

Studienbücherei



S. Brehmer · H. Apelt Analysis

I. Folgen, Reihen, Funktionen



VEB Deutscher Verlag der Wissenschaften

Mathematik für Lehrer

Band 4

Herausgegeben von:

W. Engel, S. Brehmer, M. Schneider, H. Wussing

Unter Mitarbeit von:

**G. Asser, J. Böhm, J. Flachsmeyer, G. Geise, T. Glocke,
K. Härtig, G. Kasdorf, O. Krötenheerdt, H. Lugowski,
P. H. Müller, G. Porath**

Studienbücherei

Analysis

I. Folgen, Reihen, Funktionen

S. Brehmer
H. Apelt

Mit 65 Abbildungen

Mit historischen Anmerkungen
von H.-J. Ilgauds

Vierte Auflage



VEB Deutscher Verlag
der Wissenschaften
Berlin 1985

Verlagslektor: Erika Arndt
Verlagshersteller: Birgit Burkhardt
Umschlaggestaltung: Rudolf Wendt
© 1974 und 1982 VEB Deutscher Verlag der Wissenschaften, DDR - 1080 Berlin, Postfach 1216
Lizenz-Nr. 206 - 435/56/85
Printed in the German Democratic Republic
Satz: VEB Druckhaus „Maxim Gorki“, DDR - 7400 Altenburg
Offsetdruck und buchbinderische Verarbeitung: VEB Druckerei „Thomas Müntzer“,
DDR - 5820 Bad Langensalza
LSV 1034
Bestellnummer: 570 187 i
01280

Vorwort zur dritten Auflage

Die Bände 4 und 5 der Reihe „Mathematik für Lehrer“ behandeln die Grundlagen der Analysis. Entsprechend ihrem Verwendungszweck unterscheiden sie sich in ihrer Anlage von einem Lehrbuch der Analysis für Diplommathematiker. So werden z. B. die weittragenden Hilfsmittel der Differential- und Integralrechnung erst im Teil Analysis II eingesetzt. Die Lehre von den elementaren Funktionen wird bereits im ersten Teil weitgehend entwickelt, zumal sie auch im Schulunterricht ohne Verwendung der Differential- und Integralrechnung behandelt werden muß. Die wichtigsten Eigenschaften der Potenz-, Exponential- und Logarithmusfunktionen, die im Schulunterricht bereits in der neunten und zehnten Klasse auftreten, werden sogar ohne Verwendung des Grenzwertbegriffs hergeleitet. Wir stützen uns hierbei nur auf die Existenz der kleinsten oberen Schranke und auf die Monotonieeigenschaften dieser Funktionen. Die vorliegende dritte Auflage enthält auch eine Einführung sowie eine erweiterte Behandlung der trigonometrischen Funktionen. Die Beschränkung auf elementare Hilfsmittel erfordert allerdings einen höheren Aufwand. Aus diesem Grunde werden an späteren Stellen unter Verwendung des Grenzwertbegriffs sowie der Reihenlehre andere Möglichkeiten für die Einführung der elementaren Funktionen aufgezeigt.

Trotz einer inhaltlichen Abstimmung zwischen den einzelnen Bänden der Reihe „Mathematik für Lehrer“ lassen sich gewisse Überschneidungen nicht vermeiden. Insbesondere werden im Kapitel 1 Begriffe und Sätze behandelt, die auch in anderen Bänden auftreten. Das dem Grenzwertbegriff gewidmete Kapitel 2 stellt den Hauptteil des ersten Bandes dar. In ihm werden Grenzwerte von Folgen, Reihen, Funktionen und Funktionenfolgen bzw. -reihen behandelt. Das wesentliche Anliegen des Lehrbuches besteht in einer Vertiefung bzw. Fundierung des Stoffes, der im Lehrgang Analysis der allgemeinbildenden und erweiterten polytechnischen Oberschule behandelt wird. Darüber hinaus gilt es, auch dem Lehrerstudenten einen gewissen Einblick in moderne Denkweisen der Analysis zu geben und einen ausreichenden Vorlauf für die Phase des Fachstudiums zu schaffen. Aus diesem Grunde wird in 1.5. der mehrdimensionale euklidische Raum und seine Metrik eingeführt. Die eindimensionalen Betrachtungsweisen werden regelmäßig durch kurze Ergänzungen auf den mehrdimensionalen bzw. auf den komplexen Fall ausgedehnt. Trotzdem ist, wenn man von dem nur in 1.5. behandelten Umgebungsbegriff absieht, eine zusammenhängende Lektüre des Buches unter Auslassung aller den mehrdimensionalen Fall betreffenden Abschnitte und Einführungen möglich.

Der Stoff ist vorwiegend unter systematischen Gesichtspunkten angeordnet. Die

Anordnung kann weitgehend variiert werden. So wird es sich zum Beispiel empfehlen, gewisse Sätze mit komplizierten Beweisen erst dann eingehender durcharbeiten, wenn diese explizit angewendet werden.

Der Banachsche Fixpunktsatz ist in einer so elementaren Form behandelt, wie er auch in den Schulunterricht Eingang finden könnte. Die wichtigsten Anwendungen dieses Satzes erfolgen erst im zweiten Teil in Zusammenhang mit der numerischen Auswertung grundlegender Existenzsätze (Nullstellenberechnung, Umkehrfunktionen implizit definierter Funktionen, Lösungen von Differentialgleichungen).

Die dritte Auflage von MfL 4 unterscheidet sich von den vorhergehenden Auflagen in folgenden Punkten:

- Die Aufgaben wurden gestrichen, da gesonderte Aufgabenbände erschienen sind.
- Der frühere Abschnitt 2.5. wurde in die Abschnitte 1.6. (Die trigonometrischen Funktionen) und 2.5. (Die hyperbolischen Funktionen und die Area-Funktionen) aufgegliedert.
- Der umfassende Fortsetzungssatz über monotone Funktionen in Abschnitt 1.3.5. ist in mehrere Sätze aufgegliedert. Der Abschnitt 1.4.2. über Exponentialfunktionen wurde überarbeitet.
- Der Begriff der kompakten Menge wird in 2.1.7., Definition 3, Satz 4 und Satz 5, sowie in 2.4.2., Satz 1, etwas ausführlicher dargelegt.
- Der Begriff des Limes inferior bzw. Limes superior wird in 2.1.6. in einer anderen Form eingeführt.
- Für die Bestimmung des Infimums bzw. Supremums von Mengen werden in 1.1.3. bzw. 2.1.3., Satz 8, praktikable Verfahren hergeleitet.

Für wertvolle Hinweise haben wir einer Reihe von Kollegen, insbesondere Herrn Prof. Dr. G. ASSER, sehr herzlich zu danken. Unser Dank gilt auch den Mitarbeitern des VEB Deutscher Verlag der Wissenschaften für die gute Zusammenarbeit und der Druckerei für ihre sorgfältige Arbeit.

Potsdam, im Herbst 1982

S. BREHMER
H. APELT

Inhalt

1.	Einige grundlegende Begriffsbildungen der Analysis	9
1.1.	Reelle Zahlen	9
1.1.1.	Der Körper der reellen Zahlen (Teilbereiche)	9
1.1.2.	Ordnung reeller Zahlen (Rechnen mit Ungleichungen)	10
1.1.3.	Obere und untere Schranken von Mengen reeller Zahlen	15
1.1.4.	Absolute Beträge	18
1.1.5.	Wurzeln	19
1.2.	Komplexe Zahlen	21
1.2.1.	Der Körper der komplexen Zahlen	21
1.2.2.	Das Rechnen mit komplexen Zahlen	23
1.2.3.	Quadratwurzel komplexer Zahlen	24
1.3.	Funktionen	27
1.3.1.	Der Funktionsbegriff	27
1.3.2.	Funktionen aus der Menge bzw. in die Menge der reellen bzw. komplexen Zahlen	29
1.3.3.	Verknüpfungen von Funktionen	32
1.3.4.	Ganzrationale, rationale und algebraische Funktionen	34
1.3.5.	Monotone Funktionen	39
1.4.	Potenz-, Exponential- und Logarithmusfunktionen	42
1.4.1.	Potenzen mit rationalen Exponenten	42
1.4.2.	Exponentialfunktionen	47
1.4.3.	Logarithmusfunktionen	51
1.5.	Der euklidische Raum \mathbb{R}^p	55
1.5.1.	Der zweidimensionale euklidische Raum	55
1.5.2.	Der p -dimensionale euklidische Raum	56
1.5.3.	Die Metrik des Raumes \mathbb{R}^p	60
1.5.4.	Spezielle Punkte von Teilmengen des Raumes \mathbb{R}^p	62
1.5.5.	Offene und abgeschlossene Mengen	65
1.5.6.	Normierte Räume und metrische Räume	67
1.6.	Die trigonometrischen Funktionen	67
1.6.1.	Die Funktionalgleichungen der trigonometrischen Funktionen	67
1.6.2.	Die Arkusfunktion	73
1.6.3.	Trigonometrische Darstellung komplexer Zahlen. Polar- und Zylinderkoordinaten	79
1.6.4.	Einheitskreis und Bogenlänge	82
2.	Der Grenzwertbegriff	89
2.1.	Zahlenfolgen	89
2.1.1.	Spezielle Eigenschaften von Zahlenfolgen	89
2.1.2.	Nullfolgen	92

2.1.3.	Konvergente Zahlenfolgen	98
2.1.4.	Monotone Zahlenfolgen, Intervallschachtelungen	106
2.1.5.	Darstellung des natürlichen Logarithmus und der speziellen Exponentialfunktion durch monotone Zahlenfolgen	111
2.1.6.	Häufungswerte	117
2.1.7.	Fundamentalfolgen	121
2.1.8.	Die Sätze von CANTOR und HEINE-BOREL	124
2.2.	Reihen	126
2.2.1.	Grundbegriffe der Reihenlehre	126
2.2.2.	Konvergenz von Reihen	129
2.2.3.	Alternierende Reihen	134
2.2.4.	Absolute Konvergenz	136
2.2.5.	Bedingte und unbedingte Konvergenz	142
2.2.6.	Großer Umordnungssatz	145
2.2.7.	Multiplikation unendlicher Reihen	148
2.2.8.	Die Exponentialreihe	150
2.3.	Stetigkeit und Grenzwerte von Funktionen	154
2.3.1.	Stetigkeit (Folgendefinition)	154
2.3.2.	Stetigkeit (Umgebungsdefinition)	157
2.3.3.	Stetigkeit von verknüpften Funktionen	161
2.3.4.	Grenzwerte von Funktionen	163
2.3.5.	Einseitige Grenzwerte und einseitige Stetigkeit reeller Funktionen	168
2.4.	Stetige Funktionen	170
2.4.1.	Stetigkeit in Punktmengen	170
2.4.2.	Auf abgeschlossenen Mengen stetige Funktionen	172
2.4.3.	Gleichmäßige Stetigkeit	174
2.4.4.	Nullstellen ganzzahliger Funktionen	175
2.4.5.	Der Banachsche Fixpunktsatz	181
2.5.	Die hyperbolischen Funktionen und die Area-Funktionen	184
2.5.1.	Die hyperbolischen Funktionen	184
2.5.2.	Die Area-Funktionen	187
2.6.	Folgen und Reihen von Funktionen	189
2.6.1.	Punktweise Konvergenz	189
2.6.2.	Gleichmäßige Konvergenz	191
2.6.3.	Chebyshev-Metrik	194
2.7.	Approximation von Funktionen	195
2.7.1.	Das Approximationsproblem	195
2.7.2.	Interpolation	196
2.7.3.	Gleichmäßige Approximation	198
	Namen- und Sachverzeichnis	202

1. Einige grundlegende Begriffsbildungen der Analysis

1.1. Reelle Zahlen

1.1.1. Der Körper der reellen Zahlen (Teilbereiche)

In der Menge \mathbf{R} der reellen Zahlen sind zwei zweistellige assoziative und kommutative Operationen, die *Addition* und die *Multiplikation*, definiert. Diese Operationen sind miteinander durch das Distributivgesetz verknüpft. Die Addition ist uneingeschränkt umkehrbar, d. h., für alle reellen Zahlen a, b besitzt die Gleichung $a + x = b$ (sogar genau) eine Lösung, und es gibt (genau) eine reelle Zahl 0 mit $a + 0 = a$ für alle a . Die Multiplikation ist mit einer Einschränkung umkehrbar, und zwar gibt es zu jeder von 0 verschiedenen reellen Zahl a und zu jeder reellen Zahl b (genau) eine reelle Zahl x mit $ax = b$, und es gibt genau eine reelle Zahl 1 mit $a \cdot 1 = a$ für alle a . Auf Grund dieser Eigenschaften bildet die Menge \mathbf{R} aller reellen Zahlen mit den genannten Operationen einen *Körper*.

Das Produkt einer beliebigen reellen Zahl a mit der Zahl 0 ergibt stets 0. Das Produkt zweier von 0 verschiedener reeller Zahlen a, b kann niemals 0 sein, da sonst die Gleichung $ax = 0$ die voneinander verschiedenen Lösungen b und 0 hätte.

Ein Produkt von endlich vielen Faktoren ist dann und nur dann gleich 0, wenn wenigstens einer der Faktoren gleich 0 ist. Die Sonderstellung der Zahl 0 erfordert beim Rechnen mit Variablen ganz besondere Aufmerksamkeit. Aus einer Gleichung der Form $ac = bc$ darf nicht auf $a = b$, sondern nur auf $a = b$ oder $c = 0$ geschlossen werden, und für eventuell notwendige weitere Untersuchungen ist durch Fallunterscheidung weiter zu schließen. Aus der Gleichung

$$6x(x - 3) = (x + 5)(x - 3)$$

folgt z. B. $6x = x + 5$ oder $x - 3 = 0$, d. h. $x = 1$ oder $x = 3$.

Die Menge \mathbf{N} der natürlichen Zahlen ist die kleinste Teilmenge von \mathbf{R} , die die Zahl 0 und mit jeder Zahl n auch die Zahl $n + 1$ enthält. Das hiermit im Zusammenhang stehende Beweisverfahren der vollständigen Induktion sowie den Rechtfertigungssatz für induktive Definitionen setzen wir als bekannt voraus, desgleichen das (auf induktiven Definitionen beruhende) Rechnen mit dem Summen- und Produktzeichen.

Die Menge \mathbf{Z} der ganzen Zahlen ist die Menge aller Zahlen $\pm n$, wobei n eine natürliche Zahl ist. Die Menge aller Quotienten $\frac{m}{n}$ ganzer Zahlen mit $n \neq 0$ ist die Menge \mathbf{Q} der rationalen Zahlen.

Die n -te Potenz x^n einer reellen Zahl x wird induktiv durch

$$x^0 := 1, \tag{1}$$

$$x^{n+1} := x^n x \quad (n \in \mathbf{N}) \tag{2}$$

definiert. Es gelten die Potenzgesetze

$$(xy)^m = x^m y^m, \quad (3)$$

$$x^m x^n = x^{m+n}, \quad (4)$$

$$(x^m)^n = x^{mn}. \quad (5)$$

Für reelle Zahlen x mit $x \neq 0$ wird ferner

$$x^{-n} := \frac{1}{x^n} \quad (n = 1, 2, \dots)$$

gesetzt. Die Potenzgesetze (3), (4), (5) bleiben dann im Fall $x, y \neq 0$ für alle ganzen Zahlen m, n gültig.

Die *binomische Formel* lautet

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (x, y \in \mathbb{R}, n \in \mathbb{N}). \quad (6)$$

Für zahlreiche Beweise der Analysis ist auch die für alle reellen Zahlen x, y gültige Identität

$$y^n - x^n = (y - x) \sum_{k=1}^n y^{n-k} x^{k-1} \quad (n = 1, 2, \dots) \quad (7)$$

nützlich. Der Beweis ergibt sich aus

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

denn die Summanden der letzten beiden Summen stimmen für $k \neq 0$ bzw. für $k \neq n$ überein.

1.1.2. Ordnung reeller Zahlen (Rechnen mit Ungleichungen)

In der Menge \mathbb{R} der reellen Zahlen ist eine (irreflexive) Ordnungsrelation $a < b$ definiert, d. h., es gilt stets

$$\neg a < a \quad (\text{Irreflexivität}), \quad (1)$$

$$a < b \wedge b < c \Rightarrow a < c \quad (\text{Transitivität}), \quad (2)$$

$$a < b \vee a = b \vee b < a \quad (\text{Konnexität}). \quad (3)$$

Für die Ordnungsrelation sind das *Monotoniegesetz der Addition*

$$a < b \Rightarrow a + c < b + c \quad (4)$$

und das *Monotoniegesetz der Multiplikation*

$$a < b \wedge 0 < c \Rightarrow ac < bc \quad (5)$$

erfüllt. Auf Grund dieser Eigenschaften wird \mathbf{R} zu einem *geordneten Körper*.

Der Gebrauch der Zeichen $>$, \leq , \geq sowie der Begriffe *positiv*, *negativ*, *nicht-positiv* und *nichtnegativ* wird als bekannt vorausgesetzt.

Die Menge aller nichtnegativen Zahlen eines Zahlbereichs bezeichnen wir mit einem tiefgestellten Index $+$. So ist z. B. $\mathbf{Z}_+ = \mathbf{N}$. Soll die Zahl 0 im betrachteten Zahlbereich nicht enthalten sein, so kennzeichnen wir dies durch einen hochgestellten Stern. So ist z. B. $\mathbf{N}^* = \mathbf{N} \setminus \{0\}$.

Mit Hilfe der Ordnungsrelation definieren wir die folgenden Arten von Intervallen mit den *Begrenzungspunkten* a, b , wobei wir zunächst $a, b \in \mathbf{R}$ und $a < b$ voraussetzen:

$$]a, b[:= \{x \in \mathbf{R} : a < x < b\}, \quad (6)$$

$$[a, b[:= \{x \in \mathbf{R} : a \leq x < b\}, \quad (7)$$

$$]a, b] := \{x \in \mathbf{R} : a < x \leq b\}, \quad (8)$$

$$[a, b] := \{x \in \mathbf{R} : a \leq x \leq b\}. \quad (9)$$

Lassen wir im letzten Fall zu, daß $a = b$ ist, so entartet das Intervall zur Einermenge $\{a\}$. In den anderen Fällen gehen die Intervalle im Fall $a = b$ in die leere Menge über.

Wir ergänzen nun die Menge der reellen Zahlen durch die *uneigentlichen Zahlen* $-\infty$, ∞ und vereinbaren, daß für $x \in \mathbf{R}$ stets $-\infty < x < \infty$ gelten soll. In den Fällen (6), (8) darf dann auch $a = -\infty$, in den Fällen (6), (7) auch $b = \infty$ gewählt werden.

Für $a, b \in \mathbf{R}$ heißen die Intervalle *beschränkt*, für $a = -\infty$ bzw. $b = \infty$ dagegen *unbeschränkt*. Für alle $a, b \in \mathbf{R}$ sind hiernach die Intervalle

$$]a, \infty[= \{x \in \mathbf{R} : a < x\}, \quad (10)$$

$$[a, \infty[= \{x \in \mathbf{R} : a \leq x\}, \quad (11)$$

$$]-\infty, b[= \{x \in \mathbf{R} : x < b\}, \quad (12)$$

$$]-\infty, b] = \{x \in \mathbf{R} : x \leq b\} \quad (13)$$

und $]-\infty, \infty[= \mathbf{R}$ unbeschränkt. Die Intervalle (6), (8), (10), heißen linksseitig offen, die Intervalle (7), (9), (11) linksseitig abgeschlossen. Alle Intervalle (6) bis (11) sind linksseitig beschränkt, die restlichen linksseitig unbeschränkt usw.

Es sei M eine Teilmenge von \mathbf{R} . Wir sagen, daß eine Teilmenge $M_0 \subseteq M$ *dicht* in M liegt, wenn es zu zwei Zahlen $a, b \in M$ mit $a < b$ stets ein $z \in M_0$ mit $a < z < b$ gibt. Offensichtlich gibt es dann auch ein $z_1 \in M_0$ mit $a < z_1 < z$, und folglich liegen im Intervall $]a, b[$ mit $a < b$ sogar unendlich viele Zahlen aus M_0 . So ist z. B. die Menge \mathbf{Q}_+ dicht in \mathbf{R}_+ (vgl. MFL 2, 6.2.1.).

Die folgenden Sätze über Ungleichungen dürfen als bekannt vorausgesetzt werden.

Satz 1. Für alle reellen Zahlen a, b gilt

$$a < b \Leftrightarrow a - b < 0, \quad (14)$$

$$a < b \Leftrightarrow -a > -b. \quad (15)$$

Satz 2. Eine reelle Zahl a ist genau dann positiv bzw. negativ, wenn $-a$ negativ bzw. positiv ist.

Satz 3. Das Produkt zweier reeller Zahlen ist genau dann positiv, wenn entweder beide positiv oder beide negativ sind.

Die Monotoniegesetze der Addition und Multiplikation können auf die Relation \leq übertragen werden, d. h., für alle reellen Zahlen a, b, c gilt

$$a \leq b \Rightarrow a + c \leq b + c,$$

$$a \leq b \wedge c \geq 0 \Rightarrow ac \leq bc.$$

Die Beweise werden durch Fallunterscheidung geführt.

Satz 4. Für alle reellen Zahlen a, b, c, d gilt

$$a < b \wedge c < d \Rightarrow a + c < b + d, \quad (16)$$

$$0 < a < b \wedge 0 < c < d \Rightarrow ac < bd. \quad (17)$$

Wir bemerken, daß Satz 4 auch dann gilt, wenn das Zeichen $<$ überall durch \leq ersetzt wird. Dagegen darf man aus $a < b$ und $c < d$ nicht auf $a - c < b - d$ schließen, wie das Beispiel $a = 7, b = 8, c = 2, d = 6$ zeigt.

Satz 5. Für alle reellen Zahlen a, b, c gilt

$$a < b \wedge c < 0 \Rightarrow ac > bc, \quad (18)$$

$$0 < a < b \Rightarrow \frac{1}{a} > \frac{1}{b} > 0. \quad (19)$$

Wir beweisen einige weitere Sätze über Ungleichungen.

Satz 6. Für alle reellen Zahlen a, b ist

$$ab \leq \frac{a^2 + b^2}{2}. \quad (20)$$

Beweis. Die Ungleichung (20) ist mit den folgenden Ungleichungen

$$2ab \leq a^2 + b^2, \quad (21)$$

$$0 \leq a^2 - 2ab + b^2, \quad (22)$$

$$0 \leq (a - b)^2 \quad (23)$$

äquivalent. Die letzte Ungleichung ist aber für alle reellen Zahlen a, b erfüllt, womit der Satz bewiesen ist.

Satz 7. Für alle reellen Zahlen a, b gilt

$$ab \leq \left(\frac{a+b}{2} \right)^2. \quad (24)$$

Beweis. Aus $0 \leq (a-b)^2 = a^2 - 2ab + b^2$ folgt, wenn wir auf beiden Seiten $4ab$ addieren, $4ab \leq a^2 + 2ab + b^2 = (a+b)^2$. Das ist mit der Behauptung gleichwertig.

Satz 8. Für nichtnegative reelle Zahlen x, y und für $n = 1, 2, \dots$ gilt genau dann $x^n = y^n$ bzw. $x^n < y^n$, wenn $x = y$ bzw. $x < y$ ist.

Beweis. Für $x = 0$ oder $y = 0$ ist die Behauptung richtig. Für positive x, y ist die rechts in 1.1.1.(7) stehende Summe stets positiv, so daß $y^n - x^n$ und $y - x$ beide positiv oder beide negativ oder beide gleich 0 sind. Damit ist Satz 8 bewiesen.

Satz 9. Für alle nichtnegativen reellen Zahlen x, y mit $x \neq y$ und für alle natürlichen Zahlen n mit $n > 1$ ist

$$nx^{n-1}(y-x) < y^n - x^n < ny^{n-1}(y-x). \quad (25)$$

Beweis. Aus 1.1.1.(7) folgt wegen $\sum_{k=1}^n a = na$ stets

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

Die Produkte $(y-x)(y^{n-k} - x^{n-k})$ sind für nichtnegative reelle Zahlen x, y mit $x \neq y$ wegen Satz 8 positiv, und es folgt

$$y^n - x^n - nx^{n-1}(y-x) > 0.$$

Vertauschen wir x und y , so erhalten wir

$$x^n - y^n - ny^{n-1}(x-y) > 0$$

oder

$$y^n - x^n - ny^{n-1}(y-x) < 0.$$

Aus den bewiesenen Ungleichungen können wir unsere Behauptung ablesen.

Mit ihrer Hilfe beweisen wir die folgende Verallgemeinerung von (20).

Satz 10. Für alle von 0 verschiedenen natürlichen Zahlen m, n und für alle nichtnegativen reellen Zahlen x, y mit $x \neq y$ ist

$$(m+n)x^m y^n < mx^{m+n} + ny^{m+n}. \quad (26)$$

Für $m = 0$ oder $n = 0$ oder $x = y$ geht die Ungleichung (26) in eine Gleichung über.

Beweis. Die letzte Behauptung des Satzes kann leicht unmittelbar nachgeprüft werden. Es sei nun $m, n > 0$, $x \neq y$. Da die Ungleichung (26) unverändert bleibt, wenn wir x mit y und zugleich m mit n vertauschen, genügt es, den Fall $x < y$ zu betrachten. Für $x = 0$ ist (26) erfüllt. Es sei $0 < x < y$. Wegen (25) ist

$$\frac{y^n - x^n}{y^m - x^m} < \frac{ny^{n-1}(y-x)}{mx^{m-1}(y-x)} = \frac{ny^{n-1}}{mx^{m-1}} < \frac{ny^n}{mx^m},$$

$$mx^m(y^n - x^n) < ny^n(y^m - x^m),$$

und diese Ungleichung ist mit der Behauptung (26) gleichwertig.

Eine weitere Verallgemeinerung gibt

Satz 11. Für alle von 0 verschiedenen ganzen Zahlen m, n mit $m + n \neq 0$ und für alle positiven reellen Zahlen x, y mit $x \neq y$ ist

$$\frac{1}{mn} x^m y^n < \frac{1}{n(m+n)} x^{m+n} + \frac{1}{m(m+n)} y^{m+n}. \quad (27)$$

Beweis. Es sei $m + n > 0$. Dann ist $m > 0$ oder $n > 0$, und durch Vertauschung von x bzw. m mit y bzw. n können wir erreichen, daß $m > 0$ ist. Ist auch $n > 0$, so ist (27) mit (26) äquivalent. Ist $n < 0$, so setzen wir $p = m + n$, $q = -n$ und führen (27) durch Multiplikation mit $-mn(m+n)y^{-n} = pq(p+q)y^q$ in die äquivalente Ungleichung

$$-px^{p+q} < -(p+q)x^p y^q + qy^{p+q}$$

über, deren Richtigkeit wieder aus (26) folgt. Den Fall $m + n < 0$ führen wir auf den bereits bewiesenen Fall zurück, indem wir x, y, m, n durch $x^{-1}, y^{-1}, -m, -n$ ersetzen.

Satz 12. Für alle reellen Zahlen h mit $1 + h > 0$ und für alle natürlichen Zahlen n gilt die Bernoullische Ungleichung

$$(1 + h)^n \geq 1 + nh. \quad (28)$$

Das Gleichheitszeichen gilt nur, wenn $n = 0, 1$ oder $h = 0$ ist.

Beweis. Für $h = 0$ oder $n = 0, 1$ gilt (28) mit dem Gleichheitszeichen. Im Fall $n > 0$, $h \neq 0$ setzen wir $y = 1 + h$, $x = 1$ in (25) und erhalten $nh < (1 + h)^n - 1$.

Die Ungleichung (28) tritt erstmals 1689 in der ersten von fünf als „Reihendisertationen“ bezeichneten Arbeiten von JAKOB BERNOULLI (1654–1705) über unendliche Reihen auf. Sie kann auch leicht direkt durch vollständige Induktion bewiesen werden.

1.1.3. Obere und untere Schranken von Mengen reeller Zahlen

Mit Hilfe der Ordnungsrelation kann für Mengen reeller Zahlen eine Reihe von Begriffsbildungen definiert werden, die für die Analysis von Bedeutung sind.

Definition. Eine Menge M reeller Zahlen heißt *nach oben beschränkt*, wenn eine reelle Zahl a mit folgender Eigenschaft existiert: Für alle reellen Zahlen x mit $x \in M$ gilt $x \leq a$. Jede reelle Zahl a mit dieser Eigenschaft heißt eine *obere Schranke* der Menge M :

$$a \text{ obere Schranke von } M : \Leftrightarrow \bigwedge_{x \in M} x \leq a.$$

Alle bisher bewiesenen Aussagen über reelle Zahlen gelten unverändert für den Körper \mathbb{Q} der rationalen Zahlen. Dagegen ist das nachfolgende *Grundgesetz der Stetigkeit* für den Körper \mathbb{R} der reellen, aber nicht für den Körper \mathbb{Q} der rationalen Zahlen erfüllt.

Satz. Jede nichtleere nach oben beschränkte Menge M reeller Zahlen besitzt eine kleinste obere Schranke, d. h., es gibt eine obere Schranke a derart, daß jede obere Schranke b von M die Ungleichung $a \leq b$ erfüllt.

Wegen der Gültigkeit dieses Satzes¹⁾ bildet \mathbb{R} bezüglich der Addition, Multiplikation und Ordnung einen *stetig geordneten Körper*.

Aus dem Satz ergibt sich sofort, daß jede nach oben beschränkte Menge reeller Zahlen genau eine kleinste obere Schranke besitzt. Sind nämlich a_1, a_2 kleinste obere Schranken von M , so ist $a_1 \leq a_2$ und $a_2 \leq a_1$, also $a_1 = a_2$. Die kleinste obere Schranke einer nach oben beschränkten Menge M reeller Zahlen heißt auch die *obere Grenze* oder das *Supremum* von M , in Zeichen $\sup M$.

Beispiel. Es seien b_0, b_1, b_2, \dots natürliche Zahlen, und für $k \geq 1$ sei $0 \leq b_k \leq 9$. Mit M bezeichnen wir die Menge aller rationalen Zahlen

$$d_k := b_0.b_1 \dots b_k = b_0 + \frac{b_1}{10} + \dots + \frac{b_k}{10^k}$$

mit $k \in \mathbb{N}$. Wir werden zeigen, daß es eine reelle Zahl b mit

$$b_0.b_1 \dots b_k \leq b \leq b_0.b_1 \dots b_k + 10^{-k}$$

gibt. Hierzu setzen wir $d_k' := d_k + 10^{-k}$. Wegen

$$d_k' - d_{k+1}' = \frac{1}{10^k} - \frac{1}{10^{k+1}} - (d_{k+1} - d_k) = \frac{9 - b_{k+1}}{10^{k+1}} \geq 0$$

ist

$$d_0 \leq d_1 \leq \dots \leq d_k < d_k' \leq \dots \leq d_{k+1}' \leq d_0',$$

¹⁾ Ein Beweis dieses Satzes wird in MfL 2, 6.2.1., geführt.

und es folgt $d_k \leq d_{k+m} < d'_{k+m} \leq d'_m$ für alle natürlichen Zahlen k, m . Für jede natürliche Zahl m ist also d'_m eine obere Schranke der nichtleeren Menge M . Daher besitzt M eine kleinste obere Schranke b . Es ist also stets $d_k \leq b \leq d'_k$, was wir zeigen wollten.

Wenn keine Neunerperiode auftritt, ist b nach MfL 2, 6.2.1., durch den unendlichen Dezimalbruch $b = b_0, b_1 b_2 \dots$ gegeben. Im Falle einer Neunerperiode ist diese Darstellung in der bekannten Weise zu modifizieren. So ist z. B. $234,317\bar{9} = 234,318$ zu setzen.

Die Aussage $a = \sup M$ ($a \in \mathbb{R}$) ist genau dann erfüllt, wenn a obere Schranke von M ist und wenn zu jeder reellen Zahl y , die kleiner als a ist, eine Zahl x mit $x \in M$ und $x > y$ gefunden werden kann. Dies besagt in formalisierter Schreibweise

$$a = \sup M \Leftrightarrow \bigwedge_{x \in M} x \leq a \wedge \bigvee_{y \in \mathbb{R} : y < a} (y < a \Rightarrow y < x). \quad (1)$$

Das Supremum a einer Menge M braucht nicht Element dieser Menge zu sein. Ist dies aber der Fall, so heißt a das *Maximum* der Menge M , in Zeichen $a = \max M$. Es gilt

$$a = \max M \Leftrightarrow a \in M \wedge \bigwedge_{x \in M} x \leq a, \quad (2)$$

d. h., a ist Maximum der Menge M genau dann, wenn a Element und obere Schranke von M ist.

In Analogie zu den Begriffen „nach oben beschränkt“, „obere Schranke“, „Maximum“, „kleinste obere Schranke“ bzw. „obere Grenze“ bzw. „Supremum“ führt man die Begriffe „nach unten beschränkt“, „untere Schranke“, „Minimum“, „größte untere Schranke“ bzw. „untere Grenze“ bzw. „Infimum“ ein. Die Existenz der größten unteren Schranke einer nichtleeren nach unten beschränkten Menge M reeller Zahlen kann aus dem Grundgesetz der Stetigkeit gefolgert werden, indem man die Menge $M^* := \{x : -x \in M\}$ betrachtet, die auf der Zahlengeraden spiegelbildlich zu M bezüglich O liegt. Abkürzend wird die größte untere Schranke einer nichtleeren nach unten beschränkten Menge M mit $\inf M$ (gelesen: Infimum (von) M) bezeichnet. Besitzt die Menge M ein Minimum, so wird es mit $\min M$ (gelesen: Minimum (von) M) bezeichnet.

Ist eine Menge M nicht nach oben bzw. nach unten beschränkt, so setzen wir $\sup M := \infty$ bzw. $\inf M := -\infty$.

Für jede nichtleere Menge $M \subseteq \mathbb{R}$ ist

$$I^*(M) := \begin{cases}]-\infty, b] & \text{für } b = \max M, \\]-\infty, b[& \text{für } b = \sup M \notin M \end{cases}$$

offensichtlich das kleinste linksseitig unbeschränkte Intervall, das die Menge M umfaßt. Entsprechendes gilt für das Intervall

$$I_*(M) := \begin{cases} [a, \infty[& \text{für } a = \min M, \\]a, \infty[& \text{für } a = \inf M \notin M. \end{cases}$$

Daher ist

$$I(M) := I^*(M) \cap I_*(M)$$

das *kleinste Intervall*, das die Menge M umfaßt.

Wenn es uns also gelingt, das kleinste Intervall zu finden, das eine nichtleere Menge $M \subseteq \mathbb{R}$ enthält, haben wir in dessen Begrenzungspunkten das Infimum und das Supremum von M ermittelt. Das kleinste Intervall, das die Menge

$$M = \left\{ \frac{1}{n} : n \in \mathbb{N}^* \right\}$$

umfaßt, ist z. B. das Intervall $]0, 1]$, und folglich ist $\inf M = 0$, $\sup M = \max M = 1$.

Jede nichtleere Menge natürlicher Zahlen besitzt ein Minimum. Ferner besitzt jede nach oben bzw. unten beschränkte Menge ganzer Zahlen ein Maximum bzw. ein Minimum. Ist x eine reelle Zahl, so ist die Menge aller ganzen Zahlen n mit $n \leq x$ nach oben beschränkt und besitzt ein Maximum, das wir mit $[x]$ bezeichnen:

$$[x] := \max \{n : n \in \mathbb{Z} \wedge n \leq x\}.$$

Die ganze Zahl $[x]$ heißt die *größte Ganze* der reellen Zahl x .

Jede endliche Menge reeller Zahlen besitzt ein Maximum und ein Minimum. Hiermit beweisen wir die für alle natürlichen Zahlen n , alle reellen Zahlen a_0, \dots, a_n und alle positiven reellen Zahlen b_0, \dots, b_n gültige Ungleichung

$$\min \left\{ \frac{a_0}{b_0}, \dots, \frac{a_n}{b_n} \right\} \leq \frac{a_0 + \dots + a_n}{b_0 + \dots + b_n} \leq \max \left\{ \frac{a_0}{b_0}, \dots, \frac{a_n}{b_n} \right\}. \quad (3)$$

Es gibt ein i mit

$$\frac{a_i}{b_i} \leq \frac{a_k}{b_k}, \quad a_i b_k \leq b_i a_k \quad (k = 0, \dots, n),$$

und folglich ist auch

$$a_i \sum_{k=0}^n b_k = \sum_{k=0}^n a_i b_k \leq \sum_{k=0}^n b_i a_k = b_i \sum_{k=0}^n a_k.$$

Wegen $b_i > 0$ und $\sum_{k=0}^n b_k > 0$ folgt

$$\frac{a_i}{b_i} \leq \frac{\sum_{k=0}^n a_k}{\sum_{k=0}^n b_k},$$

womit die erste Ungleichung (3) bewiesen ist. Der Beweis der zweiten Ungleichung verläuft analog.

1.1.4. Absolute Beträge

Für jede reelle Zahl a setzen wir nach KARL WEIERSTRASS (1815–1897)

$$|a| := \begin{cases} a & \text{für } a \geq 0, \\ -a & \text{für } a < 0 \end{cases} \quad (1)$$

und nennen die nichtnegative reelle Zahl $|a|$ den *Betrag* von a . Offenbar gilt

$$|a| = \max \{a, -a\}. \quad (2)$$

Für jede reelle Zahl a setzt man ferner

$$\operatorname{sgn} a := \begin{cases} 1 & \text{für } a > 0, \\ 0 & \text{für } a = 0, \\ -1 & \text{für } a < 0 \end{cases}$$

(gelesen: Signum a). Für alle a gilt dann

$$|a| = a \operatorname{sgn} a, \quad a = |a| \operatorname{sgn} a, \quad (3)$$

und für alle a, b ist

$$\operatorname{sgn} (ab) = \operatorname{sgn} a \cdot \operatorname{sgn} b. \quad (4)$$

Die Rechenregeln

$$|a| = |-a|, \quad (5)$$

$$|a - b| = |b - a|,$$

$$|ab| = |a| \cdot |b|, \quad (6)$$

$$\left| \frac{a}{b} \right| = \frac{|a|}{|b|} \quad (b \neq 0)$$

ergeben sich unmittelbar aus der Definition (1) bzw. aus (2).

Für viele Beweise der Analysis sind die beiden *Dreiecksungleichungen*

$$|a \pm b| \leq |a| + |b|, \quad (7)$$

$$||a| - |b|| \leq |a \pm b| \quad (8)$$

von besonderer Bedeutung. Es genügt, die Behauptungen (7), (8) für das obere der beiden Zeichen zu beweisen. Wegen (5) ergibt sich nämlich die zweite Behauptung, indem wir b durch $-b$ ersetzen. Zum Beweis von (7) gehen wir von den nach (2) gültigen Ungleichungen $\pm a \leq |a|$, $\pm b \leq |b|$ aus und erhalten nach Addition $\pm(a+b) \leq |a| + |b|$. Für eines der beiden Zeichen ergibt sich links $|a+b|$, womit (7) bewiesen

ist. Hieraus folgt weiterhin

$$\begin{aligned} |a| &= |(a+b) - b| \leq |a+b| + |b|, \\ |a| - |b| &\leq |a+b|. \end{aligned} \quad (9)$$

Durch Vertauschung von a und b ergibt sich ebenso

$$|b| - |a| \leq |b+a| = |a+b|. \quad (10)$$

Aus (9) und (10) folgt (8).

Die Ungleichung

$$|a-b| \leq |a-c| + |b-c| \quad (11)$$

wird ebenfalls häufig als Dreiecksungleichung bezeichnet. Sie ergibt sich wie folgt aus (7):

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

Mit Hilfe von Absolutbeträgen kann man das Maximum bzw. Minimum zweier reeller Zahlen formelmäßig angeben. Nach Definition des Absolutbetrages ist $|a-b| = a-b$ für $a \geq b$ und $|a-b| = b-a$ für $a < b$. Es folgt

$$\begin{aligned} \frac{a+b+|a-b|}{2} &= \begin{cases} a & \text{für } a \geq b, \\ b & \text{für } a < b, \end{cases} \\ \frac{a+b-|a-b|}{2} &= \begin{cases} b & \text{für } a \geq b, \\ a & \text{für } a < b, \end{cases} \end{aligned}$$

d. h., es ist

$$\max\{a, b\} = \frac{a+b+|a-b|}{2}, \quad (12)$$

$$\min\{a, b\} = \frac{a+b-|a-b|}{2}. \quad (13)$$

1.1.5. Wurzeln

In der Schulmathematik wird der nachfolgende Satz von der Existenz und Einzigkeit der n -ten Wurzel meist ohne Beweis verwendet.

Satz. Zu jeder nichtnegativen reellen Zahl a und zu jeder von 0 verschiedenen natürlichen Zahl n gibt es genau eine nichtnegative reelle Zahl b mit $b^n = a$.

Beweis. Für $n = 1$ ist a die eindeutig bestimmte Lösung der Gleichung $x^n = a$. Es sei nun $n > 1$. Wir ermitteln die natürlichen Zahlen b_0, b_1, b_2, \dots aus den Bedin-

gungen

$$\begin{aligned} b_0^n &\leq a < (b_0 + 1)^n, \\ (b_0, b_1)^n &\leq a < (b_0, b_1 + 10^{-1})^n \\ &\dots\dots\dots \\ (b_0, b_1 \dots b_k)^n &\leq a < (b_0, b_1 \dots b_k + 10^{-k})^n \end{aligned}$$

und setzen $d_k := b_0, b_1 \dots b_k$, $d_k' := d_k + 10^{-k}$. Nach 1.1.3. gibt es eine reelle Zahl b mit $d_k \leq b \leq d_k'$, also auch

$$d_k^n \leq b^n \leq d_k'^n \quad (k \in \mathbf{N}).$$

Nach Konstruktion ist andererseits $d_k^n \leq a < d_k'^n$ oder

$$-d_k'^n < -a \leq -d_k^n.$$

Addition der beiden letzten Ungleichungen ergibt

$$-(d_k'^n - d_k^n) < b^n - a \leq d_k'^n - d_k^n.$$

Mit 1.1.2.(25) folgt, wenn wir $d_k' \leq d_0'$ beachten,

$$|b^n - a| \leq d_k'^n - d_k^n < n d_k'^{n-1} (d_k' - d_k) \leq n d_0'^{n-1} 10^{-k}.$$

Wäre nun $b^n \neq a$, so könnten wir k so groß wählen, daß $10^k |b^n - a| > n d_0'^{n-1}$ ist, was zu einem Widerspruch führt. Somit ist $b^n = a$. Wegen 1.1.2., Satz 8, kann es keine von b verschiedene Lösung der Gleichung $x^n = a$ geben.

Definition. Ist a eine nichtnegative reelle Zahl, so heißt die eindeutig bestimmte nichtnegative reelle Zahl b mit $b^n = a$ ($n = 1, 2, \dots$) die n -te Wurzel von a . Sie wird mit $\sqrt[n]{a}$ bezeichnet. Die Zahl a heißt der Radikand von $\sqrt[n]{a}$. Statt $\sqrt[n]{a}$ schreibt man kürzer \sqrt{a} .

Nach Definition ist

$$(\sqrt[n]{a})^n = a$$

für alle nichtnegativen reellen Zahlen a , und nach 1.1.2., Satz 8, gilt $a = b$ bzw. $a < b$ für zwei nichtnegative reelle Zahlen a, b genau dann, wenn $\sqrt[n]{a} = \sqrt[n]{b}$ bzw. $\sqrt[n]{a} < \sqrt[n]{b}$ ist. Ferner ist stets $\sqrt[n]{a} = a$.

Man beachte vor allem, daß das Symbol $\sqrt[n]{a}$ ($n = 1, 2, \dots$) nur dann erklärt ist, wenn a nichtnegativ ist. Das ist z. B. gesichert, wenn $a = b^2$, d. h., wenn a das Quadrat einer beliebigen reellen Zahl b ist. Insbesondere gilt stets

$$\sqrt{b^2} = |b|. \quad (2)$$

So ist z. B. $\sqrt{(-3)^2} = \sqrt{9} = 3 = |-3|$.

Für ungerade natürliche Zahlen n hat die Gleichung $x^n = a$ auch im Fall $a < 0$ eine (sogar eindeutig bestimmte) reelle Lösung. Diese wird aber nicht mit $\sqrt[n]{a}$, sondern mit $-\sqrt[n]{-a}$ bezeichnet. Wir sagen also z. B. nicht, „die Gleichung $x^3 = -8$ hat die Lösung $\sqrt[3]{-8}$ “, sondern wir bezeichnen die Lösung mit $-\sqrt[3]{8}$. Die Einführung des Symbols $\sqrt[n]{a}$ für ungerade n und negative reelle Zahlen a ist unzweckmäßig, weil dies für eine Verallgemeinerung der Potenzgesetze hinderlich ist.

Diese Vereinbarung betrifft die reelle Analysis. In der komplexen Analysis, der sogenannten Funktionentheorie, und in der Algebra liegt ein anderer Sachverhalt vor. Hier ist es häufig nicht möglich, in natürlicher Weise unter allen Lösungen einer Gleichung der Form $x^n = a$ eine bestimmte auszuwählen. Das Symbol $\sqrt[n]{a}$ wird dann verwendet, um irgendeine dieser Lösungen zu bezeichnen. Bei Verwendung solcher „mehrdeutigen“ Symbole muß man sorgfältig darauf achten, daß Mißverständnisse vermieden werden.

1.2. Komplexe Zahlen

1.2.1. Der Körper der komplexen Zahlen

Die Tatsache, daß gewisse algebraische oder geometrische Problemstellungen in Zahlbereichen nicht lösbar sind, führte in der Entwicklung der Mathematik zu Erweiterungen dieser Zahlbereiche bis zum Bereich der reellen Zahlen. Eine nochmalige Erweiterung dieses Bereichs wurde zunächst dadurch motiviert, daß nicht alle Gleichungen zweiten Grades im Bereich der reellen Zahlen lösbar sind; dies gilt z. B. für die Gleichung $x^2 + 1 = 0$. Später erwies sich, daß der bei diesem Erweiterungsschritt gewonnene Zahlbereich, der Bereich der *komplexen Zahlen*, für die Mathematik und Naturwissenschaften von außerordentlich weitreichender Bedeutung war. Eine lückenlose Einführung des Begriffs „komplexe Zahl“ erfolgt an anderer Stelle. Hier begnügen wir uns damit, den nachfolgenden Satz ohne Beweis anzugeben.

Satz. *Es gibt einen Zahlbereich \mathbb{C} , dessen Elemente komplexe Zahlen heißen und der folgende Eigenschaften besitzt:*

- Der Bereich der reellen Zahlen ist ein Teilbereich des Bereichs der komplexen Zahlen, d. h., jede reelle Zahl ist auch eine komplexe Zahl.*
- Im Bereich der komplexen Zahlen sind eine Addition und eine Multiplikation definiert, und bezüglich dieser Operationen bildet \mathbb{C} einen Erweiterungskörper von \mathbb{R} .*
- Es gibt eine komplexe Zahl i mit der Eigenschaft*

$$i^2 = -1.^1) \tag{1}$$

¹⁾ Wir vermeiden es, diese Gleichung in der Form $i = \sqrt{-1}$ zu schreiben, denn wir haben das Symbol \sqrt{a} nur für *nichtnegative* reelle Radikanden a definiert.

d) Zu jeder komplexen Zahl z gibt es ein eindeutig bestimmtes Paar (x, y) reeller Zahlen mit

$$z = x + iy.^1) \quad (2)$$

Die reelle Zahl x heißt der *Realteil*, in Zeichen $\operatorname{Re} z$, die reelle Zahl y der *Imaginärteil*, in Zeichen $\operatorname{Im} z$, der komplexen Zahl z .

Wie wir aus (1) ersehen, ist im Bereich der komplexen Zahlen die Gleichung $x^2 + 1 = 0$ lösbar. Auf Grund der Eigenschaft b) kann man mit komplexen Zahlen „im wesentlichen“ ebenso rechnen wie mit reellen Zahlen. Man beachte aber, daß für komplexe Zahlen keine Ordnungsrelation definiert ist.

Wir führen noch folgende Bezeichnung ein. Die komplexe Zahl z heißt *rein imaginär*, wenn $z \neq 0$ und $\operatorname{Re} z = 0$ ist. Für jede komplexe Zahl z gilt

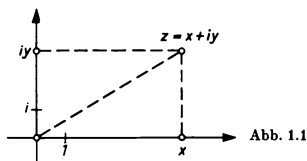
$$z = \operatorname{Re} z + i \operatorname{Im} z.$$

Die komplexe Zahl

$$\bar{z} := \operatorname{Re} z - i \operatorname{Im} z \quad (3)$$

heißt die zu z *konjugiert komplexe Zahl*.²⁾

Während wir reelle Zahlen auf einer Zahlengeraden veranschaulichen, verwenden wir für die Veranschaulichung komplexer Zahlen eine Zahlenebene, die *Gaußsche Zahlenebene* (CARL FRIEDRICH GAUSS (1777–1855)). GAUSS führte die Darstellung der komplexen Zahlen in der Ebene in seiner zweiten Abhandlung über biquadratische Reste 1831 ein.



Der komplexen Zahl z mit $z = x + iy$ ordnen wir nach Wahl eines Koordinatensystems den Punkt $P(x, y)$ zu. Es ist üblich, die Punkte der Gaußschen Zahlenebene mit den komplexen Zahlen zu identifizieren und dementsprechend mit z statt mit $P(x, y)$ zu bezeichnen (Abb. 1.1).

Diese Zuordnung ist umkehrbar eindeutig. Den reellen bzw. nicht reellen komplexen Zahlen entsprechen dabei die Punkte der x -Achse bzw. die Punkte, die nicht

¹⁾ Wegen der Kommutativgesetze für die Addition und Multiplikation können wir auch $z = x + yi = yi + x = iy + x$ schreiben.

²⁾ Neuerdings wird die konjugiert komplexe Zahl von z häufig mit z^* statt \bar{z} bezeichnet.

auf der x -Achse liegen. Den rein imaginären komplexen Zahlen entsprechen die vom Ursprung verschiedenen Punkte der y -Achse. Ist $z = x + iy$, so heißt die reelle Zahl $|z|$ mit

$$|z| := \sqrt{x^2 + y^2} \quad (4)$$

der *Betrag* von z . Offenbar ist $|z|$ der Abstand des Punktes z vom Ursprung. Für reelle komplexe Zahlen stimmt diese Definition mit der Definition von 1.1.4. überein. Es ist $|z| = 0$ genau dann, wenn $z = 0$ ist. Zwei konjugiert komplexe Zahlen z und \bar{z} liegen spiegelbildlich zur x -Achse und haben den gleichen Betrag.

1.2.2. Das Rechnen mit komplexen Zahlen

Sind z_1, z_2 komplexe Zahlen mit $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$, so gilt

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2),$$

d. h., es ist

$$\operatorname{Re}(z_1 + z_2) = \operatorname{Re} z_1 + \operatorname{Re} z_2,$$

$$\operatorname{Im}(z_1 + z_2) = \operatorname{Im} z_1 + \operatorname{Im} z_2.$$

Wegen $i^2 = -1$ ist ferner

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = x_1 x_2 - y_1 y_2 + i(x_1 y_2 + y_1 x_2),$$

d. h., es ist

$$\operatorname{Re}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Re} z_2 - \operatorname{Im} z_1 \operatorname{Im} z_2,$$

$$\operatorname{Im}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Im} z_2 + \operatorname{Im} z_1 \operatorname{Re} z_2.$$

Für $z_1 = z = x + iy$ und $z_2 = \bar{z} = x - iy$ erhalten wir insbesondere

$$z + \bar{z} = 2x, \quad z - \bar{z} = 2iy, \quad z\bar{z} = x^2 + y^2,$$

d. h., es ist

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}, \quad (1)$$

$$|z| = \sqrt{z\bar{z}}, \quad (2)$$

wobei wir das Wurzelsymbol anwenden durften, weil $z\bar{z}$ eine nichtnegative reelle Zahl ist. Wegen $\bar{\bar{z}} = z$ ist offenbar $|z| = |\bar{z}|$.

Für jede von 0 verschiedene komplexe Zahl z gilt

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{1}{|z|^2} \bar{z}. \quad (3)$$

Für $z_2 \neq 0$ ist folglich

$$\frac{z_1}{z_2} = z_1 \frac{1}{z_2} = \frac{1}{|z_2|^2} z_1 \bar{z}_2. \quad (4)$$

Für den Übergang zur konjugierten komplexen Zahl gelten die Rechenregeln

$$\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2, \quad (5)$$

$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad (6)$$

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2} \quad (z_2 \neq 0). \quad (7)$$

Für das Rechnen mit Beträgen gelten die Regeln

$$||z_1| - |z_2|| \leq |z_1 \pm z_2| \leq |z_1| + |z_2|, \quad (8)$$

$$|z_1 z_2| = |z_1| \cdot |z_2|, \quad (9)$$

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|} \quad (z_2 \neq 0). \quad (10)$$

Aus (9) folgt $|z^n| = |z|^n$ ($n \in \mathbf{N}$) durch vollständige Induktion. Wir erläutern die Rechenregeln an einigen Beispielen.

Es sei

$$a = 2 - 5i, \quad b = 3 + 4i, \quad c = \frac{1}{2}i.$$

Dann ist z. B.

$$a + b = 5 - i, \quad a - b = -1 - 9i, \quad \operatorname{Re} a = 2, \quad \operatorname{Im} a = -5,$$

$$\bar{a} = 2 + 5i, \quad |b| = \sqrt{3^2 + 4^2} = 5, \quad \frac{1}{c} = \frac{1}{|c|^2} \bar{c} = -2i, \quad ab = 26 - 7i,$$

$$\frac{a}{b} = \frac{1}{|b|^2} a \bar{b} = \frac{1}{25} (2 - 5i)(3 - 4i) = \frac{1}{25} (-14 - 23i) = -\frac{14}{25} - \frac{23}{25}i.$$

1.2.3. Quadratwurzel komplexer Zahlen

Satz 1. Die Gleichung $z^2 = c$ besitzt für $c = 0$ die eindeutig bestimmte Lösung $z_0 = 0$ und für $c \neq 0$ genau zwei Lösungen z_1, z_2 mit $z_2 = -z_1$.

Beweis. Wir unterscheiden zwei Fälle.

Fall 1. Die komplexe Zahl c ist reell. Ist $c \geq 0$ bzw. $c < 0$, so gilt $c = \sqrt{c^2}$ bzw. $c = (i\sqrt{-c^2})^2$, und die Gleichung $z^2 = c$ ist mit $z^2 = \sqrt{c^2}$ bzw. $z^2 = (i\sqrt{-c^2})^2$, also mit $(z - \sqrt{c})(z + \sqrt{c}) = 0$ bzw. $(z - i\sqrt{-c})(z + i\sqrt{-c}) = 0$ äquivalent. Diese Gleichungen sind genau dann erfüllt, wenn einer der beiden Faktoren gleich Null ist.

Die Gleichung $z^2 = c$ hat somit genau die Lösungen

$$z_0 = 0 \quad (c = 0), \quad (1)$$

$$z_{1,2} = \pm \sqrt{c} \quad (c = \operatorname{Re} c > 0), \quad (2)$$

$$z_{1,2} = \pm i \sqrt{-c} \quad (c = \operatorname{Re} c < 0). \quad (3)$$

Fall 2. Die komplexe Zahl c ist nicht reell. Es sei $c = a + ib$, $z = x + iy$ ($a, b, x, y \in \mathbb{R}$). Nach Voraussetzung ist $b \neq 0$. Dann gilt $z^2 = c$, also

$$(x + iy)^2 = a + ib,$$

d. h.

$$(x^2 - y^2) + 2ixy = a + ib$$

genau dann, wenn die Gleichungen

$$x^2 - y^2 = a, \quad (4)$$

$$2xy = b \quad (5)$$

erfüllt sind. Unsere Problemstellung ist damit auf die Lösung des (nichtlinearen) Gleichungssystems (4), (5) zurückgeführt. Im Gegensatz zur Gleichung $z^2 = c$ handelt es sich um ein Gleichungssystem im Bereich der reellen Zahlen. Wir haben somit Anlaß, an Hand eines konkreten Beispiels zu zeigen, welche Probleme bei der Lösung nichtlinearer Gleichungssysteme für reelle Zahlen zu beachten sind. Wir nehmen zuerst an, daß zwei reelle Zahlen x, y die Gleichungen (4), (5) erfüllen. Durch Quadrieren und Addieren folgt

$$\begin{array}{r} x^4 - 2x^2y^2 + y^4 = a^2 \\ 4x^2y^2 = b^2 \\ \hline x^4 + 2x^2y^2 + y^4 = a^2 + b^2 \\ (x^2 + y^2)^2 = a^2 + b^2. \end{array}$$

Somit ist

$$x^2 + y^2 = \sqrt{a^2 + b^2},$$

und wenn wir (4) addieren bzw. subtrahieren, erhalten wir

$$2x^2 = a + \sqrt{a^2 + b^2}, \quad 2y^2 = -a + \sqrt{a^2 + b^2}.$$

Es folgt

$$|x| = \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}, \quad |y| = \sqrt{\frac{-a + \sqrt{a^2 + b^2}}{2}}. \quad (6)$$

Aus $b \neq 0$ und (5) folgt $x, y \neq 0$. Setzen wir $\sigma := \operatorname{sgn} x$, so gilt $\sigma \operatorname{sgn} y = \operatorname{sgn} b$ wegen (5), woraus $\operatorname{sgn} y = \sigma \operatorname{sgn} b$ folgt. Somit ist

$$x = \sigma \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}, \quad y = \sigma \operatorname{sgn} b \sqrt{\frac{-a + \sqrt{a^2 + b^2}}{2}}. \quad (7)$$

Wegen $\sqrt{a^2 + b^2} > \pm a$ können wir umgekehrt die reellen Zahlen x, y durch (7) mit $\sigma = 1$ oder $\sigma = -1$ definieren. Dann gilt

$$\begin{aligned} x^2 - y^2 &= \frac{a + \sqrt{a^2 + b^2}}{2} - \frac{-a + \sqrt{a^2 + b^2}}{2} = a, \\ 2xy &= 2 \operatorname{sgn} b \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} \sqrt{\frac{-a + \sqrt{a^2 + b^2}}{2}} \\ &= \operatorname{sgn} b \sqrt{-a^2 + (a^2 + b^2)} = \operatorname{sgn} b \sqrt{b^2} = b, \end{aligned}$$

d. h., die Gleichungen (4), (5) sind erfüllt. Daher sind

$$z_{1,2} = \pm \left(\sqrt{\frac{\operatorname{Re} c + |c|}{2}} + i \operatorname{sgn}(\operatorname{Im} c) \sqrt{\frac{-\operatorname{Re} c + |c|}{2}} \right) \quad (\operatorname{Im} c \neq 0) \quad (8)$$

die einzigen Lösungen der Gleichung $z^2 = c$.

Damit ist Satz 1 bewiesen.

Beispiel. Die Gleichung $z^2 = -1$ hat wegen (3) die Lösungen

$$z_{1,2} = \pm i \sqrt{1} = \pm i,$$

die Gleichung $z^2 = 12 - 5i$ wegen (8) die Lösungen

$$z_{1,2} = \pm \left(\sqrt{\frac{12 + 13}{2}} - i \sqrt{\frac{-12 + 13}{2}} \right) = \pm \left(\frac{5}{2} \sqrt{2} - \frac{i}{2} \sqrt{2} \right).$$

Mit Hilfe von Satz 1 beweisen wir

Satz 2. Zu jeder positiven natürlichen Zahl p gibt es eine komplexe Zahl z_0 mit

$$z_0^p = i.$$

Beweis (vollständige Induktion).

a) Die Behauptung gilt offenbar für $p = 1$.

b) Für alle natürlichen Zahlen q mit $1 \leq q \leq p$ gebe es eine komplexe Zahl z_0 mit $z_0^q = i$. Ist $p = 4k$ bzw. $p = 4k + 2$ ($k \in \mathbb{N}$), so gilt

$$i^{p+1} = i^{4k+1} = i \quad \text{bzw.} \quad (-i)^{p+1} = (-i)^{4k} (-i)^1 = i.$$

Ist aber p ungerade, so setzen wir $q := \frac{p+1}{2}$ und bestimmen eine komplexe Zahl z_1 mit $z_1^q = i$ und eine weitere komplexe Zahl z_0 mit $z_0^2 = z_1$. Dann gilt

$$z_0^{p+1} = z_0^{2q} = z_1^q = i.$$

In allen Fällen gilt also die Behauptung auch für $p + 1$.

Damit ist Satz 2 bewiesen.

1.3. Funktionen

1.3.1. Der Funktionsbegriff

Eine Teilmenge f des kartesischen Produktes $M \times N$ zweier beliebiger Mengen M, N heit bekanntlich eine *Funktion aus M in N* , wenn aus $(x, y_1), (x, y_2) \in f$ stets $y_1 = y_2$ folgt. Die Menge aller x aus M , zu denen ein y aus N mit $(x, y) \in f$ existiert, heit der *Definitionsbereich* $D(f)$ der Funktion f ; die Menge aller y aus N , zu denen ein x aus M mit $(x, y) \in f$ existiert, heit der *Wertebereich* $W(f)$ der Funktion f . Mit $f(x)$ wird im Fall $x \in D(f)$ das (nach Definition eindeutig bestimmte) Element y mit $f \in (x, y)$ bezeichnet. Wir drcken diesen Zuordnungsproze durch

$$f: x \mapsto y = f(x) \quad (x \in D(f)) \quad (1)$$

oder krzer durch

$$y = f(x) \quad (x \in D(f))$$

aus. Ist $M = D(f)$ bzw. $N = W(f)$, so heit f eine Funktion *von M in N* bzw. *aus M auf N* . Ist zugleich $M = D(f)$ und $N = W(f)$, so heit f eine Funktion *von M auf N* . Eine Funktion f heit *konstant*, wenn es ein Element c mit $f(x) = c$ fr alle $x \in D(f)$ gibt.

Ist f eine Funktion aus M in N und M eine Produktmenge $M_1 \times \cdots \times M_n$, so heit f eine *Funktion von n Variablen*. Die Elemente x des Definitionsbereichs von f knnen in der Form $x = (x_1, \dots, x_n)$ mit $x_i \in M_i$ ($i = 1, \dots, n$) dargestellt werden. Statt $f((x_1, \dots, x_n))$ schreibt man krzer $f(x_1, \dots, x_n)$.

Zu jeder Teilmenge R von $M \times N$, die auch als *Korrespondenz* zwischen den Mengen M und N bezeichnet wird, insbesondere also zu jeder Funktion f aus M in N , existiert die durch

$$R^{-1} = \{(x, y) : (y, x) \in R\} \quad (2)$$

definierte *Umkehrkorrespondenz* oder *inverse Korrespondenz*. Ist f^{-1} wieder eine Funktion (was nicht immer der Fall ist), so heit f eine *umkehrbare Funktion*, und f^{-1} heit die *Umkehrfunktion* oder *inverse Funktion* von f . Stets ist $D(f^{-1}) = W(f)$, $W(f^{-1}) = D(f)$.

Ist M eine Teilmenge des Definitionsbereichs einer Funktion, so verstehen wir unter der *Einschrnkung $f|_M$ der Funktion f auf die Menge M* die durch

$$f|_M: x \mapsto f(x) \quad (x \in M) \quad (3)$$

definierte Funktion. Ist g eine Einschrnkung der Funktion f , so heit f eine *Fortsetzung* der Funktion g auf die Menge $D(f)$.

Eine Funktion, deren Definitionsbereich die Menge \mathbf{N} der natrlichen Zahlen (oder eine unendliche Teilmenge von \mathbf{N}) ist, heit eine *Folge*, genauer eine *Folge in der Menge M* , wenn ihr Wertebereich eine Teilmenge von M ist. Jedes Element dieses

Wertebereichs heißt ein *Glied der Folge*. Eine Folge in der Menge der reellen bzw. komplexen Zahlen heißt eine *reelle* bzw. *komplexe Zahlenfolge*. Wird der natürlichen Zahl n durch eine Folge das Glied a_n zugeordnet, so bezeichnen wir diese Folge mit $(a_n)_{n \in \mathbb{N}}$ oder auch nur mit (a_n) .

In der Renaissance kam es zu einer außerordentlich schnellen Entwicklung der Mathematik. Sie wurde hervorgerufen nicht nur durch die rechnerisch-geschäftlichen Interessen der Kaufleute, sondern in noch stärkerem Maße durch den beginnenden Einsatz von Maschinerien in der Produktion (Manufakturen, Bergwerke). Neue Erfindungen (Feuerwaffen, Buchdruck), die Konstruktion neuer Maschinen und Geräte (Uhren) führten zu einem allgemeinen neuen technischen Bewußtsein und lenkten die Aufmerksamkeit auf das wissenschaftliche Studium von Veränderungen, insbesondere von Bewegungsvorgängen. Der Himmelmechanik kam dabei besondere Bedeutung zu.

Die Mathematik der Antike und besonders die des Mittelalters hatten trotz geistreicher Ansätze kaum Brauchbares zur mathematischen Beschreibung von Bewegungsvorgängen hinterlassen. Der entstehende Frühkapitalismus forderte demnach auch mathematische Verfahren zur Lösung dieses Problems.

Einen wesentlichen Fortschritt bildeten in dieser Richtung die Untersuchungen von RENÉ DESCARTES (1596–1650). Er hatte begonnen, die Veränderung algebraischer Ausdrücke in Abhängigkeit von der Veränderung der in diese Ausdrücke eingehenden Größen zu betrachten und einen Teil des dazu notwendigen Symbolismus und der Fachsprache entwickelt. Es bestand erst mit der Entwicklung der Analysis die Notwendigkeit, auch für die Abhängigkeit veränderlicher Größen neue Begriffe zu suchen.

Das Wort „function“ verwendete 1692 erstmals GOTTFRIED WILHELM LEIBNIZ (1646–1716). Es wurde durch JOHANN BERNOULLI (1667–1748) weit bekannt gemacht. Mit diesem Wort bezeichnete man dabei stillschweigend nur solche Zusammenhänge, die durch Formeln vom Typ algebraischer Ausdrücke beschrieben werden konnten. LEONARD EULER (1707–1783) gab 1748 für diese Vorstellungen die klassische Definition: „Die Funktion der veränderlichen Größen ist ein analytischer Ausdruck, der auf beliebige Art aus veränderlichen Größen und Zahlen oder konstanten Quantitäten gebildet worden ist.“ Unter einem solchen „analytischen Ausdruck“ ist also bei EULER eine nach dem Typ algebraischer Funktionen gebildete Abhängigkeit zu verstehen. Diese Definition EULERS entsprach dem Verständnis der Mathematiker des 18. Jahrhunderts. Sie wurde nahegelegt durch das rein formale Operieren mit unendlichen Reihen. Traten wirklich einmal analytische Funktionen auf, die andere Singularitäten als die algebraischen besaßen, so wurden sie kaum beachtet. Die Mathematiker gewöhnten sich daran, „analytischer Ausdruck“ und „Funktion“ als identische Begriffe zu betrachten.

Die Diskussion eines praktischen Problems, das der Beschreibung der Oszillation der schwingenden Saite, führte seit 1750 in vielerlei Hinsicht zur Untersuchung grundlegender Fragen der Analysis und erfaßte auch den Funktionsbegriff.

Es hatte sich beim Problem der schwingenden Saite nämlich herausgestellt, daß die hier wirklich vorkommenden Abhängigkeiten in verschiedenen Abschnitten durch verschiedene analytische Ausdrücke gegeben sein können. Im Jahre 1807 zeigte nun JOSEPH FOURIER (1768–1830) sogar, daß eine gebrochene Linie (Ausgangsform der schwingenden Saite) durch eine einzige trigonometrische Reihe beschrieben werden kann. FOURIER konnte sich in seinen Untersuchungen dabei besonders auf Vorarbeiten von JOHANN BERNOULLI stützen.

Die Gleichsetzung von „Funktion“ und „analytischem Ausdruck“ mußte jetzt fallengelassen werden. Die neue notwendige Erweiterung der Definition des Funktionsbegriffes wurde zu Anfang des 19. Jahrhunderts von verschiedener Seite, so von SYLVESTER FRANÇOIS LACROIX (1765–1843), BERNARD BOLZANO (1781–1848) und NIKOLAI IVANOVITZ LOBAČEVSKIJ (1793–1856) vorgenommen. In klassischer Formulierung lautete die neue Definition 1837 bei PETER GUSTAV LEJEUNE-DIRICHLET (1805–1859): „Eine Funktion heißt y von x , wenn jedem Werte der veränderlichen Größe x innerhalb eines gewissen Intervalls ein bestimmter Wert von y entspricht; gleichviel, ob y in dem ganzen Intervalle nach demselben Gesetze von x abhängt oder nicht; ob die Abhängigkeit durch mathematische Operationen ausgedrückt werden kann oder nicht.“

Die Ausarbeitung des Funktionsbegriffes kann als eine Hauptleistung der Mathematik des 19. Jahrhunderts angesehen werden, wobei sich durch die Einführung komplexer Zahlen in den Funktionsbegriff viele neue Aspekte ergaben. Eine völlig scharfe Fassung des Begriffs, wie sie oben skizziert wurde, gelang erst bei der Überarbeitung der Mathematik mit den Prinzipien der Mengenlehre.

1.3.2. Funktionen aus der Menge bzw. in die Menge der reellen bzw. komplexen Zahlen

Eine Funktion f aus einer beliebigen Menge in die Menge \mathbf{R} der reellen bzw. in die Menge \mathbf{C} der komplexen Zahlen nennen wir eine *reellwertige* bzw. *komplexwertige* Funktion. Eine Funktion aus \mathbf{R} in \mathbf{R} bzw. aus \mathbf{C} in \mathbf{C} nennen wir kurz eine *reelle* bzw. *komplexe* Funktion. Jeder reellen Funktion f können wir nach Wahl eines kartesischen Koordinatensystems in der Ebene die Menge aller Punkte $P(x, f(x))$ mit $x \in D(f)$ zuordnen. Sie heißt der *Graph* der reellen Funktion. Eine Menge von Punkten $P(x, y)$ der Ebene ist genau dann der Graph einer Funktion, wenn jede Parallele zur y -Achse mit der Menge M höchstens einen Punkt gemein hat.

Ist M eine Teilmenge des Definitionsbereichs einer reellwertigen Funktion, so setzen wir

$$\sup_{x \in M} f(x) := \sup \left\{ y : y \in \mathbf{R} \wedge \bigvee_{x \in M} y = f(x) \right\}.$$

Entsprechend bezeichnen wir das Infimum der rechts auftretenden Menge und ihr Maximum bzw. Minimum, falls letztere existieren.

