Analysis 1 und 2

Ernst Albrecht

Vorlesungen im Wintersemester 2005/06 und Sommersemester 2006

Universität des Saarlandes

Saarbrücken

Stand: 20. Juli 2006

Inhaltsverzeichnis

KAPITEL 0

Zur Vorbereitung

1. Grundbegriffe aus der Mengenlehre

Es soll hier kurz auf die aus der Schule teilweise bekannte elementare Mengenlehre eingegangen werden, da wir deren Schreib– und Sprechweise verwenden wollen. Zunächst benötigen wir einige Bezeichnungen und Begriffe aus der mathematischen Logik.

Eine *mathematische Theorie* besteht aus grundlegenden Begriffen, definierten Begriffen, Axiomen und Sätzen.

- *Grundlegende Begriffe* sind Begriffe, die im Rahmen der Theorie nicht weiter erklärt werden.
- *Axiome* sind Aussagen über Beziehungen zwischen Begriffen, die nicht bewiesen werden.
- *definierte Begriffe* sind Begriffe, die in einer Definition mit Hilfe schon bekannter Begriffe erklärt werden.
- Ein *Satz* ist eine Aussage über Begriffe. Die Sätze einer Theorie stehen in logischer Abhängigkeit in dem Sinne, daß jeder Satz unter Verwendung von bekannten Begriffen und bereits bewiesenen Sätzen sowie der Axiome *bewiesen*, d.h. durch Anwendung von logischen Operationen aus diesen hergeleitet wird.

Von dem Axiomensystem einer mathematischen Theorie wird man folgende Eigenschaften verlangen:

- *Widerspruchsfreiheit*: Es darf im Rahmen der Theorie nicht möglich sein, einen Satz und dessen Negation zu beweisen.
- *Vollständigkeit*: Hat man sich eine Vorstellung von dem Umfang einer Theorie gebildet, so nennt man ein Axiomensystem vollständig, wenn alle „interessierenden“ Fragen im Rahmen der auf dem Axiomensystem beruhenden Theorie beantwortet werden können.
- *Unabhängigkeit*: Ein Axiomensystem heißt unabhängig, wenn kein Axiom mit Hilfe der übrigen Axiome beweisbar ist.

Wir werden jedoch bei den im Verlauf der Vorlesung aufgestellten Axiomensystemen die hierdurch entstehenden Fragen nicht behandeln können.

Wir gehen nun etwas auf die Sprache und Notation der Logik ein, soweit wir sie in dieser Vorlesung verwenden.

Die einfache Folgerung: Seien A und B zwei logische Aussagen, die wahr oder falsch sein können. Wir schreiben

$$A \Longrightarrow B\,,$$

falls gilt: Immer, wenn A richtig ist, ist auch B richtig. Für diesen logischen Sachverhalt sagen wir auch: „*Die Richtigkeit von B folgt aus der Richtigkeit von A*“ oder „*A gilt nur dann, wenn B gilt*“ oder „*B ist notwendig für A*“ oder „*A impliziert B*“.

Man nennt zwei Aussagen A und B *äquivalent* (Schreibweise: $A \Longleftrightarrow B$), falls gilt: $A \Longrightarrow B$ und $B \Longrightarrow A$. Hierfür sagen wir auch: „*B ist notwendig und hinreichend für A*“ oder „*B gilt dann und nur dann, wenn A gilt*“ oder „*B gilt genau dann, wenn A gilt*“.

Besteht die Äquivalenz von A und B aufgrund einer Definition, so schreiben wir: $A :\Longleftrightarrow B$ oder $A_{\text{def}} \Longleftrightarrow B$, wobei der Doppelpunkt bzw. $_{\text{def}}$ auf der Seite steht, die durch die Äquivalenz definiert werden soll.

Unter einem *direkten Beweis* verstehen wir eine Aufeinanderfolge von einfachen Folgerungen der Art: Gilt $A \Longrightarrow B$ und $B \Longrightarrow C$, so auch $A \Longrightarrow C$.

Beim *indirekten Beweis* wird folgendes Schema verwendet: Will man etwa die Behauptung: „Die Aussage A *ist richtig*" beweisen, so macht man zunächst die *Annahme*: A ist falsch. Aus dieser Annahme, aus Definitionen und bereits als gültig bekannten Aussagen werden dann Folgerungen gezogen, bis sich ein Widerspruch zu Definitionen, bereits als gültig erkannten Aussagen oder zur Annahme ergibt. Dies zeigt dann, daß die Aussage „A ist falsch" falsch ist. Nach dem Satz vom ausgeschlossenen Dritten muß dann die Behauptung gelten.

Das Gleichheitszeichen hat folgende Eigenschaften:

(a) $A = A$ (*Reflexivität*),
(b) $A = B \Longrightarrow B = A$ (*Symmetrie*),
(c) $A = B$ und $B = C \Longrightarrow A = C$ (*Transitivität*).

Wenn zwei Objekte A, B laut Definition übereinstimmen sollen, so schreiben wir $A := B$ oder $A_{\text{def}} = B$, wobei der Doppelpunkt bzw. $_{\text{def}}$ wieder auf der zu definierenden Seite steht.

Zur Abkürzung führen wir ferner die Quantoren der Generalisierung und der Partikularisierung ein:

Generalisierung: $\bigwedge\limits_{B} \ldots$ bzw. $\forall_B \ldots$ bzw. $\forall B \ldots$. Lies: „Für alle Objekte, die der Bedingung B genügen, gilt ...". In der Vorlesung wird meist die letzte Schreibweise verwendet.

Partikularisierung $\bigvee\limits_{B} \ldots$ bzw. $\exists_B \ldots$ bzw. $\exists B \ldots$. Lies: „Es gibt ein Objekt, das der Bedingung B genügt, so daß gilt ...". Hierbei ist „ein" im Sinn „von wenigstens ein" zu verstehen.

Nach diesen Vorbemerkungen, Vereinbarungen und Bezeichnungen wollen wir uns nun der (naiven) Mengenlehre zuwenden. Georg Cantor (3.3.1845–6.1.1918), der Begründer dieser Theorie, gab folgende „Erklärung" der Begriffe „Menge" und „Element einer Menge":

> *Unter einer „Menge" verstehen wir jede Zusammenfassung M von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denken (welche die „Elemente" von M genannt werden) zu einem Ganzen.*

Man könnte nun meinen, daß es sich hierbei um die Definition der Begriffe „Menge" und „Element einer Menge" handelt. Man stellt jedoch fest, daß die zu definierenden Begriffe „Menge" und „Element einer Menge" nur auf andere, undefinierte Begriffe (wie zum Beispiel „Zusammenfassung zu einem Ganzen") zurückgeführt werden. Nun kann man natürlich nicht eine Theorie auf einer Definition aufbauen, da diese notwendigerweise auf Begriffe zurückgreifen muß, die im Rahmen der Theorie nicht definiert sind. Man muß also beim Aufbau der Mengenlehre offenlassen, was Mengen „wirklich sind", ähnlich, wie bei der axiomatischen Begründung der Geometrie nicht gesagt wird, was Punkte und Geraden sind, sondern nur angegeben wird, welche Beziehungen zwischen ihnen bestehen.

Die (also nicht näher erläuterten) Grundbegriffe der Mengenlehre sind:

(i) *Mengen* meist durch große lateinische Buchstaben $A, B, C, \ldots, M, \ldots$ sowie durch große deutsche Buchstaben[1] oder Sonderzeichen bezeichnet wie z.B. $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ und andere.

(ii) *Elemente von Mengen*: $x \in A$ (gelesen: x ist Element der Menge A). Elemente einer Menge werden häufig (aber nicht immer) mit kleinen lateinischen, griechischen oder deutschen Buchstaben bezeichnet. Ist x nicht Element von A, so schreibt man $x \notin A$.

Unsere erste Forderung ist nun, daß Mengen durch ihre Elemente eindeutig bestimmt sein sollen.

(M1) *Zwei Mengen A und B sind gleich ($A = B$ geschrieben) genau dann, wenn sie die gleichen Elemente haben.*

Will man also die Gleichheit von zwei Mengen A und B beweisen, so muß man zeigen, daß jedes Element von A auch ein Element von B ist und jedes Element von B auch ein Element von A ist.

(M1) erlaubt es (endliche) Mengen anzugeben, indem man ihre Elemente aufzählt (*aufzählende Schreibweise*), etwa $\{1, 2, 3, 4, 5\}$ oder $\{5, 4, 3, 2, 1\}$ (diese beiden Mengen sind gleich). Weitere Beispiele $\{1, 2, 3\} \neq \{2, 3, 11\}$, $\{\alpha, \beta, \gamma\}$ usw.

0.1. Definition. Sei A eine Menge. Eine Menge B heißt *Teilmenge* von A, falls jedes Element von B auch ein Element von A ist. Wir schreiben dann $B \subseteq A$ oder $A \supseteq B$.

Ist $B \subseteq A$ und $B \neq A$, so nennen wir B eine echte Teilmenge von A und schreiben $B \subset A$.

Beispiele: $\{1\} \subseteq \{2, 1, 5\}$, $\{\beta, \alpha, \gamma\} \supseteq \{\beta, \gamma\}$. Für jede Menge A gilt offensichtlich $A \subseteq A$.

Häufig möchte man aus einer Menge alle Elemente mit einer vorgegebenen Eigenschaft aussondern und zu einer neuen Menge zusammenfassen. Sei z.B. M die Menge aller Teilnehmer an den Übungen zu dieser Vorlesung, welche in aufzählender Schreibweise durch die Angabe der Teilnehmer in einer Teilnehmerliste angegeben sei. Will man nun die Menge aller Teilnehmer dieser Vorlesung bilden, die im Hauptfach Physik studieren so könnte man aus der Gesamtteilnehmerliste natürlich auf einer neuen Liste die Namen aller derjenigen Studierenden abschreiben, die im Hauptfach Physik studieren. Rein schreibtechnisch ist es aber einfacher zu schreiben:

$$M_P = \{x\,;\, x \in M \text{ und } x \textit{ studiert im Hauptfach Physik}\}\,.$$

Eine solche Beschreibung erlaubt das folgende Aussonderungsaxiom (M2):

(M2) Sei A eine Menge und sei $\mathcal{E}$ eine Eigenschaft, die jedes Element $x \in A$ entweder hat oder nicht hat. Dann gibt es eine Menge B, deren Elemente genau diejenigen Elemente von A sind, die die Eigenschaft $\mathcal{E}$ besitzen.

Nach (M1) ist B hierdurch eindeutig bestimmt. Wir schreiben

$$B = \{x\,;\, x \in A \text{ und } x \text{ hat Eigenschaft } \mathcal{E}\}$$

oder $B = \{x \in A\,;\, x \text{ hat Eigenschaft } \mathcal{E}\}$. Offensichtlich gilt $B \subseteq A$.

[1]Im Anschluß an diese Vorbemerkungen zur Mengenlehre befindet sich eine Wiedergabe des Alphabets in deutscher Schreibschrift und des griechischen Alphabets

0.2. Beispiel. Sei A eine Menge und sei $\mathcal{E}$ die Eigenschaft $x \neq x$ für $x \in A$. Wir bilden gemäß (M2) die Menge $B := \{x\,;\, x \in A,\, x \neq x\}$. Da die Aussage $x \neq x$ für alle $x \in A$ falsch ist gibt es keine Elemente in A mit $x \neq x$. Die Menge B besitzt also keine Elemente. Wir nennen sie daher die *leere Menge* und bezeichnen sie mit $\emptyset$. Nach (M1) ist $\emptyset$ eindeutig bestimmt. Ferner ist $\emptyset$ Teilmenge von jeder Menge.

Man kann sich nun natürlich fragen, warum wir in dem Aussonderungsaxiom (M2) nur die Elemente von A zugelassen haben. Nach der Cantorschen Begriffsbildung wäre dies doch eine überflüssige Einschränkung. Es zeigt sich jedoch, daß man zu Widersprüchen kommt. So könnte man nach Cantor etwa die „Menge aller Mengen" bilden (als die „Zusammenfassung zu einem Ganzen" von allen Mengen, die ja auch „Objekte unseres Denkens" sind). Wir nehmen einmal an, daß dies möglich wäre, und bilden also die „Allmenge" $\mathfrak{A}$ aller Mengen. Dann existiert nach (M2) auch die Teilmenge

$$\mathfrak{B} := \{A \in \mathfrak{A}\,;\, A \notin A\}$$

aller Mengen, die sich nicht selbst als Element enthalten. Nun muß entweder $\mathfrak{B} \in \mathfrak{B}$ oder $\mathfrak{B} \notin \mathfrak{B}$ gelten.

Ist $\mathfrak{B} \in \mathfrak{B}$, so folgt nach Definition von $\mathfrak{B}$: $\mathfrak{B} \notin \mathfrak{B}$.

Ist $\mathfrak{B} \notin \mathfrak{B}$, so folgt nach Definition von $\mathfrak{B}$: $\mathfrak{B} \in \mathfrak{B}$.

In allen möglichen Fällen erhalten wir also einen offensichtlichen Widerspruch.

Dieser unter dem Namen *Russelsche Antinomie* bekannte Widerspruch wurde 1901 von Bertram Russel (18.5.1872–2.2.1970) entdeckt. Von Russel selbst stammt hierzu folgendes Beispiel aus dem „täglichen Leben": In einem Dorf schließt ein Friseur mit dem Gemeinderat den Vertrag, „genau diejenigen Männer des Dorfes zu rasieren, die sich nicht selbst rasieren". Bei Verletzung dieses Vertrages müsse er eine Konventionalstrafe bezahlen. Was würden Sie dem armen Friseur raten: Soll er sich selbst rasieren oder sollte er sich besser von einem Nachbarn rasieren lassen?

Geht man jedoch von (M1) und (M2) aus, so kann man zeigen, daß es eine solche „Allmenge" nicht gibt. Sei nämlich A eine beliebige Menge von Mengen und sei $B := \{x \in A\,;\, x \notin x\}$. Wenn wir zeigen können, daß $B \notin A$ gilt, so kann es auch keine Allmenge geben. Wir führen den Beweis indirekt:

Annahme: $B \in A$. Ist $B \in B$ so folgt nach Definition von B der Widerspruch $B \notin B$. Ist $B \notin B$, so müßte nach Definition von B gelten $B \in B$. In allen möglichen Fällen ergibt sich also ein Widerspruch. Die Annahme war daher falsch. Es muß also $B \notin A$ gelten.

0.3. Definition. Seien A und B zwei Mengen, dann heißt

$$A \cap B := \{x\,;\, x \in A \text{ und } x \in B\}$$

der *Durchschnitt von A mit B*.

Nach (M2) ist $A \cap B$ offensichtlich eine Menge. Die Elemente von $A \cap B$ sind genau diejenigen, die sowohl zu A als auch zu B gehören.

Beispiele. $\{1,2,3,4,5\} \cap \{1,3,5\} = \{1,3,5\}$, $\{1,2,3\} \cap \{3,4,5\} = \{3\}$, $\{2,4,6\} \cap \{1,3,5\} = \emptyset$.

In dem folgenden Lemma fassen wir einige einfache Rechenregeln für die Durchschnittsbildung zusammen, die sich alle elementar beweisen lassen.

0.4. Lemma. *A, B, C seien Mengen. Dann gelten die folgenden Aussagen:*

(a) $A \cap A = A$.

(b) $A \cap \emptyset = \emptyset$.

(c) $A \cap B = B \cap A$.
(d) $A \cap B \subseteq A$ *und* $A \cap B \subseteq B$.
(e) $(A \cap B) \cap C = A \cap (B \cap C)$.
(f) $A \subseteq B \iff A \cap B = A$.

Man möchte manchmal auch über „sehr viele“ Mengen einen Durchschnitt bilden:

0.5. Definition. Sei $\mathfrak{M}$ eine Menge, deren Elemente wieder Mengen sind. Sei $M_0 \in \mathfrak{M}$ beliebig. Dann definiert man

$$\bigcap_{M \in \mathfrak{M}} M := \{x \in M_0\,;\, x \in M \text{ für alle } M \in \mathfrak{M}\}\,.$$

Statt $\bigcap_{M \in \mathfrak{M}} M$ schreibt man manchmal auch $\bigcap\{M\,;\, M \in \mathfrak{M}\}$. Der Durchschnitt $\bigcap_{M \in \mathfrak{M}} M$ ist nach (M2) gebildet und damit eine eindeutig bestimmte Menge. Man zeigt auch leicht, daß $\bigcap_{M \in \mathfrak{M}} M$ nicht von der speziell gewählten Menge M_0 abhängt, d.h. daß für alle $M_0, M_1 \in \mathfrak{M}$ gilt:

$$\{x \in M_0\,;\, x \in M \text{ für alle } M \in \mathfrak{M}\} = \{x \in M_1\,;\, x \in M \text{ für alle } M \in \mathfrak{M}\}\,.$$

Um zeigen zu können, daß Definition 0.5 die Definition 0.3 als Spezialfall enthält, müssen wir wissen, daß es zu zwei Mengen A und B eine Menge C gibt, die A und B *als Elemente* enthält. Hierzu benötigen wir ein weiteres Axiom:

(M3) *Zu je zwei Mengen A und B gibt es eine Menge C mit $A \in C$ und $B \in C$.*

0.6. Bemerkung. Seien A und B zwei Mengen. Dann gibt es genau eine Menge, die genau A und B als Elemente enthält. Diese bezeichnen wir (in aufzählender Schreibweise) mit $\{A, B\}$.

Beweis. Sei C gemäß (M3) eine Menge mit $A \in C$ und $B \in C$. Nach (M2) können wir die Menge $\{M \in C\,;\, M = A \text{ oder } M = B\}$ bilden und erhalten eine Menge, die genau A und B als Elemente enthält. Gemäß (M1) ist diese Menge eindeutig bestimmt. □

Offensichtlich gilt nun für je zwei vorgegebene Mengen A und B:

$$A \cap B = \bigcap_{M \in \{A,B\}} M\,.$$

Eine weitere Möglichkeit aus vorgegebenen Mengen neue zu bilden liefert:

(M4) *Zu jeder Menge A gibt es eine Menge, die alle Teilmengen von A als Elemente enthält.*

0.7. Bemerkung. Sei A eine Menge. Dann gibt es genau eine Menge, die genau alle Teilmengen von A als Elemente hat. Diese Menge nennen wir die *Potenzmenge von A* und bezeichnen sie mit $\mathcal{P}(A)$.

Beweis. Sei $\mathfrak{A}$ gemäß (M4) eine Menge, welche alle Teilmengen von A als Elemente enthält. Nach (M2) können wir die Menge $\{B \in \mathfrak{A}\,;\, B \subseteq A\}$ bilden. Offensichtlich enthält diese Menge genau alle Teilmengen von A. Nach (M1) kann es höchstens eine Menge mit dieser Eigenschaft geben, womit auch die Eindeutigkeitsaussage bewiesen ist. □

Beispiele. (a) $\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$,
(b) $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$, $\mathcal{P}(\mathcal{P}\{1\}) = \{\emptyset, \{\emptyset\}, \{\{1\}\}, \{\emptyset, \{1\}\}\}$,
(c) $\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

0.8. DEFINITION. Seien A und B zwei Mengen. Die gemäß (M2) gebildete Menge

$$A \setminus B := \{x \in A\,;\, x \notin B\}$$

heißt die *Differenzmenge* zwischen A und B. Ist $B \subseteq A$, so nennt man $A \setminus B$ auch das *Komplement* von B in A und schreibt auch $\complement_A B$ statt $A \setminus B$. Ist vom Zusammenhang her klar, bezüglich welcher Menge A das Komplement gebildet wird, so schreibt man auch kurz $\complement B$.

BEISPIELE. Mit $A := \{1,2,3,4,5\}$, $B := \{3,4,5\}$, $C := \{1,2,3\}$ gilt $A \setminus B = \complement_A B = \{1,2\}$, $B \setminus A = \emptyset$, $B \setminus C = \{4,5\}$, $C \setminus B = \{1,2\}$.

Im folgenden Lemma fassen wir einige Rechenregeln für die Differenzmengenbildung und die Komplementbildung zusammen.

0.9. LEMMA. *Seien A, B, C und D Mengen mit $B \subseteq A$ und $C \subseteq A$. Dann gilt:*

(a) $\complement_A \emptyset = A \setminus \emptyset = A$, $\complement_A A = A \setminus A = \emptyset$.
(b) $A \setminus D = A \setminus (A \cap D) = \complement_A(A \cap D)$.
(c) $(A \setminus D) \cap D = \emptyset$.
(d) $B \cap \complement_A B = \emptyset$.
(e) $A \setminus (A \setminus D) = A \cap D$.
(f) $\complement_A(\complement_A B) = B$.
(g) $B \subseteq C \iff \complement_A C \subseteq \complement_A B$.

BEWEIS. Die Aussagen (a) – (d) sieht man unmittelbar ein.

Zu (e): Für alle $x \in A$ gilt: $x \in A \setminus (A \setminus D) \iff x \in A$ und $x \notin \{y \in A\,;\, y \notin D\}$ $\iff x \in A \cap D$.

(f) ist ein Spezialfall von (e).

Zu (g): Wir zeigen zunächst die Implikation „$\Longrightarrow$". Sei also $B \subseteq C$ und sei $x \in \complement_A C$ beliebig also $x \in A$ und $x \notin C$. Wegen $B \subseteq C$ ist dann auch $x \notin B$ und es folgt $x \in \complement_A B$. Da dies für alle $x \in \complement_A C$ gilt, folgt $\complement_A C \subseteq \complement_A B$. Wir müssen noch die Implikation „$\Longleftarrow$" beweisen. Sei also $\complement_A C \subseteq \complement_A B$ vorausgesetzt. Wendet man die bereits bewiesene Aussage „$\Longrightarrow$" an mit $\complement_A C$ statt B und $\complement_A B$ statt C, so erhalten wir mit (f): $B = \complement_A(\complement_A B) \subseteq \complement_A(\complement_A C) = C$. □

Um die Vereinigung von beliebigen Mengen bilden zu können, benötigen wir ein weiteres Axiom:

(M5) *Zu jeder Menge $\mathfrak{M}$ von Mengen gibt es eine Menge M mit $A \subseteq M$ für alle $A \in \mathfrak{M}$.*

0.10. DEFINITION. Sei $\mathfrak{M}$ eine Menge von Mengen und sei M eine gemäß (M5) existierende Menge mit $A \subseteq M$ für alle $A \in \mathfrak{M}$. Dann können wir nach dem Aussonderungsaxiom (M2) die Menge

$$\bigcup_{A \in \mathfrak{M}} A := \bigcup \{A\,;\, A \in \mathfrak{M}\} := \{x \in M\,;\, \exists A \in \mathfrak{M} :\, x \in A\}$$

bilden. Man überlegt sich, daß $\bigcup_{A \in \mathfrak{M}} A$ von der speziellen Menge M mit $A \subseteq M$ für alle $A \in \mathfrak{M}$ unabhängig ist. Die Menge $\bigcup_{A \in \mathfrak{M}} A$ heißt die *Vereinigung* der Mengen $A \in \mathfrak{M}$.

BEISPIELE. (a) Ist $\mathfrak{M} = \emptyset$, so ist auch $\bigcup_{A \in \mathfrak{M}} A = \emptyset$.

(b) Sind A und B zwei beliebige Mengen, so existiert nach (M3) und Bemerkung 0.6 die Menge $\{A, B\}$, die genau die Elemente A und B besitzt. Statt $\bigcup_{C\in\{A,B\}} C$ schreiben wir einfacher $A \cup B$. Die Menge $A \cup B$ ist also die Menge der Elemente, die Elemente von A oder von B sind.

(c) $\{1, 2, 3, 4\} \cup \{2, 3, 5\} = \{1, 2, 3, 4, 5\}$.

Die elementaren Beweise der folgenden Rechenregeln sind dem Leser als Übung überlassen.

0.11. Lemma. *A, B und C seien Mengen. Es gilt:*

(a) $A \cup A = A \cup \emptyset = A$.
(b) $A \subseteq A \cup B$.
(c) $A \cup B = B \cup A$.
(d) $A \cup (B \cup C) = (A \cup B) \cup C$.
(e) $A \subseteq B \iff A \cup B = B$.
(f) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
(g) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(h) $(A \setminus B) \cup (A \cap B) = A$.
(i) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.
(j) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

0.12. Definition. Seien A und B zwei Mengen. Für alle $a \in A, b \in B$ definieren wir

$$(a, b) := \{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$$

und bilden gemäß (M2) die Menge

$$A \times B := \{x \in \mathcal{P}(\mathcal{P}(A \cup B))\,;\, \exists a \in A, b \in B :\ x = (a, b)\}\,.$$

$A \times B$ heißt die *Produktmenge von A und B* und für $a \in A, b \in B$ heißt (a, b) das *geordnete Paar von a und b*.

0.13. Lemma. *Seien A und B zwei Mengen. Für alle $(a, b), (u, v) \in A \times B$ gilt:*

$$(a, b) = (u, v) \iff (a = u \text{ und } b = v).$$

Beweis. Die Beweisrichtung „$\Longleftarrow$“ ist nach Definition 0.12 offensichtlich.

Zu „$\Longrightarrow$“: Seien $a, u \in A, b, v \in B$ beliebig mit $(a, b) = (u, v)$, d.h. mit $\{\{a\}, \{a, b\}\} = \{\{u\}, \{u, v\}\}$. Wegen $\{a\} \in \{\{u\}, \{u, v\}\}$ folgt $\{a\} = \{u\}$ (und damit $a = u$) oder $\{a\} = \{u, v\}$, was $a = u = v$ zur Folge hat. In beiden Fällen gilt also $a = u$.

Zu zeigen bleibt noch $b = v$: Wegen $a = u$ ist $\{\{a\}, \{a, b\}\} = (a, b) = (a, v) = \{\{a\}, \{a, v\}\}$.

1. Fall: $v = a$. Dann ist $(a, v) = (a, a) = \{\{a\}, \{a, a\}\} = \{\{a\}\}$. Wegen $\{a, b\} \in (a, v) = (a, a) = \{\{a\}\}$ folgt also $v = a = b$.

2. Fall: $v \neq a$: Dann ist $\{a, v\} \neq \{a\}$ woraus wegen $\{a, v\} \in \{\{a\}, \{a, b\}\}$ folgt: $\{a, v\} = \{a, b\}$. Wegen $v \neq a$ ist dies nur möglich, wenn $v = b$ gilt.

In beiden Fällen haben wir also $v = b$ erhalten. □

Beispiele. (a) Sei $A := \{a, b\}, B := \{1, 2, 3\}$. Dann ist in aufzählender Schreibweise: $A \times B = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}$.

(b) Für $\mathbb{R} \times \mathbb{R} = \{z \in \mathcal{P}(\mathcal{P}(\mathbb{R}))\,;\, \exists x, y \in \mathbb{R} :\ (x, y) = z\}$ schreiben wir auch kürzer $\mathbb{R} \times \mathbb{R} = \{(x, y)\,;\, x, y \in \mathbb{R}\}$. Wir identifizieren $\mathbb{R} \times \mathbb{R}$ in der üblichen Weise mit der Ebene.

In dem folgenden Lemma fassen wir einige Rechenregeln für den Umgang mit Produktmengen zusammen.

0.14. Lemma. *Sind A, B, C, D Mengen, so gilt:*
(a) $A \times B = \emptyset \iff (A = \emptyset$ *oder* $B = \emptyset)$.
(b) $(A \cup B) \times C = (A \times C) \cup (B \times C)$.
(c) $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$.
(d) *Ist* $A \subseteq C$ *und* $B \subseteq D$, *so ist* $A \times B \subseteq C \times D$.

Beweis. (a) Ist $A = \emptyset$ oder $B = \emptyset$, so muß auch $A \times B := \{x \in \mathcal{P}(\mathcal{P}(A \cup B))\,;\, \exists a \in A, b \in B : x = (a, b)\} = \emptyset$ gelten. Ist $A \neq \emptyset$ und $B \neq \emptyset$, so gibt es wenigstens ein $a \in A$ und ein $b \in B$. Es ist dann $(a, b) \in A \times B$ und somit $A \times B \neq \emptyset$. Aus $A \times B = \emptyset$ folgt also $A = \emptyset$ oder $B = \emptyset$.
(b) – (d) rechnet man unmittelbar nach. □

0.15. Definition. Seien A und B zwei nicht leere Mengen. Unter einer *Abbildung* $f : A \to B$ von A nach B verstehen wir eine Vorschrift, die jedem Element $a \in A$ genau ein Element $f(a) \in B$ zuordnet. Zwei Abbildungen $f, g : A \to B$ heißen *gleich*, falls sie punktweise übereinstimmen, d.h. für alle $a \in A$ gilt: $f(a) = g(a)$.

A heißt der *Definitionsbereich* der Abbildung f und B der *Wertebereich*. Abbildungen mit Werten in den reellen Zahlen nennen wir auch *Funktionen*.

$f(a)$ heißt der Bildpunkt von $a \in A$ unter f. Für eine Teilmenge C von A heißt $f(C) := \{f(a)\,;\, a \in C\} = \{b \in B\,;\, \exists a \in C : b = f(a)\}$ die *Bildmenge* von C unter f. Ist $D \subseteq B$, so heißt die Menge $f^{-1}(D) := \{a \in A\,;\, f(a) \in D\}$ die *Urbildmenge* von D unter f. Die Menge

$$G(f) := \{(a, f(a))\,;\, a \in A\} := \{(a, b) \in A \times B\,; b = f(a)\}$$

nennen wir den *Graphen von f*.

0.16. Lemma. *Seien A und B zwei nicht leere Mengen.*
(a) *Eine Menge $G \subseteq A \times B$ ist genau dann Graph einer Abbildung von A nach B, wenn die beiden folgenden Bedingungen erfüllt sind.*
  (i) $\forall a \in A\, \exists b \in B : \quad (a, b) \in G$.
  (ii) *Sind* $(a, b_1), (a, b_2) \in G$ *so gilt schon* $b_1 = b_2$.
(b) *Zwei Abbildungen $f, g : A \to B$ sind genau dann gleich, wenn ihre Graphen gleich sind.*

Beweis. (a) Ist f eine Abbildung und $G = G(f) = \{(a, f(a))\,;\, a \in A\}$, so gibt es nach Definition der Abbildung zu jedem $a \in A$ genau einen Bildpunkt $f(a) \in B$. Dies zeigt, daß (i) und (ii) erfüllt sind. Ist umgekehrt $G \subseteq A \times B$ mit (i) und (ii) gegeben, so gibt es zu jedem $a \in A$ genau ein Element $f(a) \in B$ mit $(a, f(a)) \in G$. Hierdurch ist eine Abbildung $f : A \to B$ definiert, für die $G(f) = G$ gilt.

(b) „$\Longrightarrow$“ ist unmittelbar klar. Seien nun umgekehrt $f, g : A \to B$ zwei Abbildungen mit $G(f) = G(g)$. Dann gilt für alle $a \in A$: Es ist $(a, f(a)) \in G(f) = G(g)$ und damit $f(a) = g(a)$. □

Beispiele von Abbildungen bzw. Funktionen hat man schon in der Schule kennengelernt. Auch im Verlauf dieser Vorlesung werden uns viele Beispiele begegnen. Deshalb beschränken wir uns hier zunächst auf ein paar wenige, die von elementarer Natur sind.

0.17. Beispiele. Sei A, C, D nicht leere Mengen.
(a) Die Funktion $\mathrm{id}{:}A \to A$ mit $\mathrm{id}(a) := a$ für alle $a \in A$ heißt die *Identität* auf A.

(b) Sei $B \subseteq A$. Dann heißt $\chi_B : A \to \mathbb{R}$ mit

$$\chi_B(x) := \begin{cases} 1 & \text{für } x \in B \\ 0 & \text{für } x \in A \setminus B \end{cases}$$

die *charakteristische Funktion von A*.

(c) Die Abbildung $p : C \times D \to C$ mit $p(c,d) := c$ für alle $(c,d) \in C \times D$ heißt die *Projektion auf die erste Koordinate*.

Im folgenden Lemma fassen wir wieder einige Rechenregeln zusammen.

0.18. Lemma. *Sei $f : A \to B$ eine Abbildung und seien C, D Teilmengen von A sowie E, F Teilmengen von B. Dann gelten die folgenden Aussagen:*

(a) $f^{-1}(E \cap F) = f^{-1}(E) \cap f^{-1}(F)$.
(b) $f^{-1}(E \cup F) = f^{-1}(E) \cup f^{-1}(F)$.
(c) $f(C \cup D) = f(C) \cup f(D)$.
(d) $f(C \cap D) \subseteq f(C) \cap f(D)$, *wobei die Inklusion echt sein kann.*

Beweis. Für $a \in A$ gilt:

$$\begin{aligned} a \in f^{-1}(E \cap F) &\iff f(a) \in E \cap F \iff (f(a) \in E \text{ und } f(a) \in F) \\ &\iff (a \in f^{-1}(E) \text{ und } a \in f^{-1}(F)) \\ &\iff a \in f^{-1}(E) \cap f^{-1}(F) \end{aligned}$$

sowie

$$\begin{aligned} a \in f^{-1}(E \cup F) &\iff f(a) \in E \cup F \iff (f(a) \in E \text{ oder } f(a) \in F) \\ &\iff (a \in f^{-1}(E) \text{ oder } a \in f^{-1}(F)) \\ &\iff a \in f^{-1}(E) \cup f^{-1}(F)\,. \end{aligned}$$

Damit sind (a) und (b) bewiesen.

Für alle $b \in B$ gilt:

$$\begin{aligned} b \in f(C \cup D) &\iff \exists a \in C \cup D :\ f(a) = b \\ &\iff (b \in f(C) \text{ oder } b \in f(D)) \\ &\iff b \in f(C) \cup f(D) \end{aligned}$$

und damit (c). Ferner:

$$b \in f(C \cap D) \iff (\exists a \in C \cap D :\ f(a) = b) \Longrightarrow b \in f(C) \cap f(D)\,.$$

Wir zeigen durch Angabe eines Beispiels, daß die Inklusion in (d) echt sein kann. Wir definieren $f : \{1,2\} \to \{1\}$ durch $f(1) := f(2) := 1$. Es ist $\{1\} \cap \{2\} = \emptyset$ und daher $f(\{1\} \cap \{2\}) = \emptyset$ aber $f(\{1\}) \cap f(\{2\}) = \{1\} \neq \emptyset$. □

0.19. Definition. Sind $f : A \to B$ und $g : B \to C$, so definieren wir die *Komposition* (*Hintereinanderausführung*, *zusammengesetzte Funktion*) $g \circ f : A \to C$ durch $(g \circ f)(a) := g(f(a))$ für alle $a \in A$.

Die Hintereinanderausführung von Abbildungen ist assoziativ:

0.20. Lemma. *Sind $f : A \to B$, $g : B \to C$ und $h : C \to D$ Abbildungen, so gilt $(h \circ g) \circ f = h \circ (g \circ f)$.*

Beweis. Für alle $a \in A$ gilt:

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a))) = h((g \circ f)(a)) = (h \circ (g \circ f))(a)\,.$$

□

0.21. Beispiel. Sei $\mathbb{R}_+ := \{x \in \mathbb{R}\,;\, x \geq 0\}$. Wir definieren $f : \mathbb{R} \to \mathbb{R}_+$ und $g : \mathbb{R}_+ \to \mathbb{R}$ durch $f(x) := x^2$ für alle $x \in \mathbb{R}$ und $g(x) := \sqrt{x}$ für alle reellen $x \geq 0$. Dann gilt $f \circ g = \mathrm{id}_{\mathbb{R}_+}$ und $(g \circ f)(x) = |x|$ für alle $x \in \mathbb{R}$.

0.22. Definition. Eine Abbildung $f : A \to B$ heißt

(a) *injektiv*, falls für alle $x, y \in A$ aus $f(x) = f(y)$ schon $x = y$ folgt.
(b) *surjektiv*, falls $f(A) = B$ ist, d.h. falls es zu jedem $b \in B$ ein $a \in A$ mit $f(a) = b$ gibt.
(c) *bijektiv*, falls f injektiv und surjektiv ist.

0.23. Beispiel. Seien $f : \mathbb{R} \to \mathbb{R}_+$ und $g : \mathbb{R}_+ \to \mathbb{R}$ wie in Beispiel 0.21. Dann ist f surjektiv aber nicht injektiv und g injektiv aber nicht surjektiv.

0.24. Satz. *Sei $f : A \to B$ eine Abbildung. Es gilt:*

(a) *f ist genau dann bijektiv, wenn es eine Abbildung $g : B \to A$ gibt mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$.*
(b) *Ist f bijektiv, so ist die Funktion g aus (a) eindeutig bestimmt.*

Beweis. (a) Sei $f : A \to B$ bijektiv. Ist $b \in B$ beliebig, so gibt es, da f bijektiv ist, genau ein $a \in A$ mit $f(a) = b$. Wir setzen $g(b) := a$. Hierdurch wird eine Abbildung $g : B \to A$ definiert mit $g(f(a)) = a$ für alle $a \in A$ und $f(g(b)) = b$ für alle $b \in B$.

Ist umgekehrt $g : B \to A$ eine Abbildung mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$, so gilt für alle $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$: Es ist

$$a_1 = \mathrm{id}_A(a_1) = g(f(a_1)) = g(f(a_2)) = \mathrm{id}_A(a_2) = a_2\,.$$

Also ist f injektiv. Für alle $b \in B$ hat man $b = \mathrm{id}_B(b) = f(g(b))$ mit $g(b) \in A$, d.h. f ist auch surjektiv und damit bijektiv.

(b) Ist auch $h : B \to A$ eine Funktion mit $f \circ h = \mathrm{id}_B$ und $h \circ f = \mathrm{id}_A$, so folgt wegen der Assoziativität der Hintereinanderausführung (Lemma 0.20): $h = h \circ \mathrm{id}_B = h \circ (f \circ g) = (h \circ f) \circ g = \mathrm{id}_A \circ g = g$. $\square$

0.25. Definition. Sei $f : A \to B$ eine bijektive Abbildung. Die gemäß Satz 10.14 eindeutig bestimmte Abbildung $g : B \to A$ mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$ heißt die *Umkehrabbildung* (auch *inverse Abbildung*, *Umkehrfunktion* oder *inverse Funktion*). Man schreibt auch f^{-1} für g.

2. Verschiedene Alphabete

2.1. Die deutsche Schreibschrift (Sütterlin).

Großbuchstaben		Kleinbuchstaben	
A	A	a	a
B	B	b	b
C	C	c	c
D	D	d	d
E	E	e	e
F	F	f	f
G	G	g	g
H	H	h	h
I	I	i	i
J	J	j	j
K	K	k	k
L	L	l	l
M	M	m	m
N	N	n	n
O	O	o	o
P	P	p	p
Q	Q	q	q
R	R	r	r
S	S	s	s
T	T	t	t
U	U	u	u
V	V	v	v
W	W	w	w
X	X	x	x
Y	Y	y	y
Z	Z	z	z

2.2. Das griechische Alphabet.

Name	Großbuchstabe	Kleinbuchstabe
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ϵ, ε
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ, ϑ
Iota	I	ι
Kappa	K	$\kappa, \varkappa$
Lambda	Λ	λ
My	M	μ
Ny	N	ν
Xi	Ξ	ξ
Omikron	O	o
Pi	Π	π, ϖ
Rho	P	ρ, ϱ
Sigma	Σ	σ, ς
Tau	T	τ
Ypsilon	Υ	υ
Phi	Φ	ϕ, φ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

2.3. Hebräisches Alphabet.

Aus diesem Alphabet werden wir nur den den ersten Buchstaben *Aleph*: $\aleph$ verwenden.

KAPITEL 1

Der Körper $\mathbb{R}$ der reellen Zahlen

Der Körper der reellen Zahlen ist uns natürlich schon von der Schule her bekannt und vertraut. Wir wollen ihn hier nochmals (in knapper Form) axiomatisch einführen. Wir fordern die Existenz eines Körpers $\mathbb{R}$, der den (später in Definition 1.5 angegebenen) Anordnungsaxiomen (P1)-(P3) und dem Supremumsaxiom (S) (siehe weiter unten in 1.14) genügt.

1. Körper

Wir beginnen zunächst mit der Definition eines Körpers:

1.1. Definition. Ein Menge $\mathbb{K}$ heißt *Körper*, falls sie mit zwei Abbildungen

$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K} \quad (x, y) \mapsto x + y, \quad (\textit{Addition} \text{ genannt})$$

und

$$\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K} \quad (x, y) \mapsto xy, \quad (\textit{Multiplikation} \text{ genannt})$$

versehen ist, so daß die folgenden Bedingungen erfüllt sind:

(KA1) Es gilt das *Assoziativgesetz der Addition*:

$$\forall x, y, z \in \mathbb{K} : \quad (x + y) + z = x + (y + z).$$

(KA2) Existenz eines *neutralen* Elementes der Addition: Es gibt ein Element $0 \in \mathbb{K}$ mit $x + 0 = x$ für alle $x \in \mathbb{K}$.

(KA3) Existenz von bezüglich der Addition *inversen* Elementen:

$$\forall x \in \mathbb{K}\ \exists y \in \mathbb{K} : \qquad x + y = 0\,.$$

(KA4) Es gilt das *Kommutativgesetz der Addition*:

$$\forall x, y \in \mathbb{K} : \qquad x + y = y + x\,.$$

(KM1) Es gilt das *Assoziativgesetz der Multiplikation*:

$$\forall x, y, z \in \mathbb{K} : \qquad (xy)z = x(yz).$$

(KM2) Existenz eines *neutralen* Elementes der Multiplikation: Es gibt ein Element $1 \in \mathbb{K}_* := \mathbb{K} \setminus \{0\}$ mit $1 \cdot x = x$ für alle $x \in \mathbb{K}$.

(KM3) Existenz von bezüglich der Multiplikation *inversen* Elementen in $\mathbb{K}_*$:

$$\forall x \in \mathbb{K}_*\ \exists y \in \mathbb{K} : \qquad xy = 1\,.$$

(KM4) Es gilt das *Kommutativgesetz der Multiplikation*:

$$\forall x, y \in \mathbb{K} : \qquad xy = yx\,.$$

(KD) Es gilt das *erste Distributivgesetz*:

$$\forall x, y, z \in \mathbb{K} : \qquad (x + y)z = xz + yz\,.$$

1.2. Lemma. *Sei* $(\mathbb{K}, +, \cdot)$ *ein Körper.*

(a) *Das neutrale Element* 0 *der Addition ist eindeutig bestimmt. Wir nennen es das Nullelement von* $\mathbb{K}$ *oder die Null von* $\mathbb{K}$.

(b) *Das neutrale Element* 1 *der Multiplikation ist eindeutig bestimmt. Wir nennen es das Einselement von* $\mathbb{K}$ *oder die Eins von* $\mathbb{K}$.

(c) *Sei* $x \in \mathbb{K}$. *Das nach* (KA3) *existierende zu* x *bezüglich der Addition inverse Element ist eindeutig bestimmt. Wir bezeichnen es mit* $-x$. *Es ist* $-(-x) = x$. *Statt* $y + (-x)$ *schreiben wir auch* $y - x$ *für alle* $x, y \in \mathbb{K}$.

(d) *Sei* $x \in \mathbb{K}_*$. *Das nach* (KM3) *existierende zu* x *bezüglich der Multiplikation inverse Element ist eindeutig bestimmt. Wir bezeichnen es mit* x^{-1} *oder* $\frac{1}{x}$. *Es ist* $(x^{-1})^{-1} = x$. *Für alle* $y \in \mathbb{K}$ *schreiben wir auch* $\frac{y}{x}$ *statt* $y \cdot \frac{1}{x}$.

Beweis. (a) Sei auch $0' \in \mathbb{K}$ mit $x + 0' = x$ für alle $x \in \mathbb{K}$. Speziell mit $x := 0$ folgt unter Verwendung von (KA4): $0 = 0 + 0' = 0' + 0 = 0'$.

(b) Sei auch $1' \in \mathbb{K}$ mit $1' \cdot x = x$ für alle $x \in \mathbb{K}$. Speziell mit $x := 1$ folgt unter Verwendung von (KM4): $1 = 1' \cdot 1 = 1 \cdot 1' = 1'$.

(c) Seien $u, v \in \mathbb{K}$ mit $x + u = 0 = x + v$. Unter Verwendung von (KA1), (KA2) und (KA4) folgt: $u = u + 0 = u + (x + v) = (u + x) + v = (x + u) + v = 0 + v = v + 0 = v$.

(d) Seien $u, v \in \mathbb{K}$ mit $xu = 1 = xv$. Unter Verwendung von (KM1), (KM2) und (KM4) folgt: $u = 1 \cdot u = (xv)u = (vx)u = v(xu) = v \cdot 1 = 1 \cdot v = v$. □

Im folgenden Lemma fassen wir einige weitere Rechenregeln zusammen:

1.3. Lemma. *Sei* $(\mathbb{K}, +, \cdot)$ *ein Körper.*

(a) *Für alle* $u, v \in \mathbb{K}$ *hat die Gleichung* $u + x = v$ *genau eine Lösung* $x \in \mathbb{K}$. *Es ist* $x = v - u$.

(b) *Für alle* $u \in \mathbb{K}_*, v \in \mathbb{K}$ *hat die Gleichung* $ux = v$ *genau eine Lösung* $x \in \mathbb{K}$. *Es ist* $x = \frac{v}{u}$.

(c) *Es gilt das zweite Distributivgesetz:*

$$\forall x, y, z \in \mathbb{K}: \quad x(y + z) = xy + xz\,.$$

(d) $\forall x \in \mathbb{K}: \quad x \cdot 0 = 0 \cdot x = 0$.

(e) $\mathbb{K}$ *ist nullteilerfrei, d.h.: Sind* $u, v \in \mathbb{K}$ *mit* $uv = 0$, *so ist schon* $u = 0$ *oder* $v = 0$.

(f) $\forall x \in \mathbb{K}: \quad (-1) \cdot x = -x$.

(g) *Ist* $u \in \mathbb{K}$ *mit* $u^2 := u \cdot u = u$, *so ist schon* $u = 0$ *oder* $u = 1$.

(h) *Für alle* $x \in \mathbb{K}_*$ *ist auch* $x^{-1} \in \mathbb{K}_*$.

Beweis. (a) Daß $v - u$ eine Lösung der Gleichung $u + x = v$ ist rechnet man mit Hilfe von (KA1)–(KA4) unmittelbar nach. Ist $x \in \mathbb{K}$ mit $u + x = v$, so folgt (ebenfalls unter Verwendung von (KA1)–(KA4)): $x = 0 + x = (-u + u) + x = -u + (u + x) = -u + v = v - u$.

(b) ist die multiplikative Variante von (a) und wird analog bewiesen (unter Verwendung von (KM1)–(KM4)).

(c) folgt mit Hilfe von (KM4) aus (KD).

(d) Nach (KA2) ist $0 + 0 = 0$ und somit nach (KD): $0 \cdot x = 0 \cdot x + 0 \cdot x$. Nach (a) ist also $0 \cdot x = 0 \cdot x - 0 \cdot x = 0$.

(e) Seien $u, v \in \mathbb{K}$ mit $uv = 0$. Ist $u = 0$, so ist nichts zu zeigen. Ist $u \neq 0$, so folgt nach (b): Es gibt genau ein $x \in \mathbb{K}$ mit $ux = 0$ und es ist $x = \frac{1}{u} \cdot 0 = 0$ (wegen (d)).

(f) Für alle $x \in \mathbb{K}$ gilt unter Verwendung von (KA2), (KA4), (KD) und (d):

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 - 1)x = 0 \cdot x = 0$$

und somit nach Lemma 1.2: $(-1) \cdot x = -x$.

(g) Ist $p = 0$, so ist nach (d): $p^2 = 0 \cdot 0 = 0$. ist $p \neq 0$, so gibt es nach (b) genau ein $x \in \mathbb{K}$ mit $px = p$ und es ist $x = \frac{1}{p} \cdot p = 1$. Ist also $p^2 = p$ und $p \neq 0$, so muß $p = 1$ gelten.

(h) Sei $x \in \mathbb{K}_*$ beliebig. Wegen $0 \cdot x = 0$ (nach (d)) und $0 \neq 1$ nach (KM2) muß $x^{-1} \neq 0$ gelten. □

Wegen (KA2) und (KM2) besitzt ein Körper mindestens zwei Elemente (nämlich 0 und 1). Das folgende Beispiel zeigt, daß es tatsächlich Körper mit nur zwei Elementen gibt.

1.4. Beispiel. Sei $\mathbb{F}_2 := \{0,1\}$ eine Menge mit zwei Elementen ($0 \neq 1$). Wir definieren auf $\mathbb{F}_2$ eine Addition und eine Multiplikation durch

$$0+0:=0, \quad 0+1:=1, \quad 1+0:=1, \quad 1+1:=0,$$
$$0\cdot 0:=0, \quad 0\cdot 1:=0, \quad 1\cdot 0:=0, \quad 1\cdot 1:=1.$$

Dann sind (wie man durch Inspektion, also Nachrechnen aller möglichen Fälle sieht) die Körpergesetze (KA1)–(KA4), (KM1)–(KM4) und (KD) erfüllt. $(\mathbb{F}_2,+,\cdot)$ ist also ein Körper.

2. Angeordnete Körper

Das Beispiel 1.4 zeigt, daß wir weitere Eigenschaften benötigen, um den von der Schule her vertrauten Körper der reellen Zahlen zu charakterisieren.

1.5. Definition. Ein Körper $(\mathbb{K},+,\cdot)$ heißt *angeordnet*, falls es in ihm eine Teilmenge P gibt, die den folgenden Bedingungen (P1)–(P3) genügt:

(P1) Für jedes $x \in \mathbb{K}$ gilt genau eine der drei folgenden Aussagen:

$$x=0, \quad x \in P, \quad -x \in P \qquad \textit{Trichtonomie}$$

(P2) P ist abgeschlossen bezüglich der Addition, d.h. für alle $u,v \in P$ gilt $u+v \in P$.

(P3) P ist abgeschlossen bezüglich der Multiplikation, d.h. für alle $u,v \in P$ gilt $u \cdot v \in P$.

Sei $(\mathbb{K},+,\cdot)$ ein angeordneter Körper und $P \subset \mathbb{K}$ mit (P1)–(P3). Wir nennen dann P die *Menge der positiven Zahlen in* $\mathbb{K}$. $x \in \mathbb{K}$ heißt *positiv*, falls $x \in P$ und *negativ*, falls $-x \in P$. Sind $u,v \in \mathbb{K}$, so sagen wir *u ist kleiner als v* (und schreiben $u < v$), falls $v-u \in P$ sowie *u ist größer als v* (geschrieben $u > v$), falls $u - v \in P$. Ferner definieren wir für $u,v \in \mathbb{K}$:

$$u \geq v :\Longleftrightarrow \quad (u > v \text{ oder } u = v) \qquad \text{„}u \textit{ ist größer oder gleich } v\text{“}$$
$$u \leq v :\Longleftrightarrow \quad (u < v \text{ oder } u = v) \qquad \text{„}u \textit{ ist kleiner oder gleich } v\text{“}$$

Die Rechenregeln in dem folgenden Lemma rechnet man leicht nach. Wir werden sie im folgenden laufend verwenden ohne jeweils im einzelnen darauf hinzuweisen.

1.6. Lemma. *Sei $(\mathbb{K},+,\cdot)$ ein angeordneter Körper und $P \subset \mathbb{K}$ die Menge der positiven Zahlen von $\mathbb{K}$. Für alle $a,b,c,d \in \mathbb{K}$ gilt:*

(a) $a > 0 \iff a \in P, \quad a < 0 \iff -a \in P.$

(b) *Es gilt genau eine der drei Beziehungen: $a = 0$, $a > 0$, $a < 0$.*

(c) *Es gelten die folgenden Transitivitätsaussagen:*

  (i) $(a < b$ *und* $b < c) \Longrightarrow a < c.$

  (ii) $(a < b$ *und* $b \leq c) \Longrightarrow a < c.$

  (iii) $(a \leq b$ *und* $b < c) \Longrightarrow a < c.$

  (iv) $(a \leq b$ *und* $b \leq c) \Longrightarrow a \leq c.$

(d) $(a < b$ *und* $c \leq d) \Longrightarrow a + c < b + d.$
$(a \leq b$ *und* $c \leq d) \Longrightarrow a + c \leq b + d.$

(e) $(a < b$ *und* $c > 0) \Longrightarrow ac < bc.$
$(a \leq b$ *und* $c \geq 0) \Longrightarrow ac \leq bc.$

(f) *$(a < b$ und $c < 0) \Longrightarrow ac > bc$.*
$(a \leq b$ und $c \leq 0) \Longrightarrow ac \geq bc$.
(g) *Ist $a \neq 0$, so ist $a^2 = a \cdot a > 0$. Insbesondere ist $1 = 1^2 > 0$.*
(h) *Kürzungsregeln: $(a + c \leq b$ und $c \geq 0) \Longrightarrow a \leq b$,*
$(a + c < b$ und $c \geq 0) \Longrightarrow a < b$,
$(0 \leq ac \leq b$ und $c \geq 1 \Longrightarrow a \leq b$,
$(0 \leq ac < b$ und $c \geq 1 \Longrightarrow a < b$.
(i) *Regeln für die Inversenbildung: $a \leq b \Longrightarrow -a \geq -b$, $a < b \Longrightarrow -a > -b$,*
$0 < a \leq b \Longrightarrow 0 < b^{-1} \leq a^{-1}$, $0 < a < b \Longrightarrow 0 < b^{-1} < a^{-1}$.

1.7. Definition. Sei $(\mathbb{K}, +, \cdot)$ ein angeordneter Körper. Für $x \in \mathbb{K}$ heißt die durch

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

definierte Zahl der *Absolutbetrag* (oder auch kurz der *Betrag*) von x.

Wir fassen einige Eigenschaften des Absolutbetrags im folgenden Lemma zusammen:

1.8. Lemma. *Sei $(\mathbb{K}, +, \cdot)$ ein angeordneter Körper.*
(a) $\forall x \in \mathbb{K}: \quad |x| \geq 0$. *Es ist $|x| = 0$ genau dann, wenn $x = 0$.*
(b) *Für alle $u, v \in \mathbb{K}$ gilt: $|u| \leq v \iff -v \leq u \leq v$.*
(c) $\forall x \in \mathbb{K}: \quad -|x| \leq x \leq |x|$.
(d) *Für alle $u, v \in \mathbb{K}$ gilt: $|uv| = |u| \cdot |v|$. Insbesondere gilt (mit $v := -1$): $|-u| = |u|$ für alle $u \in \mathbb{K}$.*
(e) *Für alle $u, v \in \mathbb{K}$ gilt:*
  (i) $|u + v| \leq |u| + |v|$ *Dreiecksungleichung.*
  (ii) $||u| - |v|| \leq |u - v|$.

Beweis. (a) folgt unmittelbar aus der Definition des Absolutbetrags.

(b) Ist $u \geq 0$, so ist $0 \leq u = |u|$ und die behauptete Äquivalenz ist offensichtlich. Ist $u < 0$, so ist $0 < |u| = -u$. Ist also $|u| \leq v$, so ist $v > 0$ und $-u \leq v$, also $-v \leq u < 0 < v$. Gilt umgekehrt $-v \leq u \leq v$, so folgt aus der linken Ungleichung mit Lemma 1.6 (i): $|u| = -u \leq v$.

(c) folgt mit $v := |u|$ aus (b).

(d) Ist $u \geq 0$ und $v \geq 0$ (bzw. $u \leq 0$ und $v \leq 0$), so folgt nach Lemma 1.6 (e) (bzw. (f)): $uv \geq 0$ und nach der Definition des Absolutbetrags $|uv| = uv = |u| \cdot |v|$ (bzw. $|uv| = (-u)(-v) = |u| \cdot |v|$). Ist $u \geq 0$ und $v \leq 0$ (bzw. $u \leq 0$ und $v \geq 0$), so folgt nach Lemma 1.6 (f): Es ist $uv \leq 0$ und somit $|uv| = -uv = u(-v) = |u| \cdot |v|$ (bzw. $|uv| = -uv = (-u)v = |u| \cdot |v|$).

(e) Wegen $-|u| \leq u \leq |u|$ und $-|v| \leq u \leq |v|$ (nach (c)) folgt die Dreiecksungleichung durch Anwendung von Lemma 1.6 (d).

Zu (ii): Nach der Dreiecksungleichung gilt: $|u| = |u - v + v| \leq |u - v| + |v|$ und damit $|u| - |v| \leq |u - v|$. Wegen $|v| = |v - u + u| \leq |v - u| + |u|$ folgt mit (c) auch $-|u - v| = -|v - u| \leq |u| - |v|$. Mit (b) folgt als die behauptete Ungleichung. □

3. Obere Schranken und das Supremumsaxiom

1.9. Definition. Sei $(\mathbb{K}, +, \cdot)$ ein angeordneter Körper und $\emptyset \neq A \subseteq \mathbb{K}$. A heißt *nach oben* (bzw. *nach unten*) *beschränkt*, falls es ein $b \in \mathbb{K}$ gibt, so daß für alle $a \in A$ gilt: $a \leq b$ (bzw. $b \leq a$). b heißt dann eine *obere* (bzw. *untere*) *Schranke* für A.

