quad B = 1 \\ (-3)^2+1 = 10 = C(-3-2)^2 \quad &, \textit{ also} \quad C = \frac{2}{5} \end{aligned}$$

Mit diesen Werten für B *und* C *folgt*

$$\bigwedge_{t\in\mathbb{R}} (t^2+1) - (t+3) - \frac{2}{5}(t-2)^2 \stackrel{!}{=} A(t-2)(t+3)$$

und nach Differentiation

$$\bigwedge_{t\in\mathbb{R}} 2t-1-\frac{4}{5}(t-2) \stackrel{!}{=} A(t+3) + A(t-2) \; .$$

Auswertung an der Stelle 2 *liefert*

$$2\cdot 2 - 1 - \frac{4}{5}(2-2) = 3 \stackrel{!}{=} A(2+3) \quad , \textit{ also} \quad A = \frac{3}{5} \; .$$

Folglich ist

$$\begin{aligned} \int_a^x \frac{t^2+1}{(t-2)^2(t+3)}\,dt &= \frac{3}{5}\int_a^x \frac{dt}{t-2} + \int_a^x \frac{dt}{(t-2)^2} + \frac{2}{5}\int_a^x \frac{dt}{t+3} \\ &= \left(\frac{3}{5}\ln|t-2| - \frac{1}{t-2} + \frac{2}{5}\ln|t+3|\right)\Bigg|_a^x \; . \end{aligned}$$

$$\int_a^x \frac{1}{(t^2+2t+2)^2(t-2)}\,dt = ? \tag{7.63}$$

Mit dem Ansatz

$$\begin{aligned} \frac{1}{(t^2+2t+2)^2(t-2)} &= \frac{A}{t^2+2t+2} + \frac{B(t+1)}{t^2+2t+2} \\ &\quad + \frac{C}{(t^2+2t+2)^2} + \frac{D(t+1)}{(t^2+2t+2)^2} + \frac{E}{t-2} \end{aligned}$$

wird man auf

$$\begin{aligned} \bigwedge_{t\in\mathbb{R}} &(t^2+2t+2)(t-2)A + (t^2+2t+2)(t+1)(t-2)B \\ &+(t-2)C + (t+1)(t-2)D + (t^2+2t+2)^2E = 1 \end{aligned} \tag{7.64}$$

geführt. Ein Koeffizientenvergleich würde zu einem linearen Gleichungssystem für A, B, C, D, E *führen mit reeller Koeffizientenmatrix und reeller rechter Seite.* A, B, C, D, E *sind daher reell. Wertet man* (7.64) *für* $t = 2$ *aus, erhält man* $100E = 1$*, also*

$$E = \frac{1}{100} \ .$$

Das Polynom $t^2 + 2t + 2$ *hat die komplexen Nullstellen* $-1 + i$ *und* $-1 - i$*. Wertet man* (7.64) *für* $t = -1 + i$ *aus, erhält man*

$$(-3 + i)C + i(-3 + i)D = 1 \quad ,$$

und da C, D *reell sind,*

$$-3C - D = 1 \quad , \quad i(C - 3D) = 0 \ .$$

Mithin

$$D = -\frac{1}{10} \quad , \quad C = -\frac{3}{10} \ .$$

Setzt man in (7.64) $t = -1$ *ein, erhält man*

$$-3A - 3C + E = 1 \quad ,$$

also

$$A = -\frac{3}{100} \ .$$

Schließlich erhält man B*, indem man den Koeffizienten von* t^4 *auf beiden Seiten von* (7.64) *bestimmt. Dieser ist auf der linken Seite* $B + E$ *und auf der rechten Seite* 0*. Also folgt*

$$B = -E = -\frac{1}{100} \ .$$

Insgesamt folgt

$$\begin{aligned}
&\int_a^x \frac{dt}{(t^2+2t+2)^2(t-2)} = -\frac{3}{100}\int_a^x \frac{dt}{t^2+2t+2} - \frac{1}{100}\int_a^x \frac{t+1}{t^2+2t+2}\,dt \\
&- \frac{3}{10}\int_a^x \frac{dt}{(t^2+2t+2)^2} - \frac{1}{10}\int_a^x \frac{t+1}{(t^2+2t+2)^2}\,dt + \frac{1}{100}\int \frac{dt}{t-2} \\
&= \Bigg(-\frac{3}{100}\arctan(t+1) - \frac{1}{200}\ln\left(t^2+2t+2\right) - \frac{3}{20}\arctan(t+1) \\
&\qquad -\frac{3}{20}\left(\frac{t}{t^2+2t+2}\right) + \frac{1}{20}\left(\frac{1}{t^2+2t+2}\right) + \frac{1}{100}\ln|t-2|\Bigg)\Bigg|_a^x \qquad (7.65)
\end{aligned}$$

7.3 Riemann–integrierbare Funktionen

Hier soll keine vollständige Charakterisierung gegeben werden. Andererseits sollen aber auch Funktionen aufgezeigt werden, die wiewohl nicht stetig so doch Riemann-integrierbar sind. Ferner soll die Menge $\mathcal{R}(I)$ der Riemann-integrierbaren Funktionen auf dem kompakten Intervall $I := [\,a\,,\,b\,]$ — Bezeichnung gültig für den gesamten Abschnitt — als Vektorraum erkannt werden, in dem noch eine zusätzliche Multiplikation erklärt ist, welche nicht aus der Menge herausführt — die Multiplikation zweier Funktionen.

So eine Struktur nennt man eine Algebra.

Bei den Beweisen werden wir zum Teil Gebrauch vom Riemannschen Integrabilitätskriterium machen, das wir zunächst beweisen.

Satz 7.22 *Sei $f \in \mathcal{B}(I)$. Dann sind äquivalent:*

(i) $f \in \mathcal{R}(I)$

(ii) $\displaystyle\bigwedge_{J\in\mathcal{I}(I)} \quad \underline{\int_J} f(t)\,dt = \overline{\int_J} f(t)\,dt$

(iii) $\displaystyle\underline{\int_I} f(t)\,dt = \overline{\int_I} f(t)\,dt.$

(iv) *Riemannsches Integrabilitätskriterium:*
Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $z \in \mathcal{Z}(I)$ mit

$$\overline{S}(z, I, f) - \underline{S}(z, I, f) < \varepsilon\ .$$

Ist $f \in \mathcal{R}(I)$, so gilt für jedes $J \in \mathcal{I}(I)$

$$\int_J f(t)\,dt = \underline{\int_J} f(t)\,dt = \overline{\int_J} f(t)\,dt\ . \tag{7.66}$$

Beweis:
Dass (i) und (ii) äquivalent ist, folgt direkt aus Satz 7.7, ebenso (7.66).Natürlich folgt (iii) aus (ii); denn $I \in \mathcal{I}(I)$. Umgekehrt folgt (ii) aus (iii): Hierzu nehmen wir an, es gäbe ein kompaktes Intervall $[\,c\,,\,d\,] \in \mathcal{I}(I)$ mit

$$\underline{\int_c^d} f(t)\,dt < \overline{\int_c^d} f(t)\,dt\ .$$

Da — mit der Bezeichnung aus Satz 7.7 — $\underline{\mu}_f$ und $\overline{\mu}_f$ Flächeninhaltsfunktionen für f über I sind, folgt für $I = [\,a\,,\,b\,]$:

$$\underline{\int}_I f(t)\,dt = \underline{\int}_a^c f(t)\,dt + \underline{\int}_c^d f(t)\,dt + \underline{\int}_d^b f(t)\,dt$$
$$< \overline{\int}_a^c f(t)\,dt + \overline{\int}_c^d f(t)\,dt + \overline{\int}_d^b f(t)\,dt = \overline{\int}_I f(t)\,dt\ .$$

im Widerspruch zu (iii).

Aus (iii) folgt (iv); denn zu jedem $\varepsilon > 0$ existieren Zerlegungen $z^{(1)}, z^{(2)}, \in \mathcal{Z}(I)$ mit

$$\overline{S}(z^{(1)}, I, f) - \overline{\int}_I f(t)\,dt < \varepsilon/2 \quad , \quad \underline{\int}_I f(t)\,dt - \underline{S}(z^{(2)}, I, f) < \varepsilon/2\ ,$$

und für die Zerlegung $z := z^{(1)} \cup z^{(2)}$ folgt

$$\overline{S}(z, I, f) - \underline{S}(z, I, f) \le \overline{S}(z^{(1)}, I, f) - \underline{S}(z^{(2)}, I, f) =$$
$$= \overline{S}(z^{(1)}, I, f) - \overline{\int}_I f(t)\,dt + \underline{\int}_I f(t)\,dt - \underline{S}(z^{(2)}, I, f) < \varepsilon\ .$$

Dass schließlich(iv) die Bedingung (iii) impliziert, ist klar.

q.e.d.

Satz 7.23 *Mit der üblichen Addition und Multiplikation mit Skalaren ist $\mathcal{R}(I)$ ein Vektorraum über $\mathbb{R}$. Das Integral ist ein lineares Funktional auf $\mathcal{R}(I)$, d. h. die Abbildung*

$$R\ :\ \mathcal{R}(I) \longrightarrow \mathbb{R}\,,\ f \longmapsto \int_I f(t)\,dt$$

ist eine lineare Abbildung des Vektorraums $\mathcal{R}(I)$ in seinen Grundkörper.

Darüberhinaus ist R monoton: Gilt $f \le g$ für $f, g \in \mathcal{R}(I)$, d. h. $f(t) \le g(t)$ für jedes $t \in I$, so ist

$$\int_a^b f(t)\,dt \le \int_a^b g(t)\,dt\ .$$

Insbesondere gilt mit $||f||_\infty := \sup |f|$:

$$\left|\int_a^b f(t)\,dt\right| \le \int_a^b |f(t)|\,dt \le ||f||_\infty (b-a)\ .$$

Beweis:
Seien $f, g \in \mathcal{R}(I)$. Es gibt zu jedem $\varepsilon > 0$ eine Zerlegung $z \in \mathcal{Z}(I)$ mit

$$\int_I f(t)\,dt - \varepsilon \le \underline{S}(z, f, I) \le \int_I f(t)\,dt \tag{7.67}$$

$$\int_I f(t)\,dt \le \overline{S}(z, f, I) \le \int_I f(t)\,dt + \varepsilon \tag{7.68}$$

$$\int_I g(t)\,dt - \varepsilon \le \underline{S}(z, g, I) \le \int_I g(t)\,dt \tag{7.69}$$

$$\int_I g(t)\,dt \le \overline{S}(z, g, I) \le \int_I g(t)\,dt + \varepsilon \tag{7.70}$$

In der Tat können (7.67) – (7.70) mit der gleichen Zerlegung z erfüllt werden: sicher kann man die vier Ungleichungsketten mit vier eventuell verschiedenen Zerlegungen $z^{(1)}, \ldots, z^{(4)}$ erfüllen; dann sind sie aber mit $z := z^{(1)} \cup z^{(2)} \cup z^{(3)} \cup z^{(4)}$ richtig.

Nun ist

$$\begin{aligned} &\underline{S}(z, f, I) + \underline{S}(z, g, I) \le \underline{S}(z, f + g, I)\ , \\ &\overline{S}(z, f + g, I) \le \overline{S}(z, f, I) + \overline{S}(z, g, I)\ . \end{aligned}$$

Es folgt mit (7.67) – (7.70)

$$\begin{aligned} \int_I f(t)\,dt + \int_I g(t)\,dt - 2\varepsilon &\le \underline{S}(z, f + g, I) \le \underline{\int_I}(f + g)(t)\,dt \\ &\le \overline{\int_I}(f + g)(t)\,dt \le \overline{S}(z, f + g, I) \\ &\le \int_I f(t)\,dt + \int_I g(t)\,dt + 2\varepsilon\ . \end{aligned}$$

Für jedes $\varepsilon > 0$ gilt mithin

$$\begin{aligned} \int_I f(t)\,dt + \int_I g(t)\,dt - 2\varepsilon &\le \underline{\int_I}(f + g)(t)\,dt \\ &\le \overline{\int_I}(f + g)(t)\,dt \le \int_I f(t)\,dt + \int_I g(t)\,dt + 2\varepsilon\ , \end{aligned}$$

folglich

$$\int_I f(t)\,dt + \int_I g(t)dt = \underline{\int_I}(f + g)(t)\,dt = \overline{\int_I}(f + g)(t)\,dt\ .$$

Also ist $f + g \in \mathcal{R}(I)$ und

$$\int_I (f+g)(t)\,dt = \int_I f(t)\,dt + \int_I g(t)\,dt\ .$$

Ist $\alpha \in \mathbb{R}^+$ und $f \in \mathcal{R}(I)$, so ist $\alpha f \in \mathcal{R}(I)$ und

$$\int_I \alpha f(t)\,dt = \alpha \int_I f(t)\,dt\ ;$$

denn für jede Zerlegung $z \in \mathcal{Z}(I)$ sind

$$\overline{S}(z, \alpha f, I) = \alpha \overline{S}(z, f, I) \quad , \quad \underline{S}(z, \alpha f, I) = \alpha \underline{S}(z, f, I)\ .$$

Wegen 7.7(i) ist für $f \in \mathcal{R}(I)$ auch $-f \in \mathcal{R}(I)$ und

$$\int_I (-f)(t)\,dt = -\int_I f(t)\,dt\ .$$

Schließlich ist nach dem Hauptsatz (Satz 7.8) die Nullfunktion aus $\mathcal{R}(I)$ und hat Integral 0.

Damit ist $\mathcal{R}(I)$ Vektorraum und R ein lineares Funktional auf $\mathcal{R}(I)$.

Schließlich ist R monoton: Wegen der Linearität genügt es zu zeigen, dass für nicht negatives f das Integral über I nicht negativ ist. Ist also $f \geq 0$, so ist

$$R(f) = \int_I f(t)\,dt \geq (\inf_I f)|I| \geq 0\ .$$

q.e.d.

Satz 7.24 *Sind f und g Riemann-integrierbar auf I, so auch*

(i) $|f|$

(ii) $f \cdot g$

(iii) *f/g, sofern $|g(t)| \geq \delta$ ist für jedes $t \in I$ mit einer Zahl $\delta > 0$. (Man sagt in diesem Fall: g ist von 0 wegbeschränkt)*

(iv) *u, v mit $u(t) := \max\{f(t), g(t)\}$, $v(t) := \min\{f(t), g(t)\}$ für $t \in I$.*

Bei den Beweisen beachte man, dass sich die Differenz von Ober– und Untersumme zu einer Zerlegung $z \in \mathcal{Z}(I)$ schreiben läßt in der Form

$$\overline{S}(z,I,f) - \underline{S}(z,I,f) = \sum_{j=1}^{n(z)} \left(\sup_{s,t \in [z_{j-1},z_j]} |f(t) - f(s)| \right) (z_j - z_{j-1}) \ . \tag{7.71}$$

Beweis:
Sei $\varepsilon > 0$ vorgegeben, und z sei eine Zerlegung von I, für die

$$\overline{S}(z,I,f) - \underline{S}(z,I,f) < \varepsilon, \quad \overline{S}(z,I,g) - \underline{S}(z,I,g) < \varepsilon$$

gelten. Im Folgenden erstrecken sich sämtliche Summen von 1 bis $n(z)$.

(i) erhält man aus

$$\begin{aligned}\overline{S}(z,I,|f|) - \underline{S}(z,I,|f|) &= \sum_j \sup_{s,t \in [z_{j-1},z_j]} ||f|(t) - |f|(s)| \, (z_j - z_{j-1}) \\ &\le \sum_j \sup_{s,t \in [z_{j-1},z_j]} |f(t) - f(s)| \, (z_j - z_{j-1}) \\ &= \overline{S}(z,I,f) - \underline{S}(z,I,f) < \varepsilon \ .\end{aligned}$$

Zu (ii): Mit $C := \max\{||f||_\infty, ||g||_\infty\}$ ist

$$\begin{aligned}&\overline{S}(z,fg,I) - \underline{S}(z,fg,I) = \sum_j \sup_{s,t \in [z_{j-1},z_j]} |f(s)g(s) - f(t)g(t)|(z_j - z_{j-1}) \\ &\le \sum_j \sup_{s,t \in [z_{j-1},z_j]} \left(|f(s) - f(t)| \cdot |g(s)| + |g(s) - g(t)| \cdot |f(t)|\right)(z_j - z_{j-1}) \\ &\le C\left(\overline{S}(z,I,f) - \underline{S}(z,I,f)\right) + C\left(\overline{S}(z,I,g) - \underline{S}(z,I,g)\right) \le 2C\varepsilon \ .\end{aligned}$$

Zu (iii): Es genügt, $1/g \in \mathcal{R}(I)$ zu zeigen:

$$\begin{aligned}\overline{S}(z,1/g,I) - \underline{S}(z,1/g,I) &= \sum_j \sup_{s,t \in [z_{j-1},z_j]} \left| \frac{g(t) - g(s)}{g(t)g(s)} \right| (z_j - z_{j-1}) \\ &\le \frac{1}{\delta^2}\left(\overline{S}(z,g,I) - \underline{S}(z,g,I)\right) < \frac{\varepsilon}{\delta^2} \ .\end{aligned}$$

(iv) folgt aus

$$u = \frac{1}{2}(f+g) + \frac{1}{2}|f-g| \quad , \quad v = \frac{1}{2}(f+g) - \frac{1}{2}|f-g| \ .$$

q.e.d.

Sind $f \in C(I)$, $g \in \mathcal{R}(I)$ mit $g \geq 0$, so liefert die Monotonie des Integrals

$$(\min_I f) \int_I g(t)\, dt \leq \int_I f(t) g(t)\, dt \leq (\max_I f) \int_I g(t)\, dt\ . \tag{7.72}$$

Folglich ist

$$\int_I f(t) g(t)\, dt = \alpha \int g(t)\, dt \tag{7.73}$$

mit einer reellen Zahl

$$\alpha \in [\,\min_I f\,,\ \max_I f\,]\ . \tag{7.74}$$

Nach dem Zwischenwertsatz (Korollar 4.21) existiert ein $\xi \in I$ mit

$$\alpha = f(\xi)\ . \tag{7.75}$$

Das kann man noch etwas verschärfen zum Mittelwertsatz der Integralrechnung.

Satz 7.25 *Für jedes $f \in C(I)$ und $g \in \mathcal{R}(I)$ mit $g \geq 0$ gibt es ein $\xi \in \overset{\circ}{I}$ mit*

$$\int_I f(t) g(t)\, dt = f(\xi) \int_I g(t)\, dt\ . \tag{7.76}$$

Beweis:
Zu zeigen bleibt, dass ξ mit (7.76) sogar im Inneren des Intervalls existiert. Zunächst kann o. B. d. A. angenommen werden, dass $\int_I g(t)\, dt > 0$ ist.

Für die Zahl α aus (7.74), (7.75) diskutieren wir die drei Möglichkeiten

$$\min f < \alpha < \max f \tag{7.77}$$

$$\alpha = \max f \tag{7.78}$$

$$\alpha = \min f \tag{7.79}$$

Im ersten Fall gibt es ein ξ zwischen der Minimal- und der Maximalstelle von f, also sicher im Inneren von I mit (7.76).

Im zweiten Fall argumentiert man wie folgt: Da $\int_I g(t)\, dt > 0$ ist und da $\lim\limits_{\delta \to 0} \int_a^{a+\delta} g(t)\, dt = \lim\limits_{\delta \to 0} \int_{b-\delta}^{b} g(t)\, dt = 0$ sind, gibt es ein Intervall $J := [\,c\,,\ d\,]$ mit

$$J \subset \overset{\circ}{I}\ , \tag{7.80}$$

$$\int_J g(t)\,dt > 0\ . \tag{7.81}$$

Mit (7.72), (7.73) und (7.78) folgt

$$\begin{aligned}(\max_I f)\int_a^c g(t)\,dt + (\max_I f)\int_c^d g(t)\,dt + (\max_I f)\int_d^b g(t)\,dt &= \int_I f(t)g(t)\,dt \\ \le\ \max_I f\int_a^c g(t)\,dt + \max_J f\int_c^d g(t)\,dt + \max_I f\int_d^b g(t)\,dt,&\end{aligned}$$

und da $\int_c^d g(t)\,dt > 0$ ist:

$$\max_I f \le \max_J f\ .$$

Andererseits ist natürlich $\max\limits_J f \le \max\limits_I f$, also folgt

$$\max_J f = \max_I f,$$

und es gibt ein $\xi \in J \subset \overset{\circ}{I}$ mit $f(\xi) = \max\limits_J f = \max\limits_I f$.

Analog schließt man im Fall (7.79).

q.e.d.

Satz 7.26 *Ist $f \in \mathcal{B}(\mathcal{I})$ monoton, so ist f Riemann-intgrierbar.*

Beweis:
O. B. d. A. sei f monoton wachsend und $f(b) > f(a)$. Ist $\varepsilon > 0$ vorgegeben und ist $z \in \mathcal{Z}(I)$ eine "äquidistante" Zerlegung, für die

$$\bigwedge_{j=1,\dots n(z)} (z_j - z_{j-1}) = \delta < \frac{\varepsilon}{f(b) - f(a)}$$

gilt, so folgt

$$\overline{S}(z, I, f) - \underline{S}(z, I, f) = \sum_j (f(z_j) - f(z_{j-1}))(z_j - z_{j-1}) = (f(b) - f(a))\delta < \varepsilon\ .$$

q.e.d.

Satz 7.27 *Sei M eine endliche Teilmenge von I, $f \in \mathcal{B}(I)$ derart, dass für jedes Intervall $J \in \mathcal{I}(I)$ mit $J \cap M = \emptyset$ gilt: $f_{|J} \in \mathcal{R}(J)$. Dann ist $f \in \mathcal{R}(I)$.*

Beweis:
Sei $z \in \mathcal{Z}(I)$ die Zerlegung $M \cup \{a, b\}$. Ferner stehe $\fint$ für das Ober– oder Unterintegral.

Für jedes

$$\delta < \frac{1}{2} \min\{z_j - z_{j-1} \colon j = 1, \dots, n(z)\}$$

gilt mit $k := n(z)$

$$\begin{aligned}
\fint_I f(t)\,dt &= \fint_{z_0}^{z_0+\delta} f(t)\,dt + \sum_{j=1}^{k-1} \left(\fint_{z_{j-1}+\delta}^{z_j-\delta} f(t)\,dt + \fint_{z_j-\delta}^{z_j+\delta} f(t)\,dt \right) \\
&\quad + \fint_{z_{k-1}+\delta}^{z_k-\delta} f(t)\,dt + \fint_{z_k-\delta}^{z_k} f(t)\,dt \\
&= \sum_{j=1}^{k} \int_{z_{j-1}+\delta}^{z_j-\delta} f(t)\,dt + \left\{ \fint_{z_0}^{z_0+\delta} f(t)\,dt + \fint_{z_k-\delta}^{z_k} f(t)\,dt + \sum_{j=1}^{k-1} \fint_{z_j-\delta}^{z_j+\delta} f(t)\,dt \right\}
\end{aligned}$$

denn $f \in \mathcal{R}([z_{j-1} + \delta, z_j - \delta])$ für jedes j. Die Beträge der Ober– bzw. Unterintegrale in der geschweiften Klammer lassen sich abschätzen durch

$$\left| \fint_{z_0}^{z_0+\delta} f(t)\,dt \right| + \left| \fint_{z_k-\delta}^{z_k} f(t)\,dt \right| + \sum_{j=1}^{k-1} \left| \fint_{z_j-\delta}^{z_j+\delta} f(t)\,dt \right| \leq 2k||f||_\infty \cdot \delta \,.$$

Folglich ist

$$\overline{\int_I} f(t)\,dt - \underline{\int_I} f(t)\,dt \leq 4k||f||_\infty \cdot \delta \,.$$

Da dies für jedes $\delta > 0$ wahr ist, folgt $\underline{\int_I} f(t)\,dt = \overline{\int_I} f(t)\,dt$.

q.e.d.

Insbesondere liefert der Beweis als Formel zur Berechung des Integrals

$$\int_I f(t)\,dt = \lim_{\delta \searrow 0} \sum_{j=1}^{k} \int_{z_{j-1}+\delta}^{z_j-\delta} f(t)\,dt \,. \tag{7.82}$$

Nach Satz 7.27 ist zum Beispiel jede beschränkte Funktion f integrierbar, für die es eine endliche Teilmenge M von I derart gibt, dass f stetig in $I \setminus M$ ist.

Auch erkennt man aus Formel (7.82):

Korollar 7.28 *Sei $M \subset I$ eine endliche Menge. Ist $f \in \mathcal{R}(I)$ und ist $g \in \mathcal{B}(I)$ mit $g_{|I\setminus M} = f_{|I\setminus M}$, so ist $g \in \mathcal{R}(I)$ und*

$$\int_I f(t)\,dt = \int_I g(t)\,dt\ .$$

Denn $g \in \mathcal{B}(I)$, für jedes $J \in \mathcal{I}(I)$ mit $J \cap M = \emptyset$ ist $g \in \mathcal{R}(J)$, und mit $z := M \cup \{a, b\}$ ist

$$\begin{aligned}\int_I g(t)\,dt &= \lim_{\delta\searrow 0} \sum_{j=1}^{k} \int_{z_{j-1}+\delta}^{z_j-\delta} f(t)\,dt \\ &= \int_I f(t)\,dt - \lim_{\delta\searrow 0} \sum_{j=1}^{n(z)} \left(\int_{z_{j-1}}^{z_{j-1}+\delta} f(t)\,dt + \int_{z_j-\delta}^{z_j} f(t)\,dt \right) = \int_I f(t)\,dt\ .\end{aligned}$$

Korollar 7.29 *Seien $c \in (\,a\,,\,b\,)$, $f \in \mathcal{R}([\,a\,,\,c\,])$, $g \in \mathcal{R}([\,c\,,\,b\,])$ sowie α irgendeine reelle Zahl. Dann ist*

$$h\ :\ [\,a\,,\,b\,] \longrightarrow \mathbb{R}\,,\ \ x \longmapsto \begin{cases} f(t) & in\ \ [\,a\,,\,c\,) \\ g(t) & in\ \ (\,c\,,\,b\,] \\ \alpha & in\ \ c \end{cases}$$

aus $\mathcal{R}(I)$ und

$$\int_a^b h(t)\,dt = \int_a^c f(t)\,dt + \int_c^b g(t)\,dt\ .$$

Denn $h \in \mathcal{B}(I)$, für jedes $J \in \mathcal{I}(I)$ mit $J \cap \{c\} = \emptyset$ ist $h \in \mathcal{R}(J)$ — dort ist nämlich $h = f$ oder $h = g$ —, und

$$\int_I h(t)\,dt = \lim_{\delta\searrow 0} \left(\int_{a+\delta}^{c-\delta} f(t)\,dt + \int_{c+\delta}^{b-\delta} g(t)\,dt \right) = \int_a^c f(t)\,dt + \int_c^b g(t)\,dt\ .$$

7.4 Uneigentliche Integrale

Wir wollen nicht nur beschränkte Funktionen und nicht nur über kompakte Intervalle integrieren. Hierzu erklären wir für eine Teilmenge $X \subset \mathbb{R}$ den Vektorraum $\mathcal{R}^{\text{loc}}(X)$ als Raum der Funktionen f auf X, die auf jedem in X enthaltenen kompakten Intervall $J \subset X$ Riemann–integrierbar sind.

Zum Beispiel ist mit $\alpha \in \mathbb{R}$

$$f\ :\ \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}\,,\ \ x \longmapsto |x|^{-\alpha}$$

ein Element aus $\mathcal{R}^{\text{loc}}\,(\mathbb{R}\setminus\{0\})$.

Zunächst soll das Integral erklärt werden für gewisse Funktionen aus $\mathcal{R}^{\text{loc}}\,(X)$ im Fall eines halboffenen Intervalles X.