Definition 1. Eine reell- bzw. komplexwertige Funktion heißt *beschränkt*, wenn ihr Wertebereich beschränkt ist:

$$f \text{ beschränkt} : \Leftrightarrow \bigvee_{K \in \mathbf{R}_+} \bigwedge_{x \in D(f)} |f(x)| \leq K. \quad (1)$$

Entsprechend heißt eine reellwertige Funktion *nach oben* bzw. *nach unten* beschränkt, wenn ihr Wertebereich nach oben bzw. nach unten beschränkt ist:

$$f \text{ nach oben beschränkt} : \Leftrightarrow \bigvee_{K \in \mathbf{R}} \bigwedge_{x \in D(f)} f(x) \leq K, \quad (2)$$

$$f \text{ nach unten beschränkt} : \Leftrightarrow \bigvee_{K \in \mathbf{R}} \bigwedge_{x \in D(f)} f(x) \geq K.$$

Definition 2. Eine von 0 verschiedene reelle bzw. komplexe Zahl λ heißt eine *Periode* einer Funktion f aus \mathbf{R} bzw. aus \mathbf{C} , wenn für $x \in D(f)$ stets $x + \lambda \in D(f)$ ist und

$$f(x + \lambda) = f(x) \quad (3)$$

gilt. Gibt es zu einer Funktion f aus \mathbf{R} eine kleinste positive reelle Zahl λ , die die Bedingung (3) erfüllt, so heißt sie die *primitive Periode* der Funktion f .

Funktionen, die eine Periode besitzen, werden *periodische Funktionen* genannt. Wichtige Beispiele für periodische Funktionen sind bekanntlich die trigonometrischen Funktionen, die wir in 1.6. einführen werden. Die Funktionen $y = \sin x$ und $y = \cos x$ besitzen die primitive Periode 2π .

Definition 3. Ein Element x des Definitionsbereichs einer Funktion f heißt eine *c-Stelle* bzw. eine *Nullstelle* der Funktion, wenn $f(x) = c$ bzw. $f(x) = 0$ ist.

Die nachfolgende Begriffsbildung ist u. a. für die Entwicklung von Funktionen in Reihen von Bedeutung.

Definition 4. Eine reelle bzw. komplexe Funktion heißt *gerade* bzw. *ungerade*, wenn aus $x \in D(f)$ stets $-x \in D(f)$ und

$$f(-x) = f(x) \quad (4)$$

bzw.

$$f(-x) = -f(x) \quad (5)$$

folgt.

Der Graph einer geraden bzw. ungeraden reellen Funktion liegt symmetrisch zur y -Achse bzw. symmetrisch zum Ursprung. Alle Potenzfunktionen $y = x^n$ mit $n \in \mathbf{N}$ sind für gerade n gerade, für ungerade n ungerade Funktionen.

Es sei f eine beliebige reelle Funktion mit einem zum Ursprung symmetrischen Definitionsbereich. Setzen wir

$$g(x) := \frac{f(x) + f(-x)}{2}, \quad h(x) := \frac{f(x) - f(-x)}{2},$$

so gelten die Beziehungen

$$g(-x) = g(x), \quad h(-x) = -h(x),$$

und es ist

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

Wir können also f als Summe einer geraden und einer ungeraden Funktion darstellen.

Wir unterwerfen den Graphen einer reellen Funktion f gewissen geometrischen Transformationen, die wieder den Graphen einer Funktion g ergeben, und untersuchen die Beziehungen zwischen den zugehörigen Funktionen.

A. Spiegelungen

Geht der Graph von g aus dem Graphen von f durch Spiegelung an der y -Achse bzw. an der x -Achse bzw. am Ursprung hervor, so gilt

$$g(x) = f(-x) \quad (-x \in D(f))$$

bzw.

$$g(x) = -f(x) \quad (x \in D(f))$$

bzw.

$$g(x) = -f(-x) \quad (-x \in D(f))$$

(Abb. 1.2).

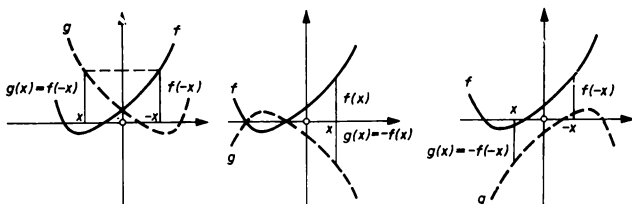


Abb. 1.2

B. Translationen

Geht der Graph von g aus dem Graphen von f durch Translation um den Vektor (a, b) hervor, so gilt

$$g(x) = f(x - a) + b \quad (x - a \in D(f))$$

(Abb. 1.3).

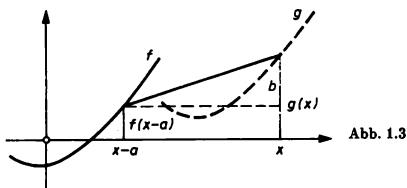


Abb. 1.3

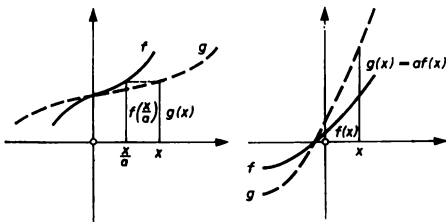


Abb. 1.4

C. *Dehnungen und Stauchungen*

Geht der Graph von g aus dem Graphen von f durch Dehnung (bzw. Stauchung) im Verhältnis $a:1$ ($a > 0$) in Richtung der x - bzw. y -Achse hervor, so gilt

$$g(x) = f\left(\frac{x}{a}\right) \quad \left(\frac{x}{a} \in D(f)\right)$$

bzw.

$$g(x) = af(x) \quad (x \in D(f))$$

(Abb. 1.4).

1.3.3. *Verknüpfungen von Funktionen*

Wir stellen einige Möglichkeiten zusammen, aus gegebenen Funktionen neue Funktionen zu bilden.

Aus der allgemeinen Theorie des Funktionsbegriffs ist bekannt, daß zu jeder Funktion f aus L in M und zu jeder Funktion g aus M in N eine mit $f \circ g$ (gelesen: f nach g) bezeichnete (möglicherweise leere) Funktion gebildet werden kann. Sie ist durch

$$(f \circ g)(x) := f(g(x)) \quad (x \in D(g) \wedge g(x) \in D(f)) \quad (1)$$

definiert und heißt eine *zusammengesetzte Funktion*. Hiernach ist

$$D(f \circ g) = \{x \in D(g) : g(x) \in D(f)\}. \quad (2)$$

Ferner gilt das *Assoziativgesetz*

$$(f \circ g) \circ h = f \circ (g \circ h). \quad (3)$$

Für reellwertige bzw. komplexwertige Funktionen können weitere Verknüpfungen definiert werden. Ist f eine reell- bzw. komplexwertige Funktion und λ eine reelle bzw. komplexe Zahl, so definieren wir die Funktion λf durch

$$(\lambda f)(x) := \lambda f(x) \quad (x \in D(f)). \quad (4)$$

Die Funktionen f und λf haben somit denselben Definitionsbereich. Die Summe bzw. Differenz zweier reell- bzw. komplexwertiger Funktionen f und g definieren wir auch für den Fall, daß die Definitionsbereiche von f und g nicht identisch sind. Wir setzen

$$(f \pm g)(x) := f(x) \pm g(x) \quad (x \in D(f) \cap D(g)). \quad (5)$$

Sind f, g reelle bzw. komplexe Funktionen, so definieren wir das Produkt von f, g durch

$$(f \cdot g)(x) := f(x) g(x) \quad (x \in D(f) \cap D(g)). \quad (6)$$

Man hat sorgfältig zwischen $f \cdot g = fg$ und $f \circ g$ zu unterscheiden. Die Definitionsbereiche von $f \pm g$ und $f \cdot g$ sind gleich dem Durchschnitt der Definitionsbereiche von f und g . Ist dieser Durchschnitt leer, so ergibt sich die „leere Funktion“.

Bei der Bildung des *Kehrwertes* einer reellen oder komplexen Funktion ist zu beachten, daß eine Division durch 0 ausgeschlossen ist. Wir setzen

$$\frac{1}{f}(x) := \frac{1}{f(x)} \quad (x \in D(f) \wedge f(x) \neq 0). \quad (7)$$

Die Nullstellen der Funktion f gehören somit nicht zum Definitionsbereich der Funktion $\frac{1}{f}$. Sie darf nicht mit der möglicherweise existierenden Umkehrfunktion f^{-1} von f verwechselt werden. Den Quotienten zweier reeller oder komplexer Funktionen definieren wir durch

$$\frac{f}{g} := f \cdot \frac{1}{g},$$

so daß

$$\frac{f}{g}(x) = \frac{f(x)}{g(x)} \quad (x \in D(f) \cap D(g), g(x) \neq 0) \quad (8)$$

gilt.

Wir führen die Menge \mathfrak{F}_D aller reellwertigen bzw. komplexwertigen Funktionen mit gemeinsamem Definitionsbereich D ein. Wir zeigen, daß die Menge \mathfrak{F}_D einen Vektorraum bildet. Für alle $f, g, h \in \mathfrak{F}_D$ und alle reellen (bzw. komplexen) Zahlen λ, μ folgt aus den Definitionen (4), (5):

A I. $f + g \in \mathfrak{F}_D$.

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

III. Die Funktion $k := g - f$ genügt der Gleichung $f + k = g$.

IV. $g + f = f + g$.

B I. $\lambda f \in \mathfrak{F}_D$.

II. $1 \cdot f = f$.

III. $(\lambda\mu)f = \lambda(\mu f)$.

IV. $(\lambda + \mu)f = \lambda f + \mu f$.

V. $\lambda(f + g) = \lambda f + \lambda g$.

Das Nullelement dieses Vektorraumes ist die *Nullfunktion*, die jedes x aus D auf 0 abbildet.

Wie in allen Vektorräumen kann nun der Begriff der *Linearkombination*

$$f = \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n$$

der Funktionen f_1, f_2, \dots, f_n mit den Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_n$ eingeführt werden. Ein System f_1, f_2, \dots, f_n von n verschiedenen Funktionen f_i mit $f_i \in \mathfrak{F}_D$ ($i = 1, 2, \dots, n$) heißt *linear unabhängig*, wenn aus

$$\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n = 0$$

stets $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$ folgt.

1.3.4. Ganzrationale, rationale und algebraische Funktionen

Für jede natürliche Zahl n heißt die Funktion

$$y = x^n \quad (x \in \mathbb{R} \text{ bzw. } x \in \mathbb{C})$$

eine reelle bzw. komplexe *Potenzfunktion*. Da wir den reellen und den komplexen Fall häufig parallel behandeln, kann x auch eine Variable für komplexe Zahlen sein. In diesem Fall setzen wir $x_1 := \operatorname{Re} x$, $x_2 := \operatorname{Im} x$.

Definition 1. Jede Linearkombination von reellen Potenzfunktionen mit reellen Koeffizienten heißt eine *reelle ganzrationale Funktion*. Jede Linearkombination von komplexen Potenzfunktionen mit komplexen Koeffizienten heißt eine *komplexe ganzrationale Funktion*.

Häufig werden ganzrationale Funktionen *Polynome* genannt. Neuerdings versteht man unter Polynomen die Elemente gewisser algebraischer Strukturen, die allerdings in sehr engen Beziehungen zu den ganzrationalen Funktionen stehen.

Zur Vorbereitung des Beweises über die lineare Unabhängigkeit des Systems der Potenzfunktionen benötigen wir den in mehrfacher Hinsicht bedeutungsvollen

Satz 1. Es sei $a_n \neq 0$ und

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \quad (x \in \mathbb{R} \text{ bzw. } x \in \mathbb{C}) \quad (1)$$

eine reelle bzw. komplexe ganzrationale Funktion. Dann gibt es im Fall $n \geq 1$ eine positive reelle Zahl R derart, daß

$$|f(x)| \geq \frac{|a_n|}{2} \cdot |x|^n \quad \text{für } |x| \geq R \quad (2)$$

und im Fall einer reellen Funktion

$$\frac{f(x)}{a_n x^n} \geq \frac{1}{2} \quad \text{für } |x| \geq R \quad (3)$$

ist.

Beweis. Für $|x| \geq 1$ ist

$$\begin{aligned} \left| \frac{f(x)}{a_n x^n} - 1 \right| &= \frac{|f(x) - a_n x^n|}{|a_n| \cdot |x|^n} \leq \frac{|a_0| + |a_1| \cdot |x| + \cdots + |a_{n-1}| \cdot |x|^{n-1}}{|a_n| \cdot |x|^n} \\ &\leq \frac{|a_0| + |a_1| + \cdots + |a_{n-1}|}{|a_n| \cdot |x|}. \end{aligned}$$

Wählen wir $R \geq 1$ und

$$R \geq 2 \frac{|a_0| + |a_1| + \cdots + |a_{n-1}|}{|a_n|},$$

so folgt aus $|x| \geq R$ stets

$$1 - \frac{|f(x)|}{|a_n| \cdot |x|^n} \leq \left| \frac{f(x)}{a_n x^n} - 1 \right| \leq \frac{1}{2},$$

und der Vergleich der linken und rechten Seite ergibt (2). Aus der zweiten Ungleichung folgt, wenn f reell ist,

$$-\left(\frac{f(x)}{a_n x^n} - 1\right) \leq \frac{1}{2},$$

und das ist die Behauptung (3).

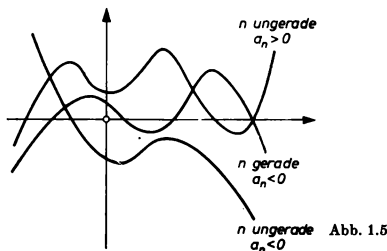


Abb. 1.5

In Abb. 1.5 sind die Graphen einiger reeller ganzrationaler Funktionen angedeutet.

Satz 2. Für jede natürliche Zahl n ist das System der reellen bzw. komplexen Potenzfunktionen k -ten Grades mit $k = 0, 1, \dots, n$ linear unabhängig.

Beweis. Es sei $y = a_0 + a_1 x + \cdots + a_n x^n$ die Nullfunktion, d. h., für alle reellen bzw. komplexen Zahlen x gelte

$$a_0 + a_1 x + \cdots + a_n x^n = 0.$$

Nehmen wir an, es sei $a_k \neq 0$ für wenigstens ein k . Ist dann m die größte natürliche Zahl mit $a_m \neq 0$, so gilt $m > 0$ und

$$a_0 + a_1x + \dots + a_mx^m = 0$$

für alle x , was ein Widerspruch zu (2) ist. Somit ist $a_0 = a_1 = \dots = a_n = 0$, und Satz 2 ist bewiesen.

Wie in allen Vektorräumen gilt der Satz, daß die Koeffizienten einer Linearkombination eines linear unabhängigen Systems eindeutig bestimmt sind.

Satz 3. Jeder von der Nullfunktion verschiedenen ganzrationalen Funktion f kann umkehrbar eindeutig ein Term $a_0 + a_1x + \dots + a_nx^n$ in der Variablen x mit $a_n \neq 0$ zugeordnet werden derart, daß

$$f(x) = a_0 + a_1x + \dots + a_nx^n \quad (a_n \neq 0)$$

für alle $x \in D(f)$ ist.

Die Behauptung ist eine unmittelbare Folgerung aus Satz 2.

Definition 2. Besitzt die ganzrationale Funktion f die Darstellung (1) mit $a_n \neq 0$, so heißt die natürliche Zahl n der Grad der Funktion f , in Zeichen $n = \deg f$.¹⁾

Die konstanten, von der Nullfunktion verschiedenen ganzrationalen Funktionen haben den Grad 0, die (nicht konstanten) linearen Funktionen den Grad 1, die (eigentlich) quadratischen Funktionen den Grad 2.

Unter einer reellen bzw. komplexen rationalen Funktion verstehen wir eine Funktion, die sich als Quotient zweier ganzrationaler Funktionen ohne gemeinsame (komplexe) Nullstellen darstellen läßt. Abgesehen von konstanten Faktoren ist diese Darstellung eindeutig bestimmt.

Während eine ganzrationale Funktion von mindestens erstem Grade nach Satz 1 niemals beschränkt ist, besitzt z. B. die Funktion

$$f(x) = \frac{1}{1+x^2} \quad (x \in \mathbb{R})$$

die Schranke 1 (Abb. 1.6). Ist f der Quotient der ganzrationalen Funktionen g, h ohne gemeinsame Nullstellen, so heißt jede Nullstelle von h eine Polstelle der rationalen Funktion f . So ist z. B. 1 eine Polstelle der rationalen Funktion

$$f(x) = \frac{x+1}{x-1}$$

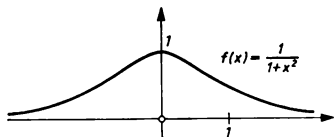


Abb. 1.6

¹⁾ degree (engl.) — Grad.

(Abb. 1.7). Eine rationale Funktion $f = \frac{g}{h}$ mit ganzrationalen Funktionen g, h ohne gemeinsame Nullstellen heißt *echt gebrochen*, wenn der Grad des Zählers kleiner als der Grad des Nenners ist. Wenn dies nicht der Fall ist, kann sie mit dem in der Algebra behandelten Verfahren der *Restdivision* als Summe einer ganzrationalen und einer echt gebrochenen rationalen Funktion dargestellt werden.

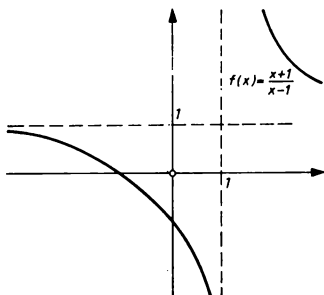


Abb. 1.7

Zur Definition der algebraischen Funktionen erweitern wir zunächst den Begriff der Potenzfunktion und den Begriff der ganzrationalen Funktion auf Funktionen von mehreren Variablen. Sind m_1, \dots, m_p beliebige natürliche Zahlen, so heißt die Funktion f mit

$$f(x_1, x_2, \dots, x_p) := x_1^{m_1} x_2^{m_2} \dots x_p^{m_p}$$

eine *reelle bzw. komplexe Potenzproduktfunktion* der p reellen bzw. komplexen Variablen x_1, x_2, \dots, x_p . Jede Linearkombination von reellen bzw. komplexen Potenzproduktfunktionen mit reellen bzw. komplexen Koeffizienten heißt eine *reelle bzw. komplexe ganzrationale Funktion*. So ist z. B.

$$y = 3x_1^3 x_2 x_3^2 + x_1 x_3^4 - 2x_1^2 x_2 + 2x_1 + x_3 + 5 \quad (x \in \mathbb{R})$$

eine ganzrationale Funktion. Jede ganzrationale Funktion F bestimmt im Bereich der reellen bzw. komplexen Zahlen die p -stellige Relation

$$\{(x_1, x_2, \dots, x_p) : F(x_1, x_2, \dots, x_p) = 0\}.$$

Wir beschränken uns jetzt auf den Fall reeller ganzrationaler Funktionen f von zwei Variablen. Die Menge

$$\{(x, y) : x, y \in \mathbb{R} \wedge F(x, y) = 0\} \quad (4)$$

stellt im allgemeinen eine Punktmenge in der Ebene dar, die sich aus einer oder aus mehreren Kurven zusammensetzt (Abb. 1.8). Wie diese Beispiele zeigen, wird durch (4) im allgemeinen keine Funktion definiert. Oft kann man jedoch die durch (4)

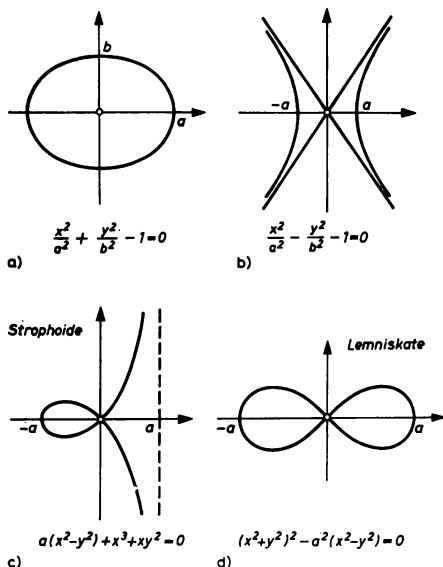


Abb. 1.8

definierten Punktmengen der Ebene in Teilmengen zerlegen, die Graphen von Funktionen sind. Funktionen, die in dieser Weise definiert werden können, nennt man *algebraisch*. Ist also f eine algebraische Funktion, so gibt es eine ganzrationale Funktion F von zwei Variablen derart, daß

$$F(x, f(x)) = 0 \quad (5)$$

für alle x mit $x \in D(f)$ ist. Eine weitere Bedingung dafür, daß eine algebraische Funktion vorliegt, können wir erst später erläutern. Sie besteht darin, daß die Funktion f (im Sinne von 3.5.3.) durch die Funktion F implizit definiert und stetig ist.

Beispiel. Für gegebene positive reelle Zahlen a, b setzen wir

$$F(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \quad (x, y \in \mathbb{R})$$

und

$$f(x) := b \sqrt[3]{1 - \frac{x^2}{a^2}} \quad \text{für } -a \leq x \leq a.$$

Dann gilt

$$F(x, f(x)) = 0 \quad \text{für } -a \leq x \leq a,$$

d. h., die Funktion f ist algebraisch. Ihr Graph ist die obere Hälfte der in Abb. 1.8a dargestellten Ellipse.

Ist f eine nichtnegative algebraische Funktion, so ist auch deren m -te Wurzel ($m = 1, 2, \dots$), d. h. die Funktion g mit $g(x) = \sqrt[m]{f(x)}$, eine algebraische Funktion.

Sind g, f zwei algebraische Funktionen, so sind auch $\lambda f, f \pm g, fg, \frac{f}{g}$ algebraische

Funktionen. Es sei aber nachdrücklich darauf hingewiesen, daß man, mit Hilfe der genannten Operationen, von rationalen Funktionen ausgehend, nicht alle in der Definition erfaßten algebraischen Funktionen konstruieren kann. Dies gilt z. B. für die algebraischen Funktionen

$$f = \{(x, y) : y^5 - 4y - x - 2 = 0 \wedge y \geq 1\}. \quad (6)$$

Der Beweis dieser Behauptung gründet sich auf einen berühmten, von NIELS HENRIK ABEL (1802–1829) gefundenen Satz, nach dem es algebraische Gleichungen fünften und höheren Grades gibt, deren Lösung nicht durch ineinandergeschachtelte Wurzeln angegeben werden kann.

1.3.5. Monotone Funktionen

Die in der Schulmathematik auftretenden reellen Funktionen sind monoton oder wenigstens stückweise monoton, d. h., der Definitionsbereich läßt sich in Intervalle zerlegen derart, daß die Einschränkung der Funktion auf diese Intervalle monoton ist. Wir geben zunächst die

Definition. Eine reelle Funktion f heißt

monoton wachsend, wenn aus $x_1 < x_2$ stets $f(x_1) \leq f(x_2)$,

monoton fallend, wenn aus $x_1 < x_2$ stets $f(x_1) \geq f(x_2)$,

streng monoton wachsend, wenn aus $x_1 < x_2$ stets $f(x_1) < f(x_2)$,

streng monoton fallend, wenn aus $x_1 < x_2$ stets $f(x_1) > f(x_2)$

folgt, wobei stets $x_1, x_2 \in D(f)$ gelte.

Eine konstante reelle Funktion ist hiernach monoton wachsend und monoton fallend zugleich.

Als Folgerung aus der Definition ergibt sich der

Satz 1. *Jede streng monotone reelle Funktion besitzt eine Umkehrfunktion.*

Beweis. Die Funktion f sei streng monoton. Dann folgt aus $x_1, x_2 \in D(f)$ und $x_1 \neq x_2$ stets $f(x_1) \neq f(x_2)$, und durch Kontraposition ergibt sich, daß aus $f(x_1) = f(x_2)$ stets $x_1 = x_2$ folgt. Somit ist f^{-1} eine Funktion.

Satz 2. *Die Umkehrfunktion einer streng monoton wachsenden (fallenden) reellen Funktion ist wieder streng monoton wachsend (fallend).*

Beweis. Es sei etwa f streng monoton wachsend und $f^{-1}(y_1) \geq f^{-1}(y_2)$. Dann folgt $y_1 = f(f^{-1}(y_1)) \geq f(f^{-1}(y_2)) = y_2$. Kontraposition ergibt: Aus $y_1 < y_2$ folgt $f^{-1}(y_1) < f^{-1}(y_2)$, d. h., f^{-1} ist streng monoton wachsend.

Eine zusammengesetzte Funktion $h(x) = f(g(x))$ zweier streng monotonen Funktionen ist wieder streng monoton, und zwar wachsend bzw. fallend, je nachdem, ob f und g das gleiche oder das entgegengesetzte Monotonieverhalten haben.

Im folgenden Abschnitt werden wir Potenzen a^r für beliebige positive Basen a und beliebige rationale Exponenten r definieren. Dabei wird sich zeigen, daß die für festes $a > 0$ mit $a \neq 1$ definierte Funktion $f(r) := a^r$ ($r \in \mathbb{Q}$) streng monoton ist, und zwar wachsend oder fallend, je nachdem, ob $a > 1$ oder $0 < a < 1$ ist. Der Definitionsbereich der Funktion f ist also nur die Menge aller rationalen Zahlen. Dieser Sachverhalt kann zwar grafisch nicht dargestellt, wohl aber plausibel gemacht werden (Abb. 1.9). Die Punkte x_0, x_1, x_2 sollen zum Definitionsbereich der Funktion gehören, der Punkt x dagegen nicht. Die Funktion soll nun auch in den „Lücken“ definiert werden.

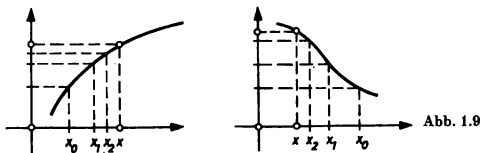


Abb. 1.9

Es sei f zunächst eine monoton wachsende Funktion, und I sei das kleinste Intervall, das den Definitionsbereich von f umfaßt. Nach Abb. 1.9 liegt es nahe, den Funktionswert in einem Punkt $x \in I$, der nicht zum Definitionsbereich gehört, als kleinste obere Schranke aller Funktionswerte $f(t)$ mit $t < x$ zu definieren. Wir setzen also

$$\bar{f}(x) := \sup \{f(t) : t \in D(f) \wedge t \leq x\}. \quad (1)$$

Für $x \in D(f)$ ist dann offensichtlich $\bar{f}(x) = f(x)$, d. h., \bar{f} ist eine Fortsetzung der Funktion f . Ferner folgt aus $x_1, x_2 \in I$ und $x_1 < x_2$ stets $\bar{f}(x_1) \leq \bar{f}(x_2)$, d. h., auch die auf das Intervall I fortgesetzte Funktion f ist monoton wachsend. Im Fall einer

monoton fallenden Funktion f hat die Funktion

$$\bar{f}(x) := \sup \{f(t) : t \in D(f) \wedge t \geq x\} \quad (x \in I) \quad (2)$$

analoge Eigenschaften.

Eine zusätzliche Forderung sichert, daß dies die einzige Möglichkeit ist, unsere Funktion zu einer auf I monotonen Funktion fortzusetzen. Wir nennen eine Teilmenge M von \mathbb{R} *relativ dicht*, wenn sie dicht liegt im kleinsten Intervall $I(M)$, das die Menge M umfaßt (vgl. 1.1.3.).

Satz 3. *Eine monotone Funktion f , deren Wertebereich relativ dicht ist, kann auf genau eine Weise zu einer monotonen Funktion \bar{f} fortgesetzt werden, die auf dem kleinsten, den Definitionsbereich von f umfassenden Intervall I definiert ist.*

Beweis. Es sei I' das kleinste Intervall, das den Wertebereich von f umfaßt. Die Funktion f sei etwa monoton wachsend, und f^* sei eine beliebige monotone Fortsetzung von f mit $D(f^*) = I$. Dann ist auch f^* monoton wachsend. Da zu $x \in I$ stets $a, b \in D(f)$ mit $a \leq x \leq b$, also $f(a) = f^*(a) \leq f(x) \leq f^*(b) = f(b)$ existieren, ist $f^*(x) \in I'$, also $W(f^*) \subseteq I'$. Speziell haben wir stets

$$W(\bar{f}) \subseteq I'. \quad (3)$$

Nehmen wir nun an, es gäbe zwei verschiedene monotone Fortsetzungen f_1, f_2 , dann gibt es ein $x_0 \in I$ mit $f_1(x_0) \neq f_2(x_0)$, etwa $f_1(x_0) < f_2(x_0)$, woraus $x_0 \notin D(f)$ folgt. Für $x_1, x_2 \in D(f)$ und $x_1 < x_0 < x_2$ ist aber

$$f(x_1) = f_1(x_1) \leq f_1(x_0) < f_2(x_0) \leq f_2(x_2) = f(x_2),$$

und für keinen einzigen Punkt $x \in D(f)$ gilt $f_1(x_0) < f(x) < f_2(x_0)$. Das widerspricht unserer Voraussetzung, nach der $W(f)$ dicht in I' liegt. Daher ist f die einzige Fortsetzung mit den gewünschten Eigenschaften.

Unter den Voraussetzungen von Satz 3 nennen wir \bar{f} die *monotone Fortsetzung* von f .

Satz 4. *Die monotone Fortsetzung einer streng monotonen Funktion mit relativ dichtem Definitions- und Wertebereich ist streng monoton.*

Beweis. Wir nehmen wieder an, daß f , also auch \bar{f} monoton wachsend ist. Es sei $x_1, x_2 \in D(\bar{f})$ und $x_1 < x_2$. Dann gibt es Elemente $a, b \in D(f)$ mit $x_1 < a < b < x_2$, und es folgt $\bar{f}(x_1) \leq \bar{f}(a) = f(a) < f(b) = \bar{f}(b) \leq \bar{f}(x_2)$, womit Satz 4 bewiesen ist.

Ein zentraler Satz der Theorie der stetigen Funktionen, der sogenannte *Zwischenwertsatz*, besagt, daß eine auf einem Intervall I definierte stetige Funktion keinen „Zwischenwert“ ausläßt, d. h., daß zu jeder Zahl c , die zwischen zwei Funktionswerten liegt, eine Zahl $x \in I$ mit $f(x) = c$ existiert (vgl. 2.4.2., Satz 3). Gleichbedeutend hiermit ist die Aussage, daß der Wertebereich der Funktion ein Intervall ist. Für streng monotone Funktionen kann ein analoger Satz bereits jetzt ohne Mühe bewiesen werden.

Satz 5. Ist der Wertebereich einer auf einem Intervall I definierten streng monotonen Funktion relativ dicht, so ist der Wertebereich ein Intervall.

Beweis. Es sei wieder I' das kleinste $W(f)$ umfassende Intervall, und g sei die Umkehrfunktion von f , also $D(g) = W(f)$ und $W(g) = I$. Die (sogar streng) monotone Fortsetzung \bar{g} von g ist also auf I' definiert, und wegen (3) ist $W(\bar{g}) \subseteq I$. Es sei nun $c \in I'$ und $x := \bar{g}(c) \in I$. Dann ist $y := f(x) \in D(g)$, also $\bar{g}(y) = g(y) = x = \bar{g}(c)$. Da \bar{g} streng monoton ist, folgt $c = y \in W(f)$, und Satz 5 ist bewiesen.

Der Nachweis, daß $W(f)$ relativ dicht ist, kann oft mit folgendem Satz geführt werden.

Satz 6. Ist der Definitionsbereich einer Funktion f relativ dicht und gibt es ein $K > 0$ mit

$$|f(x) - f(y)| \leq K |x - y| \quad (x, y \in D(f)), \quad (4)$$

so ist auch der Wertebereich relativ dicht.

Beweis. Es sei $a, b \in D(f)$ und $f(a) < f(b)$. Wir können etwa $a < b$ voraussetzen. Da $D(f)$ relativ dicht ist, können wir Elemente $x_0, \dots, x_n \in D(f)$ mit $a = x_0 < x_1 < \dots < x_n = b$ und $0 < x_{i+1} - x_i < \frac{f(b) - f(a)}{K}$ ($i = 0, \dots, n-1$) finden. Wegen $f(x_0) = f(a) < f(b) = f(x_n)$ gibt es ein j mit $0 \leq j < n$ und $f(x_j) \leq f(a) < f(x_{j+1})$. Aus (4) folgt

$$f(x_{j+1}) - f(x_j) \leq K(x_{j+1} - x_j) < f(b) - f(a),$$

und es ist $f(a) < f(x_{j+1}) < f(x_j) + (f(b) - f(a)) \leq f(a) + (f(b) - f(a)) = f(b)$. Der Funktionswert im Punkt x_{j+1} liegt also echt zwischen $f(a)$ und $f(b)$.

Aus der Bedingung (4) folgt übrigens, wie wir noch sehen werden, daß die Funktion stetig ist.

1.4. Potenz-, Exponential- und Logarithmusfunktionen

1.4.1. Potenzen mit rationalen Exponenten

Die Grundlage für den weiteren Ausbau der Lehre von den Potenzfunktionen ist der Satz von der Existenz und Einzigkeit der n -ten Wurzel (vgl. 1.1.5.). Auf seiner Grundlage definieren wir Potenzen mit rationalen Exponenten. Für alle nicht-negativen reellen Zahlen a und alle natürlichen Zahlen p, q mit $q \neq 0$ ist

$$\left(\sqrt[q]{a}\right)^p = \left(\sqrt[p]{a}\right)^{pq} = \left(\sqrt[pq]{a}\right)^p = a^p$$

und folglich

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

Jede nichtnegative rationale Zahl r kann in der Form $r = \frac{p}{q}$ mit $p, q \in \mathbf{N}$ und $q \neq 0$ dargestellt werden. Ist auch $\frac{p'}{q'} = r = \frac{p}{q}$, so gilt

$$\sqrt[q]{a^{p'}} = \sqrt[q]{a^{pq'}} = \sqrt[q]{a^{p'q}} = \sqrt[a^{p'}]{q} = a^{p'},$$

woraus

$$\sqrt[q]{a^p} = \sqrt[q']{a^{p'}} \quad \left(\frac{p'}{q'} = \frac{p}{q} \right)$$

folgt. Wir können daher für alle nichtnegativen rationalen Zahlen r mit $r = \frac{p}{q}$ durch

$$a^{p/q} = \sqrt[q]{a^p} \quad (a \in \mathbf{R}, a \neq 0; p, q \in \mathbf{N}, q \neq 0) \quad (2)$$

Potenzen mit nichtnegativer reeller Basis und mit nichtnegativen rationalen Exponenten definieren. Die Definition ist unabhängig von der Wahl der natürlichen Zahlen p, q mit $\frac{p}{q} = r$.

Wir erweitern im Fall einer positiven reellen Basis a die Definition der Potenz auf den Fall negativer rationaler Exponenten und setzen

$$a^{-r} := \frac{1}{a^r} \quad (a > 0, r \text{ rational}, r > 0).$$

Damit sind Potenzen a^r in folgenden Fällen definiert:

	Basis	Exponent
a)	reell	natürliche Zahl
b)	reell, von Null verschieden	ganz
c)	reell, nichtnegativ	rational, nichtnegativ
d)	reell, positiv	rational

Die Graphen der Funktionen $y = x^r$ sind in den genannten vier Fällen für einige Werte des Exponenten r in Abb. 1.10 dargestellt (vgl. Satz 2).

Die Potenzgesetze

$$a^r b^r = (ab)^r, \quad a^r a^s = a^{r+s}, \quad (a^r)^s = a^{rs} \quad (3)$$

gelten auch dann, wenn für die Basen a, b und die Exponenten r, s jeweils der Fall c) oder jeweils der Fall d) vorliegt.

Ist $a = 0$ bzw. $b = 0$ (Fall c)), so ist (3) sofort klar. Es sei also jetzt $a, b > 0$. Dann genügt es, die Gültigkeit von (3) für den Fall d) zu zeigen. Zuerst beweisen wir die letzte Behauptung von (3) für $s = m$, wobei m eine ganze Zahl ist. Wir unterscheiden drei Fälle.

Fall 1. Es ist $m \in \mathbf{N}$, $r = \frac{p}{q}$ ($p, q \in \mathbf{N}$, $q \neq 0$). Aus der Gültigkeit der Potenzgesetze für ganzzahlige Exponenten und aus (1), (2) folgt

$$((a^r)^m)^q = (a^r)^{mq} = ((a^{p/q})^q)^m = (a^p)^m = a^{pm} = (a^{pm/q})^q = (a^{rm})^q.$$

Aus 1.1.2., Satz 8, folgt $(a^r)^m = a^{rm}$.

Fall 2. Es ist $m \in \mathbf{N}$, $r = -\frac{p}{q}$ ($p, q \in \mathbf{N}$, $q \neq 0$). Mit $t = -r$ gilt dann

$$(a^r)^m = \left(\frac{1}{a^t}\right)^m = \frac{1}{(a^t)^m} = \frac{1}{a^{tm}} = a^{rm}.$$

Fall 3. Es ist $m = -k$ mit $k \in \mathbf{N}$, r rational. Dann ist

$$(a^r)^m = \frac{1}{(a^r)^k} = \frac{1}{a^{rk}} = a^{rm}.$$

Damit ist die letzte Behauptung von (3) für den Fall $s = m$ (m ganz) bewiesen.

Wir wählen nun eine natürliche Zahl m derart, daß mr und ms ganze Zahlen sind. Auf Grund des bereits bewiesenen Teiles unserer Behauptung gilt

$$\begin{aligned}(a^r b^r)^m &= (a^r)^m (b^r)^m = a^{rm} b^{rm} = (ab)^{rm} = ((ab)^r)^m, \\ (a^r a^s)^m &= (a^r)^m (a^s)^m = a^{rm} a^{sm} = a^{rm+sm} = a^{(r+s)m} = (a^{r+s})^m, \\ ((a^r)^s)^m &= (a^r)^{sm} = a^{rsm} = (a^{rs})^m.\end{aligned}$$

Aus 1.1.2., Satz 8, folgen nun wiederum die Behauptungen (3).

Wir leiten einige Ungleichungen her.

Satz 1. Für alle positiven reellen Zahlen x, y mit $x \neq y$ und für alle von 0, 1 verschiedenen rationalen Zahlen r gilt

$$\frac{x^{r-1}(y-x)}{r-1} < \frac{y^r - x^r}{r(r-1)} < \frac{y^{r-1}(y-x)}{r-1}. \quad (4)$$

Beweis. Wir ersetzen x bzw. y in 1.1.2.(27) durch $x^{1/n}$ bzw. $y^{1/n}$ und wählen m, n derart, daß

$$r-1 = \frac{m}{n}, \quad m = n(r-1), \quad m+n = nr$$

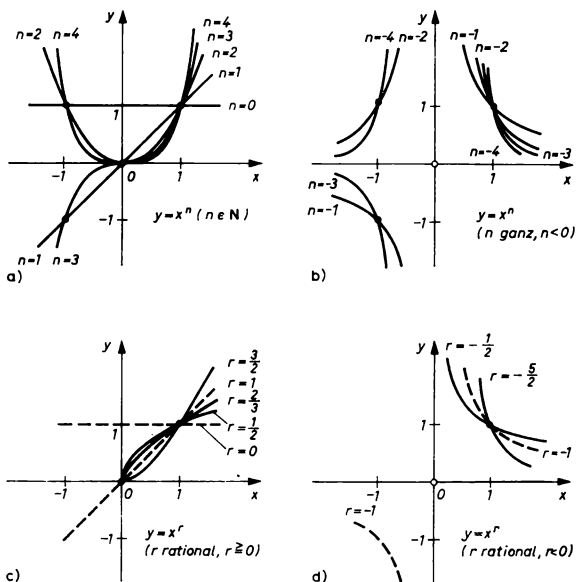


Abb. 1.10

ist. Wir erhalten

$$\frac{x^{r-1}y}{n^2(r-1)} < \frac{x^r}{n^2r} + \frac{y^r}{n^2r(r-1)},$$

$$\frac{x^{r-1}(y-x)}{r-1} < x^r \left(\frac{1}{r} - \frac{1}{r-1} \right) + \frac{y^r}{r(r-1)} = \frac{y^r - x^r}{r(r-1)},$$

und die erste Ungleichung (4) ist bewiesen. Die zweite folgt aus der ersten, indem wir x mit y vertauschen und mit -1 multiplizieren.

Satz 2. Die Funktion

$$y = x^r \quad (x \in \mathbb{R}, x > 0)$$

ist für $r > 0$ bzw. $r < 0$ streng monoton wachsend bzw. streng monoton fallend (Abb. 1.10).

Beweis. Nach (4) haben $x_1 - x_2$ und $x_1^r - x_2^r$ gleiches oder entgegengesetztes Signum, je nachdem, ob $r > 0$ oder $r < 0$ ist.

Die Bernoullische Ungleichung kann auf rationale Exponenten ausgedehnt werden.

Satz 3. Für alle reellen Zahlen h mit $1 + h > 0$ und $h \neq 0$ und für alle rationalen Zahlen r mit $r > 1$ ist

$$(1 + h)^r > 1 + rh. \quad (5)$$

Beweis. In der ersten Ungleichung (4) setzen wir $x = 1$, $y = 1 + h$, was wegen $1 + h \neq 1$, $1 + h > 0$ zulässig ist. Nach Multiplikation mit $r(r - 1)$ erhalten wir $rh < (1 + h)^r - 1$ und damit die Behauptung.

Satz 4. Genügt die positive reelle Zahl x den Ungleichungen

$$x, x^{-1} \leq Q \quad (6)$$

und die rationale Zahl r der Ungleichung

$$|r| \leq R \quad (7)$$

mit rationalem R , so gilt

$$x^r \leq Q^R, \quad (8)$$

$$|x^r - 1| \leq Q^{R+1} |r| |x - 1|. \quad (9)$$

Beweis. Ist $x = 1$ oder $r = 0$, so sind die Behauptungen richtig. Es sei $x \neq 1$, $r \neq 0$. Wegen (6) ist stets $Q \geq 1$, und aus Satz 2 folgt $Q^s \geq 1$ für alle nichtnegativen rationalen Zahlen s . Für $r > 0$ ist daher $x^r \leq Q^r \leq Q^r \cdot Q^{R-r} = Q^R$. Für $r < 0$ folgt $x^r = (x^{-1})^{|r|} \leq Q^R$, und (8) ist bewiesen. Die Behauptung (9) gilt für $r = 1$. Wenn $r \neq 0, 1$ ist, folgt aus (4), wenn wir $x^{r-1} = x^r \cdot x^{-1} \leq Q^R \cdot Q = Q^{R+1}$ berücksichtigen,

$$\begin{aligned} \frac{x-1}{r-1} &< \frac{x^r-1}{r(r-1)} < \frac{x^{r-1}(x-1)}{r-1}, \\ \left| \frac{x^r-1}{r(r-1)} \right| &\leq \max \left\{ \frac{|x-1|}{|r-1|}, |x|^{r-1} \frac{|x-1|}{|r-1|} \right\} \leq Q^{R+1} \frac{|x-1|}{|r-1|}, \end{aligned}$$

womit auch (9) bewiesen ist.

Satz 5. Für alle positiven reellen Zahlen x, y und alle rationalen Zahlen $p, q > 1$ mit

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (10)$$

gilt

$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy. \quad (11)$$

Das Gleichheitszeichen gilt genau dann, wenn $x^p = y^q$ ist.

Beweis. Es gibt natürliche Zahlen k, m, n mit

$$p = \frac{k}{m}, \quad q = \frac{k}{n}.$$

Wegen (10) ist $m + n = k$, und es folgt

$$d := \frac{x^p}{p} + \frac{y^q}{q} - xy = k \left(mx^{\frac{m+n}{m}} + ny^{\frac{m+n}{n}} - (m+n)xy \right).$$

Mit $a := x^{\frac{1}{m}}$, $b := y^{\frac{1}{n}}$ ergibt sich auf Grund von Satz 10 in 1.1.2.

$$d = k(ma^{m+n} + nb^{m+n} - (m+n)a^mb^n) \geq 0.$$

Das Gleichheitszeichen gilt genau dann, wenn $a = b$ bzw. $x^{\frac{1}{m}} = y^{\frac{1}{n}}$ bzw. $x^p = y^q$ ist.

Mit Hilfe von Grenzwertsätzen für Zahlenfolgen und der Relation 2.3.1.(5) kann Satz 5 auf den Fall reeller Exponenten p, q übertragen werden.

1.4.2. Exponentialfunktionen

Für jede positive reelle Zahl a ist die Funktion

$$f(r) = a^r \quad (r \in \mathbb{Q}) \quad (1)$$

für alle rationalen Zahlen r definiert und überall positiv.

Die Funktion (1) ist für $a > 0$, $a \neq 1$ streng monoton, denn nach 1.4.1., Satz 2, haben $a^r - 1$ und $a - 1$ für $r > 0$ gleiches Signum. Aus $r < s$ folgt somit

$$(a^s - a^r)(a - 1) = a^r(a^{s-r} - 1)(a - 1) > 0,$$

d. h., die Funktion (1) ist für $0 < a < 1$ bzw. $a > 1$ streng monoton fallend bzw. wachsend.

Da \mathbb{R} das kleinste Intervall ist, das alle rationalen Zahlen umfaßt, können wir f zu einer für alle reellen Zahlen definierten monotonen Funktion \tilde{f} fortsetzen. Wir setzen $a^x := \tilde{f}(x)$, d. h., es sei

$$a^x = \sup \{a^t : t \in \mathbb{Q} \wedge t \leq x\} \quad (x \in \mathbb{R}, a > 1), \quad (2)$$

$$a^x = \sup \{a^t : t \in \mathbb{Q} \wedge t \geq x\} \quad (x \in \mathbb{R}, 0 < a < 1). \quad (3)$$

Im Fall $a = 1$ setzen wir $1^x := 1$ für alle $x \in \mathbb{R}$.

Definition. Ist a eine von 1 verschiedene positive Zahl, so heißt die Funktion

$$y = a^x \quad (x \in \mathbb{R})$$

die *Exponentialfunktion zur Basis a* .

Satz 1. Die Exponentialfunktion zur Basis a ($a > 0$, $a \neq 1$) ist für $a > 1$ bzw. $0 < a < 1$ streng monoton wachsend bzw. fallend, und ihr Wertebereich ist die Menge aller positiven reellen Zahlen (Abb. 1.11).

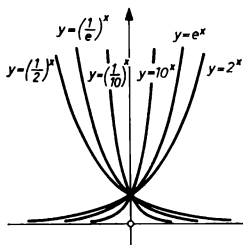


Abb. 1.11

Beweis. Es sei zunächst $a > 1$. Der Definitionsbereich der durch (1) definierten Funktion f liegt dicht in \mathbb{R} , und es ist $W(f) \subseteq]0, \infty[$. Es seien b, c zwei beliebige positive Zahlen mit $b < c$. Mit $h := a - 1$ gilt nach der Bernoullischen Ungleichung

$$f(n) = (1 + h)^n > 1 + nh, \quad f(-n) = \frac{1}{(1 + h)^n} < \frac{1}{1 + nh}$$

für alle natürlichen Zahlen $n \geq 2$. Es gibt deshalb zu b, c ein festes n mit $f(-n) < b < c < f(n)$, und folglich ist $\mathbb{R}_+^* =]0, \infty[$ das kleinste Intervall, das $W(f)$ umfaßt. Für alle rationalen Zahlen $r, s \in \llbracket -n, n \rrbracket$ ist nach Satz 3 in 1.4.1. stets

$$|f(r) - f(s)| = a^r |a^{s-r} - 1| \leq a^n a^{2n+1} |r - s| |a - 1|,$$

$$|f(r) - f(s)| \leq K |r - s| \quad (r, s \in \llbracket -n, n \rrbracket).$$

Aus Satz 6 in 1.3.5. schließen wir, daß der Wertebereich von f dicht in $\llbracket b, c \rrbracket$ ist. Da b, c beliebig gewählt war, ist $W(f)$ dicht in \mathbb{R}_+^* . Im Fall $0 < a < 1$ folgt dies aus $a^r = \left(\frac{1}{a}\right)^{-r}$.

Aus Satz 4 bzw. Satz 5 in 1.3.5. folgt nun, daß die Exponentialfunktion streng monoton bzw. daß ihr Wertebereich das Intervall \mathbb{R}_+^* ist.

Satz 2. Für jede positive Basis a gilt das Potenzgesetz

$$a^{x \cdot y} = a^x a^y \quad (x, y \in \mathbb{R}). \quad (4)$$

Beweis. Wir könnten einen Beweis direkt auf die Definition (2), (3) stützen, indem wir etwa die Annahme $a^{x+y} < a^x a^y$ im Fall $a > 1$ dadurch zum Widerspruch führen, daß wir zwei rationale Zahlen r, s mit $x < r$, $y < s$ und $a^{r+s} < a^x a^y$ konstruieren. Der Beweis gestaltet sich aber mit unseren Sätzen über monotone Funktionen viel einfacher.

Für festes $s \in \mathbb{Q}$ ist die Funktion

$$g(x) := \frac{a^{x+s}}{a^s} \quad (x \in \mathbb{R})$$

eine monotone Fortsetzung der Funktion

$$g(r) := \frac{a^{r+s}}{a^s} = a^r \quad (r \in \mathbb{Q}),$$

und folglich ist $g(x) = a^x$. Daher ist $a^{x+s} = a^x a^s$ für alle $x \in \mathbb{R}$, $s \in \mathbb{Q}$. Für festes $y \in \mathbb{R}$ ist die Funktion

$$h(x) := \frac{a^{x+y}}{a^y} \quad (x \in \mathbb{R})$$

eine monotone Fortsetzung der Funktion $h(r) := \frac{a^{r+y}}{a^y} = a^r$, woraus $h(x) = a^x$, $a^{x+y} = a^x a^y$ folgt.

Jede Exponentialfunktion $f(x) = a^x$ erfüllt also die Gleichung

$$f(x+y) = f(x)f(y) \quad (x, y \in \mathbb{R}). \quad (5)$$

Sie heißt die *Funktionalgleichung der Exponentialfunktion*.

Die Funktionalgleichung (5) wird offensichtlich von der Funktion

$$f(x) = 0 \quad (x \in \mathbb{R}) \quad (6)$$

erfüllt. Jede von der Nullfunktion verschiedene, auf \mathbb{R} definierte reellwertige Funktion f , die der Funktionalgleichung (5) genügt, wollen wir eine *verallgemeinerte Exponentialfunktion* nennen.

Satz 3. Für jede verallgemeinerte Exponentialfunktion f gilt

$$f(0) = 1, \quad (7)$$

$$f(x) > 0 \quad (x \in \mathbb{R}), \quad (8)$$

$$f(rx) = f(x)^r \quad (x \in \mathbb{R}, r \in \mathbb{Q}). \quad (9)$$

Beweis. Nach Voraussetzung gibt es ein c mit $f(c) \neq 0$, und aus der Gleichung $f(c)f(0) = f(c+0) = f(c)$ folgt (7).

Aus $f(x)f(-x) = f(x-x) = f(0) = 1$ folgt $f(x) \neq 0$ für alle x , und wegen der Beziehung $f(x) = f\left(\frac{x}{2}\right)f\left(\frac{x}{2}\right) = \left(f\left(\frac{x}{2}\right)\right)^2$ gilt sogar (8).

Wir zeigen, daß

$$f(nx) = f(x)^n \quad (10)$$

ist. Die Behauptung ist nach (7) richtig für $n = 0$, und aus (10) folgt

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

womit die Behauptung (10) durch vollständige Induktion für alle $n \in \mathbf{N}$ bewiesen ist.

Aus

$$f(-nx)f(x)^n = f(-nx)f(nx) = f(0) = 1$$

folgt $f(-nx) = f(x)^{-n}$, und (10) gilt für alle ganzen Zahlen n .

Sind m, n beliebige ganze Zahlen mit $n \neq 0$ so ist

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

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

Damit gilt (9), und Satz 3 ist bewiesen.

Satz 4. Ist eine verallgemeinerte Exponentialfunktion f in einem (beliebig kleinen) Intervall beschränkt, so ist

$$f(x) = f(1)^x \quad (x \in \mathbf{R}). \quad (11)$$

Beweis. Es sei $a := f(1)$. Die Funktion

$$g(x) := a^x f(-x)$$

ist offensichtlich wieder eine verallgemeinerte Exponentialfunktion. Nach Satz 3 ist für $r \in \mathbf{Q}$ stets

$$g(r) = f(1)^r f(-r) = f(r)f(-r) = f(0) = 1.$$

Nach Voraussetzung gibt es ein Intervall I und eine Zahl $K_1 > 0$ mit $f(x) \leq K_1$ für $x \in I$. Es sei $\llbracket b, c \rrbracket \subseteq I$ und $b < c$. Da auch a^x in diesem Intervall beschränkt ist, gibt es ein $K > 0$ mit $g(x) \leq K$ für alle $x \in \llbracket b, c \rrbracket$. Wir nehmen nun an, es sei $g(x) \neq 1$. Wir können $g(x) > 1$ voraussetzen. Andernfalls ersetzen wir x durch $-x$. Wir wählen eine natürliche Zahl n mit $g(nx) = g(x)^n > K$ und eine rationale Zahl $r \in \llbracket b - nx, c - nx \rrbracket$. Dann ist $r + nx \in \llbracket b, c \rrbracket$, und wegen $g(r) = 1$ erhalten wir den Widerspruch

$$K < g(x)^n = g(nx) = g(nx)g(r) = g(nx+r) \leq K.$$

Somit ist $a^x/(-x) = g(x) = 1$ für alle $x \in \mathbf{R}$. Multiplizieren wir mit $f(x)$, so erhalten wir die Behauptung (11).

Insbesondere ist auf Grund dieses Satzes jede streng monotone verallgemeinerte Exponentialfunktion eine eigentliche Exponentialfunktion, denn eine monotone Funktion ist auf jedem Intervall $\llbracket a, b \rrbracket \subseteq D(f)$ durch $f(a)$ oder $f(b)$ nach oben bzw. unten beschränkt.

Jetzt können die restlichen Potenzgesetze leicht bewiesen werden.

Satz 5. Für alle positiven Zahlen a, b ist

$$\begin{aligned} a^x b^x &= (ab)^x & (x \in \mathbf{R}), \\ (a^x)^y &= a^{xy} & (x, y \in \mathbf{R}). \end{aligned}$$

Beweis. Durch $f(x) := a^x b^x$ bzw. $g(x) := a^{cx}$ mit festem $c \in \mathbf{R}$ werden offensichtlich verallgemeinerte Exponentialfunktionen definiert, die im Intervall $\llbracket 0, 1 \rrbracket$ beschränkt sind. Folglich ist $a^x b^x = f(1)^x = (ab)^x$ und $a^{cx} = g(1)^x = (a^c)^x$ für alle $x \in \mathbf{R}$, und Satz 5 ist bewiesen.

Satz 6. Genügt eine auf einem Intervall $\llbracket b, c \rrbracket$ beschränkte Funktion f der Funktionalgleichung

$$f(x+y) = f(x) + f(y),$$

so ist $f(x) = xf(1)$.

Beweis. Für $a > 1$ ist die Funktion $g(x) := a^{f(x)}$ eine auf einem gewissen Intervall beschränkte verallgemeinerte Exponentialfunktion. Nach Satz 4 ist $a^{f(x)} = g(x) = g(1)^x = (a^{f(1)})^x = a^{xf(1)}$. Aus der strengen Monotonie der Exponentialfunktion folgt die Behauptung. Der Beweis kann allerdings auch ohne die Exponentialfunktion geführt werden.

1.4.3. Logarithmusfunktionen

Definition. Die Umkehrfunktion der Exponentialfunktion zur Basis a ($a > 0$, $a \neq 1$) heißt die *Logarithmusfunktion zur Basis a* . Sie wird mit

$$y = \log_a x \quad (x \in \mathbf{R}, x > 0) \tag{1}$$

bezeichnet (Abb. 1.12).

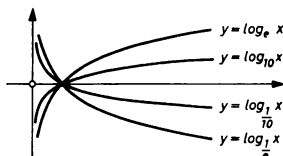


Abb. 1.12

Für jede positive reelle Zahl a mit $a \neq 1$ gilt hiernach

$$y = \log_a x \Leftrightarrow x = a^y \quad (x, y \in \mathbf{R}, x > 0). \quad (2)$$

Im Fall $a = 10$ setzt man

$$\lg x := \log_{10} x \quad (x > 0) \quad (3)$$

und nennt den Funktionswert den *dekadischen Logarithmus* der positiven reellen Zahl x .

Die von der Schule her bekannte Bedeutung der Logarithmusfunktion für numerische Rechnungen beruht auf dem *Logarithmengesetz*

$$\log_a(xy) = \log_a x + \log_a y \quad (a, x, y \in \mathbf{R}; a, x, y > 0), \quad (4)$$

das die Zurückführung des Multiplizierens auf das Addieren gestattet. Zum Beweis setzen wir

$$p := \log_a x, \quad q := \log_a y.$$

Dann ist $a^p = x$, $a^q = y$,

$$xy = a^p a^q = a^{p+q}, \quad p + q = \log_a(xy),$$

und (4) ist bewiesen.

Aus

$$(a^{x \log_a b})^x = (a^{\log_a b})^x = b^x$$

folgt das Logarithmengesetz

$$\log_a(b^x) = x \log_a b \quad (a, b, x \in \mathbf{R}; a, b > 0). \quad (5)$$

Setzen wir $g(x) := \log_a x$, so nimmt (4) die Gestalt

$$g(xy) = g(x) + g(y) \quad (x, y \in \mathbf{R}; x, y > 0) \quad (6)$$

an. Diese Gleichung heißt die *Funktionalgleichung der Logarithmusfunktion*. Eine von der Nullfunktion verschiedene Funktion g , die dieser Funktionalgleichung für alle positiven x, y genügt, nennen wir eine *verallgemeinerte Logarithmusfunktion*. Mit ähnlichen Methoden wie im vorigen Abschnitt kann gezeigt werden, daß für eine solche Funktion stets

$$g(b^r) = rg(b) \quad (b \in \mathbf{R}, b > 0, r \in \mathbf{Q})$$

ist. Wir begnügen uns mit dem

Satz. Ist eine verallgemeinerte Logarithmusfunktion g in einem (beliebig kleinen) Intervall beschränkt, so ist g die Logarithmusfunktion zur Basis

$$a := \frac{1}{b^{g(b)}}, \quad (7)$$

wobei b eine beliebige von 1 verschiedene positive Zahl ist.

Beweis. Für beliebige $b > 0$, $b \neq 1$ ist die Funktion $f(x) := g(b^x)$ auf einem Intervall beschränkt. Ferner ist

$$f(x+y) = g(b^{x+y}) = g(b^x b^y) = g(b^x) + g(b^y) = f(x) + f(y).$$

Nach 1.4.2., Satz 6, ist $f(x) = x f(1)$, d. h., es ist

$$g(b^x) = x g(b). \quad (8)$$

Da g nicht die Nullfunktion ist, muß $g(b) \neq 0$ sein. Ist a durch (7) definiert, so folgt aus (8)

$$g(a^x) = g\left(b^{x \log(b)a}\right) = \frac{x}{g(b)} \cdot g(b) = x.$$

Somit ist g die Umkehrfunktion der Exponentialfunktion zur Basis a , und der Satz ist bewiesen.

Für eine beliebige reelle Zahl α definieren wir die *allgemeine Potenzfunktion*

$$y = x^\alpha \quad (x \in \mathbb{R}, x > 0). \quad (9)$$

Aus $0 < x < y$ und $\alpha \geq 0$ folgt $\lg x < \lg y$ und

$$\lg x^\alpha = \alpha \lg x \leq \alpha \lg y, \quad x^\alpha \leq y^\alpha,$$

d. h., die allgemeine Potenzfunktion ist für $\alpha > 0$ bzw. $\alpha < 0$ streng monoton wachsend bzw. fallend.

Die Kenntnis von höheren Potenzen der natürlichen Zahlen kann bis nach Babylonien (3. Jahrtausend v. u. Z.) zurückverfolgt werden, und aus Kleinasien ist dieses Wissen auch nach Griechenland gelangt. Während die babylonische Mathematik eine stark arithmetisch-algebraische Tendenz hatte, war der vorherrschende Zug der griechisch-hellenistischen Mathematik geometrischer Art. Man betrieb neben reiner Geometrie auch geometrische Algebra. Da alle algebraischen Probleme in geometrischer Form behandelt wurden, waren höhere Potenzen als die dritte, weil keine geometrische Interpretation zur Verfügung stand, für fast alle griechisch-hellenistischen Mathematiker nicht Gegenstand von Untersuchungen. Nur gelegentlich, so bei HERON (um 130 v. u. Z.) und dem Zahlentheoretiker DIOPHANTOS (um 250), treten auch höhere Potenzen auf.

Durch Übersetzungen, hauptsächlich arabischer Werke, gelangte das vorhandene Wissen in das mittelalterliche Europa und wurde zuerst in Italien fruchtbar. Hier findet sich auch bei LUCA PACIOLI (1445–1514?) die Erweiterung der Potenz auf beliebig hohe (natürliche) Exponenten. Um 1450 begannen, aufbauend auf italienischen Vorbildern, deutsche Mathematiker die Potenzrechnung weiterzuentwickeln. Es entstand die „Coß“ (von ital. cosa, d. h. Unbekannte). Die Cossisten lösten zwei Grundaufgaben: 1. Sie trennten die Potenzrechnung von der Gleichungslehre; 2. sie schufen sich zur Bezeichnung der Potenzen eine eigene Zeichensprache. Die heutige Art des Schreibens der Potenzen findet sich ausgebildet 1637 bei DESCARTES, nachdem viele Vorstufen durchlaufen worden sind.

Es bleibt noch anzugeben, seit wann die Exponenten nicht notwendig als natürliche Zahlen aufgefaßt wurden. Rationale Exponenten verwendete NICOLE ORESME (um 1323–1382). Seine Schreibweise war für heutige Begriffe höchst ungewöhnlich und schwerfällig. Der Exponent Null tritt, ebenso wie negative ganzzahlige Exponenten, 1484 bei NICOLAS CHUQUET (gest. um 1500) auf. ORESME und CHUQUETs Werke haben jedoch auf den lebendigen Gang der Mathematik kaum Einfluß gehabt, da sie erst in der Neuzeit durch die mathematik-historische Forschung erschlossen wurden. Die modernere Potenzrechnung beginnt so eigent-

lich erst mit MICHAEL STIFEL (1487?–1567), von dem übrigens auch das Wort „Exponent“ stammt. Er gab 1544 eine Zusammenfassung bisheriger Kenntnisse der Potenzrechnung. Imaginäre und komplexe Exponenten treten dann 1740 und 1749 bei EULER in der Blütezeit der „formalen“ Analysis auf.

Nachdem bei DESCARTES und besonders bei ISAAC NEWTON (1643–1727) bereits allgemeine Exponenten auftraten, schien der Schritt nahezuliegen, auch zu variablen Exponenten überzugehen. Jedoch war dies erst möglich, nachdem die durch die Praxis geforderte Inangriffnahme der mathematischen Behandlung von Bewegungsproblemen erfolgt war und damit die Begriffe (in heutiger Sprechweise) „variable Größe“ und „Funktion“ zur Diskussion standen. Es ist deshalb kein Zufall, daß gerade LEIBNIZ, der maßgeblich an der Schaffung der Analysis beteiligt war, 1679 zu variablen Exponenten überging. Er behandelte Gleichungen wie $x^2 + z^x = b$. Die hier auftretenden Terme x^2 , z^x bezeichnete er später als „exponentialiter transcendentales“. Sie waren dann auch Gegenstand weiterer Untersuchungen von LEIBNIZ (1695) und JOHANN BERNOULLI (1697), ohne daß jedoch wesentliche Fortschritte erzielt wurden. Auch hier waren die Arbeiten EULERS bahnbrechend. Einen wirklichen Abschluß bedeuten erst die Untersuchungen NIELS HENRIK ABELS (1802–1829), der die Funktionalgleichung $f(x) \cdot f(y) = f(x + y)$ erschöpfend behandelte.

Die Geschichte der Logarithmen ist eng mit der Kenntnis geometrischer und arithmetischer Folgen verbunden. Solche Folgen waren seit der Antike bekannt. Das Vergleichen einer willkürlichen arithmetischen und einer willkürlichen geometrischen Folge findet sich in allen bedeutenden mathematischen Werken des Mittelalters. Vor allem STIFEL hat sich mit tieferen Untersuchungen beim Vergleich der beiden Folgen abgegeben und praktisch alle Grundgesetze des Rechnens mit Logarithmen erkannt.

Aber diese Erkenntnisse waren oft nur zufällig, eine praktische Bedeutung wurde ihnen kaum beigemessen. Das änderte sich entscheidend, als im 16. und 17. Jahrhundert rein rechnerische Arbeiten, insbesondere bei der Auswertung astronomischer Beobachtungen, außerordentlich an Umfang zunahmen. Die noch aus der Antike oder von den Arabern übernommenen Methoden und die später entwickelten Rechenverfahren waren zeitaufwendig. Das gesellschaftliche und wissenschaftliche Bedürfnis nach neuen Rechenhilfsmitteln führte nicht nur zur Entwicklung erster Rechenmaschinen, sondern auch zu den ersten Logarithmen. Die ersten Tafeln sind von JOST BÜRGI (1552–1632?), einem Bekannten des großen Astronomen und Mathematikers JOHANNES KEPLER (1571–1630), um 1600 berechnet worden. Sie wurden jedoch erst 1620 veröffentlicht, weil BÜRGI sich nur außerordentlich schwer zur Publikation seiner Untersuchungen entschließen konnte. BÜRGIS Tafeln enthalten die arithmetische Folge

mit $a_n = 10^n$ und die geometrische mit $a_n = 10^n \left(1 + \frac{1}{10^4}\right)^{n-1}$. Die Basis des Logarithmen-

systems (der Begriff ist BÜRGI völlig fremd) ist $1,001^{1000} \approx e$. Fast gleichzeitig mit BÜRGI beschäftigt sich der Schotte JOHN NEPER (1550–1617) mit der Aufstellung von logarithmischen Tafeln. Seine erste Tafel erschien 1614. Auf den Titel dieser Tafel geht auch das Wort „Logarithmus“ zurück. Die Neperische Tafel ist eine logarithmisch-trigonometrische Tafel mit einer Basis, die angenähert gleich e^{-1} ist.

Auch KEPLER berechnete „nach der schottischen Manier“ eigene Tafeln und wurde selbst zum größten Fürsprecher des logarithmischen Rechnens.

Sehr schnell wurden die BÜRGISchen und NEPERschen Logarithmen durch die dekadischen verdrängt, da diese erheblich einfacher zu handhaben waren. Es ist das Verdienst von HENRY BRIGGS (1556–1630), einem Professor am Gresham College in London, der mit NEPER gut bekannt war, solche Tafeln erstmals berechnet zu haben. Seine Tafeln erschienen 1637. Die folgende Zeit ist gekennzeichnet durch das Bestreben, die Genauigkeit der Tafeln zu erhöhen und ihre Form günstiger zu gestalten. Gleichzeitig damit mußten die Berechnungsmethoden für Logarithmen verbessert werden. Entscheidend gelang dies erst mit der Entdeckung der logarithmischen Reihe 1667 durch NICOLAUS MERCATOR (1620–1687), deren Theorie bald sorgfältig ausgearbeitet wurde.

Wer die logarithmische Kurve zuerst betrachtet hat, ist unbekannt. Die erste bekannte Arbeit über sie stammt von CHRISTIAN HUYGENS (1629–1695), der sich auf ungenannte Vorgänger bezieht.

1.5. Der euklidische Raum \mathbb{R}^p

1.5.1. Der zweidimensionale euklidische Raum

Die Menge $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ aller geordneten Paare reeller Zahlen kann nach Wahl eines Koordinatensystems umkehrbar eindeutig auf die Menge aller Punkte einer Ebene abgebildet werden. In der analytischen Geometrie und linearen Algebra wird in der Menge \mathbb{R}^2 eine zweistellige Operation, die Addition, definiert. Für zwei Elemente x, y von \mathbb{R}^2 mit $x = (x_1, x_2)$, $y = (y_1, y_2)$ setzt man

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2). \quad (1)$$

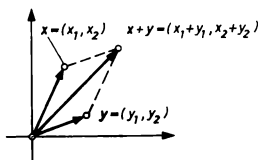


Abb. 1.13

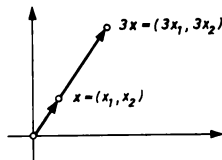


Abb. 1.14

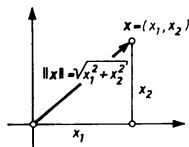


Abb. 1.15

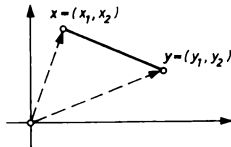


Abb. 1.16

In Abb. 1.13 wird diese Operation geometrisch veranschaulicht. Sie ist ebenso wie die Addition reeller Zahlen kommutativ, assoziativ und umkehrbar, so daß auch die weiteren Eigenschaften der Addition reeller Zahlen sinngemäß übertragen werden können. Insbesondere ist $x - y = (x_1 - y_1, x_2 - y_2)$.

Ferner setzt man

$$t(x_1, x_2) := (tx_1, tx_2) \quad (2)$$

für alle reellen Zahlen t, x_1, x_2 . Durch (2) wird offenbar eine eindeutige Abbildung von $\mathbb{R} \times \mathbb{R}^2$ auf \mathbb{R}^2 definiert. Sie ist in Abb. 1.14 veranschaulicht. Ihre Eigenschaften wollen wir hier ebenfalls als bekannt voraussetzen.

Schließlich wollen wir noch eine eindeutige Abbildung von \mathbb{R}^2 auf die Menge der nichtnegativen reellen Zahlen definieren, und zwar ordnen wir jedem Element x

von \mathbf{R}^2 den Abstand des zugeordneten Punktes vom Ursprung des Koordinatensystems zu. Wir bezeichnen diesen Abstand mit $\|\mathbf{x}\|$ (gelesen: *Norm* von \mathbf{x}).

Für $\mathbf{x} = (x_1, x_2)$ gilt nach dem pythagoreischen Lehrsatz

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2} \quad (3)$$

(Abb. 1.15). Hierdurch wird eine Funktion von \mathbf{R}^2 in \mathbf{R}_+ definiert. Sind \mathbf{x}, \mathbf{y} zwei Elemente von \mathbf{R}^2 und ist $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$, so haben die Bildpunkte von \mathbf{x} und \mathbf{y} den Abstand

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (4)$$

(Abb. 1.16). Der Abb. 1.13 entnehmen wir die Dreiecksungleichung

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|. \quad (5)$$

Sie besagt, daß im Dreieck die Summe zweier Dreiecksseiten nicht kleiner als die dritte Dreiecksseite ist. Einen von der Anschauung unabhängigen Beweis dieser Ungleichung führen wir in 1.5.2. Wir nennen die Menge \mathbf{R}^2 aller geordneten Paare reeller Zahlen in Verbindung mit den durch (1), (2) und (3) definierten Operationen bzw. Funktionen den *zweidimensionalen euklidischen Raum*¹⁾.

1.5.2. Der p -dimensionale euklidische Raum

Die Definitionen (1), (2), (3) in 1.5.1. können sinngemäß auf die Menge \mathbf{R}^3 übertragen werden, und auch hier ist eine analoge geometrische Deutung möglich. So ist z. B. $\sqrt{x_1^2 + x_2^2 + x_3^2}$ der Abstand des Bildpunktes eines Elementes \mathbf{x} aus \mathbf{R}^3 mit $\mathbf{x} = (x_1, x_2, x_3)$ vom Ursprung des Koordinatensystems (Abb. 1.17), und $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$ ist der Abstand der Bildpunkte der Elemente \mathbf{x}, \mathbf{y} aus \mathbf{R}^3 mit $\mathbf{x} = (x_1, x_2, x_3)$ und $\mathbf{y} = (y_1, y_2, y_3)$.