1.10. Beispiel. Sei $(\mathbb{K}, +, \cdot)$ ein angeordneter Körper und

$$A := \{x \in \mathbb{K}\,;\, x^2 \leq 2 := 1 + 1\}\,.$$

Dann ist $A \neq \emptyset$, denn wegen $1 = 1^2 < 1 + 1 = 2$ (nach Lemma 1.6 (g) und (d)) ist $1 \in A$. Für alle $u \in \mathbb{K}$ mit $u > 2$ gilt $2 < 2u < u^2$ und daher $u \notin A$. Also gilt für alle $a \in A$: Es ist $a \leq 2$. A ist also nach oben beschränkt und 2 ist eine obere Schranke für A.

1.11. Lemma. *Sei $(\mathbb{K}, +, \cdot)$ ein angeordneter Körper und $\emptyset \neq A \subseteq \mathbb{K}$ nach oben beschränkt. Dann gilt:*

(a) *Ist x eine obere Schranke für A und ist $y \geq x$, so ist auch y eine obere Schranke für A.*

(b) *Die Menge*

$$\mathcal{S} := \{x \in \mathbb{K}\,;\, x \text{ ist obere Schranke für } A\}$$

ist nach unten beschränkt. In $\mathcal{S}$ gibt es höchstens ein kleinstes Element.

Beweis. (a) Ist x obere Schranke für A und $y \geq x$, so gilt für alle $a \in A$: $a \leq x$ und $x \leq y$ also auch $a \leq y$, d.h. y ist obere Schranke für A.

(b) Wegen $A \neq \emptyset$ gibt es wenigstens ein $a \in A$. Für alle oberen Schranken x von A ist $a \leq x$. Die Menge $\mathcal{S}$ ist also nach unten beschränkt und a ist eine untere Schranke für $\mathcal{S}$. Sind $u, v \in \mathcal{S}$ mit $u \leq x$ und $v \leq x$ für alle $x \in \mathcal{S}$, so folgt insbesondere $u \leq v$ und $v \leq u$, d.h. $u = v$. $\square$

1.12. Bemerkungen. (a) Lemma 1.11 macht keine Aussage über die *Existenz* einer kleinsten oberen Schranke für A.

(b) Eine zu Lemma 1.11 analoge Aussage gilt natürlich auch für nach unten beschränkte Mengen und untere Schranken.

(c) Die Menge $P := \{x \in \mathbb{K}\,;\, x > 0\}$ der positiven Zahlen ist nach unten durch 0 beschränkt. P besitzt kein kleinstes Element, denn für alle $x \in P$ gilt $0 < \frac{x}{2} < x$.

1.13. Definition. Sei $\emptyset \neq A$ eine nach oben beschränkte Teilmenge eines angeordneten Körpers $\mathbb{K}$. Besitzt die Menge $\mathcal{S}$ der oberen Schranken von A ein kleinstes Element y, so nennt man y die *kleinste obere Schranke* oder das *Supremum* von A und schreibt $y = \sup A$. Existiert $y = \sup A$ und ist $y \in A$, so daß also y das größte Element von A ist, so nennt man y das *Maximum* von A und schreibt $y = \max A$.

Analog definiert man: Ist $A \neq \emptyset$ eine nach unten beschränkte Teilmenge von $\mathbb{K}$ und besitzt die Menge der unteren Schranken von A ein größtes Element y, so nennt man dieses die *größte untere Schranke* oder das *Infimum* von A und schreibt $y = \inf A$. Existiert $y = \inf A$ und gilt $y \in A$ so nennt man y das *Minimum* von A und schreibt $y = \min A$.

Wir fordern nun:

1.14. Axiom (Supremumsaxiom). *Es gibt einen angeordneten Körper $\mathbb{R}$ mit der Eigenschaft*

(S) *Für jede nicht leere nach oben beschränkte Menge $A \subseteq \mathbb{R}$ existiert das Supremum $\sup A$ in $\mathbb{R}$.*

Wir nennen $\mathbb{R}$ den Körper der reellen Zahlen.

Man kann zeigen, daß es „im wesentlichen“ nur einen angeordneten Körper gibt, der dem Supremumsaxiom genügt.

Wir definieren zur Abkürzung:

$$\mathbb{R}_+ := \{x \in \mathbb{R}\,;\, x \geq 0\} \text{ und } \mathbb{R}_+^* := \{x \in \mathbb{R}\,;\, x > 0\}\,.$$

1.15. FOLGERUNG. *Für jede nach unten beschränkte nicht leere Menge* $A \subseteq \mathbb{R}$ *existiert das Infimum* $\inf A$ *von* A *in* $\mathbb{R}$.

BEWEIS. Die Menge $B := \{b \in \mathbb{R}\,;\, -b \in A\}$ ist nicht leer, da A nicht leer ist. Ist x untere Schranke für A (wenigstens eine solche existiert nach Voraussetzung), so gilt für alle $a \in A$: $x \leq a$ also auch $-a \leq -x$. Die Menge B ist also nach oben durch $-x$ beschränkt. Nach dem Supremumsaxiom existiert also $y := \sup B$ in $\mathbb{R}$. Wir zeigen: $-y = \inf A$.

Wegen $y = \sup B$ gilt für alle $a \in A$: $-a \leq y$ also $-y \leq a$. Also ist $-y$ eine untere Schranke für A. Ist $x \in \mathbb{R}$ eine beliebige untere Schranke von A, so ist $-x$ eine obere Schranke für B und es folgt nach Definition des Supremums: $y = \sup B \leq -x$, also auch $x \leq -y$. Dies zeigt, daß $-y$ die größte untere Schranke für A ist. □

Die folgende Hilfsaussage ist häufig nützlich für die Berechnung von Suprema bzw. Infima konkreter Mengen.

1.16. LEMMA. *Sei* $\emptyset \neq A \subset \mathbb{R}$ *nach oben (bzw. nach unten) beschränkt. Für* $y \in \mathbb{R}$ *gilt* $y = \sup A$ *(bzw.* $y = \inf A$*) genau dann, wenn die beiden folgenden Bedingungen erfüllt sind:*

(a) y *ist obere (bzw. untere) Schranke von* A.
(b) $\forall \varepsilon > 0\, \exists a \in A: \quad a > y - \varepsilon \quad (\textit{bzw.}\, a < y + \varepsilon)$.

BEWEIS. (Nur für Suprema) „$\Longrightarrow$“: Sei also $y = \sup A$. Dann ist y insbesondere obere Schranke und (a) ist erfüllt. Sei nun $\varepsilon > 0$ beliebig. Da y die *kleinste* obere Schranke von A ist, kann $y - \varepsilon$ keine obere Schranke für A sein. Es gibt also ein $a \in A$ mit $a > y - \varepsilon$.

„$\Longleftarrow$“: Seien nun die Bedingungen (a) und (b) für ein gegebenes $y \in \mathbb{R}$ erfüllt. Ist $x < y$, so ist $\varepsilon := y - x > 0$ und nach Voraussetzung gibt es ein $a \in A$ mit $a > y - \varepsilon = x$. x kann also keine obere Schranke für A sein. Somit ist y die kleinste obere Schranke für A. □

Gelegentlich ist es nützlich, $\mathbb{R}$ noch um zwei „uneigentliche“ Elemente $+\infty$ und $-\infty$ zu erweitern. Wir nennen $\hat{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ die *erweiterte Zahlengerade* und definieren die folgenden Rechenvorschriften:

(a) $\forall a \in \mathbb{R}: \quad -\infty < a < \infty$.
(b) Für $-\infty \neq a \in \hat{\mathbb{R}}$ sei $a + \infty := \infty + a := \infty$.
(c) Für $\infty \neq a \in \hat{\mathbb{R}}$ sei $a - \infty := -\infty + a := -\infty$.
(d) Für $0 < a \in \hat{\mathbb{R}}$ sei $a \cdot \infty := \infty \cdot a := \infty$ und $a \cdot (-\infty) := (-\infty) \cdot a := -\infty$.
(e) Für $0 \neq a \in \mathbb{R}$ sei $\dfrac{a}{\infty} := \dfrac{a}{-\infty} := 0$

Mit dieser Erweiterung können wir Infimum und Supremum für *alle* Teilmengen A von $\mathbb{R}$ definieren. Wir definieren für $A \subseteq \mathbb{R}$:

$$\sup A := \begin{cases} \text{kleinste obere Schranke von } A & \text{falls } A \text{ nach oben beschränkt ist} \\ -\infty & \text{falls } A = \emptyset \\ +\infty & \text{falls } A \neq \emptyset \text{ und nicht nach oben beschränkt.} \end{cases}$$

und entsprechend:

$$\inf A := \begin{cases} \text{größte untere Schranke von } A & \text{falls } A \text{ nach unten beschränkt ist} \\ +\infty & \text{falls } A = \emptyset \\ -\infty & \text{falls } A \neq \emptyset \text{ und nicht nach unten beschränkt.} \end{cases}$$

Für $A, B \subseteq \mathbb{R}$ und $\lambda \in \mathbb{R}$ führen wir ferner noch folgende Bezeichnungen ein:

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

Offensichtlich hat man $\lambda A = \{\lambda\}A$ und $\lambda \pm A = \{\lambda\} \pm A$. Im folgenden Lemma fassen wir einige Rechenregeln für Infima und Suprema beliebiger nicht leerer Teilmengen von $\mathbb{R}$ zusammen.

1.17. Lemma. *Für nicht leere Teilmengen A, B von $\mathbb{R}$ gilt:*

(a) $\forall 0 > \lambda \in \mathbb{R}$*:* $\inf(\lambda A) = \lambda \cdot \sup A$ *und* $\sup(\lambda A) = \lambda \cdot \inf A$.
(b) $\forall 0 < \lambda \in \mathbb{R}$*:* $\sup(\lambda A) = \lambda \cdot \sup A$ *und* $\inf(\lambda A) = \lambda \cdot \inf A$.
(c) $A \subseteq B \implies (\sup A \leq \sup B$ *und* $\inf A \geq \inf B)$.
(d) $\sup(A \cup B) = \max\{\sup A, \sup B\}$ *und* $\inf(A \cup B) = \min\{\inf A, \inf B\}$.
(e) $\sup(A + B) = \sup A + \sup B$ *und* $\inf(A + B) = \inf A + \inf B$.
(f) $\sup(A - B) = \sup A - \inf B$ *und* $\inf(A - B) = \inf A - \sup B$.
(g) *Ist* $A \subseteq P = \{x \in \mathbb{R}\,;\ x > 0\}$*, so ist A genau dann nach oben beschränkt, wenn gilt* $\inf\{a^{-1};\ a \in A\} > 0$*. Es ist dann*

$$\sup A = \left(\inf\{a^{-1};\ a \in A\}\right)^{-1}.$$

Beweis. (a) Ist $\sup A = \infty$, so ist A nicht nach oben beschränkt und es gibt zu jedem $r \in \mathbb{R}$ ein $a \in A$ mit $a > r/\lambda$, also auch mit $\lambda a < r$. Dies zeigt, daß λA nicht nach unten beschränkt ist und damit $\inf \lambda A = -\infty = \lambda \cdot \infty = \lambda \sup A$ in diesem Fall erfüllt ist. Sei nun A nach oben beschränkt, so daß also $\sup A$ in $\mathbb{R}$ existiert. Für alle $a \in A$ folgt $a \leq \sup A$ und damit auch $\lambda a \geq y := \lambda \sup A$. y ist also eine untere Schrank für λA. Ist $x \in \mathbb{R}$ mit $x > y$ so ist $\frac{x}{\lambda} < \frac{y}{\lambda} = \sup A$. Da $\sup A$ die kleinste obere Schranke für A ist, gibt es ein $a \in A$ mit $a > \frac{x}{\lambda}$. Es folgt $\lambda a < x$. x ist somit keine obere Schranke für λA. Also ist $y = \lambda \sup A$ die größte untere Schranke für λA und die linke Gleichung in (a) ist bewiesen. Indem man die linke Gleichung anwendet auf $A' := \lambda A$ statt A und λ^{-1} statt λ erhält man auch die rechte Gleichung. Insbesondere gilt $\sup(-A) = -\inf A$ und $\inf(-A) = -\sup A$.

(b) Nach (a) (angewendet mit $-\lambda$ statt λ) gilt für $\lambda > 0$:

$$\sup \lambda A = \sup(-\lambda)(-A) = -\lambda \inf(-A) = \lambda \sup A.$$

Auch die zweite Gleichung in (b) erhält man auf diese Weise mit Hilfe von (a).

(c) Sei $A \subseteq B$. Ist B nicht nach oben beschränkt, also $\sup B = \infty$, so gilt trivialerweise $\sup A \leq \sup B$. Ist B nach oben beschränkt, so folgt für alle $a \in A$ auch $a \leq \sup B$ (wegen $A \subseteq B$). Also ist $\sup B$ eine obere Schranke für A und daher $\sup A \leq \sup B$. Die zweite Ungleichung erhält man mit Hilfe von (a), indem man die linke Ungleichung auf $-A$ und $-B$ statt auf A und B anwendet.

(d) Nach (c) ist $\sup(A \cup B)$ eine obere Schranke für A und für B und daher $\sup(A \cup B) \geq \max\{\sup A, \sup B\}$. Ist $x \in A \cup B$ beliebig, so folgt $x \in A$ (und daher $a \leq \sup B$) oder $x \in B$ (und damit $x \leq \sup B$), so daß $\max\{\sup A, \sup B\}$ eine obere Schranke für $A \cup B$ ist. Daher gilt auch $\sup(A \cup B) \leq \max\{\sup A, \sup B\}$ und es folgt die linke Gleichung. Die rechte beweist man analog oder erhält sie mit Hilfe von (a) aus der linken (angewendet auf $-A$ statt A).

(e), (f): Für alle $a \in A, b \in B$ gilt $a \leq \sup A, b \leq \sup B$ und damit auch $a + b \leq \sup A + \sup B$. Sind A und B nach oben beschränkt, so ist also auch $A + B$ nach oben beschränkt und $\sup(A + B) \leq \sup A + \sup B$. Zu beliebigem $\varepsilon > 0$ gibt es in diesem Fall nach Lemma 1.16 Elemente $a \in A, b \in B$ mit $a > \sup A - \varepsilon/2, b > \sup B - \varepsilon/2$, also auch $a + b > \sup A + \sup B + \varepsilon$. Mit Lemma 1.16 folgt $\sup(A + B) = \sup A + \sup B$. Ist A oder B unbeschränkt, z.B. $\sup A = \infty$, so gibt es zu jedem $r \in \mathbb{R}$ und jedem $b \in B$ ein $a \in A$ mit $a > r - b$, also mit $a + b > r$. Daher ist auch $A + B$ unbeschränkt. Die übrigen Gleichungen aus (e) und (f) erhält man leicht unter Verwendung von (a) aus der bereits bewiesenen ersten Gleichung in (e).

(g) als Übung. □

4. Natürliche Zahlen und vollständige Induktion

Wir führen nun die Menge $\mathbb{N}$ der natürlichen Zahlen als Teilmenge von $\mathbb{R}$ ein.

1.18. Satz. *Es gibt genau eine Teilmenge $\mathbb{N}$ von $\mathbb{R}$ mit den folgenden drei Eigenschaften.*

(a) $1 \in \mathbb{N}$.
(b) $\forall n \in \mathbb{N}: \quad n + 1 \in \mathbb{N}$.
(c) *Für jede Teilmenge B von $\mathbb{N}$, die (a) und (b) erfüllt gilt schon $B = \mathbb{N}$.*

Beweis. Wir zeigen zunächst die Existenz einer Menge $\mathbb{N}$ mit den Eigenschaften (a), (b) und (c). Hierzu definieren wir

$$\mathfrak{M} := \{A \subseteq \mathbb{R}\, ;\, A \text{ erfüllt (a) und (b)}\}.$$

Wegen $\mathbb{R} \in \mathfrak{M}$ ist $\mathfrak{M} \neq \emptyset$. Wir setzen nun

$$\mathbb{N} := \bigcap_{A \in \mathfrak{M}} A$$

und zeigen, daß $\mathbb{N}$ die Bedingungen (a), (b) und (c) erfüllt. Wegen $1 \in A$ für alle $A \in \mathfrak{M}$ ist $1 \in \bigcap_{A \in \mathfrak{M}} A = \mathbb{N}$. Ist $n \in \mathbb{N}$ so folgt $n \in A$ und damit auch $n + 1 \in A$ für alle $A \in \mathfrak{M}$, da diese Mengen (b) erfüllen. Also ist auch $n + 1 \in \bigcap_{A \in \mathfrak{M}} A = \mathbb{N}$. Ist schließlich B eine Teilmenge von $\mathbb{N}$, welche (a) und (b) erfüllt, so folgt $B \in \mathfrak{M}$ und damit auch $\mathbb{N} = \bigcap_{A \in \mathfrak{M}} A \subseteq B$. Damit ist gezeigt, daß $\mathbb{N}$ die Forderungen (a)–(c) erfüllt.

Ist $\mathbb{N}' \subseteq \mathbb{R}$ eine weitere Menge, die den Bedingungen (a)–(c) genügt, so folgt insbesondere $\mathbb{N}' \in \mathfrak{M}$ und daher (da $\mathbb{N}$ (c) erfüllt) auch $\mathbb{N} \subseteq \mathbb{N}'$. Da $\mathbb{N}$ (a) und (b) erfüllt und (c) für $\mathbb{N}'$ gilt, erhalten wir hieraus $\mathbb{N}' \subseteq \mathbb{N}$ und somit insgesamt $\mathbb{N}' = \mathbb{N}$. □

1.19. Satz (Definition durch vollständige Induktion). *Für alle $n \in \mathbb{N}$ seien Begriffe $B(n)$ gegeben, so daß gilt:*

(a) *$B(1)$ ist definiert, d.h. wir wissen, was $B(1)$ bedeutet.*
(b) *Ist $n \in \mathbb{N}$ so, daß $B(n)$ definiert ist, so ist auch $B(n + 1)$ definiert.*

Dann ist $B(n)$ für alle $n \in \mathbb{N}$ definiert.

Beweis. Es ist $A := \{n \in \mathbb{N}\, ;\, B(n) \text{ ist definiert}\}$ eine Teilmenge von $\mathbb{N}$, die den Bedingungen (a) und (b) aus Satz 1.18 genügt. Da $\mathbb{N}$ die Bedingung 1.18 (c) erfüllt, folgt $A = \mathbb{N}$. □

1.20. BEISPIELE (für Definitionen durch vollständige Induktion). (a) Wir wollen für alle $n \in \mathbb{N}$ den Begriff $B(n) = n!$ (lies n *Fakultät*) definieren. Wir setzen hierzu $1! := 1$. Ist $n \in \mathbb{N}$, so daß $n!$ definiert ist, so setzen wir $(n+1)! := n! \cdot (n+1)$. Gemäß Satz 1.19 ist $n!$ damit für alle $n \in \mathbb{N}$ definiert. Man definiert noch $0! := 1$.

(b) *Potenzen*: Sei $a \in \mathbb{R}$. Wir wollen a^n für alle $n \in \mathbb{N}$ definieren. Wir definieren zunächst $a^1 := a$. Ist nun $n \in \mathbb{N}$ so, daß a^n definiert ist, so setzen wir $a^{n+1} := a^n \cdot a$. Man definiert noch $a^0 := 1$ für alle $a \in \mathbb{R}_* = \{x \in \mathbb{R}\,;\, x \neq 0\}$ (Manchmal vereinbart man dies auch im Fall $a = 0$).

(c) *Summenzeichen*: $B(n) := \sum_{k=1}^n a_k$ soll für alle $a_1, \ldots, a_n \in \mathbb{R}$ definiert werden. Definition von $B(1)$: Für alle $a_1 \in \mathbb{R}$ setzen wir $\sum_{k=1}^1 a_k := a_1$. Sei nun $n \in \mathbb{N}$, so daß $B(n)$ definiert ist, d.h. daß für alle $a_1, \ldots, a_n \in \mathbb{R}$ der Ausdruck $\sum_{k=1}^n a_k$ definiert ist. Sind nun Elemente $a_k \in \mathbb{R}$ für $1 \leq k \leq n+1$ gegeben, so definieren wir

$$\sum_{k=1}^{n+1} a_k := \left(\sum_{k=1}^{n} a_k\right) + a_{n+1}.$$

Damit ist $B(n+1)$ definiert. Nach Satz 1.19 ist $B(n)$ nun für alle $n \in \mathbb{N}$ definiert. Für alle $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ und alle $a_0, \ldots, a_n \in \mathbb{R}$ definieren wir noch

$$\sum_{k=0}^{n} a_k := \begin{cases} a_0 & \text{falls } n = 0 \\ a_0 + \displaystyle\sum_{k=1}^{n} a_k & \text{für alle } n \in \mathbb{N}\,. \end{cases}$$

(d) Analog können wir das *Produktzeichen* einführen. Für $n = 1$ und alle $a_1 \in \mathbb{R}$ setzen wir $\prod_{k=1}^1 a_k := a_1$. Ist $n \in \mathbb{N}$, so daß $\prod_{k=1}^n a_k$ für alle $a_k \in \mathbb{R}$, $1 \leq k \leq n$, definiert ist und sind beliebige $a_k \in \mathbb{R}$ für $1 \leq k \leq n+1$ gegeben, so definieren wir

$$\prod_{k=1}^{n+1} a_k := \left(\prod_{k=1}^{n} a_k\right) \cdot a_{n+1}.$$

Nach Satz 1.19 ist $\prod_{k=1}^n a_k$ somit für alle $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{R}$, definiert. Für alle $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ und alle $a_0, \ldots, a_n \in \mathbb{R}$ definieren wir noch

$$\prod_{k=0}^{n} a_k := \begin{cases} a_0 & \text{falls } n = 0 \\ a_0 \cdot \displaystyle\prod_{k=1}^{n} a_k & \text{für alle } n \in \mathbb{N}\,. \end{cases}$$

1.21. SATZ (Beweis durch vollständige Induktion). *Für alle $n \in \mathbb{N}$ sei eine logische Aussage $A(n)$ gegeben, die wahr oder falsch sein kann. Es gelte:*

(a) *Induktionsanfang: $A(1)$ ist wahr.*
(b) *Induktionsschluß: Für alle $n \in \mathbb{N}$ gilt $A(n) \implies A(n+1)$.*

Dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr.

BEWEIS. Es ist $M := \{n \in \mathbb{N}\,;\, A(n) \text{ ist wahr}\}$ eine Teilmenge von $\mathbb{N}$, die den Bedingungen (a) und (b) aus Satz 1.18 genügt. Da $\mathbb{N}$ die Bedingung 1.18 (c) erfüllt, folgt $M = \mathbb{N}$. □

1.22. BEISPIELE (für Beweise durch vollständige Induktion). (a) *Für alle $n \in \mathbb{N}$ gilt:*

$$A(n): \quad \sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,.$$

Beweis: *Induktionsanfang*: Die Aussage $A(1)$ ist offensichtlich wahr.

Induktionsschluß: Für alle $n \in \mathbb{N}$ gilt $A(n) \implies A(n+1)$: Sei also $n \in \mathbb{N}$, so daß $A(n)$ wahr ist (*Induktionsvoraussetzung*). Nach Definition des Summenzeichens und wegen der Induktionsvoraussetzung gilt dann:

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

Also gilt $A(n+1)$. Nach Satz 1.21 gilt $A(n)$ für alle $n \in \mathbb{N}$.

(b) *Für alle $n \in \mathbb{N}$ gilt:*

$$A(n): \quad \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

Beweis: *Induktionsanfang*: Die Aussage $A(1)$ ist offensichtlich wahr.

Induktionsschluß: Für alle $n \in \mathbb{N}$ gilt $A(n) \implies A(n+1)$: Sei also $n \in \mathbb{N}$, so daß $A(n)$ wahr ist. Nach Definition des Summenzeichens und wegen der Induktionsvoraussetzung gilt dann:

$$\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3 = \left(\frac{n(n+1)}{2}\right)^2 + (n+1)^3 = \frac{(n+1)^2}{4}\left(n^2+4n+4\right) = \\ = \left(\frac{(n+1)(n+2)}{2}\right)^2.$$

Also gilt $A(n+1)$. Nach Satz 1.21 gilt $A(n)$ für alle $n \in \mathbb{N}$.

(c) Endliche geometrische Reihe: *Sei $\mathbb{K}$ ein beliebiger Körper und $1 \neq q \in \mathbb{K}$. Dann gilt mit der Vereinbarung $q^0 := 1$ für alle $n \in \mathbb{N}$ die Aussage*

$$A(n): \quad \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Beweis: *Induktionsanfang*: Die Aussage $A(1)$ ist richtig, denn

$$\sum_{k=0}^{1} q^k = 1 + q = \frac{(1-q)(1+q)}{1-q} = \frac{1-q^{n+1}}{1-q}.$$

Induktionsschluß: Für alle $n \in \mathbb{N}$ gilt $A(n) \implies A(n+1)$: Sei also $n \in \mathbb{N}$, so daß $A(n)$ wahr ist. Nach Definition des Summenzeichens und wegen der Induktionsvoraussetzung gilt dann:

$$\sum_{k=0}^{n+1} q^k = \sum_{k=0}^{n} q^k + q^{n+1} = \frac{1-q^{n+1}}{1-q} + q^{n+1} = \frac{1-q^{n+2}}{1-q}.$$

Also folgt $A(n+1)$. Nach Satz 1.21 gilt $A(n)$ für alle $n \in \mathbb{N}$.

(d) Bernoullische Ungleichung[1]: *Sei p eine reelle Zahl mit $p \geq -1$. Dann gilt für alle $n \in \mathbb{N}$ die Ungleichung*:

$$A(n): \quad 1 + np \leq (1+p)^n.$$

Beweis: *Induktionsanfang*: Die Aussage $A(1)$ ist richtig, denn für $n = 1$ ist $1 + np = 1 + p = (1+p)^n$.

Induktionsschluß: Für alle $n \in \mathbb{N}$ gilt $A(n) \implies A(n+1)$: Sei also $n \in \mathbb{N}$, so daß $A(n)$ wahr ist. Wegen $p \geq -1$ ist $1 + p \geq 0$ und daher nach Induktionsvoraussetzung und Lemma 1.6:

$$1 + (n+1)p \leq 1 + (n+1)p + np^2 = (1+p)(1+np) \leq (1+p)(1+p)^n = (1+p)^{n+1}.$$

[1] Jacob Bernoulli (22.12.1654–16.8.1705)

(e) *Für alle $x, y \in \mathbb{R}$ mit $0 \leq x < y$ und für alle $n \in \mathbb{N}$ gilt:*

$$A(n): \quad 0 \leq x^n < y^n.$$

Beweis: *Induktionsanfang*: Die Aussage $A(1)$ gilt nach Voraussetzung.
Induktionsschluß: Ist $A(n)$ erfüllt für ein $n \in \mathbb{N}$, so folgt mit Hilfe von Lemma 1.6 auch die Gültigkeit von $A(n+1)$.

Auch die folgenden Aussagen zeigt man leicht durch vollständige Induktion.

1.23. Lemma. (a) *Für alle $n \in \mathbb{N}$ ist $n \geq 1$. Insbesondere ist $1 = \min \mathbb{N}$.*
(b) *Für alle $n, m \in \mathbb{N}$ ist $n + m, n \cdot m \in \mathbb{N}$.*
(c) *Für alle $u, v \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt: $(uv)^n = u^n v^n$.*
(d) *Für alle $u \in \mathbb{R}$ und alle $m, n \in \mathbb{N}_0$ gilt (mit der Vereinbarung $u^0 := 1$):*

$$u^{n+m} = u^n u^m.$$

(e) *Seien $a, b, c \in \mathbb{R}$ mit $0 < a \leq b$ und $0 < c \leq 1 \leq d$. Dann gilt für alle $n \in \mathbb{N}$: $0 < a^n \leq b^n$ und $0 < c^n \leq 1 \leq d^n$.*

Mit Hilfe der Bernoullischen Ungleichung zeigen wir nun:

1.24. Folgerung. *Für alle $n \in \mathbb{N}$ und alle $p \in \mathbb{R}$ mit $-1 < p < \dfrac{1}{n}$ gilt*

$$(1+p)^n \leq \frac{1}{1-np}.$$

Beweis. Aus der Voraussetzung und Lemma 1.23 (a) folgt $-1 \leq -\frac{1}{n} < -p < 1$. Also ist $1 - p > 0$ und $0 < |p| < 1$, somit auch $0 < 1 - p^2 = (1-p)(1+p) < 1$. Hieraus folgt $0 < 1 + p < (1-p)^{-1}$. Mit 1.22 (e) ergibt sich

$$0 < (1+p)^n < \frac{1}{(1-p)^n}. \tag{1.1}$$

Wegen $-1 < -p$ können wir die Bernoulli–Ungleichung auf $-p$ anwenden und erhalten:

$$1 - np \leq (1-p)^n.$$

Hieraus folgt mit (1.1)

$$(1+p)^n < \frac{1}{(1-p)^n} \leq \frac{1}{1-np}$$

und damit die Behauptung. □

1.25. Definition. (a) $\mathbb{Z} := \{x \in \mathbb{R}\,;\, x \in \mathbb{N}_0 \text{ oder } -x \in \mathbb{N}\}$ heißt die *Menge der ganzen Zahlen*.
(b) $\mathbb{Q} := \{\frac{p}{q}\,;\, p \in \mathbb{Z}, q \in \mathbb{N}\}$ heißt die *Menge der rationalen Zahlen*. Die Elemente aus $\mathbb{R} \setminus \mathbb{Q}$ heißen *irrationale Zahlen*.

Der Beweis des folgenden Satzes bleibt dem Leser als Übung überlassen.

1.26. Satz. (a) *$\mathbb{Z}$ ist bezüglich der Addition eine kommutative Gruppe, d.h. $(\mathbb{Z}, +)$ erfüllt die Forderungen* (KA1)–(KA4) *aus Definition 1.1.*
(b) *$\mathbb{Q}$ ist versehen mit den von $\mathbb{R}$ induzierten Operationen der Addition und der Multiplikation ein angeordneter Körper, wobei $P_{\mathbb{Q}} := \{x \in \mathbb{Q}\,;\, x > 0\}$ die Menge der positiven rationalen Zahlen ist.*

Häufig ist die folgende Variante zu Satz 1.21 nützlich:

1.27. Satz. *Sei $n_0 \in \mathbb{Z}$ und sei für alle $n \in \mathbb{Z}$ mit $n \geq n_0$ eine logische Aussage $A(n)$ gegeben, die wahr oder falsch sein kann. Es gelte:*

(a) *Induktionsanfang: $A(n_0)$ ist wahr.*
(b) *Induktionsschluß: Für alle $n \geq n_0$ aus $\mathbb{Z}$ gilt $A(n) \implies A(n+1)$.*

Dann ist $A(n)$ für alle $n \geq n_0$ aus $\mathbb{Z}$ wahr.

Beweis. Für alle $n \in \mathbb{N}$ sei $B(n) := A(n_0 - 1 + n)$. Wegen (a) ist $B(1)$ wahr und wegen (b) gilt die Implikation $B(n) \implies B(n+1)$ für alle $n \in \mathbb{N}$. Nach Satz 1.21 ist also $B(n)$ für alle $n \in \mathbb{N}$ wahr. Damit ist für alle $n \geq n_0$ die Aussage $A(n)$ wahr (wegen $A(n) = B(n+1-n_0)$). □

1.28. Beispiel. Die Aussage

$$A(n): \qquad 4n \leq 2^n$$

ist für alle $n \in \mathbb{Z} \setminus \{1,2,3\}$ wahr.

Beweis. Für $n \leq 0$ ist dies klar und für $n = 1,2,3$ ist $A(n)$ offensichtlich falsch. Wegen $4 \cdot 4 = 16 = 2^4$ ist $A(4)$ wahr. Sei nun $4 \leq n \in \mathbb{Z}$, so daß $A(n)$ wahr ist, d.h. $4n \leq 2^n$ gilt. Dann folgt

$$4(n+1) < 4(n+n) = 2 \cdot 4n \leq 2 \cdot 2^n = 2^{n+1}.$$

Mit Satz 1.27 folgt die Gültigkeit von $A(n)$ für alle $n \geq 4$. □

Für $p, q \in \mathbb{Z}$ und $a_k \in \mathbb{R}$ für $k \geq p$ und $k \leq q$ definieren wir:

$$\sum_{k=p}^{q} a_k := \begin{cases} 0 & \text{falls } p > q \\ \sum\limits_{k=1}^{q-p+1} a_{k+p-1} & \text{falls } p \leq q \end{cases}$$

und

$$\prod_{k=p}^{q} a_k := \begin{cases} 1 & \text{falls } p > q \\ \prod\limits_{k=1}^{q-p+1} a_{k+p-1} & \text{falls } p \leq q\,. \end{cases}$$

Die folgenden Rechenregeln für den Umgang mit endlichen Summen und Produkten kann man alle mit Hilfe von vollständiger Induktion beweisen.

1.29. Lemma. *Sei $(\mathbb{K}, +, \cdot)$ ein Körper und seien $p, q \in \mathbb{Z}$ mit $p \leq q$.*

(a) *Sind $a_k, b_k \in \mathbb{K}$ für alle $k \in \mathbb{Z}$ mit $p \leq k \leq q$, so gilt für alle $x, y \in \mathbb{K}$:*

$$\sum_{k=p}^{q} (x a_k + y b_k) = x \Big(\sum_{k=p}^{q} a_k \Big) + y \Big(\sum_{k=p}^{q} b_k \Big).$$

(b) *Sei $n \in \mathbb{N}$ und seien $a_k, b_k \in \mathbb{K}$ für alle $k \in \mathbb{Z}$ mit $p \leq k \leq q+n$. Dann gilt*

$$\sum_{k=p}^{q+n} a_k = \Big(\sum_{k=p}^{q} a_k \Big) + \Big(\sum_{k=q+1}^{q+n} a_k \Big) \quad \textit{und} \quad \prod_{k=p}^{q+n} a_k = \Big(\prod_{k=p}^{q} a_k \Big) \cdot \Big(\prod_{k=q+1}^{q+n} a_k \Big).$$

Dies ist eine Verallgemeinerung des Assoziativgesetzes. Für eine noch allgemeinere Fassung sei auf Lamprecht [**16**], *S. 24 oder auf* Grauert-Lieb [**5**], *S. 13 verwiesen.*

(c) *Allgemeines Kommutativgesetz und Indextransformationen: Seien auch $p', q' \in \mathbb{Z}$ mit $q' - p' = q - p$ und sei*

$$\varphi : \{k \in \mathbb{Z}\,;\, p \le k \le q\} \to \{j \in \mathbb{Z}\,;\, p' \le j \le q'\}$$

eine bijektive Abbildung. Dann gilt für alle $a_j \in \mathbb{K}$, $p' \le j \le q'$,

$$\sum_{j=p'}^{q'} a_j = \sum_{k=p}^{q} a_{\varphi(k)} \quad \textit{und} \quad \prod_{j=p'}^{q'} a_j = \prod_{k=p}^{q} a_{\varphi(k)}\,.$$

(d) *Für alle $x \in \mathbb{R}$ ist $\sum_{k=p}^{q} x = (q-p+1)x$.*

1.30. Definition (Binomialkoeffizienten). Für alle $x \in \mathbb{R}$ und alle $n \in \mathbb{N}_0$ sei

$$\binom{x}{n} := \frac{1}{n!}\prod_{k=1}^{n}(x+1-k) = \begin{cases} 1 & \text{falls } n = 0 \\ \dfrac{x(x-1)\ldots(x+1-n)}{n!} & \text{falls } n \in \mathbb{N}\,. \end{cases}$$

Ist speziell $x = m \in \mathbb{N}$, so hat man für $m \ge n$:

$$\binom{m}{n} := \frac{1}{n!}\prod_{k=1}^{n}(m+1-k) = \frac{m(m-1)\ldots(m+1-n)}{n!}\cdot\frac{(m-n)!}{(m-n)!} = \frac{m!}{n!(m-n)!}$$

und $\binom{m}{n} = 0$ für $m < n$ (da dann das Produkt in der Definition einen Faktor 0 enthält).

Offensichtlich gilt auch $\binom{n}{n} = 1 = \binom{n}{0}$.

1.31. Lemma. *Für alle $x \in \mathbb{R}, n \in \mathbb{N}$ gilt:*

$$\binom{x}{n-1} + \binom{x}{n} = \binom{x+1}{n}.$$

Beweis. Nach Definition gilt

$$\begin{aligned}\binom{x}{n-1} + \binom{x}{n} &= \frac{1}{(n-1)!}\prod_{k=1}^{n-1}(x+1-k) + \frac{1}{n!}\prod_{k=1}^{n}(x+1-k) \\ &= \frac{n + (x+1-n)}{n!}\prod_{k=1}^{n-1}(x+1-k) \\ &= \frac{x+1}{n!}\prod_{j=2}^{n}(x+2-j) = \frac{1}{n!}\prod_{j=1}^{n}(x+1+1-j) = \binom{x+1}{n},\end{aligned}$$

wobei wir beim dritten Gleichheitszeichen Lemma 1.29 (c) mit der durch $\varphi(j) := j-1$ gegebenen Indextransformation $\varphi : \{2, \ldots, n\} \to \{1, \ldots, n-1\}$ verwendet haben. □

Mit Hilfe dieses Lemmas und unter Verwendung von $\binom{n}{0} = \binom{n}{n} = 1$ lassen sich die Binomialkoeffizienten $\binom{m}{n}$ für $n, m \in \mathbb{N}_0$ rekursiv berechnen: In dem Schema des *Pascalschen*[2] *Dreiecks* ist jede Zahl im Innern des Dreiecks die Summe der beiden darüberstehenden Zahlen.

[2]Blaise Pascal (19.6.1623–19.8.1662).

$$\begin{array}{ccccccccccccccc}
&&&&&&&1&&&&&&& \\
&&&&&&1&&1&&&&&& \\
&&&&&1&&2&&1&&&&& \\
&&&&1&&3&&3&&1&&&& \\
&&&1&&4&&6&&4&&1&&& \\
&&1&&5&&10&&10&&5&&1&& \\
&1&&6&&15&&20&&15&&6&&1& \\
\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots&\dots
\end{array}$$

1.32. Satz (Binomialsatz). *Für alle $a, b \in \mathbb{R}$ gilt (mit $a^0 := b^0 := (a+b)^0 := 1$):*

$$A(n): \qquad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Beweis durch vollständige Induktion. Für $n = 0$ und $n = 1$ ist die Aussage offensichtlich richtig. Sei nun $n \in \mathbb{N}$, so daß $A(n)$ gilt. Dann folgt unter Verwendung der Induktionsvoraussetzung:

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

Wenden wir auf die erste der beiden Summen Lemma 1.29 (c) an mit der Indextransformation $\varphi : \{0, \dots, n-1\} \to \{1, \dots, n\}$, $k \mapsto \varphi(k) := k+1$ an, so folgt wegen

$$\binom{n}{n} = \binom{n+1}{n+1} = 1 = \binom{n}{0} = \binom{n+1}{0}$$

und unter Verwendung von Lemma 1.31:

$$\begin{aligned}
(a+b)^{n+1} &= \binom{n+1}{n+1} a^{n+1} b^0 + \sum_{k=1}^{n} \binom{n}{k-1} a^k b^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} + \binom{n+1}{0} a^0 b^{n+1} \\
&= \binom{n+1}{n+1} a^{n+1} b^0 + \sum_{k=1}^{n} \left(\binom{n}{k-1} + \binom{n}{k} \right) a^k b^{n+1-k} + \binom{n+1}{0} a^0 b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}.
\end{aligned}$$

Also gilt auch die Aussage $A(n+1)$. □

Für die spezielle Wahl $a = 1, b = 1$ bzw. $a = -1, b = 1$ erhalten wir:

1.33. Folgerung. *Für alle $n \in \mathbb{N}$ gilt:*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \qquad \textit{und} \qquad \sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0.$$

1.34. Folgerung. *Für alle* x, ε *mit* $x \geq 0$ *und* $0 \leq \varepsilon \leq 1$ *und alle* $n \in \mathbb{N}$ *gilt:*

(a) $(x+\varepsilon)^n \leq x^n + \varepsilon(x+1)^n$.

(b) $(x-\varepsilon)^n \geq x^n - \varepsilon(x+1)^n$.

Beweis. (a) Nach dem Binomialsatz 1.32 gilt

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

(b) Nach dem Binomialsatz 1.32 gilt

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

□

5. Das Archimedische Axiom

1.35. Satz. (a) $\mathbb{N}$ *ist nicht nach oben beschränkt.*

(b) *Jede nicht leere Teilmenge* A *von* $\mathbb{N}$ *besitzt ein Minimum.*

Beweis. (a) Wir gehen indirekt vor und nehmen an, daß $\mathbb{N}$ doch nach oben beschränkt ist. Nach dem Supremumsaxiom existiert dann eine kleinste obere Schranke b von $\mathbb{N}$. Für alle $n \in \mathbb{N}$ gilt also: $n \leq b$. Da mit n auch $n+1$ eine natürliche Zahl ist, folgt für alle $n \in \mathbb{N}$: Es ist $n+1 \leq b$ und damit auch $n \leq b-1$ im *Widerspruch* dazu, daß b die *kleinste* obere Schranke von $\mathbb{N}$ war. Also hat unsere Annahme zu einem Widerspruch geführt. $\mathbb{N}$ kann daher nicht nach oben beschränkt sein.

(b) Wegen $1 \leq n$ für alle $n \in \mathbb{N}$ ist die Aussage im Fall $1 \in A$ trivial. Sei nun $1 \notin A$. Wir definieren die Hilfsmenge

$$H := \{n \in \mathbb{N}\,;\, \{k \in \mathbb{N}\,;\, 1 \leq k \leq n\} \cap A = \emptyset\}\,.$$

Offensichtlich ist $1 \in H$. Wäre die Implikation $n \in H \implies n+1 \in H$ richtig, so wäre $H = \mathbb{N}$ und es würde folgen $A \cap \mathbb{N} = \emptyset$ im Widerspruch dazu, daß A nach Voraussetzung eine nicht leere Teilmenge von $\mathbb{N}$ ist. Also gibt es ein $n \in H$ mit $n+1 \notin H$, d.h. mit $\{k \in \mathbb{N}\,;\, 1 \leq k \leq n\} \cap A = \emptyset$ und $\{k \in \mathbb{N}\,;\, 1 \leq k \leq n+1\} \cap A \neq \emptyset$. Es ist also $n+1 \in A$ und $n+1$ ist das gesuchte Minimum von A. □

Als Folgerung erhalten wir

1.36. Folgerung (Archimedisches[3] Axiom). *Zu jedem* $\varepsilon > 0$ *gibt es ein* $n \in \mathbb{N}$ *mit* $\frac{1}{n} < \varepsilon$.

Beweis. Wäre die Aussage falsch, so gäbe es ein $\varepsilon > 0$ mit $\varepsilon \leq \frac{1}{n}$ also auch $n \leq \frac{1}{\varepsilon}$ für alle $n \in \mathbb{N}$ im Widerspruch zu Satz 1.35. □

[3] Archimedes von Syrakus (287 v.Chr.–212 v.Chr.)

1.37. Beispiele. (a) Jede nicht leere endliche Teilmenge A von $\mathbb{R}$ ist nach oben und nach unten beschränkt und besitzt ein Minimum und ein Maximum.

(b) Für die Menge $A := \{(-1)^n + \dfrac{1}{n}\,;\, n \in \mathbb{N}\}$ gilt:

(i) $\inf A = -1$ und A hat kein Minimum.

(ii) $\sup A = \max A = \dfrac{3}{2}$.

Beweis. (a) zeigt man leicht durch vollständige Induktion nach der Anzahl der Elemente von A.

Zu (b): (i) Für alle $n \in \mathbb{N}$ gilt

$$-1 \leq (-1)^n < (-1)^n + \frac{1}{n}\,.$$

Also ist -1 eine untere Schranke von A und $-1 \notin A$. Sei nun $\varepsilon > 0$ beliebig. Nach dem Archimedischen Axiom 1.36 gibt es ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$. Es gilt

$$a := -1 + \frac{1}{2n+1} = (-1)^{2n+1} + \frac{1}{2n+1} \in A$$

und $a < -1 + \frac{1}{n} < -1 + \varepsilon$. Nach dem Kriterium 1.16 folgt $-1 = \inf A$.

(ii) Wegen $\frac{3}{2} = (-1)^2 + \frac{1}{2} \in A$ genügt es zu zeigen, daß $\frac{3}{2}$ eine obere Schranke für A ist. Für ungerades $n = 2k+1$ mit $k \in \mathbb{N}_0$ gilt

$$(-1)^n + \frac{1}{n} = -1 + \frac{1}{2k+1} \leq 0 < \frac{3}{2}$$

und für gerade $n = 2k$, $k \in \mathbb{N}$, hat man

$$(-1)^n + \frac{1}{n} = 1 + \frac{1}{2k} \leq 1 + \frac{1}{2} = \frac{3}{2}\,.$$

Also ist $\frac{3}{2}$ obere Schranke von A. □

6. Potenzen mit rationalen Exponenten

Wir können nun die Existenz und Eindeutigkeit der n-ten nicht negativen Wurzel aus einer nicht negativen reellen Zahl beweisen.

1.38. Satz. *Zu $0 \leq a \in \mathbb{R}$ und $n \in \mathbb{N}$ gibt es genau ein $x \in \mathbb{R}$ mit $x \geq 0$ und $x^n = a$. Wir nennen dann x die n–te nicht negative Wurzel von a und schreiben $x = \sqrt[n]{a}$ oder $x = a^{1/n}$.*

Beweis. Wir definieren $A := \{u \in \mathbb{R}\,;\, u \geq 0 \text{ und } u^n \leq a\}$. Wegen $0 \in A$ ist $A \neq \emptyset$. Für alle $u > a+1$ gilt nach dem Binomialsatz 1.32 (unter Verwendung von 1.22 (e)):

$$u^n > (a+1)^n = \sum_{k=0}^{n} \binom{n}{k} a^k 1^{n-k} \geq 1 + na + \sum_{k=2}^{n} \binom{n}{k} a^k \geq 1 + na > a$$

und daher $u \notin A$. Also ist $a+1$ eine obere Schranke für A. Nach dem Supremumsaxiom existiert $x := \sup A$ in $\mathbb{R}$. Genau einer der drei folgenden Fälle muß zutreffen:

(i) $x^n < a$ (ii) $x^n > a$ (iii) $x^n = a\,.$

Annahme: Es ist $x^n < a$. Dann ist

$$\varepsilon := \min\left\{\frac{a - x^n}{(x+1)^n},\, 1\right\} > 0$$

und nach Folgerung 1.34 folgt

$$(x+\varepsilon)^n \leq x^n + \varepsilon(x+1)^n \leq x^n + (a - x^n) = a$$

und somit $x+\varepsilon \in A$ im Widerspruch dazu, daß x obere Schranke zu A ist. Also trifft (i) nicht zu.

Annahme: Es ist $x^n > a$. Dann ist

$$\varepsilon := \min\left\{\frac{x^n - a}{(x+1)^n}, 1\right\} > 0.$$

Nach Lemma 1.16 gibt es also ein $u \in A$ mit $u > x - \varepsilon$. Mit 1.22 (e) und Folgerung 1.34 folgt der Widerspruch

$$a \geq u^n > (x-\varepsilon)^n \geq x^n - \varepsilon(x+1)^n \geq x^n - (x^n - a) = a.$$

Also trifft auch (ii) nicht zu und es muß (iii) gelten.

Zur Eindeutigkeit: Ist auch $v \geq 0$ mit $v^n = a$, so ist $v \in A$ und daher $v \leq x$. Wäre $v < x$, so würde nach 1.22 (e) der Widerspruch $a = v^n < x^n = a$ folgen. Also muß $v = x$ gelten. □

Im Beweis des folgenden Lemmas verwenden wir

$$\mathbb{N} = (2\mathbb{N}) \cup (2\mathbb{N}-1), \qquad \text{und } (2\mathbb{N}) \cap (2\mathbb{N}-1) = \emptyset\,.$$

Hierbei folgt die erste Aussage aus Satz 1.18, da $(2\mathbb{N}) \cup (2\mathbb{N}-1)$ die Bedingungen (a) und (b) aus 1.18 erfüllt. Die zweite Aussage rechnet man leicht nach. Ist $p = 2k-1 \in 2\mathbb{N}-1$ (also ungerade) mit $k \in \mathbb{N}$, so ist $p^2 = 4k(k-1)+1$ ebenfalls ungerade.

1.39. Lemma. *$\sqrt{2} := 2^{1/2}$ ist eine irrationale Zahl.*

Beweis. Annahme: $\sqrt{2} := 2^{1/2}$ ist rational. Dann ist $\sqrt{2} = \frac{p}{q}$ mit $p, q \in \mathbb{N}$. Wegen Satz 1.35 können wir p minimal mit dieser Eigenschaft wählen. Es folgt durch Quadrieren: $2q^2 = p^2$. Nach unserer Vorbemerkung muß als $p = 2p_1$ mit $p_1 \in \mathbb{N}$ gelten. Es folgt $p_1 < p$ und $2q^2 = 4p_1^2$ also auch $q^2 = 2p_1^2$. Hieraus folgt wie oben $q = 2q_1$ mit $q_1 \in \mathbb{N}$. Also gilt

$$\sqrt{2} = \frac{p}{q} = \frac{p_1}{q_1}$$

im Widerspruch zur minimalen Wahl von p. □

Für $0 < a \in \mathbb{R}$ haben wir bisher den Ausdruck a^x definiert für alle $x \in \mathbb{N}_0 \cup \{-1\} \cup \{\frac{1}{n}\,;\, n \in \mathbb{N}\}$. Wir wollen nun a^x für alle $x \in \mathbb{Q}$ definieren und einige Rechenregeln für den Umgang mit Potenzen festhalten. Zunächst definieren wir für alle $n \in \mathbb{N}$: $a^{-n} := \left(\frac{1}{a}\right)^n$. Wegen $a^n \cdot \left(\frac{1}{a}\right)^n = \left(a \cdot \frac{1}{a}\right)^n = 1^n = 1$ folgt $a^{-n} = (a^n)^{-1}$.

1.40. Lemma. *Für alle $a, b \in \mathbb{R}_+^*$ gilt:*

(a) $\forall n \in \mathbb{Z}: \quad (ab)^n = a^n b^n$.
(b) $\forall n \in \mathbb{N}: \quad (ab)^{1/n} = a^{1/n} b^{1/n}$.
(c) $\forall n, m \in \mathbb{Z}: \quad a^{n+m} = a^n a^m$ *und* $(a^n)^m = a^{nm}$.
(d) $\forall p \in \mathbb{Z}, q \in \mathbb{N}: \quad (a^p)^{1/q} = (a^{1/q})^p$. *Wir definieren $a^{p/q} := (a^p)^{1/q}$. Dieser Ausdruck ist wohldefiniert, d.h. sind auch $p' \in \mathbb{Z}, q' \in \mathbb{N}$ mit $p'/q' = p/q$, so ist $(a^{1/q})^p = (a^{1/q'})^{p'}$.*
(e) $\forall x \in \mathbb{Q}: \quad (ab)^x = a^x b^x$.
(f) $\forall x, y \in \mathbb{Q}: \quad a^{x+y} = a^x a^y,\ a^{xy} = (a^x)^y$.
(g) *Ist $0 < a < b$ und $0 < x \in \mathbb{Q}$, so gilt $0 < a^x < b^x$ und $0 < b^{-x} < a^{-x}$.*
(h) *Ist $0 < a < 1 < b$ und sind $x, y \in \mathbb{Q}$ mit $0 < x < y$, so gilt*

$$0 < a^y < a^x < 1 < b^x < b^y.$$

Beweis. Die Aussagen (a) und (c) zeigt man leicht durch vollständige Induktion.

Zu (b): Nach (a) und der Definition der nicht negativen n–ten Wurzel ist für alle $n \in \mathbb{N}$: $(a^{1/n}b^{1/n})^n = (a^{1/n})^n(b^{1/n})^n = ab$. Wegen der Eindeutigkeit der nicht negativen n–ten Wurzel folgt $a^{1/n}b^{1/n} = (ab)^{1/n}$.

Zu (d): Seien $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ beliebig. Unter Verwendung von (c) erhält man

$$a^p = ((a^{1/q})^q)^p = (a^{1/q})^{qp} = ((a^{1/q})^p)^q .$$

Wegen der Eindeutigkeit der q–ten nicht negativen Wurzel folgt $(a^p)^{1/q} = (a^{1/q})^p$.

Zur Wohldefiniertheit: Seien $p, p' \in \mathbb{Z}$, $q, q' \in \mathbb{N}$ gegeben mit $p/q = p'/q'$. Wegen $q'p = p'q$ folgt $(a^p)^{q'} = a^{q'p} = a^{p'q} = (a^{p'})^q$ und hieraus durch Übergang zur q–ten nicht negativen Wurzel mit den bereits bewiesenen Aussagen:

$$((a^p)^{1/q})^{q'} = ((a^p)^{q'})^{1/q} = a^{p'} .$$

Wegen der Eindeutigkeit der q'–ten nicht negativen Wurzel folgt $(a^p)^{1/q} = (a^{p'})^{1/q'}$.

(e) folgt leicht mit (a), (b) und (d).

(f) Seien $x, y \in \mathbb{Q}$ beliebig, $x = p/q, y = m/n$ mit $p, m \in \mathbb{Z}$ und $q, n \in \mathbb{N}$. Dann ist

$$x + y = \frac{pn + mq}{qn}$$

und wir erhalten mit (d), (c) und (b):

$$a^{x+y} = (a^{pn+mq})^{1/qn} = (a^{pn}a^{mq})^{1/qn} = (a^{pn})^{1/qn}(a^{mq})^{1/qn} = a^x a^y .$$

Auch die zweite Behauptung erhält man leicht mit Hilfe von (a), (b) und (d).

(g) und (h) als Übung. □

7. Reichhaltigkeit von $\mathbb{R}$

1.41. Lemma. *Zu jedem $x \in \mathbb{R}$ gibt es genau ein $\lfloor x \rfloor \in \mathbb{Z}$ (lies: x entier) mit*

$$\lfloor x \rfloor \le x < \lfloor x \rfloor + 1 .$$

Beweis. Ist $x \ge 0$, so ist $\{n \in \mathbb{N} ; n > x\}$ nach dem Archimedischen Axiom nicht leer und besitzt daher nach Satz 1.35 (b) ein Minimum $n_x \in \mathbb{N}$. Für $\lfloor x \rfloor := n_x - 1$ muß dann $\lfloor x \rfloor \le x < \lfloor x \rfloor + 1$ gelten. Ist $x < 0$, so leistet $\lfloor x \rfloor := -\min\{n \in \mathbb{N} ; n \ge -x\}$ das Gewünschte (wobei die Menge über die das Minimum gebildet wird ebenfalls nach dem Archimedischen Axiom nicht leer ist).

Zur Eindeutigkeit: Wir nehmen an, es gibt $n_j \in \mathbb{Z}$, $j = 1, 2$, mit $n_j \le x < n_j + 1$ mit $n_1 \ne n_2$. Dann muß $n_1 < n_2$ oder $n_2 < n_1$ gelten. O.B.d.A. sei $n_1 < n_2$. Dann ist $n_2 - n_1 \in \mathbb{N}$ und daher $n_2 - n_1 \ge 1$. Es folgt der Widerspruch: $n_1 \le x < n_1 + 1 \le n_2 < x$. Also gibt es nur eine ganze Zahl n mit $n \le x < n + 1$. □

1.42. Folgerung. *Zu jedem $x \in \mathbb{R}$ gibt es genau ein $\lceil x \rceil \in \mathbb{Z}$ mit $\lceil x \rceil - 1 < x \le \lceil x \rceil$.*

Beweis. Man überlegt sich, daß $-\lfloor -x \rfloor$ das Gewünschte leistet. Die Eindeutigkeit zeigt man wie im Beweis zu Lemma 1.41. □

1.43. Satz. *Sind $x, y \in \mathbb{R}$ mit $x < y$, so gibt es Zahlen $r \in \mathbb{Q}$ und $s \in \mathbb{R} \setminus \mathbb{Q}$ mit $x < r < y$ und $x < s < y$.*

Beweis. (a) Nach dem Archimedischen Axiom gibt es ein $m \in \mathbb{N}$ mit $1/m < y - x$. Es folgt $1 < my - mx$. Also ist $mx < \lfloor mx \rfloor + 1 \leq mx + 1 < my$. Dann gilt

$$r := \frac{\lfloor mx \rfloor + 1}{m} \in \mathbb{Q} \quad \text{mit } x < r < y\,.$$

(b) Nach (a) gibt es ein $r \in \mathbb{Q}$ mit $x < r < y$ und zu r und y ein $\tilde{r} \in \mathbb{Q}$ mit $r < \tilde{r} < y$. Mit

$$s := r + \frac{\sqrt{2}}{2}(\tilde{r} - r)$$

gilt dann $s \in \mathbb{R} \setminus \mathbb{Q}$ (sonst wäre ja $\sqrt{2} = 2(s - r)(\tilde{r} - r)^{-1} \in \mathbb{Q}$ im Widerspruch zu Lemma 1.39) und

$$x < r < s < r + (\tilde{r} - r) = \tilde{r} < y \quad \text{wegen } 0 < \frac{\sqrt{2}}{2} < 1.$$

$\square$

KAPITEL 2

$\mathbb{R}^N$ und der Körper $\mathbb{C}$ der komplexen Zahlen

1. Der Vektorraum $\mathbb{R}^N$

2.1. Definition. Sei $\mathbb{K}$ ein Körper und V eine nicht leere Menge, so daß gilt:

(VR1) Auf V ist eine *Addition* $+ : V \times V \to V$, $(x, y) \mapsto x+y$, gegeben, so daß $(V, +)$ eine kommutative Gruppe ist, d.h. den Bedingungen (KA1)–(KA4) aus Definition 1.1 genügt. Das neutrale Element der Addition bezeichnen wir mit $\vec{0}$ oder auch mit 0 und nennen es den *Nullvektor* von V. Für $v \in V$ sei $-v$ das zu v bezüglich der Addition inverse Element aus V.

