

Mathematik für Physiker I

Ausarbeitung einer Vorlesung
vom Wintersemester 2007/08

JOACHIM WEIDMANN

Fachbereich Informatik und Mathematik
der Universität Frankfurt

Stand 6. Februar 2008

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Zahlen (natürliche, ganze, rationale, reelle) | 5 |
| 2 | Konvergenz, Vollständigkeit von \mathbb{R} | 14 |
| 3 | Exponential- und Logarithmusfunktionen | 24 |
| 4 | Weitere elementare Funktionen | 33 |
| 4.1 | Zusammengesetzte Funktionen | 33 |
| 4.2 | Polynome | 33 |
| 4.3 | Rationale Funktionen | 34 |
| 4.4 | Trigonometrischen Funktionen | 35 |
| 5 | Komplexe Zahlen und elementare komplexe Funktionen | 41 |
| 5.1 | Komplexe Zahlen | 41 |
| 5.2 | Komplexe Exponential- und Logarithmusfunktionen | 44 |
| 6 | Unendliche Reihen | 46 |
| 6.1 | Konvergente und divergente Reihen | 46 |
| 6.2 | Konvergenzkriterien | 49 |
| 6.3 | Potenzreihen | 54 |
| 7 | Vektorräume; vom \mathbb{R}^2 zum \mathbb{C}^n | 56 |
| 7.1 | Koordinaten und Vektoren | 57 |
| 7.2 | Lineare Abbildungen | 61 |
| 7.3 | Lineare Un-/Abhängigkeit | 62 |
| 7.4 | Norm und Skalarprodukt | 63 |
| 7.5 | Konvergenz in \mathbb{R}^m und \mathbb{C}^m | 66 |

| | |
|---|------------|
| <i>INHALTSVERZEICHNIS</i> | 3 |
| 8 Eigenschaften stetiger Funktionen | 68 |
| 8.1 Grundbegriffe und Beispiele | 68 |
| 8.2 Grenzwerte von Funktionen | 71 |
| 8.3 Eigenschaften stetiger Funktionen einer Variablen | 72 |
| 8.4 Extremwerte von Funktionen | 75 |
| 9 Differentiation | 78 |
| 10 Eigenschaften differenzierbarer Funktionen | 87 |
| 10.1 Mittelwertsätze | 87 |
| 10.2 Höhere Ableitung und der Satz von Taylor | 89 |
| 10.3 Extremwerte differenzierbarer Funktionen | 94 |
| 10.4 Die Regeln von de l'Hospital | 97 |
| 11 Das Integral für Regelfunktionen | 99 |
| 11.1 Integral für Treppenfunktionen | 99 |
| 11.2 Das Integral für Regelfunktionen | 101 |
| 11.3 Regelfunktionen | 105 |
| 11.4 Berechnung von Integralen | 107 |
| 12 Integrationsmethoden | 111 |
| 12.1 Partielle Integration | 111 |
| 12.2 Substitutionsregel | 115 |
| 12.3 Hyperbelfunktionen | 117 |
| 12.4 Integration rationaler Funktionen | 119 |
| 13 Integration und Grenzübergänge | 127 |
| 13.1 Vertauschung von Integration und Differentiation mit Grenzübergängen . . . | 127 |
| 13.2 Uneigentliche Integrale | 132 |

| | |
|---|------------|
| <i>INHALTSVERZEICHNIS</i> | 4 |
| 14 Vektorräume, Dimension, Basis | 137 |
| 15 Fourierreihen | 145 |
| 15.1 Vorbemerkungen | 145 |
| 15.2 Konvergenz im Quadratmittel | 147 |
| 15.3 Gleichmäßige und punktweise Konvergenz | 154 |
| 15.4 Die Sätze von Weierstraß | 158 |
| 15.5 Übungsaufgaben | 159 |

1 Zahlen (natürliche, ganze, rationale, reelle)

Die reellen Zahlen bilden die Grundlage der Analysis. Aus vielerlei Gründen spielen aber auch die Teilmengen der rationalen (ganzen) Zahlen eine wesentliche Rolle. Für Messungen und Größenangaben würden natürlich rationale Zahlen ausreichen. Für die Analysis, wie wir sie hier betreiben wollen, und wie sie eine wesentliche Grundlage der Physik bildet, sind diese jedoch zu „lückenhaft“ in zweierlei Sinn:

- sie weisen tatsächlich Lücken auf (Zahlenstrahl),
- sie haben ganz wichtige Eigenschaften der reellen Zahlen nicht.

Hier soll allerdings nicht der Zahlaufbau im Detail diskutiert werden:

- $\mathbb{N} = \{1, 2, 3, \dots\}$ *natürliche Zahlen*,
- $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ *ganze Zahlen*,
- $\mathbb{Q} = \{\text{Brüche } \frac{p}{q} \text{ mit ganzen Zahlen } p, q, q \neq 0\}$ *rationale Zahlen*,
- $\mathbb{R} =$ *reelle Zahlen*, und schließlich
- $\mathbb{C} =$ *komplexe Zahlen*.

Wir gehen stattdessen von einem naiven Verständnis der reellen Zahlen und der Rechenoperation in \mathbb{R} aus.

Alle Rechenregeln in \mathbb{R} (deren Kenntnis i. wes. vorausgesetzt wird) beruhen auf der Gültigkeit von ganz wenigen Gesetzen (*Axiomen*), aus denen sich alle Regeln ableiten lassen:

Addition:

- (A1) $(a + b) + c = a + (b + c)$ *Assoziativität*,
- (A2) $a + b = b + a$ *Kommutativität*,
- (A3) \exists *Nullelement* 0 mit $a + 0 = a$ für jedes a , *neutrales Element der Addition*,
- (A4) für jedes a existiert das zugehörige *negative Element* $-a$ mit $a + (-a) = 0$ (entsprechend wird jede Gleichung $a + x = b$ gelöst durch $x = b - a$).

Das Nullelement ist eindeutig bestimmt: Haben 0 und $0'$ diese Eigenschaft, so gilt:

$$0 = 0 + 0' = 0' + 0 = 0'.$$

Multiplikation(M1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ Assoziativität,(M2) $a \cdot b = b \cdot a$ Kommutativität,(M3) Es existiert ein *Einselement* 1 mit $a \cdot 1 = a$, *neutrales Element der Multiplikation*.(M4) Zu jedem $a \neq 0$ gibt es das zugehörige *inverse Element* a^{-1} mit $a \cdot a^{-1} = 1$ (entsprechend wird für jedes $a \neq 0$ die Gleichung $a \cdot x = b$ durch $x = b \cdot a^{-1}$ gelöst.)

Auch das 1-Element ist eindeutig bestimmt ($\dot{\mathbf{U}}$). Entsprechendes gilt für die Lösungen x von $a + x = b$ und $a \cdot x = b$.

Die Addition und die Multiplikation werden verbunden durch das

Distributivgesetz(D) $a \cdot (b + c) = a \cdot b + a \cdot c$.

Den Mal-Punkt werden wir im folgenden meist weglassen. Zur Illustration sollen zwei wichtige Rechenregeln, die daraus folgen, bewiesen werden:

a) $0 \cdot a = 0$ für jedes a : Es gilt

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

Also ist $0 \cdot a$ die eindeutig bestimmte Lösung von $0 \cdot a + x = 0 \cdot a$ und somit 0.

b) $(-1)a = -a$ (wobei -1 bzw. $-a$ die zu 1 bzw. a gehörigen negativen Elemente sind); speziell ist also $(-1)(-1) = 1$: Es gilt

$$0 = 0 \cdot a = (1 + (-1))a = 1 \cdot a + (-1)a = a + (-1)a,$$

also ist $(-1)a$ das zu a gehörige negative Element.

Alle diese Rechenregeln gelten nicht nur in \mathbb{R} , sondern auch in den *rationalen Zahlen*:

$$\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\}.$$

Eine Menge mit diesen Operationen heißt *Körper* (engl.: *field*), wenn außerdem $1 \neq 0$ gilt (stattdessen kann man auch fordern, dass mindestens zwei Elemente enthalten sind).

In \mathbb{Z} gibt es zwar zu jedem Element das negative Element, aber nicht das inverse Element. In \mathbb{N} gibt es auch kein negatives Element. Die Gleichungen $a + x = b$ und $ax = b$ sind also nicht mehr in allen Fällen lösbar (wenn aber eine Lösung existiert, ist sie eindeutig bestimmt).

Jetzt fehlen noch zwei wichtige Eigenschaften der reellen Zahlen; zunächst die

Anordnung. Die *Kleiner-Beziehung* $a < b$ („ a kleiner als b “ oder „ a kleiner b “), auch als *Größer-Beziehung* $b > a$ („ b größer als a “ oder „ b größer a “) geschrieben, hat folgende Eigenschaften:

(O1) Es gilt immer genau eine der Beziehungen

$$a < b, \quad a = b, \quad a > b, \quad (\text{Trichotomie}),$$

(O2) Aus $a < b$ und $b < c$ folgt $a < c$ (*Transitivität*),

(O3) Aus $a < b$ folgt $a + c < b + c$ für jedes c ,

(O4) Aus $a < b$ und $c > 0$ folgt $ac < bc$.

Auf dem Zahlenstrahl entspricht $a < b$ der Tatsache, dass a links von b liegt. Eine Zahl a heißt *positiv*, wenn $a > 0$ gilt, *negativ*, wenn $a < 0$ gilt.

Man schreibt $a \leq b$, („ a kleiner oder gleich b “), wenn $a < b$ oder $a = b$ gilt. Die Rechenregeln für Ungleichungen werden als bekannt vorausgesetzt. Nur einige wichtige Regeln sollen bewiesen werden:

a) Aus $a > 0$ folgt $-a < 0$:

$$0 = a - a > 0 - a = -a. \text{ Entsprechend folgt } -a > 0 \text{ aus } a < 0 \text{ (Ü).}$$

b) Aus $a \neq 0$ folgt $a^2 > 0$:

$$(i) \quad a > 0 \Rightarrow a^2 = aa > a0 = 0,$$

$$(ii) \quad a < 0 \Rightarrow -a > 0 \Rightarrow a^2 = (-a)^2 > 0,$$

1: 16. Oktober 2007

c) Aus $a > 0$ folgt $\frac{1}{a} > 0$:

Wäre $\frac{1}{a} < 0$, so wäre wegen (O4) $1 = a \cdot \frac{1}{a} < 0$, ein Widerspruch. Wäre $\frac{1}{a} = 0$, so wäre $1 = a \cdot \frac{1}{a} = 0$, ebenfalls ein Widerspruch.

Teil c ist ein Beispiel für einen

Indirekten Beweis bzw. einen **Beweis durch Widerspruch**: Besagt der Satz, dass aus einer Voraussetzung V eine Aussage A folgt, so beweist man den Satz wie folgt: Man nimmt an, dass A nicht gilt, kurz $\neg A$ („non A “). Für den Rest des Beweises unterscheidet man zwei Varianten (die aber eigentlich fast identisch sind).

- man folgert aus $\neg A$, dass auch V nicht gilt, also ein *Widerspruch zur Voraussetzung*, oder
- aus V und $\neg A$ folgert man eine Aussage, von der man weiß, dass sie falsch ist (z. B. $2 = 1$ oder $1 < 0$, oder ...), also wieder ein *Widerspruch*.

Eine besonders wichtige Rolle spielt die Menge \mathbb{N} der *natürlichen Zahlen* $\mathbb{N} := \{1, 2, 3, \dots\}$. Dieser naiven Darstellung als Aufzählung (das folgende Glied ist jeweils gleich dem Vorhergehenden $+1$) entspricht das

Induktionsprinzip: Ist M eine Teilmenge von \mathbb{N} mit den Eigenschaften

- a) $1 \in M$, und
- b) ist $k \in M$, so ist auch $k + 1 \in M$,

so gilt $M = \mathbb{N}$.

Daraus ergibt sich eine andere wichtige Eigenschaft von \mathbb{N} , die

Wahlordnung von \mathbb{N} : Jede nichtleere Teilmenge M von \mathbb{N} enthält ein kleinstes Element m .

Beweis. Indirekt: Wir nehmen an, dass M kein kleinstes Element enthält. Sei $M' := \mathbb{N} \setminus M$ die Menge der Elemente aus \mathbb{N} , die nicht in M liegen. Dann liegt jedenfalls 1 in M' , denn sonst wäre es sicher das kleinste Element von M . Nehmen wir nun an, dass $n \in M'$ ist, so liegt auch $n + 1$ in M' , denn sonst wäre es in M und somit kleinstes Element von M . Also ist $M' = \mathbb{N}$ und somit M leer im Widerspruch zur Voraussetzung. ■

Umgekehrt folgt aus der Wohlordnung von \mathbb{N} das Induktionsprinzip.

Beweis. \mathbb{N} sei wohlgeordnet und die Eigenschaften a) und b) seien erfüllt. Für $M' := \mathbb{N} \setminus M$ ist also $M' = \emptyset$ (die leere Menge) zu beweisen. Wir nehmen an, dass M' nicht leer ist. M' besitzt also ein kleinstes Element m . Es ist $m > 1$, da $1 \in M$ gilt (Eigenschaft a). Also ist auch $m - 1$ eine natürliche Zahl; diese liegt nicht in M' , also $m - 1 \in M$. Wegen b) ist dann auch $m = (m - 1) + 1$ aus M . Das ist aber ein Widerspruch zu $m \in M'$. ■

Auf dem Induktionsprinzip beruht der **Beweis durch vollständige Induktion**: Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage (die wahr oder falsch sein kann).

Kann man zeigen, dass

- a) $A(1)$ wahr ist (*Induktionsverankerung*) und
 b) aus der Gültigkeit von $A(n)$ (*Induktionsannahme*) auch die Gültigkeit von $A(n+1)$ folgt (*Induktionsschluß* $n \Rightarrow n+1$),

so ist $A(n)$ wahr für jedes $n \in \mathbb{N}$. (Häufig, insbesondere wenn $A(n)$ eine komplizierte Aussage ist, wird vor dem Schritt b noch die Aussage $A(n)$ explizit formuliert, *Induktionsannahme*.)

Als **Beispiel** beweisen wir die Formel

$$\sum_{k=1}^n k^2 := 1^2 + 2^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

Beweis. Für $n=1$ sind offenbar beide Seiten gleich 1.

$n \Rightarrow n+1$: Gilt die Formel für n , so folgt offenbar

$$\begin{aligned} \sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^n k^2 + (n+1)^2 = \frac{1}{6}n(n+1)(2n+1) + (n+1)^2 \\ &= (n+1) \left\{ \frac{1}{6}n(2n+1) + (n+1) \right\} = \frac{1}{6}(n+1) \{2n^2 + 7n + 6\} \\ &= \frac{1}{6}(n+1) \{(n+2)(2n+3)\} = \frac{1}{6}(n+1) \left((n+1)+1 \right) \left(2(n+1)+1 \right). \end{aligned}$$

Das ist die entsprechende Formel für $n+1$ statt n , womit der Induktionsschritt vollzogen ist. ■

Hier haben wir den Ausdruck $\sum_{k=1}^n a_k$ in naiver Weise benutzt. Genauer ist dies ein schönes Beispiel für *Definition durch Induktion*: Definiert man $\sum_{k=1}^1 a_k := a_1$ und $\sum_{k=1}^{n+1} a_k := \sum_{k=1}^n a_k + a_{n+1}$, so ist offenbar $\sum_{k=1}^n a_k$ für alle $n \in \mathbb{N}$ definiert.

Mit Hilfe der vollständigen Induktion beweisen wir nun noch die häufig benutzte

Satz 1.1 (Bernoullische Ungleichung) Für $h > -1$ und alle $n \in \mathbb{N}$ gilt

$$(1+h)^n \geq 1+nh \quad (\text{mit } > \text{ für } n > 1 \text{ und } h \neq 0).$$

Beweis. ($n=1$): $(1+h)^1 = 1+h$.

$n \Rightarrow n+1$: Aus $(1+h)^n \geq 1+nh$ folgt wegen $(1+h) > 0$

$$\begin{aligned} (1+h)^{n+1} &= (1+h)^n(1+h) \geq (1+nh)(1+h) \\ &= 1+(n+1)h+nh^2 \begin{cases} \geq 1+(n+1)h & \text{für alle } h > -1, \\ > 1+(n+1)h & \text{für } h \neq 0. \end{cases} \end{aligned}$$

Satz 1.3 (Binomischer Lehrsatz) Für $x, y \in \mathbb{R}$ (auch für komplexe Zahlen x, y) und $n \in \mathbb{N}$ gilt

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

(das gilt auch für $n = 0$, dann sind beide Seiten gleich 1).

Der *Beweis* ist eine einfache Induktion unter Verwendung der obigen Resultate (Ü).

Der *Betrag* $|a|$ einer reellen Zahl a ist definiert durch

$$|a| = \text{„}a \text{ Betrag“} := \begin{cases} a & \text{für } a \geq 0, \\ -a & \text{für } a < 0. \end{cases}$$

Dies ist also gewissermaßen der Abstand des Punktes a vom Nullpunkt. $|a - b|$ ist der Abstand der beiden Zahlen a und b .

Satz 1.4 Der Betrag erfüllt die folgenden wichtigen Eigenschaften:

- a) $|a| \geq 0$ für alle $a \in \mathbb{R}$, $|a| = 0$ nur für $a = 0$,
- b) $|a \cdot b| = |a| |b|$,
- c) $|a + b| \leq |a| + |b|$ Dreiecksungleichung,
- d) $\left| |a| - |b| \right| \leq |a - b|$ Dreiecksungleichung nach unten.

Beweis. a) und b) sind leicht nachzurechnen.

c) Mit Hilfe von b) ergibt sich offenbar

$$\begin{aligned} |a + b|^2 &= (a + b)^2 = a^2 + 2ab + b^2 \leq a^2 + |2ab| + b^2 \\ &= a^2 + 2|a||b| + b^2 = (|a| + |b|)^2, \end{aligned}$$

woraus das Resultat folgt.

d) Aus c) folgt

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

also

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

Entsprechend folgt $|b| \leq |a| + |a - b|$, also

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

und somit die Behauptung. ■

Nun noch zur *Lückenhaftigkeit* von \mathbb{Q} . Es scheint selbstverständlich, dass es zu jeder positiven Zahl a eine Zahl r geben sollte mit $r^2 = a$ (r ist z. B. die Seitenlänge eines Quadrats mit der Fläche a). Wenn \mathbb{Q} den ganzen Zahlenstrahl ausfüllen würde, müßte es also insbesondere ein $r = \frac{p}{q} \in \mathbb{Q}$ geben mit $\frac{p^2}{q^2} = r^2 = 2$. Tatsächlich gilt aber:

Satz 1.5 *Es gibt keine rationale Zahl $r = \frac{p}{q}$ mit $r^2 = 2$. (r ist die Länge der Hypothenuse eines rechtwinkligen Dreiecks mit Kathetenlängen 1 bzw. die Länge der Diagonalen eines Quadrats mit Kantenlänge 1.)*

Beweis. Ohne Einschränkung können wir annehmen, dass p und q teilerfremd sind (d. h. der Bruch ist gekürzt). Aus

$$p^2 = 2q^2$$

folgt, dass p gerade ist, $p = 2p'$, also

$$2p'^2 = q^2.$$

Daraus wiederum folgt, dass q gerade ist. p und q sind also nicht teilerfremd. ■

Eine Menge A heißt *abzählbar*, wenn sich die Menge in einer Folge (a_n) anordnen läßt, d. h.

$$A = \{a_n : n \in \mathbb{N}\} = \{a_1, a_2, a_3, \dots\}.$$

die letzte Darstellung nennen wir auch eine *Abzählung* von A .

Satz 1.6 *Die Menge Q der rationalen Zahlen ist abzählbar. (Wir werden sehen, dass dies für die reellen Zahlen \mathbb{R} nicht gilt.)*

Beweis. In der (unendlichen) Matrix

$$\begin{array}{cccccc} 1 & 1/2 & 1/3 & 1/4 & 1/5 & \dots \\ 2 & 2/2 & 2/3 & 2/4 & 2/5 & \dots \\ 3 & 3/2 & 3/3 & 3/4 & 3/5 & \dots \\ 4 & 4/2 & 4/3 & 4/4 & 4/5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

kommen offenbar alle positiven rationalen Zahlen vor (sogar ∞ oft). Die Matrix kann nun so durchlaufen werden:

$$1, 1/2, 2, 3, 2/2, 1/3, 1/4, 2/3, 3/2, 4, 5, 4/2, \dots$$

(damit sollte klar sein, wie es weiter geht). Zählt man nun diese Reihenfolge entsprechend ab – wobei man Zahlen, die schon vorkamen, wegläßt (also z. B. $2/2 = 1$, $4/2 = 2$, ...) – so erhält man eine Abzählung der positiven rationalen Zahlen $\{q_1, q_2, q_3, \dots\}$. Dann erhält man mit $\{0, q_1, -q_1, q_2, -q_2, \dots\}$ offenbar eine Abzählung aller rationalen Zahlen. ■

An diesem Beweis erkennt man auch, dass jede Vereinigung abzählbar vieler abzählbarer Mengen abzählbar ist (Ü).

Wir vereinbaren die folgenden Schreibweisen für *Intervalle*:

- $[a, b]$ *abgeschlossenes Intervall* $\{x \in \mathbb{R} : a \leq x \leq b\}$
- $(a, b]$ und $[a, b)$ *halboffene Intervalle* $\{x \in \mathbb{R} : a < x \leq b\}$ bzw. $\{x \in \mathbb{R} : a \leq x < b\}$
- (a, b) *offenes Intervall* $\{x \in \mathbb{R} : a < x < b\}$

an Stellen mit „(“ bzw. „)“ ist auch $-\infty$ bzw. ∞ möglich.

2:18. Oktober 2007

2 Konvergenz, Vollständigkeit von \mathbb{R}

Alle bisher angegebenen Eigenschaften der reellen Zahlen werden auch von den rationalen Zahlen \mathbb{Q} erfüllt. Wir haben zwar gesehen, dass in \mathbb{Q} keine Zahl existiert, deren Quadrat 2 ist. Aber, ob es die in \mathbb{R} gibt, ist auch nicht offensichtlich. Mit Hilfe einer weiteren Eigenschaft, die die reellen Zahlen erfüllen, werden wir das zeigen können.

Die Eigenschaft von \mathbb{R} , die uns noch fehlt, ist die *Vollständigkeit*; hierfür gibt es mindestens 6 verschiedene äquivalente Definitionen (Supremumsaxiom, Monotonieprinzip, Gültigkeit des CAUCHYSchen Konvergenzkriteriums, Satz von BOLZANO–WEIERSTRASS, Intervallschachtelungsprinzip und das DEDEKINDSche Schnitaxiom), die wir zumindest teilweise nach und nach kennen lernen werden. Hier benutzen wir als Definition der *Vollständigkeit* das

Supremumsaxiom: Jede nicht–leere nach oben beschränkte Teilmenge M von \mathbb{R} besitzt eine kleinste obere Schranke; diese wird als *Supremum* von M , $\sup M$, bezeichnet. $s = \sup M$ hat also die Eigenschaften:

- s ist eine obere Schranke von M , d. h. $x \leq s$ für jedes $x \in M$.
- ist $t < s$, so ist es nicht obere Schranke von M (d. h. es gibt mindestens ein $x \in M$ mit $x > t$).

Natürlich ist das Axiom äquivalent zu dem entsprechenden *Infimumsaxiom*: Jede nach unten beschränkte Teilmenge M von \mathbb{R} besitzt eine größte untere Schranke, das Infimum von M , $\inf M$ (Ü).

Beispiele Das Intervall $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$ mit $a < b$ hat das Supremum b und das Infimum a . Die Menge $\left\{1 - \frac{1}{n} : n \in \mathbb{N}\right\}$ hat das Infimum 0 und das Supremum 1.

Aus dem obigen Axiom ergibt sich zunächst die

Archimedische Eigenschaft von \mathbb{R} : Zu zwei beliebigen positiven Zahlen $a, b > 0$ gibt es ein $n \in \mathbb{N}$ mit $n \cdot a > b$. Oder: Zu jeder positiven Zahl $r > 0$ gibt es ein $n \in \mathbb{N}$ mit $n > r$, oder: Zu jedem $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.

Beweis. Wir beweisen die zweite Aussage: *Annahme:* Es gibt kein $n \in \mathbb{N}$ mit $n > r$, d. h. die Menge \mathbb{N} ist durch r nach oben beschränkt, besitzt also ein Supremum $s := \sup \mathbb{N}$. Dann ist also $n + 1 \leq s$ bzw. $n \leq s - 1$ für alle $n \in \mathbb{N}$. Das bedeutet, dass $s - 1$ eine obere Schranke von \mathbb{N} ist, im Widerspruch zur Definition von s (als kleinste obere Schranke). ■

Folgerung 2.1 In jedem Intervall (a, b) mit $a < b$ (so klein das auch sein mag) liegt eine (sogar unendlich viele) rationale Zahl(en); man sagt hierfür auch \mathbb{Q} liegt dicht in \mathbb{R} . Ist $n \in \mathbb{N}$ so, dass $\frac{1}{n} < \frac{b-a}{2}$ gilt, so gibt es ein m mit $a < \frac{m}{n} < \frac{m+1}{n} < b$, woraus alles folgt.

Satz 2.2 Jede Zahl $a > 0$ hat in \mathbb{R} eine positive reelle Quadratwurzel. (Entsprechendes gilt für n -te Wurzeln, $n \in \mathbb{N}$ (Ü).)

Beweis. Dazu sei $M := \{x \in \mathbb{R} : x^2 < a\}$. Offenbar ist M nicht leer (z. B. $\min\{1, a\} \in M$) und nach oben beschränkt (z. B. durch $\max\{1, a\}$). Sei also $q := \sup M$. Wir zeigen $q^2 = a$. Dabei benutzen wir, dass offensichtlich $1 < q < 2$ gilt.

$q^2 \leq a$: Annahme: $q^2 > a$; also z. B. $q^2 > a + 10^{-k_0}$.

Wählen wir nun ein $x \in M$ mit $x > q - 10^{-k-1}$ mit $k \geq k_0$, so gilt für hinreichend großes k (beachte $q < \max\{1, a\}$)

$$x \in M \quad \text{und} \quad x^2 > q^2 - 2q10^{-k-1} + 10^{-2k-2} > q^2 - 10^{-k_0} > a,$$

ein Widerspruch.

$q^2 \geq a$: Annahme: $q^2 < a$, also z. B. $q^2 < a - 10^{-k_0}$ wählen wir ein $x \in \mathbb{R}$ mit $q < x < q + 10^{-k-1}$ für $k \geq k_0$, so gilt für hinreichend großes k (beachte $q < \max\{1, a\}$)

$$x^2 < q^2 + 2q10^{-k-1} + 10^{-2k-2} < q^2 + 10^{-k_0} < 2,$$

d. h. $x \in M$ und $x > q$ im Widerspruch zur Wahl von q als Supremum von M .

Es bleibt also nur die Möglichkeit $q^2 = a$. ■

Eine Folge (x_n) in \mathbb{R} (oder entsprechend in jeder beliebigen Menge) ist eine Zuordnung (Funktion), die jedem $n \in \mathbb{N}$ (oder \mathbb{N}_0 , evtl. auch $\{k, k+1, \dots\}$) ein Element $x_n \in \mathbb{R}$ zuordnet; man schreibt häufig auch (x_1, x_2, x_3, \dots) , konkret z. B. $(2, 4, 6, 8, \dots)$.

Eine reelle Folge (x_n) konvergiert gegen den Grenzwert (Limes) $x \in \mathbb{R}$, $x = \lim_{n \rightarrow \infty} x_n = \lim x_n$ oder $x_n \rightarrow x$ für $x \rightarrow \infty$ oder $x_n \rightarrow x$, wenn die x_n für hinreichend große n der Zahl x beliebig nahe kommen; mathematisch sauberer formuliert man das so:

Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $|x_n - x| < \varepsilon$ für $n \geq N$.

Mit sogenannten Quantoren $\exists =$ „es existiert“ und $\forall =$ „für alle“ schreibt man abgekürzt:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \quad |x_n - x| < \varepsilon.$$

Eine gegen 0 konvergente Folge wird auch als *Nullfolge* bezeichnet. (x_n) ist also genau dann eine Nullfolge, wenn gilt

$$\forall \varepsilon > 0 \quad \exists N \quad \forall n \geq N \quad |x_n| < \varepsilon.$$

Z. B. ist (x_n) mit $x_n = \frac{1}{n}$ offensichtlich eine Nullfolge.

Ist eine Folge (x_n) nicht konvergent, so heißt sie *divergent*. Gibt es zu jedem $C \in \mathbb{R}$ ein N mit $x_n > C$ für $n \geq N$, so sagt man auch, (x_n) ist *bestimmt divergent* (auch: *konvergent*) gegen ∞ . Entsprechend definiert man für $-\infty$.

Satz 2.3 *Es gilt $x_n \rightarrow x$ genau dann, wenn $(x_n - x)$ eine Nullfolge ist.*

Eine *Teilfolge* von (x_n) erhält man aus der Folge (x_n) durch Weglassen gewisser Glieder (so, dass unendliche viele Glieder übrig bleiben), z. B.

$$(x_1, x_3, x_5, \dots), (x_1, x_4, x_9, \dots, x_{n^2}, \dots)$$

(das muss allerdings nicht so einfach gesetzmäßig sein, wie in diesen Beispielen).

Da man alle Konvergenzaussagen auf Konvergenz gegen 0 zurückführen kann, ist es sinnvoll, Nullfolgen genauer zu untersuchen. Hier einige wichtige Eigenschaften, die leicht zu beweisen sind:

Satz 2.4 (Eigenschaften von Nullfolgen) a) *Ist (y_n) eine Nullfolge und gilt $|x_n| \leq y_n$, so ist auch (x_n) eine Nullfolge.*

b) *Ist (x_n) eine Nullfolge, so ist auch (cx_n) eine Nullfolge für jedes $c \in \mathbb{R}$.*

c) *Jede Teilfolge einer Nullfolge ist eine Nullfolge.*

d) *Für $|q| < 1$ ist (q^n) eine Nullfolge (für $|q| \geq 1$ ist dies offensichtlich nicht der Fall).*

e) *Sind (x_n) und (y_n) Nullfolgen, so sind auch $(x_n \pm y_n)$ Nullfolgen.*

f) *Jede Nullfolge (x_n) ist beschränkt.*

g) *Ist (x_n) eine Nullfolge und (y_n) beschränkt (d. h. es gibt ein c mit $|y_n| \leq c$ für alle n), so ist auch $(x_n y_n)$ eine Nullfolge*

h) *Ist (x_n) eine Nullfolge, so ist auch $(\sqrt[m]{|x_n|})$ eine Nullfolge (entsprechendes gilt für m -te Wurzeln, $m \in \mathbb{N}$).*

Beweis.

- a) Wir setzen $|q| = \frac{1}{1+h}$ (also $h = \frac{1}{|q|} - 1 > 0$). Dann folgt aus der Bernoullischen Ungleichung

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

also $|q^n| \leq \frac{1}{nh}$, $|q^n| \rightarrow 0$, $q^n \rightarrow 0$.

- f) Da (x_n) eine Nullfolge ist, existiert (zu $\varepsilon = 1$) ein N mit $|x_n| < 1$ für alle $n > N$. Damit folgt für alle $n \in \mathbb{N}$

$$|x_n| \leq C := \max\{1, x_1, \dots, x_N\}.$$

- h) Da (x_n) Nullfolge ist, gibt es zu jedem $\varepsilon > 0$ ein N mit $|x_n| < \varepsilon^2$ für $n \geq N$. Dann gilt aber $\sqrt{|x_n|} < \varepsilon$ für $n \geq N$.

■

3: 23.10.2007

Satz 2.5 (Aussagen über konvergente Folgen) a) *Jede konvergente Folge (x_n) ist beschränkt.*

- b) *Ist (x_n) konvergent gegen x , so konvergiert $(|x_n|)$ gegen $|x|$.*
 c) *Ist $M \subset \mathbb{R}$ und $s = \sup M$, so gibt es eine Folge (x_n) aus M mit $x_n \rightarrow x$ (es ist möglich, dass dies nur mit der konstanten Folge $x_n = s$ gilt).*
 d) *Dreifolgensatz: Gilt $x = \lim x_n = \lim y_n$ und $x_n \leq w_n \leq y_n$, so folgt $x = \lim w_n$.*
 e) *Aus $x = \lim x_n$, $y = \lim y_n$ und $\alpha, \beta \in \mathbb{R}$ folgt*

$$\lim(\alpha x_n + \beta y_n) = \alpha x + \beta y, \quad \lim x_n y_n = xy.$$

- f) *Gilt in e) außerdem $y \neq 0$, so folgt auch*

$$\lim \frac{x_n}{y_n} = \frac{x}{y}, \quad \text{insbes. } \frac{1}{y_n} \rightarrow \frac{1}{y}.$$

(Wobei zu beachten ist, dass in der Folge $\left(\frac{x_n}{y_n}\right)$ u. U. endlich viele Glieder gar nicht definiert sind, da $y_n = 0$ sein kann).

- g) *Jede Teilfolge einer konvergenten Folge ist konvergent gegen den gleichen Grenzwert.*
 h) *Aus $x_n \rightarrow x$ folgt $x_n^m \rightarrow x^m$ für $n \rightarrow \infty$ und jedes $m \in \mathbb{N}$.*

- i) Ist (x_n) eine Folge mit $x_n \geq 0$ und $x_n \rightarrow x$, so folgt $\sqrt[m]{x_n} \rightarrow \sqrt[m]{x}$ für jedes $m \in \mathbb{N}$.
Für Nullfolgen wurde dies bereits oben bewiesen.

Beweis.

- a) Ist $x = \lim x_n$, so ist $(x_n - x)$ eine Nullfolge. Da diese beschränkt ist, ist dann auch die Folge $(x_n) = \left(x + (x_n - x)\right)$ beschränkt.
e) Nur das Produkt ist interessant: Ist C eine Schranke von (x_n) , so gilt

$$|x_n y_n - xy| = |(x_n - x)y_n + (y_n - y)x| \leq C|x_n - x| + |x| |y_n - y| \rightarrow 0.$$

- f) Für hinreichend große n ist $|y_n| > \frac{1}{2}|y|$, insbesondere $y_n \neq 0$, und somit

$$\begin{aligned} \left| \frac{x_n}{y_n} - \frac{x}{y} \right| &= \left| \frac{x_n y - x y_n}{y_n y} \right| \leq \frac{2}{|y|^2} \left| (x_n - x)y + x(y - y_n) \right| \\ &\leq \frac{2}{|y|} |x_n - x| + 2 \frac{|x|}{|y|^2} |y - y_n| \rightarrow 0 \text{ für } n \rightarrow \infty. \end{aligned}$$

- h) Sei M eine Schranke von $(|x_n|)$. Dann folgt mit Hilfe der Formel $a^m - b^m = (a - b) \sum_{j=1}^m a^{m-j} b^{j-1}$

$$\begin{aligned} |x_n^m - x^m| &= |x_n - x| \left| \sum_{j=1}^m x_n^{m-j} x^{j-1} \right| \leq |x_n - x| \sum_{j=1}^m |x_n^{m-j}| |x^{j-1}| \\ &\leq |x_n - x| m M^{m-1} \rightarrow 0 \text{ für } n \rightarrow \infty. \end{aligned}$$

- i) Ist $x \neq 0$, so gilt für hinreichend große n $|x_n| > \frac{x}{2}$, und somit (mit der gleichen Formel wie in h))

$$\begin{aligned} \left| \sqrt[m]{x_n} - \sqrt[m]{x} \right| &= \left| \frac{\sqrt[m]{x_n} - \sqrt[m]{x}}{\sum_{j=1}^m (\sqrt[m]{x_n})^{m-j} (\sqrt[m]{x})^{j-1}} \sum_{j=1}^m (\sqrt[m]{x_n})^{m-j} (\sqrt[m]{x})^{j-1} \right| \\ &\leq \frac{1}{m(\sqrt[m]{1/2})^{m-1}} |x_n - x| \rightarrow 0 \text{ für } n \rightarrow \infty. \end{aligned}$$

Ist $x = 0$, so existiert zu jedem $\varepsilon > 0$ ein N mit $x_n < \varepsilon^m$ für $n \geq N$, also $\sqrt[m]{x_n} < \varepsilon$ für $n \geq N$, d. h. $\sqrt[m]{x} \rightarrow 0$. ■

Eine weitere Möglichkeit, die Vollständigkeit von \mathbb{R} zu charakterisieren ist das

Intervallschachtelungsprinzip: Eine Folge von Intervallen $[a_n, b_n]$ mit

- $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ für alle $n \in \mathbb{N}$,
- $b_n - a_n \rightarrow 0$ für $n \rightarrow \infty$

nennt man eine *Intervallschachtelung*. Für jede Intervallschachtelung $\{[a_n, b_n] : n \in \mathbb{N}\}$ gibt es genau eine Zahl a , die allen Intervallen angehört; es gilt $a = \lim a_n = \lim b_n$.

Beweis. Es gilt (z. B.) $a_n \leq b_1$ für alle n , also existiert

$$a := \sup\{a_n : n \in \mathbb{N}\}.$$

Da alle $a_n \leq b_m$ sind für beliebige m , ist $a \leq b_m$ für jedes m . a ist also untere Schranke von $\{b_n : n \in \mathbb{N}\}$, also

$$a \leq \inf\{b_n : n \in \mathbb{N}\} =: b.$$

Für jedes n gilt

$$b - a \leq b_n - a_n \rightarrow 0,$$

also $b = a$ und

$$|a - a_n| = a - a_n \leq b_n - a_n \rightarrow 0$$

$$|a - b_n| = b_n - a \leq b_n - a_n \rightarrow 0$$

für $n \rightarrow \infty$. ■

Für jede reelle Zahl > 0 läßt sich nun leicht eine (unendliche)

Dezimalbruchentwicklung angeben: Für $x > 0$ sei

$$\begin{aligned} x_0 &:= \text{größte ganze Zahl} < x, \\ x_1 &:= \max\{k \in \{0, 1, 2, \dots, 9\} : x_0 + k10^{-1} < x\}, \\ &\vdots \\ x_{n+1} &:= \max\{k \in \{0, 1, 2, \dots, 9\} : x_0 + x_110^{-1} + \dots + x_n10^{-n} + k10^{-n-1} < x\}. \end{aligned}$$

Auf diese Weise wird jedem $x > 0$ ein $x_0 \in \mathbb{N}_0$ und eine Folge (x_n) von Ziffern zugeordnet, die *nicht* ab einer Stelle identisch Null ist (eine nichtabbrechende Folge, warum?). Hiermit wird eine Intervallschachtelung definiert:

$$a_n := x_0 + x_110^{-1} + \dots + x_n10^{-n}, \quad b_n := a_n + 10^{-n}.$$

Diese enthält genau die vorgegebene Zahl x , es gilt $a_n \rightarrow x$. Man nennt deshalb

$$x = x_0, x_1x_2x_3 \dots$$

die Dezimalbruchentwicklung von x . Offensichtlich haben zwei verschiedene Zahlen zwei verschiedene Entwicklungen, und zwei verschiedene Entwicklungen stehen für verschiedene reelle Zahlen. (Man beachte z. B. $2 \sim 1,999\dots$, $1,5 \sim 1,4999\dots$)

Damit läßt sich nun leicht zeigen:

Satz 2.6 (Nichtabzählbarkeit von \mathbb{R}) Die Menge der reellen Zahlen ist nicht abzählbar (man sagt auch: überabzählbar).

Beweis. Nehmen wir an, dass \mathbb{R} abzählbar ist; dann ist sicher auch die Menge $(0, 1]$ als Teilmenge von \mathbb{R} abzählbar, das heißt, es läßt sich eine Folge angeben, in der alle Elemente von $(0, 1]$ vorkommen:

$$\begin{aligned} x_1 &= 0, x_{11} x_{12} x_{13} \dots \\ x_2 &= 0, x_{21} x_{22} x_{23} \dots \\ x_3 &= 0, x_{31} x_{31} x_{33} \dots \\ &\vdots \end{aligned}$$

Wir haben einen Widerspruch zu dieser Annahme, wenn wir eine Zahl (= unendliche Dezimalentwicklung $0, y_1 y_2 \dots$) aus $(0, 1]$ angeben können, die in dieser Folge nicht vorkommt. Dazu definieren wir

$$y_j := \begin{cases} x_{jj} + 1 & \text{falls } x_{jj} < 9, \\ 5 & \text{falls } x_{jj} = 9 \end{cases}$$

(man könnte sich auch viele andere Möglichkeiten einfallen lassen). Dann kommt $y = 0, y_1 y_2 \dots$ offensichtlich in obiger Folge nicht vor (man beachte, dass es sicher ein nichtabbrechender Dezimalbruch ist). ■

4:25.10.07

Bei Konvergenzuntersuchungen haben wir bisher stets den Grenzwert gekannt oder mitbestimmt. Dies wird aber nicht immer gelingen. Deshalb ist es nützlich, Kriterien anzugeben, mit denen man die Konvergenz nachweisen kann, ohne gleichzeitig den Grenzwert zu bestimmen.

Eine Folge heißt *monoton wachsend* (manchmal auch etwas präziser *nicht-fallend*), wenn $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt, *streng monoton wachsend* (auch *strikt monoton wachsend*), wenn $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$ gilt. Entsprechend *monoton fallend* (*nicht-wachsend*), *streng monoton fallend* (*strikt monoton fallend*); häufig läßt man hierbei das Wort *monoton* weg. Eine Folge heißt *monoton*, *streng/strikt monoton*, falls sie \dots wachsend oder \dots fallend ist.

Satz 2.7 (Monotonieprinzip) : Jede monoton wachsende nach oben beschränkte Folge (x_n) ist konvergent. Entsprechend für monoton fallend und nach unten beschränkt. (Da das Konvergenzverhalten nicht von den ersten Gliedern der Folge abhängt, genügt es natürlich, wenn die Folge ab einer bestimmten Stelle monoton ist.)

Beweis. Nach dem Supremumsaxiom existiert

$$x := \sup\{x_n : n \in \mathbb{N}\}.$$

Es bleibt $x_n \rightarrow x$ zu beweisen. Da x das Supremum ist, gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$x - \varepsilon < x_N \leq x.$$

Wegen der Monotonie gilt das dann auch für alle x_n mit $n \geq N$, also $|x - x_n| < \varepsilon$ für $n \geq N$. ■

Bei der Definition und Untersuchung der Exponentialfunktion spielt das folgende Beispiel eine wesentliche Rolle.

Beispiel 2.8 Die Folge (x_n) mit $x_n = x_n(x) := \left(1 + \frac{x}{n}\right)^n$ ist für $n \in \mathbb{N}$ mit $n > |x|$ streng wachsend, denn es gilt

$$\begin{aligned} \frac{x_{n+1}}{x_n} &= \frac{\left(1 + \frac{x}{n+1}\right)^{n+1}}{\left(1 + \frac{x}{n}\right)^n} = \left(1 + \frac{x}{n}\right) \left(\frac{1 + \frac{x}{n+1}}{1 + \frac{x}{n}}\right)^{n+1} \\ &= \left(1 + \frac{x}{n}\right) \left(1 + \frac{\frac{x}{n+1} - \frac{x}{n}}{1 + \frac{x}{n}}\right)^{n+1} = \left(1 + \frac{x}{n}\right) \left(1 - \frac{\frac{x}{n(n+1)}}{1 + \frac{x}{n}}\right)^{n+1} \\ &= \left(1 + \frac{x}{n}\right) \left(1 + \frac{-x}{(n+1)(n+x)}\right)^{n+1} = \left(1 + \frac{x}{n}\right) (1+h)^{n+1} \\ &\quad \left(\text{Bernoulli, mit } h = \frac{-x}{(n+1)(n+x)} = \frac{1 + \frac{x}{n+1}}{1 + \frac{x}{n}} - 1 > -1 \text{ für } n > |x|\right) \\ &> \left(1 + \frac{x}{n}\right) \left(1 - (n+1) \frac{x}{(n+1)(n+x)}\right) = \left(1 + \frac{x}{n}\right) \left(1 - \frac{x}{n+x}\right) \\ &= \left(1 + \frac{x}{n}\right) \left(1 + \frac{x}{n}\right)^{-1} = 1. \end{aligned}$$

Außerdem ist (x_n) beschränkt, denn wegen $\left(1 + \frac{x}{n}\right) \left(1 - \frac{x}{n}\right) = 1 - \frac{x^2}{n^2} < 1$ gilt für $n > |x|$

$$x_n = \left(1 + \frac{x}{n}\right)^n \leq \left(1 - \frac{x}{n}\right)^{-n} = \left[\left(1 - \frac{x}{n}\right)^n\right]^{-1},$$

und letztere Folge ist für große n fallend. Damit ist klar, dass die Folge (x_n) für jedes x konvergiert. Über den Grenzwert wissen wir allerdings bisher fast nichts. Wir werden später sehen, dass durch $\lim x_n = \lim \left(1 + \frac{x}{n}\right)^n$ die *Exponentialfunktion* definiert werden kann. □

Schließlich beweisen wir noch das auch in allgemeineren Situationen als Definition der Vollständigkeit benutzte *Cauchysches Konvergenzkriterium*:

Eine Folge (x_n) heißt eine *Cauchyfolge*, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit

$$|x_n - x_m| < \varepsilon \text{ für } n, m \geq N,$$

oder, mit Quantoren: $\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N |x_n - x_m| < \varepsilon$, oder, etwas unpräzise: wenn $|x_n - x_m|$ klein wird für große n, m .

Satz 2.9 (Cauchysches Konvergenzkriterium) *Jede Cauchyfolge (x_n) in \mathbb{R} ist konvergent, d. h. $\exists x \in \mathbb{R}$ mit $x_n \rightarrow x$ (die Umkehrung ist offensichtlich).*

Der Beweis wird bequemer, wenn wir vorher noch zwei andere Aussagen beweisen:

Satz 2.10 *Jede Cauchyfolge ist beschränkt.*

Beweis. Ist N so, dass $|x_n - x_m| < 1$ für $n, m \geq N$ gilt, so ist insbesondere $|x_n| \leq |x_N| + 1$ und somit

$$|x_n| \leq \max \left\{ |x_1|, |x_2|, \dots, |x_N|, |x_N| + 1 \right\}.$$

■

Die folgende Aussage kann auch als Charakterisierung der Vollständigkeit benutzt werden:

Satz 2.11 (Bolzano–Weierstraß) *Jede beschränkte Folge in \mathbb{R} enthält eine konvergente Teilfolge.*

Beweis. Da (x_n) beschränkt ist, so gibt es ein Intervall $[a, b]$, in der die Folge enthalten ist. Halbiert man das Intervall, so sind in mindestens einem der Teilintervalle I_1 unendlich viele Glieder der Folge enthalten. Sei $n_1 := \min\{n \in \mathbb{N} : x_n \in I_1\}$.

I_1 wird wieder halbiert und ein Teilintervall I_2 gewählt, in dem unendlich viele Glieder liegen, $n_2 := \min\{n \in \mathbb{N} : n > n_1, x_n \in I_2\}$ usw. Die Intervallfolge (I_k) liefert eine Intervallschachtelung, die also genau einen Punkt enthält. Gegen diesen konvergiert die Folge (x_{n_k}) . ■

Beweis des Cauchyschen Konvergenzkriteriums. Jede Cauchyfolge ist beschränkt und enthält deshalb eine konvergente Teilfolge (x_{n_k}) , $x_{n_k} \rightarrow x$. Da die Glieder (einer Cauchyfolge) für große n nahe zusammen liegen, liegen sie dann für große n nahe bei x . Dies ist genau die Konvergenz der (gesamten) Folge gegen x .

Genauer: Zu jedem $\varepsilon > 0$ gibt es:

- ein $N = N(\varepsilon)$ mit $|x_n - x_m| < \varepsilon/2$ für $n, m \geq N$,
- ein $K = K(\varepsilon) = \tilde{K}(\varepsilon, N(\varepsilon))$ mit $n_k \geq N$ und $|x_{n_k} - x| < \varepsilon/2$ für $k \geq K$.

Also gilt, mit einem $k \geq K$,

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ für } n \geq N.$$

■

3 Exponential- und Logarithmusfunktionen

Wir kommen zurück zur Folge $\left(1 + \frac{x}{n}\right)^n$ mit $x \in \mathbb{R}$. Wie kann man sich die Entstehung dieser Folge vorstellen?

Kontinuierliche Verzinsung: Nehmen wir an, dass das Kapital K mit einem Zinssatz von $p\%$ angelegt ist. Es sei $\alpha := \frac{p}{100}$. Werden die Zinsen

- jährlich gutgeschrieben, so beträgt das Kapital nach 1 Jahr $K(1 + \alpha) = K(1 + \alpha)^1$,
- halbjährlich gutgeschrieben, so beträgt das Kapital nach 1 Jahr $K\left(1 + \frac{\alpha}{2}\right)^2$,
- täglich gutgeschrieben, so beträgt das Kapital nach 1 Jahr $K\left(1 + \frac{\alpha}{365}\right)^{365}$,
- sekundlich ... $K\left(1 + \frac{\alpha}{365 \cdot 24 \cdot 3600}\right)^{365 \cdot 24 \cdot 3600}$.