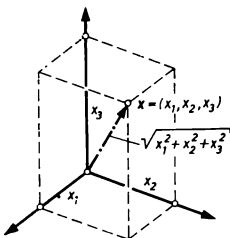


Abb. 1.17

¹⁾ In den ersten Auflagen wurde hierfür das Symbol \mathbf{R}_2 verwendet.

Wir können zwar nicht die geometrische Veranschaulichung, wohl aber die abstrakten Definitionen (1), (2), (3) von 1.5.1. und die entsprechende Terminologie auch auf den Fall der Menge \mathbf{R}^p mit $p \in \mathbf{N}$ und $p > 3$ übertragen.

Definition 1. Unter dem *p-dimensionalen euklidischen Raum* \mathbf{R}^p mit $p \geq 1$ versteht man die Menge aller geordneten *p*-Tupel reeller Zahlen in Verbindung mit den durch

$$(x_1, x_2, \dots, x_p) + (y_1, y_2, \dots, y_p) = (x_1 + y_1, x_2 + y_2, \dots, x_p + y_p), \quad (1)$$

$$t(x_1, x_2, \dots, x_p) = (tx_1, tx_2, \dots, tx_p) \quad (t \in \mathbf{R}), \quad (2)$$

$$\|(x_1, x_2, \dots, x_p)\| = \sqrt{x_1^2 + x_2^2 + \dots + x_p^2} \quad (3)$$

definierten Operationen bzw. Funktionen.

Die durch (3) für jedes x mit $x \in \mathbf{R}^p$ definierte nichtnegative reelle Zahl $\|x\|$ heißt die *Norm* von x . Das Element $(0, 0, \dots, 0)$ des Raumes \mathbf{R}_p bezeichnen wir im Fall $p > 1$ mit 0 . Offenbar ist $\|x\| = 0$ genau dann, wenn $x = 0$ ist. Die Elemente des eindimensionalen euklidischen Raumes \mathbf{R}^1 sind die reellen Zahlen, und die Norm eines Elements x mit $x \in \mathbf{R}^1$ stimmt wegen 1.1.5.(1) mit dem Betrag der reellen Zahl x überein.

Wir leiten aus der Definition des Raumes \mathbf{R}^p einige Rechenregeln ab.

Für alle t, x mit $t \in \mathbf{R}$ und $x \in \mathbf{R}^p$ gilt

$$\|tx\| = |t| \|x\|. \quad (4)$$

Satz 1. Für alle reellen Zahlen x_1, x_2, \dots, x_p und y_1, y_2, \dots, y_p gilt

$$\left| \sum_{i=1}^p x_i y_i \right| \leq \sqrt{\sum_{i=1}^p x_i^2} \sqrt{\sum_{i=1}^p y_i^2} \quad (5)$$

(Schwarzsche oder Bunjakovskijsche Ungleichung; HERMANN AMANDUS SCHWARZ (1843–1921), VIKTOR JAKOVLEVIČ BUNJAKOVSKIJ (1804–1889)).

Beweis. Es sei $x = (x_1, x_2, \dots, x_p)$ und $y = (y_1, y_2, \dots, y_p)$. Für $x = 0$ oder $y = 0$ ist (5) erfüllt. Ist $x, y \neq 0$, so setzen wir

$$a_i := \frac{x_i}{\|x\|}, \quad b_i := \frac{y_i}{\|y\|}.$$

Dann ist

$$\sum_{i=1}^p a_i^2 = \frac{1}{\|x\|^2} \sum_{i=1}^p x_i^2 = 1, \quad \sum_{i=1}^p b_i^2 = 1, \quad (6)$$

und aus (6) und 1.1.2., Satz 6, folgt

$$\pm \sum_{i=1}^p x_i y_i = \pm \sum_{i=1}^p \|x\| a_i \|y\| b_i = \|x\| \|y\| \sum_{i=1}^p (\pm a_i b_i) \leq \|x\| \|y\| \sum_{i=1}^p \frac{a_i^2 + b_i^2}{2}.$$

Wegen (6) ist die letzte Summe gleich 1, und es folgt

$$\pm \sum_{i=1}^p x_i y_i \leq \|x\| \|y\|,$$

was mit (5) gleichbedeutend ist.

Führt man die Abkürzung

$$\langle x, y \rangle := \sum_{i=1}^p x_i y_i \quad (7)$$

ein, so nimmt die Schwarzsche Ungleichung die Gestalt

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (8)$$

an. Die durch (7) definierte reelle Zahl heißt das *Skalarprodukt* von x und y .

Satz 2. Für alle $x, y, z \in \mathbb{R}^p$ gelten die Dreiecksungleichungen

$$\|x \pm y\| = \|x\| + \|y\|, \quad (9)$$

$$|\|x\| - \|y\|| \leq \|x \pm y\|, \quad (10)$$

$$\|x - z\| \leq \|x - y\| + \|y - z\|. \quad (11)$$

Beweis. Wegen (1), (3), (8) ist

$$\|x \pm y\|^2 = \sum_{i=1}^p (x_i \pm y_i)^2 = \sum_{i=1}^p x_i^2 + \sum_{i=1}^p y_i^2 \pm 2 \sum_{i=1}^p x_i y_i, \quad (12)$$

$$\|x \pm y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2.$$

Da $\|x \pm y\|$ und $\|x\| + \|y\|$ nichtnegativ sind, folgt (9). Die Ungleichung (10) wird aus (9) ganz analog abgeleitet wie die Ungleichung 1.1.4.(8) aus 1.1.4.(7). Wählen wir in (9) das Zeichen $+$ und ersetzen x durch $x - y$ sowie y durch $y - z$, so erhalten wir (11).

Satz 3. Für alle x mit $x \in \mathbb{R}^p$ und $x = (x_1, x_2, \dots, x_p)$ gilt

$$|x_i| \leq \|x\| \leq |x_1| + |x_2| + \dots + |x_p| \quad (i = 1, \dots, p). \quad (13)$$

Beweis. Die erste Ungleichung folgt unmittelbar aus (3), und aus

$$x_1^2 + x_2^2 + \dots + x_p^2 \leq (|x_1| + |x_2| + \dots + |x_p|)^2$$

folgt die zweite Behauptung.

Wir führen nun noch einige spezielle Punktmengen im euklidischen Raum \mathbb{R}^p ein.

Definition 2. Unter der *Verbindungsstrecke* zweier Punkte x_1, x_2 aus \mathbb{R}^p verstehen wir die Menge

$$S[x_1, x_2] = \{x_1 + t(x_2 - x_1) : 0 \leq t \leq 1\}.$$

In den Räumen \mathbb{R}^1 , \mathbb{R}^2 , \mathbb{R}^3 kann der Begriff der Verbindungsstrecke anschaulich gedeutet werden (Abb. 1.18).

Definition 3. Eine Teilmenge M des euklidischen Raumes \mathbb{R}^p heißt *konvex*, wenn aus $x_1, x_2 \in M$ stets $S[x_1, x_2] \subseteq M$ folgt.

In Abb. 1.19 sind eine konvexe Punktmenge M_1 und eine nichtkonvexe Menge M_2 des Raumes \mathbb{R}^2 veranschaulicht.

Definition 4. Eine Teilmenge M des euklidischen Raumes \mathbb{R}^p heißt *sternförmig*, wenn es einen Punkt x_0 in M gibt, so daß aus $x \in M$ stets $S[x_0, x] \subseteq M$ folgt (Abb. 1.20).

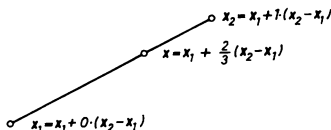


Abb. 1.18

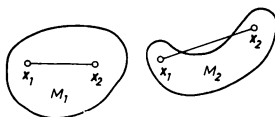


Abb. 1.19

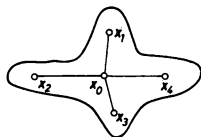


Abb. 1.20

Für das Hinausgehen über die drei Dimensionen des Anschauungsraumes auf n -dimensionale mathematische Räume bestand in der Antike und im Mittelalter beim niedrigen Stand der mathematischen Abstraktion keine Veranlassung. Erst mit der Behandlung von Integralen tauchten vereinzelt, etwa bei BLAISE PASCAL (1623–1662), Ideen über vierdimensionale geometrische Gebilde auf.

Von philosophischer Seite wurde die Möglichkeit einer n -dimensionalen Geometrie deutlich 1746 durch IMMANUEL KANT (1724–1804) in seiner Erstlingsschrift ausgesprochen. KANT hat allerdings später diese Festlegung zurückgenommen und die Ansicht vertreten, daß der dreidimensionale euklidische Raum denknötwendig sei. In den Werken von Mathematikern finden sich Betrachtungen über die Geometrie von mehr als drei Dimensionen spontan (JOSEPH LOUIS LAGRANGE (1736–1813), CARL GUSTAV JACOB JACOBI (1804–1851), CARL FRIEDRICH GAUSS (1777–1855)). Für den Aufbau einer n -dimensionalen Geometrie gab es im 19. Jahrhundert zwei Ausgangspunkte, einerseits gedankliche Kühnheit (ARTHUR CAYLEY (1821 bis 1895), HERMANN GRASSMANN (1809–1877)) und andererseits die Verbindung von Mathematik und Physik, bei der mehr als zwei unabhängige Größen gebraucht wurden.

CAYLEY und insbesondere GRASSMANN führten 1843 bzw. 1844 in höchst abstrakter Weise n -dimensionale affine Mannigfaltigkeiten ein. Diese Arbeiten wurden mit Unglauben und Mißtrauen, auch von mathematischer Seite, aufgenommen. Erst BERNHARD RIEMANN (1826 bis 1866) bewirkte mit seinem berühmten Habilitationsvortrag von 1854 über die Grundlagen der Geometrie eine Änderung dieser Haltung. Es wurde jetzt klar zwischen realem Raum, der dreidimensional und Gegenstand der Naturforschung ist, und abstrakten mathematischen „Räumen“, deren „Dimension“ beliebig hoch sein kann, unterschieden.

1.5.3. Die Metrik des Raumes \mathbb{R}^p

Aus der analytischen Geometrie ist bekannt, daß der Abstand zweier Punkte x, y der Ebene mit den Koordinaten x_1, x_2 bzw. y_1, y_2 durch

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = \|(x_1 - y_1, x_2 - y_2)\| = \|x - y\|$$

gegeben ist. Eine analoge Formel gilt für Punkte des Raumes. Wir vereinbaren daher allgemein die

Definition 1. Unter dem *Abstand* zweier Punkte x, y des Raumes \mathbb{R}^p verstehen wir die Norm der Differenz von x und y , d. h. die Zahl $\|x - y\|$.

Der Abstand zweier Punkte wird auch häufig mit $\varrho(x, y)$ bezeichnet, so daß

$$\varrho(x, y) = \|x - y\| \quad (1)$$

ist. Im Fall $n = 1$ ist hiernach

$$\varrho(x, y) = |x - y|; \quad (2)$$

ebenso wird der Abstand zweier komplexer Zahlen definiert.

Satz 1. Für alle Elemente x, y, z des Raumes \mathbb{R}^p gilt

$$\varrho(x, y) \geq 0, \quad (3)$$

$$\varrho(x, y) = 0 \text{ genau dann, wenn } x = y, \quad (4)$$

$$\varrho(x, y) = \varrho(y, x), \quad (5)$$

$$\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z). \quad (6)$$

Der Beweis dieser Behauptungen wurde in 1.5.2. geführt.

Mit Hilfe des Abstandsbegriffs führen wir einen für die Analysis grundlegenden Begriff ein.

Definition 2. Es sei x_0 ein Punkt des Raumes \mathbb{R}^p und ε eine positive reelle Zahl. Dann heißt die Menge aller Punkte x mit $x \in \mathbb{R}^p$, deren Abstand vom Punkt x_0 kleiner als ε ist, die ε -Umgebung des Punktes x_0 . Sie wird mit $U_\varepsilon(x_0)$ bezeichnet. Der Punkt x_0 heißt der *Mittelpunkt*, die Zahl ε der *Radius* der ε -Umgebung $U_\varepsilon(x_0)$.

Nach der Definition ist

$$U_\varepsilon(x_0) = \{x: x \in \mathbb{R}^p \wedge \varrho(x_0, x) < \varepsilon\}.$$

Auf der Zahlengeraden (im \mathbb{R}^1) besteht die ε -Umgebung des Punktes x_0 aus allen reellen Zahlen x , die der Bedingung

$$|x - x_0| < \varepsilon, \quad x_0 - \varepsilon < x < x_0 + \varepsilon \quad (7)$$

genügen (Abb. 1.21). Sie bildet ein *offenes Intervall*.

In der Ebene bzw. im Raum (im \mathbb{R}^2 bzw. im \mathbb{R}^3) besteht die ε -Umgebung aus allen Punkten, die im Innern des Kreises bzw. der Kugel mit dem Mittelpunkt x_0 und dem Radius ε liegen (Abb. 1.22).

In manchen Anwendungen ist es zweckmäßig, neben der ε -Umgebung auch die Menge zu betrachten, die aus ihr durch Entfernung ihres Mittelpunktes entsteht.

Definition 3. Ist x_0 ein Punkt des Raumes \mathbb{R}^p und ε eine positive reelle Zahl, so heißt die Menge

$$U_\varepsilon'(x_0) = U_\varepsilon(x_0) \setminus \{x_0\}$$

die *punktierte ε -Umgebung* von x_0 .

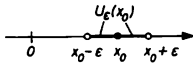


Abb. 1.21

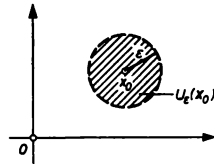


Abb. 1.22

Offenbar ist

$$U_\varepsilon'(x_0) = \{x: x \in \mathbb{R}^p \wedge 0 < \varrho(x_0, x) < \varepsilon\}. \quad (8)$$

Definition 4. Unter dem *Durchmesser* $d(M)$ einer nichtleeren Teilmenge M des Raumes \mathbb{R}^p verstehen wir die kleinste obere Schranke aller Zahlen $\varrho(x, y)$ mit $x, y \in M$.

Hiernach ist

$$d(M) := \sup \{\varrho(x, y) : x, y \in M\}. \quad (9)$$

Definition 5. Eine Teilmenge M des Raumes \mathbb{R}^p heißt *beschränkt*, wenn sie leer oder wenn ihr Durchmesser endlich ist.

Diese Bedingung kann auch anders formuliert werden.

Satz 2. Eine Teilmenge M des euklidischen Raumes \mathbb{R}^p ist beschränkt genau dann, wenn es eine reelle Zahl K gibt derart, daß

$$\|x\| \leq K \quad (10)$$

ist für alle x mit $x \in M$.

Die Zahl K heißt eine *Schranke* der Menge M :

$$K \text{ Schranke von } M \Leftrightarrow \bigwedge_{x \in M} \|x\| \leq K.$$

Beweis. Ist $\|x\| \leq K$ für alle x aus M , so gilt für zwei beliebige Elemente x, y aus M die Abschätzung $\|x - y\| \leq \|x\| + \|y\| \leq 2K$, d. h., es ist $d(M) \leq 2K$.

Ist $d(M)$ endlich und x_0 ein beliebiger Punkt aus M , so folgt aus $x \in M$ stets

$$\|x\| = \|x_0 + (x - x_0)\| \leq \|x_0\| + \|x - x_0\| \leq \|x_0\| + d(M),$$

d. h., die Zahl $\|x_0\| + d(M)$ ist eine Schranke von M .

Eine Menge reeller Zahlen ist beschränkt genau dann, wenn sie nach oben und nach unten beschränkt ist. Ist K eine Schranke, so ist $-K$ eine untere Schranke und K eine obere Schranke von M .

1.5.4. Spezielle Punkte von Teilmengen des Raumes \mathbb{R}^p

In diesem Abschnitt betrachten wir Relationen, die zwischen Elementen und Teilmengen des Raumes \mathbb{R}^p bestehen. Alle nachfolgenden Definitionen stützen sich auf den Begriff der Umgebung bzw. der punktierten Umgebung eines Punktes.

Definition 1. Ein Punkt x des Raumes \mathbb{R}^p heißt ein *innerer Punkt* einer Menge M mit $M \subseteq \mathbb{R}^p$, wenn es eine ε -Umgebung des Punktes x gibt, die ganz in der Menge M enthalten ist (Abb. 1.23):

$$x \text{ innerer Punkt von } M \Leftrightarrow \bigvee_{\varepsilon > 0} U_\varepsilon(x) \subseteq M^1)$$

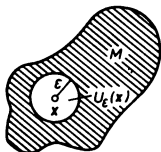


Abb. 1.23

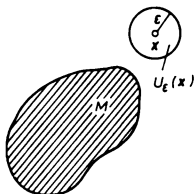


Abb. 1.24

Definition 2. Ein Punkt x des Raumes \mathbb{R}^p heißt ein *äußerer Punkt* einer Menge M mit $M \subseteq \mathbb{R}^p$, wenn es eine ε -Umgebung des Punktes x gibt, die keinen Punkt der Menge M enthält (Abb. 1.24):

$$x \text{ äußerer Punkt von } M \Leftrightarrow \bigvee_{\varepsilon > 0} U_\varepsilon(x) \cap M = \emptyset.$$

Offenbar ist x ein äußerer Punkt der Menge M genau dann, wenn x ein innerer Punkt der Komplementärmenge $\mathbb{R}^p \setminus M$ ist.

¹⁾ Für den gesamten Lehrgang Analysis vereinbaren wir, daß ε und δ (bzw. ε' , δ ; usw.) stets Variable für positive Zahlen sind. Wir drücken dies an Stelle der etwas schwerfälligen Schreibweise $\varepsilon \in \mathbb{R}_+^*$ kürzer durch $\varepsilon > 0$ aus.

Mit $\text{int } M$ (gelesen: interior M) bzw. $\text{ext } M$ (gelesen: exterior M) bezeichnen wir die Menge aller inneren bzw. äußeren Punkte der Menge M . Im allgemeinen gilt nicht $\mathbb{R}^p = \text{int } M \cup \text{ext } M$. Punkte, die weder innere noch äußere Punkte einer Menge M sind, nennen wir *Begrenzungspunkte* der Menge M . Ist x ein Begrenzungspunkt, so kann eine beliebige ε -Umgebung von x weder in der Menge M enthalten noch mit M durchschnittsfremd sein, da x sonst innerer bzw. äußerer Punkt der Menge wäre. Der Begriff „Begrenzungspunkt“ kann daher auch wie folgt definiert werden.

Definition 3. Ein Punkt x des Raumes \mathbb{R}^p heißt ein *Begrenzungspunkt* einer Menge M mit $M \subseteq \mathbb{R}^p$, wenn in jeder ε -Umgebung von x ein Punkt der Menge M und ein Punkt der Menge $\mathbb{R}^p \setminus M$ liegt:

$$x \text{ Begrenzungspunkt von } M : \Leftrightarrow \bigwedge_{\varepsilon > 0} U_\varepsilon(x) \cap M \neq \emptyset \wedge U_\varepsilon(x) \cap \mathbb{R}^p \setminus M \neq \emptyset.$$

Wir bemerken, daß in dieser Definition nichts darüber gesagt wird, ob der Begrenzungspunkt zur Menge gehört oder nicht. Beide Fälle sind möglich. Nach der Definition kann \mathbb{R}^p selbst keine Begrenzungspunkte besitzen. Ein Punkt x ist Begrenzungspunkt der Menge M genau dann, wenn er Begrenzungspunkt ihrer Komplementärmenge ist. Die Menge aller Begrenzungspunkte einer Menge M bezeichnen wir mit βM .

Punkte, die hier Begrenzungspunkte genannt werden, hießen früher *Randpunkte* der Menge. Hier werden nur diejenigen Begrenzungspunkte Randpunkte genannt, die zur Menge gehören:

$$x \text{ Randpunkt von } M : \Leftrightarrow x \in M \cap \beta M.$$

Ein Punkt x ist hiernach Begrenzungspunkt von M genau dann, wenn er Randpunkt von M oder Randpunkt der Komplementärmenge von M ist.

Definition 4. Ein Punkt x des Raumes \mathbb{R}^p heißt *Häufungspunkt* einer Menge M mit $M \subseteq \mathbb{R}^p$, wenn in jeder ε -Umgebung von x unendlich viele Punkte der Menge M liegen.

Ein äußerer Punkt kann somit nicht Häufungspunkt sein.

Ferner gilt der

Satz. Der Punkt x ist Häufungspunkt von M genau dann, wenn in jeder punktierten ε -Umgebung von x (wenigstens) ein Punkt der Menge M liegt:

$$x \text{ Häufungspunkt von } M \Leftrightarrow \bigwedge_{\varepsilon > 0} U'_\varepsilon(x) \cap M \neq \emptyset.$$

Beweis. Ist x Häufungspunkt von M , so sind in $U_\varepsilon(x)$ für alle $\varepsilon > 0$ unendlich viele, also auch in $U'_\varepsilon(x)$ sogar unendlich viele Elemente von M enthalten. Ist x nicht Häufungspunkt, so gibt es ein $\varepsilon > 0$ derart, daß $U_\varepsilon(x)$ keine oder nur endlich

viele von x verschiedene Punkte x_1, x_2, \dots, x_k der Menge M liegen. Ist dann

$$\varepsilon_0 := \frac{1}{2} \min \{\varrho(x, x_i) : i = 1, 2, \dots, k\},$$

so enthält $U'_{\varepsilon_0}(x)$ keinen Punkt der Menge M , d. h., auch die Bedingung des Satzes ist nicht erfüllt.

Abschließend geben wir die

Definition 5. Ein Punkt x des Raumes \mathbb{R}^p heißt ein *isolierter Punkt* einer Menge M mit $M \subseteq \mathbb{R}^p$, wenn $x \in M$ ist und es eine punktierte ε -Umgebung von x gibt, die kein Element von M enthält.

Aus dieser Definition ergibt sich, daß jeder isolierte Punkt des metrischen Raumes \mathbb{R}^p auch Begrenzungspunkt und Randpunkt dieser Menge ist.

Beispiel 1. Im metrischen Raum \mathbb{R}^2 sei

$$M := \{(x, y) : x^2 + y^2 < 1 \wedge x \geq 0\}.$$

Dann ist

$$\text{int } M = \{(x, y) : x^2 + y^2 < 1 \wedge x > 0\}$$

die Menge aller inneren Punkte von M und

$$\beta M = \{(x, y) : (x^2 + y^2 = 1 \wedge x \geq 0) \vee (x = 0 \wedge -1 \leq y \leq 1)\}$$

die Menge aller Begrenzungspunkte von M . Ferner ist $\text{int } M \cup \beta M$ die Menge aller Häufungspunkte und $\mathbb{R}^2 \setminus (\text{int } M \cup \beta M)$ die Menge aller äußeren Punkte von M , während isolierte Punkte von M nicht existieren (Abb. 1.25).

Beispiel 2. Im metrischen Raum \mathbb{R}^2 sei

$$M = \{(x, y) : y = \text{sgn } x\}.$$

In diesem Fall gilt

$$\text{int } M = \emptyset,$$

$$\beta M = M \cup \{(0, 1), (0, -1)\},$$

$$\text{ext } M = \mathbb{R}^2 \setminus \beta M.$$

Die Menge der Häufungspunkte ist $\beta M \setminus \{(0, 0)\}$, und $(0, 0)$ ist der einzige isolierte Punkt der Menge M (Abb. 1.26).

Beispiel 3. Im metrischen Raum \mathbb{R} sei M die Menge aller rationalen Zahlen. Dann ist $\text{int } M = \text{ext } M = \emptyset$ und $\beta M = \mathbb{R}$, denn in jeder Umgebung einer reellen Zahl x gibt es rationale und irrationale Zahlen. Ferner sind alle reellen Zahlen Häufungspunkte von M , und isolierte Punkte sind nicht vorhanden.

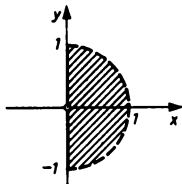


Abb. 1.25

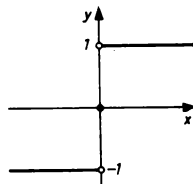


Abb. 1.26

1.5.5. Offene und abgeschlossene Mengen

Wir führen die für viele Betrachtungen in der Analysis bedeutungsvollen Begriffe der offenen bzw. abgeschlossenen Menge ein.

Definition 1. Eine Teilmenge G des Raumes \mathbb{R}^p heißt *offen*, wenn jeder Punkt von G ein innerer Punkt von G ist.

Nach dieser Definition sind offenbar die leere Menge und \mathbb{R}^p selbst offene Mengen, und eine Menge ist offen genau dann, wenn sie keinen ihrer Begrenzungspunkte enthält.

Jede ε -Umgebung eines Punktes x_0 ist eine offene Punktmenge. Ist nämlich $x_1 \in U_\varepsilon(x_0)$, so ist $\|x_1 - x_0\| < \varepsilon$, und setzen wir $\delta := \varepsilon - \|x_1 - x_0\|$, so ist $\delta > 0$. Aus $\|x - x_1\| < \delta$ folgt dann

$$\|x - x_0\| \leq \|x - x_1\| + \|x_1 - x_0\| < \delta + \|x_1 - x_0\| = \varepsilon,$$

d. h., es ist $U_\delta(x_1) \subseteq U_\varepsilon(x_0)$, und $U_\varepsilon(x_0)$ ist offen.

Definition 2. Eine Teilmenge F des Raumes \mathbb{R}^p heißt *abgeschlossen*, wenn ihre Komplementärmenge $\mathbb{R}^p \setminus F$ eine offene Menge ist.

Somit sind \mathbb{R}^p und die leere Menge offen und abgeschlossen zugleich. Im Raum \mathbb{R}^1 ist jedes offene bzw. abgeschlossene Intervall eine offene bzw. abgeschlossene Punktmenge. Das abgeschlossene Intervall $[a, b]$ ist der Spezialfall ($p = 1$) des im euklidischen Raum \mathbb{R}^p wie folgt definierten *Rechtecksbereichs* R . Ist $a_i < b_i$ mit $a_i, b_i \in \mathbb{R}$ ($i = 1, 2, \dots, p$), so setzen wir

$$R := \{(x_1, x_2, \dots, x_p) : a_i \leq x_i \leq b_i; i = 1, 2, \dots, p\}.$$

Jeder solche Rechtecksbereich ist eine abgeschlossene Punktmenge. Werden in den Ungleichungen die Gleichheitszeichen nicht zugelassen, so bezeichnen wir die entstehende Punktmenge als einen *offenen Rechtecksbereich*. Er ist eine offene Punktmenge.

Eine Menge M , die wenigstens einen ihrer Begrenzungspunkte enthält und wenigstens einen ihrer Begrenzungspunkte nicht enthält, ist weder offen noch abgeschlossen.

Satz 1. Die Vereinigung von beliebig vielen offenen Mengen des Raumes \mathbb{R}^p ist wieder eine offene Menge.

Beweis. Es sei \mathcal{G} ein Mengensystem, dessen Elemente offene Teilmengen von \mathbb{R}^p sind. Ist $x \in \bigcup \mathcal{G}$, so gibt es ein G mit $x \in G$ und folglich ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq G$, woraus $U_\varepsilon(x) \subseteq \bigcup \mathcal{G}$ folgt. Jeder Punkt von $\bigcup \mathcal{G}$ ist somit ein innerer Punkt von $\bigcup \mathcal{G}$, d. h., $\bigcup \mathcal{G}$ ist eine offene Menge.

Satz 2. Der Durchschnitt von beliebig vielen abgeschlossenen Mengen des Raumes \mathbb{R}^p ist wieder eine abgeschlossene Menge.

Beweis. Es sei \mathcal{F} ein Mengensystem, dessen Elemente abgeschlossene Teilmengen von \mathbb{R}^p sind. Ist $x \notin \bigcap \mathcal{F}$, dann gibt es ein F mit $F \in \mathcal{F}$ und $x \notin F$, d. h. mit $x \in \mathbb{R}^p \setminus F$. Da $\mathbb{R}^p \setminus F$ offen ist, gibt es ein ε mit $U_\varepsilon(x) \subseteq \mathbb{R}^p \setminus F$. Wegen $F \in \mathcal{F}$ ist $\mathbb{R}^p \setminus F \subseteq \mathbb{R}^p \setminus \bigcap \mathcal{F}$, also $U_\varepsilon(x) \subseteq \mathbb{R}^p \setminus \bigcap \mathcal{F}$. Daher ist die Komplementärmenge von $\bigcap \mathcal{F}$ offen, und Satz 2 ist bewiesen.

Satz 3. Eine Teilmenge F des Raumes \mathbb{R}^p ist genau dann abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.

Beweis. a) Es sei F abgeschlossen, und x sei ein Häufungspunkt von F . Nehmen wir an, es gelte $x \notin F$, also $x \in \mathbb{R}^p \setminus F$. Nach Definition 2 ist $\mathbb{R}^p \setminus F$ offen, und es gibt ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq \mathbb{R}^p \setminus F$. Dies bedeutet aber, daß x äußerer Punkt von F ist, während wir vorausgesetzt hatten, daß x Häufungspunkt von F ist. Somit gilt $x \in F$, die abgeschlossene Menge F enthält alle ihre Häufungspunkte.

b) Es sei F eine Menge, die alle ihre Häufungspunkte enthält. Ist dann $x \in \mathbb{R}^p \setminus F$, so ist x nicht Häufungspunkt von F . Nach dem Satz in 1.5.4. gibt es zunächst eine punktierte ε -Umgebung von x , die keinen Punkt von F enthält. Da auch $x \notin F$ ist, gilt $U_\varepsilon(x) \cap F = \emptyset$, $U_\varepsilon(x) \subseteq \mathbb{R}^p \setminus F$. Somit ist jeder Punkt x mit $x \in \mathbb{R}^p \setminus F$ ein innerer Punkt von $\mathbb{R}^p \setminus F$, d. h., $\mathbb{R}^p \setminus F$ ist offen. Daher ist F abgeschlossen, und Satz 3 ist bewiesen.

Definition 3. Eine Teilmenge M des Raumes \mathbb{R}^p heißt ein *Gebiet*, wenn sie offen ist und wenn es zu je zwei Punkten x_1, x_2 aus M stets endlich viele Punkte $x_1 = y_1, y_2, \dots, y_n = x_2$ mit $S[y_i, y_{i+1}] \subset M$ ($i = 1, 2, \dots, n-1$) gibt.

Je zwei Punkte x_1, x_2 eines Gebietes können also durch einen ganz in diesem Gebiet verlaufenden Streckenzug verbunden werden (Abb. 1.27).

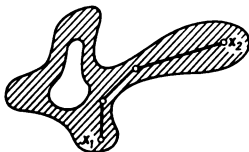


Abb. 1.27

1.5.6. Normierte Räume und metrische Räume

In der Mathematik treten häufig Mengen auf, in denen ebenso wie im Raum \mathbb{R}^p jedem Element eine Norm mit ähnlichen Eigenschaften zugeordnet werden kann.

Definition 1. Es sei E ein reeller Vektorraum. Eine Abbildung, die jedem Element $x \in E$ eine nichtnegative reelle Zahl $\|x\|$ zuordnet, heißt eine *Norm* in E , wenn für alle $x, y \in E$ und alle $\lambda \in \mathbb{R}$ die Bedingungen

$$\|\lambda x\| = |\lambda| \|x\|, \quad (1)$$

$$\|x + y\| \leq \|x\| + \|y\|, \quad (2)$$

$$\|x\| > 0, \text{ falls } x \neq 0, \quad (3)$$

erfüllt sind. Ist in E eine Norm ausgezeichnet, so heißt E ein reeller *normierter Raum*.

Neben der Dreiecksungleichung (2) gelten in jedem normierten Raum auch die abgeleiteten Dreiecksungleichungen 1.5.2., (10), (11). Beginnend mit Definition 2 in 1.5.2., können alle Begriffsbildungen und Sätze aus 1.5. wörtlich auf normierte Räume übertragen werden.

Eine weitere Verallgemeinerung geben wir in der nachfolgenden Definition. Hier wird nicht einmal vorausgesetzt, daß die Grundmenge ein reeller Vektorraum ist.

Definition 2. Unter einem *metrischen Raum* versteht man eine Menge X in Verbindung mit einer Funktion ϱ , die jedem geordneten Paar (x, y) von Elementen x, y aus dieser Menge eine reelle Zahl $\varrho(x, y)$ zuordnet und die folgende Eigenschaften besitzt: Für alle Elemente x, y, z der Menge X gilt

$$\varrho(x, y) \geq 0, \quad (1)$$

$$\varrho(x, y) = 0 \text{ genau dann, wenn } x = y, \quad (2)$$

$$\varrho(x, y) = \varrho(y, x) \quad (\text{Symmetrie}), \quad (3)$$

$$\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z) \quad (\text{Dreiecksungleichung}). \quad (4)$$

Die Elemente eines metrischen Raumes werden häufig *Punkte*, die Funktion ϱ wird die *Metrik* des Raumes genannt.

Die nachfolgenden Begriffsbildungen und Sätze können sogar auf metrische Räume übertragen werden: die Definitionen 2, 3, 4, 5 aus 1.5.3., alle Definitionen und der Satz aus 1.5.4. sowie die Definitionen 1, 2 und die Sätze 1, 2, 3 aus 1.5.5.

1.6. Die trigonometrischen Funktionen

1.6.1. Die Funktionalgleichungen der trigonometrischen Funktionen

In der Schulmathematik werden die trigonometrischen Funktionen auf Grund von rein geometrischen Betrachtungen eingeführt, ohne daß wichtige Begriffsbildungen wie Winkel, Winkelmaß, Orientierung, Stetigkeitsaxiom eine strenge Begründung

erfahren. Im Rahmen dieser Lehrbuchreihe erfolgt eine exakte Einführung auf geometrischer Basis in den Bänden MfL 6, 7. Im Rahmen des Lehrgangs Analysis geben wir eine von der geometrischen Axiomatik unabhängige Konstruktion. Geometrische Betrachtungen werden nur zur Veranschaulichung und Motivation herangezogen. Die Definitionen und Beweisführungen stützen sich aber ausschließlich auf die Eigenschaften des Körpers der reellen Zahlen.

Den nachfolgenden Satz von der *Existenz und Einzigkeit der trigonometrischen Funktionen* werden wir in mehreren Etappen, teilweise auch auf verschiedene Arten beweisen.

Satz 1. *Es gibt genau ein Paar reeller Funktionen, das für alle reellen Zahlen α, β den Funktionalgleichungen*

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta, \quad (1)$$

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta. \quad (2)$$

$$\cos(-\alpha) = \cos \alpha, \quad (3)$$

$$\sin(-\alpha) = -\sin \alpha \quad (4)$$

und weiterhin den Ungleichungen

$$0 < \sin \alpha < \alpha < \frac{\sin \alpha}{\cos \alpha} \quad (0 < \alpha < 1) \quad (5)$$

genügt.

Die Funktionalgleichungen (1), (2) heißen auch die *Additionstheoreme* der Winkelfunktionen, und (3), (4) charakterisieren bekannte Symmetrieeigenschaften. Die Ungleichungen (5) können leicht plausibel gemacht werden. Nach Abb. 1.28 ist der Sinus eines spitzen Winkels stets kleiner als die zu diesem Winkel gehörende Bogenlänge α . Es ist nicht ganz so offensichtlich, aber doch einleuchtend, daß der Tangens dieses Winkels größer als die Bogenlänge ist. Genau diese Aussagen sind in (5) fixiert.

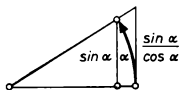


Abb. 1.28

Einen Beweis der Einzigkeitsaussage unseres Satzes werden wir in 1.6.2. führen. Ein Beweis der Existenzaussage mit elementaren Hilfsmitteln ist sehr aufwendig. Wir führen einen solchen Beweis in 1.6.4., wobei nur der Begriff der kleinsten oberen Schranke und unsere Sätze über monotone Funktionen als Hilfsmittel eingesetzt werden. Ein besonders einfacher Beweis der Existenzaussage wird in 2.2.8. mit den Hilfsmitteln der Reihenlehre geführt.

Im Rest dieses Abschnittes gehen wir von der Annahme aus, daß uns die Funk-

tionen $y = \cos x$, $y = \sin x$ noch nicht bekannt sind.¹⁾ Wir stellen uns also auf den Standpunkt, daß ein beliebiges Funktionenpaar gegeben sei, daß die Eigenschaften (1) bis (5) besitzt.

Zunächst wollen wir einige Aspekte der Funktionalgleichungen (1) bis (4) allein untersuchen. Sie sind mit den *Subtraktionstheoremen*

$$\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta, \quad (6)$$

$$\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta \quad (7)$$

äquivalent. Ersetzen wir nämlich β in den Additionstheoremen (1), (2) durch $-\beta$, so ergeben sich mit Hilfe der Symmetrieeigenschaften (3), (4) die Formeln (6), (7). Umgekehrt folgt aus (6), (7) zunächst $\cos(\alpha - \beta) = \cos(\beta - \alpha)$ und $\sin(\alpha - \beta) = -\sin(\beta - \alpha)$, und damit sind auch (3), (4) erfüllt. Ersetzen wir nun wieder β in (6), (7) durch $-\beta$, so erhalten wir (1), (2).

Es gibt Funktionen, z. B. die Funktionen

$$C(\alpha) := e^{\alpha} \cos \alpha, \quad S(\alpha) := e^{\alpha} \sin \alpha,$$

die wohl den Additionstheoremen, aber nicht den Subtraktionstheoremen genügen. Nach dem zuvor Bewiesenen sind also die letzteren aussagekräftiger.

Die Funktionalgleichungen (1) bis (4) sind offensichtlich erfüllt, wenn wir $\cos \alpha = \sin \alpha = 0$ für alle $\alpha \in \mathbb{R}$ setzen. Diesen trivialen Fall wollen wir fortan ausschließen. Nach (4) ist aber stets

$$\sin 0 = 0, \quad (8)$$

und es folgt $\cos \alpha = \cos(\alpha + 0) = \cos \alpha \cos 0$, $\sin \alpha = \sin(\alpha + 0) + \sin \alpha \cos 0$ für alle $\alpha \in \mathbb{R}$. Dies ist nur möglich, wenn

$$\cos 0 = 1 \quad (9)$$

ist. Wir zeigen nun, daß die Eigenschaften (6), (7), (9) mit (1), (2) und

$$\cos^2 \alpha + \sin^2 \alpha = 1 \quad (10)$$

äquivalent sind. Gilt nämlich (6), (7), (9), so sind (1), (2) erfüllt, und aus (6), (9) folgt mit $\beta := \alpha$ die als „trigonometrische Form des pythagoreischen Lehrsatzes“ bekannte Formel (10). Es seien umgekehrt (1), (2), (10) erfüllt. In (1), (2) ersetzen wir α, β durch $\beta, \alpha - \beta$ und erhalten

$$\cos \alpha = \cos \beta \cos(\alpha - \beta) - \sin \beta \sin(\alpha - \beta), \quad (11)$$

$$\sin \alpha = \sin \beta \cos(\alpha - \beta) + \cos \beta \sin(\alpha - \beta). \quad (12)$$

¹⁾ Zur Verdeutlichung dieses Gedankens wäre es einerseits vorteilhafter, in den Herleitungen überall $c(\alpha)$, $s(\alpha)$ statt $\cos \alpha$, $\sin \alpha$ zu schreiben. Dies hätte andererseits den großen Nachteil, daß der im folgenden entwickelte umfangreiche Formelapparat nicht sofort in der Form erscheint, wie es für die Anwendungen wünschenswert ist.

Multiplizieren wir in diesen Gleichungen mit $\cos \beta$, $\sin \beta$ bzw. mit $-\sin \beta$, $\cos \beta$ und addieren, so erhalten wir unter Beachtung von (10) die Formeln (6), (7). Auf Grund von (10) ist stets $|\cos \alpha|, |\sin \alpha| \leq 1$.

Weitere Additionstheoreme gewinnen wir, wenn wir α, β in (1), (6) durch $\frac{\alpha + \beta}{2}$, $\frac{\alpha - \beta}{2}$ ersetzen und addieren bzw. subtrahieren. Wir erhalten

$$\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}, \quad (13)$$

$$\cos \alpha - \cos \beta = 2 \sin \frac{\alpha + \beta}{2} \sin \frac{\beta - \alpha}{2}. \quad (14)$$

Aus (2), (7) folgt in gleicher Weise

$$\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}, \quad (15)$$

$$\sin \alpha - \sin \beta = 2 \cos \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}. \quad (16)$$

Mit $\frac{\alpha}{2}$ statt α, β gehen (1), (2) in die wichtigen *Halbwinkelformeln*

$$\cos \alpha = \cos^2 \frac{\alpha}{2} - \sin^2 \frac{\alpha}{2}, \quad (17)$$

$$\sin \alpha = 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} \quad (18)$$

über. Die erste kann mit (10) auch in die Form

$$\cos \alpha = 2 \cos^2 \frac{\alpha}{2} - 1 = 1 - \sin^2 \frac{\alpha}{2} \quad (19)$$

gebracht werden. Hieraus ergibt sich

$$\left| \cos \frac{\alpha}{2} \right| = \sqrt{\frac{1 + \cos \alpha}{2}}, \quad (20)$$

$$\left| \sin \frac{\alpha}{2} \right| = \sqrt{\frac{1 - \cos \alpha}{2}}. \quad (21)$$

Aus (18) kann durch vollständige Induktion die Produktdarstellung

$$\sin \alpha = 2^n \sin \frac{\alpha}{2^n} \cos \frac{\alpha}{2} \cos \frac{\alpha}{4} \cdots \cos \frac{\alpha}{2^n} \quad (22)$$

abgeleitet werden.

Jetzt soll auch die Eigenschaft (5) herangezogen werden. Unter dieser Voraussetzung gilt stets

$$|\sin \alpha| \leq |\alpha|. \quad (23)$$

In der Tat ist (23) für $0 < \alpha < 1$ nach (5) erfüllt, und für $-1 < \alpha \leq 0$ folgt dies aus (8), (4). Ist aber (23) für ein α richtig, so ist $|\sin 2\alpha| = 2 |\sin \alpha| |\cos \alpha| \leq 2 |\sin \alpha| \leq |2\alpha|$, und (23) gilt auch für 2α . Hieraus folgt unsere Behauptung.

Wenden wir (23) speziell auf (14), (16) an, so erhalten wir, wenn wir $\left| \cos \frac{\alpha + \beta}{2} \right|$, $\left| \sin \frac{\alpha + \beta}{2} \right| \leq 1$ beachten,

$$|\cos \alpha - \cos \beta| \leq |\alpha - \beta|, \quad (24)$$

$$|\sin \alpha - \sin \beta| \leq |\alpha - \beta|. \quad (25)$$

Der nachfolgende Satz charakterisiert die Kreiszahl π durch die bekannte Eigenschaft, daß $\frac{\pi}{2}$ die kleinste positive Nullstelle der Kosinusfunktion ist.

Satz 2. *Es gibt eine positive Zahl π mit*

$$\sin \alpha > 0 \quad (0 < \alpha < \pi), \quad (26)$$

$$\cos \frac{\pi}{2} = 0, \quad \sin \frac{\pi}{2} = 1. \quad (27)$$

Beweis. Eine Umstellung von (5) ergibt $0 < \alpha \cos \alpha < \sin \alpha$, woraus $\alpha^2 \cos^2 \alpha < \sin^2 \alpha = 1 - \cos^2 \alpha$ oder

$$\cos^2 \alpha < \frac{1}{1 + \alpha^2} \quad (0 < \alpha < 1)$$

folgt. Speziell ist

$$\cos^2 \frac{1}{2} < \frac{4}{5}, \quad \cos 1 = 2 \cos^2 \frac{1}{2} - 1 < \frac{7}{25},$$

also $\cos 2 = 2 \cos^2 1 - 1 < 0$. Zum Beweis der Existenz einer Nullstelle der Kosinusfunktion im Intervall $[0, 2]$ genügt nun der Nachweis, daß die Kosinusfunktion in diesem Intervall streng monoton ist. Wegen (24) ist dann nämlich der Zwischensatz erfüllt (vgl. 1.3.5., Satz 6, Satz 5).

Für $0 < \alpha < 1$ ist nach (5) stets $\cos \alpha, \sin \alpha > 0$, und folglich ist

$$\sin \alpha = 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} > 0 \quad (0 < \alpha < 2). \quad (28)$$

Ist also $0 \leq \alpha < \beta \leq 2$, so haben wir

$$\cos \alpha - \cos \beta = 2 \sin \frac{\alpha + \beta}{2} \sin \frac{\beta - \alpha}{2} > 0,$$

und die strenge Monotonie ist bewiesen. Folglich gibt es genau eine Zahl π mit $0 < \pi < 4$ und $\cos \frac{\pi}{2} = 0$. Für $0 \leq \alpha < \frac{\pi}{2}$ ist $\cos \alpha > 0$, nach (28) auch $\sin \alpha > 0$. Für $0 < \alpha < \pi$ ist demzufolge $\sin \alpha = 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} > 0$. Aus (10) folgt nun $\sin \frac{\pi}{2} = 1$, und Satz 2 ist bewiesen.

Aus den Formeln (27) können nun alle Periodizitätseigenschaften der trigonometrischen Funktionen abgeleitet werden. Mit Hilfe der Additionstheoreme ergibt sich

$$\cos \left(\alpha + \frac{\pi}{2} \right) = -\sin \alpha, \quad (29)$$

$$\sin \left(\alpha + \frac{\pi}{2} \right) = \cos \alpha. \quad (30)$$

Nochmalige Anwendung ergibt

$$\cos (\alpha + \pi) = -\cos \alpha, \quad (31)$$

$$\sin (\alpha + \pi) = -\sin \alpha. \quad (32)$$

(Abb. 1.29). Durch vollständige Induktion gewinnen wir die Formeln

$$\cos (\alpha + k\pi) = (-1)^k \cos \alpha, \quad (33)$$

$$\sin (\alpha + k\pi) = (-1)^k \sin \alpha, \quad (34)$$

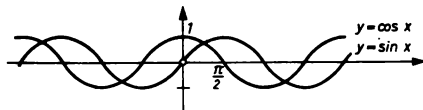


Abb. 1.29

nach denen die trigonometrischen Funktionen die Periode 2π besitzen. Für $0 < \alpha < \pi$ bzw. $\pi < \alpha < 2\pi$ ist stets $\sin \alpha > 0$ bzw. $\sin \alpha < 0$, und folglich ist keine Zahl mit $0 < \alpha < 2\pi$ eine Periode. Die Zahl 2π ist also die *primitive* Periode der Winkelfunktionen. Die durch

$$\tan \alpha := \frac{\sin \alpha}{\cos \alpha} \quad \left(\alpha \neq k\pi + \frac{\pi}{2} \right), \quad (40)$$

$$\cot \alpha := \frac{\cos \alpha}{\sin \alpha} \quad (\alpha \neq k\pi) \quad (41)$$

definierte *Tangens-* bzw. *Kotangensfunktion* besitzt dagegen die primitive Periode π (Abb. 1.30). Die beiden Funktionen sind weiterhin durch

$$\frac{1}{\tan \alpha} = \cot \alpha = \tan \left(\frac{\pi}{2} - \alpha \right)$$

und ähnliche weitere Relationen miteinander verknüpft.

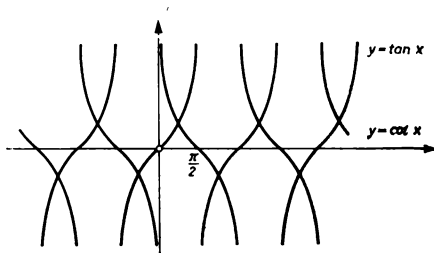


Abb. 1.30

1.6.2. Die Arkusfunktionen

In diesem Abschnitt gehen wir wieder von der Voraussetzung aus, daß ein Funktionenpaar mit den in 1.6.1. zusammengestellten Eigenschaften existiert.

Für $0 < \alpha < \pi$ ist $\sin \alpha > 0$, also

$$\cos \alpha - \cos \beta = 2 \sin \frac{\alpha + \beta}{2} \sin \frac{\beta - \alpha}{2} > 0 \quad (0 \leq \alpha < \beta \leq \pi).$$

Die Kosinusfunktion ist daher im Intervall $[0, \pi]$ streng monoton fallend. Da außerdem $|\cos \alpha - \cos \beta| \leq |\alpha - \beta|$ ist, nimmt sie in diesem Intervall jeden Wert zwischen $\cos 0 = 1$ und $\cos \pi = -1$ genau einmal an. Auf Grund der Relation $\sin \alpha = \cos \left(\frac{\pi}{2} - \alpha \right)$ ist die Sinusfunktion im Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ streng monoton wachsend und besitzt dort ebenfalls den Wertebereich $[-1, 1]$. Somit besitzt die Einschränkung der Kosinusfunktion bzw. Sinusfunktion auf das Intervall $[0, \pi]$ bzw. $\left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ eine auf dem Intervall $[0, 1]$ definierte Umkehrfunktion, die *Arkuskosinusfunktion* bzw. die *Arkussinusfunktion*. Sie werden mit

$$y = \arccos x \quad (-1 \leq x \leq 1), \quad (1)$$

$$y = \arcsin x \quad (-1 \leq x \leq 1) \quad (2)$$

bezeichnet (Abb. 1.31). Die Bezeichnungen rühren daher, daß $\alpha = \arccos x$ bzw. $\alpha = \arcsin y$ derjenige *Arkus* (Bogen oder Bogenlänge) ist, dessen Kosinus bzw. Sinus den Wert x bzw. y annimmt. Dies werden wir am Schluß dieses Abschnitts und in 1.6.4. genauer erläutern (Abb. 1.32).

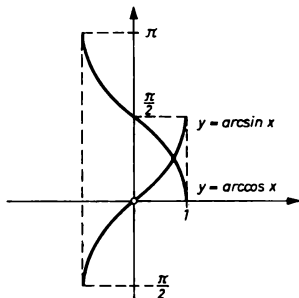


Abb. 1.31

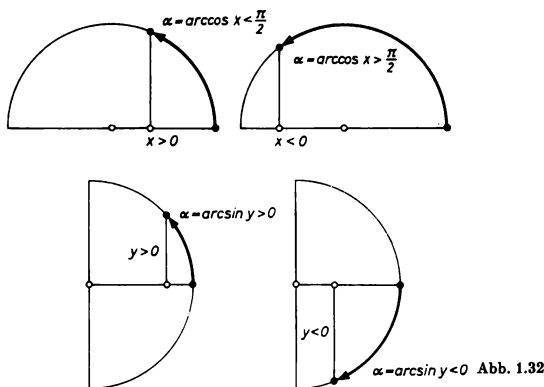


Abb. 1.32

Die Einschränkung der Tangensfunktion auf das Intervall $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ ist streng monoton wachsend, denn für $-\frac{\pi}{2} < \alpha < \beta < \frac{\pi}{2}$ ist

$$\tan \beta - \tan \alpha = \frac{\sin \beta \cos \alpha - \cos \beta \sin \alpha}{\cos \beta \cos \alpha} = \frac{\sin (\beta - \alpha)}{\cos \beta \cos \alpha} > 0.$$

Sie nimmt in diesem Intervall auch jede reelle Zahl x an, denn setzen wir

$$\alpha := \arcsin \frac{x}{\sqrt{1+x^2}},$$

so ist $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$ und

$$\sin \alpha = \frac{x}{\sqrt{1+x^2}}, \quad \cos \alpha > 0.$$

Lösen wir nach x auf, so erhalten wir

$$x = \frac{\sin \alpha}{\sqrt{1 - \sin^2 \alpha}} = \frac{\sin \alpha}{\cos \alpha} = \tan \alpha.$$

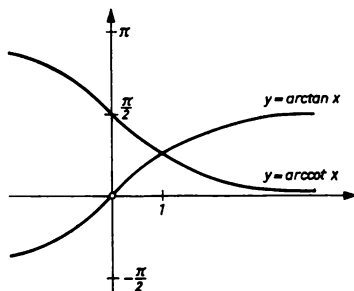


Abb. 1.33

Somit existiert eine für alle reellen Zahlen x definierte Umkehrfunktion, die *Arkustangensfunktion*. Sie wird mit

$$y = \arctan x \quad (x \in \mathbb{R}) \quad (3)$$

bezeichnet. Es ist für viele Anwendungen bedeutungsvoll, daß diese Funktion das unbeschränkte Intervall $] -\infty, \infty[$ umkehrbar eindeutig auf das beschränkte Intervall $] -\frac{\pi}{2}, \frac{\pi}{2}[$ abbildet. Mit

$$y = \operatorname{arccot} x \quad (x \in \mathbb{R}) \quad (4)$$

bezeichnen wir die Umkehrfunktion der Einschränkung der Kotangensfunktion auf das Intervall $]0, \pi[$ (Abb. 1.33). Die Existenz leiten wir aus der Relation $\cot \alpha = \tan\left(\frac{\pi}{2} - \alpha\right)$ ab.

Die Arkusfunktionen sind miteinander durch die Relationen

$$\arccos x + \arcsin x = \frac{\pi}{2}, \quad (5)$$

$$\arctan x + \operatorname{arccot} x = \frac{\pi}{2} \quad (6)$$

verknüpft, denn aus $\alpha = \arcsin x$ folgt $x = \sin \alpha = \cos\left(\frac{\pi}{2} - \alpha\right)$ und $-\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2}$, $0 \leq \frac{\pi}{2} - \alpha \leq \pi$. Daher ist $\frac{\pi}{2} - \alpha = \arccos x$, und das ist die Behauptung (5).

Ebenso wird (6) bewiesen.

Wir geben jetzt ein Verfahren zur Berechnung der Funktionswerte der Funktion $\alpha = \arccos x$, das im Prinzip bereits auf ARCHIMEDES zurückgeht. Wir erläutern es zuerst geometrisch. Wir teilen den zu α gehörenden Kreisbogen in 2^{n-1} gleiche Teile. Die Länge des zugehörigen Sehnenzuges ist offensichtlich

$$s_n := 2^n \sin \frac{\alpha}{2^n} = 2^n \sqrt{1 - \cos^2 \frac{\alpha}{2^n}}, \quad (7)$$

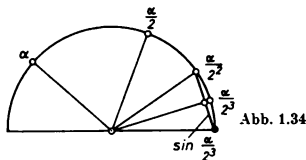


Abb. 1.34

und diese Zahl kommt nach unseren anschaulichen Vorstellungen der Bogenlänge α um so näher, je größer wir die Zahl n wählen (Abb. 1.34). Nun kann aber der Kosinus des Winkels $\frac{\alpha}{2^{n+1}}$ nach 1.6.1.(20) durch den Kosinus von $\frac{\alpha}{2^n}$ dargestellt werden.

Die Zahlen s_n können also, von x ausgehend, induktiv berechnet werden. Nunmehr lösen wir uns von der geometrischen Interpretation und beweisen den

Satz 1. Es sei $-1 \leq x \leq 1$ und $\alpha := \arccos x$. Setzen wir

$$x_0 := x, \quad x_{n+1} := \sqrt{\frac{1+x_n}{2}} \quad (n \in \mathbf{N}), \quad (8)$$

so ist

$$0 \leq \alpha - 2^n \sqrt{1-x_n^2} \leq \frac{\alpha^2}{2^{n+1}} < \frac{1}{2^{n-3}} \quad (n \geq 2). \quad (9)$$

Beweis. Im Fall $x = 1$, $\alpha = 0$ ist $x_n = 1$ für alle $n \in \mathbf{N}$, und (9) ist erfüllt. Es sei nun $-1 \leq x < 1$ und $\alpha_n := \frac{\alpha}{2^n}$. Wir zeigen, daß

$$x_n = \cos \alpha_n \quad (10)$$

ist. Für $n = 0$ ist dies richtig. Gilt (10) für ein $n \in \mathbf{N}$, so ist $\alpha_{n+1} \leq \frac{\pi}{2}$, also

$$x_{n+1} = \sqrt{\frac{1 + \cos \alpha_n}{2}} = \left| \cos \frac{\alpha_n}{2} \right| = \cos \alpha_{n+1},$$

und (10) ist durch Induktion bewiesen. Es folgt $\sqrt{1 - x_n^2} = \sin \alpha_n$. Für $n \geq 2$ ist $x_{n-1} \geq 0$ und

$$\sin \alpha_n = \sqrt{\frac{1 - \cos \alpha_{n-1}}{2}} \leq \sqrt{\frac{1 + \cos \alpha_{n-1}}{2}} = \cos \alpha_n,$$

d. h., es ist $\tan \alpha_n \leq 1$. Andererseits ist für $n \geq 2$ stets $\alpha_n \leq \frac{\pi}{4} < 1$ und

$$\sin \alpha_n < \alpha_n < \tan \alpha_n,$$

$$0 < \alpha_n - \sin \alpha_n < \tan \alpha_n - \sin \alpha_n$$

$$= \tan \alpha_n (1 - \cos \alpha_n) \leq 1 - \cos \alpha_n = 2 \sin^2 \alpha_{n+1} < 2 \alpha_{n+1}^2,$$

$$0 < \frac{\alpha}{2^n} - \sqrt{1 - x_n^2} < \frac{\alpha^2}{2^{2n+1}}.$$

Multiplizieren wir mit 2^n , so erhalten wir die ersten beiden Ungleichungen (9) in einer verschärften Form. Wegen $0 < \alpha \leq 4$ gilt auch die letzte Ungleichung (9), und Satz 1 ist bewiesen.

Im Fall $x = 0$ ist $\alpha = \frac{\pi}{2} < 2$, und wir haben die Abschätzung

$$0 \leq \pi - 2^{n+1} \sqrt{1 - x_n^2} < \frac{1}{2^{n-2}} \quad (n \geq 2). \quad (11)$$

Sie ermöglicht es, den Zahlenwert von π in jeder gewünschten Genauigkeit zu bestimmen. Für die numerische Rechnung empfiehlt sich die Herleitung einer Rekursionsformel für die Zahlen

$$\pi_n := 2^{n+1} \sqrt{1 - x_n^2}.$$

Es ist $2(1 - x_{n+1}^2) = 1 - x_n^2$, also

$$\frac{\pi_n}{\pi_{n+1}} = \sqrt{\frac{1 - x_n^2}{4(1 - x_{n+1}^2)}} = \sqrt{\frac{(1 - x_n)(1 + x_n)}{2(1 - x_n)}} = \sqrt{\frac{1 + x_n}{2}} = x_{n+1},$$

d. h., es ist

$$\pi_{n+1} = \frac{\pi_n}{x_{n+1}}.$$

Mit einem Taschenrechner wurden nach dieser Formel folgende Näherungswerte für π berechnet.

n	$x_0 := 0$ $x_{n+1} := \sqrt{\frac{1+x_n}{2}}$	$\pi_0 := 2\sqrt{1-x_0^2} = 2$ $\pi_{n+1} := \frac{\pi_n}{x_{n+1}}$
1	0,7071067811	2,828427124
2	0,9238795325	3,061467458
3	0,9807852804	3,121445152
4	0,9951847266	3,136548490
5	0,9987954562	3,140331156
6	0,9996988186	3,141277250
7	0,9999247018	3,141513801
8	0,9999811752	3,141572940
9	0,9999952938	3,141587725
10	0,9999988234	3,141591421
11	0,9999997058	3,141592345
12	0,9999999264	3,141592576
13	0,9999999816	3,141592634
14	0,9999999954	3,141592648
15	0,9999999988	3,141592652
16	0,9999999997	3,141592653
17	0,9999999999	3,141592653

Die Fehlerabschätzung (11) gilt nur bedingt, denn es treten unvermeidliche Rundungsfehler auf. Der genaue Zahlenwert für π beginnt mit

3,141592653589793 ...

Aus Satz 1 kann ein weiterer wichtiger Schluß gezogen werden.

Satz 2. *Es gibt höchstens ein Funktionenpaar, das den Bedingungen 1.6.1.(1) bis (5) genügt.*

Beweis. Wenn ein solches Funktionenpaar existiert, sind die Werte der Funktion $y = \arccos x$ für $-1 \leq x \leq 1$ nach Satz 1 eindeutig bestimmt (vgl. auch 2.1.4.(10)). Damit ist aber auch deren Umkehrfunktion, also die Funktion $y = \cos x$ ($0 \leq x \leq \pi$) eindeutig bestimmt. Die Behauptung des Satzes folgt nun aus 1.6.1.(33) und (29).

1.6.3. Trigonometrische Darstellung komplexer Zahlen. Polar- und Zylinderkoordinaten

Es sei $z = x + iy$ ($x, y \in \mathbb{R}$). Wir setzen $r := |z|$, d. h.

$$r := \sqrt{x^2 + y^2}, \quad (1)$$

und zeigen, daß es einen Winkel φ mit

$$x = r \cos \varphi, \quad y = r \sin \varphi \quad (2)$$

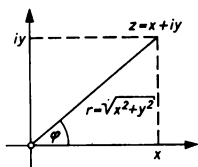


Abb. 1.35

gibt (Abb. 1.35). Im Fall $z = 0$, also $r = x = y = 0$, ist (2) für jede reelle Zahl φ erfüllt. Ist $z \neq 0$, so setzen wir

$$\varphi := \begin{cases} \arccos \frac{x}{r} & \text{für } y \geq 0, \\ -\arccos \frac{x}{r} & \text{für } y < 0. \end{cases} \quad (3)$$

Die erste Gleichung (2) ist dann erfüllt. Im Fall $y \geq 0$ ist $0 \leq \varphi \leq \pi$,

$$0 \leq \sin \varphi = \sqrt{1 - \cos^2 \varphi} = \sqrt{1 - \frac{x^2}{r^2}} = \sqrt{\frac{y^2}{r^2}} = \frac{y}{r}.$$

Im Fall $y < 0$ ist $\frac{x}{r} \neq \pm 1$, $-\pi < \varphi < 0$,

$$0 > \sin \varphi = -\sqrt{1 - \cos^2 \varphi} = -\sqrt{1 - \frac{x^2}{r^2}} = -\frac{|y|}{r} = \frac{y}{r}.$$

In beiden Fällen gilt somit (2). Aus (2) und $z = x + iy$ folgt

$$z = r(\cos \varphi + i \sin \varphi), \quad (4)$$

womit wir die *trigonometrische Darstellung* komplexer Zahlen hergeleitet haben. So ist z. B.

$$\begin{aligned} -1 + i &= \sqrt{2} \left(\cos \frac{3\pi}{4} + i \sin \frac{3\pi}{4} \right), \\ -3 - i &= 2 \left(\cos \frac{7\pi}{6} + i \sin \frac{7\pi}{6} \right), \\ -1 &= (\cos \pi + i \sin \pi). \end{aligned} \quad (5)$$

Mit Hilfe der Reihenlehre werden wir später zeigen, daß die Definition

$$e^{i\varphi} := \cos \varphi + i \sin \varphi \quad (6)$$

sinnvoll ist. Alle Zahlen $e^{i\varphi}$ liegen wegen $\cos^2 \varphi + \sin^2 \varphi = 1$ auf dem Einheitskreis. Aus (5) und $r \geq 0$ folgt also durch Übergang zu den Beträgen stets $|z| = r$. Es gibt somit genau eine nichtnegative reelle Zahl r , für die eine Gleichung der Form (4) erfüllt ist, nämlich den Betrag der komplexen Zahl z . Es ist

$$\begin{aligned} &(\cos \varphi + i \sin \varphi)(\cos \psi + i \sin \psi) \\ &= \cos \varphi \cos \psi - \sin \varphi \sin \psi + i(\cos \varphi \sin \psi + \sin \varphi \cos \psi) \\ &= \cos(\varphi + \psi) + i \sin(\varphi + \psi), \end{aligned}$$

d. h., es gilt

$$e^{i(\varphi + \psi)} = e^{i\varphi} e^{i\psi}.$$

Jede reelle Zahl φ , für die (2) bzw. (4) erfüllt ist, heißt ein *Argument* der komplexen Zahl z . Wegen der Periodizität der trigonometrischen Funktionen ist mit φ auch jede reelle Zahl $\varphi + 2k$ mit $k \in \mathbb{Z}$ ein Argument von z . Sind umgekehrt φ und ψ Argumente einer von Null verschiedenen komplexen Zahl z , so ist $z = re^{i\varphi} = re^{i\psi}$, d. h., es ist $e^{i(\varphi - \psi)} = e^{i\varphi} e^{-i\psi} = e^{i\varphi} e^{-i\psi} = e^0 = 1$. Daher ist $\cos(\varphi - \psi) = 1$ und damit $\varphi - \psi = 2k\pi$ ($k \in \mathbb{Z}$). Argumente von komplexen Zahlen z mit $z \neq 0$ sind somit bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt.

Die Darstellung wird im Fall $z \neq 0$ eindeutig, wenn wir φ auf ein halboffenes Intervall der Länge 2π , z. B. auf das Intervall $]-\pi, \pi]$ oder $[0, 2\pi[$ beschränken. Der erste Fall entspricht der Definition (3). Wir nennen die im Fall $z \neq 0$ eindeutig bestimmte reelle Zahl φ mit $-\pi < \varphi \leq \pi$, die der Gleichung (4) bzw. (6) mit $r > 0$ genügt, das *Hauptargument* der komplexen Zahl z . Die Darstellung

$$z = r(\cos \varphi + i \sin \varphi) = re^{i\varphi} \quad (r > 0, -\pi < \varphi \leq \pi) \quad (7)$$

ist daher eindeutig. Es ist stets

$$(r_1 e^{i\varphi_1}) (r_2 e^{i\varphi_2}) = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}. \quad (8)$$

Hieraus ersehen wir erneut, daß der Betrag des Produkts zweier komplexer Zahlen gleich dem Produkt ihrer Beträge ist (vgl. 1.2.2.(9)). Ferner ist die Summe zweier beliebiger Argumente der Faktoren stets ein Argument des Produkts.

Wir beweisen die *Moivreschen Formeln* (ABRAHAM DE MOIVRE (1667—1754))

$$\cos n\varphi = \sum_{k=0}^n (-1)^k \binom{n}{2k} \cos^{n-2k} \varphi \sin^{2k} \varphi, \quad (9)$$

$$\sin n\varphi = \sum_{k=0}^n (-1)^k \binom{n}{2k+1} \cos^{n-(2k+1)} \varphi \sin^{2k+1} \varphi. \quad (10)$$

Sie sind implizit bereits in einer Arbeit von MOIVRE aus dem Jahre 1707 zu finden.

Explizit treten sie erstmals 1748 bei EULER auf. Für $k > \frac{n}{2}$ verschwinden alle Summanden wegen

$$\binom{n}{j} = 0 \quad (j > n). \quad (11)$$

Durch vollständige Induktion zeigt man, daß

$$\cos n\varphi + i \sin n\varphi = e^{in\varphi} = (e^{i\varphi})^n = (\cos \varphi + i \sin \varphi)^n$$

ist. Anwendung der binomischen Formel ergibt

$$\cos n\varphi + i \sin n\varphi = \sum_{j=0}^n \binom{n}{j} i^j \sin^j \varphi \cos^{n-j} \varphi.$$

Zur Vermeidung von Fallunterscheidungen erstrecken wir die Summation bis $2n+1$, was wegen (11) zulässig ist. Wir erhalten

$$\begin{aligned} \cos n\varphi + i \sin n\varphi &= \sum_{j=0}^{2n+1} \binom{n}{j} i^j \sin^j \varphi \cos^{n-j} \varphi \\ &= \sum_{k=0}^n \binom{n}{2k} i^{2k} \sin^{2k} \varphi \cos^{n-2k} \varphi \\ &\quad + \sum_{k=0}^n \binom{n}{2k+1} i^{2k+1} \sin^{2k+1} \varphi \cos^{n-(2k+1)} \varphi. \end{aligned}$$

Aus $i^{2k} = (-1)^k$, $i^{2k+1} = (-1)^k i$ ergibt sich durch Vergleich von Real- und Imaginärteil beider Seiten die Behauptung.

Jeder Punkt der Ebene ist durch seine Koordinaten x, y bezüglich eines kartesischen Koordinatensystems bestimmt. Auf Grund von (2) kann er aber auch durch die ebenen Polarkoordinaten r, φ festgelegt werden. Wählen wir $\varphi = \varphi_0$ bzw. $r = r_0$ konstant und variieren nur φ bzw. r , so erhalten wir die *Koordinatenlinien* dieses *krümmelinigen Koordinatensystems*. Es sind im ersten Fall vom Ursprung ausgehende Halbgeraden, im zweiten Fall konzentrische Kreise mit dem Mittelpunkt O .

Eine Verallgemeinerung dieser Betrachtungen führt zu krummlinigen Koordinaten im Raum. Durch die Gleichungen

$$\left. \begin{aligned} x &= \varrho \cos \varphi, \\ y &= \varrho \sin \varphi, \\ z &= z \end{aligned} \right\} \quad (12)$$

werden *Zylinderkoordinaten* ϱ, φ, z eingeführt (Abb. 1.36). Hierbei ist $\varrho := \sqrt{x^2 + y^2}$ der Abstand des Punktes (x, y, z) von der z -Achse. Durch die Gleichungen

$$\left. \begin{aligned} x &= r \sin \vartheta \cos \varphi, \\ y &= r \sin \vartheta \sin \varphi, \\ z &= r \cos \vartheta \end{aligned} \right\} \quad (13)$$

werden *räumliche Polarkoordinaten* r, ϑ, φ eingeführt (Abb. 1.37). Hierbei ist $r := \sqrt{x^2 + y^2 + z^2}$ der Abstand des Punktes (x, y, z) vom Ursprung, der Winkel ϑ die von der positiven z -Achse aus gemessene *Poldistanz*, und der Winkel φ entspricht dem geographischen „Längengrad“ auf der Erdkugel.

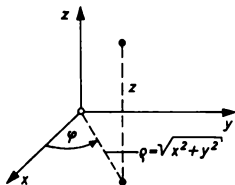


Abb. 1.36

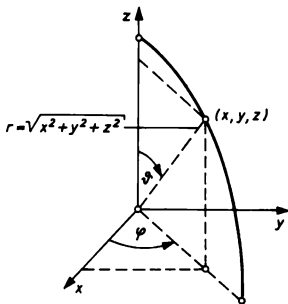


Abb. 1.37

1.6.4. Einheitskreis und Bogenlänge

In diesem Abschnitt geben wir den angekündigten Beweis für die Existenz zweier Funktionen mit den Eigenschaften 1.6.1.(1) bis (5) auf der Grundlage des Begriffs „Bogenlänge“. Wiederum erläutern wir die auftretenden Begriffsbildungen anschaulich, während die eigentlichen Definitionen und Sätze unabhängig von der geometrischen Deutung formuliert bzw. bewiesen werden.

Es sei $-1 \leq x < 1$. Mit $\mathfrak{Z}(x)$ bezeichnen wir die Menge aller *Zerlegungen*

$$\mathfrak{Z}: x = x_0 < x_1 < \dots < x_n = 1 \quad (1)$$

des Intervalls $[[x, 1]]$ in endlich viele Teilpunkte. Ferner sei

$$y_i := \sqrt{1 - x_i^2} \quad (2)$$

und

$$d(x_i, x_j) := \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (3)$$

In der anschaulichen Deutung liegen die Punkte $r_i := (x_i, \sqrt{1 - x_i^2})$ auf dem oberen Halbkreis des Einheitskreises, und $d(x_i, x_j)$ ist der Abstand der Punkte r_i, r_j (Abb. 1.38). Die Dreiecksungleichung besagt, daß

$$d(x_i, x_k) \leq d(x_i, x_j) + d(x_j, x_k)$$

ist. Nach Satz 3 in 1.5.2. ist andererseits

$$d(x_i, x_j) \leq |x_i - x_j| + |y_i - y_j|. \quad (4)$$

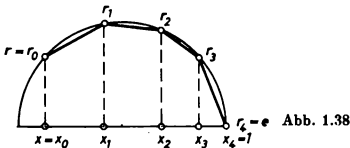


Abb. 1.38

Wegen $x_i^2 + y_i^2 = 1$ kann (3) auch in der Form

$$d(x_i, x_j) = \sqrt{2(1 - x_i x_j - y_i y_j)} \quad (5)$$

geschrieben werden.