(VR2) Auf V ist weiter eine *Multiplikation mit Skalaren* $\cdot : \mathbb{K} \times V \to V$, $(\lambda, v) \mapsto \lambda v$, gegeben, so daß die folgenden Bedingungen erfüllt sind:
- (a) $\forall v \in V : \quad 1 \cdot v = v$.
- (b) $\forall \alpha, \beta \in \mathbb{K}, v \in V : \quad (\alpha + \beta)v = \alpha v + \beta v$.
- (c) $\forall \alpha \in \mathbb{K}, u, v \in V : \quad \alpha(u + v) = \alpha u + \alpha v$.
- (d) $\forall \alpha, \beta \in \mathbb{K}, v \in V : \quad (\alpha\beta)v = \alpha(\beta v)$.

Dann heißt $(V, +, \cdot)$ ein $\mathbb{K}$–*Vektorraum*. Statt $(V, +, \cdot)$ schreiben wir auch kurz V, wenn die Addition und die Multiplikation vom Zusammenhang her klar sind. Die Elemente eines $\mathbb{K}$-Vektorraums nennt man *Vektoren*. $\mathbb{K}$ heißt der *Skalarbereich* oder der *Skalarkörper* zu V und die Elemente von $\mathbb{K}$ werden auch *Skalare* genannt.

2.2. Beispiele. (a) Sei $V := \{\vec{0}\}$ eine Menge mit genau einem Element $\vec{0}$, $\mathbb{K}$ ein beliebiger Körper. Wir definieren $\vec{0} + \vec{0} := \vec{0}$ und $\lambda\vec{0} := \vec{0}$ für alle $\lambda \in \mathbb{K}$. Dann ist $(V, +, \cdot)$ ein $\mathbb{K}$–Vektorraum.

(b) Jeder Körper $\mathbb{K}$ ist bezüglich der Körperaddition und der Körpermultiplikation auch ein $\mathbb{K}$–Vektorraum.

(c) $\mathbb{R}$ ist mit der von $\mathbb{R}$ hergegebenen Addition und Multiplikation ein $\mathbb{Q}$–Vektorraum.

(d) Sei X eine nicht leere Menge und sei $\mathbb{K}$ ein beliebiger Körper. Wir betrachten die Menge $\mathcal{F}(X, \mathbb{K})$ aller Abbildungen von X nach $\mathbb{K}$. Für $f, g \in \mathcal{F}(X, \mathbb{K})$ und $\lambda \in \mathbb{K}$ definieren wir die Abbildungen $f + g : X \to \mathbb{K}$ und $\lambda \cdot f : X \to \mathbb{K}$ durch $(f + g)(x) := f(x)+g(x)$ und $(\lambda \cdot f)(x) := \lambda f(x)$ für alle $x \in X$ (*punktweise Addition und Multiplikation mit Skalaren*). Dann ist $(\mathcal{F}(X, \mathbb{K}), +, \cdot)$ ein $\mathbb{K}$–Vektorraum. Statt $\lambda \cdot f$ schreiben wir auch λf.

Seien $N \in \mathbb{N}$ und Mengen $X_1, \ldots, X_N$ gegeben. Wir hatten in Definition 0.12 den Begriff eines geordneten Paares eingeführt. Analog kann man für $x_1 \in X_1, \ldots, x_N \in X_N$ induktiv den Begriff eines *geordneten N–Tupels* $(x_1, \ldots, x_N)$ einführen. Statt $(x_1, \ldots, x_N)$ schreiben wir auch $(x_j)_{j=1}^N$. Zwei geordnete N–Tupel $(x_1, \ldots, x_N)$ und $(y_1, \ldots, y_N)$ mit $x_1, y_1 \in X_1, \ldots, x_N, y_N \in X_N$ sind gleich genau dann, wenn gilt: $x_j = y_j$ für alle $j \in \{1, \ldots, N\}$. Die Menge

$$\prod_{j=1}^{N} X_j := X_1 \times \cdots \times X_N := \{(x_1, \ldots, x_N)\,;\, x_j \in X_j \text{ für } 1 \leq j \leq N\}$$

heißt die *Produktmenge der Mengen* $X_1, \ldots, X_N$. Ist $X_1 = X_2 = \cdots = X_N = X$, so schreiben wir auch X^N statt $\prod_{j=1}^N X$. Insbesondere ist also $\mathbb{R}^N$ die Menge aller geordneten N–Tupel $(x_1, \ldots, x_N)$ mit Einträgen $x_1, \ldots, x_N \in \mathbb{R}$.

2.3. Lemma. *Gegeben seien N Vektorräume $V_1, \ldots, V_N$ über dem Grundkörper $\mathbb{K}$. Dann wird $V := V_1 \times \cdots \times V_N$ ein $\mathbb{K}$–Vektorraum, wenn wir V mit den durch*

$$x + y := (x_1 + y_1, \ldots, x_N + y_N), \quad \lambda x := (\lambda x_1, \ldots, \lambda x_N)$$

für alle $x = (x_1, \ldots, x_N), y = (y_1, \ldots, y_N) \in V, \lambda \in \mathbb{K}$, komponentenweise definierten Operationen der Addition und der Multiplikation mit Skalaren versehen, wieder ein $\mathbb{K}$–Vektorraum.

Zum Beweis der Aussagen in 2.2 und Lemma 2.3 rechnet man einfach nach, daß die Vektorraumaxiome (VR1) und (VR2) jeweils erfüllt sind.

Im folgenden Lemma fassen wir einige Rechenregeln in $\mathbb{K}$–Vektorräumen zusammen.

2.4. Lemma. *Sei $(V, +, \cdot)$ ein $\mathbb{K}$–Vektorraum. Für alle $u, v, w \in V$ und alle $\lambda \in \mathbb{K}$ gilt:*

(a) $u + v = w + v \Longrightarrow u = w$.
(b) $0 \cdot u = \vec{0}$ *und* $\lambda \cdot \vec{0} = \vec{0}$.
(c) $(-\lambda)u = \lambda(-u) = -(\lambda u)$.
(d) $-u = (-1)u$.
(e) $\lambda u = \vec{0} \iff (\lambda = 0$ *oder* $u = \vec{0})$.

Beweis. (a) ist die Kürzungsregel in der kommutativen Gruppe $(V, +)$.

(b) Wegen

$$0 \cdot u + 0 \cdot u = (0 + 0)u = 0 \cdot u = 0 \cdot u + \vec{0}\,,$$
$$\lambda \cdot \vec{0} + \lambda \cdot \vec{0} = \lambda(\vec{0} + \vec{0}) = \lambda \cdot \vec{0} = \lambda \cdot \vec{0} + \vec{0}$$

folgen mit (a) die Behauptungen.

(c) Es ist $\lambda u + (-\lambda)u = (\lambda + (-\lambda))u = 0 \cdot u = \vec{0}$ (nach (b)) und daher $(-\lambda)u = -(\lambda u)$. Ebenso: $\lambda u + \lambda(-u) = \lambda(u + (-u)) = \lambda \cdot \vec{0} = \vec{0}$ und daher $\lambda(-u) = -(\lambda u)$.

(d) folgt aus (c) und (a) in (VR2).

(e) Nach (b) ist $0 \cdot u = \vec{0}$. Ist $\lambda u = \vec{0}$ aber $\lambda \neq 0$, so existiert λ^{-1} in $\mathbb{K}$ und nach (b) folgt: $\vec{0} = \lambda^{-1}\vec{0} = \lambda^{-1}(\lambda u) = (\lambda^{-1} \cdot \lambda)u = 1 \cdot u = u$. □

2.5. Definition. Sei $N \in \mathbb{N}$. Für $u = (u_1, \ldots, u_N), v = (v_1, \ldots, v_N) \in \mathbb{R}^N$ definieren wir

$$\langle u, v \rangle := \sum_{j=1}^N u_j v_j\,.$$

Die hierdurch definierte Abbildung $\langle \cdot, \cdot \rangle : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ heißt das *Standardskalarprodukt* in $\mathbb{R}^N$.

2.6. Lemma (Rechenregeln für das Standardskalarprodukt). *Für alle $\lambda \in \mathbb{R}$ und alle $u = (u_1, \ldots, u_N), v = (v_1, \ldots, v_N), w = (w_1, \ldots, w_N) \in \mathbb{R}^N$ gilt:*

(a) $\langle u, u \rangle \geq 0$ *und* $\langle u, u \rangle = 0$ *genau dann, wenn* $u = \vec{0}$.
(b) $\langle u, v \rangle = \langle v, u \rangle$.
(c) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.
(d) $\langle w, u + v \rangle = \langle w, u \rangle + \langle w, v \rangle$.
(e) $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle = \langle u, \lambda v \rangle$.
(f) $\langle \vec{0}, v \rangle = 0 = \langle v, \vec{0} \rangle$.

Dies rechnet man unter Verwendung der Definition des Standardskalarproduktes unmittelbar nach.

2.7. Definition. Wir definieren den *euklidischen*[1] *Betrag*

$$|\cdot| : \mathbb{R}^N \to \mathbb{R}_+ = \{x \in \mathbb{R}\,;\, x \geq 0\}$$

durch:

$$|u| := \sqrt{\langle u, u\rangle} = \Big(\sum_{j=1}^{N} u_j^2\Big)^{1/2}$$

für alle $u = (u_1, \dots, u_N) \in \mathbb{R}^N$.

Für $N = 1$ stimmt dies mit dem Absolutbetrag in $\mathbb{R}$ überein. In der Literatur wird auch $\|\cdot\|_2$ statt $|\cdot|$ verwendet.

2.8. Satz. (Eigenschaften des euklidischen Betrags).

(a) $\forall u, v \in \mathbb{R}^N: \quad |\langle u, v\rangle| \leq |u| \cdot |v|$ *(Cauchy–Schwarzsche Ungleichung*[2]*),.*
Es ist $|\langle u, v\rangle| = |u| \cdot |v|$ *genau dann, wenn* $v = \vec{0}$ *oder* $u = \lambda v$ *für ein* $\lambda \in \mathbb{R}$ *gilt.*

(b) *Für alle* $u, v \in \mathbb{R}^N, \lambda \in \mathbb{R}$ *gilt:*
  (i) $|u| \geq 0$ *und* $|u| = 0$ *genau dann, wenn* $u = \vec{0}$.
  (ii) $|\lambda u| = |\lambda| \cdot |u|$.
  (iii) $|u + v| \leq |u| + |v|$ *(Dreiecksungleichung).*
  (iv) $||u| - |v|| \leq |u - v|$.

(c) *Für alle* $u_1, \dots u_n \in \mathbb{R}^N$ *gilt* $\left|\sum_{k=1}^n u_k\right| \leq \sum_{k=1}^n |u_k|$.

Beweis. (a) Für $v = \vec{0}$ gilt nach Lemma 2.6 $\langle u, \vec{0}\rangle = 0 = |u| \cdot 0 = |u| \cdot |\vec{0}|$.

Sei nun $v \neq \vec{0}$. Dann ist nach Lemma 2.6 $\langle v, v\rangle > 0$ also auch $|v| = \sqrt{\langle v, v\rangle} > 0$. Zur Abkürzung setzen wir

$$\mu := \frac{\langle u, v\rangle}{\langle v, v\rangle}\,.$$

Dann folgt unter Verwendung von Lemma 2.6

$$0 \leq \langle u - \mu v, u - \mu v\rangle = \langle u, u\rangle - 2\mu\langle u, v\rangle + \mu^2\langle v, v\rangle = \langle u, u\rangle - \frac{\langle u, v\rangle^2}{\langle v, v\rangle}$$

und damit auch $0 \leq \langle u, u\rangle \cdot \langle v, v\rangle - \langle u, v\rangle^2$, also $\langle u, v\rangle^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle$. Durch Wurzelziehen erhält man $|\langle u, v\rangle| \leq |u| \cdot |v|$.

Ist nun $|\langle u, v\rangle| = |u| \cdot |v|$, so folgt im Fall $v \neq \vec{0}$:

$$0 \leq \langle u - \mu v, u - \mu v\rangle = \langle u, u\rangle - \frac{\langle u, v\rangle^2}{\langle v, v\rangle} = 0$$

und damit (nach Lemma 2.6(a)) $u - \mu v = \vec{0}$ also $u = \mu v$. Ist umgekehrt $u = \mu v$ für ein $\mu \in \mathbb{R}$, so gilt

$$|\langle u, v\rangle| = |\langle \mu v, v\rangle| = |\mu|\langle v, v\rangle = (\mu^2\langle v, v\rangle)^{1/2}|v| = (\langle \mu v, \mu v\rangle)^{1/2}|v| = |u| \cdot |v|.$$

(b) (i) folgt aus Lemma 2.6(a).

---

[1] Euklid von Alexandria ($\approx$ 325 v.Chr.–$\approx$ 265 v.Chr.).

[2] Für $N = 3$ findet man diese Ungleichung schon 1793 bei Joseph-Louis Lagrange (25.1.1736–10.4.1813). Für beliebiges $N \in \mathbb{N}$ findet man sie im Cours d'Analyse (1821) von Augustin Louis Cauchy (21.4.1789–23.5.1857). Eine Integralvariante dieser Ungleichung trat 1859 bei Viktor Yakovlevich Bunyakovsky (16.12.1804–12.12.1889) und später, 1885 in der Festschrift zum 70. Geburtstag von Weierstraß, bei Herrmann Amandus Schwarz (25.1.1843–30.11.1921) auf. Schwarz erkannte diese Ungleichung als ein wichtiges Handwerkszeug der Analysis.

(ii) $|\lambda u|^2 = \langle \lambda u, \lambda u\rangle = \lambda^2\langle u, u\rangle = |\lambda|^2|u|^2$. Durch Ziehen der Wurzel folgt die Behauptung.

(iii) Mit der Cauchy–Schwarzschen Ungleichung folgt unter Verwendung der Rechenregeln aus Lemma 2.6:

$$|u+v|^2 = \langle u+v, u+v\rangle = |u|^2 + 2\langle u, v\rangle + |v|^2 \leq |u|^2 + 2|u|\cdot|v| + |v|^2 = (|u|+|v|)^2.$$

Durch Ziehen der Wurzel folgt die Behauptung.

(iv) Mit der Dreiecksungleichung folgt $|u| = |u-v+v| \leq |u-v| + |v|$ und somit

$$|u| - |v| \leq |u-v|\,. \tag{2.1}$$

Ebenso: $|v| = |v-u+u| \leq |v-u| + |u| = |u-v| + |u|$ und daher $|u| - |v| \geq -|u-v|$. Zusammen mit (2.1) ergibt dies die Behauptung.

(v) folgt durch vollständige Induktion aus der Dreiecksungleichung. □

2. Der Körper $\mathbb{C}$ der komplexen Zahlen

Die quadratische Gleichung $x^2+1=0$ hat keine Lösung in $\mathbb{R}$, da Quadrate von reellen Zahlen immer nicht negativ sind. Es gibt aber einen $\mathbb{R}$ enthaltenden Körper $\mathbb{C}$, in dem diese Gleichung eine Lösung besitzt.

2.9. Satz (Der Körper $\mathbb{C}$ der komplexen Zahlen). *Wir definieren auf* $\mathbb{C} := \mathbb{R}^2$ *die Addition* $+ : \mathbb{C}\times\mathbb{C}\to\mathbb{C}$ *komponentenweise durch*

$$\forall (u,v),(x,y)\in\mathbb{C}: \qquad (u,v)+(x,y) := (u+x, v+y)$$

und eine Multiplikation $\cdot : \mathbb{C}\times\mathbb{C}\to\mathbb{C}$ *durch*

$$\forall (u,v),(x,y)\in\mathbb{C}: \qquad (u,v)\cdot(x,y) := (ux-vy, uy+vx)\,.$$

Dann ist $(\mathbb{C},+,\cdot)$ *ein Körper mit dem Einselement* $1_{\mathbb{C}} = (1,0)$ *und dem Nullelement* $(0,0)$.

Beweis. Da $\mathbb{R}^2$ mit den komponentenweisen Operationen der Addition und der Multiplikation mit Skalaren ein $\mathbb{R}$–Vektorraum ist, ist $(\mathbb{C},+)$ eine kommutative Gruppe und erfüllt die Körperbedingungen (KA1)–(KA4) aus Definition 1.1. Wir müssen also nur die Bedingungen (KM1)–(KM4) und das Distributivgesetz nachrechnen.

Zu (KM1): Für alle $z_1 = (x_1,y_1), z_2 = (x_2,y_2), z_3 = (x_3,y_3)\in\mathbb{C}$ gilt nach Definition der Multiplikation unter Verwendung des Assoziativgesetzes und der Distributivgesetze in $\mathbb{R}$:

$$\begin{aligned}(z_1z_2)z_3 =&(x_1x_2-y_1y_2, x_1y_2+y_1x_2)(x_3,y_3)\\ =&((x_1x_2-y_1y_2)x_3-(x_1y_2+y_1x_2)y_3, (x_1x_2-y_1y_2)y_3+(x_1y_2+y_1x_2)x_3)\\ =&(x_1(x_2x_3-y_2y_3)-y_1(y_2x_3+x_2y_3), x_1(x_2y_3+y_2x_3)+y_1(x_2x_3-y_2y_3))\\ =&(x_1,y_1)(x_2x_3-y_2y_3, y_2x_3+x_2y_3)\\ =&z_1(z_2z_3)\,.\end{aligned}$$

Zu (KM2): Für alle $z=(x,y)\in\mathbb{C}$ gilt nach Definition der Multiplikation in $\mathbb{C}$: $1_{\mathbb{C}}z = (1,0)(x,y) = (x,y)$.

Zu (KM3): Sei $(0,0)\neq z = (x,y)\in\mathbb{C}$ beliebig. Dann ist $x^2+y^2>0$ und mit

$$w := \Big(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\Big)$$

folgt

$$zw = (x,y)\Big(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\Big) = \Big(\frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2}, \frac{-xy}{x^2+y^2} + \frac{yx}{x^2+y^2}\Big) = (1,0).$$

Zu (KM4): Für alle $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{C}$ gilt nach Definition der Multiplikation unter Verwendung des Kommutativgesetzes in $\mathbb{R}$:

$$z_1 z_2 = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) = (x_2x_1 - y_2y_1, x_2y_1 + y_2x_1) = z_2z_1\,.$$

Zu (KD): Für alle $z_1 = (x_1, y_1), z_2 = (x_2, y_2), z_3 = (x_3, y_3) \in \mathbb{C}$ gilt unter Verwendung des Distributivgesetzes in $\mathbb{R}$:

$$\begin{aligned}(z_1 + z_2)z_3 =&(x_1 + x_2, y_1 + y_2)(x_3, y_3)\\ =&((x_1 + x_2)x_3 - (y_1 + y_2)y_3, (x_1 + x_2)y_3 + (y_1 + y_2)x_3)\\ =&(x_1x_3 + x_2x_3 - y_1y_3 - y_2y_3, x_1y_3 + x_2y_3 + y_1x_3 + y_2x_3)\\ =&(x_1x_3 - y_1y_3, x_1y_3 + y_1x_3) + (x_2x_3 - y_2y_3, x_2y_3 + y_2x_3)\\ =&z_1z_3 + z_2z_3\,.\end{aligned}$$

□

Wir definieren eine Abbildung $\phi : \mathbb{R} \to \mathbb{C}$ durch $\phi(x) := (x, 0)$ für alle $x \in \mathbb{R}$. Wegen $\phi(x) = \phi(y) \iff (x, 0) = (y, 0) \iff x = y$ ist ϕ injektiv und somit bijektiv von $\mathbb{R}$ nach $\phi(\mathbb{R}) = \{(x, 0)\,;\, x \in \mathbb{R}\}$. Für alle $x, y \in \mathbb{R}$ gilt weiter:

$$\begin{aligned}\phi(x + y) =&(x + y, 0) = (x, 0) + (y, 0) = \phi(x) + \phi(y),\\ \phi(xy) =&(xy, 0) = (x, 0)(y, 0) = \phi(x)\phi(y),\\ |\phi(x)| = |(x, 0)| = |x|.\end{aligned}$$

Ferner gilt $\phi(1) = 1_{\mathbb{C}}$ und $\phi(0) = (0, 0) = 0_{\mathbb{C}}$. ϕ ist also eine bijektive betrags- und strukturerhaltende Abbildung von $\mathbb{R}$ auf $\phi(\mathbb{R})$. Wir identifizieren daher im folgenden $\mathbb{R}$ mit der *reellen Achse* $\phi(\mathbb{R})$ in der komplexen Ebene und lassen den Index $\mathbb{C}$ beim Eins– und beim Nullelement weg. Wir schreiben also statt $(x, 0)$ wieder x. Ferner definieren wir $i := (0, 1)$. i heißt die *imaginäre Einheit von* $\mathbb{C}$ und $i\mathbb{R} := \{xi\,;\, x \in \mathbb{R}\} = \{(0, x)\,;\, x \in \mathbb{R}\}$ die *imaginäre Achse*. Mit diesen Vereinbarungen gilt:

2.10. Lemma. (a) $i^2 = -1$.

(b) *Jedes $z \in \mathbb{C}$ besitzt eine eindeutige Darstellung der Form $z = x + iy$ mit $x, y \in \mathbb{R}$.*

Beweis. (a) $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$.

(b) Ist $z = (x, y) \in \mathbb{C}$, so ist $z = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0) = x + iy$. Ist umgekehrt $z = x+iy$ mit $x, y \in \mathbb{R}$, so folgt $z = (x, 0)+(0, 1)(y, 0) = (x, 0)+(0, y) = (x, y)$. Insbesondere sind x und y durch z eindeutig bestimmt. □

Die zu $z \in \mathbb{C}$ gemäß Lemma 2.10 eindeutig bestimmten reellen Zahlen x, y mit $z = x + iy$ nennen wir *Realteil* und *Imaginärteil* von z und schreiben $\operatorname{Re} z = x, \operatorname{Im} z = y$. Ist $z = (x, y) = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ gegeben, so heißt $\bar{z} := (x, -y) = x - iy$ die zu z *konjugiert komplexe Zahl*.

Wir fassen einige einfache Rechenregeln in $\mathbb{C}$ in dem folgenden Lemma zusammen.

2.11. Lemma. *Für alle $z, w \in \mathbb{C}$ gilt:*

(a) $\operatorname{Re} z = \dfrac{1}{2}(z + \bar{z})$, $\operatorname{Im} z = \dfrac{1}{2i}(z - \bar{z})$.
(b) $\overline{z + w} = \bar{z} + \bar{w}$, $\operatorname{Re}(z + w) = \operatorname{Re} z + \operatorname{Re} w$, $\operatorname{Im}(z + w) = \operatorname{Im} z + \operatorname{Im} w$.
(c) $|z|^2 = z\bar{z}$ *und daher* $|z| = \sqrt{z\bar{z}}$.
(d) $|zw| = |z| \cdot |w|$ *und* $\overline{zw} = \bar{z}\bar{w}$.
(e) *Für $z \in \mathbb{C}$ sind die drei folgenden Aussagen äquivalent:*
 (i) $z \in \mathbb{R}$.
 (ii) $\operatorname{Im} z = 0$.

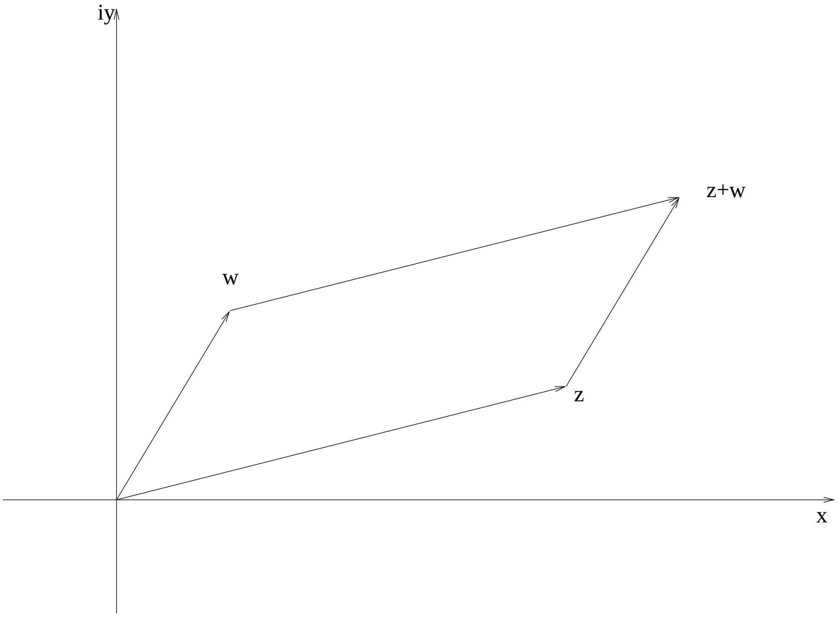


ABBILDUNG 1. Addition komplexer Zahlen

(iii) $z = \bar{z}$.

(f) *Für alle* $z \in \mathbb{C}_* = \mathbb{C} \setminus \{0\}$ *ist* $z^{-1} = \dfrac{\bar{z}}{|z|^2}$. *Insbesondere gilt für alle* $z \in \mathbb{C}$ *mit* $|z| = 1$*: Es ist* $z^{-1} = \bar{z}$.

BEWEIS. (a) und (b) rechnet man leicht nach.
(c) Für alle $z = x + iy = (x, y) \in \mathbb{C}$ (mit $x, y \in \mathbb{R}$) gilt:

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2.$$

(d) Für alle $z = x + iy, w = u + iv \in \mathbb{C}$ (mit $x, y, u, v \in \mathbb{R}$) gilt:

$$\begin{aligned}\overline{zw} =& \overline{(xu - yv) + i(xv + yu)} = (xu - yv) - i(xv + yu)\\ =&(xu - (-y)(-v)) + i(x(-v) + (-y)u) = \bar{z}\bar{w}.\end{aligned}$$

Wegen (c) folgt weiter:

$$|zw|^2 = (zw)(\overline{zw}) = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2.$$

Durch Wurzelziehen folgt die verbleibende Behauptung aus (d).
(e) ist klar und (f) folgt wegen $z \cdot \bar{z}|z|^{-2} = 1$ nach (c). □

2.12. GEOMETRISCHE INTERPRETATION. Wir veranschaulichen uns $\mathbb{C} = \mathbb{R}^2$ als Ebene mit einem festen rechtwinkligen Koordinatensystem. Hierbei ordnen wir $z = x + iy$ den Punkt mit den Koordinaten (x, y) zu. Man spricht bei dieser Interpretation von $\mathbb{C}$ auch von der *Gaußschen*[3] *Zahlenebene*. Der x–Achse entspricht hierbei die reelle Zahlengerade $\mathbb{R}$. Den rein imaginären Zahlen $i\mathbb{R} = \{iy\,; \, y \in \mathbb{R}\}$ wird die y–Achse zugeordnet. Mit dieser Interpretation von $\mathbb{C}$ als Gaußsche Zahlenebene gilt:

(a) Für $z \in \mathbb{C}$ ist $|z|$ der euklidische Abstand von z zum Ursprung.
(b) Für $z, w \in \mathbb{C}$ ist $z + w$ der Punkt, der durch Translation von z um w entsteht, d.h. der Ortsvektor von $z+w$ entsteht durch die Vektoraddition der Ortsvektoren von z und w. $z + w$ ist der dem Ursprung diagonal gegenüberliegende Eckpunkt

[3] JOHANN CARL FRIEDRICH GAUSS (30.4.1777–23.2.1855).

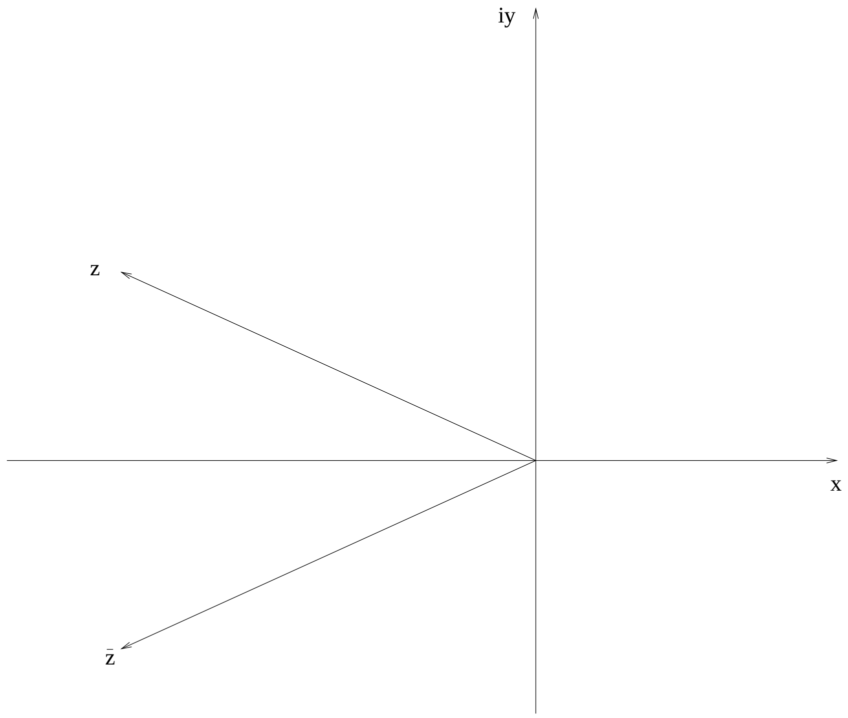


Abbildung 2. Konjugiertenbildung als Spiegelung an der reellen Achse

des von den Strecken von 0 nach z und 0 nach w aufgespannten Parallelogramms (vergl. Abb. 1).

(c) Der Abbildung $z \mapsto \bar{z}$ entspricht die Spiegelung an der reellen Achse.

(d) Zur geometrischen Veranschaulichung der Inversion $z \mapsto z^{-1}$ verwenden wir die *Spiegelung am Einheitskreis*. Sei zunächst allgemeiner K ein Kreis mit dem Mittelpunkt O und dem Radius r. Zwei Punkte P, Q in der Ebene heißen Spiegelpunkte bezüglich K falls gilt:
   (i) P und Q liegen auf der gleichen Halbgeraden durch O.
   (ii) $\overline{OQ} \cdot \overline{OP} = r^2$.

Die Fixpunkte der Spiegelung an K sind also die Punkte der Kreislinie. Liegt P außerhalb des Kreises K, so findet man den Spiegelpunkt Q, indem man die Tangenten von P an K legt. Der Schnittpunkt der Sehne durch die Berührungspunkte T_1, T_2 der Tangenten mit der Geraden durch O und P ist dann der gesuchte Spiegelpunkt. In der Tat gilt nach dem Höhensatz

$$\overline{QT_1}^2 = \overline{OQ} \cdot \overline{QP} = \overline{OQ} \cdot (\overline{OP} - \overline{OQ})$$

und nach dem Satz von Pythagoras

$$r^2 = \overline{OQ}^2 + \overline{QT_1}^2 = \overline{OQ}^2 + \overline{OQ} \cdot (\overline{OP} - \overline{OQ}) = \overline{OQ} \cdot \overline{OP}\,.$$

Ist umgekehrt Q ein von O verschiedener Punkt aus dem Inneren des Kreises, so erhält man den Spiegelpunkt, indem man in Q auf der Geraden durch O und Q die Senkrechte errichtet und in den Schnittpunkten T_1, T_2 dieser Senkrechten mit K die Tangenten an K legt. Diese Tangente schneiden sich in einem Punkt P außerhalb des Kreises, der der Spiegelpunkt zu Q ist.

Sei nun $K = \mathbb{T}$ die Einheitskreislinie in $\mathbb{C} = \mathbb{R}^2$. Wir berechnen zu $z \in \mathbb{C}_*$ den Spiegelpunkt w an $\mathbb{T}$. Wegen (i) muß $w = \alpha z$ für ein $\alpha > 0$ gelten und wegen (ii) ist $|z| \cdot |w| = 1$ also $\alpha = |z|^{-2}$ und damit $w = z|z|^{-2}$. Es folgt $z^{-1} = \bar{w}$.

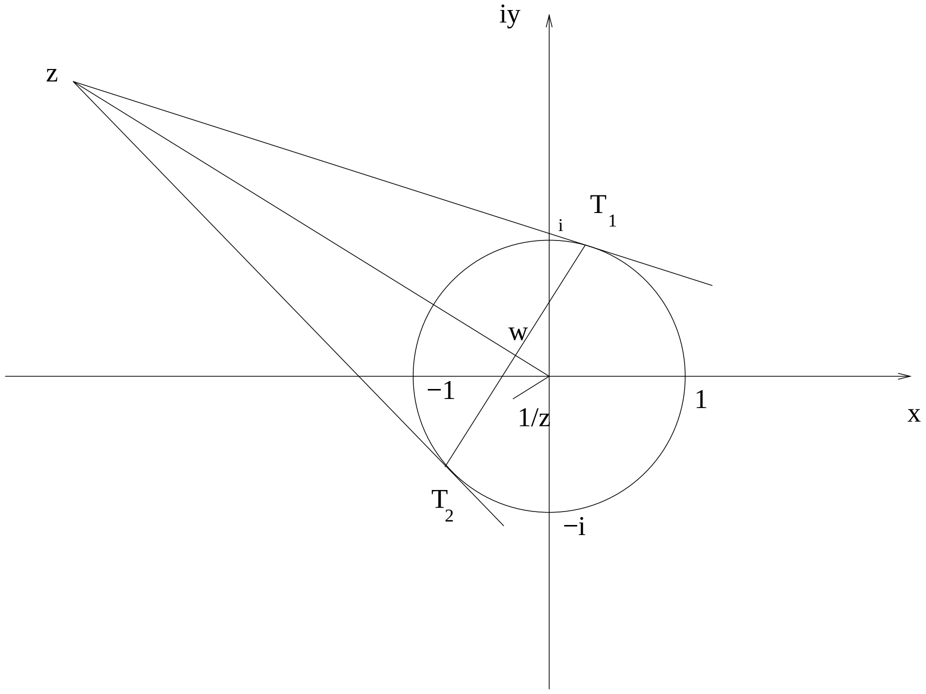


ABBILDUNG 3. Spiegelung am Einheitskreis und Inversenbildung

Man erhält also z^{-1} durch Spiegelung des Punktes z an der Einheitskreislinie und anschließende Spiegelung an der reellen Achse.

(e) Um auch die Multiplikation elementargeometrisch veranschaulichen zu können, führen wir Polarkoordinaten ein. Eine komplexe Zahl $0 \neq z \in \mathbb{C}$ ist eindeutig bestimmt durch ihren euklidischen Betrag $r = |z|$ und den Winkel φ zwischen der reellen Achse und dem Strahl von 0 durch z (im Bogenmaß gegen den Uhrzeigersinn gemessen). Sind r und φ gegeben, so ist $z = r(\cos\varphi + i\sin\varphi)$. Wir wollen dies hier elementargeometrisch verstehen. Später werden wir die trigonometrischen Funktionen auch analytisch einführen. Offenbar beschreiben (r, φ) und $(r, \varphi + 2k\pi)$ mit $k \in \mathbb{Z}$ die gleiche komplexe Zahl. φ ist jedoch eindeutig festgelegt, wenn wir zusätzlich $0 \leq \varphi < 2\pi$ fordern. Seien nun $z, w \in \mathbb{C}_*$ in der Form

$$z = r(\cos\varphi + i\sin\varphi), \qquad w = R(\cos\psi + i\sin\psi)$$

gegeben. Dann folgt unter Anwendung der Additionstheoreme für den Cosinus und den Sinus:

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

Die Abbildung $z \mapsto zw$ bedeutet also eine Streckung um den Faktor $R = |w|$ mit anschließender Drehung um den Winkel ψ (gegen den Uhrzeigersinn).

KAPITEL 3

Konvergenz von Folgen und Reihen

1. Der Grenzwertbegriff

Eine *Folge mit Werten in einer Menge X* ist eine Abbildung

$$a : \mathbb{Z}_m := \{n \in \mathbb{Z}\,;\, n \geq m\} \to X,$$

wobei m eine ganze Zahl sei. Statt $a(n)$ schreiben wir a_n. Jedem $n \geq m$ aus $\mathbb{Z}_m$ wird also eindeutig ein Wert $a_n \in X$ zugeordnet. Wir schreiben auch $a = (a_n)_{n=m}^{\infty} = (a_n)_{n\in\mathbb{Z}_m} = (a_m, a_{m+1}, a_{m+2}, \ldots)$. Ist speziell $X = \mathbb{R}$ (bzw. $X = \mathbb{C}$), so nennen wir a eine *Folge reeller* (bzw. *komplexer*) *Zahlen*. Meistens (aber nicht immer) wird in unseren Anwendungen $m = 0$ oder $m = 1$ sein.

3.1. Definition. Eine $\mathbb{R}^N$–wertige Folge $(a_n)_{n=m}^{\infty}$ heißt *konvergent*, falls es ein $u \in \mathbb{R}^N$ gibt, so daß zu jedem $\varepsilon > 0$ ein $n_0 \geq m$ aus $\mathbb{Z}$ existiert, derart daß für alle $n \geq n_0$ aus $\mathbb{Z}$ der Abstand von a_n zu u kleiner als ε ist. Kürzer: $(a_n)_{n=m}^{\infty}$ heißt konvergent, falls es ein $u \in \mathbb{R}^N$ gibt mit

$$\forall\varepsilon > 0\, \exists n_0 \in \mathbb{Z}_m\, \forall n \geq n_0, n \in \mathbb{Z} : \qquad |a_n - u| < \varepsilon\,. \tag{3.1}$$

Wir schreiben dann auch $a_n \to u$ für $n \to \infty$. Gibt es kein $u \in \mathbb{R}^N$, so daß (3.1) erfüllt ist, so nennen wir die Folge $(a_n)_{n=m}^{\infty}$ *divergent*.

Es gilt also $a_n \to u$ für $n \to \infty$ genau dann, wenn für jedes $\varepsilon > 0$ für *fast alle* $n \in \mathbb{Z}_m$ (d.h. alle bis auf endlich viele) die zugehörigen Folgenglieder a_n in der *ε–Umgebung* $U_\varepsilon(u) := \{x \in \mathbb{R}^N\,;\, |x - u| < \varepsilon\}$ von u liegen.

Für $N = 1$ haben wir insbesondere die Konvergenz in $\mathbb{R}$ und für $N = 2$ mit der Identifikation von $\mathbb{C}$ mit $\mathbb{R}^2$ die Konvergenz in $\mathbb{C}$ definiert.

3.2. Satz. *Sei $(a_n)_{n=m}^{\infty}$ eine $\mathbb{R}^N$–wertige Folge. Dann gibt es höchstens ein $u \in \mathbb{R}^N$ mit $a_n \to u$ für $n \to \infty$. Ist die Folge $(a_n)_{n=m}^{\infty}$ also konvergent, so ist u durch (3.1) eindeutig bestimmt. Wir nennen u in diesem Fall den Grenzwert oder Limes der Folge für $n \to \infty$ und schreiben $u = \lim\limits_{n\to\infty} a_n$.*

Beweis. Seien $u, v \in \mathbb{R}^N$ mit $a_n \to u$ und $a_n \to v$ für $n \to \infty$. Sei $\varepsilon > 0$ beliebig. Dann gibt es nach Definition 3.1 ein n_u und ein n_v in $\mathbb{Z}_m$, so daß gilt:

$$\forall n \geq n_u, n \in \mathbb{Z} : \quad |a_n - u| < \frac{\varepsilon}{2} \quad \text{und} \quad \forall n \geq n_v, n \in \mathbb{Z} : \quad |a_n - v| < \frac{\varepsilon}{2}\,.$$

Mit $n_0 := \max\{n_u, n_v\}$ folgt:

$$|u - v| = |u - a_{n_0} + a_{n_0} - v| \leq |u - a_{n_0}| + |a_{n_0} - v| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Für jedes $\varepsilon > 0$ ist also $|u - v| < \varepsilon$, d.h. es ist $|u - v| = 0$ und somit nach Satz 2.8 $u = v$. $\square$

3.3. Bemerkung. Durch richtiges Negieren von (3.1) erhält man: Eine $\mathbb{R}^N$–wertige Folge $(a_n)_{n=m}^\infty$ ist divergent genau dann, wenn für jedes $u \in \mathbb{R}^N$ gilt:

$$\exists \varepsilon > 0 \, \forall n_0 \in \mathbb{Z}_m \, \exists n \geq n_0, n \in \mathbb{Z} : \qquad |u - a_n| \geq \varepsilon \, . \tag{3.2}$$

3.4. Beispiele. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}(= \mathbb{R}^2)$.

(a) $\lim\limits_{n\to\infty} \dfrac{1}{n} = 0$ in $\mathbb{K}$.
(b) Für alle $q \in \mathbb{K}$ mit $|q| < 1$ ist $\lim\limits_{n\to\infty} q^n = 0$.
(c) Ist $q \in \mathbb{K}$ mit $|q| < 1$ und $s_n := \sum_{k=0}^n q^k$ für alle $n \in \mathbb{N}_0$, so gilt $\lim\limits_{n\to\infty} s_n = \dfrac{1}{1-q}$ in $\mathbb{K}$.
(d) Ist $a \in \mathbb{R}^N$ und $a_n = a$ für alle $n \in \mathbb{Z}_m$, so ist $\lim\limits_{n\to\infty} a_n = a$.
(e) Die Folge $(a_n)_{n=0}^\infty$ mit $a_n = (-1)^n$ für alle $n \in \mathbb{N}_0$ ist divergent.

Beweis. (a) Sei $\varepsilon > 0$ beliebig. Nach dem Archimedischen Axiom gibt es ein $n_0 \in \mathbb{N}_0$ mit $1/n_0 < \varepsilon$. Dann gilt für alle $n \geq n_0$ aus $\mathbb{N}_0$: $0 < 1/n \leq 1/n_0 < \varepsilon$ und daher

$$\forall n \geq n_0, n \in \mathbb{N}_0 : \qquad \Big|\frac{1}{n} - 0\Big| < \varepsilon \, .$$

Also folgt $\lim_{n\to\infty} 1/n = 0$.

(b) Im Fall $q = 0$ ist dies offensichtlich, denn für alle $n \geq 1$ aus $\mathbb{N}_0$ ist $0^n = 0$. Sei nun $q \neq 0$ vorausgesetzt. Sei $\varepsilon > 0$ beliebig. Nach der Bernoullischen Ungleichung (1.22 (d)) gilt (wegen $1/|q| > 1$) für alle $n \in \mathbb{N}_0$:

$$\Big(\frac{1}{|q|}\Big)^n = \Big(1 + \Big(\frac{1}{|q|} - 1\Big)\Big)^n \geq 1 + n\Big(\frac{1}{|q|} - 1\Big) > \frac{n(1-|q|)}{|q|}$$

und daher

$$|q|^n < \frac{|q|}{n(1-|q|)} \, .$$

Nach dem Archimedischen Axiom gibt es ein $n_0 \in \mathbb{N}_0$ mit $n_0^{-1} < \varepsilon(1 - |q|)|q|^{-1}$. Für alle $n \geq n_0$ aus $\mathbb{N}_0$ gilt dann:

$$|q^n - 0| = |q|^n < \frac{|q|}{n(1-|q|)} \leq \frac{|q|}{n_0(1-|q|)} < \varepsilon \, .$$

(c) Für alle $n \in \mathbb{N}_0$ gilt nach 1.22 (c)

$$\Big|s_n - \frac{1}{1-q}\Big| = \Big|\frac{1-q^{n+1}}{1-q} - \frac{1}{1-q}\Big| = \frac{|q|^{n+1}}{|1-q|} \leq \frac{|q|^{n+1}}{1-|q|} \, , \tag{3.3}$$

denn nach der Dreiecksungleichung ist $1 - |q| \leq |1 - q|$. Sei nun $\varepsilon > 0$ beliebig vorgegeben. Dann ist $\varepsilon' := \varepsilon(1 - |q|) > 0$. Nach (b) gibt es daher ein $n_0 \in \mathbb{N}_0$ mit $|q|^n < \varepsilon'$ für alle $n \geq n_0$. Mit (3.3) folgt also

$$\forall n \geq n_0, n \in \mathbb{N}_0 : \qquad \Big|s_n - \frac{1}{1-q}\Big| \leq \frac{|q|^{n+1}}{1-|q|} < \frac{\varepsilon'}{1-|q|} = \varepsilon \, .$$

Damit folgt die Behauptung.

(d) Sei $\varepsilon > 0$ beliebig. Dann folgt $\forall n \in \mathbb{Z}_m : \quad |a_n - a| = 0 < \varepsilon$. (3.1) ist also mit $n_0 = m$ erfüllt.

(e) Sei $u \in \mathbb{K}$ beliebig und $\varepsilon := 1$. Sei $n_0 \in \mathbb{N}_0$ beliebig.

1. Fall: $\operatorname{Re} u \geq 0$. Es ist $n := 2n_0 + 1 > n_0$ und

$$|u - (-1)^{2n_0+1}| = |u + 1| = \sqrt{(\operatorname{Re} u + 1)^2 + (\operatorname{Im} u)^2} \geq 1 \, .$$

2. Fall: $\operatorname{Re} u < 0$. Es ist $n := 2(n_0 + 1) > n_0$ und

$$|u - (-1)^{2(n_0+1)}| = |u - 1| = \sqrt{(\operatorname{Re} u - 1)^2 + (\operatorname{Im} u)^2} \geq 1\,.$$

In beiden Fällen ist also (3.2) erfüllt. □

3.5. Definition. Eine $\mathbb{R}^N$–wertige Folge $(a_n)_{n=m}^{\infty}$ heißt *beschränkt*, falls gilt:

$$\exists C > 0\, \forall n \in \mathbb{Z}_m : \qquad |a_n| \leq C\,.$$

3.6. Lemma. *Sei $(x_n)_{n=m}^{\infty}$ eine konvergente Folge in $\mathbb{R}^N$. Dann ist auch die reelle Folge $(|x_n|)_{n=m}^{\infty}$ konvergent und es gilt:*

$$\lim_{n\to\infty} |x_n| = |\lim_{n\to\infty} x_n|\,.$$

Beweis. Sei $u := \lim_{n\to\infty} x_n$ und sei $\varepsilon > 0$ beliebig. Dann gibt es eine ganze Zahl $n_0 \geq m$ mit $|x_n - u| < \varepsilon$ für alle $n \geq n_0$ aus $\mathbb{Z}$. Mit Satz 2.8 erhält man für alle $n \geq n_0$ aus $\mathbb{Z}$: $||x_n| - |u|| \leq |x_n - u| < \varepsilon$. Damit folgt die Behauptung. □

3.7. Lemma. *Jede konvergente Folge $(a_n)_{n=m}^{\infty}$ in $\mathbb{R}^N$ ist beschränkt.*

Beweis. Sei $u := \lim_{n\to\infty} a_n$ der Grenzwert der Folge $(a_n)_{n=m}^{\infty}$. Zu $\varepsilon := 1 > 0$ gibt es also ein $n_0 \in \mathbb{Z}_m$ mit

$$\forall n \geq n_0, n \in \mathbb{Z} : \qquad |a_n - u| < \varepsilon = 1.$$

Insbesondere folgt

$$\forall n \geq n_0, n \in \mathbb{Z} : \qquad |a_n| = |a_n - u + u| \leq |a_n - u| + |u| < |u| + 1.$$

Mit $C := \max\{|a_m|, \dots, |a_{n_0}|, |u| + 1\}$ folgt also $|a_n| \leq C$ für alle $n \in \mathbb{Z}_m$. □

3.8. Folgerung. *Die Folge $(n)_{n=0}^{\infty}$ ist nach Satz 1.35 unbeschränkt und daher nach Lemma 3.7 nicht konvergent.*

Bemerkung. Das Beispiel der Folge $((-1)^n)_{n=0}^{\infty}$ zeigt, daß es auch beschränkte divergente Folgen gibt.

3.9. Satz. *Die Folge $(h_n)_{n=1}^{\infty}$ mit*

$$h_n := \sum_{k=1}^{n} \frac{1}{k} \qquad \textit{für alle } n \in \mathbb{N}$$

ist unbeschränkt und daher divergent. Man nennt diese Folge auch die harmonische Reihe.

Beweis. Sei $C > 0$ beliebig. Zu zeigen ist: Es gibt ein $n \in \mathbb{N}$ mit $h_n > C$. Durch vollständige Induktion zeigt man, daß für alle $n \in \mathbb{N}$ gilt:

$$h_{2^n} = 1 + \sum_{k=1}^{n} \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{j}\,.$$

Für $2^{k-1} + 1 \leq j \leq 2^k$ gilt $\frac{1}{j} \geq \frac{1}{2^k}$. Also folgt:

$$h_{2^n} \geq 1 + \sum_{k=1}^{n} \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{2^k} = 1 + \sum_{k=1}^{n} \frac{2^{k-1}}{2^k} = 1 + \frac{n}{2}\,.$$

Da $\mathbb{N}$ nicht nach oben beschränkt ist gibt es ein $n_0 \in \mathbb{N}$ mit $n_0 > 2C$. Es folgt

$$h_{2^{n_0}} \geq 1 + \frac{n_0}{2} > 1 + C > C.$$

Also ist $(h_n)_{n=1}^{\infty}$ unbeschränkt[1]. □

2. Konvergenzkriterien und Grenzwertrechenregeln

3.10. Definition. Eine Folge $(a_n)_{n=m}^{\infty}$ reeller Zahlen heißt

(a) *streng monoton wachsend* (bzw. *monoton wachsend*), falls für alle $n_1, n_2 \in \mathbb{Z}_m$ mit $n_1 < n_2$ gilt $a_{n_1} < a_{n_2}$ (bzw. $a_{n_1} \leq a_{n_2}$).

(b) *streng monoton fallend* (bzw. *monoton fallend*), falls für alle $n_1, n_2 \in \mathbb{Z}_m$ mit $n_1 < n_2$ gilt $a_{n_1} > a_{n_2}$ (bzw. $a_{n_1} \geq a_{n_2}$).

Durch vollständige Induktion zeigt man leicht: Eine Folge $(a_n)_{n=m}^{\infty}$ reeller Zahlen ist genau dann streng monoton wachsend (bzw. monoton fallend), wenn für alle $n \in \mathbb{Z}_m$ gilt $a_n < a_{n+1}$ (bzw. $a_n \leq a_{n+1}$). Entsprechendes gilt für den monoton fallenden Fall.

3.11. Satz (Monotoniekriterium). *Jede monoton wachsende, nach oben beschränkte Folge $(a_n)_{n=m}^{\infty}$ reeller Zahlen konvergiert gegen ihr Supremum*

$$\sup_{n \in \mathbb{Z}_m} a_n := \sup\{a_n \,;\, n \in \mathbb{Z}_m\}\,.$$

Jede monoton fallende, nach unten beschränkte Folge $(a_n)_{n=m}^{\infty}$ reeller Zahlen konvergiert gegen ihr Infimum $\inf_{n \in \mathbb{Z}_m} a_n := \inf\{a_n \,;\, n \in \mathbb{Z}_m\}$.

Beweis. Wir führen den Beweis nur für den monoton wachsenden Fall. Der Beweis für monoton fallende Folgen verläuft analog. Sei also $(a_n)_{n=m}^{\infty}$ eine monoton wachsende, nach oben beschränkte Folge reeller Zahlen. Nach dem Supremumsaxiom existiert $y := \sup\{a_n \,;\, n \in \mathbb{Z}_m\}$ in $\mathbb{R}$. Sei nun $\varepsilon > 0$ beliebig. Nach Lemma 1.16 gibt es also ein $n_0 \in \mathbb{Z}_m$ mit $y - \varepsilon < a_{n_0} \leq y$. Für alle ganzen Zahlen $n \geq n_0$ gilt dann wegen der Monotonievoraussetzung:

$$y - \varepsilon < a_{n_0} \leq a_n \leq y$$

und damit $0 \leq y - a_n \leq y - a_{n_0} < \varepsilon$. Für alle $n \geq n_0$ ist also $|y - a_n| < \varepsilon$. Damit folgt $y = \lim_{n\to\infty} a_n$. □

3.12. Beispiel. *Für alle reellen $x > 0$ gilt*: $\lim\limits_{n\to\infty} x^{1/n} = 1$.

Beweis. Für $x = 1$ ist dies offensichtlich. Für $0 < x < 1$ ist die Folge $(x^{1/n})_{n=1}^{\infty}$ nach Lemma 1.40 (h) monoton wachsend und nach oben durch 1 beschränkt. Nach dem Monotoniekriterium 3.11 gilt also

$$\lim_{n\to\infty} x^{1/n} = y := \sup_{n \in \mathbb{N}} x^{1/n} \leq 1.$$

Annahme: $y < 1$. Wegen $0 < x = x^1 \leq y < 1$ folgt $y^{-1} > 1$. Nach der Bernoullischen Ungleichung folgt $y^{-n} \geq 1 + n(y^{-1} - 1)$. Die Folge y^{-n} ist also unbeschränkt. Es gibt daher ein $n_0 \in \mathbb{N}$ mit $y^{-n_0} > x^{-1}$ also auch $y^{n_0} < x$. Hieraus folgt aber $y < x^{1/n_0}$ im Widerspruch dazu, daß y obere Schranke für $\{x^{1/n} \,;\, n \in \mathbb{N}\}$ ist. Also war die Annahme falsch und es gilt $\lim_{n\to\infty} x^{1/n} = y = 1$.

Den verbleibenden Fall $x > 1$ behandelt man entweder analog oder man führt ihn vermöge Teil (e) der folgenden Rechenregeln auf den bereits bewiesenen Fall $0 < x < 1$ zurück. □

[1]Die Idee dieses Beweises findet sich schon bei Nicole Oresme (1320/25–11.7.1382).

3.13. Lemma (Grenzwertrechenregeln). *Es seien konvergente Folgen* $(a_n)_{n=p}^\infty$, $(b_n)_{n=q}^\infty$ *in* $\mathbb{R}^N$ *gegeben,* $a := \lim_{n\to\infty} a_n$, $b := \lim_{n\to\infty} b_n$, *sowie konvergente Folgen* $(\lambda_n)_{n=p}^\infty$, $(\mu_n)_{n=q}^\infty$ *in* $\mathbb{R}$, $(z_n)_{n=p}^\infty$, $(w_n)_{n=q}^\infty$ *in* $\mathbb{C}$ *mit*

$$\lambda := \lim_{n\to\infty} \lambda_n, \quad \mu := \lim_{n\to\infty} \mu_n, \quad z := \lim_{n\to\infty} z_n, \quad w := \lim_{n\to\infty} w_n\,.$$

Dann gilt mit $m := \max\{p, q\}$*:*

(a) *Die Folgen* $(a_n + b_n)_{n=m}^\infty$, $(\lambda_n + \mu_n)_{n=m}^\infty$ *und* $(z_n + w_n)_{n=m}^\infty$ *sind konvergent in* $\mathbb{R}^N$, *bzw.* $\mathbb{R}$, *bzw.* $\mathbb{C}$ *und es gilt:*

$$\lim_{n\to\infty}(a_n + b_n) = a + b, \quad \lim_{n\to\infty}(\lambda_n + \mu_n) = \lambda + \mu, \quad \lim_{n\to\infty}(z_n + w_n) = z + w\,.$$

(b) *Die Folgen* $(\mu_n a_n)_{n=m}^\infty$, $(\lambda_n\mu_n)_{n=m}^\infty$ *und* $(z_n w_n)_{n=m}^\infty$ *sind konvergent in* $\mathbb{R}^N$, *bzw.* $\mathbb{R}$, *bzw.* $\mathbb{C}$ *und es gilt:*

$$\lim_{n\to\infty}(\mu_n a_n) = \mu a, \quad \lim_{n\to\infty}(\lambda_n\mu_n) = \lambda\mu, \quad \lim_{n\to\infty}(z_n w_n) = zw\,.$$

(c) *Für alle* $\alpha \in \mathbb{R}$, $u \in \mathbb{R}^N$, $\zeta \in \mathbb{C}$ *sind die Folgen* $(\alpha a_n)_{n=p}^\infty$, $(\lambda_n u)_{n=p}^\infty$, $(\alpha\lambda_n)_{n=p}^\infty$ *und* $(\zeta z_n)_{n=p}^\infty$ *konvergent in* $\mathbb{R}^N$, *bzw.* $\mathbb{R}$, *bzw.* $\mathbb{C}$ *und es gilt:*

$$\lim_{n\to\infty}(\alpha a_n) = \alpha a, \quad \lim_{n\to\infty}(\lambda_n u) = \lambda u, \quad \lim_{n\to\infty}(\alpha\lambda_n) = \alpha\lambda, \quad \lim_{n\to\infty}(\zeta z_n) = \zeta z\,.$$

(d) *Gibt es ein* $n_0 \geq m$, *so daß für alle* $n \geq n_0$ *gilt* $\lambda_n \leq \mu_n$, *so ist* $\lambda \leq \mu$. *Ist speziell* $\lambda = \lim_{n\to\infty}\lambda_n = \lim_{n\to\infty}\mu_n = \mu$, *so gilt für jede Folge* $(x_n)_{n=m}^\infty$ *mit* $\lambda_n \leq x_n \leq \mu_n$ *für alle* $n \geq n_0$*: Es ist* $x_n \to \lambda$ *für* $n \to \infty$. (Einschachtelungsregel).