Definition 7.30 *Sei X ein halboffenes Intervall, $X := [\,a\,,\,b\,)$ mit $a \in \mathbb{R}$, $b \in \hat{\mathbb{R}}$, $a < b$ oder $X := (\,a\,,\,b\,]$ mit $a \in \hat{\mathbb{R}}$, $a < b$. Die Funktion $f \in \mathcal{R}^{\text{loc}}\,(X)$ heißt uneigentlich Riemann–integrierbar über X genau dann, wenn*

$$\int_X f(t)\,dt = \int_a^b f(t)\,dt := \lim_{R \nearrow b} \int_a^R f(t)\,dt \quad \text{im Fall } X := [\,a\,,\,b\,)\,,$$
$$\int_X f(t)\,dt = \int_a^b f(t)\,dt := \lim_{R \searrow a} \int_R^b f(t)\,dt \quad \text{im Fall } X := (\,a\,,\,b\,]$$

existiert. Offenbar ist die Menge aller uneigentlich Riemann-integrierbaren Funktionen, versehen mit üblicher Addition und Multiplikation mit reellen Zahlen, ein Vektorraum, und wir bezeichnen diesen mit $\mathcal{R}_u(X)$.

Beispiel 7.31 *Für $X := (\,0\,,\,b\,]$, $b \in \mathbb{R}$ und $f(x) := x^{-\alpha}$, $\alpha \in \mathbb{R}$, gilt — sofern der Limes existiert —*

$$\int_{(\,0\,,\,b\,]} x^{-\alpha}\,dx = \lim_{R \searrow 0} \int_R^b x^{-\alpha}\,dx = \lim_{R \searrow 0} \begin{cases} \frac{1}{1-\alpha} x^{1-\alpha}\big|_R^b & \text{, falls} \quad \alpha \neq 1 \\ \ln x\big|_R^b & \text{, falls} \quad \alpha = 1 \end{cases} .$$

Folglich ist $f \in \mathbb{R}_u(X)$, falls $\alpha < 1$, aber $f \notin \mathbb{R}_u(X)$, falls $\alpha \geq 1$.

Für $X := [\,a\,,\,\infty\,)$, $a \in \mathbb{R}$, $a > 0$ und $f(x) = x^{-\alpha}$, $\alpha \in \mathbb{R}$ erhält man

$$\int_{[\,a\,,\,\infty\,)} x^{-\alpha}\,dx = \lim_{R \to \infty} \begin{cases} \frac{1}{1-\alpha} x^{1-\alpha}\big|_a^R & \text{, falls} \quad \alpha \neq 1 \\ \ln x\big|_a^R & \text{, falls} \quad \alpha = 1 \end{cases} ,$$

und folglich $f \in \mathcal{R}_u([\,a\,,\,\infty\,))$ im Fall $\alpha > 1$, $f \notin \mathcal{R}_u([\,a\,,\,\infty\,))$ im Fall $\alpha \leq 1$.

Für die Integrale erhält man

$$\int_{(\,0\,,\,b\,]} x^{-\alpha}\,dx = \frac{1}{1-\alpha} b^{1-\alpha}, \text{ falls} \quad \alpha < 1$$
$$\int_{[\,a\,,\,\infty\,)} x^{-\alpha}\,dx = \frac{1}{\alpha-1} a^{1-\alpha}, \text{ falls} \quad \alpha > 1\,.$$

Ist X Vereinigung endlich vieler halboffener Intervalle, von denen je zwei höchstens einen Punkt gemeinsam haben, so definiert man das uneigentliche Riemann-Integral einer Funktion $f \in \mathcal{R}^{\text{loc}}\,(X)$ als Summe der uneigentlichen Riemann-Integrale über die einzelnen halboffenen Intervalle:

Definition 7.32 *Sei* $X = \bigcup_{j=1}^{k} X_j$ *mit halboffenen Intervallen* $X_1, \ldots, X_k$ *derart, dass*

$$\#(X_j \cap X_m) \leq 1 \qquad \textit{für} \quad j \neq m\,.$$

$f \in \mathcal{R}^{\text{loc}}\,(X)$ *heißt uneigentlich Riemann-integrierbar über* X $(f \in \mathcal{R}_u(X))$ *genau dann, wenn* $f_{|X_j} \in \mathcal{R}_u(X_j)$ *ist für alle* $j = 1, \ldots, k$. *Dann schreibt man*

$$\int_X f(x)\,dx := \sum_{j=1}^{k} \int_{X_j} f(x)\,dx\,.$$

Zum Beispiel ist $X := (\,0\,,\,1\,)$ Vereinigung zweier halboffener Intervalle mit nur einem gemeinsamen Punkt:

$$X = (\,0\,,\,1/\pi\,] \cup [\,1/\pi\,,\,1\,)\,,$$

aber zum Beispiel auch

$$X = (\,0\,,\,3/\pi\,] \cup [\,3/\pi\,,\,1\,)\,.$$

Man mache sich klar, dass das uneigentliche Integral über X durch 7.32 wohl definiert ist, also unabhängig ist davon, auf welche Weise X in halboffene Intervalle zerlegt wird.

Beispiel 7.33 $X = (\,0\,,\,\infty\,) = (\,0\,,\,c\,] \cup [\,c\,,\,\infty\,)$ *mit einem* $c \in \mathbb{R}^+$. *Für* f *mit*

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

ist

$$\begin{aligned}\int_R^c f(x)\,dx &= \big[-\ln x + \ln(1+x^2)\big]\,|_R^c \\ &= \ln\frac{1+c^2}{c} - \ln\frac{1+R^2}{R} \to -\infty \quad \textit{für} \quad R \searrow 0\,.\end{aligned}$$

Folglich ist $f_{|(\,0\,,\,c\,]} \notin \mathcal{R}_u((\,0\,,\,c\,])$ *und damit ist auch* $f \notin \mathcal{R}_u(X)$.

Wohl aber kann man zum Beispiel

$$\lim_{R\to\infty} \int_{1/R}^{R} f(x)\,dx = \lim_{R\to\infty} \left(\ln\frac{1+R^2}{R} - \ln\frac{1-1/R^2}{1/R}\right) = 0$$

bilden. Aber das ist bei der Frage nach der uneigentlichen Riemann–Integrierbarkeit völlig unerheblich.

Satz 7.34 *Die Menge $X \subset \mathbb{R}$ sei Vereinigung endlich vieler halboffener Intervalle, von denen je zwei höchstens einen Punkt gemeinsam haben. Ferner seien die Funktionen $f, g \in \mathcal{R}^{\text{loc}}(X)$ und $g \geq 0$.*

(i) *Ist $g \in \mathcal{R}_u(X)$ und gilt $|f| \leq g$, so ist $f \in \mathcal{R}_u(X)$*

(ii) *Ist $g \notin \mathcal{R}_u(X)$ und gilt $f \geq g$, so ist $f \notin \mathcal{R}_u(X)$*

Beweis:
Es muss nur (i) bewiesen werden und zwar für den Fall, dass X ein einziges halboffenes Intervall $(a, b]$ oder $[a, b)$ ist. Wir beschränken uns auf den zweiten Fall: $X = [a, b)$ mit $a \in \mathbb{R}$, $b > a$, $b \in \hat{\mathbb{R}}$. Zum Nachweis, dass $\lim\limits_{R \nearrow b} \int_a^R f(t)\,dt$ existiert, genügt es zu zeigen, dass zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$\left| \int_R^S f(t)\,dt \right| < \varepsilon$$

für alle $R, S \in U_\delta(b)$ mit $R < S$. Aber für alle $R < S$, $R, S \in X$ gilt

$$\left| \int_R^S f(t)\,dt \right| \leq \int_R^S |f(t)|\,dt \leq \int_R^S g(t)\,dt < \varepsilon$$

für $R, S \in U_\delta(b)$ mit geeignetem δ; denn $\lim\limits_{R \to b} \int_a^R g(t)\,dt$ existiert.

q.e.d.

Beispiel 7.35

$$\int_\pi^\infty \frac{\cos t}{t^2}\,dt$$

existiert, denn $\left|\frac{\cos t}{t^2}\right| \leq \frac{1}{t^2}$ für alle $t \in [\pi, \infty)$ und $\int_\pi^\infty \frac{1}{t^2}\,dt$ existiert nach Beispiel 7.31.

Man kann uneigentliche Integration verwenden, um die absolute Konvergenz von Reihen zu beweisen: Ist $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge, so gilt mit

$$f : [0, \infty) \longrightarrow \mathbb{R}, \ t \longmapsto a_n \quad \text{, falls} \quad t \in [n-1, n)$$

für jedes $N \in \mathbb{N}$

$$\sum_{n=1}^{N} a_n = \int_0^N f(t)\,dt \,. \tag{7.83}$$

Bezeichnet für $R \in \mathbb{R}$

$$[R] := \max\{n \in \mathbb{Z} \colon n \leq R\}$$

die sogenannte Gauß-Klammer von R, so ist für jedes $R > 0$:

$$\int_0^R f(t)\,dt = \int_0^{[R]} f(t)\,dt + (R - [R])a_{[R]+1} \;. \tag{7.84}$$

Die Formeln (7.83) und (7.84) liefern: $\sum\limits_{n=1}^{\infty} a_n$ konvergiert genau dann, wenn f uneigentlich Riemann-integrierbar ist über $[\,0\,,\,\infty\,)$. Zusammen mit Satz 7.34 erhält man das Integralkriterium für Reihen:

Satz 7.36 *Sei mit $k \in \mathbb{N}$*

$$f \;:\; [\,k\,,\,\infty\,) \;\longrightarrow\; \mathbb{R}$$

eine nicht negative monoton fallende Funktion und $(a_n)_{n\in\mathbb{N}}$ eine Folge. Es gelten:

(i) $\sum\limits_{n=1}^{\infty} a_n$ *konvergiert absolut, wenn $f \in \mathcal{R}_u([\,k\,,\,\infty\,))$ ist und die Abschätzung*

$$\bigwedge_{n \geq k} \quad |a_n| \leq f(n)$$

erfüllt ist.

(ii) $\sum\limits_{n=1}^{\infty} a_n$ *konvergiert nicht, wenn $f \notin \mathcal{R}_u([\,k\,,\,\infty\,))$ ist und die Abschätzung*

$$\bigwedge_{n \geq k} \quad a_n \geq f(n)$$

gilt.

Beispiel 7.37

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

divergiert.
Denn $f \;:\; [\,2\,,\,\infty\,) \;\longrightarrow\; \mathbb{R}\,,\;\; t \longmapsto \frac{1}{t\ln t}$ ist monoton fallend, nicht–negativ und

$$a_n := \frac{1}{n \ln n} = f(n)\;.$$

Aber für alle $R \geq 2$ *ist*

$$\int_2^R \frac{1}{t \ln t}\, dt = \int_{\ln 2}^{\ln R} \frac{1}{s}\, ds = \ln(\ln R) - \ln(\ln 2) \to \infty$$

für $R \to \infty$.

Mithin ist $f \notin \mathcal{R}_u([\,2\,,\,\infty\,))$, *und Satz* 7.36*(ii) ist anwendbar.*

Andererseits ist mit $\alpha > 1$

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^\alpha}$$

(absolut) konvergent. Denn $f \;:\; [\,2\,,\,\infty\,) \longrightarrow \mathbb{R}\,,\; t \longmapsto \frac{1}{t(\ln t)^\alpha}$ *ist monoton fallend, nicht negativ,*

$$a_n := \frac{1}{n(\ln n)^\alpha} = f(n)$$

und $f \in \mathcal{R}_u([\,2\,,\,\infty\,))$, *nämlich:*

$$\int_2^\infty f(t)\, dt = \frac{1}{\alpha - 1}(\ln 2)^{1-\alpha}\,.$$

Kapitel 8

Potenzreihen

Für die Funktionen $\exp, \sin, \cos$ hatten wir in Abschnitt 5.4 die Darstellungen

$$\exp(t) = \sum_{n=0}^{\infty} \frac{1}{n!} t^n \ , \ \cos t = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} t^{2n} \ , \ \sin t = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1}$$

für alle $t \in \mathbb{R}$ gefunden. Die Reihen, die hier auftreten, nennt man Potenzreihen. Wir haben also exp, sin und cos durch Potenzreihen dargestellt.

Es sei an Satz 5.43 erinnert: Konvergiert eine Potenzreihe

$$\sum_{n=0}^{\infty} a_n (t - t_0)^n \tag{8.1}$$

absolut für jedes $t \in U_\delta(t_0)$, $\delta > 0$, etwa gegen $f(t)$, so konvergiert die Reihe der Ableitungen

$$\sum_{n=0}^{\infty} (n+1) a_{n+1} (t - t_0)^n \tag{8.2}$$

absolut für jedes $t \in U_\delta(t_0)$ gegen die Ableitung von f. Dies kann man natürlich wieder auf die Reihe in (8.2) anwenden und erhält: $f \in C^\infty(U_\delta(t_0))$.

Umgekehrt kann man fragen, welche Funktionen aus $C^\infty(U_\delta(t_0))$ sich durch Potenzreihen darstellen lassen – eventuell sogar alle? Eine Antwort darauf werden wir mit Hilfe der Taylorformel gewinnen.

Wir wollen aber bereits jetzt feststellen: Ist f durch (8.1) in $U_\delta(t_0)$ definiert, erhält man die k-te Ableitung $f^{(k)}$ von f als

$$f^{(k)}(t) = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k} (t - t_0)^n \ , \tag{8.3}$$

wie man mit einer vollständigen Induktion verifiziert. Hieraus erhält man

$$f^{(k)}(t_0) = k! a_k \quad ,$$

also

$$a_k = \frac{1}{k!} f^{(k)}(t_0) \ . \tag{8.4}$$

Unsere Frage lautet also: Für welche C^∞–Funktionen f auf $U_\delta(t_0)$ kann man

$$\bigwedge_{t \in U_\delta(t_0)} \quad f(t) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(t_0)(t - t_0)^n \tag{8.5}$$

beweisen? Die Reihe in (8.5) nennt man die Taylorreihe von f um den Entwicklungspunkt t_0 (Brook Taylor, 1685 – 1731).

8.1 Die Taylorformel

Der Hauptsatz der Differential- und Integralrechnung liefert für eine im Intervall $I := (\, x_0 - \delta \,,\, x_0 + \delta \,)$ stetig differenzierbare Funktion f:

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t) dt \ . \tag{8.6}$$

Ist f sogar aus $C^2(I)$, so kann man das Integral partiell intergrieren. Setzt man $u(t) := (t - x)$ und $v(t) := f'(t)$, so erhält man

$$\begin{aligned} \int_{x_0}^{x} f'(t)\, dt &= \int_{x_0}^{x} u'(t) v(t)\, dt = u(t) v'(t)\Big|_{x_0}^{x} - \int_{x_0}^{x} u(t) v(t)\, dt \\ &= (x - x_0) f'(x_0) + \int_{x_0}^{x} (x - t) f''(t) dt \ . \end{aligned}$$

Eingesetzt in (8.6) liefert das

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x} (x - t) f''(t) dt \ . \tag{8.7}$$

Ist f sogar aus $C^3(I)$, liefert erneute partielle Integration

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \frac{1}{2} \int_{x_0}^{x} (x - t)^2 f'''(t) dt \ . \tag{8.8}$$

Mit vollständiger Induktion erhält man nun die Taylorformel:

Satz 8.1 *Sei I ein Intervall mit nicht leerem Inneren $\overset{\circ}{I}$ und sei $x_0 \in I$. Ist mit $n \in \mathbb{N}$ die Funktion $f \in C^{n+1}(I)$, so gilt für alle $x \in I$*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{1}{n!} \int_{x_0}^{x} (x - t)^n f^{(n+1)}(t)dt \; . \tag{8.9}$$

Beweis: Man beachte, dass (8.6), (8.7), (8.8) auch gelten, wenn I nicht offen ist und x_0 oder x Randpunkt von I ist.

Zum Beweis fehlt nur noch der Induktionsschluss: Ist (8.9) richtig für irgendein $n \in \mathbb{N}$ und ist sogar $f \in C^{n+2}(I)$, so folgt mit partieller Integration

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &\quad + \frac{1}{n!}\left(\frac{1}{n+1}(x - x_0)^{n+1} f^{(n+1)}(x_0) + \frac{1}{n+1} \int_{x_0}^{x} (x - t)^{n+1} f^{(n+2)}(t)dt \right) \\ &= \sum_{k=0}^{n+1} \frac{f^{(k)}(x_0)}{k!} + \frac{1}{(n+1)!} \int_{x_0}^{x} (x - t)^{n+1} f^{(n+2)}(t)dt \; . \end{aligned}$$

q.e.d.

Definition 8.2 *Ist im Punkt $x_0 \in D(f)$ die reelle Funktion f n–mal differenzierbar, so versteht man unter dem n–ten Taylorpolynom für f um x_0 ein Polynom $T_n := T_{n,f,x_0}$ höchstens n–ten Grades, für das und dessen Ableitungen*

$$f^{(k)}(x_0) = T_n^{(k)}(x_0) \qquad \text{für } k = 0, \ldots, n \tag{8.10}$$

gelten.

T_{n,f,x_0} und die ersten n Ableitungen von T_{n,f,x_0} stimmen also im Punkt x_0 mit f und den n Ableitungen von f überein.

Setzt man T_n an als

$$T_n(x) := \sum_{j=0}^{n} a_j (x - x_0)^j$$

mit noch zu bestimmenden Koeffizienten a_j, so liefern

$$T_n^{(k)}(x) = \sum_{j=0}^{n-k} \frac{(j+k)!}{j!} a_{j+k} (x - x_0)^j \quad , \quad k \in \{0, \ldots, n\} \quad ,$$

sowie (8.10)

$$f^{(k)}(x_0) = k! a_k \ .$$

Das n–te Taylorpolynom von f um x_0 ist also eindeutig bestimmt durch

$$\bigwedge_{x\in\mathbb{R}} \quad T_{n,f,x_0}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \ . \tag{8.11}$$

Die Taylorformel gibt also eine Darstellung des Fehlers, den man macht, wenn man den Wert von f im Punkt x durch das n–te Taylorpolynom von f um x_0 ersetzt. Diese Darstellung erlaubt, den Fehler — man nennt ihn den n–ten Taylorrest — abzuschätzen. Die Berechnung mit hinlänglicher Genauigkeit so unzugänglicher Funktionen wie der trigonometrischen Funktionen, der Exponential– oder der Logarithmusfunktion in Computern oder Taschenrechnern wird mittels der Taylorformel durchgeführt. Weitere zugänglichere Darstellungen des n–ten Taylorrests gibt folgendes Korollar: .

Korollar und Definition 8.3 *Sei I ein Intervall, $\overset{\circ}{I} \neq \emptyset$, $x_0 \in I$ und $f \in C^n(I)$ und $n \in \mathbb{N}$. Die Funktion*

$$R_n := R_{n,f,x_0} := f - T_{n,f,x_0}$$

heißt n–ter Taylorrest von f um x_0. Ist sogar $f \in C^{n+1}(I)$, so läßt der n–te Taylorrest sich darstellen auf jede der folgenden Weisen:

(i) *in integraler Form:*

$$\bigwedge_{x\in I} \quad R_n(x) = \frac{1}{n!}\int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt \quad ,$$

(ii) *nach Schlömilch (Oskar Schlömilch, 1823–1901) — ab hier bezeichne $I(x_0,x)$ das Intervall $(\min\{x_0,x\}, \max\{x_0,x\})$ — :*

$$\bigwedge_{p\in[1\,,\,n+1]} \quad \bigwedge_{x\in I} \quad \bigvee_{\xi\in I(x_0,x)} \quad R_n(x) = \frac{1}{p\cdot(n!)}\left(\frac{x-x_0}{x-\xi}\right)^p (x-\xi)^{n+1} f^{(n+1)}(\xi) \ ,$$

(iii) *nach Lagrange (Joseph Louis Lagrange, 1736–1813):*

$$\bigwedge_{x\in I} \quad \bigvee_{\xi\in I(x_0,x)} \quad R_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-x_0)^{n+1} \ ,$$

(iv) *nach Cauchy (Augustin Louis Cauchy, 1789–1857):*

$$\bigwedge_{x\in I} \quad \bigvee_{\xi\in I(x_0,x)} \quad R_n(x) = \frac{1}{n!}(x-x_0)(x-\xi)^n f^{(n+1)}(\xi)\ .$$

Beweis: Die Darstellung in integraler Form wurde in Satz 8.1 bewiesen. Die Darstellungen von Lagrange und Cauchy erhält man aus der von Schlömilch, indem man p auf $n+1$ bzw. auf 1 spezialisiert. Bleibt die Schlömilchsche Darstellung nachzuweisen: Schreibt man im Fall $x > x_0$ das Restglied $R_n(x)$ in der Form

$$R_n(x) = \frac{1}{n!}\int_{x_0}^{x}(x-t)^{p-1}(x-t)^{n+1-p}f^{(n+1)}(t)dt$$

mit $p \in [\,1\,,\ n+1\,]$, so liefert der Mittelwertsatz der Integralrechnung (Satz 7.25):

$$\begin{aligned} R_n(x) &= (x-\xi)^{n+1-p}f^{(n+1)}(\xi)\cdot\frac{1}{n!}\int_{x_0}^{x}(x-t)^{p-1}dt \\ &= \frac{1}{p\cdot(n!)}(x-x_0)^p(x-\xi)^{n+1-p}f^{(n+1)}(\xi) \end{aligned}$$

mit einem $\xi \in I(x_0, x)$.

Ist $x < x_0$, geht man analog vor, muß aber darauf achten, dass negative Zahlen nicht zur p–ten Potenz erhoben werden dürfen, wenn p nicht ganzzahlig ist. Mit der integralen Darstellung erhält man

$$\begin{aligned} R_n(x) &= \frac{(-1)^{n+1}}{n!}\int_{x}^{x_0}(t-x)^{p-1}(t-x)^{n+1-p}f^{(n+1)}(t)dt \\ &= (\xi-x)^{n+1-p}f^{(n+1)}(\xi)\cdot\frac{(-1)^{n+1}}{n!}\int_{x}^{x_0}(t-x)^{p-1}dt \\ &= \frac{1}{p\cdot(n!)}\left(\frac{x-x_0}{x-\xi}\right)^p(x-\xi)^{n+1}f^{(n+1)}(\xi)\ . \end{aligned}$$

q.e.d.

Beispiele 8.4 *Für $f = \ln$ und $x_0 = 1$ erhält man*

$$f^{(k)}(x) = (-1)^{k-1}(k-1)!x^{-k} \quad \textit{für jedes } k \in \mathbb{N} \textit{ und } x > 0,$$

wie man mit Induktion leicht zeigt.

Folglich ist $f^{(0)}(1) = 0$ *und* $f^{(k)}(1) = (-1)^{k-1}(k-1)!$ *und*

$$T_n(x) = T_{n,\ln,1}(x) = \sum_{k=1}^{n} (-1)^{k-1} \frac{1}{k} (x-1)^k \; .$$

ist das n*–te Taylorpolynom von* ln *um* 1*.*

Die Lagrangesche Restglieddarstellung liefert

$$f(x) - T_n(x) = (-1)^n \frac{1}{n+1} \xi^{-n-1} (x-1)^{n+1} \quad \textit{mit} \quad \xi \in I(x_0, x) \; .$$

Speziell für $x = \frac{3}{2}$ *erhält man*

$$\left| \ln \frac{3}{2} - T_n\left(\frac{3}{2}\right) \right| = \frac{1}{n+1} \xi^{-n-1} \left(\frac{1}{2}\right)^{n+1} \leq \frac{1}{n+1} \left(\frac{1}{2}\right)^{n+1} \qquad ,$$

denn $\xi \in (\,1\,,\,3/2\,)$*. Diese Abschätzung erlaubt die Vorhersage, dass der Wert von* T_n *an der Stelle* $3/2$ *sich von* $\ln 3/2$ *um höchstens* ε *— mit vorgegebenem* ε *— unterscheidet, wenn man den Grad* n *des Taylorpolynoms so groß wählt, dass*

$$\frac{1}{n+1} \left(\frac{1}{2}\right)^{n+1} \leq \varepsilon$$

ist. Ist etwa $\varepsilon = \frac{1}{64}$*, so unterscheidet sich* $\ln 3/2$ *von*

$$T_3\left(\frac{3}{2}\right) = \frac{1}{2} - \frac{1}{2} \cdot \left(\frac{1}{2}\right)^2 + \frac{1}{3} \cdot \left(\frac{1}{2}\right)^3 = \frac{5}{12}$$

um höchstens $\frac{1}{64}$*.*

Wir kennen sogar die Richtung der Abweichung:

$$R_3\left(\frac{3}{2}\right) = -\frac{1}{4} \xi^{-4} \left(\frac{1}{2}\right)^4 < 0 \; .$$

Also ist

$$\frac{5}{12} - \frac{1}{64} < \ln \frac{3}{2} = T_3\left(\frac{3}{2}\right) + R_3\left(\frac{3}{2}\right) < \frac{5}{12} \tag{8.12}$$

oder

$$0.4010416 < \frac{77}{192} < \ln \frac{3}{2} < \frac{80}{192} < 0.4166667 \tag{8.13}$$

Unter Verwendung des Taylorpolynoms dritter Ordnung um 1 *soll jetzt eine Annäherung an* $\ln \frac{2}{3}$ *bestimmt werden. Man beachte:* $\ln \frac{2}{3} = -\ln \frac{3}{2}$*. Wir erhalten so also auch eine zweite Annäherung an* $\ln \frac{3}{2}$*.*

Zunächst ist

$$T_3\left(\frac{2}{3}\right) = -\frac{1}{3} - \frac{1}{2}\left(-\frac{1}{3}\right)^2 + \frac{1}{3}\left(-\frac{1}{3}\right)^3 = -\frac{65}{162}\,.$$

Die Lagrangesche Restglieddarstellung liefert mit einem ξ zwischen $\frac{2}{3}$ und 1:

$$R_3\left(\frac{2}{3}\right) = -\frac{1}{4}\xi^{-4}\left(-\frac{1}{3}\right)^4 < 0\,. \tag{8.14}$$

Folglich ist $T_3(\frac{2}{3}) > \ln\frac{2}{3}$.

Ferner läßt sich der Abstand zwischen $T_3(\frac{2}{3})$ und $\ln\frac{2}{3}$ abschätzen, indem man in (8.14) das unbekannte ξ durch die Zahl $\xi_0 \in [\,\frac{2}{3}\,,\,1\,]$ ersetzt, die für $\left|-\frac{1}{4}\xi^{-4}(-\frac{1}{3})^4\right|$ den pessimistischen Wert, das ist der größte Wert, liefert:

$$\left|R_3\left(\frac{2}{3}\right)\right| < \frac{1}{4}\cdot\left(\frac{2}{3}\right)^{-4}\left(\frac{1}{3}\right)^4 = \frac{1}{64}\,.$$

Hieraus folgt:

$$-\frac{65}{162} > \ln\frac{2}{3} > -\frac{65}{162} - \frac{1}{64}\,.$$

Verwendet man die Cauchysche Restglieddarstellung

$$R_3\left(\frac{2}{3}\right) = \frac{1}{3!}\left(-\frac{1}{3}\right)\left(\frac{2}{3}-\xi\right)^3\cdot 3!(-\xi^{-4}) = -\frac{1}{3}\left(\xi-\frac{2}{3}\right)^3\xi^{-4}$$

(mit $\xi \in (\,2/3\,,\,1\,)$) und trägt man für ξ den pessimistischen Wert zwischen $\frac{2}{3}$ und 1, nämlich 1, ein, so erhält man

$$\left|R_3\left(\frac{2}{3}\right)\right| < \frac{1}{81}\,.$$

Hieraus erhält man für $\ln\frac{3}{2} = -\ln\frac{2}{3}$:

$$0.4012345 \le \frac{65}{162} < \ln\frac{3}{2} < \frac{67}{162} \le 0.4135804\,. \tag{8.15}$$

Diese Abschätzung ist schärfer als (8.13); ja, der in (8.12) gefundene Näherungswert liegt nicht einmal in dem in (8.15) beschriebenen Intervall.