Für jede Zerlegung $Z \in \mathcal{Z}(x)$ gibt die Zahl

$$S(Z) := \sum_{i=1}^n d(x_{i-1}, x_i) \quad (6)$$

die Länge des *Sehnenzuges* $r_0 r_1 \dots r_n$ an. Entsteht die Zerlegung Z' aus Z durch eine „Verfeinerung“, d. h. durch Hinzufügen weiterer Teilpunkte, so ist nach der Dreiecksungleichung stets $S(Z) \leq S(Z')$. Wir beweisen, daß die Zahl 4 eine obere Schranke für alle Summen (6) ist. Es sei Z_0 eine beliebige Zerlegung aus $\mathcal{Z}(x)$. Wir ergänzen sie zu einer Zerlegung Z des Intervalls $[-1, 1]$, die auch den Teilpunkt 0 enthält. Es gelte etwa

$$Z: -1 = x_0 < x_1 < \dots < x_m = 0 < x_{m+1} < \dots < x_{m+n} = 1.$$

Für $i = 1, \dots, m$ ist $y_{i-1} < y_i$, für $i = m+1, \dots, m+n$ dagegen $y_i < y_{i-1}$. Aus (4) und $y_0 = y_{m+n} = 0, y_m = 1$ folgt

$$\begin{aligned} S(Z_0) &\leq S(Z) \leq \sum_{i=1}^{m+n} (|x_i - x_{i-1}| + |y_i - y_{i-1}|) \\ &= \sum_{i=1}^{m+n} (x_i - x_{i-1}) + \sum_{i=1}^m (y_i - y_{i-1}) + \sum_{i=m+1}^{m+n} (y_{i-1} - y_i) \\ &= (x_{m+n} - x_0) + (y_m - y_0) + (y_m - y_{m+n}) = 4. \end{aligned}$$

Dies rechtfertigt die

Definition. Für jede Zahl x mit $-1 \leq x < 1$ heißt die kleinste obere Schranke aller Zahlen $S(Z)$ mit $Z \in \mathfrak{Z}(x)$ die *Bogenlänge* des vom Punkt $e = (1, 0)$ zum Punkt $r = (x, \sqrt{1-x^2})$ führenden Bogens des Einheitskreises.

Wir bezeichnen diese Bogenlänge mit $\sigma(x)$. Für $x = 1$ sei

$$\sigma(1) := 0.$$

Nach Definition gilt

$$\sigma(x) = \sup_{Z \in \mathfrak{Z}(x)} S(Z) \quad (-1 \leq x < 1).$$

Die in dieser Weise definierte Funktion σ mit $D(\sigma) := [-1, 1]$ wird sich als identisch mit der Arkuskosinusfunktion erweisen. Die Zahl π definieren wir jetzt als Bogenlänge des Halbkreises, d. h., wir setzen

$$\pi := \sigma(-1).$$

Die folgende Abschätzung wird zum Beweis von 1.6.1.(5) führen.

Satz 1. Für $0 < x < 1$ ist

$$\sqrt{1-x^2} < \sigma(x) \leq 2 \sqrt{\frac{1-x}{1+x}} \cdot \frac{\sqrt{1-x^2}}{x}. \quad (7)$$

Beweis. Zunächst ist $\sqrt{1-x^2} = y < \sqrt{(1-x)^2 + y^2} = d(1, x) < \sigma(x)$. Zum Beweis der zweiten Ungleichung (7) zeigen wir, daß $d(x_i, x_j)$ für $i < j$ mit der Zahl

$$a_{ij} := \sqrt{1-x_i} \sqrt{1+x_j} - \sqrt{1+x_i} \sqrt{1-x_j}$$

übereinstimmt. Wegen $x_j - x_i > x_i - x_j$ ist $(1-x_i)(1+x_j) > (1+x_i)(1-x_j)$, also auch $a_{ij} > 0$. Andererseits ist

$$\begin{aligned} a_{ij}^2 &= (1-x_i)(1+x_j) + (1+x_i)(1-x_j) - 2\sqrt{1-x_i^2}\sqrt{1-x_j^2} \\ &= 2(1-x_ix_j - y_iy_j) = d(x_i, x_j)^2. \end{aligned}$$

und folglich ist

$$\begin{aligned} d(x_i, x_j) &= a_{ij} = \sqrt{1+x_i} \sqrt{1-x_j} \left(\sqrt{\frac{1-x_i}{1+x_i}} - \sqrt{\frac{1-x_j}{1+x_j}} \right) \\ &< 2 \left(\sqrt{\frac{1-x_i}{1+x_i}} - \sqrt{\frac{1-x_j}{1+x_j}} \right) \quad (i < j). \end{aligned}$$

In (6) eingesetzt, erhalten wir

$$S(Z) < 2 \sum_{k=1}^n \left(\sqrt{\frac{1-x_{k-1}}{1+x_{k-1}}} - \sqrt{\frac{1-x_k}{1+x_k}} \right) = 2 \left(\sqrt{\frac{1-x_0}{1+x_0}} - \sqrt{\frac{1-x_n}{1+x_n}} \right) = 2 \sqrt{\frac{1-x}{1+x}}.$$

Da Z beliebig gewählt war, gilt auch die zweite Ungleichung (7). Schließlich ist

$$2 \sqrt{\frac{1-x}{1+x}} = 2 \sqrt{\frac{(1-x)(1+x)}{(1+x)(1+x)}} = 2 \frac{\sqrt{1-x^2}}{1+x} < \frac{\sqrt{1-x^2}}{x},$$

denn es ist $2x < 1+x$.

Die Aussage des nachfolgenden Satzes machen wir an Hand von Abb. 1.39 plausibel. Es sei $-1 \leq x' < 1$ und x ein Punkt des Intervalls $[[x', 1]]$. Wir spiegeln den zu x gehörenden Kreispunkt $r = (x, \sqrt{1-x^2}) = (x, y)$ an der Mittelsenkrechten der Verbindungsstrecke der Kreispunkte $e = (1, 0)$ und $r' = (x', \sqrt{1-x'^2}) = (x', y')$. Der Bildpunkt $r^* = (x^*, \sqrt{1-x^{*2}}) = (x^*, y^*)$ muß dann der Bedingung

$$d(x', x) = \|r' - r\| = \|e - r^*\| = d(1, x^*),$$

genügen. Nach (5) ist

$$2(1 - xx' - yy') = d(x', x)^2 = d(1, x^*)^2 = 2(1 - x^*),$$

d. h., es ist $x^* = xx' + yy'$. Da sich der Bogen von e nach r' aus den Bögen von e nach r und von r nach r' zusammensetzt und der letztere zum Bogen von e nach r^* kongruent ist, vermuten wir, daß stets $\sigma(x') = \sigma(x) + \sigma(x^*)$ ist. Dies ist die wesentliche Aussage von

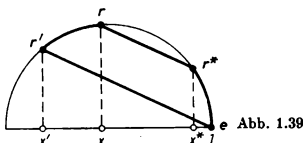


Abb. 1.39

Satz 2. Für alle x, x' mit $-1 \leq x' \leq x \leq 1$ ist

$$\sigma(x') - \sigma(x) = \sigma(xx' + \sqrt{1-x^2} \sqrt{1-x'^2}), \quad (8)$$

$$\sigma\left(\sqrt{\frac{1+x}{2}}\right) = \frac{1}{2} \sigma(x). \quad (9)$$

Beweis. Wir definieren auf dem Intervall $[[x', 1]]$ zwei Funktionen f, g durch

$$\left. \begin{aligned} f(x) &:= xx' + yy', \\ g(x) &:= xy' - yx' \end{aligned} \right\} \quad (x' \leq x \leq 1),$$

wobei wie üblich $y' := \sqrt{1-x'^2}$, $y := \sqrt{1-x^2}$ zu setzen ist. Offensichtlich ist

$$f(x') = 1, \quad f(1) = x'. \quad (10)$$

Wir zeigen, daß $g(x) \geq 0$ ist. Im Fall $0 \leq x' \leq x \leq 1$ ist $y \leq y'$, also $yx' \leq xy'$. Für $x' < 0 < x$ ist dies offensichtlich. Ist schließlich $x' \leq x \leq 0$, so ist $0 \leq y \leq y'$ und wiederum $yx' \leq xy'$. In allen Fällen ist also $g(x) \geq 0$. Ist $x_1, x_2 \in [[x', 1]]$ so haben wir

$$\begin{aligned} f(x_1) f(x_2) + g(x_1) g(x_2) &= (x_1 x' + y_1 y') (x_2 x' + y_2 y') + (x_1 y' - y_1 x') (x_2 y' - y_2 x') \\ &= x_1 x_2 x'^2 + y_1 y_2 y'^2 + x_1 x_2 y'^2 + y_1 y_2 x'^2, \end{aligned}$$

denn die restlichen Glieder heben sich heraus. Es folgt

$$f(x_1) f(x_2) + g(x_1) g(x_2) = x_1 x_2 + y_1 y_2, \quad (11)$$

und speziell ist $f(x)^2 + g(x)^2 = 1$ für alle $x \in [x', 1]$. Setzen wir also $x^* := f(x)$, so ist $y^* := \sqrt{1 - x^{*2}} = g(x)$. Aus (11) und (5) folgt nun

$$d(x_1, x_2) = d(x_1^*, x_2^*) \quad (x_1, x_2 \in [x', 1]). \quad (12)$$

Als nächstes zeigen wir, daß stets

$$x_1^* > x_2^* \quad \text{für} \quad x' \leq x_1 < x_2 \leq 1 \quad (13)$$

ist. Dies ist richtig, wenn $y_1 = y_2 = 0$ ist, denn dann ist $x' = x_1 = -1$, $x_2 = 1$ und $x_1^* = 1$, $x_2^* = -1$. Andernfalls ist $y_1 + y_2 > 0$, und wegen $y_1^2 - y_2^2 = x_2^2 - x_1^2$ ist

$$\begin{aligned} (x_1^* - x_2^*) (y_1 + y_2) &= ((x_1 - x_2) x' + (y_1 - y_2) y') (y_1 + y_2) \\ &= (x_1 - x_2) (y_1 + y_2) x' + (y_1^2 - y_2^2) y' \\ &= (x_1 - x_2) (y_1 + y_2) x' + (x_2^2 - x_1^2) y' \\ &= (x_2 - x_1) ((x_1 + x_2) y' - (y_1 + y_2) x') \\ &= (x_2 - x_1) (y_1^* + y_2^*), \end{aligned} \quad (14)$$

woraus $x_1^* - x_2^* \geq 0$ folgt. Nun ist aber $d(x_1^*, x_2^*) = d(x_1, x_2) > 0$, und folglich gilt $x_1^* > x_2^*$. Damit ist (13) bewiesen.

Es sei nun $-1 \leq x' \leq x \leq 1$. Für $x' = x$ oder $x = 1$ ist die Behauptung (8) trivial. Im Fall $-1 \leq x' < x < 1$ wählen wir eine beliebige Zerlegung Z' des Intervalls $[x', 1]$. Durch Hinzufügen des Teilpunktes x entstehe die Zerlegung

$$Z'_0: x' = x_0 < x_1 < \dots < x_m = x < x_{m+1} < \dots < x_{m+n} = 1.$$

Dann ist

$$Z: x = x_m < x_{m+1} < \dots < x_{m+n} = 1$$

eine Zerlegung von $[x, 1]$. Nach (13) ist $x_i^* < x_{i-1}^*$, und wegen $x_0^* = f(x') = 1$ ist

$$Z^*: x^* = x_m^* < \dots < x_1^* < x_0^* = 1$$

eine Zerlegung von $[x^*, 1]$. Aus (12) folgt

$$\begin{aligned} S(Z') &\leq S(Z'_0) = \sum_{i=1}^m d(x_i^*, x_{i-1}^*) + \sum_{i=1}^n d(x_{m+i-1}, x_{m+i}) \\ &= S(Z^*) + S(Z) \leq \sigma(x^*) + \sigma(x). \end{aligned}$$

Da Z' beliebig gewählt war, ist $\sigma(x') \leq \sigma(x^*) + \sigma(x)$.

Sind umgekehrt zwei beliebige Zerlegungen

$$Z: x = x_0 < x_1 < \dots < x_m = 1,$$

$$Z^*: x^* = \bar{x}_0 < \bar{x}_1 < \dots < \bar{x}_n = 1$$

von $[x, 1]$ bzw. $[x^*, 1]$ vorgegeben, so ist $f(\bar{x}_n) = f(1) = x'$,

$$f(\bar{x}_0) = f(x^*) = x^* x' + y^* y' = (xx' + yy') x' + (xy' - yx') y' = xx'^2 + xy'^2 = x,$$

und

$$Z': x' = f(\bar{x}_n) < \dots < f(\bar{x}_0) = x_0 < x_1 < \dots < x_m = 1$$

ist eine Zerlegung aus $\mathcal{B}(x')$. Ferner ist

$$S(Z^*) + S(Z) = \sum_{i=1}^n d(f(\bar{x}_i); f(\bar{x}_{i-1})) + \sum_{k=1}^m d(x_{k-1}, x_k) = S(Z') \leq \sigma(x').$$

Dies ist nur möglich, wenn auch $\sigma(x^*) + \sigma(x) \leq \sigma(x')$ ist. Damit ist (8) bewiesen.

Im Spezialfall $x := \sqrt{\frac{1+x'}{2}}$ ist $y = \sqrt{1-x^2} = \sqrt{\frac{1-x'}{2}}$ und

$$x^* = \sqrt{\frac{1+x'}{2}} x' + \sqrt{\frac{1-x'}{2}} (1-x'^2) = \sqrt{\frac{1+x'}{2}} x' + (1-x') \sqrt{\frac{1+x'}{2}} = \sqrt{\frac{1+x'}{2}} = x,$$

d. h., es ist $\sigma(x') = \sigma(x) + \sigma(x^*) = 2\sigma(x)$. Mit x statt x' erhalten wir (9), und Satz 2 ist bewiesen.

Satz 3. Die Funktion σ besitzt eine auf dem Intervall $[0, \pi]$ definierte Umkehrfunktion C .

Beweis. Aus $-1 \leq x' \leq x \leq 1$ und $x^* := xx' + yy'$ folgt $\sigma(x') - \sigma(x) = \sigma(x^*) \geq 0$. Wir zeigen, daß aus $\sigma(x') = \sigma(x)$ stets $x' = x$ folgt. Es ist dann nämlich $\sigma(x^*) = 0$, $x^* = 1$, also

$$xx' + yy' = 1,$$

$$xy' - yx' = 0.$$

Multiplizieren wir mit x' bzw. y' und addieren, so erhalten wir $x = x'$. Die Funktion σ ist daher streng monoton fallend und besitzt eine Umkehrfunktion C mit $C(0) = 1$, $C(\pi) = -1$. Zum Beweis unseres Satzes brauchen wir nur noch zu zeigen, daß $D(C) = W(\sigma)$ dicht in $[0, \pi]$ liegt. Hierzu wiederum genügt der Nachweis, daß für $\alpha, \beta \in D(C)$ und $\alpha \leq \beta$ stets $\frac{\alpha}{2}, \beta - \alpha \in D(C)$ ist. Durch Induktion ergibt sich nämlich leicht, daß alle Zahlen der Form $\pi - \frac{k}{2^n}\pi$ ($k = 0, \dots, 2^n$) in $D(C)$ liegen, woraus unsere Behauptung folgt.

Es sei $x = C(\alpha)$, $x' = C(\beta)$. Dann ist $-1 \leq x' \leq x \leq 1$ und $\beta - \alpha = \sigma(x') - \sigma(x) = \sigma(xx' + yy') \in W(\sigma) = D(C)$. Nach (9) ist schließlich

$$\frac{\alpha}{2} = \frac{1}{2} \sigma(x) = \sigma\left(\sqrt{\frac{1+x}{2}}\right) \in W(\sigma) = D(C),$$

und Satz 3 ist bewiesen.

Mit $x' = -1$, $x = C(\alpha)$ geht (8) in $\pi - \alpha = \sigma(-1) - \sigma(x) = \sigma(-x)$ über, d. h., es ist $C(\pi - \alpha) = -x$, also

$$C(\pi - \alpha) = -C(\alpha). \quad (15)$$

Mit $\alpha := \frac{\pi}{2}$ folgt hieraus

$$C\left(\frac{\pi}{2}\right) = 0. \quad (16)$$

Für die Funktion $S(\alpha) := \sqrt{1 - C^2(\alpha)}$ ($0 \leq \alpha \leq \pi$) gilt dann

$$S(\pi - \alpha) = S(\alpha), \quad (17)$$

$$S\left(\frac{\pi}{2}\right) = 1. \quad (18)$$

Jede reelle Zahl α kann auf genau eine Weise in der Form

$$\alpha = \alpha_0 + k\pi \quad (0 \leq \alpha_0 < \pi, k \in \mathbb{Z}) \quad (19)$$

dargestellt werden. Wir definieren nun die trigonometrischen Funktionen durch periodische Fortsetzung der Funktionen C, S . Ist α in der Form (19) dargestellt, so setzen wir

$$\cos \alpha := (-1)^k C(\alpha_0), \quad (20)$$

$$\sin \alpha := (-1)^k S(\alpha_0).$$

Für zwei reelle Zahlen α, β seien $\alpha = \alpha_0 + k\pi, \beta = \beta_0 + m\pi$ die Darstellungen gemäß (19). Wir unterscheiden die Fälle $\beta_0 \leq \alpha_0$ und $\alpha_0 < \beta_0$. Im ersten Fall setzen wir $x' := C(\alpha_0)$, $x := C(\beta_0)$. Dann ist

$$\alpha_0 - \beta_0 = \sigma(x') - \sigma(x) = \sigma(xx' + yy'),$$

also

$$C(\alpha_0 - \beta_0) = xx' + yy' = C(\alpha_0) C(\beta_0) + S(\alpha_0) S(\beta_0).$$

Wegen $(-1)^{k-m} = (-1)^{k+m}$ folgt

$$\begin{aligned} \cos(\alpha - \beta) &= \cos((\alpha_0 - \beta_0) + (k - m)\pi) = (-1)^{k-m} C(\alpha_0 - \beta_0) \\ &= (-1)^{k+m} (C(\alpha_0) C(\beta_0) + S(\alpha_0) S(\beta_0)) \\ &= \cos \alpha \cos \beta + \sin \alpha \sin \beta. \end{aligned} \quad (22)$$

Im zweiten Fall ist $0 < \pi - (\beta_0 - \alpha_0) < \pi$, und aus (15) folgt

$$\begin{aligned} \cos(\alpha - \beta) &= \cos((\pi - (\alpha_0 - \beta_0)) + (k - m - 1)\pi) \\ &= (-1)^{k-m-1} C(\pi - (\alpha_0 - \beta_0)) = (-1)^{k-m} C(\beta_0 - \alpha_0), \end{aligned}$$

d. h., es gilt wiederum (22). Speziell ist

$$\begin{aligned} \cos\left(\alpha - \frac{\pi}{2}\right) &= \cos \alpha \cos \frac{\pi}{2} + \sin \alpha \sin \frac{\pi}{2} = \sin \alpha, \\ \sin\left(\alpha - \frac{\pi}{2}\right) &= \cos(\alpha - \pi) = \cos \alpha \cos \pi + \sin \alpha \sin \pi = -\cos \alpha. \end{aligned}$$

Es folgt

$$\begin{aligned} \sin(\alpha - \beta) &= \cos\left(\alpha - \beta - \frac{\pi}{2}\right) = \cos\left(\alpha - \frac{\pi}{2}\right) \cos \beta + \sin\left(\alpha - \frac{\pi}{2}\right) \sin \beta \\ &= \sin \alpha \cos \beta - \cos \alpha \sin \beta. \end{aligned}$$

Damit sind die Subtraktionstheoreme bewiesen.

Es sei schließlich $0 < \alpha < 1$, $x := \cos \alpha$. Wegen $\frac{\pi}{2} = \sigma(0) > d(1, 0) = \sqrt{2} > 1$ ist dann $0 < \alpha < \frac{\pi}{2}$, und aus (7) folgt

$$\sin \alpha = S(\alpha) = \sqrt{1 - x^2} < \sigma(x) = \alpha < \frac{\sqrt{1 - x^2}}{x} = \frac{\sin \alpha}{\cos \alpha}.$$

Damit gilt auch 1.6.1.(5), und der Existenzbeweis für die trigonometrischen Funktionen ist geführt.

2. Der Grenzwertbegriff

2.1. Zahlenfolgen

2.1.1. Spezielle Eigenschaften von Zahlenfolgen

In 1.3.1. haben wir den allgemeinen Begriff der Folge eingeführt. Wir betrachten einige spezielle Zahlenfolgen $(a_n)_{n \in \mathbf{N}}$. Es sei

$$a_n := 2(-1)^n n, \quad (1)$$

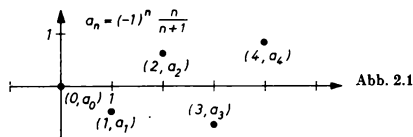
$$a_n := \begin{cases} 0 & \text{für gerade } n, \\ n & \text{für ungerade } n, \end{cases} \quad (2)$$

$$a_n := \frac{n}{n+1}, \quad (3)$$

$$a_n := 2i, \quad (4)$$

$$a_n := \frac{1}{n+1} + ni. \quad (5)$$

Durch (1) bis (3) sind reelle, durch (4), (5) komplexe Zahlenfolgen definiert. In anderer Schreibweise gibt man z. B. die Folgen (2) bzw. (4) durch $(0, 1, 0, 3, 0, 5, \dots)$ bzw. $(2i, 2i, 2i, \dots)$ an. Für die Veranschaulichung von reellen Zahlenfolgen gibt es zwei Möglichkeiten. Wir können sie durch ihre Graphen veranschaulichen. Der Graph einer Zahlenfolge besteht aus allen Punkten (n, a_n) mit $n \in \mathbf{N}$. Auf jeder Parallelen zur y -Achse durch einen Punkt $(n, 0)$ mit $n \in \mathbf{N}$ liegt genau ein Punkt des Graphen (Abb. 2.1).



Eine zweite Möglichkeit der Veranschaulichung von reellen Zahlenfolgen besteht darin, daß man den zu n gehörigen Funktionswert a_n der Zahlenfolge auf der Zahlengeraden markiert und durch das Symbol a_n kennzeichnet.

Diese zweite Möglichkeit der Veranschaulichung besteht auch für komplexe Zahlenfolgen, wenn man von der Zahlengeraden zur Gaußschen Zahlenebene übergeht. In Abb. 2.2 ist die Folge (5) veranschaulicht.

Die Begriffe „nach oben“ bzw. „nach unten beschränkt“ für reelle Zahlenfolgen sowie „beschränkt“ für alle Zahlenfolgen werden ebenso wie für Funktionen definiert, denn Zahlenfolgen sind ja spezielle Funktionen. Die Folgen (1), (5) sind z. B. nicht beschränkt, die Folge (2) ist nach unten, aber nicht nach oben beschränkt, und die Folgen (3), (4) sind beschränkt.

Eine konstante Folge wird auch *stationär* genannt. So ist z. B. die Folge (4) stationär.

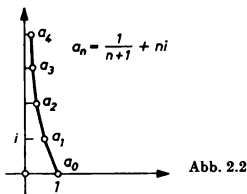


Abb. 2.2

Wir haben bisher stets vorausgesetzt, daß Folgen die Menge aller natürlichen Zahlen als Definitionsbereich besitzen. In vielen Fällen ist es zur Vereinfachung der Schreibweise zweckmäßig, auch solche Funktionen als Folgen zu bezeichnen, deren Definitionsbereich die Menge aller natürlichen (oder auch ganzen) Zahlen n mit $n \geq n_0$ ist, wobei der *Anfangsindex* n_0 eine beliebige natürliche (bzw. ganze) Zahl ist. Wenn Mißverständnisse zu befürchten sind, bezeichnen wir solche Folgen ausführlicher mit $(a_n)_{n \geq n_0}$.

Der Begriff der Monotonie kann ebenso wie für Funktionen definiert werden. Da der Definitionsbereich von Folgen „diskret“ ist, kann die Monotonie auch wie folgt definiert werden.

Definition 1. Eine reelle Zahlenfolge (a_n) heißt

monoton wachsend, wenn $a_n \leq a_{n+1}$,

monoton fallend, wenn $a_n \geq a_{n+1}$,

streng monoton wachsend, wenn $a_n < a_{n+1}$,

streng monoton fallend, wenn $a_n > a_{n+1}$

für alle natürlichen Zahlen n ist.

Durch vollständige Induktion zeigt man, daß aus $a_n \leq a_{n+1}$ für alle natürlichen Zahlen n stets $a_n \leq a_{n+k}$ für alle natürlichen Zahlen n und k folgt. Analoge Aussagen gelten in den anderen Fällen. Die Definitionen sind daher mit der früher gegebenen Definition äquivalent.

Beispiel 1. Für eine beliebige positive reelle Zahl x mit $x \neq 1$ betrachten wir die Zahlenfolge mit den Gliedern

$$a_n := n(\sqrt[n]{x} - 1) \quad (n \geq 1).$$

Ersetzen wir x, y, m, n in 1.1.2.(26) der Reihe nach durch $\frac{n(n+1)}{2}, x, 1, n, 1$, so erhalten wir

$$(n+1) \sqrt[n+1]{x} < n \sqrt[n]{x} + 1.$$

Wir subtrahieren auf beiden Seiten $n+1$. Es folgt

$$a_{n+1} = (n+1) \left(\sqrt[n+1]{x} - 1 \right) < n \left(\sqrt[n]{x} - 1 \right) = a_n; \quad (6)$$

die Folge ist streng monoton fallend.

Beispiel 2. Für eine beliebige von 0 verschiedene reelle Zahl x betrachten wir die Zahlenfolge mit den Gliedern

$$a_n := \left(1 + \frac{x}{n} \right)^n \quad (n > -x).$$

Wir nehmen an, für $n > -x$ wäre

$$a_n = \left(\frac{n+x}{n} \right)^n \geq \left(\frac{n+1+x}{n+1} \right)^{n+1} = a_{n+1}.$$

Multiplikation mit $\left(\frac{n}{n+x} \right)^{n+1}$ und Anwendung der Bernoullischen Ungleichung ergibt

$$\begin{aligned} \frac{n}{n+x} &\geq \left(\frac{n+1+x}{n+1} \cdot \frac{n}{n+x} \right)^{n+1} = \left(1 - \frac{x}{(n+1)(n+x)} \right)^{n+1} \\ &> 1 - (n+1) \frac{x}{(n+1)(n+x)} = 1 - \frac{x}{n+x} = \frac{n}{n+x}, \end{aligned}$$

was ein Widerspruch ist. Für $n > -x$ gilt somit

$$0 < a_n = \left(1 + \frac{x}{n} \right)^n < \left(1 + \frac{x}{n+1} \right)^{n+1} = a_{n+1}; \quad (7)$$

die Folge ist streng monoton wachsend.

Die beiden folgenden Definitionen sind naheliegend.

Definition 2. Sind (a_n) und (b_n) Zahlenfolgen, so heißen die Zahlenfolgen $(a_n + b_n)$, $(a_n - b_n)$, $(a_n b_n)$ die *Summe*, die *Differenz* bzw. das *Produkt* der beiden Zahlenfolgen. Ist $b_n \neq 0$ für alle n mit $n \in \mathbb{N}$, so heißt die Zahlenfolge $\left(\frac{a_n}{b_n} \right)$ der *Quotient* der beiden Zahlenfolgen.

Definition 3. Ist $(n_k)_{k \in \mathbb{N}}$ eine streng monoton wachsende Zahlenfolge mit $n_k \in \mathbb{N}$ für alle $k \in \mathbb{N}$, so heißt die Zahlenfolge $(a_{n_k})_{k \in \mathbb{N}}$ eine *Teilfolge* der Zahlenfolge (a_n) .

So ist z. B. $(1, 3, 5, \dots, 2n+1, \dots)$ eine Teilfolge der Folge $(0, 1, 2, \dots, n, \dots)$.

2.1.2. Nullfolgen

Wir betrachten in diesem Abschnitt eine spezielle Klasse von reellen Zahlenfolgen, die für den weiteren Aufbau der Analysis von besonderer Bedeutung ist, und zwar wollen wir den Begriff der Nullfolge einführen. Eine wesentliche Eigenschaft einer Nullfolge besteht darin, daß ihre Glieder a_n mit wachsendem n der Zahl 0 „beliebig“ nahekommen. Um diesen Sachverhalt zu präzisieren, ordnen wir für eine vorgegebene Folge (a_n) jeder positiven reellen Zahl ε eine Teilmenge $M(\varepsilon)$ der Menge der natürlichen Zahlen zu. Wir setzen

$$M(\varepsilon) := \{n : |a_n| \geq \varepsilon\}. \quad (1)$$

Ist $n \in M(\varepsilon)$, d. h., ist $|a_n| \geq \varepsilon$, so liegt der Punkt (n, a_n) in dem in Abb. 2.3 schraffierten Bereich. Für unsere weiteren Betrachtungen ist es wichtig zu entscheiden, ob $M(\varepsilon)$ eine endliche oder eine unendliche Menge ist, d. h., ob in dem in Abb. 2.3 schraffierten Bereich endlich oder unendlich viele Punkte des Graphen der Folge liegen.

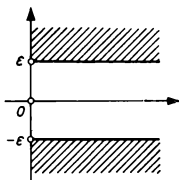


Abb. 2.3

Beispiel 1. Gegeben sei die Folge $\left(\frac{(-1)^n + n}{10n + 1}\right)$. Gemäß (1) setzen wir

$$M(\varepsilon) = \left\{n : \left|\frac{(-1)^n + n}{10n + 1}\right| \geq \varepsilon\right\} = \{n : (-1)^n + n \geq \varepsilon(10n + 1)\}.$$

Es ist

$$M(2) = \emptyset, \quad M(1) = \{0\}, \quad M\left(\frac{1}{8}\right) = \{0, 2\}, \quad M\left(\frac{1}{9}\right) = \{0, 2, 4, 6, 8\}.$$

Diese Mengen sind endlich. Dagegen ist $M\left(\frac{1}{10}\right)$ die Menge aller geraden natürlichen Zahlen, also nicht endlich. Dies bedeutet, daß die Glieder a_n der Folge mit wachsendem n der Zahl 0 nicht „beliebig“ nahekommen, denn es gibt „beliebig große“ natürliche Zahlen n mit $|a_n| \geq \frac{1}{10}$.

Beispiel 2. Gegeben sei die Folge $\left(\frac{(-1)^n \cdot 5}{n + 1}\right)$. Dann ist

$$M(\varepsilon) = \left\{n : \frac{5}{n + 1} \geq \varepsilon\right\} = \left\{n : n \leq \frac{5}{\varepsilon} - 1\right\}.$$

Für jede positive reelle Zahl ε ist die Menge $M(\varepsilon)$ endlich. Mit Ausnahme von höchstens endlich vielen Punkten liegen die Punkte (n, a_n) innerhalb des sogenannten ε -Streifens in Abb. 2.4. Da $M(\varepsilon)$ stets eine endliche Menge natürlicher Zahlen ist, besitzt jede dieser Mengen eine obere Schranke. Wenn n größer als diese obere Schranke ist, gilt $|a_n| < \varepsilon$. Da die positive Zahl ε beliebig klein gewählt werden kann, kommen die Glieder a_n der Folge mit wachsendem n der Zahl 0 beliebig nahe.

Wir können nun den Begriff der Nullfolge wie folgt präzisieren.

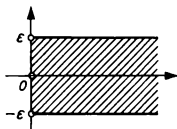


Abb. 2.4

Definition 1. Die reelle Zahlenfolge (a_n) heißt eine *Nullfolge*, wenn für jede positive reelle Zahl ε die Menge aller natürlichen Zahlen n mit $|a_n| \geq \varepsilon$ endlich ist.

Die Folge in Beispiel 2 ist eine Nullfolge, die Folge in Beispiel 1 dagegen nicht.

Als unmittelbare Folgerung aus Definition 1 ergibt sich, daß jede Teilfolge einer Nullfolge wieder eine Nullfolge ist. Aus der Definition ersehen wir ferner sofort, daß die Eigenschaft einer Folge (a_n) , Nullfolge zu sein, erhalten bleibt, wenn man endlich viele Folgenglieder durch beliebige andere reelle Zahlen ersetzt. Bei der Untersuchung einer gegebenen Folge auf die Nullfolgeeigenschaft können wir also stets endlich viele Folgenglieder unbeachtet lassen.

In der Mathematik sagt man, „für fast alle n gilt die Aussage $H(n)$ “, wenn die Menge der natürlichen Zahlen, für die $H(n)$ nicht erfüllt ist, endlich ist. Wir können damit den Begriff der Nullfolge auch wie folgt definieren.

Definition 1'. Die reelle Zahlenfolge (a_n) heißt eine *Nullfolge*, wenn für jede positive Zahl ε gilt: Für fast alle n ist $|a_n| < \varepsilon$.

In Definition 1 können wir das Wort „endlich“ auch durch „beschränkt“ ersetzen, denn eine Menge natürlicher Zahlen ist endlich genau dann, wenn sie in \mathbf{N} beschränkt ist. Sie ist aber auch endlich genau dann, wenn sie in \mathbf{R} beschränkt ist. Dies ist eine einfache Folgerung aus dem archimedischen Axiom. Ist nämlich eine reelle Zahl a eine Schranke einer Teilmenge M der natürlichen Zahlen, so ist auch die natürliche Zahl $[a + 1]$ eine Schranke von M , und folglich ist M in dem (endlichen) Abschnitt der natürlichen Zahlen von 0 bis $[a + 1]$ enthalten. Umgekehrt besitzt jede endliche Menge natürlicher Zahlen ein Maximum und damit eine Schranke in \mathbf{N} , also auch in \mathbf{R} . Ist die Menge $\{n: |a_n| \geq \varepsilon\}$ endlich, so bezeichnen wir eine beliebige Schranke dieser Menge mit $n_0(\varepsilon)$ bzw. $N(\varepsilon)$, je nachdem, ob sie eine natürliche oder eine reelle Zahl ist. Man deutet hiermit an, daß diese Schranken von der Wahl der positiven Zahl ε abhängen. Dies bedeutet allerdings nicht, daß es sich um ein Funktionssymbol handelt, denn mit jeder Zahl n_0 bzw. N ist auch jede größere Zahl eine Schranke.

Die Ermittlung einer Schranke n_0 aus \mathbf{N} hat den Vorzug, daß die häufig verwendete Sprechweise „von der Stelle n_0 an gilt $a_n < \varepsilon$ “ möglich wird. Andererseits stößt man bei Beispielen und Anwendungen gewöhnlich auf reelle Schranken, und es ist überflüssig, noch eine größere natürliche Zahl als Schranke zu fixieren.

Wir werden im folgenden alle Definitionen und Sätze mit reellen Schranken formulieren, weisen aber darauf hin, daß überall „reelle Zahl $N(\varepsilon)$ “ bzw. „ $N \in \mathbf{R}$ “ durch „natürliche Zahl $n_0(\varepsilon)$ “ bzw. „ $n_0 \in \mathbf{N}$ “ ersetzt werden darf.

Satz 1. Die reelle Zahlenfolge (a_n) ist eine Nullfolge genau dann, wenn es für jede positive reelle Zahl ε eine reelle Zahl $N(\varepsilon)$ gibt derart, daß aus $n \geq N(\varepsilon)$ stets $|a_n| < \varepsilon$ folgt.

In der Tat existiert eine solche Zahl $N(\varepsilon)$ genau dann, wenn $|a_n| < \varepsilon$ für fast alle $n \in \mathbf{N}$ ist.

Die in Satz 1 angegebene charakteristische Eigenschaft von Nullfolgen dient häufig zur Definition dieses Begriffs. Dieses Kriterium für die Nullfolgeeigenschaft kann in Zeichen wie folgt formuliert werden:

$$(a_n) \text{ Nullfolge} \Leftrightarrow \bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbf{R}} \bigwedge_{n \in \mathbf{N}} (n \geq N \Rightarrow |a_n| < \varepsilon). \quad (2)$$

Daß N von ε abhängig ist, kommt hierbei durch die Reihenfolge der Quantifikatoren zum Ausdruck. Eine weitere Charakterisierung einer Nullfolge ergibt sich, wenn wir die Implikation „aus $n \geq N(\varepsilon)$ folgt $|a_n| < \varepsilon$ “ durch ihre Kontraposition ersetzen.

Satz 2. Die reelle Zahlenfolge (a_n) ist eine Nullfolge genau dann, wenn es für jede positive reelle Zahl ε eine reelle Zahl $N(\varepsilon)$ gibt derart, daß aus $|a_n| \geq \varepsilon$ stets $n < N(\varepsilon)$ folgt.

In formalisierter Schreibweise besagt dies:

$$(a_n) \text{ Nullfolge} \Leftrightarrow \bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbf{R}} \bigwedge_{n \in \mathbf{N}} (|a_n| \geq \varepsilon \Rightarrow n < N). \quad (3)$$

Da es hierbei, wie oben bemerkt, auf endlich viele Folgenglieder nicht ankommt, genügt es, die Implikation „aus $|a_n| \geq \varepsilon$ folgt $n < N(\varepsilon)$ “ nur für alle natürlichen Zahlen n mit $n \geq n_0$ zu beweisen. Ferner kann $n < N(\varepsilon)$ durch $n \leq N(\varepsilon)$ ersetzt werden.

Zur Untersuchung der Frage, ob eine Folge (a_n) eine Nullfolge ist, versucht man also aus der Ungleichung $|a_n| \geq \varepsilon$ eine Ungleichung der Form $n \leq N(\varepsilon)$ herzuleiten. Wenn dies gelingt, bedeutet dies, daß die Menge $\{n: |a_n| \geq \varepsilon\}$ beschränkt ist.

Beispiel 3. Gegeben sei die Folge $\left(\frac{n+5}{n^2+n+1}\right)$. Aus $\left|\frac{n+5}{n^2+n+1}\right| \geq \varepsilon$ und $n \geq 1$ folgt $\varepsilon \leq \frac{n+5n}{n^2} = \frac{6}{n}$, $n \leq \frac{6}{\varepsilon}$. Für jedes ε mit $\varepsilon > 0$ besitzt also die reelle Zahl $N(\varepsilon) := \frac{6}{\varepsilon}$ die Eigenschaft, daß aus $\left|\frac{n+5}{n^2+n+1}\right| \geq \varepsilon$ und $n \geq 1$ stets $n \leq N(\varepsilon)$ folgt. Die gegebene Folge ist nach Satz 2 eine Nullfolge.

Beispiel 4. Gegeben sei die Folge (a^n) mit $a \in \mathbb{R}$. Wir unterscheiden drei Fälle.

Fall 1. Es ist $a = 0$. Dann ist (a^n) eine Nullfolge.

Fall 2. Es ist $0 < |a| < 1$. Dann gibt es eine positive reelle Zahl x mit $|a| = \frac{1}{1+x}$, und für $n \geq 1$ folgt aus der Bernoullischen Ungleichung

$$|a^n| = \frac{1}{(1+x)^n} \leq \frac{1}{1+nx} < \frac{1}{nx}.$$

Aus $\varepsilon \leq |a^n|$ und $n \geq 1$ folgt somit $\varepsilon < \frac{1}{nx}$, $n < \frac{1}{\varepsilon x} = N(\varepsilon)$, die Folge ist eine Nullfolge.

Fall 3. Es ist $|a| \geq 1$. Dann ist $|a^n| \geq 1$ für alle natürlichen Zahlen n , d. h., die Menge $M(1)$ ist unendlich, die Folge ist keine Nullfolge.

Es gilt somit der

Satz 3. Die reelle Zahlenfolge (a^n) ist genau dann eine Nullfolge, wenn $|a| < 1$ ist.

Beispiel 5. Gegeben sei die Folge (na^n) mit $0 < |a| < 1$. Mit $|a| = \frac{1}{1+x}$ folgt (für $n \geq 2$)

$$|na^n| = \frac{n}{(1+x)^n} = \frac{n}{1 + \binom{n}{1}x + \binom{n}{2}x^2 + \dots + x^n} < \frac{n}{\binom{n}{2}x^2} = \frac{2}{(n-1)x^2}.$$

Aus $\varepsilon \leq na^n$ und $n \geq 2$ folgt somit $n-1 \leq \frac{2}{\varepsilon x^2}$, $n \leq \frac{2}{\varepsilon x^2} + 1$, die Folge ist eine Nullfolge.

In den obigen Beispielen ist es uns gelungen, für die Folgenglieder a_n eine Abschätzung $|a_n| \leq \frac{c}{n}$ (im letzten Beispiel $|a_n| \leq \frac{c}{n-1}$) herzuleiten. Aus $\varepsilon \leq |a_n|$ folgte dann stets $n \leq \frac{c}{\varepsilon}$ (bzw. $n \leq \frac{c}{\varepsilon} + 1$). Abgesehen von dem konstanten Faktor c haben wir damit die Nullfolgeneigenschaft der gegebenen Folge auf die Nullfolgeneigenschaft der Folge $\left(\frac{1}{n+1}\right)$ zurückgeführt. Wir werden dieses Prinzip der Untersuchung von Folgen auf die Nullfolgeneigenschaft in einem der folgenden Sätze verallgemeinern.

Satz 4. Die reelle Zahlenfolge (a_n) ist genau dann eine Nullfolge, wenn die Folge $(|a_n|)$ eine Nullfolge ist.

Der Beweis folgt unmittelbar aus der Definition der Nullfolge.

Satz 5. Jede Nullfolge ist beschränkt.

Beweis. Die Menge aller reellen Zahlen $|a_n|$ mit $n \in \{m: |a_m| \geq 1\}$ ist für eine Nullfolge endlich und besitzt damit ein Maximum K . Die größte der beiden Zahlen 1 und K ist dann eine Schranke für alle Glieder der Folge.

Satz 6. Die Summe, die Differenz und das Produkt zweier Nullfolgen sind wieder Nullfolgen.

Beweis. Für jedes ε mit $\varepsilon > 0$ sind die Mengen $\left\{n: |a_n| \geq \frac{\varepsilon}{2}\right\}$ und $\left\{n: |b_n| \geq \frac{\varepsilon}{2}\right\}$ endlich und haben folglich eine gemeinsame (von ε abhängige) Schranke N . Für alle n mit $n > N$ gilt somit

$$|a_n \pm b_n| \leq |a_n| + |b_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Nach Satz 2 sind also $(a_n \pm b_n)$ Nullfolgen. Daß das Produkt zweier Nullfolgen wieder eine Nullfolge ist, ergibt sich als Teilaussage von Satz 7. Wir betonen aber ausdrücklich, daß der Quotient zweier Nullfolgen im allgemeinen nicht wieder eine Nullfolge ist. Dies zeigt etwa das Beispiel der Folgen $\left(\frac{1}{n}\right)$, $\left(\frac{1}{n^2}\right)$ mit dem Quotienten (n) .

Satz 7. Ist (a_n) eine Nullfolge und (b_n) eine beschränkte Zahlenfolge, so ist ihr Produkt $(a_n b_n)$ eine Nullfolge.

Beweis. Es sei K eine positive Schranke der Zahlenfolge (b_n) . Dann existiert zu jedem ε mit $\varepsilon > 0$ eine Schranke N der (endlichen) Menge $\left\{n: |a_n| \geq \frac{\varepsilon}{K}\right\}$. Für alle n mit $n > N$ folgt

$$|a_n b_n| = |a_n| \cdot |b_n| \leq K |a_n| < K \cdot \frac{\varepsilon}{K} = \varepsilon,$$

d. h., $(a_n b_n)$ ist eine Nullfolge. Insbesondere ist (ca_n) für jede reelle Zahl c eine Nullfolge, wenn dies für (a_n) der Fall ist.

Für die Untersuchung von Zahlenfolgen auf die Nullfolgeeigenschaft ist das folgende *Vergleichskriterium* nützlich.

Satz 8. Ist die reelle Zahlenfolge (b_n) eine Nullfolge und gibt es zur reellen Zahlenfolge (a_n) reelle Zahlen n_0 und K derart, daß

$$|a_n| \leq K |b_n| \tag{4}$$

für alle n mit $n \geq n_0$ ist, so ist auch (a_n) eine Nullfolge.

Beweis. Für jedes ε mit $\varepsilon > 0$ besitzt die Menge $\left\{n: |b_n| \geq \frac{\varepsilon}{K+1}\right\}$ eine obere Schranke N , und für alle natürlichen Zahlen n mit $n > N$, $n \geq n_0$ gilt

$$|a_n| \leq K |b_n| < K \frac{\varepsilon}{K+1} < \varepsilon,$$

d. h., (a_n) ist eine Nullfolge.

Satz 9. Die stationäre Zahlenfolge (a_n) mit $a_n = c$ für alle n ist genau dann eine Nullfolge, wenn $c = 0$ ist.

Beweis. Ist $c = 0$, so ist (a_n) eine Nullfolge. Ist $c \neq 0$, so ist die Menge $\left\{n: |a_n| \geq \frac{|c|}{2}\right\}$ unendlich, d. h., die Folge (a_n) ist keine Nullfolge.

Satz 10. Für alle positiven reellen Zahlen r gilt: Die reelle Zahlenfolge (a_n) ist genau dann eine Nullfolge, wenn die Zahlenfolge $(|a_n|^r)$ eine Nullfolge ist.

Beweis. Ist (a_n) eine Nullfolge, so ist die Menge

$$\{n: |a_n| \geq \varepsilon^{1/r}\} = \{n: |a_n|^r \geq \varepsilon\}$$

für alle ε mit $\varepsilon > 0$ endlich, d. h., $(|a_n|^r)$ ist eine Nullfolge. Ist umgekehrt $(|a_n|^r)$ eine Nullfolge, so ist die Menge

$$\{n: |a_n|^r \geq \varepsilon^r\} = \{n: |a_n| \geq \varepsilon\}$$

für alle positiven ε endlich, d. h., (a_n) ist eine Nullfolge.

Die Definitionen und Sätze dieses Abschnittes können unverändert auf komplexe Zahlenfolgen und mit Ausnahme von Satz 3 und Satz 10 sogar auf Folgen im euklidischen Raum \mathbb{R}^p übertragen werden, wobei nur $|a_n|$ durch $\|a_n\|$ zu ersetzen ist. Es gilt dann

Satz 11. Eine Folge (a_n) im euklidischen Raum \mathbb{R}^p mit Gliedern $a_n = (\alpha_{1n}, \dots, \alpha_{pn})$ ist eine Nullfolge genau dann, wenn die (reellen) Zahlenfolgen $(\alpha_{jn})_{n \in \mathbb{N}}$ ($j = 1, \dots, p$) Nullfolgen sind.

Beweis. Nach 1.5.2., Satz 3, ist stets

$$|\alpha_{jn}| \leq \|a_n\| \leq |\alpha_{1n}| + \dots + |\alpha_{pn}| \quad (j = 1, \dots, p). \quad (5)$$

Ist die Folge (a_n) , nach Satz 4 also auch die Folge $(\|a_n\|)$ eine Nullfolge, so sind nach Satz 8 alle Folgen $(\alpha_{jn})_{n \in \mathbb{N}}$ und damit $(\alpha_{jn})_{n \in \mathbb{N}}$ ($j = 1, \dots, p$) Nullfolgen. Umgekehrt folgt aus Satz 6 und der zweiten Ungleichung (5), daß mit $(\alpha_{jn})_{n \in \mathbb{N}}$ ($j = 1, \dots, p$) auch $(\|a_n\|)$ und damit (a_n) eine Nullfolge ist. Speziell gilt

Satz 12. Eine komplexe Zahlenfolge (z_n) ist eine Nullfolge genau dann, wenn die Folgen (x_n) und (y_n) der Realteile x_n und Imaginärteile y_n von z_n Nullfolgen sind.

Beispiel 6. Die Folge mit den Gliedern

$$z_n := \left(\frac{1}{2^n}\right)^{3/5} + \frac{i}{n+1}$$

ist eine Nullfolge.

2.1.3. Konvergente Zahlenfolgen

In 2.1.2. haben wir den Begriff der Nullfolge eingeführt. Liegt eine Nullfolge vor, so sagt man auch, die Folge „strebe gegen 0“ oder „konvergiere gegen 0“. Wir wollen nun allgemeiner Folgen (a_n) betrachten, die gegen eine beliebige reelle Zahl a „konvergieren“. Wir führen den Begriff der Konvergenz auf den Begriff der Nullfolge zurück.

Definition 1. Eine reelle Zahlenfolge (a_n) heißt *konvergent*, wenn es eine reelle Zahl a gibt derart, daß die Folge $(a_n - a)^{1)}$ eine Nullfolge ist. Eine Zahlenfolge, die nicht konvergent ist, heißt *divergent*.

Nach 2.1.2.(2) ist (a_n) genau dann konvergent, wenn

$$\bigvee_{a \in \mathbb{R}} \bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbb{N}} \bigwedge_{n \in \mathbb{N}} (n \geq N \Rightarrow |a_n - a| < \varepsilon) \quad (1)$$

gilt. Wir zeigen, daß es zu jeder Zahlenfolge (a_n) höchstens eine reelle Zahl a gibt derart, daß $(a_n - a)$ eine Nullfolge ist. Sind nämlich $(a_n - a)$ und $(a_n - a')$ Nullfolgen, so ist auch ihre Differenz, d. h. die stationäre Folge mit den Gliedern $a' - a$, eine Nullfolge. Nach 2.1.2., Satz 9, ist dies nur für $a' - a = 0$, $a' = a$ möglich.

Zu jeder konvergenten Folge (a_n) gibt es somit genau eine reelle Zahl a , für die $(a_n - a)$ eine Nullfolge ist. Sie heißt der *Grenzwert* oder *Limes* der konvergenten Folge.

Definition 2. Die reelle Zahl a heißt der *Grenzwert* oder der *Limes* der reellen Zahlenfolge (a_n) , in Zeichen

$$a = \lim_{n \rightarrow \infty} a_n,$$

wenn $(a_n - a)$ eine Nullfolge ist.

Man sagt dann auch, „die Zahlenfolge (a_n) konvergiert gegen a “, und schreibt

$$a_n \rightarrow a \quad \text{für } n \rightarrow \infty \quad (\text{gelesen: } a_n \text{ gegen } a \text{ für } n \text{ gegen } \infty).$$

¹⁾ Wegen 2.1.2., Satz 4, können wir $a_n - a$ auch durch $|a_n - a|$ ersetzen.

Mitunter wird die Konvergenz der Folge (a_n) mit dem Grenzwert a auch kurz durch die Sprechweise „ a_n strebt gegen a “ gekennzeichnet.

Auf Grund der entsprechenden Eigenschaft für Nullfolgen ergibt sich aus Definition 2, daß jede Teilfolge einer konvergenten Folge wieder konvergent ist und denselben Grenzwert besitzt.

Beispiel 1. Gegeben sei die Zahlenfolge $\left(\frac{n^2 + 1}{2n^2 + 3}\right)$. Schreiben wir das Folgenglied in der Form

$$\frac{1 + \frac{1}{n^2}}{2 + \frac{3}{n^2}} \quad (n > 0)$$

und beachten, daß $\frac{1}{n^2}$ und $\frac{3}{n^2}$ Glieder von Nullfolgen sind, so gelangen wir zu der Vermutung, daß die Folge den Grenzwert $\frac{1}{2}$ besitzt. In der Tat ist

$$\left| \frac{n^2 + 1}{2n^2 + 3} - \frac{1}{2} \right| = \left| \frac{2(n^2 + 1) - (2n^2 + 3)}{2(2n^2 + 3)} \right| = \left| -\frac{1}{2(2n^2 + 3)} \right| < \frac{1}{n^2},$$

und die Folge mit den rechts stehenden Gliedern ist eine Nullfolge. Somit gilt

$$\lim_{n \rightarrow \infty} \frac{n^2 + 1}{2n^2 + 3} = \frac{1}{2}.$$

Beispiel 2. Gegeben sei die Zahlenfolge $(\sqrt{n^2 + 6n + 1} - n)$. Es ist nicht ohne weiteres zu erkennen, daß diese Zahlenfolge einen Grenzwert besitzt. Die Umformungen

$$\begin{aligned} \sqrt{n^2 + 6n + 1} - n &= \frac{(\sqrt{n^2 + 6n + 1} - n)(\sqrt{n^2 + 6n + 1} + n)}{\sqrt{n^2 + 6n + 1} + n} \\ &= \frac{n^2 + 6n + 1 - n^2}{\sqrt{n^2 + 6n + 1} + n} = \frac{6 + \frac{1}{n}}{\sqrt{1 + \frac{6}{n} + \frac{1}{n^2}} + 1} \quad (n \geq 1) \end{aligned}$$

lassen vermuten, daß als Grenzwert nur $\frac{6}{\sqrt{1} + 1} = 3$ in Frage kommt. In der Tat ist

$$\begin{aligned} |\sqrt{n^2 + 6n + 1} - n - 3| &= \left| \frac{(\sqrt{n^2 + 6n + 1} - (n + 3))(\sqrt{n^2 + 6n + 1} + n + 3)}{\sqrt{n^2 + 6n + 1} + n + 3} \right| \\ &= \left| \frac{n^2 + 6n + 1 - (n^2 + 6n + 9)}{\sqrt{n^2 + 6n + 1} + n + 3} \right| \leq \frac{8}{n} \quad (n \geq 1), \end{aligned}$$

und die rechte Seite strebt gegen 0 für $n \rightarrow \infty$. Auf Grund von 2.1.2., Satz 8, ist unsere Vermutung bestätigt.

Beispiel 3. Wir zeigen, daß

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n q^k = \frac{1}{1-q} \quad (2)$$

für alle reellen (oder komplexen) Zahlen q mit $|q| < 1$ gilt. Nach 1.1.1.(7) ist

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

Es folgt

$$\left| \sum_{k=0}^n q^k - \frac{1}{1-q} \right| = \left| \frac{1 - q^{n+1}}{1-q} - \frac{1}{1-q} \right| = \frac{|q|^{n+1}}{|1-q|},$$

und die rechte Seite strebt nach 2.1.2., Satz 3, für $n \rightarrow \infty$ gegen 0.

Beispiel 4. Wir betrachten die Zahlenfolge $(\sqrt[n]{a})_{n \geq 1}$, wobei a eine positive reelle Zahl ist. Aus 1.4.1.(9) folgt mit $Q = \max \left\{ a, \frac{1}{a} \right\}$, $R = 1$, daß

$$|a^{1/n} - 1| \leq Q^n \frac{1}{n} |a - 1|$$

ist. Die rechte Seite strebt gegen 0 für $n \rightarrow \infty$, und daher ist

$$\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1 \quad (a > 0). \quad (3)$$

Beispiel 5. Wir zeigen, daß

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1 \quad (4)$$

ist. Hierzu setzen wir $a_n := \sqrt[n]{n} - 1$ und zeigen, daß (a_n) eine Nullfolge ist. Für $n \geq 2$ folgt aus der binomischen Formel

$$n = (a_n + 1)^n > \binom{n}{2} a_n^2 = \frac{n(n-1)}{2} a_n^2,$$

$$|a_n|^2 = a_n^2 < \frac{2}{n-1}.$$

Wegen 2.1.2., Satz 8 und Satz 10, folgt die Behauptung.

Einige der Sätze über Nullfolgen können sinngemäß auf konvergente Zahlenfolgen übertragen werden.

Satz 1. *Jede konvergente reelle Zahlenfolge ist beschränkt.*

Beweis. Gilt $a_n \rightarrow a$ für $n \rightarrow \infty$, so ist die Menge der reellen Zahlen a_n mit $n \in \{m: |a_m - a| \geq 1\}$ endlich und besitzt daher eine Schranke K . Für alle n mit $n \notin \{m: |a_m - a| \geq 1\}$ gilt

$$|a_n| \leq |a_n - a| + |a| < 1 + |a|.$$

Das Maximum der beiden Zahlen K und $|a| + 1$ ist daher eine Schranke für alle Glieder der Folge.

Die offenbar divergente Folge $(1, -1, 1, -1, \dots, (-1)^n, \dots)$ ist beschränkt. Dies zeigt, daß die Umkehrung von Satz 1 nicht gilt.

Satz 2. *Aus $a_n \rightarrow a$ und $b_n \rightarrow b$ für $n \rightarrow \infty$ folgt*

$$a_n \pm b_n \rightarrow a \pm b, \quad (5)$$

$$a_n b_n \rightarrow ab, \quad (6)$$

$$|a_n| \rightarrow |a| \quad (7)$$

für $n \rightarrow \infty$. Ist ferner $b \neq 0$, so gibt es eine natürliche Zahl n_0 mit $b_n \neq 0$ für $n \geq n_0$, und es gilt

$$\frac{a_n}{b_n} \rightarrow \frac{a}{b} \quad (n \geq n_0) \quad (8)$$

für $n \rightarrow \infty$.

Beweis. Die Folge $((a_n \pm b_n) - (a \pm b)) = ((a_n - a) \pm (b_n - b))$ ist als Summe bzw. Differenz zweier Nullfolgen wieder eine Nullfolge, und damit ist (5) bewiesen. Die Folgen $(a_n(b_n - b))$ und $(b(a_n - a))$ sind nach 2.1.2., Satz 7, Nullfolgen, und daher ist auch ihre Summe $(a_n b_n - ab)$ eine Nullfolge, womit (6) bewiesen ist. Wegen $||a_n| - |a|| \leq |a_n - a|$ ist $(|a_n| - |a|)$ eine Nullfolge, und (7) ist bewiesen.

Ist $b \neq 0$, so gibt es ein n_0 mit $|b_n - b| < \frac{|b|}{2}$ für $n \geq n_0$. Dann ist $|b| - |b_n| \leq |b_n - b| < \frac{|b|}{2}$, $|b_n| > \frac{b}{2} > 0$ für $n \geq n_0$. Es folgt

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \left| \frac{b - b_n}{b_n b} \right| \leq |b_n - b| \cdot \frac{2}{b^2} \quad (n \geq n_0)$$

und damit $\frac{1}{b_n} \rightarrow \frac{1}{b}$ für $n \rightarrow \infty$. Durch Anwendung von (6) erhalten wir hieraus (8).

Sind die Folgen (a_n) und (b_n) konvergent, so gelten nach Satz 2 die folgenden Grenzwertsätze, wobei im letzten Fall noch vorausgesetzt werden muß, daß (b_n) keine Nullfolge ist:

$$\lim_{n \rightarrow \infty} (a_n \pm b_n) = \lim_{n \rightarrow \infty} a_n \pm \lim_{n \rightarrow \infty} b_n, \quad (5')$$

$$\lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} a_n \lim_{n \rightarrow \infty} b_n, \quad (6')$$

$$\lim_{n \rightarrow \infty} |a_n| = |\lim_{n \rightarrow \infty} a_n|, \quad (7')$$

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}. \quad (8')$$

Mit Hilfe dieser Grenzwertsätze kann die Berechnung von Grenzwerten häufig auf bereits bekannte Grenzwerte zurückgeführt werden.

Beispiel 6. Wir berechnen den Grenzwert der Zahlenfolgen (a_n) mit

$$a_n = \frac{n^3 + 3 \sqrt[n]{n^3}}{3n^3 - 2n + 5}.$$

Offenbar ist

$$a_n = \frac{1 + \frac{3}{n^3} (\sqrt[n]{n})^3}{3 - \frac{2}{n} + \frac{5}{n^3}}.$$

Auf Grund der Grenzwertsätze erhalten wir, wenn wir (4) beachten, für die Zählerfolge den Grenzwert 1 und für die Nennerfolge den Grenzwert 3, d. h., es ist

$$\lim_{n \rightarrow \infty} \frac{n^3 + 3 \sqrt[n]{n^3}}{3n^3 - 2n + 5} = \frac{1}{3}.$$

Satz 3. Gilt für zwei reelle Zahlenfolgen $a_n \rightarrow a$ und $b_n \rightarrow b$ für $n \rightarrow \infty$ und gibt es eine natürliche Zahl n_0 mit $a_n \leq b_n$ für alle natürlichen Zahlen n mit $n \geq n_0$, so ist $a \leq b$.

Beweis. Angenommen, es ist $a > b$. Für alle n mit $n \geq n_0$ ist dann

$$a - b \leq a - b + (b_n - a_n) = (a - a_n) - (b - b_n) \leq |a - a_n| + |b - b_n|.$$

Wir können n so groß wählen, daß $|a - a_n| < \frac{a-b}{2}$ und $|b - b_n| < \frac{a-b}{2}$ ist. Dann folgt $a - b < \frac{a-b}{2} + \frac{a-b}{2} = a - b$, womit wir zu einem Widerspruch gelangt sind. Somit ist $a \leq b$.

Wir betonen ausdrücklich, daß aus der verschärften Voraussetzung $a_n < b_n$ für alle n mit $n \geq n_0$ nur auf $a \leq b$ und nicht auf $a < b$ geschlossen werden darf, wie das Beispiel $a_n := \frac{n}{n+1}$, $b_n := \frac{n+1}{n}$ zeigt.

Satz 4. Gilt $a_n \rightarrow a$ und $b_n \rightarrow a$ für $n \rightarrow \infty$ und gibt es eine natürliche Zahl n_0 mit $a_n \leq c_n \leq b_n$ für $n \geq n_0$, so gilt auch $c_n \rightarrow a$ für $n \rightarrow \infty$.

Beweis. Für $n \geq n_0$ ist

$$|c_n - a| \leq |c_n - a_n| + |a_n - a| \leq |b_n - a_n| + |a_n - a|,$$

und die Behauptung folgt aus dem Vergleichskriterium für Nullfolgen.

Satz 5. Ist (a_n) eine reelle konvergente Zahlenfolge mit nichtnegativen Gliedern, so gilt für positive rationale Zahlen r stets

$$\lim_{n \rightarrow \infty} a_n^r = \left(\lim_{n \rightarrow \infty} a_n \right)^r. \quad (9)$$

Insbesondere ist

$$\lim_{n \rightarrow \infty} \sqrt[m]{a_n} = \sqrt[m]{\lim_{n \rightarrow \infty} a_n} \quad (m = 1, 2, \dots). \quad (10)$$

Beweis. Es sei $a = \lim_{n \rightarrow \infty} a_n$. Für $a = 0$ gilt die Behauptung nach 2.1.2., Satz 10. Ist $a \neq 0$, dann gibt es eine natürliche Zahl n_0 derart, daß für alle natürlichen Zahlen n mit $n \geq n_0$ stets $\frac{1}{2} \leq \frac{a_n}{a} < 2$ ist. Aus 1.4.1.(9) folgt, wenn wir $Q = 2$, $R = r$ setzen,

$$|a_n^r - a^r| = a^r \left| \left(\frac{a_n}{a} \right)^r - 1 \right| \leq a^r 2^{r+1} r \left| \frac{a_n}{a} - 1 \right| = a^{r-1} 2^{r+1} r |a_n - a|.$$

Da $|a_n - a|$ nach Voraussetzung das Glied einer Nullfolge ist, gilt dies auch für $|a_n^r - a^r|$. Damit ist der Satz bewiesen.

Definition 3. Eine reelle Zahlenfolge (a_n) heißt *bestimmt divergent* mit dem uneigentlichen Grenzwert ∞ bzw. $-\infty$, in Zeichen

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} a_n = -\infty,$$

wenn es zu jeder positiven Zahl ω^1) eine reelle Zahl $N(\omega)$ gibt derart, daß aus $n \geq N(\omega)$ stets $a_n > \omega$ bzw. $a_n < -\omega$ folgt.

Für eine Zahlenfolge mit positiven Gliedern gilt offensichtlich $\lim_{n \rightarrow \infty} a_n = \infty$ genau dann, wenn $\left(\frac{1}{a_n}\right)$ eine Nullfolge ist.

Beispiel 7. Wir zeigen, daß die Zahlenfolge $(a + bn - n^2)$ für alle reellen Zahlen a, b bestimmt divergent mit dem uneigentlichen Grenzwert $-\infty$ ist. Aus $a + bn - n^2 \geq -\omega$ folgt nämlich $n^2 - bn \leq \omega + a$, $\left(n - \frac{b}{2}\right)^2 \leq \omega + a + \frac{b^2}{4}$, $n \leq \frac{b}{2} + \sqrt{\omega + a + \frac{b^2}{4}}$, und durch Kontraposition ergibt sich die Behauptung mit $N(\omega) = \frac{b}{2} + \sqrt{\omega + a + \frac{b^2}{4}}$.

Die Definitionen 1 bzw. 2 für die Konvergenz bzw. den Grenzwert können unverändert auf komplexe Folgen und Folgen im Raum \mathbb{R}^p übertragen werden. Die Sätze 1, 2 bleiben dann weiterhin für komplexe Zahlenfolgen und mit Ausnahme von (6), (8) auch für Folgen im Raum \mathbb{R}^p gültig.

Satz 6. Die Folge (a_n) mit $a_n = (\alpha_{1n}, \alpha_{2n}, \dots, \alpha_{pn}) \in \mathbb{R}^p$ ist konvergent genau dann, wenn alle Koordinatenfolgen (α_{jn}) ($j = 1, 2, \dots, p$) konvergent sind, und die Koordinaten α_j des Grenzwertes der Folge (a_n) sind die Grenzwerte der entsprechenden Koordinatenfolgen.

Beweis. Die Folge (a_n) ist konvergent, und $a = (\alpha_1, \dots, \alpha_p)$ ist ihr Grenzwert genau dann, wenn die Folge $(a_n - a)$ mit den Gliedern $a_n - a = (\alpha_{1n} - \alpha_1, \dots, \alpha_{pn} - \alpha_p)$ konvergent ist. Dies ist nach 2.1.2., Satz 11, genau dann der Fall, wenn alle Folgen $(\alpha_{jn} - \alpha_j)_{n \in \mathbb{N}}$ Nullfolgen sind, d. h., wenn

$$\alpha_j = \lim_{n \rightarrow \infty} \alpha_{jn} \quad (j = 1, \dots, p)$$

ist. Damit ist Satz 6 bewiesen.

Satz 7. Für konvergente komplexe Zahlenfolgen (z_n) mit $z_n = x_n + iy_n$ ($x_n, y_n \in \mathbb{R}$) gilt

$$\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} x_n + i \lim_{n \rightarrow \infty} y_n, \quad (11)$$

$$\overline{\lim_{n \rightarrow \infty} z_n} = \lim_{n \rightarrow \infty} \overline{z_n}. \quad (12)$$

¹⁾ Wir wählen hier die Bezeichnung ω statt ε , um anzudeuten, daß ω beliebig groß gewählt werden kann. Bei der Betrachtung von Nullfolgen kam es dagegen darauf an, daß für ε eine beliebig kleine positive Zahl gewählt werden kann.

Beweis. Es existieren die Grenzwerte

$$z := \lim_{n \rightarrow \infty} z_n, \quad x := \lim_{n \rightarrow \infty} x_n, \quad y := \lim_{n \rightarrow \infty} y_n,$$

und aus Satz 2 und Satz 6 folgt

$$z = \lim_{n \rightarrow \infty} (x_n + iy_n) = \lim_{n \rightarrow \infty} x_n + \lim_{n \rightarrow \infty} iy_n = \lim_{n \rightarrow \infty} x_n + i \lim_{n \rightarrow \infty} y_n = x + iy,$$

$$\bar{z} = x - iy = \lim_{n \rightarrow \infty} (x_n - iy_n) = \lim_{n \rightarrow \infty} \bar{z}_n,$$

womit Satz 7 bewiesen ist.

Der Begriff der Konvergenz kann sogar in beliebigen metrischen Räumen X definiert werden.

Definition 4. Eine Folge (a_n) in einem metrischen Raum X heißt *konvergent*, wenn es in X ein Element a gibt derart, daß die Zahlenfolge $(\varrho(a_n, a))$ eine Nullfolge ist.

Sind die Folgen $(\varrho(a_n, a))$ und $(\varrho(a_n, b))$ Nullfolgen, so ist auch die Folge $(\varrho(a_n, a) + \varrho(a_n, b))$ eine Nullfolge. Wegen

$$0 \leq \varrho(a, b) \leq \varrho(a_n, a) + \varrho(a_n, b)$$

ist dann auch die stationäre Folge mit den Gliedern $\varrho(a, b)$ eine Nullfolge, d. h., es ist $\varrho(a, b) = 0$ und damit $a = b$. Es gibt somit höchstens ein a mit $a \in X$, das die in der Definition 1 geforderte Eigenschaft besitzt. Dies rechtfertigt die

Definition 5. Das Element a heißt der *Grenzwert* oder der *Limes* der Folge (a_n) , in Zeichen $\lim_{n \rightarrow \infty} a_n = a$, wenn $(\varrho(a_n, a))$ eine Nullfolge ist.

Zum Schluß dieses Abschnitts soll noch ein nützliches Kriterium für die Ermittlung der kleinsten oberen bzw. größten unteren Schranke einer reellen Zahlenmenge angegeben werden.

Satz 8. Für eine nichtleere Menge M reeller Zahlen gilt

$$a = \sup M \quad \text{bzw.} \quad a = \inf M$$

genau dann, wenn folgende Bedingungen erfüllt sind:

- a ist eine obere bzw. untere Schranke von M .
- Es gibt eine Folge von Elementen $a_n \in M$ mit $a_n \rightarrow a$.

Beweis. Es genügt, den erstgenannten Fall zu beweisen. Die Bedingungen a), b) seien erfüllt. Ist dann b eine beliebige obere Schranke von M , so gilt $a_n \leq b$, also auch

$$a = \lim_{n \rightarrow \infty} a_n \leq b.$$

Folglich ist a die kleinste obere Schranke von M .

Ist umgekehrt $a = \sup M$, so unterscheiden wir zwei Fälle.

Fall 1. Es ist $a < \infty$, d. h., M ist nach oben beschränkt. Für kein $n \in \mathbf{N}^*$ ist dann $a - \frac{1}{n}$ eine obere Schranke von M , und folglich existieren Elemente $a_n \in M$ mit $a - \frac{1}{n} < a_n \leq a$. Nach Satz 4 gilt dann $a_n \rightarrow a$.

Fall 2. Es ist $a = \infty$. Dann ist M nicht nach oben beschränkt, und zu jedem $n \in \mathbf{N}$ gibt es ein Element $a_n \in M$ mit $a_n > n$. Somit gilt $a_n \rightarrow a = \infty$.

Die Idee, daß man beispielsweise zur Flächenberechnung des Kreises ein „Polygon“ mit „unendlich vielen“ Seiten ausnutzen könnte, trat in der frühen griechischen Mathematik auf. In diesem Sinne ist der Begriff der „Grenze“ geometrischen Ursprungs. Implizit tritt er in völlig korrekter Weise in der griechisch-hellenistischen Mathematik in der Exhaustionsmethode (das Wort stammt erst aus dem 17. Jahrhundert) auf. Diese Vorstellungen wurden im 16. Jahrhundert erneut im Zusammenhang mit praktischen Fragen der Flächen- und Volumenberechnung aufgegriffen.