(e) *Ist* $z = \lim_{n\to\infty} z_n \neq 0$, *so gibt es ein* $n_1 \geq p$ *in* $\mathbb{Z}$ *mit* $z_n \neq 0$ *für alle* $n \geq n_1$ *und die Folge* $\left(\frac{w_n}{z_n}\right)_{n=m_0}^\infty$ *(mit* $m_0 := \max\{n_1, q\}$*) ist konvergent mit* $\lim_{n\to\infty}\frac{w_n}{z_n} = \frac{w}{z}$.

Beweis. (a) Wegen $\mathbb{R} = \mathbb{R}^1$ und $\mathbb{C} = \mathbb{R}^2$ genügt es, die erste Aussage zu beweisen. Sei also $\varepsilon > 0$ beliebig. Wegen $a_n \to a$, $b_n \to b$ für $n \to \infty$ gibt es ein $n_a \geq p$ und ein $n_b \geq q$ in $\mathbb{Z}$ mit

$$\forall n \geq n_a, n \in \mathbb{Z}: \qquad |a_n - a| < \frac{\varepsilon}{2} \quad \text{und} \quad \forall n \geq n_b, n \in \mathbb{Z}: \qquad |b_n - b| < \frac{\varepsilon}{2}\,.$$

Dann folgt für alle $n \geq n_0 := \max\{n_a, n_b\}$ aus $\mathbb{Z}$ mit der Dreiecksungleichung:

$$|a_n + b_n - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also gilt $a_n + b_n \to a + b$ für $n \to \infty$.

(b) Sei $\varepsilon > 0$ beliebig. Für alle $n \geq m = \max\{p, q\}$ schätzt man mit Hilfe der Dreiecksungleichung ab:

$$|\mu_n a_n - \mu a| \leq |\mu_n - \mu| \cdot |a_n| + |\mu| \cdot |a_n - a|\,.$$

Nach Lemma 3.7 ist die Folge $(a_n)_{n=p}^\infty$ beschränkt. Es gibt also ein $M > 0$ mit $|a_n| \leq M$ für alle $n \in \mathbb{Z}_p$. Wegen $\mu_n \to \mu$ gibt es zu $\varepsilon' := \frac{\varepsilon}{2M}$ ein $n_1 \geq m$ aus $\mathbb{Z}$ mit

$$\forall n \geq n_1, n \in \mathbb{Z}: |\mu_n - \mu| < \varepsilon' = \frac{\varepsilon}{2M}\,.$$

Wegen $a_n \to a$ für $n \to \infty$ gibt es ein $n_2 \geq m$ aus $\mathbb{Z}$ mit

$$\forall n \geq n_2, n \in \mathbb{Z}: |a_n - a| < \varepsilon'' := \frac{\varepsilon}{2|\mu| + 1}\,.$$

Für alle ganzen Zahlen $n \geq n_0 := \max\{n_1, n_2\}$ folgt dann

$$|\mu_n a_n - \mu a| \leq |\mu_n - \mu| \cdot |a_n| + |\mu| \cdot |a_n - a| < \frac{\varepsilon}{2M} \cdot M + \frac{|\mu|\varepsilon}{2|\mu| + 1} < \varepsilon\,.$$

Also gilt $\mu_n a_n \to \mu a$ für $n \to \infty$. Die übrigen Aussagen in (b) zeigt man analog.

(c) ist ein Spezialfall von (b). (Beachte Beispiel 3.4 (d)).

(d) Annahme: $\lambda > \mu$. Dann gibt es nach (a) zu $\varepsilon := \lambda - \mu > 0$ ein $n_1 \geq m$ in $\mathbb{Z}$, so daß für alle $n \geq n_1$ gilt:

$$0 < \lambda - \mu \leq \mu_n - \lambda_n - (\mu - \lambda) = |(\mu_n - \lambda_n) - (\mu - \lambda)| < \varepsilon = \lambda - \mu\,.$$

Dies ist jedoch ein Widerspruch. Also war die Annahme falsch und es folgt $\lambda \leq \mu$. Sei nun speziell $\lambda = \mu$ vorausgesetzt und gelte $\lambda_n \leq x_n \leq \mu_n$ also auch $\lambda_n - \lambda \leq x_n - \lambda \leq \mu_n - \lambda$ für alle $n \geq n_0$. Sei $\varepsilon > 0$ beliebig. Dann gibt es nach Voraussetzung $n_\lambda \in \mathbb{Z}_p$, $n_\mu \in \mathbb{Z}_q$ mit

$$\forall n \geq n_\lambda, n \in \mathbb{Z}: \quad |\lambda_n - \lambda| < \varepsilon \quad \text{und} \quad \forall n \geq n_\mu, n \in \mathbb{Z}: \quad |\mu_n - \lambda| < \varepsilon\,.$$

Dann folgt

$$\forall n \geq n_2 := \max\{n_0, n_\lambda, n_\mu\}: \qquad |x_n - \lambda| \leq \max\{|\lambda_n - \lambda|, |\mu_n - \lambda|\} < \varepsilon\,.$$

Also gilt $x_n \to \lambda = \mu$ für $n \to \infty$.

(e) Wegen $z \neq 0$ ist $\varepsilon_0 := |z|/2 > 0$. Es gibt also ein $n_1 \in \mathbb{Z}_p$ mit

$$\forall n \geq n_1, n \in \mathbb{Z}: \qquad |z_n - z| < \varepsilon_0 = \frac{|z|}{2}\,.$$

Es folgt dann für alle $n \geq n_1$: $|z| \leq |z_n| + |z - z_n| \leq |z_n| + |z|/2$ und somit $|z_n|^{-1} \leq 2|z|^{-1}$. Sei nun $\varepsilon > 0$ beliebig. Wegen $z_n \to z$ gibt es zu $\varepsilon' := \frac{\varepsilon}{2}|z|^2 > 0$ ein n_2 in $\mathbb{Z}_p$ mit

$$\forall n \geq n_2, n \in \mathbb{Z}: \qquad |z_n - z| < \varepsilon' = \frac{\varepsilon}{2}|z|^2\,.$$

Damit folgt für alle $n \geq n_0 := \max\{n_1, n_2\}$:

$$\Big|\frac{1}{z_n} - \frac{1}{z}\Big| = \frac{|z - z_n|}{|z_n| \cdot |z|} \leq \frac{2|z - z_n|}{|z|^2} < \varepsilon\,.$$

Also gilt $\dfrac{1}{z_n} \to \dfrac{1}{z}$ für $n \to \infty$. Die Behauptung folgt nun mit Hilfe von (b). □

3.14. Lemma (Ungleichung zwischen dem arithmetischen und geometrischen Mittel). *Für alle* $0 \leq u, v \in \mathbb{R}$ *gilt*

$$\sqrt{uv} \leq \frac{u+v}{2}\,.$$

Beweis. Es gilt $0 \leq (u - v)^2 = u^2 + v^2 - 2uv$. Addition von $4uv$ und anschließende Division durch 4 ergibt mit der ersten binomischen Formel

$$uv \leq \Big(\frac{u+v}{2}\Big)^2,$$

woraus man nach Ziehen der Wurzel die Behauptung erhält. □

3.15. Beispiel. (a) Sei $a > 0$ und $0 < x_0 \in \mathbb{R}$. Wir definieren eine Folge $(x_n)_{n=0}^\infty$ induktiv durch

$$x_{n+1} := \frac{1}{2}\Big(x_n + \frac{a}{x_n}\Big) \quad \text{für alle } n \in \mathbb{N}_0\,. \tag{3.4}$$

Dann gilt $x_n \to \sqrt{a}$ für $n \to \infty$.

(b) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.

Beweis. (a) Induktiv zeigt man $x_n > 0$ für alle $n \in \mathbb{N}_0$. Nach der Ungleichung zwischen dem geometrischen und dem arithmetischen Mittel folgt für alle $n \in \mathbb{N}_0$:

$$x_{n+1} = \frac{1}{2}\Big(x_n + \frac{a}{x_n}\Big) \geq \sqrt{x_n \cdot \frac{a}{x_n}} = \sqrt{a}\,,$$

d.h. es ist $x_n \geq \sqrt{a}$ für alle $n \in \mathbb{N}$ also auch $x_n^2 - a \geq 0$. Hiermit folgt

$$x_n - x_{n+1} = x_n - \frac{1}{2}\Big(x_n + \frac{a}{x_n}\Big) = \frac{x_n^2 - a}{2x_n} \geq 0\,.$$

Damit ist gezeigt, daß die Folge $(x_n)_{n=1}^\infty$ monoton fallend und nach unten durch $\sqrt{a}$ beschränkt ist. Nach dem Monotoniekriterium konvergiert sie also gegen ein $y \geq \sqrt{a}$. Mit $x_n \to y$ gilt auch $x_{n+1} \to y$ für $n \to \infty$. Hiermit folgt nach Lemma 3.13 durch Übergang zum Grenzwert für $n \to \infty$ in der Rekursionsformel (3.4)

$$\sqrt{a} \leq y = \frac{1}{2}\Big(y + \frac{a}{y}\Big)$$

also $y^2 = a$ und $y = \sqrt{a}$.

(b) Sei $\varepsilon > 0$ beliebig. Wegen $\sqrt[n]{\frac{2}{\varepsilon}} \to 1$ für $n \to \infty$ (nach Beispiel 3.12) gibt es ein $n_0 \in \mathbb{N}$ mit

$$\forall n \geq n_0, n \in \mathbb{N}: \quad \sqrt[n]{\frac{2}{\varepsilon}}\Big(1 + \frac{\varepsilon}{2}\Big) < 1 + \varepsilon\,.$$

Nach der Bernoullischen Ungleichung 1.22 (d) gilt für alle $n \in \mathbb{N}$:

$$\Big(1 + \frac{\varepsilon}{2}\Big)^n \geq 1 + n \cdot \frac{\varepsilon}{2}$$

also auch

$$1 + \frac{\varepsilon}{2} \geq \sqrt[n]{1 + n \cdot \frac{\varepsilon}{2}} \geq \sqrt[n]{n \cdot \frac{\varepsilon}{2}} = \sqrt[n]{n} \cdot \sqrt[n]{\frac{\varepsilon}{2}}$$

und damit für alle $n \geq n_0$:

$$1 \leq \sqrt[n]{n} \leq \Big(1 + \frac{\varepsilon}{2}\Big)\sqrt[n]{\frac{2}{\varepsilon}} < 1 + \varepsilon\,.$$

Dies zeigt $|\sqrt[n]{n} - 1| < \varepsilon$ für alle $n \geq n_0$. Also gilt $\sqrt[n]{n} \to 1$ für $n \to \infty$. □

Für reelle Folgen definieren wir auch *uneigentliche Grenzwerte* wie folgt:

3.16. Definition. Sei $(x_n)_{n=m}^\infty$ eine Folge in $\mathbb{R}$. Wir sagen x_n *konvergiert gegen* ∞ (bzw. $-\infty$) und schreiben $x_n \to \infty$ (bzw. $x_n \to -\infty$) für $n \to \infty$, falls es zu jedem $c > 0$ ein $n_0 \geq m$ aus $\mathbb{Z}$ gibt, so daß für alle $n \geq n_0$, $n \in \mathbb{Z}$, gilt: $x_n > c$ (bzw. $x_n < -c$).

Beispiele. (a) $n \to \infty$ für $n \to \infty$.

(b) $h_n = \sum_{k=1}^{n} \frac{1}{k} \to \infty$ für $n \to \infty$ (vergl. den Beweis zu Satz 3.9).

3.17. Bezeichnungen. Für $a, b \in \mathbb{R}$ mit $a \leq b$ nennen wir die Punktmenge

- $[a, b] := \{x \in \mathbb{R}\,;\, a \leq x \leq b\}$ ein *abgeschlossenes Intervall.*
- $[a, a] = \{a\}$ ein *ausgeartetes* abgeschlossenes Intervall.

Für $a, b \in \mathbb{R}$ mit $a < b$ nennen wir

- $(a, b) := \{x \in \mathbb{R}\,;\, a < x < b\}$ ein *offenes Intervall,*

- $[a, b) := \{x \in \mathbb{R}\,;\, a \leq x < b\}$ und $(a, b] := \{x \in \mathbb{R}\,;\, a < x \leq b\}$ ein *halboffenes Intervall.*

Für $a \in \mathbb{R}$ heißen

- $(a, \infty) := \{x \in \mathbb{R}\,;\, x > a\}$ und $(-\infty, a) := \{x \in \mathbb{R}\,;\, x < a\}$ *offene uneigentliche Intervalle,*
- $[a, \infty) := \{x \in \mathbb{R}\,;\, x \geq a\}$ und $(-\infty, a] := \{x \in \mathbb{R}\,;\, x \leq a\}$ *abgeschlossene uneigentliche Intervalle.*

3.18. Intervallschachtelung. *Sei $([a_n, b_n])_{n=0}^\infty$ eine Folge von abgeschlossenen Intervallen mit den folgenden Eigenschaften:*

(I1) *$(a_n)_{n=0}^\infty$ ist monoton wachsend und $(b_n)_{n=0}^\infty$ ist monoton fallend, d.h. für alle $n \in \mathbb{N}_0$ gilt $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$.*

Dann ist $\displaystyle\bigcap_{n=0}^\infty [a_n, b_n] \neq \emptyset$.

Gilt zusätzlich

(I2) *$b_n - a_n \to 0$ für $n \to \infty$,*

so ist

$$\bigcap_{n=0}^\infty [a_n, b_n] = \{u\}$$

für genau ein $u \in \mathbb{R}$ und für alle Folgen $(x_n)_{n=0}^\infty$ mit $a_n \leq x_n \leq b_n$ für alle $n \in \mathbb{N}_0$ gilt $x_n \to u$ für $n \to \infty$.

Beweis. Induktiv sieht man $a_0 \leq a_n \leq a_{n+1} \leq b_{n+1} \leq b_n \leq b_0$ für alle $n \in \mathbb{N}_0$. Nach dem Monotoniekriterium 3.11 sind daher die Folgen $(a_n)_{n=0}^\infty$ und $(b_n)_{n=0}^\infty$ konvergent und es gilt (unter Verwendung von Lemma 3.13 (d)):

$$\lim_{n\to\infty} a_n = a := \sup_{n\in\mathbb{N}_0} a_n \leq \lim_{n\to\infty} b_n = b := \inf_{n\in\mathbb{N}_0} b_n\,.$$

Insbesondere hat man $a_n \leq a \leq b \leq b_n$ für alle $n \in \mathbb{N}_0$, d.h. $[a, b] \subseteq [a_n, b_n]$ für alle $n \in \mathbb{N}_0$ und damit

$$\emptyset \neq [a, b] \subseteq \bigcap_{n=0}^\infty [a_n, b_n]\,.$$

Gilt nun zusätzlich $b_n - a_n \to 0$ für $n \to \infty$, so ist $a = b$ und für jede Folge $(x_n)_{n=0}^\infty$ mit $a_n \leq x_n \leq b_n$ für alle $n \in \mathbb{N}_0$ folgt nach der Einschachtelungsregel 3.13 (d): $x_n \to u := a = b$ für $n \to \infty$. Insbesondere ist $u := a = b$ der einzige Punkt in $\bigcap_{n=0}^\infty [a_n, b_n]$. $\square$

3.19. Definition. Sei $N \in \mathbb{N}$. Eine Folge $(x_n)_{n=m}^\infty$ in $\mathbb{R}^N$ heißt *Cauchy–Folge*[2], falls gilt:

(CF) $\forall \varepsilon > 0\, \exists n_0 \geq m\, \forall n, k \in \mathbb{Z}$ mit $n, k \geq n_0 : \qquad |x_n - x_k| < \varepsilon.$

Man überlegt sich leicht, daß (CF) äquivalent ist zu

(CF′) $\forall \varepsilon > 0\, \exists n_0 \geq m\, \forall n \in \mathbb{Z}$ mit $n \geq n_0\, \forall p \in \mathbb{N}_0 : \qquad |x_n - x_{n+p}| < \varepsilon.$

3.20. Satz. *Jede konvergente $\mathbb{R}^N$-wertige Folge ist eine Cauchy–Folge.*

[2] Augustin Louis Cauchy (21.8.1789–23.5.1857)

Beweis. Sei also $(x_n)_{n=m}^{\infty}$ eine konvergente Folge in $\mathbb{R}^N$ mit $u := \lim_{n\to\infty} x_n$ und sei $\varepsilon > 0$ beliebig. Zu $\varepsilon' = \varepsilon/2 > 0$ gibt es dann ein $n_0 \geq m$ in $\mathbb{Z}$ mit

$$\forall n \geq n_0, n \in \mathbb{Z}: \qquad |x_n - u| < \varepsilon' = \varepsilon/2\,.$$

Dann gilt für alle $n, k \geq n_0$ aus $\mathbb{Z}$:

$$|x_n - x_k| \leq |x_n - u| + |u - x_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also ist (CF) erfüllt. □

Umgekehrt gilt (zunächst für $N = 1$):

3.21. Satz (Cauchy–Kriterium in $\mathbb{R}$). *Jede reelle Cauchy–Folge ist konvergent.*

Beweis. Sei also $(x_n)_{n=m}^{\infty}$ eine Cauchy–Folge in $\mathbb{R}$. Zu jedem $p \in \mathbb{N}_0$ gibt es dann ein $n_p \geq m$ aus $\mathbb{Z}$ mit

$$\forall n, k \in \mathbb{Z} \text{ mit } n, k \geq n_p: \qquad |x_n - x_k| < \varepsilon_p := 2^{-p-1}\,.$$

Durch vollständige Induktion sieht man, daß hierbei die Folge $(n_p)_{p=0}^{\infty}$ so gewählt werden kann, daß $n_p \leq n_{p+1}$ für alle $p \in \mathbb{N}_0$ gilt. Wir setzen nun

$$I_p := [x_{n_p} - 2^{-p}, x_{n_p} + 2^{-p}] = \{x \in \mathbb{R}\,;\, |x - x_{n_p}| \leq 2^{-p}\}\,.$$

Dann ist $I_{p+1} \subset I_p$ für alle $p \in \mathbb{N}_0$, denn für alle $x \in I_{p+1}$ folgt

$$|x - x_{n_p}| \leq |x - x_{n_{p+1}}| + |x_{n_{p+1}} - x_{n_p}| < 2^{-p-1} + 2^{-p-1} = 2^{-p}.$$

Die Intervallfolge $(I_p)_{p=0}^{\infty}$ erfüllt also die Bedingung (I1) und wegen

$$x_{n_p} + 2^{-p} - (x_{n_p} - 2^{-p}) = 2^{1-p} \to 0 \text{ für } p \to \infty$$

auch die Bedingung (I2). Nach dem Intervallschachtelungsprinzip 3.18 gibt es also genau ein $u \in \bigcap_{p=0}^{\infty} I_p$. Sei nun $\varepsilon > 0$ beliebig. Dann gibt es ein p_0 mit $2^{1-p_0} < \varepsilon$. Für alle $n \geq n_{p_0}$, $n \in \mathbb{Z}$, gilt dann:

$$|x_n - u| \leq |x_n - x_{n_{p_0}}| + |x_{n_{p_0}} - u| < 2^{-p_0-1} + 2^{-p_0} < 2^{1-p_0} < \varepsilon\,.$$

Also gilt $x_n \to u$ für $n \to \infty$. □

Wir wollen das Cauchy–Kriterium auch für $\mathbb{R}^N$–wertige und komplexwertige Folgen herleiten. Hierzu zeigen wir zunächst:

3.22. Satz. *Eine $\mathbb{R}^N$–wertige Folge ist genau dann konvergent, wenn sie komponentenweise konvergiert, d.h.: Ist $(x_n)_{n=m}^{\infty}$ eine Folge in $\mathbb{R}^N$ und $u = (u_1, \ldots, u_N) \in \mathbb{R}^N$, so gilt $x_n \to u$ in $\mathbb{R}^N$ für $n \to \infty$ genau dann, wenn für alle $j = 1, \ldots, N$ für die j–te Komponente $x_{j,n}$ von x_n gilt $x_{j,n} \to u_j$ in $\mathbb{R}$ für $n \to \infty$.*

Beweis. „$\Longrightarrow$“: Sei also $x_n \to u$ für $n \to \infty$ und sei $\varepsilon > 0$ beliebig. Dann gibt es ein $n_0 \geq m$ aus $\mathbb{Z}$ mit $|x_n - u| < \varepsilon$ für alle $n \geq n_0$ aus $\mathbb{Z}$. Insbesondere gilt für die j–te Komponente ($j = 1, \ldots, N$) für alle $n \geq n_0$ aus $\mathbb{Z}$:

$$|x_{j,n} - u_j| \leq \Big(\sum_{k=1}^{N} (x_{k,n} - u_k)^2\Big)^{1/2} = |x_n - u| < \varepsilon\,.$$

Also folgt $x_{j,n} \to u_j$ in $\mathbb{R}$ für $n \to \infty$ und alle $j = 1, \ldots, N$.

„$\Longleftarrow$“: Es gelte nun für $j = 1, \ldots, N$: $x_{j,n} \to u_j$ in $\mathbb{R}$ für $n \to \infty$. Sei $\varepsilon > 0$ beliebig. Dann ist $\varepsilon' := \frac{\varepsilon}{\sqrt{N}} > 0$. Zu jedem $j \in \{1, \ldots, N\}$ gibt es also ein $n_j \geq m$ aus $\mathbb{Z}$ mit

$$\forall n \geq n_j, n \in \mathbb{Z}: \qquad |x_{j,n} - u_j| < \varepsilon' = \frac{\varepsilon}{\sqrt{N}}\,.$$

Dann gilt für alle $n \geq n_0 := \max\{n_1, \ldots, n_N\}$ aus $\mathbb{Z}$:

$$|x_n - u| = \Big(\sum_{k=1}^{N}(x_{k,n} - u_k)^2\Big)^{1/2} < \Big(\sum_{k=1}^{N}\varepsilon'^2\Big)^{1/2} = \varepsilon'\sqrt{N} = \varepsilon\,.$$

Also folgt $x_n \to u$ in $\mathbb{R}^N$ für $n \to \infty$. $\square$

3.23. Satz (Cauchy–Kriterium in $\mathbb{R}^N$). *Für eine Folge $(x_n)_{n=m}^{\infty}$ in $\mathbb{R}^N$ sind die folgenden Aussagen äquivalent:*

(a) *$(x_n)_{n=m}^{\infty}$ ist konvergent in $\mathbb{R}^N$.*
(b) *$(x_n)_{n=m}^{\infty}$ ist eine Cauchy-Folge in $\mathbb{R}^N$.*
(c) *Für alle $j \in \{1, \ldots, N\}$ ist $(x_{j,n})_{n=m}^{\infty}$ eine Cauchy–Folge in $\mathbb{R}$.*
(d) *Für alle $j \in \{1, \ldots, N\}$ ist $(x_{j,n})_{n=m}^{\infty}$ in $\mathbb{R}$ konvergent.*

Beweis. Die Äquivalenzen (c) $\iff$ (d) $\iff$ (a) und die Implikation (a) $\Longrightarrow$ (b) gelten nach den Sätzen 3.20, 3.21 und 3.22. Es ist also nur die Implikation (b) $\Longrightarrow$ (c) zu zeigen. Sei also $(x_n)_{n=m}^{\infty}$ eine Cauchy–Folge in $\mathbb{R}^N$ und sei $\varepsilon > 0$ beliebig. Dann gibt es ein $n_0 \geq m$ aus $\mathbb{Z}$, so daß für alle ganzen Zahlen $n, k \geq n_0$ gilt $|x_n - x_k| < \varepsilon$. Wegen $|x_{j,n} - x_{j,k}| \leq |x_n - x_k| < \varepsilon$ für alle $j = 1, \ldots, N$ und alle $n, k \geq n_0$ aus $\mathbb{Z}$ ist dann auch $(x_{j,n})_{n=m}^{\infty}$ eine Cauchy–Folge in $\mathbb{R}$ für alle $j = 1, \ldots, N$. $\square$

Da wir $\mathbb{C}$ mit $\mathbb{R}^2$ identifizieren können erhalten wir:

3.24. Folgerung (Cauchy–Kriterium in $\mathbb{C}$). *Für eine Folge $(z_n)_{n=m}^{\infty}$ in $\mathbb{C}$ sind die folgenden Aussagen äquivalent:*

(a) *$(z_n)_{n=m}^{\infty}$ ist konvergent in $\mathbb{C}$.*
(b) *$(z_n)_{n=m}^{\infty}$ ist eine Cauchy-Folge in $\mathbb{C}$.*
(c) *$(\operatorname{Re} z_n)_{n=m}^{\infty}$ und $(\operatorname{Im} z_n)_{n=m}^{\infty}$ sind Cauchy–Folgen in $\mathbb{R}$.*
(d) *Die Folgen $(\operatorname{Re} z_n)_{n=m}^{\infty}$ und $(\operatorname{Im} z_n)_{n=m}^{\infty}$ sind in $\mathbb{R}$ konvergent.*

3. Häufungswerte und der Satz von Bolzano–Weierstraß

3.25. Definition. Sei X eine nicht leere Menge und sei $(x_n)_{n=m}^{\infty}$ eine Folge von Elementen aus X. Ist $(n_k)_{k=1}^{\infty}$ eine streng monoton wachsende Folge ganzer Zahlen mit $m \leq n_1 < n_2 < \cdots < n_k < n_{k+1} < \ldots$, so heißt die Folge $(x_{n_k})_{k=1}^{\infty}$ eine *Teilfolge von* $(x_n)_{n=m}^{\infty}$.

3.26. Satz. *Für eine Folge $(x_n)_{n=m}^{\infty}$ in $\mathbb{R}^N$ sind äquivalent:*

(a) *Es gibt ein $u \in \mathbb{R}^N$, so daß für jede Teilfolge $(x_{n_k})_{k=1}^{\infty}$ von $(x_n)_{n=m}^{\infty}$ gilt $x_{n_k} \to u$ für $k \to \infty$.*
(b) *Jede Teilfolge $(x_{n_k})_{k=1}^{\infty}$ von $(x_n)_{n=m}^{\infty}$ ist konvergent.*
(c) *$(x_n)_{n=m}^{\infty}$ ist konvergent.*

Beweis. (a) $\Longrightarrow$ (b) ist offensichtlich und auch (b) $\Longrightarrow$ (c) ist klar, da $(x_n)_{n=m}^{\infty}$ eine Teilfolge von sich ist (mit $n_k := m - 1 + k$ für alle $k \in \mathbb{N}$). Zu zeigen ist also nur (c) $\Longrightarrow$ (a). Sei also $(x_n)_{n=m}^{\infty}$ konvergent, $u := \lim_{n\to\infty} x_n$ und sei $(x_{n_k})_{k=1}^{\infty}$ eine beliebige Teilfolge von $(x_n)_{n=m}^{\infty}$. Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es ein $n_0 \geq m$ aus $\mathbb{Z}$ mit

$$\forall n \geq n_0, n \in \mathbb{Z}: \qquad |x_n - u| < \varepsilon\,.$$

Induktiv zeigt man, daß wegen der strengen Monotonie der Folge $(n_k)_{k=1}^{\infty}$ gilt: $n_k \geq m - 1 + k$ für alle $k \in \mathbb{N}$. Für alle $k \geq k_0 := n_0 - m + 1$ gilt dann $n_k \geq m - 1 + k \geq n_0$ und daher $|x_{n_k} - u| < \varepsilon$. Also konvergiert $(x_{n_k})_{k=1}^{\infty}$ gegen u für $n \to \infty$. $\square$

Beispiele. Die Folge $((-1)^n)_{n=0}^{\infty}$ ist divergent, besitzt aber konvergente Teilfolgen wie zum Beispiel die gegen 1 konvergente Folge $((-1)^{2k})_{k=1}^{\infty}$ und die gegen -1 konvergente Folge $((-1)^{2k-1})_{k=1}^{\infty}$.

Die Folge $(n)_{n=0}^{\infty}$ besitzt keine konvergente Teilfolge.

3.27. Definition. Ein Punkt $b \in \mathbb{R}^N$ heißt *Häufungswert* einer Folge $(x_n)_{n=m}^{\infty}$ in $\mathbb{R}^N$, falls diese eine gegen b konvergente Teilfolge besitzt.

Beispiele. Die Punkte 1 und -1 sind Häufungswerte der Folge $((-1)^n)_{n=0}^{\infty}$.

Die Folge $(n)_{n=0}^{\infty}$ besitzt keine Häufungswerte, da sie keine konvergente Teilfolge besitzt.

Häufungswerte lassen sich auch wie folgt charakterisieren:

3.28. Satz. *Sei $(x_n)_{n=m}^{\infty}$ eine Folge in $\mathbb{R}^N$. Ein Punkt $b \in \mathbb{R}^N$ ist genau dann Häufungswert der Folge $(x_n)_{n=m}^{\infty}$, wenn gilt:*

$$\forall \varepsilon > 0\ \forall p \in \mathbb{Z}_m\ \exists q \geq p, q \in \mathbb{Z} : \qquad x_q \in U_\varepsilon(b) = \{x \in \mathbb{R}^N;\ |x - b| < \varepsilon\}\,.$$

Beweis. „$\Longrightarrow$“: Sei also b ein Häufungswert der Folge $(x_n)_{n=m}^{\infty}$ und sei $\varepsilon > 0$ und $p \in \mathbb{Z}$ beliebig mit $p \geq m$. Nach Definition des Häufungswertes gibt es eine gegen b konvergente Teilfolge $(x_{n_k})_{k=1}^{\infty}$. Für diese existiert also ein $k_0 \in \mathbb{N}$, so daß für alle $k \geq k_0$ aus $\mathbb{N}$ gilt: $|a_{n_k} - b| < \varepsilon$. Sei $k := \max\{k_0, p - m + 1\}$. Dann ist $q := n_k \geq m - 1 + k \geq p$ und $a_q = a_{n_k} \in U_\varepsilon(b)$.

„$\Longleftarrow$“: Sei nun die Bedingung des Satzes erfüllt. Wir konstruieren induktiv eine Teilfolge $(x_{n_k})_{k=1}^{\infty}$ von $(x_n)_{n=m}^{\infty}$, für die die beiden folgenden Bedingungen für alle $k \in \mathbb{N}$ erfüllt sind (mit $n_0 := m$):

(a) $n_k > n_{k-1}$.
(b) $|b - x_{n_k}| < k^{-1}$.

Induktionsanfang: Nach Voraussetzung gibt es zu $\varepsilon = 1 > 0$ ein $n_1 > m$ aus $\mathbb{Z}$ mit $|b - x_{n_1}| < 1$. Damit sind die Bedingungen (a) und (b) für $k = 1$ erfüllt.

Seien nun schon n_j für $j \leq k$ aus $\mathbb{Z}$ gefunden, so daß die Bedingungen (a) und (b) für $1 \leq j \leq k$ erfüllt sind. Nach Voraussetzung gibt es zu $\varepsilon := (k+1)^{-1} > 0$ und $p := n_k + 1$ ein $n_{k+1} \geq p > n_k$ mit $|b - x_{n_{k+1}}| < \varepsilon = (k+1)^{-1}$. Damit sind (a) und (b) auch für $k+1$ erfüllt.

Für die so konstruierte Teilfolge gilt nun $|b - x_{n_k}| < 1/k \to 0$ für $k \to \infty$ und somit $b - x_{n_k} \to 0$ in $\mathbb{R}^N$, also auch $x_{n_k} \to b$ für $k \to \infty$. □

3.29. Satz (Bolzano–Weierstraß[3]). *Jede beschränkte Folge reeller Zahlen hat wenigstens einen Häufungswert, also auch wenigstens eine konvergente Teilfolge.*

Beweis. Sei also $(x_n)_{n=m}^{\infty}$ eine beliebige beschränkte Folge in $\mathbb{R}$. Dann gibt es ein $M > 0$ mit $|x_n| \leq M$ für alle $n \geq m$ aus $\mathbb{Z}$. Für alle $n \geq m$ aus $\mathbb{Z}$ gilt dann mit $C_n := \{x_j\,;\ n \leq j \in \mathbb{Z}\}$:

$$-M \leq x_n \leq \sup C_n \leq M \text{ und } C_{n+1} \subseteq C_n \text{ also auch } \sup C_{n+1} \leq \sup C_n\,.$$

Die Folge $(\sup C_n)_{n=m}^{\infty}$ ist also monoton fallend und nach unten durch $-M$ beschränkt. Nach dem Monotoniekriterium 3.11 gilt also

$$\sup C_n \to y := \inf\{\sup C_n\,;\ m \leq n \in \mathbb{Z}\} \quad \text{für } n \to \infty.$$

[3] Bernard Placidus Johann Nepomuk Bolzano(5.10.1781–18.12.1848), Karl Theodor Wilhelm Weierstrass (31.10.1815–19.2.1897)

Wir zeigen, daß der Punkt y ein Häufungswert der Folge $(x_n)_{n=m}^\infty$ ist: Seien hierzu $\varepsilon > 0$ und $m \le p \in \mathbb{Z}$ beliebig vorgegeben. Dann gibt es ein $n_0 \ge m$ aus $\mathbb{Z}$ mit

$$\forall n \ge n_0, n \in \mathbb{Z}: \qquad y \le \sup C_n < y + \varepsilon\,.$$

Zu $k := \max\{p, n_0\}$ gibt es nach Lemma 1.16 ein $k \le q \in \mathbb{Z}$ mit

$$y - \varepsilon \le \sup C_k - \varepsilon < x_q \le \sup C_k < y + \varepsilon\,.$$

Es folgt $q \ge p$ und $x_q \in U_\varepsilon(y)$. Nach Satz 3.28 ist y also ein Häufungswert der Folge $(x_n)_{n=m}^\infty$. □

3.30. Bemerkung. Der im Beweis zum Satz 3.29 von Bolzano–Weierstraß berechnete Häufungswert $y = \lim_{n\to\infty} \sup C_n$ ist der größte Häufungswert der Folge $(x_n)_{n=m}^\infty$.

Beweis. Sei $s > y$ beliebig. Dann ist $\varepsilon := (s-y)/2 > 0$. Es gibt also ein $p \ge m$ aus $\mathbb{Z}$ mit $\sup C_p < y + \varepsilon$. Für alle $q \ge p$ ist daher $x_q \notin U_\varepsilon(s)$. Nach Satz 3.28 kann s kein Häufungswert von $(x_n)_{n=m}^\infty$ sein. □

Analog zeigt man:

3.31. Bemerkung. Sei $(x_n)_{n=m}^\infty$ eine beschränkte reelle Folge. Dann existiert der Grenzwert

$$\lim_{n\to\infty} \inf\{x_k\,;\, n \le k \in \mathbb{Z}\}$$

und ist der kleinste Häufungswert der Folge $(x_n)_{n=m}^\infty$.

3.32. Definition. Sei $(x_n)_{n=m}^\infty$ eine beschränkte Folge in $\mathbb{R}$. Dann heißt der gemäß 3.30 (bzw. 3.31) definierte Wert der *Limes superior* (bzw. *Limes inferior*) von $(x_n)_{n=m}^\infty$ und wird mit $\limsup_{n\to\infty} x_n$ (bzw. $\liminf_{n\to\infty} x_n$) bezeichnet. Es ist also

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k\,;\, n \le k \in \mathbb{Z}\} \quad \text{und} \quad \liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k\,;\, n \le k \in \mathbb{Z}\}\,.$$

Wir definieren allgemeiner für beliebige Folgen $(x_n)_{n=m}^\infty$ in $\mathbb{R}$:

$$\limsup_{n\to\infty} x_n := \begin{cases} +\infty & \text{falls } (x_n)_{n=m}^\infty \text{ nicht nach oben beschränkt ist} \\ -\infty & \text{falls } x_n \to -\infty \text{ für } n \to \infty \\ \lim_{n\to\infty} \sup\{x_k\,;\, n \le k \in \mathbb{Z}\} & \text{sonst} \end{cases}$$

und ebenso

$$\liminf_{n\to\infty} x_n := \begin{cases} +\infty & \text{falls } x_n \to +\infty \text{ für } n \to \infty \\ -\infty & \text{falls } (x_n)_{n=m}^\infty \text{ nicht nach unten beschränkt ist} \\ \lim_{n\to\infty} \inf\{x_k\,;\, n \le k \in \mathbb{Z}\} & \text{sonst}\,. \end{cases}$$

3.33. Satz. *Für eine beschränkte Folge $(x_n)_{n=m}^\infty$ in $\mathbb{R}$ sind die folgenden Aussagen äquivalent:*

(a) $(x_n)_{n=m}^\infty$ *ist konvergent.*
(b) $(x_n)_{n=m}^\infty$ *besitzt genau einen Häufungswert.*
(c) $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

Ist (b) erfüllt, so stimmt der einzige Häufungswert von $(x_n)_{n=m}^\infty$ mit dem Grenzwert der Folge überein.

Beweis. (a) $\Longrightarrow$ (b) aus Satz 3.26 und der Definition des Häufungswertes.

(b) $\iff$ (c) ist klar, da $\liminf_{n\to\infty} x_n$ der kleinste und $\limsup_{n\to\infty} x_n$ der größte Häufungswert von $(x_n)_{n=m}^\infty$ ist.

(b) $\Longrightarrow$ (a): Sei u der einzige Häufungswert der Folge $(x_n)_{n=m}^\infty$. Wir machen die Annahme, daß die Folge $(x_n)_{n=m}^\infty$ nicht gegen den einzigen Häufungswert u konvergiert, d.h. daß ein $\varepsilon > 0$ existiert mit

$$\forall m \le p \in \mathbb{Z}\ \exists q \ge p, q \in \mathbb{Z}: \qquad |x_q - u| \ge \varepsilon\,.$$

Insbesondere gibt es ein $n_0 := m < m+1 \le n_1 \in \mathbb{Z}$ mit $|x_{n_1} - u| \ge \varepsilon$. Sind für $1 \le j \le k$ die Zahlen $n_j \in \mathbb{Z}$ schon konstruiert mit $n_j > n_{j-1}$ und $|x_{n_j} - u| \ge \varepsilon$ für $1 \le j \le k$, so gibt es wegen der gemachten Annahme eine ganze Zahl $n_{k+1} \ge n_k + 1 > n_k$ mit $|x_{n_{k+1}} - u| \ge \varepsilon$. Auf diese Weise haben wir induktiv eine Teilfolge $(x_{n_k})_{k=1}^\infty$ von $(x_n)_{n=m}^\infty$ konstruiert mit $|x_{n_k} - u| \ge \varepsilon$ für alle $k \in \mathbb{N}$. Als Teilfolge einer beschränkten Folge ist auch die Folge $(x_{n_k})_{k=1}^\infty$ beschränkt, besitzt also nach dem Satz von Bolzano–Weierstraß eine konvergente Teilfolge $(x_{n_{k_j}})_{j=1}^\infty$. Sei $v := \lim_{j\to\infty} x_{n_{k_j}}$. Da $(x_{n_{k_j}})_{j=1}^\infty$ insbesondere auch Teilfolge der Ausgangsfolge $(x_n)_{n=m}^\infty$ ist, ist v ein Häufungspunkt von $(x_n)_{n=m}^\infty$, muß also mit u übereinstimmen. Also gilt $|x_{n_{k_j}} - u| \to 0$ für $j \to \infty$ im Widerspruch zu $|x_{n_{k_j}} - u| \ge \varepsilon$ für alle $j \in \mathbb{N}$. Also war die Annahme falsch und es folgt die Konvergenz von $(x_n)_{n=m}^\infty$ gegen u für $n \to \infty$. □

3.34. Satz (Satz von Bolzano–Weierstraß in $\mathbb{R}^N$). *Jede beschränkte Folge in $\mathbb{R}^N$ hat wenigstens eine konvergente Teilfolge und daher wenigstens einen Häufungswert.*

Beweis. Sei also $(x_n)_{n=m}^\infty$ eine beliebige beschränkte Folge in $\mathbb{R}^N$. Dann gibt es ein $M > 0$ mit $|x_n| \le M$ für alle $m \le n \in \mathbb{Z}$. Insbesondere folgt für alle Komponenten $x_{j,n}$ von $x_n = (x_{1,n}, \dots, x_{N,n})$:

$$\forall m \le n \in \mathbb{Z} \quad \forall j \in \{1, \dots, N\}: \qquad |x_{j,n}| \le |x_n| \le M\,.$$

Nach dem Satz 3.29 von Bolzano–Weierstraß (in $\mathbb{R}$) hat also die Folge $(x_{1,n})_{n=m}^\infty$ eine gegen ein $u_1 \in \mathbb{R}$ konvergente Teilfolge $(x_{1,n_1(k)})_{k=1}^\infty$. Dann sind für $j = 1, \dots, N$ die Folgen $(x_{j,n_1(k)})_{k=1}^\infty$ als Teilfolgen der beschränkten Folgen $(x_{j,n})_{n=m}^\infty$ wieder beschränkte reelle Folgen. Insbesondere besitzt nach 3.29 die Folge $(x_{2,n_1(k)})_{k=1}^\infty$ eine gegen ein $u_2 \in \mathbb{R}$ konvergente Teilfolge $(x_{2,n_2(k)})_{k=1}^\infty$. Als Teilfolge der gegen u_1 konvergenten Folge $(x_{1,n_1(k)})_{k=1}^\infty$ ist nach Satz 3.26 auch $(x_{1,n_2(k)})_{k=1}^\infty$ gegen u_1 konvergent und die Folgen $(x_{j,n_2(k)})_{k=1}^\infty$ sind als Teilfolgen der beschränkten Folgen $(x_{j,n_1(k)})_{k=1}^\infty$ wieder beschränkte reelle Folgen. Fahren wir in dieser Weise fort, so erhalten wir nach N Schritten eine streng monoton wachsende Folge $(n_N(k))_{k=1}^\infty$, so daß für $j = 1, \dots, N$ die zugehörigen Teilfolgen der Folgen $(x_{j,n})_{n=m}^\infty$ gegen ein u_j in $\mathbb{R}$ konvergieren. Nach Satz 3.23 ist also die Teilfolge $(x_{n_N(k)})_{k=1}^\infty$ von $(x_n)_{n=m}^\infty$ konvergent in $\mathbb{R}^N$. □

Eng verwandt zum aber verschieden vom Begriff des Häufungswertes ist der Begriff des Häufungspunktes.

3.35. Definition. Sei $\emptyset \ne X \subseteq \mathbb{R}^M$. Ein Punkt $a \in \mathbb{R}^M$ heißt

- *Häufungspunkt der Menge X*, falls es eine gegen a konvergente Folge in $X \setminus \{a\}$ gibt.
- *isolierter Punkt von X*, falls es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(a) \cap X = \{a\}$.

3.36. Beispiele. (a) Die Folge $((-1)^n)_{n=1}^\infty$ hat die Häufungswerte 1 und -1. Die Menge $\{(-1)^n\,;\, n \in \mathbb{N}\}$ ihrer Werte besteht nur aus den isolierten Punkten 1 und -1, hat also keine Häufungspunkte.

(b) Für die Menge $\mathbb{B}_N := \{x \in \mathbb{R}^N\,;\, |x| < 1$ sind alle $x \in \mathbb{R}^N$ mit $|x| \leq 1$ Häufungspunkte. Die Häufungspunkte einer Menge X müssen also nicht notwendig Elemente der Menge X sein.

Wir geben zunächst eine äquivalente Beschreibung der Häufungspunkte:

3.37. Lemma. *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$. Ein Punkt $a \in \mathbb{R}^M$ ist genau dann ein Häufungspunkt der Menge X, wenn gilt:*

$$\forall \varepsilon > 0: \qquad (X \setminus \{a\}) \cap U_\varepsilon(a) \neq \emptyset\,.$$

Beweis. Ist a Häufungspunkt der Menge X, so gibt es eine gegen a konvergente Folge $(x_n)_{n=1}^\infty$ aus $X \setminus \{a\}$. Für alle $\varepsilon > 0$ existiert also ein n_0 so, daß $x_n \in U_\varepsilon(a)$ für alle $n \geq n_0$. Insbesondere ist $U_\varepsilon(a) \cap (X \setminus \{a\}) \neq \emptyset$.

Ist umgekehrt die Bedingung des Lemmas erfüllt, so gibt es zu $\varepsilon_n := 1/n$ wenigstens einen Punkt $x_n \in U_{1/n}(a) \cap (X \setminus \{a\})$. Die so erhaltene Folge $(x_n)_{n=1}^\infty$ aus $X \setminus \{a\}$ konvergiert offensichtlich gegen a für $n \to \infty$. □

3.38. Satz (Satz von Bolzano–Weierstraß für beschränkte Mengen in $\mathbb{R}^M$). *Ist $\emptyset \neq X \subset \mathbb{R}^M$ beschränkt und hat X unendlich viele Elemente, so besitzt X wenigstens einen Häufungspunkt.*

Beweis. Da X nicht endlich ist, gibt es eine Folge $(x_n)_{n=1}^\infty$ von paarweise verschiedenen Punkten in X. Diese ist beschränkt, da X beschränkt ist, besitzt also nach dem Satz von Bolzano–Weierstraß für Folgen wenigstens eine konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Sei $a := \lim_{k\to\infty} x_{n_k}$. Da die Glieder der Folge $(x_n)_{n=1}^\infty$ paarweise verschieden sind, kann $x_{n_k} = a$ für höchstens ein $k = k_0 \in \mathbb{N}$ gelten. Die Folge $(x_{n_{k_0+j}})_{j=1}^\infty$ ist also eine gegen a konvergente Folge aus $X \setminus \{a\}$, d.h. a ist ein Häufungspunkt von X. □

Die Probleme eines Häufungspunktes schildern anschaulich die folgenden Verse[4]:

Der Häufungspunkt

Ein Häufungspunkt, welchselb'ger zwar
In herrlicher Umgebung war,
Der aber dennoch unzufrieden
Mit seinem Schicksal war hienieden,
Bedachte seine Lage neulich
Und sprach betrübt: "Es ist abscheulich!
Mir sagt wohl jeder, der mich kennt,
Ich hätt zum Einsiedler Talent.
Wie gern möcht ich zurück mich ziehn,
Die lästige Umgebung fliehn;
Allein in jedem Kreis mit ρ
Sind immer noch Begleiter do!"
Des Punktes Standpunkt leuchtet ein;
Auch ich möcht solch ein Punkt nicht sein!
Ging ich z.B. zur Toilette,

[4] Aus: *Häufungspunkte* -Mathematischer Reimsalat für die studierende Jugend von 3 bis 17 Semestern, von Dr. h.c. N^2. Herausgegeben vom Mathematischen Verein an der Universität Berlin, Berlin 1927.

In jeder nächsten Näh ich hätte
Unendlich viel Bekannte
und liebe Anverwandte.

Dies aber wäre mir furchtbar peinlich.

4. Konvergenzkriterien für unendliche Reihen

Sei $(a_n)_{n=m}^{\infty}$ eine Folge in $\mathbb{R}^N$. Die Folge $(s_n)_{n=m}^{\infty}$ der n–ten *Partialsummen*

$$s_n := \sum_{k=m}^{n} a_k \qquad \text{für } m \leq n \in \mathbb{Z}$$

nennt man eine *unendliche Reihe* und bezeichnet sie mit $\sum_{k=m}^{\infty} a_k$. Ist die Partialsummenfolge $(s_n)_{n=m}^{\infty}$ konvergent gegen einen Grenzwert u in $\mathbb{R}^N$, so schreiben wir

$$\sum_{k=m}^{\infty} a_k = u$$

und sagen die Reihe $\sum_{k=m}^{\infty} a_k$ *ist konvergent.* Die Reihe $\sum_{k=m}^{\infty} a_k$ heiß *divergent*, falls sie nicht konvergent ist, d.h. falls die Folge ihrer Partialsummen divergiert.

Da unendliche Reihen insbesondere Folgen sind, lassen sich die im vorigen Kapitel bewiesenen Aussagen sinngemäß auch auf unendliche Reihen anwenden.

Umgekehrt läßt sich auch jede Folge $(a_n)_{n=m}^{\infty}$ als unendliche Reihe darstellen, denn es gilt für alle $m \leq n \in \mathbb{Z}$:

$$a_n = a_m + \sum_{k=m+1}^{n} (a_k - a_{k-1}) = \sum_{k=m}^{n} b_k$$

mit $b_m := a_m$ und $b_k := a_k - a_{k-1}$ für $m < k \in \mathbb{Z}$.

Einige Beispiele von unendlichen Reihen haben wir schon kennengelernt.

3.39. Beispiele. (a) Die *harmonische Reihe*[5] $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ ist nach Satz 3.9 divergent, da die Folge ihrer Partialsummen unbeschränkt ist.

(b) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $q \in \mathbb{K}$. Die *geometrische Reihe*[6] $\sum_{k=0}^{\infty} q^k$ ist nach Beispiel 3.4 (c) für $|q| < 1$ konvergent und es gilt

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}\,.$$

(c) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = 1.$

Beweis. Zu (c): Es gilt für die Folge der Partialsummen:

$$s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = 1 - \frac{1}{n+1} \to 1 \text{ für } n \to \infty\,.$$

□

[5] Für alle $n \in \mathbb{N}$ ist das $n+1$–te Glied der harmonischen Reihe gerade das harmonische Mittel des n–ten und des $n+2$–ten Glieds.

[6] Für alle $n \in \mathbb{N}_0$ ist das $n+1$–te Glied der geometrischen Reihe gerade das geometrische Mittel des n–ten und des $n+2$–ten Glieds.

Die folgenden Rechenregeln für unendliche Reihen erhält man durch Anwendung der Grenzwertrechenregeln aus Lemma 3.13 auf die Folgen der Partialsummen.

3.40. Lemma. (a) *Seien* $\sum_{n=m}^{\infty} a_n$ *und* $\sum_{n=m}^{\infty} b_n$ *zwei konvergente Reihen in* $\mathbb{R}^N$. *Dann ist auch die Reihe* $\sum_{n=m}^{\infty}(a_n + b_n)$ *konvergent und es gilt:*

$$\sum_{n=m}^{\infty}(a_n + b_n) = \sum_{n=m}^{\infty} a_n + \sum_{n=m}^{\infty} b_n \,.$$

Für alle $\lambda \in \mathbb{R}$ *ist auch die Reihe* $\sum_{n=m}^{\infty} \lambda a_n$ *konvergent und es gilt*

$$\sum_{n=m}^{\infty} \lambda a_n = \lambda \sum_{n=m}^{\infty} a_n \,.$$

(b) *Seien* $\sum_{n=m}^{\infty} z_n$ *und* $\sum_{n=m}^{\infty} w_n$ *zwei konvergente Reihen in* $\mathbb{C}$. *Dann ist auch die Reihe* $\sum_{n=m}^{\infty}(z_n + w_n)$ *konvergent und es gilt:*

$$\sum_{n=m}^{\infty}(z_n + w_n) = \sum_{n=m}^{\infty} z_n + \sum_{n=m}^{\infty} w_n \,.$$

Für alle $\lambda \in \mathbb{C}$ *ist auch die Reihe* $\sum_{n=m}^{\infty} \lambda z_n$ *konvergent und es gilt*

$$\sum_{n=m}^{\infty} \lambda z_n = \lambda \sum_{n=m}^{\infty} z_n \,.$$

Auch das Cauchy–Kriterium für unendliche Reihen erhält man aus dem für Folgen (Satz 3.23).

3.41. Satz (Cauchy–Kriterium für unendliche Reihen). *Sei* $(x_n)_{n=m}^{\infty}$ *eine Folge in* $\mathbb{R}^N$. *Dann sind die folgenden Aussagen äquivalent:*

(a) *Die Reihe* $\sum_{n=m}^{\infty} x_n$ *ist konvergent in* $\mathbb{R}^N$.

(b) $\forall \varepsilon > 0 \quad \exists m \leq n_0 \in \mathbb{Z} \quad \forall n, k \in \mathbb{Z}$ *mit* $n_0 \leq k \leq n$: $\quad \left|\sum_{j=k}^{n} x_j\right| < \varepsilon$.

(c) $\forall \varepsilon > 0 \quad \exists m \leq n_0 \in \mathbb{Z} \quad \forall n_0 \leq n \in \mathbb{Z} \quad \forall p \in \mathbb{N}_0 : \quad \left|\sum_{j=n}^{n+p} x_j\right| < \varepsilon$.

(d) *Die Reihe* $\sum_{n=m}^{\infty} x_n$ *ist komponentenweise konvergent, d.h. für* $j = 1, \ldots, N$ *ist die Reihe* $\sum_{n=m}^{\infty} x_{j,n}$ *der* j*–ten Komponenten in* $\mathbb{R}$ *konvergent.*

Insbesondere gilt dieses Kriterium natürlich in den wichtigen Spezialfällen $N = 1$ für $\mathbb{R}$ und $N = 2$ für $\mathbb{C}$.

3.42. Folgerung. *Ist* $\sum_{n=m}^{\infty} x_n$ *eine konvergente Reihe in* $\mathbb{R}^N$, *so ist die Folge* $(x_n)_{n=m}^{\infty}$ *ihrer Reihenglieder eine Nullfolge, d.h. es gilt* $|x_n| \to 0$ *für* $n \to \infty$.

Beweis. Sei $\varepsilon > 0$ beliebig. Nach Satz 3.41 gibt es ein $m \leq n_0 \in \mathbb{Z}$ mit

$$\forall n_0 \leq n \in \mathbb{Z} \quad \forall p \in \mathbb{N}_0 : \qquad \left|\sum_{j=n}^{n+p} x_j\right| < \varepsilon \,.$$

Speziell mit $p = 0$ erhalten wir $|x_n| < \varepsilon$ für alle ganzen Zahlen $n \geq n_0$. Also gilt $|x_n| \to 0$ für $n \to \infty$. □

3.43. Bemerkungen. (a) Wie das Beispiel 3.39 (a) der harmonischen Reihe zeigt, ist die Bedingung, daß die Reihenglieder eine Nullfolge bilden, notwendig aber nicht hinreichend für die Konvergenz einer Reihe.

(b) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ ist für $q \in \mathbb{K}$ mit $|q| \geq 1$ divergent, denn dann ist $|q|^n \geq 1$ für alle $n \in \mathbb{N}_0$ und damit $(q^n)_{n=0}^{\infty}$ keine Nullfolge.