Neben der Integralformel (8.9) gibt es einen weiteren Schlüssel zum Nachweis des Taylorschen Satzes, allerdings mit Lagrangescher Restglieddarstellung; das ist der verallgemeinerte Mittelwertsatz der Differentialrechnung. Wir wollen diese Methode benutzen, um den Taylorschen Satz in qualitativer Form unter etwas schwächeren Differenzierbarkeitsannahmen an f zu gewinnen.

Satz 8.5 *(verallgemeinerter Mittelwertsatz) Sei $I := [\, a \,,\, b \,]$ ein kompaktes Intervall. Sind $f, g \in C(I)$ differenzierbar in $\overset{\circ}{I}$, dann gibt es ein $\xi \in (\, a \,,\, b \,)$ mit*

$$(f(b) - f(a))g'(\xi) \quad = \quad (g(b) - g(a))f'(\xi) \; . \tag{8.16}$$

Hat g' keine Nullstelle, so erhält man aus (8.16):

$$\bigvee_{\xi \in (\, a \,,\, b \,)} \quad \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \; . \tag{8.17}$$

Beweis: Man wendet den Satz von Rolle (Lemma 5.32) auf die Funktion

$$G \;:\; I \;\longrightarrow\; \mathbb{R} \,,\;\; x \;\longmapsto\; (f\,(b) - f\,(a))\, g(x) - (g\,(b) - g\,(a))\, f(x)$$

an.

q.e.d.

Ist nun F in einer Umgebung U von $x_0 = 0$ $(n-1)$–mal differenzierbar und gelten $F^{(j)}(0) = 0$ für $j = 0, \ldots, n-1$, so erhält man aus (8.17), indem man a durch 0, b durch $x \in U$, f durch F und g durch $g(t) := t^n$ ersetzt

$$\frac{F(x)}{x^n} = \frac{F'(\xi_1)}{n\xi_1^{n-1}} = \frac{F''(\xi_2)}{n(n-1)\xi_2^{n-2}} = \ldots = \frac{F^{(n-1)}(\xi_{n-1})}{n!\xi_{n-1}}$$

mit ξ_1 zwischen 0 und $x, \ldots, \xi_{j+1}$ zwischen 0 und ξ_j. Ist nun F in 0 sogar n–mal differenzierbar mit $F^{(n)}(0) = 0$, so folgt

$$\lim_{x \to 0} \frac{F(x)}{x^n} = \lim_{\xi_{n-1} \to 0} \frac{1}{n!} \frac{F^{(n-1)}(\xi_{n-1})}{\xi_{n-1}} = \frac{1}{n!} F^{(n)}(0) = 0 \; .$$

Unter F kann man sich das n–te Taylorsche Restglied vorstellen, und indem man 0 durch x_0 ersetzt erhält man:

Satz 8.6 *Ist f n–mal differenzierbar in $x_0 \in D(f)$, so gilt*

$$\lim_{x \to x_0} \frac{1}{(x - x_0)^n} \left(f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j \right) = 0 \; .$$

Aus dem verallgemeinerten Mittelwertsatz erhält man auch die Regel von de l'Hospital (Guillaume François Antoine de l'Hospital, 1661–1704).

Satz 8.7 *Seien f, g reelle Funktionen und $x_0 \in \hat{\mathbb{R}}$ ein Häufungspunkt in $\hat{\mathbb{R}}$ eines Intervalls $I \subset D(f'/g')$. Es möge eine der beiden folgenden Voraussetzungen erfüllt sein:*

(i) $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$,

(ii) $\hat{\mathbb{R}} - \lim_{x\to x_0} f(x) \in \{-\infty, \infty\} \quad \wedge \quad \hat{\mathbb{R}} - \lim_{x\to x_0} g(x) \in \{-\infty, \infty\}$.

Existiert $\hat{\mathbb{R}} - \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$, *so existiert auch* $\hat{\mathbb{R}} - \lim_{x\to x_0} \frac{f(x)}{g(x)}$ *und*

$$\hat{\mathbb{R}} - \lim_{x\to x_0} \frac{f(x)}{g(x)} = \hat{\mathbb{R}} - \lim_{x\to x_0} \frac{f'(x)}{g'(x)} \ .$$

Der Beweis unter Voraussetzung (i) ist eine direkte Anwendung von Satz 8.5. Der Beweis unter Voraussetzung (ii) soll hier nicht durchgeführt werden.

Beispiel 8.8

$$\lim_{x\to 0} \frac{\arctan x}{\tan x} = \lim_{x\to 0} \frac{\frac{1}{1+x^2}}{1+\tan^2 x} = 1 \ .$$

Der Grenzwert ließ sich auf diese Weise wegen $\lim_{x\to 0} \arctan x = \lim_{x\to 0} \tan x = 0$ *bestimmen.*

Mit $\alpha > 0$ *ist*

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = \lim_{x\to\infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha}\frac{1}{x^\alpha} = 0 \quad .$$

Der Grenzwert ließ sich auf diese Weise wegen $\lim_{x\to\infty} \ln x = \lim_{x\to\infty} x^\alpha = \infty$ *bestimmen.*

Natürlich kennt man das n–te Taylorpolynom für exp, sin, cos etwa um 0:

$$T_{n,\exp}(x) = \sum_{j=0}^{n} \frac{1}{j!} x^j \ ,$$

$$T_{n,\sin}(x) = \sum_{j=0}^{k} \frac{(-1)^j}{(2j+1)!} x^{2j+1} \quad \text{mit} \quad 2k+1 \in \{n-1, n\} \ ,$$

$$T_{n,\cos}(x) = \sum_{j=0}^{k} \frac{(-1)^j}{(2j)!} x^{2j} \quad \text{mit} \quad 2k \in \{n-1, n\} \ ,$$

und in den Übungen gibt es weitere Anwendungsbeispiele.

Ist nun $f \in C^\infty(I)$, I ein offenes Intervall mit $x_0 \in I$, so kann man zu f das Taylorpolynom einer jeden Ordnung bestimmen und eine Potenzreihe für f aufschreiben:

Definition 8.9 *Sei f in einer Umgebung eines Punktes $x_0 \in \mathbb{R}$ beliebig oft differenzierbar. Die Reihe*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

heißt Taylorreihe für f um den Entwicklungspunkt x_0.

Es erhebt sich die Frage, für welche $x \in \mathbb{R}$ die Taylorreihe konvergiert und ob dann $f(x)$ ihr Grenzwert ist.

Tatsächlich treten folgende Situationen wirklich ein:

(i) Die Taylorreihe von f um x_0 konvergiert für jedes $x \in D(f)$ gegen $f(x)$ (wie in den Fällen von $f = \exp$, $f = \sin$ oder $f = \cos$ und $x_0 = 0$).

(ii) Die Taylorreihe von f um x_0 konvergiert für jedes x in einer Umgebung $U_\delta(x_0)$ gegen $f(x)$. Aber für $|x - x_0| > \delta$ ist sie divergent, obwohl f dort auch definiert und beliebig oft differenzierbar ist (wir werden später sehen, dass eine Taylorreihe, die für ein $\tilde{x} \neq x_0$ konvergiert, für alle x aus einer Umgebung $U_\delta(x_0)$ mit $\delta := |\tilde{x} - x_0|$ konvergiert).

(iii) Mit einer Zahl $\delta > 0$ konvergiert die Taylorreihe für jedes $x \in U_\delta(x_0)$, aber nicht gegen $f(x)$ — es sei denn $x = x_0$.

(iv) Die Taylorreihe konvergiert für kein $x \neq x_0$.

Für den Fall (iii) geben wir ein wichtiges Beispiel an:

Beispiel 8.10 *Sei*

$$f \,:\, \mathbb{R} \longrightarrow \mathbb{R}\,, \; x \longmapsto \begin{cases} \exp(-1/x^2) & \textit{für} \;\; x \neq 0 \\ 0 & \textit{für} \;\; x = 0 \end{cases}$$

Wir zeigen, dass f in der Tat eine C^∞–Funktion ist und bestimmen die Ableitungen von f in 0.

Aus dem Mittelwertsatz der Differentialrechnung (Korollar 5.34*) erhält man:*

Ist f in einer Umgebung U von $x_0 \in \mathbb{R}$ definiert, in $U \setminus \{x_0\}$ differenzierbar, in x_0 stetig und existiert

$$\lim_{x \to x_0} f'(x) =: \alpha \quad ,$$

so ist f auch in x_0 differenzierbar und $f'(x_0) = \alpha$.

Die Funktion f ist stetig in 0, denn

$$\lim_{x\to 0} \exp(-1/x^2) = \lim_{t\to\infty} \exp(-t) = 0 \ .$$

Ferner ist

$$f'(x) = 2x^{-3} \exp(-1/x^2)$$

für $x \neq 0$ und

$$\lim_{x\to 0} f'(x) = \lim_{t\to\infty} \frac{\pm 2t^{3/2}}{\exp(t)} = 0$$

(siehe Satz 5.37). Daher ist f auch in 0 differenzierbar mit $f'(0) = 0$ und f' ist stetig in 0.

Per Induktion zeigen wir nun:

Zu jedem $n \in \mathbb{N}$ gibt es ein Polynom P_n, so dass für alle $x \in \mathbb{R} \setminus \{0\}$ gilt

$$f^{(n)}(x) = P_n\left(\frac{1}{x}\right) \exp(-1/x^2) \ . \tag{8.18}$$

Dass dies für $n = 1$ richtig ist, ist soeben dargelegt worden. Gilt für irgendein $n \in \mathbb{N}$ die Formel (8.18), so erhält man für $(n+1)$ und $x \in \mathbb{R} \setminus \{0\}$

$$\begin{aligned} f^{(n+1)}(x) = \left(f^{(n)}\right)'(x) &= \left[-\frac{1}{x^2} P_n'(\frac{1}{x}) + \frac{2}{x^3} P_n\left(\frac{1}{x}\right)\right] \exp(-1/x^2) \\ &= P_{n+1}\left(\frac{1}{x}\right) \exp(-1/x^2) \end{aligned}$$

mit dem Polynom P_{n+1}, gegeben durch

$$P_{n+1}(t) = -t^2 P_n'(t) + 2t^3 P_n(t) \ .$$

Damit ist (8.18) bewiesen.

Für x gegen 0 ist nach Satz 5.37

$$\begin{aligned} \lim_{x\searrow 0} f^{(n)}(x) &= \lim_{t\to\infty} P_n(t^{1/2}) \exp(-t) = \lim_{x\nearrow 0} f^{(n)}(x) \\ &= \lim_{t\to\infty} P_n(-t^{1/2}) \exp(-t) = 0 \ . \end{aligned}$$

Daher sind alle Ableitungen von f in 0 stetig ergänzbar durch 0, folglich

$$f \in C^\infty(\mathbb{R}) \quad , \quad f^{(n)}(0) = 0 \ . \tag{8.19}$$

Taylorreihe für f um 0 ist

$$\sum_{k=0}^{\infty} 0 \cdot x^k \quad ,$$

und diese konvergiert für alle $x \in \mathbb{R}$ gegen 0.

Aber

$$f(x) > 0 \quad \textit{für alle} \quad x \neq 0 \, . \tag{8.20}$$

Für den Fall (ii) dient

Beispiel 8.11 *Die Taylorreihe für f mit $f(x) = (1+x^2)^{-1}$ um 0 ist*

$$\sum_{k=0}^{\infty} (-1)^k x^{2k} \, .$$

Diese Reihe konvergiert für jedes $x \in (\,-1\,,\,1\,)$ gegen $f(x)$ (geometrische Reihe!), aber für $|x| \geq 1$ ist sie divergent.

Man beachte, dass es unnötig ist, tatsächlich alle Ableitungen von f zu berechnen: Mit

$$a_n := \begin{cases} 0 & \textit{für ungerades} \quad n \\ (-1)^{n/2} & \textit{für gerades} \quad n \end{cases}$$

ist ja für $x \in (\,-1\,,\,1\,)$

$$f(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} = \sum_{n=0}^{\infty} a_n x^n \quad ,$$

und Formel (8.4) liefert

$$f^{(n)}(0) = n! a_n = \begin{cases} 0 & \textit{für ungerades} \quad n \\ (-1)^{n/2} n! & \textit{für gerades} \quad n \end{cases}$$

Beispiel 8.12 *Die Taylorreihe für g mit*

$$g(x) = -\frac{2x}{(1+x^2)^2}$$

ist

$$\sum_{k=1}^{\infty} (-1)^k (2k) x^{2k-1} \quad ,$$

und $g(x) = \sum\limits_{k=1}^{\infty} (-1)^k (2k) x^{2k-1}$ in $(\,-1\,,\,1\,)$. Denn $g = f'$ mit $f(x) = \frac{1}{1+x^2}$. Nach Satz 5.43 erhält man die Ableitung der in $(\,-1\,,\,1\,)$ konvergenten Potenzreihe als in $(\,-1\,,\,1\,)$ konvergente Reihe der Ableitungen.

Aber man kann Taylorreihen auch integrieren. Analog zu Satz 5.43 gilt

Satz 8.13 *Wenn mit einem $\delta > 0$ die Reihe*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

für alle $x \in (\, x_0 - \delta \,,\, x_0 + \delta \,)$ absolut konvergiert, so auch die Reihe

$$\sum_{k=1}^{\infty} \frac{a_{k-1}}{k} (x - x_0)^k \;,$$

und zwar ist

$$\int_{x_0}^{x} \left(\sum_{k=0}^{\infty} a_k (t - x_0)^k \right) dt = \sum_{k=1}^{\infty} \frac{a_{k-1}}{k} (x - x_0)^k \;.$$

Beweis: Wegen des Majorantenkriteriums konvergiert

$$\sum_{k=0}^{\infty} \frac{1}{k+1} a_k (x - x_0)^k$$

absolut für alle $x \in (\, x_0 - \delta \,,\, x_0 + \delta \,)$, denn für jedes $k \in \mathbb{N}_0$ und $x \in (\, x_0 - \delta \,,\, x_0 + \delta \,)$ ist

$$\left| \frac{1}{k+1} a_k (x - x_0)^k \right| \leq \left| a_k (x - x_0)^k \right| \;.$$

Damit ist aber

$$\sum_{k=1}^{\infty} \frac{1}{k} a_{k-1} (x - x_0)^k = (x - x_0) \sum_{k=0}^{\infty} \frac{1}{k+1} a_k (x - x_0)^k$$

absolut konvergent für alle $x \in (\, x_0 - \delta \,,\, x_0 + \delta \,)$. Den Grenzwert bezeichnen wir mit $F(x)$. F ist nach Satz 5.43 differenzierbar und

$$F'(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \quad , \quad F(x_0) = 0 \;.$$

Hieraus folgt die Behauptung.

q.e.d.

Beispiel 8.14 *Wegen Beispiels* 8.11 *ist für alle* $x \in (-1, 1)$

$$\arctan x = \int_0^x \frac{1}{1+t^2} dt = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1} x^{2k+1} . \tag{8.21}$$

Rechts steht also die Taylorreihe von arctan *um* 0, *und diese konvergiert in* $(-1, 1)$.

Beispiel 8.15 *Für* $x \in (0, 2)$ *ist*

$$\frac{1}{x} = \frac{1}{1+(x-1)} = \sum_{j=0}^{\infty} (-1)^j (x-1)^j \tag{8.22}$$

(geometrische Reihe).

Daher gilt für $x \in (0, 2)$

$$\ln x = \int_1^x \frac{1}{t} dt = \sum_{k=1}^{\infty} \frac{1}{k} (-1)^{k-1} (x-1)^k . \tag{8.23}$$

Auf den rechten Seiten von (8.22) *bzw.* (8.23) *stehen jeweils die Taylorreihen für* $\frac{1}{x}$ *bzw.* $\ln x$ *um* 1, *und diese konvergieren im Intervall* $(0, 2)$.

Eine Antwort auf die Frage, ob eine Funktion f in einer Umgebung von x_0 durch ihre Taylorreihe um x_0 dargestellt wird, kann man natürlich mit Hilfe des Restglieds finden.

Es ist ja für jedes $n \in \mathbb{N}$

$$f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!} (x-x_0)^j + R_{n,f}(x) \quad ,$$

falls $f \in C^\infty(U)$ und U eine Umgebung von x_0 ist. Also konvergiert die Taylorreihe gegen $f(x)$, wenn $R_{n,f}(x)$ für n gegen ∞ gegen 0 strebt. Verwendet man noch die integrale, die Lagrangesche oder die Cauchysche Restglieddarstellung aus Korollar 8.3, so erhält man

Satz 8.16 *Sei* U *eine Umgebung von* $x_0 \in \mathbb{R}$ *und* $f \in C^\infty(U)$. *Für ein* $x \in U$ *ist*

$$f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!} (x-x_0)^j \quad ,$$

wenn (i), (ii) oder (iii) gilt:

(i) $\lim\limits_{n\to\infty} \frac{1}{n!} \int\limits_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt = 0$

(ii) $\lim\limits_{n\to\infty} \frac{1}{(n+1)!} \left(|x-x_0|^{n+1} \sup\limits_{\theta\in(0,1)} \cdot \left| f^{(n+1)}\left((1-\theta)x + +\theta x_0\right)\right| \right) = 0.$

(iii) $\lim\limits_{n\to\infty} \frac{1}{n!} |x-x_0|^{n+1} \sup\limits_{\theta\in(0,1)} \left(\theta^n \cdot \left| f^{(n+1)}\left(1-\theta)x + \theta x_0\right)\right|\right) = 0$

Beispiel 8.17 *(Die Binomialreihe) Die Taylorreihe für*

$$f \ : \ (-1\,,\infty\,) \longrightarrow \mathbb{R}\,, \ x \longmapsto (1+x)^\alpha \quad , \quad \alpha \in \mathbb{R}$$

um 0 *nennt man die Binomialreihe. Ist* $\alpha := N \in \mathbb{N}$*, so ist* f *ein Polynom:*

$$f(x) = \sum_{n=0}^{N} \binom{N}{n} x^n$$

nach der binomischen Formel und rechts steht natürlich die in diesem Fall abbrechende Taylorreihe von f*.*

Ist $\alpha = -1$*, so hat man die geometrische Reihe als Taylorreihe, und für* $\alpha = -N$*,* $N \in \mathbb{N}$ *erhält man die Taylorreihe durch Differentiation der geometrischen Reihe.*

Wir wollen nun $\alpha \notin \mathbb{N}$ *und* $\alpha > 0$ *annehmen. Wegen*

$$\begin{aligned} f^{(1)}(x) &= \alpha(1+x)^{\alpha-1}\ , \\ f^{(2)}(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2}\ , \\ &\vdots \\ f^{(k)}(x) &= \left[\prod_{j=0}^{k-1}(\alpha-j)\right](1+x)^{\alpha-k} \end{aligned}$$

erhält man mit

$$\binom{\alpha}{k} := \frac{1}{k!}\prod_{j=0}^{k-1}(\alpha-j) \quad , \quad \binom{\alpha}{0} := 1 \tag{8.24}$$

die "Binomialreihe"

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

als Taylorreihe für f *um* 0*. Man beachte:* (8.24) *ist tatsächlich eine Verallgemeinerung der Binomialkoeffizienten auf beliebige reelle Zahlen als oberen Parameter.*

Sind $\alpha, k \in \mathbb{N}_0$ mit $k \leq \alpha$, so stimmt (8.24) *tatsächlich mit den bereits definierten Binomialkoeffizienten überein.*

Ist $|x| < 1$ so erhält man mit Cauchyscher Restglieddarstellung für das n–te Restglied

$$R_n(x) = (n+1)\begin{pmatrix} \alpha \\ n+1 \end{pmatrix} x(x-\xi)^n(1+\xi)^{\alpha-n-1} \quad ,$$

also mit $\xi := \theta x$, $0 < \theta < 1$:

$$\begin{aligned} |R_n(x)| &\leq 2^\alpha (n+1)|x| \begin{pmatrix} \alpha \\ n+1 \end{pmatrix} |x-\xi|^n (1+\xi)^{-n-1} \\ &\leq \sup_{\theta\in(0,1)} 2^\alpha |x|^{n+1}(n+1)\begin{pmatrix} \alpha \\ n+1 \end{pmatrix}(1-\theta)^n(1+\theta x)^{-n-1} \\ &= 2^\alpha |x|^{n+1}(n+1)\begin{pmatrix} \alpha \\ n+1 \end{pmatrix} \to 0 \quad \textit{für} \quad n \to \infty \quad , \end{aligned}$$

denn für positives α ist $\begin{pmatrix} \alpha \\ n+1 \end{pmatrix}$ ist eine beschränkte Folge.

Ist $\alpha \notin \mathbb{Z}$, aber negativ, so konvergiert die Binomialreihe für $x \in (-1, 1)$ ebenfalls gegen $f(x)$. Denn mit einer Konstanten c, einer natürlichen Zahl $k > \alpha$ und mit $g(x) := (1+x)^{\alpha+k}$ ist $f = cg^{(k)}$. Das Konvergenzresultat folgt aus Satz 5.43.

8.2 Potenzreihen in $\mathbb{C}$

Eine Potenzreihe

$$\sum_{j=0}^{\infty} a_j(x-x_0)^j$$

ist eine Folge von Polynomen

$$\left(\sum_{j=0}^{n} a_j(x-x_0)^j\right)_{n\in\mathbb{N}} \quad ,$$

und der Wert eines Polynoms kann auch für Argumente in den komplexen Zahlen bestimmt werden. Das Ergebnis ist dann ebenfalls eine komplexe Zahl. Um dann von Potenzreihen in $\mathbb{C}$ zu sprechen, muss man in den komplexen Zahlen einen Konvergenzbegriff einführen.

Die komplexen Zahlen sind in der linearen Algebra eingeführt worden. Zur Erinnerung: Der Körper der komplexen Zahlen besteht aus reellen Zahlenpaaren

$\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] \in \mathbb{R}^2$ zusammen mit der üblichen Vektoraddition und der Multiplikation

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1y_1 - x_2y_2 \\ x_1y_2 + x_2y_1 \end{bmatrix} . \tag{8.25}$$

Da man (8.25) auch unter Verwendung des Matrizenkalküls schreiben kann

$$\begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad ,$$

kann man die Multiplikation zweier komplexer Zahlen als eine "Drehsteckung" interpretieren: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ mit $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ zu multiplizieren, bedeutet im Fall $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, den Vektor $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ mit den Faktor $\delta := (|x_1|^2 + |x_2|^2)^{1/2}$ zu strecken und dann mit der orthogonalen Matrix (mit Determinante 1) $\begin{bmatrix} \delta^{-1}x_1 & -\delta^{-1}x_2 \\ \delta^{-1}x_2 & \delta^{-1}x_1 \end{bmatrix}$ zu drehen. Wie aus der linearen Algebra bekannt, ist $\mathbb{C}$ mit den Operationen $+$, $\times$ ein Körper.

Um den Körper $\mathbb{C}$ vom reellen Vektorraum $\mathbb{R}^2$ zu unterscheiden, schreibt man die Zahlenpaare in $\mathbb{C}$ in der Gestalt

$$x := x_1 + ix_2 \mathrel{\hat{=}} x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

Bei der Multiplikation kann man dann das Distributiv- und Assoziativgesetz verwenden und muss nur $i \cdot i = -1$ bedenken. Man schreibt dann auch wieder $\cdot$ statt $\times$ oder läßt den $\cdot$ wie in $\mathbb{R}$ meistens fort:

$$\begin{aligned}(x_1 + ix_2)(y_1 + iy_2) =& x_1y_1 + ix_1y_2 + ix_2y_1 + (-1)x_2y_2 \\ =& (x_1y_1 - x_2y_2) + i(x_1y_2 + x_2y_1)\end{aligned}$$

Ist $z = x + iy$ mit $x, y \in \mathbb{R}$, so nennt man x den Realteil und y den Imaginärteil von z:

$$x =: \operatorname{Re} z \quad , \quad y =: \operatorname{Im} z .$$

Der Körper $\mathbb{R}$ der reellen Zahlen läßt sich in $\mathbb{C}$ einbetten vermöge

$$\iota \ : \ \mathbb{R} \longrightarrow \mathbb{C} \, , \ x \longmapsto x + i \cdot 0 \quad ,$$

und man nennt ein Element aus $\mathbb{C}$ mit Imaginärteil 0 auch eine reelle Zahl.

Für $z := x + iy$, $x, y \in \mathbb{R}$, heißt

$$\overline{z} := x - iy$$

die zu z konjugiert komplexe Zahl; sie entsteht geometrisch durch Spiegelung von z an der reellen Achse $\{t + i0\colon\, t \in \mathbb{R}\}$.

Dann ist

$$\mathrm{Re}\, z = \frac{1}{2}(z + \overline{z}) \quad , \quad \mathrm{Im}\, z = \frac{1}{2i}(z - \overline{z}) = -\frac{i}{2}(z - \overline{z}) \; .$$

Ferner für $z, w \in \mathbb{C}$

$$\overline{zw} = \overline{z} \cdot \overline{w} \quad , \quad \overline{\overline{z}} = z \quad , \quad \mathrm{Re}\, z = \mathrm{Re}\, \overline{z} \quad , \quad \mathrm{Im}\, \overline{z} = -\mathrm{Im}\, z \; . \tag{8.26}$$

In der linearen Algebra sind Skalarprodukt und Euklidische Norm in $\mathbb{R}^2$ definiert worden. Mit dem kanonischen Skalarprodukt

$$\langle \begin{bmatrix} x \\ y \end{bmatrix} , \begin{bmatrix} u \\ v \end{bmatrix} \rangle = xu + yv$$

erhält man mit $z := x + iy$, $w := u + iv$

$$\langle \begin{bmatrix} x \\ y \end{bmatrix} , \begin{bmatrix} u \\ v \end{bmatrix} \rangle = \mathrm{Re}\,(z\overline{w}) = \mathrm{Re}\,(\overline{z\overline{w}}) = \mathrm{Re}\,(\overline{z}w) \; . \tag{8.27}$$

Insbesondere ist $z \cdot \overline{z} \in \mathbb{R}$. Daher

$$x^2 + y^2 = z\overline{z} \quad ,$$

und wir bezeichnen

$$|\cdot| \; : \; \mathbb{C} \longrightarrow \mathbb{R} \, , \; z \longmapsto \sqrt{z \cdot \overline{z}} = (x^2 + y^2)^{1/2} \quad \text{, falls} \quad z = x + iy$$

als Betragsfunktion in $\mathbb{C}$.

Hierfür gilt:

Satz 8.18 *Für alle $a, b \in \mathbb{C}$ gelten:*

$$|a| \geq 0 \quad , \quad |a| = 0 \Leftrightarrow a = 0 \tag{8.28}$$

$$|ab| = |a||b| \tag{8.29}$$

$$|a + b| \leq |a| + |b| \tag{8.30}$$

$$|\overline{a}| = |a| \tag{8.31}$$

Man vergleiche Satz 1.9. Im Gegensatz zur Voraussetzung dort, ist $\mathbb{C}$ kein geordneter Körper.
Beweis: Man erhält (8.28) und (8.30) direkt aus Resultaten der linearen Algebra, indem man (8.27) beachtet, man erhält (8.29) aus (8.26):

$$|ab|^2 = ab \cdot \overline{ab} = ab\overline{a}\overline{b} = (a \cdot \overline{a}) \cdot (b \cdot \overline{b}) = |a|^2|b|^2 \quad ,$$

und schließlich liefert (8.26) auch (8.31).

q.e.d.