Eine wissenschaftliche Darstellung fand diese Methode 1635 bei BONAVENTURA CAVALIERI (1598–1647). Im Jahre 1655 gab JOHN WALLIS (1616–1703) dann eine arithmetische Fassung des Begriffs der Grenze, indem er die Grenze a einer Zahlenfolge (a_n) als eine Zahl einführt, für die $a - a_n$ bei wachsendem n beliebig klein wird. Auch LEIBNIZ versuchte, aber durch Stetigkeitsbetrachtungen, 1687 den Begriff des Grenzwertes zu erfassen, ohne völlig das Problem meistern zu können.

NEWTON hatte in seinen Werken zur Infinitesimalrechnung den Begriff der Grenze auch nur undeutlich ausgesprochen. Er verwandte solche Begriffe wie „erstes und letztes Verhältnis“ einer „Fluxion“.

Erst JEAN BAPTISTE LE ROND D'ALEMBERT (1717–1783) ersetzte diese Begriffe in einigen Artikeln der „Encyclopédie“ durch die Idee des Grenzwerts. Bei ihm ist eine Größe der Grenzwert einer anderen, wenn sich die zweite der ersten beliebig nähert.

Eine ebenfalls weitgehend einwandfreie Darstellung des Grenzwertbegriffes gab 1786 SIMON L'HUILIER (1750–1840) in seiner „Exposition élémentaire des principes des calculs supérieurs“.

D'ALEMBERTS und L'HUIERS neue Begründung der Infinitesimalrechnung fand jedoch durchaus nicht den Beifall der Zeitgenossen. So lehnte LAGRANGE die Theorie der Grenzwerte ab und ersetzte sie durch eine neue Methode, die auf der (angenommenen) Existenz der Taylorreihe für eine „beliebige“ Funktion beruhte.

Der an sich unhaltbare Zustand, sich bei Grenzwertbetrachtungen auf die Anschauung zu stützen, hielt weiter an. Erst mit AUGUSTIN-LOUIS CAUCHYS (1789–1857) „Résumé des leçons données à l'école polytechnique sur le calcul infinitesimal“ von 1823, in dem sich eine einwandfreie Definition und ein Existenzbeweis für das bestimmte Integral finden, sowie durch die Arbeiten BERNARD BOLZANOS (1791–1848), wurde dieser Zustand endgültig beseitigt.

CAUCHY erkannte die Notwendigkeit, die Existenz einer Flächenmaßzahl arithmetisch (d. h. über Zahlenfolgen) zu beweisen, und führte den Beweis in der Hauptsache auch einwandfrei durch. Er stützt sich wesentlich auf das notwendige grundlegende „Konvergenzkriterium von CAUCHY“ (vgl. 2.1.7.), von dem die „Notwendigkeit“ seinerzeit auch leicht bewiesen wurde, daß es „hinreichend“ ist, aber erst nach der Schaffung einer strengen Theorie der Irrationalzahlen.

2.1.4. Monotone Zahlenfolgen, Intervallschachtelungen

Die Untersuchung von reellen monotonen Zahlenfolgen auf Konvergenz gestaltet sich sehr einfach, denn es gilt der

Satz 1. *Eine monotone Zahlenfolge ist genau dann konvergent, wenn sie beschränkt ist. Ist sie nicht beschränkt, so ist sie bestimmt divergent.*

Beweis. Die Zahlenfolge (a_n) sei etwa monoton wachsend. Dann ist das erste Glied der Folge eine untere Schranke für den Wertebereich $\{a_n: n \in \mathbf{N}\}$ der Folge. Wir unterscheiden zwei Fälle.

Fall 1. Die Folge ist nach oben beschränkt. Mit a bezeichnen wir die kleinste obere Schranke des Wertebereichs der Folge. Ist $\varepsilon > 0$, so ist $a - \varepsilon$ keine obere Schranke des Wertebereichs, d. h., es gibt ein n_0 mit $n_0 \in \mathbf{N}$ und $a_{n_0} > a - \varepsilon$. Für alle n mit $n \geq n_0$ gilt dann

$$|a_n - a| = a - a_n = a - a_{n_0} - (a_n - a_{n_0}) \leq a - a_{n_0} < \varepsilon,$$

d. h., $(a_n - a)$ ist eine Nullfolge. Somit ist

$$\lim_{n \rightarrow \infty} a_n = \sup \{a_n: n \in \mathbf{N}\} \quad (a_{n+1} \geq a_n). \quad (1)$$

Fall 2. Die Folge ist nicht nach oben beschränkt. Dann gibt es zu jeder positiven reellen Zahl ω eine natürliche Zahl n_0 mit $a_{n_0} > \omega$, und wegen der Monotonie gilt dann auch $a_n > \omega$ für alle n mit $n \geq n_0$. Somit ist $\lim_{n \rightarrow \infty} a_n = \infty$, d. h., die Beziehung

(1) ist auch für nicht beschränkte monoton wachsende Zahlenfolgen erfüllt (vgl. 1.1.3.).

Für monoton fallende Zahlenfolgen gilt analog

$$\lim_{n \rightarrow \infty} a_n = \inf \{a_n: n \in \mathbf{N}\} \quad (a_{n+1} \leq a_n). \quad (2)$$

Wir bemerken, daß eine beschränkte Zahlenfolge auch dann konvergent ist, wenn sie, abgesehen von endlich vielen Gliedern, monoton ist.

Beispiel 1. Die Zahlenfolge mit den Gliedern $a_n = n(\sqrt[n]{x} - 1)$ ist nach Beispiel 1 in 2.1.1. streng monoton fallend, und 0 ist für $x > 1$ eine untere Schranke dieser Folge. Daher existiert

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) \quad (x > 1). \quad (3)$$

Für $0 < x < 1$ folgt, wenn wir 2.1.3.(3) beachten,

$$\begin{aligned} \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) &= -\lim_{n \rightarrow \infty} \sqrt[n]{x} n \left(\sqrt[n]{\frac{1}{x}} - 1 \right) = -\lim_{n \rightarrow \infty} \sqrt[n]{x} \lim_{n \rightarrow \infty} n \left(\sqrt[n]{\frac{1}{x}} - 1 \right) \\ &= -\lim_{n \rightarrow \infty} n \left(\sqrt[n]{\frac{1}{x}} - 1 \right), \end{aligned}$$

und dieser Grenzwert existiert, weil $x^{-1} > 1$ ist. Für $x = 1$ ist die Zahlenfolge stationär, und ihr Grenzwert ist 0. Wir werden diese Folgen im nächsten Abschnitt weiter untersuchen.

Wenn auch die Existenz des Grenzwertes einer monotonen Zahlenfolge häufig sehr einfach bewiesen werden kann — nämlich durch Angabe einer Schranke für ihre Glieder —, kann die Berechnung des Grenzwertes bzw. seine näherungsweise Berechnung mit Angabe einer Fehlerschranke mit großen Schwierigkeiten verbunden sein. Das letztgenannte Problem ist wesentlich leichter zu lösen, wenn eine sogenannte Intervallschachtelung vorliegt.

Definition 1. Zwei reelle Zahlenfolgen (a_n) , (b_n) bilden eine *Intervallschachtelung*, wenn folgende Bedingungen erfüllt sind:

- a) Die Folge (a_n) ist monoton wachsend.
- b) Die Folge (b_n) ist monoton fallend.
- c) Für alle natürlichen Zahlen n ist $a_n \leq b_n$.
- d) Die Folge $(b_n - a_n)$ ist eine Nullfolge.

Wir bezeichnen eine Intervallschachtelung auch mit $(a_n | b_n)$.

Satz 2. Ist $(a_n | b_n)$ eine Intervallschachtelung, so gibt es genau eine reelle Zahl a mit $a_n \leq a \leq b_n$ ($n \in \mathbb{N}$), und zwar ist

$$a = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n. \quad (4)$$

Beweis. Für alle natürlichen Zahlen n, k gilt $a_n \leq a_{n+k} \leq b_{n+k} \leq b_n$, und hieraus folgt $a_m \leq b_n$ für alle natürlichen Zahlen m, n . Daher ist b_n für alle n eine obere Schranke der Folge (a_n) , und umgekehrt ist a_n für alle n eine untere Schranke der Folge (b_n) .

Nach Satz 1 sind beide Folgen konvergent. Ihre Grenzwerte seien a bzw. b . Wegen $a - b = (a - a_n) + (b_n - b) + (a_n - b_n)$ ist die stationäre Folge mit dem Glied $a - b$ als Summe dreier Nullfolgen eine Nullfolge und somit $a - b = 0$, d. h., (4) ist erfüllt. Wegen 2.1.3., Satz 4, kann es keine von a verschiedene Zahl a mit $a_n \leq a \leq b_n$ für alle $n \in \mathbb{N}$ geben.

Beispiel 2. Wir zeigen, daß $(a_n | b_n)$ mit

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

eine Intervallschachtelung ist.

Nach 2.1.1., Beispiel 2, sind die Folgen mit den Gliedern

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

streng monoton wachsend. Wegen

$$\left(1 + \frac{1}{n}\right) \left(1 - \frac{1}{n+1}\right) = 1$$

ist $b_n c_{n+1} = 1$. Die Folge mit den Gliedern b_n ist somit als Folge der Kehrwerte einer streng monoton wachsenden Folge mit positiven Gliedern streng monoton fallend. Damit sind die Eigenschaften a), b) bewiesen. Es ist

$$b_n - a_n = \left(1 + \frac{1}{n}\right)^n \cdot \frac{1}{n} > 0, \quad (5)$$

und die Bedingung c) für eine Intervallschachtelung ist erfüllt. Aus $a_n < b_n \leq b_1$ folgt, daß die Folge (a_n) beschränkt ist. Nach (5) ist $b_n - a_n = a_n \cdot \frac{1}{n}$, und wegen 2.1.2., Satz 7, ist $(b_n - a_n)$ eine Nullfolge. Damit sind alle Eigenschaften einer Intervallschachtelung nachgewiesen.

Das Prinzip der Intervallschachtelung gehört zu den frühen Methoden der Mathematik. Bereits Mathematiker der Antike verwenden es insbesondere bei Quadraturen, wobei es hier meist als heuristisches Verfahren ohne strenge Fassung auftritt. Völlig einwandfrei findet es sich erst bei BOLZANO und KARL WEIERSTRASS (1815–1897) als eine Grundmöglichkeit für einen korrekten Aufbau der Lehre von den reellen Zahlen.

Der durch die Intervallschachtelung im Beispiel 2 bestimmte Grenzwert ist eine für die Analysis außerordentlich wichtige reelle Zahl. Sie wird nach EULER mit dem Buchstaben e bezeichnet. Es gilt also

$$e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^{n+1}.$$

Für die ersten Glieder der Intervallschachtelung erhalten wir

$$\begin{aligned} \left(1 + \frac{1}{1}\right)^1 &= 2 & \left(1 + \frac{1}{1}\right)^2 &= 4 \\ \left(1 + \frac{1}{2}\right)^2 &= \frac{9}{4} = 2,25 & \left(1 + \frac{1}{2}\right)^3 &= \frac{27}{8} = 3,37 \dots \\ \left(1 + \frac{1}{3}\right)^3 &= \frac{64}{27} = 2,37 \dots & \left(1 + \frac{1}{3}\right)^4 &= \frac{256}{81} = 3,16 \dots \\ \left(1 + \frac{1}{4}\right)^4 &= \frac{625}{256} = 2,44 \dots & \left(1 + \frac{1}{4}\right)^5 &= \frac{3125}{1024} = 3,05 \dots \end{aligned}$$

Die Dezimalbruchentwicklung der Eulerschen Zahl e beginnt mit

$$e = 2,71828\ 1828\ 4590 \dots$$

Die Logarithmusfunktion zur Basis e wird mit

$$\ln x := \log_e x$$

bezeichnet und heißt aus einem noch zu erläuternden Grunde der *natürliche Logarithmus*.

Beispiel 3. Es sei $-1 \leq x \leq 1$. Wir betrachten die induktiv definierte Zahlenfolge

$$x_0 := x, \quad x_{n+1} := \sqrt{\frac{1+x_n}{2}}. \quad (6)$$

Für $-1 \leq x \leq 1$ ist stets $0 \leq \frac{1+x}{2} \leq \sqrt{\frac{1+x}{2}} \leq 1$, und folglich ist die Definition (6) sinnvoll. Ferner ist stets

$$x_n \leq \frac{1+x_n}{2} = x_{n+1}^2 \leq x_{n+1} \leq 1. \quad (7)$$

Die Folge (x_n) ist also konvergent, und die Folge (x_n^2) besitzt wegen (7) denselben Grenzwert c . Somit ist $c = c^2$. Für $n \geq 2$ ist stets $x_n > 0$, und folglich ist $c = c^2 > 0$, d. h., es ist

$$\lim_{n \rightarrow \infty} x_n = 1. \quad (8)$$

Wir zeigen, daß durch

$$a_n := 2^n \sqrt{1-x_n^2}, \quad b_n := 2^n \frac{\sqrt{1-x_n^2}}{x_n} = \frac{a_n}{x_n} \quad (9)$$

eine Intervallschachtelung definiert wird. Damit der Nenner x_n positiv wird, haben wir im Fall $x = -1$ bzw. $-1 < x \leq 0$ den Anfangsindex $n = 2$ bzw. $n = 1$ zu wählen. Wegen

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

ist

$$a_{n+1} \cdot x_{n+1} = 2^{n+1} \sqrt{\frac{1-x_n}{2}} \sqrt{\frac{1+x_n}{2}} = 2^n \sqrt{1-x_n^2} = a_n.$$

Wegen $0 \leq x_{n+1}$ ist also $a_n \leq a_{n+1}$.

Andererseits ist

$$a_{n+1}x_n \leq a_{n+1} \cdot \frac{1+x_n}{2} = a_{n+1}x_{n+1}^2 = a_n x_{n+1}.$$

Für $n \geq 2$ ist also

$$b_{n+1} = \frac{a_{n+1}}{x_{n+1}} \leq \frac{a_n}{x_n} = b_n.$$

Wegen $b_n - a_n = b_n(1 - x_n) \geq 0$ ist $a_n \leq b_n$. Schließlich ist

$$0 \leq b_n - a_n = b_n(1 - x_n) \leq b_2(1 - x_n),$$

und wegen (8) ist $(b_n - a_n)$ eine Nullfolge. Der Vergleich mit 1.6.2., Satz 1, zeigt, daß der gemeinsame Grenzwert der Folgen (a_n) , (b_n) der Arkuskosinus von x ist. Wir haben also

$$\arccos x = \lim_{n \rightarrow \infty} 2^n \sqrt{1 - x_n^2} = \lim_{n \rightarrow \infty} 2^n \frac{\sqrt{1 - x_n^2}}{x_n}, \quad (10)$$

wobei die Folge (x_n) durch (6) definiert ist. Man kann die Arkuskosinusfunktion auch durch (6), (10) definieren und Satz 2 in 1.6.2. (mit $\sigma(x) = \arccos x$) mit Hilfe konvergenter Zahlenfolgen beweisen. Dies ist eine weitere Möglichkeit, die trigonometrischen Funktionen mit elementaren Hilfsmitteln einzuführen. Wir gehen hierauf nicht näher ein.

2.1.5. Darstellung des natürlichen Logarithmus und der speziellen Exponentialfunktion durch monotone Zahlenfolgen

In 2.1.4., Beispiel 1, haben wir gezeigt, daß die Folge mit den Gliedern

$$a_n = n(\sqrt[n]{x} - 1) \quad (1)$$

für alle positiven reellen Zahlen x konvergent ist. Wir können daher eine Funktion L durch

$$L(x) := \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) \quad (x > 0) \quad (2)$$

definieren. Ersetzen wir in 1.4.1.(4) die Zahlen y, r durch $1, \frac{1}{n}$ ($n \geq 2$), so erhalten wir

$$\begin{aligned} \frac{x^{(1/n)-1}(1-x)}{\frac{1}{n}-1} &< \frac{-x^{1/n}+1}{\frac{1}{n}\left(\frac{1}{n}-1\right)} < \frac{1-x}{\frac{1}{n}-1}, \\ \sqrt[n]{x} \frac{x-1}{x} &< n(\sqrt[n]{x}-1) < x-1. \end{aligned}$$

Der Grenzübergang $n \rightarrow \infty$ ergibt, wenn wir 2.1.3., Satz 3, Satz 2, und 2.1.3.(3) anwenden,

$$\frac{x-1}{x} \leq L(x) \leq x-1 \quad (x > 0). \quad (3)$$

Hieraus folgt

$$\bigwedge_{x \in \mathbb{R}} (x > 1 \Rightarrow L(x) > 0). \quad (4)$$

Satz 1. Die Funktion (2) genügt der Funktionalgleichung der Logarithmusfunktionen und ist streng monoton wachsend.

Beweis. Unter Beachtung von 2.1.3.(5), (6), (3) gilt, wenn x, y positive reelle Zahlen sind,

$$\begin{aligned} L(xy) - L(y) &= \lim_{n \rightarrow \infty} n(\sqrt[n]{xy} - 1) - \lim_{n \rightarrow \infty} n(\sqrt[n]{y} - 1) = \lim_{n \rightarrow \infty} n(\sqrt[n]{xy} - \sqrt[n]{y}) \\ &= \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) \sqrt[n]{y} = \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) \cdot \lim_{n \rightarrow \infty} \sqrt[n]{y} = L(x) \cdot 1, \end{aligned}$$

d. h., es ist

$$L(x \cdot y) = L(x) + L(y) \quad (x, y > 0). \quad (5)$$

Für $0 < x < y$ ist, wenn wir (4) berücksichtigen,

$$L(y) = L\left(x \cdot \frac{y}{x}\right) = L(x) + L\left(\frac{y}{x}\right) > L(x),$$

und Satz 1 ist bewiesen.

Nach dem Satz in 1.4.3. ist L eine Logarithmusfunktion.

Satz 2. Für jede reelle Zahl x existiert der Grenzwert

$$E(x) := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n, \quad (6)$$

und stets ist

$$E(x) E(-x) = 1. \quad (7)$$

Beweis. Für $x = 0$ ist die Behauptung trivial, und für jede reelle Zahl x mit $x \neq 0$ ist die Folge mit den Gliedern

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

nach 2.1.1., Beispiel 2, streng monoton wachsend.

Wir zeigen, daß die gegebene Folge für jede reelle Zahl x beschränkt ist. Für $x < 0$ und $n > -x$ ist $0 < 1 + \frac{x}{n} \leq 1$ und folglich

$$\left(1 + \frac{x}{n}\right)^n \leq 1.$$

Für $0 < x \leq 1$ ist

$$\left(1 + \frac{x}{n}\right)^n \leq \left(1 + \frac{1}{n}\right)^n < e.$$

Für $x > 1$ wählen wir eine natürliche Zahl k mit $k > x$ und setzen $y := \frac{x}{k}$. Dann ist $y < 1$, und nach dem zuvor Bewiesenen ist

$$\left(1 + \frac{x}{n}\right)^n \leq \left(1 + \frac{x}{nk}\right)^{nk} = \left[\left(1 + \frac{y}{n}\right)^n\right]^k < e^k.$$

Damit ist die Behauptung bewiesen, und die Folge ist für jede reelle Zahl x konvergent, so daß wir E durch (6) definieren können. Es ist

$$E(x) \cdot E(-x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \left(1 - \frac{x}{n}\right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{x^2}{n^2}\right)^n.$$

Für $n > |x|$ folgt aus der Bernoullischen Ungleichung

$$1 \geq \left(1 - \frac{x^2}{n^2}\right)^n \geq 1 - n \frac{x^2}{n^2} = 1 - \frac{x^2}{n}$$

und damit

$$\lim_{n \rightarrow \infty} \left(1 - \frac{x^2}{n^2}\right)^n = 1,$$

womit Satz 2 bewiesen ist.

Nun können wir den Zusammenhang der Funktionen L , E mit der Eulerschen Zahl e aufdecken.

Satz 3. *Es ist*

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \quad (x \in \mathbb{R}), \quad (8)$$

$$\ln x = \lim_{n \rightarrow \infty} n(\sqrt[n]{x} - 1) \quad (x \in \mathbb{R}, x > 0). \quad (9).$$

Beweis. Wir ersetzen x in (3) für $n > -x$ durch $1 + \frac{x}{n}$ und erhalten

$$\frac{\frac{x}{n}}{1 + \frac{x}{n}} \leq L\left(1 + \frac{x}{n}\right),$$

$$\frac{x}{1 + \frac{x}{n}} \leq nL\left(1 + \frac{x}{n}\right) = L\left(\left(1 + \frac{x}{n}\right)^n\right) < L(E(x)),$$

denn die Folge mit den Gliedern $\left(1 + \frac{x}{n}\right)^n$ und die Funktion L sind streng monoton wachsend. Der Grenzübergang $n \rightarrow \infty$ liefert

$$x \leq L(E(x)) \quad (x \in \mathbb{R}).$$

Hierin können wir x auch durch $-x$ ersetzen. Da L eine Logarithmusfunktion ist und (7) gilt, folgt

$$-x \leq L(E(-x)) = L\left(\frac{1}{E(x)}\right) = L(1) - L(E(x)) = -L(E(x)),$$

$$L(E(x)) \leq x \quad (x \in \mathbb{R}).$$

Somit ist $L(E(x)) = x$, $E(x) = L^{-1}(x)$. Speziell ist $L(e) = L(E(1)) = 1$, d. h., es ist $L(x) = \log_e x = \ln x$.

Da L die Umkehrfunktion von E ist, gilt $E(x) = e^x$. Damit ist Satz 3 bewiesen.

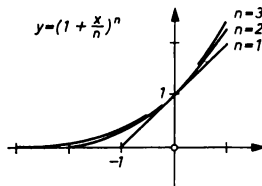


Abb. 2.5

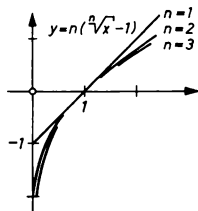


Abb. 2.6

In Abb. 2.5 bzw. 2.6 wird deutlich gemacht, wie sich die Graphen der Funktionen

$$y = \left(1 + \frac{x}{n}\right)^n \quad (n = 1, 2, \dots)$$

bzw.

$$y = n(\sqrt[n]{x} - 1) \quad (n = 1, 2, \dots)$$

den Graphen der Funktionen $y = e^x$ bzw. $y = \ln x$ nähern. Da jede Teilfolge einer konvergenten Folge wieder konvergent ist, können wir von (9) zu der Gleichung

$$\ln x = \lim_{n \rightarrow \infty} 2^n \left(\sqrt[2^n]{x} - 1\right) \quad (x \in \mathbb{R}, x > 0)$$

übergehen. Sie hat den Vorteil, daß die Glieder dieser Folge neben den elementaren Rechenoperationen allein durch wiederholtes Quadratwurzelziehen berechnet werden können. Wegen der unvermeidlichen Rundungsfehler zeigt sich aber nach anfänglichem Annähern an den Grenzwert $\ln x$ immer stärker werdendes Divergieren von diesem Grenzwert.

In der Ungleichung (3) ersetzen wir x durch a^x mit $a > 0$ und erhalten die später benötigte Ungleichung

$$\frac{a^x - 1}{a^x} \leq \ln(a^x) \leq a^x - 1,$$

$$1 - a^{-x} \leq x \ln a \leq a^x - 1 \quad (a, x \in \mathbb{R}, a > 0). \quad (10)$$

Wir wenden die Darstellung (8) der speziellen Exponentialfunktion mit der Basis e auf eine technische Problemstellung an. Der Querschnitt einer eine Last L tragenden Säule soll so gestaltet werden, daß der Druck in allen Höhenstufen den konstanten Wert c besitzt. Der Querschnitt muß dann offenbar nach unten zunehmen, da das Gewicht des darüberliegenden Säulenteils die Last vergrößert (Abb. 2.7).

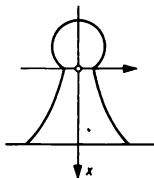


Abb. 2.7

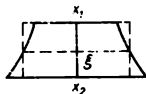


Abb. 2.8

Sind $F(x)$ die Gesamtlast und $q(x)$ der Querschnitt in der Höhe x , so muß also

$$\frac{F(x)}{q(x)} = \frac{F(0)}{q(0)} = \frac{L}{q(0)} = c$$

sein. Wie Abb. 2.8 zeigt, ist das Volumen des Säulenteils zwischen den Höhenstufen x_2 und x_1 gleich $q(\xi)(x_2 - x_1)$, wobei ξ eine passend gewählte Zahl zwischen x_1 und x_2 ist. Ist ρ das (konstante) spezifische Gewicht des Säulenmaterials, so ist also

$$F(x_2) - F(x_1) = \rho q(\xi)(x_2 - x_1),$$

$$c(q(x_2) - q(x_1)) = \rho q(\xi)(x_2 - x_1).$$

Mit $a = \frac{\rho}{c}$ folgt

$$q(x_2) - q(x_1) = a q(\xi)(x_2 - x_1) \quad (x_1 < \xi < x_2). \quad (11)$$

Die Querschnittszunahme ist also proportional zur Höhendifferenz und zu einem „mittleren“ Querschnitt.

Um Näherungslösungen für den gesuchten Querschnitt zu finden, teilen wir das Intervall $[0, x]$ in n gleiche Teile und setzen

$$x_m = m \frac{x}{n} \quad (m = 0, 1, \dots, n).$$

1. Näherungslösung. Wir ersetzen x_1, x_2, ξ in (11) durch x_{m-1}, x_m, x_{m-1} , d. h., wir berücksichtigen nur den in Abb. 2.9 schraffierten Teil des Säulenabschnitts. Es wird sich dann offensichtlich ein zu kleiner Querschnitt $q_n(x)$ ergeben. Aus

$$q_n(x_m) - q_n(x_{m-1}) = a q_n(x_{m-1})(x_m - x_{m-1})$$

oder

$$q_n \left(m \frac{x}{n} \right) - q_n \left((m-1) \frac{x}{n} \right) = a q_n \left((m-1) \frac{x}{n} \right) \frac{x}{n}$$

folgt

$$q_n \left(m \frac{x}{n} \right) = q_n \left((m-1) \frac{x}{n} \right) \left(1 + \frac{ax}{n} \right) \quad (m = 1, \dots, n).$$

Ausführlich geschrieben gilt

$$q_n \left(\frac{x}{n} \right) = q(0) \left(1 + \frac{ax}{n} \right),$$

$$q_n \left(2 \frac{x}{n} \right) = q_n \left(\frac{x}{n} \right) \left(1 + \frac{ax}{n} \right),$$

.....

$$q_n(x) = q_n \left(n \frac{x}{n} \right) = q_n \left((n-1) \frac{x}{n} \right) \left(1 + \frac{ax}{n} \right),$$

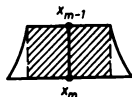


Abb. 2.9

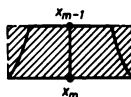


Abb. 2.10

woraus durch Multiplikation und Kürzen der Näherungswert

$$q_n(x) = q(0) \left(1 + \frac{ax}{n} \right)^n \quad (12)$$

folgt. Durch Grenzübergang erhalten wir

$$\lim_{n \rightarrow \infty} q_n(x) = q(0) e^{ax}, \quad (13)$$

und dieser Wert kann nicht größer als der exakte Wert sein.

2. Näherungslösung. Wir ersetzen x_1, x_2, ξ in (11) durch x_{m-1}, x_m, x_m . Wie Abb. 2.10 zeigt, ergibt sich ein zu großer Querschnitt q_n . Aus

$$q_n(x_m) - q_n(x_{m-1}) = a q_n(x_m) (x_m - x_{m-1}),$$

$$q_n \left(m \frac{x}{n} \right) - q_n \left((m-1) \frac{x}{n} \right) = a q_n \left(m \frac{x}{n} \right) \frac{x}{n}$$

folgt

$$q_n \left(m \frac{x}{n} \right) \left(1 - \frac{ax}{n} \right) = q_n \left((m-1) \frac{x}{n} \right).$$

Ähnlich wie im ersten Fall erhalten wir

$$q_n(x) \left(1 - \frac{ax}{n} \right)^n = q(0), \quad q_n(x) = \frac{q(0)}{\left(1 - \frac{ax}{n} \right)^n}$$

$$\lim_{n \rightarrow \infty} q_n(x) = \frac{q(0)}{e^{-ax}} = q(0) e^{ax},$$

also wiederum (13), und dieser Wert kann nicht kleiner als der exakte Wert sein. Wenn es also überhaupt eine Funktion q gibt derart, daß zu x_1, x_2 mit $x_1 < x_2$ stets ein ξ mit (11) existiert, so kann dies nur die Funktion

$$q(x) = q(0) e^{ax}$$

sein. Mit dem Mittelwertsatz der Differentialrechnung werden wir zeigen, daß diese Funktion tatsächlich die geforderte Bedingung erfüllt.

Ein ähnliches Verhalten wie bei der Querschnittsfunktion finden wir bei zahlreichen physikalischen und biologischen Prozessen. So ist z. B. die Temperaturabnahme eines erhitzten Körpers im Zeitraum von t_1 bis t_2 proportional zur Temperatur $T(\tau)$ in einem Zeitpunkt τ mit $t_1 < \tau < t_2$ und proportional zur Zeitdifferenz, d. h., es gilt

$$T(t_2) - T(t_1) = aT(\tau)(t_2 - t_1) \quad (t_1 < \tau < t_2).$$

Für den Zerfall einer (von der Zeit t abhängigen) Menge $M(t)$ eines radioaktiven Stoffes gilt

$$M(t_2) - M(t_1) = aM(\tau)(t_2 - t_1) \quad (t_1 < \tau < t_2). \quad (14)$$

In diesen beiden Beispielen ist $a < 0$. Die zur Zeit t vorhandene Holzmenge $M(t)$ eines Waldes (oder die Bakterienmenge in einer Nährlösung) genügt ebenfalls der Bedingung (14) mit $a > 0$. In allen diesen Beispielen gelangen wir wie oben zu der Funktion $T(t) = T(0) e^{at}$ bzw. $M(t) = M(0) e^{at}$. Wegen des zuletzt genannten Beispiels bezeichnet man das in dieser Weise gefundene Resultat als das *Gesetz des organischen Wachstums* und die Logarithmusfunktion zur Basis e als den *natürlichen Logarithmus*.

2.1.6. Häufungswerte

Wir führen einen Begriff ein, der als eine Verallgemeinerung des Begriffs „Grenzwert einer Zahlenfolge“ aufgefaßt werden kann.

Definition 1. Die reelle Zahl a heißt ein *Häufungswert* der reellen Zahlenfolge (a_n) , wenn es zu jeder positiven reellen Zahl ε unendlich viele natürliche Zahlen n mit $|a_n - a| < \varepsilon$ gibt.

Eine andere häufig anzutreffende Definition lautet wie folgt:

Definition 1'. Die reelle Zahl a heißt ein *Häufungswert* der reellen Zahlenfolge (a_n) , wenn eine konvergente Teilfolge (a_{k_n}) mit dem Grenzwert a existiert.

Wir beweisen die Äquivalenz der Definitionen. Die reelle Zahl a erfülle die in Definition 1 geforderte Bedingung. Dann ermitteln wir eine Folge natürlicher Zahlen $k_1 < k_2 < \dots$ mit $|a_{k_n} - a| < \frac{1}{n}$. Offensichtlich gilt dann $a_{k_n} \rightarrow a$.

Gibt es umgekehrt eine Teilfolge (a_{k_n}) mit $a_{k_n} \rightarrow a$, so gibt es zu $\varepsilon > 0$ ein n_0 mit $|a_{k_n} - a| < \varepsilon$ für $n \geq n_0$. In jeder ε -Umgebung von a liegen also unendlich viele Glieder der Folge (a_n) .

Besitzt die reelle Zahlenfolge (a_n) eine Teilfolge mit dem uneigentlichen Grenzwert ∞ bzw. $-\infty$, so nennen wir ∞ bzw. $-\infty$ einen *uneigentlichen Häufungswert* der Folge (a_n) .

Aus Definition 1' ergibt sich sofort

Satz 1. *Jede konvergente reelle Zahlenfolge besitzt genau einen Häufungswert, und zwar den Grenzwert der Folge.*

Eine Folge, die mehr als einen Häufungswert besitzt, kann daher nicht konvergent sein. Die Folge $((-1)^n)$ hat z. B. die Häufungswerte 1 und -1 .

Satz 2. *Jede reelle Zahlenfolge enthält eine monotone Teilfolge.*

Beweis. Es sei M die Menge aller natürlichen Zahlen k mit $a_{k+n} \leq a_k$ für alle n mit $n \in \mathbf{N}$.

Fall 1. Die Menge M enthält unendlich viele Zahlen k_1, k_2, \dots , wobei $k_1 < k_2 < \dots$ vorausgesetzt werden kann. Nach Definition von M ist $a_{k_1} \geq a_{k_2} \geq \dots$, die Teilfolge $(a_{k_1}, a_{k_2}, \dots)$ ist monoton (fallend).

Fall 2. Die Menge M ist endlich und besitzt daher eine obere Schranke K . Wir wählen $k_1 > K$. Wegen $k_1 \notin M$ gibt es eine (kleinste) natürliche Zahl k_2 mit $k_2 > k_1$ und $a_{k_2} > a_{k_1}$. Ist k_n für eine natürliche Zahl n schon bestimmt, so gibt es ebenso eine natürliche Zahl k_{n+1} mit $k_{n+1} > k_n$ und $a_{k_{n+1}} > a_{k_n}$, und die so induktiv definierte Folge $(a_{k_1}, a_{k_2}, \dots)$ ist (streng) monoton (wachsend).

Aus dem Konvergenzkriterium für monotone Zahlenfolgen ergibt sich als Folgerung der

Satz 3. *Jede beschränkte reelle Zahlenfolge enthält eine konvergente Teilfolge und besitzt damit einen Häufungswert.*

Dieser von BOLZANO und WEIERSTRASS stammende Satz ist für viele Anwendungen in der Analysis von fundamentaler Bedeutung.

Satz 4. *Eine reelle Zahlenfolge ist genau dann konvergent, wenn sie beschränkt ist und höchstens einen Häufungswert besitzt.*

Beweis. Ist die Folge konvergent, so ist sie beschränkt und besitzt nach Satz 1 genau einen Häufungswert.

Besitzt die Folge höchstens einen Häufungswert und ist sie beschränkt, so besitzt sie nach Satz 3 auch wenigstens einen Häufungswert, nämlich den Grenzwert einer konvergenten Teilfolge. Sie besitzt somit genau einen Häufungswert a . Wir zeigen, daß a Grenzwert der gegebenen Zahlenfolge ist. Nehmen wir an, zu einer positiven reellen Zahl ε gäbe es unendlich viele natürliche Zahlen n_0, n_1, \dots mit $n_0 < n_1 < \dots$ und $|a_{n_k} - a| \geq \varepsilon$ ($k \in \mathbf{N}$). Dann enthält die Folge $(a_{n_k})_{k \in \mathbf{N}}$ eine konvergente Teilfolge mit einem Grenzwert, der wegen $|a_{n_k} - a| \geq \varepsilon$ von a verschieden ist. Die Folge (a_n) hätte dann aber zwei Häufungswerte.

Satz 5. *Besitzen zwei reelle Zahlenfolgen (a_n) und (a'_n) den gemeinsamen Grenzwert a , so besitzt auch die Mischfolge*

$$(a_0, a'_0, a_1, a'_1, \dots, a_n, a'_n, \dots)$$

den Grenzwert a .

Beweis. Die Mischfolge ist beschränkt und besitzt neben a keinen weiteren Häufungswert.

Es sei (a_n) eine reelle Zahlenfolge und

$$a_n' := \inf \{a_k : k \geq n\}, \quad (1)$$

$$a_n'' := \sup \{a_k : k \geq n\}. \quad (2)$$

Für alle natürlichen Zahlen n ist dann

$$-\infty \leq a_n' \leq a_n \leq a_n'' \leq \infty, \quad (3)$$

$$a_n' \leq a_{n+1}', \quad (4)$$

$$a_n'' \geq a_{n+1}''. \quad (5)$$

Ist die Folge (a_n) nach unten bzw. nach oben beschränkt, so ist also (a_n') bzw. (a_n'') eine monoton wachsende bzw. fallende reelle Zahlenfolge. Wir bezeichnen ihren (eigentlichen oder uneigentlichen) Grenzwert als *Limes inferior* bzw. *Limes superior* der Folge (a_n) und setzen

$$\liminf_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} \inf \{a_k : k \geq n\} \quad (6)$$

bzw.

$$\limsup_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} \sup \{a_k : k \geq n\}. \quad (7)$$

Neben diesen Beziehungen sind auch die Schreibweisen

$$\varliminf_{n \rightarrow \infty} a_n := \liminf_{n \rightarrow \infty} a_n, \quad \overline{\lim}_{n \rightarrow \infty} a_n := \limsup_{n \rightarrow \infty} a_n$$

üblich. Ist die Folge (a_n) nicht nach unten bzw. nicht nach oben beschränkt, so ist $a_n' = -\infty$ bzw. $a_n'' = \infty$ für alle $n \in \mathbb{N}$. In diesem Fall bezeichnen wir auch $-\infty$ bzw. ∞ als Limes inferior bzw. Limes superior der Folge (a_n) .

Satz 6. *Der Limes inferior bzw. Limes superior einer reellen Zahlenfolge ist der kleinste bzw. größte Häufungswert der Folge.¹⁾*

Beweis. Es sei a' der Limes inferior der Folge (a_n) . Ist $a' = -\infty$, d. h., ist die Folge nicht nach unten beschränkt, so gibt es zu jedem $n \in \mathbb{N}$ einen Index k_n mit $a_{k_n} < -n$. Die Folge (a_n) hat also den uneigentlichen Häufungswert $-\infty$. Ist aber die Folge (a_n) nach unten beschränkt, so ist stets $a_n' \in \mathbb{R}$, und auf Grund der Definition dieser Zahlen als Infimum gibt es eine Folge natürlicher Zahlen k_n mit $k_n \geq n$ und

$$a_0' \leq a_{k_0} < a_0' + 1,$$

$$a_1' \leq a_{k_1} < a_1' + \frac{1}{2},$$

$$a_2' \leq a_{k_2} < a_2' + \frac{1}{3},$$

.....

¹⁾ Dieser Häufungswert kann auch uneigentlich sein.

Hieraus folgt

$$\lim_{n \rightarrow \infty} a_{k_n} = \lim_{n \rightarrow \infty} a'_n = a',$$

und die Folge (a_n) hat den Häufungswert a' . Da die Folge (a'_n) monoton wachsend ist, können nur die Fälle $a' \in \mathbf{R}$ oder $a' = \infty$ auftreten.

Analog zeigt man, daß a'' ein eigentlicher oder uneigentlicher Häufungswert der Folge (a_n) ist.

Es sei nun a ein beliebiger Häufungswert der Folge (a_n) . Dann gibt es eine Teilfolge (a_{k_n}) mit $a_{k_n} \rightarrow a$. Nach (3) ist $a'_{k_n} \leq a_{k_n} \leq a''_{k_n}$, und der Grenzübergang $n \rightarrow \infty$ liefert $a' \leq a \leq a''$. Damit ist gezeigt, daß a' der kleinste und a'' der größte Häufungswert der Folge (a_n) ist.

Beispiel. Für die Folge

$$(a_n) = \left(1, -2, \frac{1}{3}, -4, \frac{1}{5}, \dots\right)$$

gilt stets $a'_n = -\infty$, also auch $a' = -\infty$ und

$$(a''_n) = \left(1, \frac{1}{3}, \frac{1}{3}, \frac{1}{5}, \frac{1}{5}, \dots\right)$$

und damit $a'' = 0$. Die Folge besitzt neben $-\infty$ und 0 keine weiteren Häufungswerte.

Satz 7. Eine reelle Zahlenfolge (a_n) ist genau dann konvergent oder bestimmt divergent, wenn ihr Limes inferior und ihr Limes superior übereinstimmen, und in diesem Fall ist

$$\lim_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n. \quad (8)$$

Beweis. Ist $a' = a''$, so folgt (8) aus (3). Ist die Folge (a_n) konvergent oder bestimmt divergent, so besitzt sie genau einen Häufungswert, d. h., es ist $a' = a''$.

Die Definition des Häufungswertes und Satz 1 können unverändert auf komplexe Zahlenfolgen bzw. auf Folgen im euklidischen Raum \mathbf{R}^p oder noch allgemeiner auf Folgen in einem metrischen Raum übertragen werden. Wir zeigen, daß der Satz von BOLZANO-WEIERSTRASS auch für komplexe Zahlenfolgen und für Folgen im euklidischen Raum \mathbf{R}^p gilt.

Satz 8. Jede beschränkte Folge im Raum \mathbf{C} bzw. im euklidischen Raum \mathbf{R}^p enthält eine konvergente Teilfolge und besitzt damit einen Häufungswert.

Beweis. Es sei $\mathbf{a}_n = (\alpha_{1n}, \dots, \alpha_{pn})$, und die Folge (\mathbf{a}_n) sei beschränkt. Wegen $|\alpha_{in}| \leq \|\mathbf{a}_n\|$ ($i = 1, \dots, p$; $n \in \mathbf{N}$) sind dann auch die Folgen (α_{in}) beschränkt.

Wir wählen aus der beschränkten Zahlenfolge (α_{1n}) eine konvergente Teilfolge und aus der Folge (\mathbf{a}_n) die entsprechende Teilfolge aus. Bezeichnen wir diese der Einfach-

heit halber wieder mit (a_n) , so konvergiert die Folge der ersten Koordinaten, und wir können wieder eine Teilfolge auswählen, für die auch die Folge der zweiten Koordinaten konvergiert. Die Konvergenz der Folge der ersten Koordinaten bleibt hierbei erhalten. In dieser Weise fortfahrend, erhalten wir nach dem p -ten Schritt eine Teilfolge der Ausgangsfolge, in der alle Koordinatenfolgen konvergieren, die also selbst konvergent ist. Analog verläuft der Beweis für komplexe Zahlenfolgen.

2.1.7. Fundamentalfolgen

Für reelle monotone Zahlenfolgen haben wir in 2.1.4. ein sehr einfaches Konvergenzkriterium gefunden. In vielen Fällen konnten wir Grenzwerte auch mit Hilfe der Grenzwertsätze ermitteln. Im allgemeinen Fall sind wir bisher auf die Definition 1 in 2.1.3. angewiesen. Diese Definition enthält aber einen für die Anwendungen schwerwiegenden Mangel. Um sie auf eine Zahlenfolge (a_n) anwenden zu können, muß man bereits von einer reellen Zahl a wissen oder zumindest vermuten, daß sie der Grenzwert der Folge ist. Erst dann kann man ja untersuchen, ob die Folge $(a_n - a)$ eine Nullfolge ist. Von CAUCHY stammt die Erkenntnis, daß die Entscheidung, ob eine Folge konvergent ist oder nicht, unabhängig von der Kenntnis des Grenzwertes getroffen werden kann. Wir führen den folgenden Begriff ein.

Definition 1. Eine reelle Zahlenfolge (a_n) heißt eine *Fundamentalfolge* oder *Cauchyfolge*, wenn es zu jeder positiven Zahl ε eine reelle Zahl $N(\varepsilon)$ gibt, derart, daß für alle natürlichen Zahlen m, n aus $m, n \geq N(\varepsilon)$ stets $|a_m - a_n| < \varepsilon$ folgt.

Die eine Fundamentalfolge charakterisierende Bedingung kann analog 2.1.2.(2) formalisiert wie folgt angegeben werden:

$$\bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbb{N}} \bigwedge_{m \in \mathbb{N}} \bigwedge_{n \in \mathbb{N}} (m, n \geq N \Rightarrow |a_m - a_n| < \varepsilon).$$

Wir beweisen einige Sätze über Fundamentalfolgen.

Satz 1. Jede konvergente reelle Zahlenfolge ist eine Fundamentalfolge.

Beweis. Die Folge (a_n) sei konvergent mit dem Grenzwert a . Ist N eine Schranke der Menge $\left\{n: |a_n - a| \geq \frac{\varepsilon}{2}\right\}$, so folgt aus $m, n \geq N$ stets

$$|a_m - a_n| \leq |a_m - a| + |a_n - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

die Folge ist eine Fundamentalfolge.

Satz 2. Jede Fundamentalfolge reeller Zahlen ist beschränkt.

Beweis. Nach Definition gibt es zu $\varepsilon = 1$ eine reelle Zahl $N(1)$ mit $|a_m - a_n| < 1$ für alle m, n mit $m, n \geq N(1)$. Wir wählen eine feste Zahl $m > N(1)$. Für $n \geq m$ gilt dann

$$|a_n - a_0| \leq |a_m - a_0| + |a_m - a_n| < |a_m - a_0| + 1,$$

die Zahlenfolge ist beschränkt.

Satz 3. *Jede Fundamentalfolge reeller Zahlen ist konvergent.*

Beweis. Nach dem Satz von BOLZANO-WEIERSTRASS und Satz 2 enthält jede Fundamentalfolge (a_n) eine konvergente Teilfolge (a_{n_k}) . Es sei $a = \lim_{k \rightarrow \infty} a_{n_k}$. Dann gibt es zu jeder positiven reellen Zahl ε reelle Zahlen N_1 bzw. N_2 mit $|a_{n_k} - a| < \frac{\varepsilon}{2}$ für $n_k \geq N_1$ bzw. $|a_m - a_n| < \frac{\varepsilon}{2}$ für $m, n \geq N_2$. Ist N das Maximum der Zahlen N_1, N_2 , so folgt aus $n, n_k \geq N$ stets

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d. h., die gegebene Fundamentalfolge besitzt den Grenzwert a .

Eine Zusammenfassung der Sätze 1 und 3 ergibt das wichtige, notwendige und hinreichende

Konvergenzkriterium von CAUCHY. *Eine reelle Zahlenfolge ist konvergent genau dann, wenn sie eine Fundamentalfolge ist.*

Beispiel 1. Wir betrachten die Folge mit den Gliedern

$$a_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n.$$

Für alle natürlichen Zahlen n mit $n \geq 1$ ist

$$a_{n+1} - a_n = \frac{1}{n+1} - \ln(n+1) + \ln n.$$

Nach 2.1.5.(3) gilt

$$\begin{aligned} \frac{\frac{n+1}{n} - 1}{\frac{n+1}{n}} &\leq \ln\left(\frac{n+1}{n}\right) \leq \frac{n+1}{n} - 1, \\ \frac{1}{n+1} &\leq \ln\left(\frac{n+1}{n}\right) \leq \frac{1}{n}. \end{aligned} \tag{1}$$

Es folgt

$$0 \leq \ln \frac{n+1}{n} - \frac{1}{n+1} = a_n - a_{n+1} \leq \frac{1}{n} - \frac{1}{n+1}.$$

Entsprechend ist

$$0 \leq a_{n+1} - a_{n+2} \leq \frac{1}{n+1} - \frac{1}{n+2},$$

$$\dots\dots\dots$$

$$0 \leq a_{n+k-1} - a_{n+k} \leq \frac{1}{n+k-1} - \frac{1}{n+k} \quad (k = 1, 2, \dots).$$

Addition ergibt

$$0 \leq a_n - a_{n+k} \leq \frac{1}{n} - \frac{1}{n+k} \leq \frac{1}{n},$$

$$|a_{n+k} - a_n| \leq \frac{1}{n}.$$

Sind die natürlichen Zahlen m, n größer als $\frac{1}{\varepsilon}$, so gilt also

$$|a_m - a_n| < \varepsilon,$$

d. h., es liegt eine Fundamentalfolge vor. Ihr Grenzwert wird als *Eulersche Konstante* C bezeichnet. EULER führte die Berechnung von C bis auf 16 Stellen durch.

Beispiel 2. Es sei (a_n) eine beschränkte Zahlenfolge und q eine reelle Zahl mit $|q| < 1$. Wir zeigen, daß die Zahlenfolge mit den Gliedern

$$s_n := \sum_{k=0}^n a_k q^k$$

eine Fundamentalfolge ist. Es gilt, wenn K eine Schranke der Folge (a_n) ist,

$$\begin{aligned} |s_{n+p} - s_n| &= \left| \sum_{k=n+1}^{n+p} a_k q^k \right| \leq \sum_{k=n+1}^{n+p} |a_k| |q|^k \\ &\leq K \sum_{k=n+1}^{n+p} |q|^k = K |q|^{n+1} \sum_{k=0}^{p-1} |q|^k = K |q|^{n+1} \frac{1 - |q|^p}{1 - |q|} \\ &\leq K |q|^{n+1} \frac{1}{1 - |q|}. \end{aligned}$$

Nach 2.1.2., Satz 3, können wir zu jedem positiven ε ein $N(\varepsilon)$ ermitteln, so daß die rechte Seite für natürliche Zahlen mit $n \geq N(\varepsilon)$ stets kleiner als ε wird. Die Folge (s_n) ist somit eine Fundamentalfolge.

Der Begriff der Fundamentalfolge und die Sätze 1 und 2 können unverändert auf komplexe Zahlenfolgen, auf Folgen im Raum \mathbf{R}^p sowie auf Folgen in normierten bzw. metrischen Räumen übertragen werden. Für komplexe Zahlenfolgen und Folgen im Raum \mathbf{R}^p bleiben auch Satz 3 und das Konvergenzkriterium von CAUCHY unverändert gültig.

Metrische Räume, in denen dieses Kriterium hinreichend für die Konvergenz ist, erhalten einen besonderen Namen.

Definition 2. Ein metrischer Raum X bzw. ein normierter Raum E heißt *vollständig*, wenn jede Fundamentalfolge in diesem Raum konvergent ist.¹⁾ Gibt es in X eine Fundamentalfolge, die nicht konvergent ist, so heißt der Raum *unvollständig*.

Die normierten Räume \mathbf{R} , \mathbf{R}^p und \mathbf{C} sind hiernach vollständig. Der normierte Raum \mathbf{Q} der rationalen Zahlen mit der Norm $|x|$ ist dagegen nicht vollständig. So besitzt z. B. die Fundamentalfolge $\left(1 + \frac{1}{n}\right)^n$ keinen rationalen Grenzwert.

Die folgenden Begriffsbildungen spielen in der modernen Analysis eine hervorragende Rolle.

Definition 3. Eine Teilmenge M eines metrischen Raumes X heißt *präkompakt*, wenn jede Folge in M eine Teilfolge besitzt, die eine Fundamentalfolge ist.

Satz 4. Eine Teilmenge M des Raumes \mathbf{R} oder \mathbf{C} oder \mathbf{R}^p ist genau dann präkompakt, wenn sie beschränkt ist.

Beweis. Die Menge M sei beschränkt. Da jede Folge (a_n) in M beschränkt ist, besitzt sie nach 2.1.6., Satz 8, eine konvergente Teilfolge. Folglich ist M präkompakt.

Ist die Menge M nicht beschränkt, so gibt es eine Folge von Elementen $a_n \in M$ mit $\|a_n\| \rightarrow \infty$. Dieselbe Eigenschaft hat aber dann jede Teilfolge dieser Folge. Es gibt also keine beschränkte Teilfolge der Folge (a_n) . Dies besagt nach Satz 2, daß M nicht präkompakt ist, und Satz 4 ist bewiesen.

Definition 4. Eine Teilmenge M eines metrischen Raumes X heißt *kompakt*, wenn jede Folge in M eine gegen ein Element $a \in M$ konvergierende Teilfolge besitzt.

Satz 5. Eine Teilmenge M eines vollständigen metrischen Raumes ist genau dann kompakt, wenn sie präkompakt und abgeschlossen ist.

Beweis. Die Menge M sei kompakt. Dann ist sie auch präkompakt, denn jede konvergente Folge ist eine Fundamentalfolge. Ist a ein Häufungspunkt von M , so gibt es eine Folge (a_n) in M , die gegen a konvergiert. Nach Definition der Kompaktheit besitzt sie eine gegen ein Element $a' \in M$ konvergierende Teilfolge. Dies ist nur für $a = a' \in M$ möglich, und folglich ist M abgeschlossen.

Ist M aber präkompakt und abgeschlossen, so besitzt jede Folge (a_n) in M eine Teilfolge (a_{n_k}) , die eine Fundamentalfolge ist. Da X vollständig ist, besitzt sie einen Grenzwert a , und a ist ein Häufungswert von M . Somit ist $a \in M$, und M ist kompakt.

2.1.8. Die Sätze von Cantor und Heine-Borel

In diesem Abschnitt beschäftigen wir uns mit zwei Sätzen, die für weiterführende Betrachtungen der Analysis bedeutsam sind. Der erste wird nach GEORG CANTOR (1845–1918) benannt und gibt eine Verallgemeinerung des Satzes 2 in 2.1.4. über Intervallschachtelungen.

Satz 1. Es sei (F_1, F_2, \dots) eine Folge abgeschlossener, nichtleerer Mengen des euklidischen Raumes \mathbf{R}^p mit $F_1 \supseteq F_2 \supseteq \dots$ und $\lim_{n \rightarrow \infty} d(F_n) = 0$. Dann gibt es genau einen Punkt x , der in allen Mengen F_n enthalten ist.

Beweis.²⁾ Wir wählen aus jeder Menge F_n ein Element x_n . Wegen $F_{n+p} \subseteq F_n$ ($n, p \in \mathbf{N}$) gilt dann $d(x_n, x_{n+p}) \leq d(F_n)$. Hieraus ergibt sich, daß (x_n) eine Fundamentalfolge ist, denn zu jedem $\varepsilon > 0$ gibt es ein $N(\varepsilon)$ mit $d(F_n) < \varepsilon$ für $n \geq N(\varepsilon)$. Wegen der Vollständigkeit des

¹⁾ Wesentlich ist hierbei, daß der Grenzwert ebenfalls ein Element des Raumes ist.

²⁾ Hier wird von dem Auswahlprinzip Gebrauch gemacht.

Raumes \mathbb{R}^p existiert ein Element x mit $x \in \mathbb{R}^p$ und $x = \lim_{n \rightarrow \infty} x_n$. Nehmen wir an, es gäbe ein n mit $x \notin F_n$. Dann ist x in der offenen Menge $\mathbb{R}^p \setminus F_n$ enthalten, und es gibt eine Umgebung $U_\varepsilon(x)$, die keinen Punkt von F_n , damit aber auch keinen Punkt von F_{n+p} ($p \in \mathbb{N}$) enthält. Es folgt $\varrho(x_{n+p}, x) \geq \varepsilon$ für alle natürlichen Zahlen p , was wegen $x_n \rightarrow x$ nicht möglich ist. Somit gilt $x \in F_n$ für alle natürlichen Zahlen n . Ist auch $y \in F_n$ für alle n mit $n \in \mathbb{N}$, so gilt stets $\varrho(x, y) \leq d(F_n)$, woraus $\varrho(x, y) = 0$, $x = y$ folgt.

Der zweite Satz wird in der mathematischen Literatur häufig als *Überdeckungssatz von Heine-Borel* (EDUARD HEINE (1821–1881), EMILE BOREL (1871–1956)) bezeichnet. Bevor wir ihn formulieren, geben wir die

Definition 1. Ein Mengensystem \mathcal{G} von offenen Mengen G eines metrischen Raumes X heißt eine *Überdeckung* der Menge M mit $M \subseteq X$, wenn M in der Vereinigungsmenge des Mengensystems \mathcal{G} enthalten ist.

In Abb. 2.11 bildet z. B. das Mengensystem $\mathcal{G} = \{G_1, G_2, G_3, G_4, G_5\}$ eine Überdeckung der Menge M .

Satz 2. Zu jeder Überdeckung \mathcal{G} einer abgeschlossenen beschränkten Menge F des euklidischen Raumes \mathbb{R}^p gibt es ein endliches Teilsystem $\{G_1, G_2, \dots, G_n\}$ mit $G_j \in \mathcal{G}$ ($j = 1, 2, \dots, n$), das ebenfalls eine Überdeckung von F ist.

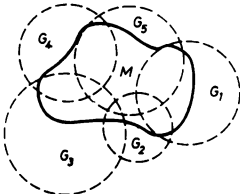


Abb. 2.11

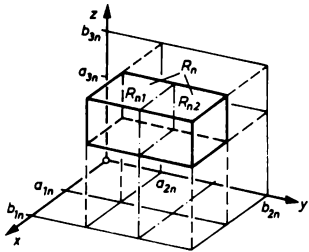


Abb. 2.12

Beweis. Es sei G die Vereinigung des Mengensystems \mathcal{G} . Nach Voraussetzung gilt $F \subseteq G$. Mit \mathfrak{M} bezeichnen wir das Mengensystem aller derjenigen Teilmengen von G , zu deren Überdeckung endlich viele Mengen des Mengensystems \mathcal{G} ausreichen. Der Satz ist bewiesen, wenn die Annahme $F \notin \mathfrak{M}$ zu einem Widerspruch führt.

Mit Hilfe des sogenannten *Halbierungsverfahrens* konstruieren wir eine Folge (R_n) von Rechtecksbereichen R_n . Da die Menge F beschränkt ist, besitzt die Menge aller Koordinaten der Punkte x mit $x \in F$ eine Schranke K . Wir setzen

$$R_0 = \{(x_1, \dots, x_p) : |x_j| \leq K, j = 1, \dots, p\}.$$

Wegen $F \subseteq R_0$ ist nach unserer Annahme $R_0 \cap F = F \notin \mathfrak{M}$.

Ist für eine natürliche Zahl n der Rechtecksbereich

$$R_n = \{(x_1, x_2, \dots, x_p) : a_{jn} \leq x_j \leq b_{jn}, j = 1, 2, \dots, p\}$$

schon definiert und ist $R_n \cap F \notin \mathfrak{M}$, so bestimmen wir die maximale Kantenlänge, d. h. das Maximum der Zahlen $b_{jn} - a_{jn}$, und den kleinsten Index j_0 , für den dieses Maximum angenom-

men wird. Wir zerlegen R_n in zwei Teilbereiche R_{n1} und R_{n2} , indem wir das Intervall $[a_{jn}, b_{jn}]$ halbieren und alle anderen Intervalle $[a_{jn}, b_{jn}]$ ($j \neq j_0$) unverändert lassen (Abb. 2.12).

Von den Mengen $R_{n1} \cap F$ und $R_{n2} \cap F$ kann wegen

$$R_n \cap F = (R_{n1} \cap F) \cup (R_{n2} \cap F)$$

höchstens eine im Mengensystem \mathfrak{M} enthalten sein. Ist $R_{n1} \cap F \notin \mathfrak{M}$, so setzen wir $R_{n+1} = R_{n1}$, andernfalls $R_{n+1} = R_{n2}$. Dann ist R_{n+1} ein Rechtecksbereich und $R_{n+1} \cap F \notin \mathfrak{M}$. Damit ist die Folge (R_n) induktiv definiert. Da bei jedem Konstruktionsschritt die bzw. eine längste Kante des Rechtecksbereichs halbiert wird, beträgt die maximale Kantenlänge des Rechtecksbereichs R_{n+p} höchstens die Hälfte der maximalen Kantenlänge des Rechtecksbereichs R_n . Somit bildet die Folge der maximalen Kantenlängen, damit aber auch die Folge der Durchmesser der Rechtecksbereiche R_n eine Nullfolge.

Die Mengen F_n mit $F_n = R_n \cap F$ sind abgeschlossen und nicht leer, es gilt $d(F_n) \rightarrow 0$ für $n \rightarrow \infty$, und wegen $R_n \subseteq R_{n+1}$ gilt $F_n \subseteq F_{n+1}$. Nach Satz 1 gibt es genau einen Punkt x mit $x \in F_n$ ($n \in \mathbb{N}$). Wegen $F_n \subseteq F$ ist $x \in F$, und da \mathcal{G} eine Überdeckung von F ist, gibt es eine offene Menge G_0 mit $x \in G_0$ und $G_0 \in \mathcal{G}$. Nach Definition der offenen Menge gibt es ein ε mit $U_\varepsilon(x) \subseteq G_0$.

Wir bestimmen eine natürliche Zahl n derart, daß der Durchmesser von R_n kleiner als ε ist. Da $x \in F_n \subseteq R_n$ gilt, muß R_n in $U_\varepsilon(x)$ enthalten sein, und es folgt $F_n \subseteq R_n \subseteq U_\varepsilon(x) \subseteq G_0$. Zur Überdeckung der Menge F_n genügt also eine einzige Menge des Mengensystems \mathcal{G} . Das ist wegen $F_n \notin \mathfrak{M}$ ein Widerspruch zur Definition von \mathfrak{M} , und Satz 2 ist bewiesen.

In Verallgemeinerung von Satz 2 kann man beweisen, daß eine Teilmenge F eines metrischen Raumes X genau dann kompakt ist, wenn aus jeder Überdeckung von F ein endliches Teilsystem ausgewählt werden kann, das ebenfalls eine Überdeckung von F ist.

2.2. Reihen

2.2.1. Grundbegriffe der Reihenlehre

Bevor die Grundbegriffe der Analysis im vorigen Jahrhundert exakt formuliert wurden, war es üblich, eine unendliche Reihe als eine „Summe“ von unendlich vielen Gliedern aufzufassen und mit ihnen ebenso wie mit endlichen Summen zu rechnen. Wir wollen an zwei Beispielen zeigen, daß eine solche Erweiterung des Summenbegriffs nicht möglich ist, weil sie zu Widersprüchen führt.

Beispiel 1. Es sei

$$s := 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \dots \quad (1)$$

Dann ist

$$\frac{s}{2} = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \dots \quad (2)$$

Subtrahieren wir (2) von (1), so folgt

$$\frac{s}{2} = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \dots \quad (3)$$

Subtrahieren wir nun (2) von (3), so erhalten wir

$$0 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots,$$

$$0 = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots > 1 - \frac{1}{2} = \frac{1}{2},$$

und wir sind zu einem Widerspruch gelangt.

Beispiel 2. Es sei

$$s := 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots. \quad (4)$$

Dann ist

$$s = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6}\right) + \dots > 1 - \frac{1}{2} = \frac{1}{2}.$$

Eine Umordnung der Summanden ergibt

$$s = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + + - \dots. \quad (5)$$

Aus (4) folgt

$$\frac{s}{2} = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \dots. \quad (6)$$

Addition von (5) und (6) ergibt

$$\frac{3}{2}s = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} - \frac{1}{2} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} - \frac{1}{4} + + + - \dots,$$

und das ist wiederum (4) in veränderter Anordnung der Summanden. Somit ist $s = \frac{3}{2}s$, was wegen $s > \frac{1}{2}$ nicht möglich ist.

Diese Beispiele machen deutlich, daß bei der Begriffsbildung und beim Rechnen mit „unendlichen Summen“ sehr sorgfältig vorgegangen werden muß. Wir werden den Begriff der Summe nicht auf unendlich viele Summanden ausdehnen, sondern den Begriff der Reihe auf den Folgenbegriff zurückführen.

Definition 1. Ist (a_n) eine Zahlenfolge, so heißt die Zahlenfolge (s_n) mit den Gliedern

$$s_n := \sum_{k=0}^n a_k, \quad (7)$$

d. h. die Zahlenfolge

$$(a_0, a_0 + a_1, a_0 + a_1 + a_2, \dots),$$

die *Folge der Partialsummen* der Folge (a_n) .

Das Wesen der Reihenlehre besteht darin, das Verhalten, insbesondere das Konvergenzverhalten der Partialsummenfolge, nicht auf die Untersuchung ihrer Glieder s_n , sondern auf die Untersuchung der Glieder a_n derjenigen Folge zurückzuführen, deren Partialsummenfolge sie ist. In der Reihenlehre sind also stets zwei Folgen nebeneinander zu betrachten, deren Glieder durch (7) miteinander verknüpft sind. Wir geben die folgende

Definition 2. Unter der (unendlichen) *Reihe mit den Gliedern a_n* verstehen wir die Folge der Partialsummen $\sum_{k=0}^n a_k$ der Folge (a_n) .

Diese Reihe wird mit

$$\sum_{n=0}^{\infty} a_n \quad \text{oder} \quad a_0 + a_1 + a_2 + \dots + a_n + \dots$$

bezeichnet.

Eine Reihe ist entweder durch ihre Glieder oder aber auch durch ihre Partialsummen bestimmt. Aus den Partialsummen kann man nämlich ihre Glieder berechnen, denn aus (7) folgt

$$a_0 = s_0, \quad a_{n+1} = s_{n+1} - s_n \quad (n \in \mathbf{N}). \quad (8)$$

Hiernach kann jede gegebene Folge (s_n) als Partialsummenfolge der Reihe mit den durch (8) definierten Gliedern a_n ($n \in \mathbf{N}$) aufgefaßt werden.

Ist nicht 0, sondern n_0 der Anfangsindex der Folge (a_n) , so wird die Reihe mit

$$\sum_{n=n_0}^{\infty} a_n$$

bezeichnet. Der Summationsbuchstabe n kann durch jeden von a verschiedenen Buchstaben ersetzt werden. Wir werden im folgenden alle allgemeinen Aussagen der Reihenlehre nur für den Fall $n_0 = 0$ formulieren. Die Übertragung auf einen beliebigen anderen Anfangsindex bereitet keine Schwierigkeiten.

Wir führen einige weitere Bezeichnungen ein.

Definition 3. Für jede natürliche Zahl m heißt die Reihe $\sum_{n=0}^{\infty} a_n$ ein *Rest* und die (endliche) Summe $\sum_{n=m+1}^{m+p} a_n$ ($p \geq 1$) ein *Ausschnitt* der Reihe $\sum_{n=0}^{\infty} a_n$.

Offenbar kann jeder Ausschnitt in der Form $s_{m+p} - s_m$ dargestellt werden.

Der Begriff der Reihe kann unverändert auf den euklidischen Raum \mathbf{R}^p übertragen werden.

Die Geschichte der Behandlung der unendlichen Reihen bringt wie kaum ein anderes mathematisches Gebiet die unterschiedlichen Auffassungen von mathematischer Strenge zum Ausdruck. Der Freude über die gefundenen neuen Methoden folgte eine Periode der recht zügellosen Anwendung und Verwendung von unendlichen Reihen, wobei der ursprünglich noch geometrisch-analytischen Spekulation später mehr formale Aspekte abgewonnen wurden. Erst mit GAUSS beginnt das neue mathematische Zeitalter, in dem neben formaler Gewandtheit auch Strenge der Untersuchung verlangt wird.

Unendliche Reihen traten erstmals bei ARCHIMEDES (287?–212 v. u. Z.) auf. In seiner „Quadratur der Parabel“ lieferte er auch erstmals die Summation einer solchen Reihe mit dem Quotienten $1/4$.

Die Lehre von den unendlichen Reihen wurde im modernen Sinn von NICOLAUS MERCATOR eröffnet, der 1667 die logarithmische Reihe entdeckte. Allerdings traten auch in mittelalterlichen Schriften (von ORESME, ALVARUS THOMAS (um 1500)) gelegentlich unendliche Reihen auf, ohne daß erkennbar ist, woher die Mathematiker des Mittelalters ihre Kenntnisse bezogen. Auch war NEWTON seit 1666 im Besitz der Binominalreihe, ohne, wie es damals häufig der Fall war, die Ergebnisse im Druck zu veröffentlichen. In schneller Folge wurden jetzt weitere Reihen entdeckt. Die Untersuchung dieser Reihen wurden dann zu einer der wichtigsten Quellen der Schaffung der Analysis bei LEIBNIZ und NEWTON. Bei der Erforschung von unendlichen Reihen beschäftigte man sich jedoch kaum mit dem Konvergenzverhalten, sondern operierte unbefangen mit ihnen ebenso wie mit endlichen Ausdrücken. Ein Beispiel für den Umgang mit sogar unbestimmt divergenten Reihen bietet die Behandlung der berühmten „Reihe“ von GUIDO GRANDI (1671–1742)

$$1 - 1 + 1 - 1 + \dots,$$

der der Wert $1/2$ zugeordnet wird, und diese Behauptung wird sogar von LEIBNIZ und EULER trotz der Einsprüche einiger Mathematiker aufrechterhalten.

Durch GAUSS, ABEL, BOLZANO, CAUCHY wurde die Konvergenztheorie der unendlichen Reihen auf sichere Fundamente gestellt. Erinnert sei an die Untersuchung von GAUSS zur hypergeometrischen Reihe (1812), ABELS Arbeiten über die binomische Reihe, CAUCHYS „Cours d'analyse“ (1821) und BOLZANOS „Functionenlehre“ und „Paradoxien des Unendlichen“.

2.2.2. Konvergenz von Reihen

Die Konvergenz von Reihen wird auf die Konvergenz von Folgen zurückgeführt.

Definition 1. Eine Reihe heißt *konvergent* bzw. *divergent* bzw. *bestimmt divergent*¹⁾, wenn die Folge ihrer Partialsummen konvergent bzw. divergent bzw. bestimmt divergent ist. Besitzt die Folge der Partialsummen den eigentlichen oder uneigentlichen Grenzwert s , so heißt s die *Summe* der Reihe, und man setzt

$$s = \sum_{n=0}^{\infty} a_n := \lim_{k \rightarrow \infty} \sum_{n=0}^k a_n. \quad (1)$$

Es ist ein bedauerlicher und nur aus der historischen Entwicklung zu verstehender Umstand, daß das Symbol $\sum_{n=0}^{\infty} a_n$ damit eine doppelte Bedeutung erlangt hat, während sonst in der Mathematik auf eindeutige, unmißverständliche Bezeichnungen großer Wert gelegt wird. Nach unserer in 2.2.1. gegebenen Definition 2 bedeutet dieses

¹⁾ Die Begriffe treten erstmals 1668 bei dem englischen Mathematiker JAMES GREGORY (1638–1675) in seinem Werk „Vera circuli et Hyperbolae Quadratura“ auf, haben aber dort einen etwas anderen Sinn.

Symbol die Folge der Partialsummen der Folge (a_n) , während es im Fall der Konvergenz nach (1) zugleich den Grenzwert der Partialsummenfolge bezeichnet. Allerdings wird aus dem Zusammenhang gewöhnlich hervorgehen, welche der beiden Bedeutungen gemeint ist. Solange wir es mit konvergenten oder bestimmt divergenten Reihen zu tun haben, werden wir das Symbol gewöhnlich in der zweiten Bedeutung, d. h. als Summe der Reihe, verstehen.

Zur Vereinfachung schreiben wir im folgenden häufig $\sum a_n$ an Stelle der ausführlichen Bezeichnung $\sum_{n=0}^{\infty} a_n$.

Alle Sätze über die Konvergenz von Zahlenfolgen (a_n) können auf Reihen $\sum a_n$ übertragen werden, indem man in den jeweiligen Kriterien a_n durch $s_n = \sum_{k=0}^n a_k$ ersetzt. Die wichtigsten sich hierbei ergebenden Sätze stellen wir noch einmal zusammen.

Satz 1. *Konvergiert die Reihe $\sum a_n$, so konvergiert die Reihe $\sum ca_n$ für jede reelle bzw. komplexe Zahl c , und es ist $\sum ca_n = c \sum a_n$.*

Beweis. Sind s_n bzw. s'_n die Partialsummen der Reihe $\sum a_n$ bzw. $\sum ca_n$, so ist $s'_n = cs_n$, und aus $s_n \rightarrow s$ folgt $s'_n \rightarrow cs$.