Durch Anwendung des Monotoniekriteriums 3.11 für Folgen und des Lemmas 3.7 auf die Partialsummen erhalten wir das Monotoniekriterium für Reihen:

3.44. Satz (Monotoniekriterium). *Eine Reihe $\sum_{n=m}^{\infty} x_n$ mit $0 \leq x_n \in \mathbb{R}$ für alle $m \leq n \in \mathbb{Z}$ konvergiert genau dann, wenn die Folge der Partialsummen nach oben beschränkt ist.*

3.45. Satz (Leibniz–Kriterium[7] für alternierende Reihen). *Sei $(x_n)_{n=0}^{\infty}$ eine monoton fallende Folge nicht negativer reeller Zahlen mit $x_n \to 0$ für $n \to \infty$. Dann ist die Reihe $\sum_{n=0}^{\infty}(-1)^n x_n$ konvergent in $\mathbb{R}$.*

Beweis. Für die Partialsummen $s_n := \sum_{k=0}^{n}(-1)^k x_k$ gilt für alle $p \in \mathbb{N}_0$:

$$s_{2(p+1)} - s_{2p} = x_{2(p+1)} - x_{2p+1} \leq 0$$

und

$$s_{2p+3} - s_{2p+1} = x_{2p+2} - x_{2p+3} \geq 0,$$

da die Folge $(x_n)_{n=0}^{\infty}$ monoton fallend ist. Die Folge $(s_{2p})_{p=0}^{\infty}$ ist also monoton fallend und $(s_{2p+1})_{p=0}^{\infty}$ ist monoton wachsend. Ferner gilt für alle $p \in \mathbb{N}_0$:

$$s_{2p+1} - s_{2p} = -x_{2p+1} \leq 0$$

und daher

$$s_1 \leq s_{2p+1} \leq s_{2p} \leq s_0 \,.$$

Nach dem Monotoniekriterium 3.11 für Folgen existieren also die Grenzwerte

$$S := \lim_{p\to\infty} s_{2p} = \inf_{p\in\mathbb{N}_0} s_{2p} \quad \text{und} \quad S_* := \lim_{p\to\infty} s_{2p+1} = \sup_{p\in\mathbb{N}_0} s_{2p+1}$$

in $\mathbb{R}$. Wegen

$$S - S_* = \lim_{p\to\infty} s_{2p} - \lim_{p\to\infty} s_{2p+1} = \lim_{p\to\infty}(s_{2p} - s_{2p+1}) = \lim_{p\to\infty} x_{2p+1} = 0$$

folgt $S = S_*$. Sei nun $\varepsilon > 0$ beliebig. Dann gibt es $p_0, p_1 \in \mathbb{N}_0$, so daß

$$\forall p \geq p_0 : S - \varepsilon < s_{2p+1} \leq S\,, \qquad \forall p \geq p_1 : S \leq s_{2p} < S + \varepsilon\,.$$

Damit ist gezeigt, daß für alle $n \geq n_0 := \max\{2p_0+1, 2p_1\}$ aus $\mathbb{N}_0$ gilt: $|s_n - S| < \varepsilon$. Also ist die Reihe konvergent und es gilt $\sum_{n=0}^{\infty}(-1)^n x_n = S$. □

3.46. Beispiele. (a) Die *alternierende harmonische Reihe* $\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}$ konvergiert. Wir werden später sehen, daß der Grenzwert $\log 2$ ist.

(b) Der Reihengrenzwert $\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$ existiert. Man kann zeigen, daß er mit $\pi/4$ übereinstimmt. Dieser Wert wurde schon von Leibniz angegeben.

[7] Gottfried Wilhelm von Leibniz (1.7.1646–14.11.1716)

3.47. Definition. Eine Reihe $\sum_{n=m}^{\infty} x_n$ in $\mathbb{R}^N$ heißt *absolut konvergent*, falls die Reihe der euklidischen Beträge $\sum_{n=m}^{\infty} |x_n|$ in $\mathbb{R}$ konvergent ist.

Das Beispiel der alternierenden harmonischen Reihe zeigt, daß es konvergente Reihen gibt, die nicht absolut konvergent sind. Andererseits gilt:

3.48. Satz. *Jede absolut konvergente Reihe $\sum_{n=m}^{\infty} x_n$ in $\mathbb{R}^N$ ist auch konvergent und es gilt*

$$\Big|\sum_{n=m}^{\infty} x_n\Big| \leq \sum_{n=m}^{\infty} |x_n|$$

Beweis. Da die Reihe $\sum_{n=m}^{\infty} |x_n|$ konvergiert, gibt es nach Satz 3.41 zu beliebig vorgegebenem $\varepsilon > 0$ ein $m \leq n_0 \in \mathbb{Z}$ mit

$$\forall n_0 \leq n \in \mathbb{Z} \quad \forall p \in \mathbb{N}_0 : \qquad \sum_{k=n}^{n+p} |x_k| < \varepsilon\,.$$

Mit der Dreiecksungleichung folgt also

$$\forall n_0 \leq n \in \mathbb{Z} \quad \forall p \in \mathbb{N}_0 : \qquad \Big|\sum_{k=n}^{n+p} x_k\Big| \leq \sum_{k=n}^{n+p} |x_k| < \varepsilon\,.$$

Nach dem Cauchy-Kriterium 3.41 ist also die Reihe $\sum_{n=m}^{\infty} x_n$ konvergent in $\mathbb{R}^N$.

Ferner gilt nach Lemma 3.6 und Lemma 3.13 (d):

$$\Big|\sum_{k=m}^{\infty} x_k\Big| = \lim_{n\to\infty} \Big|\sum_{k=m}^{n} x_k\Big| \leq \lim_{n\to\infty} \sum_{k=m}^{n} |x_k| = \sum_{k=m}^{\infty} |x_k|.$$

□

3.49. Satz (Majorantenkriterium). *Sei $\sum_{n=m}^{\infty} c_n$ eine in $\mathbb{R}$ konvergente Reihe mit $c_n \geq 0$ für alle $m \leq n \in \mathbb{Z}$. Ist $\sum_{n=m}^{\infty} x_n$ eine Reihe in $\mathbb{R}^N$ mit $|x_n| \leq c_n$ für alle $m \leq n \in \mathbb{Z}$, so ist $\sum_{n=m}^{\infty} x_n$ absolut konvergent und damit auch konvergent in $\mathbb{R}^N$ und es gilt:*

$$\Big|\sum_{n=m}^{\infty} x_n\Big| \leq \sum_{n=m}^{\infty} |x_n| \leq \sum_{n=m}^{\infty} c_n\,.$$

Man nennt die Reihe $\sum_{n=m}^{\infty} c_n$ dann eine konvergente Majorante zu $\sum_{n=m}^{\infty} x_n$.

Beweis. Für $m \leq n \in \mathbb{Z}$ gilt:

$$\sum_{k=m}^{n} |x_k| \leq \sum_{k=m}^{n} c_k \leq \sum_{k=m}^{\infty} c_k.$$

Nach dem Monotoniekriterium ist $\sum_{k=m}^{\infty} |x_k|$ konvergent und $\sum_{k=m}^{\infty} |x_k| \leq \sum_{k=m}^{\infty} c_k$. Der Rest der Behauptungen folgt nun mit Satz 3.48. □

3.50. Beispiel. Die Reihe $\sum_{n=1}^{\infty} n^{-p}$ ist für alle (zunächst noch rationalen) $p \geq 2$ konvergent.

Beweis. Nach Beispiel 3.39 ist die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ und damit auch $\sum_{k=1}^{\infty} \frac{2}{k(k+1)}$ konvergent. Für alle $p \geq 2$ und alle $k \in \mathbb{N}$ gilt

$$\frac{1}{k^p} \leq \frac{1}{k^2} \leq \frac{2}{k(k+1)}\,.$$

$\sum_{k=1}^{\infty} \frac{2}{k(k+1)}$ ist also eine konvergente Majorante zu $\sum_{n=1}^{\infty} n^{-p}$ für $p \geq 2$. □

Später werden wir sehen, daß $\sum_{n=1}^{\infty} n^{-p}$ für alle reellen $p > 1$ konvergiert.

3.51. Bemerkung. Aus dem Majorantenkriterium 3.49 erhält man das folgende Divergenzkriterium: Sei $\sum_{n=m}^{\infty} a_n$ eine divergente Reihe in $\mathbb{R}$ mit $a_n \geq 0$ für alle $m \leq n \in \mathbb{Z}$. Ist $\sum_{n=m}^{\infty} c_n$ eine Reihe in $\mathbb{R}$ mit $a_n \leq c_n$ für alle $m \leq n \in \mathbb{Z}$, so ist auch $\sum_{n=m}^{\infty} c_n$ divergent (andernfalls wäre $\sum_{n=m}^{\infty} c_n$ ja eine konvergente Majorante zu der divergenten Reihe $\sum_{n=m}^{\infty} a_n$ im Widerspruch zum Majorantenkriterium). Man nennt dann die Reihe $\sum_{n=m}^{\infty} a_n$ auch eine *divergente Minorante* zu $\sum_{n=m}^{\infty} c_n$.

3.52. Beispiele. (a) Für $0 < p \leq 1$ ist $\sum_{n=1}^{\infty} n^{-p}$ divergent, denn die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{2n}$ ist eine divergente Minorante.

(b) Die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^2+1}$ ist divergent, denn $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ ist eine divergente Minorante.

3.53. Bemerkung. Seien $p, q \in \mathbb{Z}$ mit $p < q$ und sei $\sum_{n=p}^{\infty} x_n$ eine Reihe in $\mathbb{R}^N$. Genau dann ist die Reihe $\sum_{n=p}^{\infty} x_n$ konvergent (bzw. absolut konvergent), wenn die Reihe $\sum_{n=q}^{\infty} x_n$ konvergiert (bzw. absolut konvergiert). In diesem Fall gilt $\sum_{n=p}^{\infty} x_n = \sum_{n=p}^{q-1} x_n + \sum_{n=q}^{\infty} x_n$. Das Verhalten von endlich vielen Gliedern der Folge $(x_n)_{n=p}^{\infty}$ hat keine Auswirkungen auf das Konvergenzverhalten der zugehörigen Reihe.

3.54. Satz (Quotientenkriterium). *Sei $(x_n)_{n=m}^{\infty}$ eine Folge in $\mathbb{R}^N$, für die ein $n_0 \geq m$ in $\mathbb{Z}$ existiert mit $x_n \neq 0$ für alle $n \geq n_0$.*

(a) *Gibt es ein $q \in (0,1)$, so daß*

$$\forall n_0 \leq n \in \mathbb{Z}: \qquad \frac{|x_{n+1}|}{|x_n|} \leq q\,, \tag{3.5}$$

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

(b) *Gilt $\forall n_0 \leq n \in \mathbb{Z}: \quad \dfrac{|x_{n+1}|}{|x_n|} \geq 1$, so ist die Reihe $\sum_{n=m}^{\infty} x_n$ divergent.*

Beweis. (a) Für $k \geq n_1 := \max\{0, n_0\}$ folgt durch vollständige Induktion

$$|x_k| \leq q^{k-n_1}|x_{n_1}| = q^k \frac{|x_{n_1}|}{q^{n_1}}\,. \tag{3.6}$$

Wegen $0 < q < 1$ konvergiert (nach Beispiel 3.39) die geometrische Reihe $\sum_{k=0}^{\infty} q^k$, also (nach Lemma 3.40 und Bemerkung 3.53) auch die Reihe $\sum_{k=n_1}^{\infty} q^k \frac{|x_{n_1}|}{q^{n_1}}$. Wegen (3.6) ist $\sum_{k=n_1}^{\infty} x_k$ nach dem Majorantenkriterium 3.49 und damit nach 3.53 auch die Reihe $\sum_{n=m}^{\infty} x_n$ absolut konvergent.

(b) Durch vollständige Induktion sieht man

$$\forall n_0 \leq n \in \mathbb{Z}: \qquad |x_n| \geq |x_{n_0}| > 0\,.$$

Die Folge $(x_n)_{n=m}^{\infty}$ ist also keine Nullfolge. Nach Folgerung 3.42 ist $\sum_{n=m}^{\infty} x_n$ also divergent. $\square$

Bei der Anwendung des Quotientenkriteriums ist darauf zu achten, daß q unabhängig von n sein muß. Ist etwa $x_n = n^{-1}$ und $y_n = n^{-2}$ für alle $n \in \mathbb{N}$, so ist

$$\forall n \in \mathbb{N}: \qquad \frac{|x_{n+1}|}{|x_n|} = \frac{n}{n+1} < 1 \text{ und } \frac{|y_{n+1}|}{|y_n|} = \frac{n^2}{(n+1)^2} < 1$$

aber $\sum_{n=1}^{\infty} n^{-1}$ ist divergent und $\sum_{n=1}^{\infty} n^{-2}$ ist konvergent.

3.55. Bemerkung. Die Forderung (3.5) in Satz 3.54 (a) ist erfüllt, wenn gilt

$$\limsup_{n\to\infty} \frac{|x_{n+1}|}{|x_n|} < 1 . \tag{3.7}$$

Insbesondere ist sie also erfüllt, wenn der Grenzwert $\lim_{n\to\infty} \frac{|x_{n+1}|}{|x_n|}$ existiert und echt kleiner als 1 ist.

Beweis. Sei also (3.7) erfüllt. Dann gibt es ein $q \in \mathbb{R}$ mit $\limsup_{n\to\infty} \frac{|x_{n+1}|}{|x_n|} < q < 1$. Nach Definition des Limes superior existiert ein $m \leq k_0 \in \mathbb{Z}$ mit

$$\sup\left\{\frac{|x_{n+1}|}{|x_n|}\,;\, n \geq k_0\right\} < q\,,$$

so daß also (3.5) für $k_0 \leq n \in \mathbb{Z}$ erfüllt ist. □

3.56. Satz (Einführung von exp, sin und cos). (a) *Für alle $z \in \mathbb{C}$ ist die Reihe $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ absolut konvergent. Durch*

$$\exp(z) := \sum_{n=0}^{\infty} \frac{1}{n!} z^n \qquad \textit{für } z \in \mathbb{C}$$

wird also eine komplexwertige Funktion $\exp : \mathbb{C} \to \mathbb{C}$ *definiert. Für alle $x \in \mathbb{R}$ ist auch* $\exp(x) \in \mathbb{R}$*. Es ist* $\exp(0) = 1$.

(b) *Für alle $z \in \mathbb{C}$ sind die Reihen*

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} \quad \textit{und} \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}$$

absolut konvergent. Durch

$$\cos(z) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} \quad \textit{und} \quad \sin(z) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \quad \textit{für } z \in \mathbb{C}$$

werden also komplexwertige Funktionen $\cos : \mathbb{C} \to \mathbb{C}$ *und* $\sin : \mathbb{C} \to \mathbb{C}$ *definiert. Für alle $x \in \mathbb{R}$ sind auch* $\cos(x)$ *und* $\sin(x)$ *reell. Es ist* $\cos(0) = 1$ *und* $\sin(0) = 0$.

(c) *Für alle $z \in \mathbb{C}$ gilt* $\cos(-z) = \cos(z)$, *und* $\sin(-z) = -\sin(z)$.

(d) *Für alle $z \in \mathbb{C}$ gilt* $\exp(iz) = \cos(z) + i\sin(z)$.

(e) *Für alle $z \in \mathbb{C}$ gilt*

$$\cos(z) = \frac{1}{2}\left(\exp(iz) + \exp(-iz)\right) \quad \textit{und} \quad \sin(z) = \frac{1}{2i}\left(\exp(iz) - \exp(-iz)\right) .$$

Beweis. (a) Für $z = 0$ ist die Konvergenz klar und es gilt $\exp(0) = 1$ (wobei wir 0^0 als 1 vereinbart haben). Für $0 \neq z \in \mathbb{C}$ gilt mit $a_n := \frac{1}{n!} z^n$: Es ist $a_n \neq 0$ für alle $n \in \mathbb{N}_0$ und

$$\frac{|a_{n+1}|}{|a_n|} = \left|\frac{z^{n+1} n!}{(n+1)! z^n}\right| = \frac{|z|}{n+1} \to 0 \quad \text{für } n \to \infty .$$

Also ist die Reihe $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ nach dem Quotientenkriterium und Bemerkung 3.55 absolut konvergent. Ist $z = x \in \mathbb{R}$, so sind für alle $n \in \mathbb{N}$ die Partialsummen $\sum_{n=0}^{k} \frac{1}{n!} x^n$ reell, also auch deren Grenzwert für $k \to \infty$.

(b) Für $z = 0$ ist die Konvergenz trivial und man hat $\cos(0) = 1$, $\sin(0) = 0$. Sei nun $z \neq 0$. Wir setzen

$$a_n := \frac{(-1)^n}{(2n)!} z^{2n} \quad \text{und} \quad b_n := \frac{(-1)^n}{(2n+1)!} z^{2n+1}, \qquad n \in \mathbb{N}_0.$$

Dann ist $a_n \neq 0$, $b_n \neq 0$ für alle $n \in \mathbb{N}_0$ und es gilt

$$\frac{|a_{n+1}|}{|a_n|} = \left|\frac{z^{2(n+1)}(2n)!}{(2(n+1))!z^{2n}}\right| = \frac{|z|^2}{(2n+1)(2n+2)} \to 0 \quad \text{für } n \to \infty$$

sowie

$$\frac{|b_{n+1}|}{|b_n|} = \left|\frac{z^{2n+3}(2n+1)!}{(2n+3))!z^{2n+1}}\right| = \frac{|z|^2}{(2n+2)(2n+3)} \to 0 \quad \text{für } n \to \infty\,.$$

Wie in (a) folgt die absolute Konvergenz der Reihen nun mit dem Quotientenkriterium und Bemerkung 3.55 und wie in (a) sieht man auch, daß $\sin(x)$ und $\cos(x)$ für reelles x ebenfalls reell sind.

(c) folgt unmittelbar aus den Reihendarstellungen der Cosinus– und der Sinusfunktion.

(d) Für alle $z \in \mathbb{C}$ gilt (mit Lemma 3.40 (b))

$$\begin{aligned}
\cos(z) + i\sin(z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \\
&= \lim_{k\to\infty} \sum_{n=0}^{k} \left(\frac{(-1)^n}{(2n)!} z^{2n} + i \frac{(-1)^n}{(2n+1)!} z^{2n+1} \right) \\
&= \lim_{k\to\infty} \sum_{n=0}^{k} \left(\frac{(iz)^{2n}}{(2n)!} + \frac{(iz)^{2n+1}}{(2n+1)!} z^{2n+1} \right) \\
&= \lim_{k\to\infty} \sum_{n=0}^{2k+1} \frac{1}{n!} (iz)^n \\
&= \exp(iz)\,.
\end{aligned}$$

(e) Nach (d) und (c) gilt

$$\begin{aligned}
\exp(iz) + \exp(-iz) &= \cos(z) + i\sin(z) + \cos(-z) + i\sin(-z) \\
&= \cos(z) + i\sin(z) + \cos(z) - i\sin(z) \\
&= 2\cos(z)
\end{aligned}$$

und ebenso

$$\begin{aligned}
\exp(iz) - \exp(-iz) &= \cos(z) + i\sin(z) - \cos(-z) - i\sin(-z) \\
&= \cos(z) + i\sin(z) - \cos(z) + i\sin(z) \\
&= 2i\sin(z)\,.
\end{aligned}$$

Hieraus folgt (e). □

3.57. Satz (Wurzelkriterium). *Sei $(x_n)_{n=m}^{\infty}$ eine Folge in $\mathbb{R}^N$.*

(a) *Gibt es ein $q \in [0,1)$ und ein $n_0 \in \mathbb{N}$ mit $n_0 \geq m$, so daß*

$$\forall n_0 \leq n \in \mathbb{N}: \qquad \sqrt[n]{|x_n|} \leq q\,, \tag{3.8}$$

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

(b) *Gibt es ein $n_0 \in \mathbb{N}$ mit $n_0 \geq m$, so daß*

$$\forall n_0 \leq n \in \mathbb{N}: \qquad \sqrt[n]{|x_n|} \geq 1\,, \tag{3.9}$$

so ist die Reihe $\sum_{n=m}^{\infty} x_n$ divergent.

Beweis. (a) Aus (3.8) folgt $|x_n| \leq q^n$ für alle $n \geq n_0$. Wegen $0 \leq q < 1$ konvergiert die geometrische Reihe $\sum_{k=0}^{\infty} q^k$ (nach Beispiel 3.39). Also ist $\sum_{k=n_0}^{\infty} x_k$ nach dem Majorantenkriterium 3.49 und damit nach 3.53 auch die Reihe $\sum_{n=m}^{\infty} x_n$ absolut konvergent.

(b) Wegen (3.9) ist $|x_n| \geq 1$ für alle $n \geq n_0$. Die Folge $(x_n)_{n=m}^{\infty}$ ist also keine Nullfolge. Nach Folgerung 3.42 ist $\sum_{n=m}^{\infty} x_n$ daher divergent. □

Wie bei der Anwendung des Quotientenkriteriums ist darauf zu achten, daß q unabhängig von n sein muß. Ist etwa $x_n = n^{-1}$ und $y_n = n^{-2}$ für alle $n \in \mathbb{N}$, so ist

$$\forall n \in \mathbb{N}: \qquad \sqrt[n]{|x_n|} = \frac{1}{\sqrt[n]{n}} < 1 \quad \text{und} \ \sqrt[n]{|y_n|} = \frac{1}{\sqrt[n]{n^2}} < 1$$

aber $\sum_{n=1}^{\infty} n^{-1}$ ist divergent und $\sum_{n=1}^{\infty} n^{-2}$ ist konvergent.

Analog wie Bemerkung 3.55 zeigt man:

3.58. Bemerkung. Die Forderung (3.8) in Satz 3.57 (a) ist erfüllt, wenn gilt

$$\limsup_{n\to\infty} \sqrt[n]{|x_n|} < 1\,.$$

Insbesondere ist sie also erfüllt, wenn der Grenzwert $\lim_{n\to\infty} \sqrt[n]{|x_n|}$ existiert und echt kleiner als 1 ist.

5. Dezimalbruchentwicklung und Überabzählbarkeit von $\mathbb{R}$

3.59. Satz (d-adische Entwicklung reeller Zahlen). *Sei $2 \leq d \in \mathbb{N}$. Wir nennen die Elemente von $M_d := \{k \in \mathbb{N}_0\,;\, k \leq d-1\}$ d–adische Ziffern.*

(a) *Sei $m \in \mathbb{N}_0$ und sei $(a_n)_{n=-m}^{\infty}$ eine beliebige Folge d–adischer Ziffern. Dann konvergiert $\pm\sum_{n=-m}^{\infty} a_n d^{-n}$ gegen eine reelle Zahl.*

(b) *Ist umgekehrt $x \in \mathbb{R}$ gegeben, so gibt es genau eine Folge $(a_n)_{n=-m}^{\infty}$ d–adischer Ziffern mit den folgenden Eigenschaften:*
   (i) $x = \pm\sum_{n=-m}^{\infty} a_n d^{-n}$.
   (ii) $m = \min\{n \in \mathbb{N}_0\,;\, |x| < d^{n+1}\}$.
   (iii) *Mit* $x_n := \sum_{k=-m}^{n} a_k d^{-k}$ *gilt* $x_n \leq |x| < x_n + d^{-n}$.

   Wir nennen die Darstellung $x = \pm\sum_{n=-m}^{\infty} a_n d^{-n}$ dann die d–adische Entwicklung von x. Man schreibt üblicherweise $x = \pm a_{-m} \ldots a_0, a_1 a_2 a_3 \ldots$

Beweis. (a) Es ist $|a_n d^{-n}| = a_n d^{-n} \leq (d-1)d^{-n}$ für alle $-m \leq n \in \mathbb{Z}$. Wegen $0 < d^{-1} < 1$ ist also $\sum_{n=-m}^{\infty}(d-1)d^{-n}$ eine konvergente Majorante für $\pm\sum_{n=-m}^{\infty} a_n d^{-n}$. Nach dem Majorantenkriterium folgt die Konvergenz der Reihe.

(b) Es genügt die Behauptung für $x \geq 0$ zu zeigen. Wegen $d^n \to \infty$ für $n \to \infty$ gibt es ein $k \in \mathbb{N}_0$ mit $x < d^{k+1}$. Insbesondere existiert $m := \min\{n \in \mathbb{N}_0\,;\, |x| < d^{n+1}\}$. Wir zeigen nun induktiv, daß es genau eine Folge $(a_n)_{n=-m}^{\infty}$ d–adischer Ziffern gibt mit (iii).

Induktionsanfang: Es ist $0 = 0 \cdot d^m < 1 \cdot d^m < \cdots < (d-1)d^m < d^{m+1}$. Wegen $0 \leq x < d^{m+1}$ gibt es genau ein $a_{-m} \in M_d$ mit $x_{-m} := a_{-m}d^m \leq x < (a_{-m}+1)d^m = x_{-m} + d^m$.

Induktionsvoraussetzung: Seien nun schon eindeutig bestimmte $a_{-m}, \ldots, a_\nu \in M_d$ so gefunden, daß (iii) für $-m \leq n \leq \nu$ erfüllt ist.

Induktionsschluß: Nach Induktionsvoraussetzung gilt

$$x_\nu = \sum_{n=-m}^{\nu} a_n d^{-n} \leq x < x_\nu + d^{-\nu}.$$

Nun ist

$$x_\nu = x_\nu + 0 \cdot d^{-\nu-1} < x_\nu + 1 \cdot d^{-\nu-1} < \cdots < x_\nu + (d-1)d^{-\nu-1} < x_\nu + d^{-\nu}\,.$$

Es gibt also genau ein $a_{\nu+1} \in M_d$ mit

$$x_{\nu+1} := x_\nu + a_{\nu+1} d^{-\nu-1} \le x < x_\nu + (a_{\nu+1}+1) d^{-\nu-1} = x_{\nu+1} + d^{-\nu-1}.$$

Damit ist die induktive Konstruktion vollendet und es ist gezeigt, daß die d–adischen Ziffern a_n für $-m \le n \in \mathbb{Z}$ eindeutig durch (iii) bestimmt sind.

Nach Konstruktion ist $|x - x_n| < d^{-n}$ für alle $-m \le n \in \mathbb{Z}$ erfüllt. Damit folgt $\lim_{n\to\infty} |x - x_n| = 0$, d.h. es gilt (i). □

Für $d = 10$ haben wir die aus der Schule bekannte Dezimalbruchentwicklung wiedergefunden. In der Informatik sind auch die Basen $d = 2, 8, 16$ gebräuchlich. Die Babylonier rechneten im *Sexagesimalsystem* mit $d = 60$, was heute auch noch in der Zeitmessung (Stunden, Minuten, Sekunden) gemacht wird.

Satz 3.59 zeigt auch, daß sich jede reelle Zahl x beliebig genau durch die rationalen Zahlen x_n approximieren läßt. Wir wollen zeigen, daß es dennoch „mehr“ irrationale Zahlen als rationale Zahlen gibt. Um dies zu präzisieren, führen wir die folgenden Begriffe ein:

3.60. Definition. Zwei nicht leere Mengen M_1 und M_2 heißen *von gleicher Mächtigkeit*, falls es eine bijektive Abbildung $\varphi : M_1 \to M_2$ gibt. Wir schreiben dann $|M_1| = |M_2|$. Wir setzen $|\emptyset| := 0$. Eine nicht leere Menge M heißt

- *endlich*, falls $|M| = |\{n \in \mathbb{N}\,;\, n \le k\}|$ für ein $k \in \mathbb{N}$ gilt. Wir schreiben dann $|M| = k$. Auch die leere Menge $\emptyset$ wird zu den endlichen Mengen gerechnet.
- *abzählbar unendlich*, falls $|M| = \aleph_0 := |\mathbb{N}|$ gilt. (Der Buchstabe $\aleph$ (*Aleph*) ist der erste Buchstabe des hebräischen Alphabets).
- *abzählbar*, falls M endlich oder abzählbar unendlich ist.
- *überabzählbar*, falls M nicht abzählbar ist.

3.61. Satz. (a) *Jede Teilmenge einer abzählbaren Menge ist abzählbar.*
(b) $\mathbb{N} \times \mathbb{N}$ *ist abzählbar unendlich.*
(c) *Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.*

Beweis. (a) Die Teilmengen von endlichen Mengen sind wieder endliche Mengen und daher abzählbar. Sei nun M eine abzählbar unendliche Menge und sei $L \subseteq M$ beliebig. Ist L endlich, so ist L insbesondere abzählbar. Sei nun L eine nicht endliche Teilmenge von M. Da M abzählbar unendlich ist, gibt es eine Bijektion $\varphi : M \to \mathbb{N}$. Wir konstruieren nun induktiv eine Folge $(\psi(k))_{k=1}^\infty$ in L, welche für alle $k \in \mathbb{N}$ die folgenden Bedingungen erfüllt:

(i) Für alle $j \in \mathbb{N}$ mit $j < k$ gilt $\varphi(\psi(j)) < \varphi(\psi(k))$.
(ii) Für alle $u \in L \setminus \{\psi(j)\,;\, 1 \le j \le k\}$ gilt $\varphi(\psi(k)) < \varphi(u)$.
(iii) $\{u \in L\,;\, \varphi(u) \le \varphi(\psi(k))\} = \{\psi(j)\,;\, 1 \le j \le k\}$.

Mit $\psi(1) := \varphi^{-1}(\min\{\varphi(L)\})$ sind diese drei Bedingungen für $k = 1$ erfüllt. Das Minimum existiert hierbei, da $\{\varphi(L)\}$ eine nicht leere Teilmenge von $\mathbb{N}$ ist.

Seien nun schon für ein $n \in \mathbb{N}$ die Elemente $\psi(1), \dots, \psi(n)$ so konstruiert, daß die Bedingungen (i)–(iii) für alle $k \in \{1, \dots, n\}$ erfüllt sind. Dann ist

$$L_n := L \setminus \{\psi(j)\,;\, 1 \le j \le n\} \ne \emptyset\,,$$

da L eine nicht endliche Menge ist. Daher existiert $\psi(n+1) := \varphi^{-1}(\min\{\varphi(L_n)\})$. Man überprüft leicht, daß die Bedingungen (i)–(iii) mit dieser Wahl von $\psi(n+1)$ auch für $k = n+1$ erfüllt sind. Damit ist die induktive Konstruktion vollzogen.

Wegen (i) ist die Abbildung $\psi : \mathbb{N} \to L$, $k \mapsto \psi(k)$, injektiv. Ist $u \in L$ beliebig, so folgt $\varphi(u) \le \varphi(\psi(\varphi(u)))$ (da die Abbildung $k \mapsto \varphi(\psi(k))$ von $\mathbb{N}$ in sich wegen (i) streng

$$\begin{array}{cccccccccccccc}
(1,1) & & (1,2) & & (1,3) & & (1,4) & & (1,5) & & (1,6) & & (1,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(2,1) & & (2,2) & & (2,3) & & (2,4) & & (2,5) & & (2,6) & & (2,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(3,1) & & (3,2) & & (3,3) & & (3,4) & & (3,5) & & (3,6) & & (3,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(4,1) & & (4,2) & & (4,3) & & (4,4) & & (4,5) & & (4,6) & & (4,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(5,1) & & (5,2) & & (5,3) & & (5,4) & & (5,5) & & (5,6) & & (5,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(6,1) & & (6,2) & & (6,3) & & (6,4) & & (6,5) & & (6,6) & & (6,7) & \dots \\
& \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \\
(7,1) & & (7,2) & & (7,3) & & (7,4) & & (7,5) & & (7,6) & & (7,7) & \dots \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

ABBILDUNG 1. Abzählung der Elemente von $\mathbb{N} \times \mathbb{N}$

monoton wachsend ist). Wegen (iii) gibt es also ein $j \in \{1, \dots, \varphi(u)\}$ mit $\psi(j) = u$. Damit ist gezeigt, daß $\psi : \mathbb{N} \to L$ bijektiv ist. L ist also abzählbar unendlich und damit insbesondere abzählbar.

(b) Bei der Darstellung der Elemente von $\mathbb{N}\times\mathbb{N}$ nach dem in Abbildung 1 angedeuteten Schema numerieren wir die Elemente den eingezeichneten Diagonalen folgend: Die j–te Diagonale enthält die j Elemente $(1, j), (2, j-1), \dots, (j, 1)$. Für $k = 1, \dots, j$ erhält das k–te Element der j–ten Diagonale (also $(k, j+1-k)$) unter Verwendung von Beispiel 1.22 die Nummer

$$\varphi(k, j+1-k) := k + \sum_{p=1}^{j-1} p = \frac{j(j-1)}{2} + k\,.$$

Hierdurch ist eine bijektive Abbildung $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ gegeben, da jede natürliche Zahl n eine eindeutige Darstellung der Form $\dfrac{j(j-1)}{2} + k$ für ein $j \in \mathbb{N}$ und ein $k \in \mathbb{N}$ mit $1 \le k \le j$ besitzt.

(c) Sei also $M = \bigcup_{k=1}^{n} A_k$ mit $n \in \mathbb{N} \cup \{\infty\}$ und abzählbaren Mengen A_k. Wir setzen $B_1 := A_1$ und für $k \in \mathbb{N}$ mit $k+1 \le n$:

$$B_{k+1} := A_{k+1} \setminus \bigcup_{j=1}^{k} A_j\,.$$

Die so definierten Mengen B_k sind paarweise disjunkt und als Teilmengen der abzählbaren Mengen A_k wieder abzählbar. Sie erfüllen (wie man induktiv sieht) für $k \in \mathbb{N}$ mit $k \le n$:

$$\bigcup_{j=1}^{k} B_j = \bigcup_{j=1}^{k} A_j\,.$$

Also hat man

$$M = \bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{n} B_j$$

mit paarweise disjunkten, abzählbaren Mengen B_j, $j \in \mathbb{N}$, $j \leq n$. Ist $B_j \neq \emptyset$, so gibt es also eine Bijektion φ_j von B_j auf eine Teilmenge der Gestalt $N_j := \{p \in \mathbb{N}\,;\, p \leq m_j\}$ für ein $m_j \in \mathbb{N} \cup \{\infty\}$. Zu jedem $x \in M$ gibt es also genau ein $j(x) \in \{j \in \mathbb{N}\,;\, k \leq n\}$ mit $x \in B_{j(x)}$. Durch $x \mapsto \Phi(x) := (j(x), \varphi_{j(x)}(x))$ ist nun eine injektive Abbildung von M nach $\mathbb{N} \times \mathbb{N}$ gegeben. Als Teilmenge der abzählbaren Menge $\mathbb{N} \times \mathbb{N}$ ist $\Phi(M)$ abzählbar. Es gibt also ein $m \in \mathbb{N} \cup \{\infty\}$ und eine Bijektion $\Psi : \Phi(M) \to \{p \in \mathbb{N}\,;\, p \leq m\}$. Die Hintereinanderausführung $\Psi \circ \Phi : M \to \{p \in \mathbb{N}\,;\, p \leq m\}$ ist also auch eine Bijektion. Also ist M abzählbar. □

3.62. Folgerung. *Die Mengen* $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-n\,;\, n \in \mathbb{N}\}$ *und* $\mathbb{Q} = \bigcup_{q=1}^{\infty} \left\{\frac{p}{q}\,;\, p \in \mathbb{Z}\right\}$ *sind als abzählbare Vereinigungen von abzählbaren Mengen abzählbar.*

3.63. Satz. *Die nicht endlichen Mengen* $\mathbb{C}$, $\mathbb{R}$, $\mathbb{R} \setminus \mathbb{Q}$ *und das Intervall* $(0,1)$ *sind nicht abzählbar.*

Beweis. Wegen $(0,1) \subset \mathbb{R} \subset \mathbb{C}$ und $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$ genügt es nach Lemma 3.61 und Folgerung 3.62 zu zeigen, daß das Intervall $(0,1)$ nicht abzählbar sind.

Annahme: $(0,1)$ ist abzählbar unendlich, d.h. es gibt eine Bijektion $\varphi : \mathbb{N} \to (0,1)$. Wir betrachten die Dezimalbruchentwicklungen der Zahlen $\varphi(n)$ für alle $n \in \mathbb{N}$:

$$\begin{aligned}
\varphi(1) &= 0, a_{1,1}a_{1,2}a_{1,3}a_{1,4}a_{1,5}\ldots \\
\varphi(2) &= 0, a_{2,1}a_{2,2}a_{2,3}a_{2,4}a_{2,5}\ldots \\
\varphi(3) &= 0, a_{3,1}a_{3,2}a_{3,3}a_{3,4}a_{3,5}\ldots \\
&\vdots \\
\varphi(n) &= 0, a_{n,1}a_{n,2}a_{n,3}a_{n,4}a_{n,5}\ldots \\
&\vdots
\end{aligned}$$

Sei nun $x = 0, b_1b_2b_3b_4b_5 \cdots \in (0,1)$ definiert durch

$$b_k := \begin{cases} 0 & \text{falls } a_{k,k} = 1 \\ 1 & \text{falls } a_{k,k} \neq 1. \end{cases}$$

Dann folgt $x \neq \varphi(n)$ für alle $n \in \mathbb{N}$ im Widerspruch zu $x \in (0,1) = \varphi(\mathbb{N})$. Also kann $(0,1)$ doch nicht abzählbar sein. □

6. Umordnung von Reihen

3.64. Beispiel. Wir betrachten die alternierende harmonische Reihe $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$ (vergl. Beispiel 3.46 (a)). Nach dem Leibniz-Kriterium 3.45 ist diese konvergent gegen einen Grenzwert s (der, wie wir später sehen werden, mit $\log 2$ übereinstimmt) und nach dem Beweis zu 3.45 gilt

$$s \geq s_1 = 1 - \frac{1}{2} = \frac{1}{2} > 0. \tag{3.10}$$

Wir definieren eine Abbildung $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ durch:

$$\varphi(n) := \begin{cases} 2m & \text{falls } n = 3m \text{ für ein } m \in \mathbb{N}_0 \\ 4m+1 & \text{falls } n = 3m+1 \text{ für ein } m \in \mathbb{N}_0 \\ 4m+3 & \text{falls } n = 3m+2 \text{ für ein } m \in \mathbb{N}_0 \,. \end{cases}$$

Man überzeugt sich leicht, daß φ bijektiv ist. Wir betrachten nun die Reihe

$$\sum_{n=0}^{\infty} b_n \quad \text{mit } b_n := \frac{(-1)^{\varphi(n)}}{\varphi(n)+1}\,.$$

Diese Reihe hat also die gleichen Glieder wie die Ausgangsreihe, nur in anderer Anordnung.

$$\begin{aligned} \sum_{n=0}^{3m+2} b_n &= \sum_{k=0}^{m} (b_{3k} + b_{3k+1} + b_{3k+2}) \\ &= \sum_{k=0}^{m} \Big(\frac{1}{2k+1} - \frac{1}{4k+2} - \frac{1}{4k+4} \Big) \\ &= \sum_{k=0}^{m} \frac{1}{2k+1} - \frac{1}{2} \sum_{k=0}^{m} \frac{1}{2k+1} - \frac{1}{2} \sum_{k=0}^{m} \frac{1}{2k+2} \\ &= \frac{1}{2} \sum_{k=0}^{m} \frac{1}{2k+1} - \frac{1}{2} \sum_{k=0}^{m} \frac{1}{2k+2} \\ &= \frac{1}{2} \sum_{j=0}^{2m+1} \frac{(-1)^j}{j+1} \to \frac{s}{2} \text{ für } m \to \infty\,. \end{aligned}$$

Wegen $s > 0$ (nach(3.10)) und $b_n \to 0$ für $n \to \infty$ folgt hieraus

$$\sum_{n=0}^{\infty} b_n = \frac{s}{2} \neq s = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}\,.$$

Dieses Beispiel zeigt, daß das allgemeine Kommutativgesetz für unendliche Reihen nicht mehr gilt. Für absolut konvergente Reihen erhalten wir jedoch noch:

3.65. Satz (Umordnungssatz). *Sei $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe (in $\mathbb{R}$, $\mathbb{C}$ oder allgemein in $\mathbb{R}^N$, $N \in \mathbb{N}$) mit dem Reihengrenzwert S. Dann konvergiert auch für jede bijektive Abbildung $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ die umgeordnete Reihe $\sum_{n=0}^{\infty} a_{\varphi(n)}$ gegen S.*

Beweis. Sei also $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ eine beliebige bijektive Abbildung und sei $\varepsilon > 0$ beliebig. Wegen der absoluten Konvergenz der Reihe gibt es ein $n_0 \in \mathbb{N}_0$ mit

$$\sum_{n=n_0+1}^{\infty} |a_n| = \sum_{n=0}^{\infty} |a_n| - \sum_{n=0}^{n_0} |a_n| < \frac{\varepsilon}{2}\,.$$

Da φ bijektiv ist, gibt es ein $n_1 \in \mathbb{N}_0$ mit

$$\{0, 1, \ldots, n_0\} \subseteq \{\varphi(0), \varphi(1), \ldots, \varphi(n_1)\}\,.$$

Für alle $n \in \mathbb{N}$ setzen wir $m(n) := \max_{1 \le k \le n} \varphi(n)$. Dann gilt für alle $n \in \mathbb{N}_0$ mit $n \ge n_1$:

$$\begin{aligned}\left|\sum_{k=0}^{n} a_{\varphi(k)} - S\right| &\le \left|\sum_{k=0}^{n} a_{\varphi(k)} - \sum_{k=0}^{n_0} a_k\right| + \left|\sum_{k=0}^{n_0} a_k - S\right| \\ &\le \sum_{j=n_0+1}^{m(n)} |a_j| + \left|\sum_{k=n_0+1}^{\infty} a_k\right| \\ &\le \sum_{j=n_0+1}^{\infty} |a_j| + \sum_{k=n_0+1}^{\infty} |a_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.\end{aligned}$$

Also konvergiert auch die umgeordnete Reihe $\sum_{n=0}^{\infty} a_{\varphi(n)}$ gegen S. □

3.66. Bemerkung. Man kann zeigen: Ist $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{R}$, welche konvergiert aber nicht absolut konvergent ist, so gibt es zu jedem $s \in \mathbb{R} \cup \{-\infty, \infty\}$ eine bijektive Abbildung $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$ mit $\sum_{n=0}^{m} a_{\varphi(n)} \to s$ für $m \to \infty$. Es gibt auch solche bijektive Abbildungen $\varphi : \mathbb{N}_0 \to \mathbb{N}_0$, für die $\sum_{n=0}^{\infty} a_{\varphi(n)}$ auch nicht im uneigentlichen Sinn konvergiert.

3.67. Satz (Cauchy–Produkt von Reihen). *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei absolut-konvergente Reihen in $\mathbb{R}$ oder $\mathbb{C}$. Für alle $n \in \mathbb{N}_0$ definieren wir*

$$c_n := \sum_{k=0}^{n} a_k b_{n-k}\,.$$

Dann ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt:

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right) \cdot \left(\sum_{n=0}^{\infty} b_n\right)\,. \tag{3.11}$$

Beweis. Seien S_a und S_b die Grenzwerte der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ und für alle $n \in \mathbb{N}_0$ sei

$$s_n^a := \sum_{\nu=0}^{n} a_\nu, \quad s_n^b := \sum_{\nu=0}^{n} b_\nu, \quad \text{und} \quad s_n^c := \sum_{\nu=0}^{n} c_\nu \quad \text{sowie}$$
$$\sigma_n^a := \sum_{\nu=0}^{n} |a_\nu|, \quad \sigma_n^b := \sum_{\nu=0}^{n} |b_\nu|, \quad \text{und} \quad \sigma_n^c := \sum_{\nu=0}^{n} |c_\nu|\,.$$

Wegen der absoluten Konvergenz der Reihen existieren die Grenzwerte

$$\Sigma_a := \lim_{n\to\infty} \sigma_n^a \quad \text{und} \quad \Sigma_b := \lim_{n\to\infty} \sigma_n^b$$

und es gilt $\sigma_n^a \le \Sigma_a$ und $\sigma_n^b \le \Sigma_b$ für alle $n \in \mathbb{N}_0$. Für $n \in \mathbb{N}_0$ sei weiter

$$\Delta_n := \{(j,k) \in \mathbb{N}_0 \times \mathbb{N}_0\,;\, j+k \le n\} \quad \text{und} \quad \Gamma_n := \{(j,k) \in \mathbb{N}_0 \times \mathbb{N}_0\,;\, j \le n, k \le n\}\,.$$

Offensichtlich gilt $\Delta_n \subseteq \Gamma_n$ für alle $n \in \mathbb{N}_0$ sowie $\Gamma_m \subseteq \Delta_n$ für alle $m, n \in \mathbb{N}_0$ mit $2m \le n$. Für alle $n \in \mathbb{N}_0$ gilt nach der Dreiecksungleichung

$$\sigma_n^c = \sum_{\nu=0}^{n} |c_\nu| \le \sum_{\nu=0}^{n}\sum_{k=0}^{\nu} |a_{\nu-k}| \cdot |b_k| = \sum_{(j,k)\in\Delta_n} |a_j| \cdot |b_k| \le \sum_{(j,k)\in\Gamma_n} |a_j| \cdot |b_k| = \sigma_n^a \sigma_n^b \le \Sigma_a \Sigma_b\,.$$

Nach dem Majorantenkriterium 3.49 ist also die Reihe $\sum_{\nu=0}^{\infty} |c_\nu|$ konvergent, d.h. $\sum_{\nu=0}^{\infty} c_\nu$ ist absolut konvergent und damit insbesondere konvergent gegen einen Grenzwert S_c.

Nach den Grenzwertrechenregeln ist auch die Folge $(\sigma_n^a \sigma_n^b)_{n=0}^\infty$ konvergent und damit eine Cauchy-Folge. Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es also ein $n_0 \in \mathbb{N}_0$, so daß für alle $n \geq n_0$ gilt: $0 \leq \sigma_n^a \sigma_n^b - \sigma_{n_0}^a \sigma_{n_0}^b < \varepsilon$. Wegen

$$s_n^c = \sum_{\nu=0}^{n} \sum_{k=0}^{\nu} a_{\nu-k} b_k = \sum_{(j,k)\in\Delta_n} a_j b_k$$

erhalten wir für alle $n \geq 2n_0$:

$$\begin{aligned} |s_n^a s_n^b - s_n^c| &= \left| \sum_{(j,k)\in\Gamma_n} a_j b_k - \sum_{(j,k)\in\Delta_n} a_j b_k \right| = \left| \sum_{(j,k)\in\Gamma_n\setminus\Delta_n} a_j b_k \right| \leq \\ &\leq \sum_{(j,k)\in\Gamma_n\setminus\Delta_n} |a_j b_k| \leq \sum_{(j,k)\in\Gamma_n\setminus\Gamma_{n_0}} |a_j b_k| = \sigma_n^a \sigma_n^b - \sigma_{n_0}^a \sigma_{n_0}^b < \varepsilon\,. \end{aligned}$$

Also konvergiert $s_n^a s_n^b - s_n^c$ gegen 0 für $n \to \infty$. Wegen $s_n^c \to S_c$ und $s_n^a s_n^b \to S_a S_b$ folgt hieraus $S_c = S_a S_b$ und damit (3.11). □

In den Übungen wird gezeigt: Ist

$$a_n := b_n := \frac{(-1)^n}{\sqrt{n+1}} \quad \text{und} \quad c_n := \sum_{k=0}^{n} a_k b_{n-k}$$

für alle $n \in \mathbb{N}_0$, so sind die Reihen $\sum_{n=0}^\infty a_n$ und $\sum_{n=0}^\infty b_n$ konvergent, aber die Reihe $\sum_{n=0}^\infty c_n$ ist divergent. Auf die Voraussetzung der absoluten Konvergenz der Reihen $\sum_{n=0}^\infty a_n$ und $\sum_{n=0}^\infty b_n$ in Satz 3.67 kann also nicht verzichtet werden.

3.68. Satz (Funktionalgleichung der Exponentialfunktion). *Für alle $z, w \in \mathbb{C}$ gilt* $\exp(z+w) = \exp(z)\exp(w)$.

Beweis. Seien $z, w \in \mathbb{C}$ beliebig. Nach Satz 3.56 sind die Reihen $\sum_{n=0}^\infty \frac{1}{n!} z^n$ und $\sum_{n=0}^\infty \frac{1}{n!} w^n$ absolut konvergent. Nach Satz 3.67 gilt also (mit absoluter Konvergenz):

$$\begin{aligned} \exp(z)\exp(w) &= \left(\sum_{n=0}^\infty \frac{1}{n!} z^n \right) \left(\sum_{n=0}^\infty \frac{1}{n!} w^n \right) = \sum_{n=0}^\infty \left(\sum_{k=0}^{n} \frac{z^{n-k} w^k}{(n-k)!k!} \right) = \\ &= \sum_{n=0}^\infty \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^{n-k} w^k = \sum_{n=0}^\infty \frac{1}{n!} (z+w)^n = \exp(z+w)\,, \end{aligned}$$

wobei das vorletzte Gleichheitszeichen aus dem Binomialsatz 1.32 folgt. □

3.69. Folgerungen. (a) *Für alle $z \in \mathbb{C}$ ist $\exp(z) \neq 0$ und $\dfrac{1}{\exp(z)} = \exp(-z)$.*

(b) *Für alle $x \in \mathbb{R}$ ist $\exp(x) > 0$.*

(c) *Für alle $n \in \mathbb{Z}$ ist $\exp(n) = e^n$ mit $e := \exp(1)$.*

Man nennt die Konstante $e = \exp(1)$ die Eulersche[8] Zahl.

Beweis. (a) Nach Satz 3.68 und Satz 3.56 gilt für alle $z \in \mathbb{C}$:

$$\exp(z)\exp(-z) = \exp(0) = 1\,.$$

Insbesondere ist $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$.

[8] Leonhard Euler (15.4.1707–18.9.1983)

(b) Sei $x \in \mathbb{R}$ beliebig. Nach (a) ist $\exp(x) \neq 0$ und nach Satz 3.56 reell. Wegen

$$\exp(x) = \exp\left(\frac{x}{2} + \frac{x}{2}\right) = \exp\left(\frac{x}{2}\right) \exp\left(\frac{x}{2}\right)$$

ist $\exp(x)$ ein Quadrat, also positiv.

(c) Mit Hilfe der Funktionalgleichung der Exponentialfunktion zeigt man dies leicht durch vollständige Induktion. $\square$

3.70. Folgerungen. *Für alle $z, w \in \mathbb{C}$ gelten die Additionstheoreme:*

(a) $\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$.
(b) $\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w)$.
(c) $\sin(z)^2 + \cos(z)^2 = 1$.
(d) $\forall x \in \mathbb{R}: \qquad |\cos(x)| \leq 1\,, \quad |\sin(x)| \leq 1$.

Beweis. (a), (b): Nach Satz 3.56 (e) gilt für alle $z, w \in \mathbb{C}$

$$\cos(z+w) = \frac{1}{2}\left(\exp(iz+iw) + \exp(-iz-iw)\right)$$

und

$$\sin(z+w) = \frac{1}{2i}\left(\exp(iz+iw) - \exp(-iz-iw)\right) .$$

Wendet man nun auf die Ausdrücke $\exp(iz+iw)$ und $\exp(-iz-iw)$ die Funktionalgleichung der Exponentialfunktion an und verwendet anschließend Satz 3.56 (d), so folgen die Behauptungen.

(c) folgt wegen $\cos(0) = 1$ aus (b) mit $w = -z$ und (d) folgt aus (c). $\square$

KAPITEL 4

Stetige Funktionen und Grenzwerte bei Funktionen

1. Stetigkeit von Funktionen

Naiv versteht man unter einer stetigen Funktion eine Funktion, die keine Sprünge macht und keinen plötzlichen Veränderungen unterliegt, besser: Eine Funktion f auf einer Menge X in $\mathbb{R}$ ($\mathbb{C}$ oder $\mathbb{R}^N$) mit Werten in $\mathbb{R}$ ($\mathbb{C}$ oder $\mathbb{R}^N$) wird man stetig in einem Punkt a nennen, wenn $f(x)$ *nahe bei* $f(a)$ liegt, wenn nur x *hinreichend nahe bei* a liegt. Da wir in $\mathbb{R}$, $\mathbb{C}$ oder allgemeiner im $\mathbb{R}^N$ ($N \in \mathbb{N}$) einen Abstandsbegriff zur Verfügung haben, können wir diese naive Vorstellung präzisieren.

Für alle $y \in \mathbb{R}^N$ und $\varepsilon > 0$ sei im folgenden

$$U_\varepsilon(y) := \{x \in \mathbb{R}^N\,;\, |x-y| < \varepsilon\}\,.$$

4.1. Definition. Seien $N, M \in \mathbb{N}$ und sei $\emptyset \neq X \subseteq \mathbb{R}^M$. Eine Funktion $f : X \to \mathbb{R}^N$ heißt *stetig in einem Punkt* $a \in X$, falls es zu jedem $\varepsilon > 0$ ein $\delta = \delta(a,\varepsilon) > 0$ gibt, so daß für alle $x \in X$ mit $|x-a| < \delta$ gilt: $|f(x) - f(a)| < \varepsilon$.

Formal können wir dies auch so ausdrücken: $f : X \to \mathbb{R}^N$ ist stetig in $a \in X$ genau dann, wenn gilt

$$\forall \varepsilon > 0\, \exists \delta = \delta(a,\varepsilon) > 0\, \forall x \in X : \quad |x-a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon\,. \tag{4.1}$$

(4.1) ist offensichtlich äquivalent zu

$$\forall \varepsilon > 0\, \exists \delta = \delta(a,\varepsilon) > 0\, \forall x \in X \cap U_\delta(a) : \quad f(x) \in U_\varepsilon(f(a)) \tag{4.2}$$

und zu

$$\forall \varepsilon > 0\, \exists \delta = \delta(a,\varepsilon) > 0 : \quad f\big(U_\delta(a) \cap X\big) \subseteq U_\varepsilon(f(a))\,. \tag{4.3}$$

Eine Funktion $f : X \to \mathbb{R}^N$ heißt *stetig auf* X, falls f in allen Punkten $a \in X$ stetig ist.

$f : X \to \mathbb{R}^N$ heißt *unstetig in einem Punkt* $a \in X$, falls f in a nicht stetig ist. Durch richtiges Negieren der Aussagen (4.1)–(4.3) sieht man, daß dies äquivalent zu jeder der drei folgenden Aussagen ist:

$$\exists \varepsilon > 0\, \forall \delta > 0\, \exists x \in X : \quad |x-a| < \delta \text{ und } |f(x) - f(a)| \geq \varepsilon\,. \tag{4.4}$$

$$\exists \varepsilon > 0\, \forall \delta > 0\, \exists x \in X \cap U_\delta(a) : \quad f(x) \notin U_\varepsilon(f(a))\,. \tag{4.5}$$

$$\exists \varepsilon > 0\, \forall \delta > 0 : \quad f\big(U_\delta(a) \cap X\big) \not\subseteq U_\varepsilon(f(a))\,. \tag{4.6}$$

4.2. Beispiele. (a) Konstante Funktionen sind stetig, d.h. ist $b \in \mathbb{R}^N$ und $\emptyset \neq X \subseteq \mathbb{R}^M$, so ist die Funktion $f : X \to \mathbb{R}^N$ mit $f(x) := b$ für alle $x \in X$ auf X stetig.

(b) id: $\mathbb{R}^N \to \mathbb{R}^N$ mit $\mathrm{id}(x) := x$ für alle $x \in \mathbb{R}^N$ ist auf $\mathbb{R}^N$ stetig.

(c) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $f(x) := x^2$ für alle $x \in \mathbb{K}$. Dann ist die hierdurch definierte Funktion $f : \mathbb{K} \to \mathbb{K}$ auf $\mathbb{K}$ stetig.

(d) Sei $\emptyset \neq X \subseteq \mathbb{R}^M$. Man sagt, daß eine Funktion $f : X \to \mathbb{R}^N$ auf X einer *Lipschitz–Bedingung*[1] genügt, falls es eine reelle Konstante $L \geq 0$ gibt mit

$$\text{(L)} \qquad \forall x, y \in X : \qquad |f(x) - f(y)| \leq L|x - y| .$$

L heißt dann eine *Lipschitz–Konstante* für f. Genügt f auf X einer Lipschitz–Bedingung, so ist f auf X auch stetig.

(e) Der euklidische Betrag $|\cdot| : \mathbb{R}^M \to \mathbb{R}$ ist auf $\mathbb{R}^M$ stetig.

(f) Die kanonischen Projektionen (auch Koordinatenfunktionen genannt) $\pi_k : \mathbb{R}^M \to \mathbb{R}$ mit $\pi_k(x) := x_k$ für alle $x = (x_1, \ldots, x_M) \in \mathbb{R}^M$ sind auf $\mathbb{R}^M$ stetig.

(g) Die Abbildungen

$$\mathrm{Re} : \mathbb{C} \to \mathbb{R}, \quad x + iy \mapsto x, \qquad \mathrm{Im} : \mathbb{C} \to \mathbb{R}, \quad x + iy \mapsto y,$$
$$|\cdot| : \mathbb{C} \to \mathbb{R}, \quad x + iy \mapsto \sqrt{x^2 + y^2}, \qquad ^{-} : \mathbb{C} \mapsto \mathbb{C}, \quad x + iy \mapsto x - iy,$$

sind stetig auf $\mathbb{C}$.

(h) Die *Heavyside–Funktion* $h : \mathbb{R} \to \mathbb{R}$ mit

$$h(x) := \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

ist in 0 unstetig und in allen Punkten $a \in \mathbb{R} \setminus \{0\}$ stetig.

(i) Die Funktion $\chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ mit

$$\chi_{\mathbb{Q}}(x) := \begin{cases} 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & \text{für } x \in \mathbb{Q} \end{cases}$$

ist in keinem Punkt von $\mathbb{R}$ stetig.

(j) Die Funktion $\sqrt{\cdot} : [0, \infty) \to \mathbb{R}$, $x \mapsto \sqrt{x}$ ist auf $[0, \infty)$ stetig.

(k) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $\mathbb{K}_* = \mathbb{K} \setminus \{0\}$. Die Funktion $f : \mathbb{K}_* \to \mathbb{K}$ mit $f(x) := 1/x$ für alle $x \in \mathbb{K}_*$ ist auf $\mathbb{K}_*$ stetig und die Funktionen $s, p : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ mit $s(x, y) := x + y$ und $p(x, y) := xy$ sind auf $\mathbb{K} \times \mathbb{K}$ stetig. (Hierbei identifizieren wir $\mathbb{R} \times \mathbb{R}$ mit $\mathbb{R}^2$ und $\mathbb{C} \times \mathbb{C}$ mit $\mathbb{R}^4$).