Hat man in $\mathbb{C}$ eine Betragsfunktion erklärt, kann man auch den Abstand zweier Punkte (Synonym für Zahlen) aus $\mathbb{C}$,

$$d(z,w) := |z-w| \quad \text{für } z,w \in \mathbb{C} \tag{8.32}$$

erklären und erhält für d die Metrikeigenschaften

Satz 8.19 *Für alle $a,b,c \in \mathbb{C}$ gelten*

$$d(a,b) \geq 0 \quad , \quad d(a,b) = 0 \Leftrightarrow a = b \tag{8.33}$$

$$d(a,b) = d(b,a) \tag{8.34}$$

$$d(a,c) \leq d(a,b) + d(b,c)\ . \tag{8.35}$$

Definition 8.20 *Für $a \in \mathbb{C}$, $\delta > 0$ heißen*

$$\begin{aligned} U_\delta(a) &:= \{z \in \mathbb{C}\colon d(z,a) < \delta\} \\ K_\delta(a) &:= \{z \in \mathbb{C}\colon d(z,a) \leq \delta\} \\ S_\delta(a) &:= \{z \in \mathbb{C}\colon d(z,a) = \delta\} \end{aligned}$$

Offene bzw. abgeschlossene δ–Umgebung um a (oder offene bzw. abgeschlossene δ–Kreisscheibe um a) bzw. δ–Kreislinie um a.

Eine Teilmenge $M \subset \mathbb{C}$ heißt offen, wenn zu jedem $a \in M$ eine offene δ–Kreisscheibe um a existiert, die in M liegt:

$$\bigwedge_{a \in M} \ \bigvee_{\delta > 0} \ U_\delta(a) \subset M\ .$$

Ein Punkt $z \in \mathbb{C}$ heißt Häufungspunkt von M, wenn in jeder δ–Kreisscheibe um z ein von z verschiedener Punkt von M liegt, also wenn

$$\bigwedge_{\delta > 0} \ (U_\delta(z) \setminus \{z\}) \cap M \neq \emptyset\ .$$

Eine Teilmenge $M \subset \mathbb{C}$ heißt abgeschlossen, wenn jeder Häufungspunkt von M in M liegt.

Die Menge $U_\delta(a)$ ist offen. $K_\delta(a)$ ist abgeschlossen.

Man sieht, dass all diese Begriffe völlig analog zu den entsprechenden Begriffen in $\mathbb{R}$ definiert sind. Insbesondere ist eine Teilmenge $M \subset \mathbb{C}$ abgeschlossen, wenn $\mathbb{C} \setminus M$ offen ist.

Anders als in $\mathbb{R}$ führt man noch und nur einen uneigentlichen Häufungspunkt ∞ einer Teilmenge von $\mathbb{C}$ ein. (Es wird wohl keine Probleme machen, hier das gleiche Symbol wie in $\mathbb{R}$ zu benutzen).

Definition 8.21 *Wir nennen ∞ einen uneigentlichen Häufungspunkt der Teilmenge $M \subset \mathbb{C}$, wenn außerhalb einer jeden Kreisscheibe (um 0) Punkte aus M liegen, also wenn*

$$\bigwedge_{\delta>0} \quad \bigvee_{m\in M} \quad |m| > \frac{1}{\delta} \, .$$

$\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ *nennen wir die erweiterte komplexe Zahlenebene, und für $\delta > 0$ heißt*

$$U_\delta(\infty) := \{z \in \mathbb{C}\colon |z| > 1/\delta\} = \mathbb{C} \setminus K_{1/\delta}(0)$$

eine δ–Umgebung von ∞.

Anders als bei $\mathbb{R}$ haben wir zu $\mathbb{C}$ nur einen uneigentlichen Häufungspunkt hinzugefügt.

Die Grenzwertdefinition läßt sich aber wieder analog zu $\mathbb{R}$ durchführen:

Definition 8.22 *Sei $A \subset \mathbb{R}$ oder $A \subset \mathbb{C}$, a eigentlicher oder uneigentlicher Häufungspunkt von A und $f : A \longrightarrow \mathbb{C}$ eine Funktion von A nach $\mathbb{C}$. Wir nennen $w \in \hat{\mathbb{C}}$ einen Grenzwert in $\hat{\mathbb{C}}$ von f, wenn*

$$\bigwedge_{\varepsilon>0} \quad \bigvee_{\delta>0} \quad f\left((U_\delta(a) \setminus \{a\}) \cap A\right) \subset U_\varepsilon(w) \, .$$

Wir schreiben

$$w = \hat{\mathbb{C}} - \lim_{x\to a} f(x) \, .$$

Ist $w \neq \infty$, also $w \in \mathbb{C}$, so nennen wir f konvergent bei a, schreiben

$$w := \lim_{x\to a} f(x)$$

und nennen w den Grenzwert (—ohne das Adjektiv "uneigentlich" —) von f bei a. Ist f nicht konvergent bei a, so nennt man f divergent bei a.

Im Fall

$$\hat{\mathbb{C}} - \lim_{x \to a} f(x) = \infty$$

sagt man: f divergiert bei a gegen ∞.

Wie in $\hat{\mathbb{R}}$ zeigt man, dass Grenzwerte in $\hat{\mathbb{C}}$ eindeutig bestimmt sind. Auch gilt Satz 2.25 für $\hat{\mathbb{C}}$–Grenzwerte. Aber Vorsicht: für $b \in \mathbb{C} \setminus \{0\}$ ist $\frac{b}{0} := \infty$ in $\hat{\mathbb{C}}$ definiert, aber <u>nicht</u> $\infty + \infty$!

f konvergiert bei a genau dann gegen $w \in \mathbb{C}$, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ bei a gegen $\operatorname{Re} w \in \mathbb{R}$ und $\operatorname{Im} w \in \mathbb{R}$ konvergieren. Denn es gelten die Abschätzungen

$$\bigwedge_{z \in \mathbb{C}} \quad |\operatorname{Re} z|, |\operatorname{Im} z| \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z| \quad , \tag{8.36}$$

und damit folgt

$$\begin{aligned} |\operatorname{Re} f(x) - \operatorname{Re} w|\,, |Imf(x) - \operatorname{Im} w| &\leq |f(x) - w| \\ &\leq |\operatorname{Re} f(x) - \operatorname{Re} w| + |\operatorname{Im} f(x) - \operatorname{Im} w|\ . \end{aligned} \tag{8.37}$$

Ist also in Definition 8.20 die Menge A Teilmenge von $\mathbb{R}$, so bedeutet Konvergenz von f gegen w in $\mathbb{C}$ nichts anderes als Konvergenz sowohl von $\operatorname{Re} f$ als auch von $\operatorname{Im} f$ in $\mathbb{R}$ gegen $\operatorname{Re} w$ bzw. $\operatorname{Im} w$.

Im Fall der Divergenz gilt: f divergiert genau dann gegen ∞ bei a, wenn $|\operatorname{Re} f|$ <u>oder</u> $|\operatorname{Im} f|$ gegen ∞ divergieren.

Die angesprochenen Funktionen können natürlich auch Folgen in $\mathbb{C}$ (komplexe Folgen) sein, also Abbildungen von $\mathbb{N}$ nach $\mathbb{C}$. Man schreibt eine solche Folge dann wieder als $(a_n)_{n \in \mathbb{N}}$. Die Reihe der a_j ist die Folge $\left(\sum\limits_{j=1}^{n} a_j\right)_{n \in \mathbb{N}}$. Sie wird wieder mit $\displaystyle\sum_{j=1}^{\infty} a_j$ bezeichnet. Wie früher wird diese Bezeichnung auch für den eventuell existierenden Grenzwert benutzt.

Wie bei reellen Folgen kann man komplexe Cauchyfolgen definieren.

Satz und Definition 8.23 *Eine komplexe Zahlenfolge $(a_n)_{n \in \mathbb{N}}$ heißt Cauchyfolge oder Cauchy-konvergent, wenn die folgenden äquivalenten Bedingungen gelten*

(i) $\lim\limits_{n \to \infty} \left(\sup\limits_{k \in \mathbb{N}} d(a_{n+k}, a_n) \right) = 0$

(ii) *Es gibt eine Nullfolge $(\nu_n)_{n\in\mathbb{N}}$ (in $\mathbb{R}$), so dass*

$$d(a_{n+k}, a_n) \le \nu_n$$

ist für alle $n, k \in \mathbb{N}$.

(iii) $\bigwedge\limits_{\varepsilon>0} \bigvee\limits_{n\in\mathbb{N}} \bigwedge\limits_{k\in\mathbb{N}} d(a_{n+k}, a_n) < \varepsilon$

(iv) $\bigwedge\limits_{\varepsilon>0} \bigvee\limits_{n_\varepsilon\in\mathbb{N}} \bigwedge\limits_{n\in\mathbb{N}} \bigwedge\limits_{m\in\mathbb{N}} (n, m \ge n_\varepsilon \Rightarrow d(a_n, a_m) < \varepsilon)$.

Wegen der Abschätzung (8.37) ist eine komplexe Folge $(a_n)_{n\in\mathbb{N}}$ genau dann eine Cauchyfolge, wenn die beiden reellen Folgen $(\operatorname{Re} a_n)_{n\in\mathbb{N}}$, $(\operatorname{Im} a_n)_{n\in\mathbb{N}}$ Cauchyfolgen sind. Da wie oben bemerkt Konvergenz von $(a_n)_{n\in\mathbb{N}}$ äquivalent ist mit Konvergenz von $(\operatorname{Re} a_n)$ und $(\operatorname{Im} a_n)$, liefert der Vollständigkeitssatz 3.34:

Satz 8.24 *$\mathbb{C}$ ist vollständig, d. h. für eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ sind äquivalent*

(i) *$(a_n)_{n\in\mathbb{N}}$ ist Cauchyfolge*

(ii) *$(a_n)_{n\in\mathbb{N}}$ ist konvergent (gegen einen Grenzwert $a \in \mathbb{C}$).*

Insbesondere gilt für Reihen:

Korollar und Definition 8.25 *Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$, so ist die Reihe $\sum\limits_{j=1}^{\infty} a_j$ genau dann konvergent, wenn (i) oder (ii) gilt:*

(i) *Es gibt eine Nullfolge $(\nu_n)_{n\in\mathbb{N}}$ und $n_0 \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \ge n_0$ und $k \in \mathbb{N}_0$:*

$$\left|\sum_{j=n}^{n+k} a_j\right| \le \nu_n$$

(ii) *Zu jedem $\varepsilon > 0$ gibt es einen Index n, so dass für alle $k \in \mathbb{N}_0$:*

$$\left|\sum_{j=n}^{n+k} a_j\right| < \varepsilon$$

Die Reihe $\sum\limits_{j=1}^{\infty} a_j$ heißt absolut konvergent, wenn $\sum\limits_{j=1}^{\infty} |a_j|$ konvergent ist. Absolut konvergente Reihen in $\mathbb{C}$ sind konvergent.

Wir wenden uns nun Potenzreihen

$$\sum_{j=0}^{\infty} a_j (z - z_0)^j$$

zu. Genauer bezeichnet man eine solche Reihe als Potenzreihe um z_0.

Eine Potenzreihe um z_0 konvergiert entweder für alle $z \in \mathbb{C}$ (in $\mathbb{C}$) oder nur für $z = z_0$, oder es gibt eine offene Kreisscheibe $U_\delta(z_0)$ um z_0, in der die Reihe konvergiert, während sie außerhalb der abgeschlossenen Kreisscheibe divergiert.

Satz und Definition 8.26 *Vorgelegt sei die komplexe Potenzreihe*

$$\sum_{j=0}^{\infty} a_j (z - z_0)^j \ .$$

(i) *Konvergiert die Reihe für ein $\tilde{z} \neq z_0$, so konvergiert sie sogar absolut für alle $z \in U_\delta(z_0)$ mit $\delta := |\tilde{z} - z_0|$.*

Es tritt einer der drei folgenden Fäle ein:

(ii) *Die Reihe konvergiert absolut für alle $z \in \mathbb{C}$. In diesem Fall sagen wir: die Potenzreihe hat Konvergenzradius ∞.*

(iii) *Es gibt eine reelle Zahl $\delta > 0$, so dass die Reihe für alle $z \in U_\delta(z_0)$ absolut konvergiert, aber für jedes $z \in \mathbb{C} \setminus K_\delta(z_0)$ divergiert. In diesem Fall sagen wir, die Potenzreihe hat Konvergenzradius δ.*

(iv) *Die Reihe divergiert für jedes $z \neq z_0$. In diesem Fall nennen wir 0 den Konvergenzradius der Reihe.*

Beweis:
Zu (i): Da $\sum\limits_{j=0}^{\infty} a_j(\hat{z} - z_0)^j$ konvergiert, ist nach Korollar 8.25(i) (mit $k := 1$) die Folge $(a_n(\hat{z} - z_0)^n)_{n\in\mathbb{N}}$ eine Nullfolge. Mithin ist

$$\lim_{n\to\infty} |a_n|\delta^n = 0 \quad ,$$

und insbesondere ist $(|a_n|\delta^n)_{n\in\mathbb{N}}$ beschränkt etwa durch $\gamma > 0$.

Für $z \in U_\delta(z_0)$ ist $0 \leq |\hat{z} - z_0|^{-1}|z - z_0| =: q < 1$ und folglich

$$\bigwedge_{n\in\mathbb{N}} \quad |a_n(z - z_0)^n| = (|a_n|\delta^n)q^n \leq \gamma q^n \ .$$

Nach dem Majorantenkriterium (Satz 3.35) konvergiert $\sum\limits_{j=0}^{\infty} |a_j(z-z_0)^j|$, d. h. die vorgelegte Reihe konvergiert absolut.

Um die Alternative "entweder (ii) oder (iii) oder (iv)" zu beweisen, nehmen wir an, die vorgelegte Reihe konvergiere für wenigstens ein $z \neq z_0$. Wenn die Reihe sogar für alle $z \in \mathbb{C}$ konvergiert, so konvergiert sie nach (i) auch absolut für alle $z \in \mathbb{C}$. Nehmen wir also an, die vorgelegt Reihe konvergiere für ein $z \neq z_0$, aber nicht für alle $z \in \mathbb{C}$. Setze

$$M := \{r \in \mathbb{R}\colon \sum_{j=0}^{\infty} a_j(z-z_0)^j \quad \text{konvergiert für ein} \quad z \in S_r(z_0)\}$$

Dann ist M nach oben beschränkt, denn anderenfalls konvergierte die Reihe für jedes $z \in \mathbb{C}$ nach (i). Ferner ist

$$\delta := \sup M > 0\ .$$

Ist nun $z \in U_\delta(z_0)$, so gibt es ein $r > |z-z_0|$ mit $r \in M$. Also konvergiert $\sum\limits_{j=0}^{\infty} a_j(\hat{z}-z_0)^j$ für ein $\hat{z} \in S_r(z_0)$. Nach (i) konvergiert $\sum\limits_{j=0}^{\infty} a_j(z-z_0)^j$ absolut. Da δ das Supremum von M ist, divergiert die Reihe für $z \in \mathbb{C} \setminus K_\delta(z_0)$.

q.e.d.

Um den Konvergenzradius zu bestimmen, benötigen wir zwei Kriterien für die Konvergenz von Reihen

Satz 8.27 *(Wurzelkriterium) Vorgelegt sei eine Reihe* $\sum\limits_{n=0}^{\infty} b_n$ *in* $\mathbb{C}$.

(i) *Wenn*

$$\limsup_{n\to\infty} |b_n|^{1/n} < 1$$

ist, so konvergiert die vorgelegte Reihe absolut.

(ii) *Wenn*

$$\limsup_{n\to\infty} |b_n|^{1/n} > 1$$

ist, so divergiert die Reihe.

Beweis:

Zu (ii): Ist $\limsup\limits_{n\to\infty} |b_n|^{1/n} > 1$, so ist b_n keine Nullfolge, und die Reihe kann nicht konvergieren.
Zu (i): Sei

$$s := \limsup_{n\to\infty} |b_n|^{1/n} < 1 \; .$$

Fixiere $q \in (\, s \,,\, 1 \,)$. Dann existiert $n_0 \in \mathbb{N}$ mit

$$|b_n|^{1/n} \le q \quad \text{für} \quad n \ge n_0 \; .$$

Für $n \ge n_0$ ist dann $|b_n| \le q^n$ und $\sum\limits_{n=0}^{\infty} |b_n|$ konvergiert nach dem Majorantenkriterium 3.35.

q.e.d.

Analog zeigt man das Quotientenkriterium:

Satz 8.28 *(Quotientenkriterium) Vorgelegt sei eine Reihe* $\sum\limits_{n=0}^{\infty} b_n$ *in* $\mathbb{C}$.

(i) *Wenn*

$$\limsup_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| < 1$$

ist, so konvergiert die vorgelegte Reihe absolut.

(ii) *Wenn*

$$\liminf_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| > 1$$

ist, so divergiert die Reihe.

Achtung: In 8.27(ii) steht $\limsup$, in 8.28(ii) steht $\liminf$.
Beweis: Es gibt $q \in (\, 0 \,,\, 1 \,)$ im Fall (i) bzw. $p \in (\, 1 \,,\, \infty \,)$ im Fall (ii), sowie $N \in \mathbb{N}$ mit

$$\left| \frac{b_{n+1}}{b_n} \right| \le q \quad \text{bzw.} \quad \left| \frac{b_{n+1}}{b_n} \right| \ge p \qquad \text{für } n \ge N$$

Hieraus folgt für $n \ge N$

$$|b_{n+k}| \le q^k |b_N| \quad \text{bzw.} \quad |b_{n+k}| \ge p^k |b_N| \; .$$

Das Majorantenkriterium 3.35 liefert absolute Konvergenz im ersten und Divergenz im zweiten Fall.

q.e.d.

Nun kann man den Konvergenzradius aus den Koeffizienten einer Potenzreihe bestimmen.

Satz 8.29 *Der Konvergenzradius* δ *einer Potenzreihe* $\sum_{j=0}^{\infty} a_j(z-z_0)^j$ *ist*

$$\delta = \begin{cases} \infty & \text{, falls } \limsup_{n\to\infty} |a_n|^{1/n} = 0 \\ \frac{1}{s} & \text{, falls } \limsup_{n\to\infty} |a_n|^{1/n} = s \in \mathbb{R}^+ \\ 0 & \text{, falls } \limsup_{n\to\infty} |a_n|^{1/n} = \infty \end{cases} \tag{8.38}$$

Denn

$$\limsup_{n\to\infty} |a_n(z-z_0)^n|^{1/n} = |z-z_0| \limsup_{n\to\infty} |a_n|^{1/n} < 1 \quad ,$$

falls $z \neq z_0$ und $|z-z_0| < \delta$ und δ wie in (8.38). Ist andererseits $|z-z_0| > \delta$, so erhält man

$$\limsup_{n\to\infty} |a_n(z-z_0)^n|^{1/n} > 1$$

Beispiel 8.30 $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ *hat Konvergenzradius* ∞ *— diese Reihe konvergiert ja für jedes* $z \in \mathbb{R}$. *Daher folgt*

$$\limsup_{n\to\infty} \left(\frac{1}{n!}\right)^{1/n} = 0 \; .$$

$\sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{1}{n}(z-1)^n$ *hat Konvergenzradius* 1, *denn*

$$\limsup_{n\to\infty} \left|(-1)^n \frac{1}{n}\right|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{n^{\frac{1}{n}}} = 1 \; .$$

Definition 8.31

$$\exp \; : \; \mathbb{C} \longrightarrow \mathbb{C} \, , \; z \longmapsto \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

heißt (komplexe) Exponentialfunktion. Die Schreibweise $e^z := \exp z$ *ist üblich!*

In den Übungen ist nachgewiesen worden: Für Folgen $(b_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^+$ ist

$$\liminf_{n\to\infty} \frac{b_{n+1}}{b_n} \le \liminf_{n\to\infty} b_n^{1/n} \le \limsup_{n\to\infty} b_n^{1/n} \le \limsup_{n\to\infty} \frac{b_{n+1}}{b_n} \ .$$

Konvergiert also $\left(\frac{b_{n+1}}{b_n}\right)_{n\in\mathbb{N}}$, so ist

$$\limsup_{n\to\infty} b_n^{1/n} = \lim_{n\to\infty} b_n^{1/n} = \lim_{n\to\infty} \frac{b_{n+1}}{b_n} \ .$$

Letzterer Grenzwert ist manchmal leichter zu berechnen, als der von $(b_n^{1/n})$.

Beispiel 8.32 $\sum\limits_{n=0}^{\infty} \begin{pmatrix} \alpha \\ n \end{pmatrix} z^n$ *hat Konvergenzradius* 1*, denn*

$$\lim_{n\to\infty} \frac{\left|\begin{pmatrix} \alpha \\ n+1 \end{pmatrix}\right|}{\left|\begin{pmatrix} \alpha \\ n \end{pmatrix}\right|} = \lim_{n\to\infty} \left|\frac{\alpha - n}{n+1}\right| = 1 \ .$$

Zwischen komplexer Exponentialfunktion und den trigonometrischen Funktionen besteht eine enge Beziehung:

Satz 8.33 *(Moivre-Formel, Abraham de Moivre, 1667–1754) Für alle $t \in \mathbb{R}$ ist*

$$e^{it} = \cos t + i \sin t$$

Beweis:

$$\operatorname{Re} e^{it} = \lim_{n\to\infty} \operatorname{Re} \sum_{j=0}^{2n} \frac{(it)^j}{j!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(-1)^k t^{2k}}{(2k)!} = \cos t \ ,$$

$$\operatorname{Im} e^{it} = \lim_{n\to\infty} \operatorname{Im} \sum_{j=0}^{2n+1} \frac{(it)^j}{j!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(-1)^k t^{2k+1}}{(2k+1)!} = \sin t \ .$$

q.e.d.

Man würde gern für die komplexe Exponentialfunktion ein Additionstheorem herleiten, wie es für die reelle Exponentialfunktion gilt:

$$\bigwedge_{z,w\in\mathbb{C}} \quad e^{(z+w)} = e^z e^w \tag{8.39}$$

Dann kann man sich zum Beispiel die Additionstheoreme für sin und cos besser merken:

$$\begin{aligned}\bigwedge_{t,s\in\mathbb{R}} \quad & \cos(t+s)+i\sin(t+s) = e^{i(t+s)} = e^{it}\cdot e^{is} \\ & = (\cos t + i\sin t)(\cos s + i\sin s) \\ & = (\cos t\cos s - \sin t\sin s) + i(\sin t\cos s + \cos t\sin s)\ .\end{aligned}$$

Formel (8.39) werden wir aus dem Satz über das "Cauchyprodukt" gewinnen:

Satz 8.34 *$\sum\limits_{n=0}^{\infty} a_n$ und $\sum\limits_{n=0}^{\infty} b_n$ seien absolut konvergente Reihen in $\mathbb{C}$. Dann konvergiert das "Cauchy-Produkt" $\sum\limits_{n=0}^{\infty}\left(\sum\limits_{k=0}^{n} a_k b_{n-k}\right)$ absolut, und zwar ist*

$$\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) = \left(\sum_{n=0}^{\infty} a_n\right)\cdot\left(\sum_{n=0}^{\infty} b_n\right)\ . \tag{8.40}$$

Beweis: Es genügt, Konvergenz von $\sum\limits_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right)$ gegen die rechte Seite von (8.40) zu zeigen. Die absolute Konvergenz erhält man dann, indem man den gleichen Schluss auf die Reihen $\sum\limits_{n=0}^{\infty} |a_n|$ und $\sum\limits_{n=0}^{\infty} |b_n|$ anwendet.

Zunächst schreiben wir

$$\begin{aligned}\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) &= \lim_{N\to\infty}\left(\sum_{k=0}^{N} a_k\cdot\sum_{l=0}^{N} b_l\right) \\ &= \lim_{N\to\infty}\sum_{k=0}^{N}\sum_{l=0}^{N} a_k b_l =: p\end{aligned}$$

und veranschaulichen uns diese Summation anhand des folgenden Matrixschemas:

$$\begin{array}{cccccccc}
a_0b_0 & a_0b_1 & a_0b_2 & a_0b_3 & . & . & . & a_0b_N \\
a_1b_0 & a_1b_1 & a_1b_2 & a_1b_3 & . & . & . & a_1b_N \\
a_2b_0 & a_2b_1 & a_2b_2 & a_2b_3 & . & . & . & a_2b_N \\
a_3b_0 & a_3b_1 & a_3b_2 & a_3b_3 & . & . & . & a_3b_N \\
. & . & . & . & & & . & . \\
. & . & . & . & & . & & . \\
. & . & . & . & . & & & . \\
a_Nb_0 & a_Nb_1 & a_Nb_2 & a_Nb_3 & . & . & . & a_Nb_N
\end{array}$$

Die Doppelsumme

$$P_N := \sum_{k=0}^{N} \sum_{l=0}^{N} a_k b_l$$

erhält man, indem man sämtliche Zahlen in obiger Matrix aufsummiert. Die N–te Partialsumme

$$C_N := \sum_{n=0}^{N} \left(\sum_{k=0}^{n} a_k b_{n-k} \right)$$

des Cauchyprodukts ist die Summe der Matrixelemente oberhalb und einschließlich der Diagonalen von $a_N b_0$ bis $b_0 a_N$. $P_N - C_N$ ist dann die Summe der Matrixelemente unterhalb dieser Diagonalen. Mit $[N/2] := \max\{K \in N_0\colon K \leq N/2\}$ gilt:

$$|P_N - C_N| = \left| \sum_{n=N+1}^{2N} \left(\sum_{k=n-N}^{N} a_k b_{n-k} \right) \right| \leq \sum_{n=N+1}^{2N} \sum_{k=n-N}^{N} |a_k||b_{n-k}|$$
$$\leq \left(\sum_{n=1}^{N} |a_n| \right) \left(\sum_{n=1}^{N} |b_n| \right) - \left(\sum_{n=1}^{[N/2]} |a_n| \right) \left(\sum_{n=1}^{[N/2]} |b_n| \right) \to 0$$

für $N \to \infty$ wegen der absoluten Konvergenz der beiden vorgelegten Reihen.

q.e.d.

Auf die Exponentialfunktion angewendet, liefert Satz 8.34 für $z, w \in \mathbb{C}$

$$e^{z+w} = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{z^k}{k!} \frac{w^{n-k}}{(n-k)!} \right)$$
$$= \left(\sum_{n=0}^{\infty} \frac{z^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{w^n}{n!} \right) = e^z \cdot e^w \, .$$

Formel (8.39) ist also in der Tat richtig, und wir können folgern

$$\bigwedge_{x,y \in \mathbb{R}} \quad e^{(x+iy)} = e^x (\cos y + i \sin y) \, , \tag{8.41}$$

$$\bigwedge_{z \in \mathbb{C}} \quad e^{\overline{z}} = \overline{(e^z)} \, , \tag{8.42}$$

$$\bigwedge_{z \in \mathbb{C}} e^z = e^{z+2\pi i} \, , \tag{8.43}$$

die komplexe Exponentialfunktion ist also "$2\pi i$–periodisch".

In vielen Anwendungen ist es üblich, die Lage eines Punktes der Ebene in "Polarkoordinaten" anzugeben:

Satz 8.35 *Zu jedem $x \in \mathbb{R}^2 \setminus \{0\}$ gibt es genau eine Zahl $r > 0$ und genau eine Zahl $\varphi \in (\,-\pi\,,\,\pi\,]$ so, dass*

$$x = \begin{bmatrix} r\cos\varphi \\ r\sin\varphi \end{bmatrix} .$$

r, φ heißen die Polarkoordinaten des Punktes x.