Obwohl es schwer ist, diese Zahlen explizit zu berechnen ist offensichtlich, dass das Kapital immer mehr anwächst, je schneller die Zinsen gutgeschrieben werden (das haben wir aber auch bereits im vorhergehenden Kapitel bewiesen). Im Grenzfall erhält man

$$\text{nach 1 Jahr : } \lim_{n \rightarrow \infty} K\left(1 + \frac{\alpha}{n}\right)^n, \quad \text{nach } t \text{ Jahren } \lim_{n \rightarrow \infty} K\left(1 + \frac{\alpha t}{n}\right)^n$$

für den Fall der „kontinuierlichen Verzinsung“. (Ein realistischeres Modell wäre z. B. das Wachstum einer Bakterienkultur.)

Ein physikalisches (dort auch realistisches) Modell liefert der

Radioaktive Zerfall: Von N Teilchen zerfallen in einer Zeiteinheit $\beta \cdot N$ Teilchen ($0 \leq \beta < 1$). Wie stark ist der Zerfall während eines (großen) Zeitintervalls t ? Zunächst würde man erwarten, dass

- über das ganze Intervall gerechnet $N(1 - \beta t)$ Teilchen übrig sind (das ist sicher nicht gut; wenn $t > \frac{1}{\beta}$ ist, käme ein negativer Bestand heraus), deshalb ist es sinnvoll,
- wenn man das Intervall in n Teile zu zerlegen und nach jeweils einem Zeitintervall $\frac{t}{n}$ auswertet (das scheint sinnvoll, weil nach der Zeit $\frac{t}{n}, 2\frac{t}{n}, \dots$ natürlich nicht mehr N Teilchen für den Zerfall zur Verfügung stehen), dann erhält man nach der Zeit t einen Bestand von $N\left(1 - \frac{\beta t}{n}\right)^n$,
- und im Grenzfall ($n \rightarrow \infty$), die einzig wirklich sinnvolle Betrachtungsweise, da sich die Anzahl „kontinuierlich“ verkleinert

$$\lim_{n \rightarrow \infty} N\left(1 - \frac{\beta t}{n}\right)^n.$$

Es ist also offenbar sinnvoll, die Funktion

$$\exp : \mathbb{R} \rightarrow \mathbb{R}, \quad \exp(x) := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \quad (x \in \mathbb{R}),$$

die *Exponentialfunktion*, zu betrachten. (Zum Namen „Exponentialfunktion“ später mehr).

Etwas zum Begriff der Funktion: Eine *Funktion* $f : A \rightarrow B$ ist eine Abbildung (Zuordnung), die jedem $x \in A$ genau ein Element $f(x) \in B$ zuordnet. Veranschaulicht werden reelle Funktionen häufig mit Hilfe ihres *Graphen* G_f . Das ist Punktmenge

$$F_f := \{(x, f(x)) : x \in A\} \subset \mathbb{R}^2,$$

die in vielen Fällen leicht graphisch dargestellt werden kann.

Die obige Funktion \exp ist nach den früheren Überlegungen tatsächlich für alle $x \in \mathbb{R}$ wohldefiniert.

Satz 3.1 *Für die oben definierte Funktion \exp gilt*

- a) $\exp(0) = 1$,
- b) $\exp(x + y) = \exp(x) \exp(y)$ für alle $x, y \in \mathbb{R}$ (die Funktionalgleichung für \exp); diese Identität macht deutlich, dass durch $\exp(x)$ der Zuwachs im Zeitintervall x beschrieben wird.
- c) $\exp(x) > 0$ für alle $x \in \mathbb{R}$,
- d) $\exp(-x) = \frac{1}{\exp(x)}$.

Beweis. a) Offensichtlich, da $\left(1 + \frac{0}{n}\right)^n = 1$ für alle n .

b) Nach Definition von \exp gilt (wieder mit der Formel $a^n - b^n = (a - b) \sum_{k=1}^n a^{n-k} b^{k-1}$)

$$\begin{aligned} & \exp(x + y) - \exp x \exp y \\ &= \lim_{n \rightarrow \infty} \left[\left(1 + \frac{x + y}{n}\right)^n - \left(1 + \frac{x}{n}\right)^n \left(1 + \frac{y}{n}\right)^n \right] = \lim_{n \rightarrow \infty} (c_n^n - (a_n b_n)^n) \\ & \quad \text{mit } a_n := 1 + \frac{x}{n}, \quad b_n := 1 + \frac{y}{n}, \quad c_n := 1 + \frac{x + y}{n}, \\ &= \lim_{n \rightarrow \infty} (c_n - a_n b_n) \sum_{k=1}^n c_n^{n-k} (a_n b_n)^{k-1}. \end{aligned}$$

Wegen

$$\begin{aligned} |a_n b_n| &\leq \left(1 + \frac{|x|}{n}\right) \left(1 + \frac{|y|}{n}\right), \\ |c_n| &\leq 1 + \frac{|x|}{n} + \frac{|y|}{n} \leq 1 + \frac{|x|}{n} + \frac{|y|}{n} + \frac{|x||y|}{n^2} = \left(1 + \frac{|x|}{n}\right) \left(1 + \frac{|y|}{n}\right), \\ c_n - a_n b_n &= -\frac{xy}{n^2}, \end{aligned}$$

folgt

$$\begin{aligned} |c_n^{n-k} (a_n b_n)^{k-1}| &\leq \left(1 + \frac{|x|}{n}\right)^{n-1} \left(1 + \frac{|y|}{n}\right)^{n-1} \\ &\leq \left(1 + \frac{|x|}{n}\right)^n \left(1 + \frac{|y|}{n}\right)^n \leq \exp(|x|) \exp(|y|) \end{aligned}$$

also für obigen [...]–Ausdruck

$$[\dots] \leq \frac{|x||y|}{n^2} n \cdot \exp(|x|) \cdot \exp(|y|) \longrightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Damit folgt die Behauptung.

c) Da $\left(1 + \frac{x}{n}\right)^n$ für große n positiv ist, ist $\exp(x) > 0$, da $\left(\left(1 + \frac{x}{n}\right)^n\right)$ wachsend ist. d) $\exp(-x) = \exp(x)^{-1}$. ■

Man definiert die *Eulersche Zahl* durch

$$e := \exp(1) = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Damit gilt für jede natürliche Zahl n

$$e^n = (\exp(1))^n = \exp(1 + \dots + 1) = \exp(n)$$

und somit auch $e^n = \exp(n)$ für alle $n \in \mathbb{Z}$ (Beweis!). Wegen

$$\left[\exp\left(\frac{m}{n}\right)\right]^n = \exp\left(\frac{m}{n}n\right) = \exp(m) = e^m$$

ist

$$\exp\left(\frac{m}{n}\right) = \sqrt[n]{e^m}, \quad \text{dafür schreibt man } e^{\frac{m}{n}}.$$

Es gilt also für jede rationale Zahl q

$$\exp(q) = e^q.$$

Deshalb schreibt man für alle $x \in \mathbb{R}$: $e^x := \exp(x)$. (Mit der Stetigkeit der Funktion \exp wird dann folgen: Ist $x \in \mathbb{R}$ und (x_n) eine Folge aus \mathbb{Q} mit $x_n \rightarrow x$, so gilt $e^x = \lim_{n \rightarrow \infty} e^{x_n}$)
Hiervon rührt der Name *Exponentialfunktion*.

Satz 3.2 (Eigenschaften der Exponentialfunktion) Für die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \exp(x) = e^x$ gilt

- a) $e^0 = 1$.
- b) $e^{x+y} = e^x e^y$ (Exponentialgesetz=Funktionalgleichung der Exponentialfunktion).
- c) $e^x > 1 + x$.
- d) e^x ist streng monoton wachsend: aus $x < y$ folgt $e^x < e^y$.
- e) Für $|x| < 1$ gilt $|e^x - 1| \leq \frac{|x|}{1 - |x|}$, insbesondere $e^x \rightarrow 1 = e^0$ für $x \rightarrow 0$.
- f) e^x wächst für $x \rightarrow \infty$ schneller als jede Potenz (d. h. für jedes n ist $e^x > x^n$ für hinreichend große x , z. B. $x > 4n^2$).
- g) $e^x \rightarrow 0$ für $x \rightarrow -\infty$, schneller als jede negative Potenz $|x|^{-n}$.

Beweis. a), b) sind bekannt

c) Für $x > -1$ gilt (wegen Monotonie der Folge $\left(1 + \frac{x}{n}\right)^n$ für $n > |x|$)

$$1 + x = \left(1 + \frac{x}{1}\right)^1 \leq \left(1 + \frac{x}{n}\right)^n \leq e^x.$$

Ist $x \leq -1$, so ist $e^x > 0 \geq 1 + x$.

d) Wegen $e^z > 1$ für $z > 0$ und $e^x > 0$ für alle x gilt für $x < y$

$$e^y - e^x = e^x(e^{y-x} - 1) > 0.$$

e) Für $|x| < 1$ fällt $\left(1 - \frac{x}{n}\right)^{-n}$ monoton gegen $e^x = (e^{-x})^{-1}$. Also ist $e^x \leq \left(1 - \frac{x}{1}\right)^{-1} = \frac{1}{1-x}$ und somit

$$e^x - 1 \leq \frac{1}{1-x} - 1 = \frac{x}{1-x}.$$

Für $x \geq 0$ ist das die Behauptung. Für $-1 < x < 0$ gilt nach Teil c, da $1 - |x| < 1$ ist,

$$1 - e^x \leq -x = |x| \leq \frac{|x|}{1 - |x|}.$$

f) Für $x > 4n^2$ ist $\sqrt{x} > 2n$, $\frac{x}{2n} > \sqrt{x}$, also

$$e^x > \left(1 + \frac{x}{2n}\right)^{2n} > (1 + \sqrt{x})^{2n} > (\sqrt{x})^{2n} = x^n.$$

g) Folgt sofort aus Teil f und $e^{-x} = (e^x)^{-1}$. ■

Eine Funktion f nennen wir (umgangssprachlich ausgedrückt) *stetig*, wenn

- sich der Wert $f(x)$ mit x kontinuierlich ändert, oder
- man den Graphen in einem Zug zeichnen kann.

Beides ist aber mathematisch kaum brauchbar. Also genauer:

Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *stetig im Punkt* x_0 , wenn für jede Folge (x_n) mit $x_n \in I$ und $x_n \rightarrow x_0$ gilt $f(x_n) \rightarrow f(x_0)$ (*Folgendefinition, Folgenstetigkeit*). f heißt *stetig*, wenn es in jedem Punkt von I stetig ist.

Satz 3.3 Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ ist stetig.

Beweis. Sei $x_0 \in \mathbb{R}$, (x_n) eine Folge mit $x_n \rightarrow x_0$ bzw. $x_n - x_0 \rightarrow 0$. Dann gilt $|x_n - x_0| < 1$ für hinreichend großes n , und somit nach Satz 3.2.e

$$\begin{aligned} |\exp x_0 - \exp x_n| &= \exp x_0 \left| 1 - \exp(x_n - x_0) \right| \\ &\leq \exp x_0 \frac{|x_n - x_0|}{1 - |x_n - x_0|} \rightarrow 0 \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Also ist \exp stetig in x_0 . ■

Der folgende Satz liefert eine äquivalente Definition ($\varepsilon - \delta$ -Definition) der Stetigkeit.

Satz 3.4 Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann stetig in x_0 , wenn gilt ($\varepsilon - \delta$ -Stetigkeit)

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad \text{mit } |x - x_0| < \delta \quad \text{gilt } |f(x) - f(x_0)| < \varepsilon.$$

Beweis. \Rightarrow : Sei f (folgen-)stetig in x_0 im Sinne der obigen Definition. Annahme, dies gilt nicht, d. h. (schrittweise Umkehr der Quantoren)

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x \in I \quad \text{mit } |x - x_0| < \delta \quad \text{und } |f(x) - f(x_0)| \geq \varepsilon.$$

Insbesondere existiert für dieses $\varepsilon > 0$ und jedes $n \in \mathbb{N}$ ($\delta_n = \frac{1}{n}$) ein $x_n \in I$ mit $|x_n - x_0| < \delta_n = \frac{1}{n}$ und $|f(x_n) - f(x_0)| \geq \varepsilon$, also gilt $x_n \rightarrow x_0$ und $f(x_n) \not\rightarrow f(x_0)$, ein Widerspruch zur Stetigkeit.

\Leftarrow : Sei f $\varepsilon - \delta$ -stetig in x_0 . Sei (x_n) eine Folge aus I mit $x_n \rightarrow x_0$. Zu $\varepsilon > 0$ sei $\delta = \delta(\varepsilon)$ gemäß Voraussetzung gewählt. Dann gibt es ein $N = N(\varepsilon)$ mit $|x_n - x_0| < \delta$ und somit $|f(x_n) - f(x_0)| < \varepsilon$ für $n \geq N$, also $f(x_n) \rightarrow f(x_0)$. ■

Als **Übung** beweise man mit Hilfe der obigen Resultate (ohne die Folgenstetigkeit zu verwenden) die $\varepsilon - \delta$ -Stetigkeit von \exp .

Eine Funktion heißt *injektiv*, wenn gilt

$$\text{aus } x_1 \neq x_2 \text{ folgt } f(x_1) \neq f(x_2)$$

bzw.

$$\text{aus } f(x_1) = f(x_2) \text{ folgt } x_1 = x_2.$$

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$ (oder $x \mapsto x^3$) ist injektiv¹; dagegen ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ nicht injektiv. Anschaulich bedeutet Injektivität von f , dass jede Parallele zur x -Achse den Graphen von f höchstens 1 mal schneidet.

Eine Funktion (Abbildung) $f : M \rightarrow N$ heißt *surjektiv*, wenn N die Bildmenge von f ist, $N = f(M) := \{f(x) : x \in M\}$, oder: Wenn zu jedem $y \in N$ (mindestens) ein $x \in M$ existiert mit $f(x) = y$.

Wieder ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$ surjektiv, während $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ nicht surjektiv ist. (Surjektivität kann ggf. immer erzwungen werden, indem N durch die Bildmenge ersetzt wird; im Fall von $f(x) = x^2$ z. B. $N = [0, \infty)$.)

Eine Abbildung $f : M \rightarrow N$ heißt *bijektiv*, wenn sie injektiv *und* surjektiv ist. (Eine injektive Abbildung kann stets bijektiv gemacht werden, indem N durch die Bildmenge ersetzt wird.) Bei einer bijektiven Abbildung existiert also zu jedem $y \in N$ *genau ein* $x \in M$ mit $f(x) = y$. Diese Abbildung

$$f^{-1} : N \rightarrow M, \quad y \mapsto x \text{ mit } f(x) = y$$

heißt die *Umkehrabbildung* (auch *Inverse*) von f (*Achtung*: f^{-1} darf nicht mit $\frac{1}{f}$ verwechselt werden.)

Satz 3.5 (Umkehrfunktion) Sei $I \subset \mathbb{R}$ ein Intervall. Ist $f : I \rightarrow \mathbb{R}$ streng monoton (wachsend oder fallend), so ist f injektiv. (Es existiert also die Umkehrfunktion $f^{-1} : f(I) \rightarrow I$ und diese ist ebenfalls streng monoton wachsend bzw. fallend.)

Beweis. Die Injektivität ist offensichtlich. Die Monotonie der Umkehrfunktion kann leicht indirekt gezeigt werden (Übung). ■

¹aus $0 = x^3 - y^3 = (x - y)(x^2 + xy + y^2) = \frac{1}{2}(x - y)(x^2 + y^2 + (x + y)^2)$ folgt $x - y = 0$ oder $x = y$.

Satz 3.6 Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow (0, \infty)$ ist bijektiv. Die Umkehrabbildung

$$\ln : (0, \infty) \rightarrow \mathbb{R}, \quad y \mapsto x \quad \text{mit } y = \exp(x) \quad (\ln = \text{logarithmus naturalis})$$

heißt Logarithmusfunktion (auch \log oder \log_e). Die Logarithmusfunktion hat die folgenden Eigenschaften:

- a) \ln ist streng wachsend.
- b) $\ln 1 = 0$, $\ln e = 1$.
- c) $\ln(a \cdot b) = \ln a + \ln b$ (Funktionalgleichung des Logarithmus)
und entsprechend $\ln \frac{a}{b} = \ln a - \ln b$ und $\ln x^n = n \ln x$.
- d) $\ln(1+x) \leq x$ für alle $x > -1$ bzw. $\ln x \leq x - 1$ für $x > 0$.
- e) \ln wächst für $x \rightarrow \infty$ langsamer als jede positive Potenz $x \mapsto x^\alpha$. Genauer: Für $x > \exp(4n^2)$ ist $\ln x < \sqrt[n]{x} = x^{1/n}$.

Beweis. Wir wissen, dass \exp injektiv ist (streng monoton). Es bleibt die Surjektivität bezüglich $(0, \infty)$ zu beweisen, d. h. zu jedem $y > 0$ ist ein x mit $\exp x = y$ zu finden (Mit Hilfe des Zwischenwertsatzes wird dies viel einfacher gehen). Dazu sei

$$A := \{x \in \mathbb{R} : e^x < y\}.$$

A ist nicht leer, denn wegen $e^{-n} \rightarrow 0$ für $n \rightarrow \infty$ gibt es ein n mit $e^{-n} < y$, d. h. $-n \in A$. A ist auch nach oben beschränkt, weil für jedes $x \in A$ gilt $1+x \leq e^x < y$, also $x < y-1$.

Also existiert $a := \sup A$. Wir zeigen $e^a = y$: Zu jedem $n \in \mathbb{N}$ gibt es ein x_n aus A mit

$$a - \frac{1}{n} < x_n \leq a,$$

also, da $a + \frac{1}{n}$ nicht in A liegt,

$$e^{a-\frac{1}{n}} < e^{x_n} < y \leq e^{a+\frac{1}{n}}.$$

Für $n \rightarrow \infty$ folgt (beachte $\left|e^{\pm\frac{1}{n}} - 1\right| \leq \frac{\frac{1}{n}}{1 - \frac{1}{n}} \rightarrow 0$ für $n \rightarrow \infty$)

$$e^{a-\frac{1}{n}} = e^a e^{-\frac{1}{n}} \rightarrow e^a, \quad e^{a+\frac{1}{n}} = e^a e^{\frac{1}{n}} \rightarrow e^a,$$

also (nach dem Dreifolgensatz) $e^a \leq y \leq e^a$, d. h. $e^a = y$.

Es bleiben die weiteren Eigenschaften zu zeigen:

- a) Die strenge Monotonie von \ln folgt leicht (indirekter Beweis) aus der strengen Monotonie von \exp .
- b) Folgt aus $\exp(0) = 1$, $\exp(1) = e$.
- c) Sei $a = \exp \alpha$, $b = \exp \beta$. Dann folgt

$$\begin{aligned}\ln(a \cdot b) &= \ln(\exp(\alpha + \beta)) = \alpha + \beta = \ln a + \ln b, \\ \ln \frac{a}{b} &= \ln(\exp(\alpha - \beta)) = \alpha - \beta = \ln a - \ln b.\end{aligned}$$

- d) Aus $\exp(x) \geq 1 + x$ für alle x folgt $\ln(1 + x) < \ln(\exp(x)) = x$ für alle $x > 0$, also

$$\ln x \leq x - 1 \quad \text{für alle } x > 0.$$

■

6: 1. 11 2007**Satz 3.7** $\ln : (0, \infty) \rightarrow \mathbb{R}$ ist stetig

Beweis. Ist (x_n) aus $(0, \infty)$ mit $x_n \rightarrow x \in (0, \infty)$, so gilt $\frac{|x_n - x|}{x} \rightarrow 0$ und somit (beim ersten \leq -Zeichen beachte man die Monotonie von \ln)

$$|\ln x_n - \ln x| = \left| \ln \frac{x_n}{x} \right| = \left| \ln \frac{x + x_n - x}{x} \right| \leq \ln \left(1 + \left| \frac{x_n - x}{x} \right| \right) \leq \frac{|x_n - x|}{x} \rightarrow 0$$

für $n \rightarrow \infty$.

■

Für rationales $y = \frac{p}{q}$ ($p \in \mathbb{N}$, $q \in \mathbb{Z}$) und $x > 0$ ist

$$x^y = \sqrt[q]{x^p} = \sqrt[q]{(\exp \ln x)^p} = \exp\left(\frac{1}{q} p \ln x\right) = \exp(y \ln x)$$

(man beachte, dass das auch für $p < 0$ richtig ist). Deshalb definieren wir für alle $y \in \mathbb{R}$ und $x > 0$

$$x^y := \exp(y \ln x) \quad \left(\text{also} = \lim_{n \rightarrow \infty} x^{y_n} \text{ mit } y_n \in \mathbb{Q}, y_n \rightarrow y \right)$$

Es gilt dann, wie für ganzzahlige und rationale Exponente,

$$(xy)^r = x^r y^r, \quad (x^r)^s = x^{rs}, \quad x^{r+s} = x^r x^s,$$

insbesondere $x^0 = 1$, $x^{-r} = 1/x^r$.

Die Funktion \log_{10} (*Zehnerlogarithmus, Logarithmus zur Basis 10* oder *Dekadischer Logarithmus*) ist die Umkehrfunktion von

$$\mathbb{R} \rightarrow (0, \infty), \quad x \mapsto 10^x = \exp(x \ln 10).$$

Es gilt offenbar

$$\log_{10}(x) = \frac{\ln x}{\ln 10}.$$

Der Zehnerlogarithmus spielt z. B. in der Chemie (pH-Wert) eine Rolle.

4 Weitere elementare Funktionen

4.1 Zusammengesetzte Funktionen

Einige wichtige *Konstruktionsprinzipien* für kompliziertere Funktionen aus einfachen Funktionen $f, g : D \rightarrow \mathbb{R}$ sind

$$\begin{aligned} f + g : D &\rightarrow \mathbb{R}, & x &\mapsto f(x) + g(x) \\ f \cdot g : D &\rightarrow \mathbb{R}, & x &\mapsto f(x) \cdot g(x) \\ \lambda f : D &\rightarrow \mathbb{R}, & x &\mapsto \lambda f(x) \quad (\lambda \in \mathbb{R}) \\ -f : D &\rightarrow \mathbb{R}, & x &\mapsto -f(x) \\ \frac{f}{g} : D \setminus \{x : g(x) = 0\} &\rightarrow \mathbb{R}, & x &\mapsto \frac{f(x)}{g(x)}. \end{aligned}$$

Eine weitere Möglichkeit, aus zwei Funktionen eine neue Funktion zu konstruieren, ist die *Hintereinanderausführung (Komposition)*: Seien $f : D \rightarrow M$ und $g : M \rightarrow N$. Dann ist die *Hintereinanderausführung* $g \circ f$ („ g nach f “) definiert durch

$$g \circ f : D \rightarrow N, \quad x \mapsto g(f(x)) \quad \text{für } x \in D.$$

Wichtige Spezialfälle sind:

- die konstante Funktion auf D , $c : D \rightarrow \mathbb{R}$, $x \mapsto c$ für alle $x \in D$, insbesondere die
- *Nullfunktion* auf D ist $0 : D \rightarrow \mathbb{R}$, $x \mapsto 0$ für alle $x \in D$.
Dementsprechend bedeutet $f \neq 0$ lediglich, dass es mindestens ein $x \in D$ gibt mit $f(x) \neq 0$.
- die *identische Funktion* $\text{id} : D \rightarrow \mathbb{R}$, $x \mapsto x$ für alle $x \in D$.

Mit diesen Bezeichnungen gilt z.B. $\ln \circ \exp$ die *Identität* auf \mathbb{R} , $\exp \circ \ln$ die Identität auf $(0, \infty)$

$$\ln \circ \exp = \text{id}_{\mathbb{R}}, \quad \text{quad} \quad \exp \circ \ln = \text{id}_{(0, \infty)}$$

4.2 Polynome

Besonders einfache Funktionen sind die *Monome*

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^n \quad (n \in \mathbb{N}_0).$$

Durch Zusammensetzen mit Hilfe obiger Konstruktionsprinzipien erhält man *Polynome*

$$x \mapsto a_0 + a_1x + \dots + a_nx^n =: p(x).$$

Ist $a_n \neq 0$, so heißt n der *Grad* von p .

Satz 4.1 Sind zwei Polynome $p(x) = a_0 + a_1x + \dots + a_nx^n$ und $q(x) = b_0 + b_1x + \dots + b_mx^m$ als Funktionen gleich (d. h. $p(x) = q(x)$ für alle $x \in \mathbb{R}$), so gilt

$$n = m \quad \text{und} \quad a_k = b_k \quad \text{für} \quad k = 0, \dots, n$$

(die Umkehrung ist offensichtlich).

Beweis. Offensichtlich ist $r(x) := p(x) - q(x) = c_0 + c_1x + \dots + c_lx^l$ mit $c_j = a_j - b_j$ ein Polynom vom Grad $l \leq \max\{n, m\}$ und $r(x) = 0$ für alle $x \in \mathbb{R}$.

Aus $r(0) = 0$ folgt $c_0 = 0$, also $a_0 = b_0$. Also ist

$$0 = \frac{r(x)}{x} = c_1 + c_2x + \dots + c_lx^{l-1} \quad \text{für} \quad x \neq 0,$$

und somit für $x_k \rightarrow 0$

$$0 = \lim_{k \rightarrow \infty} \frac{r(x_k)}{x_k} = c_1.$$

So erhält man schließlich $c_0 = c_1 = \dots = c_l = 0$, also $a_j = b_j$ für alle j . ■

Summen und Produkte von Polynomen sind offenbar Polynome. Der Grad der Summe ist \leq dem Maximum der beiden Grade, der Grad des Produkts ist die Summe beider Grade.

4.3 Rationale Funktionen

Eine *rationale Funktion* ist eine Funktion der Form $x \mapsto \frac{p(x)}{q(x)}$ mit Polynom p und q ; der Definitionsbereich ist die Menge $D = \{x \in \mathbb{R} : q(x) \neq 0\}$.

Hat in einer rationalen Funktion der Zähler einen Grad \geq dem Grad des Nenners, so kann eine *Polynomdivision mit Rest* durchgeführt werden. Diese beschreibt man am besten durch ein Beispiel:

$$\begin{array}{r} (3x^4 + 2x^3 \quad -x - 7) : (x^2 + 3) = (3x^2 + 2x - 9) \\ \underline{-(3x^4 \quad + 9x^2)} \\ 2x^3 - 9x^2 - x - 7 \\ \underline{-(2x^3 \quad + 6x)} \\ -9x^2 - 7x - 7 \\ \underline{-(-9x^2 \quad - 27)} \\ -7x + 20 \end{array}$$

Das *Ergebnis* ist

$$3x^2 + 2x - 9 \text{ Rest } 5x + 20.$$

Der Rest ist offenbar ein Polynom vom Grad $<$ als der Grad des Nenners. Für die entsprechende rationale Funktion gilt also

$$\frac{3x^4 + 2x^3 - x - 7}{x^2 + 3} = 3x^2 + 2x - 9 + \frac{5x + 20}{x^2 + 3}.$$

Ist x_0 eine Nullstelle des Polynoms $p(x)$, so liefert die Division $p(x) : (x - x_0)$ den Rest 0: Ist nämlich $p(x) : (x - x_0) = q(x)$ Rest a , so gilt

$$p(x) = q(x)(x - x_0) + a,$$

und wegen $p(x_0) = 0$ folgt daraus $a = 0$. Haben Zähler und Nenner die gleiche Nullstelle x_0 , so kann also durch $(x - x_0)$ „gekürzt“ werden; d. h. man dividiert Zähler und Nenner durch $(x - x_0)$. Dies kann so lange durchgeführt werden, bis Zähler und Nenner keine gemeinsame Nullstelle mehr haben. Es bleiben dann nur Definitionslücken, in denen der rationalen Funktion kein vernünftiger Wert zugeordnet werden kann (Zähler $\neq 0$, Nenner = 0).

Satz 4.2 *Sind f und g stetig, so sind auch die nach obigen Konstruktionsprinzipien zusammengesetzten Funktionen stetig. Insbesondere sind alle Polynome und rationalen Funktionen (in ihrem Definitionsbereich) stetig.*

Der *Beweis* ist mit Hilfe der Folgendefinition einfach.

4.4 Trigonometrischen Funktionen

Diese wollen wir hier geometrisch einführen.

Der *Cosinus* eines Winkels α ist im rechtwinkligen Dreieck mit einem Winkel α definiert als der Quotient

$$\frac{\text{Länge der Ankathete}=\text{Kathete an der der Winkel } \alpha \text{ liegt}}{\text{Länge der Hypotenuse}}$$

Entsprechend ist der *Sinus* des Winkels α definiert als

$$\frac{\text{Länge der Gegenkathete}=\text{Kathete gegenüber dem Winkel } \alpha}{\text{Länge der Hypotenuse}}$$

Aufgrund der Ähnlichkeit aller rechtwinkligen Dreiecke mit einem Winkel α hängen diese Definitionen tatsächlich nur vom Winkel ab.

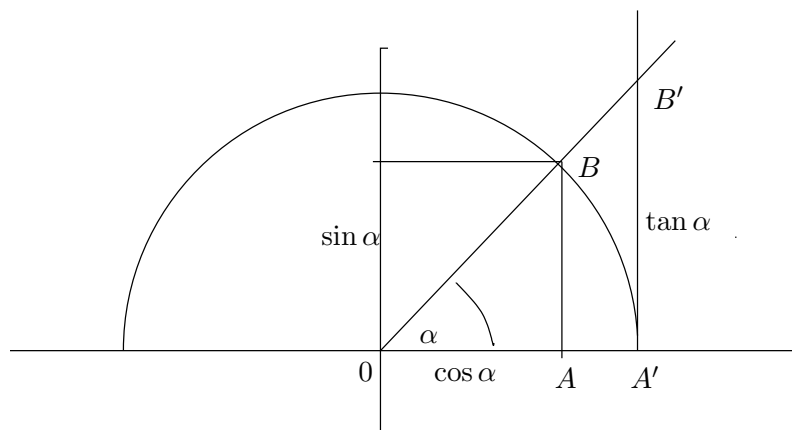


Abbildung 1: Sinus, Cosinus, Tangens

Insbesondere kann das Dreieck im Einheitskreis so gewählt werden, dass der Winkel α im Nullpunkt liegt, eine Kathete auf der x -Achse und die Hypotenuse ein Radius ist.

Der Winkel kann einerseits in Grad gemessen werden (voller Winkel = 360° , gestreckter Winkel = 180° , rechter Winkel = 90° .) Mathematisch stellt es sich aber als günstiger heraus, den Winkel im *Bogenmaß* zu messen, der Länge des dem Winkel entsprechenden Bogens auf dem Einheitskreis (voller Winkel = 2π [π wird dadurch definiert, zum Zahlenwert später], gestreckter Winkel = π , rechter Winkel = $\pi/2$).

Der Strahl von 0 aus, der mit der positiven Achse den Winkel α bildet (in mathematisch positivem Sinne, d. h. entgegen dem Uhrzeigersinn, gemessen) schneidet also der Einheitskreis im Punkt $(\cos \alpha, \sin \alpha)$. Die Betrachtung am Einheitskreis hat auch den Vorteil, dass die Funktionen \cos und \sin für alle $\alpha \in \mathbb{R}$ und nicht nur für $\alpha \leq 90^\circ \sim \pi/2$ definiert werden können. Für negative α wird der Winkel im mathematisch negativen Sinn (im Uhrzeigersinn) aufgetragen.

Offensichtlich sind \cos und \sin 2π -periodisch, d. h. es gilt

$$\cos(\alpha + 2\pi) = \cos \alpha \quad \text{und} \quad \sin(\alpha + 2\pi) = \sin \alpha.$$

Die Definition am Einheitskreis liefert auch sofort die wichtige Identität

Satz 4.3 *Es gilt*

$$\cos^2 \alpha + \sin^2 \alpha = 1 \quad \text{für alle } \alpha \in \mathbb{R}$$

Ebenso leicht erhält man einige wichtige Funktionswerte

$$\begin{aligned} \cos 0 = 1, \quad \cos \frac{\pi}{2} = \cos \frac{3\pi}{2} = 0, \quad \cos \pi = -1, \\ \sin 0 = \sin \pi = 0, \quad \sin \frac{\pi}{2} = 1, \quad \sin \frac{3\pi}{2} = -1. \end{aligned}$$

Mit etwas elementarer Geometrie (und insbesondere dem Satz von Pythagoras) erhält man auch:

$$\begin{aligned} \sin \frac{\pi}{4} &= \sin \frac{3\pi}{4} = \frac{\sqrt{2}}{2}, & \sin \frac{5\pi}{4} &= \sin \frac{7\pi}{4} = -\frac{\sqrt{2}}{2}, \\ \cos \frac{\pi}{4} &= \cos \frac{7\pi}{4} = \frac{\sqrt{2}}{2}, & \cos \frac{3\pi}{4} &= \cos \frac{5\pi}{4} = -\frac{\sqrt{2}}{2}, \\ \sin \frac{\pi}{6} &= \sin \frac{5\pi}{6} = \frac{1}{2}, & \sin \frac{7\pi}{6} &= \sin \frac{11\pi}{6} = -\frac{1}{2}, \\ \cos \frac{\pi}{6} &= \cos \frac{11\pi}{6} = \frac{\sqrt{3}}{2}, & \cos \frac{5\pi}{6} &= \cos \frac{7\pi}{6} = -\frac{\sqrt{3}}{2}. \end{aligned}$$

Als **Übung** bestimme man $\sin \frac{\pi}{3}$, $\sin \frac{2\pi}{3}$, $\sin \frac{4\pi}{3}$, $\sin \frac{5\pi}{3}$ und die entsprechenden Cosinuswerte?

Weiter sieht man der Definition am Einheitskreis sofort an, (und damit wollen wir uns zunächst begnügen)

Satz 4.4 *sin und cos sind stetig.*

Beweis. Die Änderung der sin- bzw. cos-Werte ist immer kleiner als die Änderung von α

$$|\sin \alpha - \sin \beta| \leq |\alpha - \beta|, \quad |\cos \alpha - \cos \beta| \leq |\alpha - \beta|$$

(in der $\varepsilon - \delta$ -Definition der Stetigkeit kann δ stets gleich ε gewählt werden). ■

Ebenfalls klar sind folgende Eigenschaften:

$$\begin{aligned} \sin &\text{ ist streng wachsend in } \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \text{ fallend in } \left[\frac{\pi}{2}, \frac{3\pi}{2}\right] \\ \cos &\text{ ist streng fallend in } [0, \pi], \text{ wachsend in } [\pi, 2\pi]. \end{aligned}$$

Schließlich lassen sich auch die folgenden Symmetrieeigenschaften leicht ablesen:

$$\begin{aligned} \cos(-\varphi) &= \cos \varphi, & \sin(-\varphi) &= -\sin \varphi, \\ \cos(\varphi + \pi) &= -\cos \varphi, & \sin(\varphi + \pi) &= -\sin \varphi \\ \cos\left(\frac{\pi}{2} - \varphi\right) &= \sin \varphi, & \sin\left(\frac{\pi}{2} - \varphi\right) &= \cos \varphi. \end{aligned}$$

Von besonderer Bedeutung sind schließlich die *Additionstheoreme*, die sich am Einheitskreis (zumindest für Winkel im Bereich $\left[0, \frac{\pi}{2}\right]$) sehr leicht elementargeometrisch beweisen lassen:

Satz 4.5 (Additionstheoreme für cos und sin)

$$\begin{aligned} \cos(\varphi + \psi) &= \cos \varphi \cos \psi - \sin \varphi \sin \psi \\ \sin(\varphi + \psi) &= \sin \varphi \cos \psi + \cos \varphi \sin \psi \end{aligned}$$

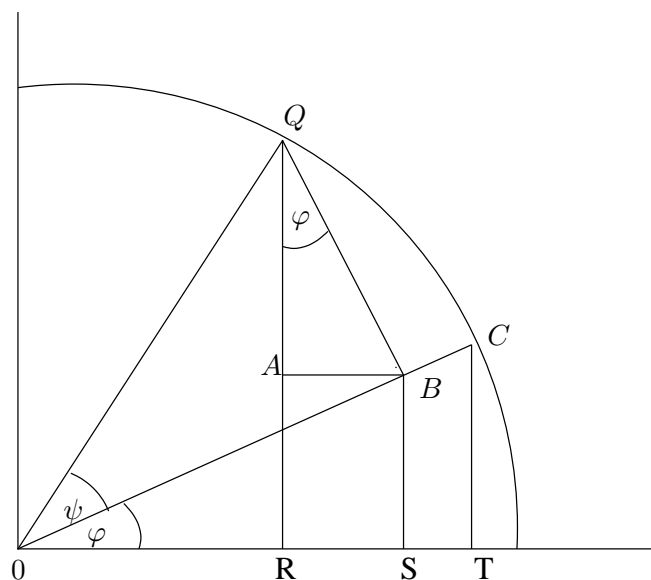


Abbildung 2: Additionstheoreme für Sinus und Cosinus

Beweis. Aus der Skizze ergibt sich zunächst

$$\sin \varphi = \overline{CT}, \quad \cos \varphi = \overline{OT},$$

$$\sin \psi = \overline{QB}, \quad \cos \psi = \overline{OB}$$

$$\sin(\varphi + \psi) = \overline{QR}, \quad \cos(\varphi + \psi) = \overline{OR}.$$

Weiter ist

$$\overline{QA} = \overline{QB} \cos \varphi = \sin \psi \cos \varphi,$$

$$\overline{AR} = \overline{BS} = \overline{OB} \sin \varphi = \cos \psi \sin \varphi,$$

$$\overline{OS} = \overline{OB} \cos \varphi = \cos \psi \cos \varphi,$$

$$\overline{RS} = \overline{AB} = \overline{QB} \sin \varphi = \sin \psi \sin \varphi,$$

also

$$\cos(\varphi + \psi) = \overline{OR} = \overline{OS} - \overline{RS} = \cos \psi \cos \varphi - \sin \psi \sin \varphi,$$

$$\sin(\varphi + \psi) = \overline{QR} = \overline{QA} + \overline{AR} = \sin \psi \cos \varphi + \cos \psi \sin \varphi.$$

■

Satz 4.6 (Halbwinkelformeln)

$$1 - \cos \varphi = 2 \sin^2 \frac{\varphi}{2}, \quad 1 + \cos \varphi = 2 \cos^2 \frac{\varphi}{2}$$

Beweis.

$$\cos \varphi = \cos \left(\frac{\varphi}{2} + \frac{\varphi}{2} \right) = \cos^2 \frac{\varphi}{2} - \sin^2 \frac{\varphi}{2} = 1 - 2 \sin^2 \frac{\varphi}{2}$$

und entsprechend

$$\cos \varphi = \cos \left(\frac{\varphi}{2} + \frac{\varphi}{2} \right) = 2 \cos^2 \frac{\varphi}{2} - 1,$$

woraus die beiden Gleichungen folgen. ■

Die folgenden beiden Ungleichungen werden bei der Differentiation von \sin und \cos wichtig sein:

Satz 4.7 Für alle φ gilt

a) $|\sin \varphi| \leq |\varphi|,$

b) $1 - \frac{\varphi^2}{2} \leq \cos \varphi \leq 1.$

(Im ersten Fall gilt $<$ für alle $\varphi \neq 0$. Im zweiten Fall gilt an beiden Stellen $<$ für $0 < \varphi < 2\pi$.)

Beweis. a) Der Punkt auf dem Einheitskreis, der durch den Winkel φ beschrieben wird, habe die Koordinaten (x, y) . Dann gilt (da die Sekante immer kürzer ist als der entsprechende Kreisbogen)

$$|\sin \varphi| = |y| \leq \sqrt{(1-x)^2 + y^2} \leq |\varphi| \quad \text{bzw. } < |\varphi| \quad \text{für } \varphi \neq 0.$$

b) Mit der ersten Halbwinkelformel und Teil a) folgt

$$0 \leq 1 - \cos \varphi = 2 \sin^2 \frac{\varphi}{2} \leq 2 \left(\frac{\varphi}{2} \right)^2 = \frac{1}{2} \varphi^2,$$

(„ $<$ “ statt „ \leq “ an beiden Stellen für $0 < \varphi < 2\pi$) und somit die Behauptung. ■

Offensichtlich sind \sin und \cos als Funktionen auf \mathbb{R} *nicht injektiv*, d. h. es gibt keine globale Umkehrfunktion. Wir haben aber gesehen, dass die Funktion Sinus im Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ streng wachsend ist, dass also das Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$ injektiv in das Intervall $[-1, 1]$ abgebildet wird (aus dem Zwischenwertsatz wird folgen, dass diese Abbildung auch surjektiv ist; von der geometrischen Anschauung her ist dies offensichtlich).

Es existiert also die Umkehrfunktion

$$\text{Arcus Sinus, } \arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

die ebenfalls bijektiv ist (arc Sius, d. h. dem Sinus wird der entsprechende Bogen zugeordnet). Den Graphen erhält man, wie immer, durch Spiegelung an der Winkelhalbierenden ($x = y$).

Die Funktion Cosinus ist im Intervall $[0, \pi]$ streng fallend und bildet $[0, \pi]$ bijektiv auf $[-1, 1]$ ab. Die entsprechende Umkehrfunktion ist

$$\text{Arcus Cosinus, } \arccos : [-1, 1] \rightarrow [0, \pi].$$

Die Funktion *Tangens* ist im rechtwinkligen Dreieck definiert durch Gegenkathete/Ankathete, also

$$\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi : k \in \mathbb{Z} \right\} \rightarrow \mathbb{R}, \quad \tan \varphi := \frac{\sin \varphi}{\cos \varphi}.$$

Aus dem Verhalten von \sin und \cos leitet man leicht ab, dass \tan in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ streng wachsend ist und $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ bijektiv auf \mathbb{R} abbildet. Die Umkehrfunktion ist

$$\text{Arcus Tangens, } \arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right);$$

auch sie ist streng wachsend und bijektiv.

Cotangens ist im rechtwinkligen Dreieck definiert durch Ankathete/Gegenkathete, also

$$\cot : \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\} \rightarrow \mathbb{R}, \quad \cot \varphi = \frac{\cos \varphi}{\sin \varphi} = \frac{1}{\tan \varphi}$$

($\cot \varphi = 0$ dort wo $\tan \varphi$ nicht definiert ist). \cot ist streng fallend in $(0, \pi)$ und bildet $(0, \pi)$ bijektiv auf \mathbb{R} ab. Die Umkehrfunktion ist

$$\text{Arcus Cotangens, } \operatorname{arccot} : \mathbb{R} \rightarrow (0, \pi),$$

ebenfalls streng fallend und bijektiv.

Natürlich könnte man ebenso gut \sin und \tan auf jedem der Intervalle $\left(k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}\right)$ ($k \in \mathbb{Z}$) und \cos und \cot auf jedem der Intervalle $(k\pi, k\pi + \pi)$ betrachten und würde entsprechende Umkehrfunktionen erhalten. Die oben beschriebenen Funktionen werden deshalb oft als *Hauptzweig* bezeichnet (ihr Wert als der Hauptwert).

Der Leser versuche auf der Grundlage der obigen Aussagen Skizzen der Graphen dieser Funktionen anzufertigen und seine Ergebnisse mit geeigneten Lehrbüchern zu vergleichen.

5 Komplexe Zahlen und elementare komplexe Funktionen

5.1 Komplexe Zahlen

Jede positive reelle Zahl hat eine positive und eine negative Quadratwurzel (die Null hat die einzige Quadratwurzel Null). Die negativen Zahlen haben keine Quadratwurzel in \mathbb{R} . Wir erweitern deshalb den Körper \mathbb{R} , indem wir ein Element i (die *imaginäre Einheit*) mit der Eigenschaft $i^2 = -1$ hinzunehmen und die Menge der *komplexen Zahlen*

$$\mathbb{C} := \{x + iy : x, y \in \mathbb{R}\}$$

betrachten. Für die Elemente in \mathbb{C} definieren wir

$$\text{Addition : } (x + iy) + (u + iv) = (x + u) + i(y + v),$$

$$\text{Multiplikation : } (x + iy)(u + iv) = (xu + yv) + i(yu - xv).$$

Es ist leicht zu sehen, dass mit dem

$$\text{Einselement } 1 = 1 + i0$$

und dem

$$\text{Nullelement } 0 = 0 + i0$$

alle Körperaxiome (wie bei \mathbb{R} , ohne die Anordnung) erfüllt sind (nur das Distributivgesetz ist etwas mühsam nachzurechnen). Dabei ist das *negative Element* zu $x + iy \in \mathbb{C}$

$$-(x + iy) = -x - iy = (-x) + i(-y),$$

das *inverse Element* zu $x + iy (\neq 0, \text{ d. h. } (x, y) \neq (0, 0) \text{ bzw. } x^2 + y^2 > 0)$

$$(x + iy)^{-1} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2}i.$$

Identifiziert man die reellen Zahlen x mit den komplexen Zahlen $x + i0$, so kann man \mathbb{R} als *Unterkörper* von \mathbb{C} auffassen (die Teilmenge der Elemente $x + i0$ ist bezüglich der Operationen $+$ und \times abgeschlossen). Entsprechend wird \mathbb{C} als *Körpererweiterung* von \mathbb{R} bezeichnet.

Betrachtet man die komplexen Zahlen $x + iy$ als Punkt (x, y) der Ebene \mathbb{R}^2 , so hat man also auf \mathbb{R}^2 eine Körperstruktur. Diese Betrachtungsweise ist im folgenden häufig nützlich (*komplexe Ebene*).

Zu einer komplexen Zahl $x + iy$ definiert man:

– *Realteil* von $z = (x + iy)$, $\text{Re } z := x$,

- Imaginärteil von $z = (x + iy)$, $\text{Im } z := y$ (nicht iy),
- die zu $z = x + iy$ konjugiert komplexe Zahl $\bar{z} := x - iy$,
- den Betrag von $z = x + iy$, $|z| := \sqrt{x^2 + y^2}$ (= Länge des Vektors (x, y) in \mathbb{R}^2).

Offenbar gilt:

$$\begin{aligned} \text{Re } z &= \frac{1}{2}(z + \bar{z}), & \text{Im } z &= \frac{1}{2i}(z - \bar{z}) \\ |z| &= \sqrt{z\bar{z}} \\ \overline{z+w} &= \bar{z} + \bar{w} & \text{und } \overline{z\bar{w}} &= \bar{z} \bar{w} \\ |zw| &= |z| |w| & \text{und } |z+w| &\leq |z| + |w|. \end{aligned}$$

Nur bei $|zw| = |z| |w|$ ist etwas zu zeigen. Der Leser möge die einfache Rechnung durchführen (weiter unten folgt das noch bequemer).

Eine Folge (z_n) aus \mathbb{C} heißt

- konvergent gegen z , wenn $|z_n - z| \rightarrow 0$ gilt für $n \rightarrow \infty$,
- eine Cauchyfolge, wenn zu jedem $\varepsilon > 0$ ein N existiert mit $|z_n - z_m| < \varepsilon$ für $n, m \geq N$.

Satz 5.1 a) Es gilt $z_n = x_n + iy_n \rightarrow z = x + iy$ genau dann, wenn $x_n \rightarrow x$ und $y_n \rightarrow y$ gilt.

b) $(z_n) = (x_n + iy_n)$ ist genau dann eine Cauchyfolge, wenn (x_n) und (y_n) Cauchyfolgen sind.

c) \mathbb{C} ist vollständig, d. h. jede Cauchyfolge in \mathbb{C} ist konvergent.

Der Beweis kann dem Leser überlassen werden.

Definiert man für $z = x + iy \in \mathbb{C} \setminus \{0\}$ das Argument von $z = \arg(z)$ gleich dem Winkel zwischen der positiven x -Achse und dem Vektor (x, y) (im mathematisch positiven Sinn, entgegen dem Uhrzeigersinn, gerechnet), also

$$\cos \alpha = \frac{x}{|z|}, \quad \sin \alpha = \frac{y}{|z|} \quad \text{bzw.} \quad \alpha = \arccos \frac{x}{|z|} = \arcsin \frac{y}{|z|},$$

so kann z geschrieben werden in der Form

$$z = |z|(\cos \alpha + i \sin \alpha)$$

(natürlich ist $\arg(z)$ nur bis auf additive Vielfache von 2π eindeutig bestimmt, für $z = 0$ ist $\arg(z)$ nicht definiert, bzw. beliebig).

Diese Schreibweise bringt nichts für die Addition, macht aber die Multiplikation komplexer Zahlen besonders durchsichtig.

Satz 5.2 Für

$$z = |z|(\cos \varphi + i \sin \varphi), \quad w = |w|(\cos \psi + i \sin \psi)$$

gilt

$$zw = |z| |w| \left(\cos(\varphi + \psi) + i \sin(\varphi + \psi) \right)$$

d. h. bei der Multiplikation werden die Beträge multipliziert, die Argumente addiert.

Der Beweis folgt aus

$$zw = |z| |w| \left\{ (\cos \varphi \cos \psi - \sin \varphi \sin \psi) + i(\cos \varphi \sin \psi + \sin \varphi \cos \psi) \right\}$$

mit Hilfe der Additionstheoreme.