Beweis. (a) Für alle $a \in X$ ist die Bedingung (4.1) offensichtlich für beliebiges $\varepsilon > 0$ sogar für alle $\delta > 0$ erfüllt.

(b) Für alle $a \in \mathbb{R}^N$ ist die Bedingung (4.1) offensichtlich für beliebiges $\varepsilon > 0$ mit $\delta = \delta(\varepsilon) := \varepsilon$ erfüllt.

(c) Seien $a \in \mathbb{K}$ beliebig und sei $\varepsilon > 0$ beliebig vorgegeben. Für alle $x \in \mathbb{K}$ gilt

$$|f(x) - f(a)| = |x^2 - a^2| = |x - a| \cdot |x + a| .$$

Wählen wir also

$$\delta := \min\left\{1, \frac{\varepsilon}{2|a| + 1}\right\},$$

so folgt für alle $x \in \mathbb{K}$ mit $|x - a| < \delta$: Es ist $|x + a| \leq |x - a| + |2a| < 2|a| + 1$ und daher $|f(x) - f(a)| = |x - a| \cdot |x + a| \leq |x - a|(2|a| + 1) < \varepsilon$. Also ist (4.1) für alle $a \in \mathbb{R}^N$ erfüllt.

(d) Sei nun (L) für $f : X \to \mathbb{R}^N$ erfüllt mit einer Lipschitz–Konstante $L \geq 0$ und sei $a \in X$ beliebig. Für jedes $\varepsilon > 0$ gilt dann mit $\delta := \varepsilon \cdot (L + 1)$: Für alle $x \in X$ mit $|x - a| < \delta$ ist

$$|f(x) - f(a)| \leq L|x - a| \leq \frac{L}{L + 1}\varepsilon < \varepsilon .$$

[1] Rudolf Otto Sigismund Lipschitz (14.5.1832–7.10.1903).

(e)–(g) Man rechnet leicht nach, daß die Abbildungen in (e)–(g) Lipschitz–stetig sind mit Lipschitz–Konstante $L = 1$. Damit folgt die Stetigkeit der zu untersuchenden Funktionen aus (d).

(h) Ist $0 \neq a \in \mathbb{R}$ und $\varepsilon > 0$ beliebig, so folgt mit $\delta := \delta(a) := |a|$ für alle $x \in \mathbb{R}$ mit $|x - a| < \delta = |a|$: Es hat x das gleiche Vorzeichen wie a und somit gilt nach Definition von h: $|h(x) - h(a)| = 0 \leq \varepsilon$. Also ist (4.1) für a und h erfüllt und h ist in a stetig.

Für $a = 0$ gilt mit $\varepsilon = 1$: Für alle $\delta > 0$ ist $x := -\delta/2 \in U_\delta(0)$ und $|h(x) - h(0)| = 1 = \varepsilon$, d.h. es ist $h(x) \notin U_\varepsilon(h(0))$. Somit ist (4.5) für h erfüllt, d.h. h ist in 0 unstetig.

(i) Sei $a \in \mathbb{R}$ beliebig. Wir wählen $\varepsilon = 1$. Ist $x \in \mathbb{Q}$, so gibt es nach Satz 1.43 zu jedem $\delta > 0$ ein irrationales $x \in U_\delta(a)$ und im Fall $a \in \mathbb{R} \setminus \mathbb{Q}$ ein rationales $x \in U_\delta(a)$. In beiden Fällen erhalten wir $|\chi_{\mathbb{Q}}(x) - \chi_{\mathbb{Q}}(a)| = 1 = \varepsilon$ und somit $\chi_{\mathbb{Q}}(x) \notin U_\varepsilon(\chi_{\mathbb{Q}}(a))$. Also ist (4.5) für $\chi_{\mathbb{Q}}$ und alle $a \in \mathbb{R}$ erfüllt, d.h. $\chi_{\mathbb{Q}}$ ist in allen Punkten $a \in \mathbb{R}$ unstetig.

(j) Im Punkt $a = 0$ gilt für beliebig vorgegebenes $\varepsilon > 0$ mit $\delta := \varepsilon^2$ für alle $x \geq 0$ mit $|x - 0| = x < \delta = \varepsilon^2$ wegen der Monotonie der Wurzel: $|\sqrt{x} - \sqrt{0}| = \sqrt{x} < \varepsilon$. Also ist (4.1) für $a = 0$ erfüllt und damit $\sqrt{\cdot}$ in 0 stetig.

Sei nun $a > 0$. In diesem Fall erhalten wir für alle $x \geq 0$ mit $|x - a| < \delta := \varepsilon\sqrt{a}$ unter Verwendung der dritten binomischen Formel:

$$|\sqrt{x} - \sqrt{a}| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{|x - a|}{\sqrt{a}} < \varepsilon\,.$$

Wieder ist (4.1) für a erfüllt und daher $\sqrt{\cdot}$ in a stetig.

(k) Sei $0 \neq a \in \mathbb{K}$ beliebig und $\varepsilon > 0$ beliebig. Für alle $x \in \mathbb{K}_*$ gilt dann:

$$|f(x) - f(a)| = \left|\frac{1}{x} - \frac{1}{a}\right| = \frac{|a - x|}{|x| \cdot |a|}\,.$$

Wählen wir also $\delta := \min\{|a|/2, \varepsilon|a|^2/2\}$, so folgt für alle $x \in \mathbb{K}_*$ mit $|x - a| \leq \delta$: Es ist $|x| \geq |a|/2$ und daher

$$|f(x) - f(a)| = \frac{|a - x|}{|x| \cdot |a|} \leq \frac{2|a - x|}{|a|^2} < \varepsilon\,.$$

f ist daher in allen $a \in \mathbb{K}_*$ stetig (da (4.1) erfüllt ist).

Um die Stetigkeit von s zu beweisen, beachten wir, daß für alle $(u, v), (x, y) \in \mathbb{K} \times \mathbb{K}$ unter Verwendung der Dreiecksungleichung und der Cauchy-Schwarzschen Ungleichung (vergl. Satz 2.8 (a)) gilt:

$$\begin{aligned} |s(x, y) - s(u, v)| =& |x + y - (u + v)| \leq |x - u| + |y - v| \\ \leq& \sqrt{2}(|x - u|^2 + |y - v|^2)^{1/2} = \sqrt{2}|(x, y) - (u, v)|\,. \end{aligned}$$

s erfüllt also eine Lipschitz–Bedingung mit der Lipschitz–Konstante $\sqrt{2}$ und ist daher nach (d) auf ganz $\mathbb{K} \times \mathbb{K}$ stetig.

Um die Stetigkeit der Produktbildung zu beweisen, schätzen wir zunächst wie folgt ab: Für alle $(x, y), (u, v) \in \mathbb{K} \times \mathbb{K}$ gilt

$$|p(x, y) - p(u, v)| = |xy - uv| \leq |xy - uy| + |uy - uv| = |y| \cdot |x - u| + |u| \cdot |y - v|.$$

Ferner $|y| = |v + (y - v)| \leq |v| + |y - v| \leq |v| + |(x, y) - (u, v)|$.

Sei nun $(u, v) \in \mathbb{K} \times \mathbb{K}$ beliebig und $\varepsilon > 0$ beliebig vorgegeben. Für alle $(x, y) \in \mathbb{K} \times \mathbb{K}$ mit

$$|(x, y) - (u, v)| < \delta := \min\left\{1, \frac{\varepsilon}{|u| + |v| + 1}\right\}$$

gilt dann: Wegen $|x - u| \leq |(x, y) - (u, v)| < \delta$ und $|y - v| \leq |(x, y) - (u, v)| < \delta$ ist

$$|p(x, y) - p(u, v)| \leq |y| \cdot |x - u| + |u| \cdot |y - v| < (|v| + 1 + |u|)\delta \leq \varepsilon\,.$$

Damit ist gezeigt, daß (4.1) für alle $(u, v) \in \mathbb{K} \times \mathbb{K}$ erfüllt ist. p ist also auf ganz $\mathbb{K} \times \mathbb{K}$ stetig. $\square$

4.3. Definition. Seien X, Y zwei nicht leere Mengen und sei $\emptyset \neq A \subset X$. Ist $f : X \to Y$ eine Abbildung, so heißt die durch

$$(f|_A)(a) := f(a) \qquad \text{für alle } a \in A$$

definierte Abbildung $f|_A : A \to Y$ die *Einschränkung* oder *Restriktion* von f auf A.

Das folgende Lemma zeigt, daß Stetigkeit eine lokale Eigenschaft ist und daß diese Eigenschaft sich auf Restriktionen vererbt.

4.4. Lemma. *Sei $\emptyset \neq A \subset X \subseteq \mathbb{R}^M$ und sei $f : X \to \mathbb{R}^N$ eine Funktion.*

(a) *Ist f in einem Punkt $a \in A$ stetig, so ist auch $f|_A$ in a stetig.*
(b) *Ist f auf X stetig, so ist auch $f|_A$ auf A stetig.*
(c) *Gibt es zu jedem $x \in X$ ein $\eta > 0$, so daß die Restriktion $f|_{U_\eta(x)\cap X}$ auf $U_\eta(x) \cap X$ stetig ist, so ist f auf X stetig.*

Beweis. Zu (a): Sei f in dem Punkt $a \in A$ stetig und sei $\varepsilon > 0$ beliebig. Dann gibt es ein $\delta > 0$, so daß $f(U_\delta(a) \cap X) \subseteq U_\varepsilon(f(a))$, also gilt auch $f(U_\delta(a) \cap A) \subseteq f(U_\delta(a) \cap X) \subseteq U_\varepsilon(f(a))$. Damit ist gezeigt, daß $f|_A$ in a stetig ist.

(b) folgt durch Anwendung von (a) in jedem Punkt von A.

Zu (c): Sei die Bedingung in (c) erfüllt und seien $x \in X$ und $\varepsilon > 0$ beliebig vorgegeben. Nach Voraussetzung gibt es ein $\eta > 0$, so daß $f|_{U_\eta(x)\cap X}$ stetig ist. Insbesondere gibt es ein $\delta_1 > 0$ mit $f(U_{\delta_1}(x) \cap U_\eta(x) \cap X) = (f|_{U_\eta(x)})(U_{\delta_1}(x) \cap U_\eta(x) \cap X) \subseteq U_\varepsilon(f(x))$. Für $\delta := \min\{\eta, \delta_1\}$ ist $U_\delta(x) \subseteq U_\eta(x)$ und daher $f(U_\delta(x) \cap X) = f(U_{\delta_1}(x) \cap U_\eta(x) \cap X) \subseteq U_\varepsilon(f(a))$. Damit ist gezeigt, daß f in allen $x \in X$ stetig ist. $\square$

Häufig verwendet man das folgende Kriterium um die Stetigkeit oder die Unstetigkeit einer gegebenen Funktion nachzuweisen.

4.5. Satz (Folgenkriterium). *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$ und sei $f : X \to \mathbb{R}^N$ eine $\mathbb{R}^N$–wertige Funktion auf X $(N, M \in \mathbb{N})$.*

(a) *f ist genau dann in einem Punkt $a \in X$ stetig, wenn für <u>jede</u> gegen a konvergente Folge $(x_n)_{n=1}^\infty$ in X gilt: $f(x_n) \to f(a)$ für $n \to \infty$.*
(b) *f ist genau dann in einem Punkt $a \in X$ unstetig, wenn es eine gegen a konvergente Folge $(x_n)_{n=1}^\infty$ in X gibt, für die die Folge $(f(x_n))_{n=1}^\infty$ der Bildpunkte <u>nicht</u> gegen $f(a)$ konvergiert für $n \to \infty$.*
(c) *f ist genau dann auf X stetig, wenn für jede konvergente Folge $(x_n)_{n=1}^\infty$ in X mit $\lim_{n\to\infty} x_n \in X$ auch die Folge $(f(x_n))_{n=1}^\infty$ der Bildpunkte konvergent ist und falls gilt:*

$$f\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} f(x_n).$$

Beweis. (a) „$\Longrightarrow$“: Sei f in a stetig und sei $(x_n)_{n=1}^\infty$ eine beliebige gegen a konvergente Folge aus X. Sei $\varepsilon > 0$ beliebig. Da f in a stetig ist gibt es ein $\delta > 0$ mit

$$\forall x \in X : \qquad |x - a| < \delta \Longrightarrow |f(x) - f(a)| < \varepsilon\,. \tag{4.7}$$

Wegen $x_n \to a$ für $n \to \infty$ gibt es ein $n_0 \in \mathbb{N}$ mit $|x_n - a| < \delta$ für alle $n \geq n_0$. Wegen (4.7) folgt also $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$ und somit $f(x_n) \to f(a)$ für $n \to \infty$.

„$\Longleftarrow$“: Gelte nun für jede gegen a konvergente Folge $(x_n)_{n=1}^\infty$, daß die Folge $(f(x_n))_{n=1}^\infty$ gegen $f(a)$ konvergiert für $n \to \infty$. Wir machen die *Annahme*, daß f in a unstetig ist, d.h. daß ein $\varepsilon_0 > 0$ existiert mit

$$\forall \delta > 0 \, \exists x \in U_\delta(a) \cap X : \qquad |f(x) - f(a)| \geq \varepsilon_0 \, .$$

Insbesondere gibt es zu $\delta := 1/n$ ein $x_n \in U_\delta(a) \cap X$ mit $|f(x_n) - f(a)| \geq \varepsilon_0$. Die Folge $(f(x_n))_{n=1}^\infty$ konvergiert daher nicht gegen $f(a)$ obwohl $|x_n - a| < 1/n \to 0$ für $n \to \infty$. Dies steht im Widerspruch zur Voraussetzung. Die Annahme war also falsch. f muß in a stetig sein.

(b) folgt durch Negieren aus (a) und (c) erhält man durch Anwendung von (a) in allen Punkten von X. $\square$

Mit Hilfe des Folgenkriteriums für die Stetigkeit erhalten wir nun leicht die folgenden Rechenregeln für stetige Funktionen aus den Grenzwertrechenregeln 3.13.

4.6. Lemma (Rechenregeln für stetige Funktionen). *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$, $a \in X$, $\lambda \in \mathbb{R}$ und $\mu \in \mathbb{C}$.*

(a) *Sind die Funktionen $f, g : X \to \mathbb{R}^N$ stetig in a (bzw. auf ganz X), so gilt dies auch für die Funktionen $f + g$ und λf.*
(b) *Sind die Funktionen $f : X \to \mathbb{R}$ und $g : X \to \mathbb{R}^N$ stetig in a (bzw. auf ganz X), so gilt dies auch für die Funktion $fg : x \mapsto f(x)g(x)$.*
(c) *Sind die Funktionen $f, g : X \to \mathbb{C}$ stetig in a (bzw. auf ganz X), so gilt dies auch für die Funktionen $f + g$, fg und μf.*
(d) *Sei $\mathbb{K} := \mathbb{R}$ oder $\mathbb{K} := \mathbb{C}$ und sei $f : X \to \mathbb{K}$ eine in a stetige Funktion mit $f(a) \neq 0$. Dann gibt es ein $\eta > 0$, so daß $f(x) \neq 0$ für alle $x \in U_\eta(a) \cap X$ und die durch*
$$h(x) := \frac{1}{f(x)} \qquad \text{für alle } x \in U_\eta(a) \cap X$$
definierte Funktion $h : U_\eta(a) \cap X \to \mathbb{K}$ ist in a stetig.
Ist $f : X \to \mathbb{K}$ auf ganz X stetig mit $f(X) \subseteq \mathbb{K}_*$, so ist auch die Funktion
$$\frac{1}{f} : X \to \mathbb{K} \, , \qquad x \mapsto \frac{1}{f(x)}$$
auf ganz X stetig.
(e) *Sei $f : X \to \mathbb{R}^N$ stetig in a (bzw. auf ganz X). Ist $Y \subseteq \mathbb{R}^N$ mit $f(X) \subseteq Y$ und ist $g : Y \to \mathbb{R}^K$ eine in $f(a)$ (bzw. auf ganz Y) stetige Funktion, so ist auch die Hintereinanderausführung*
$$g \circ f : X \to \mathbb{R}^K \, , \qquad x \mapsto g(f(x))$$
in a (bzw. auf ganz X) stetig.

Beweis. Die globalen Stetigkeitsaussagen folgen jeweils durch Anwendung der lokalen Stetigkeitsaussagen in allen Punkten $a \in X$. Es müssen daher nur die lokalen Aussagen bewiesen werden.

(a), (b) und (c) folgen direkt mit Hilfe der entsprechenden Grenzwertrechenregeln aus Lemma 3.13.

Zu (d): Wegen $f(a) \neq 0$ ist $\varepsilon := |f(a)| > 0$. Da f in a stetig ist, gibt es ein $\delta > 0$ mit $|f(x) - f(a)| < \varepsilon = |f(a)|$ für alle $x \in U_\delta(a) \cap X$. Es folgt

$$\forall x \in U_\delta(a) \cap X : \qquad |f(a)| - |f(x)| \leq |f(x) - f(a)| < |f(a)|$$

und hieraus $f(x) \neq 0$ für alle $x \in U_\delta(a) \cap X$. Für jede gegen a konvergente Folge $(x_n)_{n=1}^\infty$ aus $U_\delta(a) \cap X$ gilt nach Satz 4.5 $f(x_n) \to f(a)$ für $n \to \infty$ und daher nach Lemma 3.13 (e) auch $1/f(x_n) \to 1/f(a)$ für $n \to \infty$. Nach dem Folgenkriterium 4.5 ist h in a stetig.

Zu (e): Für jede gegen a konvergente Folge $(x_n)_{n=1}^\infty$ aus X konvergiert die Folge $(f(x_n))_{n=1}^\infty$ nach Satz 4.5 gegen $f(a)$ für $n \to \infty$. Da auch g in $f(a)$ stetig ist folgt ebenfalls nach 4.5: $(g \circ f)(x_n) = g(f(x_n)) \to g(f(a)) = (g \circ f)(a)$ für $n \to \infty$. Nach dem Folgenkriterium 4.5 ist $g \circ f$ also in a stetig. $\square$

Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Sind $n_0 \in \mathbb{N}_0$ und $a_0, \dots, a_n \in \mathbb{K}$ gegeben, so nennt man die durch

$$p(x) := \sum_{k=0}^{n} a_k x^k \qquad \text{für alle } x \in \mathbb{K}$$

definierte Funktion eine *Polynomfunktion*. Ihr *Grad* $\deg(p)$ ist definiert als

$$\deg(p) := \sup\{k \in \mathbb{N}_0\,;\, k \le n, a_k \neq 0\}\,.$$

Ist also $a_k = 0$ für alle $k = 0, \dots, n$ so ist $\deg(p) = -\infty$, da das Supremum über die leere Menge als $-\infty$ definiert ist. Ist wenigstens einer der Koeffizienten a_k von Null verschieden, so ist $\deg(p) \in \mathbb{N}_0$.

Sind p, q zwei Polynomfunktionen und ist $q \not\equiv 0$, so nennt man die Funktion

$$\frac{p}{q} : \mathbb{K} \setminus q^{-1}(\{0\}) \to \mathbb{K}, \quad x \mapsto \frac{p(x)}{q(x)}$$

eine *rationale Funktion*. Die Menge

$$D_{p/q} := \mathbb{K} \setminus q^{-1}(\{0\}) = \{x \in \mathbb{K}\,;\, q(x) \neq 0\}$$

heißt ihr *natürlicher Definitionsbereich*.

4.7. Folgerung. *Jede Polynomfunktion ist auf $\mathbb{K}$ stetig. Jede rationale Funktion ist auf ihrem natürlichen Definitionsbereich stetig.*

Beweis. Nach Beispiel 4.2 (a) und (b) sind konstante Funktionen und die Funktion $x \mapsto x$ auf $\mathbb{K}$ stetig. Die Aussage für Polynome beweist man hieraus durch vollständige Induktion nach dem Grad mit Hilfe von Lemma 4.6 (a)–(c). Die Stetigkeit der rationalen Funktionen folgt dann mit Teil (d) von Lemma 4.6. $\square$

4.8. Bezeichnungen. Sei $\emptyset \neq X \subseteq \mathbb{R}^M$ und sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Wir definieren

$$\begin{aligned} C(X, \mathbb{R}^N) :=& \{f\,;\, f \text{ stetige } \mathbb{R}^N\text{–wertige Funktion auf } X\} \\ C(X, \mathbb{K}) :=& \{f\,;\, f \text{ stetige } \mathbb{K}\text{–wertige Funktion auf } X\} \\ \mathbb{K}[Z] :=& \{p\,;\, p \text{ Polynomfunktion auf } \mathbb{K} \text{ mit Koeffizienten in } \mathbb{K}\}\,. \end{aligned}$$

4.9. Folgerung. *Mit den Voraussetzungen und Bezeichnungen von 4.8 gilt: Versehen mit den punktweisen Operationen der Addition und der Multiplikation mit Skalaren sind $C(X, \mathbb{K})$ und $\mathbb{K}[Z]$ $\mathbb{K}$–Vektorräume und $C(X, \mathbb{R}^N)$ ein $\mathbb{R}$-Vektorraum.*

Zum Beweis rechnet man mit Hilfe von Lemma 4.6 einfach nach, daß die Vektorraumaxiome (VR1) und (VR2) in Definition 2.1 erfüllt sind.

4.10. Satz. *Die Funktionen* exp, sin *und* cos *sind auf ganz $\mathbb{C}$ (also auch auf ganz $\mathbb{R}$) stetig.*

Beweis. Sei $z_0 \in \mathbb{C}$ und $\varepsilon > 0$ beliebig. Für alle $z \in \mathbb{C}$ mit $|z - z_0| < 1$ gilt unter Verwendung der Funktionalgleichung der Exponentialfunktion (Satz 3.68):

$$\begin{aligned}|\exp(z) - \exp(z_0)| =& |\exp(z_0)| \cdot |\exp(z - z_0) - 1| = |\exp(z_0)| \cdot \left| \sum_{k=1}^{\infty} \frac{(z - z_0)^k}{k!} \right| \leq \\ \leq& |\exp(z_0)| \cdot |z - z_0| \sum_{k=1}^{\infty} \frac{|z - z_0|^{k-1}}{k!} \leq \\ \leq& |\exp(z_0)| \cdot |z - z_0| \sum_{k=1}^{\infty} \frac{1}{k!} = |\exp(z_0)| \cdot |z - z_0|(e - 1)\,.\end{aligned}$$

Wählen wir also

$$\delta := \min \left\{ 1, \frac{\varepsilon}{|\exp(z_0)|(e-1)} \right\},$$

so folgt $|\exp(z) - \exp(z_0)| < \varepsilon$ für alle $z \in \mathbb{C}$ mit $|z - z_0| \leq \delta$. Somit ist exp in allen Punkten $z_0 \in \mathbb{C}$ stetig. Wegen

$$\cos(z) = \frac{\exp(iz) + \exp(-iz)}{2} \quad \text{und} \quad \sin(z) = \frac{\exp(iz) - \exp(-iz)}{2i}$$

folgt nach Lemma 4.6 (c) auch die Stetigkeit der Funktionen cos und sin auf ganz $\mathbb{C}$. □

2. Eigenschaften stetiger Funktionen

In diesem Abschnitt wollen wir einige wichtige Eigenschaften stetiger Funktionen und deren Anwendungen kennenlernen. Im folgenden sei $[a, b]$ ein abgeschlossenes Intervall mit $a < b$.

4.11. Beispiel. Die Funktion $f : [0, 2] \cap \mathbb{Q} \to \mathbb{R}$ mit $f(x) = x^2$ für alle $x \in [0, 2] \cap \mathbb{Q}$ ist stetig auf $[0, 2] \cap \mathbb{Q}$. Es gilt $0 = f(0) < \sqrt{2} < 4 = f(2)$, aber es gibt kein $x \in [0, 2] \cap \mathbb{Q}$ mit $f(x) = 2$. Der Grund ist, daß der Definitionsbereich von f *Lücken* besitzt: $\sqrt{2}$ ist zwar in $[0, 2]$, aber nicht rational.

Es gilt jedoch:

4.12. Satz (Zwischenwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ eine auf $[a, b]$ stetige Funktion und sei $y \in \mathbb{R}$ ein Wert zwischen $f(a)$ und $f(b)$, d.h. es gelte $f(a) \leq y \leq f(b)$ oder $f(a) \geq y \geq f(b)$. Dann gibt es ein $x_0 \in [a, b]$ mit $f(x_0) = y$.*

Beweis. Die Aussage scheint anschaulich klar, bedarf aber doch eines Beweises. Wir geben einen konstruktiven Beweis mit Hilfe von Intervallschachtelung an. Wir beschränken uns auf den Fall $f(a) \leq y \leq f(b)$. (Den Fall $f(a) \geq y \geq f(b)$ zeigt man entweder analog oder man beachtet $-f(a) \leq -y \leq -f(b)$ und wendet den ersten Fall auf die ebenfalls stetige Funktion $-f$ an).

Wir konstruieren induktiv zwei Folgen $(a_n)_{n=0}^{\infty}$ und $(b_n)_{n=0}^{\infty}$ in $[a, b]$ so, daß für alle $n \in \mathbb{N}$ gilt:

(i)$_n$ $b_n - a_n = \dfrac{b - a}{2^{n-1}}$.

(ii)$_n$ $a_{n-1} \leq a_n < b_n \leq b_{n-1}$.

(iii)$_n$ $f(a_n) \leq y \leq f(b_n)$.

Induktionsanfang: Mit $a_0 := a_1 := a$ und $b_0 := b_1 := b$ sind die Aussagen (i)$_1$-(iii)$_1$ erfüllt.

Induktionsschluß: Seien nun schon a_n und b_n mit (i)$_n$-(iii)$_n$ konstruiert. Wir setzen $m_n := (a_n + b_n)/2$. Ist $f(m_n) \leq y$, so definieren wir $a_{n+1} := m_n$ und $b_{n+1} := b_n$. Im Fall $f(m) > y$ setzen wir $a_{n+1} := a_n$ und $b_{n+1} := m_n$. In beiden Fällen sind die Forderungen (i)$_{n+1}$-(iii)$_{n+1}$ erfüllt.

Damit haben wir eine Intervallschachtelung $([a_n, b_n])_{n=0}^{\infty}$ erhalten. Nach dem Intervallschachtelungssatz 3.18 gibt es genau einen Punkt $x_0 \in \bigcap_{n=0}^{\infty}[a_n, b_n]$ und es gilt $a_n \to x_0$ und $b_n \to x_0$ für $n \to \infty$. Da f nach Voraussetzung auf $[a, b]$ und damit insbesondere in x_0 stetig ist gilt nach dem Folgenkriterium 4.5

$$\lim_{n\to\infty} f(a_n) = f(x_0) = \lim_{n\to\infty} f(b_n)\,.$$

Wegen $f(a_n) \leq y \leq f(b_n)$ folgt aus der Einschachtelungsregel in Lemma 3.13 (d) schon $f(x_0) = y$. □

4.13. Folgerung (Nullstellensatz). *Sei $f : [a, b] \to \mathbb{R}$ eine auf $[a, b]$ stetige Funktion mit $f(a)f(b) < 0$ (d.h. f hat in $[a, b]$ einen Vorzeichenwechsel). Dann gibt es ein $x_0 \in (a, b)$ mit $f(x_0) = 0$.*

4.14. Folgerung. *Hat $p \in \mathbb{R}[Z]$ ungeraden Grad, so hat p wenigstens eine reelle Nullstelle.*

Beweis. Sei also p eine Polynomfunktion von der Gestalt $x \mapsto p(x) = \sum_{k=0}^{2n+1} c_k x^k$ für ein $n \in \mathbb{N}_0$ mit $c_0, \dots, c_{2n+1} \in \mathbb{R}$ und $c_{2n+1} \neq 0$. Durch Einsetzen und Ausmultiplizieren sieht man, daß $x \mapsto p(x)p(-x)$ eine Polynomfunktion der Gestalt

$$p(x)p(-x) = -c_{2n+1}^2 x^{4n+2} + \sum_{k=0}^{4n+1} d_k x^k$$

mit gewissen $d_0, \dots, d_{4n+1} \in \mathbb{R}$. Nach den Grenzwertrechenregeln folgt für $j \in \mathbb{N}$, $j \to \infty$:

$$\frac{p(j)p(-j)}{j^{4n+2}} = -c_{2n+1}^2 + \sum_{k=0}^{4n+1} d_k j^{k-4n-2} \to -c_{2n+1}^2 < 0\,.$$

Es gibt daher ein $j \in \mathbb{N}$ mit $\frac{p(j)p(-j)}{j^{4n+2}} < 0$, also auch mit $p(j)p(-j) < 0$. Da p auf $[-j, j]$ stetig ist, muß p nach dem Nullstellensatz 4.13 eine Nullstelle in $(-j, j)$ besitzen. □

Wir benötigen noch einige topologische Grundbegriffe:

4.15. Definition. Eine Menge $X \subseteq \mathbb{R}^M$ heißt

- *offen*, falls es zu jedem $x \in X$ ein $\varepsilon > 0$ gibt mit
$$U_\varepsilon(x) = \{y \in \mathbb{R}^M\,;\ |y - x| < \varepsilon\} \subseteq X\,.$$
- *abgeschlossen*, falls für jede konvergente Folge $(x_n)_{n=1}^{\infty}$ mit $x_n \in X$ für alle $n \in \mathbb{N}$ gilt: $\lim_{n\to\infty} x_n \in X$.

4.16. Beispiele. (a) $\emptyset$ und $\mathbb{R}^M$ sind offene und abgeschlossene Teilmengen des $\mathbb{R}^M$.

(b) Sei $x_0 \in \mathbb{R}^M$ und $r > 0$ beliebig. Dann ist $U_r(x_0)$ eine offene aber nicht abgeschlossene Teilmenge des $\mathbb{R}^M$.
(c) Sei $x_0 \in \mathbb{R}^M$ und $r > 0$ beliebig. Dann ist $B_r(x_0) := \{x \in \mathbb{R}^M\,;\ |x - x_0| \leq r\}$ eine abgeschlossene aber nicht offene Teilmenge des $\mathbb{R}^M$.
(d) $\mathbb{Q}$ ist weder offen noch abgeschlossen in $\mathbb{R}$.

(e) $(0,1]$ ist weder offen noch abgeschlossen in $\mathbb{R}$.

Beweis. (a) ist offensichtlich.

(b) Sei $x \in U_r(x_0)$ beliebig. Dann ist $\varepsilon := r - |x - x_0| > 0$ und für alle $y \in U_\varepsilon(x)$ gilt nach der Dreiecksungleichung: $|y - x_0| \leq |y - x| + |x - x_0| < \varepsilon + |x - x_0| = r$. Es folgt $U_\varepsilon(x) \subseteq U_r(x_0)$. Die Menge $U_r(x_0)$ ist daher offen.

Ist $u \in \mathbb{R}^M$ ein Vektor mit $|u| = 1$ (z.B. $u = (1, 0, \ldots, 0)$), so gilt

$$x_n := x_0 + (1 - \frac{1}{n})ru \in U_r(x_0) \quad \text{für alle } n \in \mathbb{N}$$

aber $x_0 + ru = \lim_{n\to\infty} x_n \notin U_r(x_0)$. Daher ist $U_r(x_0)$ nicht abgeschlossen.

(c) Sei u wie im Beweis zu (b). Dann ist $x := x_0 + ru \in B_r(x_0)$ und für jedes $\varepsilon > 0$ ist $y := x + \frac{1}{2}\varepsilon u \in U_\varepsilon(x)$ aber $y \notin B_r(x_0)$, denn $|x_0 - y| = |x_0 - x_0 - ru - \frac{1}{2}\varepsilon u| = (r + \frac{1}{2}\varepsilon)|u| = r + \frac{1}{2}\varepsilon > r$. Also ist $B_r(x_0)$ nicht offen.

Sei nun $(x_n)_{n=1}^\infty$ eine beliebige konvergente Folge mit $x_n \in B_r(x_0)$ für alle $n \in \mathbb{N}$. Nach Lemma 3.6 in Verbindung mit den Grenzwertrechenregeln für Folgen folgt

$$|\lim_{n\to\infty} x_n - x_0| = \lim_{n\to\infty} |x_n - x_0| \leq r.$$

Also ist $\lim_{n\to\infty} x_n \in B_r(x_0)$ und $B_r(x_0)$ ist abgeschlossen.

(d) zeigt man in naheliegender Weise mit Satz 1.43 und (e) rechnet man leicht nach. □

4.17. Satz (von der Beschränktheit stetiger Funktionen auf abgeschlossenen, beschränkten Mengen). *Sei $\emptyset \neq X \subset \mathbb{R}^M$ abgeschlossen und beschränkt und sei $f : X \to \mathbb{R}^N$ stetig auf X. Dann ist die Funktion f beschränkt auf X, d.h. es gibt ein $C > 0$ mit $|f(x)| \leq C$ für alle $x \in X$.*

Beweis. *Annahme*: f ist nicht beschränkt, d.h.

$$\forall C > 0\, \exists x \in X : \qquad |f(x)| > C\,.$$

Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in X$ mit $|f(x_n)| > n$. Wegen der Beschränktheit von X ist die hierdurch gegebene Folge $(x_n)_{n=1}^\infty$ beschränkt und besitzt daher nach dem Satz 3.34 von Bolzano–Weierstraß eine konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Da X auch abgeschlossen ist folgt $x_0 := \lim_{k\to\infty} x_{n_k} \in X$. Nun ist f nach Voraussetzung auf X und damit insbesondere auch in x_0 stetig. Nach dem Folgenkriterium 4.5 gilt $f(x_{n_k}) \to f(x_0)$ für $k \to \infty$ im Widerspruch dazu, daß die Folge $(f(x_{n_k}))_{k=1}^\infty$ wegen $|f(x_{n_k})| > n_k \geq k$ unbeschränkt ist und daher nach Lemma 3.7 nicht konvergent sein kann. Unsere Annahme war also falsch. f muß somit beschränkt sein. □

4.18. Satz. *Sei $\emptyset \neq X \subset \mathbb{R}^M$ abgeschlossen und beschränkt und sei $f : X \to \mathbb{R}^N$ eine stetige $\mathbb{R}^N$–wertige Funktion auf X. Dann ist auch die Bildmenge $f(X)$ beschränkt und abgeschlossen in $\mathbb{R}^N$.*

Beweis. In Satz 4.17 haben wir soeben gezeigt, daß $f(X)$ beschränkt ist. Wir zeigen, daß $f(X)$ auch abgeschlossen ist. Sei also $(y_n)_{n=1}^\infty$ eine beliebige in $\mathbb{R}^N$ konvergente Folge mit $y_n \in f(X)$ für alle $n \in \mathbb{N}$. Wir müssen zeigen, daß auch der Grenzwert $y := \lim_{n\to\infty} y_n$ in $f(X)$ enthalten ist. Zu jedem $n \in \mathbb{N}$ gibt es ein $x_n \in X$ mit $f(x_n) = y_n$. Da X beschränkt ist, ist auch die Folge $(x_n)_{n=1}^\infty$ beschränkt und besitzt nach dem Satz 3.34 von Bolzano–Weierstraß eine konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Wegen der Abgeschlossenheit von X ist $x := \lim_{k\to\infty} x_{n_k} \in X$. Wegen der Stetigkeit von f gilt nach dem Folgenkriterium 4.5 und Satz 3.26:

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} y_{n_k} = y$$

und somit $y \in f(X)$. $\square$

4.19. Satz (Satz von der Annahme des Maximums und Minimums). *Sei $\emptyset \neq X \subset \mathbb{R}^M$ abgeschlossen und beschränkt und sei $f : X \to \mathbb{R}$ eine stetige reellwertige Funktion auf X. Dann gibt es $u, v \in X$ mit $f(u) \le f(x) \le f(v)$ für alle $x \in X$. Die Bildmenge $f(X)$ besitzt also ein Minimum und ein Maximum und es gilt:*

$$f(u) = \min f(X) =: \min_{x\in X} f(x) \quad \textit{und} \quad f(v) = \max f(X) =: \max_{x\in X} f(x)\,.$$

Beweis. Nach dem Satz 4.17 von der Beschränktheit ist $f(X)$ beschränkt und nicht leer (wegen $X \neq \emptyset$). Nach dem Supremumsaxiom existieren $m := \inf f(X)$ und $M := \sup f(X)$ in $\mathbb{R}$. Nach Lemma 1.16 gibt es also zu jedem $n \in \mathbb{N}$ Punkte $x_n, y_n \in X$ mit

$$m \le f(x_n) < m + \frac{1}{n} \quad \text{und} \quad M - \frac{1}{n} < f(y_n) \le M\,. \tag{4.8}$$

Da die Menge X und damit auch die Folgen $(x_n)_{n=1}^\infty$ und $(y_n)_{n=1}^\infty$ in $\mathbb{R}^N$ beschränkt sind, besitzen diese nach dem Satz 3.34 von Bolzano–Weierstraß konvergente Teilfolgen $(x_{n_1(k)})_{k=1}^\infty$ und $(y_{n_2(k)})_{k=1}^\infty$. Da X auch abgeschlossen ist, folgt $u := \lim_{k\to\infty} x_{n_1(k)} \in X$ und $v := \lim_{k\to\infty} y_{n_2(k)} \in X$. Wegen (4.8) und dem Folgenkriterium 4.5 erhalten wir

$$m = \lim_{k\to\infty} f(x_{n_1(k)}) = f(u)\,, \quad M = \lim_{k\to\infty} f(y_{n_2(k)}) = f(v)$$

und damit die Behauptung. $\square$

Aus Satz 4.19 erhalten wir mit Hilfe des Zwischenwertsatzes 4.12:

4.20. Folgerung. *Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a,b] \to \mathbb{R}$ eine auf $[a,b]$ stetige reellwertige Funktion. Sind $u, v \in [a,b]$ wie in Satz 4.19, so gilt $f([a,b]) = [f(u), f(v)]$.*

Stetige Bilder von abgeschlossenen, beschränkten Intervallen sind also wieder abgeschlossene, beschränkte Intervalle.

3. π

4.21. Satz. $\sin(\mathbb{R}) = \cos(\mathbb{R}) = [-1, 1]$.

Beweis. Wegen $\sin(x)^2 + \cos(x)^2 = 1$ ist $\sin(\mathbb{R}) \cup \cos(\mathbb{R}) \subseteq [-1,1]$. Ferner ist $\cos(0) = 1$ und wegen

$$(-1)^{2k-1}\frac{2^{2(2k-1)}}{(2(2k-1))!} + (-1)^{2k}\frac{2^{2(2k)}}{(2(2k))!} = \frac{2^{4k}}{(4k)!} - \frac{2^{4k-2}}{(4k-2)!} = \frac{2^{4k-2}}{(4k)!}(4 - (4k-1)4k) < 0$$

für alle $k \in \mathbb{N}$ folgt

$$\cos(2) = \sum_{n=0}^\infty (-1)^n \frac{2^{2n}}{(2n)!} < 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = -\frac{1}{3} < 0.$$

Nach dem Zwischenwertsatz 4.12 gibt es daher ein $t \in (0,2)$ mit $\cos(t) = 0$, d.h.

$$M := \{s \in (0,2)\,;\, \cos(s) = 0\} \neq \emptyset\,.$$

Nach dem Supremumsaxiom existiert $t_0 = \inf M$. Wegen der Stetigkeit der Cosinusfunktion und wegen $\cos(0) = 1$ ist $t_0 > 0$ und $0 < \cos(t) \le 1$ für alle $t \in [0, t_0)$. Ferner gilt wegen $t_0 = \inf M$ nach Lemma 1.16:

$$\forall n \in \mathbb{N}\, \exists t_n \in M : \qquad t_0 \le t_n < t_0 + \frac{1}{n}\,.$$

Es folgt $t_n \to t_0$ für $n \to \infty$ und daher wegen der Stetigkeit der Cosinusfunktion

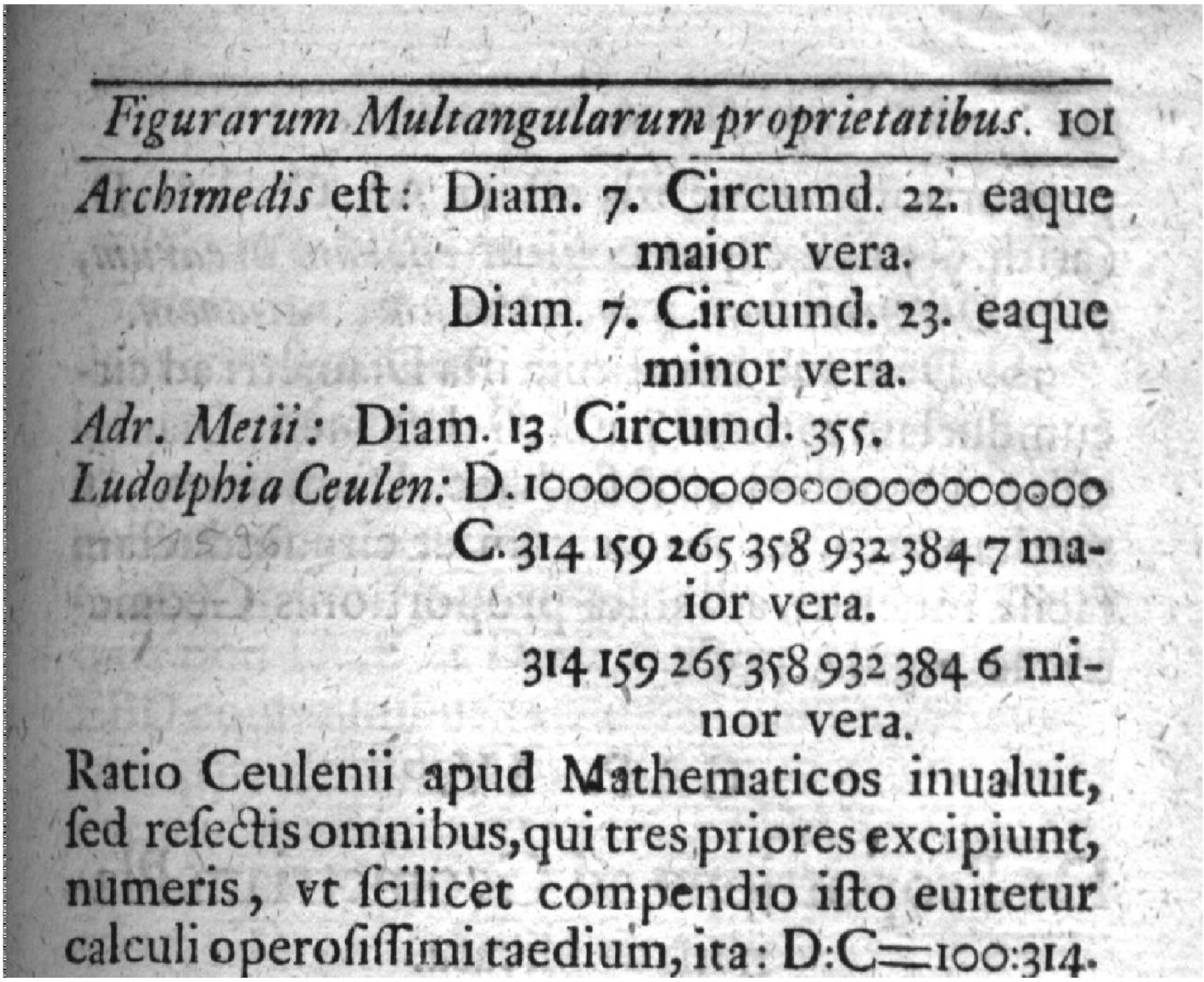
$$\cos(t_0) = \lim_{n\to\infty} \cos(t_n) = 0\,, \text{d.h. } t_0 = \min M\,.$$

t_0 ist also die kleinste positive Nullstelle der Cosinusfunktion. Wegen $\sin(x)^2 + \cos(x)^2 = 1$ ist $\sin(t_0) \in \{-1, +1\}$. Unter Beachtung von $\sin(-t_0) = -\sin(t_0)$ folgt nach dem Zwischenwertsatz 4.12, daß $\sin([-t_0, t_0]) = [-1, 1]$ und daher auch $\sin(\mathbb{R}) = [-1, 1]$ gilt. Aus dem Additionstheorem für den Cosinus folgt

$$\cos(2t_0) = \cos(t_0 + t_0) = \cos(t_0)^2 - \sin(t_0)^2 = -1$$

und hieraus unter abermaliger Verwendung des Zwischenwertsatzes $\cos([0, 2t_0]) = [-1, 1]$, also auch $\cos(\mathbb{R}) = [-1, 1]$. $\square$

4.22. Definition. Wir definieren $\pi := 2t_0 = 2\inf\{t \in (0, 2)\,;\, \cos(t) = 0\}$. [2]



Figurarum Multangularum proprietatibus. 101

Archimedis eſt: Diam. 7. Circumd. 22. eaque maior vera.
Diam. 7. Circumd. 23. eaque minor vera.

Adr. Metii: Diam. 13 Circumd. 355.

Ludolphi a Ceulen: D. 100000000000000000000
C. 314 159 265 358 932 384 7 maior vera.
314 159 265 358 932 384 6 minor vera.

Ratio Ceulenii apud Mathematicos inualuit, ſed reſectis omnibus, qui tres priores excipiunt, numeris, vt ſcilicet compendio iſto euitetur calculi operoſiſſimi taedium, ita: D:C=100:314.

Abbildung 1. Ausschnitt aus Io. Augusto Ernesti: *Initia Doctrinae Solidioris, Pars Prima: Arithmeticam, Geometriam, Psychologiam et Ontologiam*, Leipzig, 1734.

[2] Archimedes von Syrakus (287 v.Chr.–212 v.Chr.) zeigte $223/71 < \pi < 22/7$. Adriaan Metius (9.12.1571–1635) publizierte diese auf seinen Vater zurückgehende (aber auch schon zuvor bekannte) Näherung 355/113 im Jahr 1585. Ludolph van Ceulen (28.1.1540–31.12.1610) berechnete π sogar auf 35 Stellen genau. Deshalb findet man in der älteren deutschen Literatur häufig auch die Bezeichnung *Ludolphsche Zahl* für π. Zur Geschichte von π siehe auch [1]

4.23. Folgerungen. (a) $\sin(\frac{\pi}{2}) = 1$, $\sin(-\frac{\pi}{2}) = -1$.

(b) *Für alle* $x \in \mathbb{R}$ *gilt* $\sin(x + \frac{\pi}{2}) = \cos(x)$ *und* $\cos(x + \frac{\pi}{2}) = -\sin(x)$.

(c) *Für alle* $x \in \mathbb{R}$ *gilt* $\sin(x + \pi) = -\sin(x)$ *und* $\cos(x + \pi) = -\cos(x)$.

(d) *Für alle* $x \in \mathbb{R}$ *gilt* $\sin(x + 2\pi) = \sin(x)$ *und* $\cos(x + 2\pi) = \cos(x)$.

(e) $\frac{\pi}{2}$ *und* $\frac{3}{2}\pi$ *sind die einzigen Nullstellen von* cos *in* $[0, 2\pi]$.

(f) $0, \pi, 2\pi$ *sind die einzigen Nullstellen von* sin *in* $[0, 2\pi]$.

Beweis. (a) Für $0 < x < \sqrt{6}$ und alle $k \in \mathbb{N}$ gilt

$$\begin{aligned}\frac{x^{2\cdot 2k+1}}{(2\cdot 2k+1)!} - \frac{x^{2(2k+1)+1}}{(2(2k+1)+1)!} &= \frac{x^{4k+1}}{(4k+1)!} - \frac{x^{4k+3}}{(4k+3)!} \\ &= \frac{x^{4k+1}}{(4k+3)!}((4k+2)(4k+3) - x^2) > 0\end{aligned}$$

und daher

$$\begin{aligned}\sin(x) &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ &= x - \frac{x^3}{3!} + \sum_{k=1}^{\infty} \left(\frac{x^{2\cdot 2k+1}}{(2\cdot 2k+1)!} - \frac{x^{2(2k+1)+1}}{(2(2k+1)+1)!} \right) > \left(1 - \frac{x^2}{6}\right) > 0\,.\end{aligned}$$

Wegen $\sin(\pi/2)^2 = 1$ und $0 < \frac{\pi}{2} < 2$ folgt $\sin(\pi/2) = 1$.

Die Aussagen (b), (c) und (d) erhält man mit Hilfe der Additionstheoreme unter Beachtung von $\sin(\pi/2) = 1$ und $\cos(\pi/2) = 0$.

(e) Nach dem Beweis zu Satz 4.21 gilt für alle $x \in [0, \pi/2)$: $0 < \cos(x) \leq 1$. Wegen $\cos(x) = \cos(-x)$ folgt $0 < \cos(x) \leq 1$ für alle $x \in (-\pi/2, 0]$ und hieraus mit (c) und (d): $-1 \leq \cos(x) < 0$ für $x \in (\pi/2, 3\pi/2)$ sowie $0 < \cos(x) \leq 1$ für alle $x \in (3\pi/2, 2\pi)$. Damit folgt die Behauptung.

(f) erhält man mit Hilfe von (b) aus (e). □

4.24. Definition. Für eine reelle Zahl T heißt eine Funktion $f : \mathbb{R} \to \mathbb{R}$ *periodisch mit der Periode* T, falls für alle $x \in \mathbb{R}$ gilt: $f(x + T) = f(x)$.

4.25. Bemerkung. Ist $f : \mathbb{R} \to \mathbb{R}$ periodisch von der Periode $T \in \mathbb{R}$ und von der Periode $T_1 \in \mathbb{R}$, so ist f auch periodisch von der Periode $T + T_1$ und von der Periode nT für alle $n \in \mathbb{Z}$.

Beweis. Für alle $x \in \mathbb{R}$ gilt: $f(x+T+T_1) = f((x+T)+T_1) = f(x+T) = f(x)$ und $f(x - T) = f((x - T) + T)) = f(x)$. Hiermit erhält man durch vollständige Induktion auch die zweite Behauptung. □

4.26. Satz. 2π *ist die kleinste positive Periode von* sin *und* cos.

Beweis. Ist $0 < T \leq 2\pi$ eine Periode für sin, so folgt $\sin(T) = \sin(0+T) = \sin(0) = 0$ und daher nach Folgerung 4.23 (f): $T = \pi$ oder $T = 2\pi$. Wegen $\sin(-\frac{\pi}{2}) = -1 \neq 1 = \sin(-\frac{\pi}{2} + \pi)$ ist π keine Periode der Sinusfunktion. Also muß $T = 2\pi$ gelten. Wegen Folgerung 4.23 (b) haben die Cosinusfunktion und die Sinusfunktion die gleichen Perioden. Damit folgt die Behauptung auch für die Cosinusfunktion. □

4.27. Folgerung. *Für die Nullstellenmengen $N_{\sin}$ und $N_{\cos}$ von* $\sin : \mathbb{R} \to \mathbb{R}$ *und* $\cos : \mathbb{R} \to \mathbb{R}$ *gilt:*

$$N_{\sin} = \pi\mathbb{Z} = \{n\pi \,;\, n \in \mathbb{Z}\} \quad \textit{und} \quad N_{\cos} = \{\frac{\pi}{2} + n\pi \,;\, n \in \mathbb{Z}\}\,.$$

Beweis. Dies folgt wegen der 2π–Periodizität der Sinus– und der Cosinusfunktion aus den Folgerungen 4.23 (e) und (f). □

4. Stetigkeit der Umkehrfunktionen

Wir beginnen mit

4.28. Satz. *Sei $\emptyset \neq X \subset \mathbb{R}^M$ beschränkt und abgeschlossen und sei $f : X \to \mathbb{R}^N$ stetig und injektiv, also bijektiv als Abbildung von X nach $f(X)$. Dann ist die nach Satz 10.14 existierende Umkehrabbildung $f^{-1} : f(X) \to X \subset \mathbb{R}^M$ ebenfalls stetig.*

Beweis. Sei $y \in f(X)$ beliebig und sei $(y_n)_{n=1}^{\infty}$ eine beliebige gegen y konvergente Folge aus $f(X)$ und sei u ein beliebiger Häufungspunkt der wegen der Beschränktheit von X beschränkten Folge $(f^{-1}(y_n))_{n=1}^{\infty}$. Dann gibt es eine gegen u konvergente Teilfolge $(f^{-1}(y_{n_k}))_{k=1}^{\infty}$. Wegen der Stetigkeit von f gilt nach dem Folgenkriterium 4.5 (a)

$$f(u) = \lim_{j\to\infty} f\left(f^{-1}(y_{n_j})\right) = \lim_{j\to\infty} y_{n_1(j)} = y\,.$$

Wegen der Injektivität von f ist $u = f^{-1}(y)$ der einzige Häufungswert von $(f^{-1}(y_n))_{n=1}^{\infty}$. Nach Satz 3.33 folgt $f^{-1}(y) = \lim_{n\to\infty} f^{-1}(y_n)$ und hieraus nach dem Folgenkriterium 4.5 (a) die Stetigkeit von f^{-1} in y. □

4.29. Definition. Sei $\emptyset \neq X \subseteq \mathbb{R}$. Eine reellwertige Funktion $f : X \to \mathbb{R}$ heißt

- *monoton wachsend* (bzw. *monoton fallend*), falls für alle $x, y \in X$ mit $x \leq y$ gilt: $f(x) \leq f(y)$ (bzw. $f(x) \geq f(y)$).
- *monoton*, falls sie monoton wachsend oder monoton fallend ist.
- *streng monoton wachsend* (bzw. *streng monoton fallend*), falls für alle $x, y \in X$ mit $x < y$ gilt: $f(x) < f(y)$ (bzw. $f(x) > f(y)$).
- *streng monoton*, falls f auf X streng monoton wachsend oder streng monoton fallend ist.

4.30. Lemma. *Jede streng monotone Funktion ist injektiv.*

Beweis. Sind $a, b \in X$ mit $a \neq b$, so ist $a < b$ oder $a > b$. In beiden Fällen folgt aus der strengen Monotonie $f(a) \neq f(b)$. Also ist f injektiv. □

4.31. Satz. (a) *Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : I := [a, b] \to \mathbb{R}$ eine streng monoton wachsende (bzw. fallende) stetige Funktion. Dann ist*

$$f([a,b]) = [f(a), f(b)] \qquad (\textit{bzw.} \qquad f([a,b]) = [f(b), f(a)]\,) \tag{4.9}$$

und die Abbildung $f : [a, b] \to f([a, b])$ ist bijektiv. Die Umkehrfunktion $f^{-1} : f([a, b]) \to [a, b]$ ist ebenfalls stetig und streng monoton wachsend (bzw. fallend).

(b) *Sei $-\infty \leq a < b \leq \infty$ und sei $f : I := (a, b) \to \mathbb{R}$ eine streng monoton wachsende (bzw. fallende) stetige Funktion. Dann ist*

$$f((a,b)) = (m_f, M_f) \qquad \textit{mit } m_f := \inf_{x\in I} f(x) \quad \textit{und} \quad M_f := \sup_{x\in I} f(x) \tag{4.10}$$

und die Abbildung $f : (a, b) \to f((a, b))$ ist bijektiv. Die Umkehrfunktion $f^{-1} : f((a, b)) \to (a, b)$ ist ebenfalls stetig und streng monoton wachsend (bzw. fallend).

(c) *Sei* $-\infty \le a < b < \infty$ *und sei* $f : I := (a,b] \to \mathbb{R}$ *eine streng monoton wachsende (bzw. fallende) stetige Funktion. Dann ist*

$$f((a,b]) = (m_f, f(b)] \qquad (\textit{bzw.} \qquad f((a,b]) = [f(b), M_f)\,) \tag{4.11}$$

mit $m_f := \inf_{x\in I} f(x)$ *und* $M_f := \sup_{x\in I} f(x)$ *und die Abbildung* $f : (a,b] \to f((a,b])$ *ist bijektiv. Die Umkehrfunktion* $f^{-1} : f((a,b]) \to (a,b]$ *ist ebenfalls stetig und streng monoton wachsend (bzw. fallend).*

(d) *Sei* $-\infty < a < b \le \infty$ *und sei* $f : I := [a,b) \to \mathbb{R}$ *eine streng monoton wachsende (bzw. fallende) stetige Funktion. Dann ist*

$$f([a,b)) = [f(a), M_f) \qquad (\textit{bzw.} \qquad f([a,b)) = (m_f, f(a)]\,) \tag{4.12}$$

mit $m_f := \inf_{x\in I} f(x)$ *und* $M_f := \sup_{x\in I} f(x)$ *und die Abbildung* $f : [a,b) \to f([a,b))$ *ist bijektiv. Die Umkehrfunktion* $f^{-1} : f([a,b)) \to [a,b)$ *ist ebenfalls stetig und streng monoton wachsend (bzw. fallend).*

Beweis. In jedem der vier Fälle (a)–(d) folgt nach Lemma 4.30 aus der strengen Monotonie von f die Injektivität von f und daher die Bijektivität von f als Abbildung von I auf $f(I)$. Man überprüft unmittelbar, daß die nach Satz 10.14 existierende Umkehrabbildung $f^{-1} : f(I) \to I \subset \mathbb{R}$ wieder streng monoton wachsend (bzw. fallend) ist.