Man erhält für $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$:

$$r = |x| = (x_1^2 + x_2^2)^{1/2}$$

sowie

$$\varphi = \begin{cases} \arcsin \frac{x_2}{r} & ,\text{ falls } \ x_1 \geq 0 \\ \pi - \arcsin \frac{x_2}{r} & ,\text{ falls } \ x_1 < 0 \wedge x_2 \geq 0 \\ -\pi - \arcsin \frac{x_2}{r} & ,\text{ falls } \ x_1 < 0 \wedge x_2 < 0 \end{cases}$$

In $\mathbb{C}$ läßt sich dies schreiben als

Korollar 8.36 *Zu jedem $z \in \mathbb{C} \setminus \{0\}$ existiert genau ein "Radius" $r > 0$ sowie ein "Winkel" $\varphi \in (\,-\pi\,,\,\pi\,]$ mit der Eigenschaft*

$$z = re^{i\varphi}$$

(r, φ) heißen Polarkoordinaten für z. Ist $w \in \mathbb{C} \setminus \{0\}$ mit Polarkoordinaten s, ψ, so ist

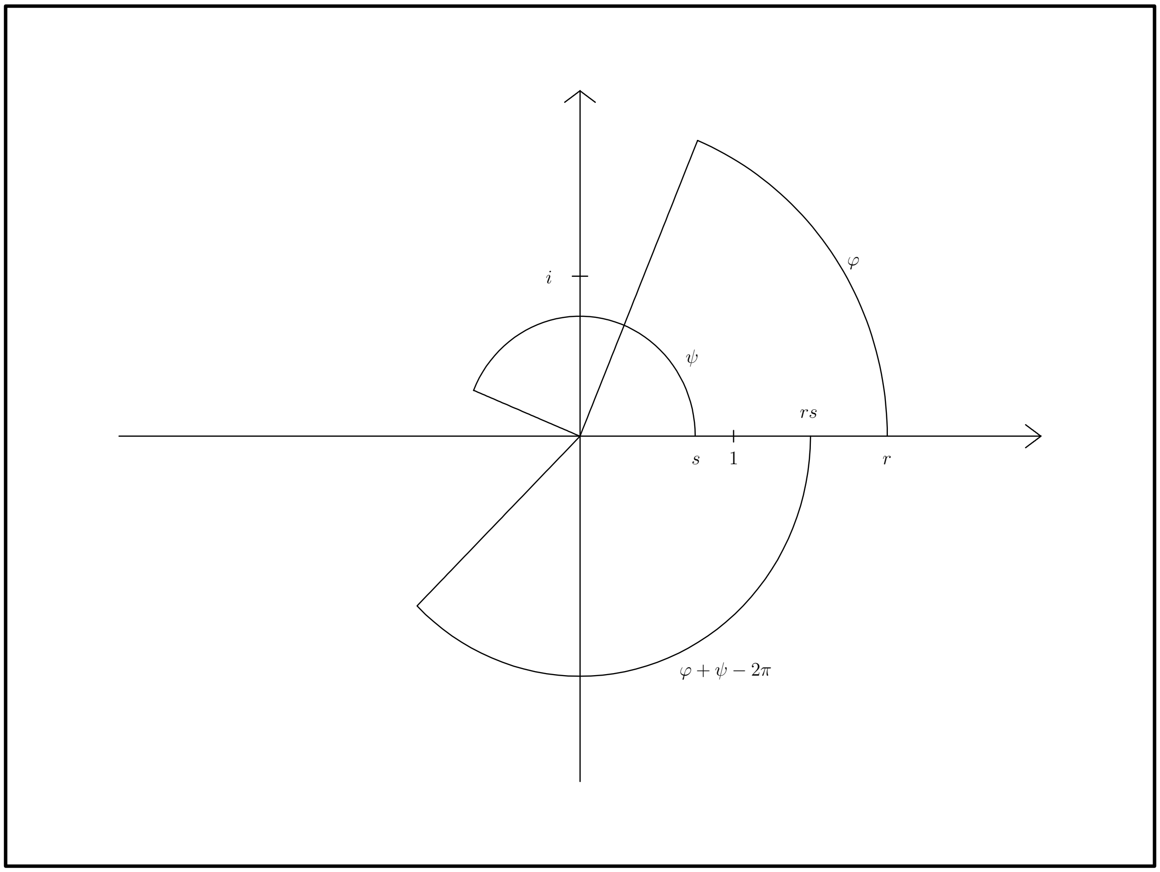
$$zw = rse^{i(\varphi+\psi)} \quad , \quad \frac{z}{w} = \frac{r}{s} e^{i(\varphi-\psi)} .$$

Die Polarkoordinaten von zw bzw. $\frac{z}{w}$ sind also $(\,rs\,,\,\theta\,)$ bzw. $(\,\frac{r}{s}\,,\,\tau\,)$, wobei die Winkel $\theta, \tau \in (\,-\pi\,,\,\pi\,]$ durch die Beziehungen

$$\theta - (\varphi + \psi) \in \{-2\pi, 0, 2\pi\} \quad , \quad \tau - (\varphi - \psi) \in \{-2\pi, 0, 2\pi\}$$

definiert sind.

Erläuterung der Multiplikation in $\mathbb{C}$:



Aus der Darstellung der Multiplikation mittels Polarkoordinaten erkennt man:

Korollar 8.37 *Zu jedem $a \in \mathbb{C} \setminus \{0\}$ und $n \in \mathbb{N}$ gibt es genau n komplexe Zahlen $z^{(0)}, \ldots, z^{(n-1)}$, die die Gleichung*

$$z^n = a$$

lösen.

Ist $a = 1$, so nennt man diese n–Zahlen "die n–ten Einheitswurzeln $w_0^{(n)}, \ldots, w_{n-1}^{(n)}$.

Diese sind

$$w_j^{(n)} = e^{\left(2\pi \frac{j}{n}\right) \cdot i}$$

mit den Polarkoordinaten $1, \varphi_j^{(n)}$, wobei

$$\varphi_j^{(n)} = \begin{cases} 2\pi j/n & \text{, falls} \quad 2\pi j/n \le \pi \\ 2\pi j/n - 2\pi & \text{, falls} \quad 2\pi j/n > \pi \end{cases} \quad .$$

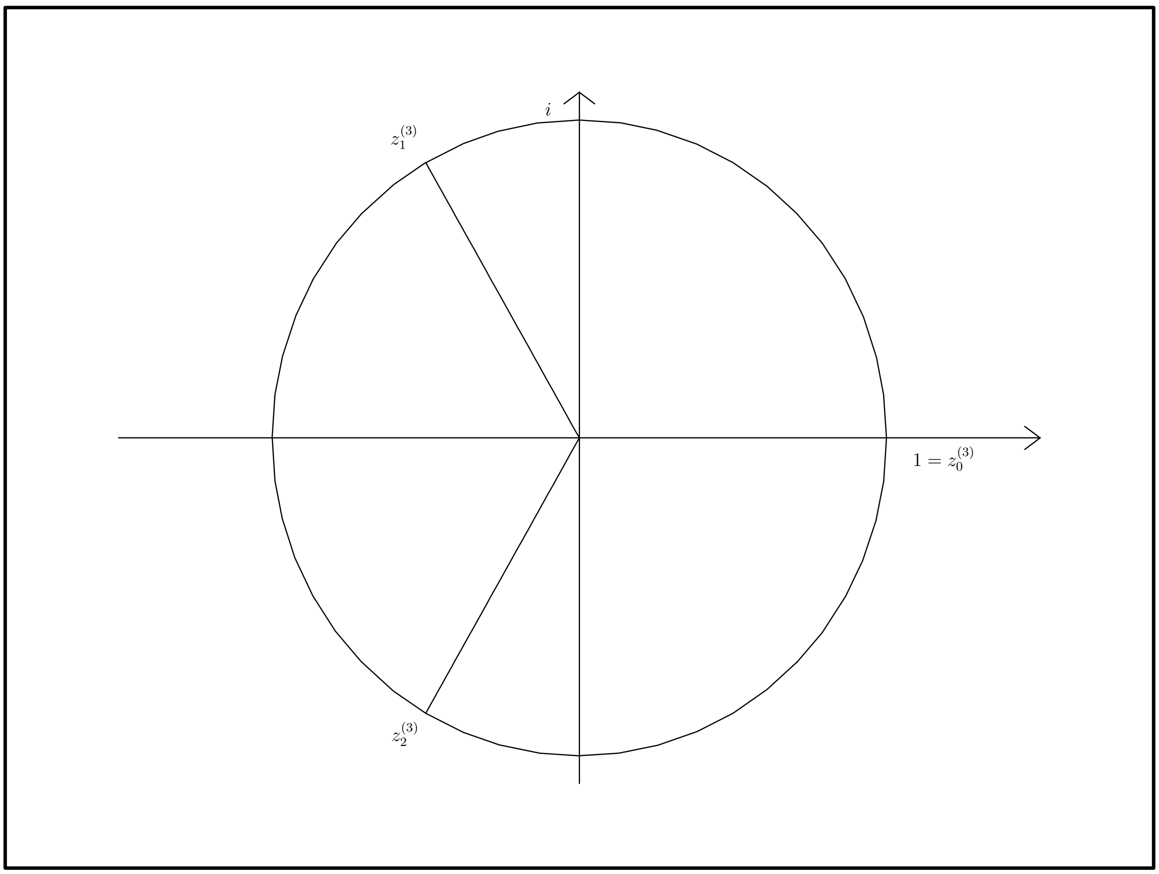
Ist mit $r > 0$ und $\psi \in (-\pi, \pi]$

$$a = re^{i\psi} \quad ,$$

so erhält man $z^{(0)}, \ldots, z^{(n-1)}$ durch

$$z^{(0)} := r^{1/n} e^{i\psi/n} \quad , \quad z^{(j)} = z^{(0)} \cdot w_j^{(n)} \quad , \quad j = 1, \ldots, n-1 \; .$$

Folgende Abbildung zeigt die dritten Einheitswurzeln:



Kapitel 9

Strukturen der Analysis

Wir haben gesehen, dass man Begriffe wie Konvergenz, offene und abgeschlossene Mengen (und anderes: Stetigkeit, Kompaktheit) nicht nur in $\mathbb{R}$, sondern auch in $\mathbb{C}$ einführen kann. Allerdings wurde dabei die Körperstruktur von $\mathbb{C}$ nicht benötigt, sondern lediglich die Tatsache, dass man zwischen zwei Punkten aus $\mathbb{C}$ ein Abstand definiert ist mit dem in Satz 8.19 angegebenen grundlegenden Eigenschaften. Für die Behandlung einer großen Klasse mathematischer Fragestellungen ist es wesentlich, den Abstandsbegriff in möglichst großer Allgemeinheit zu definieren, zum Beispiel zwischen Punkten der $\mathbb{R}^N$, $\mathbb{C}^N$, zwischen stetigen oder integrierbaren Funktionen u. v. a. Mengen, zwischen deren Elementen ein Abstand definiert ist, nennt man "metrische Räume".

Spezielle metrische Räume lassen sich aus normierten Räumen gewinnen. Das sind Vektorräume, für deren Elemente eine Länge, "Norm" genannt, definiert ist mit Eigenschaften, die denen des Betrags ähneln. Während man für Funktionen zwischen metrischen Räumen den Begriff der Stetigkeit definieren kann, kann auf Funktionen zwischen (gewissen) normierten Räumen der Differenzierbarkeitsbegriff verallgemeinert werden.

Diese Verallgemeinerung ist Ziel dieses Kapitels.

9.1 Metrische Räume

Definition 9.1 *Unter einem metrischen Raum versteht man eine nicht leere Menge X zusammen mit einer Abbildung*

$$d : X \times X \longrightarrow \mathbb{R}$$

so, dass gelten:

$$\bigwedge_{x,y\in X} d(\,x\,,\,y\,) = 0 \Leftrightarrow x = y\;, \tag{9.1}$$

$$\bigwedge_{x,y\in X} d(\,x\,,\,y\,) = d(\,y\,,\,x\,)\;, \tag{9.2}$$

$$\bigwedge_{x,y,z\in X} d(\,x\,,\,z\,) \le d(\,x\,,\,y\,) + d(\,y\,,\,z\,)\;. \tag{9.3}$$

Die Funktion d nennt man eine Metrik oder Abstandsfunktion auf X; für den metrischen Raum schreibt man $(X\,,\,d)$.

Wendet man (9.3) mit $z := x$ an, so erhält man

$$\bigwedge_{x,y\in X} d(\,x\,,\,x\,) \le d(\,x\,,\,y\,) + d(\,y\,,\,x\,)$$

und zusammen mit (9.1) und (9.2)

$$\bigwedge_{x,y\in X} d(\,x\,,\,y\,) \ge 0\;. \tag{9.4}$$

Die Eigenschaften (9.1) – (9.4) sind genau das, was man erwartet, wenn man sich ”Abstand” als ”Länge” des ”kürzesten Wegs zwischen zwei Punkten” vorstellt. Der kürzeste Weg zwischen zwei verschiedenen Punkten ist nicht negativ und genau dann 0, wenn beide Punkte zusammenfallen (9.1), (9.4). Die Länge des kürzesten Wegs zwischen zwei Punkten hängt nicht davon ab, ob man hin oder zurück geht (9.2). Der direkte Weg ist nie länger als ein Umweg (9.3). Die Eigenschaft (9.1) zusammen mit (9.4) nennt man ”Positivität”, (9.2) ist die ”Symmetrie” und (9.3) die ”Dreiecksungleichung”.

Beispiele 9.2 *(für metrische Räume)*

$(\,\mathbb{R}\,,\,d\,)$ *mit* $d(\,x\,,\,y\,) := |x-y|$ *für* $x, y \in \mathbb{R}$*;*

$(\,\mathbb{C}\,,\,d\,)$ *mit* $d(\,z\,,\,w\,) := |z-w|$ *für* $z, w \in \mathbb{C}$*;*

$(\,\mathbb{R}^N\,,\,d_2\,)$ *mit* $d_2(\,x\,,\,y\,) := \left(\sum_{n=1}^{N} |x_n - y_n|^2\right)^{1/2}$ *für* $x, y \in \mathbb{R}^N$*;*

$(\,\mathbb{R}^N\,,\,d_p\,)$ *mit* $d_p(\,x\,,\,y\,) := \left(\sum_{n=1}^{N} |x_n - y_n|^p\right)^{1/p}$ *für* $x, y \in \mathbb{R}^N$ *mit* $p \ge 1$.

$(\,\mathbb{R}^N\,,\,d_\infty\,)$ *mit* $d_\infty(\,x\,,\,y\,) := \max\limits_{n=1,\ldots,N} \;|x_n - y_n|$ *für* $x, y \in \mathbb{R}^N$.
Analog: $(\,\mathbb{C}^N\,,\,d_p\,)$ *für* $p \in [\,1\,,\,\infty\,]$.

$(C(I), d_\infty)$ *mit* $d_\infty(f, g) := \sup_{t \in I} |f(t) - g(t)|$ *für* $f, g \in C(I)$, I*: ein kompaktes Intervall.*

$(C(I), d_2)$ *mit* $d_2(f, g) := \left(\int_I |f - g|^2 (t) dt\right)^{1/2}$ *für* $f, g \in C(I)$, I*: ein kompaktes Intervall.*

(F, d) : *F sei die Menge aller französischen Orte mit Bahnanschluss, und d(x,y) bezeichnet die Eisenbahnstrecke von x nach y: Dies ist die direkte Strecke falls ein Ort auf der Strecke vom anderen Ort nach Paris liegt oder die Summe der Strecken des einen Orts nach Paris und des anderen Orts nach Paris.*

In der Tat ist etwa $(\mathbb{R}^N, d_p)$ mit $p \in (1, \infty)$ ein metrischer Raum. Die kritische Eigenschaft ist die Dreiecksungleichung (**??**). Die folgt aber mit der Minikowski-Ungleichung (6.23).

Die Metriken bis auf die letzte in Beispiel 9.2 sind aus Normen gewonnen. Der Normbegriff wird in der nächsten Sektion genau eingeführt. Eine Metrik, die natürlich nicht aus einer Norm gewonnen werden kann, läßt sich auf jeder beliebigen nicht leeren Menge X einführen:

$$d : X \times X \longrightarrow \mathbb{R}, \ (x, y) \longmapsto \begin{cases} 1 & \text{, falls } \ x \neq y \\ 0 & \text{, falls } \ x = y \end{cases}$$

ist eine Metrik auf X, die "diskrete Metrik".

Weitere Beispiele für metrische Räume erhält man aus

Satz 9.3 *Ist* (X, d) *ein metrischer Raum und* $M \subset X$ *eine nicht–leere Teilmenge von* X*, so ist auch* $(M, d|_{M \times M})$ *ein metrischer Raum.*

Definition 9.4 *Sei* (X, d) *ein metrischer Raum,* $z \in X$ *und* $\delta > 0$*. Dann heißen*

$$U_\delta(z) := \{x \in X\colon d(x, z) < \delta\}$$

$$K_\delta(z) := \{x \in X\colon d(x, z) \leq \delta\}$$

offene bzw. abgeschlossene δ*–Umgebung (oder* δ*–Kugel) um* z*.*

Ferner nennen wir

$$S_\delta(z) := \{x \in X\colon d(x, z) = \delta\}$$

die δ*–Sphäre um* z*.*

Also ist $K_\delta(z) = U_\delta(z) \cup S_\delta(z)$.

Beachten Sie: Es ist möglich, dass $S_\delta(z) = \emptyset$ ist, z. B. im Fall der diskreten Metrik und $\delta = \frac{1}{2}$. Ist $S_\delta(z) = \emptyset$, so ist $K_\delta(z) = U_\delta(z)$. Die Mengen $U_\delta(z)$ und $K_\delta(z)$ enthalten stets den Punkt z. Im Fall der diskreten Metrik und $\delta = \frac{7}{8}$ enthalten sie nur den Punkt z. Im Fall der diskreten Metrik ist

$$S_1(z) = X \setminus \{z\} \ , \ K_1(z) = X \ , \ U_1(z) = \{z\} \ .$$

Nach Einführung der Kugeln kann man jetzt die Begriffe "offene Menge" , "Häufungspunkt" definieren. Diese Definitionen sind völlig analog zu den entsprechenden Definitionen aus dem ersten Semester, z. B. 4.14

Definition 9.5 *Sei $(X \, , \, d)$ metrischer Raum und $M \subset X$. M heißt offen, wenn zu jedem $m \in M$ ein $\delta > 0$ existiert mit $U_\delta(m) \subset M$:*

$$\bigwedge_{m \in M} \ \bigvee_{\delta > 0} \ U_\delta(m) \subset M \ .$$

Ein Punkt $\hat{x} \in X$ heißt Häufungspunkt von M, wenn in jeder Umgebung $U_\delta(\hat{x})$ wenigstens ein von $\hat{x}$ verschiedener Punkt aus M liegt:

$$\bigwedge_{\delta > 0} \ (U_\delta(\hat{x}) \setminus \{\hat{x}\}) \cap M \neq \emptyset$$

M heißt abgeschlossen, wenn jeder Häufungspunkt von M in M liegt.

Zum Beispiel ist $M := (-\sqrt{2} \, , \, \sqrt{2}\,) \cap \mathbb{Q}$ sowohl offene als auch abgeschlossene Teilmenge des metrischen Raums $(\mathbb{Q} \, , \, d)$ mit $d(x \, , \, y) = |x - y|$. Ist $X := (0 \, , \, 1) \subset \mathbb{R}$, $d(x \, , \, y) := |x - y|$, so ist $(0 \, , \, 1)$ sowohl offene als auch abgeschlossene Teilmenge von $(X \, , \, d)$.

Man beachte: Für jedes $\delta > 0$, $x \in X$ ist $U_\delta(x)$ offen, $K_\delta(x)$ abgeschlossen. Ist nämlich $z \in U_\delta(x)$ und $\rho := \delta - d(x \, , \, z) > 0$, so gilt für jedes $y \in U_\rho(z)$

$$d(x \, , \, y) \leq d(x \, , \, z) + d(z \, , \, y) < d(x \, , \, z) + \rho = \delta \ ;$$

also ist $U_\rho(z) \subset U_\delta(x)$. Ist z ein Häufungspunkt von $K_\delta(x)$, so gibt es zu jedem $\varepsilon > 0$ ein $y \in K_\delta(x)$ mit $y \in U_\varepsilon(z)$. Also ist

$$d(x \, , \, z) \leq d(x \, , \, y) + d(y \, , \, z) \leq \delta + \varepsilon \ .$$

Da $d(x \, , \, z) \leq \delta + \varepsilon$ für alle $\varepsilon > 0$ gilt, ist $d(x \, , \, z) \leq \delta$, also $z \in K_\delta(x)$.

Analog zu den Sätzen 4.14, 4.15 erhält man

Satz 9.6 *Sei* (X, d) *ein metrischer Raum.*

(i) $M \subset X$ *ist genau dann offen, wenn* $X \setminus M$ *abgeschlossen ist.*

(ii) $\emptyset$ *und* X *sind offen,* $\emptyset$ *und* X *sind abgeschlossen.*

(iii) *Ist* $\mathbb{U}$ *eine Familie offener,* $\mathbb{A}$ *eine Familie abgeschlossener Mengen, so sind*

$$V := \bigcup_{U \in \mathbb{U}} U$$

offen und

$$C := \bigcap_{A \in \mathbb{A}} A$$

abgeschlossen.

(iv) *Sind* U_1, U_2 *offen,* A_1, A_2 *abgeschlossen, so sind* $U_1 \cup U_2$ *offen und* $A_1 \cap A_2$ *abgeschlossen.*

Aufgrund des Satzes 9.6 kann man einführen

Satz und Definition 9.7 *Ist* $M \subset X$, (X, d) *metrischer Raum, so heißen*

$$\overset{\circ}{M} := \bigcup_{U \in \mathbb{U}(M)} U \quad \text{bzw.} \quad \overline{M} := \bigcap_{A \in \mathbb{A}(M)} A$$

offener Kern bzw. Abschluss von M. *Hier sind die Familien* $\mathbb{U}(M)$ *und* $\mathbb{A}(M)$ *definiert durch*

$$\mathbb{U}(M) := \{U \subset X\colon U \quad \text{offen} \quad \wedge \quad U \subset M\}\,,$$

$$\mathbb{A}(M) := \{A \subset X\colon A \quad \text{abgeschlossen} \quad \wedge \quad A \supset M\}\,.$$

$\overset{\circ}{M}$ *ist die größte offene Menge, die in* M *enthalten ist;* $\overline{M}$ *ist die kleinste* M *umfassende abgeschlossene Menge. Insbesondere ist*

$$\overline{M} = \{x \in X\colon x \quad \text{ist Häufungspunkt von} \quad M\} \cup M\,. \tag{9.5}$$

Beweis: Tatsächlich ist $\overset{\circ}{M}$ größte offene von M umfasste Teilmenge von X: $\overset{\circ}{M}$ ist nämlich nach Satz 9.6 offen und nach Definition in M enthalten; ist schließlich U eine offene Teilmenge von M, so ist U eine der Mengen in $\mathbb{U}(M)$ und damit Teilmenge von $\overset{\circ}{M}$. Entsprechend ist und $\overline{M}$ die kleinste M enthaltende abgeschlossene

Teilmenge von X. Die Familie $\mathbb{A}(M)$ ist nämlich nicht leer: $X \in \mathbb{A}(M)$, $\overline{M}$ enthält M und ist nach Satz 9.6 abgeschlossen. Ist schließlich A abgeschlossene, M enthaltende Menge, so ist A eine der Mengen aus $\mathbb{A}(M)$, und damit ist $\overline{M} \subset A$.

Bezeichnet nun $\tilde{M}$ die Menge auf der rechten Seite von Gleichung (9.5), so ist sicher $\tilde{M} \subset \overline{M}$, denn jeder Häufungspunkt von M muss zu $\overline{M}$ gehören. $\tilde{M}$ ist aber abgeschlossen. Ist $\hat{x}$ Häufungspunkt von $\tilde{M}$, o.B.d.A. $\hat{x} \notin M$, und ist $\varepsilon > 0$ vorgelegt, so gibt es in $U_{\varepsilon/2}(\hat{x})$ einen Punkt $\tilde{x} \neq \hat{x}$ von $\tilde{M}$. In $U_{\varepsilon/2}(\tilde{x})$ gibt es aber einen Punkt $x \in M$: dieser kann $\tilde{x}$ selbst sein oder ein von $\tilde{x}$ verschiedener Punkt, falls $\tilde{x}$ Häufungspunkt von M ist. Dann aber ist

$$x \in U_{\varepsilon/2}(\tilde{x}) \subset U_\varepsilon(\hat{x}) \ .$$

q.e.d.

Man beachte: In vielen Fällen ist $\overline{U_\delta(x)} = K_\delta(x)$, aber nicht in allen Fällen (diskrete Metrik, $X = \mathbb{Q}$ mit üblichem Abstand).

Wir sind nun in der Lage, Konvergenz zu definieren und starten — anders als in der Analysis I — mit der Konvergenz von Folgen:

Definition 9.8 *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(\, X \,,\, d \,)$ heißt konvergent gegen einen Punkt $x \in X$, wenn*

$$\lim_{n\to\infty} d(\, x_n \,,\, x \,) = 0$$

ist. Wir schreiben dann

$$\lim_{n\to\infty} x_n = x$$

In der Tat ist der Grenzwert einer Folge $(x_n)_{n\in\mathbb{N}}$ in X eindeutig bestimmt: Sind nämlich sowohl $x \in X$ als auch $\tilde{x} \in X$ Grenzwerte von (x_n), so ist für alle $n \in \mathbb{N}$:

$$d(\, x \,,\, \tilde{x} \,) \leq d(\, x \,,\, x_n \,) + d(\, x_n \,,\, \tilde{x} \,) \ .$$

Da die rechte Seite gegen 0 strebt, folgt $d(\, x \,,\, \tilde{x} \,) = 0$, also $x = \tilde{x}$.

Wie in der Analysis I können auch Cauchyfolgen definiert werden.

Definition 9.9 *Eine Folge $(x_n)_{n\in\mathbb{N}}$ im metrischen Raum $(\, X \,,\, d \,)$ heißt Cauchyfolge (oder Cauchy–konvergent), wenn eine (reelle) Nullfolge $(\nu_n)_{n\in\mathbb{N}}$ existiert mit*

$$\bigwedge_{n,k\in\mathbb{N}} \quad d(\, x_n \,,\, x_{n+k} \,) \leq \nu_n$$

Satz 9.10 *Jede im metrischen Raum* (X , d) *konvergente Folge* $(x_n)_{n\in\mathbb{N}}$ *ist Cauchyfolge.*

Aber nicht umgekehrt!

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ gegen x in (X , d) konvergente Folge und setze

$$\nu_n := \sup_{k\in\mathbb{N}_\circ} d(x_{n+k} , x) .$$

Dann ist $(\nu_n)_{n\in\mathbb{N}}$ monoton fallend,

$$\lim_{n\to\infty} \nu_n = \limsup_{n\to\infty} d(x_n , x) = 0 \quad ,$$

und für alle $n \in \mathbb{N}$, $k \in \mathbb{N}$ gilt

$$d(x_n , x_{n+k}) \leq d(x_n , x) + d(x_{n+k} , x) \leq 2\nu_n \to 0 .$$

q.e.d.