Eine überraschende Konsequenz hieraus ist, dass man sofort für jede komplexe Zahl $z \neq 0$ genau k k -te Wurzeln angeben kann:

Satz 5.3 Ist

$$z = r(\cos \alpha + i \sin \alpha),$$

so ist

$$w = r^{1/k} \left(\cos \frac{\alpha}{k} + i \sin \frac{\alpha}{k} \right)$$

eine k -te Wurzel aus z . k k -te Wurzeln erhält man durch

$$w_j = r^{1/k} \left(\cos \frac{\alpha + j2\pi}{k} + i \sin \frac{\alpha + j2\pi}{k} \right), \quad j = 0, \dots, k-1$$

(für $j = k$ erhält man wieder w_0). Dies sind alle k -ten Wurzeln.

Beweis. Offensichtlich ist $w^k = z$. Dass dies alle k -ten Wurzeln sind, erhält man so: Ist

$$w = s(\cos \varphi + i \sin \varphi)$$

eine k -te Wurzel von $z = r(\cos \alpha + i \sin \alpha)$, so gilt

$$s^k (\cos k\varphi + i \sin k\varphi) = w^k = z = r(\cos \alpha + i \sin \alpha).$$

Das ist nur möglich, wenn $k\varphi = \alpha \pmod{2\pi}$ gilt. ■

8:8.11.07

5.2 Komplexe Exponential- und Logarithmusfunktionen

Wir können nun die *komplexe Exponentialfunktion* (eine Fortsetzung der reellen Exponentialfunktion) erklären, die wir dann weiterhin mit \exp bezeichnen.

Satz 5.4 *Die Funktion*

$$F : \mathbb{C} \rightarrow \mathbb{C}, \quad F(z) = F(x + iy) := e^x(\cos y + i \sin y) \quad \text{für } z = x + iy \in \mathbb{C}$$

ist eine Fortsetzung von \exp auf ganz \mathbb{C} . F erfüllt die Funktionalgleichung der Exponentialfunktion

$$F(z)F(w) = F(z + w).$$

Beweis. Für $z = x + iy$, $w = u + iv$ gilt

$$\begin{aligned} F(z)F(w) &= e^x e^u (\cos y + i \sin y)(\cos v + i \sin v) \\ &= e^{x+u} (\cos(y+v) + i \sin(y+v)) = F(z+w). \end{aligned}$$

Für $x \in \mathbb{R}$ gilt $F(x) = \exp x$, d. h. F ist tatsächlich eine Fortsetzung der reellen Exponentialfunktion, und erfüllt auch für komplexe Argumente die Funktionalgleichung. ■

Der Spezialfall $x = 0$, also $z = iy$ mit $y \in \mathbb{R}$,

$$e^{iy} = \cos y + i \sin y \quad \text{bzw.} \quad \cos y = \frac{1}{2}(e^{iy} + e^{-iy}) \quad \text{und} \quad \sin y = \frac{1}{2i}(e^{iy} - e^{-iy})$$

wird als *Eulersche Formel*(n) bezeichnet. Insbesondere gilt also

$$|e^{iy}| = 1 \quad \text{für } y \in \mathbb{R},$$

wobei die Zahl e^{iy} auf dem Einheitskreis erreicht wird, indem man vom Punkt 1 aus in mathematisch positiver Richtung um die Bogenlänge y voranschreitet. Es gilt auch

$$\begin{aligned} |e^z| &= e^{\operatorname{Re} z}, \\ e^z &= e^w \Leftrightarrow z = w + 2\pi ik, \\ \overline{e^z} &= \overline{e^{x+iy}} = e^x(\cos y - i \sin y) = e^x(\cos(-y) + i \sin(-y)) \\ &= e^{x-iy} = e^{\bar{z}}. \end{aligned}$$

Eine besonders überraschende Identität ist

$$e^{i\pi} = -1$$

(man beachte: e und π sind transzendent und i rein imaginär!).

Wie für die reelle Exponentialfunktion gilt auch für die komplexe Exponentialfunktion $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$.

Da e^x für $x \in \mathbb{R}$ alle positiven Werte genau ein mal durchläuft, und $\cos y + i \sin y$ für $0 \leq y < 2\pi$ alle Punkte des Einheitskreises genau ein mal trifft, ist offenbar

$$\exp : \{z = x + iy : x \in \mathbb{R}, 0 \leq y < 2\pi\} \rightarrow \mathbb{C} \setminus \{0\}$$

bijektiv. Das gleiche gilt für jeden zur reellen Achse parallelen Streifen der Breite 2π .

Satz 5.5 $\exp : \{z = x + iy : x \in \mathbb{R}, -\pi < y \leq \pi\} \rightarrow \mathbb{C} \setminus \{0\}$ ist bijektiv, und die Umkehrfunktion \ln , der komplexe Logarithmus,

$$\ln : \mathbb{C} \setminus \{0\} \rightarrow \{z = x + iy : x \in \mathbb{R}, -\pi < y \leq \pi\}$$

ist explizit gegeben durch

$$\ln(w) = \ln(u + iv) = \ln|w| + i \arg(w) \quad \text{mit} \quad -\pi < \arg(w) \leq \pi.$$

Da auch ein anderer Streifen, z. B.

$$\left\{ z = x + iy : x \in \mathbb{R}, \quad c < y \leq c + \pi \right\}$$

oder

$$\left\{ z = x + iy : x \in \mathbb{R}, \quad c \leq y < c + \pi \right\}$$

gewählt werden kann, auf dem die Exponentialfunktion invertierbar ist (und je nach Situation ist dies auch sinnvoll), nennt man die oben definierte Funktion den *Hauptzweig des Logarithmus*.

Mit Hilfe des Logarithmus können z. B. auch Wurzelfunktionen geschlossen dargestellt werden:

$$\sqrt{z} = \exp\left(\frac{1}{2} \ln z\right), \quad \text{oder} \quad z^w = \exp(w \ln z) \quad \text{für} \quad z \neq 0.$$

Je nach Wahl des Zweiges von \ln erhält man verschiedene Zweige der jeweiligen Funktion.

6 Unendliche Reihen

6.1 Konvergente und divergente Reihen

Das *Paradoxon* von Achilles und der Schildkröte: Die Schildkröte legt in der Minute 10m zurück, Achilles 100m; die Schildkröte hat 100m Vorsprung.

- Um den Punkt zu erreichen, an dem die Schildkröte zur Zeit 0 war, braucht Achilles 1 Minute. Während dieser Zeit hat die Schildkröte 10m zurückgelegt, hat also 10m Vorsprung.
- Um den Punkt zu erreichen, wo die Schildkröte jetzt ist, braucht Achilles $1/10$ Minute, die Schildkröte hat dann noch einen Meter Vorsprung.
- usw.

Es scheint, dass Achilles die Schildkröte nicht erreicht. Was natürlich unsinnig ist, denn nach 2 Minuten hat Achilles bereits 200 Meter zurückgelegt, die Schildkröte nur 20, so dass Achilles bereits 80m Vorsprung hat. Genauer: Nach $10/9$ Minuten hat Achilles $111\frac{1}{9}$ Meter zurückgelegt, die Schildkröte $11\frac{1}{9}$, d. h. Achilles hat die Schildkröte eingeholt. Die $10/9$ Minuten sind gerade die Summe der oben betrachteten unendlich vielen Zeitintervalle:

$$1 + \frac{1}{10} + \frac{1}{100} + \frac{1}{1000} \dots = 1,111\dots$$

Sei (x_n) eine reelle oder komplexe Folge (entsprechendes werden wir mit Folgen in \mathbb{R}^n , \mathbb{C}^n oder noch allgemeiner machen). Man möchte der *unendlichen Summe* (=Reihe)

$$\sum_{n=1}^{\infty} x_n := x_1 + x_2 + x_3 + \dots$$

einen Wert zuordnen. Das ist bisher nicht definiert. Wir können eigentlich nur zwei Zahlen, oder induktiv endlich viele, addieren. Hier muß also etwas neues definiert werden:

Die m -te *Partialsumme* der Reihe $\sum_{n=1}^{\infty} x_n$ ist

$$s_m := \sum_{n=1}^m x_n.$$

Diese Folge (s_m) der *Partialsummen* kann man nun konvergent oder divergent sein (für die Folge $x_n \equiv 0$ ist auch $s_m \equiv 0$ also ist (s_m) konvergent; für die Folge $x_n \equiv 1$ ist $s_m = m$, also ist (s_m) divergent).

Man sagt, die Reihe $\sum_{n=1}^{\infty} x_n$ *konvergiert/divergiert*, wenn die Folge (s_m) der Partialsummen konvergiert/divergiert. Ist die Reihe konvergent, so heißt

$$\sum_{n=1}^{\infty} x_n := \lim_{m \rightarrow \infty} s_m$$

der Wert oder die *Summe* der Reihe $\sum_{n=1}^{\infty} x_n$. Natürlich betrachtet man entsprechend auch Reihen

$$\sum_{n=k}^{\infty} x_n \text{ mit beliebigen } k \in \mathbb{Z}$$

(mit $k = -\infty$ muß man vorsichtiger sein!).

Für die Untersuchung weiterer Reihen ist es wichtig, einige konvergente und divergente Reihen zu kennen:

Beispiel 6.1 Die *geometrische Reihe*

$$\sum_{n=0}^{\infty} q^n \text{ konvergiert genau dann, wenn } |q| < 1 \text{ ist.}$$

Für $|q| \geq 1$ ist sie also divergent. Für die Partialsummen gilt (Beweis z. B. durch Induktion)

$$s_m = \sum_{n=0}^m q^n = \frac{1 - q^{m+1}}{1 - q} \text{ für } q \neq 1, \quad s_m = m + 1 \text{ für } q = 1.$$

Für $|q| < 1$ gilt $q^{m+1} \rightarrow 0$ für $m \rightarrow \infty$, also

$$s_m \rightarrow \frac{1}{1 - q} =: \sum_{n=0}^{\infty} q^n.$$

In allen anderen Fällen ist (s_m) offenbar divergent (Ü). □

Beispiel 6.2 Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ ist konvergent und hat die Summe 1. Dies erkennt man mit Hilfe eines kleinen Tricks:

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

und somit

$$\begin{aligned} s_m &= \sum_{n=1}^m \frac{1}{n(n+1)} = \sum_{n=1}^m \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= \left(1 - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \dots + \left(\frac{1}{m} - \frac{1}{m+1} \right) \\ &\quad \text{(ein vollständiger Beweis nutzt Induktion)} \\ &= 1 - \frac{1}{m+1} \rightarrow 1 \text{ für } m \rightarrow \infty. \end{aligned}$$

□

Beispiel 6.3 Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ist konvergent mit einer Summe ≤ 2 (später: $= \pi^2/6$). Die Folge der Partialsummen ist offenbar monoton wachsend (Beweis!) und sie ist nach oben beschränkt, wie man durch Vergleich mit der Reihe $\sum \frac{1}{n(n+1)}$ erkennt:

$$\begin{aligned} s_m &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{m^2} \\ &\leq 1 + \left(\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{(m-1)m} \right) = 1 + \left(1 - \frac{1}{m} \right) < 2, \end{aligned}$$

also $\lim s_m \leq 2$ (wie sieht man < 2 ? (Ü)). □

Beispiel 6.4 Die *harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent, da die Folge der Partialsummen unbeschränkt ist (der Beweis ist ein Spezialfall des *Verdichtungskriteriums*; später):

$$\begin{aligned} s_{2^m} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \dots \\ &\quad + \left(\frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \dots + \frac{1}{2^m} \right) \\ &> 1 + \frac{1}{2} + 2 \frac{1}{4} + 4 \frac{1}{8} + \dots + 2^{m-1} \frac{1}{2^m} = 1 + \frac{m}{2}, \end{aligned}$$

die Glieder der Folge (s_m) werden also beliebig groß. □

Einige wichtige Eigenschaften konvergenter Folgen:

Satz 6.5 a) Aus $\sum_{n=0}^{\infty} x_n = x$, $\sum_{n=0}^{\infty} y_n = y$ folgt $\sum_{n=0}^{\infty} (ax_n + by_n) = ax + by$ für alle $a, b \in \mathbb{R}$;
insbesondere ist die Reihe $\sum_{n=0}^{\infty} (ax_n + by_n)$ konvergent.

b) Ist die Reihe $\sum x_n$ konvergent, so ist (x_n) eine Nullfolge (notwendige Bedingung für die Konvergenz einer Reihe); ist (x_n) nicht Nullfolge, so ist die Reihe sicher divergent).

c) Die Abänderung endlich vieler Glieder in einer Reihe ändert nichts am Konvergenzverhalten (natürlich wird sich im Allgemeinen der Wert der Reihe ändern).

Beweis. a) Offenbar ist, wenn (s_m) , (t_m) und (u_m) die Folgen der Partialsummen von $\sum x_n$, $\sum y_n$ und $\sum (ax_n + by_n)$ sind

$$u_m = as_m + bt_m.$$

Daraus folgt die Behauptung.

b) Ist (s_m) konvergent, also eine Cauchyfolge, so muß gelten

$$x_{m+1} = s_{m+1} - s_m \rightarrow 0 \text{ für } m \rightarrow \infty.$$

c) Sind die Folgen (x_n) und (y_n) gleich für $n \geq N$, so gilt für die Folgen (s_m) und (t_m) der Partialsummen

$$t_m = s_m + (t_N - s_N) \text{ für alle } m \geq N,$$

d. h. die Folgen (s_m) und (t_m) unterscheiden sich für $m \geq N$ nur um die Konstante $t_N - s_N$. ■

6.2 Konvergenzkriterien

Wie kann man zum Konvergenz einer Reihe beweisen, ohne gleichzeitig die Summe zu bestimmen (was i. Allg. viel schwerer sein wird)? Das folgende Resultat ergibt sich sofort aus dem Monotonieprinzip für Folgen; es wird deshalb auch gelegentlich als *Monotoniekriterium* bezeichnet.

Satz 6.6 *Sind alle $x_n \geq 0$ (möglicherweise mit endlich vielen Ausnahmen; vgl. Teil c des obigen Satzes), so ist $\sum_{n=0}^{\infty} x_n$ genau dann konvergent, wenn die Folge der Partialsummen (nach oben) beschränkt ist.*

Beweis. Die Folge (s_m) der Partialsummen ist für große m monoton wachsend. Mit dem Monotonieprinzip für Folgen folgt die Behauptung. ■

Eine Reihe $\sum x_n$ heißt *absolut konvergent*, wenn $\sum |x_n|$ konvergiert.

Satz 6.7 *Aus absoluter Konvergenz folgt Konvergenz.*

Beweis. Seien (s_m) und (t_m) die Partialsummenfolgen von $\sum x_n$ bzw. $\sum |x_n|$. Dann gilt $|s_m - s_\ell| \leq |t_m - t_\ell|$. Wenn (t_m) eine Cauchyfolge ist, ist also auch (s_m) eine Cauchyfolge. ■

Absolut konvergente Reihen können beliebig umgeordnet werden, ohne dass sich am Konvergenzverhalten oder an der Summe etwas ändert. Dabei nennen wir (y_k) eine Umordnung von (x_n) , wenn die Folge (y_k) bis auf die Reihenfolge die gleichen Terme enthält wie (x_n) ; dann ist natürlich auch (x_n) eine Umordnung von (y_k) .

Satz 6.8 (Umordnungssatz) *Ist $\sum x_n$ absolut konvergent, so ist auch jede Umordnung absolut konvergent und hat die gleiche Summe.*

Beweis. Sei (y_k) eine Umordnung von (x_n) . Um zu zeigen, dass $\sum y_k$ absolut konvergiert, genügt es zu zeigen, dass für jedes $m \in \mathbb{N}$ gilt

$$\sum_{k=1}^m |y_k| \leq S := \sum_{n=1}^{\infty} |x_n|.$$

Dazu sei $N(m)$ so, dass $\{y_1, \dots, y_m\} \subset \{x_1, \dots, x_{N(m)}\}$ gilt. Dann ist offenbar

$$\sum_{k=1}^m |y_k| \leq \sum_{n=1}^{N(m)} |x_n| \leq S.$$

Also ist $\sum y_k$ absolut konvergent.

Sei nun $s := \sum_{n=1}^{\infty} x_n$ und zu jedem $\varepsilon > 0$ sei $N = N(\varepsilon) \in \mathbb{N}$ so, dass

$$\sum_{n=N+1}^{\infty} |x_n| < \frac{\varepsilon}{3}, \quad \text{also} \quad \left| s - \sum_{n=1}^N x_n \right| < \frac{\varepsilon}{3}$$

gilt, und $M = M(N) \in \mathbb{N}$ so, dass $\{x_1, \dots, x_N\} \subset \{y_1, \dots, y_M\}$ gilt. Dann gilt für $m \geq M$ (da die beiden ersten Terme auf der rechten Seite nur x_n mit $n > N$ enthalten)

$$\begin{aligned} \left| \sum_{k=1}^m y_k - s \right| &\leq \left| \sum_{k=1}^m y_k - \sum_{k=1}^M y_k \right| + \left| \sum_{k=1}^M y_k - \sum_{n=1}^N x_n \right| + \left| \sum_{n=1}^N x_n - s \right| \\ &= \left| \sum_{k=M+1}^m y_k \right| + \left| \sum_{k=1}^M y_k - \sum_{n=1}^N x_n \right| + \left| \sum_{n=1}^N x_n - s \right| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Also gilt $\lim_{m \rightarrow \infty} \sum_{k=1}^m y_k = s = \sum_{n=1}^{\infty} x_n$. ■

9:13.11.07

Bemerkung 6.9 *Ist eine reelle Reihe konvergent aber nicht absolut konvergent, so sind die Teilreihen aus positiven bzw. negativen Termen divergent und zu jedem $c \in \mathbb{R}$ gibt es eine Umordnung der Reihe mit Summe c (Ü).*

Satz 6.10 (Majorantenkriterium) *Ist $\sum_{n=0}^{\infty} y_n$ konvergent und gilt $|x_n| \leq y_n$ für $n \geq N$, so ist $\sum_{n=0}^{\infty} x_n$ absolut konvergent. Die Reihe $\sum_{n=0}^{\infty} y_n$ heißt eine Majorante der Reihe $\sum_{n=0}^{\infty} x_n$; den Satz kann man also so formulieren: Besitzt eine Reihe eine konvergente Majorante, so ist sie (absolut) konvergent.*

Umgekehrt: Ist $\sum y_n$ divergent und gilt $x_n \geq |y_n|$ für $n \geq N$, so ist $\sum x_n$ divergent; d. h.: Besitzt $\sum x_n$ eine divergente Minorante, so ist $\sum x_n$ divergent.

Beweis. Seien (s_m) und (t_m) die Partialsummenfolgen von $\sum |x_n|$ bzw. $\sum y_n$. Da $\sum y_n$ konvergiert, ist (t_m) eine Cauchyfolge, d. h. $|t_m - t_l|$ ist klein für große m und l . Wegen

$$|s_m - s_l| = \left| \sum_{n=l+1}^m x_n \right| \leq \sum_{n=l+1}^m |x_n| \leq \sum_{n=l+1}^m y_n = |t_m - t_l|$$

ist dann auch (s_m) eine Cauchyfolge und somit konvergent. ■

Beispiel 6.11 $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ konvergiert für jedes $z \in \mathbb{C}$ (wir werden später sehen, dass diese (Potenz-)Reihe die Exponentialfunktion in \mathbb{R} oder \mathbb{C} darstellt, *Exponentialreihe*). Dazu sei $N \in \mathbb{N}$ mit $N > 2|z|$, $M := \frac{|2z|^N}{N!}$. Dann gilt für $n \geq N$

$$\left| \frac{z^n}{n!} \right| = \left| \frac{z^N}{N!} \right| \cdot \frac{|z|}{N+1} \cdots \frac{|z|}{n} \leq \left| \frac{z^N}{N!} \right| \left(\frac{1}{2} \right)^{n-N} = \frac{|2z|^N}{N!} \left(\frac{1}{2} \right)^n = M \left(\frac{1}{2} \right)^n,$$

d. h. die Reihe $\sum_{n=0}^{\infty} M \left(\frac{1}{2} \right)^n$ ist eine konvergente Majorante. □

Satz 6.12 (Quotientenkriterium) a) *Existiert ein $N \in \mathbb{N}$ und ein $q \in (0, 1)$ mit und*

$$x_n \neq 0 \quad \text{und} \quad \left| \frac{x_{n+1}}{x_n} \right| \leq q \quad \text{für } n \geq N,$$

so ist die Reihe $\sum_{n=0}^{\infty} x_n$ absolut konvergent.

b) *Existiert ein $N \in \mathbb{N}$ mit $x_n \neq 0$ und $\left| \frac{x_{n+1}}{x_n} \right| \geq 1$ für $n \geq N$, so ist die Reihe divergent.*

Beweis. a) Für $n \geq N$ gilt

$$|x_n| \leq q|x_{n-1}| \leq q^2|x_{n-2}| \leq \dots \leq q^{n-N}|x_N|.$$

also ist die konvergente Reihe $\frac{x_N}{q^N} \sum q^n$ eine Majorante der Reihe $\sum x_n$.

b) Die Folge (x_n) ist offenbar keine Nullfolge. ■

Beispiel 6.13 Auch aus diesem Kriterium folgt leicht die Konvergenz der Exponentialreihe $\sum \frac{z^n}{n!}$. Mit $x_n = \frac{z^n}{n!}$ gilt in diesem Fall

$$\left| \frac{x_{n+1}}{x_n} \right| = \frac{|z^{n+1}|}{(n+1)!} \frac{n!}{|z^n|} = \frac{|z|}{n+1} < \frac{1}{2} \quad \text{für } n \geq 2|z|.$$

□

Satz 6.14 (Wurzelkriterium) a) Existiert ein $N \in \mathbb{N}$ und ein $q \in (0, 1)$ mit

$$\sqrt[n]{|x_n|} \leq q \text{ für } n \geq N,$$

so ist $\sum x_n$ absolut konvergent.

b) Ist $\sqrt[n]{|x_n|} \geq 1$ für $n \geq N$, so ist die Reihe $\sum x_n$ divergent.

Beweis. a) Für $n \geq N$ ist $|x_n| = \sqrt[n]{|x_n|}^n \leq q^n$, d. h. die konvergente geometrische Reihe ist eine Majorante.

b) In diesem Fall ist $|x_n| \geq 1$ für $n \geq N$, also ist die Reihe divergent ■

Bemerkung 6.15 In den beiden vorangehenden Kriterien bleibt ein Bereich, für den die Kriterien keine Aussage liefern. Gilt z. B. nur

$$\left| \frac{x_{n+1}}{x_n} \right| < 1 \text{ für alle } n \text{ bzw. } \sqrt[n]{|x_n|} < 1 \text{ für alle } n,$$

so kann die entsprechende Reihe konvergieren (wie in den obigen Sätzen) oder divergieren: für $x_n = \frac{1}{n}$ gilt

$$\left| \frac{x_{n+1}}{x_n} \right| = \frac{n}{n+1} < 1 \text{ und } \sqrt[n]{|x_n|} = \sqrt[n]{\frac{1}{n}} < 1,$$

aber die Reihe ist divergent. Für $x_n = \frac{1}{n^2}$ gilt ebenfalls $\left| \frac{x_{n+1}}{x_n} \right| < 1$ und $\sqrt[n]{|x_n|} < 1$, während die Reihe konvergiert.

Satz 6.16 (Leibniz–Kriterium für alternierende Reihen) Ist (x_n) eine monoton fallende Nullfolge (nicht notwendig streng monoton), so ist die (alternierende) Reihe $\sum_{n=0}^{\infty} (-1)^n x_n$ konvergent. Für die Summe s und die Partialsummen s_m gilt $|s - s_m| \leq x_{m+1}$ für alle $m \in \mathbb{N}$.

Beweis. Für die Partialsummen gilt offenbar

$$|s_m - s_l| \leq |x_{\min(m,l)+1}| \rightarrow 0 \text{ für } m, l \rightarrow \infty,$$

also ist (s_m) eine Cauchyfolge. Für $l \rightarrow \infty$ folgt die zweite Behauptung. ■

Beispiel 6.17 Die *alternierende harmonische Reihe*

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$$

ist konvergent. Später werden wir sehen, dass die Summe dieser Reihe $= \ln 2$ ($\approx 0,7$) ist.

Die Folgenglieder dürfen aber beliebig langsam gegen 0 konvergieren. So ist z. B. auch die Reihe

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

konvergent. □

Für Reihen $\sum x_n$ mit positiven monoton fallenden Folgen (x_n) ist das *Verdichtungskriterium* sehr nützlich.

Satz 6.18 (Verdichtungskriterium) *Ist (x_n) eine positive monoton fallende Folge, so ist $\sum_{n=0}^{\infty} x_n$ genau dann konvergent, wenn $\sum_{k=0}^{\infty} 2^k x_{2^k}$ konvergiert. (Entsprechendes gilt für $\sum_{n=0}^{\infty} x_n$ und $\sum_{k=0}^{\infty} p^k x_{p^k}$ mit $p \in \mathbb{N}$, $p \geq 2$ (Ü).)*

Beweis. Wegen der Monotonie ist

$$2^{k-1} x_{2^k} < x_{2^{k-1}+1} + x_{2^{k-1}+2} + \dots + x_{2^k} \leq 2^{k-1} x_{2^{k-1}}.$$

Daraus folgt, dass die Partialsummen von $\sum x_n$ genau dann beschränkt sind, wenn die Partialsummen $\sum 2^k x_{2^k}$ bzw. $\frac{1}{2} \sum 2^k x_{2^k} = \sum 2^{k-1} x_{2^k}$ beschränkt sind. Daraus folgt die Behauptung. ■

Beispiel 6.19 $\sum \frac{1}{n^\alpha}$ ist genau dann konvergent, wenn $\alpha > 1$ ist. Die verdichtete Reihe ist nämlich

$$\sum 2^k \frac{1}{(2^k)^\alpha} = \sum (2^{1-\alpha})^k = \sum q^k$$

mit $q = 2^{1-\alpha}$ (< 1 für $\alpha > 1$, ≥ 1 für $\alpha \leq 1$). □

6.3 Potenzreihen

Das Wurzelkriterium liefert eine außerordentlich wichtige Aussage über die Konvergenz von *Potenzreihen* (z_0 heißt der *Entwicklungspunkt* der Reihe)

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n \quad (z \in \mathbb{C}).$$

Als Beispiel haben wir oben bereits die Exponentialreihe kennen gelernt. Es stellt sich die Frage, für welche $x \in \mathbb{R}$ bzw. $z \in \mathbb{C}$ die Reihe konvergiert, d. h. auf welcher Teilmenge von \mathbb{R} bzw. \mathbb{C} durch die Reihe eine Funktion gegeben ist.

Wir benötigen dazu den *Limes-Superior* („größter Limes“) einer reellen Folge (x_n) , dieser ist definiert als

$$\limsup_{n \rightarrow \infty} x_n = \lim_{k \rightarrow \infty} \left\{ \sup_{n \geq k} x_n \right\},$$

wobei die Werte $\pm\infty$ zugelassen sind. Die Folge $(\sup_{n \geq k} x_n)_{k \in \mathbb{N}}$ ist offenbar fallend (nicht-wachsend), wobei der Fall $\sup_{n \geq k} x_n = \infty$ erlaubt ist. Ist $x > \limsup_{n \rightarrow \infty} x_n$, so gibt es ein $N \in \mathbb{N}$ mit $x > x_n$ für $n \geq N$; ist $x < \limsup_{n \rightarrow \infty} x_n$, so ist $x < x_n$ für mindestens unendlich viele n .

Satz 6.20 (Konvergenzradius) Die Potenzreihe $\sum a_n(z - z_0)^n$

a) konvergiert absolut für alle $z \in \mathbb{C}$ mit

$$|z - z_0| < \varrho := \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$$

$$(\varrho = \infty \text{ falls } \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0, \quad \varrho = 0 \text{ falls } \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \infty),$$

b) divergiert für alle $z \in \mathbb{C}$ mit $|z - z_0| > \varrho$.

c) für z mit $|z - z_0| = \varrho$ ist keine allgemeine Aussage möglich.

Die Zahl ϱ heißt aus naheliegenden Gründen *Konvergenzradius* der Potenzreihe. (Es ist überraschend, dass der Konvergenzbereich einer Potenzreihe stets eine Kreisscheibe ist, $\varrho = 0$ und ∞ zugelassen.)

Beweis. a) Ist $|z - z_0| < \varrho$, so gibt es ein ϱ' mit $|z - z_0| < \varrho' < \varrho$, also $\frac{1}{|z - z_0|} > \frac{1}{\varrho'} > \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$. Deshalb gibt es ein N mit $\frac{1}{\varrho'} > \sqrt[n]{|a_n|}$ für alle $n \geq N$ (vgl. obige Anmerkung) und somit

$$\sqrt[n]{|a_n| |z - z_0|^n} = |z - z_0| \sqrt[n]{|a_n|} < \frac{|z - z_0|}{\varrho'} =: q < 1 \quad \text{für } n \geq N.$$

Mit dem Wurzelkriterium folgt die absolute Konvergenz der Potenzreihe.

b) Ist $|z - z_0| > \varrho$, also $\frac{1}{|z - z_0|} < \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$, so gilt $\frac{1}{|z - z_0|} < \sqrt[n]{|a_n|}$ bzw. $|z - z_0| > \frac{1}{\sqrt[n]{|a_n|}}$ für unendlich viele n (vgl. obige Anmerkung), also

$$|a_n(z - z_0)^n| = |a_n||z - z_0|^n > 1 \quad \text{für unendlich viele } n.$$

Also ist die Reihe divergent.

c) Die Reihe $\sum \frac{1}{n} z^n$ hat den Konvergenzradius 1, sie konvergiert für $z = -1$ (vgl. das folgende Kriterium über alternierende Reihen) und divergiert für $z = 1$. – Die Reihe $\sum \frac{1}{n^2} z^n$ hat den Konvergenzradius 1 und konvergiert für alle z mit $|z| = 1$. – Die Reihe $\sum z^n$ hat den Konvergenzradius 1 und divergiert für alle z mit $|z| = 1$. ■

10:15.11.07

7 Vektorräume; vom \mathbb{R}^2 zum \mathbb{C}^n

In der Physik spielen verschiedene Größen eine Rolle:

- z. B. die Zeit, Masse, Energie, Temperatur, die durch eine Zahl (Maßzahl) allein zusammen mit einer physikalischen Dimension (Sekunde, Gramm, kWh, . . .) beschrieben werden, *skalare Größen*,
- z. B. die Kraft, Geschwindigkeit, Beschleunigung, Drehimpuls, die durch eine Richtung (in der Ebene \mathbb{R}^2 oder im Raum \mathbb{R}^3) und einem Betrag (positive Zahl) und natürlich wieder einer physikalischen Dimension beschrieben werden (Newton, m/sec , m/sec^2), *vektorielle* oder *gerichtete Größen*.

Zu der ersten Kategorie ist hier nichts zu sagen. Die zweite beschreibt man bequem durch einen Pfeil, und dieser hat eine Richtung (wie oben verlangt) und eine Länge, die den oben genannten Betrag veranschaulicht. Skalare Größen können einfach addiert werden. Vektorielle Größen werden mit Hilfe eines Parallelogramms addiert: die beiden zu addierenden Größen bilden zwei Seiten, die vom gleichen Eckpunkt ausgehen; die Diagonale stellt dann die Summe dar.

Diese *Vektoraddition* spielt auch in anderem Zusammenhang eine Rolle: Eine *Translation* (zunächst in der Ebene oder im Raum) ist eine Abbildung $t : X \rightarrow X$ bei der (grob gesagt) jeder Punkt in eine feste Richtung und um eine feste Strecke verschoben wird. Eine Translation t wird eindeutig beschrieben durch einen Punkt P und dessen Bildpunkt $P' = t(P)$. Veranschaulichen kann man t durch den *Translationsvektor* $\vec{v} := \overrightarrow{PP'}$. Hintereinanderausführung zweier Translationen $t_{\vec{v}}$ und $t_{\vec{w}}$ um die Vektoren \vec{v} und \vec{w} ist gleich der Translation $t_{\vec{w} \circ \vec{v}}$ um den Vektor $\vec{v} + \vec{w} = \vec{w} + \vec{v}$, wobei letzterer offenbar aus dem zur Addition dienenden Parallelogramm hervorgeht.

Die identische Abbildung ist die Translation um den Nullvektor $\vec{0}$ (seine Richtung hat keinen Sinn, der Betrag ist Null). Die Inverse zur Translation t um den Vektor \vec{v} ist die Translation t^{-1} um den Vektor $-\vec{v}$.

Eine Translation kann mit einer Zahl a *multipliziert* werden, indem die Länge des Vektors entsprechend multipliziert wird; $(-a)\vec{v}$ ist dabei der zu $a\vec{v}$ gehörige „negative Vektor“ $-(a\vec{v})$.

Geometrische Grundkenntnisse werden im folgenden stillschweigend vorausgesetzt und benutzt (so z. B. Parallelenaxiom, Stufenwinkel, Winkelsumme im Dreieck). Wir betrachten hier auch nur die euklidische Geometrie, obwohl in der modernen Physik nicht-euklidische Geometrien durchaus eine wesentliche Rolle spielen.

7.1 Koordinaten und Vektoren

Einen Punkt in der Ebene, und entsprechend im (3-dimensionalen) Raum, kann man wie folgt beschreiben: Man legt einen (im Prinzip beliebigen) Punkt als *Nullpunkt* 0 fest. Durch diesen legt man eine Gerade g_1 auf der man einen Einheitspunkt E_1 im Abstand 1 von 0 wählt. Dann zieht man durch 0 eine weitere auf g_1 senkrecht stehende Gerade g_2 , auf der man wieder einen Einheitspunkt E_2 wählt (üblicherweise so, dass E_2 links liegt, wenn man auf g_1 von 0 in Richtung E_1 geht). Im Raum kommt entsprechend noch eine weitere Achse dazu, wobei E_3 dann oberhalb der Ebene liegt (*Drei-Finger-Regel* der rechten Hand).

Nun kann man jeden Punkt P der Ebene (des Raums) durch ein *Koordinationspaar* (x_1, x_2) (Koordinatentripel (x_1, x_2, x_3)) beschreiben, wobei man x_j erhält, indem man von P aus das Lot auf g_j fällt. Wir schreiben kurz

$$P = x = (x_1, x_2) \quad \text{bzw.} \quad P = x = (x_1, x_2, x_3).$$

Ist nun \vec{x} ein Vektor, so können wir diesen im Nullpunkt ansetzen und erhalten einen Punkt

$$x = (x_1, x_2) \quad \text{bzw.} \quad x = (x_1, x_2, x_3) \quad \text{mit} \quad \vec{x} = 0\vec{x}.$$

Diesen *Ortsvektor* \vec{x} von x beschreiben wir durch die

$$\text{Koordinatenspalte} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Es wird sich später zeigen, dass die Spaltendarstellung für Vektoren günstiger ist. Bezeichnen wir mit

$$\vec{e}_1 = \vec{OE}_1, \quad \vec{e}_2 = \vec{OE}_2, \dots$$

so kann \vec{x} geschrieben werden in der Form

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 \quad \text{bzw.} \quad \vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

Das a -Fache des Vektors \vec{x} ist (wenn wir die Vektoren mit den entsprechenden Translationen identifizieren) gegeben durch

$$a\vec{x} = \begin{pmatrix} a x_1 \\ a x_2 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} a x_1 \\ a x_2 \\ a x_3 \end{pmatrix}.$$

Zwei Vektoren \vec{x} und \vec{y} werden addiert durch

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}.$$

Im folgenden werden wir in der Regel auf die Pfeile verzichten, wenn Verwechslungen nicht möglich sind (im Druck werden Punkte des Raumes und Vektoren meist durch die Drucktypen unterschieden).

Wir haben damit einfachste Beispiele von *Vektorräumen* kennen gelernt.

Ein *Vektorraum* $(V, +, \cdot)$ ist eine Menge V , auf der eine Addition (von Elementen aus V) und eine Multiplikation mit Skalaren (aus einem Körper \mathbb{K} , hier immer \mathbb{R} oder \mathbb{C}) definiert sind so, dass gilt:

- $(V, +)$ ist eine abelsche (kommutative) Gruppe
- $a(x + y) = ax + ay$, $(a + b)x = ax + bx$, $a(bx) = (ab)x$ und $1x = x$.

Diese Eigenschaften sind offenbar für \mathbb{R}^2 und \mathbb{R}^3 , und für entsprechend definierte \mathbb{R}^n erfüllt mit $\mathbb{K} = \mathbb{R}$; entsprechend für \mathbb{C}^n mit $\mathbb{K} = \mathbb{C}$.

Weitere Beispiele, mit in natürlicher Weise definierter Addition und skalarer Multiplikation sind:

- $P_n = \{\text{Polynome vom Grad } \leq n\}$,
- $P = \{\text{Polynome (beliebigen Grades)}\}$,
- $C(I) = \{\text{stetige Funktionen } I \rightarrow \mathbb{R} \text{ bzw. } I \rightarrow \mathbb{C}\}$.

Eine *Gerade* in \mathbb{R}^2 (oder \mathbb{R}^3 , oder...) kann auf mindestens zwei Arten beschrieben werden (und diese werden wir auch in allgemeineren Vektorräumen benutzen):

- Die Gerade g durch die Punkte x und y ($x \neq y$) ist die Menge der Punkte

$$g = \{x + t(y - x) : t \in \mathbb{R}\},$$

- die Gerade g durch den Punkt x mit Richtung des Vektors v ($\neq 0$) ist die Menge der Punkte

$$g = \{x + tv : t \in \mathbb{R}\};$$

im vorigen Fall kann natürlich $v := y - x$ gewählt werden.

Zwei Darstellungen

- $\{x + t(y - x) : t \in \mathbb{R}\}$ und $\{\tilde{x} + t(\tilde{y} - \tilde{x}) : t \in \mathbb{R}\}$ bzw.

$$- \{x + tv : t \in \mathbb{R}\} \text{ und } \{\tilde{x} + t\tilde{v} : t \in \mathbb{R}\}$$

beschreiben genau dann die gleiche Gerade, wenn t_0 und a_0 existieren mit

$$- \tilde{x} = x + t_0(y - x) \text{ und } \tilde{y} - \tilde{x} = a_0(y - x) \text{ bzw.}$$

$$- \tilde{x} = x + t_0v \text{ und } \tilde{v} = a_0v.$$

Die *Verbindungsstrecke* zweier Punkte mit Ortsvektoren x und y ist

$$\{x + t(y - x) : 0 \leq t \leq 1\} = \{\alpha x + \beta y : \alpha, \beta \geq 0, \alpha + \beta = 1\}.$$

Der *Schnitt* zweier Geraden g und h ist die Menge der gemeinsamen Punkte (offenbar die leere Menge oder genau ein Punkt oder die ganze Gerade $g = h$). Für

$$g = \{x + ty : t \in \mathbb{R}\} \quad \text{und} \quad h = \{\tilde{x} + s\tilde{y} : s \in \mathbb{R}\}$$

bedeutet das

$$x + ty = \tilde{x} + s\tilde{y},$$

also, in der Ebene \mathbb{R}^2 ,

$$ty_1 - s\tilde{y}_1 = \tilde{x}_1 - x_1$$

$$ty_2 - s\tilde{y}_2 = \tilde{x}_2 - x_2.$$

Das ist ein *lineares Gleichungssystem* für die Unbekannten s und t . (Geometrisch ist klar, dass dieses System genau dann eindeutig lösbar ist [d. h. es gibt genau einen gemeinsamen Punkt], wenn g und h nicht parallel sind. Sind sie parallel und verschieden, so gibt es keinen gemeinsamen Punkt [keine Lösung]; sind sie gleich, so gibt es natürlich unendlich viele gemeinsame Punkte.)

Wie löst man nun ein solches Gleichungssystem?

$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Wenn nicht alle $a_{ij} = 0$ sind, können wir o. E. annehmen, dass $a_{11} \neq 0$ ist (sonst ggf. vertauschen der Gleichungen und/oder Variablen). Addiert man das $-\frac{a_{21}}{a_{11}}$ -fache der ersten Gleichung zur zweiten, so erhält man das neue System

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ \left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right)x_2 &= b_2 - \frac{a_{21}}{a_{11}}b_1. \end{aligned}$$

Ist $a_{11}a_{22} - a_{21}a_{12} \neq 0$, so kann dieses System von unten her aufgelöst werden²:

$$\begin{aligned}x_2 &= \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}} = \frac{1}{D}(a_{11}b_2 - a_{21}b_1), \\x_1 &= \frac{1}{D}(a_{22}b_1 - a_{12}b_2),\end{aligned}$$

wobei

$$D := a_{11}a_{22} - a_{12}a_{21}$$

die *Determinante* $\det A$ des Systems bzw. der Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

ist. Die Zähler erhält man als Determinante der Matrix, in der die 1. bzw. 2. Spalte von A durch $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ ersetzt wird

$$x_1 = (\det A)^{-1} \det \begin{pmatrix} b_1 & a_{12} \\ b_1 & a_{22} \end{pmatrix}, \quad x_2 = (\det A)^{-1} \det \begin{pmatrix} a_{11} & b_1 \\ a_{12} & b_2 \end{pmatrix}.$$

Das werden wir später als einen Spezialfall der *Cramerschen Regel* erkennen.

Satz 7.1 a) *Ist die Determinante von A ungleich Null, so hat das System für jede rechte Seite genau eine Lösung, die durch die Cramersche Regel gegeben wird.*

b) *Ist die Determinante gleich 0, so hat das System entweder keine Lösung, oder die Menge der Lösungen enthält mindestens eine Gerade.*

Teil b) wird hier noch nicht bewiesen. Wir werden dies später allgemeiner nachholen. Hier folgt es aus der geometrischen Überlegung von weiter oben. Die Determinante ist gleich 0, wenn die Vektoren $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$, im obigen konkreten Fall $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ und $\begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}$, die gleiche Richtung (bis auf Vorzeichen) haben. Die entsprechenden Geraden sind dann entweder gleich oder verschieden und parallel. (Den Fall, dass eine oder beide Spalten gleich $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ sind überlege sich der Leser.)

11:20.11.07

²Rechnung für x_1 ist:

$$\begin{aligned}x_1 &= -\frac{a_{12}}{a_{11}}x_2 + \frac{b_1}{a_{11}} = -\frac{a_{12}}{a_{11}} \frac{a_{11}b_2 - a_{21}b_1}{D} + \frac{b_1}{a_{11}} \\ &= \frac{1}{a_{11}D} \left(-a_{12}a_{11}b_2 + a_{12}a_{21}b_1 + a_{11}a_{22}b_1 - a_{12}a_{21}b_1 \right) = \dots\end{aligned}$$

7.2 Lineare Abbildungen

Eine wichtige Rolle in Vektorräumen spielen die *linearen Abbildungen*. Sei V ein Vektorraum (im Augenblick meinen wir damit insbesondere \mathbb{R}^2 oder allgemeiner \mathbb{R}^n oder \mathbb{C}^n bzw. allgemein \mathbb{K}^n). Eine *lineare Abbildung* $L : V \rightarrow V$ ist eine Abbildung mit den Eigenschaften

$$L(ax + by) = aLx + bLy \quad \text{für alle } a, b \in \mathbb{K}, \quad x, y \in V$$

(\mathbb{K} ist hier \mathbb{R} im Fall von \mathbb{R}^n , \mathbb{C} im Fall \mathbb{C}^n). Ist im Fall \mathbb{K}^n

$$e_j := (0, \dots, 0, 1, 0, \dots, 0)^t$$

der j -te Einheitsvektor (mit der 1 an der j -ten Stelle; t steht für transponiert, d. h. die e_j sind Spaltenvektoren) und

$$a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{nj} \end{pmatrix} := a_j := Le_j,$$

so ist für beliebiges $x \in (\mathbb{R}^2, \mathbb{R}^n, \mathbb{C}^n)$

$$Ax = (a_1 : a_2 : \dots : a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix},$$

wobei

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix}$$

ist. Das heißt, jede lineare Abbildung wird auf diese Weise durch eine Matrix dargestellt. Umgekehrt ist offenbar jede so definierte Abbildung linear.

Spezialfälle sind

– die Nullabbildung $0x := 0$ für alle x ; sie wird durch die *Nullmatrix* $0 := \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$ erzeugt,

– die identische Abbildung $Ix := x$ für alle x ; sie wird durch die *Einheitsmatrix* $E := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ erzeugt.

Als Spezialfall in \mathbb{R}^2 wollen wir die *Drehung um den Winkel φ* (im mathematische positiven Sinn, falls $\varphi > 0$ ist, d. h. um $|\varphi|$ im mathematischen negativen Sinn, falls $\varphi < 0$ ist) betrachten. Wir wissen, dass wir \mathbb{R}^2 auch als die komplexe Ebene auffassen können. Dort wissen wir, dass die Drehung um den Winkel φ der Multiplikation mit $w(\varphi) = \cos \varphi + i \sin \varphi$ entspricht.

Die Zahl $z = u + iv$ geht dabei über in

$$w(\varphi)z = u \cos \varphi - v \sin \varphi + i(u \sin \varphi + v \cos \varphi).$$

Da Realteil der ersten Komponente, Imaginärteil der zweiten Komponente entspricht, ist also im obigen Sinn ($a_j := Le_j$)

$$a_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad a_2 = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix},$$

d. h. es gilt:

Satz 7.2 Die Drehung in \mathbb{R}^2 um den Winkel φ entgegen dem Uhrzeigersinn wird durch die Matrix

$$A_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

erzeugt.

7.3 Lineare Un-/Abhängigkeit

dies ist eine für Vektorräume grundlegende Begriffsbildung: Vektoren v_1, \dots, v_p aus einem \mathbb{K} -Vektorraum heißen *linear unabhängig*, wenn gilt

$$\text{aus } \sum_{j=1}^p c_j v_j = 0 \text{ mit } c_j \in (\mathbb{K}) \text{ folgt } c_1 = c_2 = \dots = c_p = 0,$$

in anderen Worten, wenn sich der Nullvektor nur als triviale Linearkombination³ der Vektoren v_1, \dots, v_p darstellen läßt. Andernfalls heißen die Vektoren v_1, \dots, v_p *linear abhängig*.

Sind v_1, \dots, v_p linear abhängig, so gibt es also Zahlen c_1, \dots, c_p , die nicht alle gleich 0 sind, mit $\sum_{j=1}^p c_j v_j = 0$. Durch Auflösen nach einem v_j mit $c_j \neq 0$ folgt also, dass sich (mindestens) ein v_j als Linearkombination der anderen v_j darstellen läßt.

Ein *maximales linear unabhängiges System* in V hat offenbar die Eigenschaften:

³Ein Ausdruck der Form $\sum_{j=1}^p c_j v_j$ heißt eine Linearkombination der Vektoren v_1, \dots, v_p .

- jedes Element aus V läßt sich als Linearkombination dieses Systems darstellen, (gäbe es ein Element v , das sich nicht darstellen lässt, so wäre $\{v_1, \dots, v_p, v\}$ ein linear unabhängiges System, im Widerspruch zur Maximalität von $\{v_1, \dots, v_p\}$).
- wenn ein Element weggelassen wird, kann nicht mehr jedes Element dargestellt werden. (bereits das weggelassene Element lässt sich nicht mehr darstellen).
- Außerdem ist die Darstellung eindeutig.
(Sind $v = \sum_{j=1}^p c_j v_j$ und $v = \sum_{j=1}^p c'_j v_j$ zwei Darstellungen von v , so folgt $0 = \sum_{j=1}^p (c_j - c'_j) v_j$, also $c_j = c'_j$ für alle j).

Mann nennt deshalb ein solches System eine *Basis* von V .

Wir werden später sehen, dass alle diese maximalen Systeme (bzw. Basen) in einem Vektorraum V gleich viele Elemente enthalten. Die Anzahl dieser Elemente nennt man deshalb die *Dimension* von V ($\dim V$).

In \mathbb{R}^n bzw. \mathbb{C}^n bilden offenbar die Einheitsvektoren e_1, \dots, e_n eine solche Basis, d. h. beide Räume haben die (reelle bzw. komplexe) Dimension n .

In P_n bilden die Monome x^0, x^1, \dots, x^n eine Basis; in P bilden die Monome x^0, x^1, x^2, \dots , eine (allerdings unendliche) Basis.

7.4 Norm und Skalarprodukt

Um von Stetigkeit (sowohl im Sinne der Folgendefinition, als auch im Sinne der $\varepsilon - \delta$ -Definition) von Funktionen $\mathbb{K} \rightarrow \mathbb{K}^n$ oder $\mathbb{K}^m \rightarrow \mathbb{K}^n$ usw. reden zu können, brauchen wir noch einen Abstands begriff (*Betrag, Norm, Metrik*):

Den *Abstand* der Punkte $x = (x_1, x_2)$ und $y = (y_1, y_2)$ erhält man (nach Pythagoras) mit

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad \text{bzw.} \quad = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

Die *Norm* eines Vektors x

$$\|x\| = \sqrt{x_1^2 + x_2^2} \quad \text{bzw.} \quad = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

ist also nichts anderes, als der Abstand des durch diesen Vektor beschriebenen Punktes vom Nullpunkt (in \mathbb{C}^m ist hier x_j^2 durch $|x_j|^2$ zu ersetzen).