Zu (a): Wegen der strengen Monotonie ist $f(a) = \min f([a,b]) < f(b) = \max f([a,b])$ (bzw. $f(a) = \max f([a,b]) > f(b) = \min f([a,b])$). Nach Folgerung 4.20 gilt also (4.9). Nach Satz 4.28 ist f^{-1} stetig auf $f([a,b])$.

Zu (b), (c) und (d): In jedem der drei Fälle ist

$$I = \bigcup\{[\alpha,\beta]\,;\, \alpha,\beta \in I, \alpha < \beta\}$$

und damit auch

$$f(I) = \bigcup\{f([\alpha,\beta])\,;\, \alpha,\beta \in I, \alpha < \beta\}\,.$$

Hieraus erhält man leicht die Aussagen (4.10)–(4.12). Zum Nachweis der Stetigkeit genügt es nach Lemma 4.4 die Stetigkeit von f^{-1} auf jedem beschränkten abgeschlossenen Teilintervall des Intervalls $f(I)$ zu zeigen. Seien also $u, v \in f(I)$ mit $u < v$. Dann ist $f^{-1}([u,v]) = [f^{-1}(u), f^{-1}(v)]$ (bzw. $f^{-1}([u,v]) = [f^{-1}(v), f^{-1}(u)]$) ein beschränktes abgeschlossenes Intervall in $\mathbb{R}$ und es gilt $f^{-1}|_{[u,v]} = (f|_{f^{-1}([u,v])})^{-1}$. Wegen (a) ist $f^{-1}|_{[u,v]}$ stetig. □

4.32. Satz. *Sei I ein nicht ausgeartetes Intervall (also wie in (a), (b), (c) oder (d) in Satz 4.31). Genau dann ist eine stetige Funktion $f : I \to \mathbb{R}$ injektiv, wenn sie streng monoton ist.*

Beweis. (a) Wir zeigen zunächst: *Ist $f : I \to \mathbb{R}$ eine stetige Funktion, für die Punkte* $t_1, t_2, t_3 \in I$ existieren mit

$$t_1 < t_2 < t_3 \quad \textit{und } f(t_2) < m := \min\{f(t_1), f(t_3)\} \tag{4.13}$$

oder mit

$$t_1 < t_2 < t_3 \quad \textit{und } f(t_2) > M := \max\{f(t_1), f(t_3)\}\,, \tag{4.14}$$

so ist f nicht injektiv.

Nach dem Zwischenwertsatz 4.12 gibt es dann nämlich Punkte $s_1 \in [t_1, t_2)$ und $s_2 \in (t_2, t_3]$ mit $f(s_1) = m = f(s_2)$ bzw. mit $f(s_1) = M = f(s_2)$.

(b) Zum Beweis des Satzes genügt es wegen Lemma 4.30 die Richtung „$\Longrightarrow$“ zu beweisen. Sei also $f : I \to \mathbb{R}$ stetig und injektiv.

Annahme: f ist nicht streng monoton. Dann gibt es Punkte $x_1, x_2, x_3, x_4 \in I$ mit

$$x_1 < x_2, x_3 < x_4 \quad \text{und} \quad f(x_1) \geq f(x_2) \text{ sowie } f(x_3) \leq f(x_4)\,.$$

Wegen der Injektivität von f ist dann $f(x_1) > f(x_2)$ und $f(x_3) < f(x_4)$. Sei $\alpha := \min\{x_1, x_2, x_3, x_4\}$ und $\beta := \max\{x_1, x_2, x_3, x_4\}$. Dann ist $\alpha < \beta$. Wegen der Injektivität von f gilt $f(\alpha) < f(\beta)$ oder $f(\alpha) > f(\beta)$.

Wir betrachten zunächst den Fall $f(\alpha) < f(\beta)$. Ist $f(x_2) \geq f(\beta)$ und damit $f(\alpha) < f(\beta) \leq f(x_2) < f(x_1)$, so ist (4.14) mit $t_1 = \alpha, t_2 = x_1$ und $t_3 = \beta$ erfüllt. Ist hingegen $f(x_2) < f(\beta)$, so gilt (4.13) mit $t_1 = x_1, t_2 = x_2$ und $t_3 = \beta$. In beiden Situationen erhalten wir wegen (a) einen Widerspruch zur Voraussetzung der Injektivität von f.

Sei nun $f(\alpha) > f(\beta)$. Ist $f(x_3) \geq f(\alpha)$, also $f(\beta) < f(\alpha) \leq f(x_3) < f(x_4)$, so gilt (4.14) mit $t_1 = \alpha, t_2 = x_4$ und $t_3 = \beta$, und im verbleibenden Fall $f(x_3) < f(\alpha)$ ist (4.13) mit $t_1 = \alpha, t_2 = x_3$ und $t_3 = x_4$ erfüllt. Nach (a) erhalten wir jeweils einen Widerspruch zur Voraussetzung der Injektivität von f.

Da die Annahme in allen möglichen Fällen zu einem Widerspruch führte, muß f doch streng monoton wachsend oder streng monoton fallend sein. $\square$

4.33. Beispiel. Sei $n \in \mathbb{N}$. Die Funktion $\sqrt[n]{\cdot} : [0, \infty) \to [0, \infty)$ ist auf $[0, \infty)$ stetig und streng monoton wachsend.

Beweis. Die Funktion $f_n : \mathbb{R} \to \mathbb{R}$ mit $f_n(x) := x^n$ ist nach Folgerung 4.7 stetig auf $\mathbb{R}$, also auch auf $[0, \infty)$. Nach Lemma 1.40 sind f_n und $\sqrt[n]{\cdot}$ streng monoton wachsend auf $[0, \infty)$. Ferner ist $\sup f_n([0, \infty)) \geq \sup\{k^n \,;\, k \in \mathbb{N}\} = \infty$. Mit Satz 4.31 (d) folgt die Stetigkeit von $\sqrt[n]{\cdot} : [0, \infty) \to [0, \infty)$. $\square$

4.34. Satz (Der natürliche Logarithmus). *Die Exponentialfunktion* $\exp : \mathbb{R} \to \mathbb{R}$ *ist auf* $\mathbb{R}$ *streng monoton wachsend und bildet* $\mathbb{R}$ *bijektiv auf* $(0, \infty)$ *ab. Nach Satz 4.31 (a) besitzt* $\exp : \mathbb{R} \to (0, \infty)$ *also eine stetige und streng monoton wachsende Umkehrfunktion. Diese nennt man den natürlichen Logarithmus und bezeichnet sie mit* $\log$ *(oder auch mit* $\ln$*). Es gilt weiter*

(a) $\forall a, b \in (0, \infty) : \quad \log(ab) = \log(a) + \log(b)$.
(b) $\log(e) = 1$ *und* $\log(1) = 0$.
(c) $\log(x) > 0$ *für* $x > 1$ *und* $\log(x) < 0$ *für* $x < 1$.

Beweis. Nach Folgerung 3.69 ist $\exp(\mathbb{R}) \subseteq (0, \infty)$. Wir beachten:

$$\forall x > 0 : \qquad \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \geq 1 + x > 1\,. \tag{4.15}$$

Für alle $a, b \in \mathbb{R}$ mit $a < b$ gilt also $\exp(b - a) - 1 > 0$ und daher

$$\exp(b) - \exp(a) = \exp(a)(\exp(b - a) - 1) > 0\,.$$

Somit ist $\exp$ streng monoton wachsend. Ferner folgt mit (4.15):

$$\sup \exp(\mathbb{R}) \geq \sup \exp(\mathbb{N}) \geq \sup(1 + \mathbb{N}) = \infty\,.$$

und wegen

$$\exp(-n) = \frac{1}{\exp n} \leq \frac{1}{1 + n} \to 0 \text{ für } n \to \infty$$

ist $\inf \exp(\mathbb{R}) = 0$. Nach Satz 4.31 (a) folgt $\exp(\mathbb{R}) = (0, \infty)$ und die Stetigkeit und strenge Monotonie der Umkehrfunktion. Die Funktionalgleichung (a) der so erhaltenen Logarithmusfunktion folgt aus der Funktionalgleichung der Exponentialfunktion. (b) ist offensichtlich und (c) folgt wegen der strengen Monotonie aus (b). $\square$

Wir können nun auch endlich *reelle Exponenten* einführen:
Sei $x > 0$. Wegen $\exp \circ \log = \mathrm{id}_{(0,\infty)}$ gilt

$$x = \exp(\log(x)) .$$

Induktiv zeigt man mit Hilfe der Funktionalgleichung der Exponentialfunktion:

$$\forall n \in \mathbb{N} : \qquad x^n = \exp(n \log(x))$$

Hieraus folgt leicht $x^n = \exp(n \log(x))$ für alle $n \in \mathbb{Z}$. Wegen

$$\forall k \in \mathbb{N} : \qquad \left(\exp\left(\frac{1}{k}\log(x)\right)\right)^k = \exp(\log(x)) = x$$

und der Eindeutigkeit der nicht negativen k–ten Wurzel folgt weiter

$$\forall k \in \mathbb{N} : \qquad \sqrt[k]{x} = \exp\left(\frac{1}{k}\log(x)\right) .$$

Durch vollständige Induktion sieht man schließlich:

$$\forall p \in \mathbb{Z}\, \forall q \in \mathbb{N} : \qquad x^{p/q} = \exp\left(\frac{p}{q}\log(x)\right) .$$

Wir definieren daher für alle $\alpha \in \mathbb{R}$:

$$x^\alpha := \exp(\alpha \log(x)) .$$

Insbesondere folgt für den Spezialfall $x := e$:

$$e^\alpha = \exp(\alpha \log(e)) = \exp(\alpha) .$$

Wegen der Stetigkeit der Exponentialfunktion und der Logarithmusfunktion sind auch die Abbildungen $x \mapsto x^\alpha = \exp(\alpha \log(x))$ und $\alpha \mapsto x^\alpha = \exp(\alpha \log(x))$ stetig auf $(0, \infty)$ bzw. auf $\mathbb{R}$.

5. Gleichmäßige Stetigkeit

4.35. Definition. Sei $\emptyset \neq X \subseteq \mathbb{R}^M$. Eine Funktion $f : X \to \mathbb{R}^N$ heißt *gleichmäßig stetig auf* X, falls gilt

$$\forall \varepsilon > 0\, \exists \delta = \delta(\varepsilon) > 0\, \forall x, y \in X : \qquad |x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon . \tag{4.16}$$

δ ist hierbei also unabhängig von x und y.

4.36. Bemerkungen. (a) Offensichtlich ist jede gleichmäßig stetige Funktion insbesondere auch stetig. In den folgenden Beispielen werden wir sehen, daß es stetige Funktionen gibt, die nicht gleichmäßig stetig sind.

(b) Durch korrektes Negieren von (4.16) erhält man: f ist nicht gleichmäßig stetig genau dann, wenn gilt

$$\exists \varepsilon > 0\, \forall \delta > 0\, \exists x, y \in X : \qquad |x - y| < \delta \quad \text{und} \quad |f(x) - f(y)| \geq \varepsilon . \tag{4.17}$$

4.37. Beispiele. (a) Sei $\emptyset \neq X \subseteq \mathbb{R}^M$. Genügt eine Funktion $f : X \to \mathbb{R}^N$ einer Lipschitz–Bedingung mit einer Lipschitz–Konstante L, so ist f schon gleichmäßig stetig auf X.

(b) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, ist auf $\mathbb{R}$ stetig (vergl. Beispiel 4.2 (c)) aber nicht gleichmäßig stetig.

(c) Die Funktion $\sqrt{\cdot} : [0, \infty)$, $x \mapsto \sqrt{x}$, ist auf $[0, \infty)$ gleichmäßig stetig, genügt aber auf $[0, \infty)$ keiner Lipschitz–Bedingung.

Beweis. (a) Sei $\varepsilon > 0$ beliebig. Nach Voraussetzung gilt

$$\forall x, y \in X : |f(x) - f(y)| \le L|x - y|\,.$$

Wählen wir als $\delta := \varepsilon/(L+1)$, so folgt für alle $x, y \in X$ mit $|x-y| < \delta$: Es ist $|f(x)-f(y)| \le L\delta < \varepsilon$. (4.16) ist also erfüllt und f ist daher gleichmäßig stetig.

(b) Für alle $x, y \in \mathbb{R}$ gilt: $f(x) - f(y) = x^2 - y^2 = (x+y)(x-y)$. Sei $\varepsilon_0 := 1$ und $\delta > 0$ beliebig. Dann gibt es ein $n \in \mathbb{N}$ mit $n\delta > 1 = \varepsilon_0$. Für $x := n + \delta/2$ und $y := n$ gilt dann $|x - y| = \delta/2 < \delta$ und

$$|f(x) - f(y)| = |(x+y)(x-y)| = \left(2n + \frac{\delta}{2}\right)\frac{\delta}{2} > n\delta > 1 = \varepsilon_0\,.$$

Also ist (4.17) erfüllt und f daher nicht gleichmäßig stetig auf $\mathbb{R}$.

(c) Sei $L > 0$ beliebig. Wegen $|\sqrt{x} - \sqrt{0}| = \sqrt{x} = \dfrac{x}{\sqrt{x}} > |x - 0|L$ für alle $x > 0$ mit $x < L^{-2}$ kann $\sqrt{\cdot}$ auf $[0, \infty)$ keiner Lipschitz–Bedingung mit Lipschitz–Konstante L genügen. Auch $L = 0$ ist trivialerweise nicht möglich.

Wir zeigen die gleichmäßige Stetigkeit von $\sqrt{\cdot}$ auf $[0, \infty)$: Sei $\varepsilon > 0$ beliebig. Mit $\delta := \varepsilon^2$ gilt für alle $x, y \ge 0$ mit $|x - y| < \delta$ (wobei wir o.B.d.A. $y \ge x$ annehmen können)

$$(\sqrt{y} - \sqrt{x})^2 \le (\sqrt{y} - \sqrt{x})(\sqrt{y} + \sqrt{x}) = y - x = |x - y| < \delta = \varepsilon^2$$

und daher wegen der strengen Monotonie der Wurzel: $|\sqrt{x} - \sqrt{y}| = \sqrt{y} - \sqrt{x} < \sqrt{\delta} = \varepsilon$. Also ist (4.16) erfüllt und $\sqrt{\cdot}$ daher gleichmäßig stetig auf $[0, \infty)$. □

Es gilt jedoch:

4.38. Satz. *Sei $\emptyset \ne X \subset \mathbb{R}^M$ abgeschlossen und beschränkt. Dann ist jede auf X stetige Funktion $f : X \to \mathbb{R}^N$ schon gleichmäßig stetig auf X.*

Beweis. Annahme: Es gibt eine stetige Funktion $f : X \to \mathbb{R}^N$, welche auf X stetig aber nicht gleichmäßig stetig ist, für welche also ein $\varepsilon_0 > 0$ existiert mit

$$\forall \delta > 0\, \exists x, y \in X : \qquad |x - y| < \delta \quad \text{und} \quad |f(x) - f(y)| \ge \varepsilon_0\,.$$

Insbesondere gibt es für jedes $n \in \mathbb{N}$ zu $\delta_n := 1/n$ Punkte $x_n, y_n \in X$ mit $|x_n - y_n| < \delta_n = 1/n$ und $|f(x_n) - f(y_n)| \ge \varepsilon_0$. Da X beschränkt ist, besitzt die Folge $(x_n)_{n=1}^\infty$ eine konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Da X auch abgeschlossen ist, folgt $x_0 := \lim_{k\to\infty} x_{n_k} \in X$. Wegen

$$|y_{n_k} - x_0| \le \underbrace{|y_{n_k} - x_{n_k}|}_{<1/n_k \le 1/k \to 0} + \underbrace{|x_{n_k} - x_0|}_{\to 0} \qquad \text{für } k \to \infty$$

folgt $y_{n_k} \to x_0$ für $k \to \infty$. Da f in x_0 stetig ist, gilt nach dem Folgenkriterium 4.5 $\lim_{k\to\infty} f(y_{n_k}) = f(x_0) = \lim_{k\to\infty} f(x_{n_k})$. Dies führt zu dem Widerspruch $\varepsilon_0 \le |f(y_{n_k}) - f(x_{n_k})| \to 0$ für $k \to \infty$. Die Annahme war daher falsch. Jede auf X stetige Funktion $f : X \to \mathbb{R}^N$ muß also auf X gleichmäßig stetig sein. □

6. Grenzwerte bei Funktionen

4.39. Lemma. *Sei $\emptyset \ne X \subseteq \mathbb{R}^M$ und sei $a \in \mathbb{R}^M$ ein Häufungspunkt der Menge X. Sei $f : X \to \mathbb{R}^N$ eine Funktion auf X mit Werten in $\mathbb{R}^N$.*

(a) *Ist $y \in \mathbb{R}^N$, so ist die durch*

$$g_y(x) := \begin{cases} f(x) & \text{für } x \in X \setminus \{a\} \\ y & \text{für } x = a \end{cases}$$

definierte Funktion $g_y : X \cup \{a\} \to \mathbb{R}^N$ *genau dann stetig in* a*, wenn gilt*

$$\forall \varepsilon > 0\, \exists \delta > 0 \forall x \in X \setminus \{a\} : \qquad |x - a| < \delta \implies |f(x) - y| < \frac{\varepsilon}{2}\,. \tag{4.18}$$

(b) *Es gibt höchstens ein* $y \in \mathbb{R}^N$*, so daß die Funktion* g_y *aus (a) in* a *stetig ist.*

Beweis. (a) folgt in naheliegender Weise aus der Definition der Stetigkeit.

(b) Sind y_1 und y_2 zwei Punkte aus $\mathbb{R}^N$, so daß die Funktionen g_{y_1} und g_{y_2} in a stetig sind, so gilt nach (a) für beide Funktionen die Bedingung (4.18) für $y = y_1$ bzw. $y = y_2$. Ist also $\varepsilon > 0$ beliebig vorgegeben, so gibt es $\delta_1, \delta_2 > 0$ mit

$$\forall x \in X \setminus \{a\} : \qquad |x - a| < \delta_j \implies |f(x) - y_j| < \varepsilon \quad (j = 1, 2)\,. \tag{4.19}$$

Wir setzen $\delta := \min\{\delta_1, \delta_2\}$. Da a ein Häufungspunkt von X ist gibt es einen Punkt $x_0 \in U_\delta(a) \cap (X \setminus \{a\})$. Wegen (4.19) folgt dann

$$|y_1 - y_2| \le |y_1 - f(x_0)| + |f(x_0) - y_2| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Da die äußere Ungleichung für alle $\varepsilon > 0$ gilt, muß $y_1 = y_2$ gelten. □

4.40. Definition. Sei $\emptyset \ne X \subseteq \mathbb{R}^M$ und sei $a \in \mathbb{R}^M$ ein Häufungspunkt der Menge X. Ist $f : X \to \mathbb{R}^N$ eine $\mathbb{R}^N$–wertige Funktion, so sagen wir f *hat in* a *einen Grenzwert*, falls ein $y \in \mathbb{R}^N$ existiert, so daß die durch

$$g(x) := \begin{cases} f(x) & \text{für } x \in X \setminus \{a\} \\ y & \text{für } x = a \end{cases}$$

definierte Funktion $g : X \cup \{a\} \to \mathbb{R}^N$ stetig in a ist. Nach Lemma 4.39 ist y hierdurch eindeutig bestimmt. Wir nennen y *den Grenzwert von* f *in* a und schreiben

$$y = \lim_{x \to a} f(x) \quad \text{oder} \quad f(x) \to y \text{ für } x \to a\,.$$

4.41. Satz. *Sei* $\emptyset \ne X \subseteq \mathbb{R}^M$ *und sei* $a \in \mathbb{R}^M$ *ein Häufungspunkt der Menge* X*. Ist* $f : X \to \mathbb{R}^N$ *eine* $\mathbb{R}^N$*–wertige Funktion und* $y \in \mathbb{R}^N$*, so sind die folgenden drei Aussagen äquivalent:*

(a) $f(x) \to y$ *für* $x \to a$.

(b) $\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in X \setminus \{a\} : \quad |x - a| < \delta \implies |f(x) - y| < \varepsilon$.

(c) *Für jede Folge* $(x_n)_{n=1}^\infty$ *aus* $X \setminus \{a\}$ *mit* $\lim_{n\to\infty} x_n = a$ *gilt* $f(x_n) \to y$ *für* $n \to \infty$.

Beweis. Die Äquivalenz von (a) und (b) folgt aus der Definition 4.40 in Verbindung mit Lemma 4.39.

Die Implikation (a) $\implies$ (c) folgt mit Satz 4.5 aus der Definition 4.40.

Zu zeigen bleibt noch die Implikation (c) $\implies$ (b): Wir gehen indirekt vor und nehmen an, daß (c) erfüllt ist aber nicht (b), d.h.

$$\exists \varepsilon_0 > 0\, \forall \delta > 0\, \exists x \in X \setminus \{a\} \text{ mit } |x - a| < \delta \text{ und } |f(x) - y| \ge \varepsilon_0\,.$$

Zu $\delta_n := 1/n$ gibt es daher ein $x_n \in X \setminus \{a\}$ mit $|x_n - a| < \delta_n = 1/n$ und $|f(x_n) - y| \ge \varepsilon_0$. Dann folgt $x_n \to a$ für $n \to \infty$ aber $f(x_n) \not\to y$ im Widerspruch zu (c). □

Bemerkung. Es hat keinen Sinn einen Grenzwert in einem isolierten Punkt zu definieren, da jede Funktion in einem isolierten Punkt ihres Definitionsbereichs stetig ist. Der „Grenzwert“ wäre in diesem Fall also nicht eindeutig bestimmt.

Als Folgerung erhalten wir aus Satz 4.41:

4.42. Lemma. *Sei $\emptyset \neq Y \subseteq X \subseteq \mathbb{R}^M$ und sei $a \in \mathbb{R}^M$ ein Häufungspunkt von Y und damit auch von X. Ist $f : X \to \mathbb{R}^M$ eine Funktion für die der Grenzwert $\lim_{x\to a} f(x)$ existiert, so existiert auch der Grenzwert $\lim_{x\to a}(f|_Y)(x)$ und stimmt mit $\lim_{x\to a} f(x)$ überein.*

Beweis. Sei $(x_n)_{n=1}^{\infty}$ eine beliebige Folge aus $Y \setminus \{a\}$ mit $x_n \to a$ für $n \to \infty$. Wegen der Aussage $(a) \implies (c)$ in Satz 4.41 folgt $f(x_n) \to \lim_{x\to a} f(x)$. Da auch die Aussage $(c) \implies (a)$ in Satz 4.41 gilt, erhalten wir die Existenz von $\lim_{x\to a}(f|_Y)(x)$ und $\lim_{x\to a}(f|_Y)(x) = \lim_{x\to a} f(x)$. □

4.43. Beispiele. (a) Sei $f : (0,1] \to \mathbb{R}$ definiert durch

$$f(x) := \begin{cases} 1 & \text{für } x \in \mathbb{Q} \cap (0,1] \\ 0 & \text{für } x \in (0,1] \setminus \mathbb{Q}\,. \end{cases}$$

Dann hat f in dem Häufungspunkt 0 von $(0,1]$ keinen Grenzwert.

(b) $\displaystyle\lim_{z\to 0} \frac{\sin(z)}{z} = 1.$

(c) $\displaystyle\lim_{z\to 0} \frac{\exp(z)-1}{z} = 1$

Beweis. (a) Wegen

$$\lim_{n\to\infty} f\left(\frac{1}{n}\right) = 1 \neq 0 = \lim_{n\to\infty} f\left(\frac{\sqrt{2}}{n}\right)$$

kann f nach Satz 4.41 keinen Grenzwert in 0 haben.

(b) Sei $\varepsilon > 0$ beliebig. Nach den Rechenregeln für unendliche Reihen gilt für alle $z \in \mathbb{C}$ mit $0 < |z| < 1$:

$$\begin{aligned}\left|\frac{\sin(z)}{z} - 1\right| &= \left|\frac{1}{z}\sum_{n=0}^{\infty}(-1)^n \frac{z^{2n+1}}{(2n+1)!} - 1\right| = \left|\sum_{n=1}^{\infty}(-1)^n \frac{z^{2n}}{(2n+1)!}\right| = \\ &= \left|z\sum_{n=1}^{\infty}(-1)^n \frac{z^{2n-1}}{(2n+1)!}\right| \leq |z|\sum_{n=1}^{\infty}\frac{1}{(2n+1)!} < |z|e\,.\end{aligned}$$

Für alle $z \in \mathbb{C}$ mit $0 < |z-0| = |z| < \delta := \min\{1, \varepsilon/e\}$ gilt also

$$\left|\frac{\sin(z)}{z} - 1\right| < \varepsilon\,.$$

Mit Satz 4.41 folgt die Behauptung.

(c) Sei $\varepsilon > 0$ beliebig. Nach den Rechenregeln für unendliche Reihen gilt für alle $z \in \mathbb{C}$ mit $0 < |z| < 1$:

$$\left|\frac{\exp(z)-1}{z} - 1\right| = \left|\frac{1}{z}\sum_{n=1}^{\infty}\frac{z^n}{n!} - 1\right| = \left|\sum_{n=2}^{\infty}\frac{z^{n-1}}{n!}\right| \leq |z|\sum_{n=2}^{\infty}\frac{1}{n!} = |z|(e-2)\,.$$

Für alle $z \in \mathbb{C}$ mit $0 < |z-0| = |z| < \delta := \min\{1, \varepsilon/e\}$ folgt

$$\left|\frac{\exp(z)-1}{z} - 1\right| < \delta(e-2) < \varepsilon\,.$$

Wie in (b) folgt mit Satz 4.41 die Behauptung. □

Aus Satz 4.41 erhalten wir auch unmittelbar die folgende Charakterisierung der Stetigkeit in einem Punkt durch den Grenzwertbegriff.

4.44. Folgerung. *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$ und sei $a \in X$ ein in X liegender Häufungspunkt der Menge X. Eine $\mathbb{R}^N$–wertige Funktion $f : X \to \mathbb{R}^N$ ist genau dann in a stetig, wenn $\lim_{x\to a} f(x) = f(a)$.*

Mit den Grenzwertrechenregeln für Folgen, Satz 4.41 und den Rechenregeln für stetige Funktionen zeigt man in naheliegender Weise die folgenden Rechenregeln für Grenzwerte bei Funktionen:

4.45. Lemma. *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$ und sei $a \in \mathbb{R}^M$ ein Häufungspunkt der Menge X. $f, g : X \to \mathbb{R}^N$ und $k, h : X \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) seien Funktionen, für die die Grenzwerte*

$$\lim_{x\to a} f(x)\,, \quad \lim_{x\to a} g(x)\,, \quad \lim_{x\to a} h(x)\,, \quad \text{und } \lim_{x\to a} k(x)$$

in $\mathbb{R}^N$ bzw. $\mathbb{K}$ existieren. Dann gilt:

(a) *Für alle $\alpha, \beta \in \mathbb{R}$ existiert der Grenzwert $\lim_{x\to a}(\alpha f(x) + \beta g(x))$ und es gilt*
$$\lim_{x\to a}(\alpha f(x) + \beta g(x)) = \alpha \lim_{x\to a} f(x) + \beta \lim_{x\to a} g(x)\,.$$
(b) *Der Grenzwert $\lim_{x\to a}(hk)(x)$ existiert und stimmt mit $\lim_{x\to a} h(x) \cdot \lim_{x\to a} k(x)$ überein. Für alle $\alpha, \beta \in \mathbb{K}$ existiert der Grenzwert $\lim_{x\to a}(\alpha h(x) + \beta k(x))$ und es gilt*
$$\lim_{x\to a}(\alpha h(x) + \beta k(x)) = \alpha \lim_{x\to a} h(x) + \beta \lim_{x\to a} k(x)\,.$$
(c) *Ist speziell $\mathbb{K} = \mathbb{R}$, so existiert der Grenzwert $\lim_{x\to a}(hf)(x)$ und stimmt mit $\lim_{x\to a} h(x) \cdot \lim_{x\to a} f(x)$ überein.*
(d) *Ist speziell $\mathbb{K} = \mathbb{R}$ und gibt es ein $\eta > 0$, so daß für alle $x \in X$ mit $0 < |x - a| < \eta$ gilt $h(x) \leq k(x)$, so ist $\lim_{x\to a} h(x) \leq \lim_{x\to a} k(x)$.*
(e) *Ist $\lim_{x\to a} h(x) \neq 0$, so ist a Häufungspunkt von $X_h := \{x \in X\,;\, h(x) \neq 0\}$. Es existiert dann der Grenzwert $\lim_{x\to a}(1/h)(x)$ und stimmt mit $(\lim_{x\to a} h(x))^{-1}$ überein. Hierbei ist $1/h$ die Funktion*
$$\frac{1}{h} : X_h \to \mathbb{K}\,, \qquad x \mapsto \frac{1}{h}(x) := \frac{1}{h(x)}\,.$$

4.46. Lemma (Grenzwerte bei Hintereinanderausführung von Funktionen). *Sei $\emptyset \neq X \subseteq \mathbb{R}^M$, sei $a \in \mathbb{R}^M$ ein Häufungspunkt von X und sei $\emptyset \neq Y \subseteq \mathbb{R}^N$. Sind $f : X \to \mathbb{R}^N$ und $g : Y \to \mathbb{R}^K$ Funktionen mit $f(X) \subseteq Y$, existiert der Grenzwert $y := \lim_{x\to a} f(x)$ und ist g in dem Punkt y stetig, so existiert auch der Grenzwert $\lim_{x\to a}(g \circ f)(x)$ und es gilt:*

$$\lim_{x\to a}(g \circ f)(x) = g\big(\lim_{x\to a} f(x)\big)\,.$$

Beweis. Sei $(x_n)_{n=1}^\infty$ eine beliebige Folge aus $X \setminus \{a\}$ mit $x_n \to a$ für $n \to \infty$. Da der Grenzwert $y := \lim_{x\to a} f(x)$ nach Voraussetzung existiert folgt nach Satz 4.41: $\lim_{n\to\infty} f(x_n) = y$. Wegen der Stetigkeit von g in y gilt nach Satz 4.5 auch $(g \circ f)(x_n) = g(f(x_n)) \to g(y)$ für $n \to \infty$. Mit Satz 4.41 folgt nun die Behauptung. □

In $\mathbb{R}$ können wir auch einseitige Grenzwerte einführen:

4.47. Definition (Einseitige Grenzwerte). Sei $\emptyset \neq X \subseteq \mathbb{R}$ und sei $f : X \to \mathbb{R}^N$ eine Funktion auf X. Ist $a \in \mathbb{R}$ ein Häufungspunkt von $X \cap (a, \infty)$ (bzw. von $X \cap (-\infty, a)$), so heißt $y \in \mathbb{R}^N$ *rechtsseitiger* (bzw. *linksseitiger*) *Grenzwert von f in a*, falls die Einschränkung $f|_{X\cap(a,\infty)}$ (bzw. $f|_{X\cap(-\infty,a)}$) in a den Grenzwert y besitzt. Wir schreiben dann

$$y = \lim_{x\to a^+} f(x) \text{ oder } y = \lim_{x\searrow a} f(x) \quad \left(\text{bzw. } y = \lim_{x\to a^-} f(x) \text{ oder } y = \lim_{x\nearrow a} f(x)\right).$$

4.48. Beispiel. Die Heavyside–Funktion (vergl. Beispiel 4.2) $h : \mathbb{R} \to \mathbb{R}$ mit

$$h(x) := \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

ist in 0 unstetig. Es gilt $\lim_{x\to 0^-} h(x) = 0$ und $\lim_{x\to 0^+} h(x) = 1$.

Beweis. 0 ist sowohl Häufungspunkt von $(0, \infty)$ als auch von $(-\infty, 0)$. Da h auf $(-\infty, 0)$ identisch verschwindet und auf $(0, \infty)$ identisch 1 ist, folgt $\lim_{x\to 0^-} h(x) = 0$ und $\lim_{x\to 0^+} h(x) = 1$. □

4.49. Satz. *Sei $\emptyset \neq X \subseteq \mathbb{R}$ und sei $f : X \to \mathbb{R}^N$ eine Funktion auf X. Ist $a \in \mathbb{R}$ sowohl Häufungspunkt von $(-\infty, a)\cap X$ als auch von $(a,\infty)\cap X$ so gilt für $y \in \mathbb{R}^N$: f hat in a genau dann den Grenzwert y, wenn die rechtsseitigen und linksseitigen Grenzwerte von f in a existieren und mit y übereinstimmen.*

Beweis. Die Implikation „$\Longrightarrow$“ folgt unmittelbar aus Lemma 4.42.

„$\Longleftarrow$“: Gelte nun $\lim_{x\to a^-} f(x) = y = \lim_{x\to a^+} f(x)$ und sei $\varepsilon > 0$ beliebig vorgegeben. Nach Satz 4.41 gibt es $\delta_-, \delta_+ > 0$, so daß für alle $x \in (-\infty, a) \cap (X \setminus \{a\})$ (bzw. $x \in (a,\infty)\cap(X\setminus\{a\})$) mit $|x-a| < \delta_-$ (bzw. mit $|x-a| < \delta_+$) gilt $|f(x)-y| < \varepsilon$. Dann folgt für alle $x \in X \setminus \{a\}$ mit $|x - a| < \delta := \min\{\delta_-, \delta_+\}$: $|f(x) - y| < \varepsilon$. Wieder mit Satz 4.41 folgt $y = \lim_{x\to a} f(x)$. □

4.50. Definition. Sei $\emptyset \neq X \subseteq \mathbb{R}$ und sei $f : X \to \mathbb{R}^N$ eine Funktion auf X. f heißt in einem Punkt $a \in X$ *linksseitig* (bzw. *rechtsseitig*) *stetig*, falls die eingeschränkte Funktion $f|_{X\cap(-\infty,a]}$ (bzw. $f|_{X\cap[a,\infty)}$) in a stetig ist.

Mit Folgerung 4.44 erhalten wir unmittelbar:

4.51. Folgerung. *Sei $\emptyset \neq X \subseteq \mathbb{R}$ und sei $f : X \to \mathbb{R}^N$ eine Funktion auf X. $a \in X$ sei ein Häufungspunkt von $(-\infty, a) \cap X$ (bzw. von $(a,\infty) \cap X$). Genau dann ist f in a linksseitig (bzw. rechtsseitig) stetig, wenn gilt $\lim_{x\to a^-} f(x) = f(a)$ (bzw. $\lim_{x\to a^+} f(x) = f(a)$).*

Auch die folgenden Aussagen ergeben sich unmittelbar aus Folgerung 4.44, Satz 4.49 und Folgerung 4.51.

4.52. Folgerung. *Sei $\emptyset \neq X \subseteq \mathbb{R}$ und sei $a \in X$, so daß a sowohl Häufungspunkt von $(-\infty, a)\cap X$ als auch von $(a,\infty)\cap X$ (und damit auch von X) ist. Für eine Funktion $f : X \to \mathbb{R}^N$ sind die folgenden vier Aussagen äquivalent:*

(a) *f ist in a stetig.*
(b) $\lim_{x\to a} f(x) = f(a)$.
(c) $\lim_{x\to a^-} f(x) = f(a) = \lim_{x\to a^+} f(x)$.
(d) *f ist in a linksseitig und rechtsseitig stetig.*

Wir führen nun noch Grenzwerte für $x \to \infty$ und $x \to -\infty$ sowie uneigentliche Grenzwerte ein:

4.53. Definition. Sei $\emptyset \neq X \subseteq \mathbb{R}$ so daß wenigstens eine Folge $(x_n)_{n=1}^{\infty}$ in X existiert mit $x_n \to \infty$ (bzw. $x_n \to -\infty$) für $n \to \infty$. Sei $f : X \to \mathbb{R}^N$ eine Funktion. Ein Punkt $y \in \mathbb{R}^n$ heißt *Grenzwert von f für $x \to \infty$ (bzw. für $x \to -\infty$)*, falls gilt:

$$\forall \varepsilon > 0\, \exists C = C(\varepsilon) > 0\, \forall x \in X : \quad x > C \Longrightarrow |f(x) - y| < \varepsilon$$
$$(\text{bzw. } x < -C \Longrightarrow |f(x) - y| < \varepsilon)\,.$$

Man überlegt sich leicht, daß dies genau dann der Fall ist, wenn für jede Folge $(x_n)_{n=1}^{\infty}$ aus X mit $x_n \to \infty$ (bzw. $x_n \to -\infty$) für $n \to \infty$ gilt: $f(x_n) \to y$ für $n \to \infty$. Auch die Grenzwertrechenregeln übertragen sich sinngemäß.

4.54. Definition. Sei $\emptyset \neq X \subseteq \mathbb{R}^M$ und sei $a \in \mathbb{R}^M$ ein Häufungspunkt der Menge X. Für eine reellwertige Funktion $f : X \to \mathbb{R}$ definieren wir: $\lim_{x\to a} f(x) = \infty$ oder $f(x) \to \infty$ für $x \to a$, falls

$$\forall C > 0\, \exists \delta = \delta(C) > 0\, \forall x \in X \setminus \{a\} : \quad |x - a| < \delta \Longrightarrow f(x) > C\,.$$

Man überlegt sich, daß dies äquivalent ist zu: Für jede gegen a konvergente Folge $(x_n)_{n=1}^{\infty}$ aus $X \setminus \{a\}$ gilt $f(x_n) \to \infty$ für $n \to \infty$.

Analog definiert man Konvergenz gegen $-\infty$ für $x \to a$.

Auch in der Situation von Definition 4.53 kann man uneigentliche Grenzwerte einführen. Man definiert z.B.: $f(x) \to \infty$ für $x \to \infty$, falls gilt

$$\forall C > 0\, \exists K > 0\, \forall x \in X : \quad x > K \Longrightarrow f(x) > C\,.$$

Dies ist äquivalent zu: $f(x_n) \to \infty$ für jede Folge $(x_n)_{n=1}^{\infty}$ aus X mit $x_n \to \infty$ für $n \to \infty$.

Analog definiert man Konvergenz gegen $-\infty$ für $x \to \infty$ und Konvergenz gegen ∞ (bzw. $-\infty$) für $x \to -\infty$.

4.55. Beispiele (Wachstumsverhalten der Exponentialfunktion).

(a) $\displaystyle\lim_{x\to\infty} \frac{\exp(x)}{x^k} = \infty$ für alle $k \in \mathbb{N}_0$.
(b) $\displaystyle\lim_{x\to\infty} \exp(-x)x^k = 0$ für alle $k \in \mathbb{N}_0$.

Beweis. (a) Sei also $C > 0$ beliebig vorgegeben. Für alle $x > K := C \cdot (k+1)!$ gilt

$$\frac{\exp(x)}{x^k} = \frac{1}{x^k} \sum_{n=0}^{\infty} \frac{x^n}{n!} > \frac{1}{x^k} \cdot \frac{x^{k+1}}{(k+1)!} = \frac{x}{(k+1)!} > C\,.$$

Damit folgt die Behauptung.

(b) Sei $\varepsilon > 0$ beliebig. Für alle $x > C := \frac{(k+1)!}{\varepsilon}$ gilt:

$$|\exp(-x)x^k| = \frac{x^k}{\exp(x)} = \frac{x^k}{\sum_{n=0}^{\infty} \frac{x^n}{n!}} \leq \frac{x^k(k+1)!}{x^{k+1}} < \varepsilon\,.$$

Nach Definition 4.53 folgt die Behauptung. □

4.56. Beispiele (Wachstumsverhalten der Logarithmusfunktion).

(a) Es gilt $\log(x) \to \infty$ für $x \to \infty$ und $\log(x) \to -\infty$ für $x \to 0^+$.
(b) Für alle $\alpha > 0$ gilt

$$\text{(i) } \lim_{x\to\infty} \frac{\log(x)}{x^\alpha} = 0. \qquad \text{(ii) } \lim_{x\to 0} x^\alpha \log(x) = 0.$$

Beweis. (a) Da $\log : (0, \infty) \to \mathbb{R}$ bijektiv ist, gibt es zu beliebig vorgegebenem $C > 0$ ein $K > 0$ und ein $\delta > 0$ mit $\log(K) = C$ und $\log(\delta) = -C$. Wegen der strengen Monotonie der Logarithmusfunktion gilt dann $\log(x) > \log(K) = C$ für alle reellen $x > K$ und $\log(x) < \log(\delta) = -C$ für alle $x \in (0, \delta)$. Hiermit folgen die Behauptungen.

(b) Nach Definition von x^α gilt:

$$\frac{\log(x)}{x^\alpha} = \frac{1}{\alpha} \exp(-\alpha \log(x)) \cdot \alpha \cdot \log(x)\,. \tag{4.20}$$

Nach Beispiel 4.55 (b) gibt es ein $y_0 > 0$, so daß

$$\forall y > y_0 : \quad |\exp(-y) \cdot y| < \alpha\varepsilon\,. \tag{4.21}$$

Nach (a) gibt es zu $C := y_0/\alpha$ ein $x_0 > 0$, so daß für alle $x > x_0$ gilt $\log(x) > y_0/\alpha$. Wegen (4.20) und (4.21) gilt dann

$$\forall x > x_0 : \left| \frac{\log(x)}{x^\alpha} - 0 \right| < \varepsilon\,.$$

Hiermit ist (i) bewiesen. Der Beweis zu (ii) wird in den Übungen ausgeführt. □

Der folgende Satz wird uns gelegentlich hilfreich sein:

4.57. Satz. *Sei $-\infty \le a < b \le \infty$ und sei $f : (a, b) \to \mathbb{R}$ eine stetige Funktion, für die die Grenzwerte $\lim_{x\to a} f(x)$ und $\lim_{x\to b} f(x)$ in $\hat{\mathbb{R}}$ existieren und gleich sind. Dann existiert das Maximum oder das Minimum von $f((a, b))$ in $\mathbb{R}$.*

Beweis. Sei $\ell := \lim_{x\to a} f(x) = \lim_{x\to b} f(x)$. Ist f konstant auf (a, b), so ist $f \equiv \ell \in \mathbb{R}$ und die Behauptung ist trivial. In allen anderen Fällen gibt es ein $x_0 \in (a, b)$ mit $f(x_0) \ne \ell$.

Im Fall $f(x_0) < \ell$ gibt es wegen $\lim_{x\to a} f(x) = \lim_{x\to b} f(x) = \ell$ zu einem fest gewählten $c \in (f(x_0), \ell)$ Punkte $\alpha, \beta \in (a, b)$ mit $\alpha < \beta$ und mit

$$\forall x \in (a, \alpha) \cup (\beta, b) : \quad f(x) \ge c\,. \tag{4.22}$$

Die Funktion f nimmt aber auf $[\alpha, \beta]$ nach Satz 4.19 ihr Minimum in einem Punkt $u \in [\alpha, \beta]$ an. Insbesondere ist $x_0 \in [\alpha, \beta]$ und damit $f(u) \le f(x_0) < c$. Wegen (4.22) folgt also $f(u) = \min f((a, b))$.

Den Fall $f(x_0) > \ell$ behandelt man analog oder man geht zur Funktion $-f$ über. Diese besitzt wie schon gezeigt auf (a, b) ein Minimum, welches in einem Punkt $v \in (a, b)$ angenommen wird. Dann ist $f(v) = -\min(-f((a, b))) = \max f((a, b))$. □

KAPITEL 5

Differentialrechnung in einer Veränderlichen

1. Differenzierbare Funktionen

Die Einführung des Begriffes der Differenzierbarkeit ist durch die beiden folgende Fragestellungen motiviert:

- **Die Tangentenaufgabe**: Gegeben sei eine Funktion $f : D \to \mathbb{R}^N$ mit Definitionsbereich $D \subseteq \mathbb{R}$. Gesucht ist die Tangente an den Graphen von f im Punkt $(a, f(a))$ für ein $a \in D$.
- **Bestimmung der Momentangeschwindigkeit eines bewegten Körpers**: Man bestimme die Momentangeschwindigkeit eines nicht notwendig gleichförmig bewegten Körpers an einem vorgegebenen Zeitpunkt t.

Beide Aufgaben führen zu der folgenden Problemstellung: Wie kann man dem für $x \neq a$ aber x „nahe“ bei a definierten Ausdruck

$$\frac{1}{x-a}(f(x) - f(a))$$

in $x = a$ einen Sinn geben? Für uns ist dies kein Problem mehr, da wir den Grenzwertbegriff zur Verfügung haben.

5.1. Definition. Sei D eine nicht leere Teilmenge von $\mathbb{R}$ und sei $a \in D$ ein Häufungspunkt von D. Eine Funktion $f : D \to \mathbb{R}^N$ heißt im Punkt a *differenzierbar*, falls der Grenzwert

$$f'(a) := \frac{df}{dx}(a) := \lim_{x\to a} \frac{1}{x-a}(f(x) - f(a)) \tag{5.1}$$

in $\mathbb{R}^N$ existiert. $f'(a)$ heißt die *Ableitung* der Funktion f im Punkt a.

Häufig schreiben wir den Grenzwert (5.1) auch in der folgenden Form:

$$f'(a) := \lim_{h\to 0} \frac{1}{h}(f(a+h) - f(a)) \tag{5.2}$$

Man überlegt sich leicht: Genau dann existiert der Grenzwert in (5.2), wenn der in (5.1) existiert. Beide Grenzwerte stimmen in diesem Fall überein.

5.2. Definition. Sei D eine nicht leere Teilmenge von $\mathbb{R}$, so daß jeder Punkt aus D ein Häufungspunkt von D ist. Dies trifft z.B. zu, wenn D eine Vereinigung von nicht ausgearteten Intervallen ist. Eine Funktion $f : D \to \mathbb{R}^N$ heißt *auf D differenzierbar*, falls f in allen Punkten $a \in D$ differenzierbar ist. In diesem Fall wird durch

$$a \mapsto f'(a) = \frac{df}{dx}(a) = \lim_{x\to a} \frac{1}{x-a}(f(x) - f(a))$$

eine Funktion $f' = \frac{df}{dx} : D \to \mathbb{R}^N$ definiert. f' heißt die *abgeleitete Funktion* oder *Ableitung* von f. Ist f' in einem Punkt $a \in D$ differenzierbar, so heißt

$$f''(a) := (f')'(a) = \lim_{x\to a} \frac{1}{x-a}(f'(x) - f'(a))$$

die *zweite Ableitung* von f im Punkt a. Wir definieren induktiv höhere Ableitungen von f wie folgt:

$$f^{(0)} := f\,, \quad f^{(1)} := \frac{d^1 f}{dx^1} := f'.$$

Ist $f^{(n)} = \frac{d^n f}{dx^n}$ schon für ein $n \in \mathbb{N}_0$ auf ganz D definiert und ist die Funktion $f^{(n)}$ auf D differenzierbar in allen Punkten von D, so heißt die durch

$$x \mapsto f^{(n+1)}(x) := \frac{d^{n+1} f}{dx^{n+1}(x)} := (f^{(n)})'(x) = \Big(\frac{d}{dx}\Big(\frac{d^n f}{dx^n}\Big)\Big)(x)$$

definierte Funktion $f^{(n+1)} : D \to \mathbb{R}^N$ die $(n+1)$*–te Ableitung* von f. Die Funktion f heißt dann auch $n+1$ *mal differenzierbar* auf D.

Aus Definition 5.1 und Satz 4.41 erhalten wir die folgende Charakterisierung der Differenzierbarkeit in einem Punkt:

5.3. Satz. *Sei D eine nicht leere Teilmenge von $\mathbb{R}$ und sei $a \in D$ ein Häufungspunkt von D. Für eine Funktion $f : D \to \mathbb{R}^N$ und einen Punkt $y \in \mathbb{R}^N$ sind die folgenden Aussagen äquivalent:*

(a) *f ist in a differenzierbar und $f'(a) = y$.*

(b) $\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in D \setminus \{a\} : \quad |x - a| < \delta \Longrightarrow \Big|\frac{1}{x-a}(f(x) - f(a)) - y\Big| < \varepsilon.$

(c) *Für jede Folge $(x_n)_{n=1}^\infty$ aus $D \setminus \{a\}$ mit $x_n \to a$ für $n \to \infty$ gilt*

$$y = \lim_{n\to\infty} \frac{1}{x_n - a}\big(f(x_n) - f(a)\big)\,.$$

5.4. Beispiele. (a) Sei D eine nicht leere Teilmenge von $\mathbb{R}$ und sei $a \in D$ ein Häufungspunkt von D. Ist $f : D \to \mathbb{R}^N$ eine konstante Funktion, d.h. f ist gegeben durch $x \mapsto f(x) = c$ für ein festes $c \in \mathbb{R}^N$, so ist $f'(a) = 0$.

(b) Sei $n \in \mathbb{N}$. Dann ist die durch $f_n(x) := x^n$ für alle $x \in \mathbb{R}$ auf $\mathbb{R}$ definierte Funktion auf ganz $\mathbb{R}$ differenzierbar und es gilt $f_n'(a) = na^{n-1}$ für alle $a \in \mathbb{R}$.

(c) Für alle $\lambda \in \mathbb{C}$ ist die Funktion $f_\lambda : \mathbb{R} \to \mathbb{C}$ mit $f_\lambda(x) := \exp(\lambda x)$ für alle $x \in \mathbb{R}$ auf ganz $\mathbb{R}$ differenzierbar und es gilt

$$\forall x \in \mathbb{R} : \qquad f_\lambda'(x) = \lambda \exp(\lambda x)\,.$$

Insbesondere ist die Exponentialfunktion beliebig oft auf $\mathbb{R}$ differenzierbar und es gilt $\exp^{(n)}(x) = \exp(x)$ für alle $n \in \mathbb{N}, x \in \mathbb{R}$.

(d) Die Funktion $f : \mathbb{R}\setminus\{0\} \to \mathbb{R}$ mit $f(x) := 1/x$ ist auf ganz $\mathbb{R}\setminus\{0\}$ differenzierbar und es gilt

$$f'(x) = \frac{-1}{x^2} \quad \text{für alle } x \in \mathbb{R}\setminus\{0\}\,.$$

(e) Die Betragsfunktion $|\cdot| : \mathbb{R} \to \mathbb{R}$ ist im Punkt 0 nicht differenzierbar.

Beweis. (a) Dies ist unmittelbar klar wegen

$$\frac{1}{x-a}(f(x) - f(a)) = \frac{1}{x-a}(c - c) = 0 \to 0 \quad \text{für } x \to a.$$

(b)Für reelles $x \neq a$ gilt

$$\frac{f(x) - f(a)}{x - a} = \frac{x^n - a^n}{x - a} = \sum_{j=0}^{n-1} x^j a^{n-1-j}.$$

Mit den Grenzwertrechenregeln folgt also

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=\lim_{x\to a}\sum_{j=0}^{n-1}x^j a^{n-1-j}=\sum_{j=0}^{n-1}a^j a^{n-1-j}=n\cdot\cdot a^{n-1}.$$

(c) Für $\lambda=0$ ist die Behauptung trivial. Für $\lambda\neq 0$ und alle reellen $h\neq 0$ gilt:

$$\frac{\exp(\lambda(a+h))-\exp(\lambda a)}{h}=\exp(\lambda a)\frac{\exp(\lambda h)-1}{h}\,.$$

Zu beliebig vorgegebenem $\varepsilon>0$ gibt es nach Beispiel 4.43 (c) ein $\delta>0$, so daß für alle $u\in\mathbb{R}\setminus\{0\}\subset\mathbb{C}\setminus\{0\}$ mit $|u|<\delta$ gilt

$$\Big|\frac{\exp(u)-1}{u}-1\Big|<\frac{\varepsilon}{|\lambda\exp(\lambda a)|}\,.$$

Also erhalten wir für alle $0\neq h\in\mathbb{R}$ mit $|h|<\frac{\delta}{|\lambda|}$:

$$\Big|\frac{\exp(\lambda(a+h))-\exp(\lambda a)}{h}-\lambda\exp(\lambda a)\Big|=|\lambda\exp(\lambda a)|\cdot\Big|\frac{\exp(\lambda h)-1}{\lambda h}-1\Big|<\varepsilon\,.$$

Hiermit folgt die Behauptung.

(d) Für $0\neq x\in\mathbb{R}$ gilt nach den Grenzwertrechenregeln

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0}\frac{1}{h}\Big(\frac{1}{x+h}-\frac{1}{x}\Big)=\lim_{h\to 0}\frac{-1}{(x+h)x}=\frac{-1}{x^2}\,.$$

(e) Wegen

$$\frac{|x|-0}{x-0}=\begin{cases}1 & \text{für } x>0\\ -1 & \text{für } x<0\end{cases}$$

haben wir

$$\lim_{x\to 0^+}\frac{|x|-0}{x-0}=1\quad\text{und}\quad\lim_{x\to 0^-}\frac{|x|-0}{x-0}=-1\,.$$

Nach Satz 4.49 hat also der Differenzenquotient $\frac{|x|-0}{x-0}$ in 0 keinen Grenzwert. □

Die folgende Aussage erhält man in naheliegender Weise mit Hilfe von Lemma 4.42.

5.5. Lemma. *Sei $f:D\to\mathbb{R}^N$ eine Funktion mit Definitionsbereich $D\subseteq\mathbb{R}$ und sei $a\in D$ ein Häufungspunkt von D.*

(a) *Sei f in a differenzierbar. Ist $Y\subseteq D$ mit $a\in Y$ und ist a auch ein Häufungspunkt von Y, so ist die Funktion $f|_Y$ in a differenzierbar und es gilt $(f|_Y)'(a)=f'(a)$.*

(b) *Sei $\varepsilon>0$ beliebig. Genau dann ist f in a differenzierbar, wenn die Restriktion $f|_{U_\varepsilon(a)\cap D}$ in a differenzierbar ist. Es gilt dann $f'(a)=(f|_{U_\varepsilon(a)\cap D})'(a)$.*

5.6. Satz. *Sei $f:D\to\mathbb{R}^N$ eine Funktion mit Definitionsbereich $D\subseteq\mathbb{R}$ und sei $a\in D$ ein Häufungspunkt von D. Ist f in a differenzierbar, so ist f in a auch stetig.*

Beweis. Nach den Grenzwertrechenregeln gilt:

$$\begin{aligned}\lim_{x\to a}(f(x)-f(a))&=\lim_{x\to a}(x-a)\frac{1}{x-a}(f(x)-f(a))=\\&=\big(\lim_{x\to a}(x-a)\big)\cdot\lim_{x\to a}\frac{1}{x-a}(f(x)-f(a))=0\cdot f'(a)=0.\end{aligned}$$

Mit Folgerung 4.44 folgt die Behauptung. □

Warnung: Für die Differenzierbarkeit ist die Stetigkeit einer Funktion zwar *notwendig* aber *nicht hinreichend*. So ist die Betragsfunktion $|\cdot| : \mathbb{R} \to \mathbb{R}$ zwar auf $\mathbb{R}$ stetig aber nicht in 0 differenzierbar. Es gibt sogar Funktionen $f : \mathbb{R} \to \mathbb{R}$, die überall stetig aber nirgends differenzierbar sind.

Wir fassen nun die wichtigsten Rechenregeln für die Differenzierbarkeit zusammen:

5.7. Satz (Rechenregeln für die Differentiation). *Sei $D \subseteq \mathbb{R}$ und sei $a \in D$ ein Häufungspunkt von D.*

(a) *Sind $f, g : D \to \mathbb{R}^N$ in a differenzierbar, so ist für alle $\alpha, \beta \in \mathbb{R}$ auch die Funktion $\alpha f + \beta g$ in a differenzierbar und es gilt:*

$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a).$$

(b) *Sind $f, g : D \to \mathbb{C}$ in a differenzierbar, so ist für alle $\alpha, \beta \in \mathbb{C}$ auch die Funktion $\alpha f + \beta g$ in a differenzierbar und es gilt:*

$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a).$$

(c) (Produktregel). (i) *Sind $f, g : D \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) in a differenzierbar, so ist auch die Funktion fg in a differenzierbar und es gilt:*

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

(ii) Sind $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}^N$ in a differenzierbar, so ist auch die Funktion fg in a differenzierbar und es gilt:

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

(d) (Kettenregel). *Ist $f : D \to \mathbb{R}$ in a differenzierbar und ist $f(D) \subseteq Y \subseteq \mathbb{R}$ so, daß $f(a)$ ein Häufungspunkt von Y ist, so ist für jede in $f(a)$ differenzierbare Funktion $g : Y \to \mathbb{R}^N$ die Funktion $g \circ f : D \to \mathbb{R}^N$ in a differenzierbar und es gilt:*

$$(g \circ f)'(a) = f'(a)g'(f(a))\,.$$

(e) (Quotientenregel). *Sind $f, g : D \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) in a differenzierbar und ist $g(a) \neq 0$, so ist die auf $D_g := \{x \in D\,;\, g(x) \neq 0\}$ definierte Funktion $f/g : D_g \to \mathbb{K}$ in a differenzierbar und es gilt*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}\,. \tag{5.3}$$

Beweis. Die Aussagen (a) und (b) folgen in nahe liegender Weise aus den Grenzwertrechenregeln.