Beispiele 9.11 *Eine Folge* $(z^{(n)})_{n\in\mathbb{N}}$ *in* $(\mathbb{C}^N , d_2)$ *mit der Metrik*

$$d_2(u , v) := \left(\sum_{j=1}^{N} |u_j - v_j|^2 \right)^{1/2}$$

konvergiert genau dann gegen $z \in \mathbb{C}^N$, *wenn die* N *Folgen* $(z_1^{(n)})_{n\in\mathbb{N}}, \ldots , (z_N^{(n)})_{n\in\mathbb{N}}$ *in* $\mathbb{C}$ *jeweils (in* $\mathbb{C}$*) gegen* $z_1, \ldots , z_N$ *konvergieren:*

$$\bigwedge_{j=1,\ldots,N} \quad \lim z_j^{(n)} = z_j$$

Auch ist $(z^{(n)})$ *genau dann Cauchyfolge, wenn* $(z_1^{(n)}), \ldots , (z_N^{(n)})$ *Cauchyfolgen in* $\mathbb{C}$ *sind.*

Ist I *ein kompaktes Intervall und* $(f_n)_{n\in\mathbb{N}}$ *eine Folge in* $(C(I) , d_\infty)$, *so bedeutet "Konvergenz in* $(C(I) , d_\infty)$ *gegen* $f \in C(I)$*":*

$$\lim_{n\to\infty} \left(\sup_{t\in I} |f_n(t) - f(t)| \right) = 0 \tag{9.6}$$

Äquivalent hierzu sind

$$\bigwedge_{\varepsilon>0} \quad \bigvee_{N\in\mathbb{N}} \quad \bigwedge_{n\geq N} \quad \sup_{t\in I} |f_n(t) - f(t)| < \varepsilon \tag{9.7}$$

$$\bigwedge_{\varepsilon>0} \bigvee_{N\in\mathbb{N}} \bigwedge_{n\geq N} \bigwedge_{t\in I} |f_n(t) - f(t)| < \varepsilon \tag{9.8}$$

$$\bigvee_{\textit{Nullfolge}\,(\nu_n)_{n\in\mathbb{N}}} \bigwedge_{t\in I} \bigwedge_{n\in\mathbb{N}} |f_n(t) - f(t)| \leq \nu_n \tag{9.9}$$

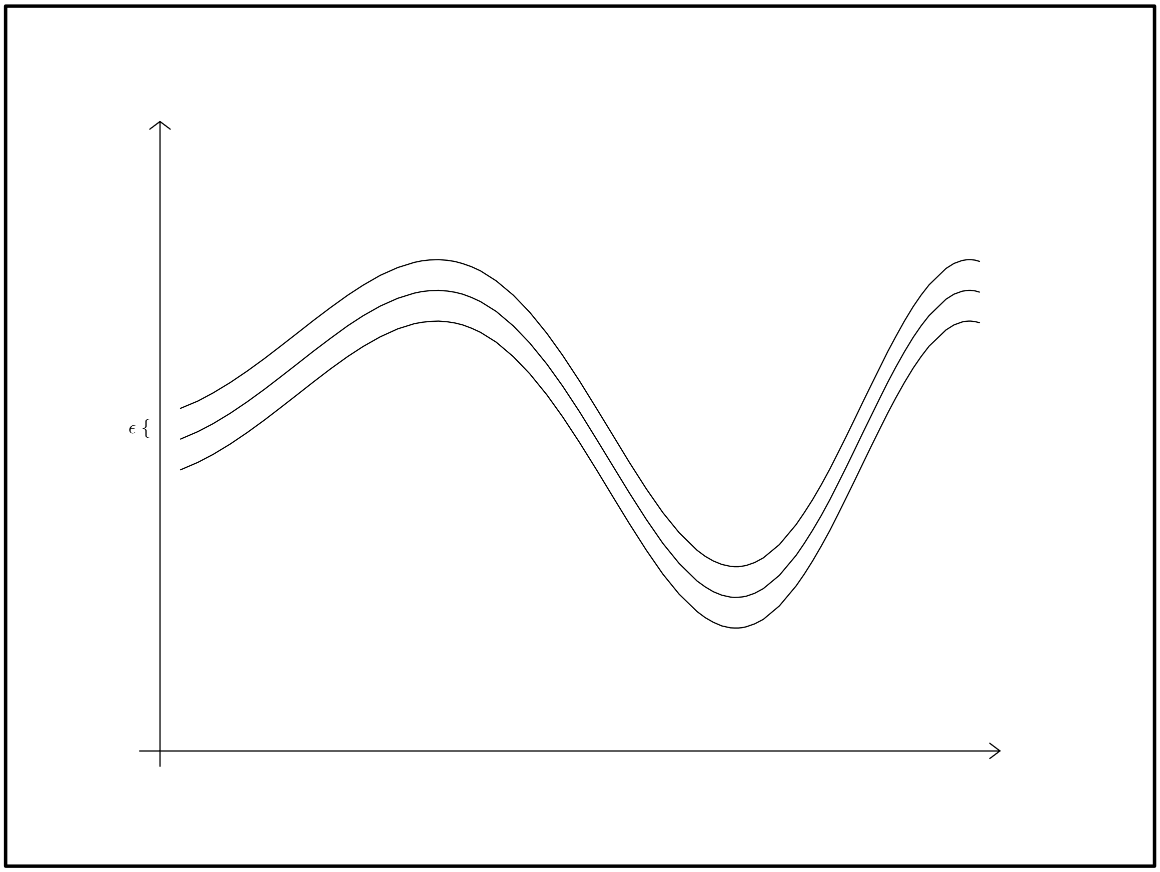
Gemäß (9.8) *darf bei vorgelegtem* $\varepsilon > 0$ *der Index* N*, ab dem* $|f_n(t) - f(t)| < \varepsilon$ *ist, nicht von* t *abhängen. Gilt* (9.8) *für eine Folge von Funktionen auf* I*, so spricht man von "gleichmäßiger Konvergenz bezüglich* t*". Zu unterscheiden ist hiervon die "punktweise Konvergenz"; hier rückt der Allquantor über* t *nach vorn:* $(f_n)_{n\in\mathbb{N}}$ *konvergiert genau dann punktweise gegen* f*, wenn*

$$\bigwedge_{t\in I} \bigwedge_{\varepsilon>0} \bigvee_{N\in\mathbb{N}} \bigwedge_{n\geq N} |f_n(t) - f(t)| < \varepsilon \,. \tag{9.10}$$

Natürlich impliziert (9.8) *die Bedingung* (9.10)*: punktweise Konvergenz folgt aus gleichmäßiger Konvergenz. Aber der umgekehrte Schluss ist falsch.*

Will man sich 9.8 *veranschaulichen, so zeichne man um den Graphen der Grenzfunktion* f *einen* ε*–Schlauch:*

$$\Sigma_\varepsilon(f) := \{(t\,,\,s) \in \mathbb{R}^2 \colon\, t \in I \;\wedge\; f(t) - \varepsilon < s < f(t) + \varepsilon\}$$



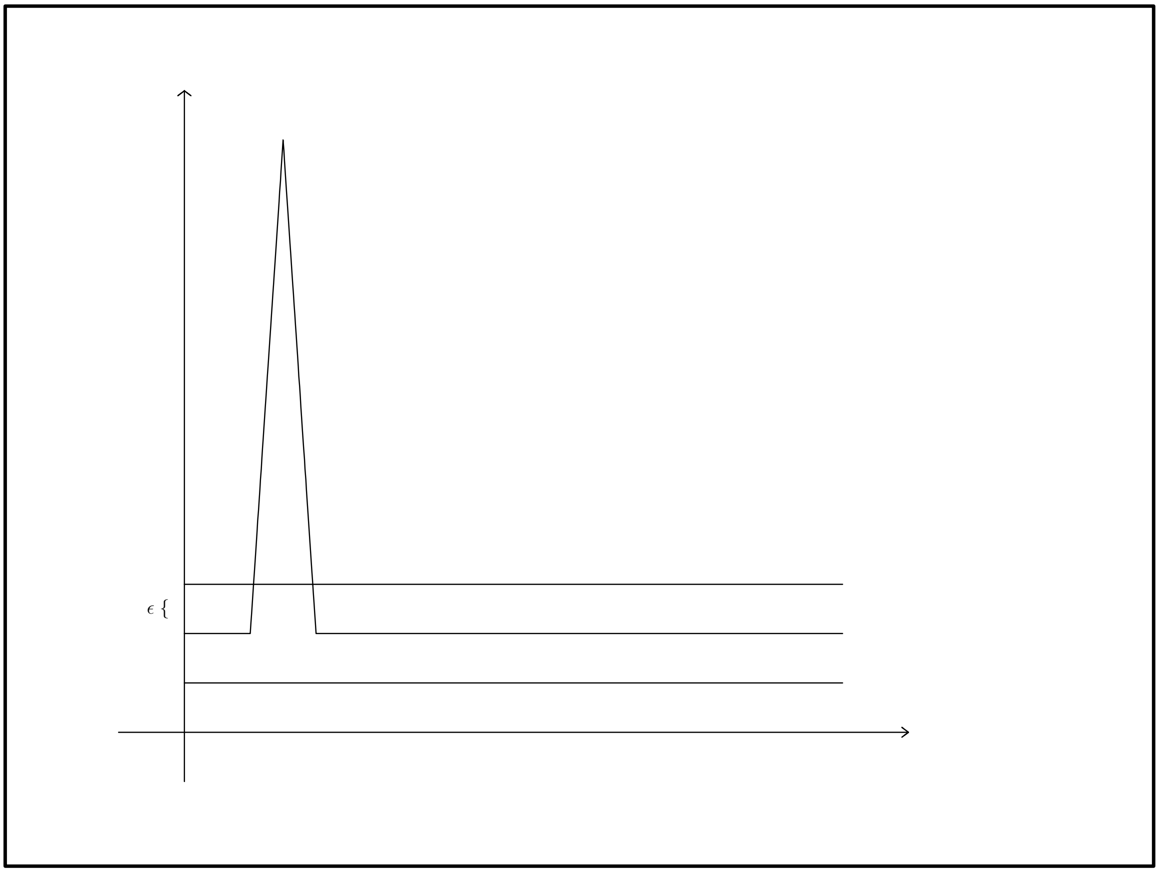
Ein ε–Schlauch

Von einem gewissen Index N an müssen die Graphen der Funktionen (f_n) in diesem Schlauch verlaufen.

Ein konkretes Beispiel: Seien $I := [\,0\,,\,1\,]$, $\alpha \in \mathbb{R}$ vorgegeben und für $n \in \mathbb{N}$, $t \in I$:

$$f_n(t) := \alpha + \max\{0; 1 - |4nt - 3|\}\ .$$

f_n ist also identisch gleich α in $[\,0\,,\,\frac{1}{2n}\,] \cup [\,\frac{1}{n}\,,\,1\,]$, und zwischen $\frac{1}{2n}$ und $\frac{1}{n}$ besteht der Graph aus den zwei gleichen Schenkeln des gleichschenkligen Dreiecks mit Höhe 1 über einer Strecke der Länge $\frac{1}{2n}$ auf der durch $y = \alpha$ gegebenen Geraden als Basis.



Punktweise Konvergenz impliziert nicht gleichmäßige Konvergenz

Ist $t > 0$, *so gibt es ein* $N \in \mathbb{N}$ *mit* $f_n(t) = 0$ *für alle* $n \geq N$. *Ist* $t = 0$, *so ist* $f_n(t) = 0$ *für alle* $n \in \mathbb{N}$. *Daher konvergiert* $(f_n)_{n\in\mathbb{N}}$ *punktweise gegen die Nullfunktion* $f \; : \; I \longrightarrow \mathbb{R}\,, \; t \longmapsto 0$. *Aber* $(f_n)_{n\in\mathbb{N}}$ *konvergiert nicht gleichmäßig gegen* f, *denn der Graph von* f_n *hält sich etwa für* $\varepsilon = 1/2$ *für kein* $n \in \mathbb{N}$ *im* ε*–Schlauch um* f *auf.*

Definition 9.12 *Ein metrischer Raum* $(\, X \,,\, d\,)$ *heißt vollständig, wenn jede Cauchyfolge* $(x_n)_{n\in\mathbb{N}}$ *in* $(\, X \,,\, d\,)$ *konvergiert.*

Beispiele 9.13 *Beispiele für vollständige metrische Räume sind* $(\,\mathbb{R}\,,\, d\,)$ *mit der üblichen Metrik* $d(\, x \,,\, y\,) := |x - y|$, *ferner* $(\,\mathbb{C}^N \,,\, d_2\,)$, *denn hier sind Konvergenz bzw. Cauchy-Konvergenz äquivalent zur Konvergenz bzw. Cauchykonvergenz der einzelnen "Komponentenfolgen".*

Auch $(\,\hat{\mathbb{R}}\,,\, d\,)$ *mit* $d(\, s \,,\, t\,) := |\arctan s - \arctan t|$, *wobei* $\arctan \pm\infty := \pm\pi/2$, *ist vollständiger metrischer Raum, und die im ersten Semester eingeführte* $\hat{\mathbb{R}}$*–Konvergenz entspricht der Konvergenz im metrischen Raum* $(\hat{\mathbb{R}}\,,\,d)$.

Wir werden später zeigen, dass auch $(C(I)\,,\,d_\infty)$ *vollständig ist.*

Weitere Beispiele erhält man aus:

Satz 9.14 *Sei* $(X\,,\,d)$ *metrischer Raum und* $Y \subset X$ *nicht leer.*

Dann gelten:

(i) *Ist* $(Y\,,\,d_{|Y\times Y})$ *vollständiger metrischer Raum, so ist* Y *abgeschlossene Teilmenge von* X

(ii) *Ist* Y *abgeschlossene Teilmenge von* X *und ist* $(X\,,\,d)$ *vollständig, so ist* $(Y\,,\,d_{|Y\times Y})$ *vollständig.*

Beachten Sie: (ii) ist <u>nicht</u> einfach die Umkehrung von (i)!

Beweis:
Zu (i): Ist y ein Häufungspunkt von Y, so gibt es eine Folge $(y_n)_{n\in\mathbb{N}}$ in Y, die in $(X\,,\,d)$ gegen y konvergiert: wähle nämlich zu $1/n (n \in \mathbb{N})$ jeweils $y_n \in Y$ mit $d(y_n\,,\,y) < 1/n$. Da $(y_n)_{n\in\mathbb{N}}$ in $(X\,,\,d)$ konvergiert, ist $(y_n)_{n\in\mathbb{N}}$ Cauchyfolge (Satz 9.10). Da (y_n) Folge in Y ist und die Metrik in $(Y\,,\,d_{|Y\times Y})$ genau die aus $(X\,,\,d)$ ist, ist (y_n) auch Cauchyfolge im metrischen Raum $(Y\,,\,d_{|Y\times Y})$. Als solche besitzt sie – $(Y\,,\,d_{|Y\times Y})$ ist vollständig – einen Grenzwert $\hat{y}$ in Y. Dieser muss aber zugleich Grenzwert in X sein. Also ist $\hat{y} = y$, d. h. $y \in Y$. Y ist also abgeschlossen.

Zu (ii): Ist $(y_n)_{n\in\mathbb{N}}$ Cauchyfolge in $(Y\,,\,d_{|Y\times Y})$, so auch in $(X\,,\,d)$, und da X vollständig ist, besitzt sie einen Grenzwert $x \in X$. Läge x nicht in Y, wäre x Häufungspunkt von Y; aber Y ist abgeschlossen nach Voraussetzung. Damit liegt x in jedem Fall in Y, und die Cauchyfolge $(y_n)_{n\in\mathbb{N}}$ besitzt Grenzwert in Y. Y ist also vollständig.

q.e.d.

Wenn man die reellen Zahlen konstruktiv aus den rationalen Zahlen gewinnen will, benutzt man eine Methode die nun in der abstrakteren Situation der metrischen Räume vorgeführt werden soll.

Definition 9.15 *Sei* $(X\,,\,d)$ *ein metrischer Raum. Ein metrischer Raum* $(\tilde{X}\,,\,\tilde{d})$ *heißt "Vervollständigung von* $(X\,,\,d)$*", wenn gelten*

(i) *$(\tilde{X}\,,\,\tilde{d}\,)$ ist vollständig.*

(ii) *Es gibt eine isometrische Funktion*

$$\varphi\,:\,X\,\longrightarrow\,\tilde{X}\quad ,$$

— das ist eine Funktion mit der Eigenschaft

$$\bigwedge_{x,y\in X}\quad \tilde{d}(\,\varphi(x)\,,\,\varphi(y)\,)=d(\,x\,,\,y\,)\quad — ,$$

so dass $\tilde{X}$ der Abschluß von $\varphi(X)$ ist:

$$\overline{\varphi(X)}=\tilde{X}\ .$$

Die Funktion φ nennt man die Einbettung von X in $\tilde{X}$ und $\varphi(X)$ die Realisierung von X in $\tilde{X}$.

Sind S, T Teilmengen eines metrischen Raumes $(\,M\,,\,d\,)$ und ist $S\subset T\subset\overline{S}$, so sagt man: S ist dicht in T. Laut Definition 9.15 liegt die Realisierung von X also dicht in der Vervollständigung von X.

Satz 9.16 *(Vervollständigungssatz)*

(i) *Zu jedem metrischen Raum $(\,X\,,\,d\,)$ gibt es eine Vervollständigung $(\,\tilde{X}\,,\,\tilde{d}\,)$.*

(ii) *Sind $(\,\tilde{X}\,,\,\tilde{d}\,)$ und $(\,\hat{X}\,,\,\hat{d}\,)$ Vervollständigungen von $(\,X\,,\,d\,)$, so gibt es eine bijektive isometrische Funktion $\Phi\,:\,\tilde{X}\,\longrightarrow\,\hat{X}$, welche die Realisierungen von X in $\tilde{X}$ bzw. $\hat{X}$ aufeinander abbildet.*

Beweis:
Zu i): Es wird die übliche Konstruktion vorgestellt. Zunächst sei Z die Menge aller Cauchyfolgen in X. Auf dieser Menge Z wird eine Äquivalenzrelation $\sim$ eingeführt. Zwei Cauchyfolgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ heißen äquivalent – in Zeichen $(a_n)\sim(b_n)_{n\in\mathbb{N}}$, wenn

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

ist.

Dies ist in der Tat eine Äquivalenzrelation; denn für alle Cauchyfolgen $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(z_n)_{n\in\mathbb{N}}\in Z$ gelten

$$(x_n)_{n\in\mathbb{N}}\sim(x_n)_{n\in\mathbb{N}}\ ,\tag{9.11}$$

denn $d(x_n, x_n) = 0 \to 0$ für $n \to \infty$;

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \Rightarrow (y_n)_{n\in\mathbb{N}} \sim (x_n)_{n\in\mathbb{N}} , \tag{9.12}$$

denn $d(y_n, x_n) = d(x_n, y_n) \to 0$ wegen $(x_n) \sim (y_n)$;

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \wedge (y_n)_{n\in\mathbb{N}} \sim (z_n)_{n\in\mathbb{N}} \Rightarrow (x_n)_{n\in\mathbb{N}} \sim (z_n)_{n\in\mathbb{N}} \quad , \tag{9.13}$$

denn

$$d(x_n, z_n) \le d(x_n, y_n) + d(y_n, z_n) \to 0 .$$

Durch die Äquivalenzrelation $\sim$ wird Z in Äquivalenzklassen eingeteilt. Sei M eine nicht-leere Teilmenge von Z, deren sämtliche Elemente jeweils paarweise äquivalent sind und die in dem Sinn maximal ist, dass $Z\backslash M$ kein Element enthält, das zu einem (und damit wegen (9.13) zu jedem) Element von Z äquivalent ist. So eine Menge M nennt man eine Äquivalenzklasse. Jedes Element von Z liegt in genau einer Äquivalentklasse und wird auch Repräsentant der Äquivalenzklasse genannt. Die Menge Z ist dann disjunkte Vereinigung aller Äquivalenzklassen. Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, so bezeichnet

$$[(x_n)_{n\in\mathbb{N}}] := \{(y_n)_{n\in\mathbb{N}} \in Z\colon (y_n)_{n\in\mathbb{N}} \sim (x_n)_{n\in\mathbb{N}}\} \tag{9.14}$$

die Äquivalenzklasse, die $(x_n)_{n\in\mathbb{N}}$ als Repräsentanten besitzt. $[(x_n)_{n\in\mathbb{N}}]$ ist also die von $(x_n)_{n\in\mathbb{N}}$ repräsentierte Äquivalenzklasse.

Beachte: Ist $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$, so ist $[(x_n)_{n\in\mathbb{N}}] = [(y_n)_{n\in\mathbb{N}}]$.

Dass all dies richtig ist, folgt aus (9.11), (9.12) und (9.13).

Wir definieren nun

$$\tilde{X} := Z/\sim := \{[(x_n)_{n\in\mathbb{N}}]\colon (x_n)_{n\in\mathbb{N}} \in Z\} .$$

Die in $\tilde{X}$ zusammengefassten Objekte sind also Mengen (genauer: Äquivalenzklassen) von Cauchfolgen in X.

In $\tilde{X}$ wird eine Metrik eingeführt:

$$\tilde{d}([(x_n)_{n\in\mathbb{N}}], [(y_n)_{n\in\mathbb{N}}]) := \lim_{n\to\infty} d(x_n, y_n) \tag{9.15}$$

Aus der Dreiecksungleichung erhält man leicht die in jedem metrischen Raum (X, d) gültige Ungleichung

$$\bigwedge_{a,b,c,d\in X} |d(a, c) - d(b, d)| \le d(a, b) + d(c, d) . \tag{9.16}$$

Hieraus erkennt man, dass der Grenzwert auf der rechten Seite von (9.15) existiert; die Folge $(d(x_n\,,\,y_n))_{n\in\mathbb{N}}$ ist nämlich Cauchyfolge in $\mathbb{R}$ nach (9.16) und damit konvergent:

$$|d(x_{n+k}\,,\,y_{n+k}) - d(x_n\,,\,y_n)| \le d(x_{n+k}\,,\,x_n) + d(y_{n+k}\,,\,y_n)\;.$$

Auch ist die Festlegung (9.15) unabhängig von der Wahl des Repräsentanten; ist nämlich $(x_n) \sim (a_n)$,$(y_n) \sim (b_n)$, so gilt

$$|d(a_n\,,\,b_n) - d(x_n\,,\,y_n)| \le d(a_n\,,\,x_n) + d(b_n\,,\,y_n) \to 0 \quad \text{für} \quad n\to\infty\;.$$

Auch ist $\tilde{d}$ eine Metrik: Sind $[(x_n)]$, $[(y_n)]$, $[(z_n)] \in \tilde{X}$, so gelten

$$\begin{aligned}
\tilde{d}(\,[(x_n)]\,,\,[(y_n)]\,) = 0 &\Leftrightarrow \lim_{n\to\infty} d(x_n\,,\,y_n) = 0 \Leftrightarrow (x_n) \sim (y_n) \Leftrightarrow [(x_n)] = [(y_n)],\\
\tilde{d}(\,[(x_n)]\,,\,[(y_n)]\,) &= \lim_{n\to\infty} d(x_n\,,\,y_n) = \lim_{n\to\infty} d(y_n\,,\,x_n) = \tilde{d}(\,[(y_n)]\,,\,[(x_n)]\,),\\
\tilde{d}(\,[(x_n)]\,,\,[(z_n)]\,) &= \lim_{n\to\infty} d(x_n\,,\,z_n) \le \lim_{n\to\infty} (d(x_n\,,\,y_n) + d(y_n\,,\,z_n))\\
&= \tilde{d}(\,[(x_n)]\,,\,[(y_n)]\,) + \tilde{d}(\,[(y_n)]\,,\,[(z_n)]\,)\;.
\end{aligned}$$

Sind nun $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen in X mit den Grenzwerten x bzw. y, so sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ bekanntlich Cauchyfolgen und

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \Leftrightarrow \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n\;.$$

Der Raum

$$\tilde{X}_0 = \{[(x_n)_{n\in\mathbb{N}}]\colon \lim_{n\to\infty} x_n \;\text{ existiert}\}$$

ist mithin metrischer Unterraum von $\tilde{X}$ und es gibt eine isometrische Abbildung

$$\varphi \;:\; X \longrightarrow \tilde{X}\,,\;\; x \longmapsto [(x)_{n\in\mathbb{N}}]$$

mit Wertebereich $\tilde{X}_0$.

φ bildet also x auf die Äquivalenzklasse aller gegen x konvergenten Folgen ab, und ein Repräsentant dieser Klasse ist die konstante Folge, deren jedes Glied x ist. Mit

$$\psi \;:\; \tilde{X}_0 \longrightarrow X\,,\;\; [(x_n)_{n\in\mathbb{N}}] \longmapsto \lim_{n\to\infty} x_n$$

sind dann $\psi\circ\varphi$ bzw. $\varphi\circ\psi$ die Identität auf X bzw. auf $\tilde{X}_0$. In der Tat ist φ isometrisch, denn für $x, y \in X$ ist

$$\tilde{d}(\,\varphi(x)\,,\,\varphi(y)\,) = \lim_{n\to\infty} d(x\,,\,y) = d(x\,,\,y)\;.$$

Wir zeigen, dass $\varphi(X) = \tilde{X}_0$ dicht liegt in $\tilde{X}$. Ist $(x_n)_{n\in\mathbb{N}}$ Repräsentant einer Äquivalenzklasse $\tilde{x} \in \tilde{X}$, also

$$[(x_n)_{n\in\mathbb{N}}] = \tilde{x} \quad ,$$

so gilt mit $\tilde{x}^m := \varphi(x_m) = [(x_m)_{n\in\mathbb{N}}]$

$$\tilde{d}(\,\tilde{x}^m\,,\,\tilde{x}\,) = \lim_{n\to\infty} d(\,x_m\,,\,x_n\,) = \lim_{k\to\infty} d(\,x_m\,,\,x_{m+k}\,) \le \sup_{k\in\mathbb{N}} d(\,x_m\,,\,x_{m+k}\,) \to 0$$

für $m \to \infty$; denn $(x_n)_{n\in\mathbb{N}}$ ist Cauchyfolge. Also ist $\tilde{X}_0$ in der Tat dicht in $\tilde{X}$.

Bleibt die Vollständigkeit von $\tilde{X}$ zu zeigen. Hierzu sei $(\tilde{x}^n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\tilde{X}$ (jedes $\tilde{x}^n$ ist also selbst eine Äquivalenzklasse von Cauchyfolgen in X). Sei $(\nu_n)_{n\in\mathbb{N}}$ eine Nullfolge (in $\mathbb{R}$) mit

$$\tilde{d}(\,\tilde{x}^n\,,\,\tilde{x}^{n+k}\,) \le \nu_n$$

für alle $n, k \in \mathbb{N}$. Da $\tilde{X}_0$ dicht ist in $\tilde{X}$ gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in X$ mit

$$\tilde{d}(\,\tilde{x}^n\,,\,\varphi(x_n)\,) \le \frac{1}{n}\ .$$

Die Folge $(x_n)_{n\in\mathbb{N}}$ ist dann eine Cauchyfolge in X; denn für $n, k \in \mathbb{N}$ ist

$$\begin{aligned} d(\,x_{n+k}\,,\,x_n\,) &= \tilde{d}(\,\varphi(x_{n+k})\,,\,\varphi(x_n)\,) \\ &\le \tilde{d}(\,\varphi(x_{n+k})\,,\,\tilde{x}_{n+k}\,) + d(\,\tilde{x}_{n+k}\,,\,\tilde{x}_n\,) + d(\,\tilde{x}_n\,,\,\varphi(x_n)\,) \\ &\le \frac{1}{n+k} + \nu_n + \frac{1}{n} \le \nu_n + \frac{2}{n} \to 0 \quad (\text{für} \quad n \to \infty)\ . \end{aligned} \tag{9.17}$$

Mit

$$\tilde{x} := [(x_n)_{n\in\mathbb{N}}] \in \tilde{X}$$

gilt dann

$$\begin{aligned} \tilde{d}(\,\tilde{x}^n\,,\,\tilde{x}\,) &\le \tilde{d}(\,\tilde{x}^n\,,\,\varphi(x_n)\,) + \tilde{d}(\,\varphi(x_n)\,,\,\tilde{x}\,) \\ &\le \frac{1}{n} + \lim_{m\to\infty} d(\,x_n\,,\,x_m\,) \le \frac{3}{n} + \nu_n \to 0 \quad (\text{für} \quad n \to \infty)\ . \end{aligned}$$

Die letzte Ungleichung folgt mit (9.17).

Zu (ii): Seien φ bzw. ψ die Einbettungen von X in $\tilde{X}$ bzw. $\hat{X}$ und setze für $\tilde{y} \in \varphi(X) =: \tilde{X}_0$

$$\Phi(\tilde{y}) := \psi(\varphi^{-1}(\tilde{y})) \in \hat{X}_0 := \psi(X) \subset \hat{X}\ .$$

Dann gilt für $\tilde{y}, \tilde{z} \in \tilde{X}_0$:

$$\hat{d}(\,\Phi(\tilde{y})\,,\,\Phi(\tilde{z})\,) = d(\,\varphi^{-1}(\tilde{y})\,,\,\varphi^{-1}(\tilde{z})\,) = \tilde{d}(\,\varphi(\varphi^{-1}(\tilde{y})\,,\,\varphi(\varphi^{-1}(\tilde{z}))\,) = \tilde{d}(\,\tilde{y}\,,\,\tilde{z}\,)$$

Sei nun $\tilde{x} \in \tilde{X} \setminus \tilde{X}_0$. Dann existiert eine Folge $(\tilde{y}_n)_{n\in\mathbb{N}}$ in $\tilde{X}_0$, die in $\tilde{X}$ gegen $\tilde{x}$ konvergiert. Diese ist insbesondere Cauchyfolge und damit ist $(\Phi(\tilde{y}_n))_{n\in\mathbb{N}}$ Cauchyfolge in $\hat{X}$. Da $\hat{X}$ vollständig ist, ist $(\Phi(\tilde{y}_n))_{n\in\mathbb{N}}$ in $\hat{X}$ konvergent, etwa gegen $\hat{x}$.