Satz 2. *Konvergieren die Reihen $\sum a_n$ und $\sum b_n$, so konvergiert die Reihe $\sum (a_n \pm b_n)$, und es ist*

$$\sum (a_n \pm b_n) = \sum a_n \pm \sum b_n.$$

Beweis. Sind s'_n, s''_n bzw. s_n die Partialsummen der Reihen $\sum a_n, \sum b_n$ bzw. $\sum (a_n \pm b_n)$, so ist $s_n = s'_n \pm s''_n$, und aus $s'_n \rightarrow s', s''_n \rightarrow s''$ folgt $s_n \rightarrow s' \pm s''$.

Man drückt den Inhalt des hiermit bewiesenen Satzes häufig wie folgt aus: Konvergente Reihen dürfen gliedweise addiert bzw. subtrahiert werden.

Satz 3. *Konvergiert die Reihe $\sum a_n$, so konvergiert auch jede Reihe, deren Glieder durch Zusammenfassen von Gliedern der Reihe $\sum a_n$ (bei Erhaltung ihrer Reihenfolge) entstehen, und beide Reihen besitzen dieselbe Summe.*

Genauer besagt dies folgendes: Ist $0 = n_0 < n_1 < n_2 < \dots$, so ist

$$\sum_{n=0}^{\infty} a_n = \sum_{k=0}^{\infty} \left(\sum_{j=n_k}^{n_{k+1}-1} a_j \right)$$

oder, ausführlicher geschrieben,

$$a_0 + a_1 + a_2 + \dots = (a_0 + \dots + a_{n_1-1}) + (a_{n_1} + \dots + a_{n_2-1}) + \dots \quad (2)$$

Der Beweis der Behauptung ergibt sich unmittelbar daraus, daß die Partialsummen der neu gebildeten Reihe eine Teilfolge der Folge der Partialsummen der ursprünglichen Reihe bilden. In konvergenten Reihen kann man also wegen (2) nach Belieben Reihenausschnitte in Klammern einschließen. Dagegen darf man vorhandene

Klammern im allgemeinen nicht weglassen, wie das Beispiel der Reihen

$$(1 - 1) + (1 - 1) + \dots \quad \text{und} \quad 1 - 1 + 1 - 1 + \dots$$

zeigt. Die erste Reihe hat die Summe 0, die zweite ist divergent.

Satz 4. *Besitzen die Reihen mit den Gliedern a_0, a_1, a_2, \dots bzw. a'_0, a'_1, a'_2, \dots die Summe s bzw. s' , so besitzt die Reihe mit den Gliedern $a_0, a'_0, a_1, a'_1, a_2, a'_2, \dots$ die Summe $s + s'$.*

Beweis. Für die Partialsummen s_n'' der „Mischreihe“ gilt

$$s_{2k}'' = \sum_{n=0}^k a_n + \sum_{n=0}^{k-1} a'_n, \quad s_{2k+1}'' = \sum_{n=0}^k a_n + \sum_{n=0}^k a'_n,$$

woraus $s_{2k}'' \rightarrow s + s'$, $s_{2k+1}'' \rightarrow s + s'$ für $n \rightarrow \infty$ und damit $s_n'' \rightarrow s + s'$ für $n \rightarrow \infty$ folgt.

Wir formulieren nun ein wichtiges notwendiges Kriterium für die Konvergenz einer Reihe.

Satz 5. *Ist eine Reihe konvergent, so bilden ihre Glieder eine Nullfolge.*

Beweis. Sind a_n bzw. s_n die Glieder bzw. die Partialsummen der Reihe, so gilt

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_{n+1} = \lim_{n \rightarrow \infty} (s_{n+1} - s_n) = \lim_{n \rightarrow \infty} s_{n+1} - \lim_{n \rightarrow \infty} s_n = 0.$$

Nach dem hiermit bewiesenen Satz ist die Bedingung $\lim_{n \rightarrow \infty} a_n = 0$ eine notwendige Bedingung für die Konvergenz der Reihe $\sum a_n$. Sie ist aber nicht hinreichend. Dies zeigt

Beispiel 1. Die Glieder der *harmonischen Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$$

bilden eine Nullfolge. Die Reihe ist aber nicht konvergent, denn ist K eine beliebige positive reelle Zahl, ist $m > 2K$ ($m \in \mathbb{N}$) und $n_0 = 2^m$, so gilt für natürliche Zahlen $n \geq n_0$ stets

$$\begin{aligned} s_n &\geq s_{n_0} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots + \frac{1}{2^m} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots \\ &\quad + \left(\frac{1}{2^{m-1}+1} + \dots + \frac{1}{2^m}\right) \\ &> 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2^2} + 2^2 \cdot \frac{1}{2^3} + \dots + 2^{m-1} \cdot \frac{1}{2^m} = \frac{m}{2} > K, \end{aligned}$$

und hieraus folgt

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

Die Divergenz der harmonischen Reihe findet man in dem Werk „Questions super Geometriam Euclidis“ von ORESME nachgewiesen. Andere Beweise, die aber alle komplizierter sind, stammen 1650 von PIETRO MENGOLI (1625–1686) und 1689 von JOHANN und JAKOB BERNOULLI.

Beispiel 2. Wir untersuchen das Konvergenzverhalten der *geometrischen Reihe*

$$\sum_{n=0}^{\infty} q^n$$

mit $q \in \mathbb{C}$. Für $|q| \geq 1$ bilden die Glieder der Reihe keine Nullfolge, die Reihe ist divergent. Für $|q| < 1$ gilt

$$s_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q},$$

und wegen $q^n \rightarrow 0$ für $n \rightarrow \infty$ folgt

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1 - q}. \quad (3)$$

Auf Grund des Konvergenzverhaltens der geometrischen Reihe kann eine auf den ersten Blick geradezu paradox erscheinende Aussage über die Menge aller rationalen Zahlen bewiesen werden. Sie beruht auf der Tatsache, daß diese Menge abzählbar ist, daß also eine Folge (r_n) existiert, deren Wertebereich alle rationalen Zahlen enthält. Wählen wir eine beliebige positive reelle Zahl ε , so können wir für jede der Zahlen r_n die Umgebung $U_{\varepsilon^{n+1}}(r_n)$ bilden. Da die rationalen Zahlen in der Menge der reellen Zahlen dicht liegen, liegt die Vermutung nahe, daß durch diese Umgebungen alle reellen Zahlen „überdeckt“ werden. Es ist überraschend, daß dies nicht immer der Fall ist und daß es im Gegenteil sogar möglich ist, die Summe aller Durchmesser¹⁾ dieser Umgebungen kleiner als jede noch so kleine positive reelle Zahl ε' zu machen. Hieraus folgt dann, daß nicht alle reellen Zahlen überdeckt sein können.

Der Durchmesser der Umgebung von r_n ist $2\varepsilon^{n+1}$, und für die Summe aller dieser Durchmesser erhalten wir, wenn $\varepsilon < 1$ gewählt wird,

$$\sum_{n=0}^{\infty} 2\varepsilon^{n+1} = 2 \frac{\varepsilon}{1 - \varepsilon}.$$

Wählen wir ε kleiner als $\frac{\varepsilon'}{2 + \varepsilon'}$ ($\varepsilon' > 0$), so folgt

$$\sum_{n=0}^{\infty} 2\varepsilon^{n+1} = 2 \frac{\varepsilon}{1 - \varepsilon} < \varepsilon',$$

was zu beweisen war.

¹⁾ Genauer: die Summe der von ihnen gebildeten Reihe.

Während das in Satz 5 gegebene Konvergenzkriterium nur hinreichend ist, ist das *Konvergenzkriterium von Cauchy* notwendig und hinreichend. Wegen $s_{n+k} - s_n = a_{n+1} + a_{n+2} + \dots + a_{n+k}$ lautet es wie folgt:

Satz 6. *Die Reihe $\sum a_n$ ist konvergent genau dann, wenn zu jeder positiven reellen Zahl ε eine reelle Zahl $N(\varepsilon)$ angegeben werden kann derart, daß für alle natürlichen Zahlen n, k mit $k \geq 1$ und $n > N(\varepsilon)$ stets*

$$|a_{n+1} + \dots + a_{n+k}| < \varepsilon$$

gilt.

Die Divergenz der harmonischen Reihe kann mit diesem Kriterium wie folgt bewiesen werden. Für alle natürlichen Zahlen n mit $n \geq 1$ gilt

$$\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} \geq n \cdot \frac{1}{2n} = \frac{1}{2}.$$

Daher kann die Reihe nicht konvergieren.

Eine Übertragung des Konvergenzkriteriums für monotone Folgen führt zu dem

Satz 7. *Eine Reihe mit nichtnegativen reellen Gliedern ist genau dann konvergent, wenn die Folge ihrer Partialsummen beschränkt ist.*

Beweis. Die Folge der Partialsummen ist monoton wachsend.

Abschließend bemerken wir, daß eine Reihe offenbar genau dann konvergiert, wenn einer (und dann auch jeder) ihre Reste konvergent ist, und in diesem Fall gilt

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^m a_n + \sum_{n=m+1}^{\infty} a_n. \quad (4)$$

Läßt man in einer Reihe endlich viele Glieder fort, so bleibt zwar das Konvergenzverhalten — im allgemeinen aber nicht die Summe der Reihe — unverändert. Das Konvergenzverhalten und die Summe bleiben jedoch erhalten, wenn man in die Folge der Glieder der Reihe beliebig viele (aber natürlich zwischen je zwei aufeinanderfolgende Glieder höchstens endlich viele) Nullen einfügt oder beliebig viele Glieder a_n mit $a_n = 0$ wegläßt.

Mit Ausnahme von Satz 6 und Satz 7 können alle Definitionen und Sätze auf Reihen übertragen werden, deren Glieder Elemente eines normierten Raumes E sind. Satz 6 gilt auch in vollständigen normierten Räumen.

Bis zu CAUCHY gab es nur wenige Versuche, wirklich exakt die Konvergenz von Reihen zu bestimmen. Das Cauchysche Konvergenzkriterium kann, freilich in völlig anderer Form, bereits bei EULER festgestellt werden. GAUSS gab 1812 mit seiner Untersuchung über die hypergeometrische Reihe die erste systematische Untersuchung der Konvergenz von Reihen.

In CAUCHYS berühmten Lehrbüchern, etwa „Cours d'Analyse“ von 1821, findet man zusammenfassend, auf dem Grenzwertbegriff von D'ALEMBERT basierend, den noch heute üblichen Aufbau der Infinitesimalrechnung. Dort sind auch alle einfachen Konvergenzkriterien vorhanden.

2.2.3. Alternierende Reihen

Eine Reihe mit reellen Gliedern heißt *alternierend*, wenn je zwei aufeinanderfolgende Glieder verschiedene Vorzeichen haben. Für solche Reihen gibt es ein einfaches, von LEIBNIZ stammendes *hinreichendes Konvergenzkriterium* (Brief an JOHANN BERNOULLI vom 10. 1. 1714).

Satz 1. Eine alternierende Reihe konvergiert, wenn die Absolutbeträge ihrer Glieder eine monotone Nullfolge bilden.

Beweis. In der alternierenden Reihe $\sum a_n$ sei o. B. d. A. $a_0 > 0$. Für alle natürlichen Zahlen n ist dann $a_{2n} > 0$, $a_{2n+1} < 0$. Nach Voraussetzung ist $a_0 \geq |a_1| \geq a_2 \geq |a_3| \geq \dots$. Wir setzen

$$b_n := s_{2n+1}, \quad c_n := s_{2n},$$

wobei s_n die Partialsummen der gegebenen Reihe sind. Wir zeigen, daß $(b_n | c_n)$ eine Intervallschachtelung ist (Abb. 2.13). Für alle natürlichen Zahlen n ist

$$\begin{aligned} b_{n+1} - b_n &= (s_{2n+3} - s_{2n+2}) - (s_{2n+2} - s_{2n+1}) \\ &= a_{2n+3} + a_{2n+2} = a_{2n+2} - |a_{2n+3}| \geq 0, \end{aligned}$$

$$\begin{aligned} c_{n+1} - c_n &= (s_{2n+2} - s_{2n+1}) + (s_{2n+1} - s_{2n}) \\ &= a_{2n+2} + a_{2n+1} = a_{2n+2} - |a_{2n+1}| \leq 0, \end{aligned}$$

$$c_n - b_n = s_{2n} - s_{2n+1} = |a_{2n+1}| > 0,$$

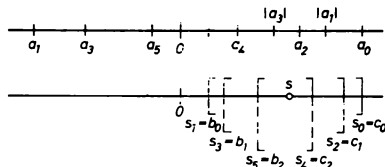


Abb. 2.13

und da $(|a_n|)$ eine Nullfolge ist, gilt dasselbe für ihre Teilfolge $(|a_{2n+1}|)$, d. h., $(c_n - b_n)$ ist eine Nullfolge. Die Folgen (s_{2n}) , (s_{2n+1}) besitzen somit denselben Grenzwert s , und s ist auch der Grenzwert ihrer Mischfolge (s_n) , d. h., es ist $s = \sum a_n$, und der Satz ist bewiesen.

Auf Grund dieses Satzes erkennen wir unmittelbar, daß die Reihen

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - + \dots,$$

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - + \dots$$

konvergent sind. Wie wir später sehen werden, sind ihre Summen gleich $\ln 2$ bzw. $\frac{\pi}{4}$.

Für die Konvergenz einer Reihe ist nach 2.2.2., Satz 5, notwendig, daß die Glieder, also auch die Absolutbeträge der Glieder, eine Nullfolge bilden. Diese Bedingung ist aber auch bei alternierenden Reihen nicht hinreichend. Die Bedeutung der im Satz geforderten Monotonie zeigt das folgende Beispiel. Andererseits ist die Bedingung der Monotonie der Nullfolge $(|a_n|)$ zwar hinreichend, aber nicht notwendig. Es gibt konvergente alternierende Reihen, für die $(|a_n|)$ keine monotone Nullfolge ist.

Beispiel. Die Reihe $\sum a_n$ sei durch

$$a_{2n} := \frac{1}{\sqrt{n+2}-1}, \quad a_{2n+1} := -\frac{1}{\sqrt{n+2}+1}$$

definiert. Nehmen wir an, die Reihe sei konvergent. Nach 2.2.2., Satz 3, können wir je zwei aufeinanderfolgende Glieder dieser Reihe zusammenfassen, d. h., es ist

$$\sum_{n=0}^{\infty} a_n = \sum_{k=0}^{\infty} (a_{2k} + a_{2k+1}).$$

Wegen

$$a_{2n} + a_{2n+1} = \frac{\sqrt{n+2}+1 - \sqrt{n+2}-1}{(\sqrt{n+2}-1)(\sqrt{n+2}+1)} = \frac{2}{n+1}$$

ist daher $\sum a_n = \sum \frac{2}{n+1}$. Dies ist ein Widerspruch, denn die harmonische Reihe ist divergent. Die betrachtete alternierende Reihe ist somit nicht konvergent, obwohl ihre Glieder eine Nullfolge (aber eben keine monotone Nullfolge) bilden.

Ein wichtiges Problem der numerischen Mathematik besteht darin, den Fehler abzuschätzen, den man begeht, wenn man die Summe einer Reihe durch eine ihrer Partialsummen approximiert (annähert). Dies ist in vielen Fällen ein sehr schwieriges Problem. Dagegen ist die Fehlerabschätzung bei alternierenden Reihen, die dem Leibnizschen Konvergenzkriterium genügen, sehr einfach. Dies zeigt

Satz 2. Die n -te Partialsumme einer alternierenden Reihe, deren Glieder dem Betrag nach eine monotone Nullfolge bilden, unterscheidet sich von der Summe der Reihe höchstens um den Absolutbetrag des $(n+1)$ -ten Gliedes der Reihe.

Beweis. Mit den Bezeichnungen des Beweises von Satz 1 ist im Fall $a_0 > 0$ stets $s_{2n+1} = b_n \leq s \leq c_n \leq s_{2n}$, und folglich gilt

$$\begin{aligned} |s_{2n} - s| &= s_{2n} - s \leq s_{2n} - s_{2n+1} = |a_{2n+1}|, \\ |s_{2n+1} - s| &= s - s_{2n+1} \leq s_{2n+2} - s_{2n+1} = a_{2n+2}, \end{aligned}$$

d. h., es ist

$$|s_n - s| \leq |a_{n+1}|. \quad (1)$$

2.2.4. Absolute Konvergenz

Aus jeder gegebenen Reihe können wir eine neue Reihe bilden, deren Glieder die Absolutbeträge der Glieder der gegebenen Reihe sind. Für Reihen mit nichtnegativen Gliedern stimmt die so gebildete Reihe natürlich mit der ursprünglichen Reihe überein. Wir vereinbaren die

Definition. Die Reihe $\sum a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum |a_n|$ konvergent ist.

Auch die Bedeutung dieses Konvergenzbegriffs wurde von CAUCHY erkannt. Für Reihen mit nichtnegativen Gliedern gilt $|a_n| = a_n$, d. h., eine Reihe mit nichtnegativen Gliedern ist genau dann absolut konvergent, wenn sie konvergent ist. Für Reihen mit beliebigen Gliedern ist die absolute Konvergenz nur eine hinreichende Bedingung für die Konvergenz. Dies besagt

Satz 1. Ist die Reihe $\sum |a_n|$ konvergent, so konvergiert auch die Reihe $\sum a_n$.

Beweis. Es ist stets

$$\begin{aligned} |a_{n+1} + a_{n+2} + \dots + a_{n+k}| &\leq |a_{n+1}| + |a_{n+2}| + \dots + |a_{n+k}| \\ &= ||a_{n+1}| + |a_{n+2}| + \dots + |a_{n+k}||. \end{aligned}$$

Der Betrag eines Ausschnittes der Reihe ist also niemals größer als der Betrag des entsprechenden Ausschnittes der Reihe der Absolutbeträge. Die Behauptung ergibt sich nun unmittelbar aus dem Cauchyschen Konvergenzkriterium.

Ein einfaches Kriterium für die absolute Konvergenz gibt

Satz 2. Eine Reihe ist absolut konvergent genau dann, wenn die Partialsummen der Reihe der Absolutbeträge eine beschränkte Folge bilden.

Der Beweis folgt sofort aus 2.2.2., Satz 7.

Für die absolute Konvergenz leiten wir zwei wichtige Vergleichskriterien her. Das sogenannte *erste Majorantenkriterium* lautet:

Satz 3. Existiert zu einer Reihe $\sum a_n$ eine konvergente Reihe mit reellen Gliedern c_n derart, daß von einer Stelle p an stets

$$|a_n| \leq c_n \quad (n \geq p) \tag{1}$$

gilt, so ist die Reihe $\sum a_n$ absolut konvergent. Existiert eine divergente Reihe $\sum d_n$ mit reellen Gliedern derart, daß von einer Stelle p an stets

$$|a_n| \geq |d_n| \quad (n \geq p) \tag{2}$$

gilt, so ist die Reihe $\sum a_n$ nicht absolut konvergent.

Beweis. Die Folge der Partialsummen der Reihe $\sum a_n$ ist im ersten Fall wegen

$$|a_p| + |a_{p+1}| + \dots + |a_{p+k}| \leq c_p + c_{p+1} + \dots + c_{p+k}$$

beschränkt, im zweiten Fall wegen

$$|a_p| + |a_{p+1}| + \cdots + |a_{p+k}| \geq |d_p| + |d_{p+1}| + \cdots + |d_{p+k}|$$

nicht beschränkt, denn wegen Satz 1 ist $\sum |d_n|$ nicht konvergent. Die Behauptungen folgen somit aus Satz 2.

Wir formulieren nun das *zweite Majorantenkriterium*.

Satz 4. *Existiert zu einer Reihe $\sum a_n$ mit $a_n \neq 0$ ($n \in \mathbf{N}$) eine konvergente Reihe mit positiven reellen Gliedern c_n derart, daß von einer Stelle p an stets*

$$\frac{|a_{n+1}|}{|a_n|} \leq \frac{c_{n+1}}{c_n} \quad (n \geq p) \quad (3)$$

gilt, so ist die Reihe $\sum a_n$ absolut konvergent. Existiert eine divergente Reihe mit Gliedern $d_n \neq 0$ derart, daß von einer Stelle p an stets

$$\frac{|a_{n+1}|}{|a_n|} \geq \frac{|d_{n+1}|}{|d_n|} \quad (n \geq p) \quad (4)$$

gilt, so ist die Reihe $\sum a_n$ nicht absolut konvergent.

Beweis. Für die Folgen (d_n) , (a_n) , (c_n) gelte

$$\begin{aligned} \frac{|d_{p+1}|}{|d_p|} &\leq \frac{|a_{p+1}|}{|a_p|} \leq \frac{c_{p+1}}{c_p}, \\ \frac{|d_{p+2}|}{|d_{p+1}|} &\leq \frac{|a_{p+2}|}{|a_{p+1}|} \leq \frac{c_{p+2}}{c_{p+1}}, \\ &\dots\dots\dots \\ \frac{|d_{p+k}|}{|d_{p+k-1}|} &\leq \frac{|a_{p+k}|}{|a_{p+k-1}|} \leq \frac{c_{p+k}}{c_{p+k-1}}. \end{aligned}$$

Multiplikation ergibt

$$\frac{|d_{p+k}|}{|d_p|} \leq \frac{|a_{p+k}|}{|a_p|} \leq \frac{c_{p+k}}{c_p},$$

d. h., für $n = p + k > p$ ist

$$\frac{|a_p|}{|d_p|} \cdot |d_n| \leq |a_n| \leq \frac{|a_p|}{c_p} c_n. \quad (5)$$

Ist die Reihe $\sum c_n$ konvergent, so gilt dies auch für

$$\sum_{n=0}^{\infty} \frac{|a_p|}{c_p} c_n = \frac{|a_p|}{c_p} \sum_{n=0}^{\infty} c_n.$$

Nach dem ersten Majorantenkriterium ist also $\sum a_n$ absolut konvergent. Ist die Reihe $\sum |d_n|$ divergent, so gilt dies auch für

$$\sum_{n=0}^{\infty} \frac{|a_p|}{|d_p|} |d_n| = \frac{|a_p|}{|d_p|} \sum_{n=0}^{\infty} |d_n|.$$

Aus dem ersten Majorantenkriterium und aus (5) folgt die Divergenz der Reihe $\sum |a_n|$.

Wir leiten aus jedem der beiden Majorantenkriterien einen für die Anwendungen besonders wichtigen Spezialfall ab. Diese Spezialfälle stützen sich auf das Konvergenzverhalten der geometrischen Reihe. Wählen wir in Satz 3 die geometrische Reihe als Vergleichsreihe, so erhalten wir das sogenannte *Wurzelkriterium* (CAUCHY 1821).

Satz 5. Die Reihe $\sum a_n$ ist absolut konvergent, wenn es eine reelle Zahl q gibt derart, daß von einer Stelle p an stets

$$\sqrt[n]{|a_n|} \leq q < 1 \quad (n \geq p) \quad (6)$$

ist. Sie ist divergent, wenn für unendlich viele natürliche Zahlen n die Ungleichung

$$\sqrt[n]{|a_n|} \geq 1 \quad (7)$$

erfüllt ist.

Beweis. Aus (6) folgt $|a_n| \leq q^n$, d. h., es gilt (1) mit $c_n := q^n$. Nach Satz 3 ergibt sich die erste Behauptung. Gilt dagegen (7) für unendlich viele n , so bilden die Glieder der Reihe keine Nullfolge, und die Reihe ist divergent.

Wir bemerken, daß die in Satz 5 angegebenen beiden Fälle nicht alle Möglichkeiten ausschöpfen und daß die angegebenen Bedingungen hinreichend, aber nicht notwendig sind.

Beispiel 1. Für die Reihe mit den Gliedern

$$a_n := \left(\frac{1}{\ln(n+3)} \right)^n$$

gilt

$$\sqrt[n]{|a_n|} = \frac{1}{\ln(n+3)} \leq \frac{1}{\ln 3} < \frac{1}{\ln e} = 1.$$

Nach Satz 5 ist die Reihe konvergent.

Beispiel 2. Für die Reihe $\sum a_n$ mit den Gliedern

$$a_n := \frac{1}{n^r} \quad (r \geq 1)$$

gilt

$$\sqrt[n]{|a_n|} = \frac{1}{(\sqrt[n]{n})^r} < 1 \quad (n > 1).$$

Es kann aber keine Zahl q angegeben werden, für die (6) erfüllt ist. Daher ist Satz 5 nicht anwendbar.

Für $r = 1$ ergibt sich die divergente harmonische Reihe. Um zu zeigen, daß die Reihe für $r > 1$ stets konvergent ist, ermitteln wir für die Partialsummen der Reihe eine Schranke. Ist n eine beliebige positive natürliche Zahl, dann wählen wir eine positive natürliche Zahl k mit $n \leq 2^{k+1} - 1$, und es gilt

$$\begin{aligned} s_n &\leq s_{2^{k+1}-1} = 1 + \left(\frac{1}{2^r} + \frac{1}{3^r}\right) + \left(\frac{1}{4^r} + \cdots + \frac{1}{7^r}\right) + \cdots \\ &\quad + \left(\frac{1}{(2^k)^r} + \cdots + \frac{1}{(2^{k+1}-1)^r}\right) \\ &< 1 + 2 \frac{1}{2^r} + 2^2 \frac{1}{(2^2)^r} + \cdots + 2^k \frac{1}{(2^k)^r} \\ &= 1 + \frac{1}{(2^1)^{r-1}} + \frac{1}{(2^2)^{r-1}} + \cdots + \frac{1}{(2^k)^{r-1}} \\ &< \sum_{m=0}^{\infty} \left(\frac{1}{2^{r-1}}\right)^m = \frac{1}{1 - 2^{1-r}}. \end{aligned}$$

Damit ist die Behauptung bewiesen.

Als Folgerung aus Satz 5 ergibt sich

Satz 6. Die Reihe $\sum a_n$ ist absolut konvergent, wenn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1, \quad (8)$$

und divergent, wenn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1 \quad (9)$$

ist.

Beweis. Wir setzen

$$K := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}. \quad (10)$$

Ist $K < 1$, so gibt es ein q mit $K < q < 1$, und nach Definition des Limes superior gibt es höchstens endlich viele natürliche Zahlen n mit $\sqrt[n]{|a_n|} > q$. Die absolute Konvergenz folgt somit aus Satz 5. Ist $K > 1$, so gibt es unendlich viele natürliche Zahlen n mit $\sqrt[n]{|a_n|} \geq 1$, und die Divergenz folgt wiederum aus Satz 5. Damit ist Satz 6 bewiesen.

Gilt dagegen in (8), (9) das Gleichheitszeichen, so können keine Schlüsse auf das Konvergenzverhalten gezogen werden. So sind z. B. die Reihen mit den Gliedern $\frac{1}{n^2}$ bzw. $(-1)^n \frac{1}{n}$ bzw. $\frac{1}{n}$ absolut konvergent bzw. konvergent und nicht absolut konvergent bzw. divergent, und für alle drei Reihen gilt $K = 1$.

Wählen wir in Satz 4 die geometrische Reihe als Vergleichsreihe, so erhalten wir das sogenannte *Quotientenkriterium* (CAUCHY 1821).

Satz 7. Die Reihe $\sum a_n$ ist absolut konvergent, wenn es eine reelle Zahl q gibt derart, daß von einer Stelle p an stets

$$\frac{|a_{n+1}|}{|a_n|} \leq q < 1 \quad (n \geq p) \quad (11)$$

ist. Sie ist divergent, wenn von einer Stelle p an stets

$$\frac{|a_{n+1}|}{|a_n|} \geq 1 \quad (n \geq p) \quad (12)$$

ist.¹⁾

Beweis. Aus (11) folgt

$$\frac{|a_{n+1}|}{|a_n|} \leq q = \frac{q^{n+1}}{q^n},$$

und die erste Behauptung folgt aus Satz 4 mit $c_n = q^n$. Gilt dagegen (12), so ist $|a_{n+1}| \geq |a_n| > 0$ für $n \geq p$, und die Glieder der Reihe bilden keine Nullfolge.

Als Folgerung aus Satz 7 ergibt sich

Satz 8. Konvergiert die Folge mit den Gliedern $\frac{|a_{n+1}|}{|a_n|}$, so ist die Reihe $\sum a_n$ absolut konvergent, wenn

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} < 1,$$

und divergent, wenn

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} > 1$$

ist.

Beweis. Der Grenzwert sei gleich K . Ist $K < 1$, so gibt es ein q mit $K < q < 1$, und von einer Stelle p an ist die Voraussetzung (11) von Satz 7 erfüllt. Ist dagegen $K > 1$, so gilt (12), und Satz 8 ist bewiesen.

Ist der Grenzwert gleich 1, so können wiederum keine Schlüsse auf das Konvergenzverhalten gezogen werden.

¹⁾ Offenbar können (11), (12) nur gelten, wenn $a_n \neq 0$ für $n \geq p$ ist.

Beispiel 3. Wir betrachten die (für $x = 0$ stets konvergenten) Reihen

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots, \quad (13)$$

$$\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = \binom{\alpha}{0} + \binom{\alpha}{1} x + \binom{\alpha}{2} x^2 + \dots \quad (\alpha \in \mathbb{R}), \quad (14)$$

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots, \quad (15)$$

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots, \quad (16)$$

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots, \quad (17)$$

$$\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots, \quad (18)$$

$$\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots \quad (19)$$

und bestimmen von 0 verschiedene reelle (oder komplexe) Zahlen x , für welche diese Reihen konvergieren bzw. divergieren. Wir bezeichnen die Glieder der Reihen jeweils mit a_n . In (13) gilt

$$\frac{|a_{n+1}|}{|a_n|} = |x| \cdot \frac{n}{n+1}.$$

Ist $|x| < 1$, so ist (11) mit $q := |x|$ stets erfüllt, die Reihe ist für alle reellen (oder komplexen) Zahlen x mit $|x| < 1$ absolut konvergent. Ist $|x| > 1$, so gilt von einer Stelle p an stets (12), die Reihe ist divergent. Für $x = 1$ ist die Reihe offenbar divergent, für $x = -1$ konvergent.

Ist α in (14) eine natürliche Zahl, so sind höchstens endlich viele Glieder der Reihe von 0 verschieden, und die Reihe konvergiert absolut für alle reellen Zahlen x . In allen anderen Fällen gilt

$$\frac{|a_{n+1}|}{|a_n|} = |x| \frac{|\alpha - n|}{n+1}.$$

Da der rechts stehende Bruch für $n \rightarrow \infty$ gegen 1 strebt, ergibt sich wiederum, daß die Reihe für $|x| < 1$ absolut konvergiert und für $|x| > 1$ divergiert.

Ohne Beweis vermerken wir, daß die Reihe, falls α positiv ist, auch für $x = -1$ konvergiert und, falls $\alpha > -1$ ist, auch für $x = 1$ konvergiert.

In (15) gilt

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|x|}{n+1},$$

und da rechts für alle reellen Zahlen x die Glieder einer Nullfolge stehen, ist (11) stets erfüllt. Die Reihe (15) ist für alle Zahlen x absolut konvergent. Da die Glieder einer konvergenten Reihe stets eine Nullfolge bilden, haben wir damit zugleich bewiesen, daß die Zahlenfolge $\left(\frac{x^n}{n!}\right)$ für jede reelle Zahl x eine Nullfolge ist. Dies bedeutet, daß die Folge $(n!)$ wesentlich stärker gegen ∞ strebt als die Folge (x^n) , obwohl für große x die ersten Folgenglieder x^n erheblich größer als die entsprechenden Folgenglieder $n!$ sind.

Das Konvergenzverhalten der Reihen (16) bis (19) könnten wir ebenfalls mit Hilfe von Satz 7 bestimmen. Betrachten wir jedoch die Folgen der Absolutbeträge der Glieder der Reihen (15) bis (19), so erkennen wir, daß die zu den Reihen (16) bis (19) gebildeten Folgen Teilfolgen der entsprechenden Folge (15) sind. Daher sind auch die Reihen (16) bis (19) für alle reellen (oder komplexen) Zahlen x absolut konvergent.

Die Reihen (13) bis (19) sind erste Beispiele für sogenannte *Potenzreihen*, die an späterer Stelle systematisch untersucht werden.

Alle Definitionen und Sätze dieses Abschnittes können unverändert auf Reihen in vollständigen normierten Räumen z. B. im Raum \mathbb{R}^p übertragen werden.

2.2.5. Bedingte und unbedingte Konvergenz

Wir untersuchen in diesem Abschnitt den Einfluß einer Umordnung der Glieder einer Reihe auf ihre Summe bzw. auf ihr Konvergenzverhalten. Dabei nennen wir eine Reihe $\sum b_n$ eine *Umordnung* der Reihe $\sum a_n$, wenn eine umkehrbare eindeutige Funktion φ von \mathbb{N} auf \mathbb{N} existiert derart, daß $b_n = a_{\varphi(n)}$ ist. Bilden wir z. B. zur Reihe $\sum a_n$ die Reihe

$$\begin{aligned} \sum b_n &= a_0 + a_2 + a_1 + \\ &\quad a_4 + a_6 + a_3 + \\ &\quad \dots\dots\dots + \\ &\quad a_{4k} + a_{4k+2} + a_{2k+1} + \dots, \end{aligned}$$

d. h., setzen wir

$$\left. \begin{aligned} \varphi(3k) &:= 4k, \\ \varphi(3k+1) &:= 4k+2, \\ \varphi(3k+2) &:= 2k+1 \end{aligned} \right\} \quad (k \in \mathbb{N})$$

und $b_n := a_{\varphi(n)}$, so ist $\sum b_n$ eine Umordnung der Reihe $\sum a_n$. Diese Umordnung wurde in den Reihen 2.2.1.(4) bzw. (5) betrachtet, und die dortige Annahme, daß jede Umordnung auf den Summenwert ohne Einfluß bleibt, hat zu einem Widerspruch geführt. Es gibt also in der Tat Reihen, deren Konvergenzverhalten sich bei einer Umordnung der Glieder ändert.

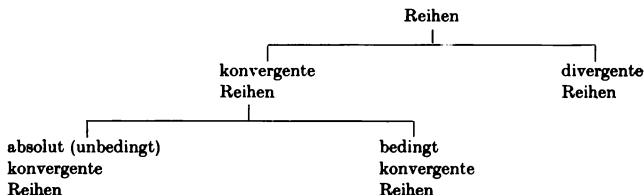
Definition. Eine Reihe heißt *bedingt konvergent*, wenn sie konvergent ist und wenn es eine Umordnung der Reihe gibt, die einen anderen Summenwert besitzt

oder nicht konvergent ist. Sie heißt *unbedingt konvergent*, wenn sie konvergent ist und wenn der (endliche) Summenwert bei jeder Umordnung unverändert bleibt.

Der folgende Satz stellt einen überraschenden Zusammenhang zwischen den Begriffen „unbedingt konvergent“ und „absolut konvergent“ her. Er geht auf DIRICHLET zurück und wird gelegentlich der *kleine Umordnungssatz* genannt.

Satz. Eine Reihe ist absolut konvergent genau dann, wenn sie unbedingt konvergent ist.

Nach diesem Satz können wir Reihen wie folgt klassifizieren:



Wir führen nun den

Beweis. Wir setzen zuerst voraus, daß die Reihe $\sum a_n$ absolut konvergent ist. Ferner sei φ eine umkehrbar eindeutige Funktion von \mathbf{N} auf \mathbf{N} und $b_n = a_{\varphi(n)}$. Zu jeder positiven reellen Zahl ε gibt es dann wegen der absoluten Konvergenz der Reihe $\sum a_n$ eine natürliche Zahl p mit

$$\sum_{j=p+1}^{p+q+1} |a_j| < \varepsilon \quad (q \in \mathbf{N}). \quad (1)$$

Wir bestimmen eine natürliche Zahl N so, daß

$$\{0, 1, \dots, p\} \subseteq \{\varphi(0), \varphi(1), \dots, \varphi(n)\}$$

für alle n mit $n \geq N$ ist. Offenbar ist dann $N \geq p$. Das Maximum der Zahlen $\varphi(0), \varphi(1), \dots, \varphi(n)$ sei n^* . Dann ist $n \leq n^*$, und für alle n mit $n \geq N$ gilt

$$\left| \sum_{j=0}^n b_j - \sum_{j=0}^n a_j \right| = \left| \sum_{j=0}^n a_{\varphi(j)} - \sum_{j=0}^n a_j \right| \leq |a_{p+1}| + \dots + |a_{n^*}| < \varepsilon,$$

denn nach Definition von N treten alle Summanden a_j mit $j \leq p$ auch unter den Zahlen $a_{\varphi(0)}, a_{\varphi(1)}, \dots, a_{\varphi(n)}$ auf, und für die verbleibenden Summanden a_j gilt $j \geq p+1$, so daß die Abschätzung (1) angewendet werden kann. Daher ist

$$\lim_{n \rightarrow \infty} \left| \sum_{j=0}^n b_j - \sum_{j=0}^n a_j \right| = 0,$$

Woraus

$$\sum_{j=0}^{\infty} b_j = \sum_{j=0}^{\infty} a_j$$

folgt. Somit ist jede absolut konvergente Reihe auch unbedingt konvergent. Daß eine unbedingt konvergente Reihe auch absolut konvergent ist, zeigen wir durch den Beweis der Kontraposition dieser Aussage. Die Reihe $\sum a_n$ sei nicht absolut konvergent, d. h., die Reihe $\sum |a_n|$ sei divergent.¹⁾ Ist $\sum a_n$ nicht konvergent, so ist diese Reihe erst recht nicht unbedingt konvergent. Ist sie konvergent, so können wir o. B. d. A. voraussetzen, daß alle Glieder der Reihe von 0 verschieden sind. Wir setzen

$$c_n := \frac{|a_n| + a_n}{2} = \begin{cases} a_n & \text{für } a_n > 0, \\ 0 & \text{für } a_n < 0, \end{cases} \quad (2)$$

$$d_n := \frac{|a_n| - a_n}{2} = \begin{cases} -a_n & \text{für } a_n < 0, \\ 0 & \text{für } a_n > 0. \end{cases}$$

Dann ist stets $c_n, d_n \geq 0$ und $|a_n| = c_n + d_n$, $a_n = c_n - d_n$. Nehmen wir an, eine der Reihen $\sum c_n$, $\sum d_n$ wäre konvergent. Wegen $c_n = a_n + d_n$ bzw. $d_n = c_n - a_n$ und der vorausgesetzten Konvergenz der Reihe $\sum a_n$ müßte dann auch die zweite der Reihen konvergent sein. Dann wäre aber auch die Reihe $\sum |a_n| = \sum c_n + \sum d_n$ im Widerspruch zu unserer Voraussetzung konvergent. Somit sind die Reihen $\sum c_n$, $\sum d_n$ divergent, und zwar als Reihen mit nichtnegativen Gliedern bestimmt divergent. Wegen (2) und der Konvergenz der Reihe $\sum a_n$ bilden ihre Glieder jedoch Nullfolgen.

Wir skizzieren zunächst den weiteren Beweisgedanken. Wir wählen eine beliebige reelle Zahl x und setzen die Glieder der Folgen (c_n) und $(-d_n)$ in der folgenden Weise zu einer neuen Folge (b_n) zusammen. Zunächst wählen wir so viele Glieder der Folge (c_n) , bis deren Summe erstmals die Zahl x übertrifft. Dann fügen wir so viele Glieder der Folge $(-d_n)$ an, bis die Summe aller so ausgewählten Glieder die Zahl x erstmals unterschreitet. Nun addieren wir wieder so viele Glieder der Folge (c_n) , bis x übertrifft wird, und so fort. Die Partialsummen s_n der in dieser Weise gebildeten Folge (b_n) pendeln also stets über die reelle Zahl x hinweg. Da die Folgen (c_n) , $(-d_n)$ Nullfolgen sind, ist auch die Folge $(s_n - x)$ eine Nullfolge. Die Reihe $\sum b_n$ konvergiert gegen die beliebig gewählte reelle Zahl x . In der Folge (b_n) sind unendlich viele Glieder von null verschieden. Wir betrachten die Teilfolge $(b_{n'})$, die nur aus den von null verschiedenen Gliedern der Folge (b_n) besteht. Wegen (2) ist dann die Reihe $\sum b_{n'}$ eine Umordnung der Reihe $\sum a_n$, und wegen $\sum b_{n'} = \sum b_n = x$ ist die Reihe $\sum a_n$ nur bedingt konvergent.²⁾

¹⁾ Im Fall einer Reihe mit komplexen Gliedern a_n bzw. einer Reihe mit Gliedern $a_n \in \mathbb{R}^p$ folgt dann aus 2.1.2.(5), daß wenigstens eine der Reihen $\sum |\alpha_{jn}|$ ($j = 1, \dots, p$) divergent sein muß. Die nachfolgenden Schlüsse können dann auf die Reihe $\sum \alpha_{jn}$ angewendet werden, so daß der Beweis auch für diese Reihe gültig bleibt.

²⁾ Über die Aussage des Satzes hinaus folgt aus dem Beweis: Zu jeder reellen Zahl x und zu jeder bedingt konvergenten Reihe gibt es eine Umordnung dieser Reihe, deren Summe gleich x ist. Dies ist eine Teilaussage des sogenannten *Riemannschen Umordnungssatzes*.

Zur Ausführung des Beweisgedankens definieren wir induktiv vier Zahlenfolgen (x_n) , (y_n) , (p_n) , (q_n) . Mit p_0 bzw. q_0 bezeichnen wir die kleinste natürliche Zahl mit der Eigenschaft

$$c_0 + c_1 + \dots + c_{p_0} > x \quad \text{bzw.} \quad c_0 + c_1 + \dots + c_{p_0} - d_0 - d_1 - \dots - d_{q_0} < x$$

und setzen

$$x_0 := \sum_{j=0}^{p_0} c_j, \quad y_0 := x_0 - \sum_{j=0}^{q_0} d_j.$$

Nach Definition von p_0 , q_0 ist dann

$$x_0 - c_{p_0} \leq x < x_0, \quad y_0 < x \leq y_0 + d_{q_0}.$$

Für eine natürliche Zahl n seien p_n , q_n , x_n , y_n schon definiert, und es gelte

$$x_n - c_{p_n} \leq x < x_n, \quad y_n < x \leq y_n + d_{q_n}.$$

Mit p_{n+1} bzw. q_{n+1} bezeichnen wir die kleinste natürliche Zahl¹⁾ mit der Eigenschaft

$$x_n + \sum_{j=p_n+1}^{p_{n+1}} c_j > x \quad \text{bzw.} \quad y_n + \sum_{j=p_n+1}^{p_{n+1}} c_j - \sum_{k=q_n+1}^{q_{n+1}} d_k < x$$

und setzen

$$x_{n+1} := y_n + \sum_{j=p_n+1}^{p_{n+1}} c_j, \quad y_{n+1} := x_{n+1} - \sum_{j=q_n+1}^{q_{n+1}} d_j.$$

Nach Definition von p_{n+1} , q_{n+1} ist dann

$$x_{n+1} - c_{p_{n+1}} \leq x < x_{n+1}, \quad y_{n+1} < x \leq y_{n+1} + d_{q_{n+1}}. \quad (3)$$

Damit sind die Folgen (p_n) , (q_n) , (x_n) , (y_n) induktiv definiert. Wegen (3) und $c_n, d_n \rightarrow 0$ für $n \rightarrow \infty$ gilt $x_n \rightarrow x$, $y_n \rightarrow x$ für $n \rightarrow \infty$. Wir bilden die Reihe $\sum b_n$, deren Glieder durch die Folge

$$(c_0, \dots, c_{p_0}, -d_0, \dots, -d_{q_0}, c_{p_0+1}, \dots, c_{p_1}, -d_{q_0+1}, \dots, -d_{q_1}, \dots)$$

gegeben sind. Die Folgen (x_j) , (y_j) sind konvergente Teilfolgen der Partialsummenfolge (s_n) dieser Reihe. Zu jeder natürlichen Zahl n mit $n > p_0 + q_0 + 1$ gibt es genau eine natürliche Zahl j mit

$$p_j + q_j + 1 < n \leq p_{j+1} + q_j + 1$$

oder

$$p_{j+1} + q_j + 1 < n \leq p_{j+1} + q_{j+1} + 1.$$

Ferner gilt $j \rightarrow \infty$ für $n \rightarrow \infty$. Nach unserer Konstruktion ist stets $y_j \leq s_n \leq x_{j+1}$ bzw. $y_j \leq s_n \leq x_j$. Daher konvergiert auch die Folge (s_n) gegen x , es ist $\sum b_n = x$. Die bereits oben definierte Reihe $\sum b_n'$ ist dann eine Umordnung der Reihe $\sum a_n$, und unser Satz ist bewiesen.

2.2.6. Großer Umordnungssatz

In vielen Anwendungen der Analysis treten sogenannte *Doppelreihen* der Form

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn}$$

¹⁾ Die Existenz dieser Zahlen folgt aus der Divergenz der Reihen $\sum c_n$, $\sum d_n$.

auf. Hierbei handelt es sich um die folgende Problemstellung. Gegeben ist eine *Doppelfolge* $(a_{mn})_{m,n \in \mathbb{N}}$. Für jede natürliche Zahl m soll die Konvergenz der Reihe

$$\sum_{n=0}^{\infty} a_{mn}$$

untersucht werden, und wenn diese Reihe stets eine endliche Summe Z_m besitzt, soll auch geprüft werden, ob die Reihe mit den Gliedern Z_m konvergiert. Ist dies der Fall, so setzt man

$$\sum_{m=0}^{\infty} Z_m := \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn}.$$

Oft ist auch die Frage zu entscheiden, ob die Reihenfolge der beiden Grenzübergänge vertauscht werden darf, d. h., ob

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{mn}$$

ist. Schließlich ist es oft zweckmäßig, für die Indexpaare (m, n) eine „lineare“ Anordnung $((m_1, n_1), (m_2, n_2), \dots, (m_j, n_j), \dots)$, d. h. eine umkehrbar eindeutige Funktion φ von \mathbb{N} auf $\mathbb{N} \times \mathbb{N}$ mit $\varphi(j) = (m_j, n_j)$ vorzugeben und zu untersuchen, ob auch die „einfache“ Reihe

$$\sum_{j=0}^{\infty} a_{m_j, n_j}$$

denselben Grenzwert wie die beiden oben genannten Doppelreihen besitzt.

Die beiden wichtigen Spezialfälle von linearen Anordnungen sind die *Anordnung nach Quadraten* und die *Anordnung nach Diagonalen*, die aus den folgenden beiden Schemata entnommen werden können:

$$\begin{array}{ccccccc} a_{00} & a_{01} & a_{02} & \dots & . & . & . \\ & \uparrow & \uparrow & \uparrow & & & \\ a_{10} \rightarrow & a_{11} & a_{12} & \dots & . & . & . \\ & & \uparrow & \uparrow & & & \\ a_{20} \rightarrow & a_{21} \rightarrow & a_{22} & \dots & . & . & . \\ & & & \uparrow & & & \\ . & \rightarrow & . & \rightarrow & . & \rightarrow & . \\ . & . & . & . & . & . & . \end{array}$$

$$\begin{array}{ccccccc} a_{00} & a_{01} & a_{02} & a_{03} & & & \\ & \nearrow & \nearrow & \nearrow & & & \\ a_{10} & a_{11} & a_{12} & . & . & . & . \\ & \nearrow & \nearrow & & & & \\ a_{20} & a_{21} & . & . & . & . & . \\ & \nearrow & & & & & \\ a_{30} & . & . & . & . & . & . \\ . & . & . & . & . & . & . \end{array}$$

Im ersten Schema ist

$$a_{00}, a_{10}, a_{11}, a_{01}, a_{20}, a_{21}, a_{22}, a_{12}, a_{02}, \dots,$$

im zweiten ist

$$a_{00}, a_{10}, a_{01}, a_{20}, a_{11}, a_{02}, a_{20}, a_{21}, a_{12}, a_{03} \dots$$

die lineare Anordnung.

Ein hinreichendes (zum Teil auch notwendiges) Konvergenzkriterium liefert der in der mathematischen Literatur gewöhnlich als *großer Umordnungssatz* bezeichnete

Satz. *Es sei (a_{mn}) eine Doppelfolge, und*

$$a_{m_1 n_1}, a_{m_2 n_1}, a_{m_3 n_1}, \dots$$

sei eine lineare Anordnung der Doppelfolge. Existiert dann eine reelle Zahl K derart, daß stets

$$\sum_{m=0}^p \sum_{n=0}^p |a_{mn}| \leq K \quad (p \in \mathbf{N}) \quad (1)$$

ist, so konvergieren alle Reihen

$$\sum_{n=0}^{\infty} a_{mn} \quad (m \in \mathbf{N}), \quad \sum_{m=0}^{\infty} a_{mn} \quad (n \in \mathbf{N}), \quad \sum_{j=0}^{\infty} a_{m_j n_j} \quad (2a, b, c)$$

absolut, und es gilt

$$\sum_{j=0}^{\infty} a_{m_j n_j} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{mn}. \quad (3)$$

Beweis. Wegen (1) ist K offenbar eine obere Schranke für die Summe der Absolutbeträge beliebig vieler Glieder der Reihen (2a) und (2b). Dasselbe gilt aber auch für die Reihe (2c), denn ist p das Maximum der Zahlen m_j, n_j mit $j \leq k$, so gilt

$$\sum_{j=0}^k |a_{m_j n_j}| \leq \sum_{m=0}^p \sum_{n=0}^p |a_{mn}| \leq K.$$

Die Reihen (2) konvergieren daher sämtlich absolut. Wir setzen

$$Z_m := \sum_{n=0}^{\infty} a_{mn},$$

$$S := \sum_{j=0}^{\infty} a_{m_j n_j}, \quad S^* := \sum_{j=0}^{\infty} |a_{m_j n_j}|,$$

$$S_p := \sum_{m=0}^p \sum_{n=0}^p a_{mn}, \quad S_p^* := \sum_{m=0}^p \sum_{n=0}^p |a_{mn}|.$$

Die erste Behauptung (3) lautet dann

$$S = \sum_{m=0}^{\infty} Z_m. \quad (4)$$

Da die Reihe mit den Gliedern a_{m,n_j} absolut, also auch unbedingt konvergiert, können wir eine spezielle lineare Anordnung wählen. Ohne Beschränkung der Allgemeinheit sei dies die Anordnung nach Quadraten. Die Folge (S_p) bzw. (S_p^*) ist unter dieser Voraussetzung eine Teilfolge der Partialsummenfolge der Reihe $\sum a_{m,n_j}$ bzw. $\sum |a_{m,n_j}|$, und es gilt

$$S_p \rightarrow S, \quad S_p^* \rightarrow S^* \quad \text{für } p \rightarrow \infty. \quad (5)$$

Für alle p, k mit $k \geq 1$ ist ferner

$$S_p^* + \left| \sum_{m=0}^p \sum_{n=p+1}^{p+k} a_{mn} \right| \leq \sum_{m=0}^p \sum_{n=0}^{p+k} |a_{mn}| \leq S_{p+k}^*,$$

also

$$\left| \sum_{m=0}^p \sum_{n=p+1}^{p+k} a_{mn} \right| \leq S_{p+k}^* - S_p^*.$$

Der Beweis von (4) ergibt sich nun wie folgt. Es ist

$$\begin{aligned} \left| S - \sum_{m=0}^p Z_m \right| &= \left| S - \sum_{m=0}^p \left(\lim_{k \rightarrow \infty} \sum_{n=0}^{p+k} a_{mn} \right) \right| = \left| \lim_{k \rightarrow \infty} \left(S - \sum_{m=0}^p \sum_{n=0}^{p+k} a_{mn} \right) \right| \\ &= \lim_{k \rightarrow \infty} \left| S - \sum_{m=0}^p \sum_{n=0}^{p+k} a_{mn} \right|^{1)} \\ &\leq \lim_{k \rightarrow \infty} \left(\left| S - \sum_{m=0}^p \sum_{n=0}^p a_{mn} \right| + \left| \sum_{m=0}^p \sum_{n=p+1}^{p+k} a_{mn} \right| \right) \\ &\leq \lim_{k \rightarrow \infty} (|S - S_p| + S_{p+k}^* - S_p^*) = |S - S_p| + S^* - S_p^*. \end{aligned}$$

Für $p \rightarrow \infty$ strebt die rechte Seite wegen (5) gegen 0, und damit ist (4) bewiesen. Analog wird die zweite Behauptung (3) bewiesen.

Der große Umordnungssatz ist auch in vollständigen normierten Räumen gültig.

2.2.7. Multiplikation unendlicher Reihen

Gegeben seien zwei konvergente Reihen $\sum a_n, \sum b_n$ mit reellen (oder komplexen) Gliedern. Unter gewissen Voraussetzungen kann das Produkt der beiden Summen in Form einer Doppelreihe oder auch einer einfachen Reihe angegeben werden. Ein hinreichendes Kriterium hierfür ist die absolute Konvergenz der beiden Reihen

¹⁾ Vgl. 2.1.3.(7').

$\sum a_n, \sum b_n$. Dann gibt es nämlich reelle Zahlen K_1, K_2 mit

$$\sum_{n=0}^{\infty} |a_n| \leq K_1, \quad \sum_{n=0}^{\infty} |b_n| \leq K_2,$$

und für alle natürlichen Zahlen p gilt

$$\sum_{m=0}^p \sum_{n=0}^p |a_m b_n| = \sum_{m=0}^p |a_m| \sum_{n=0}^p |b_n| \leq K_1 K_2,$$

d. h., für die Doppelfolge $(a_m b_n)$ ist die Voraussetzung 2.2.6.(1) des großen Umordnungssatzes erfüllt. Aus

$$\sum_{m=0}^p a_m \sum_{n=0}^p b_n = \sum_{m=0}^p \sum_{n=0}^p a_m b_n,$$

dem großen Umordnungssatz und 2.1.3.(6) folgt somit

$$\sum_{m=0}^{\infty} a_m \sum_{n=0}^{\infty} b_n = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_m b_n = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_m b_n. \quad (1)$$

Eine für die Anwendungen besonders wichtige Formel ergibt sich, wenn man im großen Umordnungssatz die Anordnung nach Diagonalen wählt.

Wenn wir beachten, daß die Summe der beiden Indizes aller in derselben Diagonalen stehenden Glieder $a_m b_n$ konstant ist, erhalten wir den

Satz. Sind die Reihen $\sum a_n, \sum b_n$ absolut konvergent, so gilt

$$\sum_{m=0}^{\infty} a_m \sum_{n=0}^{\infty} b_n = \sum_{k=0}^{\infty} \sum_{j=0}^k a_j b_{k-j}. \quad (2)$$

Die in (2) rechts stehende (einfache) Reihe mit den Gliedern $\sum_{j=0}^k a_j b_{k-j}$ heißt die *Cauchysche Produktreihe* der Reihen $\sum a_n, \sum b_n$.

Beispiel 1. Für alle x mit $|x| < 1$ gilt

$$\begin{aligned} \left(\sum_{n=0}^{\infty} x^n \right) \left(\sum_{n=0}^{\infty} (-1)^n x^n \right) &= \sum_{k=0}^{\infty} \sum_{j=0}^k x^j (-1)^{k-j} x^{k-j} = \sum_{k=0}^{\infty} x^k \sum_{j=0}^k (-1)^{k-j} \\ &= \sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2}, \end{aligned}$$

denn es ist

$$\sum_{j=0}^k (-1)^{k-j} = \begin{cases} 1 & \text{für } k = 2n \quad (n \in \mathbf{N}), \\ 0 & \text{für } k = 2n + 1 \quad (n \in \mathbf{N}). \end{cases}$$

Ein weiteres wichtiges Beispiel behandeln wir im nächsten Abschnitt.

Die Ergebnisse dieses Abschnitts machen zusammen mit dem kleinen Umordnungssatz in 2.2.5. die eigentliche Bedeutung der absoluten Konvergenz deutlich. Mit absolut konvergenten Reihen darf man (im Gegensatz zu der negativen Feststellung am Anfang von 2.2.1.) ebenso wie mit endlichen Summen uneingeschränkt kommutativ, distributiv und assoziativ rechnen.

2.2.8. Die Exponentialreihe

Für alle reellen (bzw. komplexen) Zahlen x definieren wir $\exp x$ (gelesen: Exponent x) durch

$$\exp x := \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \cdots \quad (x \in \mathbf{R} \text{ bzw. } \mathbf{C}). \quad (1)$$

In 2.2.4., Beispiel 3, wurde bewiesen, daß diese Reihe für alle reellen bzw. komplexen Zahlen x absolut konvergent ist. Sie wird *Exponentialreihe* genannt. Sie wurde von NEWTON im Jahre 1666 durch Umkehrung der Mercatorschen Reihe für den natürlichen Logarithmus gefunden. Wir drücken das Produkt von $\exp x$ und $\exp y$ durch die Cauchysche Produktreihe aus. Es ist

$$\begin{aligned} \exp x \exp y &= \sum_{m=0}^{\infty} \frac{x^m}{m!} \sum_{n=0}^{\infty} \frac{y^n}{n!} = \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{x^j}{j!} \frac{y^{k-j}}{(k-j)!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} x^j y^{k-j} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} x^j y^{k-j} \\ &= \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} = \exp(x+y). \end{aligned}$$

Die reelle bzw. komplexe Funktion (1) genügt somit der Funktionalgleichung der Exponentialfunktion, es ist

$$\exp(x+y) = \exp x \exp y \quad (x, y \in \mathbf{R} \text{ bzw. } \mathbf{C}). \quad (2)$$

Wegen

$$\exp x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots > 1 \quad (x \in \mathbf{R}, x > 0) \quad (3)$$

folgt aus $x < y$ stets

$$\exp y = \exp(x + (y-x)) = \exp x \exp(y-x) > \exp x,$$

d. h., die reelle Funktion

$$\exp x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots \quad (x \in \mathbf{R}) \quad (4)$$

ist streng monoton wachsend und genügt der Funktionalgleichung der Exponentialfunktion. Daher ist (4) eine Exponentialfunktion, deren Basis a wir ermitteln wollen. Nach (4) ist

$$a^x - 1 = \exp x - 1 \leq x + x^2 + \dots = \frac{x}{1-x} \quad (x \in \mathbb{R}, 0 < x < 1), \quad (5)$$

und aus dem Satz über die Partialsummen alternierender Reihen (2.2.3., Satz 2) folgt

$$1 - a^{-x} = x - \frac{x^2}{2!} + \frac{x^3}{3!} - \dots \geq x - \frac{x^2}{2} \quad (x \in \mathbb{R}, 0 < x < 1). \quad (6)$$

Mit 2.1.5.(10) ergibt sich

$$x - \frac{x^2}{2} \leq 1 - a^{-x} \leq x \ln a \leq a^x - 1 \leq \frac{x}{1-x} \quad (0 < x < 1).$$

Für $x = \frac{1}{n+1}$ folgt

$$1 - \frac{1}{2(n+1)} \leq \ln a \leq 1 + \frac{1}{n},$$

und der Grenzübergang $n \rightarrow \infty$ liefert $\ln a = 1$, $a = e$. Somit ist $\exp x = e^x$ ($x \in \mathbb{R}$), und wir gewinnen mit

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots \quad (x \in \mathbb{R}) \quad (7)$$

eine neue wichtige Darstellung der speziellen Exponentialfunktion. Für die Zahl e ergibt sich mit $x = 1$ die sehr gut konvergierende Reihendarstellung

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!}. \quad (8)$$

Hiermit beweisen wir den

Satz. Die Eulersche Zahl e ist irrational.

Beweis. Nehmen wir an, es wäre $e = \frac{p}{q}$ ($p, q \in \mathbb{N}$). Wegen $2 < e < 3$ ist $q > 1$, und aus (8) folgt

$$\frac{p}{q} = \sum_{k=0}^{\infty} \frac{1}{k!} = \sum_{k=1}^{\infty} \frac{1}{(q+k)!}.$$

Multiplizieren wir beide Seiten mit $q!$, so steht links eine ganze Zahl m . Da die rechte Seite positiv ist, ist $m \geq 1$. Es folgt

$$\begin{aligned} 1 \leq m &= \sum_{k=1}^q \frac{q!}{(q+k)!} < \frac{1}{q+1} + \frac{1}{(q+1)^2} + \cdots + \frac{1}{(q+1)^k} + \cdots \\ &= \frac{1}{q+1} \cdot \frac{1}{1 - \frac{1}{q+1}} = \frac{1}{q} < 1. \end{aligned}$$

Das ist ein Widerspruch, und folglich ist e irrational.

Erheblich schwieriger als der eben geführte Beweis ist der Nachweis, daß e eine *transzendente Zahl* ist, d. h., daß es keine ganzrationale Funktion f mit rationalen Koeffizienten a_0, a_1, \dots, a_n ($a_n \neq 0$) und $f(e) = 0$ gibt.

Nach (7) gilt $e^x = \exp x$ für reelle x , während das Symbol e^x für nichtreelle komplexe Zahlen x nicht definiert ist. Wir definieren es ebenso wie im Reellen, d. h., wir setzen

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (x \in \mathbb{C}), \quad (9)$$

und damit lautet die am Anfang bewiesene Funktionalgleichung

$$e^{x+y} = e^x e^y \quad (x, y \in \mathbb{C}). \quad (10)$$

Die durch (9) definierte Funktion heißt die *komplexe Exponentialfunktion*. Wegen $e^x \cdot e^{-x} = e^0 = 1$ ist stets $e^x \neq 0$, die komplexe Exponentialfunktion besitzt keine Nullstelle. Wir vermerken weiterhin, daß stets

$$\overline{e^x} = e^{\bar{x}} \quad (x \in \mathbb{C}) \quad (11)$$

gilt. Dies folgt aus der Darstellung (9) mit Hilfe von 2.2.2.(1), 2.1.3.(12) und 1.2.2.(5), (6).

Wir führen jetzt den angekündigten Beweis für die Existenz der trigonometrischen Funktionen mit den Hilfsmitteln der Reihenlehre. Für alle $x \in \mathbb{R}$ ist

$$\begin{aligned} \sum_{k=0}^{2n+1} \frac{(ix)^k}{k!} &= \sum_{k=0}^n \frac{(ix)^{2k}}{(2k)!} + \sum_{k=0}^n \frac{(ix)^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

Für $n \rightarrow \infty$ konvergiert die linke Seite nach (9) gegen e^{ix} . Auf der rechten Seite ergeben sich die absolut konvergenten Reihen 2.2.4.(16), (17). Setzen wir also

$$\cos x := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \quad (x \in \mathbb{R}) \quad (12)$$

bzw.

$$\sin x := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \dots \quad (x \in \mathbb{R}), \quad (13)$$

so erhalten wir die *Eulersche Relation*

$$e^{ix} = \cos x + i \sin x \quad (x \in \mathbb{R}),$$

die EULER wohl bereits 1740 gefunden hatte.

Die Entwicklung der Sinus- und Kosinusfunktion in unendliche Reihen war seit NEWTONS „De analysi per aequationes“ von 1666/69 bekannt. Die Reihen selbst zur Definition der Funktionen, und zwar für komplexe Argumente, verwendet zu haben ist das Verdienst von EULER 1749.

Die Relationen $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$ können unmittelbar aus (12), (13) abgelesen werden. Die Additionstheoreme folgen aus

$$\begin{aligned} \cos(x+y) + i \sin(x+y) &= e^{i(x+y)} = e^{ix} e^{iy} \\ &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) \\ &\quad + i(\sin x \cos y + \cos x \sin y) \end{aligned}$$

durch Vergleich von Real- und Imaginärteil. Es bleibt nur noch 1.6.1.(5) zu beweisen. Für $0 < x < 1$ bilden die Absolutbeträge der Glieder der Reihen (12), (13) streng monoton fallende Nullfolgen. Aus dem Satz über die Partialsummen alternierender Reihen folgt daher

$$\left. \begin{aligned} 0 &< 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \dots = \cos x, \\ 0 &< x - \frac{x^3}{3!} + - \dots = \sin x \end{aligned} \right\} \quad (0 < x < 1).$$

Ferner ist

$$\begin{aligned} \sin x - x \cos x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \dots - \left(x - \frac{x^3}{2!} + \frac{x^5}{4!} - + \dots \right) \\ &= x^3 \left(\frac{1}{2!} - \frac{1}{3!} \right) - x^5 \left(\frac{1}{4!} - \frac{1}{5!} \right) + - \dots, \end{aligned}$$

und auf Grund desselben Satzes ist $\sin x - x \cos x > 0$.

Zusammenfassend ergibt sich

$$0 < \sin x < x < \frac{\sin x}{\cos x} \quad (0 < x < 1),$$

und die Existenzaussage von Satz 1 in 1.6.1. ist bewiesen.

Nach 1.6.3.(7) kann nun jede komplexe Zahl $z \neq 0$ auf genau eine Weise in der Form

$$z = re^{i\varphi} \quad (r > 0, -\pi < \varphi \leq \pi) \quad (14)$$

dargestellt werden.

2.3. Stetigkeit und Grenzwerte von Funktionen

2.3.1. Stetigkeit (Folgendefinition)

Im folgenden lernen wir einen fundamentalen Begriff der Analysis kennen. Der Begriff der *Stetigkeit* einer Funktion kann auf verschiedene Weisen definiert werden. In diesem vorbereitenden Abschnitt führen wir diesen Begriff auf die Konvergenz von Folgen zurück. Im nächsten Abschnitt werden wir eine zweite äquivalente Definition geben, die sich auf den Umgebungsbegriff stützt. Es sei f eine reelle Funktion. Dann können wir zu jeder Folge (x_n) mit $x_n \in D(f)$ ($n \in \mathbb{N}$) die Folge der zugehörigen Funktionswerte $f(x_n)$ bilden. In der Theorie der Grenzwerte und der Stetigkeit wird ein Zusammenhang zwischen der Konvergenz der Folgen (x_n) und $f((x_n))$ hergestellt.

Beispiel 1. Wir betrachten die Funktion $f(x) = \operatorname{sgn} x$ ($x \in \mathbb{R}$) und bilden zu den Nullfolgen

$$\left(\frac{1}{n}\right), \quad \left(-\frac{1}{n}\right), \quad \left(\frac{(-1)^n}{n}\right) \quad (n > 0)$$

die Folgen der zugehörigen Funktionswerte, d. h. die Folgen mit den Gliedern 1, -1 , $(-1)^n$.

Die ersten beiden Folgen sind konvergent, haben aber die verschiedenen Grenzwerte 1 und -1 , die dritte Folge ist unbestimmt divergent. Aus der Konvergenz der vorgegebenen Folgen kann somit nicht auf die Konvergenz der Folgen der Funktionswerte geschlossen werden.

Beispiel 2. Für eine beliebige natürliche Zahl m betrachten wir die reelle (oder komplexe) Potenzfunktion $f(x) = x^m$. Es sei (x_n) eine beliebige konvergente Folge mit dem Grenzwert a . Dann besitzt die Folge eine Schranke K , und es gilt auch $|a| \leq K$. Im Fall $m = 0$ ist

$$|f(x_n) - f(a)| = |x_n^0 - a^0| = 0,$$

und für $m = 1, 2, \dots$ gilt

$$\begin{aligned} |f(x_n) - f(a)| &= |x_n^m - a^m| = \left| (x_n - a) \sum_{j=1}^m x_n^{j-1} a^{m-j} \right| \\ &\leq |x_n - a| \sum_{j=1}^m K^{j-1} K^{m-j} = |x_n - a| m \cdot K^{m-1}. \end{aligned}$$

Rechts stehen die Glieder einer Nullfolge, d. h., es gilt

$$\lim_{n \rightarrow \infty} x_n^m = a^m. \quad (1)$$

Aus der Konvergenz der vorgegebenen Folge folgt somit im Gegensatz zu Beispiel 1 stets die Konvergenz der Folge der zugehörigen Funktionswerte, und zwar konvergiert sie stets gegen den zum Grenzwert der vorgegebenen Folge gehörenden Funktionswert.

Nach diesen Vorbetrachtungen formulieren wir die

Definition 1. Eine reelle (bzw. komplexe) Funktion f heißt *im Punkt* $a \in D(f)$ oder *an der Stelle* $a \in D(f)$ *stetig*, wenn aus $x_n \in D(f)$ ($n \in \mathbb{N}$) und $\lim_{n \rightarrow \infty} x_n = a$ stets folgt, daß $\lim_{n \rightarrow \infty} f(x_n)$ existiert und

$$\lim_{n \rightarrow \infty} f(x_n) = f(a) \quad (2)$$

ist. Ist f an der Stelle $a \in D(f)$ nicht stetig, so sagen wir, f sei *an der Stelle* a *unstetig*.

Die Funktion $y = \operatorname{sgn} x$ ist an der Stelle 0 unstetig, die Potenzfunktionen sind dagegen an allen Stellen ihres Definitionsbereichs, d. h. für alle reellen (bzw. komplexen) Zahlen, x stetig.

Der Begriff der stetigen Funktion tritt in der mathematischen Literatur recht spät auf, da man zum größten Teil von vornherein stetige Funktionen auf Grund des engen Funktionsbegriffs betrachtete und den Ausnahmen nicht besonderen Wert beimaß. In BOLZANOS „Rein analytischen Beweis ...“ von 1817 findet sich die erste genaue Definition: „Nach einer richtigen Erklärung ... versteht man unter der Redensart, daß eine Function $f(x)$ für alle Werthe von x , die inner- oder außerhalb gewisser Grenzen liegen, nach dem Gesetze der Stetigkeit sich ändere, nur so viel, daß, wenn z irgend ein solcher Werth ist, der Unterschied $f(x + \omega) - f(x)$ kleiner als jede gegebene Größe gemacht werden könne, wenn man ω so klein, als man nur immer will, annehmen kann.“ CAUCHYS Definition von 1821, mit der der Begriff der stetigen Funktion fest im Lehrgebäude der Mathematik verankert wurde, unterscheidet sich nur unwesentlich von der BOLZANOS.

Ist eine Funktion f an der Stelle a stetig, so folgt aus der Definition sofort, daß jede Einschränkung von f , deren Definitionsbereich den Punkt a enthält, ebenfalls an der Stelle a stetig ist. Ebenso leicht erkennt man, daß f an der Stelle a stetig ist, wenn a ein isolierter Punkt des Definitionsbereichs von f ist.¹⁾

Beispiel 3. Die Funktion

$$y = \ln x \quad (x \in \mathbb{R}, x > 0) \quad (3)$$

ist in jedem Punkt a mit $a > 0$ stetig. Zum Beweis gehen wir von 2.1.5.(3), (9) aus. Hiernach ist stets

$$\begin{aligned} \frac{\frac{x}{a} - 1}{\frac{x}{a}} &\leq \ln \frac{x}{a} \leq \frac{x}{a} - 1 \\ \frac{x - a}{x} &\leq \ln x - \ln a \leq \frac{x - a}{a}. \end{aligned}$$

¹⁾ In manchen Lehrbüchern wird der Begriff der Stetigkeit nur für Häufungspunkte des Definitionsbereichs definiert.

Wählen wir eine Folge positiver Zahlen x_n , die gegen die positive reelle Zahl a konvergiert, so erhalten wir mit 2.1.3., Satz 4, durch Grenzübergang

$$0 \leq \lim_{n \rightarrow \infty} (\ln x_n - \ln a) \leq 0,$$

und hieraus folgt die Behauptung.