Wir definieren allgemein: Eine *Norm* $\|\cdot\|$ auf einem Vektorraum V ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ mit den Eigenschaften:

(N1) $\|x\| \geq 0$ für $x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$ (*Positivität*)

(gilt hier nur die erste Eigenschaft, so spricht man von einer *Halbnorm*)

(N2) $\|ax\| = |a| \|x\|$ für alle $x \in V$, $a \in \mathbb{K}$ (*positive Homogenität*)

(N3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$ (*Dreiecksungleichung*)

Nur die dritte Eigenschaft ist für obige Norm in \mathbb{R}^2 oder \mathbb{R}^3 nicht völlig trivial (ist aber geometrisch offensichtlich); allgemein wird sie weiter unten bewiesen.

Ein *Skalarprodukt* auf einem Vektorraum V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ mit den Eigenschaften:

(S1) $\langle x, x \rangle \geq 0$ für alle $x \in V$; $\langle x, x \rangle = 0 \Leftrightarrow x = 0$,

(gilt hier nur die erste Eigenschaft, so redet man von einem *Semi-Skalarprodukt*)

(S2) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ für alle $x, y \in V$,

(S3) $\langle x, ay \rangle = a \langle x, y \rangle$ für alle $x, y \in V$, $a \in \mathbb{K}$,⁴

(S4) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ für alle $x, y, z \in V$.

Insbesondere gilt $\langle ax, y \rangle = \overline{a} \langle x, y \rangle$ und $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

Ein *Skalarprodukt* auf \mathbb{K}^m ist definiert durch

$$\langle x, y \rangle := \sum_{j=1}^m \overline{x_j} y_j \quad \text{für } x = (x_1, \dots, x_m), y = (y_1, \dots, y_m)$$

(im Falle von \mathbb{R}^m ist der Querstrich natürlich bedeutungslos). In diesem Fall gilt

$$\|x\| = \langle x, x \rangle^{1/2}$$

Tatsächlich liefert auf diese Weise jedes Skalarprodukt eine Norm. Die ersten zwei Eigenschaften sind offensichtlich (Ü). Für die dritte benötigen wir:

Satz 7.3 (Schwarzsche Ungleichung) *Ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V und $\|x\| = \langle x, x \rangle^{1/2}$, so gilt*

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (\text{Schwarzsche Ungleichung}).$$

Das Gleichheitszeichen gilt genau dann, wenn x und y linear abhängig sind (d. h. $x = ay$ oder $y = ax$); die Gleichung $\langle x, y \rangle = \|x\| \|y\|$, also ohne Betragsstriche, gilt genau dann, wenn ein positives a existiert mit $x = ay$ oder $y = ax$.

Beweis. Sei o. E. $y \neq 0$ (sonst trivial, oder vertauschen). Für alle $t \in \mathbb{R}$ gilt offensichtlich

$$\begin{aligned} 0 &\leq \|x - ty\|^2 = \|x\|^2 - 2t \operatorname{Re} \langle x, y \rangle + t^2 \|y\|^2, \\ 0 &\leq t^2 - \frac{2 \operatorname{Re} \langle x, y \rangle}{\|y\|^2} t + \frac{\|x\|^2}{\|y\|^2}. \end{aligned}$$

⁴Häufig wird auch die Homogenität im ersten Term gefordert, $\langle ax, y \rangle = a \langle x, y \rangle$; zusammen mit (S2) folgt dann $\langle x, ay \rangle = \overline{a} \langle x, y \rangle$.

Dieses quadratische Polynom in t hat also keine oder höchstens eine reelle Nullstelle. Da die Lösungen der quadratischen Gleichung $t^2 + at + b = 0$ gegeben sind durch

$$t_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b},$$

gibt es genau dann höchstens eine reelle Lösung, wenn $\frac{a^2}{4} - b \leq 0$ gilt. In unserem konkreten Fall bedeutet das

$$\frac{(\operatorname{Re} \langle x, y \rangle)^2}{\|y\|^4} - \frac{\|x\|^2}{\|y\|^2} \leq 0,$$

$$|\operatorname{Re} \langle x, y \rangle| \leq \|x\| \|y\|.$$

Im reellen Fall ($\mathbb{K} = \mathbb{R}$) ist man damit fertig.

Im komplexen Fall ($\mathbb{K} = \mathbb{C}$) ersetzt man x durch ax mit $|a| = 1$ so, dass $\langle ax, y \rangle = |\langle x, y \rangle|$ gilt. Man erhält damit

$$|\langle x, y \rangle| = \langle ax, y \rangle = |\operatorname{Re} \langle ax, y \rangle| \leq \|ax\| \|y\| = \|x\| \|y\|.$$

Der Rest des Beweises sei hier weggelassen! ■

12:22.11.07

Hiermit kann nun leicht die Dreiecksungleichung bewiesen werden:

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Wie kann man das *Skalarprodukt geometrisch* deuten? Dazu schauen wir uns den Fall \mathbb{R}^2 an und benutzen dazu einen Rückgriff auf die komplexe Multiplikation: Interpretieren wir Vektoren $x = (x_1, x_2)$ und $y = (y_1, y_2)$ als komplexe Zahlen $x_1 + ix_2$ und $y_1 + iy_2$, so gilt für den Winkel φ zwischen den Vektoren $x = (x_1, x_2)$ und $y = (y_1, y_2)$

$$\begin{aligned} \varphi &= \arg(x_1 + ix_2) - \arg(y_1 + iy_2) = \arg\left(\frac{x_1 + ix_2}{y_1 + iy_2}\right) \\ &= \arg\frac{(x_1 + ix_2)(y_1 - iy_2)}{y_1^2 + y_2^2} = \arg\frac{(x_1y_1 + x_2y_2) + i(x_2y_1 - x_1y_2)}{y_1^2 + y_2^2}, \\ &= \arccos \frac{\operatorname{Re} \dots}{|\dots|} = \arccos \frac{x_1y_1 + x_2y_2}{y_1^2 + y_2^2} \cdot \frac{\sqrt{y_1^2 + y_2^2}}{\sqrt{x_1^2 + x_2^2}} = \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}, \\ \cos \varphi &= \frac{\langle x, y \rangle}{\|x\| \|y\|} \end{aligned}$$

oder

$$\langle x, y \rangle = \|x\| \|y\| \cos \varphi.$$

Dies läßt sich auch in \mathbb{R}^3 entsprechend ausrechnen.

In höheren Dimensionen bzw. in anderen Räumen und bezüglich beliebiger Skalarprodukte wird der Winkel zwischen zwei Vektoren durch die obige Identität definiert.

7.5 Konvergenz in \mathbb{R}^m und \mathbb{C}^m

Mit Hilfe der Norm auf \mathbb{K}^m (und entsprechend auf \mathbb{C}^m) kann dort völlig analog wie in \mathbb{R} (oder $\mathbb{C} \sim \mathbb{R}^2$) mit Hilfe von $|\cdot|$ definiert werden:

- Konvergenz einer Folge (x_n) aus \mathbb{K}^m gegen ein $x \in \mathbb{R}^m$, ($\|x_n - x\| \rightarrow 0$ für $n \rightarrow \infty$),
- Cauchyfolge (x_n) in \mathbb{K}^m ($\|y_n - x_k\|$ klein für große n und k)..

Um Folgenglieder und Elemente aus \mathbb{K}^m besser unterscheiden zu können, schreiben wir Elemente aus \mathbb{K}^m in der Form

$$x = (\xi_1, \xi_2, \dots, \xi_m), \quad y = (\eta_1, \eta_2, \dots, \eta_m), \dots$$

Offensichtlich gilt wieder (mit $x_n = (\xi_{n,1}, \dots, \xi_{n,m})$)

- $x_n \rightarrow x \Leftrightarrow \xi_{n,j} \rightarrow \xi_j$ für $j = 1, \dots, m$,
- (x_n) ist Cauchyfolge $\Leftrightarrow (\xi_{n,j})_{n \in \mathbb{N}}$ sind Cauchyfolgen für alle $j = 1, \dots, m$.

Eine Teilmenge M von \mathbb{K}^m heißt *offen*, wenn zu jedem $x \in M$ ein $\varepsilon = \varepsilon(x) > 0$ existiert so, dass die Kugel $K_\varepsilon(x) := \{y \in \mathbb{K}^m : \|x - y\| < \varepsilon\}$ mit Radius ε ganz in M liegt. M heißt *abgeschlossen*, wenn $\mathbb{K} \setminus M$ offen ist.

Achtung: Es gibt Mengen, die weder offen noch abgeschlossen sind!! Die Intervalle (a, b) sind offen, $[a, b]$ sind abgeschlossen, $[a, b)$ und $(a, b]$ sind weder offen noch abgeschlossen.

Satz 7.4 $M \subset \mathbb{K}^m$ ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten Folge aus M in M liegt.

Beweis. \Rightarrow : Sei M abgeschlossen. Angenommen, es gibt eine Folge (x_n) aus M mit $x_n \rightarrow x$ und $x \in \mathbb{K}^m \setminus M$. Dann gibt es ein $\varepsilon > 0$ mit $K_\varepsilon(x) \subset \mathbb{K}^m \setminus M$ (da $\mathbb{K}^m \setminus M$ offen ist). Also existiert ein $N(\varepsilon)$ mit $x_n \in K_\varepsilon(x) \subset \mathbb{K}^m \setminus M$ für $n \geq N(\varepsilon)$ im Widerspruch zu (x_n) aus M .

\Leftarrow : Angenommen, M ist *nicht* abgeschlossen, d. h. $\mathbb{K} \setminus M$ ist *nicht* offen. Dann gibt es ein $x \in \mathbb{K} \setminus M$ mit $K_\varepsilon(x) \cap M \neq \emptyset$ für jedes $\varepsilon > 0$. Also gibt es eine Folge (x_n) mit $x_n \in K_{1/n}(x) \cap M$. Für diese Folge gilt $x_n \rightarrow x$, (x_n) aus M , $x \notin M$ im Widerspruch zur Voraussetzung. ■

Es gilt das Analogon zum Satz von BOLZANO–WEIERSTRASS:

Satz 7.5 *Jede beschränkte Folge in \mathbb{K}^m enthält eine konvergente Teilfolge. – Ist $M \subset \mathbb{K}^m$ abgeschlossen und beschränkt, so enthält jede Folge aus M eine konvergente Teilfolge mit Limes in M (man nennt deshalb eine beschränkte und abgeschlossene Teilmenge von \mathbb{R}^m auch kompakt).*

Beweis. Der zweite Teil folgt unmittelbar aus dem ersten.

Sei $(x_n) = ((\xi_{n,1}, \xi_{n,2}, \dots, \xi_{n,m}))$ eine beschränkte Folge in \mathbb{K}^m . Dann sind die Folgen $(\xi_{n,j})$ beschränkte Folgen in \mathbb{K} . Also existiert

- eine Teilfolge $(n_k^{(1)})_k$ von (n) so, dass $(\xi_{n_k^{(1)},1})_k$ konvergiert, $\xi_{n_k^{(1)},1} \rightarrow \xi_1$,
- eine Teilfolge $(n_k^{(2)})_k$ von $(n_k^{(1)})_k$ so, dass $(\xi_{n_k^{(2)},2})_k$ konvergiert, $\xi_{n_k^{(2)},2} \rightarrow \xi_2$,
und schließlich
- eine Teilfolge $(n_k^{(m)})_k$ von $(n_k^{(m-1)})_k$ so, dass $(\xi_{n_k^{(m)},m})_k$ konvergiert, $\xi_{n_k^{(m)},m} \rightarrow \xi_m$.

Also gilt $\xi_{n_k^{(m)},j} \rightarrow \xi_j$ für alle j und somit $x_{n_k^{(m)}} \rightarrow x = (\xi_1, \dots, \xi_m)$. ■

8 Eigenschaften stetiger Funktionen

8.1 Grundbegriffe und Beispiele

Im folgenden sei $D \subset \mathbb{R}, \mathbb{C}, \mathbb{R}^m, \mathbb{C}^m$ (eventuell auch allgemeiner eine Menge Ω , auf der eine Metrik definiert ist) der Definitionsbereich einer Funktion.

$$f : D \rightarrow \mathbb{R}, \mathbb{C}, \mathbb{R}^m, \mathbb{C}^m, \dots$$

Konkrete **Beispiele** für Funktionen sind:

Mit $D \subset \mathbb{R}$, der Zeitachse,

$X : D \rightarrow \mathbb{R}, X(t) =$ Ort eines Körpers zur Zeit t auf der Geraden,

$X : D \rightarrow \mathbb{R}^2$ oder $\mathbb{R}^3, X(t) =$ Ort des Körpers zur Zeit t in der Ebene oder im Raum,

$v : D \rightarrow \mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, v(t)$ Geschwindigkeit eines Körpers zur Zeit t , ggf. vektoriell.

$u : D \rightarrow \mathbb{R}, u(t) =$ Auslenkung einer Seite an der Stelle x ,

Mit $D \subset \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$,

$u : D \rightarrow \mathbb{R}, u(x, t) =$ Auslenkung einer Seite an der Stelle x zur Zeit t ,

Mit $D \subset \mathbb{R}^3$,

$T : D \rightarrow \mathbb{R}, T(x) =$ Temperatur an der Stelle x (Raum \times Temperaturachse)

Mit $D \subset \mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$,

$T : D \rightarrow \mathbb{R}, T(x, t) =$ Temperatur an der Stelle x zur Zeit t .

Man wird naiv annehmen, dass alle diese Funktionen stetig sind, d. h. dass sich bei geringen Änderungen von x und/oder t die Funktionswerte wenig ändern.

Wichtige explizite **Beispielfunktionen** sind:

1. Die *konstante Funktion* $f(x) = c$ für alle $x \in D$.
2. Die *identische Funktion* $\text{id}(x) = x$ für alle $x \in D$.
3. Die *Betragsfunktion* auf \mathbb{R}, \mathbb{R}^m oder $\mathbb{C}^m, f(x) = |x|$ oder $\|\cdot\|$ (gelegentlich auch abs für Absolutbetrag).
4. Die *Quadratwurzel* $f : [0, \infty) =: \mathbb{R}_+ \rightarrow \mathbb{R}, f(x) = \sqrt{x}$, (gelegentlich auch sqrt für Squareroot).
5. *Polynome* in $m \geq 1$ Variablen,

$$f : \mathbb{R}^m \rightarrow \mathbb{R}, f(x_1, \dots, x_m) = \sum a_{\alpha_1, \dots, \alpha_m} x_1^{\alpha_1} \dots x_m^{\alpha_m}$$

mit $a_{\alpha_1, \dots, \alpha_m} \in \mathbb{R}$; dabei wird nur über endlich viele m -Tupel $(\alpha_1, \dots, \alpha_m) \in \mathbb{N}_0^m$ summiert; die maximale Summe $\alpha_1 + \dots + \alpha_m$ mit $a_{\alpha_1, \dots, \alpha_m} \neq 0$ ist der *Grad* des Polynoms.

Entsprechend erhält man Polynome $\mathbb{C}^m \rightarrow \mathbb{C}$, wenn man $a_{\alpha_1, \dots, \alpha_m} \in \mathbb{C}$ zulässt.

6. *Rationale Funktionen* in $m \geq 1$ Variablen,

$$f : \{x \in \mathbb{R}^m : q(x) \neq 0\} \rightarrow \mathbb{R}, \quad f(x) = \frac{p(x)}{q(x)} \quad (p, q \text{ Polynome}).$$

Entsprechend im komplexen Fall.

7. Die *Dirichletfunktionen*

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 1 & \text{falls } x \text{ rational} \\ 0 & \text{falls } x \text{ irrational} \end{cases}$$

8. Eine *modifizierte Dirichletfunktion*

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} \frac{1}{q} & \text{falls } x = \frac{p}{q} \text{ gekürzt,} \\ 0 & \text{falls } x \text{ irrational.} \end{cases}$$

9. Das *größte Ganze* (auch *Gauß-Klammer* $[\cdot]$)

$$\text{int} : \mathbb{R} \rightarrow \mathbb{R}, \quad \text{int}(x) = [x] = \text{größte ganze Zahl} \leq x.$$

Zur Erinnerung: $f : D \rightarrow X$ heißt *stetig* in $x_0 \in D$, wenn

- zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, dass für alle $x \in D$ mit $|x - x_0| < \delta$ gilt $|f(x) - f(x_0)| < \varepsilon$ (evtl. $\|\cdot\|$ statt $|\cdot|$), oder
- für jede Folge (x_n) aus D mit $x_n \rightarrow x_0$ gilt $f(x_n) \rightarrow f(x_0)$.

Die ersten 6 Beispiele sind offenbar stetig (wo sie definiert sind).

Die Dirichletfunktion ist in keinem Punkt stetig, da in jeder Umgebung eines rationalen (irrationalen) Punktes irrationale (rationale) Punkte liegen.

Die modifizierte Dirichletfunktion ist in allen irrationalen Punkten stetig, in allen rationalen Punkten unstetig (Die Unstetigkeit in rationalen Punkte ist offensichtlich. Ist $x_n = p_n/q_n$ eine rationale Folge, die gegen eine irrationale Zahl konvergiert, so gilt $q_n \rightarrow \infty$, also $f(x_n) \rightarrow 0$).

Die Funktion int ist in allen ganzen Zahlen unstetig, sonst stetig.

13:27.11.07

Es sei hier daran erinnert, dass alle Funktionen die mit Hilfe der früher beschriebenen Konstruktionsprinzipien aus stetigen Funktionen zusammengesetzt werden, stetig sind.

$f : D \rightarrow X$ heißt in $x_0 \in D$

- *partiell stetig*, wenn die Einschränkung auf achsenparallele Geraden durch x_0 in x_0 stetig ist, d. h. $t \mapsto f(x_0 + te_j)$ ist stetig an der Stelle $t = 0$ und für alle $j = 1, \dots, m$.
- *richtungsstetig*, wenn die Einschränkung auf jede Gerade durch x_0 in x_0 stetig ist, d. h. $t \mapsto f(x_0 + tv)$ ist stetig an der Stelle $t = 0$ für alle $v \in \mathbb{R}^m$.

Offensichtlich gilt:

Satz 8.1 *Stetig \implies richtungsstetig \implies partiell stetig.*

Es gibt Funktionen (z. B. $\mathbb{R}^2 \rightarrow \mathbb{R}$), die richtungsstetig sind, aber nicht stetig, und solche die partiell stetig sind, aber nicht richtungsstetig:

Beispiel 8.2 a) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ wie folgt definiert:

$$f(x) = f(x_1, x_2) = 1 \quad \text{für } x_2 \leq 0 \quad \text{und für } x_2 \geq x_1^2,$$

$$f(x) = 0 \quad \text{für } x_2 = \frac{1}{2}x_1^2,$$

linear interpoliert in x_2 -Richtung im Zwischenbereich.

Der Leser fertige sich eine Skizze. An dieser erkennt er sofort: Dieses f ist richtungsstetig, aber nicht stetig in 0. Außerdem stelle er die Funktion im Bereich $0 < x_2 < x_1^2$ explizit analytisch dar.

b) Sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ wie folgt definiert:

$$f(x) = 1 \quad \text{für } x_2 \leq 0 \quad \text{und für } x_2 \geq x_1$$

$$f(x) = 0 \quad \text{für } x_2 = \frac{1}{2}|x_1|,$$

dazwischen linear interpoliert.

Dieses g ist partiell stetig, aber nicht richtungsstetig in 0.

Beide Funktionen sind übrigens stetig in allen von 0 verschiedenen Punkten.. □

Bemerkung 8.3 In der Literatur findet man häufig das Beispiel

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{für } (x_1, x_2) \neq (0, 0), \\ 0 & \text{für } (x_1, x_2) = (0, 0). \end{cases}$$

Der Leser überzeuge sich davon, dass diese Funktion in $(0, 0)$ partiell stetig ist, aber nicht richtungsstetig (also auch nicht stetig). – Später erkennen wir, dass diese Funktion partiell differenzierbar ist, aber nicht (total) differenzierbar.

8.2 Grenzwerte von Funktionen

a heißt ein *Berührungspunkt* von D wenn gilt

$$\forall \varepsilon > 0 \quad \exists x \in D \quad \text{mit } |x - a| < \varepsilon,$$

oder gleichwertig

$$\exists \text{ Folge } (x_n) \text{ aus } D \text{ mit } x_n \rightarrow a \text{ für } n \rightarrow \infty.$$

Offenbar gilt z. B.:

- Jedes $a \in D$ ist Berührungspunkt,
- die Randpunkte jedes Intervalls I sind Berührungspunkte von I .

Sei f eine auf D definierte Funktion. A heißt der *Grenzwert der Funktion f* im Berührungspunkt a von D , kurz $A = \lim_{x \rightarrow a} f(x)$, wenn gilt

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D \quad \text{mit } |x - a| < \delta \quad \text{gilt } |f(x) - A| < \varepsilon,$$

oder gleichwertig

$$\forall \text{ Folge } (x_n) \text{ aus } D \text{ mit } x_n \rightarrow a \quad \text{gilt } f(x_n) \rightarrow A.$$

Wie man leicht sieht, gilt:

- Ist $a \in D$ und existiert $\lim_{x \rightarrow a} f(x)$, so gilt $f(a) = \lim_{x \rightarrow a} f(x)$.
- Ist $a \in D$ und f stetig in a , so existiert $\lim_{x \rightarrow a} f(x)$ und es gilt $f(a) = \lim_{x \rightarrow a} f(x)$.
- Ist $a \notin D$ und existiert $A := \lim_{x \rightarrow a} f(x)$, so ist

$$\hat{f}: D \cup \{a\} \rightarrow \mathbb{R}^n \text{ (oder } \dots), \quad \hat{f}(x) = \begin{cases} f(x) & \text{für } x \in D, \\ A & \text{für } x = a \end{cases}$$

stetig in a .

Ist $D \subset \mathbb{R} = \mathbb{R}^1$, so betrachtet man auch *einseitige Grenzwerte*

$$\lim_{x \nearrow a} f(x) = \lim_{\substack{x \rightarrow a \\ x < a}} f(x) \quad \text{und} \quad \lim_{x \searrow a} f(x) = \lim_{\substack{x \rightarrow a \\ x > a}} f(x).$$

Auch diese existieren genau dann, wenn $f(x_n)$ für jede Folge (x_n) aus D mit $x_n \nearrow a$ bzw. $x_n \searrow a$ konvergiert.

Für die Funktion $[\cdot]$ (Gauß-Klammer) existieren in jedem Punkt beide einseitigen Grenzwerte. In den ganzen Zahlen sind sie verschieden, der Limes im obigen Sinne existiert nicht. In nicht ganzzahligen Punkten existiert der Limes im obigen Sinn.

Beispiel 8.4 Die Funktion

$$f(x) = \begin{cases} \sin \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

ist stetig in allen $x \neq 0$, unstetig in $x = 0$. In 0 existieren weder die einseitigen Grenzwerte noch der beidseitige. \square

Beispiel 8.5 Die Funktion

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$$

ist überall stetig. In jedem Punkt existiert der beidseitige Limes und ist gleich dem Funktionswert an dieser Stelle. \square

8.3 Eigenschaften stetiger Funktionen einer Variablen

Zunächst Aussagen, die nur für Funktionen von $I \subset \mathbb{R}^1$ nach \mathbb{R}^1 Sinn machen:

Satz 8.6 (Zwischenwertsatz) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $f(a) \neq f(b)$. Dann gibt es zu jedem c zwischen $f(a)$ und $f(b)$ (mindestens) ein $x_0 \in (a, b)$ mit $f(x_0) = c$.

Beweis. Sei o. E. $f(a) < f(b)$ (sonst?)

$$M := \{x \in [a, b] : f(x) < c\}.$$

Dann ist $M \neq \emptyset$, da $a \in M$ ist, und M ist durch b nach oben beschränkt. Also existiert $x_0 := \sup M$.

Dann existiert eine Folge (x_n) aus M mit $x_n \rightarrow x_0$ und somit wegen der Stetigkeit von f , da $f(x_n) < c$ gilt,

$$f(x_0) = \lim_{n \rightarrow \infty} f(x_n) \leq c.$$

Da andererseits $x_0 + \frac{1}{n}$ für alle $n \in \mathbb{N}$ nicht in M liegt, gilt $f\left(x_0 + \frac{1}{n}\right) \geq c$ für alle n , und somit

$$f(x_0) = \lim_{n \rightarrow \infty} f\left(x_0 + \frac{1}{n}\right) \geq c,$$

und damit $f(x_0) = c$.

Wegen $f(a) < c < f(b)$ ist $x_0 \in (a, b)$. ■

Beispiel 8.7 Jedes Polynom ungeraden Grades hat mindestens eine (reelle) Nullstelle, und entsprechend für jedes $c \in \mathbb{R}$ mindestens eine c -Stelle.

Ist

$$p(x) = \sum_{j=0}^n a_j x^j \quad \text{mit } n \text{ ungerade und } a_n \neq 0$$

so gilt für große $|x|$

$$p(x) = x^n \sum_{j=0}^n a_j x^{j-n} = x^n (a_0 x^{-n} + a_1 x^{1-n} + \dots + a_n).$$

Für hinreichend große $|x|$ ist $|a_0 x^{-n} + a_1 x^{1-n} + \dots + a_{n-1} x^{-1}| < \frac{1}{2}|a_n|$, d. h. $p(x) \rightarrow \pm\infty$ für $x \rightarrow -\infty$, falls $a_n > 0$ ist, gegen $\mp\infty$ für $x \rightarrow \infty$, falls $a_n < 0$ ist. Damit folgt die Behauptung aus dem Zwischenwertsatz. \square

Wir wissen, dass eine Funktion f eine *Umkehrfunktion* f^{-1} besitzt, wenn sie injektiv ist. Dies ist für eine Funktion auf einem reellen Intervall insbesondere dann erfüllt, wenn sie streng monoton ist.

Satz 8.8 (Umkehrfunktion) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und streng monoton, $A := f(a)$, $B := f(b)$. Dann bildet f das Intervall $[a, b]$ bijektiv auf $[A, B]$ (falls f wachsend ist) bzw. $[B, A]$ (falls f fallend ist) ab. Die Umkehrfunktion

$$f^{-1} : [A, B] \rightarrow [a, b] \quad \text{bzw.} \quad [B, A] \rightarrow [a, b],$$

$$y \mapsto x, \quad \text{falls } f(x) = y \text{ ist,}$$

ist ebenfalls stetig und streng monoton (wachsend oder fallend, wie f).

Beweis. O.E. sei f streng wachsend. Aus $a < x < b$ folgt $A < f(x) < B$ für alle $x \in (a, b)$, also insbesondere $A < B$; f ist injektiv. Und aus dem Zwischenwertsatz folgt, dass f jeden Wert zwischen A und B annimmt, d. h.

$$f : [a, b] \rightarrow [A, B] \text{ ist bijektiv}$$

und es existiert $f^{-1} : [A, B] \rightarrow [a, b]$.

f^{-1} ist streng monoton: Sei $y < z$. Wäre $f^{-1}(y) \geq f^{-1}(z)$, so würde aus der Monotonie von f folgen

$$y = f^{-1}(f(y)) \geq f^{-1}(f(z)) = z,$$

ein Widerspruch.

f^{-1} ist stetig in $y_0 \in [A, B]$:

(α) $\left(f^{-1}\left(y_0 + \frac{1}{n}\right)\right)$ ist fallend und durch $f^{-1}(y_0)$ nach unten beschränkt, also konvergent gegen ein $x_0 \in [a, b]$. Mit der Stetigkeit von f folgt daraus

$$f(x_0) = \lim_{n \rightarrow \infty} f\left(f^{-1}\left(y_0 + \frac{1}{n}\right)\right) = \lim_{n \rightarrow \infty} \left(y_0 + \frac{1}{n}\right) = y_0,$$

d. h. es gilt $f^{-1}(y_0) = x_0$ und somit

$$f^{-1}\left(y_0 + \frac{1}{n}\right) \rightarrow x_0 = f^{-1}(y_0).$$

Entsprechend folgt

$$f^{-1}\left(y_0 - \frac{1}{n}\right) \rightarrow f^{-1}(y_0).$$

für $y_0 = A$ hat nur die erste aussage Sinn, für $y_0 = B$ nur die zweite (rechtsseitige bzw. linksseitige Stetigkeit).

(β) Sei nun (y_n) eine Folge aus $[A, B]$ mit $y_n \rightarrow y_0$. Für jedes $\varepsilon > 0$ existiert nach (α) ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$f^{-1}(y_0) - f^{-1}\left(y_0 - \frac{1}{N}\right) < \frac{\varepsilon}{2} \quad \text{und} \quad f^{-1}\left(y_0 + \frac{1}{N}\right) - f^{-1}(y_0) < \frac{\varepsilon}{2}.$$

Außerdem gibt es ein N' mit

$$|y_n - y_0| < \frac{1}{N} \quad \text{für } n \geq N'.$$

und somit für $n \geq N'$

$$\begin{aligned} \left|f^{-1}(y_n) - f^{-1}(y_0)\right| &\leq \left|f^{-1}\left(y_0 + \frac{1}{N}\right) - f^{-1}\left(y_0 - \frac{1}{N}\right)\right| \\ &= \left|f^{-1}\left(y_0 + \frac{1}{N}\right) - f^{-1}(y_0)\right| + \left|f^{-1}(y_0) - f^{-1}\left(y_0 - \frac{1}{N}\right)\right| < \varepsilon, \end{aligned}$$

d. h. $f^{-1}(y_n) \rightarrow f^{-1}(y_0)$.

Ein eleganterer, wenn auch vielleicht etwas begrifflicherer Beweis der Stetigkeit geht wie folgt: Sei $y_0 = f(x_0)$. Es soll gezeigt werden:

$$\forall \varepsilon > 0 \quad \text{mit} \quad \left|f^{-1}(y) - f^{-1}(y_0)\right| < \varepsilon \quad \text{für } |y - y_0| < \delta.$$

Sei also $\varepsilon > 0$,

$$y_- := f(x_0 - \varepsilon), \quad y_+ := f(x_0 + \varepsilon), \quad \delta := \min\{y_+ - y_0, y_0 - y_-\} > 0.$$

Dann ist $[y_0 - \delta, y_0 + \delta] \subset [y_-, y_+]$ ganz in $f([x_0 - \varepsilon, x_0 + \varepsilon])$ enthalten, d. h. $f^{-1}([y_0 - \delta, y_0 + \delta]) \subset [x_0 - \varepsilon, x_0 + \varepsilon]$ bzw.

$$\left|f^{-1}(y) - f^{-1}(y_0)\right| < \varepsilon \quad \text{für } |y - y_0| < \delta.$$

■

Entsprechend gilt für beliebige Intervalle

Satz 8.9 Sei I ein beliebiges Intervall (offen, abgeschlossen, halboffen, Halbachse,...), $f : I \rightarrow \mathbb{R}$ stetig und streng monoton. Dann ist $J = f(I) := \{f(x) : x \in I\}$ ebenfalls ein Intervall, $f : I \rightarrow J$ ist bijektiv, $f^{-1} : J \rightarrow I$ ist stetig und streng monoton.

Auf den *Beweis* kann verzichtet werden.

Beispiele für die Anwendung dieser Sätze sind

- k -te Wurzeln als Umkehrfunktion der k -ten Potenz, $\sqrt[k]{\cdot} : [0, \infty) \rightarrow [0, \infty)$ für gerades k ,
- k -te Wurzeln $\sqrt[k]{\cdot} : \mathbb{R} \rightarrow \mathbb{R}$ für ungerades k ,
- \ln als Umkehrfunktion der Exponentialfunktion, $\ln : (0, \infty) \rightarrow \mathbb{R}$.

8.4 Extremwerte von Funktionen

Eine Funktion $f : D \rightarrow \mathbb{R}$ ($\mathbb{C}, \mathbb{R}^m, \mathbb{C}^m$) heißt *beschränkt*, wenn ein C existiert mit $|f(x)| \leq C$ für alle $x \in D$. Für $f : D \rightarrow \mathbb{R}$ kann entsprechend *nach oben beschränkt* und *nach unten beschränkt* definiert werden.

Eine Funktion $f : D \rightarrow \mathbb{R}$ *nimmt in* $x_0 \in D$ *ihr (globales) Maximum an*, wenn gilt

$$f(x) \leq f(x_0) \quad \text{für alle } x \in D,$$

entsprechend (globales) *Minimum*, falls

$$f(x) \geq f(x_0) \quad \text{für alle } x \in D.$$

Das Maximum bzw. Minimum heißt *strikt*, falls gilt

$$f(x) < f(x_0) \quad \text{bzw.} \quad f(x) > f(x_0) \quad \text{für alle } x \in D \setminus \{x_0\}$$

Man sagt, in x_0 liegt ein *lokales Extremum* oder ein *striktes lokales Extremum* (Maximum oder Minimum) vor, wenn dies wenigstens für ein kleines Intervall um x_0 gilt.

14:29.11.07

Man beachte, dass eine beschränkte reellwertige Funktion zwar immer ein Infimum und ein Supremum besitzt. Dies ist nicht notwendig ein Minimum bzw. Maximum:

Beispiel 8.10 $\text{id} : (0, 1) \rightarrow \mathbb{R}$ hat kein Minimum und kein Maximum. Aber auch die Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) \begin{cases} 1/2 & \text{für } x = 0 \text{ und } x = 1, \\ x & \text{für } 0 < x < 1 \end{cases}$$

hat kein Minimum und kein Maximum. (In Bezug auf den folgenden Satz ist festzustellen, dass im ersten Fall $(0, 1)$ nicht kompakt ist, während im zweiten Fall f nicht stetig ist; beide Voraussetzungen sind also wesentlich.) \square

Satz 8.11 (Satz vom Maximum) Sei $K \subset \mathbb{R}^m$ oder \mathbb{C}^m kompakt, $f : K \rightarrow \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt sein Maximum und sein Minimum an (an jeweils mindestens einer Stelle).

Beweis. a) *Beschränktheit:* Nehmen wir an, dass f unbeschränkt ist, d. h.

$$\exists \text{ Folge } (x_n) \text{ aus } K \text{ mit } |f(x_n)| \geq n.$$

Da K kompakt ist, existiert eine Teilfolge (x_{n_k}) von (x_n) und ein $x_0 \in K$ mit $x_{n_k} \rightarrow x_0$ für $k \rightarrow \infty$. Mit der Stetigkeit von f folgt $f(x_{n_k}) \rightarrow f(x_0)$ im Widerspruch zur Unbeschränktheit der Folge $(f(x_{n_k}))$.

b) *Existenz des Maximums:* Nach Teil a ist die Bildmenge $\{f(x) : x \in K\}$ beschränkt, besitzt also ein Supremum $\sup f$. Es existiert also eine Folge (x_n) aus K mit

$$f(x_n) \rightarrow \sup\{f(x) : x \in K\} = \sup f.$$

Da K kompakt ist, existiert wieder eine Teilfolge (x_{n_k}) von (x_n) und ein x_{\max} mit

$$x_{n_k} \rightarrow x_{\max} \text{ für } k \rightarrow \infty \text{ und } f(x_{\max}) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \sup f.$$

f nimmt also in x_{\max} sein Maximum an.

Entsprechend findet man ein x_{\min} , indem f sein Minimum annimmt. \blacksquare

Als wichtige Folgerung ergibt sich:

Korollar 8.12 Sei $K \subset \mathbb{R}^m$ oder \mathbb{C}^m kompakt, $f : K \rightarrow \mathbb{R}$ stetig, $f(x) \neq 0$ für alle $x \in K$. Dann existiert ein $c > 0$ mit $|f(x)| \geq c$ für alle $x \in K$. (Wenn f überall ungleich 0 ist, hat der Betrag eine strikt positive untere Schranke.)

Beweis. $|f| : K \rightarrow \mathbb{R}$, $|f|(x) = |f(x)|$ ist stetig, nimmt also sein Minimum $\min |f|$ in einem Punkt $x_{\min} \in K$ an. Also gilt für alle $x \in K$

$$|f(x)| \geq \min |f| = |f(x_{\min})| =: c > 0.$$

■

Schließlich ist noch interessant, wie sich kompakte Teilmengen unter stetigen Abbildungen verhalten:

Satz 8.13 *Sei $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ stetig. Ist $K \subset \mathbb{R}^m$ kompakt (also beschränkt und abgeschlossen), so ist auch $f(K) = \{f(x) : x \in K\}$ kompakt.*

Beweis. Es ist zu zeigen, dass jede Folge (y_ℓ) aus $f(K)$ eine konvergente Teilfolge enthält, deren Limes in $f(K)$ liegt. Zu jedem $y_\ell \in f(K)$ gibt es (mindestens) ein $x_\ell \in K$ mit $f(x_\ell) = y_\ell$. Da K kompakt ist, gibt es eine Teilfolge (x_{l_k}) mit $x_{l_k} \rightarrow x \in K$ für $k \rightarrow \infty$. Wegen Stetigkeit von f gilt

$$y_{l_k} = f(x_{l_k}) \rightarrow f(x) =: y \in f(K).$$

■

Bemerkung 8.14 Die eleganteste und in jedem Kontext vernünftige Definition von *kompakt* ist: K heißt *kompakt*, wenn jede Überdeckung von K mit offenen Mengen endlich viele Mengen enthält, die K überdecken. Mit der früher angegebenen Definition der Stetigkeit („Urbild jeder offenen Menge ist stetig“) kann dann der obige Satz besonders einfach bewiesen werden.

9 Differentiation

Zur Motivation betrachten wir die Begriffe „Geschwindigkeit“ und „Tangente“ an einer Kurve.

Ein Körper bewege sich im Raum \mathbb{R}^3 (entsprechend in der Ebene \mathbb{R}^2 oder auf der Geraden \mathbb{R}). Zur Zeit t sei er im Punkt $x(t) \in \mathbb{R}^3$. Im Zeitintervall $[t_0, t]$ ist die *mittlere Geschwindigkeit* gleich dem *Differenzenquotienten*

$$v_{t_0,t} = \frac{x(t) - x(t_0)}{t - t_0}.$$

Dies ist ein Vektor, der von $x(t_0)$ in Richtung $x(t)$ zeigt, dessen Länge gleich $\frac{\|x(t) - x(t_0)\|}{|t - t_0|}$ ist. Mit der tatsächlichen Geschwindigkeit zu irgendeinem Zeitpunkt zwischen t_0 und t hat das wenig zu tun, da der Körper vielleicht gar nicht den direkten Weg nimmt, sondern einen unbekanntem Umweg.

Wenn man das Intervall $[t_0, t]$ (oder entsprechend $[t, t_0]$) immer mehr verkleinert, wird der Vektor von $x(t_0)$ nach $x(t)$ vermutlich immer besser mit dem wirklichen Weg des Körpers übereinstimmen, und die Geschwindigkeit $v_{t_0,t}$ wird immer besser mit der tatsächlichen Geschwindigkeit für die Zeitpunkte zwischen t_0 und t übereinstimmen.

Wenn für $t \rightarrow t_0$ (für $t < t_0$ und $t > t_0$) der Grenzwert

$$v_{t_0} := \lim_{t \rightarrow t_0} \frac{x(t) - x(t_0)}{t - t_0}$$

existiert, nennt man diesen die (*Momentan-*)*Geschwindigkeit* zur Zeit t_0 . Dies ist wieder ein Vektor, dessen Richtung die Bewegungsrichtung zur Zeit t_0 angibt, während der Betrag die (skalare) Geschwindigkeit angibt. Ist $v_{t_0} = 0$, so befindet sich der Körper zur Zeit t_0 in Ruhe; eine Bewegungsrichtung gibt es nicht (der Nullvektor hat keine Richtung).

Das *Tangentenproblem* ist mit obigem Problem eng verwandt. Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ eine Funktion. Wir betrachten den Graphen von f . Für $x_0 \in I$ und $x \in I$ mit $x \neq x_0$ gibt der *Differenzenquotient*

$$\frac{f(x) - f(x_0)}{x - x_0} \cong \frac{\Delta f}{\Delta x}$$

die *Steigung* der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$ an (*Steigung* = Tangens des Steigungswinkels). Existiert der Limes für $x \rightarrow x_0$,

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

so sollte dies also die Steigung der Tangente an die Kurve im Punkt $(x_0, f(x_0))$ sein; wir nennen dies die Steigung des Graphen im Punkt $(x_0, f(x_0))$.

Dies führt uns zur *Definition*: Eine Funktion $f : I \rightarrow \mathbb{R}^m$ heißt im Punkt $x_0 \in I$ *differenzierbar*, wenn der Grenzwert

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad (f'(t_0) \text{ statt } f'(t_0), \text{ wenn } t \text{ die Zeit ist})$$

existiert; $f'(x_0)$ heißt die *Ableitung* von f an der Stelle x_0 . Man schreibt dafür auch

$$\frac{d}{dx} f(x_0), \quad \left. \frac{d}{dx} f(x) \right|_{x=x_0}, \quad \frac{df}{dx}(x_0).$$

Diese Schreibweise hat auch die Bezeichnung „Differentialquotient“ zur Folge: aus den oben angegebenen Größen Δf und Δx wird im *Infinitesimalen* (im „unendlich kleinen“) df bzw. dx , die man als Differentiale bezeichnet, womit man zum *Differentialquotienten* kommt.

Die Vorstellung, dass $x(t)$ den Ort eines Körpers zur Zeit t und v_{t_0} die Geschwindigkeit zum Zeitpunkt t_0 beschreiben, legt nahe, dass für t nahe bei t_0 gelten sollte

$$x(t) \sim x(t_0) + (t - t_0)v_{t_0}$$

(offen ist zunächst was \sim genau bedeutet, d. h. wie genau diese „Gleichung“ gilt). Der folgende Satz beschreibt dies genauer; dies führt zu einer äquivalenten Definition der Differenzierbarkeit, die im Gegensatz zu obiger Definition fast unverändert bei Funktionen mehrerer Variablen verwendet werden kann.

Satz 9.1 (Lineare Approximierbarkeit) *Eine Funktion $f : I \rightarrow \mathbb{R}^m$ ist in $x_0 \in I$ genau dann differenzierbar, wenn f bei x_0 linear approximierbar ist, d. h. wenn gilt*

$$f(x) = f(x_0) + (x - x_0)y + \varphi_{x_0}(x)$$

mit $y \in \mathbb{R}^m$ und $\frac{\varphi_{x_0}(x)}{x - x_0} \rightarrow 0$ für $x \rightarrow x_0$; y ist dann die Ableitung von f im Punkt x_0 (die Abweichung in obiger Gleichung ist also für x nahe x_0 „viel kleiner“ als $|x - x_0|$).⁵

Beweis. \Rightarrow : Ist f in x_0 differenzierbar, so gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \rightarrow f'(x_0) \quad \text{für } x \rightarrow x_0;$$

für $\varphi_{x_0}(x) := f(x) - f(x_0) - (x - x_0)f'(x_0)$ gilt also

$$\frac{\varphi_{x_0}(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \rightarrow 0 \quad \text{für } x \rightarrow x_0,$$

⁵Bei Funktionen $f : A \rightarrow \mathbb{R}^m$ mit $A \subset \mathbb{R}^n$ wird man definieren: f heißt in $x_0 \in A$ (total) *differenzierbar*, wenn eine lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ existiert und ein $\varphi_{x_0} : A \rightarrow \mathbb{R}^m$ mit $\varphi_{x_0}(x)/\|x - x_0\| \rightarrow 0$ für $x \rightarrow x_0$ so, dass gilt

$$f(x) = f(x_0) + L(x - x_0) + \varphi_{x_0}(x).$$

L heißt dann die Ableitung von f im Punkt x_0 .

d. h. die lineare Approximierbarkeit ist erfüllt mit $y = f'(x_0)$.

\Leftarrow : Ist umgekehrt f linear approximierbar bei x_0 , so gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = y + \frac{\varphi_{x_0}(x)}{x - x_0} \rightarrow y \quad \text{für } x \rightarrow x_0,$$

d. h. f ist in x_0 differenzierbar mit $f'(x_0) = y$. ■

Eine Funktion $f : I \rightarrow \mathbb{R}^m$ heißt *differenzierbar* (ohne weiteren Zusatz), wenn es in jedem Punkt aus I differenzierbar ist.

Der folgende Satz ermöglicht es, dass wir bei theoretischen Überlegungen statt $f : I \rightarrow \mathbb{R}^m$ nur die Funktion $f : I \rightarrow \mathbb{R}$ betrachten müssen:

Satz 9.2 $f : I \rightarrow \mathbb{R}^m$ habe die Form $x \mapsto f(x) = (f_1(x), \dots, f_m(x))$. Dann ist f genau dann differenzierbar (in x_0), wenn jedes f_j ($j = 1, \dots, m$) (in x_0) differenzierbar ist. Es gilt $f'(x_0) = (f'_1(x_0), \dots, f'_m(x_0))$.

Der *Beweis* ist offensichtlich.

Satz 9.3 Ist $f : I \rightarrow \mathbb{R}^m$ differenzierbar (in x_0), so ist f stetig (in x_0).

Beweis. Aus der linearen Approximierbarkeit folgt

$$\begin{aligned} \lim_{x \rightarrow x_0} f(x) &= f(x_0) + \lim_{x \rightarrow x_0} (x - x_0)y + \lim_{x \rightarrow x_0} \frac{\varphi_{x_0}(x)}{x - x_0}(x - x_0) \\ &= f(x_0) + 0 \cdot y + 0 \cdot 0 = f(x_0). \end{aligned}$$

d. h. f ist stetig in x_0 . ■

Beispiel 9.4 Die *konstante Funktion* $f(x) = c$ für alle $x \in I$ hat offenbar in jedem Punkt die Ableitung 0, da alle Differenzquotienten verschwinden. Für $c \in \mathbb{R}$ bedeutet das, dass der Graph horizontal verläuft. Für $c \in \mathbb{R}^m$ mit $m > 1$ ist der Vektor $f'(x) = 0$ für alle x . □

Beispiel 9.5 Die *Identität* $\text{id} : \mathbb{R} \rightarrow \mathbb{R}$ hat in jedem Punkt die Ableitung 1, da alle Differenzquotienten gleich 1 sind. □

Beispiel 9.6 Die Funktion $f(x) = x^2$ hat für jedes $x \in \mathbb{R}$ die Ableitung $2x$, denn

$$\lim_{z \rightarrow x} \frac{f(z) - f(x)}{z - x} = \lim_{z \rightarrow x} \frac{z^2 - x^2}{z - x} = \lim_{z \rightarrow x} \frac{z - x}{z - x} (z + x) = \lim_{z \rightarrow x} (z + x) = 2x.$$

(Mit der Produktregel wird dies, wie auch die Ableitung von x^m für $m > 2$ aus der Ableitung von x folgen.) \square

Beispiel 9.7 Die *Exponentialfunktion* $\exp : \mathbb{R} \rightarrow \mathbb{R}$ hat in jedem $x \in \mathbb{R}$ die Ableitung $\exp(x)$. Zum Beweis sei an die Ungleichungen

$$1 + x \leq e^x \quad \text{und} \quad |e^x - 1| \leq \frac{|x|}{1 - |x|} \quad \text{für } |x| < 1$$

erinnert, die in Kapitel 3 (Satz 3.2) bewiesen wurde. Daraus wurde in Aufgabe 7 gefolgert

$$\frac{1}{1 + |x|} \leq \frac{e^x - 1}{x} \leq \frac{1}{1 - |x|}.$$

Also folgt

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1 \quad (= e^0),$$

und allgemeiner

$$\lim_{z \rightarrow x} \frac{e^z - e^x}{z - x} = \lim_{z \rightarrow x} e^x \frac{e^{z-x} - 1}{z - x} = e^x \lim_{y \rightarrow 0} \frac{e^y - 1}{y} = e^x.$$

\square

Beispiel 9.8 *Sinus und Cosinus*: Wir wissen $|\sin x| \leq |x|$ für alle x , d. h.

$$\sin x \leq x \quad \text{für } x \geq 0 \quad \text{und} \quad \sin x \geq x \quad \text{für } x \leq 0, \quad \text{also} \quad \frac{\sin x}{x} \leq 1 \quad \text{für alle } x.$$

Außerdem gilt

$$|\tan x| \geq |x|, \quad \text{also} \quad \frac{\tan x}{x} \geq 1 \quad \text{für} \quad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}.$$

(Die Ungleichung $|\tan x| \geq |x|$ sieht man so: Man zerlegt den Kreisbogen zum Winkel x in kleine Teile und projiziert die Sekanten vom Nullpunkt aus auf die Gerade durch $(1, 0)$ parallel zur y -Achse. Die Projektion jeder Sekante ist größer als die entsprechende Sekante. Damit ist die Projektion des Bogens x (diese ist $= \tan x$) größer als die Länge x des Bogens, da die Länge des Kreisbogens der Limes der Summe der Sekantenlängen bei immer feinerer Zerlegung ist.)