Man würde nun gern $\Phi(\tilde{x}) := \hat{x}$ definieren. Hierzu muss man zeigen, dass $\hat{x}$ unabhängig ist von der Wahl der speziellen $\tilde{x}$ approximierenden Folge $(\tilde{y}_n)_{n\in\mathbb{N}}$. Ist $(\tilde{z}_n)_{n\in\mathbb{N}}$ eine weitere in $\tilde{X}$ gegen $\tilde{x}$ konvergierende Folge, so konvergiert auch

$$\tilde{u}_n := \begin{cases} \tilde{y}_{n/2} & \text{, falls } n \text{ gerade} \\ \tilde{z}_{(n+1)/2} & \text{, falls } n \text{ ungerade} \end{cases}$$

gegen $\tilde{x}$, ist also eine Cauchyfolge und damit ist $(\Phi(\tilde{u}_n))$ konvergent. Der Grenzwert muss wieder $\hat{x}$ sein, denn $(\Phi(\tilde{u}_{2n}))_{n\in\mathbb{N}}$ konvergiert ja gegen $\hat{x}$. Damit ist auch $(\Phi(\tilde{z}_n))_{n\in\mathbb{N}} = (\Phi(\tilde{u}_{2n-1}))_{n\in\mathbb{N}}$ gegen $\hat{x}$ konvergent.

Die Definition

$$\Phi \;:\; \tilde{X} \longrightarrow \hat{X}\,,\;\; \tilde{x} \longmapsto \begin{cases} \psi(\varphi^{-1}(\tilde{x})) & \text{, falls } \tilde{x} \in \tilde{X}_0 \\ \lim\limits_{n\to\infty} \psi(\varphi^{-1}(\tilde{y}_n)) & \text{, falls } \tilde{x} \in \tilde{X} \setminus \tilde{X}_0 \text{ und} \\ & \lim\limits_{n\to\infty} \tilde{y}_n = \tilde{x},\; \tilde{y}_n \in \tilde{X}_0 \end{cases}$$

ist sinnvoll und für alle $\tilde{x}, \tilde{z} \in \tilde{X}$ gilt

$$\hat{d}\,(\,\Phi(\tilde{x})\,,\,\Phi(\tilde{z})\,) = \lim_{n\to\infty} \hat{d}(\,\Phi(\tilde{x}_n)\,,\,\Phi(\tilde{z}_n)\,) = \lim_{n\to\infty} \tilde{d}(\,\tilde{x}_n\,,\,\tilde{z}_n\,) = \tilde{d}(\,\tilde{x}\,,\,\tilde{z}\,) \quad ,$$

wobei $(\tilde{x}_n)$, $(\tilde{z}_n)$ gegen $\tilde{x}$ bzw. $\tilde{z}$ konvergierende Folgen in $\tilde{X}_0$ sind. Zum Nachweis der ersten bzw. letzten Gleichheit benötigt man die Ungleichung (9.16).

Φ ist also isometrisch und damit injektiv.

Φ ist aber auch surjektiv. Ist $\hat{x} \in \hat{X}$, so gibt es eine Folge $(\hat{x}_n)_{n\in\mathbb{N}}$ in $\hat{X}_0$, die gegen $\hat{x}$ konvergiert. Die Folge $(\varphi(\psi^{-1}(\hat{x}_n)))_{n\in\mathbb{N}} =: (\tilde{x}_n)_{n\in\mathbb{N}}$ ist dann Folge in $\tilde{X}_0$, die Cauchykonvergent, damit konvergent gegen ein $\tilde{x} \in \tilde{X}$ ist, und $\Phi(\tilde{x}) = \hat{x}$.

q.e.d.

Wir führen nun stetige Funktionen zwischen metrischen Räumen und den Grenzwert an einer Stelle ein.

Definition 9.17 $(X\,,\,d)$, $(Y\,,\,\tilde{d})$ *seien metrische Räume und* $f \;:\; X \longrightarrow Y$ *eine Funktion von* X *nach* Y *(mit Definitionsbereich* X*).* f *heißt stetig im Punkt* $x \in X$*, wenn*

$$\bigwedge_{\varepsilon>0} \; \bigvee_{\delta>0} \; f(U_\delta(x)) \subset U_\varepsilon(f(x)) \;;$$

f *heißt stetig in* $M \subset X$*, wenn* f *stetig ist in jedem Punkt* $x \in M$*;* f *heißt stetig, wenn* f *stetig ist in* X*.*

Eine reelle Funktion $f : D(f) \subset \mathbb{R} \longrightarrow \mathbb{R}$ ist also stetig, wenn sie als Funktion vom metrischen Raum $(D(f), d)$ in den metrischen Raum $(\mathbb{R}, d)$ stetig ist (mit $d(s, t) := |s - t|$).

Definition 9.18 *(X, d), $(Y, \tilde{d})$ seien metrische Räume und $f : X \longrightarrow Y$ eine Funktion von X nach Y. Ist $\tilde{x}$ ein Häufungspunkt von X, so heißt $y \in Y$ Grenzwert von f bei $\tilde{x}$, wenn*

$$\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} f(U_\delta(\tilde{x}) \setminus \{\tilde{x}\}) \subset U_\varepsilon(y) .$$

Man schreibt dann

$$\lim_{x\to\tilde{x}} f(x) := y .$$

Man erkennt leicht, dass der Grenzwert eindeutig bestimmt ist.

Satz 9.19 *Seien (X, d), $(Y, \tilde{d})$ metrische Räume, $f : X \longrightarrow Y$ eine Funktion und $\hat{x} \in X$. Folgende Aussagen sind äquivalent:*

(i) *f ist stetig in $\hat{x}$*

(ii) $\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} \bigwedge_{x\in X} (d(x, \hat{x}) < \delta \Rightarrow \tilde{d}(f(x), f(\hat{x})) < \varepsilon)$

(iii) $\bigwedge_{\varepsilon>0} \bigvee_{\delta>0} U_\delta(\hat{x}) \subset f^{-1}(U_\varepsilon(f(\hat{x})))$

(iv) *$\lim_{x\to\hat{x}} f(x) = f(\hat{x})$, wenn $\hat{x}$ Häufungspunkt von X ist*

(v) *Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in X mit $\lim_{n\to\infty} x_n = \hat{x}$ konvergiert $(f(x_n))_{n\in\mathbb{N}}$ in $(Y, \tilde{d})$ gegen $f(\hat{x})$.*

Korollar 9.20 *Sind (mit metrischen Räumen (X, d_x), (Y, d_y), (Z, d_z)) die Funktionen $f : X \longrightarrow Y$, $g : Y \longrightarrow Z$ stetig in $\hat{x}$ bzw. $f(\hat{x})$, so ist $g \circ f$ stetig in $\hat{x}$.*

Beweis: Die Äquivalenzen von (i), (ii), (iii) und (iv) zeigt man analog zum Beweis von Satz 4.2. Auch der Schluss von (iv) auf (v) verläuft analog zum Schluss von (iv) auf (vi) im Beweis von Satz 4.2.

Schließlich folgert man (iv) aus (v), mittels indirekten Beweises: Wäre (iv) falsch und $\hat{x}$ ein Häufungspunkt von X, so gäbe es ein $\varepsilon > 0$ und zu jedem $n \in \mathbb{N}$ ein $x_n \in X$ mit

$$x_n \in U_{1/n}(\hat{x}) \wedge f(x_n) \notin U_\varepsilon(f(\hat{x})) \ .$$

Die so gewonnene Folge (x_n) konvergiert in (X , d) gegen $\hat{x}$, aber $(f(x_n))$ konvergiert in $(Y , \tilde{d})$ nicht gegen $f(\hat{x})$ im Widerspruch zu (v).

q.e.d.

Satz 9.21 *Seien* (X , d), $(Y , \tilde{d})$ *metrische Räume und* $f : X \longrightarrow Y$ *eine Funktion. Dann sind äquivalent*

(i) f *ist stetig*

(ii) $f^{-1}(U)$ *ist offen für jede offene Teilmenge* U *von* Y

(iii) $f^{-1}(A)$ *ist abgeschlossen für jede abgeschlossene Teilmenge* A *von* Y

Beweis:
i $\Rightarrow$ ii: Ist $x \in f^{-1}(U)$ und $y = f(x)$, so existiert ein $\varepsilon > 0$ mit $U_\varepsilon(y) \subset U$. Nach Satz 9.19(*iii*) gibt es $\delta > 0$ mit

$$U_\delta(x) \subset f^{-1}(U_\varepsilon(f(x)) \subset f^{-1}(U) \quad ,$$

d. h. $f^{-1}(U)$ ist offen.

ii $\Rightarrow$ i: Ist $x \in X$ und $\varepsilon > 0$ vorgegeben, so ist $f^{-1}(U_\varepsilon(f(x)))$ wegen (ii) offen und $x \in f^{-1}(U_\varepsilon(f(x)))$. Mithin existiert $\delta > 0$, so dass $U_\delta(x) \subset f^{-1}(U_\varepsilon(f(x)))$, d. h. f ist stetig in x. Das liefert (i), denn x war beliebiger Punkt aus X.

q.e.d.

Viele der Eigenschaften reeller stetiger Funktionen lassen sich auf stetige Funktionen zwischen metrischen Räumen übertragen:

Definition 9.22 *Ein metrischer Raum* (X , d) *heißt kompakt, wenn jede Folge* $(x_n)_{n\in\mathbb{N}}$ *in* X *eine (gegen ein* $x \in X$*) konvergente Teilfolge* $(x_{\pi(n)})_{n\in\mathbb{N}}$ *besitzt.*

Eine Teilmenge $M \subset X$ *heißt kompakt, wenn der metrische Raum* $(M , d_{|M\times M})$ *kompakt ist.*

Satz 9.23 *Ist* (X , d) *kompakter metrischer Raum und* $f : X \longrightarrow Y$ *stetige Abbildung von* X *in einen weiteren metrischen Raum* (Y , d_Y)*, so ist der Wertebereich* $W(f) = f(X)$ *kompakt.*

Stetige Funktionen bilden also kompakte Mengen in kompakte Mengen ab.

Der Beweis von Satz 9.23 läßt sich fast wörtlich vom Beweis von Satz 4.12 übertragen.

Entsprechend Korollar 4.13 hat man:

Korollar 9.24 *Ist* (X , d) *kompakter metrischer Raum und* $f : X \longrightarrow \mathbb{R}$ *stetig, so nimmt* f *sein Minimum und sein Maximum an.*

Wir haben definiert, wann eine Teilmenge von $\mathbb{R}$ zusammenhängend ist, und haben dies in Satz 4.18 als topologische Eigenschaft erkannt. Wir orientieren uns an Definiton 4.19 und definieren:

Definition 9.25 *Ein metrischer Raum* (X , d) *heißt zusammenhängend, wenn sich* X *nicht darstellen läßt als punktfremde Vereinigung zweier nicht leerer offener Mengen, wenn also gilt:*

$$\bigwedge_{U,V\subset X,\, U,V \text{ offen}} (X = U \cup V \ \wedge \ U \cap V = \emptyset \Longrightarrow U = \emptyset \ \vee \ V = \emptyset) \, .$$

$M \subset X$ *heißt zusammenhängend, wenn der metrische Raum* $(M , d_{|M\times M})$ *zusammenhängend ist.*

Satz 9.26 *Ist* $f : X \longrightarrow Y$ *stetige Abbildung eines zusammenhängenden metrischen Raumes* (X , d) *in einen metrischen Raum* (Y , d)*, so ist der Wertebereich* $W(f) := f(X)$ *zuammenhängend.*

Beweis: Wäre $f(X)$ nicht zusammenhängend, gäbe es nicht leere offene Teilmengen U, V von $f(X)$ (offen relativ zu $f(X)$) mit $f(X) = U \cup V$, $U \cap V = \emptyset$.

Dann aber ist

$$X = f^{-1}(U) \cup f^{-1}(V) \ , \ \emptyset = f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V),$$

und $f^{-1}(U)$, $f^{-1}(V)$ sind nicht leere und nach 9.21(ii) offene Mengen. X wäre also nicht zusammenhängend.

q.e.d.

Definition 9.27 *Sind* (X, d) *und* (Y, d_Y) *metrische Räume so heißt*

$$C(X, Y) := \{ f : X \longrightarrow Y : X \text{ ist stetig}\}$$

Raum der stetigen Funktionen von X *nach* Y*.*

$$C_b(X, Y) := \{f \in C(X, Y) \colon \bigvee_{C>0} \quad \bigwedge_{x_1, x_2 \in X} \quad d_Y(f(x_1), f(x_2)) \leq C\}$$

heißt Raum der beschränkten stetigen Funktionen von X *nach* Y*.*

$C_b(X, Y)$ *wird zum metrischen Raum vermöge*

$$d_\infty : C_b(X, Y) \times C_b(X, Y) \longrightarrow \mathbb{R}, \ (f, g) \longmapsto \sup\{d_Y(f(x), g(x)) \colon x \in X\}$$

In der Tat ist dieses Supremum endlich: Gilt nämlich $d_Y(f(x_1), f(x_2)) \leq d_Y(g(x_1), g(x_2)) \leq C$ für alle $x_1, x_2 \in X$, so erhält man mit einem festen Punkt $\hat{x} \in X$:

$$\bigwedge_{x \in X} \quad d_Y(f(x), g(x)) \leq d_Y(f(x), f(\hat{x})) + d_Y(f(\hat{x}), g(\hat{x})) + d_Y(g(\hat{x}), g(x))$$
$$\leq 2c + d_Y(f(\hat{x}), g(\hat{x})) =: c_1 < \infty .$$

Ist (X, d) kompakt, so ist $C(X, Y) = C_b(X, Y)$: Fixiere nämlich $\hat{x} \in X$ und definiere für $f \in C(X, Y)$

$$F : X \longrightarrow \mathbb{R}, \ x \longmapsto d(f(x), f(\hat{x})) .$$

F ist stetig und nimmt damit sein Maximum an:

$$\max_{x \in X} F(x) =: \gamma .$$

Dann ist für $x_1, x_2 \in X$

$$d(f(x_1), f(x_2)) \leq d(f(x_1), f(\hat{x})) + d(f(\hat{x}), f(x_2)) \leq 2\gamma .$$

Wir möchten nun zeigen

Satz 9.28 *Sind* (X, d)*,* (Y, d_Y) *metrische Räume und ist* (Y, d_Y) *vollständig, so ist auch* $(C_b(X, Y), d_\infty)$ *ein vollständiger metrischer Raum.*

Beweis: Sei f_n eine Cauchyfolge in $(C_b(X, Y), d_\infty)$; es gibt also eine Nullfolge $(\nu_n)_{n \in \mathbb{N}}$ so, dass für alle $n, m \in \mathbb{N}$ mit $m \geq n$ gilt

$$d_\infty(f_n, f_m) = \sup_{x \in X} d_Y(f_n(x), f_m(x)) \leq \nu_n . \tag{9.18}$$

Diese Abschätzung impliziert, dass für jedes $x \in X$ die Folge $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchyfolge in $(Y\,,\,d_Y)$ ist. Aber Y ist vollständig, $(f_n(x))_{n\in\mathbb{N}}$ konvergiert also in Y für jedes feste $x \in X$. ("punktweise" Konvergenz), und wir definieren

$$f\ :\ X \longrightarrow Y\ ,\ x \longmapsto \lim_{n\to\infty} f_n(x)\ .$$

Der Nachweis, dass f stetig ist, verlangt die Vertauschung zweier Grenzprozesse: Ist $x \in X$ und $(x_k)_{k\in\mathbb{N}}$ eine in X gegen x konvergente Folge, so haben wir $\lim_{k\to\infty} f(x_k) = f(x)$ zu zeigen. Mit anderen Worten: zu zeigen ist

$$\lim_{k\to\infty}\lim_{n\to\infty} f_n(x_k) = \lim_{n\to\infty}\lim_{k\to\infty} f_n(x_k)\ . \tag{9.19}$$

Aus der Abschätzung (9.18) gewinnt man

$$\bigwedge_{x\in X} \quad d_Y(\,f_n(x)\,,\,f(x)\,) = \lim_{m\to\infty} d_Y(\,f_n(x)\,,\,f_m(x)\,) \le \nu_n \quad , \tag{9.20}$$

und (9.20) — die gleichmäßige Konvergenz von $f_n(x)$ gegen $f(x)$ — liefert eine hinreichende Bedingung für (9.19):
Für alle $n, k \in \mathbb{N}$ ist nämlich

$$\begin{aligned}
&d_Y(\,f(x)\,,\,f(x_k)\,)\\
&\quad\le d_Y(\,f(x)\,,\,f_n(x)\,) + d_Y(\,f_n(x)\,,\,f_n(x_k)\,) + d_Y(\,f_n(x_k)\,,\,f(x_k)\,)\\
&\quad\le 2\nu_n + d_Y(\,f_n(x)\,,\,f_n(x_k)\,)\ .
\end{aligned}$$

Hieraus erhält man für alle $n \in \mathbb{N}$

$$\limsup_{k\to\infty} d_Y(\,f(x)\,,\,f(x_k)\,) \le 2\nu_n \quad ,$$

also

$$\limsup_{k\to\infty} d_Y(\,f(x)\,,\,f(x_k)\,) = 0 \quad ,$$

mithin

$$\lim_{k\to\infty} f(x_k) = f(x)\ .$$

Auch ist $f \in C_b(\,X\,,\,Y\,)$, denn für $x_1, x_2 \in X$ ist

$$\begin{aligned}
&d_Y(\,f(x_1)\,,\,f(x_2)\,)\\
&\quad\le d_Y(\,f(x_1)\,,\,f_9(x_1)\,) + d_Y(\,f_9(x_1)\,,\,f_9(x_2)\,) + d_Y(\,f_9(x_2)\,,\,f(x_2)\,)\\
&\quad\le 2\nu_9 + C_9 \quad ,
\end{aligned}$$

wobei

$$C_n := \sup_{x_1,x_2\in X} d_Y(\, f_n(x_1)\,,\, f_n(x_2)\,)\ .$$

Schließlich ist $\lim\limits_{n\to\infty} f_n = f$, denn (9.20) impliziert

$$d_\infty(\, f_n\,,\, f\,) = \sup_{x\in X} d_Y(\, f_n(x)\,,\, f(x)\,) \le \nu_n \to 0\ .$$

q.e.d.

Hinter dem Nachweis von (9.18) steht ein Prinzip zur Vertauschung von Grenzprozessen, das wir herausarbeiten wollen:

Satz 9.29 *(über die Vertauschung von Grenzprozessen).*

Sei $(\, X\,,\, d\,)$ metrischer Raum und $(x_{nm})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ eine Doppelfolge, d. i. eine Funktion von $\mathbb{N}\times\mathbb{N}$ nach X, $(n,m)\to x_{nm}$. Es mögen gelten

(i) *Für jedes feste $n\in\mathbb{N}$ konvergiert $(x_{nm})_{m\in\mathbb{N}}$ gegen $x_{n\infty}\in X$:*

$$\bigwedge_{n\in\mathbb{N}} \lim_{m\to\infty} x_{nm} = x_{n\infty}$$

(ii) *Für jedes feste $m\in\mathbb{N}$ konvergiert $(x_{nm})_{n\in\mathbb{N}}$ gegen $x_{\infty m}\in X$:*

$$\bigwedge_{m\in\mathbb{N}} \lim_{n\to\infty} x_{nm} = x_{\infty m}$$

(iii) *$(x_{n\infty})_{n\in\mathbb{N}}$ konvergiert gegen $x_{\infty\infty}\in X$:*

$$\lim_{n\to\infty} x_{n\infty} = x_{\infty\infty}$$

(iv) *Der Grenzübergang in (i) oder der in (ii) ist gleichmäßig bezüglich der anderen Variablen: Es gibt eine Nullfolge $(\nu_k)_{k\in\mathbb{N}}$ mit*

$$\bigwedge_{(n,m)\in\mathbb{N}\times\mathbb{N}} d(\, x_{nm}\,,\, x_{n\infty}\,) \le \nu_m \tag{9.21}$$

oder

$$\bigwedge_{n,m\in\mathbb{N}\times\mathbb{N}} d(\, x_{nm}\,,\, x_{\infty m}\,) \le \nu_n \tag{9.22}$$

Dann ist

$$\lim_{m\to\infty} x_{\infty m} = x_{\infty\infty}\ .$$

Beweis: Im Fall (9.21) ist für alle $n, m \in \mathbb{N}$

$$d(\, x_{n\infty} \, , \, x_{nm} \,) \leq \nu_m$$

Mithin ist für alle $m \in \mathbb{N}$

$$d(\, x_{\infty\infty} \, , \, x_{\infty m} \,) = \lim_{n\to\infty} d(\, x_{n\infty} \, , \, x_{nm} \,) \leq \nu_m \quad ,$$

also strebt $x_{\infty m}$ gegen $x_{\infty\infty}$ für m gegen ∞.

Im Fall (9.22) gehen wir wie beim Beweis von 9.18 vor:

Für alle $n, m \in \mathbb{N}$ ist

$$\begin{aligned} d(\, x_{\infty\infty} \, , \, x_{\infty m} \,) &\leq d(\, x_{\infty\infty} \, , \, x_{n\infty} \,) + d(\, x_{n\infty} \, , \, x_{nm} \,) + d(\, x_{nm} \, , \, x_{\infty m} \,) \\ &\leq d(\, x_{\infty\infty} \, , \, x_{n\infty} \,) + d(\, x_{n\infty} \, , \, x_{nm} \,) + \nu_n \; . \end{aligned}$$

Daher folgt
$0 \leq \limsup_{m\to\infty} d(\, x_{\infty\infty} \, , \, x_{\infty m} \,) \leq d(\, x_{\infty\infty} \, , \, x_{n\infty} \,) + \nu_n$.
Da die rechte Seite n–tes Glied einer Nullfolge ist, folgt $\lim\limits_{m\to\infty} d(\, x_{\infty\infty} \, , \, x_{\infty m} \,) = 0$.

q.e.d.

9.2 Normierte Räume

Bei einer Reihe von Beispielen metrischer Räume, die im vergangenen Abschnitt angegeben worden sind, handelt es sich um Vektorräume über dem Körper $\mathbb{K} := \mathbb{R}$ oder $\mathbb{K} := \mathbb{C}$, und die angegebenen Metriken hatten die Eigenschaften, translationsinvariant zu sein und im folgenden Sinn Streckungen zu respektieren: streckt man zwei Vektoren mit einem Faktor $\lambda \in \mathbb{K}$, so wird ihr Abstand um den Faktor $|\lambda|$ gestreckt. Diese beiden Bedingungen bedeuten für die Metrik d:

$$d(\, x \, , \, y \,) = d(\, 0 \, , \, y - x \,) \quad ; \quad d(\, \lambda x \, , \, \lambda y \,) = |\lambda| \, d(\, x \, , \, y \,) \; .$$

Der Abstand zweier Punkte berechnet sich also als Abstand des Differenzvektors zum Nullpunkt oder als "Länge des Differenzvektors".

Definition 9.30 *Ein normierter Raum ist ein reeller oder komplexer Vektorraum X — $\mathbb{K}$ bezeichne den Grundkörper — , auf dem eine Norm erklärt ist. Das ist eine Abbildung $\| \cdot \| \; : \; X \longrightarrow \mathbb{R}$ mit den Eigenschaften*

(i) $\bigwedge\limits_{x\in X} (\|x\| = 0 \Leftrightarrow x = 0)$,

(ii) $\bigwedge\limits_{x\in X} \bigwedge\limits_{\lambda\in\mathbb{K}} \|\lambda x\| = |\lambda| \, \|x\|$,

(iii) $\bigwedge\limits_{x,y\in X} \|x+y\| \le \|x\| + \|y\|$.

Wir schreiben meistens X *oder* $(X, \|\cdot\|)$ *statt der korrekteren Schreibweise* $(X, +, \cdot, \|\cdot\|)$.

Indem man in (iii) $y := -x$ setzt und (ii) beachtet, erhält man

$$\bigwedge_{x\in X} \|x\| \ge 0 \, . \tag{9.23}$$

(iii) nennt man wieder "die Dreiecksungleichung"; wegen (i) zusammen mit (9.23) nennt man $\|\cdot\|$ "positiv definiert", und (ii) nennt man die Homogenität der Norm.

Der Beweis des folgenden Satzes ist sehr einfach, und wir wollen ihn hier nicht ausführen:

Satz 9.31 *Durch*

$$d \, : \; X \times X \longrightarrow \mathbb{R} \, , \; (x,y) \longmapsto \|x-y\|$$

ist im normierten Raum $(X, \|\cdot\|)$ *eine Metrik definiert, die natürliche Metrik des normierten Raumes* $(X, \|\cdot\|)$.

Alle Begriffe, die in Abschnitt 9.1 für metrische Räume eingeführt worden sind, sind auch auf normierten Räumen erklärt, in denen man sich die natürliche Metrik als Metrik eingeführt denkt.

Beispiele 9.32 *Mit* $p \in [1, \infty]$ *für* $x \in \mathbb{K}^N$ *sei*

$$\|x\|_p := \begin{cases} \left(\sum\limits_{n=1}^{N} |x_n|^p\right)^{1/p} & , \textit{ falls } \; p < \infty \\ \max\limits_{n\in 1,\dots,N} |x_n| & , \textit{ falls } \; p = \infty \, . \end{cases}$$

$\|\cdot\|_p$ *ist eine Norm auf* $\mathbb{K}^N$. *Insbesondere die Dreiecksungleichung ist genau die Minkowski-Ungleichung (Korollar* 6.25*). Statt* $(\mathbb{K}^N, \|\cdot\|_p)$ *schreibt man oft*

$$l_p(N) \quad \textit{oder} \quad l_p(N, \mathbb{K}) \, .$$

Ist X ein normierter und M ein metrischer Raum, so ist

$$C_b(M , X)$$

ein normierter Raum, wenn man zunächst eine Vektorraumstruktur auf $C_b(M , X)$ durch

$$\alpha f + \beta g \; : \; M \longrightarrow X \, , \; m \longmapsto \alpha f(m) + \beta g(m)$$

für $\alpha, \beta \in \mathbb{K}$, $f, g \in C_b(M , X)$ definiert und dann

$$\|f\|_\infty := \sup_{m \in M} \|f(m)\| \quad \text{für} \quad f \in C_b(M , X)$$

als Norm einführt. Hier ist $\| \cdot \|$ die Norm in X.

Definition 9.33 *Ist der normierte Raum X mit seiner natürlichen Metrik vollständig, so nennt man X einen Banachraum (Stefan Banach, 1892 - 1945).*

Beispiele für Banachräume sind die Räume $lp(N)$, $N \in \mathbb{N}$ sowie $C_b(M , X)$, falls M metrischer Raum und X Banachraum ist (siehe Satz 9.28.

Andererseits ist (mit einem kompakten Intervall I) der normierte Raum $(C(I) \, , \, \| \cdot \|_2)$ mit

$$\|f\|_2 := \left(\int_I |f(t)|^2 \, dt \right)^{1/2}$$

<u>kein</u> Banachraum. Dazu gab es eine Übungsaufgabe.

Typische Abbildungen zwischen Vektorräumen sind die linearen Abbildungen. Hat man es mit normierten Räumen zu tun, interessieren besonders stetige lineare Abildungen.