Beispiel 4. Die trigonometrischen Funktionen $y = \cos x$, $y = \sin x$ sind in allen Punkten $x \in \mathbb{R}$ stetig, denn nach 1.6.1.(24), (25) ist

$$|\cos x_n - \cos x| \leq |x_n - x|,$$

$$\sin x_n - \sin x| \leq |x_n - x|.$$

Beim Nachweis der Stetigkeit können wir uns auf monotone Folgen beschränken.

Satz 1. Eine reelle Funktion f ist in einem Häufungspunkt $a \in D(f)$ stetig genau dann, wenn (2) für alle streng monotonen Folgen (x_n) mit dem Grenzwert a und mit $x_n \in D(f)$ ($n \in \mathbb{N}$) gilt.

Beweis. Ist f an der Stelle a stetig, so ist die Bedingung von Satz 1 nach Definition 1 offensichtlich erfüllt. Die Umkehrung beweisen wir durch Kontraposition. Die Funktion f sei an der Stelle a unstetig. Dann gibt es eine Folge (x_n) mit $x_n \in D(f)$ und $x_n \rightarrow a$, für die $\{f(x_n)\}$ nicht den Grenzwert $f(a)$ besitzt. Es gibt daher ein $\varepsilon > 0$ derart, daß $|f(x_n) - f(a)| \geq \varepsilon$ für unendlich viele natürliche Zahlen n erfüllt ist. Wir können daher zu einer Teilfolge (x_n') von (x_n) mit $|f(x_n') - f(a)| \geq \varepsilon$ übergehen. Diese Teilfolge enthält eine monotone Teilfolge (x_n'') . Wegen $x_n'' \rightarrow a$ für $n \rightarrow \infty$ und $|f(x_n'') - f(a)| \geq \varepsilon$ ist stets $x_n'' \neq a$, und es gibt eine streng monotone Teilfolge (x_n^*) von (x_n'') mit $|f(x_n^*) - f(a)| \geq \varepsilon$. Die Folge $\{f(x_n^*)\}$ besitzt somit nicht den Grenzwert $f(a)$, und Satz 1 ist bewiesen.

Beispiel 5. Für jede positive reelle Zahl a ist die Funktion

$$y = a^x \quad (x \in \mathbb{R}) \quad (4)$$

an jeder Stelle b ihres Definitionsbereichs stetig. Zum Beweis wählen wir eine streng monoton wachsende Folge (x_n') und eine streng monoton fallende Folge (x_n'') mit $x_n' \rightarrow b$, $x_n'' \rightarrow b$ für $n \rightarrow \infty$. Ferner wählen wir rationale Zahlen r_n' , r_n'' mit $x_{n-1}' < r_n' < x_n'$ und $x_n'' < r_n'' < x_{n-1}''$, so daß auch $r_n' \rightarrow b$, $r_n'' \rightarrow b$ für $n \rightarrow \infty$ gilt. Ist etwa $a \geq 1$, so folgt

$$a^{r_n'} \leq a^{x_n'} \leq a^b \leq a^{x_n''} \leq a^{r_n''},$$

so daß wir nur noch zu zeigen haben, daß $(a^{r_n''} - a^{r_n'})$ eine Nullfolge ist. Dies beweisen wir mit Hilfe von 1.4.1.(8), (9). Ist R eine (rationale) Schranke von $r_n'' - r_n'$, so ist

$$|a^{r_n''} - a^{r_n'}| = a^{r_n'} |a^{r_n'' - r_n'} - 1| \leq a^{r_n'} a^{R+1} |a - 1| |r_n'' - r_n'|,$$

und hieraus folgt die Behauptung, weil $(r_n'' - r_n')$ eine Nullfolge ist. Die Funktion (4) ist somit an jeder Stelle b stetig.

Dieses Ergebnis zeigt, daß die Exponentialfunktion für irrationale Exponenten x auch durch

$$a^x := \lim_{n \rightarrow \infty} a^{r_n} \quad \left(\lim_{n \rightarrow \infty} r_n = x \right) \quad (5)$$

mit $r_n \in \mathbb{Q}$ definiert werden kann.

Die obige Definition für die Stetigkeit kann unverändert auf Funktionen aus einem euklidischen Raum \mathbb{R}^q in einen euklidischen Raum \mathbb{R}^q (oder noch allgemeiner auf Funktionen aus einem metrischen Raum X in einen metrischen Raum Y) übertragen werden.

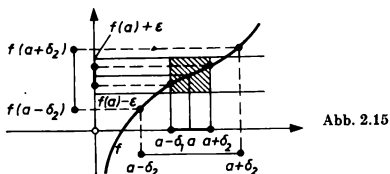
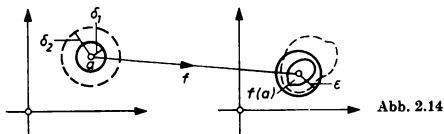
Definition 1'. Eine Funktion f aus einem metrischen Raum X in einen metrischen Raum Y heißt *im Punkt a oder an der Stelle a stetig*, wenn $a \in D(f)$ ist und wenn aus $x_n \in D(f)$ ($n \in \mathbb{N}$) und $\lim_{n \rightarrow \infty} x_n = a$ stets folgt, daß $\lim_{n \rightarrow \infty} f(x_n)$ existiert und $\lim_{n \rightarrow \infty} f(x_n) = f(a)$ ist.

Auch Satz 1 bleibt für Funktionen aus \mathbb{R} in einen metrischen Raum X gültig.

2.3.2. Stetigkeit (Umgebungsdefinition)

Das nachfolgende Kriterium für die Stetigkeit einer Funktion wird häufig als Definition für diesen Begriff verwendet. Diese Definition hat gegenüber der Folgendefinition den Vorteil, daß sie auf allgemeinere Raumtypen übertragen werden kann. Ihr Inhalt kann geometrisch wie folgt geschildert werden. Es sei $a \in D(f)$. Dann wählen wir eine beliebige positive reelle Zahl ε , und betrachten die ε -Umgebung von $f(a)$. (In Abb. 2.14 bzw. Abb. 2.15 ist der Sachverhalt für den Fall einer komplexen bzw. reellen Funktion angedeutet.)

Für verschiedene positive Zahlen δ betrachten wir nun die Menge aller Bildpunkte $f(x)$ mit $x \in D(f)$ und $x \in U_\delta(a)$. Im allgemeinen werden diese Bildpunkte nicht sämtlich in der vorgegebenen ε -Umgebung von $f(a)$ liegen. Das Kriterium für die Stetig-



keit einer Funktion f an der Stelle a besteht nun gerade darin, daß für ein hinreichend kleines δ alle diese Bildpunkte in der vorgegebenen ε -Umgebung liegen.

Satz 1. Eine reelle (bzw. komplexe) Funktion f ist im Punkt $a \in D(f)$ stetig genau dann, wenn es zu jeder positiven reellen Zahl ε eine positive reelle Zahl δ gibt derart, daß aus

$$x \in D(f) \quad \text{und} \quad |x - a| < \delta$$

stets

$$|f(x) - f(a)| < \varepsilon$$

folgt:

$$f \text{ stetig in } a \in D(f) \Leftrightarrow$$

$$\bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x \in D(f)} (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon). \quad (1)$$

Beweis. Wir betrachten die rechte Aussage der Äquivalenz (1). Ist diese Aussage wahr, so wählen wir eine beliebige Folge (x_n) mit $x_n \in D(f)$ und $\lim_{n \rightarrow \infty} x_n = a$. Es sei ε eine positive reelle Zahl. Wir wählen gemäß (1) eine positive reelle Zahl δ mit der Eigenschaft

$$\bigwedge_{x \in D(f)} (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon). \quad (2)$$

Nach der Definition für die Konvergenz von Folgen gibt es ein N mit $|x_n - a| < \delta$ für $n \geq N$. Wegen (2) ist dann $|f(x_n) - f(a)| < \varepsilon$ für $n \geq N$, d. h., es ist $\lim_{n \rightarrow \infty} f(x_n) = f(a)$, die Funktion f ist an der Stelle a stetig.

Ist die rechte Aussage in (1) falsch, so ist ihre Negation wahr, d. h., es gibt eine positive Zahl ε mit der Eigenschaft

$$\bigwedge_{\delta > 0} \bigvee_{x \in D(f)} (|x - a| < \delta \wedge |f(x) - f(a)| \geq \varepsilon).$$

Wir setzen nun $\delta := \frac{1}{n}$ und bestimmen x_n so, daß $x_n \in D(f)$, $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - f(a)| \geq \varepsilon$ ist. Dann gilt $\lim_{n \rightarrow \infty} x_n = a$, aber nicht $\lim_{n \rightarrow \infty} f(x_n) = f(a)$, d. h., die Funktion f ist an der Stelle a unstetig. Damit ist Satz 1 bewiesen.

Die Aussage (1) ist logisch äquivalent mit der Aussage

$$f \text{ stetig in } a \in D(f) \Leftrightarrow$$

$$\bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x \in D(f)} (|f(x) - f(a)| \geq \varepsilon \Rightarrow |x - a| \geq \delta). \quad (3)$$

In manchen Fällen lassen sich Stetigkeitsuntersuchungen in dieser Form besser durchführen.

Beispiel 1. Wir zeigen, daß die komplexe Exponentialfunktion in jedem Punkt a ihres Definitionsbereichs stetig ist. Für den reellen Fall wurde der Beweis schon in 2.3.1., Beispiel 4, geführt.

Für $|x| < 1$ ist

$$|e^x - 1| = \left| \frac{x}{1!} + \frac{x^2}{2!} + \dots \right| \leq |x| \left(1 + \frac{1}{2!} + \dots \right), \quad (4)$$

$$|e^x - 1| \leq |x| e \quad (|x| < 1).$$

Für $|x - a| < 1$ ist daher

$$|e^x - e^a| = |e^a(e^{x-a} - 1)| \leq |e^a| \cdot e \cdot |x - a|,$$

$$|e^x - e^a| \leq |e^{a+1}| \cdot |x - a|. \quad (5)$$

Wählen wir zu vorgegebenem $a \in \mathbf{C}$ und $\varepsilon > 0$ ein δ mit $0 < \delta < 1$ und $\delta < \varepsilon |e^{-a-1}|$, so folgt aus $|x - a| < \delta$ stets

$$|e^x - e^a| < |e^{a+1}| \cdot \varepsilon |e^{-a-1}| = \varepsilon$$

und damit die Behauptung.

Stetigkeitsuntersuchungen können häufig mit dem folgenden Vergleichskriterium durchgeführt werden.

Satz 2. Es seien f, g reelle (bzw. komplexe) Funktionen mit $D(g) \subseteq D(f)$. Gibt es zu $a \in D(g)$ ein K mit

$$|g(x) - g(a)| \leq K |f(x) - f(a)| \quad (x \in D(g)), \quad (6)$$

so folgt aus der Stetigkeit von f an der Stelle a auch die Stetigkeit von g an der Stelle a .

Beweis. Aus $x_n \rightarrow a$ folgt $f(x_n) \rightarrow f(a)$ und damit $g(x_n) \rightarrow g(a)$, d. h., g ist in a stetig.

Wir bemerken, daß es in Satz 2 genügt, die Gültigkeit von (6) für alle x aus einer hinreichend kleinen Umgebung von a nachzuweisen. Für reelle stetige Funktionen ist stets

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)|,$$

und dies besagt nach Satz 2, daß mit f auch die durch $f_0(x) := |f(x)|$ definierte Funktion f_0 in a stetig ist. Allgemeiner gilt

Satz 3. Ist f eine im Punkt $a \in D(f)$ stetige komplexe Funktion, so sind ihr Absolutbetrag, ihr Realteil und ihr Imaginärteil, d. h. die Funktionen

$$\begin{aligned} f_0(x) &:= |f(x)| & (x \in D(f)), \\ f_1(x) &:= \operatorname{Re} f(x) & (x \in D(f)), \\ f_2(x) &:= \operatorname{Im} f(x) & (x \in D(f)), \end{aligned}$$

im Punkt a stetig.

Beweis. Für $i = 0, 1, 2$ ist

$$|f_i(x) - f_i(a)| \leq |f(x) - f(a)| \quad (a, x \in D(f)),$$

und die Behauptung folgt aus Satz 2.

Die Sätze 1, 2 und 3 können wir wieder verallgemeinern.

Satz 1'. Eine Funktion f aus einem metrischen Raum X in einen metrischen Raum Y ist im Punkt $a \in D(f)$ stetig genau dann, wenn es zu jeder positiven reellen Zahl ε eine positive reelle Zahl δ gibt derart, daß aus

$$x \in D(f) \quad \text{und} \quad \varrho_X(x, a) < \delta$$

stets

$$\varrho_Y(f(x), f(a)) < \varepsilon$$

folgt.

Der Beweis von Satz 1 kann unverändert übertragen werden, wenn wir überall $|x - a|$ durch $\varrho_X(x, a)$ und $|f(x) - f(a)|$ durch $\varrho_Y(f(x), f(a))$ ersetzen.

Satz 2'. Es sei f bzw. g eine Funktion aus einem metrischen Raum X in einen metrischen Raum Y bzw. Z mit $D(g) \subseteq D(f)$. Gibt es zu $a \in D(g)$ ein K mit

$$\varrho_Z(g(x), g(a)) \leq K \varrho_Y(f(x), f(a)) \quad (a \in D(g)), \quad (6')$$

so folgt aus der Stetigkeit von f an der Stelle a auch die Stetigkeit von g an der Stelle a .

Der Beweis von Satz 2 bleibt unverändert, und wiederum genügt es, die Gültigkeit von (6') für eine hinreichend kleine Umgebung von a nachzuweisen.

Satz 3'. Es sei f eine Funktion aus einem metrischen Raum X in den Raum \mathbb{R}^p und

$$f_0(x) = \|f(x)\|, \quad f(x) = (f_1(x), \dots, f_p(x)) \quad (x \in D(f)).$$

Die Funktion f ist an der Stelle $a \in D(f)$ stetig genau dann, wenn die Funktionen f_1, \dots, f_p an der Stelle a stetig sind. Mit f ist auch f_0 an der Stelle a stetig.

Beweis. Für alle $x \in D(f)$ und $i = 1, \dots, p$ gilt

$$|f_i(x) - f_i(a)| \leq \|f(x) - f(a)\| \leq \sum_{j=1}^p |f_j(x) - f_j(a)|.$$

Ist f an der Stelle a stetig, so sind nach Satz 2' auch die Funktionen f_i ($i = 1, \dots, p$) an der Stelle a stetig. Sind umgekehrt alle Funktionen f_i an der Stelle a stetig, so können wir zu vorgegebenem $\varepsilon > 0$ Zahlen $\delta_i > 0$ mit $|f_i(x) - f_i(a)| < \frac{\varepsilon}{p}$ für $\varrho(x, a) < \delta_i$ bestimmen. Ist δ das Minimum der Zahlen $\delta_1, \dots, \delta_p$, so folgt aus $\varrho(x, a) < \delta$ stets $\|f(x) - f(a)\| < \varepsilon$, und f ist an der Stelle a stetig. Die Stetigkeit von f_0 an der Stelle a folgt nun aus

$$|f_0(x) - f_0(a)| = \left| \|f(x)\| - \|f(a)\| \right| \leq \|f(x) - f(a)\|$$

und Satz 2'.

Stetigkeitsuntersuchungen für Funktionen zweier Variabler lassen sich mitunter mit Hilfe von ebenen Polarkoordinaten bequem durchführen.

Beispiel 2. Es sei

$$f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \in \mathbb{R}^2 \setminus (0, 0), \\ 0 & \text{für } (x, y) = (0, 0), \end{cases}$$

$$g(x, y) := \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \in \mathbb{R}^2 \setminus (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Wir prüfen, ob diese Funktionen im Punkt $(0, 0)$ stetig sind. Für alle $(x, y) \neq (0, 0)$ bestimmen wir reelle Zahlen r, φ mit $x = r \cos \varphi$, $y = r \sin \varphi$ ($r > 0$). Dann ist

$$f(x, y) = r^3 \cos \varphi \sin \varphi \frac{\cos^2 \varphi - \sin^2 \varphi}{\cos^2 \varphi + \sin^2 \varphi} = \frac{r^3}{4} \sin 4\varphi,$$

und es folgt

$$|f(x, y) - f(0, 0)| \leq r^3.$$

Ist $\varepsilon > 0$ vorgegeben und setzen wir $\delta := \sqrt[3]{\varepsilon}$, so folgt aus

$$r = \|(x, y) - (0, 0)\| < \delta$$

stets

$$|f(x, y) - f(0, 0)| \leq r^3 < \delta^3 = \varepsilon.$$

Die Funktion f ist somit an der Stelle $(0, 0)$ stetig. Für die Funktion g gilt dagegen

$$g(x, y) = \cos^2 \varphi - \sin^2 \varphi = \cos 2\varphi.$$

Hieraus kann man sofort die Unstetigkeit der Funktion an der Stelle $(0, 0)$ ablesen,

denn setzen wir $x_n := \frac{1}{n} \cos \varphi$, $y_n := \frac{1}{n} \sin \varphi$, so gilt

$$\lim_{n \rightarrow \infty} g(x_n, y_n) = \cos 2\varphi.$$

Wir können daher gegen $(0, 0)$ konvergierende Folgen mit den Gliedern (x_n, y_n) angeben, für die die Folge der Funktionswerte gegen eine beliebig vorgegebene reelle Zahl zwischen -1 und $+1$ konvergiert.

2.3.3. Stetigkeit von verknüpften Funktionen

In 1.3.3. haben wir verschiedene Verknüpfungen von Funktionen kennengelernt. Wir prüfen, welche Folgerungen aus der Stetigkeit der verknüpften Funktionen gezogen werden können. Hierfür gilt der

Satz 1. Sind die Funktionen f, g an der Stelle a stetig, so sind die Funktionen λf , $f \pm g$, $f \cdot g$ an der Stelle a stetig. Ist darüber hinaus $g(a) \neq 0$, so ist auch $\frac{f}{g}$ an der Stelle a stetig.

Beweis. Aus $\lim_{n \rightarrow \infty} x_n = a$ und $x_n \in D(\lambda f)$ bzw. $x_n \in D(f \pm g)$ bzw. $x_n \in D(f \cdot g)$ folgt

$$\lim_{n \rightarrow \infty} \lambda f(x_n) = \lambda \lim_{n \rightarrow \infty} f(x_n) = \lambda f(a)$$

bzw.

$$\lim_{n \rightarrow \infty} (f(x_n) \pm g(x_n)) = \lim_{n \rightarrow \infty} f(x_n) \pm \lim_{n \rightarrow \infty} g(x_n) = f(a) \pm g(a)$$

bzw.

$$\lim_{n \rightarrow \infty} f(x_n) g(x_n) = \lim_{n \rightarrow \infty} f(x_n) \cdot \lim_{n \rightarrow \infty} g(x_n) = f(a) \cdot g(a).$$

Ist $g(a) \neq 0$, $\lim_{n \rightarrow \infty} x_n = a$ und $x_n \in D(h)$ mit $h = \frac{f}{g}$, so folgt aus der Konvergenz der Folge $(g(x_n))$, daß $g(x_n) \neq 0$ für $n \geq N_0$ ist. Somit gilt

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \frac{\lim_{n \rightarrow \infty} f(x_n)}{\lim_{n \rightarrow \infty} g(x_n)} = \frac{f(a)}{g(a)}.$$

Damit ist Satz 1 bewiesen.

Satz 2. Ist die Funktion f an der Stelle $a \in D(f)$ und ist die Funktion g an der Stelle $f(a) \in D(g)$ stetig, so ist die zusammengesetzte Funktion $g \circ f$ an der Stelle a stetig.

Beweis. Aus $\lim_{n \rightarrow \infty} x_n = a$ und $x_n \in D(g \circ f)$ folgt $x_n \in D(f)$ und $f(x_n) \in D(g)$. Somit ist $\lim_{n \rightarrow \infty} f(x_n) = f(a)$ wegen der Stetigkeit von f im Punkt a , und aus der Stetigkeit von g im Punkt $f(a)$ folgt

$$\lim_{n \rightarrow \infty} (g \circ f)(x_n) = \lim_{n \rightarrow \infty} g(f(x_n)) = g(f(a)) = (g \circ f)(a).$$

Für streng monotone Funktionen f gilt ein überraschender Satz. Ohne jede Voraussetzung über die Stetigkeit von f kann nämlich auf die Stetigkeit der Umkehrfunktion f^{-1} in jedem Punkt ihres Definitionsbereichs geschlossen werden, wenn der Definitionsbereich von f eine zusätzliche Bedingung erfüllt.

Satz 3. Ist der Definitionsbereich einer streng monotonen Funktion f ein Intervall (oder eine abgeschlossene Menge), so ist die Umkehrfunktion von f in jedem Punkt ihres Definitionsbereichs stetig.

Beweis. Die Funktion f sei o. B. d. A. streng monoton wachsend, und es sei $g := f^{-1}$, $b \in D(g) = W(f)$ und $a := g(b)$. Um die Stetigkeit von g im Punkt b zu beweisen, genügt es nach 2.3.1., Satz 1, zu zeigen, daß für jede streng monotone Folge

(y_n) mit $y_n \in D(g)$ und $y_n \rightarrow b$ die Folge $(g(y_n))$ gegen $g(b)$ konvergiert. Es sei etwa (y_n) streng monoton wachsend. Aus $y_n < y_{n+1} < b$ folgt $g(y_n) < g(y_{n+1}) < g(b)$, und es existiert der Grenzwert a^* der streng monoton wachsenden Folge $(g(y_n))$. Ferner ist $g(y_n) < a^* \leq g(b) = a$. Es ist $g(y_n), g(b) \in D(f)$, und da $D(f)$ ein Intervall (oder eine abgeschlossene Menge) ist, folgt $a^* \in D(f)$. Nehmen wir an, es wäre $a^* < a$. Dann ist $f(a^*) < f(a) = b$, und wegen $y_n \rightarrow b, y_n < b$ gibt es ein n mit $f(a^*) < y_n$. Hieraus folgt $a^* = g(f(a^*)) < g(y_n)$, was der Definition von a^* widerspricht. Somit ist $a^* = a = g(b)$. Ganz analog wird gezeigt, daß für jede streng monoton fallende Folge (y_n) mit $y_n \in D(g)$ und $y_n \rightarrow b$ die Folge $(g(y_n))$ gegen $g(b)$ strebt. Damit ist Satz 3 bewiesen.¹⁾

2.3.4. Grenzwerte von Funktionen

Wir betrachten in diesem Abschnitt eine Abschwächung der Forderungen, die wir bei der Definition der Stetigkeit einer Funktion in einem Punkt a gestellt haben. Die Verallgemeinerung besteht darin, daß wir den Funktionswert an der Stelle a nicht in Betracht ziehen und nicht einmal fordern, daß a zum Definitionsbereich der Funktion gehört. Allerdings muß a ein Häufungspunkt dieses Definitionsbereichs sein. Wir fragen, ob es eine reelle Zahl b gibt derart, daß die durch

$$g(x) := \begin{cases} f(x) & \text{für } x \neq a, \\ b & \text{für } x = a \end{cases}$$

definierte Funktion g an der Stelle a stetig ist. In diesem Fall sagen wir, die Funktion f besitze an der Stelle a den Grenzwert b . Ist a ein Häufungspunkt des Definitionsbereichs von f , so besitzt also f an der Stelle a einen Grenzwert, wenn f dort stetig ist oder wenn f durch eine „Korrektur“ des Funktionswertes an der Stelle a (wenn f an der Stelle a nicht definiert war, durch eine geeignete Definition dieses Funktionswertes) zu einer an der Stelle a stetigen Funktion gemacht werden kann. Wir formulieren diesen Sachverhalt in einer unmittelbaren Definition, die nicht auf den Begriff der Stetigkeit zurückgreift.

Definition. Eine reelle (bzw. komplexe) Funktion f hat in einem Häufungspunkt a von $D(f)$ den Grenzwert b , in Zeichen

$$\lim_{x \rightarrow a} f(x) = b, \quad (1)$$

wenn aus $x_n \in D(f), x_n \neq a (n \in \mathbf{N})$ und $\lim_{n \rightarrow \infty} x_n = a$ stets

$$\lim_{n \rightarrow \infty} f(x_n) = b$$

folgt.

¹⁾ Wenn man die Stetigkeit einer Funktion nur für Häufungspunkte des Definitionsbereichs erklärt, gilt der Satz nicht in der vorliegenden Form.

Neben dieser „Folgendefinition“ können wir analog zu 2.3.2. den Begriff des Grenzwertes auch mit einer „Umgebungsdefinition“ charakterisieren, und zwar gilt der

Satz 1. Eine reelle (bzw. komplexe) Funktion f hat in einem Häufungspunkt a von $D(f)$ den Grenzwert b genau dann, wenn es zu jeder positiven reellen Zahl ε eine positive reelle Zahl δ gibt derart, daß aus

$$x \in D(f) \quad \text{und} \quad 0 < |x - a| < \delta$$

stets

$$|f(x) - b| < \varepsilon$$

folgt:

$$\lim_{x \rightarrow a} f(x) = b \quad (a \text{ Häufungspunkt von } D(f)) : \Leftrightarrow \\ \bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x \in D(f)} (0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon).$$

Der Beweis verläuft analog zum Beweis von 2.3.2., Satz 1. Ferner kann 2.3.2., Satz 2, sinngemäß übertragen werden. Ist a Häufungspunkt von $D(f)$ und von $D(g)$, so folgt aus

$$\lim_{x \rightarrow a} f(x) = b$$

und

$$|g(x) - c| \leq K |f(x) - b| \quad (x \in D(g) \setminus \{a\} \subseteq D(f)) \quad (2)$$

stets

$$\lim_{x \rightarrow a} g(x) = c.$$

An Stelle von (1) schreiben wir gelegentlich „ $f(x) \rightarrow b$ für $x \rightarrow a$ “. Die Aussage „die Funktion f ist an der Stelle a stetig“ können wir, wenn a kein isolierter Punkt des Definitionsbereichs von f ist, auch in der Form

$$f(a) = \lim_{x \rightarrow a} f(x) \quad (3)$$

schreiben.

Satz 2. Es sei a ein Häufungspunkt des Definitionsbereichs der zusammengesetzten Funktion $y = f(g(x))$. Aus

$$\lim_{x \rightarrow a} g(x) = b \quad (4)$$

und der Stetigkeit von f an der Stelle b folgt

$$\lim_{x \rightarrow a} f(g(x)) = \lim_{t \rightarrow b} f(t). \quad (5)$$

Der Beweis ergibt sich unmittelbar aus der obigen Definition.

Beispiel 1. Es sei $f(x) = x^n$ ($n \in \mathbf{N}$, $x \in \mathbf{R}$). Da f in allen Punkten a mit $a \in \mathbf{R}$ stetig ist und jede reelle Zahl a ein Häufungspunkt des Definitionsbereichs ist, besitzt f in jedem Punkt a einen Grenzwert, und zwar ist

$$\lim_{x \rightarrow a} x^n = a^n.$$

Beispiel 2. Es sei $f(x) = (\operatorname{sgn} x)^2$. Für alle x mit $x \neq 0$ ist

$$|f(x) - 1| = |(\operatorname{sgn} x)^2 - 1| = 0$$

und folglich

$$\lim_{x \rightarrow 0} f(x) = 1,$$

während $(\operatorname{sgn} 0)^2 = 0$ ist. Diese Funktion ist also an der Stelle $x = 0$ unstetig, aber sie besitzt dort einen Grenzwert, der vom Funktionswert verschieden ist.

Beispiel 3. Für jede reelle Zahl a ist

$$\lim_{x \rightarrow a} \frac{x^2 - a^2}{x - a} = 2a,$$

denn es ist

$$\left| \frac{x^2 - a^2}{x - a} - 2a \right| = |(x + a) - 2a| = |x - a| \quad (x \neq a).$$

Beispiel 4. Für die reelle (bzw. komplexe) Exponentialfunktion gilt

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1, \quad (6)$$

denn für $0 < |x| < 1$ ist

$$\left| \frac{e^x - 1}{x} - 1 \right| = \left| \frac{x}{2!} + \frac{x^2}{3!} + \cdots \right| < |x| \left(\frac{1}{2!} + \frac{1}{3!} + \cdots \right) < |x| \cdot e.$$

Beispiel 5. Es ist

$$\lim_{x \rightarrow 1} \frac{\ln x}{x - 1} = 1. \quad (7)$$

Beweis. Nach 2.1.5.(3) liegt der Quotient von $\ln x$ und $x - 1$ für $x > 0$, $x \neq 1$ stets zwischen den Zahlen x^{-1} und 1, und da aus $x \rightarrow 1$ stets $x^{-1} \rightarrow 1$ folgt, gilt (7).

Beispiel 6. Es ist

$$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \ln a \quad (a > 0). \quad (8)$$

Beweis. Für $a = 1$ ist die Behauptung trivial. Für $a \neq 1$ erhalten wir auf Grund von Satz 2 und Beispiel 4

$$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \ln a \cdot \lim_{x \rightarrow 0} \frac{e^{x \ln a} - 1}{x \ln a} = \ln a \cdot \lim_{y \rightarrow 0} \frac{e^y - 1}{y} = \ln a.$$

Beispiel 7. Es ist

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \quad (9)$$

Beweis. Nach 1.6.1.(5) ist $x \cdot \cos x < \sin x < x$ für $0 < x < 1$, woraus

$$\cos x < \frac{\sin x}{x} < 1$$

folgt. Dies bleibt richtig, wenn wir x durch $-x$ ersetzen. Für $x_n \neq 0$, $x_n \rightarrow 0$ gilt $\cos x_n \rightarrow 1$, und mit 2.1.3., Satz 4, ergibt sich die Behauptung.

Die obige Folgendefinition des Grenzwertes ist so allgemein gefaßt, daß sie auch die folgenden Sonderfälle umfaßt:

a) Im Fall einer reellwertigen Funktion kann der Grenzwert b auch ∞ oder $-\infty$ sein.

b) Ist $D(f) \subseteq \mathbb{R}$, so bezeichnen wir ∞ bzw. $-\infty$ als *uneigentlichen* Häufungspunkt des Definitionsbereichs, wenn dieser nicht nach oben bzw. nicht nach unten beschränkt ist. Dann kann auch a in der obigen Definition durch ∞ oder $-\infty$ ersetzt werden.

Wir erläutern dies an einigen Beispielen.

Beispiel 8. Es sei $f(x) = \frac{1}{x^2}$ ($x \neq 0$). Diese Funktion ist an der Stelle $x = 0$ nicht definiert, aber für jede Nullfolge (x_n) mit $x_n \neq 0$ gilt

$$\lim_{n \rightarrow \infty} \frac{1}{x_n^2} = \infty,$$

d. h., es ist

$$\lim_{x \rightarrow 0} \frac{1}{x^2} = \infty.$$

Beispiel 9. Es sei $f(x) = \frac{x^2}{1+x^2}$ ($x \in \mathbb{R}$). Für jede Folge (x_n) mit $\lim_{n \rightarrow \infty} x_n = \infty$ ist $\left(\frac{1}{x_n}\right)$ eine Nullfolge und folglich

$$\lim_{n \rightarrow \infty} \frac{x_n^2}{1+x_n^2} = \lim_{n \rightarrow \infty} \frac{1}{\frac{1}{x_n^2} + 1} = 1,$$

d. h., es ist

$$\lim_{x \rightarrow \infty} \frac{x^2}{1+x^2} = 1.$$

Beispiel 10. Wir betrachten die reelle ganzrationale Funktion

$$f(x) = a_0 + a_1x + \dots + a_nx^n \quad (x \in \mathbb{R}).$$

Ist $n \geq 1$, so ist

$$\lim_{x \rightarrow \infty} f(x) = \begin{cases} \infty & \text{für } a_n > 0, \\ -\infty & \text{für } a_n < 0 \end{cases} \quad (10)$$

sowie

$$\lim_{x \rightarrow -\infty} f(x) = \begin{cases} \infty & \text{für } (-1)^n a_n > 0, \\ -\infty & \text{für } (-1)^n a_n < 0. \end{cases} \quad (11)$$

Beweis. Nach 1.3.4.(3) haben $f(x)$ und a_nx^n für hinreichend große $|x|$ stets dasselbe Signum, und aus 1.3.4.(2) folgt

$$\begin{aligned} \lim_{x \rightarrow \infty} f(x) &= \lim_{x \rightarrow \infty} a_nx^n = \operatorname{sgn} a_n \lim_{x \rightarrow \infty} |a_nx^n|, \\ \lim_{x \rightarrow -\infty} f(x) &= \lim_{x \rightarrow -\infty} a_nx^n = (-1)^n \operatorname{sgn} a_n \lim_{x \rightarrow -\infty} |a_nx^n|. \end{aligned}$$

Hieraus können die Behauptungen abgelesen werden.

Aus den Grenzwertsätzen für Zahlenfolgen (vgl. 2.1.3.(5') bis (8') und 2.1.6. Satz 4) ergeben sich die folgenden Grenzwertsätze für Funktionen, die allerdings noch einer zusätzlichen Erläuterung bedürfen:

$$\lim_{x \rightarrow a} (f(x) \pm g(x)) = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} g(x), \quad (12)$$

$$\lim_{x \rightarrow a} (f(x) \cdot g(x)) = \lim_{x \rightarrow a} f(x) \cdot \lim_{x \rightarrow a} g(x), \quad (13)$$

$$\lim_{x \rightarrow a} |f(x)| = \left| \lim_{x \rightarrow a} f(x) \right|, \quad (14)$$

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} g(x)} \quad (15)$$

$$\lim_{x \rightarrow a} f(x)^r = \left[\lim_{x \rightarrow a} f(x) \right]^r \quad (f(x) > 0, r \in \mathbb{R}). \quad (16)$$

Die Beziehung (12) ist z. B. wie folgt zu interpretieren. Ist a ein Häufungspunkt von $D(f \pm g)$ (bzw. ist $D(f \pm g)$ im Fall $a = \pm\infty$ nicht nach oben bzw. nach unten beschränkt), existieren die rechts stehenden Grenzwerte, und sind sie endlich, so existiert auch der links stehende Grenzwert, und es gilt (12). Analoges gilt für (13) und (15), wobei in (15) noch gefordert werden muß, daß der im Nenner stehende Grenzwert von 0 verschieden ist. Auch in (14) kann aus der Existenz des rechts stehenden Grenzwertes auf die Existenz des links stehenden Grenzwertes geschlossen werden, aber nicht umgekehrt.

Abschließend nehmen wir wiederum einige Verallgemeinerungen vor. So können wir in der obigen Definition an Stelle von reellen Funktionen auch Funktionen aus einem metrischen Raum X in einen metrischen Raum Y setzen. Dann gilt

Satz 1'. Eine Funktion f aus einem metrischen Raum X in einen metrischen Raum Y hat in einem Häufungspunkt $a \in D(f)$ den Grenzwert b genau dann, wenn es zu jeder positiven reellen Zahl ε eine positive reelle Zahl δ gibt derart, daß aus

$$x \in D(f) \quad \text{und} \quad 0 < \varrho_X(x, a) < \delta$$

stets

$$\varrho_Y(f(x), b) < \varepsilon$$

folgt.

Auch die Folgerung aus Satz 1 bleibt richtig, wenn

$$\varrho_Z(g(x), c) \leq K \varrho_Y(f(x), b) \quad (x \in D(g) \setminus \{a\} \subseteq D(f)) \quad (1)$$

an Stelle von (1) gilt.

Die Beziehung (12) gilt auch für Funktionen in \mathbb{R}^p . Für Funktionen in \mathbb{R}^p gilt ferner

$$\lim_{x \rightarrow a} \|f(x)\| = \left\| \lim_{x \rightarrow a} f(x) \right\| \quad (14')$$

an Stelle von (14).

2.3.5. Einseitige Grenzwerte und einseitige Stetigkeit reeller Funktionen

Es sei f eine Funktion, deren Definitionsbereich in \mathbb{R} enthalten ist. Dann können wir für jede reelle Zahl a die Einschränkung der Funktion f auf die Menge

$$\{x: x \in D(f) \wedge x \geq a\}$$

bzw.

$$\{x: x \in D(f) \wedge x \leq a\}$$

betrachten und diese Einschränkung an der Stelle a auf Stetigkeit bzw. Existenz eines Grenzwertes untersuchen. Ist diese Einschränkung an der Stelle a stetig, so sagen wir, die Funktion f sei an der Stelle a *rechts-* bzw. *linksseitig stetig*. Besitzt die Einschränkung an der Stelle a einen Grenzwert b , so sagen wir, die Funktion f besitze an der Stelle a einen *rechts-* bzw. *linksseitigen* Grenzwert. Die Bezeichnung für einseitige Grenzwerte ist in der Literatur nicht einheitlich. Für rechtsseitige Grenzwerte

an der Stelle a werden die Bezeichnungen

$$\lim_{\substack{x \rightarrow a \\ x > a}} f(x) = b, \quad \lim_{x \rightarrow a+0} f(x) = b, \quad \lim_{x \downarrow a} f(x) = b,$$

für linksseitige Grenzwerte

$$\lim_{\substack{x \rightarrow a \\ x < a}} f(x) = b, \quad \lim_{x \rightarrow a-0} f(x) = b, \quad \lim_{x \uparrow a} f(x) = b$$

benutzt. Wir werden die letzte dieser Schreibweisen verwenden. Eine Funktion f ist nach unserer Definition in einem Häufungspunkt a ihres Definitionsbereichs rechts- bzw. linksseitig stetig genau dann, wenn

$$\lim_{x \downarrow a} f(x) = f(a)$$

bzw.

$$\lim_{x \uparrow a} f(x) = f(a)$$

ist. Stimmen rechts- und linksseitiger Grenzwert mit dem Funktionswert an der Stelle a überein, so ist die Funktion stetig.

Beispiel 1. Gegeben sei die Funktion $y = \operatorname{sgn} x$ ($x \in \mathbb{R}$). Für $x > 0$ bzw. $x < 0$ ist $|\operatorname{sgn} x - 1| = 0$ bzw. $|\operatorname{sgn} x - (-1)| = 0$ und folglich

$$\lim_{x \downarrow 0} \operatorname{sgn} x = 1 \quad \text{und} \quad \lim_{x \uparrow 0} \operatorname{sgn} x = -1,$$

d. h., rechts- und linksseitiger Grenzwert stimmen an der Stelle 0 nicht überein.

Beispiel 2. Gegeben sei die Funktion

$$f(x) = \frac{1}{1 + 2^{1/x}} \quad (x \in \mathbb{R}, x \neq 0).$$

Wegen

$$\lim_{x \downarrow 0} 2^{1/x} = \infty, \quad \lim_{x \uparrow 0} 2^{1/x} = \lim_{x \downarrow 0} \frac{1}{2^{1/x}} = 0$$

ist

$$\lim_{x \downarrow 0} \frac{1}{1 + 2^{1/x}} = 0, \quad \lim_{x \uparrow 0} \frac{1}{1 + 2^{1/x}} = 1.$$

Für monotone Funktionen können in diesem Zusammenhang besonders einfache Aussagen formuliert werden.

Satz. Für jede auf einem Intervall $[a, b]$ definierte monotone Funktion f existieren die einseitigen Grenzwerte

$$f(x-0) := \lim_{t \uparrow x} f(t) \quad (a < x \leq b), \quad (1)$$

$$f(x+0) := \lim_{t \downarrow x} f(t) \quad (a \leq x < b). \quad (2)$$

Beweis. Die Funktion f sei etwa monoton wachsend. Wir wählen zwei monoton wachsende Folgen (x_n) , (x'_n) mit $x_n, x'_n < x$ und $x_n \rightarrow x$, $x'_n \rightarrow x$. Dann sind die Folgen $(f(x_n))$, $(f(x'_n))$ monoton wachsend und durch $f(x)$ nach oben beschränkt. Folglich existieren die Grenzwerte

$$y := \lim_{n \rightarrow \infty} f(x_n), \quad y' := \lim_{n \rightarrow \infty} f(x'_n).$$

Nach unseren Voraussetzungen über die Folgen (x_n) , (x'_n) existiert eine Teilfolge (x_{k_n}) mit $x'_n \leq x_{k_n} < x$, woraus $f(x'_n) \leq f(x_{k_n})$ folgt. Der Grenzübergang $n \rightarrow \infty$ ergibt $y' \leq y$. Ebenso zeigt man, daß $y \leq y'$ ist. Daher ist $y = y'$ der linksseitige Grenzwert von f an der Stelle x . Analog werden die Behauptungen in den anderen Fällen bewiesen.

Für monoton wachsende Funktionen gilt offensichtlich stets

$$f(x-0) \leq f(x) \leq f(x+0), \quad (3)$$

und f ist an der Stelle x mit $a < x < b$ genau dann stetig, wenn $f(x-0) = f(x+0)$ ist. Ist $f(x-0) < f(x+0)$, so heißt x eine *Sprungstelle* der monoton wachsenden Funktion f .

Für zwei verschiedene Sprungstellen x_1, x_2 sind die Intervalle $]f(x_1-0), f(x_1+0)[$, $]f(x_2-0), f(x_2+0)[$ disjunkt. Da jedes solche Intervall mindestens eine rationale Zahl enthält, kann man hieraus schließen, daß die Menge der Sprungstellen einer monotonen Funktion höchstens abzählbar ist.

2.4. Stetige Funktionen

2.4.1. Stetigkeit in Punktmengen

Der Definitionsbereich einer nichtleeren reellen (bzw. komplexen) Funktion kann stets in zwei disjunkte Punktmengen zerlegt werden, in die Menge der *Stetigkeitspunkte* und in die Menge der *Unstetigkeitspunkte* der Funktion. Jede dieser beiden Teilmengen kann leer sein. So besitzen z. B. die Potenzfunktionen keine Unstetigkeitspunkte. Die Funktion

$$f(x) = \begin{cases} 0 & \text{für } x \text{ rational,} \\ 1 & \text{für } x \text{ irrational} \end{cases} \quad (1)$$

besitzt dagegen keine Stetigkeitspunkte, denn zu jeder reellen Zahl a gibt es eine Folge rationaler Zahlen und eine Folge irrationaler Zahlen mit dem Grenzwert a . Die (stationären) Folgen der zugehörigen Funktionswerte haben dann die verschiedenen Grenzwerte 0 bzw. 1, d. h., die Funktion f ist an der Stelle a unstetig. Die Funktion $y = \operatorname{sgn} x$ ($x \in \mathbb{R}$) hat nur den Punkt 0 als Unstetigkeitspunkt.

Wir verallgemeinern den zunächst nur für einzelne Punkte des Definitionsbereichs definierten Begriff der Stetigkeit einer Funktion.

Definition. Eine reelle (bzw. komplexe) Funktion f heißt *in der Menge M stetig*, wenn $M \subseteq D(f)$ und wenn f in allen Punkten von M stetig ist. Die Funktion f heißt *stetig*, wenn sie in ihrem ganzen Definitionsbereich stetig ist.

Aus den bereits bewiesenen Sätzen folgt sofort, daß alle Potenzfunktionen, die trigonometrischen Funktionen, die spezielle Exponentialfunktion und die natürliche Logarithmusfunktion stetig sind. Wir betrachten weitere Beispiele.

Beispiel 1. Für alle nichtnegativen rationalen Zahlen r sind die Funktionen

$$y = x^r \quad (x \in \mathbb{R}, x \geq 0) \quad (2)$$

auf Grund von 2.1.3., Satz 5, stetig.

Beispiel 2. Die Stetigkeit der Exponentialfunktion

$$y = a^x \quad (x \in \mathbb{R}) \quad (3)$$

ergibt sich wegen $\ln a^x = x \ln a$, $a^x = e^{x \ln a}$ erneut aus der Stetigkeit der speziellen Exponentialfunktion und aus 2.3.3., Satz 2 und 1.

Beispiel 3. Für jede reelle Zahl α ist die Funktion

$$y = x^\alpha \quad (x \in \mathbb{R}, x > 0) \quad (4)$$

stetig, denn wegen $\ln x^\alpha = \alpha \ln x$ ist

$$x^\alpha = e^{\alpha \ln x}. \quad (5)$$

Beispiel 4. Als Umkehrfunktion von (3) ist die Funktion

$$y = \log_a x \quad (x \in \mathbb{R}, x > 0) \quad (6)$$

nach 2.3.3., Satz 3, stetig. Die Behauptung folgt auch aus der Identität

$$\log_a x = \frac{\ln x}{\ln a}. \quad (7)$$

Eine unmittelbare Folgerung aus 2.3.3., Satz 1, ist

Satz 2. Jede Linearkombination von stetigen Funktionen, die einen gemeinsamen Definitionsbereich besitzen, ist eine stetige Funktion.

Beispiel 5. Alle ganzrationalen Funktionen sind stetig. Ferner sind alle rationalen Funktionen als Quotient zweier stetiger Funktionen stetig. Es ist inkorrekt, wenn man die möglicherweise vorhandenen Polstellen, d. h. die Nullstellen des Nenners (die nicht zugleich Nullstellen des Zählers sind), als „Unstetigkeitsstellen“ bezeichnet. Eine Unstetigkeitsstelle einer Funktion gehört stets zu ihrem Definitionsbereich, und dies ist für die Polstellen einer ganzrationalen Funktion nicht der Fall.

2.4.2. Auf abgeschlossenen Mengen stetige Funktionen

Nach 1.5.5., Satz 3, ist eine Menge F genau dann abgeschlossen, wenn sie alle ihre Häufungspunkte enthält. Die abgeschlossenen Intervalle $\llbracket a, b \rrbracket$ sind einfache Beispiele für abgeschlossene Mengen. Funktionen, die auf abgeschlossenen beschränkten Mengen stetig sind, haben wichtige Eigenschaften, die für zahlreiche Anwendungen bedeutungsvoll sind. Als erstes beweisen wir den Satz vom Maximum und Minimum.

Satz 1. Der Wertebereich $W(f)$ einer auf einer beschränkten und abgeschlossenen Menge definierten stetigen reellen Funktion f ist beschränkt und abgeschlossen. Er besitzt demzufolge ein Maximum und ein Minimum (Abb. 2.16).

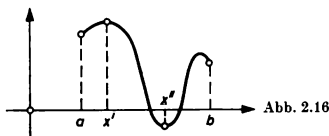


Abb. 2.16

Beweis. Nehmen wir an, der Wertebereich wäre nicht beschränkt bzw. nicht abgeschlossen. Dann gibt es eine Folge von Elementen $y_n \in W(f)$ mit $y_n \rightarrow y$ und mit $y = \pm\infty$ bzw. $y \notin W(f)$. Wir wählen Elemente $x_n \in D(f)$ mit $y_n = f(x_n)$. Die Folge (x_n) ist beschränkt und besitzt nach dem Satz von BOLZANO-WEIERSTRASS eine konvergente Teilfolge (x_{n_k}) . Ihr Grenzwert x ist Element der abgeschlossenen Menge $D(f)$, und aus der Stetigkeit der Funktion f folgt

$$f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{k \rightarrow \infty} y_{n_k} = y.$$

Das widerspricht unserer Voraussetzung, und folglich ist $W(f)$ beschränkt und abgeschlossen. Zu $y := \sup W(f)$ gibt es nach 2.1.3., Satz 8, eine gegen y konvergierende Folge von Elementen $y_n \in W(f)$. Wörtlich wie in dem soeben geführten Beweis zeigt man, daß ein Element $x \in D(f)$ mit $f(x) = y$ existiert, d. h., es ist $y \in W(f)$, und y ist das Maximum des Wertebereichs. Ganz analog wird der Beweis für die Existenz des Minimums geführt.

Mit Hilfe der in 2.1.7., Definitionen 3 und 4, eingeführten Begriffsbildungen kann dieser Satz wie folgt verallgemeinert werden.

Satz 1'. Der Wertebereich einer auf einer kompakten Teilmenge eines metrischen Raumes X definierten stetigen Funktion f in einen metrischen Raum Y ist kompakt.

Beweis. Es sei (y_n) eine Folge in $W(f)$ und $y_n = f(x_n)$. Da $D(f)$ kompakt ist, gibt es eine gegen ein Element $x \in D(f)$ konvergierende Teilfolge (x_{n_k}) . Aus der Stetigkeit von f folgt

$$f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{k \rightarrow \infty} y_{n_k}.$$

Nach 2.1.7., Definition 4, ist $W(f)$ kompakt.

Der nachfolgende *Satz von Bolzano* macht eine wichtige Aussage über die Existenz einer Nullstelle einer Funktion. Diesem Satz ist wesentlich BOLZANOS berühmte Arbeit „Rein analytischer Beweis, daß zwischen zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege“ von 1817 gewidmet. Sie ist für die Theorie der reellen Funktionen von grundlegender Bedeutung.

Satz 2. Nimmt eine auf einem abgeschlossenen beschränkten Intervall stetige reelle Funktion in den Endpunkten dieses Intervalls Funktionswerte mit entgegengesetzten Vorzeichen an, so besitzt die Funktion im Innern dieses Intervalls eine Nullstelle.

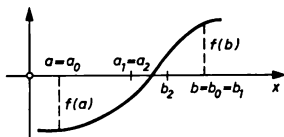


Abb. 2.17

Beweis. Wir konstruieren mit Hilfe des Halbierungsverfahrens eine Intervallschachtelung $(a_n | b_n)$. Ohne Beschränkung der Allgemeinheit sei $f(a) < 0$ und $f(b) > 0$ (Abb. 2.17). Wir setzen $a_0 := a$ und $b_0 := b$. Dann ist $a_0 < b_0$, $f(a_0) \leq 0$, $f(b_0) > 0$ und $b_0 - a_0 = \frac{b-a}{2^0}$.

Die Zahlen $a_0, a_1, \dots, a_n; b_0, b_1, \dots, b_n$ seien schon konstruiert, und es gelte

$$\left. \begin{aligned} a_0 \leq a_1 \leq \dots \leq a_n < b_n \leq \dots \leq b_1 \leq b_0, \\ f(a_n) \leq 0, \quad f(b_n) > 0 \quad \text{und} \quad b_n - a_n = \frac{b-a}{2^n}. \end{aligned} \right\} \quad (1)$$

Wir halbieren das Intervall $[[a_n, b_n]]$. Im Fall

$$f\left(\frac{a_n + b_n}{2}\right) \leq 0 \quad \text{bzw.} \quad f\left(\frac{a_n + b_n}{2}\right) > 0$$

setzen wir

$$a_{n+1} := \frac{a_n + b_n}{2} \quad \text{bzw.} \quad a_{n+1} := a_n$$

und

$$b_{n+1} := b_n \quad \text{bzw.} \quad b_{n+1} := \frac{a_n + b_n}{2}.$$

Die Aussagen (1) gelten dann mit $n + 1$ statt n , und die Folgen (a_n) , (b_n) sind induktiv definiert. Sie bilden wegen (1) eine Intervallschachtelung und besitzen daher einen gemeinsamen Grenzwert x . Wegen der Stetigkeit von f besitzen die Folgen $(f(a_n))$, $(f(b_n))$ den gemeinsamen Grenzwert $f(x)$. Als Grenzwert einer Folge nichtpositiver bzw. positiver Zahlen kann $f(x)$ weder positiv noch negativ sein. Somit ist $f(x) = 0$, und der Satz ist bewiesen. Der Beweis ist *konstruktiv*, d. h., er liefert ein Verfahren, wie man eine Nullstelle einer Funktion mit den geforderten Eigenschaften numerisch mit jeder geforderten Genauigkeit ermitteln kann.

Mit Hilfe des Satzes von BOLZANO beweisen wir nun den *Zwischenwertsatz* für beliebige stetige Funktionen.

Satz 3. *Jede auf einem abgeschlossenen beschränkten Intervall stetige reelle Funktion nimmt jede reelle Zahl, die zwischen den Funktionswerten in den Endpunkten des Intervalls liegt, in wenigstens einem Punkt dieses Intervalls als Funktionswert an.*

Beweis. Die Funktion f sei in $\llbracket a, b \rrbracket$ stetig, und es sei etwa $f(a) < c < f(b)$. Wir setzen $g(x) := f(x) - c$. Dann ist g eine stetige Funktion mit $g(a) = f(a) - c < 0 < f(b) - c = g(b)$. Nach Satz 2 gibt es ein x mit $a < x < b$ und $0 = g(x) = f(x) - c$.

Satz 4. *Jede reelle stetige Funktion, deren Definitionsbereich ein Intervall ist, nimmt mit zwei Werten y_1, y_2 auch jeden Zwischenwert y an, und folglich ist ihr Wertebereich ein Intervall.*

Beweis. Es gibt Punkte x_1, x_2 mit $f(x_1) = y_1, f(x_2) = y_2$, und es sei etwa $x_1 \leq x_2$. Die Behauptung folgt nun, wenn wir Satz 3 auf die Einschränkung von f auf das Intervall $\llbracket x_1, x_2 \rrbracket$ anwenden.

Hiernach und nach 2.3.3., Satz 3, besitzt jede auf einem Intervall definierte stetige streng monotone Funktion eine auf einem Intervall definierte stetige streng monotone Umkehrfunktion.

2.4.3. Gleichmäßige Stetigkeit

Den folgenden Begriff werden wir in der Integralrechnung benötigen.

Definition. Eine reelle oder komplexe Funktion heißt *gleichmäßig* stetig, wenn es zu jeder positiven reellen Zahl ε eine positive reelle Zahl δ gibt derart, daß aus $x', x'' \in D(f)$ und $|x' - x''| < \delta$ stets $|f(x') - f(x'')| < \varepsilon$ folgt.

Wir vergleichen die Definition der Stetigkeit und der gleichmäßigen Stetigkeit:

$$f \text{ stetig} \Leftrightarrow \bigwedge_{x' \in D(f)} \bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x'' \in D(f)} (|x' - x''| < \delta \Rightarrow |f(x') - f(x'')| < \varepsilon), \quad (1)$$

f gleichmäßig stetig : \Leftrightarrow

$$\bigwedge_{\varepsilon > 0} \bigvee_{\delta > 0} \bigwedge_{x \in D(f)} \bigwedge_{x' \in D(f)} (|x' - x''| < \delta \Rightarrow |f(x') - f(x'')| < \varepsilon). \quad (2)$$

Die unterschiedliche Reihenfolge der Quantifikatoren ist hierbei von wesentlicher Bedeutung. Während (1) aus (2) gefolgert werden kann, ist das Umgekehrte nicht der Fall. In (1) kann nämlich die zu bestimmende Zahl δ neben ε auch von der betrachteten Stelle x' abhängen. In (2) darf dagegen δ nur von ε abhängig sein. Als Beispiel betrachten wir die Funktion $f(x) = x^{-1}$ ($0 < x \leq 1$). Setzen wir etwa

$\varepsilon := \frac{1}{2}$ und ist $\delta > 0$ beliebig vorgegeben, so wählen wir eine natürliche Zahl n mit $n\delta > 1$ und setzen

$$x' := \frac{1}{n}, \quad x'' := \frac{1}{n+1}.$$

Dann ist $|x' - x''| < \frac{1}{n} < \delta$, aber $|f(x') - f(x'')| = |n - (n+1)| = 1 > \varepsilon$, und die

Negation von (2) ist wahr. Die Funktion f ist im betrachteten Intervall zwar stetig, aber nicht gleichmäßig stetig. Es gilt jedoch der

Satz. Jede reelle oder komplexe stetige Funktion, deren Definitionsbereich eine beschränkte abgeschlossene Menge ist, ist auch gleichmäßig stetig.

Beweis. Nehmen wir an, die Funktion f sei nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$ derart, daß zu jedem $\delta > 0$ Elemente $x', x'' \in D(f)$ mit $|x' - x''| < \delta$ und $|f(x') - f(x'')| \geq \varepsilon$ gefunden werden können. Zu $\delta := \frac{1}{n}$ wählen wir Elemente $x_n', x_n'' \in D(f)$ mit

$$|x_n' - x_n''| < \frac{1}{n} \quad (3)$$

und $|f(x_n') - f(x_n'')| \geq \varepsilon$. Die Folge (x_n') besitzt auf Grund unserer Voraussetzungen eine gegen ein Element $x \in D(f)$ konvergierende Teilfolge (x_{n_k}') . Die Folge (x_{n_k}'') hat dann wegen (3) ebenfalls den Grenzwert x , und aus der Stetigkeit von f ergibt sich der Widerspruch

$$\varepsilon \leq \lim_{k \rightarrow \infty} |f(x_{n_k}') - f(x_{n_k}'')| = |f(x) - f(x)| = 0.$$

Der Satz und sein Beweis können unverändert auf reell- oder komplexwertige Funktionen übertragen werden, deren Definitionsbereich eine kompakte Teilmenge eines metrischen Raumes ist.

2.4.4. Nullstellen ganzrationaler Funktionen

Die Frage nach der Existenz und Anzahl der Nullstellen einer ganzrationalen Funktion ist eines der klassischen Probleme der Mathematik. Wir formulieren und beweisen einige der Sätze, die zu diesem umfangreichen Problemkreis gehören.

Wir betrachten zuerst reelle ganzrationale Funktionen. Bekanntlich besitzt jede nichtkonstante lineare Funktion $f(x) = a_0 + a_1x$ ($a_1 \neq 0$) (genau) eine Nullstelle. Allgemeiner gilt der

Satz 1. Jede reelle ganzrationale Funktion von ungeradem Grad besitzt (mindestens) eine reelle Nullstelle.

Beweis. Wegen 2.3.4.(10), (11) (vgl. Abb. 1.5) nimmt eine ganzrationale Funktion f vom Grad $n = 2k + 1$ positive und negative Funktionswerte an. Da sie überall stetig ist, liegt zwischen zwei Stellen a, b mit $f(a)f(b) < 0$ eine Nullstelle der Funktion.

Eine Satz 1 entsprechende Aussage gilt aber nicht für reelle ganzrationale Funktionen von geradem Grad. So kann z. B. eine quadratische Funktion $f(x) = a_0 + a_1x + a_2x^2$ ($a_2 \neq 0$) zwei verschiedene reelle Nullstellen, genau eine oder auch gar keine reelle Nullstelle besitzen (Abb. 2.18).

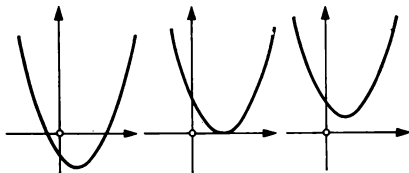


Abb. 2.18

Die Situation ändert sich aber grundlegend, wenn wir zu komplexen ganzrationalen Funktionen übergehen. So besitzt z. B. die Funktion f mit $f(x) = x^2 + 1$ keine reelle Nullstelle, aber die komplexen Nullstellen $\pm i$. In Satz 1 in 1.2.3. haben wir allgemeiner gezeigt, daß die Gleichung $x^2 = c$ mit $c \in \mathbb{C}$ mindestens eine Lösung im Bereich der komplexen Zahlen besitzt.

Eine ganzzahlige Nullstelle x_0 einer ganzrationalen Funktion mit ganzzahligen Koeffizienten kann durch Probieren gefunden werden, denn aus

$$a_0 + x_0(a_1 + a_2x_0 + \dots + a_nx_0^{n-1}) = 0$$

ersehen wir, daß x_0 stets Teiler des absoluten Gliedes a_0 ist.

Hat man eine Nullstelle x_0 einer ganzrationalen Funktion f gefunden, so kann die Ermittlung weiterer Nullstellen auf die Untersuchung einer ganzrationalen Funktion zurückgeführt werden, deren Grad um 1 kleiner als der Grad von f ist. Wegen

$$f(x) = a_0 + a_1x + \dots + a_nx^n, \quad (1)$$

$$0 = f(x_0) = a_0 + a_1x_0 + \dots + a_nx_0^n \quad (2)$$

ist nämlich

$$f(x) - f(x_0) = a_1(x - x_0) + \dots + a_n(x^n - x_0^n) \quad (3)$$

und alle Summanden lassen sich nach 1.1.1.(7) als Produkt von $x - x_0$ und einer ganzrationalen Funktion darstellen. Daher ist

$$f(x) = (x - x_0)g(x), \quad (4)$$

und $g(x)$ kann nach dem bekannten Verfahren der Restdivision bestimmt werden. Eine Verallgemeinerung dieses Sachverhalts liefert

Satz 2. Zu jeder ganzrationalen Funktion f und zu jedem x_0 mit $x_0 \in D(f)$ gibt es eine eindeutig bestimmte positive natürliche Zahl p und eine eindeutig bestimmte ganzrationale Funktion g mit

$$f(x) = f(x_0) + (x - x_0)^p g(x) \quad \text{und} \quad g(x_0) \neq 0. \quad (5)$$

Beweis. Es sei p die größte natürliche Zahl, für die $(x - x_0)^p$ Teiler von $f(x) - f(x_0)$ ist. Dann ist $p > 0$ und

$$f(x) - f(x_0) = (x - x_0)^p g(x). \quad (6)$$

Nehmen wir an, es wäre $g(x_0) = 0$. Dann hat g einen positiven Grad, und analog zu (4) erkennen wir, daß es eine ganzrationale Funktion h mit $g(x) = (x - x_0)h(x)$ gibt. Dann ist $f(x) - f(x_0) = (x - x_0)^{p+1}h(x)$, und dies widerspricht der Definition von p . Somit gilt (5). Es sei auch $f(x) - f(x_0) = (x - x_0)^q h(x)$ und $h(x_0) \neq 0$. Ohne Beschränkung der Allgemeinheit sei $q \leq p$. Wäre $q < p$, so folgte $(x - x_0)^{p-q}g(x) = h(x)$. Für $x = x_0$ ergäbe sich $0 = h(x_0)$, was nicht möglich ist. Es folgt $p = q$, $g(x) = h(x)$, und Satz 2 ist bewiesen.

Wir beweisen nun eines der klassischen Resultate der Mathematik von grundlegender Bedeutung, den sogenannten *Fundamentalsatz der klassischen Algebra*. Er wurde erstmalig von GAUSS in seiner im Alter von 21 Jahren verfaßten Dissertation bewiesen.

Satz 3. Jede komplexe ganzrationale Funktion von positivem Grad besitzt eine (komplexe) Nullstelle.

Beweis. Es sei $n \in \mathbf{N}$, $n \geq 1$, $a_n \neq 0$ und

$$f(x) = a_0 + a_1x + \cdots + a_nx^n \quad (x \in \mathbf{C}). \quad (7)$$

Wir setzen

$$m := \inf \{|f(x)| : x \in D(f)\}.$$

Nach 1.3.4. (2) gibt es eine positive reelle Zahl R mit

$$|f(x)| \geq m + 1 \quad \text{für} \quad |x| \geq R. \quad (8)$$

Die Funktion $y = |f(x)|$ ist nach 2.3.2., Satz 3, auf der abgeschlossenen beschränkten Menge $\{x : |x| \leq R\}$ stetig und besitzt dort nach dem Satz vom Maximum und Minimum (2.4.2., Satz 1) ein Minimum, d. h., es gibt ein x_0 mit $|x_0| \leq R$ und mit

$$|f(x_0)| \leq |f(x)| \quad \text{für} \quad |x| \leq R. \quad (9)$$

Nach Definition des Infimums gibt es ein x_1 mit $|f(x_1)| \leq m + \frac{1}{2}$, da sonst $m + \frac{1}{2}$ eine untere Schranke für alle Zahlen $|f(x)|$ mit $x \in D(f)$ wäre. Aus (8) folgt dann $|x_1| < R$. Wegen (9), (8) ist

$$|f(x_0)| \leq |f(x_1)| < m + 1 \leq |f(x)| \quad (|x| \geq R).$$

Somit ist

$$|f(x_0)| \leq |f(x)| \quad (x \in \mathbf{C}). \quad (10)$$

Nehmen wir an, es wäre $f(x_0) \neq 0$. Nach Satz 2 kann f in der Form (5) dargestellt werden. Wir setzen

$$h(z) := \frac{f(x_0 + z)}{f(x_0)}, \quad q(z) := \frac{g(x_0 + z)}{f(x_0)}, \quad q(0) = a + ib \quad (a, b \in \mathbf{R}).$$

Nach (5) ist $h(z) = 1 + z^p q(z)$, und wegen (10) ist $|h(z)| \geq 1$. Somit gilt

$$1 \leq |h(z)|^2 = |1 + z^p q(z)|^2 = 1 + 2 \operatorname{Re}(z^p q(z)) + |z^p q(z)|^2,$$

$$0 \leq 2 \operatorname{Re}(z^p q(z)) + |z^p q(z)|^2$$

für alle z mit $z \in \mathbf{C}$. Ersetzen wir z durch $\frac{z}{n}$ ($n = 1, 2, \dots$) und multiplizieren mit n^p , so erhalten wir

$$0 \leq 2 \operatorname{Re}\left(z^p q\left(\frac{z}{n}\right)\right) + \frac{1}{n^p} \left|z^p q\left(\frac{z}{n}\right)\right|^2.$$

Da die Funktion q ganzrational, also stetig ist, folgt, wenn wir n gegen ∞ gehen lassen,

$$0 \leq 2 \operatorname{Re}(z^p q(0)) = 2 \operatorname{Re}(z^p(a + ib)).$$

Wählen wir nun für z der Reihe nach die komplexen Zahlen $1, z_0, z_0^2, z_0^3$, wobei z_0 der Gleichung $z_0^p = i$ genügt (vgl. 1.2.3., Satz 2), so erhalten wir

$$0 \leq \operatorname{Re}(a + ib) = a,$$

$$0 \leq \operatorname{Re}(i(a + ib)) = -b,$$

$$0 \leq \operatorname{Re}(i^2(a + ib)) = -a,$$

$$0 \leq \operatorname{Re}(i^3(a + ib)) = b.$$

Daher ist $q(0) = a + ib = 0$, was wegen $g(x_0) \neq 0$ ein Widerspruch ist. Somit ist $f(x_0) = 0$, und der Fundamentalsatz ist bewiesen.

Mit Hilfe des Fundamentalsatzes beweisen wir den Satz von der Zerlegung in Linearfaktoren.

Satz 4. Jede komplexe ganzrationale Funktion f vom Grad n ($n > 0$) kann — von der Reihenfolge der Faktoren abgesehen — auf genau eine Weise in der Form

$$f(x) = a(x - x_1)^{p_1}(x - x_2)^{p_2} \dots (x - x_k)^{p_k} \quad (11)$$

dargestellt werden. Hierbei ist a eine von Null verschiedene komplexe Zahl, x_1, x_2, \dots, x_k sind voneinander verschiedene komplexe Zahlen, und p_1, p_2, \dots, p_k sind positive natür-

liche Zahlen mit

$$p_1 + p_2 + \cdots + p_k = n. \quad (12)$$

Beweis. Ist $n = 1$, so gilt $f(x) = a_0 + a_1x$, also $f(x) = a(x - x_1)^1$ mit $a := a_1$, $x_1 := -a_1^{-1}a_0$, und diese Darstellung ist durch f eindeutig bestimmt.

Es gebe für alle ganzrationalen Funktionen mit $1 \leq \deg f \leq n - 1$ genau eine Darstellung der Form (11). Ist dann $\deg f = n$ und x_1 eine Nullstelle von f , so kann f nach Satz 2 auf genau eine Weise in der Form

$$f(x) = (x - x_1)^{p_1} g(x) \quad (g(x_1) \neq 0)$$

dargestellt werden. Hierbei ist $g(x)$ eine konstante Funktion, oder auf $g(x)$ kann die Induktionsvoraussetzung angewendet werden. In beiden Fällen ergibt sich wegen $g(x_1) \neq 0$ die geforderte Darstellung, und diese ist eindeutig bestimmt. Damit ist Satz 4 bewiesen.

Offenbar ist a der Koeffizient von x^n in der Darstellung (7), und x_1, x_2, \dots, x_k sind Nullstellen von f . Da $f(x) \neq 0$ für $x \neq x_i$ ($i = 1, 2, \dots, k$) ist, besitzt f keine weiteren Nullstellen. Die Zahl p_i heißt die *Vielfachheit* der Nullstelle x_i . Zählen wir jede Nullstelle x_i so oft, wie ihre Vielfachheit angibt, so gilt der

Satz 5. Jede komplexe ganzrationale Funktion n -ten Grades besitzt genau n komplexe Nullstellen.

Für komplexe ganzrationale Funktionen mit reellen Koeffizienten können wir noch eine zusätzliche Aussage machen.

Satz 6. Besitzt eine komplexe ganzrationale Funktion f mit reellen Koeffizienten a_0, a_1, \dots, a_n eine nichtreelle Nullstelle x_0 der Vielfachheit p , so ist auch die konjugiert komplexe Zahl \bar{x}_0 eine Nullstelle von der gleichen Vielfachheit.

Beweis. Wegen 1.2.2.(5), (6) und $\overline{a_i} = a_i$ ($i = 0, 1, \dots, n$) ist

$$\overline{f(x)} = \overline{a_0 + a_1x + \cdots + a_nx^n} = a_0 + a_1\bar{x} + \cdots + a_n\bar{x}^n = f(\bar{x}).$$

Aus (11) folgt

$$\begin{aligned} f(\bar{x}) &= a_n(\bar{x} - x_1)^{p_1} (\bar{x} - x_2)^{p_2} \cdots (\bar{x} - x_k)^{p_k}, \\ f(x) &= \overline{f(\bar{x})} = a_n(x - \bar{x}_1)^{p_1} (x - \bar{x}_2)^{p_2} \cdots (x - \bar{x}_k)^{p_k}. \end{aligned}$$

Ist etwa x_1 nichtreell, so ist $\bar{x}_1 \neq x_1$, und da die Darstellung (11) bis auf die Reihenfolge der Faktoren eindeutig bestimmt ist, muß es ein j mit $x_j = \bar{x}_1$ und $p_j = p_1$ geben. Damit ist Satz 6 bewiesen.

Für reelle ganzrationale Funktionen gilt der Satz von der Zerlegung in Linearfaktoren und quadratische Faktoren ohne reelle Nullstellen.

Satz 7. Jede reelle ganzrationale Funktion f von positivem Grad kann – von der Reihenfolge der Faktoren abgesehen – auf genau eine Weise in der Form

$$f(x) = a(f_1(x))^{p_1} \cdots (f_k(x))^{p_k} \quad (13)$$

dargestellt werden, wobei die Funktionen f_1, f_2, \dots, f_k paarweise verschieden sind und für alle j mit $j = 1, 2, \dots, k$ entweder

$$f_j(x) = x - x_j \quad (x_j \in \mathbb{R}) \quad (14)$$

oder

$$f_j(x) = ((x - b_j)^2 + c_j^2) \quad (b_j, c_j \in \mathbb{R}; c_j \neq 0) \quad (15)$$

gilt.

Beweis. Ist die Summe der Vielfachheit der reellen Nullstellen von f gleich dem Grad der Funktion, so gilt die Darstellung (11), d. h., es gilt (14) für $j = 1, 2, \dots, k$. Andernfalls besitzt f — als komplexe Funktion aufgefaßt — eine nichtreelle Nullstelle, etwa x_1 . Dann ist nach Satz 6 auch \bar{x}_1 eine Nullstelle, und f kann in der Form

$$f(x) = (x - x_1)^{p_1} (x - \bar{x}_1)^{p_1} g(x)$$

mit $g(x_1) \neq 0, g(\bar{x}_1) \neq 0$ dargestellt werden.