Aus $\cos x \leq \frac{\tan x}{x} \cos x = \frac{\sin x}{x} \leq 1$ für x nahe 0 und $\cos x \rightarrow 1$ für $x \rightarrow 0$ folgt

$$\lim_{x \rightarrow 0} \frac{\sin x - \sin 0}{x - 0} = \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1, \quad \sin' 0 = 1 \quad (= \cos 0).$$

Wir wissen auch $\cos x \geq 1 - \frac{1}{2}x^2$, woraus folgt (wegen $\cos x - 1 \leq 0$)

$$0 \geq \lim_{x \rightarrow 0} \frac{\cos x - \cos 0}{|x|} = \lim_{x \rightarrow 0} \frac{\cos x - 1}{|x|} \geq - \lim_{x \rightarrow 0} \frac{1}{2}|x| = 0, \quad \cos' 0 = 0 (= -\sin 0).$$

Mit Hilfe der Additionstherme folgt

$$\begin{aligned} \sin'(x) &= \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} = \lim_{h \rightarrow 0} \frac{\sin x \cosh + \cos x \sinh - \sin x}{h} \\ &= \lim_{h \rightarrow 0} \left\{ \sin x \frac{\cosh - 1}{h} + \cos x \frac{\sinh}{h} \right\} = \cos x \\ \cos'(x) &= \lim_{h \rightarrow 0} \frac{\cos(x+h) - \cos x}{h} = \lim_{h \rightarrow 0} \frac{\cos x \cosh - \sin x \sinh - \cos x}{h} \\ &= \lim_{h \rightarrow 0} \left\{ \cos x \frac{\cosh - 1}{h} - \sin x \frac{\sinh}{h} \right\} = -\sin x. \end{aligned}$$

□

15: 4.12.07

Beispiel 9.9 Die *Betragsfunktion* $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}$, $\text{abs}(x) = |x|$ ist für $x \neq 0$ differenzierbar,

$$\text{abs}'(x) = \begin{cases} -1 & \text{für } x < 0, \\ 1 & \text{für } x > 0. \end{cases}$$

Im Punkt $x = 0$ ist sie nicht differenzierbar. Dies ist das einfachste Beispiel einer Funktion, die stetig aber nicht differenzierbar ist. Offenbar existieren in 0 jedoch die „einseitigen Ableitungen“, von links = -1, von rechts = +1. □

Um kompliziertere (insbesondere zusammengesetzte) Funktionen zu differenzieren, werden einige *Differentiationsregeln* benötigt:

Satz 9.10 Seien $f, g : I \rightarrow \mathbb{R}$ differenzierbar (in $x_0 \in I$). Dann gilt:

a) $af + bg$ ist differenzierbar (in x_0) und es gilt

$$(af + bg)'(x) = af'(x) + bg'(x) \quad (\text{Linearität der Differentiation}),$$

b) $f \cdot g$ ist differenzierbar (in x_0) und es gilt

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x) \quad (\text{Produktregel}),$$

c) dort wo $g(x) \neq 0$ ist, gilt $\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)}$ und

$$\left(\frac{f}{g}\right)'(x) = \frac{f'g - fg'}{g^2} \quad (\text{Quotientenregel}).$$

Beweis. a) Dies ist offensichtlich eine Folge der Linearität des Limes (\ddot{U}).

b) Offenbar folgt aus der Stetigkeit von f und der Differenzierbarkeit von f und g für $x \rightarrow x_0$

$$\begin{aligned} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x)(g(x) - g(x_0)) + (f(x) - f(x_0))g(x_0)}{x - x_0} \\ &= f(x) \frac{g(x) - g(x_0)}{x - x_0} + \frac{f(x) - f(x_0)}{x - x_0} g(x_0) \rightarrow f(x_0)g'(x_0) + f'(x_0)g(x_0). \end{aligned}$$

c) Entsprechend folgt zunächst

$$\frac{1}{x - x_0} \left(\frac{1}{g(x)} - \frac{1}{g(x_0)} \right) = \frac{g(x_0) - g(x)}{x - x_0} \frac{1}{g(x)g(x_0)} \rightarrow -\frac{g'(x_0)}{g^2(x_0)}$$

und daraus mit Hilfe der Produktregel

$$\left(\frac{f}{g} \right)' = \left(f \cdot \frac{1}{g} \right)' = f' \frac{1}{g} + f \left(\frac{1}{g} \right)' = f' \frac{1}{g} - f \frac{g'}{g^2} = \frac{f'g - fg'}{g^2}.$$

■

Satz 9.11 Seien $f : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ differenzierbar und $g(J) \subset I$. Dann ist auch $f \circ g : J \rightarrow \mathbb{R}$ differenzierbar mit

$$(f \circ g)'(x) = f'(g(x))g'(x) \quad (\text{Kettenregel})$$

(das entsprechende gilt punktweise, wenn g in x_0 und f in $g(x_0)$ differenzierbar sind). Merksregel: Äußere Ableitung \times innere Ableitung.

Beweis. Sei zunächst $g'(x_0) \neq 0$. Da $g'(x_0) = \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0}$ ist, muß dann $g(x) \neq g(x_0)$ sein für $x \neq x_0$ und x nahe bei x_0 (und nur solche x brauchen wir). Also gilt für x nahe bei x_0

$$\frac{f(g(x)) - f(g(x_0))}{x - x_0} = \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)} \frac{g(x) - g(x_0)}{x - x_0}.$$

Wegen $g(x) \rightarrow g(x_0)$ für $x \rightarrow x_0$ und der Differenzierbarkeit von g in x_0 und f in $g(x_0)$ konvergiert dieser Ausdruck für $x \rightarrow x_0$ gegen $f'(g(x_0))g'(x_0)$.

Sei nun $g'(x_0) = 0$. Da f im Punkt $g(x_0)$ differenzierbar ist gilt für y nahe $g(x_0)$

$$\left| \frac{f(y) - f(g(x_0))}{y - g(x_0)} \right| \leq L, \quad \text{also} \quad \left| f(y) - f(g(x_0)) \right| \leq L|y - g(x_0)|$$

und somit

$$\left| \frac{f(g(x)) - f(g(x_0))}{x - x_0} \right| \leq \frac{L|g(x) - g(x_0)|}{|x - x_0|} \rightarrow 0 \quad \text{für } x \rightarrow x_0,$$

d. h. es gilt auch in diesem Fall $(f \circ g)'(x_0) = 0 = f'(g(x_0))g'(x_0)$. ■

Folgerung: Unter entsprechenden Voraussetzungen gilt z. B.

$$(f \circ g \circ h)'(x) = f'(g(h(x)))g'(h(x))h'(x),$$

und entsprechend für mehrere hintereinander ausgeführte Funktionen.

Satz 9.12 Ist $f : I \rightarrow \mathbb{R}$ differenzierbar und streng monoton mit $f'(x) \neq 0$, so ist $f^{-1} : f(I) \rightarrow \mathbb{R}$ differenzierbar mit

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} \quad \text{bzw.} \quad (f^{-1})'(f(x)) = \frac{1}{f'(x)}.$$

Beweis. Dies folgt sofort aus der Tatsache, dass bei der Spiegelung an der Diagonalen durch den 1. und 3. Quadranten (und dies führt den Graphen der Funktion in den der Umkehrfunktion über) die Katheten der Steigungsdreiecke vertauscht werden. Damit geht die Steigung (= Quotient der Kathetenlängen) in den Kehrwert über. ■

Nun sind wir in der Lage, die uns bisher bekannten Funktionen zu differenzieren:

Beispiel 9.13 Für $f_n(x) := x^n$ ($n \in \mathbb{N}_0$) gilt

$$f'_0(x) = 0, \quad f'_n(x) = nx^{n-1} \quad \text{für } n \in \mathbb{N}.$$

Beweis durch Induktion: $n = 0$, $n = 1$ o. k. (auch $n = 2$ schon bekannt),

$n \Rightarrow n + 1$: $f'_{n+1}(x) = (x f_n)'(x) = 1 \cdot f_n(x) + x f'_n(x) = x^n + x n x^{n-1} = (n + 1)x^n$.⁶

Also gilt für ein Polynom $p(x) = \sum_{j=0}^n a_j x^j$

$$p'(x) = \sum_{j=1}^n a_j j x^{j-1} = \sum_{j=0}^{n-1} a_{j+1} (j+1) x^j.$$

□

⁶Man bekommt die Ableitung von x^n auch sofort aus

$$\lim_{z \rightarrow x} \frac{z^n - x^n}{z - x} = \frac{(z-x)(z^{n-1} + z^{n-2} + \dots + x^{n-1})}{z-x} = nx^{n-1}$$

Beispiel 9.14 Für $g_n(x) = \frac{1}{x^n} = x^{-n}$ gilt $g'_n(x) = -n \frac{1}{x^{n+1}} = -nx^{-n-1}$. Das erhält man sofort aus $g_n(x) = \frac{1}{f_n(x)}$ mit Hilfe obiger Regel. Für alle x^n ($n \in \mathbb{Z}$) gilt die gleiche Formel für die Ableitung⁷. \square

Beispiel 9.15 $\ln : (0, \infty) \rightarrow \mathbb{R}$, Umkehrfunktion von $\exp : \mathbb{R} \rightarrow \mathbb{R}$. Mit $\exp' = \exp$ folgt

$$\ln'(x) = (\exp^{-1})'(x) = \frac{1}{\exp'(\ln x)} = \frac{1}{\exp(\ln x)} = \frac{1}{x}.$$

 \square

Beispiel 9.16 $\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, Umkehrfunktion von $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1]$. In den Randpunkten $\pm \frac{\pi}{2}$ gilt $\sin' \left(\pm \frac{\pi}{2}\right) = \cos \left(\pm \frac{\pi}{2}\right) = 0$. Deshalb können wir hier nur

$$\arcsin : (-1, 1) \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \text{ als Umkehrfunktion von } \sin : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow (-1, 1)$$

differenzieren. In der folgenden Rechnung beachte man, dass wegen $\cos x > 0$ für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ die positive Quadratwurzel die richtige Wahl ist:

$$\begin{aligned} \arcsin'(x) &= (\sin^{-1})'(x) = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} \\ &= \frac{1}{+\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{+\sqrt{1 - x^2}}. \end{aligned}$$

 \square

Beispiel 9.17 $\arccos : (-1, 1) \rightarrow (0, \pi)$ als Umkehrfunktion von $\cos : (0, \pi) \rightarrow (-1, 1)$. Wegen $\sin x > 0$ für $x \in (0, \pi)$ ist hier wieder $+\sqrt{\quad}$ die richtige Wahl:

$$\arccos'(x) = \frac{1}{-\sin(\arccos x)} = \frac{-1}{+\sqrt{1 - \cos^2(\arccos x)}} = \frac{-1}{\sqrt{1 - x^2}}.$$

 \square

Beispiel 9.18 $\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ als Umkehrfunktion von $\tan = \frac{\sin}{\cos} : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}$.

Dazu müssen wir erst nachtragen:

$$\tan'(x) = \left(\frac{\sin}{\cos}\right)'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

⁷Auch dies erhält man elementar aus

$$\frac{1}{z-x} \left(\frac{1}{z^n} - \frac{1}{x^n}\right) = \frac{1}{z-x} \frac{x^n - z^n}{z^n x^n} = \frac{x^{n-1} + x^{n-2} + \dots + z^n}{z^n x^n}$$

Damit folgt

$$\begin{aligned} \arctan'(x) &= (\tan^{-1})'(x) = \frac{1}{\tan'(\arctan x)} = \cos^2(\arctan x) \\ &= \left(\frac{\cos^2(\arctan x) + \sin^2(\arctan x)}{\cos^2(\arctan x)} \right)^{-1} = \left(1 + \tan^2(\arctan x) \right)^{-1} \\ &= \frac{1}{1+x^2}. \end{aligned}$$

□

Beispiel 9.19 Mit $\cot'(x) = -\frac{1}{\sin^2 x}$ berechnet man für $\operatorname{arc cot} : \mathbb{R} \rightarrow (0, \pi)$ als Umkehrfunktion von $\tan : (0, \pi) \rightarrow \mathbb{R}$ völlig analog

$$\begin{aligned} \operatorname{arc cot}'(x) &= \frac{1}{\cot'(\operatorname{arc cot} x)} = -\sin^2(\operatorname{arc cot} x) \\ &= -\left(\frac{\sin^2(\operatorname{arc cot} x) + \cos^2(\operatorname{arc cot} x)}{\sin^2(\operatorname{arc cot} x)} \right)^{-1} = -\left(1 + \cot^2(\operatorname{arc cot} x) \right)^{-1} \\ &= -\frac{1}{1+x^2}. \end{aligned}$$

□

Beispiel 9.20 *Allgemeine Potenzfunktionen:*

- $f_1 : (0, \infty) \rightarrow \mathbb{R}, x \mapsto x^a = \exp(a \ln x), f_1'(x) = \frac{a}{x} \exp(a \ln x) = ax^{a-1}.$
- $f_2 : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto a^x = \exp(x \ln a) \ (a > 0), f_2'(x) = (\ln a) \exp(x \ln a) = a^x \ln a.$
- $f_3 : (0, \infty) \rightarrow \mathbb{R}, x \mapsto x^x = \exp(x \ln x), f_3'(x) = (1 + \ln x)x^x.$

Man beachte, dass also auch für die allgemeine Potenz $x \mapsto x^a$ die bereits bekannte Formel ax^{a-1} für die Ableitung gilt. □

10 Eigenschaften differenzierbarer Funktionen

Wir wissen bisher nur, dass differenzierbare Funktionen auf jeden Fall stetig sind. Mit Hilfe verschiedener Versionen des Mittelwertsatzes werden weitere wichtige Eigenschaften folgen.

10.1 Mittelwertsätze

Zunächst beweisen wir eine notwendige Bedingung für das Vorliegen eines Extremums (Maximum oder Minimum) einer differenzierbaren Funktion:

Satz 10.1 *Sei I ein offenes Intervall. Hat $f : I \rightarrow \mathbb{R}$ bei $x_0 \in I$ ein lokales Extremum und ist f bei x_0 differenzierbar, so ist $f'(x_0) = 0$.*

Beweis. Liegt bei x_0 ein lokales Maximum vor, so gilt für $x > x_0$, x nahe bei x_0 ,

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0, \quad \text{also} \quad f'(x_0) = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

und entsprechend für $x < x_0$, x nahe bei x_0 ,

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0, \quad \text{also} \quad f'(x_0) = \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

Insgesamt folgt also $f'(x_0) = 0$. Entsprechend verfährt man bei einem lokalen Minimum. ■

16: 6.12.07 Ein Spezialfall des Mittelwertsatzes ist der

Satz 10.2 (Satz von Rolle) *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und in (a, b) differenzierbar und $f(a) = f(b)$. Dann existiert (mindestens) ein $\xi_0 \in (a, b)$ mit $f'(\xi_0) = 0$.*

Beweis. Ist $f(x) \equiv f(a) = f(b)$ im ganzen Intervall, so ist $f'(x) = 0$ für alle $x \in (a, b)$.

Andernfalls gibt es $x \in (a, b)$ mit $f(x) > f(a)$ oder $f(x) < f(a)$. Also nimmt f in einem $\xi_0 \in (a, b)$ sein Maximum oder sein Minimum an. Nach obigem Satz ist dort $f'(\xi_0) = 0$. ■

Satz 10.3 (Mittelwertsatz) *Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und in (a, b) differenzierbar, so gibt es (mindestens) ein $\xi_0 \in (a, b)$ mit*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi_0) \quad \text{bzw.} \quad f(b) - f(a) = (b - a)f'(\xi_0).$$

D.h. es gibt mindestens einen Punkt in (a, b) , wo die Tangente die gleiche Steigung hat wie die Sekante zwischen $(a, f(a))$ und $(b, f(b))$.

Beweis. Wir betrachten die Funktion

$$g : [a, b] \rightarrow \mathbb{R}, \quad g(x) := f(x) - (x - a) \frac{f(b) - f(a)}{b - a} \quad \text{für } x \in [a, b].$$

g ist in $[a, b]$ stetig, in (a, b) differenzierbar, und es gilt $g(a) = f(a) = g(b)$. Also gibt es nach dem Satz von Rolle ein ξ_0 mit

$$0 = g'(\xi_0) = f'(\xi_0) - \frac{f(b) - f(a)}{b - a}.$$

Das ist die Behauptung. ■

Korollar 10.4 Sei $f : [a, b] \in \mathbb{R}$ in $[a, b]$ stetig und in (a, b) differenzierbar.

- a) f ist genau dann monoton wachsend (bzw. fallend), wenn für alle $x \in (a, b)$ gilt $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$).
- b) Ist $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$, so ist f strikt wachsend (bzw. fallend).

Beweis. a) \Leftarrow : Für $a \leq x < y \leq b$ existiert ein $\xi \in (x, y)$ mit

$$f(y) - f(x) = (y - x)f'(\xi).$$

Ist $f'(\xi) \geq 0$, so folgt, dass f wachsend ist; ist $f'(\xi) \leq 0$, so ist f fallend.

\Rightarrow : Sei f wachsend. Gibt es ein ξ mit $f'(\xi) < 0$, d. h. der Differenzenquotient $\frac{1}{z - \xi}(f(z) - f(\xi))$ ist negativ für z nahe bei ξ , ein Widerspruch dazu, dass f wachsend ist.

Entsprechend wird die Aussage für fallend bewiesen.

- b) Dies folgt sofort aus dem Mittelwertsatz. ■

Korollar 10.5 Sei $f : (a, b) \rightarrow \mathbb{R}$. f ist genau dann differenzierbar mit Ableitung $f'(x) = 0$ für alle $x \in (a, b)$, wenn f konstant ist.

Beweis. \Rightarrow : Folgt direkt aus dem Mittelwertsatz.

\Leftarrow : Ist offensichtlich, da die konstante Funktion Ableitung 0 hat. ■

Satz 10.6 (Verallgemeinerter Mittelwertsatz) Seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig und in (a, b) differenzierbar, $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann gibt es ein $\xi_0 \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi_0)}{g'(\xi_0)}.$$

(Für $g(x) = x$ ist dies der obige Mittelwertsatz.)

Beweis. ⁸ Wir betrachten

$$h(x) := f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a))$$

(dabei beachte man, dass auf Grund der Voraussetzung $g'(x) \neq 0$ für alle $x \in (a, b)$ und des Mittelwertsatzes

$$g(b) - g(a) = (b - a)g'(\xi_0) \neq 0$$

gilt).

h ist in (a, b) differenzierbar und es gilt $h(a) = 0 = h(b)$. Nach dem Satz von Rolle gibt es also ein ξ_0 mit

$$0 = h'(\xi_0) = f'(\xi_0) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(\xi_0),$$

das ist die Behauptung. ■

10.2 Höhere Ableitung und der Satz von Taylor

Ist $f : I \rightarrow \mathbb{K}$ differenzierbar, so ist $f' : I \rightarrow \mathbb{K}$ wieder eine Funktion, die eventuell erneut differenzierbar sein kann. Die Ableitung $(f')'$ von f' nennt man die *zweite Ableitung* von f und schreibt dafür f'' oder $f^{(2)}$. Entsprechend definiert man die *dritte Ableitung* $f''' = f^{(3)}$ und allgemein die *n-te Ableitung* $f^{(n)}$, $f^{(n+1)} := (f^{(n)})'$ für $n \geq 1$. Man schreibt dafür auch

$$f^{(n)} = \frac{d^n f}{dx^n} = \frac{d^n}{dx^n} f = \left(\frac{d}{dx}\right)^n f.$$

Satz 10.7 (Leibniz-Regel) Sind f und g auf I n mal differenzierbar, so ist auch $f \cdot g$ n mal differenzierbar und es gilt (mit $f^{(0)} = f$)

$$(f \cdot g)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)} \quad (\text{Leibniz-Regel}).$$

⁸Man kann den Satz nicht dadurch beweisen, dass man die Aussage des Mittelwertsatzes

$$f(b) - f(a) = (b - a)f'(\xi_0), \quad g(b) - g(a) = (b - a)g'(\xi_0)$$

durcheinander dividiert, da hier davon auszugehen ist, dass es sich um verschiedene ξ_0 handelt.

Beweis. Induktion: $n = 0$: $(f \cdot g)^{(0)} = f \cdot g = \sum_{k=0}^0 \binom{0}{k} f^{(k)} g^{(0-k)}$.

$n \Rightarrow n + 1$: Aus der Formel für die n -te Ableitung folgt

$$\begin{aligned} (f \cdot g)^{(n+1)} &= \frac{d}{dx} (f \cdot g)^{(n)} = \frac{d}{dx} \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)} \\ &= \sum_{k=0}^n \binom{n}{k} \left\{ f^{(k+1)} g^{(n-k)} + f^{(k)} g^{(n+1-k)} \right\} \\ &= \binom{n}{0} f^{(0)} g^{(n+1)} + \left[\binom{n}{0} + \binom{n}{1} \right] f^{(1)} g^{(n)} + \dots + \left[\binom{n}{n-1} + \binom{n}{n} \right] f^{(n)} g^{(1)} \\ &\quad + \binom{n}{n} f^{(n+1)} g^{(0)} \\ &= \binom{n+1}{0} f^{(0)} g^{(n+1)} + \binom{n+1}{1} f^{(1)} g^{(n)} + \dots + \binom{n+1}{n} f^{(n)} g^{(1)} \\ &\quad + \binom{n+1}{n+1} f^{(n+1)} g^{(0)} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)}. \end{aligned}$$

9

■

Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt *stetig differenzierbar*, wenn f in I differenzierbar ist und f' in I stetig ist. f heißt *n mal stetig differenzierbar*, wenn es n mal differenzierbar ist und $f^{(n)}$ stetig ist (die vorhergehenden Ableitungen sind dann auch stetig). Mit $C^n(I)$ bezeichnen wir die Menge (offenbar ein Vektorraum) der n mal stetig differenzierbaren Funktionen auf I .

Ein für die Analysis fundamentaler Satz ist der folgende Satz von Taylor. Die Idee ist dabei, eine mindestens $n + 1$ mal differenzierbare Funktion durch ein Polynom vom Grad n zu approximieren, dessen Funktionswert und Ableitungen der Ordnung $\leq n$ im Entwicklungspunkt x_0 mit denen der Funktion übereinstimmen. Der Satz von Taylor ermöglicht die beliebig genaue Berechnung der Funktionswerte der wichtigsten Funktionen (z. B. \exp , \cos , \sin , ...). Für ein Polynom der Ordnung n liefert dieser Prozess immer genau dieses Polynom.

Satz 10.8 (Taylor) Sei $f \in C^{n+1}(I)$, $x_0 \in I$. Das Taylorpolynom der Ordnung n von f

⁹Man kann die Leibniz-Regel auch auf die (sehr ähnlich aussehende) binomische Formel zurückführen, wenn man beachtet, dass die Produktregel in der Form $(f \cdot g)' = (D_1 + D_2)(f \cdot g)$ geschrieben werden kann, wobei D_1 und D_2 nur den ersten bzw. zweiten Faktor im Produkt differenziert. Anwendung der binomischen Formel auf $(f \cdot g)^{(n)} = (D_1 + D_2)^n (f \cdot g)$ liefert die Leibniz-Regel.

zum Entwicklungspunkt x_0 sei

$$T_{f,n}(x) := f(x_0) + f'(x_0)(x - x_0) + \dots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} = \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j.$$

Dann gilt für $x \in I$

$$f(x) = T_{f,n}(x) + R_n(x)$$

mit dem Restglied $R_n(x)$

$$R_n(x) = f^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n+1)!}$$

mit einem ξ zwischen x_0 und x , d. h. $\xi = x_0 + \vartheta(x - x_0)$ mit $0 < \vartheta < 1$. (Als Merkgel kann gelten: Das Restglied $R_n(x)$ sieht so aus wie das $(n+1)$ -te Glied aussehen würde, nur ist in $f^{(n+1)}(\cdot)$ der Punkt x_0 durch ξ ersetzt. Leider ist in dieser Formel nicht bekannt, welches ξ zu wählen ist. Um das Restglied abzuschätzen muß also eine für alle ξ zwischen x_0 und x gültige Abschätzung gesucht werden.)

Beweis. (Der folgende Beweis ist zwar einfach und kurz, wirkt aber etwas undurchsichtig; einen natürlichen Beweis, der auch eine weitere Form des Restglieds liefert, kann mit Hilfe der Integralrechnung gegeben werden.) Für festes $x \in I$ sei

$$F(t) := f(x) - \sum_{j=0}^n f^{(j)}(t) \frac{(x-t)^j}{j!}, \quad G(t) := \frac{(x-t)^{n+1}}{(n+1)!}.$$

Man rechnet leicht nach (Ableitung nach t ; die restlichen Terme annullieren sich)

$$F'(t) = -f^{(n+1)}(t) \frac{(x-t)^n}{n!}, \quad G'(t) = -\frac{(x-t)^n}{n!}.$$

Für $t = x_0$ bzw. x liefert der verallgemeinerte Mittelwertsatz ein ξ zwischen x_0 und x mit (beachte $F(x) = G(x) = 0$)

$$\frac{F(x_0)}{G(x_0)} = \frac{F(x) - F(x_0)}{G(x) - G(x_0)} = \frac{F'(\xi)}{G'(\xi)} = f^{(n+1)}(\xi), \quad (*)$$

also

$$F(x_0) = G(x_0) f^{(n+1)}(\xi),$$

und somit

$$f(x) - \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j = F(x_0) = G(x_0) f^{(n+1)}(\xi) = f^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n+1)!}.$$

folgt daraus die Behauptung. ■

Folgerung 10.9 Ist f ein Polynom vom Grad m , $f(x) = \sum_{j=0}^m c_j x^j$, so gilt $T_{f,n}(x) = \sum_{j=0}^n c_j x^j$ für $n < m$ und $T_{f,n} = f$ für $n \geq m$.

Ist eine Funktion beliebig oft differenzierbar, so kann man aus dem Satz von Taylor eine Potenzreihe gewinnen, die formale *Taylorreihe*

$$T_f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j.$$

Es ist aber zunächst nicht klar, ob diese Reihe konvergiert, und wenn sie konvergiert, ob sie gegen $f(x)$ konvergiert. Sei konvergiert offenbar nur dann (bzw. nur dort) gegen die vorgegebene Funktion, wenn (bzw. wo) das Restglied gegen 0 konvergiert. Dass dies durchaus nicht immer der Fall ist, zeigt das folgende Beispiel:

Beispiel 10.10 Sei $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = \begin{cases} 0 & \text{für } x = 0, \\ \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0. \end{cases}$

Außerhalb 0 ist f offensichtlich beliebig oft differenzierbar, und die n -te Ableitung hat die Gestalt

$$f^{(n)}(x) = p_n\left(\frac{1}{x}\right) \exp\left(-\frac{1}{x^2}\right)$$

mit geeigneten Polynomen p_n . Für $x \rightarrow 0$ konvergieren alle diese Funktionen gegen 0.

Im Nullpunkt ist zunächst

$$f'(0) = \lim_{x \rightarrow 0} \frac{1}{x} \left(\exp\left(-\frac{1}{x^2}\right) - 0 \right) = 0$$

und so folgt weiter

$$f^{(n)}(0) = \lim_{x \rightarrow 0} \frac{1}{x} \left(p_{n-1}\left(\frac{1}{x}\right) \exp\left(-\frac{1}{x^2}\right) - 0 \right) = 0 \text{ für alle } n.$$

Die Taylorreihe von f zum Entwicklungspunkt 0 enthält also nur Nullen, konvergiert also gegen die Nullfunktion. Also konvergiert die Taylorreihe dieser Funktion nur im Entwicklungspunkt gegen die Funktion. \square

17:11.12.07

Beispiel 10.11 Taylorentwicklung der Exponentialfunktion Für \exp gilt bekanntlich $\exp^{(n)} = \exp$. Für die Taylorentwicklung mit Entwicklungspunkt $x_0 = 0$ erhält man also

$$\exp(x) = \sum_{j=0}^n \frac{1}{j!} x^j + \frac{x^{n+1}}{(n+1)!} \exp(\vartheta x) \quad \text{mit } 0 < \vartheta < 1.$$

Es gilt $\frac{x^{n+1}}{(n+1)!} \rightarrow 0$ für alle $x \in \mathbb{R}$, d. h. \exp ist als Taylorreihe darstellbar:

$$\exp(x) = \sum_{j=0}^{\infty} \frac{1}{j!} x^j \quad \text{für alle } x \in \mathbb{R}.$$

Will man z. B. \exp in $[-1, 1]$ auf 2 Stellen genau berechnen, so genügt es, das Taylorpolynom 5-ten Grades zu benutzen,

$$p_5(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120},$$

denn für das Restglied gilt dann

$$\frac{x^6}{720} \cdot \exp(\vartheta x) < \frac{e}{720} < 0,0038.$$

Für $x \in [-1, 0]$ ist die Abschätzung noch verbesserungsfähig: Da die Reihe dann alternierend ist, ist der Fehler höchstens gleich dem ersten weggelassenen Glied, also $\leq \frac{|x|^6}{720} \leq \frac{1}{720} \leq 0,0014$. \square

Beispiel 10.12 Taylorentwicklung der Logarithmusfunktion

Wir betrachten $f(x) = \ln(1+x)$ für $x > -1$. Es gilt

$$f'(x) = \frac{1}{1+x}, \quad f''(x) = \frac{-1}{(1+x)^2}, \quad f'''(x) = \frac{2}{(1+x)^3}, \dots, \quad f^{(n)}(x) = (-1)^{n-1} \frac{(n-1)!}{(1+x)^n},$$

also

$$f(0) = 0, \quad f'(0) = 1, \dots, \quad f^{(n)}(0) = (-1)^{n-1} (n-1)!.$$

Damit folgt für die Taylorentwicklung (mit $0 < \vartheta < 1$)

$$\ln(1+x) = f(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - + \dots, + \frac{(-1)^{n-1}}{n}x^n + \frac{(-1)^n}{n+1} \frac{x^{n+1}}{(1+\vartheta x)^{n+1}}.$$

Für $0 \leq x \leq 1$ sieht man sofort, dass das Restglied für $n \rightarrow \infty$ gegen 0 konvergiert. Dort wird also \ln tatsächlich durch die Taylorreihe dargestellt. Insbesondere ist

$$\ln 2 = \ln(1+1) = \sum_{j=1}^{\infty} (-1)^{j-1} \frac{1}{j} \sim 0,736544\dots$$

Der Konvergenzradius ist also mindestens 1. Für $x = -1$ ist die Reihe sicher divergent (= harmonische Reihe). Also ist der Konvergenzradius der Taylorreihe = 1, d. h. die Reihe konvergiert für $x \in (-1, 1]$, aber wogegen konvergiert sie in $(-1, 0)$?. Das Restglied ist für $-1 < x < 0$ nicht leicht abzuschätzen. Man geht deshalb einen anderen Weg um zu zeigen,

dass die Taylorreihe auch für $-1 < x < 0$ gegen $\ln(1+x)$ konvergiert. Dazu wird aber etwas Integralrechnung benötigt. ¹⁰ \square

Beispiel 10.13 Taylorentwicklung von Cosinus und Sinus

$$\cos^{(n)}(x) = \begin{cases} (-1)^k \cos x & n=2k, \\ (-1)^{k+1} \sin x & n=2k+1, \end{cases} \quad \sin^{(n)}(x) = \begin{cases} (-1)^k \sin x & n=2k, \\ (-1)^k \cos x & n=2k+1. \end{cases}$$

Daraus folgt mit $x_0 = 0$

$$\begin{aligned} \cos x &= \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + (-1)^{n+1} \frac{x^{2n+1}}{(2n+1)!} \sin(\vartheta_1 x), \\ \sin x &= \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + (-1)^{n+1} \frac{x^{2n+2}}{(2n+2)!} \cos(\vartheta_2 x). \end{aligned}$$

Da in diesen Entwicklungen das jeweils nächste Glied verschwindet, kann allerdings auch das nächste Restglied gewählt werden, also

$$(-1)^{n+1} \frac{x^{2n+2}}{(2n+2)!} \cos(\tilde{\vartheta}_1 x) \quad \text{bzw.} \quad (-1)^{n+2} \frac{x^{2n+3}}{(2n+3)!} \sin(\tilde{\vartheta}_2 x).$$

Den Restgliedern ist anzusehen, dass sie für alle $x \in \mathbb{R}$ für $n \rightarrow \infty$ gegen 0 konvergieren, d. h. die Taylorreihen stellen Cosinus und Sinus für alle $x \in \mathbb{R}$ dar.

Um z. B. $\sin x$ für $0 < x < 10^{-1} \sim 6^\circ$ mit einem Fehler $< 10^{-6}$ zu berechnen reichen die ersten zwei Terme: $\sin x \sim x - \frac{x^3}{6}$. \square

10.3 Extremwerte differenzierbarer Funktionen

Wir können nun (notwendige und hinreichende) Kriterien für das Vorliegen von **Extrema** (Maxima und Minima) angeben.

Satz 10.14 (Bedingungen für Extrema) Sei $f : (a, b) \rightarrow \mathbb{R}$ stetig, $x_0 \in (a, b)$.

¹⁰Man benutzt $\ln'(x) = \frac{1}{x}$ und somit für $|x| < 1$ (vgl. Hauptsatz der Differential- und Integralrechnung)

$$\begin{aligned} \ln(1+x) &= \ln(1+x) - \ln 1 = \int_0^x \frac{1}{1+t} dt = \int_0^x \sum_{n=0}^{\infty} (-t)^n dt \\ &\stackrel{!}{=} \sum_{n=0}^{\infty} \int_0^x (-t)^n dt = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{(n+1)} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}. \end{aligned}$$

An der Stelle $\stackrel{!}{=}$ wird die gleichmäßige Konvergenz der Reihe für $|t| \leq x < 1$ benutzt, und damit die Vertauschung des Grenzübergangs mit dem Integral (hierzu bald mehr!).

Entsprechend wird man die Taylorreihe von \arctan und arccot erhalten, indem man ihre Ableitungen $\pm \frac{1}{1+x^2}$ als geometrische Reihe schreibt und integriert.

a) Notwendige Bedingungen:

- (i) Liegt bei x_0 ein lokales Extremum vor (Maximum oder Minimum) und ist f in x_0 differenzierbar, so ist $f'(x_0) = 0$.
- (ii) Liegt bei x_0 ein lokales Extremum vor und ist f in x_0 zwei mal differenzierbar (d. h. f muß zumindest in einer ganzen Umgebung von x_0 ein mal differenzierbar sein und f' in x_0 differenzierbar), so ist
 - $f''(x_0) \leq 0$, falls es sich um ein Maximum handelt,
 - $f''(x_0) \geq 0$, falls es sich um ein Minimum handelt.

Das Beispiel $f(x) = x^3$ zeigt, dass beide Bedingungen keinesfalls hinreichend sind.

b) Hinreichende Bedingung: Ist f in einer Umgebung von x_0 differenzierbar mit $f'(x_0) = 0$ und in x_0 zweimal differenzierbar mit

- $f''(x_0) < 0$, so liegt bei x_0 ein striktes lokales Maximum vor, bzw.
- $f''(x_0) > 0$, so liegt bei x_0 ein striktes lokales Minimum vor.

Beweis. a) (i) Ist bereits bekannt, vgl. Satz 10.1.

(ii) Folgt direkt aus Teil b: Ist nämlich $f''(x_0) > 0$, so liegt ein Minimum vor, ist $f''(x_0) < 0$, so liegt ein Maximum vor.

b) $f''(x_0) < 0$: Da $f'(x_0) = 0$ ist und

$$0 > f''(x_0) = \lim_{x \rightarrow x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f'(x)}{x - x_0}$$

gilt, ist

- $f'(x) < 0$ für x nahe x_0 und $x > x_0$
- $f'(x) > 0$ für x nahe x_0 und $x < x_0$.

Das heißt, dass f (in einer Umgebung von x_0) links von x_0 strikt wächst und rechts von x_0 strikt fällt. Also liegt in x_0 ein striktes lokales Maximum vor.

Entsprechend liefert $f''(x_0) > 0$ ein striktes lokales Minimum. ■

Bemerkung 10.15 a) Ist $f'(x_0) = 0$ und $f''(x_0) = 0$, so liefert der obige Satz keine Entscheidung. Ist in diesem Fall $f'''(x_0) \neq 0$, so liegt kein Extremwert vor. Ist $f'''(x_0) = 0$

und $f^{(4)}(x_0) > 0$ bzw. < 0 , so liegt bei x_0 ein striktes lokales Minimum bzw. Maximum vor. Ist entsprechend

$$f'(x_0) = \dots = f^{(2n-1)}(x_0) = 0 \quad \text{und} \quad f^{(2n)}(x) > 0 \quad \text{bzw.} \quad < 0,$$

so liegt bei x_0 ein striktes lokales Minimum bzw. Maximum vor.

b) Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und hinreichend oft differenzierbar, so liefert der obige Satz nur die lokalen Extrema im Inneren (a, b) des Intervalls $[a, b]$. Um das globale Maximum bzw. Minimum auf $[a, b]$ zu finden muß man das größte Maximum bzw. das kleinste Minimum mit den Randwerten $f(a)$ und $f(b)$ vergleichen.

Beispiel 10.16 $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^n e^x$, $n \in \mathbb{N}$. Man erkennt: für $x \rightarrow \infty$ geht f in jedem Fall gegen ∞ , für $x \mapsto -\infty$ gegen 0. Weiter gilt $f(0) = 0$, für positive x ist $f(x) > 0$ und für negative x ist $f(x) > 0$ für gerades n , $f(x) < 0$ für ungerades n . Daraus kann bereits einiges über das Verhalten von f abgelesen werden.

Mögliche Extremwerte liegen dort, wo $f'(x) = 0$ ist.

$$f'(x) = (nx^{n-1} + x^n)e^x, \quad f''(x) = (n(n-1)x^{n-2} + 2nx^{n-1} + x^n)e^x.$$

Für $n = 1$ ist $x = -1$ die einzige Nullstelle von f' . Dann ist $f''(-1) = e^{-1} > 0$, d. h. bei $x = -1$ liegt ein striktes lokales Minimum vor.

Für $n \geq 2$ hat f' die Nullstellen $x = 0$ und $x = -n$.

Ist $n \geq 2$ ungerade, folgt bereits aus obiger Überlegung, dass bei 0 kein Extremum vorliegen kann, bei $x = -n$ muß ein globales Minimum vorliegen; tatsächlich gilt (beachte $n \geq 3$)

$$\begin{aligned} f''(-n) &= \left(n(n-1)(-n)^{n-2} + 2n(-n)^{n-1} + (-n)^n \right) e^{-n} \\ &= \left(-(n-1)n^{n-1} + 2nn^{n-1} - n^n \right) e^{-n} = \left(n^n - (n-1)n^{n-1} \right) e^{-n} > 0. \end{aligned}$$

Ist $n \geq 2$ gerade so ist auf Grund der obigen Überlegung 0 ein striktes lokales Minimum, obwohl für $n \geq 4$ $f''(0) = 0$ ist (für $n = 2$ ist $f''(0) = 2 > 0$) (die Berechnung der 3. und 4. Ableitung wollen wir uns sparen). Für $x = -n$ gilt in diesem Fall

$$f''(-n) = \left(n(n-1)n^{n-2} - 2nn^{n-1} + n^n \right) e^{-n} = \left((n-1)n^{n-1} - n^n \right) e^{-n} < 0,$$

dort liegt also ein striktes lokales Maximum vor. \square

10.4 Die Regeln von de l'Hospital

Häufig muß man Grenzwerte der Form $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ bestimmen, wobei $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ gilt. Da führt oft der folgende Satz zum Ziel.

Satz 10.17 (Regel von de l'Hospital) a) Seien $f, g : (a, b) \rightarrow \mathbb{R}$ differenzierbar mit $\lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a+} g(x) = 0$ und $g'(x) \neq 0$ für x in einer rechtsseitigen Umgebung von a . Dann gilt

$$\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)},$$

falls der letzte Grenzwert existiert. Entsprechendes gilt für den linksseitigen Grenzwert.

b) Seien $f, g : (a, \infty) \rightarrow \mathbb{R}$ differenzierbar mit $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = 0$ und $g'(x) \neq 0$ für x nahe ∞ . Dann gilt

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)},$$

falls der letzte Grenzwert existiert. Entsprechendes gilt für den Grenzwert $x \rightarrow -\infty$.

Zusatz: Die entsprechenden Aussagen gelten für $\lim f(x) = \pm\infty$ und $\lim g(x) = \pm\infty$. Der Beweis ist wesentlich komplizierter (vgl. z. B. W. Walter, *Analysis 1*, §10.11).

Beweis. a) Mit $f(a) := 0$ und $g(a) := 0$ erfüllen f und g in $[a, x]$ für jedes $x \in (a, b)$ die Voraussetzung des verallgemeinerten Mittelwertsatzes, d. h. es gilt

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\vartheta_x)}{g'(\vartheta_x)} \quad \text{mit einem } \vartheta_x \in (a, x).$$

Für $x \rightarrow a$ gilt auch $\vartheta_x \rightarrow a$ und somit

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(\vartheta_x)}{g'(\vartheta_x)} = \lim_{y \rightarrow a} \frac{f'(y)}{g'(y)}.$$

b) Wegen $\lim_{x \rightarrow 0+} f\left(\frac{1}{x}\right) = \lim_{x \rightarrow 0+} g\left(\frac{1}{x}\right) = 0$ gilt

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} &= \lim_{x \rightarrow 0+} \frac{f\left(\frac{1}{x}\right)}{g\left(\frac{1}{x}\right)} = \lim_{x \rightarrow 0+} \frac{-\frac{1}{x^2} f'\left(\frac{1}{x}\right)}{-\frac{1}{x^2} g'\left(\frac{1}{x}\right)} \\ &= \lim_{x \rightarrow 0+} \frac{f'\left(\frac{1}{x}\right)}{g'\left(\frac{1}{x}\right)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}. \end{aligned}$$

■

Beispiel 10.18 Nach geeigneter Darstellung als Bruch kann der obige Satz angewendet werden:

$$\begin{aligned}\lim_{x \rightarrow \infty} x \left(\frac{\pi}{2} - \arctan x \right) &= \lim_{x \rightarrow \infty} \frac{\frac{\pi}{2} - \arctan x}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{-\frac{1}{1+x^2}}{-\frac{1}{x^2}} = \lim_{x \rightarrow \infty} \frac{x^2}{1+x^2} = 1\end{aligned}$$

□

Beispiel 10.19 Unter Verwendung des Zusatzes folgt

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x^\alpha} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \lim_{x \rightarrow \infty} \frac{1}{\alpha x^\alpha} = 0 \quad \text{für } \alpha > 0.$$

□

11 Das Integral für Regelfunktionen

Zwei Motivationen für das Integral:

1. **Das Flächenproblem:** Gegeben sei eine Funktion $f : [a, b] \rightarrow \mathbb{R}$. Kann man, bzw. unter welchen Bedingungen an f kann man der Fläche

$$F := \left\{ (x, y) \in \mathbb{R}^2 : a \leq x \leq b, 0 \leq y \leq f(x) \right\}$$

unter dem Graphen von f einen Flächeninhalt zuordnen (Flächenstücke unterhalb der x -Achse, dort wo f negativ ist, zwischen dem Graphen und der x -Achse sollen dabei negativ gerechnet werden). Dies ist offensichtlich (elementargeometrisch) z. B. für konstantes f , oder für lineares f ; entsprechend für Treppenfunktionen (stückweise konstant) oder stückweise lineare Funktionen. Wie sieht es aber für allgemeinere f aus, d. h. wie kann man den Flächeninhalt von „krummlinig“ begrenzten Flächen bestimmen?

2. **Längs eines Weges geleistete Arbeit:** Ein Körper bewegt sich unter dem Einfluß einer Kraft $k(x)$ (hier noch geradlinig) von a nach b . Ist die Kraft konstant $= k$, so ist die geleistete Arbeit $= k(b - a)$. Ist sie über Teilstücke $[x_j, x_{j+1}]$ mit $a = x_0 < x_1 < \dots < x_n = b$ konstant $= k_j$, so ist die geleistete Arbeit $= \sum_{j=0}^{n-1} k_j(x_{j+1} - x_j) = \sum_{j=1}^n k_{j-1}(x_j - x_{j-1})$. Was ist aber, wenn $k(x)$ sich kontinuierlich (oder noch allgemeiner) ändert?

11.1 Integral für Treppenfunktionen

Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt eine *Treppenfunktion*, wenn ein $N \in \mathbb{N}$ und Zahlen $a_0 < a_1 < \dots < a_N$ aus \mathbb{R} existieren mit

- f ist jeweils konstant in (a_{k-1}, a_k) ,
- $f(x) = 0$ für $x < a_0$ und für $x > a_k$;

die Werte in den Punkten a_k spielen im folgenden keine Rolle (bei allgemeineren Integralbegriffen wird das i. allg. anders sein).

Eine Treppenfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt eine *Treppenfunktion auf $[a, b]$* , wenn f außerhalb $[a, b]$ verschwindet.

Die *charakteristische Funktion* χ_I eines Intervalls I ($= 1$ für $x \in I$, $= 0$ für $x \notin I$) ist eine Treppenfunktion.

Zwei Treppenfunktionen f, g heißen *fast überall gleich* (f. ü.), $f(x) = g(x)$ f. ü., wenn $f(x) = g(x)$ gilt bis auf endlich viele Ausnahmepunkte x . Es gilt z. B. ¹¹

$$\chi_{(0,1]}(x) + \chi_{[1,3]}(x) = \chi_{(0,3)}(x) \quad \text{f. ü. .}$$

Offensichtlich gilt:

- Ist f eine Treppenfunktion und $c \in \mathbb{R}$, so sind $|f|$ und cf Treppenfunktionen.
- Sind f und g Treppenfunktionen, so sind $f + g$ und $f \cdot g$ Treppenfunktionen.
- Insbesondere bilden die Treppenfunktionen einen Vektorraum über \mathbb{R} (entsprechend natürlich über \mathbb{C} , wenn komplexe Werte zugelassen werden).

Das *Integral* einer Treppenfunktion

$$f = \sum_{k=1}^N c_k \chi_{I_k} \quad \text{f. ü.} \quad I_k = (a_{k-1}, a_k), \quad a_0 < a_1 < \dots < a_N$$

wird definiert durch

$$\int f := \sum_{k=1}^N c_k (a_k - a_{k-1}).$$

Es ist offensichtlich, dass dieser Wert nicht von der speziellen Darstellung der Treppenfunktion abhängt; genauer gilt für Treppenfunktionen f, g und $c_1, c_2 \in \mathbb{R}$

- $f = g$ f. ü. $\implies \int f = \int g$,
- $\int (c_1 f + c_2 g) = c_1 \int f + c_2 \int g$ (*Linearität*),
- aus $f \geq g$ f. ü. folgt $\int f \geq \int g$, insbesondere ist $\int f \geq 0$ für $f \geq 0$
- ist f eine Treppenfunktion auf $[a, b]$, so gilt

$$\left| \int f \right| \leq \int |f| \leq (b - a) \max \left\{ |f(x)| : a \leq x \leq b \right\}.$$

Ist f eine Treppenfunktion und $[a, b]$ ein Intervall, so schreiben wir

$$\int_a^b f = \int f \chi_{[a,b]}.$$

Insbesondere ist

$$\int_a^a f = 0 \quad \text{für jedes } a \in \mathbb{R},$$

¹¹mit $\xi_A(x) = 1$ für $x \in A$, 0 für $x \notin A$, *charakteristische Funktion von A*

und für $a \leq b \leq c$ gilt

$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Die Menge $B[a, b]$ der beschränkten Funktionen auf $[a, b]$ bildet einen Vektorraum, und durch

$$\|f\|_\infty := \sup \left\{ |f(x)| : x \in [a, b] \right\}$$

ist eine Norm auf $B[a, b]$, die *Supremumsnorm*, definiert (die Eigenschaften $\|f\|_\infty = 0 \Leftrightarrow f = 0$, $\|cf\|_\infty = |c|\|f\|_\infty$ und $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ sind offensichtlich). Die Konvergenz $(f_n \rightarrow f)$ bezüglich der ∞ -Norm heißt auch *gleichmäßige Konvergenz* ($f_n \rightarrow f$ glm.); sie kann auch so beschrieben werden:

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) \text{ mit } |f_n(x) - f(x)| < \varepsilon \text{ für } n \geq N \text{ und alle } x$$

(entscheidend für „gleichmäßig“ ist hier, dass N nicht von x abhängt).¹²

11.2 Das Integral für Regelfunktionen

Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt eine *Regelfunktion* auf $[a, b]$, wenn es eine Folge (φ_n) von Treppenfunktionen gibt mit

$$\|f - \varphi_n\|_\infty \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Auf Grund des folgenden Satzes kann man für Regelfunktionen f das Integral definieren durch

$$\int_a^b f := \lim_{n \rightarrow \infty} \int_a^b \varphi_n \quad (\varphi_n \text{ Treppenfunktion mit } \|f - \varphi_n\|_\infty \rightarrow 0).$$

Satz 11.1 *Ist f eine Regelfunktion auf $[a, b]$, (φ_n) eine Folge von Treppenfunktionen mit $\|f - \varphi_n\|_\infty \rightarrow 0$ für $n \rightarrow \infty$, so existiert der Grenzwert*

$$\lim_{n \rightarrow \infty} \int_a^b \varphi_n$$

und ist unabhängig von der Wahl der Folge (φ_n) von Treppenfunktionen mit $\|f - \varphi_n\|_\infty \rightarrow 0$.

¹²man vergleiche hiermit die *punktweise Konvergenz* $f_n \rightarrow f$,

$$\forall x \forall \varepsilon > 0 \exists N = N(x, \varepsilon) \text{ mit } |f_n(x) - f(x)| < \varepsilon \text{ für } n \geq N.$$

Die Folge (f_n) auf $[0, 1]$ mit $f_n(x) = x^n$ konvergiert punktweise (aber nicht gleichmäßig) gegen f mit $f(x) = 0$ für $0 \leq x < 1$ und $f(1) = 1$.