Satz 9.34 *Seien X, Y normierte Räume über dem gleichen Körper $\mathbb{K}$ und $\varphi : X \longrightarrow Y$ eine lineare Abbildung (Es gilt also $\varphi(\alpha u + \beta v) = \alpha\varphi(u) + \beta\varphi(v)$ für alle $u, v \in X$, $\alpha\beta \in \mathbb{K}$). Dann sind äquivalent*

(i) *φ ist stetig*

(ii) *Es gibt einen Punkt $x \in X$, in dem φ stetig ist*

(iii) *φ ist stetig in 0*

(iv) $\bigvee\limits_{C>0} \; \bigwedge\limits_{x \in X} \; \|\varphi(x)\|_Y \leq C\|x\|_X .$

(v) $\sup\{\|\varphi(x)\|_Y \colon x \in S_1(0)\} < \infty$.

Beweis:
(i) $\Rightarrow$ (ii) ist klar.
(ii) $\Rightarrow$ (iii): Ist $\varepsilon > 0$ vorgegeben, so existiert $\delta > 0$ mit $\varphi(U_\delta(x)) \subset U_\varepsilon(\varphi(x))$. Dann aber ist

$$\begin{aligned}\varphi(U_\delta(0)) &= \varphi(U_\delta(x) - x) = \varphi(U_\delta(x)) - \varphi(x) \subset U_\varepsilon(\varphi(x)) - \varphi(x) \\ &= U_\varepsilon(0)\ .\end{aligned}$$

(iii) $\Rightarrow$ (iv): Zu $\varepsilon := 1$ existiert $\delta > 0$ mit $\varphi(U_\delta(0)) \subset U_1(\varphi(0)) = U_1(0)$. Ist nun $x \in X$, $x \neq 0$, so ist

$$\hat{x} := \frac{\delta}{2}\|x\|^{-1}x \in U_\delta(0)\ .$$

Folglich ist

$$\begin{aligned}\|\varphi(x)\| &= \|\varphi((\frac{2}{\delta}\|x\|)\hat{x})\| \\ &= \frac{2}{\delta}\|x\|\|\varphi(\hat{x})\| \leq \frac{2}{\delta}\|x\|\ .\end{aligned}$$

Also gilt (iv) mit $C := 2/\delta$.
(iv) $\Rightarrow$ (v) ist unmittelbar klar
(v) $\Rightarrow$ (i): Mit

$$\gamma := \sup\{\|\varphi(x)\|_Y \colon x \in S_1(0) \subset X\}\ .$$

gilt für alle $z \in X$

$$\|\varphi(z)\| \leq \gamma\|z\|\ .$$

Dies ist richtig für $z = 0$, und für $z \neq 0$ ist

$$\|\varphi(z)\| = \|z\| \cdot \|\varphi(\|z\|^{-1}z)\| \leq \gamma\|z\|\ .$$

Strebt nun die Folge $(x_n)_{n\in\mathbb{N}}$ in X gegen ein vorgelegtes $x \in X$, so gilt

$$\|\varphi(x_n) - \varphi(x)\| = \|\varphi(x_n - x)\| \leq \gamma\|x_n - x\| \longrightarrow 0\ .$$

q.e.d.

Korollar 9.35 *Sei X normierter Raum und $N \in \mathbb{N}$. Jede lineare Abbildung von $l_2(N)$ nach X ist stetig.*

Beweis: Ist $e_1, \ldots, e_N$ ein Orthonormalsystem in $l_2(N)$ und $\varphi : l_2(N) \longrightarrow X$ linear, so gilt für jedes $u := \sum\limits_{n=1}^{N} \mu_n e_n$:

$$\begin{aligned}
\|\varphi(u)\| &= \|\sum_{n=1}^{N} \mu_n \varphi(e_n)\| \leq \sum_{n=1}^{N} |\mu_n| \, \|\varphi(e_n)\| \\
&\leq \left(\sum_{n=1}^{N} |\mu_n|^2\right)^{1/2} \left(\sum_{n=1}^{N} \|\varphi(e_n)\|^2\right)^{1/2} \\
&= C\|u\|
\end{aligned}$$

mit $C := \left(\sum\limits_{n=1}^{N} \|\varphi(e_n)\|^2\right)^{1/2}$.

q.e.d.

Man kann auf einem gegebenen Vektorraum verschiedene Normen einführen. Zum Beispiel ist $l_p(N)$ mit $p \in [1, \infty]$ stets der Vektorraum $\mathbb{K}^N$, aber mit den verschiedenen Normen $\| \cdot \|_p$. Möglicherweise erzeugen aber alle diese Normen den gleichen Konvergenzbegriff und die gleichen offenen oder abgeschlossenen Mengen. Dies ist der Fall, wenn sie äquivalent sind:

Definition 9.36 *Seien X ein Vektorraum über $\mathbb{K}$ und $\| \cdot \|$ und $| \cdot |$ Normen auf X, also Abbildungen mit den Eigenschaften aus Definition 9.31($i - iii$). Man nennt die beiden Normen äquivalent, wenn Konstanten c_-, $c_+ > 0$ existieren, so dass für alle $x \in X$ gilt*

$$c_- \|x\| \leq |x| \leq c_+ \|x\| \; . \tag{9.24}$$

In der Tat wird hierdurch eine Äquivalenzrelation auf der Menge der Normen auf X erklärt. Zum Beispiel impliziert (9.24):

$$\bigwedge_{x \in X} \quad \frac{1}{c_+} |x| \leq \|x\| \leq \frac{1}{c_-} |x| \quad ,$$

und gilt mit einer weiteren Norm $\|| \cdot \||$

$$\bigvee_{\gamma_-, \gamma_+ > 0} \quad \bigwedge_{x \in X} \quad \gamma_- |x| \leq \||x\|| \leq \gamma_+ |x| \quad ,$$

so folgt für alle $x \in X$

$$c_-\gamma_-\|x\| \leq |\!|\!|x|\!|\!| \leq c_+\gamma_+\|x\|$$

Nun sind auf endlich dimensionalen Vektorräumen alle Normen einander äquivalent:

Satz 9.37 *Ist X ein Vektorraum endlicher Dimension mit den Normen $\|\cdot\|_0$ und $\|\cdot\|_1$, so sind $\|\cdot\|_0$ und $\|\cdot\|_1$ äquivalent.*

Beweis: $\|\cdot\|_i$ sei eine der beiden Normen $\|\cdot\|_0$ oder $\|\cdot\|_1$ und N sei die Dimension des Raums X, welcher durch $\|\cdot\|_i$ normiert sei. Es gibt dann eine lineare Bijektion $\varphi : l_2(N) \longrightarrow X$, und diese ist nach Korollar 9.35 stetig. Setze

$$\hat{\gamma}_i := \sup\{\|\varphi(\xi)\|\textbf{:}\ {}_{\}}i\xi \in S_1(0) \subset l_2(N)\ .$$

Nun ist $S_1(0) \subset l_2(N)$ kompakt: Ist $(\xi^{(n)})_{n\in\mathbb{N}}$ eine Folge in $l_2(N)$ mit

$$\|\xi^{(n)}\| := \left(\sum_{j=1}^{N}\left|\xi_i^{(n)}\right|^2\right)^{1/2} = 1 \quad ,$$

so enthält $(\xi^{(n)})$ eine Teilfolge $(\xi^{\pi(n)})_{n\in\mathbb{N}}$ derart, dass $(\xi_i^{\pi(n)})$ für jedes $i = 1,\ldots,N$ konvergiert. Damit konvergiert auch $(\xi^{\pi(n)})_{n\in\mathbb{N}}$ in $l_2(N)$, und da $S_1(0)$ abgeschlossen ist, liegt der Grenzwert in $S_1(0)$.

Durch $f : S_1(0) \longrightarrow \mathbb{R}$, $\xi \longmapsto \|\varphi(\xi)\|_i$ wird eine stetige Abildung definiert, und diese nimmt an einer Stelle $\hat{\xi} \in S_1(0)$ ihr Minimum an. Da $\varphi(\hat{\xi}) \neq 0$ ist, folgt

$$\check{\gamma}_i := \min\{\|\varphi(\xi)\|_i\textbf{:}\ \xi \in S_1(0) \subset l_2(N)\} > 0\ .$$

Ist nun $x \in X \setminus \{0\}$, so gibt es ein $\xi \in l_2(N) \setminus \{0\}$ mit $\varphi(\xi) = x$.

Aus

$$\|x\|_i = \|\varphi(\xi)\|_i = \|\xi\|\ \ \|\varphi(\frac{\xi}{|\xi|})\|_i$$

folgt

$$\check{\gamma}_i\|\xi\| \leq \|x\|_i \leq \hat{\gamma}_i\|\xi\| \quad ,$$

und damit

$$\|x\|_0 \leq \hat{\gamma}_0\|\xi\| \leq \frac{\hat{\gamma}_0}{\check{\gamma}_1}\|x\|_1\ ,$$

$$\|x\|_1 \leq \frac{\hat{\gamma}_1}{\check{\gamma}_0}\|x\|_0\ ;$$

folglich:

$$\frac{\check{\gamma}_1}{\hat{\gamma}_0}\|x\|_0 \leq \|x\|_1 \leq \frac{\hat{\gamma}_1}{\check{\gamma}_0}\|x\|_0 \ .$$

q.e.d.

Satz 9.38 *Sind X, Y normierte Räume, mit Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$, so ist $\mathcal{L}(\,X\,,\,Y\,) := \{\,\varphi\ :\ X\ \longrightarrow\ Y : \varphi$ stetig und linear$\}$ ein normierter Raum mit der Norm*

$$\|\varphi\| := \sup_{x\in S_1(0)\subset X} \|\varphi(x)\|_Y \ . \tag{9.25}$$

Ist Y Banachraum, so auch $\mathcal{L}(\,X\,,\,Y\,)$.

Beweis: Der Beweis soll nur skizziert werden. In der Tat ist $\mathcal{L}(\,X\,,\,Y\,)$ ein Untervektorraum von $C(X,Y)$. Man weise selbst nach, dass durch (9.25) tatsächlich eine Norm auf $\mathcal{L}(\,X\,,\,Y\,)$ gegeben ist.

Wir zeigen, dass $\mathcal{L}(\,X\,,\,Y\,)$ vollständig ist, falls Y es ist. Hierzu benötigt man die Ungleichung

$$\bigwedge_{\phi\in\mathcal{L}(X,Y)} \ \bigwedge_{x\in X} \|\phi(x)\| \leq \|\phi\|\cdot\|x\| \ . \tag{9.26}$$

Diese Ungleichung ist für $x = 0$ trivial und folgt aus der Definition der Norm von ϕ, wenn $\|x\| = 1$ ist. Für beliebiges $x \neq 0$ hat man schließlich

$$\|\phi(x)\| = \|\|x\|\phi(\frac{x}{\|x\|})\| = \|x\|\cdot\|\phi(\frac{x}{\|x\|})\| \leq \|\phi\|\cdot\|x\| \ .$$

Sei nun $(\phi_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{L}(X,Y)$ und $(\nu_n)_{n\in\mathbb{N}}$ eine Nullfolge mit $\|\phi_{n+k} - \phi_n\| \leq \nu_n$ für alle $n, k \in \mathbb{N}$. Dann ist für jedes $x \in X$ und $n, k \in \mathbb{N}$

$$\|\phi_{n+k}(x) - \phi_n(x)\| = \|(\phi_{n+k} - \phi_n)(x)\| \leq \nu_n\|x\| \ .$$

Da Y vollständig ist, ist durch

$$\phi\ :\ X\ \longrightarrow\ Y\ ,\ \ x\ \longmapsto\ \lim_{n\to\infty}\phi_n(x)$$

eine Funktion von X nach Y definiert. Man kann zeigen, dass diese linear und stetig ist und der Grenzwert von $(\phi_n)_{n\in\mathbb{N}}$ im normierten Raum $\mathcal{L}(X,Y)$.

q.e.d.

9.3 Differentation in Banachräumen

In der linearen Algebra haben Sie lineare Abbildungen kennen gelernt. Für endlichdimensionale Räume sind alle linearen Abbildungen auch stetig. Wir definieren in Analogie zu Definition **??**:

Definition 9.39 *Sei Ω eine offene Teilmenge eines Banachraums X und*

$$f \;:\; \Omega \subset X \;\longrightarrow\; Y$$

eine Funktion in dem Banachraum Y. f heißt differenzierbar in $\tilde{x} \in \Omega$, wenn eine stetige lineare Abbildung $A \in \mathcal{L}(\,X\,,\,Y\,)$ existiert, so dass für alle $x \in \Omega$ gilt:

$$f(\tilde{x}+h) = f(\tilde{x}) + Ah + \|h\|_X \rho(h) \tag{9.27}$$

mit einer in 0 stetigen Funktion $\rho \;:\; U_\delta(0) \;\longrightarrow\; Y$, die $\delta(0) = 0$ erfüllt.

Aus (**??**) erkennt man direkt:

Satz 9.40 *Sind X, Y Banachräume, $\Omega \subset X$ offen und $f \;:\; \Omega \subset X \;\longrightarrow\; Y$ differenzierbar in $\tilde{x}$, so ist f stetig in $\tilde{x}$.*

Satz und Definition 9.41 *Mit den Bezeichnungen aus Definiton 9.39:*
Sind sowohl A als auch $\tilde{A}$ Ableitung von f in $\tilde{x}$, so ist $A = \tilde{A}$. Wir schreiben in Zukunft $f'(\tilde{x}) := A$.

Beweis: Wir verwenden (9.27) mit $h := r\xi$, $\xi \in S_1(0)$ und $0 \leq r \leq \delta$: Sind A und $\tilde{A}$ Ableitung von f in $\tilde{x}$, so folgt für alle $r \in (\,0\,,\,\delta\,)$ und $\xi \in S_1(0)$

$$rA\xi + r\rho(r\xi) = r\tilde{A}\xi + r\tilde{\rho}(r\xi) \quad ,$$

mit $\rho(r\xi)$, $\tilde{\rho}(r\xi) \to 0$ für $r \to 0$. Division durch r und Grenzübergang r gegen 0 liefert

$$A\xi = \tilde{A}\xi \quad .$$

Da dies für jedes $\xi \in S_1(0)$ gilt, folgt $A = \tilde{A}$.

q.e.d.

Beispiele 9.42

(i) *X, Y seien Banachräume, $\hat{y} \in Y$. Dann ist die konstante Funktion*

$$f \,:\, X \longrightarrow Y\,,\; x \longmapsto \hat{y}$$

differenzierbar in jedem Punkt $x \in X$ und $f'(x) = 0$, die Nullabbildung.

(ii) *Ist $A \in \mathcal{L}(X\,, Y\,)$ und $f \,:\, X \longrightarrow Y\,,\; x \longmapsto Ax$, so ist f differenzierbar (in jedem Punkt $x \in X$) und $f'(x) = A$.*

(iii) *Ist X Euklidischer Raum mit Skalarprodukt $\langle \cdot\,, \cdot \rangle$, so ist*

$$f \,:\, X \times X \longrightarrow \mathbb{R}\,,\; (x,y) \longmapsto \langle\, x\,, y\,\rangle$$

differenzierbar in jedem Punkt $(\,\hat{x}\,, \hat{y}\,) \in X \times X$.

Man beachte hier: Mit $\|(\,a\,, b\,)\| := (\|a\|^2 + \|b\|^2)^{1/2}$ für $(\,a\,, b\,) \in X \times X$ ist auch $X \times X$ ein Banachraum, und es gilt für alle $(x, y),\; h := (\xi, \eta) \in X \times X$

$$\begin{aligned} f(\,x + \xi\,, y + \eta\,) &= \langle\, x\,, y\,\rangle + \langle\, \xi\,, y\,\rangle + \langle\, x\,, \eta\,\rangle + \langle\, \xi\,, \eta\,\rangle \\ &= f(\,x\,, y\,) + A(\,\xi\,, \eta\,) + \|(\,\xi\,, \eta\,)\| \frac{\langle\, \xi\,, \eta\,\rangle}{\|(\,\xi\,, \eta\,)\|} \quad , \end{aligned}$$

wobei

$$A(\,\xi\,, \eta\,) := \langle\, \xi\,, y\,\rangle + \langle\, x\,, \eta\,\rangle$$

und

$$\begin{aligned} \lim_{(\,\xi\,, \eta\,) \to 0} \left| \frac{\langle\, \xi\,, \eta\,\rangle}{\|(\,\xi\,, \eta\,)\|} \right| &= \left| \frac{\langle\, \xi\,, \eta\,\rangle}{(\|\xi\|^2 + \|\eta\|^2)^{1/2}} \right| \le \frac{\|\xi\|\|\eta\|}{(\|\xi\|^2 + \|\eta\|^2)^{1/2}} \\ &\le \frac{1}{2}\|(\,\xi\,, \eta\,)\| \to 0\,. \end{aligned}$$

Also ist

$$f'(x,y)(\,\xi\,, \eta\,) = \langle\, x\,, \eta\,\rangle + \langle\, \xi\,, y\,\rangle\,.$$

(iv) *X wie oben und $f \,:\, X \longrightarrow \mathbb{R}\,,\; x \longmapsto \|x\|^2$. Dann ist für $h \in X$*

$$f(x+h) = \|x+h\|^2 = \|x\|^2 + 2\langle\, x\,, h\,\rangle + \|h\|^2 = f(x) + 2\langle\, x\,, h\,\rangle + \|h\|^2\,.$$

Daher $f'(x)h = 2\langle\, x\,, h\,\rangle$ für alle $x \in X,\; h \in X$.

Der folgende Satz erlaubt, Ableitungen — sofern sie existieren — zu berechnen:

Satz und Definition 9.43 *Seien X, Y Banachräume über $\mathbb{K}$, $\Omega \subset X$ offen und $f : \Omega \subset X \longrightarrow Y$ sei differenzierbar in $x \in \Omega$. Dann gilt für jedes $\xi \in X$*

$$f'(x)\xi = \lim_{t \to 0} \frac{f(x+t\xi) - f(x)}{t} \ .$$

Den Grenzwert

$$(\frac{\partial}{\partial \xi} f)(x) := \lim_{t \to 0} \frac{f(x+t\xi) - f(x)}{t}$$

nennt man die Richtungsableitung von f in x zum Richtungsvektor ξ.

Beweis: Für genügend kleines $\delta > 0$ und $|t| < \delta$ ist

$$f(x+t\xi) = f(x) + t f'(x)\xi + \|t\xi\| \rho(t\xi)$$

mit $\lim_{t \to 0} \rho(t\xi) = 0$. Daraus folgt die Behauptung.

q.e.d.

Beispiele 9.44

(i) $f \ : \ \mathbb{R}^3 \longrightarrow \mathbb{R} \, , \ x \longmapsto \exp(x_1 x_2 \cos x_3)$
Ist f differenzierbar in x, so wird die Ableitung $f'(x)$ beschrieben durch eine 1×3–Matrix (a_1, a_2, a_3), und zwar ist

$$a_j = f'(x)e^j \ \textit{mit} \ e_i^j := \delta_{ij} := \begin{cases} 0 & , \textit{falls } i \neq j \\ 1 & , \textit{falls } i = j \end{cases} \ .$$

Nach Satz 9.43 *ist*

$$f'(x)e^j = \lim_{t \to 0} \frac{1}{t} \left[f(x+te^j) - f(x) \right] = \varphi_j'(0)$$

mit

$$\varphi_j \ : \ (-\delta \, , \, \delta) \longrightarrow \mathbb{R} \, , \ t \longmapsto f(x+te^j) \, .$$

Wegen $\varphi_1(t) = \exp\left((x_1+t)x_2 \cos x_3\right)$, $\varphi_2(t) = \exp\left(x_1(x_2+t)\cos x_3\right)$ und $\varphi_3(t) = \exp\left(x_1 x_2 \cos(x_3+t)\right)$ folgt

$$\begin{aligned} f'(x)e^1 &= x_2 \cos x_3 f(x) \, , \\ f'(x)e^2 &= x_1 \cos x_3 f(x) \, , \\ f'(x)e^3 &= -x_1 x_2 \sin x_3 f(x) \, , \\ f'(x)\xi &= (f'(x)e^1, f'(x)e^2, f'(x)e^3) \cdot \xi \end{aligned}$$

(ii) $f\ :\ \mathbb{R}^3 \longrightarrow \mathbb{R}^2\,,\ x \longmapsto \begin{bmatrix} \exp\quad (x_1\cos x_2 \sin x_3) \\ \sin(x_1+x_2)\cos x_3 \end{bmatrix}$

$f'(x)e^j = \begin{bmatrix} \varphi'_{1j}(0) \\ \varphi'_{2j}(0) \end{bmatrix} \quad mit \quad \varphi_{ij}(t) := f_i(x+te^j) \quad ,$

also

$$f'(x)e^1 = \begin{bmatrix} \cos x_2 \sin x_3 \exp(x_1\cos x_2 \sin x_3) \\ \cos(x_1+x_2)\cos x_3 \end{bmatrix}$$

$$f'(x)e^2 = \begin{bmatrix} -x_1 \sin x_2 \sin x_3 \exp(x_1\cos x_2 \sin x_3) \\ \cos(x_1+x_2)\cos x_3 \end{bmatrix}$$

$$f'(x)e^3 = \begin{bmatrix} x_1 \cos x_2 \cos x_3 \exp(x_1\cos x_2 \sin x_3) \\ -\sin x_3 \cos(x_1+x_2) \end{bmatrix}$$

Dann wird für $x \in \mathbb{R}^3$ die Ableitung bezüglich der kanonischen Basis durch die 2×3–Matrix

$$J(x) := (f'(x)e^1\,,\ f'(x)e^2\,,\ f'(x)e^3)$$

beschrieben.

Satz und Definition 9.45 *Ist $\Omega \subset \mathbb{R}^N$ offen und $f\ :\ \Omega \in \mathbb{R}^N \longrightarrow \mathbb{R}^M$ differenzierbar in $x \in \Omega$, so ist für $\xi \in \mathbb{R}^N$*

$$f'(x)\xi = \begin{bmatrix} \partial_1 f_1(x), & \ldots, & \partial_N f_1(x) \\ \partial_1 f_2(x), & \ldots, & \partial_N f_2(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_M(x), & \ldots, & \partial_N f_M(x) \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_N \end{bmatrix} \tag{9.28}$$

mit

$$\partial_i f_j(x) := f'_{ji}(0) \quad mit \quad f_{ji}(t) := f_j(x+te^i)\ .$$

Die i-te Spalte dieser Matrix

$$\partial_i f(x) := \begin{bmatrix} \partial_i f_1(x) \\ \vdots \\ \partial_i f_M(x) \end{bmatrix}$$

heißt i–te partielle Ableitung von f in x, und es gilt

$$\partial_i f(x) := \lim_{t\to 0} \frac{1}{t}\,[f(x+te) - f(x)] \quad .$$

Die Matrix in (9.28) heißt "Jacobimatrix von f in x" (Carl Gustav Jacob Jacobi, 1804 - 1851) und wird mit $J_f(x)$ abgekürzt.

Wir wissen also nun, wie man die Ableitung zu berechnen hat, sofern sie existiert. Unsere erste Aufgabe im folgenden Semester wird sein, an der Jacobimatrix ablesen zu lernen, ob die vorgelegte Funktion in der Tat differenzierbar in x ist. Dann nämlich vermittelt die Jacobimatrix die Ableitung von f in x.

Wir beschließen diese Vorlesung mit der Einführung des "Gradienten". Das ist ein Vektor, der für die Ableitung von f in x eine Interpretation gibt, sofern f eine reellwertige Abbildung auf $\mathbb{R}^N$ ist.

Bildet f eine Umgebung von $\hat{x} \in \mathbb{R}^N$ nach $\mathbb{R}$ ab, und ist in $\hat{x}$ differenzierbar, so gilt für alle $\xi \in \mathbb{R}^N$

$$f'(\hat{x})\xi = (\partial_1 f(\hat{x}), \partial_2 f(\hat{x}), \ldots, \partial_N f(\hat{x})) \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_N \end{bmatrix} \tag{9.29}$$

Diese Multiplikation einer $(1 \times N)$–Matrix mit einem Vektor $\xi \in \mathbb{R}^N$ ergibt ein Element aus $\mathbb{R}^1$, also eine Zahl. Nun – und das ist direkt einsichtig – läßt sich (9.29) auch schreiben in der Form

$$f'(\hat{x})\xi = \langle \begin{pmatrix} \partial_1 f(\hat{x}) \\ \vdots \\ \partial_N f(\hat{x}) \end{pmatrix} , \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_N \end{pmatrix} \rangle \quad ,$$

wobei $\langle \cdot , \cdot \rangle$ das übliche Skalarprodukt in $\mathbb{R}^N$ bezeichnet.

Definition 9.46 *Ist $f : \Omega$ (offen) $\subset \mathbb{R}^N \longrightarrow \mathbb{R}$ differenzierbar in $\hat{x} \in \Omega$, so heißt der Vektor*

$$\nabla f(\hat{x}) := \begin{bmatrix} \partial_1 f(\hat{x}) \\ \partial_2 f(\hat{x}) \\ \vdots \\ \partial_N f(\hat{x}) \end{bmatrix}$$

der Gradient von f in $\hat{x}$.

Wir haben gesehen, dass $f'(\hat{x})\xi$ nichts anderes ist, als die Richtungsableitung von f in Richtung von ξ. Ist nun ξ ein Einheitsvektor, so liefert die Cauchy-Schwarzsche Ungleichung (Korollar 6.24 mit $p = q = 2$) die Abschätzung

$$f'(\hat{x})\xi \leq |\nabla f(\hat{x})| \ . \tag{9.30}$$

Ist $\nabla f(\hat{x}) \neq 0$, so gilt in Abschätzung (9.30) genau dann Gleichheit, wenn $\xi = |\nabla f(\hat{x})|^{-1} \nabla f(\hat{x})$ ist. Also weist der Vektor $\nabla f(\hat{x})$ in die Richtung der größten Richtungsableitung, und sein Betrag ist die größte Richtungsableitung.

Literaturverzeichnis

[B:] **Begleitende Literatur:**

[1] M. Barner, F. Flohr: Analysis I, II; De Gruyter, Berlin (2000, 1996)

[2] C. Blatter: Analysis I, II; Springer; Berlin (1991, 1992)

[3] T. Bröcker: Analysis I, II; Spektrum, Heidelberg (1995)

[4] H. Heuser: Lehrbuch der Analysis, Teil 1, 2; Teubner, Stuttgart (2000)

[5] S. Hildebrandt: Analysis I, II; Springer; Berlin (2002, 2003)

[6] W. Kaballo: Einführung in die Analysis I, II; Spektrum, Heidelberg (2000, 1997)

[7] K. Königsberger: Analysis I, II; Springer, Berlin (2001, 2002)

[8] S. Lang: Analysis; Inter European Ed., Amsterdam (1977)

[9] W. Rudin: Analysis; Physik – Verlag, Weinheim (1980)

[10] K. T. Smith: Primer of Modern Analysis; Springer, New York (1983)

[11] W. Walter: Analysis I, II; Springer, Berlin (2001, 2002)
Dieses Buch zeigt auch die historische Entwicklungen der Analysis auf

[E:] **Ergänzende Literatur:**

[12] H. Amann, J. Escher: Analysis I, II: Birkhäuser, Basel (1998, 1999)

[13] R. Courant, M. Robbins: Was ist Mathematik; Springer, Berlin (1973)

[14] H. Ebbinghaus u.a.: Zahlen; Springer, Berlin (1992)

[15] H. König: Analysis; Birkhäuser, Basel (1984)

[16] H. L. Royden: Real Analysis; McMillan, New York (1968)

[17] W. Rudin: Real and Complex Analysis; McGraw–Hill, New York (1987)

[18] K. R. Stromberg: Introduction to Classical Real Analysis; Wadsworth, Belmond, Ca. (1981)

[H:] **Historisches:**

[19] R. Courant: Vorlesung über Differential– und Integralrechnung, 1. Funktionen einer Veränderlichen; Springer, Berlin (1971)

[20] R. Dedekind: Vorlesung über Differential– und Integralrechnung; Vieweg, Braunschweig (1985)
Die Vorlesungen stammen aus den Jahren 1861/1862

[21] O. Toeplitz: Die Entwicklung der Infinitesimalrechnung; Springer, Berlin