Wegen

$$(x - x_1)(x - \bar{x}_1) = x^2 - 2x \operatorname{Re} x_1 + |x_1|^2 = (x - \operatorname{Re} x_1)^2 + (\operatorname{Im} x_1)^2$$

gilt, wenn wir $b_1 := \operatorname{Re} x_1$ und $c_1 := \operatorname{Im} x_1$ setzen,

$$f(x) = ((x - b_1)^2 + c_1^2)^{p_1} g(x).$$

Wenden wir auf $g(x)$ dieselbe Betrachtung an, so erhalten wir in endlich vielen Schritten die Darstellung (13), wobei wegen $g(x_1), g(\bar{x}_1) \neq 0$ nicht noch einmal der gleiche Faktor $f_1(x)$ auftreten kann.

Die Eindeutigkeit der Darstellung (13) ergibt sich aus der Eindeutigkeit der Darstellung (11).

Obwohl die Existenz komplexer Nullstellen ganzrationaler Funktionen von positivem Grad gesichert ist, stellt ihre numerische Ermittlung in vielen Fällen ein außerordentlich schwieriges Problem dar, das im allgemeinen nur näherungsweise lösbar ist.

Die Kenntnis der Auflösung linearer Gleichungen kann bis in das 2. Jahrtausend v. u. Z. zurückverfolgt werden. Auch quadratische Gleichungen werden schon sehr früh, etwa in der babylonischen Mathematik, im Zusammenhang mit praktischen Problemen behandelt, beispielsweise bei der Berechnung von Dammbauten.

In der griechisch-hellenistischen Mathematik konnten auch allgemeine quadratische Gleichungen einwandfrei mit geometrischen Methoden gelöst werden, wobei dann auf Grund dieser Methode fast ausschließlich nur positive Lösungen zugelassen werden. Den Arabern verdankt man die geometrische Auflösung der Gleichungen dritten Grades durch Kegelschnitte. Trotz vieler Versuche gelang jedoch eine arithmetische Lösung der kubischen Gleichungen nicht.

In der Renaissance, als die Bedeutung der Mathematik für praktische Probleme deutlicher hervortrat und ein allgemeines Interesse an mathematischen Fragen hervorrief, wandte man sich auch diesem schwierigen Problem erneut zu.

SCIPIO DEL FERRO (1465?–1526) gelang dann die Lösung dieses Problems. Unabhängig von ihm fand auch NICCOLO TARTAGLIA (1500?–1557) eine Lösung. TARTAGLIA hatte sein Verfahren HIERONIMO CARDANO (1501–1576) nach langem Widerstreben mitgeteilt und um

Geheimhaltung gebeten. Als CARDANO später feststellte, daß bereits DEL FERRO eine Lösung besessen hatte, fühlte er sich an das gegebene Versprechen nicht mehr gebunden und veröffentlichte das Verfahren. Daher heißen die entsprechenden Formeln unberechtigterweise auch „Cardanische Formeln“.

Im Zusammenhang mit der Lösung der Gleichungen dritten Grades treten beim „casus irreducibilis“ auch erstmals komplexe Lösungen als Wurzeln auf.

Die innerlogische Entwicklung der Mathematik führte dann 1545 LUDOVICO FERRARI (1522–1565) auch zur Lösung der allgemeinen Gleichung vierten Grades, und man versuchte nun, auch Gleichungen noch höheren Grades allgemein zu behandeln und Lösungen in Radikalen zu finden, d. h. durch ineinandergeschachtelte Wurzeln anzugeben.

Wie wir bereits in 1.3.4. bemerkt haben, wurde von ABEL im Jahre 1826 bewiesen, daß Gleichungen fünften und höheren Grades im allgemeinen durch Radikale nicht lösbar sind. Bereits vor ABEL hatte PAOLO RUFFINI (1765–1822) eine im wesentlichen vollständige Lösung dieses Problems gegeben.

2.4.5. Der Banachsche Fixpunktsatz

Für die Behandlung zahlreicher mathematischer Probleme, die mit der Auflösung von linearen und nichtlinearen Gleichungen bzw. Gleichungssystemen zusammenhängen, ist ein nach dem polnischen Mathematiker STEFAN BANACH (1892–1945) benanntes konstruktives Verfahren von universeller Bedeutung entwickelt worden. Wir betrachten zuerst eine stetige Funktion f , die ein beschränktes abgeschlossenes Intervall $I = [a, b]$ in sich abbildet. Setzen wir $g(x) := f(x) - x$, so genügt die stetige Funktion g wegen $a \leq f(a)$ und $f(b) \leq b$ den Bedingungen $g(b) \leq 0 \leq g(a)$ und besitzt nach dem Nullstellensatz eine Nullstelle c . Somit ist $f(c) = g(c) + c = c$. Der Punkt c heißt wegen $f(c) = c$ ein *Fixpunkt* der Abbildung f (Abb. 2.19).

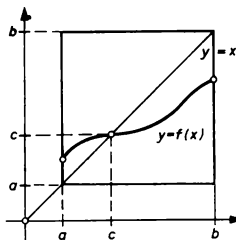


Abb. 2.19

Wir machen nun eine zusätzliche Voraussetzung, die uns neben der Existenz auch die Einzigkeit eines Fixpunktes sichert. Wir fordern, daß der Anstieg einer Kurvensekante dem Betrag nach eine positive reelle Zahl q mit $q < 1$ nicht überschreitet, d. h., daß

$$|f(x_1) - f(x_2)| \leq q |x_1 - x_2| \quad (x_1, x_2 \in I) \quad (1)$$

ist. Da die Gerade $y = x$ den Anstieg 1 besitzt, hat der Graph von f mit dieser Geraden höchstens einen, nach dem eben Bewiesenen also genau einen Punkt gemein. Die Bedingung (1) ist so stark, daß nicht einmal die Beschränktheit des Intervalls I gefordert zu werden braucht.

Eine Funktion f mit der Eigenschaft (1) heit eine *kontrahierende Abbildung*, weil der Abstand zweier Funktionswerte stets kleiner als der Abstand ihrer (voneinander verschiedenen) Argumente ist. Der nachfolgende *Fixpunktsatz* von BANACH wird aus diesem Grunde auch das *Prinzip der kontrahierenden Abbildungen* genannt.

Satz. *Es sei f eine auf einer abgeschlossenen Menge $F \subseteq \mathbb{R}$ definierte stetige Funktion mit folgenden Eigenschaften:*

- a) *Die Funktion f bildet F in sich ab, d. h., es ist $W(f) \subseteq F$.*
- b) *Die Funktion f ist kontrahierend, d. h., es gibt ein q mit $0 < q < 1$ derart, da*
stets

$$|f(x_1) - f(x_2)| \leq q |x_1 - x_2| \quad (x_1, x_2 \in F) \quad (2)$$

gilt.

Dann besitzt f genau einen Fixpunkt a in F . Ist a_0 ein ganz beliebig gewählter Punkt aus F und setzen wir

$$a_{n+1} := f(a_n) \quad (n \in \mathbb{N}), \quad (3)$$

so ist

$$a = \lim_{n \rightarrow \infty} a_n. \quad (4)$$

Ferner ist

$$|a - a_n| \leq q \frac{|a_n - a_{n-1}|}{1 - q}, \quad (5)$$

$$|a - a_n| \leq q^n \frac{|a_0 - a_1|}{1 - q}. \quad (6)$$

Beweis. Aus (2), (3) folgt

$$|a_{n+1} - a_n| = |f(a_n) - f(a_{n-1})| \leq q |a_n - a_{n-1}|,$$

und vollständige Induktion ergibt

$$|a_{n+k+1} - a_{n+k}| \leq q |a_{n+k} - a_{n+k-1}| \leq \dots \leq q^k |a_{n+1} - a_n|. \quad (7)$$

Speziell ist

$$|a_{k+1} - a_k| \leq q^k |a_0 - a_1|. \quad (8)$$

Aus der Dreiecksungleichung und aus (7), (8) folgt

$$|a_{n+p+1} - a_n| \leq \sum_{k=0}^p |a_{n+k+1} - a_{n+k}| \leq |a_{n+1} - a_n| \sum_{k=0}^p q^k, \quad (9)$$

$$|a_{n+p+1} - a_n| \leq |a_0 - a_1| q^n \sum_{k=0}^p q^k \leq |a_0 - a_1| \frac{q^n}{1 - q}. \quad (10)$$

Die Folge (a_n) ist daher eine Fundamentalfolge, die einen Grenzwert a besitzt. Da F abgeschlossen ist, gilt $a \in F$. Aus der Stetigkeit von f folgt nun

$$a = \lim_{n \rightarrow \infty} a_{n+1} = \lim_{n \rightarrow \infty} f(a_n) = f\left(\lim_{n \rightarrow \infty} a_n\right) = f(a).$$

Ist auch $b = f(b)$, so ist

$$|a - b| \leq |f(a) - f(b)| \leq q |a - b|,$$

was nur für $|a - b| = 0$, $a = b$ möglich ist. Die Funktion f besitzt somit genau einen Fixpunkt.

Wir beweisen die Behauptungen (5), (6), die eine Fehlerabschätzung ermöglichen. Aus

$$|a - a_n| \leq |a - a_{n+p+1}| + |a_{n+p+1} - a_n| \leq |a - a_{n+p+1}| + |a_{n+1} - a_n| \sum_{k=0}^p q^k$$

ergibt sich durch Grenzübergang $p \rightarrow \infty$ die Behauptung

$$\begin{aligned} |a - a_n| &\leq |a_{n+1} - a_n| \cdot \frac{1}{1 - q} \\ &\leq |a_n - a_{n-1}| \frac{q}{1 - q} \leq \dots \leq |a_1 - a_0| \frac{q^n}{1 - q}, \end{aligned}$$

und der Fixpunktsatz ist bewiesen.

Die Ungleichung (5) gestattet es, nach jedem „Iterationsschritt“ (d. h. nach jeder Berechnung von a_{n+1} aus a_n nach (3)) den jeweiligen Fehler abzuschätzen (a-posteriori-Abschätzung). Die Ungleichung (6) ermöglicht es dagegen, vor Beginn (falls man Rundungsfehler nicht in Betracht zieht) eine Maximalzahl von Iterationsschritten zur Erreichung einer vorgeschriebenen Genauigkeit festzulegen (a-priori-Abschätzung). Das Verfahren ist *selbstkorrigierend*, d. h., man kann nach jedem Teilschritt ohne Bedenken von einem gerundeten Näherungswert ausgehen, da sich die Abschätzung (5) nur auf den unmittelbar vorangehenden Näherungswert stützt.

Beispiel. Wir suchen eine Näherungslösung der kubischen Gleichung

$$3x^3 - 100x + 1 = 0.$$

Setzen wir

$$f(x) := \frac{3x^3 + 1}{100}. \quad (11)$$

so ist a genau dann eine Lösung der kubischen Gleichung, wenn $f(a) = a$, d. h., wenn a ein Fixpunkt von f ist. Für $|x| \leq 1$ ist

$$|f(x)| \leq \frac{3 + 1}{100} \leq 1,$$

d. h., die Funktion f bildet das abgeschlossene Intervall $[-1, 1]$ in sich ab. Wegen

$$\begin{aligned} |f(x_1) - f(x_2)| &= \left| \frac{3x_1^3 - 3x_2^3}{100} \right| \\ &= \frac{3}{100} |x_1^3 + x_1x_2 + x_2^3| \cdot |x_1 - x_2| \leq \frac{3}{100} \cdot 3 |x_1 - x_2|, \\ |f(x_1) - f(x_2)| &\leq \frac{1}{10} |x_1 - x_2| \quad (|x_1|, |x_2| \leq 1) \end{aligned}$$

ist die Abbildung kontrahierend mit dem *Kontraktionsfaktor* $q = 0,1$. Die Voraussetzungen des Banachschen Fixpunktsatzes sind erfüllt. Mit dem *Startwert* $a_0 = 0$ erhalten wir

$$\begin{aligned} a_1 &= f(0) = 0,01, \\ a_2 &= f(0,01) = 0,01000003. \end{aligned}$$

Die Fehlerabschätzung (6) ergibt

$$|a - a_2| \leq \frac{0,01 \cdot 0,01}{1 - 0,1} = 0,000\bar{1}.$$

Bereits nach dem zweiten Iterationsschritt haben wir damit einen brauchbaren Näherungswert gewonnen.

Der Ansatz (11) muß an dieser Stelle als Kunstgriff erscheinen. Mit den Methoden der Differentialrechnung werden wir später allgemeine Kriterien für die Anwendbarkeit des Banachschen Fixpunktsatzes herleiten.

Der Banachsche Fixpunktsatz bleibt unverändert gültig, wenn wir \mathbf{R} durch einen beliebigen vollständigen metrischen Raum ersetzen. Wie gewohnt hat man überall $|x - y|$ durch $\varrho(x, y)$ zu ersetzen.

2.5. Die hyperbolischen Funktionen und die Area-Funktionen

2.5.1. Die hyperbolischen Funktionen

Mit Hilfe der (speziellen) Exponentialfunktion definieren wir ein Funktionenpaar, für das zahlreiche Formeln gelten, die in analoger Weise bei der Kosinus- und Sinusfunktion aufgetreten sind. Wir setzen

$$\cosh x := \frac{e^x + e^{-x}}{2} \quad (x \in \mathbf{R}) \quad (1)$$

(gelesen: cosinus hyperbolicus x) und

$$\sinh x := \frac{e^x - e^{-x}}{2} \quad (x \in \mathbb{R}) \quad (2)$$

(gelesen: sinus hyperbolicus x).

Aus der Reihendarstellung für die Exponentialfunktion ergibt sich

$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, \quad (3)$$

$$\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}. \quad (4)$$

Somit ist

$$\cosh 0 = 1, \quad \sinh 0 = 0, \quad (5)$$

$$\cosh(-x) = \cosh x, \quad \sinh(-x) = -\sinh x. \quad (6)$$

Die Funktion $y = \cosh x$ ist eine gerade und $y = \sinh x$ eine ungerade Funktion. Die Additionstheoreme

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y, \quad (7)$$

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y \quad (8)$$

ergeben sich durch Einsetzen von (1), (2) auf beiden Seiten dieser Gleichungen. In gleicher Weise erhalten wir die Identität

$$\cosh^2 x - \sinh^2 x = 1. \quad (9)$$

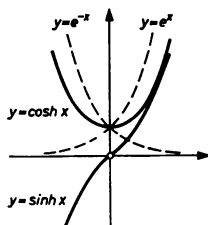


Abb. 2.20

Die Herleitung der zu 1.6.1.(6), (7) und (13) bis (16) analogen Formeln überlassen wir dem Leser.

Die Analogie zwischen den trigonometrischen und hyperbolischen Funktionen erstreckt sich nicht auf die Graphen dieser Funktionen. So sind die letzteren z. B. nicht beschränkt und (im Reellen) nicht periodisch (Abb. 2.20). Der Graph der Funktion

$y = \cosh x$ wird als *Kettenlinie* bezeichnet, weil eine an zwei Punkten befestigte Kette unter dem Einfluß der Schwerkraft die Gestalt dieser Kurve annimmt.

In GALILEO GALILEIS (1564–1642) berühmten „Discorsi ...“ von 1638 findet sich noch die Behauptung, die „Kettenlinie“ sei eine Parabel. JOACHIM JUNGIIUS (1587–1657) widerlegte durch Berechnungen und Experimente die Behauptung GALILEIS, konnte aber auch nicht angeben, welcher Art die Kurve ist. 1690 stellt JAKOB BERNOULLI in der Zeitschrift „Acta eruditorum“ öffentlich die Frage nach der Funktionsgleichung der „Kettenlinie“. LEIBNIZ, JOHANN BERNOULLI und HUYGENS lösen dann 1691 fast gleichzeitig diese Aufgabe.

Die Funktion $y = \cosh x$ ist im Intervall $[0, \infty[$ streng monoton wachsend, im Intervall $] -\infty, 0]$ streng monoton fallend, denn es gilt

$$\cosh x_1 - \cosh x_2 = \frac{e^{x_1} + e^{-x_1}}{2} - \frac{e^{x_2} + e^{-x_2}}{2} = \frac{1}{2} (e^{x_1} - e^{x_2}) (1 - e^{-(x_1+x_2)}).$$

Für $x_1 < x_2 \leq 0$ sind beide Klammerausdrücke negativ. Die Funktion $y = \cosh x$ ist somit im Intervall $] -\infty, 0]$ streng monoton fallend. Als gerade Funktion ist sie dann im Intervall $[0, \infty[$ streng monoton wachsend.

Die Funktion $y = \sinh x$ ist streng monoton wachsend, denn für $x_1 < x_2$ folgt wegen $e^{x_1} - e^{x_2} < 0$ stets

$$\sinh x_1 - \sinh x_2 = \frac{e^{x_1} - e^{-x_1}}{2} - \frac{e^{x_2} - e^{-x_2}}{2} = \frac{1}{2} (e^{x_1} - e^{x_2}) (1 + e^{-(x_1+x_2)}) < 0.$$

In Analogie zu den trigonometrischen Funktionen definieren wir

$$\tanh x := \frac{\sinh x}{\cosh x} \quad (x \in \mathbf{R}), \quad (10)$$

$$\coth x := \frac{\cosh x}{\sinh x} \quad (x \in \mathbf{R}, x \neq 0). \quad (11)$$

Für $x_1 < x_2$ ist

$$\begin{aligned} \tanh x_1 - \tanh x_2 &= \frac{\sinh x_1}{\cosh x_1} - \frac{\sinh x_2}{\cosh x_2} \\ &= \frac{\sinh x_1 \cosh x_2 - \cosh x_1 \sinh x_2}{\cosh x_1 \cosh x_2} \\ &= \frac{\sinh(x_1 - x_2)}{\cosh x_1 \cosh x_2} < 0. \end{aligned}$$

Die Funktion $y = \tanh x$ ist streng monoton wachsend. Für $x_1 < x_2 < 0$ und $0 < x_1 < x_2$ gilt

$$\coth x_1 - \coth x_2 = \frac{1}{\tanh x_1} - \frac{1}{\tanh x_2} = \frac{\tanh x_2 - \tanh x_1}{\tanh x_1 \tanh x_2} > 0.$$

Die Funktion $y = \coth x$ ist im Intervall $] -\infty, 0[$ und im Intervall $] 0, \infty[$ streng monoton fallend (Abb. 2.21).

Eine Begründung für die Bezeichnung „hyperbolische“ Funktionen werden wir später geben. Sie hängt damit zusammen, daß diese Funktionen in ähnlicher Beziehung zur Hyperbel wie die trigonometrischen Funktionen zum Kreis stehen.

Die hyperbolischen Funktionen traten in ihren Grundlagen implizit u. a. bereits bei NEWTON und DE MOIVRE auf. Die Theorie dieser Funktionen begründet zu haben, ist ein Verdienst von VINCENZO RICCATI (1707–1775). Er verwendete geometrische Betrachtungen. Im Jahre 1768 kam LAMBERT dann auf die Idee, sie für die Trigonometrie nutzbar zu machen. Seitdem sind sie Gemeingut der Mathematik geworden.

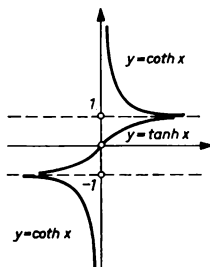


Abb. 2.21

2.5.2. Die Area-Funktionen

Nach den in 2.5.1. durchgeführten Monotonie-Untersuchungen besitzen die Funktionen

$$y = \cosh x \quad (x \in \mathbb{R}, x \geq 0),$$

$$y = \sinh x \quad (x \in \mathbb{R}),$$

$$y = \tanh x \quad (x \in \mathbb{R}),$$

$$y = \coth x \quad (x \in \mathbb{R}, x \neq 0)$$

die folgenden Umkehrfunktionen (Abb. 2.22, 2.23):

$$y = \operatorname{arcosh} x \quad (x \geq 1), \quad (1)$$

$$y = \operatorname{arsinh} x \quad (x \in \mathbb{R}), \quad (2)$$

$$y = \operatorname{artanh} x \quad (-1 < x < 1), \quad (3)$$

$$y = \operatorname{arcoth} x \quad (|x| > 1) \quad (4)$$

(gelesen: area cosinus hyperbolicus x usw.).

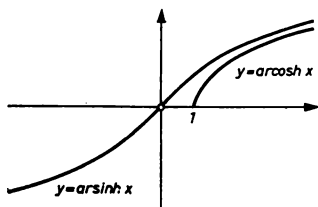


Abb. 2.22

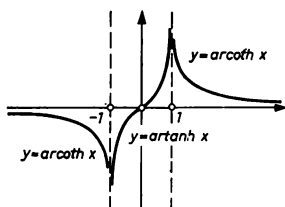


Abb. 2.23

Für die Funktionen (1) bis (4) können explizite Darstellungen gefunden werden. Es gilt $y = \operatorname{arcosh} x$ genau dann, wenn $y \geq 0$ und $x = \cosh y$, d. h.

$$\begin{aligned} 2x &= e^y + e^{-y}, \\ e^{2y} - 2xe^y + 1 &= 0 \end{aligned}$$

ist. Dies kann als quadratische Gleichung für e^y aufgefaßt werden. Das Produkt der Lösungen ist nach dem Vietschen Wurzelsatz gleich 1, und da $y \geq 0$, $e^y \geq 1$ gilt, muß die größere der beiden Lösungen, d. h. die Lösung

$$e^y = x + \sqrt{x^2 - 1},$$

gewählt werden. Somit ist

$$\operatorname{arcosh} x = \ln(x + \sqrt{x^2 - 1}). \quad (5)$$

Der Definitionsbereich dieser Funktion ist das Intervall $[1, \infty[$. Ähnliche Überlegungen führen zur Darstellung

$$\operatorname{arsinh} x = \ln(x + \sqrt{x^2 + 1}) \quad (x \in \mathbb{R}). \quad (6)$$

Es gilt $y = \operatorname{artanh} x$ genau dann, wenn $x = \tanh y$, d. h.

$$\begin{aligned} x &= \frac{e^y - e^{-y}}{e^y + e^{-y}} \quad (|x| < 1), \\ e^{2y} &= \frac{1+x}{1-x}, \quad e^y = \sqrt{\frac{1+x}{1-x}}, \quad y = \ln \sqrt{\frac{1+x}{1-x}} \end{aligned}$$

ist. Es folgt

$$\operatorname{artanh} x = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (|x| < 1). \quad (7)$$

Die Gleichung $y = \operatorname{arcoth} x$ ist mit

$$x = \frac{e^y + e^{-y}}{e^y - e^{-y}} \quad (|x| > 1),$$

$$e^{2y} = \frac{x+1}{x-1}, \quad e^y = \sqrt{\frac{x+1}{x-1}}, \quad y = \ln \sqrt{\frac{x+1}{x-1}}$$

äquivalent, d. h., es ist

$$\operatorname{arcoth} x = \frac{1}{2} \ln \frac{x+1}{x-1} \quad (|x| > 1). \quad (8)$$

Auf die Wahl der Bezeichnung „Area“ kommen wir später zurück.

2.6. Folgen und Reihen von Funktionen

2.6.1. Punktweise Konvergenz

Bei unseren bisherigen Untersuchungen sind schon mehrfach Folgen bzw. Reihen aufgetreten, deren Glieder nicht konstant, sondern von einer Variablen x abhängig waren. In dieser Weise haben wir z. B. durch 2.1.5.(8), (9), 2.2.8.(7) die spezielle Exponentialfunktion bzw. den natürlichen Logarithmus sowie durch 2.2.8.(12), (13) die trigonometrischen Funktionen eingeführt. In 2.6. wollen wir einige allgemeine

Untersuchungen über Folgen $(f_n)_{n \in \mathbb{N}}$ bzw. Reihen $\sum_{n=0}^{\infty} f_n$, deren Glieder f_n Funktionen sind, durchführen. Da jede Reihe durch die Folge ihrer Partialsummen bestimmt ist und umgekehrt zu jeder Folge eine Reihe existiert, deren Partialsummen die Glieder der gegebenen Folge sind, genügt es, einen dieser Fälle zu betrachten. Wir werden unsere Definitionen und Sätze je nach ihrer Bedeutung für die Anwendungen für Folgen oder für Reihen formulieren. Die sinngemäße Übertragung auf den anderen Fall bereitet keine grundsätzlichen Schwierigkeiten.

Wie bei Folgen und Reihen mit konstanten Gliedern ist auch hier die Frage der Konvergenz das Kernproblem.

Definition. Eine Folge (f_n) reeller oder komplexer Funktionen heißt (*punktweise*) *konvergent* in der Menge D ($D \subseteq D(f_n)$ für alle n), wenn es eine in D definierte Funktion f , die *Grenzfunktion* der Folge, gibt derart, daß

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad (1)$$

für alle x aus D ist.

Da es höchstens eine Funktion mit der Eigenschaft (1) gibt, ist die Grenzfunktion einer konvergenten Folge eindeutig bestimmt.

Sind die Funktionen f_n bzw. g_n beispielsweise durch

$$f_n(x) := \left(1 + \frac{x}{n}\right)^n \quad (x \in \mathbb{R}), \quad (2)$$

$$g_n(x) := n(\sqrt[n]{x} - 1) \quad (x \in \mathbb{R}, x > 0) \quad (3)$$

definiert, so ist die Folge (f_n) bzw. (g_n) in der Menge aller reellen Zahlen bzw. in der Menge aller positiven reellen Zahlen konvergent und besitzt als Grenzfunktion die spezielle Exponentialfunktion bzw. den natürlichen Logarithmus.

Gehen wir auf die Definition des Grenzwertes einer Folge zurück, so können wir die punktweise Konvergenz formalisiert durch

$$\bigwedge_{x \in D} \bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbb{N}} \bigwedge_{n \in \mathbb{N}} (n \geq N \Rightarrow |f(x) - f_n(x)| < \varepsilon) \quad (4)$$

definieren, wobei die ersten beiden Quantifikatoren auch vertauscht werden können.

Die Grenzfunktion einer (punktweise) konvergenten Folge stetiger Funktionen f_n braucht nicht wieder stetig zu sein.

Beispiel. Für jede positive natürliche Zahl n sei $D(f_n)$ die Menge aller nichtnegativen reellen Zahlen, und der Graph von f_n sei für $0 \leq x \leq \frac{1}{n}$ die Verbindungsstrecke der Punkte $(0, 1)$, $(\frac{1}{n}, 0)$ und falle für $x > \frac{1}{n}$ mit der x -Achse zusammen (Abb. 2.24), d. h., es sei

$$f_n(x) = \begin{cases} -nx + 1 & \text{für } 0 \leq x \leq \frac{1}{n}, \\ 0 & \text{für } x > \frac{1}{n}. \end{cases}$$

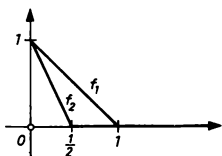


Abb. 2.24

Die Funktionen f_n sind stetig, und die Folge (f_n) ist in der Menge der nichtnegativen reellen Zahlen (punktweise) konvergent. Die Grenzfunktion f ist offensichtlich durch

$$f(x) = \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{für } x > 0 \end{cases}$$

gegeben. Sie ist an der Stelle 0 unstetig.

2.6.2. Gleichmäßige Konvergenz

Die Frage nach hinreichenden Bedingungen für die Stetigkeit der Grenzfunktion einer konvergenten Folge stetiger Funktionen führte zu einer Verschärfung des Konvergenzbegriffs.

Definition. Eine Folge (f_n) reeller oder komplexer Funktionen heißt *gleichmäßig konvergent* in der Menge D ($D \subseteq D(f_n)$ für alle n), wenn es eine in D definierte Funktion f mit folgender Eigenschaft gibt: Für alle $\varepsilon > 0$ gibt es ein N derart, daß $|f(x) - f_n(x)| < \varepsilon$ für alle x mit $x \in D$ und für alle n mit $n \geq N$ ist.

In formalisierter Schreibweise lautet die definierende Aussage

$$\bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbb{N}} \bigwedge_{x \in D} \bigwedge_{n \in \mathbb{N}} (n \geq N \Rightarrow |f(x) - f_n(x)| < \varepsilon), \quad (1)$$

und der Vergleich mit 2.6.1.(4) zeigt, daß nur eine unterschiedliche Reihenfolge der Quantifikatoren vorliegt. Der wesentliche Unterschied besteht darin, daß die gemäß (1) existierende Zahl N bei gleichmäßiger Konvergenz nur von ε , nicht aber von der Stelle x mit $x \in D$ abhängig ist. Wir haben hier einen ähnlichen Sachverhalt wie bei der gleichmäßigen Stetigkeit (vgl. 2.4.3.). Während 2.6.1.(4) aus (1) gefolgert werden kann, ist das Umgekehrte nicht der Fall. Jede gleichmäßig konvergente Folge ist punktweise konvergent, aber das folgende Beispiel zeigt, daß nicht jede punktweise konvergente Folge gleichmäßig konvergent ist.

Beispiel 1. Es sei f_n wie im Beispiel in 2.6.1. definiert. Wir setzen $\varepsilon := \frac{1}{2}$ und bestimmen zu jeder positiven reellen Zahl N eine natürliche Zahl n mit $n \geq N$ und setzen $x := \frac{1}{2n}$. Dann ist

$$|f(x) - f_n(x)| = \left| 0 - \frac{1}{2} \right| = \frac{1}{2} \geq \varepsilon.$$

Die Aussage

$$\bigvee_{\varepsilon > 0} \bigwedge_{N \in \mathbb{N}} \bigvee_{x \in D} \bigvee_{n \in \mathbb{N}} (n \geq N \wedge |f(x) - f_n(x)| \geq \varepsilon),$$

d. h. die Negation von (1), ist wahr. Die Folge ist konvergent, aber nicht gleichmäßig konvergent.

Beispiel 2. Es seien $f_n(x)$ die Partialsummen der Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} \cos kx \quad (x \in \mathbb{R}). \quad (2)$$

Wegen

$$|f_n(x)| \leq \sum_{k=1}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n+1}$$

ist die Reihe mit den Partialsummen $f_n(x)$ nach dem Vergleichskriterium für alle x konvergent, und die Folge $f_n(x)$ konvergiert punktweise gegen die Funktion

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} \cos kx.$$

Für alle x ist

$$|f(x) - f_n(x)| \leq \sum_{k=n+1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{n+1},$$

und die Bedingung (1) ist mit $N := \frac{1}{\varepsilon}$ erfüllt. Die Folge (f_n) ist gleichmäßig konvergent in \mathbb{R} .

Aus der gleichmäßigen Konvergenz können weitreichende Folgerungen abgeleitet werden. Ein erstes Beispiel hierfür gibt der

Satz 1. *Jede in einer Menge D gleichmäßig konvergente Folge stetiger Funktionen besitzt eine in der Menge D stetige Grenzfunktion.*

Beweis. Es sei f die in D definierte Grenzfunktion der Folge (f_n) und $x_0 \in D$. Zu vorgegebenem $\varepsilon > 0$ wählen wir eine reelle Zahl N derart, daß $|f(x) - f_n(x)| < \frac{\varepsilon}{3}$ für alle x mit $x \in D$ und für alle n mit $n \geq N$ ist. Zu einer fest gewählten natürlichen Zahl n mit $n \geq N$ bestimmen wir unter Berücksichtigung der Stetigkeit der Funktion f_n eine positive reelle Zahl δ derart, daß $|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}$ für alle x mit $x \in D$ und $|x - x_0| < \delta$ ist. Für diese reellen Zahlen x gilt somit

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Die Existenz der Zahl δ bei vorgegebenen x_0, ε mit dieser Eigenschaft besagt aber gerade, daß f in allen Punkten von D stetig ist, und damit ist unser Satz bewiesen.

Wir wollen die Annäherung der Folgenglieder einer gleichmäßig konvergenten Folge stetiger reeller Funktionen f_n an die Grenzfunktion f geometrisch interpretieren. Zu vorgegebenem $\varepsilon > 0$ betrachten wir den sogenannten „ ε -Schlauch“ um den

Graphen von f (Abb. 2.25). Er besteht aus allen Punkten (x, y) mit $x \in D$ und $|y - f(x)| \leq \varepsilon$. Wählen wir $N = N(\varepsilon)$ entsprechend der Definition der gleichmäßigen Stetigkeit, so gilt $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in D$ und für alle n mit $n \geq N$. Die Graphen aller Funktionen f_n mit $n \geq N$ liegen also innerhalb des ε -Schlauches. Lassen wir ε gegen 0 streben, so nähern sich die Graphen der Funktionen f_n mit $n \geq N(\varepsilon)$ „gleichmäßig“, d. h. in ihrem ganzen Funktionsverlauf der Grenzfunktion f .

Besitzt eine punktweise konvergente Folge von stetigen Funktionen f_n eine stetige Grenzfunktion, so kann hieraus nicht auf gleichmäßige Konvergenz geschlossen werden.

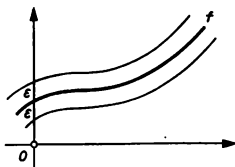


Abb. 2.25

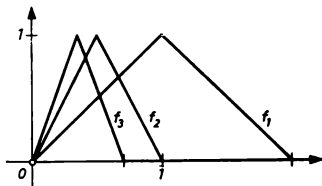


Abb. 2.26

Beispiel 3. Die Funktionen f_n seien für $n \geq 1$ wie folgt definiert:

$$f_n(x) = \begin{cases} nx & \text{für } 0 \leq x \leq \frac{1}{n}, \\ -nx + 2 & \text{für } \frac{1}{n} < x \leq \frac{2}{n}, \\ 0 & \text{für alle anderen } x \end{cases}$$

(Abb. 2.26). Da der „Zacken“ dieser Funktionen für wachsende n immer dichter an die y -Achse heranrückt, gilt

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = 0 \quad (x \in \mathbb{R}).$$

Ist $0 < \varepsilon < 1$, so liegt keine der Funktionen f_n im ε -Schlauch um den Graphen von f . Die Folge ist nicht gleichmäßig konvergent, obwohl die Grenzfunktion stetig ist.

Für die Untersuchung von Reihen auf gleichmäßige Konvergenz ist das Kriterium von Weierstraß von besonderer Bedeutung.

Satz 2. Gibt es zu einer Reihe, deren Glieder reelle oder komplexe Funktionen f_n mit dem Definitionsbereich D sind, eine konvergente Reihe $\sum a_n$ mit $|f_n(x)| \leq a_n$ für alle $x \in D$ und für alle $n > n_0$, so ist die Reihe $\sum f_n$ gleichmäßig konvergent.

Beweis. Es seien s_n die Partialsummen der Reihe $\sum f_n$. Wegen

$$|s_{n+k}(x) - s_n(x)| \leq a_{n+1} + \dots + a_{n+k}$$

und der Konvergenz der Reihe $\sum a_n$ ist die Folge $(s_n(x))$ für jedes $x \in D$ eine Fundamentalfolge, und es existiert

$$s(x) := \lim_{n \rightarrow \infty} s_n(x) \quad (x \in D).$$

Zu vorgegebenem $\varepsilon > 0$ wählen wir ein N mit $a_{n+1} + a_{n+2} + \dots < \varepsilon$ für $n \geq N$. Dann ist

$$|s(x) - s_n(x)| \leq a_{n+1} + a_{n+2} + \dots < \varepsilon$$

für $n \geq N$ und $x \in D$. Daher ist die Folge (s_n) oder, was definitionsgemäß gleichbedeutend ist, die Reihe $\sum f_n$ gleichmäßig konvergent.

Beispiel 4. Die Reihe

$$\sum_{n=0}^{\infty} \frac{\sin(10^{2n}x)}{10^n} \quad (x \in \mathbb{R})$$

ist wegen

$$\left| \frac{\sin(10^{2n}x)}{10^n} \right| \leq \frac{1}{10^n}$$

und Satz 2 in \mathbb{R} gleichmäßig konvergent, und nach Satz 1 ist die Funktion f mit

$$f(x) = \sum_{n=0}^{\infty} \frac{\sin(10^{2n}x)}{10^n} \quad (x \in \mathbb{R}) \quad (3)$$

überall stetig.

2.6.3. Čebyšev-Metrik

Die mit der gleichmäßigen Konvergenz zusammenhängenden Begriffe und Sätze können sehr übersichtlich formuliert werden, wenn man den Begriff des metrischen Raumes heranzieht. Wir beschränken uns auf den Fall, daß der Definitionsbereich D der betrachteten Funktionen eine abgeschlossene beschränkte Menge reeller oder komplexer Zahlen ist. Mit C_D bezeichnen wir die Menge aller auf D stetigen Funktionen. Wir setzen

$$\varrho(f, g) := \max_{x \in D} |f(x) - g(x)| \quad (f, g \in C_D), \quad (1)$$

wobei wir auf Grund von 2.4.2., Satz 1, \max an Stelle von \sup schreiben durften. Offensichtlich ist $\varrho(f, g) \geq 0$, und $\varrho(f, g) = 0$ gilt genau dann, wenn $f(x) = g(x)$ für alle $x \in D$, d. h., wenn $f = g$ ist. Ferner ist $\varrho(f, g) = \varrho(g, f)$. Sind f, g, h aus C_D , so gilt stets

$$|f(x) - h(x)| \leq |f(x) - g(x)| + |g(x) - h(x)| \leq \varrho(f, g) + \varrho(g, h) \quad (x \in D),$$

und folglich ist auch

$$\varrho(f, h) \leq \varrho(f, g) + \varrho(g, h).$$

Die Menge C_D bildet also in Verbindung mit der durch (1) definierten Funktion ϱ einen metrischen Raum (vgl. 1.5.6.). Die Funktion ϱ wird auch die Čebyšev-Metrik (ПАФНУТИ ЛВОНИЉ ЧЕБЫШЕВ (1821–1894)) im Raum C_D genannt.

Satz 1. Eine Folge von Funktionen f_n aus C_D ist in D gleichmäßig konvergent genau dann, wenn es eine Funktion f aus C_D mit

$$\lim_{n \rightarrow \infty} \varrho(f, f_n) = 0 \quad (2)$$

gibt.

Beweis. Es gelte (2). Dann gibt es zu vorgegebenem $\varepsilon > 0$ ein N mit $\varrho(f, f_n) < \varepsilon$ für $n \geq N$. Für alle $x \in D$ und für $n \geq N$ ist dann

$$|f(x) - f_n(x)| \leq \varrho(f, f_n) < \varepsilon,$$

und die Folge ist in D gleichmäßig konvergent. Ist umgekehrt die Folge (f_n) in D gleichmäßig konvergent mit der Grenzfunktion f , so gibt es zu vorgegebenem $\varepsilon > 0$ ein N mit

$$|f(x) - f_n(x)| \leq \frac{\varepsilon}{2}$$

für alle $x \in D$ und für $n \geq N$. Es folgt

$$\varrho(f, f_n) \leq \frac{\varepsilon}{2} < \varepsilon$$

für $n \geq N$, d. h., es gilt (2). Damit ist Satz 1 bewiesen.

Eine Folge (f_n) ist demnach in der abgeschlossenen beschränkten Punktmenge D gleichmäßig konvergent genau dann, wenn sie in der Čebyšev-Metrik konvergent ist.

Satz 2. Der Raum C_D ist bezüglich der Čebyšev-Metrik vollständig (vgl. 2.1.7., Definition 2).

Beweis. Es sei (f_n) eine Fundamentalfolge in C_D , d. h., für alle $\varepsilon > 0$ existiere ein N mit $\varrho(f_m, f_n) < \varepsilon$ für $m, n \geq N$. Für alle $x \in D$ und $m, n \geq N$ ist dann $|f_m(x) - f_n(x)| \leq \varrho(f_m, f_n) < \varepsilon$, und folglich ist $(f_n(x))$ eine reelle (bzw. komplexe) Fundamentalfolge. Für jedes $x \in D$ existiert daher der Grenzwert

$$f(x) := \lim_{n \rightarrow \infty} f_n(x).$$

Zu vorgegebenem $\varepsilon > 0$ wählen wir ein N mit

$$|f_m(x) - f_n(x)| \leq \varrho(f_m, f_n) < \frac{\varepsilon}{2} \quad (x \in D; m, n \geq N).$$

Der Grenzübergang $m \rightarrow \infty$ ergibt

$$|f(x) - f_n(x)| \leq \frac{\varepsilon}{2} < \varepsilon \quad (x \in D; n \geq N),$$

und dies besagt, daß die Folge (f_n) in D gleichmäßig gegen die Funktion f konvergiert. Die Grenzfunktion f ist nach 2.6.2., Satz 1, stetig, und damit ist Satz 2 bewiesen.

2.7. Approximation von Funktionen

2.7.1. Das Approximationsproblem

In den Anwendungen der Mathematik ist man wiederholt darauf angewiesen, eine gegebene Funktion f durch eine Funktion bzw. durch eine Folge von Funktionen f_n einer gewissen Klasse von Funktionen möglichst gut anzunähern, zu approxi-

mieren. In den wichtigsten Fällen sind die Näherungsfunktionen f_n ganzrational bzw. Linearkombinationen der trigonometrischen Funktionen $\cos kx$, $\sin kx$ ($k \in \mathbf{N}$). Für die „Güte“ einer Approximation können verschiedene Kriterien entwickelt werden, von denen wir einige aufzählen.

a) Die Näherungsfunktionen f_n sollen mit der gegebenen Funktion f an möglichst vielen gegebenen Stellen, den *Stützstellen* übereinstimmen. Eine solche Approximation bezeichnet man als *Interpolation* durch Funktionen der gegebenen Klasse.

b) Die Näherungsfunktionen f_n sollen die Funktion f in der Umgebung eines festen Punktes a möglichst gut annähern. Dieses Problem werden wir für die Klasse der ganzrationalen Funktionen mit Hilfe der *Taylorischen Formel* (BROOK TAYLOR (1685–1731)) im Kapitel Differentialrechnung lösen.

c) Die Näherungsfunktionen f_n sollen die Funktion f gleichmäßig approximieren. (Hiermit ist gleichbedeutend, daß die Abweichung $\varrho(f, f_n)$ der Näherungsfunktionen f_n von der Funktion f in der Čebyšev-Metrik sehr klein sein soll.)

Dieses Problem griff zuerst ČEBYŠEV auf. Seine Approximationstheorie ist erwachsen aus der Beschäftigung mit dem sogenannten Wattschen Parallelogramm. An der Dampfmaschine von JAMES WATT (1736–1819) befand sich ein Mechanismus, der die Aufgabe hatte, dem Endpunkt der Kolbenstange eine geradlinige Bewegung zu erteilen, damit der Kolben im Zylinder sich weder lockert noch klemmt. Der Wattsche Mechanismus ermöglichte nur eine in einem engen Bereich annähernd gerade Bewegung, so daß die technischen Anforderungen an eine Geradföhrung etwa erfüllt waren. Es entstand die mathematische Aufgabe, entweder eine exakte Geradföhrung zu finden (gelöst 1864) oder eine Methode zu geben, um die passendsten Elemente für die Genauigkeit des Spiels des Parallelogramms zu bestimmen. Das letzte erfordert, das Polynom zu finden, welches sich in einem gegebenen Intervall am besten einer gegebenen Funktion nähert. Dieser Frage ist ČEBYŠEVs Arbeit von 1853 „Théorie des mécanismes connus sous le nom de parallélogrammes“ gewidmet, die die erste systematische Darstellung zur Approximationstheorie ist. Diese Theorie ist also eindeutig durch ein praktisches Problem angeregt worden. ČEBYŠEV bemerkte selbst, daß ihm auch bei seinem Approximationsproblem (allerdings im Zusammenhang mit geographischen Projektionen) EULER (1777) vorausgegangen war.

2.7.2. Interpolation

Das erste der in 2.7.1. aufgeworfenen Probleme wird gelöst durch den

Satz 1. Zu jeder reellen (oder komplexen) Funktion f und zu jedem System von $n + 1$ verschiedenen Stützstellen x_k ($k = 0, 1, \dots, n$; $x_k \in D(f)$) gibt es genau eine ganzrationale Funktion P von höchstens n -tem Grad derart, daß

$$P(x_k) = f(x_k) \quad (k = 0, 1, \dots, n) \quad (1)$$

ist.

Beweis. Wenn eine Funktion P mit

$$P(x) = \sum_{j=0}^n c_j x^j \quad (2)$$

existiert, die den Bedingungen (1) genügt, sind die Gleichungen

$$\sum_{j=0}^n c_j x_k^j = f(x_k) \quad (k = 0, 1, \dots, n) \quad (3)$$

erfüllt. Als Gleichungssystem für die c_j hat (3) die Koeffizientenmatrix

$$\mathfrak{A} := \begin{pmatrix} 1 & x_0^1 & \dots & x_0^n \\ 1 & x_1^1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n^1 & \dots & x_n^n \end{pmatrix},$$

deren Determinante unter der Bezeichnung *Vandermondesche Determinante* (ALEXANDER THEOPHILE VANDERMONDE (1735–1796)) bekannt ist. In der Determinantentheorie wird bewiesen, daß

$$|\mathfrak{A}| = (x_1 - x_0)(x_2 - x_0) \cdots (x_n - x_0) \cdot \\ (x_2 - x_1) \cdots (x_n - x_1) \cdot \\ (x_n - x_{n-1})$$

ist. Da alle Stützstellen voneinander verschieden sind, ist die Koeffizientendeterminante ungleich 0, und es gibt genau ein System von Zahlen c_0, c_1, \dots, c_n , für das (3) erfüllt ist. Definieren wir also P durch (2), so ist (1) erfüllt, und unser Satz ist bewiesen.

Beispiel. Die Sinusfunktion soll durch eine quadratische Funktion so approximiert werden, daß beide Funktionen an den Stellen $0, \frac{\pi}{6}, \frac{\pi}{2}$ übereinstimmen.

Die Näherungsfunktion $P(x) = c_0 + c_1 x + c_2 x^2$ ist aus den Bedingungen

$$\begin{aligned} c_0 + 0 + 0 &= 0, \\ c_0 + c_1 \frac{\pi}{6} + c_2 \left(\frac{\pi}{6}\right)^2 &= \frac{1}{2}, \\ c_0 + c_1 \frac{\pi}{2} + c_2 \left(\frac{\pi}{2}\right)^2 &= 1 \end{aligned}$$

zu ermitteln. Dieses Gleichungssystem hat die Lösung $c_0 = 0, c_1 = \frac{7}{2\pi}, c_2 = -\frac{3}{\pi^2}$, d. h., es ist

$$P(x) = \frac{7}{2\pi} x - \frac{3}{\pi^2} x^2.$$

Für die Lösung des Interpolationsproblems gibt es auch explizite Formeln. Von LAGRANGE stammt die Formel

$$P(x) = \sum_{k=0}^n f(x_k) l_k(x), \quad (4)$$

wobei die Funktionen l_k durch

$$l_k(x) := \frac{(x - x_0)(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)}{(x_k - x_0)(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)}$$

definiert sind. Für $x = x_k$ stimmen Zähler und Nenner überein, d. h., es ist $l_k(x_k) = 1$. Für $x = x_j$ ($j = 0, 1, \dots, n; j \neq k$) verschwindet der Zähler, d. h., es ist $l_k(x_j) = 0$ für $j \neq k$. In der Summe

$$P(x_j) = \sum_{k=0}^n f(x_k) l_k(x_j)$$

verschwinden daher alle Summanden, in denen $k \neq j$ ist, und wegen $l_j(x_j) = 1$ gilt $P(x_j) = f(x_j)$. Da alle Funktionen l_k ganzrationale Funktionen vom Grad n sind, ist die Linearkombination (4) dieser Funktionen eine ganzrationale Funktion höchstens n -ten Grades, die das Interpolationsproblem löst. Bereits zuvor hatten wir bewiesen, daß diese Lösung eindeutig bestimmt ist.

2.7.3. Gleichmäßige Approximation

Über die gleichmäßige Approximation von stetigen Funktionen wurden von WEIERSTRASS zwei berühmte Sätze bewiesen.

Satz 1. *Zu jeder auf einem abgeschlossenen beschränkten Intervall stetigen Funktion f gibt es eine Folge von ganzrationalen Funktionen, die auf diesem Intervall gleichmäßig gegen f konvergiert.*

Beweis. Die Funktion f sei auf dem abgeschlossenen Intervall $[a, b]$ stetig. Wir führen den Beweis auf den Fall $a = 0, b = 1$ zurück, indem wir

$$g(x) := f(a + x(b - a)) \quad (1)$$

setzen. Die Funktion g ist im Intervall $[0, 1]$ stetig, und es ist

$$f(x) = g\left(\frac{x - a}{b - a}\right) \quad (a \leq x \leq b),$$

und wenn die Behauptung für g gilt, ist sie auch für f bewiesen.

Fortan sei also $a = 0, b = 1$. Wir führen die sogenannten *Bernsteinschen Polynome* (SERGEJ NATANOWIČ BERNŠTEJN (1880–1968))

$$B_n(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \quad (n \in \mathbb{N}) \quad (2)$$

der Funktion f ein und beweisen zunächst einige Hilfsformeln. Für $n \geq 2$ gilt, wenn wir zur Abkürzung $y := 1 - x$ setzen,

$$\sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = (x+y)^n = 1, \quad (3)$$

$$\begin{aligned} \sum_{k=0}^n k \binom{n}{k} x^k y^{n-k} &= \sum_{k=1}^n n \binom{n-1}{k-1} x^k y^{n-k} = nx \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-k-1} \\ &= nx(x+y)^{n-1} = nx, \end{aligned} \quad (4)$$

$$\begin{aligned} \sum_{k=0}^n k^2 \binom{n}{k} x^k y^{n-k} &= \sum_{k=1}^n n k \binom{n-1}{k-1} x^k y^{n-k} = nx \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} x^k y^{n-k-1} \\ &= nx \sum_{k=0}^{n-1} k \binom{n-1}{k} x^k y^{n-k-1} + nx \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-k-1}. \end{aligned}$$

Wenden wir (4) bzw. (3) mit $n-1$ statt n auf die vorletzte bzw. letzte Summe an, so erhalten wir

$$\sum_{k=0}^n k^2 \binom{n}{k} x^k y^{n-k} = n(n-1)x^2 + nx. \quad (5)$$

Wegen (3) ist

$$\begin{aligned} f(x) - B_n(x) &= \sum_{k=0}^n \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k y^{n-k}, \\ |f(x) - B_n(x)| &\leq \sum_{k=0}^n \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k y^{n-k}. \end{aligned} \quad (6)$$

Zu vorgegebenem $\varepsilon > 0$ wählen wir auf Grund der gleichmäßigen Stetigkeit von f ein $\delta > 0$ mit $|f(x_1) - f(x_2)| < \frac{\varepsilon}{2}$ für $|x_1 - x_2| < \delta$. Es sei $x \in [0, 1]$, und M_1 bzw. M_2 sei die Menge aller natürlichen Zahlen k mit $0 \leq k \leq n$ und

$$\left| x - \frac{k}{n} \right| < \delta \quad \text{bzw.} \quad \left| x - \frac{k}{n} \right| \geq \delta.$$

Für $k \in M_1$ ist $\left| f(x) - f\left(\frac{k}{n}\right) \right| < \frac{\varepsilon}{2}$, und folglich ist

$$\sum_{k \in M_1} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k y^{n-k} \leq \frac{\varepsilon}{2} \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = \frac{\varepsilon}{2}. \quad (7)$$

Für $k \in M_2$ ist $1 \leq \frac{1}{\delta^2} \left(x - \frac{k}{n} \right)^2$, und wenn K das Maximum von f für $0 \leq x \leq 1$ ist, erhalten wir unter Berücksichtigung von (3), (4), (5) die Abschätzung

$$\begin{aligned} \sum_{k \in M_2} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k y^{n-k} &\leq 2K \sum_{k \in M_2} \binom{n}{k} x^k y^{n-k} \\ &\leq 2K \sum_{k=0}^n \frac{1}{\delta^2} \left(x - \frac{k}{n} \right)^2 \binom{n}{k} x^k y^{n-k} \\ &= \frac{2K}{n^2 \delta^2} \sum_{k=0}^n (n^2 x^2 - 2nkx + k^2) \binom{n}{k} x^k y^{n-k} \\ &= \frac{2K}{n^2 \delta^2} [n^2 x^2 - 2nx \cdot nx + n(n-1)x^2 + n_2] \end{aligned}$$

$$\begin{aligned}
&= \frac{2K}{n\delta^2} x(1-x) = \frac{2K}{n\delta^2} \left(\frac{1}{4} - \left(x - \frac{1}{2}\right)^2 \right) \\
&\leq \frac{2K}{n\delta^2} \cdot \frac{1}{4} = \frac{K}{2n\delta^2}.
\end{aligned} \tag{8}$$

Wählen wir $N > \frac{K}{\varepsilon\delta^2}$, so gilt für $n \geq N$ und $0 \leq x \leq 1$ wegen (6), (7), (8) stets

$$|f(x) - B_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

und Satz 1 ist bewiesen.

Der zweite Satz von WEIERSTRASS ermöglicht die gleichmäßige Approximation von stetigen periodischen Funktionen durch trigonometrische Funktionensysteme. Unter einem *trigonometrischen Polynom* versteht man jede reelle Funktion T , die sich in der Form

$$T(x) = a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \quad (x \in \mathbb{R})$$

mit $a_k, b_k \in \mathbb{R}$ darstellen läßt. Jede Linearkombination von trigonometrischen Polynomen ist daher wieder ein trigonometrisches Polynom. Aus den Additionstheoremen folgt, daß mit T auch die durch $S(x) := T(x+c)$ definierte Funktion S ein trigonometrisches Polynom ist. Ersetzen wir α, β in 1.6.1.(13) bis (16) durch $(m+n)x$ bzw. $(m-n)x$, so erkennen wir, daß auch das Produkt zweier trigonometrischer Polynome wieder ein trigonometrisches Polynom ist.

Nach diesen Vorbereitungen formulieren wir

Satz 2. *Zu jeder reellen stetigen Funktion f mit der Periode 2π gibt es eine Folge von trigonometrischen Polynomen, die in \mathbb{R} gleichmäßig gegen die Funktion f konvergiert.*

Beweis. Es genügt, zu zeigen, daß zu jedem $\varepsilon > 0$ ein trigonometrisches Polynom T mit $|f(x) - T(x)| < \varepsilon$ für $x \in \mathbb{R}$ existiert. Zunächst sei f eine gerade Funktion. Die Funktion g mit $g(t) := f(\arccos t)$ ($-1 \leq t \leq 1$) ist stetig, und nach Satz 1 existiert zu vorgegebenem $\varepsilon > 0$ eine ganzrationale Funktion h mit $|g(t) - h(t)| \leq \varepsilon$ für $-1 \leq t \leq 1$. Nach Definition ist $f(x) = g(\cos x)$ zunächst für $0 \leq x \leq \pi$. Wegen $f(-x) = f(x)$, $\cos(-x) = \cos x$ und der Periodizität von $f(x)$ und $\cos x$ gilt $f(x) = g(\cos x)$ für alle $x \in \mathbb{R}$. Setzen wir $T(x) := h(\cos x)$, so ist T ein trigonometrisches Polynom, und es gilt

$$|f(x) - T(x)| = |g(\cos x) - h(\cos x)| < \varepsilon \quad (x \in \mathbb{R}).$$

Die Behauptung des Satzes ist damit für gerade Funktionen bewiesen.

Als nächstes betrachten wir Funktionen, die sich in der Form $f(x) = g(x) \sin^2 x$ mit einer stetigen Funktion g der Periode 2π darstellen lassen. Zu den (geraden) Funktionen

$$g_1(x) := \frac{g(x) + g(-x)}{2} \sin^2 x,$$

$$g_2(x) := \frac{g(x) - g(-x)}{2} \sin x$$

existieren nach dem Bewiesenen zu vorgegebenem $\varepsilon > 0$ trigonometrische Polynome T_1, T_2 mit

$$|g_i(x) - T_i(x)| < \frac{\varepsilon}{2} \quad (x \in \mathbb{R}; i = 1, 2),$$

und wegen $f(x) = g_1(x) + g_2(x) \sin x$ ist

$$\begin{aligned} |f(x) - T_1(x) - T_2(x) \sin x| &\leq |g_1(x) - T_1(x)| + |g_2(x) \sin x - T_2(x) \sin x| \\ &\leq |g_1(x) - T_1(x)| + |g_2(x) - T_2(x)| < \varepsilon \quad (x \in \mathbb{R}). \end{aligned}$$

Es sei schließlich f eine beliebige Funktion, die den Voraussetzungen des Satzes genügt. Dann gibt es zu den Funktionen

$$f_1(x) := f(x) \sin^2 x, \quad f_2(x) := f\left(x - \frac{\pi}{2}\right) \sin^2 x$$

nach dem zuletzt Bewiesenen zu vorgegebenem $\varepsilon > 0$ trigonometrische Polynome S_1, S_2 mit

$$|f_i(x) - S_i(x)| < \frac{\varepsilon}{2} \quad (x \in \mathbb{R}; i = 1, 2).$$

Wegen $f(x) = f_1(x) + f_2\left(x + \frac{\pi}{2}\right)$ folgt

$$\begin{aligned} \left| f(x) - S_1(x) - S_2\left(x + \frac{\pi}{2}\right) \right| &\leq |f_1(x) - S_1(x)| + \left| f_2\left(x + \frac{\pi}{2}\right) - S_2\left(x + \frac{\pi}{2}\right) \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (x \in \mathbb{R}), \end{aligned}$$

und Satz 2 ist bewiesen.

Namen- und Sachverzeichnis

- Abbildung, kontrahierende 182
ABEL, N. H. 39, 54, 129, 181
abgeschlossene Menge 65
absolut konvergente Reihe 136
 er Betrag 18
Abstand 60
Addition geordneter Paare 55
 — reeller Zahlen 9
Additionstheoreme der trigonometrischen
 Funktionen 68, 70
 — der hyperbolischen Funktionen 185
D'ALEMBERT, J. B. LE ROND 106, 133
algebraische Funktion 38
alternierende Reihe 134
Anfangsindex 90
a-posteriori-Abschätzung 183
a-priori-Abschätzung 183
Approximation von Funktionen 195
ARCHIMEDES 76, 129
Area-Funktionen 187
Argument einer komplexen Zahl 80
Arkus 74
Arkuskosinusfunktion 73
Arkuskotangensfunktion 75
Arkussinusfunktion 73
Arkustangensfunktion 75
Assoziativgesetz 32
äußerer Punkt 62
- BANACH, S. 181
Banachscher Fixpunktsatz 182
bedingt konvergente Reihe 142
Begrenzungspunkt 11, 63
BERNOULLI, JAKOB 14, 129, 186
Bernoullische Gleichung 14, 46
BERNOULLI, JOHANN 28, 54, 129, 134, 186
BERNŠTEJN, S. N. 198
Bernštejsche Polynome 198
beschränkte Funktion 29
 — Menge 61
-s Intervall 11
bestimmt divergente Zahlenfolge 103
- Betrag, absoluter 18
 — einer komplexen Zahl 23
binomische Formel 10
Bogenlänge 84
BOLZANO, B. 28, 106, 109, 129, 155, 173
BOREL, E. 125
BRIGGS, H. 54
BUNJAKOVSKIJ, V. J. 57
Bunjakovskische Ungleichung 57
BÜRGI, J. 54
- CANTOR, G. 124
CARDANO, H. 180, 181
CAUCHY, A.-L. 106, 121, 129, 133, 136, 138,
 140, 155
Cauchyfolge 121
Cauchysche Produktreihe 149
-s Konvergenzkriterium 122, 133
CAVALIERI, B. 106
CAYLEY, A. 59
ČEBYŠEV, P. L. 194, 196
Čebyšev-Metrik 194
CHUQUET, N. 53
c-Stelle 30
- Definitionsbereich 27
Dehnung 32
dekadischer Logarithmus 52
DESCARTES, R. 28, 53, 54
DIOPHANTOS 53
DIRICHLET, P. G. LEJEUNE- 28, 143
divergente Reihe 129
 — Zahlenfolge 98
Doppelfolge 146
Doppelreihe 145
Dreiecksungleichung 18, 56, 58, 67
Durchmesser einer Menge 61
- echt gebrochene rationale Funktion 37
Einschränkung einer Funktion 27
EULER, L. 28, 54, 81, 109, 123, 129, 133, 153,
 196

Eulersche Konstante 123

— Relation 153

— Zahl 109, 151

euklidischer Raum 56

ε -Umgebung 60

—, punktierte 61

ε -Streifen 93

Exponentialfunktion 48, 49

—, komplexe 152

Exponentialreihe 150

Fehlerabschätzung bei alternierenden Reihen
135

FERRARI, L. 181

DEL FERRO, 180, 181

Fixpunkt 181

Fixpunktsatz von BANACH 182

Folge 27

—, konvergente 105

— der Partialsummen 128

—, stationäre 90

Fortsetzung einer Funktion 27

—, monotone 41

FOUBIER, J. 28

Fundamentalfolge 121

Fundamentalsatz der klassischen Algebra 177
Funktion 27

—, algebraische 38

—, beschränkte 29

—, echt gebrochene 37

—, ganzrationale 34, 37

—, gerade 30

—, gleichmäßig stetige 174

—, inverse 27

—, komplexe 29

—, komplexwertig 29

—, konstante 27

—, lineare 176

—, linksseitig stetige 169

—, monotone 39

—, periodische 30

—, quadratische 176

—, rationale 36

—, rechtsseitig stetige 169

—, reelle 29

—, reellwertig 29

—, in einer Menge stetige 171

—, in einem Punkt stetige 155, 157

—, umkehrbare 27

—, ungerade 30

—, in einem Punkt unstetige 155

— von n Variablen 27

—, zusammengesetzte 32

—en, hyperbolische 184

—en, trigonometrische 67

Funktionalgleichung der Exponentialfunk-
tion 49, 150

— der Logarithmusfunktion 52

—en der trigonometrischen Funktionen 68

Funktionenfolge, gleichmäßige konvergente
191

—, punktweise konvergente 189

Funktionensystem, linear unabhängiges 34

GALILEI, G. 186

ganzrationale Funktion 34, 37

GAUSS, C. F. 22, 59, 129, 133, 177

Gaußsche Zahlenebene 22

Gebiet 66

geometrische Reihe 132

geordneter Körper 11

gerade Funktion 30

Gesetz des organischen Wachstums 117

gleichmäßig stetige Funktion 174

— konvergente Funktionenfolge 191

Glied einer Folge 28

Grad einer Funktion 36

GRANDI, G. 129

Graph einer Funktion 29

GRASSMANN, H. 59

GREGORY, J. 129

Grenze, obere 15

—, untere 16

Grenzfunktion 189

Grenzwert einer Folge 105

— einer Funktion 163

—, linksseitiger 168

—, rechtsseitiger 168

— einer Zahlenfolge 98

Grenzwertsätze für Folgen 102

— für Funktionen 167

größtes Ganzes 17

Grundgesetz der Stetigkeit 15

Halbierungsverfahren 125

Halbwinkelformeln 70

harmonische Reihe 131

Hauptargument einer komplexen Zahl 80

Häufungspunkt 63

—, uneigentlicher 166

Häufungswert 117

HEINE, E. 125

Heine-Borelscher Überdeckungssatz 125

HERON 53

L'HUILLIER, S. 106

HUYGENS, CHR. 54, 186

hyperbolische Funktionen 184

- Imaginärteil 22
Infimum 16
innerer Punkt 62
inverse Funktion 27
— Korrespondenz 27
Interpolation 196
Intervall, offenes 60
Intervallschachtelung 108
Irreflexivität 10
isolierter Punkt 64
Iterationsschritt 183
- JACOBI, C. G. J. 59
JUNGIUS, J. 186
- KANT, I. 59
Kehrwert 33
KEPLER, J. 54
Kettenlinie 186
kompakte Menge 124
Komplementärmenge 62
komplexe Exponentialfunktion 152
— Funktion 29
— Zahl 21
— —; trigonometrische Darstellung 80
— Zahlenfolge 28
komplexwertige Funktion 29
konjugiert komplexe Zahl 22
Konnexität 10
konstante Funktion 27
kontrahierende Abbildung 182
Kontraktionsfaktor 184
konvergente Folge 105
— Reihe 129
— Zahlenfolge 98
Konvergenzkriterium von CAUCHY 122, 133
— von LEIBNIZ 134
— von WEIERSTRASS 193
— für monotone Zahlenfolgen 106
konvexe Menge 59
Koordinatenlinien 81
Koordinatensystem, krummliniges 81
Körper, geordneter 11
—, stetig geordneter 15
— der komplexen Zahlen 21
— der reellen Zahlen 9
Korrespondenz 27
Kosinusfunktion 69
Kotangensfunktion 73
Kriterien für absolute Konvergenz 136
- LACROIX, S. F. 28
LAGRANGE, J. L. 59, 106, 198
Lagrangesche Interpolationsformel 198
- LAMBERT, J. H. 187
LEIBNIZ, G. W. 28, 54, 106, 129, 134, 186
Leibnizsches Konvergenzkriterium 134
Limes einer Folge 105
— inferior 119
— superior 119
— einer Zahlenfolge 98
linear unabhängiges Funktionensystem 34
-e Funktion 176
LOBAČEVSKIJ, N. I. 28
Logarithmengesetz 52
Logarithmus, dekadischer 52
— natürlicher 109, 117
Logarithmusfunktion 51
- Majorantenkriterium, erstes 136
—, zweites 137
Maximum 16
Menge, abgeschlossene 65
—, beschränkte 61
—, nach oben beschränkte 15
—, nach unten beschränkte 16
—, in einer Menge dichte 11
—, relativ dichte 41
—, kompakte 124
—, konvexe 59
—, offene 65
—, präkompakte 124
—, sternförmige 59
— der ganzen Zahlen 9
— der natürlichen Zahlen 9
— der rationalen Zahlen 9
— der reellen Zahlen 9
- MENGOLE, P. 132
MERCATOR, N. 54, 129
Metrik 67
metrischer Raum 67
— —, vollständiger 124
Minimum 16
Mittelpunkt einer ε -Umgebung 60
DE MOIRVE, A. 81, 187
Moivresche Formel 81
monotone Fortsetzung 41
— Funktion 39
— Zahlenfolge 90
Monotoniegesetz der Addition 10
— der Multiplikation 11
Multiplikation reeller Zahlen 9
- Näherungsverfahren, selbstkorrigierendes 183
natürlicher Logarithmus 109, 117
NEPER, J. 54
NEWTON, I. 54, 106, 129, 150, 153, 187
 n -te Potenz 9

- n -te Wurzel 20
- Norm 56, 57, 67
- normierter Raum 67
 - —, vollständiger 124
- Nullfolge 93
- Nullfunktion 33
- Nullstelle 30
- obere Grenze 15
 - Schranke 15
- offene Menge 65
- r Rechtecksbereich 65
- s Intervall 60
- Ordnung reeller Zahlen 10
- ORESME, N. 53, 129, 132
- PACIOLI, L. 53
- Partialsummenfolge 128
- PASCAL, B. 59
- Periode 29
 - , primitive 29, 72
- periodische Funktion 30
- Periodizitätseigenschaften der trigonometrischen Funktionen 72
- Polarkoordinaten, Ebene 81
 - , räumliche 82
- Poldistanz 82
- Polstelle 36
- Polynom 34
 - , trigonometrisches 200
- e, Bernsteinsche 198
- Potenzfunktion 34
 - , allgemeine 53
- Potenzgesetz 10, 48
- Potenzproduktfunktion 37
- Potenzreihe 142
- präkompakte Menge 124
- primitive Periode 29, 72
- Prinzip der kontrahierenden Abbildungen 182
- Produktreihe, Cauchysche 149
- Punkt, äußerer 62
 - , innerer 62
 - , isolierter 64
- punktierte ε -Umgebung 61
- punktweise konvergente Funktionenfolge 189
- quadratische Funktion 176
- Quotientenkriterium 140
- Radikand 20
- Radius einer ε -Umgebung 60
- Randpunkt 63
- rationale Funktion 36
- Raum, euklidischer 56
- Raum, metrischer 67
 - , normierter 67
- Realteil 22
- Rechtecksbereich 65
 - , offener 65
- reelle Funktion 29
 - Zahlenfolge 28
- reellwertige Funktion 29
- Reihe 128
 - , absolut konvergente 136
 - , alternierende 134
 - , bedingt konvergente 142
 - , bestimmt divergente 129
 - , divergente 129
 - , geometrische 132
 - , harmonische 131
 - , konvergente 129
 - , unbedingt konvergente 143
- rein imaginäre komplexe Zahl 22
- Restdivision 37
- RICCATI, V. 187
- RIEMANN, B. 59
- Riemannscher Umordnungssatz 144
- RUFFINI, P. 167, 181
- Satz von BOLZANO 173
 - von BOLZANO-WEIERSTRASS 118, 120
 - von CANTOR 124
 - vom Maximum und Minimum 172
 - von WEIERSTRASS 198, 200
- Schranke 61
 - , obere 15
 - , untere 16
- SCHWARZ, H. A. 57
- Schwarzsche Ungleichung 57
- Sehnenzug 83
- selbstkorrigierendes Näherungsverfahren 183
- Signum 19
- Sinusfunktion 69
- Skalarprodukt 58
- Spiegelung 30
- Sprungstelle 170
- Startwert 184
- stationäre Folge 90
- Stauchung 32
- sternförmige Menge 59
- stetig geordneter Körper 15
- Stetigkeit einer Funktion 154, 157, 161
 - in Punktmengen 171
- Stetigkeitspunkt 170
- STIFEL, M. 54
- Stützstelle 196
- Subtraktionstheoreme der trigonometrischen Funktionen 69

Summe einer Reihe 129
Supremum 15
Symmetrie 67

Tangensfunktion 73
TARTAGLIA, N. 180
TAYLOR, B. 196
Taylorsche Formel 196
Teilfolge 91
THOMAS, A. 129
Translation 31
Transitivität 10
transzendente Zahl 152
trigonometrische Darstellung komplexer
 Zahlen 80
 — Funktionen 67
 —s Polynom 200

Überdeckung 125
Überdeckungssatz von HEINE-BOREL 125
Umkehrfunktion 27, 33
Umkehrkorrespondenz 27
Umordnung einer Reihe 142
Umordnungssatz, großer 147
 —, kleiner 143
 —, Riemannscher 144
unbedingt konvergente Reihe 143
unbeschränktes Intervall 11
uneigentliche Zahlen 11
-r Häufungspunkt 166
ungerade Funktion 30
Ungleichung, Bernoullische 14, 46
 —, Bunjakovskische 57
 —, Schwarzsche 57
Unstetigkeitspunkt 170
Unstetigkeitsstelle 172

untere Grenze 16
 — Schranke 16

VANDERMONDE, A. T. 197
Vandermondesche Determinante 197
Verbindungsstrecke 58
Vergleichskriterium für Nullfolgen 96
vollständiger metrischer Raum 124
 — normierter Raum 124

WALLIS, J. 106
WATT, J. 196
WEIERSTRASS, K. 18, 109, 198, 200
Weierstraßsches Konvergenzkriterium 193
Wertebereich 27
Wurzelkriterium 138

Zahl, Eulersche 151
 —, komplexe 21
 —, —; trigonometrische Darstellung 80
 —, konjugiert komplexe 22
 —, rein imaginäre komplexe 22
 —, transzendente 152
 -en, uneigentliche 11
Zahlenfolge 28
 —, bestimmt divergente 103
 —, divergente 98
 —, konvergente 98
 —, monotone 90
Zerlegung eines Intervalls 82
 — in Linearfaktoren 178
 — — — und quadratische Faktoren ohne
 reelle Nullstellen 179
zusammengesetzte Funktion 32
Zwischenwertsatz 41, 174
Zylinderkoordinaten 82