Beweis. Existenz des Grenzwerts: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit

$$\|f - \varphi_n\|_\infty < \frac{\varepsilon}{2(b-a)} \quad \text{für } n \geq N.$$

Dann gilt für $n, m \geq N$

$$\|\varphi_n - \varphi_m\|_\infty \leq \|\varphi_n - f\|_\infty + \|f - \varphi_m\|_\infty < \frac{\varepsilon}{b-a},$$

und somit

$$\left| \int \varphi_n - \int \varphi_m \right| = \left| \int (\varphi_n - \varphi_m) \right| \leq (b-a) \|\varphi_n - \varphi_m\|_\infty < \varepsilon.$$

Also ist $(\int \varphi_n)$ eine Cauchyfolge in \mathbb{R} und somit konvergent.

Unabhängigkeit des Grenzwerts von der Folge: Gilt auch $\|f - \psi_n\|_\infty \rightarrow 0$, so folgt wie im ersten Teil des Beweises

$$\|\varphi_n - \psi_n\|_\infty \leq \|\varphi_n - f\|_\infty + \|f - \psi_n\|_\infty \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

also

$$\left| \int \varphi_n - \int \psi_n \right| = \left| \int (\varphi_n - \psi_n) \right| \leq (b-a) \|\varphi_n - \psi_n\|_\infty \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

d. h. $\lim_{n \rightarrow \infty} \int \varphi_n = \lim_{n \rightarrow \infty} \int \psi_n$. ■

Wir schreiben

$$\int_a^b f(x) dx \quad (\text{oder beliebige Variable für } x) \quad \text{statt} \quad \int_a^b f.$$

Die Bezeichnung geht auf Gottfried Wilhelm Leibniz (1675) zurück (\int erinnert an ein stilisiertes S für Summe, dx an kleine x -Differenzen: $\sum_{k=1}^N f(x_k) \cdot \Delta x$). Die Schreibweise hat große Vorteile

- \int und dx sind wie Klammern, Anfang und Ende des Integralausdrucks (die Schreibweise $\int dx f(x)$ ist oft verwirrend, weil u. U. nicht erkennbar ist, wo das Integral endet),
- das dx macht immer deutlich, bezüglich welcher Variablen integriert wird, z. B. in $\int f(x, y) dx$.

19:18.12.07

Alle wichtigen Eigenschaften des Integrals für Treppenfunktionen bleiben beim Grenzübergang zu Regelfunktionen erhalten:

Satz 11.2 (Eigenschaften des Integrals) a) Linearität: Sind f_1, f_2 Regelfunktionen über $[a, b]$ und $c_1, c_2 \in \mathbb{R}$, so ist auch $c_1 f_1 + c_2 f_2$ eine Regelfunktion und es gilt

$$\int_a^b (c_1 f_1 + c_2 f_2) = c_1 \int_a^b f_1 + c_2 \int_a^b f_2.$$

b) Monotonie: Sind f, g Regelfunktionen über $[a, b]$ mit $f(x) \leq g(x)$ für alle $x \in [a, b]$, so gilt

$$\int_a^b f \leq \int_a^b g.$$

Speziell gilt $\int_a^b h \geq 0$, falls $h \geq 0$ gilt.

c) Abschätzung des Integrals: Ist f Regelfunktion über $[a, b]$, so ist auch $|f|$ Regelfunktion und es gilt

$$\left| \int_a^b f \right| \leq \int_a^b |f| \leq (b-a) \|f\|_\infty.$$

d) Sind f, g Regelfunktionen, so ist auch $f \cdot g$ Regelfunktion und es gilt

$$\left| \int_a^b fg \right| \leq \|f\|_\infty \int_a^b |g| \quad \text{und} \quad \left| \int_a^b fg \right| \leq \|g\|_\infty \int_a^b |f|.$$

Beweis. a) Sind (φ_n) und (ψ_n) Folgen von Treppenfunktionen mit

$$\|f_1 - \varphi_n\|_\infty \quad \text{und} \quad \|f_2 - \psi_n\|_\infty \rightarrow 0,$$

so folgt die Behauptung aus

$$\|c_1 f_1 + c_2 f_2 - c_1 \varphi_n - c_2 \psi_n\|_\infty \rightarrow 0$$

und der Linearität des Integrals für Treppenfunktionen

$$\int_a^b (c_1 \varphi_n + c_2 \psi_n) = c_1 \int_a^b \varphi_n + c_2 \int_a^b \psi_n.$$

b) Es genügt zu zeigen, dass aus $h \geq 0$ folgt $\int h \geq 0$ (falls h Regelfunktion ist). Da $h := g - f \geq 0$ ist folgt dann aus Teil a) $\int g - \int f = \int (g - f) = \int h \geq 0$.

Sei also $h \geq 0$ und (φ_n) eine Folge von Treppenfunktionen mit $\|h - \varphi_n\|_\infty \rightarrow 0$. Dann sind auch $|\varphi_n|$ Treppenfunktionen, und es gilt

$$\left| |\varphi_n(x)| - h(x) \right| \leq \left| \varphi_n(x) - h(x) \right| \leq \|\varphi_n - h\|_\infty \rightarrow 0,$$

also

$$\int_a^b h = \lim_{n \rightarrow \infty} \int_a^b |\varphi_n| \geq 0.$$

c) Ist (φ_n) eine Folge von Treppenfunktionen mit $\|f - \varphi_n\|_\infty \rightarrow 0$, so gilt auch $\| |\varphi_n| - |f| \|_\infty \rightarrow 0$ für $n \rightarrow \infty$. $|f|$ ist also eine Regelfunktion mit

$$\pm \int_a^b f \leq \int_a^b |f| \leq \int_a^b \|f\|_\infty = (b-a)\|f\|_\infty,$$

woraus die Behauptung folgt.

d) Sind (φ_n) und (ψ_n) Folgen von Treppenfunktionen mit $\|f - \varphi_n\|_\infty \rightarrow 0$ und $\|g - \psi_n\|_\infty \rightarrow 0$, so ist auch $(\varphi_n \psi_n)$ eine Folge von Treppenfunktionen und mit $C = \sup_{n \in \mathbb{N}} \|\psi_n\|_\infty$ gilt

$$\begin{aligned} \left| fg(x) - \varphi_n \psi_n(x) \right| &\leq \left| f(x)(g(x) - \psi_n(x)) \right| + \left| (f(x) - \varphi_n(x))\psi_n(x) \right| \\ &\leq \|f\|_\infty \|g - \psi_n\|_\infty + \|f - \varphi_n\|_\infty \|\psi_n\|_\infty \leq \|f\|_\infty \|g - \psi_n\|_\infty + C \|f - \varphi_n\|_\infty. \end{aligned}$$

also $\|fg - \varphi_n \psi_n\|_\infty \rightarrow 0$ für $n \rightarrow \infty$. ■

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Regelfunktion und $[c, d]$ ein Teilintervall von $[a, b]$,

$$g : [c, d] \rightarrow \mathbb{R} \text{ mit } g(x) = f(x) \text{ für } x \in [c, d].$$

Dann ist g eine Regelfunktion auf $[c, d]$ und man definiert

$$\int_c^d f := \int_c^d g.$$

Satz 11.3 *Unter obigen Voraussetzungen gilt*

a) $\int_c^d f = \int_a^b f \chi_{[c,d]}$

b) Für $a \leq c \leq b$ gilt $\int_a^b f = \int_a^c f + \int_c^b f$.

Den *Beweis* überlassen wir dem Leser. (Ü)

Schließlich definiert man noch für $b < a$

$$\int_a^b f := - \int_b^a f;$$

das wird sich später als nützlich erweisen.

11.3 Regelfunktionen

Welche Funktionenklassen sind Regelfunktionen?

Satz 11.4 *Ist $f : [a, b] \rightarrow \mathbb{R}$ monoton (wachsend oder fallend), so ist f Regelfunktion (also im obigen Sinne integrierbar).*

Beweis. Sei o. E. f wachsend, $h_n := \frac{1}{n}(f(b) - f(a))$,

$$y_k := f(a) + k \cdot h_n \quad \text{für } k = 0, 1, \dots, n.$$

$$a_k := \sup \left\{ x \in [a, b] : f(x) \leq y_k \right\} \quad \text{für } k = 0, 1, \dots, n.$$

Für

$$\varphi_n(x) := \sum_{k=1}^n y_k \chi_{(a_{k-1}, a_k)}(x) = y_k \quad \text{für } x \in (a_{k-1}, a_k)$$

gilt dann

$$\|f - \varphi_n\|_\infty \leq h_n = \frac{f(b) - f(a)}{n} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

■

Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt von *beschränkter Variation*, wenn sie als Differenz zweier monoton wachsender Funktionen geschrieben werden kann (im komplexen Fall in der Form $f_1 - f_2 + if_3 - if_4$ mit vier wachsenden Funktionen f_1, f_2, f_3 und f_4). Offenbar ist jede Funktion von beschränkter Variation eine Regelfunktion, da sie Differenz zweier Regelfunktionen ist.

Man kann zeigen, dass eine Funktion genau dann von beschränkter Variation ist, wenn eine Konstante C existiert so, dass für jede Zerlegung $a = a_0 < a_1 < \dots < a_N = b$ von $[a, b]$ gilt $\sum_{j=1}^N |f(a_j) - f(a_{j-1})| \leq C$; das kleinste C dieser Art heißt die *Variation* von f .

Um auch stetige Funktionen zu erfassen, brauchen wir eine wichtige Verschärfung des Stetigkeitsbegriffes, die wir gleich allgemeiner formulieren, als wir sie hier benötigen.

Sei $D \subset \mathbb{R}^m$; $f : D \rightarrow \mathbb{R}^n$ (ebenso mit \mathbb{K}^m und \mathbb{K}^n) heißt *gleichmäßig stetig*, wenn gilt

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in D \quad \text{mit } |x - y| < \delta \quad \text{gilt } |f(x) - f(y)| < \varepsilon$$

(die Gleichmäßigkeit steckt hier darin, dass δ nur von ε abhängt, nicht von x und y).

Satz 11.5 *Ist $D \subset \mathbb{R}^m$ kompakt und $f : D \rightarrow \mathbb{R}^n$ stetig, so ist f gleichmäßig stetig.*

Beweis. Wäre f nicht gleichmäßig stetig, so würde gelten

$$\exists \varepsilon > 0 \forall \delta > 0 \exists x, y \in D \text{ mit } |x - y| < \delta \text{ und } |f(x) - f(y)| > \varepsilon;$$

also existieren Folgen (x_n) und (y_n) aus D mit

$$|x_n - y_n| \rightarrow 0 \quad \text{und} \quad |f(x_n) - f(y_n)| > \varepsilon.$$

Da D kompakt ist, gibt es eine Teilfolge (n_k) von (n) mit

$$x_{n_k} \rightarrow x, \quad y_{n_k} \rightarrow y, \quad \text{also} \quad x = y \quad (\text{warum?}).$$

Da f stetig ist gilt

$$f(x_{n_k}) \rightarrow f(x) \quad \text{und} \quad f(y_{n_k}) \rightarrow f(y) = f(x),$$

also

$$\left| f(x_{n_k}) - f(y_{n_k}) \right| \rightarrow 0$$

im Widerspruch zu $|f(x_{n_k}) - f(y_{n_k})| > \varepsilon$. ■

Satz 11.6 *Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so ist f eine Regelfunktion.*

Beweis. Nach dem vorhergehenden Satz ist f gleichmäßig stetig. Zu jedem $n \in \mathbb{N}$ gibt es also ein $k_n \in \mathbb{N}$ mit

$$|f(x) - f(y)| \leq \frac{1}{n} \quad \text{für} \quad |x - y| < \frac{b - a}{k_n}.$$

Mit $a_{n,k} := a + k \frac{b - a}{k_n}$ für $k = 0, \dots, y_k$ ist dann

$$\varphi_n(x) := f(a_{n,k}) \quad \text{für} \quad a_{n,k-1} \leq x < a_{n,k}, \quad k = 1, \dots, k_n,$$

eine Folge von Treppenfunktionen mit

$$\|\varphi_n - f\|_\infty \leq \frac{1}{n},$$

also ist f eine Regelfunktion. ■

Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *stückweise stetig*, wenn es eine Unterteilung $a = a_0 < a_1 < \dots < a_n = b$ von $[a, b]$ gibt so, dass

$$f_k : (a_{k-1}, a_k) \rightarrow \mathbb{R}, \quad f_k(x) = f(x) \quad \text{für} \quad x \in (a_{k-1}, a_k)$$

stetig sind und die einseitigen Grenzwerte

$$f(a_k+) = \lim_{x \rightarrow a_k+} f(x) \quad \text{und} \quad f(a_k-) = \lim_{x \rightarrow a_k-} f(x)$$

existieren (d. h. wenn sich die Funktionen $f_k : (a_{k-1}, a_k) \rightarrow \mathbb{R}$ zu stetigen Funktionen $\overline{f}_k : [a_{k-1}, a_k]$ fortsetzen lassen. Der Leser überzeuge sich selbst davon, dass auch stückweise stetige Funktionen auf $[a, b]$ Regelfunktionen sind (Ü).

Entsprechend ist natürlich jede beschränkte stückweise monotone Funktion eine Regelfunktion.

In der Lehrbuchliteratur wird meist nicht das Integral für Regelfunktionen, sondern das sogenannte *Riemann-Integral* eingeführt. Dabei ist eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ genau dann *Riemann-integrierbar*, wenn zu jedem $\varepsilon > 0$ Treppenfunktionen φ und ψ existieren mit $\varphi \leq f \leq \psi$ und $\int_a^b \psi - \int_a^b \varphi < \varepsilon$. Die Klasse der Riemann-integrierbaren Funktionen ist etwas größer als die der Regelfunktionen. Eine Riemann-integrierbare Funktion, die nicht Regelfunktion ist, ist z. B. (Ü: man zeige, dass für jede Treppenfunktion $\varphi : [a, b] \rightarrow \mathbb{R}$ gilt $\|f - \varphi\|_\infty \geq 1$.)

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 0 & \text{für } x = 0, \\ \sin \frac{1}{x} & \text{für } x \neq 0. \end{cases}$$

Dagegen ist die Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 0 & \text{für } x = 0, \\ x \sin \frac{1}{x} & \text{für } x \neq 0 \end{cases}$$

eine Regelfunktion. (Ü: auf $[0, \varepsilon]$ ist $|f(x) - 0| < \varepsilon$, auf $[\varepsilon, 1]$ ist f stetig und deshalb durch Treppenfunktionen beliebig gut approximierbar.)

Die Dirichletfunktion auf $[a, b]$ ist keine Regelfunktion (sie ist übrigens auch nicht Riemann-integrierbar), während die modifizierte Dirichletfunktion eine Regelfunktion ist.

20:20.12.07

11.4 Berechnung von Integralen

Satz 11.7 Ist $f : [a, b] \rightarrow \mathbb{R}$ eine Regelfunktion, $c \in [a, b]$. Dann ist

$$F : [a, b] \rightarrow \mathbb{R}, \quad F(x) := \int_c^x f(t) dt$$

stetig.

Beweis. Für $y \geq x$ gilt

$$|F(y) - F(x)| = \left| \int_c^y f - \int_c^x f \right| = \left| \int_a^y f - \int_a^x f \right| = \left| \int_x^y f \right| \leq \|f\|_\infty (y - x).$$

■

Satz 11.8 (Mittelwertsatz der Integralrechnung) a) Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so existiert ein $\xi \in (a, b)$ mit

$$\int_a^b f(x) \, dx = f(\xi)(b - a).$$

b) Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und $\varphi : [a, b] \rightarrow \mathbb{R}$ eine positive (negative) Regelfunktion, so existiert ein $\xi \in (a, b)$ mit

$$\int_a^b f(x)\varphi(x) \, dx = f(\xi) \int_a^b \varphi(x) \, dx.$$

(Wenn φ das Vorzeichen wechselt, gilt diese Aussage i. Allg. nicht (Ü))

Beweis. a) Folgt aus Teil b mit $\varphi(x) = 1$.

b) Sei zunächst φ positiv. Ist $\min f = \max f$, so ist f konstant und die Behauptung ist richtig für jedes $\xi \in [a, b]$. Ist $\min f < \max f$, so gilt

$$\begin{aligned} \min f \int_a^b \varphi(x) \, dx &< \int_a^b f(x)\varphi(x) \, dx < \max f \int_a^b \varphi(x) \, dx, \\ \min f &< \frac{\int_a^b f(x)\varphi(x) \, dx}{\int_a^b \varphi(x) \, dx} < \max f, \end{aligned}$$

und aus dem Zwischenwertsatz für f folgt die Behauptung.

Entsprechend wird die Behauptung für negatives φ bewiesen. ■

Eine Funktion $F : I \rightarrow \mathbb{R}$ heißt *Stammfunktion* einer Funktion $f : I \rightarrow \mathbb{R}$, wenn F differenzierbar ist mit

$$F'(x) = f(x) \quad \text{für alle } x \in I.$$

Satz 11.9 Wenn f eine Stammfunktion F besitzt, so ist die Menge aller Stammfunktionen von f gegeben durch

$$\{F + c : c \in \mathbb{R}\}.$$

Beweis. G ist zunächst genau dann eine Stammfunktion, $G' = F'$ bzw. $(G - F)' = 0$, d. h. wenn $G - F$ konstant ist. ■

Satz 11.10 Ist $f : I \rightarrow \mathbb{R}$ stetig, so besitzt f Stammfunktionen. Für jedes $c \in I$ erhält man eine Stammfunktion als das unbestimmte Integral

$$F(x) = \int_c^x f(t) dt.$$

Beweis. Nach dem Mittelwertsatz der Integralrechnung gilt mit $\xi_{x,h}$ zwischen x und $x + h$

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt = \frac{1}{h} f(\xi_{x,h})h = f(\xi_{x,h}).$$

Da f stetig ist und $\xi_{x,h}$ für $h \rightarrow 0$ gegen x konvergiert, ist also F differenzierbar mit $F'(x) = f(x)$. ■

Satz 11.11 (Hauptsatz der Differential- und Integralrechnung) a) Für jede Stammfunktion F von $f \in C(I)$ und alle $a, b \in I$ gilt

$$\int_a^b f(x) dx = F(b) - F(a) =: F(x) \Big|_a^b.$$

b) Ist f stetig differenzierbar, so gilt für $a, x \in I$

$$f(x) = f(a) + \int_a^x f'(t) dt.$$

Beweis. a) $F_0(x) := \int_a^x f(t) dt$ ist eine Stammfunktion von f , und jede Stammfunktion F von f hat dann die Form $F(x) = F_0(x) + c$. Also ist

$$F(b) - F(a) = F_0(b) - F_0(a) = F_0(b) = \int_a^b f(x) dx.$$

b) Folgt aus Teil a mit f' statt f und x statt b . ■

Wir können jetzt also Integrale berechnen für alle Funktionen, die wir als Ergebnis einer Differentiation erhalten haben:

Beispiel 11.12 $\int_a^b e^{\alpha x} dx = \frac{1}{\alpha} e^{\alpha x} \Big|_a^b$ für $\alpha \neq 0$, $\int_a^b 1 dx = x \Big|_a^b = b - a$ für $\alpha = 0$ ($a, b \in \mathbb{R}$). \square

Beispiel 11.13 $\int_a^b x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} \Big|_a^b$ für $\alpha \neq -1$ (für $\alpha < 0$ darf 0 nicht in $[a, b]$ liegen, da sonst x^α in $[a, b]$ nicht stetig ist). \square

Beispiel 11.14 $\int_a^b \frac{1}{x} dx = \ln|b| - \ln|a| = \ln \frac{b}{a}$ für $a, b \in \mathbb{R}$ mit $a \cdot b > 0$.

Wegen $\ln' x = \frac{1}{x}$ für $x > 0$ gilt dies zunächst für $0 < a \leq b$.

Für $a \leq x \leq b < 0$ gilt $\frac{d}{dx} \ln(-x) = -\frac{1}{-x} = \frac{1}{x}$ ¹³, also

$$\int_a^b \frac{1}{x} dx = \ln(-x) \Big|_a^b = \ln(-b) - \ln(-a) = \ln\left(\frac{-b}{-a}\right) = \ln\left(\frac{b}{a}\right).$$

\square

Beispiel 11.15 $\int_a^b \cos x dx = \sin x \Big|_a^b$, $\int_a^b \sin x dx = -\cos x \Big|_a^b$ für $a, b \in \mathbb{R}$. \square

Beispiel 11.16 $\int_a^b \frac{1}{1+x^2} dx = \arctan x \Big|_a^b$ für $a, b \in \mathbb{R}$. \square

Beispiel 11.17 $\int_a^b \frac{1}{\sqrt{1-x^2}} dx = \arcsin x \Big|_a^b$ für $a, b \in (-1, 1)$. \square

¹³Allgemein: $\frac{d}{dx} \ln|x| = \frac{1}{x}$ für $x \neq 0$.

12 Integrationsmethoden

Da die Integration (genauer das Aufsuchen einer Stammfunktion) die Umkehrung der Differentiation ist, zieht jede Differentiationsregel eine Integrationsregel nach sich. So ergibt sich im folgenden aus der Produktregel die Regel der *partiellen Integration* und aus der Kettenregel die *Substitutionsregel*.

12.1 Partielle Integration

Satz 12.1 (Partielle Integration) Sind $u, v \in C^1[a, b]$, so gilt

$$\int_a^b u(x)v'(x) \, dx = u(x)v(x) \Big|_a^b - \int_a^b u'(x)v(x) \, dx.$$

Beweis. Nach der Produktregel gilt

$$(uv)' = u'v + uv',$$

d. h. uv ist eine Stammfunktion von $u'v + uv'$. Nach dem Hauptsatz gilt also

$$uv \Big|_a^b = \int_a^b (u'v + uv') = \int_a^b u'v + \int_a^b uv'.$$

Das ist die Behauptung. ■

Bemerkung 12.2 In der Anwendung sieht das so aus: Ist $f \in C^1[a, b]$, $g \in C[a, b]$ und G eine Stammfunktion von g , so gilt

$$\int_a^b f(x)g(x) \, dx = f(x)G(x) \Big|_a^b - \int_a^b f'(x)G(x) \, dx.$$

Es wird also das Integral über ein Produkt (unter geeigneten Voraussetzungen) umgewandelt in einen explizit gegebenen Term fG und ein Integral über ein anderes Produkt, das hoffentlich leichter zu berechnen ist bzw. weiter umgeformt werden kann.

Beispiel 12.3 $\int_a^b xe^x \, dx = xe^x \Big|_a^b - \int_a^b e^x \, dx = (xe^x - e^x) \Big|_a^b,$
 $\int_a^b x^2e^x \, dx = x^2e^x \Big|_a^b - \int_a^b 2xe^x \, dx = (x^2 - 2x + 2)e^x \Big|_a^b.$

Allgemeiner kann auf diese Weise jedes Integral der Form

$$\int_a^b p(x)e^x dx \quad \text{mit einem Polynom } p$$

berechnet werden.

Analog können Integrale der Form $\int p(x) \cos x dx$ und $\int p(x) \sin x dx$ behandelt werden. \square

Beispiel 12.4 Für $0 \notin [a, b]$ gilt

$$\int_a^b \ln |x| dx = \int_a^b 1 \cdot \ln |x| dx = x \ln |x| \Big|_a^b - \int_a^b x \frac{1}{x} dx = x(\ln |x| - 1) \Big|_a^b.$$

Entsprechend erhält man allgemein mit einem Polynom

$$p(x) = \sum_{j=0}^n a_j x^j \quad \text{und} \quad P(x) = \sum_{j=0}^n \frac{1}{j+1} a_j x^{j+1} =: xq(x),$$

$$\int_a^b p(x) \ln x dx = P(x) \ln x \Big|_a^b - \int_a^b \frac{1}{x} P(x) dx = xq(x) \ln x \Big|_a^b - \int_a^b q(x) dx.$$

\square

Bei folgendem Beispiel (und in vielen anderen Fällen) hilft ein **Trick**: Durch geeignete, evtl. mehrfache partielle Integration erhält man

$$\int_a^b f(x)g(x) dx = h(x) \Big|_a^b + c \int_a^b f(x)g(x) dx$$

mit einem $c \neq 1$, also

$$\int_a^b f(x)g(x) dx = \frac{1}{1-c} h(x) \Big|_a^b.$$

Beispiel 12.5 $\int \cos^2 x dx$: Mit $u(x) = v'(x) = \cos x$ erhält man

$$\begin{aligned} \int_a^b \cos^2 x dx &= \int_a^b \cos x \cos x dx = \cos x \sin x \Big|_a^b + \int_a^b \sin x \sin x dx \\ &= \cos x \sin x \Big|_a^b + \int_a^b (1 - \cos^2 x) dx = \cos x \sin x \Big|_a^b + (b-a) - \int_a^b \cos^2 x dx, \end{aligned}$$

also

$$\int_a^b \cos^2 x \, dx = \frac{1}{2}(b-a) + \frac{1}{2} \cos x \sin x \Big|_a^b.$$

Speziell erhält man als Stammfunktion von $\cos^2 x$

$$\frac{1}{2}x + \frac{1}{2} \cos x \sin x.$$

(*Bemerkung:* Hätte man nach der ersten Zeile der obigen Rechnung weiter partiell integriert, hätte man die triviale Identität

$$\int_a^b \cos^2 x \, dx = \cos x \sin x \Big|_a^b - \cos x \sin x \Big|_a^b + \int_a^b \cos^2 x \, dx$$

erhalten, die nicht weitergeholfen hätte.)

Entsprechend kann $\int_a^b \sin^2 x \, dx$ berechnet werden und man erhält $\frac{1}{2}x - \frac{1}{2} \cos x \sin x$ als Stammfunktion von $\sin^2 x$. \square

Die Formel der partiellen Integration, zusammen mit dem verallgemeinerten Mittelwertsatz der Integralrechnung erlaubt nun einen alternativen Beweis des Satzes von Taylor, der auch eine andere Form des Restgliedes liefert:

Satz 12.6 (Taylor) Sei $f \in C^{n+1}(a, b)$, $x_0 \in (a, b)$. Dann gilt

$$f(x) = \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + R_n(x)$$

mit

$$R_n(x) = \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

mit einem ξ zwischen x und x_0 . Die erste Form des Restgliedes ist die Taylorsche Form, die zweite die Lagrangesche Form.

Beweis. Für $n = 0$ (d. h. wenn f mindestens einmal stetig differenzierbar ist) gilt

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt = f(x_0) + \int_{x_0}^x f'(t)(x-t)^0 dt.$$

Ist f zweimal stetig differenzierbar ($n = 1$), so erhält man mittels partieller Integration

$$\begin{aligned} f(x) &= f(x_0) + \int_{x_0}^x 1 \cdot f'(t) \, dt = f(x_0) - (x-t)f'(t) \Big|_{x_0}^x + \int_{x_0}^x (x-t)f''(t) \, dt \\ &= f(x_0) + f'(t)(x-x_0) + \int_{x_0}^x f''(t)(x-t)^1 \, dt. \end{aligned}$$

Für $n = 0$ und $n = 1$ haben wir damit die Formel

$$f(x) = \sum_{j=0}^n f^{(j)}(x_0) \frac{(x-x_0)^j}{j!} + \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt$$

mit dem Taylorschen Restglied $R_n(x)$ bewiesen. Es bleibt also noch der Induktionsschritt $n \Rightarrow n+1$ durchzuführen. Partielle Integration des Restgliedterms $R_n(x)$ liefert:

$$\begin{aligned} f(x) &= \sum_{j=0}^n f^{(j)}(x_0) \frac{(x-x_0)^j}{j!} + \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt \\ &= \sum_{j=0}^n f^{(j)}(x_0) \frac{(x-x_0)^j}{j!} - f^{(n+1)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \Big|_{x_0}^x + \int_{x_0}^x f^{(n+2)}(t) \frac{(x-t)^{n+1}}{(n+1)!} \, dt \\ &= \sum_{j=0}^{n+1} f^{(j)}(x_0) \frac{(x-x_0)^j}{j!} + R_{n+1}(x). \end{aligned}$$

Damit ist die Taylorsche Formel mit einem Restglied

$$R_n(x) = \int_{x_0}^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, dt$$

bewiesen.

Die frühere (*Lagrange'sche*) Form des Restglieds erhält nun durch Anwendung des verallgemeinerten Mittelwertsatzes der Integralrechnung (man beachte, dass $\varphi(t) := (x-t)^n$ für t zwischen x_0 und x konstantes Vorzeichen hat:

$$R_n(x) = f^{(n+1)}(\xi) \int_{x_0}^x \frac{(x-t)^n}{n!} \, dt = f^{(n+1)}(\xi) \frac{(x-x_0)^{n+1}}{(n+1)!}$$

mit ξ zwischen x_0 und x . ■

12.2 Substitutionsregel

Der Kettenregel für die Differentiation entspricht die Substitutionsregel für die Integration:

Satz 12.7 (Substitutionsregel) Ist $f \in C[a, b]$ und $u : [\alpha, \beta] \rightarrow [a, b]$ stetig differenzierbar mit $u(\alpha) = a$ und $u(\beta) = b$, so gilt

$$\int_{u(\alpha)}^{u(\beta)} f(x) \, dx = \int_{\alpha}^{\beta} f(u(t))u'(t) \, dt \quad \text{bzw.} \quad \int_a^b f(x) \, dx = \int_{u^{-1}(a)}^{u^{-1}(b)} f(u(t))u'(t) \, dt.$$

Um eine Stammfunktion von f zu bestimmen, kann man also eine Stammfunktion von $(f \circ u)u'$ bestimmen und in dieser t durch $u^{-1}(x)$ ersetzen (Resubstitution).

Beweis. Ist F eine Stammfunktion von f , so ist $F \circ u$ eine Stammfunktion von $(f \circ u)u'$, denn nach der Kettenregel gilt

$$\frac{d}{dt}F(u(t)) = F'(u(t))u'(t) = f(u(t))u'(t).$$

Mit dem Hauptsatz folgt daraus

$$\int_{u(\alpha)}^{u(\beta)} f(x) \, dx = F(x) \Big|_{u(\alpha)}^{u(\beta)} = F(u(x)) \Big|_{\alpha}^{\beta} = \int_a^b f(u(t))u'(t) \, dt.$$

Die letzte Aussage erhält man, indem man in der zweiten Formel b durch x ersetzt. ■

Diese Formel kann auf zwei Arten eingesetzt werden:

1. Art: Man versucht in einem Integral

$$\int_a^b g(x) \, dx$$

den Integranden in der Form

$$g(x) = f(u(x)) \cdot u'(x)$$

zu schreiben, wobei man eine Stammfunktion F von f kennt. Dann gilt nämlich

$$\int_a^b g(x) \, dx = \int_a^b f(u(x))u'(x) \, dx = \int_a^b \frac{d}{dx}F(u(x)) \, dx = F(u(x)) \Big|_a^b.$$

Beispiel 12.8 $\int_a^b x e^{-\frac{1}{2}x^2} dx^2$. Mit $f(y) = e^{-y}$, $F(y) = -e^{-y}$ und $u(x) = \frac{1}{2}x^2$ gilt

$$x e^{-\frac{1}{2}x^2} = f(u(x))u'(x),$$

also

$$\int_a^b x e^{-\frac{1}{2}x^2} dx = F(u(x)) \Big|_a^b = -e^{-\frac{1}{2}x^2} \Big|_a^b = e^{-\frac{1}{2}a^2} - e^{-\frac{1}{2}b^2}.$$

□

Beispiel 12.9 $\int_a^b f(\cos x) \sin x dx$ und $\int_a^b f(\sin x) \cos x dx$ können entsprechend behandelt werden. (Ü) □

Beispiel 12.10 *Logarithmische Ableitung:*

$$\int_a^b \frac{u'(x)}{u(x)} dx = \ln |u(b)| - \ln |u(a)| = \ln \frac{u(b)}{u(a)}$$

wenn $[a, b]$ keine Nullstelle von u enthält.

Ein Spezialfall ist $\int_a^b \tan x dx = \int_a^b \frac{\sin x}{\cos x} dx$; mit $u(x) = \cos x$ folgt (falls $[a, b]$ keine Nullstelle von \cos enthält, also keinen der Punkte $\frac{\pi}{2} + l\pi$ mit $k \in \mathbb{Z}$)

$$\int_a^b \tan x dx = \ln |\cos x| \Big|_a^b = \ln |\cos a| - \ln |\cos b| = \ln \frac{\cos a}{\cos b}.$$

□

2. Art: Man versucht in $\int_a^b f(x) dx$ die Variable x durch $u(t)$ zu ersetzen, wobei $u : [\alpha, \beta] \rightarrow [a, b]$ eine C^1 -Funktion mit $u(\alpha) = a$, $u(\beta) = b$ ist. Die Hoffnung dabei ist, dass ein einfaches integrierbares Integral entsteht:

$$\int_a^b f(x) dx = \int_\alpha^\beta f(u(t))u'(t) dt.$$

Formales Vorgehen: man ersetzt (*substituiert*) x durch $u(t)$ und dx durch $u'(t) dt$.

Beispiel 12.11 $\int_a^b \sin \sqrt{x} \, dx$ kann mit der *Substitution* $x = u(t) = t^2$ umgeformt werden.

$$\int_a^b \sin \sqrt{x} \, dx = \int_{\sqrt{a}}^{\sqrt{b}} (\sin t) 2t \, dt = 2 \int_{\sqrt{a}}^{\sqrt{b}} t \sin t \, dt,$$

ein Integral, das mit Hilfe partieller Integration leicht zu berechnen ist (vgl. Beispiel 12.3).

Oder allgemeiner:

$$\int_a^b f(\sqrt{x}) \, dx = 2 \int_{\sqrt{a}}^{\sqrt{b}} t f(t) \, dt.$$

□

Beispiel 12.12 $\int f(\ln x) \, dx$ kann mit der Substitution $x = u(t) = e^t$ vereinfacht werden:

$$\int_a^b f(\ln x) \, dx = \int_{\ln a}^{\ln b} f(t) e^t \, dt.$$

□

Beispiel 12.13 Bei Integralen $\int_a^b f(\sqrt{1-x^2}) \, dx$ führt die Substitution $x = u(t) = \sin t$ zu einer Vereinfachung:

$$\int_a^b f(\sqrt{1-x^2}) \, dx = \int_{\arcsin a}^{\arcsin b} f(\cos t) \cos t \, dt.$$

□

12.3 Hyperbelfunktionen

Um auch Integrale mit Termen $\sqrt{1+x^2}$ behandeln zu können, sind die *hyperbolischen* Funktionen \cosh „Cosinus hyperbolicus“ und \sinh „Sinus hyperbolicus“ nützlich. Sie können an der *Hyperbel* $x^2 - y^2 = 1$ ähnlich beschrieben werden wie die trigonometrischen Funktionen \cos und \sin am Einheitskreis $x^2 + y^2 = 1$ (vgl. W. Walter, Analysis 1, §7). Sie sind analytisch gegeben durch (vgl. Eulersche Gleichungen für \cos , \sin)

$$\cosh x = \frac{1}{2}(e^x + e^{-x}), \quad \sinh x = \frac{1}{2}(e^x - e^{-x}).$$

Offensichtlich gilt

$$\begin{aligned}\cosh(-x) &= \cosh x, & \sinh(-x) &= -\sinh(x), \\ \cosh x &\geq 1 \text{ für alle } x \in \mathbb{R} \quad (\text{Minimum bei } x = 0), \\ \cosh^2 x - \sinh^2 x &= 1 \quad (\text{nachrechnen!}), \\ \cosh x &= \sqrt{1 + \sinh^2 x}, \\ \cosh' x &= \sinh x, & \sinh' x &= \cosh x.\end{aligned}$$

Die Umkehrfunktionen

$$\begin{aligned}\cosh^{-1} : [1, \infty) &\rightarrow [0, \infty) \quad (\text{Area cosinus hyperbolicus}) \\ \sinh^{-1} : \mathbb{R} &\rightarrow \mathbb{R} \quad (\text{Area sinus hyperbolicus})\end{aligned}$$

lassen sich leicht explizit berechnen: Für $t \geq 0$ gilt

$$x = \cosh t = \frac{1}{2}(e^t + e^{-t}) \Rightarrow e^t x = \frac{1}{2}e^{2t} + \frac{1}{2} \Rightarrow e^{2t} - 2xe^t + 1 = 0, \quad (e^t)^2 - 2xe^t + 1 = 0.$$

Dies ist eine quadratische Gleichung für e^t , also

$$\begin{aligned}e^t &= x \pm \sqrt{x^2 - 1} = x + \sqrt{x^2 - 1} \quad (\text{da } e^t > 1 \text{ für } t \geq 0) \\ \cosh^{-1} x &= t = \ln(x + \sqrt{x^2 - 1}) \quad \text{für } x \geq 1.\end{aligned}$$

Entsprechend erhält man:

$$\begin{aligned}x = \sinh t &= \frac{1}{2}(e^t - e^{-t}) \Rightarrow (e^t)^2 - 2xe^t - 1 = 0 \\ \Rightarrow e^t &= x \pm \sqrt{x^2 + 1} = x + \sqrt{x^2 + 1} \quad (\text{da } e^t > 0),\end{aligned}$$

und somit

$$\sinh^{-1} = \ln \left(x + \sqrt{x^2 + 1} \right).$$

Insbesondere folgt $(\cosh^{-1})'(x) = (x^2 - 1)^{-1/2}$ für $x > 0$ und $(\sinh^{-1})'(x) = (x^2 + 1)^{-1/2}$ für $x \in \mathbb{R}$ (vgl. auch folgendes Beispiel).

Beispiel 12.14 Integrale der Form $\int f(\sqrt{1+x^2}) dx$ können mit der Substitution $x = u(t) = \sinh t$ umgeformt werden:

$$\int_a^b f(\sqrt{1+x^2}) dx = \int_{\sinh^{-1} a}^{\sinh^{-1} b} f(\cosh t) \cosh t dt,$$

also, was wir gerade durch Differentiation erkannt haben,

$$\int_a^b \frac{1}{\sqrt{1+x^2}} dx = \int_{\sinh^{-1} a}^{\sinh^{-1} b} \frac{1}{\cosh t} \cosh t dt = \sinh^{-1} b - \sinh^{-1} a = \ln \frac{b + \sqrt{1+b^2}}{a + \sqrt{1+a^2}}.$$

□

12.4 Integration rationaler Funktionen

Abschließend soll nun noch auf die Integration rationaler Funktionen $f(x) = \frac{p(x)}{q(x)}$ mit Polynome p und q eingegangen werden (im Zeitalter der Computeralgebra kommt diesem Abschnitt allerdings keine besondere Bedeutung mehr zu).

Wir werden im folgenden sehen, dass sich die Integration reeller rationaler Funktionen auf die folgenden drei „einfachen“ Fälle zurückführen läßt (für komplexe rationale Funktionen vgl. die Bemerkung im Anschluß an Satz 12.16):

1. Polynome $\sum_{j=0}^n a_j x^j$: Ein solches Polynom hat bekanntlich die Stammfunktion

$$\sum_{j=0}^n \frac{a_j}{j+1} x^{j+1} = \sum_{j=1}^{n+1} \frac{a_{j-1}}{j} x^j.$$

2. $(x - x_0)^{-\ell}$ mit $x_0 \in \mathbb{R}$ und $\ell \in \mathbb{N}$: Für diese Funktion ist eine Stammfunktion gegeben durch

$$\begin{aligned} \frac{1}{1-\ell} (x - x_0)^{1-\ell} & \quad \text{für } \ell > 1, \\ \ln|x - x_0| & \quad \text{für } \ell = 1. \end{aligned}$$

3. $\frac{Bx + C}{(x^2 + 2bx + c)^\ell}$ mit $B, C, b, c \in \mathbb{R}$, $\ell \in \mathbb{N}$ und $c > b^2$ (letzteres bedeutet, dass der Nenner keine reelle Nullstelle hat; er ist für alle x strikt positiv)¹⁴: Dieser Term läßt sich offenbar wie folgt umschreiben:

$$\frac{Bx + C}{(x^2 + 2bx + c)^\ell} = \frac{B}{2} \frac{2x + 2b}{(x^2 + 2bx + c)^\ell} + \frac{C - Bb}{(x^2 + 2bx + c)^\ell}.$$

Hier hat der erste Term die Stammfunktion (der Zähler ist die Ableitung des Klammerausdrucks im Nenner):

$$\begin{aligned} \frac{B}{2(1-\ell)} (x^2 + 2bx + c)^{1-\ell} & \quad \text{für } \ell > 1, \\ \frac{B}{2} \ln(x^2 + 2bx + c) & \quad \text{für } \ell = 1. \end{aligned}$$

Es bleibt also eine Stammfunktion zu finden für Funktionen der Form

$$\frac{1}{(x^2 + 2bx + c)^\ell} = \frac{1}{\left((x+b)^2 + (c-b^2)\right)^\ell} = \frac{1}{(c-b^2)^\ell} \frac{1}{\left(\left(\frac{x+b}{\sqrt{c-b^2}}\right)^2 + 1\right)^\ell}.$$

¹⁴Wegen $x^2 + 2bx + c = (x+b)^2 + (c-b^2) = (x+b+i\sqrt{c-b^2})(x+b-i\sqrt{c-b^2})$ könnten diese Terme auf Terme der Form $(x-z_0)^{-\ell}$ mit $z_0 \in \mathbb{C} \setminus \mathbb{R}$ zurückgeführt werden, die man entsprechend wie $(x-x_0)^{-1}$ behandeln könnte, mit dem Unterschied, dass man zunächst komplexe Stammfunktionen findet, die man zu reellen Funktionen zusammenfassen kann.

Für $\ell = 1$ ist eine Stammfunktion von $\frac{1}{x^2 + 2bx + c}$ gegeben durch

$$\frac{1}{\sqrt{c-b^2}} \arctan \frac{x+b}{\sqrt{c-b^2}}.$$

22:10.01.08

Für $\ell \geq 2$ erhält man aus folgender Rekursion

$$\begin{aligned} \int \frac{1}{(x^2 + 2bx + c)^{\ell-1}} dx &= \int \frac{1}{(x^2 + 2bx + c)^{\ell-1}} \cdot 1 dx \\ &\quad (\text{mit } x+b \text{ als Stammfunktion von } 1) \\ &= \frac{x+b}{(x^2 + 2bx + c)^{\ell-1}} + 2(\ell-1) \int \frac{(x+b)^2}{(x^2 + 2bx + c)^\ell} dx \\ &\quad (\text{mit } (x+b)^2 = (x^2 + 2bx + c) - (c-b^2)) \\ &= \frac{x+b}{(x^2 + 2bx + c)^{\ell-1}} + 2(\ell-1) \int \frac{1}{(x^2 + 2bx + c)^{\ell-1}} dx \\ &\quad - 2(\ell-1) \int \frac{c-b^2}{(x^2 + 2bx + c)^\ell} dx \end{aligned}$$

durch Auflösung nach dem letzten Term:

$$\int \frac{1}{(x^2 + 2bx + c)^\ell} dx = \frac{1}{2(\ell-1)(c-b^2)} \left\{ [2(\ell-1) - 1] \int \frac{1}{(x^2 + 2bx + c)^{\ell-1}} dx + \frac{x+b}{(x^2 + 2bx + c)^{\ell-1}} \right\}$$

Für $\ell = 2$ erhält man damit eine Stammfunktion von $\frac{1}{(x^2 + 2bx + c)^\ell}$

$$\frac{1}{2(c-b^2)} \frac{x+b}{x^2 + 2bx + c} + \frac{1}{2(c-b^2)^{3/2}} \arctan \frac{x+b}{\sqrt{c-b^2}}.$$

Hiermit erhält man insgesamt:

— Eine Stammfunktion von $\frac{Bx+C}{x^2 + 2bx + c}$ ist

$$\frac{B}{2} \ln(x^2 + 2bx + c) + \frac{C-Bb}{\sqrt{c-b^2}} \arctan \frac{x+b}{\sqrt{c-b^2}}.$$

— Eine Stammfunktion von $\frac{Bx+C}{(x^2 + 2bx + c)^2}$ ist

$$-\frac{B}{2} (x^2 + 2bx + c)^{-1} + \frac{C-Bb}{2(c-b^2)} \frac{x+b}{x^2 + 2bx + c} + \frac{C-Bb}{2(c-b^2)^{3/2}} \arctan \frac{x+b}{\sqrt{c-b^2}}.$$

(Für $\ell > 2$ kann man rekursiv ebenfalls Stammfunktionen bestimmen; das ist aber für die Praxis kaum von Bedeutung.)

Es bleibt also zu zeigen, dass jede (reelle) rationale Funktion auf Terme dieser Art zurückgeführt werden kann. Hierzu benutzen wir den

Satz 12.15 (Fundamentalsatz der Algebra) Sei $p(z) = \sum_{k=0}^n a_k z^k$ ein (i. allg. komplexes) Polynom vom Grad n ($\deg p = n$, $a_n \neq 0$). Dann existieren $z_j \in \mathbb{C}$ und $\ell_j \in \mathbb{N}$ (die Nullstellen des Polynoms p und deren Vielfachheiten) mit

$$p(z) = a_n \prod_{j=1}^k (z - z_j)^{\ell_j}, \quad \sum_{j=1}^k \ell_j = n.$$

Beweis. (Der folgende einfache Beweis geht auf R. ARGAND (1768–1822) zurück; er wurde 1814 veröffentlicht und 1820 von Cauchy vereinfacht. Ein völlig anderer Beweis wird üblicherweise in der Funktionentheorie mit Hilfe des Satzes von Liouville gegeben.)

a) Wir zeigen zunächst, dass p mindestens ein Nullstelle hat. Wegen

$$\begin{aligned} |p(z)| &= |z|^n \left| a_n + z^{-1} a_{n-1} + \dots + z^{-n} a_0 \right| \\ &\geq |z|^n \left(|a_n| - \frac{1}{|z|} \left| a_{n-1} + z^{-1} a_{n-2} + \dots + z^{1-n} a_0 \right| \right) \end{aligned}$$

für $z \neq 0$, existiert ein $r > 0$ mit

$$|p(z)| \geq \frac{|a_n|}{2} |z|^n > |a_0| = |p(0)| \quad \text{für } |z| \geq r.$$

Also nimmt $|p(z)|$ in der Kreisscheibe $|z| \leq r$ sein Minimum an (eine stetige Funktion auf einer beschränkten abgeschlossenen Teilmenge von \mathbb{R}^2 bzw. \mathbb{C} nimmt ihr Minimum an).

Wir zeigen $p(z_0) = 0$, d. h. z_0 ist Nullstelle von p . Dazu nehmen wir $p(z_0) \neq 0$ an; nach Multiplikation des Polynoms mit einer geeigneten komplexen Zahl vom Betrag 1 können wir sogar annehmen, dass $p(z_0) > 0$ ist. Offenbar ist $p(z_0 + w)$ ein Polynom vom Grad n in w mit konstantem Glied $p(z_0)$,

$$p(z_0 + w) = p(z_0) + b_k w^k + w^{k+1} \tilde{p}(w)$$

mit $1 \leq k \leq n$, $b_k \neq 0$ und einem Polynom \tilde{p} vom Grad $n - k - 1$ (dabei ist $\tilde{p} = 0$, falls $k = n$ ist; in diesem Fall ist der Rest des Beweises etwas einfacher).

Wählen wir nun w_0 so, dass $b_k w_0^k = -p(z_0)$ gilt ($w_0 := \sqrt[k]{-p(z_0)/b_k}$), so gilt für $t \in \mathbb{R}$

$$p(z_0 + t w_0) = p(z_0) + b_k t^k w_0^k + t^{k+1} w_0^{k+1} \tilde{p}(t w_0) = (1 - t^k) p(z_0) + t^{k+1} w_0^{k+1} \tilde{p}(t w_0).$$

Für hinreichend kleine $t > 0$ gilt (beachte $p(z_0) > 0$)

$$|t^{k+1}\tilde{p}(tw_0)| < t^k p(z_0), \quad \text{also } |p(z_0 + tw_0)| < p(z_0).$$

Das ist ein Widerspruch zur Wahl von z_0 als Minimum von $|p(z)|$.

b) Ist z_0 eine Nullstelle von p , so liefert Polynomdivision $p(z) = p_{n-1}(z)(z - z_0)$ (der Rest ist 0, da z_0 Nullstelle von p ist) mit einem Polynom $p_{n-1}(z)$ vom Grad $n - 1$ mit höchstem Koeffizienten a_n . Iteration dieses Verfahrens liefert nach n Schritten die Behauptung. ■

Satz 12.16 (Komplexe Partialbruchzerlegung) Sei $r(z) = \frac{q(z)}{p(z)}$ mit Polynomen q und p , $p(\cdot)$ wie im obigen Satz. Dann gibt es eindeutig bestimmte Zahlen $A_{ij} \in \mathbb{C}$ und ein Polynom $h(\cdot)$ mit

$$\begin{aligned} r(z) = h(z) &+ \frac{A_{11}}{z - z_1} + \frac{A_{12}}{(z - z_1)^2} + \dots + \frac{A_{1\ell_1}}{(z - z_1)^{\ell_1}} \\ &\quad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ &+ \frac{A_{k1}}{z - z_k} + \frac{A_{k2}}{(z - z_k)^2} + \dots + \frac{A_{k\ell_k}}{(z - z_k)^{\ell_k}}. \end{aligned}$$

Bemerkung Zur Integration (komplexer) rationaler Funktionen kann man sich also offenbar auf die obigen Fälle 1 und 2 beschränken; dabei muß man dann allerdings die komplexe Logarithmusfunktion beherrschen.

Beweis von Satz 12.16. Existenz: $h(\cdot)$ ergibt sich durch Polynomdivision mit Rest:

$$r(z) = h(z) + \frac{q_1(z)}{p(z)}, \quad \deg q_1 < \deg p.$$

Also können wir o. E. voraussetzen

$$r(z) = \frac{q(z)}{p(z)} \quad \text{mit } \deg q < \deg p.$$

Wir führen den Beweis durch Induktion nach $n = \deg p$.

Für $\mathbf{n} = \mathbf{1}$, d.h. $r(z) = \frac{c}{z - z_1}$, ist nichts zu beweisen.

$\mathbf{n} - \mathbf{1} \Rightarrow \mathbf{n}$: Die Behauptung sei also richtig für $\deg p \leq n - 1$ und somit $\deg q \leq n - 2$. Habe jetzt p den Grad n , also q den Grad $\leq n - 1$; z_0 sei eine Nullstelle von p der Ordnung $\ell \geq 1$, d. h. es gilt

$$p(z) = (z - z_0)^\ell \tilde{p}(z) \quad \text{mit } \deg \tilde{p} \leq n - 1, \tilde{p}(z_0) \neq 0, \ell \geq 1.$$

Daraus folgt (für z mit $p(z) \neq 0$), da z_0 Nullstelle von $q(z)\tilde{p}(z_0) - \tilde{p}(z)q(z_0)$ ist,

$$\frac{q(z)}{\tilde{p}(z)} - \frac{q(z_0)}{\tilde{p}(z_0)} = \frac{q(z)\tilde{p}(z_0) - \tilde{p}(z)q(z_0)}{\tilde{p}(z)\tilde{p}(z_0)} = \frac{(z - z_0)\tilde{q}(z)}{\tilde{p}(z)}$$

mit einem Polynom \tilde{q} mit $\deg \tilde{q} \leq n - 2$, also

$$\frac{q(z)}{p(z)} = \frac{q(z)}{(z - z_0)^\ell \tilde{p}(z)} = \frac{q(z_0)}{\tilde{p}(z_0)} \frac{1}{(z - z_0)^\ell} + \frac{\tilde{q}(z)}{(z - z_0)^{\ell-1} \tilde{p}(z)}.$$

Auf den letzten Term kann die Induktionsvoraussetzung angewandt werden.

Eindeutigkeit: Seien zwei solche Darstellungen mit A_{ij}, A'_{ij} gegeben. Multiplikation mit $(z - z_i)^{\ell_i}$ liefert, wenn man $z = z_i$ setzt, $A_{i\ell_i} = A'_{i\ell_i}$ (weil für $z = z_i$ alle anderen Terme verschwinden). Anschließend verfährt man entsprechend mit $(z - z_i)^{\ell_i-1}, (z - z_i)^{\ell_i-2}$, und erhält $A_{i,\ell_i-1} = A'_{i,\ell_i-1}, \dots$. Schließlich ergibt sich die Eindeutigkeit von $h(\cdot)$. ■

Satz 12.17 (Reelle Partialbruchzerlegung) Sind p und q reelle Polynome, so erhält man anstelle der Resultate der beiden vorhergehenden Sätze auch:

$$a) \quad p(x) = a_n \prod_{j=1}^r (x - x_j)^{\ell_j} \prod_{j=1}^k (x^2 + 2b_j x + c_j)^{m_j}$$

mit $x_j, b_j, c_j \in \mathbb{R}, \sum_{j=1}^r \ell_j + \sum_{j=1}^k 2m_j = n$ und $b_j^2 < c_j$.

b) Die Partialbruchzerlegung von $r(x)$ hat die Form

$$\begin{aligned} r(x) &= \frac{q(x)}{p(x)} = h(x) + \frac{A_{11}}{x - x_1} + \dots + \frac{A_{1\ell_1}}{(x - x_1)^{\ell_1}} \\ &\quad \vdots \\ &+ \frac{A_{r1}}{(x - x_r)} + \dots + \frac{A_{r\ell_r}}{(x - x_r)^{\ell_r}} \\ &+ \frac{B_{11}x + C_{11}}{x^2 + 2b_1x + c_1} + \dots + \frac{B_{1m_1}x + C_{1m_1}}{(x^2 + 2b_1x + c_1)^{m_1}} \\ &\quad \vdots \\ &+ \frac{B_{k1}x + C_{k1}}{x^2 + 2b_kx + c_k} + \dots + \frac{B_{km_k}x + C_{km_k}}{(x^2 + 2b_kx + c_k)^{m_k}} \\ &= h(x) + \sum_{j=1}^r \sum_{\ell=1}^{\ell_j} \frac{A_{j\ell}}{(x - x_j)^\ell} + \sum_{j=1}^k \sum_{m=1}^{m_j} \frac{B_{jm}x + C_{jm}}{(x^2 + 2b_jx + c_j)^m} \end{aligned}$$

mit einem reellen Polynom h und $A_{j\ell}, B_{jm}, C_{jm} \in \mathbb{R}$.

Alle Größen sind eindeutig bestimmt.

23:15.01.2008

Beweis. a) Ergibt sich aus Satz 12.15 durch Zusammenfassen der Terme $(x-z_j)$ und $(x-\bar{z}_j) = \overline{x-z_j}$.

b) Zusammenfassung der Terme mit $x-z_j$ und $x-\bar{z}_j$ in Satz 12.16 ergibt (mit $x^2+2bx+c = (x-z_j)(x-\bar{z}_j)$ und einem Polynom g mit $\deg g < 2\ell$, wobei wir hier auf den Index j verzichten)

$$\begin{aligned} \frac{g(x)}{(x^2+2bx+c)^\ell} &= \frac{1}{(x^2+2bx+c)} \frac{g(x)}{(x^2+2bx+c)^{\ell-1}} \\ &\quad (\text{Polynomdivision mit } (x^2+2bx+c)^{\ell-1}) \\ &= \frac{A_1x+C_1}{x^2+2bx+c} + \frac{g_1(x)}{(x^2+2bx+c)^\ell} \quad \text{mit } \deg g_1 \leq 2(\ell-1) \\ &\quad (\text{Polynomdivision mit } (x^2+2bx+c)^{\ell-2}) \\ &= \frac{A_1x+C_1}{x^2+2bx+c} + \frac{A_2x+C_2}{(x^2+2bx+c)^2} + \frac{g_2(x)}{(x^2+2bx+c)^\ell} \\ &\quad \text{mit } \deg g_2 \leq 2(\ell-2) \\ &= \dots \end{aligned}$$

Dies liefert nach $\ell-1$ Schritten die Behauptung. ■

Beispiel 12.18 $\frac{x^4+2}{x^3-x}$.

Die Nullstellen des Nenners $x^3-x = x(x^2-1)$ sind: $x_1 = 0$, $x_2 = 1$, $x_3 = -1$.

Polynomdivision mit Rest ergibt:

$$(x^4+2) : (x^3-x) = x + \frac{x^2+2}{x(x^2-1)} \quad (h(x) = x).$$

Die Partialbruchzerlegung von $\frac{(x^2+2)}{x(x^2-1)}$ hat die Form:

$$\frac{x^2+2}{x(x^2-1)} = \frac{A_1}{x} + \frac{A_2}{x-1} + \frac{A_3}{x+1}.$$

Multiplikation mit dem Nenner $x(x^2-1)$ der linken Seite ergibt

$$x^2+2 = A_1x^2 - A_1 + A_2x^2 + A_2x + A_3x^2 - A_3x.$$

Koeffizientenvergleich ergibt:

$$x^2 : \quad 1 = A_1 + A_2 + A_3,$$

$$x^1 : \quad 0 = A_2 - A_3,$$

$$x^0 : \quad 2 = -A_1.$$

Die letzte Gleichung liefert $A_1 = -2$; die zweite Gleichung liefert $A_2 = A_3$; zusammen mit der ersten Gleichung folgt $A_2 = A_3 = \frac{3}{2}$, also:

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

und somit für das gesuchte Integral

$$\begin{aligned} \int \frac{x^4 + 2}{x(x^2 - 1)} dx &= \frac{x^2}{2} - 2 \ln|x| + \frac{3}{2} (\ln|x-1| + \ln|x+1|) \\ &= \frac{x^2}{2} + \ln \frac{(|x-1||x+1|)^{3/2}}{|x|^2}. \end{aligned}$$

□

Beispiel 12.19 $\frac{1}{1+x^4}$.

Die Nullstellen des Nenners $1+x^4$ sind:

$$\begin{aligned} z_1 &= \frac{1}{\sqrt{2}}(1+i), & z_2 &= \bar{z}_1 = \frac{1}{\sqrt{2}}(1-i), \\ z_3 &= \frac{1}{\sqrt{2}}(-1+i), & z_4 &= \bar{z}_3 = \frac{1}{\sqrt{2}}(-1-i), \end{aligned}$$

und somit

$$\begin{aligned} (x-z_1)(x-z_2) &= x^2 - \sqrt{2}x + 1 = x^2 + 2b_1x + c_1, & b_1 &= -\frac{1}{\sqrt{2}}, & c_1 &= 1, \\ (x-z_3)(x-z_4) &= x^2 + \sqrt{2}x + 1 = x^2 + 2b_2x + c_2, & b_2 &= \frac{1}{\sqrt{2}}, & c_2 &= 1, \end{aligned}$$

also

$$1+x^4 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1).$$

Die Partialbruchzerlegung hat also die Form

$$\frac{1}{1+x^4} = \frac{B_1x + C_1}{x^2 - \sqrt{2}x + 1} + \frac{B_2x + C_2}{x^2 + \sqrt{2}x + 1}.$$

Multiplikation der Gleichung mit dem Nenner $(1+x^4)$ der linken Seite liefert in geeigneter Anordnung der Terme:

$$\begin{aligned} 1 &= B_1x^3 + \sqrt{2}B_1x^2 + B_1x \\ &\quad + C_1x^2 + \sqrt{2}C_1x + C_1 \\ &+ B_2x^3 - \sqrt{2}B_2x^2 + B_2x \\ &\quad + C_2x^2 - \sqrt{2}C_2x + C_2. \end{aligned}$$

Koeffizientenvergleich ergibt:

$$\begin{aligned} x^3 : \quad 0 &= B_1 + B_2, \\ x^2 : \quad 0 &= \sqrt{2}B_1 - \sqrt{2}B_2 + C_1 + C_2, \\ x^1 : \quad 0 &= B_1 + B_2 + \sqrt{2}C_1 - \sqrt{2}C_2, \\ x^0 : \quad 1 &= C_1 + C_2. \end{aligned}$$

Hieraus erhält man leicht

$$C_1 = C_2 = \frac{1}{2}, \quad -B_2 = B_1 = -\frac{1}{2\sqrt{2}},$$

und somit die explizite Partialbruchzerlegung

$$\frac{1}{1+x^4} = \frac{-\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 - \sqrt{2}x + 1} + \frac{\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 + \sqrt{2}x + 1}.$$

Mit der oben unter Nr. 3 angegebenen Formel erhalten wir also für die Stammfunktion:

$$\begin{aligned} \int \frac{1}{1+x^4} dx &= \frac{B_1}{2} \ln(x^2 + 2b_1x + c_1) + \frac{C_1 - B_1b_1}{\sqrt{c_1 - b_1^2}} \arctan \frac{x + b_1}{\sqrt{c_1 - b_1^2}} \\ &+ \frac{B_2}{2} \ln(x^2 + 2b_2x + c_2) + \frac{C_2 - B_2b_2}{\sqrt{c_2 - b_2^2}} \arctan \frac{x + b_2}{\sqrt{c_2 - b_2^2}} \\ &= -\frac{1}{4\sqrt{2}} \ln(x^2 - \sqrt{2}x + 1) + \frac{1}{2\sqrt{2}} \arctan \sqrt{2}\left(x - \frac{1}{\sqrt{2}}\right) \\ &\quad + \frac{1}{4\sqrt{2}} \ln(x^2 + \sqrt{2}x + 1) + \frac{1}{2\sqrt{2}} \arctan \sqrt{2}\left(x + \frac{1}{\sqrt{2}}\right). \end{aligned}$$

□

13 Integration und Grenzübergänge

13.1 Vertauschung von Integration und Differentiation mit Grenzübergängen

Eine wichtige Frage in jeder Integrationstheorie ist die folgende: Kann aus $f_n \rightarrow f$ auf $\int f_n \rightarrow \int f$ geschlossen werden? Ist (ggf. unter welchen Bedingungen) der Grenzübergang mit der Integration vertauschbar?

Satz 13.1 Sind $f_n : [a, b] \rightarrow \mathbb{R}$ ($n \in \mathbb{N}$) Regelfunktionen. $f : [a, b] \rightarrow \mathbb{R}$ mit $\|f - f_n\|_\infty \rightarrow 0$ für $n \rightarrow \infty$, so ist f Regelfunktion und es gilt

$$\int_a^b f_n(x) \, dx \rightarrow \int_a^b f(x) \, dx.$$

Beweis. Sei $\varepsilon > 0$. Dann existiert ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$\|f - f_n\|_\infty < \frac{\varepsilon}{2} \quad \text{für } n \geq N.$$

Für alle n seien $(\varphi_{n,m})_m$ Folgen von Treppenfunktionen auf $[a, b]$ mit $\|f_n - \varphi_{n,m}\|_\infty \rightarrow 0$ für $m \rightarrow \infty$. Zu jedem $n \in \mathbb{N}$ existiert ein $m(n) \in \mathbb{N}$ mit

$$\|f_n - \varphi_{n,m(n)}\|_\infty < \frac{\varepsilon}{2},$$

also

$$\|f - \varphi_{n,m(n)}\|_\infty \leq \|f - f_n\|_\infty + \|f_n - \varphi_{n,m(n)}\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{für } n \geq N.$$

Somit ist f Regelfunktion, und es gilt

$$\begin{aligned} \left| \int_a^b f(x) \, dx - \int_a^b f_n(x) \, dx \right| &= \left| \int_a^b (f(x) - f_n(x)) \, dx \right| \leq \int_a^b |f(x) - f_n(x)| \, dx \\ &\leq (b-a) \|f - f_n\|_\infty \rightarrow 0. \end{aligned}$$

Das ist die Behauptung. ■

Als Folgerung erhalten wir aus dem vorhergehenden Satz:

Korollar 13.2 Sind $f_n : [a, b] \rightarrow \mathbb{R}$ Regelfunktionen und konvergiert die Reihe

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

gleichmäßig auf $[a, b]$, so ist f eine Regelfunktion mit

$$\int_a^b f(x) \, dx = \sum_{n=0}^{\infty} \int_a^b f_n(x) \, dx.$$

Korollar 13.3 Eine Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ mit Konvergenzradius $\varrho > 0$ konvergiert für $|x-x_0| \leq r$ mit $r < \varrho$ gleichmäßig. Insbesondere können also Potenzreihen im Inneren des Konvergenzintervalls gliedweise integriert werden,

$$\int_a^b \sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-x_0)^{n+1} \Big|_a^b$$

für a, b im Konvergenzbereich.

Beweis. Nur die gleichmäßige Konvergenz für $|x-x_0| < r$ ist zu beweisen. Wegen $\varrho = \left(\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}\right)^{-1}$ gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n(x-x_0)^n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} |x-x_0| \leq \frac{r}{\varrho} < 1 \quad \text{für } |x-x_0| < \varrho,$$

also die gleichmäßige Konvergenz für $|x-x_0| \leq r$. ■

Es läßt sich aber auch ein Resultat über die gliedweise Differentiation von Folgen und Reihen ableiten:

Satz 13.4 Sei $I \subset \mathbb{R}$ ein offenes Intervall. Sind f_n ($n \in \mathbb{N}$) in I stetig differenzierbar mit

$$f_n(x) \rightarrow f(x) \quad \text{für alle } x \in I \quad (\text{punktweise konvergent}),$$

$$f'_n(x) \rightarrow g(x) \quad \text{gleichmäßig auf jedem kompakten Teilintervall von } I,$$

so ist f in I stetig differenzierbar mit $f' = g$, d. h. es gilt

$$\frac{d}{dx} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{d}{dx} f_n(x) = \lim_{n \rightarrow \infty} f'_n(x).$$

Ist I abgeschlossen und gilt $f'_n(x) \rightarrow g(x)$ gleichmäßig auf ganz I , so gilt auch die Aussage auf ganz I .

Beweis. Für $x_0 \in I$ gilt nach dem Hauptsatz

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(t) dt.$$

Für $n \rightarrow \infty$ folgt daraus mit den Voraussetzungen und dem obigen Satz

$$f(x) = f(x_0) + \int_{x_0}^x g(t) dt.$$

Da g (als gleichmäßiger Grenzwert stetiger Funktionen) stetig ist, ist f stetig differenzierbar mit $f' = g$. ■

Auch hier folgt das entsprechende Resultat für Reihen:

Korollar 13.5 Seien $f_n : I \rightarrow \mathbb{R}$ und $f(x) = \sum_{n=1}^{\infty} f_n(x)$ für $x \in I$. Sind die f_n stetig differenzierbar und ist $\sum_{n=1}^{\infty} f'_n(x)$ in jedem kompakten Teil von I gleichmäßig konvergent, so ist auch f stetig differenzierbar mit

$$\frac{d}{dx} \sum_{n=1}^{\infty} f_n(x) = f'(x) = \sum_{n=1}^{\infty} f'_n(x).$$

Speziell gilt: Hat die Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

einen Konvergenzradius $\rho > 0$, so ist f in $(x_0 - \rho, x_0 + \rho)$ beliebig oft differenzierbar und es gilt

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}, \quad f''(x) = \sum_{n=2}^{\infty} n(n-1) a_n (x - x_0)^{n-2} \quad \text{usw.}$$

Zum Beweis ist lediglich anzumerken, dass die „abgeleitete Reihe“ den gleichen Konvergenzradius ρ hat wie die ursprüngliche Reihe: Dies folgt aus

$${}^{n-1}\sqrt{n|a_n|} = {}^{n-1}\sqrt{n} \cdot {}^{n-1}\sqrt{|a_n|} = \left(\sqrt[n]{n} \cdot \sqrt[n]{|a_n|} \right)^{\frac{n}{n-1}} \quad \text{und} \quad \sqrt[n]{n} \rightarrow 1.$$

Damit wird nun der Beweis für die Konvergenz der **Logarithmusreihe** (Fußnote in Kapitel 10) nachvollziehbar. Entsprechend können jetzt noch weitere Entwicklungen angegeben werden:

Beispiel 13.6 Arcus-Tangens-Reihe: Aus

$$\arctan' x = \frac{1}{1+x^2} \quad \text{für } |x| < 1$$

und (geometrische Reihe)

$$\frac{1}{1+t^2} = \sum_{n=0}^{\infty} (-1)^n t^{2n}, \quad \text{gleichmäßig für } |t| \leq |x| < 1$$

folgt

$$\arctan x = \int_0^x \frac{1}{1+t^2} dt = \int_0^x \sum_{n=0}^{\infty} (-1)^n t^{2n} dt = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad \text{für } |x| < 1.$$

□

Beispiel 13.7 Arcus–Sinus–Reihe: Aus

$$\arcsin' x = \frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-\frac{1}{2}}$$

und der Binominalreihe

$$(1+t)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} t^n, \quad \text{gleichmäßig für } |t| \leq |x| < 1$$

folgt

$$\arcsin x = \int_0^x (1+t^2)^{-\frac{1}{2}} dt = \int_0^x \sum_{n=0}^{\infty} \binom{-1/2}{n} t^{2n} = \sum_{n=0}^{\infty} \binom{-1/2}{n} \frac{x^{2n+1}}{2n+1} \quad \text{für } |x| < 1.$$

□

Riemann–Summen

Häufig ist es nützlich zu wissen, dass das Integral durch sogenannte Riemann–Summen approximiert werden kann. Eine *Zerlegung* Z des Intervalls $[a, b]$ mit *Stützstellen* ist gegeben durch

$$\{x_0, x_1, \dots, x_k; x_1^*, \dots, x_k^*\} \quad \text{mit } x_j < x_{j+1} \quad \text{und } a = x_0 \leq x_1^* \leq x_1 \leq x_2^* \leq \dots \leq x_k^* \leq x_k = b.$$

$[a, b]$ wird dabei zerlegt in Intervalle (x_{j-1}, x_j) mit Stützstellen x_j^* im jeweiligen Intervall. Die *Feinheit* der Zerlegung ist $\eta(Z) := \max\{|x_j - x_{j-1}| : j = 1, \dots, k\}$.

Die durch die Zerlegung Z definierte Riemann–Summe für das Integral $\int_a^b f(x) dx$ ist

$$S_Z(f) := \sum_{j=1}^k f(x_j^*)(x_j - x_{j-1}).$$

Im folgenden Satz wird eine Folge (Z_n) von Zerlegungen

$$Z_n = \{x_{n,0}, x_{n,1}, \dots, x_{n,k(n)}; x_{n,1}^*, \dots, x_{n,k(n)}^*\}$$

betrachtet mit $\eta(Z_n) \rightarrow 0$ für $n \rightarrow \infty$.

Satz 13.8 (Riemann–Summen) Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Regelfunktion, (Z_n) eine Folge von Zerlegungen des Intervalls $[a, b]$ mit $\eta(Z_n) \rightarrow 0$ für $n \rightarrow \infty$. Dann gilt

$$S_{Z_n}(f) \rightarrow \int_a^b f(x) dx.$$

(Dieser Satz gilt völlig gleichlautend auch für die etwas größere Klasse der Riemann–integrierbaren Funktionen.)

Beweis. Sei $\varepsilon > 0$ vorgegeben. Da f Regelfunktion ist, existiert eine Treppenfunktion φ mit

$$\|f - \varphi\|_\infty < \frac{\varepsilon}{3(b-a)}.$$

Für dieses φ gilt offensichtlich

$$\int_a^b |f(x) - \varphi(x)| dx < \frac{\varepsilon}{3} \quad \text{und} \quad |S_{Z_n}(f) - S_{Z_n}(\varphi)| < \frac{\varepsilon}{3} \quad \text{für alle } n.$$

Außerdem gibt es für diese Treppenfunktion φ ein N mit

$$\left| S_{Z_n}(\varphi) - \int_a^b \varphi(x) dx \right| < \frac{\varepsilon}{3} \quad \text{für } n \geq N$$

(da dies „offenbar“ für alle Zerlegungen mit hinreichender Feinheit gilt ($\ddot{\text{U}}$)). Damit folgt insgesamt

$$\begin{aligned} & \left| \int_a^b f(x) dx - S_{Z_n}(f) \right| \\ & \leq \left| \int_a^b f(x) dx - \int_a^b \varphi(x) dx \right| + \left| \int_a^b \varphi(x) dx - S_{Z_n}(\varphi) \right| + |S_{Z_n}(\varphi) - S_{Z_n}(f)| \\ & < \varepsilon \quad \text{für } n \geq N. \end{aligned}$$

Also gilt $S_{Z_n}(f) \rightarrow \int_a^b f(x) dx$. ■

Beispiel 13.9 $\int_0^1 e^x dx$, $Z_n := \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n} = 1; \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}$.

$$\begin{aligned} S_{Z_n}(\exp) &= \sum_{k=1}^n e^{k/n} \cdot \frac{1}{n} = \frac{1}{n} \sum_{k=1}^n (e^{1/n})^k = \frac{1}{n} \frac{1 - e^{\frac{n+1}{n}}}{1 - e^{1/n}} \\ &= \frac{1/n}{1 - e^{1/n}} \left(1 - e^{\frac{n+1}{n}} \right) \rightarrow -1(1 - e) = e - 1. \end{aligned}$$

Das liefert natürlich auch der Hauptsatz:

$$\int_0^1 e^x dx = e^x \Big|_0^1 = e - 1.$$

□

13.2 Uneigentliche Integrale

Integrale der Form $\int_0^{\infty} e^{-x} dx$ oder $\int_0^1 \frac{1}{\sqrt{x}} dx$ sind nicht definiert, da das Intervall unbeschränkt ist bzw. der Integrand am Rand unbeschränkt ist. Man kann das in diesen Fällen so versuchen:

$$\begin{aligned} - \int_0^{\infty} e^{-x} dx &:= \lim_{b \rightarrow \infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow \infty} \left(-e^{-x} \right)_0^b = \lim_{b \rightarrow \infty} \left(-e^{-b} + e^{-0} \right) = 1, \\ - \int_0^1 \frac{1}{\sqrt{x}} dx &:= \lim_{a \rightarrow 0+} \int_a^1 \frac{1}{\sqrt{x}} dx = \lim_{a \rightarrow 0+} 2\sqrt{x} \Big|_a^1 = \lim_{a \rightarrow 0+} 2 \left(1^{1/2} - a^{1/2} \right) = 2. \end{aligned}$$

Allgemeiner betrachten wir die folgende Situation: $f : (a, b) \rightarrow \mathbb{R}$ sei über jedem Intervall $[c, d] \subset (a, b)$ eine Regelfunktion. Das Integral $\int_a^b f(x) dx$ heißt

- bei b *uneigentlich*, wenn
 - $b = \infty$ ist, oder/und
 - f über $[c, b]$ mit $c \in (a, b)$ keine Regelfunktion ist,
- bei a *uneigentlich*, wenn ... (entsprechend).

Ist $\int_a^b f(x) dx$ uneigentlich bei b , so definiert man das *uneigentliche Integral* $\int_a^b f(x) dx$ durch

$$\lim_{c \rightarrow b-} \int_a^c f(x) dx, \quad \text{falls der Limes existiert.}$$

Entsprechend, falls das Integral bei a uneigentlich ist:

$$\lim_{c \rightarrow a+} \int_c^b f(x) dx, \quad \text{falls der Limes existiert.}$$

24: 17.01.08

Ist das Integral bei a **und** bei b uneigentlich, so existiert das uneigentliche Integral, wenn die beiden uneigentlichen Integrale $\int_a^c f(x) dx$ **und** $\int_c^b f(x) dx$ für $c \in (a, b)$ existieren, und man setzt

$$\int_a^b f(x) dx = \lim_{\alpha \rightarrow a} \int_{\alpha}^c f(x) dx + \lim_{\beta \rightarrow b} \int_c^{\beta} f(x) dx.$$

Die obigen Beispiele waren bei $+\infty$ bzw. 0 uneigentlich, und die uneigentlichen Integrale existieren.

Beispiel 13.10 $\int_1^{\infty} x^{\alpha} dx$ ($\alpha \in \mathbb{R}$) ist uneigentlich bei ∞ ,

$$\int_1^{\infty} x^{\alpha} dx = \lim_{c \rightarrow \infty} \int_1^c x^{\alpha} dx = \lim_{c \rightarrow \infty} \left\{ \begin{array}{l} \frac{1}{\alpha+1} x^{\alpha+1} \Big|_1^c \\ \ln x \Big|_1^c \\ \frac{1}{\alpha+1} x^{\alpha+1} \Big|_1^c \end{array} \right\} = \begin{cases} \text{divergent für } \alpha > -1, \\ \text{divergent für } \alpha = -1, \\ \frac{-1}{\alpha+1} \text{ für } \alpha < -1. \end{cases}$$

□

Beispiel 13.11 $\int_0^1 x^{\alpha} dx$ ist für $\alpha < 0$ uneigentlich bei 0,

$$\int_0^1 x^{\alpha} dx = \lim_{c \rightarrow 0} \left\{ \begin{array}{l} \frac{1}{\alpha+1} x^{\alpha+1} \Big|_c^1 \\ \ln x \Big|_c^1 \\ \frac{1}{\alpha+1} x^{\alpha+1} \Big|_c^1 \end{array} \right\} = \begin{cases} \frac{1}{\alpha+1} & \text{für } \alpha > -1, \\ \text{divergent} & \text{für } \alpha = -1, \\ \text{divergent} & \text{für } \alpha < -1, \end{cases}$$

□

Zur Untersuchung der Existenz uneigentlicher Integrale, die man nicht explizit berechnen kann, ist folgender Satz nützlich:

Satz 13.12 Das bei b uneigentliche Integral $\int_a^b g(x) dx$ existiere. $f : [a, b] \rightarrow \mathbb{R}$ sei über jedem Intervall $[a, c]$ mit $c \in (a, b)$ Regelfunktion und es gelte $|f(x)| \leq g(x)$ für alle $x \in [a, b]$. Dann existiert das uneigentliche Integral $\int_a^b f(x) dx$ und es gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Beweis. Es genügt zu zeigen, dass für jede Folge (b_n) mit $b_n \rightarrow b-$ die Folge $\left(\int_a^{b_n} f(x) dx \right)$ eine Cauchyfolge (also konvergent) ist:

$$\left| \int_a^{b_n} f(x) dx - \int_a^{b_m} f(x) dx \right| = \left| \int_{b_n}^{b_m} f(x) dx \right| \leq \int_{b_n}^{b_m} g(x) dx \rightarrow 0 \quad \text{für } n, m \rightarrow \infty,$$

da g uneigentlich integrierbar ist. ■

Beispiel 13.13 Die Γ -Funktion (*Gamma-Funktion*) $\Gamma : (0, \infty) \rightarrow \mathbb{R}$ ist definiert durch

$$\Gamma(\alpha) := \int_0^{\infty} x^{\alpha-1} e^{-x} dx.$$

Dieses Integral ist

- für $\alpha < 1$ uneigentlich bei 0,
- für alle α uneigentlich bei ∞ .

$\int_0^1 x^{\alpha-1} e^{-x} dx$ existiert wegen $0 \leq x^{\alpha-1} e^{-x} \leq x^{\alpha-1}$ für $0 < x \leq 1$ ($\alpha - 1 > -1$).

$\int_1^{\infty} x^{\alpha-1} e^{-x} dx$ existiert, weil für $x > 1$ gilt

$$x^{\alpha-1} e^{-x} = \left(x^{\alpha-1} e^{-\frac{x}{2}} \right) e^{-\frac{x}{2}} \quad \text{und} \quad x^{\alpha-1} e^{-\frac{x}{2}} \leq c \quad \text{für} \quad x \text{ groß.}$$

□

Satz 13.14 *Es gilt*

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \quad \text{für alle } \alpha > 0,$$

$$\Gamma(n + 1) = n! \quad \text{für } n \in \mathbb{N}_0.$$

Beweis. Man rechnet leicht nach:

$$\begin{aligned} \Gamma(\alpha + 1) &= \lim_{n \rightarrow \infty} \int_{1/n}^n x^{\alpha} e^{-x} dx = \lim_{n \rightarrow \infty} \left\{ -x^{\alpha} e^{-x} \Big|_{1/n}^n + \int_{1/n}^n \alpha x^{\alpha-1} e^{-x} dx \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ -n^{\alpha} e^{-n} + n^{-\alpha} e^{-\frac{1}{n}} + \alpha \int_{1/n}^n x^{\alpha-1} e^{-x} dx \right\} \\ &= 0 + 0 + \alpha \Gamma(\alpha) = \alpha \Gamma(\alpha). \\ \Gamma(0 + 1) &= \Gamma(1) = \int_0^x e^{-x} dx = 1 = 0! \end{aligned}$$

Die verbleibende Aussage folgt daraus induktiv: $\Gamma(n + 1) = n\Gamma(n) = n(n - 1)! = n!$. ■

Bemerkung 13.15 Die Γ -Funktion ist stetig (sogar beliebig oft differenzierbar) auf $(0, \infty)$, also eine stetige Interpolation der Fakultät. Auf den etwas technischen Beweis soll hier verzichtet werden. (vgl. W. Walter *Analysis 1*, §12.8).

Schließlich soll noch ein nützlicher Zusammenhang mit der Konvergenz von Reihen bewiesen werden:

Satz 13.16 (Integralkriterium) Sei $f : [1, \infty) \rightarrow \mathbb{R}$ positiv und nicht-wachsend. Dann gilt

$$\sum_{k=1}^{\infty} f(k) \text{ konvergent} \Leftrightarrow \int_1^{\infty} f(x) dx \text{ existiert.}$$

Beweis. Die Folge (s_n) mit $s_n := \sum_{k=1}^n f(k)$ und die Funktion $F(x) := \int_1^x f(t) dt$ sind nicht-fallend. Wegen

$$f(k+1) \leq f(t) \leq f(k) \text{ für } k \leq t \leq k+1$$

gilt

$$f(k+1) \leq \int_k^{k+1} f(t) dt \leq f(k),$$

also

$$\sum_{k=2}^n f(k) \leq \int_1^n f(t) dt \leq \sum_{k=1}^{n-1} f(k)$$

bzw.

$$s_n - f(1) = \sum_2^n f(t) dt \leq F(n) \leq s_{n-1}.$$

Also ist (s_n) genau dann beschränkt, wenn $(F(x))$ beschränkt ist, also (s_n) genau dann konvergent, wenn $\int_1^{\infty} f(x) dx$ existiert. ■

Beispiel 13.17 $\sum_{n=1}^{\infty} n^{\alpha}$ konvergent $\Leftrightarrow \alpha < -1$, denn $\int_1^{\infty} x^{\alpha} dx$ existiert genau dann, wenn $\alpha < -1$ ist. □

Beispiel 13.18 $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^s}$ konvergent $\Leftrightarrow s > 1$.

a) $s \leq 1$: $\frac{1}{x \ln x}$ ist auf $(1, \infty)$ fallend, denn $x \ln x$ ist wachsend. Da außerdem

$$\int_2^b \frac{1}{x \ln x} dx = \ln(\ln x) \Big|_2^b \rightarrow \infty \text{ für } b \rightarrow \infty$$

gilt, ist $\sum \frac{1}{n \ln n}$ divergent, und damit folgt die Aussage auch für $s < 1$.

b) $s > 1$: $\frac{1}{x(\ln x)^s}$ ist auf $(1, \infty)$ fallend, denn $x(\ln x)^s$ ist wachsend. Da außerdem

$$\int_2^b \frac{1}{x(\ln x)^s} dx = \frac{1}{1-s} (\ln x)^{1-s} \Big|_2^b \rightarrow \frac{-1}{1-s} (\ln 2)^{1-s} \quad \text{für } b \rightarrow \infty$$

gilt, ist die Reihe konvergent für $s > 1$.

□

14 Vektorräume, Dimension, Basis

Wir **erinnern** uns: Ein *Vektorraum* $V = (V, +, \cdot)$ über dem Körper \mathbb{K} (\mathbb{K} -*Vektorraum*) ist eine nicht-leere Menge V , auf der eine *Addition* (von Elementen aus V) und eine *Multiplikation* (von Elementen aus V mit Elementen aus \mathbb{K})

$$+ : V \times V \rightarrow V, \quad (u, v) \mapsto u + v,$$

$$\cdot : \mathbb{K} \times V \rightarrow V, \quad (\alpha, u) \mapsto \alpha \cdot u \text{ (auch } \times \text{ statt } \cdot)$$

definiert sind so, dass $(V, +)$ eine *abelsche (kommutative) Gruppe* ist, d. h.

- es existiert ein *neutrales Element* bezüglich der Addition (*Nullelement*) Θ mit $u + \Theta = \Theta + u = u$ für alle $u \in V$,
- zu jedem $u \in V$ existiert ein bezüglich der Addition *inverses Element* (das zu u negative Element) \tilde{u} mit $\tilde{u} + u = 0$; statt \tilde{u} schreibt man auch $-u$, statt $v + \tilde{u} = v + (-u)$ auch $v - u$,
- es gilt das *Assoziationsgesetz* $(u + v) + w = u + (v + w)$ für alle $u, v, w \in V$,
- die Addition ist kommutativ, d. h. es gilt $u + v = v + u$ für alle $u, v \in V$.

Bezüglich der beiden Operationen $+$ und \cdot gilt für $\alpha, \beta \in \mathbb{K}$ und $u, v \in V$:

- $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$,
- $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$,
- $(\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot u)$,
- $1 \cdot u = u$.

Bemerkung 14.1 *In der Schreibweise unterscheiden wir nicht zwischen den verschiedenen Bedeutungen von $+$ (in V und in \mathbb{K}) bzw. von \cdot (in \mathbb{K} und zwischen \mathbb{K} und V). Ebenso unterscheiden wir i. allg. nicht zwischen der Null 0 aus \mathbb{K} und der Null Θ aus V (nur gelegentlich schreiben wir Θ für die Null aus V . Den „-“Punkt lassen wir meist weg.)*

Einige Eigenschaften:

- (1) $0 \cdot v = \Theta$ und $\alpha \cdot \Theta = \Theta$ für alle $\alpha \in \mathbb{K}, v \in V$:
 $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v; \alpha \cdot \Theta = \alpha(\Theta + \Theta) = \alpha \cdot \Theta + \alpha \cdot \Theta.$

- (2) $\alpha \cdot u = \Theta \Leftrightarrow \alpha = 0$ oder $u = \Theta$:
 \Rightarrow : ist $\alpha \neq 0$, so ist $u = 1 \cdot u = \frac{1}{\alpha} \cdot (\alpha \cdot u) = \frac{1}{\alpha} \Theta = \Theta$.
 \Leftarrow : Eigenschaft (1)
- (3) $(-\alpha) \cdot u = \alpha \cdot (-u) = -(\alpha \cdot u)$, insbesondere $(-1) \cdot u = -u$:
 $\alpha \cdot u + (-\alpha) \cdot u = (\alpha - \alpha) \cdot u = 0 \cdot u = \Theta$,
 $\alpha \cdot u + \alpha(-u) = \alpha \cdot (u - u) = \alpha \cdot \Theta = \Theta$.
- (4) $\alpha \cdot (u - v) = \alpha u - \alpha v$: $\alpha \cdot (u - v) = \alpha u + \alpha(-v) = \alpha u - \alpha v$.

Wichtige Vektorräume sind z. B. (dabei dürfte in der Regel klar sein, was das Nullelement bzw. das negative Element ist).

- $\mathbb{R}^n = \{(x_1, \dots, x_n) : x_j \in \mathbb{R}\}$ ist ein \mathbb{R} -Vektorraum mit

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n),$$

$$\alpha \cdot (x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n).$$

- $\mathbb{C}^n = \{(z_1, \dots, z_n) : z_j \in \mathbb{C}\}$ ist entsprechend ein \mathbb{C} -Vektorraum.
- Die reellen/komplexen Polynome $P_{\mathbb{R}}/P_{\mathbb{C}}$ bzw. die reellen/komplexen Polynome vom Grad $\leq n$ $P_{\mathbb{R},n}/P_{\mathbb{C},n}$ sind \mathbb{R} -/ \mathbb{C} -Vektorräume.
- Die Menge $\text{Abb}(X, \mathbb{K})$ der Abbildungen von einer (beliebigen) Menge X nach \mathbb{K} ($= \mathbb{R}$ oder \mathbb{C}) bildet ein \mathbb{R} -/ \mathbb{C} -Vektorraum mit

$$(f + g)(x) = f(x) + g(x), \quad (\alpha \cdot f)(x) = \alpha f(x).$$

Das Nullelement ist die *Nullfunktion* $f(x) = 0$ für alle $x \in X$; das zu f negative Element ist $(-f)(x) = -f(x)$ für alle $x \in X$.

Speziell:

- $\text{Abb}(\mathbb{N}, \mathbb{K})$ mit $\mathbb{K} = \mathbb{R}$ bzw. \mathbb{C} ist der Raum der reellen bzw. komplexen Folgen,
 - $\text{Abb}(\mathbb{R}, \mathbb{K})$, $\text{Abb}((a, b), \mathbb{K})$, ... ist der Raum der \mathbb{K} -wertigen Funktionen auf \mathbb{R} , (a, b) , ...
- $C[a, b]$ der Raum der reellen ($C_{\mathbb{R}}[a, b]$) bzw. komplexen ($C_{\mathbb{C}}[a, b]$) stetigen Funktionen auf $[a, b]$.
- $C^n[a, b]$, $C^n(\dots)$ der Raum der n mal stetig differenzierbaren Funktionen auf $[a, b]$, ...
- usw.

Sei V ein \mathbb{K} -Vektorraum. Eine Teilmenge $U \subset V$ heißt ein *Untervektorraum* (auch *Teilraum* oder *linearer Teilraum* von V), wenn $(U, +, \cdot)$ mit den auf V definierten Operationen $+$ und \cdot ein Vektorraum ist (d. h. wenn $+$ und \cdot nicht aus U herausführen; U ist bezüglich $+$ und \cdot „abgeschlossen“). Was ist zu zeigen, wenn man beweisen will, dass eine Teilmenge U eines Vektorraums ein Untervektorraum ist? Der folgende Satz besagt, dass man nur diese „Abgeschlossenheit“ bezüglich $+$ und \cdot zeigen muss.

Satz 14.2 *Sei V ein \mathbb{K} -Vektorraum, U eine Teilmenge von V . U ist genau dann ein Untervektorraum von V , wenn gilt:*

$$U \neq \emptyset, \quad \text{und aus } u, v \in U, \alpha \in \mathbb{K} \text{ folgt } u + v \in U \quad \text{und } \alpha u \in U.$$

Beweis. \Rightarrow : Offensichtlich.

\Leftarrow : Es ist nur die Existenz des Nullelements Θ (in U) bezüglich $+$ und des zu u negativen Elements $-u$ (des inversen Elements (in U) bezüglich $+$) zu zeigen:

Für beliebiges $u \in U$ (ein solches existiert wegen $U \neq \emptyset$) ist $0 \cdot u = \Theta \in U$.

Für $u \in U$ ist $(-1)u \in U$ und $(-1)u = -u$. ■

Für Elemente $x_1, \dots, x_n \in V$ heißt jedes Element der Form

$$\sum_{j=1}^n \alpha_j x_j \quad \text{mit} \quad \alpha_j \in \mathbb{K}$$

eine *Linearkombination* der Elemente x_1, \dots, x_n .

Die *lineare Hülle* $L(W)$ einer Teilmenge W von V ist der kleinste Teilraum von V , der W enthält. Offenbar gilt

Satz 14.3 *Ist V ein \mathbb{K} -Vektorraum und $W \subset V$, so gilt*

$$\begin{aligned} L(W) &= \cap \left\{ Z : Z \text{ Teilraum von } V, W \subset Z \right\} \\ &= \left\{ \sum_{j=1}^n \alpha_j x_j : \alpha_j \in \mathbb{K}, x_j \in W, n \in \mathbb{N} \right\}. \end{aligned}$$

Beweis. Die erste Gleichung ist offensichtlich. Die zweite Gleichung ergibt sich aus:

$$- \left\{ \sum_{j=1}^n \alpha_j x_j : \dots \right\} \text{ ist ein Teilraum von } V \text{ (Ü; vgl. obigen Satz), der } W \text{ enthält, also} \\ L(W) \subset \{ \dots \},$$

- Ist W in einem Teilraum Z enthalten, so sind auch alle Linearkombinationen von Elementen aus W in Z enthalten, also $\{\dots\} \subset L(W)$.

■

Korollar 14.4 W ist genau dann ein Teilraum, wenn $W = L(W)$ gilt.

Elemente x_1, \dots, x_n aus V heißen *linear unabhängig*, wenn aus $\sum_{j=1}^n \alpha_j x_j = 0$ mit $\alpha_j \in \mathbb{K}$ folgt $\alpha_j = 0$ für $j = 1, \dots, n$. Andernfalls heißen die Elemente x_1, \dots, x_n *linear abhängig*. Sind x_1, \dots, x_n linear abhängig, so läßt sich eines der Elemente als Linearkombination der anderen darstellen: Es gibt $(\alpha_1, \dots, \alpha_n) \neq (0, 0, \dots, 0)$ mit $\sum_{j=1}^n \alpha_j x_j = 0$; da mindestens ein $\alpha_{j_0} \neq 0$ ist, folgt

$$x_{j_0} = \sum_{\substack{j=1 \\ j \neq j_0}}^n \frac{\alpha_j}{\alpha_{j_0}} x_j.$$

Beliebig viele Elemente $x_\alpha (\alpha \in A)$ heißen *linear unabhängig*, wenn je endlich viele linear unabhängige sind. Sind die x_α linear abhängig, so läßt sich eines der x_α als Linearkombination von endlich vielen anderen darstellen.

25: 22.1.08

Bei linear unabhängigen Elementen wird die lineare Hülle echt kleiner, wenn man ein Element wegläßt. Sind sie linear abhängig, so kann man gewisse (mindestens aber eines) weglassen, ohne die lineare Hülle zu verkleinern.

Eine Teilmenge E von V heißt ein *erzeugendes System*, wenn $L(E) = V$ gilt.

Ein *minimales erzeugendes System* E (also ein erzeugendes System, bei dem Weglassen eines beliebigen Elements v dazu führt, dass $L(E \setminus \{v\})$ nicht mehr gleich V ist) ist also linear unabhängig. Man nennt ein solches System eine *Basis* von V (auch *hamel-Basis*, nach G.HAMEL). Eine Basis erzeugt also den ganzen Raum und es ist kein überflüssiges Element dabei.

Einige **Basen** lassen sich leicht angeben:

- in \mathbb{R}^m oder \mathbb{C}^m ist $\{e^j : j = 1, \dots, n\}$ mit $e^j := \{0, \dots, 0, 1, 0, \dots, 0\}$ mit der 1 an jeder j -ten Stelle einer Basis (linear unabhängig und erzeugend, \ddot{U}).
- $\{p_j : j = 0, \dots, n\}$ in $P_{\mathbb{K},n}$ bzw. $\{p_j : j \in \mathbb{N}_0\}$ in $P_{\mathbb{K}}$ mit $p_j(x) = x^j$ sind Basen (linear unabhängig und erzeugend, \ddot{U}).

So einfach ist es nicht immer:

Beispiel 14.5 \mathbb{R} kann natürlich auch als Vektorraum über \mathbb{Q} aufgefaßt werden (d. h. man erlaubt nur Multiplikation mit rationalen Zahlen. Dann sind z. B. die Elemente

$1, \sqrt{2}$ und $\sqrt{3}$ linear unabhängig.

Aus $a \cdot 1 + b\sqrt{2} + c \cdot \sqrt{3} = 0$ mit $a, b, c \in \mathbb{Q}$ folgt nämlich

$$a = -b\sqrt{2} - c\sqrt{3}, \quad a^2 = 2b^2 + 3c^2 + 2bc\sqrt{6}.$$

Das würde bedeuten, dass $\sqrt{6}$ rational ist, ein Widerspruch.

Man sieht außerdem leicht ein, dass es kein endliches oder abzählbares erzeugendes System E im Vektorraum \mathbb{R} über \mathbb{Q} gibt. Zusammen mit der Abzählbarkeit von \mathbb{Q} würde die Abzählbarkeit von $L_{\mathbb{Q}}(E) = \mathbb{R}$ folgen. \square

Hilfssatz 14.6 Ist $\{u^1, \dots, u^n\}$ linear unabhängig und $\{u^1, \dots, u^n, u\}$ linear abhängig, so ist $u \in L\{u^1, \dots, u^n\}$.

Beweis. Nach Voraussetzung existiert $\{\alpha_1, \dots, \alpha_n, \alpha\} \neq \{0, \dots, 0\}$ mit

$$\sum_{j=1}^n \alpha_j u^j + \alpha u = 0.$$

Wäre $\alpha = 0$, so wäre $\sum_{j=1}^n \alpha_j u^j = 0$ mit $\{\alpha_1, \dots, \alpha_n\} \neq \{0, \dots, 0\}$ im Widerspruch zur linearen Unabhängigkeit von $\{u^1, \dots, u^n\}$. Also ist $\alpha \neq 0$ und die Gleichung kann nach u aufgelöst werden. \blacksquare

Hilfssatz 14.7 Wenn V ein erzeugendes System mit n Elementen enthält, dann sind beliebige $n+1$ Elemente aus V linear abhängig. (Die Anzahl der Elemente eines erzeugenden Systems liefert also stets eine obere Schranke für die Anzahl linear unabhängiger Elemente.)

Beweis. Sei $E = \{u^1, \dots, u^n\}$ ein erzeugendes System, $v^1, \dots, v^{n+1} \in V$. Wegen $L(E) = V$ gibt es $\alpha_{ij} \in \mathbb{K}$ mit

$$v^j = \sum_{i=1}^n \alpha_{ij} u^i \quad \text{für } j = 1, \dots, n+1.$$

Es ist zu zeigen, dass $(x_1, \dots, x_{n+1}) \neq (0, \dots, 0)$ aus \mathbb{K}^{n+1} existiert mit

$$\sum_{j=1}^{n+1} x_j v^j = 0.$$

Diese Gleichung ist äquivalent zu

$$0 = \sum_{j=1}^{n+1} x_j \left(\sum_{i=1}^n \alpha_{ij} u^i \right) = \sum_{i=1}^n \left(\sum_{j=1}^{n+1} x_j \alpha_{ij} \right) u^i,$$

was sicher dann gilt, wenn (x_1, \dots, x_{n+1}) das Gleichungssystem

$$Ax = 0 \quad \text{mit } A = (\alpha_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, n+1}}$$

löst. Es ist also die Existenz einer nicht-trivialen Lösung dieses Systems zu zeigen. Eine solche existiert, da es sich um (nur) n Gleichungen mit $n + 1$ Unbekannten handelt (beim Gaußschen Eliminationsverfahren stehen in der letzten (untersten) Gleichung 2 Unbekannte x_j ; mindestens eine davon kann beliebig (insbesondere $\neq 0$) gewählt werden). ■

Damit können wir festhalten: Eine Teilmenge B eines Vektorraumes V ist genau dann eine Basis, wenn sie

- ein linear unabhängiges erzeugendes System ist, oder
- ein maximales linear unabhängiges System ist, oder
- ein minimales erzeugendes System ist.

Satz 14.8 *Ist der \mathbb{K} -Vektorraum $V \neq \{0\}$ endlich erzeugt (d. h. es gibt ein endliches erzeugendes System), so gilt:*

- a) V besitzt eine endliche Basis;
- b) Je zwei Basen haben gleich viele Elemente;
- c) Ist $M = \{u^1, \dots, u^r\}$ linear unabhängig, so ist M entweder eine Basis, oder es gibt Elemente $u^{r+1}, \dots, u^n \in V$ so, dass $\{u^1, \dots, u^n\}$ eine Basis ist (d. h. jedes linear unabhängige System in V kann zu einer Basis ergänzt werden, Basisergänzungssatz).

Beweis. a) Da $V \neq \{0\}$ ist, gibt es ein $u^1 \neq 0$ und $\{u^1\}$ ist linear unabhängig. Nach Teil c) kann es zu einer Basis ergänzt werden.

b) Seien B und B' Basen mit n bzw. n' Elementen. Da B erzeugendes System ist, sind je $n + 1$ Elemente linear unabhängig, also muß $n' \leq n$ gelten. Entsprechend folgt $n \leq n'$.

c) Ist $L(M) = V$, so ist M eine Basis. Ist $L(M) \neq V$, so gibt es ein $u^{r+1} \in V \setminus L(M)$ so, dass $\{u^1, \dots, u^{r+1}\}$ linear unabhängig ist (sonst wäre nach obigem Hilfssatz u^{r+1} Linearkombination von u^1, \dots, u^r). Das Verfahren kann so fortgesetzt werden und bricht nach Hilfssatz 14.7 nach endlich vielen Schritten ab. ■

Bemerkung 14.9 Ein entsprechender Satz gilt in nichtendlich-erzeugten Räumen. Der Beweis ist dann nicht konstruktiv (Zornsches Lemma). Trotzdem können in konkreten Fällen oft leicht Basen angegeben werden (z. B. im Raum P der Polynome)

Ist V endlich erzeugt, so wird die Anzahl der Elemente einer Basis als *Dimension* ($\dim V$) bezeichnet. (Nicht endlich erzeugte Räume heißen *unendlichdimensional*, wobei u. U. noch „verschiedene Unendlich“ unterschieden werden, insbesondere „abzählbar unendlich“ und „überabzählbar“.)

$\mathbb{R}^n, \mathbb{C}^n$ haben offenbar die Dimension n ; $P_{\mathbb{K},n}$ hat die Dimension $n + 1$.

Gelegentlich ist es wichtig anzugeben auf welchen Skalarenkörper sich die Dimension bezieht,

$$\dim_{\mathbb{K}} \mathbb{K}^n = n, \quad \dim_{\mathbb{R}} \mathbb{C} = 2, \quad \dim_{\mathbb{Q}} \mathbb{R} = \infty.$$

Sind U, W Untervektorräume von V , so ist $U \cap W$ ebenfalls ein Untervektorraum (\ddot{U}), $U \cup W$ jedoch im allgemeinen nicht (Beispiel?, \ddot{U}). Für beliebige Teilräume U, W nennt man

$$U + W := \{u + w : u \in U, w \in W\}$$

die *Summe* der Teilräume U und W . Dies ist offensichtlich der kleinste Teilraum von V , der U und W enthält, d. h. $U + W = L\{U \cup W\}$.

Satz 14.10 (Dimensionsformel) Sind U und W endlichdimensionale Teilräume von V (der Raum V darf durchaus unendlichdimensional sein), so gilt

$$\dim U + \dim W = \dim(U + W) + \dim(U \cap W).$$

Beweis. Sei $\{u^1, \dots, u^r\}$ Basis von $U \cap W$. Dann existieren nach dem Basisergänzungssatz (Satz 14.8 c)

– v^1, \dots, v^s so, dass $\{u^1, \dots, u^r, v^1, \dots, v^s\}$ Basis von U ist,

– w^1, \dots, w^t so, dass $\{u^1, \dots, u^r, w^1, \dots, w^t\}$ Basis von W ist.

Wir zeigen, dass

$$\{u^1, \dots, u^r, v^1, \dots, v^s, w^1, \dots, w^t\}$$

eine Basis von $U + W$ ist. Jedenfalls gilt

$$L\{u^1, \dots, w^t\} \supset U \text{ und } W, \text{ also } L\{U \cup W\}.$$

Andererseits ist offenbar

$$L\{u^1, \dots, w^t\} \subset U + W.$$

Es bleibt die lineare Unabhängigkeit zu zeigen: Ist

$$\sum_{j=1}^r \alpha_j u^j + \sum_{j=1}^s \beta_j v^j + \sum_{j=1}^t \gamma_j w^j = 0 \text{ mit } \{\alpha_j, \beta_j, \gamma_j\} \neq \{0, \dots, 0\},$$

so ist sicher ein β_j oder ein γ_j von Null verschieden (da $\{u^1, \dots, u^r\}$ linear unabhängig ist). Also ist z. B. im Fall $\gamma_j \neq 0$ eine nichttriviale Linearkombination der w^j gleich einer Linearkombination der u^j und v^j , also in U enthalten und somit in $U \cap W$, ein Widerspruch zur Konstruktion von $\{w^1, \dots, w^t\}$. Entsprechend wird der Fall $\beta_j \neq 0$ ausgeschlossen.

Die gewünschte Formel folgt sofort: links steht $r + s + r + t = 2r + s + t$; rechts steht $r + s + t + r = 2r + s + t$. ■

Ist $U \cap W = \{0\}$, so heißt $V' := U + W$ die *direkte Summe* von U und W (auch $U \oplus W$ oder $U \dot{+} W$).

Ist $V = U \dot{+} W$, so heißt U ein *Komplement* von W bzw. W ein *Komplement* von U . In diesem Fall gilt also

$$\dim V = \dim U + \dim W.$$

Satz 14.11 *Es gilt $V = U \dot{+} W$ genau dann, wenn zu jedem $v \in V$ eindeutig bestimmte $u \in U$ und $w \in W$ existieren mit $v = u + w$.*

Beweis. \Rightarrow : Sei $V = U + W$ und $U \cap W = \{0\}$. Dann läßt sich jedes $v \in V$ schreiben in der Form $v = u + w$ mit $u \in U$ und $w \in W$. Gäbe es eine zweite Darstellung $v = u' + w'$ mit $u \neq u'$ (also auch $w \neq w'$), so wäre $u - u' = w' - w \in U \cap W = \{0\}$, ein Widerspruch.

\Leftarrow : Nur $U \cap W = \{0\}$ ist zu zeigen: Ein $x \in U \cap W \setminus \{0\}$ ließe sich offensichtlich nicht eindeutig in der Form $u + w$ mit $u \in U$ und $w \in W$ darstellen (z. B. $x = x + 0$ oder $0 + x$ oder $\frac{1}{2}x + \frac{1}{2}x$). ■

15 Fourierreihen

15.1 Vorbemerkungen

Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt p -periodisch, wenn gilt

$$f(x+p) = f(x) \quad \text{für alle } x \in \mathbb{R}.$$

Offenbar gilt dann auch $f(x+kp) = f(x)$ für alle $k \in \mathbb{Z}$. Die kleinste Zahl $p > 0$, für die dies gilt, wird als *die Periode* von f bezeichnet.

Ist f periodisch mit Periode p und $L > 0$, so ist

$$\tilde{f}(x) := f\left(\frac{p}{L}x\right)$$

periodisch mit Periode L . Es bedeutet deshalb keine wesentliche Einschränkung, wenn im folgenden nur 2π -periodische Funktionen betrachtet werden.

Offensichtlich ist jedes *trigonometrische Polynom* (auf Grund der Additionstheoreme sind dies tatsächlich Polynome in $\sin x$ und $\cos x$)

$$\begin{aligned} f(x) &:= \frac{a_0}{2} + \sum_{k=1}^n \{a_k \cos(kx) + b_k \sin(kx)\} \\ &= \frac{1}{2} \left\{ a_0 + \sum_{k=1}^n [(a_k - ib_k)e^{ikx} + (a_k + ib_k)e^{-ikx}] \right\} = \sum_{k=-n}^n c_k e^{ikx} \end{aligned}$$

2π -periodisch. Dabei gilt offenbar $c_0 = \frac{a_0}{2}$ bzw. $a_0 = 2c_0$ und

$$c_k = \begin{cases} \frac{1}{2}(a_k - ib_k) & \text{für } k > 0, \\ \frac{1}{2}(a_{-k} + ib_{-k}) & \text{für } k < 0, \end{cases}$$

$$a_k = c_k + c_{-k}, \quad b_k = i(c_k - c_{-k}) \quad \text{für } k > 0.$$

Die Periode ist genau dann 2π , wenn die Menge der k mit $a_k \neq 0$ oder $b_k \neq 0$ (bzw. $c_k \neq 0$ oder $c_{-k} \neq 0$) teilerfremd ist; ist k_0 der größte gemeinsame Teiler dieser k , so ist $2\pi/k_0$ die Periode.

Satz 15.1 *Ist f ein solches trigonometrisches Polynom, so sind die Koeffizienten a_k, b_k bzw. c_k eindeutig bestimmt durch*

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) \, dx & \text{für } k = 0, 1, 2, \dots, n, \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) \, dx & \text{für } k = 1, 2, \dots, n, \\ c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} \, dx & \text{für } k = -n, -n+1, \dots, n-1, n. \end{aligned}$$

Beweis. Diese Formeln ergeben sich sofort aus den folgenden Beziehungen, wenn man in den Integralen $f(x)$ durch obige Summendarstellung ersetzt:

$$\begin{aligned}
 1. \quad & \int_0^{2\pi} \cos(kx) \sin(k'x) \, dx = 0 \quad \text{für alle } k \in \mathbb{N}_0, k' \in \mathbb{N}, \\
 2. \quad & \int_0^{2\pi} \cos(kx) \cos(k'x) \, dx = 0 \quad \text{für } k, k' \in \mathbb{N}_0, k \neq k', \\
 3. \quad & \int_0^{2\pi} \sin(kx) \sin(k'x) \, dx = 0 \quad \text{für } k, k' \in \mathbb{N}, k \neq k', \\
 4. \quad & \int_0^{2\pi} [\sin(kx)]^2 \, dx = \pi \quad \text{für alle } k \in \mathbb{N}, \\
 5. \quad & \int_0^{2\pi} [\cos(kx)]^2 \, dx = \pi \quad \text{für alle } k \in \mathbb{N}, \\
 6. \quad & \int_0^{2\pi} [\cos(0x)]^2 \, dx = 2\pi, \\
 & \text{bzw.} \\
 7. \quad & \int_0^{2\pi} e^{-ikx} e^{ik'x} \, dx = \delta_{k,k'} 2\pi, \quad \text{für } k, k' \in \mathbb{Z},
 \end{aligned}$$

wobei $\delta_{k,k'}$ das *Kroneckersymbol* ist ($= 1$ für $k = k'$, $= 0$ für $k \neq k'$), die wie folgt leicht zu beweisen sind:

1. Da der Integrand 2π -periodisch ist, ändert sich das Integral nicht, wenn über $(-\pi, \pi)$ statt $(0, 2\pi)$ integriert wird. Da der Integrand ungerade ist, folgt die Behauptung.

2. Auf Grund des Additionstheorems für den Kosinus ist der Integrand gleich

$$\frac{1}{2} \{ \cos(k+k')x + \cos(k-k')x \}.$$

Wegen $k+k' \neq 0$ und $k-k' \neq 0$ verschwindet also das Integral.

Ebenso erhält man in den Fällen 3., 4. und 5. für den Integranden

$$\begin{aligned}
 & \frac{1}{2} \{ \cos(k-k')x - \cos(k+k')x \}, \\
 & \frac{1}{2} \{ \cos(k-k)x - \cos(k+k)x \} = \frac{1}{2} \{ \cos(0x) - \cos(2kx) \}, \text{ bzw.} \\
 & \{ \cos(k+k)x - \cos(k-k)x \} = \frac{1}{2} \{ \cos(2kx) + \cos(0x) \}
 \end{aligned}$$

und somit für die Integrale $0, \pi$ und π .

Die Aussage 6 ist offensichtlich.

Die Aussage 7 ergibt sich durch einfache Integration. ■

Da die Darstellung in der Form $\sum c_k e^{ikx}$ einfacher zu handhaben ist, arbeiten wir ab jetzt mit dieser weiter (natürlich lassen sich alle Aussagen sofort auf die andere Darstellung übertragen).

Der obige Satz gilt offenbar auch, wenn f die Form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left\{ a_k \cos(kx) + b_k \sin(kx) \right\} = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

hat, wobei die Reihe gleichmäßig konvergiert (Beweis!).

Wir stellen uns die *Frage*: Hat jede 2π -periodische Funktion die Form $\sum_{k=-n}^n c_k e^{ikx}$, oder kann sie durch solche Funktionen beliebig gut approximiert werden? Was heißt hier „jede“ Funktion, und in welchem Sinn soll approximiert werden?

15.2 Konvergenz im Quadratmittel

Im folgenden sei R der Raum der 2π -periodischen Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$, die auf einem/jedem Periodenintervall Regelfunktionen sind. Für jedes $f \in R$ sind dann die *Fourierkoeffizienten* von f

$$c_k = c_k(f) := \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx, \quad k \in \mathbb{Z}$$

wohldefiniert. Wir können also die (*formale*) *Fourierreihe*

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

angeben. Im folgenden sei

$$s_n(x) = s_{f,n}(x) := \sum_{k=-n}^n c_k e^{ikx}, \quad n \in \mathbb{N}_0$$

die n -te Partialsumme der Fourierreihe von f . Es stellt sich also die

Frage: Können Bedingungen an $f \in R$ angegeben werden, unter denen gilt

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \quad \text{d. h.} \quad s_n(x) \longrightarrow f(x)?$$

Dabei kann man an gleichmäßige Konvergenz denken, an punktweise Konvergenz, oder an eine u. U. noch schwächere Konvergenz, z. B. die im folgenden erklärte Konvergenz im Quadratmittel.

Wir sagen: Eine Folge (f_n) aus R konvergiert im Quadratmittel (oder auch konvergiert im L^2 -Sinn) gegen eine Funktion $f \in R$, wenn gilt

$$\int_0^{2\pi} |f_n(x) - f(x)|^2 dx \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Durch

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx \quad \text{und} \quad \|f\| := \langle f, f \rangle^{1/2}$$

sind ein *Semiskalarprodukt* und eine *Halbnorm* auf R erklärt (Skalarprodukt bzw. Norm, wenn man Funktionen f und g mit $\int |f - g| dx = 0$ identifiziert; Beweis!) und es gilt offensichtlich:

(f_n) konvergiert genau dann im Quadratmittel gegen f , wenn (f_n) im Sinne dieser (Halb-) Norm gegen f konvergiert, d. h. $\|f_n - f\| \rightarrow 0$.

Bemerkung 15.2 Aus der gleichmäßigen Konvergenz folgt natürlich die Konvergenz im Quadratmittel (\ddot{U}). Die Umkehrung gilt nicht. Zum Beispiel konvergiert die Folge (f_n) mit

$$f_n(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq 2\pi - \frac{1}{n}, \\ n(x + \frac{1}{n} - 2\pi) & \text{für } 2\pi - \frac{1}{n} \leq x \leq 2\pi, \end{cases} \quad (2\pi\text{-periodisch})$$

im Quadratmittel gegen 0, nicht aber gleichmäßig (\ddot{U}); man mache sich dies auch an einer Skizze klar. Aus der Konvergenz im Quadratmittel folgt aber auch nicht die punktweise Konvergenz (Aufgabe 14 am Ende dieses Abschnitts).

Im folgenden seien die Funktionen e_k definiert durch

$$e_k : \mathbb{R} \rightarrow \mathbb{C}, \quad e_k(x) = e^{ikx} \quad \text{für } k \in \mathbb{Z}, x \in \mathbb{R}.$$

Dann gilt offenbar

$$c_k = c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx = \langle e_k, f \rangle$$

und (vgl. obige Eigenschaft 7)

$$\langle e_j, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-j)x} dx = \delta_{jk},$$

d. h. $\{e_k : k \in \mathbb{Z}\}$ ist ein *Orthonormalsystem* (ONS) bezüglich des Skalarprodukts $\langle \cdot, \cdot \rangle$.

Satz 15.3 (Entwicklungssatz) Sei $f \in R$; c_k und $s_n = s_{f,n}$ seien wie oben erklärt. Dann gilt:

a) $\|f - s_n\|^2 = \|f\|^2 - \sum_{k=-n}^n |c_k|^2,$

b) Es gilt die Besselsche Ungleichung

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|^2.$$

Insbesondere gilt das Riemann–Lebesgue–Lemma: $c_k \rightarrow 0$ für $k \rightarrow \infty$.

c) Die Folge (s_n) konvergiert genau dann im Quadratmittel gegen f , d. h. $\|s_n - f\| \rightarrow 0$ für $n \rightarrow \infty$, wenn gilt

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|^2 \quad (\text{Parsevalsche Gleichung}).$$

(Die Gültigkeit der Parsevalschen Gleichung für Treppenfunktionen $f \in R$ wird in Hilfsatz 15.9 bewiesen; daraus folgt dann in Satz 15.10 die Konvergenz im Quadratmittel für Funktionen aus R .)

Beweis. a) Man rechnet leicht nach:

$$\begin{aligned} \|f - s_n\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} \left| f(x) - \sum_{k=-n}^n c_k e^{ikx} \right|^2 dx \\ &= \frac{1}{2\pi} \left\{ \int_0^{2\pi} |f(x)|^2 dx - \sum_{k=-n}^n c_k \int_0^{2\pi} \overline{f(x)} e^{ikx} dx \right. \\ &\quad \left. - \sum_{k=-n}^n \overline{c_k} \int_0^{2\pi} f(x) e^{-ikx} dx + 2\pi \sum_{k=-n}^n |c_k|^2 \right\} \\ &= \|f\|^2 - \sum_{k=-n}^n \left\{ c_k \overline{c_k} + \overline{c_k} c_k - |c_k|^2 \right\} = \|f\|^2 - \sum_{k=-n}^n |c_k|^2. \end{aligned}$$

b) Dies folgt aus Teil a, da stets $\|f - s_n\|^2 \geq 0$ gilt.

c) Dies folgt direkt aus Teil a. ■

Diesen Satz kann man, bei wörtlich gleichem Beweis, auch abstrakt formulieren:

Satz 15.4 (allgemeiner Entwicklungssatz) Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und Norm $\|\cdot\| = \langle \cdot, \cdot \rangle^{1/2}$,¹⁵ $\{e_k : k \in \mathbb{N}\}$ eine orthonormale Folge, $f \in V$,

$$c_k := \langle e_k, f \rangle, \quad s_n := \sum_{k=1}^{\infty} c_k e_k.$$

Dann gilt:

¹⁵Ein solcher Raum wird in der Mathematik als *Prähilbertraum* bezeichnet, wenn er bezüglich der Norm $\|\cdot\|$ vollständig ist als *Hilbertraum* nach D. HILBERT.

- a) $\|f - s_n\|^2 = \|f\|^2 - \sum_{k=1}^n |c_k|^2,$
- b) $\sum_{k=1}^n |c_k|^2 \leq \|f\|^2$ (Besselsche Ungleichung),
- c) $\|s_n - f\| \rightarrow 0 \Leftrightarrow \sum_{k=1}^{\infty} |c_k|^2 = \|f\|^2$ (Parsevalsche Gleichung)

Der *Beweis* geht wie oben. Dabei sieht die Rechnung in Teil a wie folgt aus:

$$\begin{aligned} \|f - s_n\|^2 &= \left\langle f - \sum_{k=1}^n c_k e_k, f - \sum_{k=1}^n c_k e_k \right\rangle \\ &= \|f\|^2 - \sum_{k=1}^n c_k \langle f, e_k \rangle - \sum_{k=1}^n \bar{c}_k \langle e_k, f \rangle + \sum_{k=1}^n |c_k|^2 \\ &= \|f\|^2 - \sum_{k=1}^n |c_k|^2 - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |c_k|^2 = \|f\|^2 - \sum_{k=1}^n |c_k|^2. \end{aligned}$$

Korollar 15.5 Für die Koeffizienten a_k und b_k ergibt sich hieraus

$$\|f\|^2 \geq \left| \frac{a_0}{2} \right|^2 + \frac{1}{2} \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2),$$

Der Parsevalschen Gleichung entspricht auch hier die Gleichheit.

Der *Beweis* ergibt sich aus $c_0 = a_0/2$, $c_{\pm k} = (a_k \pm ib_k)/2$.

Es ist unser Ziel zu zeigen, dass für jedes $f \in R$ die Folge (s_n) im Quadratmittel gegen f konvergiert. Teil c) von Satz 15.3 sagt aus, dass hierfür die Parsevalsche Gleichung zu beweisen ist, was wir in Satz 15.9 zunächst für Treppenfunktionen tun werden. Dazu werden noch drei leider etwas technische Hilfssätze benötigt.

Hilfssatz 15.6 Für alle $t \in \mathbb{R} \setminus \{2k\pi : k \in \mathbb{Z}\}$ gilt

$$\sum_{k=1}^n \cos(kt) = \frac{\sin(n + \frac{1}{2})t}{2 \sin(\frac{1}{2}t)} - \frac{1}{2}.$$

Beweis. Aus $\cos(kt) = \frac{1}{2}(e^{ikt} + e^{-ikt})$ folgt mit Hilfe der Summenformel für die geometrische Reihe

$$\begin{aligned} \frac{1}{2} + \sum_{k=1}^n \cos(kt) &= \frac{1}{2} \sum_{k=-n}^n e^{ikt} = \frac{1}{2} e^{-int} \sum_{k=0}^{2n} e^{ikt} \\ &= \frac{1}{2} e^{-int} \frac{1 - e^{(2n+1)it}}{1 - e^{it}} = \frac{1}{2} \frac{e^{-i(n + \frac{1}{2})t} - e^{i(n + \frac{1}{2})t}}{e^{-i\frac{t}{2}} - e^{i\frac{t}{2}}}; \end{aligned}$$

das ist die Behauptung. ■

27:29.01.08

Hilfssatz 15.7 *Es gilt*

$$\frac{\pi - x}{2} = \sum_{k=1}^{\infty} \frac{\sin(kx)}{k} \quad \text{für } 0 < x < 2\pi;$$

für jedes $\delta > 0$ ist dabei die Konvergenz gleichmäßig in $[\delta, 2\pi - \delta]$.

Beweis. Aus Hilfssatz 15.6 folgt für $x \in (0, 2\pi)$ und $n \in \mathbb{N}$

$$\begin{aligned} \sum_{k=1}^n \frac{\sin(kx)}{k} &= \sum_{k=1}^n \int_{\pi}^x \cos(kt) dt = \int_{\pi}^x \sum_{k=1}^n \cos(kt) dt \\ &= \int_{\pi}^x \frac{\sin(n + \frac{1}{2})t}{2 \sin(\frac{1}{2}t)} dt - \frac{1}{2}(x - \pi) =: F_n(x) + \frac{\pi - x}{2} \end{aligned}$$

mit

$$\begin{aligned} F_n(x) &= \int_{\pi}^x \frac{1}{2 \sin(\frac{1}{2}t)} \sin(n + \frac{1}{2})t dt \\ &= -\frac{1}{2 \sin(\frac{1}{2}t)} \frac{\cos(n + \frac{1}{2})t}{n + \frac{1}{2}} \Big|_{\pi}^x - \frac{1}{n + \frac{1}{2}} \int_{\pi}^x \frac{\cos(\frac{1}{2}t)}{4 \sin^2(\frac{1}{2}t)} \cos(n + \frac{1}{2})t dt. \end{aligned}$$

Für jedes $\delta > 0$ konvergiert $F_n(\cdot)$ in $[\delta, 2\pi - \delta]$ gleichmäßig gegen 0, weil dort das im Nenner stehende $\sin(\frac{1}{2}t)$ gleichmäßig von 0 entfernt ist. ■

Hilfssatz 15.8 *Es gilt*

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} = \frac{(x - \pi)^2}{4} - \frac{\pi^2}{12} \quad \text{für } 0 \leq x \leq 2\pi.$$

Speziell erhält man für $x = 0$: $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Beweis. Aus Hilfssatz 15.7 folgt für $x, y \in (0, 2\pi)$ durch Integration (wegen der gleichmäßigen Konvergenz in $[y, x]$ bzw. $[x, y]$ können Summe und Integral vertauscht werden)

$$\begin{aligned} \frac{(x-\pi)^2}{4} - \frac{(y-\pi)^2}{4} &= \int_y^x \frac{t-\pi}{2} dt = - \int_y^x \sum_{k=1}^{\infty} \frac{\sin(kt)}{k} dt \\ &= \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} - \sum_{k=1}^{\infty} \frac{\cos(ky)}{k^2}, \end{aligned}$$

also

$$\frac{(x-\pi)^2}{4} = \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} + c.$$

Aus

$$\frac{\pi^3}{6} = \int_0^{2\pi} \frac{(x-\pi)^2}{4} dx = \int_0^{2\pi} \left(\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} + c \right) dx = 2\pi c$$

folgt $c = \pi^2/12$ und damit die obige Behauptung für alle $x \in (0, 2\pi)$. Für $x = 0$ und $x = 2\pi$ folgt die Behauptung durch Grenzübergang, da beide Seiten stetig sind (man beachte, dass die linke Seite wegen der gleichmäßigen Konvergenz der Reihe stetig ist). ■

Satz 15.9 Ist $f \in R$ so, dass die Einschränkung von f auf $[0, 2\pi]$ eine Treppenfunktion ist, so gilt $s_n \rightarrow f$ im Quadratmittel.

Beweis. a) Wir betrachten zunächst den Spezialfall

$$f_a(x) := \begin{cases} 1 & \text{für } 0 \leq x \leq a, \\ 0 & \text{für } a < x < 2\pi, \end{cases} \quad 2\pi\text{-periodisch fortgesetzt.}$$

In diesem Fall berechnet man für $c_k = c_k(f_a)$:

$$\begin{aligned} c_0 &= \frac{a}{2\pi}, \\ c_k &= \frac{1}{2\pi} \int_0^a e^{-ikx} dx = \frac{i}{2\pi k} (e^{-ika} - 1) \quad \text{für } k \neq 0, \\ |c_k|^2 &= \frac{1}{4\pi^2 k^2} (1 - e^{ika})(1 - e^{-ika}) = \frac{1 - \cos(ka)}{2\pi^2 k^2} \quad \text{für } k \neq 0 \end{aligned}$$

und somit (unter Verwendung von Hilfssatz 15.8)

$$\begin{aligned} \sum_{k=-\infty}^{\infty} |c_k|^2 &= \frac{a^2}{4\pi^2} + \sum_{k=1}^{\infty} \frac{1 - \cos(ka)}{\pi^2 k^2} = \frac{a^2}{4\pi^2} + \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} - \frac{1}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos(ka)}{k^2} \\ &= \frac{a^2}{4\pi^2} + \frac{1}{6} - \frac{1}{\pi^2} \left\{ \frac{(\pi-a)^2}{4} - \frac{\pi^2}{12} \right\} = \frac{a}{2\pi} \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f_a(x)|^2 dx = \|f_a\|^2. \end{aligned}$$

Mit Teil c von Satz 15.3 folgt die Behauptung in diesem Spezialfall.

b) Sei jetzt $f \in R$ und $f|_{[0,2\pi]}$ eine Treppenfunktion. Dann läßt sich f schreiben in der Form $f = \sum_{j=1}^{\ell} d_j f_j$, wobei die f_j von der in Teil a des Beweises betrachteten Gestalt sind. Sind $s_{j,n}$ die Partialsummen der zu f_j gehörigen Fourierreihen, so gilt nach Teil a

$$\|f_j - s_{j,n}\| \rightarrow 0 \text{ für } n \rightarrow \infty, j = 1, \dots, \ell$$

und somit aus der Dreiecksungleichung für $\|\cdot\|$

$$\|f - s_n\| = \left\| \sum_{j=1}^{\ell} (d_j f_j - d_j s_{j,n}) \right\| \leq \sum_{j=1}^{\ell} |d_j| \|f_j - s_{j,n}\| \rightarrow 0,$$

d. h. es gilt $s_n \rightarrow f$ im Quadratmittel. ■

Satz 15.10 Für jedes $f \in R$ konvergiert die Fourierreihe im Quadratmittel gegen f . Insbesondere gilt die Parsevalsche Gleichung

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx.$$

Beweis. Da $f \in R$ ist, gibt es zu jedem $\varepsilon > 0$ ein $f_\varepsilon \in R$ so, dass $f_\varepsilon|_{[0,2\pi]}$ eine Treppenfunktion ist und

$$|f_\varepsilon(x) - f(x)| < \varepsilon \quad \text{für alle } x \in \mathbb{R}.$$

Dann folgt für $g := f - f_\varepsilon$

$$\|g\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f_\varepsilon(x) - f(x)|^2 dx \leq \frac{\varepsilon^2}{2\pi} 2\pi = \varepsilon^2.$$

Sind s_n , $s_{g,n}$ und $s_{\varepsilon,n}$ die Partialsummen der Fourierreihen zu f , g bzw. f_ε , so gilt

$$f = f_\varepsilon + g \quad \text{und} \quad s_n = s_{\varepsilon,n} + s_{g,n},$$

und somit

$$\begin{aligned} \|f - s_n\| &= \|f_\varepsilon - s_{\varepsilon,n} + g - s_{g,n}\| \leq \|f_\varepsilon - s_{\varepsilon,n}\| + \|g - s_{g,n}\| \\ &\quad (\text{nach dem Entwicklungssatz 15.3}) \\ &\leq \|f_\varepsilon - s_{\varepsilon,n}\| + \|g\| \\ &\leq \|f_\varepsilon - s_{\varepsilon,n}\| + \varepsilon \leq 2\varepsilon \text{ für große } n. \end{aligned}$$

Das ist die Konvergenz $s_n \rightarrow f$ im Quadratmittel, woraus die anderen Aussagen mit Satz 15.3 folgen. ■

15.3 Gleichmäßige und punktweise Konvergenz

Der Satz 15.10 ist zwar sehr allgemein gültig, liefert allerdings nur eine „recht schwache“ Konvergenzaussage. Gleichmäßige Konvergenz wäre schöner. Natürlich kann man das nur für stetige Funktionen aus R erwarten; es ist aber bekannt, dass dies auch für stetige Funktionen i. allgem. nicht gilt. (Es gibt sogar stetige Funktionen, deren Fourierreihen nicht überall konvergieren; Konstruktionen gehen auf P. du Bois Raymond, H. Lebesgue, H. A. Schwarz, A. Haar und L. Fejer zurück [vgl. z. B. A. Zygmund: Trigonometric series **I**, VIII.1]. Andererseits weiß man nach L. Carleson, Acta Mathematica **116** (1966), dass die Fourierreihe jeder stetigen (sogar jeder L_2 -) Funktion „fast überall“ konvergiert.) Für etwas „glattere“ Funktionen können wir allerdings gleichmäßige Konvergenz sehr leicht beweisen.

Satz 15.11 Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch, stetig und stückweise stetig differenzierbar (d. h. es gibt eine Unterteilung $0 = t_0 < t_1 < \dots < t_r = 2\pi$ von $[0, 2\pi]$ so, dass $f|_{[t_{j-1}, t_j]}$ stetig differenzierbar ist). Dann konvergiert die Fourierreihe von f gleichmäßig gegen f .

Beweis. Sei

$$\varphi : \mathbb{R} \rightarrow \mathbb{C}, \quad \varphi(x) := \begin{cases} f'(x) & \text{für } x \neq t_j + 2k\pi, \\ 0 & \text{für } x = t_j + 2k\pi. \end{cases}$$

Dann ist offensichtlich $\varphi \in R$. Seien γ_k die Fourierkoeffizienten von φ . Nach Satz 15.10 gilt

$$\sum_{k=-\infty}^{\infty} |\gamma_k|^2 = \|\varphi\|^2 < \infty.$$

Für die Fourierkoeffizienten c_k von f gilt

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx = \frac{1}{2\pi} \sum_{j=1}^r \int_{t_{j-1}}^{t_j} f(x) e^{-ikx} dx \\ &= \frac{i}{2k\pi} \sum_{j=1}^r f(x) e^{-ikx} \Big|_{t_{j-1}}^{t_j} - \frac{i}{2k\pi} \sum_{j=1}^r \int_{t_{j-1}}^{t_j} \varphi(x) e^{-ikx} dx \\ &= 0 - \frac{i}{2k\pi} \int_0^{2\pi} \varphi(x) e^{-ikx} dx = -\frac{i}{k} \gamma_k \end{aligned}$$

und somit (wegen $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$)

$$|c_k| \leq \frac{1}{2}(k^{-2} + |\gamma_k|^2),$$

d. h. $\sum |c_k|$ ist konvergent und somit (wegen $|e^{ikx}| = 1$)

$$\sum c_k e^{ikx}$$

gleichmäßig konvergent gegen eine stetige Funktion g .

Wegen

$$\|f - g\| \leq \|f - s_{f,n}\| + \|s_{f,n} - g\| \longrightarrow 0$$

ist $\|f - g\| = 0$. Da $\|\cdot\|$ nur eine Halbnorm ist, folgt daraus noch nicht direkt, dass $f = g$ gilt. Wäre aber $f \neq g$, so gäbe es ein x_0 mit $f(x_0) - g(x_0) \neq 0$, also ein $\varepsilon > 0$ und ein Intervall I um x_0 mit $|f(x) - g(x)| \geq \varepsilon$ für $x \in I$. Das widerspricht aber $\|f - g\| = 0$; also gilt $f = g$. ■

Beispiel 15.12 Sei

$$f(x) = |x| \quad \text{für } |x| \leq \pi, \quad 2\pi\text{-periodisch fortgesetzt.}$$

Nach Satz 15.11 konvergiert in diesem Fall die Fourierreihe gleichmäßig (dies werden wir auch gleich an den Fourierkoeffizienten sehen; dass die Fourierreihe gleichmäßig gegen f konvergiert, folgt aber nicht aus der Abschätzung der Koeffizienten, sondern aus obigem Satz).

Da f gerade ist, haben wir natürlich eine reine *Kosinusreihe* (Beweis!)

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx)$$

mit

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \, dx = \pi, \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \cos(kx) \, dx = \frac{2}{\pi} \int_0^{\pi} x \cos(kx) \, dx \\ &= \frac{2}{\pi} \left\{ \frac{1}{k} x \sin(kx) \Big|_0^{\pi} - \frac{1}{k} \int_0^{\pi} \sin(kx) \, dx \right\} = \frac{2}{\pi} k^{-2} \cos(kx) \Big|_0^{\pi} \\ &= \begin{cases} 0 & \text{falls } k \text{ gerade ist,} \\ -4(k^2\pi)^{-1} & \text{falls } k \text{ ungerade ist,} \end{cases} \end{aligned}$$

also

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos(2n-1)x}{(2n-1)^2}.$$

Für $|x| \leq \pi$ gilt somit

$$\frac{\pi^2}{8} - \frac{\pi}{4}|x| = \cos x + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \dots$$

und somit speziell für $x = 0$

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots,$$

eine überraschende Formel für π^2 (vgl. frühere Formel $\pi^2/6 = \sum_{j=1}^{\infty} 1/k^2$ in Hilfssatz 15.8). \square

28:31.01.08

Wenn f nicht stetig ist, kann man zwar keine gleichmäßige Konvergenz der Fourierreihe mehr erwarten, eventuell aber wenigstens punktweise Konvergenz in gewissen Regularitätspunkten.

Satz 15.13 Sei $f \in R$ und es existieren

$$\begin{aligned} f_{\pm}(x_0) &= \lim_{h \rightarrow 0_{\pm}} f(x_0 + h), \\ f'_{\pm}(x_0) &= \lim_{h \rightarrow 0_{\pm}} \frac{1}{h} (f(x_0 + h) - f_{\pm}(x_0)). \end{aligned}$$

Dann gilt

$$s_n(x_0) \longrightarrow \frac{1}{2} (f_+(x_0) + f_-(x_0)).$$

Beweis. Man berechnet

$$\begin{aligned} s_n(x_0) &= \sum_{k=-n}^n e^{ikx_0} \frac{1}{2\pi} \int_0^{2\pi} e^{-iky} f(y) \, dy \\ &= \frac{1}{2\pi} \int_{x_0-\pi}^{x_0+\pi} \sum_{k=-n}^n e^{ik(x_0-y)} f(y) \, dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-n}^n e^{-iky} f(x_0 + y) \, dy \\ &= \frac{1}{2\pi} \int_0^{\pi} \sum_{k=-n}^n e^{-iky} (f(x_0 + y) + f(x_0 - y)) \, dy \\ &\quad \left(\text{mit } \sum_{k=-n}^n e^{-iky} = 1 + 2 \sum_{k=1}^n \cos ky = \frac{\sin(n + \frac{1}{2})y}{\sin \frac{1}{2}y}, \text{ vgl. Hilfssatz 15.6} \right) \\ &= \frac{1}{2\pi} \left\{ \int_0^{\pi} \left(1 + 2 \sum_{k=1}^n \cos ky \right) (f_+(x_0) + f_-(x_0)) \, dy \right. \\ &\quad \left. + \int_0^{\pi} \sin \left[\left(n + \frac{1}{2} \right) y \right] \underbrace{\frac{f(x_0 + y) - f_+(x_0) + f(x_0 - y) - f_-(x_0)}{\sin(\frac{1}{2}y)}}_{=g(y)} \, dy \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \left\{ \int_0^\pi \{f_+(x_0) + f_-(x_0)\} dy \right. \\
&\quad \left. + \frac{1}{2i} \int_0^\pi \left[e^{iny} \left(e^{\frac{i}{2}y} g(y) \right) - e^{-iny} \left(e^{-\frac{i}{2}y} g(y) \right) \right] dy \right\} \\
&\longrightarrow \frac{1}{2} \{f_+(x_0) + f_-(x_0)\},
\end{aligned}$$

wobei im letzten Schritt das Riem-Lebesgue-Lemma (vgl. Entwicklungssatz) benutzt wurde, und dass auf Grund der Voraussetzung die 2π -periodische Funktion

$$h(y) := \begin{cases} \exp(\pm \frac{i}{2}y)g(y) & \text{für } y \in [0, 2\pi], \\ 0 & \text{für } y \in (-\pi, 0) \end{cases}$$

in R liegt. ■

Beispiel 15.14 Die Sägefunktion ist definiert durch

$$f(x) = x \quad \text{in } (-\pi, \pi], \quad 2\pi\text{-periodisch fortgesetzt auf } \mathbb{R}.$$

Es ist dann

$$\begin{aligned}
c_k &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx = \frac{1}{2\pi} \int_{-\pi}^\pi x e^{-ikx} dx \\
&= \begin{cases} 0 & \text{für } k = 0 \\ \frac{i}{2k\pi} x e^{-ikx} \Big|_{-\pi}^\pi - \frac{i}{2k\pi} \int_{-\pi}^\pi e^{-ikx} dx = (-1)^k \frac{i}{k} & \text{für } k \neq 0. \end{cases}
\end{aligned}$$

Für die Fourierreihe der Sägefunktion erhalten wir also

$$\sum_{k=1}^{\infty} (-1)^k \frac{i}{k} (e^{ikx} - e^{-ikx}) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \sin kx,$$

eine reine Sinusreihe, da f ungerade ist. Das hätte man natürlich schon früher erkennen und gleich die Koeffizienten b_k der Sinusreihe berechnen können (Ü). Nach obigem Satz konvergiert sie für alle $x \in \mathbb{R} \setminus \{(2k+1)\pi : k \in \mathbb{Z}\}$ gegen $f(x)$; in den Punkten $(2k+1)\pi$ konvergiert sie gegen 0, den Mittelwert der beiden Grenzwerte. Tatsächlich tritt an den Sprungstellen das sogenannte GIBBSche Phänomen des „Überschwingens“ auf: die Partialsummen der Fourierreihe schwingt an der Sprungstelle noch um 18% mehr aus, als der Sprung ausmacht.

Konkrete Fourierentwicklungen und deren graphische Darstellung kann man sich unter Fourierentwicklungen

anschauen. Insbesondere kann hier auch das GIBBSche Phänomen beobachtet werden. \square

15.4 Die Sätze von Weierstraß

Für beliebige stetige 2π -periodische Funktionen kann man zwar nicht die gleichmäßige Konvergenz der Fourierreihe beweisen, trotzdem läßt sich aber jede derartige Funktion gleichmäßig durch trigonometrische Polynome approximieren.

Satz 15.15 (Weierstraßscher Approximationssatz, periodisch)

Sei $f \in R$ stetig. Dann gibt es zu jedem $\varepsilon > 0$ ein trigonometrisches Polynom p mit

$$|p(x) - f(x)| < \varepsilon \quad \text{für alle } x \in \mathbb{R}.$$

In anderen Worten: Es gibt eine Folge (p_n) von trigonometrischen Polynomen mit $p_n \rightarrow f$ gleichmäßig.

Beweis. Da f stetig ist, ist es gleichmäßig stetig und es gibt zu jedem $\varepsilon > 0$ eine stetige stückweise lineare Funktion f_ε mit $|f(x) - f_\varepsilon(x)| < \frac{\varepsilon}{2}$ für alle $x \in \mathbb{R}$ (Beweis!). Nach Satz 15.11 konvergiert die Fourierreihe von f_ε gleichmäßig gegen f_ε . Es gibt also ein trigonometrisches Polynom p mit $|p(x) - f_\varepsilon(x)| < \frac{\varepsilon}{2}$ für alle $x \in \mathbb{R}$ (eine hinreichend hohe Partialsumme der Fourierreihe von f_ε) und somit $|p(x) - f(x)| < \varepsilon$ für alle $x \in \mathbb{R}$. ■

Bemerkung 15.16 Es kann auch explizit eine Folge trigonometrischer Polynome angegeben werden, die gleichmäßig gegen eine stetige Funktion $f \in R$ konvergiert: Nach einem Satz von FEJÉR konvergiert $S_n(x) := \frac{1}{n+1} \sum_{k=0}^n s_{f,n}(x)$ (das Fejer-Mittel der Partialsummen) gleichmäßig gegen f , vgl. K. Königsberger: *Analysis I*, S. 350.

Satz 15.17 (Weierstraßscher Approximationssatz) Sei

$f : [a, b] \rightarrow \mathbb{C}$ stetig. Dann gibt es zu jedem $\varepsilon > 0$ ein Polynom p mit $|f(x) - p(x)| < \varepsilon$ für alle $x \in [a, b]$. In anderen Worten: Es gibt eine Folge (p_n) von Polynomen mit $p_n \rightarrow f$ gleichmäßig in $[a, b]$.

Beweis. Ohne Einschränkung nehmen wir an, dass $[a, b] = [0, 1]$ gilt (andernfalls betrachtet man $\hat{f}(x) := f(a + x(b - a))$). Indem man definiert

$$g(x) := \begin{cases} f(x) & \text{für } x \in [0, 1], \\ f(1) + \frac{x-1}{2\pi-1}(f(0) - f(1)) & \text{für } x \in (1, 2\pi) \end{cases},$$

kann f zu einer stetigen 2π -periodischen Funktion g fortgesetzt werden (man beachte $g(0) = g(2\pi)$). Nach Satz 15.15 gibt es ein trigonometrisches Polynom $q(x) = \sum_{k=-n}^n c_k e^{ikx}$ mit

$$|g(x) - q(x)| \leq \frac{\varepsilon}{2} \quad \text{für alle } x \in \mathbb{R}.$$

Jeder Term $c_k e^{ikx}$ läßt sich auf $[0, 1]$ bis auf $\varepsilon/4n$ durch ein Polynom p_k approximieren (z. B. ein geeignetes Stück der jeweiligen Taylorreihe); exakt für $k = 0$ mit $p_0 = c_0$. Mit dem Polynom $p(x) = \sum_{k=-n}^n p_k(x)$ gilt also für $x \in [0, 1]$

$$|f(x) - p(x)| \leq |g(x) - q(x)| + |q(x) - p(x)| \leq \frac{\varepsilon}{2} + \sum_{k=-n}^n |c_k e^{ikx} - p_k(x)| \leq \varepsilon,$$

d. h. $f : [0, 1] \rightarrow \mathbb{C}$ wird durch Polynome beliebig gut approximiert. ■

Bemerkung 15.18 Im Falle $[a, b] = [0, 1]$ ist die Folge der BERNSTEIN-Polynome

$$(B_n f)(x) = \sum_{j=1}^n f\left(\frac{j}{n}\right) \binom{n}{j} x^j (1-x)^{n-j}$$

eine Folge von Polynomen, die gleichmäßig gegen f konvergiert (vgl. z. B. HÄMMERLIN, Numerische Mathematik §4.2).

29:05.02.08

15.5 Übungsaufgaben

15.1 Man berechne die Fourierreihe der Funktion $f(x) := |\sin x|$.

Anleitung: Man beachte, dass $f(x)$ und $\cos kx$ gerade und $\sin kx$ ungerade Funktionen sind.

15.2 Ein trigonometrisches Polynom $\sum_{k=-n}^n c_k e^{ikx}$ hat genau dann die (minimale) Periode 2π , wenn die Menge $\{k \in \mathbb{N} : |c_k| + |c_{-k}| \neq 0\}$ teilerfremd ist.

15.3 Man zeige, dass die Funktionenfolge

$$f_n(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq 2\pi - 1/n, \\ n(x + 1/n - 2\pi) & \text{für } 2\pi - 1/n \leq x \leq 2\pi \end{cases}$$

im Quadratmittel konvergiert, aber nicht gleichmäßig.

15.4 Man konstruiere eine Folge (f_n) von Funktionen auf $[0, 2\pi]$, die aus immer schmaler werdenden „wandernden Buckeln“ bestehen so, dass (f_n) im Quadratmittel (gegen 0) konvergiert, aber nicht punktweise konvergent ist.

15.5 Sei

$$f(x) = \begin{cases} 0 & \text{für } 0 \leq x < \pi, \\ 1 & \text{für } \pi \leq x < 2\pi, \end{cases} \quad 2\pi\text{-periodisch.}$$

a) Man bestimme die Fourierreihe von f (im wesentlichen eine reine Sinusreihe).

b) Mit Hilfe der Parsevalschen Gleichung zeige man

$$\frac{\pi^2}{8} = \sum_{k=1}^{\infty} (2k-1)^{-2}.$$

c) Die Summe *aller* k^{-2} läßt sich daraus berechnen.

15.6 Man beweise $\sum_{k=1}^{\infty} \frac{1}{k^6} = \frac{\pi^6}{945}$.

Anleitung: Man integriere $\sum_{k=1}^{\infty} k^{-2} \cos kt = \frac{(t-\pi)^2}{4} - \frac{\pi^2}{12}$ von π bis x und wende die Vollständigkeitsrelation an.

15.7 Sei $f : [1, \infty) \rightarrow \mathbb{C}$ stetig und $\lim_{x \rightarrow \infty} f(x)$ existiere. Dann gibt es für jedes $\varepsilon > 0$ ein $n \in \mathbb{N}$ und Konstanten d_0, d_1, \dots, d_n mit

$$|f(x) - \sum_{j=0}^n d_j x^{-j}| < \varepsilon \quad \text{für alle } x \in [1, \infty).$$

15.8 Zu jedem $f \in C[0, 1]$ gibt es eine Folge gerader Polynome, die gleichmäßig auf $[0, 1]$ gegen f konvergiert.

15.9 Direkter Beweis des Riemannschen Lemmas: Für f mit $f(x) = 1$ in $[a, b]$ und $f(x) = 0$ sonst berechne man $c_k(f)$ und zeige $c_k(f) \rightarrow 0$ für $k \rightarrow \infty$. Damit beweise man $c_k(f) \rightarrow 0$ zunächst für Funktionen aus R , für die $f|_{[0, 2\pi)}$ Treppenfunktionen sind und dann für alle $f \in R$.

Literatur

- [1] Baumeister, J.: Vorlesungsskript WS 04/05, <http://math.uni-frankfurt.de/baumeist>
- [2] Fischer, H. und Kaul, H.: Mathematik für Physiker I, Teubner–Verlag 2005
- [3] Fischer, G.: Lineare Algebra, Vieweg 2005
- [4] Foster, O.: Analysis 1, Vieweg 2006
- [5] Heuser, H.: Lehrbuch der Analysis, Teil 1, Teubner–Verlag 2006
- [6] Kerner, H. und v. Wahl, W.: Mathematik für Physiker, Springer 2006
- [7] Koecher, M.: Lineare Algebra und Analytische Geometrie, Springer 1997
- [8] Walter, W.: Analysis 1, Springer 2004