(c) Da g in a differenzierbar ist, ist g nach Satz 5.6 auch in a stetig und es gilt $g(a) = \lim_{x\to a} g(x)$ (nach Folgerung 4.44). Mit den Grenzwertrechenregeln erhalten wir daher:

$$\begin{aligned}\lim_{x\to a} \frac{1}{x-a}\big((fg)(x) - (fg)(a)\big) &= \lim_{x\to a}\left(\frac{f(x)-f(a)}{x-a}g(x) + f(a)\frac{1}{x-a}(g(x)-g(a))\right)\\ &=\left(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right)\lim_{x\to a} g(x) + f(a)\lim_{x\to a}\frac{1}{x-a}(g(x)-g(a))\\ &= f'(a)g(a) + f(a)g'(a)\,.\end{aligned}$$

(d) Die Funktion $h : Y \to \mathbb{R}^N$ mit

$$h(y) := \begin{cases} \frac{1}{y-f(a)}\big(g(y) - g(f(a))\big) & \text{für } y \in Y \setminus \{f(a)\} \\ g'(f(a)) & \text{für } y = f(a) \end{cases}$$

ist wegen der Differenzierbarkeit von g im Punkt $f(a)$ nach Definition des Grenzwerts bei Funktionen in $f(a)$ stetig und es gilt für alle $x \in D$

$$g(f(x)) - g(f(a)) = (f(x) - f(a))h(f(x)) .$$

Unter Verwendung von Folgerung 4.44 und den Grenzwertrechenregeln erhalten wir

$$\begin{aligned}(g \circ f)'(a) &= \lim_{x\to a} \frac{1}{x-a}\big(g(f(x)) - g(f(a))\big) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} h(f(x)) \\ &= \Big(\lim_{x\to a} \frac{f(x)-f(a)}{x-a}\Big) \lim_{x\to a} h(f(x)) = f'(a)g'(f(a)) .\end{aligned}$$

(e) Wegen $g(a) \neq 0$ und der Stetigkeit von g in a (nach Satz 5.6) gibt es ein $\eta > 0$, so daß $U_\eta(a) \cap D \subseteq D_g$ und so daß die Funktion

$$\frac{1}{g} : D_g \to \mathbb{K}, \quad x \mapsto \frac{1}{g(x)},$$

in a stetig ist. Mit den Grenzwertrechenregeln gilt für $x \in U_\eta(a) \cap D$

$$\frac{1}{x-a}\Big(\frac{1}{g(x)} - \frac{1}{g(a)}\Big) = \frac{g(a)-g(x)}{x-a} \cdot \frac{1}{g(x)g(a)} \to -\frac{g'(a)}{g(a)^2} \quad \text{für } x \to a .$$

Wegen Lemma 5.5 ist $1/g$ also in a differenzierbar und es gilt

$$\Big(\frac{1}{g}\Big)'(a) = -\frac{g'(a)}{g(a)^2} .$$

Die Formel (5.3) folgt nun mit Hilfe der Produktregel. □

5.8. Beispiele. (a) Alle Polynomfunktionen p mit Koeffizienten in $\mathbb{R}$ oder $\mathbb{C}$ sind auf $\mathbb{R}$ differenzierbar und ihre Ableitungsfunktionen sind wieder Polynomfunktionen.

(b) Jede rationale Funktion $f = p/q$ mit Polynomen p, q mit Koeffizienten in $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ ist auf $D_q := \{x \in \mathbb{R}\,;\, q(x) \neq 0\}$ differenzierbar und die Ableitungsfunktion $f' : D_q \to \mathbb{K}$ ist wieder eine rationale Funktion, deren Nenner Nullstellen höchstens in $\mathbb{R} \setminus D_q$ besitzt.

(c) Die Funktionen $\sinh, \cosh, \tanh$ sind auf ganz $\mathbb{R}$ differenzierbar. Für die Ableitungsfunktionen gilt:

$$\forall x \in \mathbb{R} : \qquad \sinh'(x) = \cosh(x), \quad \cosh'(x) = \sinh(x), \quad \tanh'(x) = \frac{1}{\cosh(x)^2} .$$

Die Funktion $\coth : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ ist auf $\mathbb{R} \setminus \{0\}$ differenzierbar und es gilt

$$\forall 0 \neq x \in \mathbb{R} : \quad \coth'(x) = \frac{-1}{\sinh(x)^2} .$$

(d) Die trigonometrischen Funktionen $\sin, \cos : \mathbb{R} \to \mathbb{R}$ sind auf ganz $\mathbb{R}$ differenzierbar und es gilt

$$\forall x \in \mathbb{R} : \qquad \sin'(x) = \cos(x) \qquad \cos'(x) = -\sin(x).$$

Die Tangensfunktion und die Cotangensfunktion sind in ihren natürlichen Definitionsbereichen $D_t := \mathbb{R} \setminus \{n\pi + \pi/2\,;\, n \in \mathbb{Z}\}$ bzw. $D_c := \mathbb{R} \setminus \{n\pi\,;\, n \in \mathbb{Z}\}$ differenzierbar. Für die Ableitungsfunktionen gilt:

$$\forall x \in D_t : \qquad \tan'(x) = \frac{1}{\cos(x)^2}$$

und

$$\forall x \in D_c : \qquad \cot'(x) = \frac{-1}{\sin(x)^2} .$$

Beweis. (a) folgt mit Satz 5.7 (a), (b) aus Beispiel 5.4 (b). Teil (b) erhält man mit Hilfe der Quotientenregel 5.7 (e) aus (a).

(c) Wegen $\sinh(x) = \frac{1}{2}(\exp(x) - \exp(-x))$ und $\cosh(x) = \frac{1}{2}(\exp(x) + \exp(-x))$ erhält man unter Verwendung von Beispiel 5.4 (c) und den Rechenregeln 5.7 (a):

$$\sinh'(x) = \frac{1}{2}(\exp(x) - (-1)\exp(-x)) = \cosh(x)$$

und

$$\cosh'(x) = \frac{1}{2}(\exp(x) + (-1)\exp(-x)) = \sinh(x)$$

für alle $x \in \mathbb{R}$. Wegen $\tanh(x) = \sinh(x)/\cosh(x)$ folgt mit der Quotientenregel für alle $x \in \mathbb{R}$:

$$\tanh'(x) = \frac{\cosh(x)\sinh'(x) - \sinh(x)\cosh'(x)}{\cosh(x)^2} = \frac{\cosh(x)^2 - \sinh(x)^2}{\cosh(x)^2} = \frac{1}{\cosh(x)^2}$$

Die Aussage für $\coth(x) = \cosh(x)/\sinh(x)$ erhält man analog.

(d) Wir beachten

$$\forall x \in \mathbb{R}: \qquad \sin(x) = \frac{1}{2i}(\exp(ix) - \exp(-ix)), \quad \cos(x) = \frac{1}{2}(\exp(ix) + \exp(-ix)).$$

Unter Verwendung von Beispiel 5.4 (c) und den Rechenregeln 5.7 (a) erhalten wir hieraus

$$\sin'(x) = \frac{1}{2i}(i\exp(ix) - (-i)\exp(-ix)) = \frac{1}{2}(\exp(ix) + \exp(-ix)) = \cos(x)$$

und

$$\cos'(x) = \frac{1}{2}(i\exp(ix) + (-i)\exp(-ix)) = \frac{i}{2}(\exp(ix) - \exp(-ix)) = -\sin(x)\,.$$

Die Aussagen für $\tan(x) = \sin(x)/\cos(x)$ und $\cot(x) = \cos(x)/\sin(x)$ erhält man hieraus mit Hilfe der Quotientenregel. □

Ist $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall und ist $f : I \to \mathbb{R}^N$ auf I n mal differenzierbar (für ein $n \in \mathbb{N}$), so ist f und für $j = 1, \ldots, n-1$ die j–te Ableitung $f^{(j)}$ nach Satz 5.6 stetig auf I. Ist auch die n–te Ableitung $f^{(n)}$ auf I stetig, so sagen wir *f ist auf I n mal stetig differenzierbar*. Die Menge der auf I n mal stetig differenzierbaren Funktionen mit Werten in $\mathbb{R}^N$ bezeichnen wir mit $C^n(I, \mathbb{R}^N)$. Entsprechend sei $C^n(I, \mathbb{C})$ der Raum der auf I n mal stetig differenzierbaren Funktionen mit Werten in $\mathbb{C}$ (wobei wir beachten, daß $\mathbb{C}$ der $\mathbb{R}^2$ mit einer Körperstruktur ist). Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, so schreiben wir statt $C^n(I, \mathbb{K})$ auch kurz $C^n(I)$, wenn der zugrundeliegende Körper aus dem Zusammenhang klar ist. Ist $f : I \to \mathbb{R}^N$ auf I beliebig oft differenzierbar, so ist $f \in \bigcap_{n=0}^{\infty} C^n(I, \mathbb{R}^N)$. Wir schreiben auch $C^\infty(I, \mathbb{R}^N) := \bigcap_{n=0}^{\infty} C^n(I, \mathbb{R}^N)$.

Mit den Rechenregeln für die Differentiation und für stetige Funktionen folgt: Für alle $n \in \mathbb{N}_0 \cup \{\infty\}$ ist $C^n(I, \mathbb{R}^N)$ ein Untervektorraum des Vektorraums aller auf I stetigen Funktionen.

Aus den Beispielen 5.4 und 5.8 und den Rechenregeln folgt unter Verwendung der Rechenregeln insbesondere $\exp, \sin, \cos, \sinh, \cosh, \tanh \in C^\infty(\mathbb{R})$.

Wir wollen nun die Frage der Differenzierbarkeit der Umkehrfunktion einer injektiven differenzierbaren Funktion untersuchen.

5.9. Satz. *Sei $f : I \to \mathbb{R}$ eine auf einem nicht ausgearteten Intervall I definierte injektive, stetige Funktion. Ist f in einem Punkt $x_0 \in I$ differenzierbar und ist $f'(x_0) \neq 0$, so ist auch die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ in $f(x_0)$ differenzierbar und es gilt*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}\,. \tag{5.4}$$

Beweis. Nach Satz 4.32 ist f streng monoton und nach Satz 4.31 ist $f(I)$ wieder ein nicht ausgeartetes Intervall und die Funktion $f^{-1} : f(I) \to I \subseteq \mathbb{R}$ ist stetig. Insbesondere ist $f(x_0)$ ein Häufungspunkt von $f(I)$. Sei nun $(y_n)_{n=1}^{\infty}$ eine beliebige gegen $f(x_0)$ konvergente Folge aus $f(I) \setminus \{f(y_0)\}$. Wegen der Stetigkeit der Umkehrfunktion f^{-1} folgt nach dem Folgenkriterium 4.5 $x_n := f^{-1}(y_n) \to f^{-1}(f(x_0)) = x_0$ für $n \to \infty$ und wegen der Injektivität von f^{-1} auch $x_n \neq x_0$ für alle $n \in \mathbb{N}$. Wegen $f'(x_0) \neq 0$ erhalten wir nach Satz 5.3 mit Hilfe der Grenzwertrechenregeln:

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(f(x_0))}{y_n - f(x_0)} = \lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \frac{1}{f'(x_0)} .$$

Mit Satz 5.3 folgt also die Differenzierbarkeit von f^{-1} in $f(x_0)$ und (5.4). □

5.10. Beispiel. Die Funktion

$$d : \mathbb{R} \to \mathbb{R} , \quad x \mapsto x^3$$

ist auf $\mathbb{R}$ streng monoton wachsend, differenzierbar, daher auch stetig und bildet $\mathbb{R}$ bijektiv auf $\mathbb{R}$ ab. Für die Ableitungsfunktion gilt $f'(x) = 3x^2$ für alle $x \in \mathbb{R}$. Insbesondere ist $f'(0) = 0$. Die Umkehrfunktion $f^{-1} : \mathbb{R} \to \mathbb{R}$ ist nach Satz 4.28 also stetig. Sie ist aber in 0 nicht differenzierbar, denn $n^{-3} \to 0$ für $n \to \infty$, aber die Folge der Differenzenquotienten

$$\frac{f^{-1}(n^{-3}) - f^{-1}(0)}{n^{-3} - 0} = \frac{n^{-1}}{n^{-3}} = n^2$$

ist unbeschränkt und daher divergent. Dieses Beispiel zeigt, daß man in Satz 5.9 auf die Voraussetzung $f'(x_0) \neq 0$ nicht verzichten kann.

5.11. Folgerung. (a) *Der natürliche Logarithmus* $\log : (0, \infty) \to \mathbb{R}$ *ist auf ganz* $(0, \infty)$ *differenzierbar und es gilt* $\log'(x) = x^{-1}$ *für alle reellen* $x > 0$.

(b) *Die Funktion*

$$f : (0, \infty) \to \mathbb{R}, \quad x \mapsto f(x) := x^x$$

ist auf $(0, \infty)$ *differenzierbar und es gilt*

$$f'(x) = (1 + \log(x))x^x \quad \textit{für alle reellen } x > 0.$$

(c) *Sei* $\alpha \in \mathbb{R}$. *Die Funktion*

$$g : (0, \infty) \to \mathbb{R}, \quad x \mapsto g(x) := x^{\alpha}$$

ist auf $(0, \infty)$ *differenzierbar und für alle reellen* $x > 0$ *gilt* $g'(x) = \alpha x^{\alpha-1}$.

Beweis. (a) Die Abbildung $\exp : \mathbb{R} \to \mathbb{R}$ ist auf $\mathbb{R}$ differenzierbar mit $\exp'(y) = \exp(y) \neq 0$ für alle $y \in \mathbb{R}$ und injektiv mit $\exp(\mathbb{R}) = (0, \infty)$ (vergl. Satz 4.34 und Beispiel 5.4). Sei nun $x > 0$ beliebig und $y := \log(x)$. Nach Satz 5.9 ist $\log = \exp^{-1}$ in $x = \exp(y)$ differenzierbar und es gilt

$$\log'(x) = \frac{1}{\exp'(y)} = \frac{1}{x} .$$

(b) Nach Definition von x^x gilt $x^x = \exp(x \cdot \log(x))$ für alle $x \in \mathbb{R}$. Aus (a) erhalten wir also unter Verwendung der Kettenregel und der Produktregel die Differenzierbarkeit von f und für alle $x \in (0, \infty)$:

$$f'(x) = \Big(x \cdot \frac{1}{x} + 1 \cdot \log(x)\Big) \exp(x \cdot \log(x)) = (1 + \log(x))x^x .$$

(c) Wegen $x^{\alpha} = \exp(\alpha \log(x))$ für alle $x \in (0, \infty)$ folgt auch dies mit Hilfe der Kettenregel aus (a). □

Zusammen mit Beispiel 5.8 (b) folgt $\log \in C^\infty((0,\infty))$

5.12. Definition. Sei $f : I \to \mathbb{R}^N$ eine auf einem nicht ausgearteten Intervall definierte Funktion und sei $x_0 \in I$ ein innerer Punkt von I (d.h. es gibt ein $\eta > 0$ mit $U_\eta(x_0) \subseteq I$). Die Funktion f heißt in x_0 *linksseitig* bzw. *rechtsseitig differenzierbar* falls der einseitige Grenzwert

$$f'(x_0^-) := \lim_{x \to x_0^-} \frac{1}{x - x_0}(f(x) - f(x_0)) \quad \text{bzw. } f'(x_0^+) \lim_{x \to x_0^+} \frac{1}{x - x_0}(f(x) - f(x_0))$$

existiert. $f'(x_0^-)$ bzw. $f'(x_0^+)$ heißt dann die *linksseitige* bzw. *rechtsseitige Ableitung* von f in x_0.

Aus dieser Definition und Satz 4.49 folgt unmittelbar:

5.13. Folgerung. *Sei $f : I \to \mathbb{R}^N$ eine auf einem Intervall I definierte Funktion und sei x_0 ein innerer Punkt von I. Genau dann ist f in x_0 differenzierbar, wenn f in x_0 links– und rechtsseitig differenzierbar ist und $f'(x_0^-) = f'(x_0^+)$ gilt. In diesem Fall ist $f'(x_0) = f'(x_0^-) = f'(x_0^+)$.*

5.14. Beispiel. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := |x|$ für alle $x \in \mathbb{R}$ ist in 0 linksseitig und rechtsseitig differenzierbar mit $f'(x_0^-) = -1 \neq 1 = f'(x_0^+)$.

2. Der Mittelwertsatz der Differentialrechnung mit ersten Anwendungen

5.15. Definition. Sei $\emptyset \neq D \subseteq \mathbb{R}$ und sei $f : D \to \mathbb{R}$ eine reellwertige Funktion auf D. Ein Punkt $x_0 \in D$ heißt *lokale Maximumstelle* (bzw. *lokale Minimumstelle*) von f, falls es ein $\varepsilon > 0$ gibt mit

$$\forall x \in D \cap U_\varepsilon(x_0) : f(x) \leq f(x_0) \quad (\text{bzw. } f(x) \geq f(x_0)). \tag{5.5}$$

$f(x_0)$ heißt dann ein *lokales Maximum* (bzw. *lokales Minimum*) von f. Ist x_0 lokale Maximums– oder Minimumstelle von f, so nennen wir x_0 auch eine *lokale Extremstelle* und $f(x_0)$ ein *lokales Extremum* von f. Gilt in (5.5) das $<$–Zeichen bzw. das $>$–Zeichen, so nennen wir x_0 eine *isolierte* lokale Maximums– bzw. Minimumstelle von f.

Das Auffinden lokaler Extremstellen wird erleichtert durch:

5.16. Satz. *Sei $I \subseteq \mathbb{R}$ ein Intervall und $x_0 \in I$ ein innerer Punkt von I. Hat f in x_0 eine lokale Extremstelle und ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.*

Beweis. Sei zunächst x_0 eine lokale Maximumstelle. Daher und, weil x_0 ein innerer Punkt von I ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq I$ und $f(x) \leq f(x_0)$ für alle $x \in U_\varepsilon(x_0)$. Insbesondere erhalten wir

$$\forall x \in (x_0 - \varepsilon, x_0) : \quad \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

und daher (mit Folgerung 5.13)

$$f'(x_0) = f'(x_0^-) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

sowie

$$\forall x \in (x_0, x_0 + \varepsilon) : \quad \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

und somit (ebenfalls mit Folgerung 5.13)

$$f'(x_0) = f'(x_0^+) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

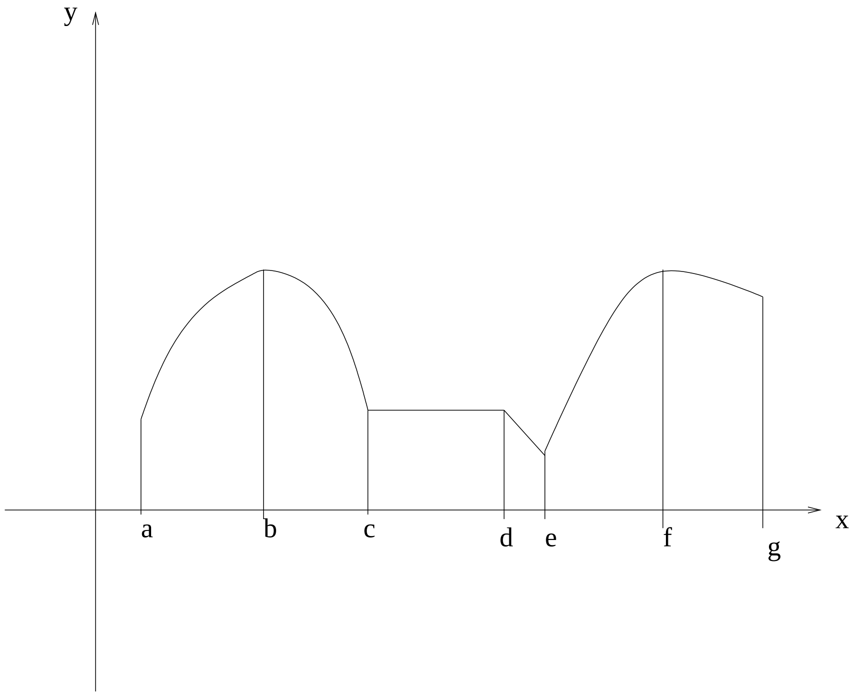


Abbildung 1. Lokale Maximumstellen liegen vor in b allen Punkten aus $(c,d]$ und in f. Davon sind b und f isolierte lokale Maximumstellen. Die Punkte a, alle Punkte aus $[c,d)$, e und g sind lokale Minimumstellen, davon a,e,g isoliert.

Also muß $f'(x_0)=0$ gelten. □

5.17. Beispiel. Die Funktion $f:\mathbb{R}\to\mathbb{R}$ mit $f(x):=x^3$ für alle $x\in\mathbb{R}$ ist auf ganz $\mathbb{R}$ differenzierbar und die Ableitungsfunktion $x\mapsto f'(x)=3x^2$ hat in 0 eine Nullstelle, aber 0 ist keine lokale Extremstelle, da $f(x)<0=f(0)$ für alle $x<0$ und $f(x)>0=f(0)$ für alle $x>0$ gilt. Das Vorliegen der Nullstelle der Ableitung einer differenzierbaren Funktion ist also notwendig, aber nicht hinreichend für das Vorliegen einer lokalen Extremstelle.

Als Folgerung aus Satz 5.16 und Satz 4.57 erhalten wir

5.18. Folgerung. *Sei $-\infty\le a<b\le\infty$ und sei $f:(a,b)\to\mathbb{R}$ eine differenzierbare Funktion, für die die Grenzwerte $\lim_{x\to a}f(x)$ und $\lim_{x\to b}f(x)$ in $\hat{\mathbb{R}}$ existieren und gleich sind. Dann gibt es ein $x_0\in(a,b)$ mit $f'(x_0)=0$.*

Beweis. Da f als differenzierbare Funktion insbesondere stetig ist, nimmt f nach Satz 4.57 in einem Punkt $x_0\in(a,b)$ ihr Maximum oder ihr Minimum an. Insbesondere ist x_0 eine lokale Extremstelle und es folgt $f'(x_0)=0$ nach Satz 5.16. □

Eine unmittelbare Folgerung hieraus ist:

5.19. Folgerung (Satz von Rolle[1]). *Sei $[a,b]\subseteq\mathbb{R}$ ein abgeschlossenes Intervall mit $-\infty<a<b<+\infty$ und sei $f:[a,b]\to\mathbb{R}$ eine auf $[a,b]$ stetige und auf (a,b) differenzierbare Funktion mit $f(a)=f(b)$. Dann gibt es ein $x_0\in(a,b)$ mit $f'(x_0)=0$.*

Als erste Folgerung aus dem Satz von Rolle erhalten wir den häufig sehr nützlichen Mittelwertsatz der Differentialrechnung:

[1] Michel Rolle (21.4.1652–8.11.1719)

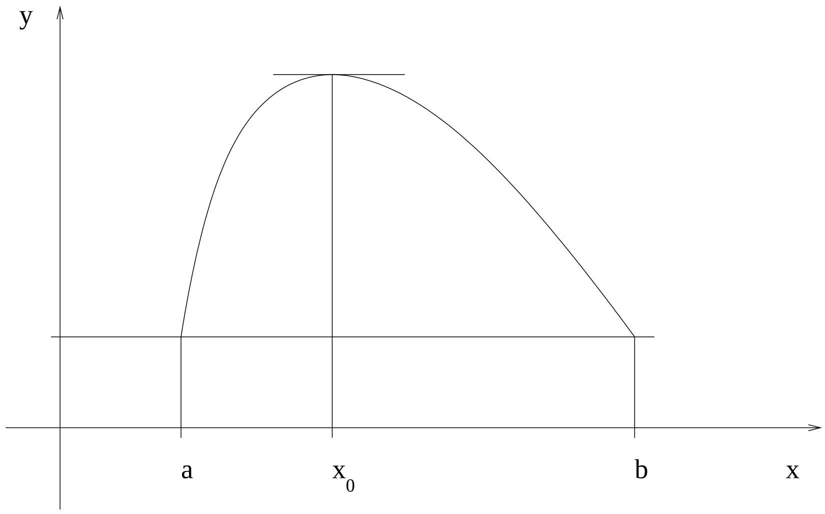


ABBILDUNG 2. Zum Satz von Rolle

5.20. SATZ (Mittelwertsatz der Differentialrechnung). *Sei $[a,b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall mit $-\infty < a < b < +\infty$ und sei $f : [a,b] \to \mathbb{R}$ eine auf $[a,b]$ stetige und auf (a,b) differenzierbare Funktion. Dann gibt es ein $x_0 \in (a,b)$ mit*

$$f'(x_0) = \frac{f(b)-f(a)}{b-a}\,.$$

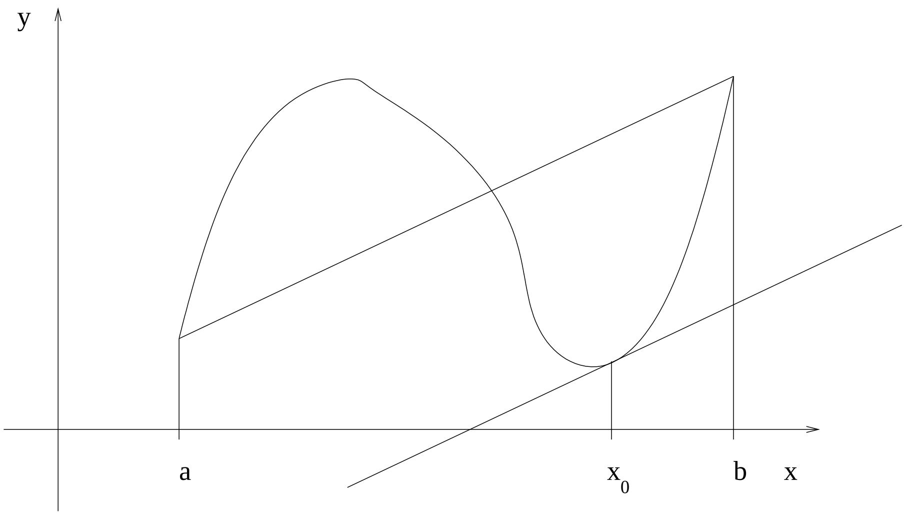


ABBILDUNG 3. Zum Mittelwertsatz der Differentialrechnung

BEWEIS. Die Funktion

$$h : [a,b] \to \mathbb{R}\,, \quad x \mapsto h(x) := f(x) - (x-a)\frac{f(b)-f(a)}{b-a}\,,$$

ist auf $[a,b]$ stetig, (a,b) differenzierbar und erfüllt $h(a) = f(a) = h(b)$. Nach dem Satz 5.19 von Rolle gibt es also ein $x_0 \in (a,b)$ mit

$$0 = h'(x_0) = f'(x_0) - \frac{f(b)-f(a)}{b-a}\,.$$

Hieraus folgt die Behauptung. □

Eine allgemeinere Fassung dieses Satzes läßt sich ebenfalls aus dem Satz von Rolle herleiten:

5.21. Satz (Verallgemeinerter Mittelwertsatz der Differentialrechnung). *Sei $[a,b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall mit $-\infty < a < b < +\infty$ und seien $f, g : [a,b] \to \mathbb{R}$ zwei auf $[a,b]$ stetige und auf (a,b) differenzierbare Funktionen. Dann gibt es ein $x_0 \in (a,b)$ mit*

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0)\,.$$

Beweis. Die Funktion $h : [a,b] \to \mathbb{R}$ mit

$$h(x) := (f(b) - f(a))g(x) - (g(b) - g(a))f(x) \quad \text{für alle } x \in [a,b]$$

ist auf $[a,b]$ stetig, auf (a,b) differenzierbar und erfüllt

$$h(a) = f(b)g(a) - f(a)g(b) = h(b)\,.$$

Nach dem Satz von Rolle gibt es also ein $x_0 \in (a,b)$ mit

$$0 = h'(x_0) = (f(b) - f(a))g'(x_0) - (g(b) - g(a))f'(x_0)\,.$$

Hieraus folgt die Behauptung. □

Wählt man in Satz 5.21 speziell $g(x) = x$ für alle $x \in [a,b]$, so sieht man daß der Mittelwertsatz der Differentialrechnung 5.20 tatsächlich ein Spezialfall des verallgemeinerten Mittelwertsatzes 5.21 ist.

In den nachstehenden drei Folgerungen sei stets $a, b \in \mathbb{R}$ mit $a < b$.

5.22. Folgerung. *Sei $f : [a,b] \to \mathbb{R}$ auf $[a,b]$ stetig und auf (a,b) differenzierbar mit $f' \equiv 0$ auf (a,b). Dann ist f auf $[a,b]$ konstant.*

Beweis. Für alle $x, y \in [a,b]$ mit $x < y$ gilt nach dem Mittelwertsatz 5.20 (angewendet auf die Einschränkung von f auf das Intervall $[x,y]$): Es gibt ein $x_0 \in (x,y)$ mit

$$f(x) - f(y) = (x - y)f'(x_0) = 0.$$

Daher ist f auf $[a,b]$ konstant. □

Wir werden häufig die folgende Variante dieser Folgerung anwenden:

5.23. Folgerung. (a) *Sei $f = (f_1, \ldots, f_N) : [a,b] \to \mathbb{R}^N$ auf $[a,b]$ stetig und auf (a,b) differenzierbar mit $f' \equiv 0$ auf (a,b). Dann ist f auf $[a,b]$ konstant.*

(b) *Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall und sei $f = (f_1, \ldots, f_N) : I \to \mathbb{R}^N$ auf I differenzierbar mit $f' \equiv 0$ auf I. Dann ist f auf I konstant.*

Beweis. (a) folgt, indem wir Folgerung 5.22 anwenden auf die Komponentenfunktionen f_j, $j = 1, \ldots, N$.

(b) Nach Satz 5.6 ist f auf I stetig. Für alle $x, y \in I$ erhalten wir durch Anwendung von (a) auf die Einschränkung von f auf $[x,y]$ die Aussage $f(x) = f(y)$. Also ist f auf I konstant. □

Der verallgemeinerte Mittelwertsatz der Differentialrechnung findet auch Anwendung im Beweis der folgenden Regeln von de l'Hospital[2].

[2] Guillaume François Antoine Marquis de L'Hopital (1661–2.2.1704).

5.24. Satz (Regeln von de l'Hospital). *Sei* $-\infty \le a < b \le \infty$ *und seien* $f, g : (a,b) \to \mathbb{R}$ *zwei auf* (a,b) *differenzierbare Funktionen mit* $g'(x) \neq 0$ *für alle* $x \in (a,b)$*. Es gelte*

(a) $\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x)$

oder

(b) $\lim_{x\to a} f(x) = \pm\infty$ *und* $\lim_{x\to a} g(x) = \pm\infty$.

Existiert der Grenzwert

$$\ell := \lim_{x\to a} \frac{f'(x)}{g'(x)} \qquad \text{in } \hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$$

so existiert auch der Grenzwert

$$\lim_{x\to a} \frac{f(x)}{g(x)} \ \text{ in } \hat{\mathbb{R}} \text{ und es gilt } \lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} .$$

Beweis. Wir behandeln zunächst den Fall (a). Hätte g eine Nullstelle u in (a,b), so würde nach Folgerung 5.18 ein $v \in (a,u)$ existieren mit $g'(v) = 0$ im Widerspruch zur Voraussetzung. Also ist $g(x) \neq 0$ für alle $x \in (a,b)$. Ebenso sieht man, daß es keine Punkte $x, y \in (a,b)$ geben kann mit $x \neq y$ und $g(x) = g(y)$.

Ist $\ell \in \mathbb{R} \cup \{-\infty\}$ und $C > \ell$ beliebig, so gibt es ein $\alpha_1 = \alpha_1(C) > a$ mit

$$\forall x \in (a, \alpha_1): \qquad \frac{f'(x)}{g'(x)} < C .$$

Nach dem verallgemeinerten Mittelwertsatz 5.21 der Differentialrechnung gibt es also für $a < x < y < \alpha_1$ ein $z = z(x,y) \in (a, \alpha_1)$ mit

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)} < C .$$

Für $x \to a$ erhalten wir hieraus $f(y)/g(y) \le C$ für alle $y \in (a, \alpha_1)$. Insbesondere gilt im Fall $\ell = -\infty$:

$$\lim_{y\to a} \frac{f(y)}{g(y)} = -\infty = \ell .$$

Ist nun $\ell \in \mathbb{R} \cup \{\infty\}$ und $c < \ell$ beliebig, so gibt es ein $\alpha_2 = \alpha_2(c) > a$ mit

$$\forall x \in (a, \alpha_2): \qquad \frac{f'(x)}{g'(x)} > c .$$

Analog wie im ersten Fall sieht man mit Hilfe des verallgemeinerten Mittelwertsatzes der Differentialrechnung,

$$\forall y \in (a, \alpha_2): \qquad \frac{f(y)}{g(y)} \ge c .$$

Insbesondere gilt im Fall $\ell = \infty$:

$$\frac{f(x)}{g(x)} \to \infty = \lim_{t\to a} \frac{f'(t)}{g'(t)} \quad \text{für } x \to a.$$

Im Fall $\ell \in \mathbb{R}$ gilt für beliebig vorgegebenes $\varepsilon > 0$ mit $\alpha := \min\{\alpha_1(\ell+\varepsilon/2), \alpha_2(\ell-\varepsilon/2)\}$ für alle $x \in (a, \alpha)$ nach dem bisher Gezeigten:

$$\left| \frac{f(x)}{g(x)} - \ell \right| \le \frac{\varepsilon}{2} < \varepsilon .$$

Damit folgt auch in diesem Fall

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} .$$

In der Situation (b) behandeln wir nur den Fall

$$\lim_{x\to a} f(x) = +\infty \text{ und } \lim_{x\to a} g(x) = +\infty \tag{5.6}$$

Die übrigen Fälle lassen sich leicht darauf zurückführen.

Ist $\ell \in \mathbb{R}\cup\{-\infty\}$ und $C > \ell$ beliebig, so gibt es zu $C_1 := \max\{C-1, (C+\ell)/2\} \in (\ell, C)$ nach Voraussetzung ein $\alpha_1 = \alpha_1(C) > a$ mit

$$\forall x \in (a, \alpha_1]: \qquad g(x) > 0,\ f(x) > 0 \text{ und } \frac{f'(x)}{g'(x)} < C_1\,.$$

Wegen (5.6) gibt es ein $\gamma_1 = \gamma_1(C)$ mit

$$\forall x \in (a, \gamma_1): \qquad g(x) > g(\alpha_1) \quad\text{und}\quad f(x) > f(\alpha_1)\,. \tag{5.7}$$

Nach dem verallgemeinerten Mittelwertsatz 5.21 der Differentialrechnung gibt es für alle $x \in (a, \gamma_1)$ ein $z(x) \in (x, \alpha_1)$ mit

$$0 < \frac{f(x) - f(\alpha_1)}{g(x) - g(\alpha_1)} = \frac{f'(z(x))}{g'(z(x))} < C_1\,. \tag{5.8}$$

Insbesondere muß $C_1 > 0$ sein. Da $C > \ell$ beliebig war, muß $\ell \geq 0$ gelten. Wegen (5.6) gibt es ein $\delta_1 = \delta_1(C) \in (a, \gamma_1)$ mit

$$\forall x \in (a, \delta_1): \qquad 0 < \frac{1 - \dfrac{g(\alpha_1)}{g(x)}}{1 - \dfrac{f(\alpha_1)}{f(x)}} < \frac{C}{C_1} \tag{5.9}$$

und es folgt mit (5.8) und (5.9) für alle $x \in (a, \delta_1)$:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(\alpha_1)}{g(x) - g(\alpha_1)} \cdot \frac{1 - \dfrac{g(\alpha_1)}{g(x)}}{1 - \dfrac{f(\alpha_1)}{f(x)}} = \frac{f'(z(x))}{g'(z(x))} < C_1 \cdot \frac{C}{C_1} = C.$$

Wegen $\ell \geq 0$ folgt im Grenzfall $\ell = 0$ hieraus

$$\lim_{x\to a} \frac{f(x)}{g(x)} = 0 = \lim_{x\to a} \frac{f'(x)}{g'(x)}\,.$$

Ist nun $\ell \in (0, \infty]$ und $0 < c < \ell$ beliebig, so gibt es zu $c_1 = \min\{c+1, (c+\ell)/2\} \in (c, \ell)$ ein $\alpha_2 = \alpha_2(c_1) > a$ mit

$$\forall x \in (a, \alpha_1): \qquad \frac{f'(x)}{g'(x)} > c_1\,.$$

Wie im Fall $\ell < \infty$ gibt es ein $\gamma_2 \in (a, \alpha_2)$, so daß (5.7) erfüllt ist. Nach dem verallgemeinerten Mittelwertsatz der Differentialrechnung gibt es für alle $x \in (a, \gamma_2)$ ein $z = z(x, \alpha_1) \in (x, \alpha_2)$ mit

$$\frac{f(x) - f(\alpha_2)}{g(x) - g(\alpha_2)} = \frac{f'(z(x))}{g'(z(x))} > c_1\,. \tag{5.10}$$

Wegen (5.6) und $0 < c/c_1 < 1$ gibt es ein $\delta_2 = \delta_2(c) \in (a, \gamma_2)$ mit

$$\forall x \in (a, \delta_2): \qquad \frac{1 - \dfrac{g(\alpha_2)}{g(x)}}{1 - \dfrac{f(\alpha_2)}{f(x)}} > \frac{c}{c_1} \tag{5.11}$$

und es folgt mit (5.10) und (5.11) für alle $x \in (a, \delta_2)$:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(\alpha_2)}{g(x) - g(\alpha_2)} \cdot \frac{1 - \dfrac{g(\alpha_2)}{g(x)}}{1 - \dfrac{f(\alpha_2)}{f(x)}} = \frac{f'(z(x))}{g'(z(x))} > c_1 \cdot \frac{c}{c_1} = c.$$

Insbesondere gilt im Fall $\ell = \infty$

$$\frac{f(x)}{g(x)} \to \infty = \lim_{t \to a} \frac{f'(t)}{g'(t)} \quad \text{für } x \to a.$$

Im Fall $\ell \in (0, \infty)$ gilt für beliebig vorgegebenes $\varepsilon > 0$ mit $\delta := \min\{\delta_1(\ell + \varepsilon/2), \delta_2(c)\}$ mit $0 < c < \ell$ und $c \geq \ell - \varepsilon/2$ für alle $x \in (a, \delta)$:

$$\left| \frac{f(x)}{g(x)} - \ell \right| < \varepsilon .$$

Damit folgt auch in diesem Fall

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)} .$$

□

Eine analoge Aussage gilt natürlich auch für Grenzwerte am rechten Intervallende und für beidseitige Grenzwerte.

Die nachstehende Folgerung aus dem Mittelwertsatz der Differentialrechnung ist häufig nützlich bei der Suche nach Abschätzungen.

5.25. Folgerung. *Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Es gebe Konstanten $m, M \in \mathbb{R}$ mit*

$$\forall x \in (a, b) : \qquad m \leq f'(x) \leq M .$$

Dann gilt für alle $x, y \in [a, b]$ mit $x < y$:

$$(y - x)m \leq f(y) - f(x) \leq M(y - x) .$$

Insbesondere ist f dann Lipschitz–stetig mit der Lipschitz–Konstanten $L = \max\{|m|, |M|\}$.

Beweis. Dies erhält man unmittelbar, indem man für alle $x, y \in [a, b]$ mit $x < y$ den Mittelwertsatz anwendet auf die Einschränkung von f auf $[x, y]$. □

5.26. Satz. *Sei $\alpha \in \mathbb{C}$ und sei $f : \mathbb{R} \to \mathbb{C}$ eine differenzierbare Funktion mit*

$$\forall x \in \mathbb{R} : \qquad f'(x) = \alpha f(x) .$$

Dann gilt

$$\forall x \in \mathbb{R} : \qquad f(x) = f(0)e^{\alpha x}. \tag{5.12}$$

Beweis. Die Funktion

$$h : \mathbb{R} \to \mathbb{C} , \qquad x \mapsto h(x) := f(x)e^{-\alpha x}$$

ist nach Beispiel 5.4 (b) und der Produktregel 5.7 (c) auf $\mathbb{R}$ differenzierbar und es gilt für alle $x \in \mathbb{R}$:

$$h'(x) = f'(x)e^{-\alpha x} - \alpha f(x)e^{-\alpha x} = (f'(x) - \alpha f(x))e^{-\alpha x} = 0 .$$

Nach Folgerung 5.23 ist h auf $\mathbb{R}$ konstant. Es gilt also für alle $x \in \mathbb{R}$:

$$h(x) = h(0) = f(0)e^{-\alpha \cdot 0} = f(0)$$

und damit (5.12). □

5.27. Satz. *Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ auf $[a, b]$ stetig und auf (a, b) differenzierbar.*

(a) *Genau dann ist f auf $[a, b]$ monoton wachsend (bzw. fallend), wenn für alle $x \in (a, b)$ gilt $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$).*
(b) *Gilt $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$, so ist f auf $[a, b]$ streng monoton wachsend (bzw. fallend).*

Beweis. Zu "$\Longrightarrow$" in (a): Sei also f auf $[a, b]$ monoton wachsend und sei $x \in (a, b)$ beliebig. Dann gilt

$$\forall \xi \in (a, x) : \qquad \frac{f(x) - f(\xi)}{x - \xi} \geq 0$$

und daher auch

$$f'(x) = \lim_{\xi \to x^-} \frac{f(x) - f(\xi)}{x - \xi} \geq 0\,.$$

Im monoton fallenden Fall verfährt man analog.

Zu (b) und zu "$\Longleftarrow$" in (a): Sei nun $f'(x) \geq 0$ (bzw. $f'(x) > 0$) für alle $x \in (a, b)$ und seien $x_1, x_2 \in [a, b]$ beliebig mit $x_1 < x_2$. Nach dem Mittelwertsatz der Differentialrechnung 5.20 (angewendet auf die Einschränkung von f auf $[x_1, x_2]$) gibt es also ein $x_0 \in (x_1, x_2) \subseteq (a, b)$ mit

$$f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}\,.$$

Wegen $x_2 - x_1 > 0$ und $f'(x_0) \leq 0$ (bzw. $f'(x_0) < 0$) folgt also $f(x_1) \leq f(x_2)$ (bzw. $f(x_1) < f(x_2)$). Die Fälle $f'(x) \geq 0$ (bzw. $f'(x) > 0$) auf (a, b) behandelt man analog. □

Bemerkung. Die Funktion $f : [-1, 1] \to \mathbb{R}$, $x \mapsto x^3$, ist auf $[-1, 1]$ streng monoton wachsend. f' besitzt aber in 0 eine Nullstelle. Dies zeigt, daß in Satz 5.27 (b) die Umkehrung nicht gilt.

Sei nun $I \subseteq \mathbb{R}$ ein beliebiges nicht zu einem Punkt ausgeartetes Intervall. Indem wir Satz 5.27 auf jedes abgeschlossene Intervall $[x_1, x_2]$ mit $x_1, x_2 \in I$ und mit $x_1 < x_2$ anwenden (unter Beachtung, daß nach Satz 5.6 jede auf I differenzierbare Funktion auch auf I stetig ist), erhalten wir:

5.28. Folgerung. *Sei $f : I \to \mathbb{R}$ eine auf einem nicht zu einem Punkt ausgearteten Intervall differenzierbare Funktion. Dann gilt mit $a := \inf I$ und $b := \sup I$:*

(a) *Genau dann ist f auf I monoton wachsend (bzw. fallend), wenn für alle $x \in (a, b)$ gilt $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$).*
(b) *Gilt $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$, so ist f auf I streng monoton wachsend (bzw. fallend).*

Wir erhalten nun ein hinreichendes Kriterium für das Vorliegen einer lokalen Extremstelle.

5.29. Satz. *Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall, sei $f : I \to \mathbb{R}$ eine auf I differenzierbare Funktion und sei x_0 ein Punkt aus I.*

(a) *Gibt es ein $\delta > 0$ mit $U_\delta(x_0) := (x_0 - \delta, x_0 + \delta) \subseteq I$ und gilt $f'(x) \leq 0$ auf $(x_0 - \delta, x_0)$ sowie $f'(x) \geq 0$ auf $(x_0, x_0 + \delta)$, so ist x_0 eine lokale Minimumstelle. Gilt sogar $f'(x) < 0$ auf $(x_0 - \delta, x_0)$ sowie $f'(x) > 0$ auf $(x_0, x_0 + \delta)$, so ist diese isoliert.*

(b) *Gibt es ein* $\delta > 0$ *mit* $U_\delta(x_0) := (x_0 - \delta, x_0 + \delta) \subseteq I$ *und gilt* $f'(x) \geq 0$ *auf* $(x_0 - \delta, x_0)$ *sowie* $f'(x) \leq 0$ *auf* $(x_0, x_0 + \delta)$, *so ist* x_0 *eine lokale Maximumstelle. Gilt sogar* $f'(x) > 0$ *auf* $(x_0 - \delta, x_0)$ *sowie* $f'(x) < 0$ *auf* $(x_0, x_0 + \delta)$, *so ist diese isoliert.*

Beweis. (a) Nach Folgerung 5.28 ist f auf $(x_0 - \delta, x_0]$ monoton fallend (bzw. streng monoton fallend) und auf $[x_0, x_0 + \delta)$ monoton wachsend (bzw. streng monoton wachsend). Insbesondere gilt also für alle $x \in (x_0 - \delta, x_0 + \delta)$ die Ungleichung $f(x) \geq f(x_0)$ (bzw. $f(x) > f(x_0)$). x_0 ist daher eine lokale Minimumstelle (bzw. eine isolierte lokale Minimumstelle).

(b) zeigt man analog. □

Eine weitere hinreichende Bedingung für das Vorliegen einer isolierten lokalen Extremstelle ist durch den folgenden Satz gegeben.

5.30. Satz. *Sei* $\emptyset \neq I \subseteq \mathbb{R}$ *ein offenes Intervall,* $f : I \to \mathbb{R}$ *auf* I *differenzierbar und sei* $x_0 \in I$ *mit* $f'(x_0) = 0$. *Existiert in* x_0 *die zweite Ableitung und ist* $f''(x_0) \neq 0$, *so ist* x_0 *eine isolierte lokale Extremstelle und zwar eine isolierte lokale Maximumstelle, falls* $f''(x_0) < 0$ *und eine isolierte lokale Minimumstelle im Fall* $f''(x_0) > 0$.

Beweis. Ist $f''(x_0) > 0$, so erhält man

$$0 < f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f'(x)}{x - x_0}.$$

Es gibt daher ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subseteq I$, so daß gilt:

$$\forall x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} : \quad \frac{f'(x)}{x - x_0} > 0\,.$$

Es folgt

$$\begin{aligned} f'(x) < 0 \quad &\text{für alle } x \in (x_0 - \delta, x_0) \text{ und} \\ f'(x) > 0 \quad &\text{für alle } x \in (x_0, x_0 + \delta)\,. \end{aligned}$$

Mit Satz 5.29 folgt: x_0 ist isolierte lokale Minimumstelle. Analog sieht man, daß im Fall $f''(x_0) < 0$ in x_0 eine lokale Maximumstelle vorliegt. □

Bemerkung. Die Bedingung $f''(x_0) \neq 0$ im vorstehenden Satz ist hinreichend, aber nicht notwendig für das Vorliegen einer isolierten lokalen Extremstelle. So gilt für die Funktion

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x) := x^4,$$

es ist $f'(0) = f''(0) = 0$. Wegen $f'(x) = 4x^3 < 0$ für $x < 0$ und $f'(x) = 4x^3 > 0$ für $x > 0$ besitzt f nach Satz 5.29 in 0 eine isolierte lokale Minimumstelle.

Beispiele. (a) 0 ist nach Satz 5.30 isolierte lokale Minimumstelle der Funktion

$$p : \mathbb{R} \to \mathbb{R}, \quad x \mapsto p(x) := x^4 - 7x^3 + 2x^2 + 2002\,,$$

denn $p'(0) = 0$ und $p''(0) = 4 > 0$.

(b) Für alle $n \in \mathbb{Z}$ ist $2n\pi$ isolierte lokale Maximumstelle und $(2n+1)\pi$ isolierte lokale Minimumstelle von $\cos : \mathbb{R} \to \mathbb{R}$, denn

$$\begin{aligned} \cos'(2n\pi) &= -\sin(2n\pi) = -\sin(0) = 0 \text{ und} \\ \cos''(2n\pi) &= -\cos(2n\pi) = -\cos(0) = -1 < 0 \end{aligned}$$

und

$$\cos'((2n+1)\pi) = -\sin((2n+1)\pi) = -\sin(\pi) = 0 \text{ und}$$
$$\cos''((2n+1)\pi) = -\cos((2n+1)\pi) = -\cos(\pi) = 1 > 0\,.$$

3. Konvexe Funktionen

5.31. Definition. Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt *konvex*, falls gilt:

$$(5.13)\quad \forall x_1, x_2 \in I \text{ mit } x_1 < x_2\, \forall \alpha \in (0,1): \quad f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)\,.$$

Gilt hierbei stets "$<$" statt "$\le$", so heißt f *streng konvex*. f heißt *konkav* bzw. *streng konkav*, falls die Funktion $-f$ konvex bzw. streng konvex ist.

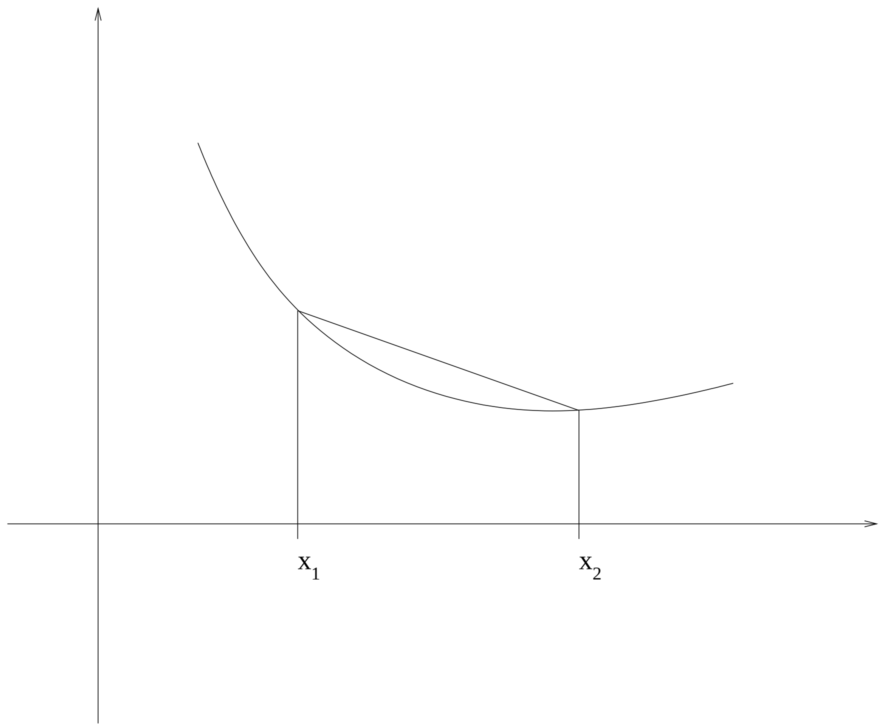


Abbildung 4. Die Ungleichungen (5.13) und (5.16) besagen, daß die Punkte der Sehne zwischen $(x_1, f(x_1))$ und $(x_2, f(x_2))$ stets oberhalb des Graphen von f liegen müssen.

Konvexe Funktionen lassen sich auch wie folgt charakterisieren:

5.32. Satz. *Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall. Für eine Funktion $f : I \to \mathbb{R}$ sind die folgenden Aussagen äquivalent:*

(a) *f ist konvex (bzw. streng konvex) auf I.*
(b) *Für alle $x_1, x_2, x \in I$ mit $x_1 < x < x_2$ gilt:*

$$(5.14)\qquad f(x) \le \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)$$

bzw.

$$f(x) < \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)\,.$$

(c) *Für alle $x_1, x_2, x \in I$ mit $x_1 < x < x_2$ gilt:*

$$\frac{f(x)-f(x_1)}{x-x_1} \leq \frac{f(x_2)-f(x)}{x_2-x} \tag{5.15}$$

bzw.

$$\frac{f(x)-f(x_1)}{x-x_1} < \frac{f(x_2)-f(x)}{x_2-x}.$$

(d) *Für alle $x_1, x_2, x \in I$ mit $x_1 < x < x_2$ gilt:*

$$f(x) \leq f(x_1) + \frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1) \tag{5.16}$$

bzw.

$$f(x) < f(x_1) + \frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1)$$

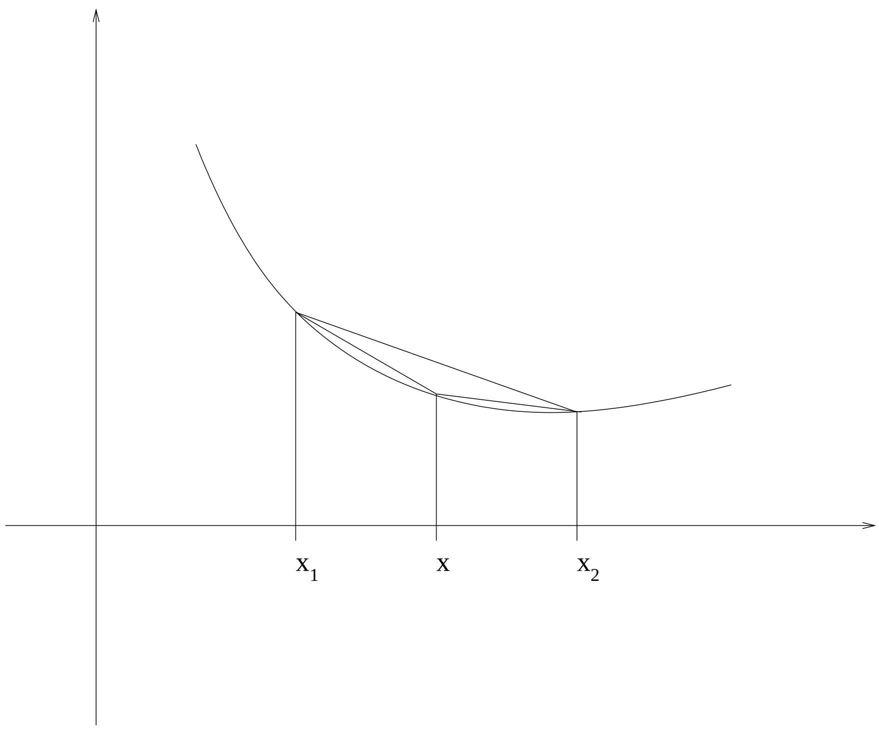


Abbildung 5. Ungleichung (5.15) besagt, daß die Steigungen aufeinanderfolgender Sekantenabschnitte monoton wachsen

Beweis. "(a)$\Longleftrightarrow$(b)" Wählen wir in (5.13) $\alpha := \dfrac{x_2-x}{x_2-x_1}$, so folgt durch elementare Rechnung $1-\alpha = \dfrac{x-x_1}{x_2-x_1}$ und $x = \alpha x_1 + (1-\alpha)x_2$. Aus (5.13) folgt also (5.14). Ist umgekehrt $x_1, x_2 \in I$ mit $x_1 < x_2$ und $\alpha \in (0,1)$ gegeben, so folgt mit $x := \alpha x_1 + (1-\alpha)x_2$ durch Auflösen nach α: $\alpha = \dfrac{x_2-x}{x_2-x_1}$ und $1-\alpha = \dfrac{x-x_1}{x_2-x_1}$. Aus (5.14) folgt also (5.13). Man sieht auch, daß genau dann in (5.13) stets das $<$–Zeichen steht, wenn es stets in (5.14) steht.

"(b)$\Longleftrightarrow$(d)" ist offensichtlich, da, wie man durch elementare Umformungen sieht, die rechten Seiten von (5.14) und (5.16) übereinstimmen.

"(b)$\Longleftrightarrow$(c)" Indem man die Ungleichung (5.15) mit $(x-x_1)(x_2-x)$ durchmultipliziert sieht man, daß diese äquivalent ist zu

$$(x_2 - x)(f(x) - f(x_1)) \leq (x - x_1)(f(x_2) - f(x))$$

was äquivalent ist zu

$$(x_2 - x_1)f(x) \leq (x_2 - x)f(x_1) + (x_2 - x)f(x_2)$$

was offensichtlich äquivalent ist zu (5.14). Die Umformungen zeigen auch, daß genau dann in (5.15) das $<$–Zeichen steht, wenn es in (5.14) steht. □

5.33. Satz. *Sei $f : I \to \mathbb{R}$ eine differenzierbare Funktion auf einem nicht zu einem Punkt ausgearteten Intervall $I \subseteq \mathbb{R}$. Dann gilt:*

(a) *Genau dann ist f auf I konvex (bzw. streng konvex), wenn $f'(x)$ monoton (bzw. streng monoton) wachsend auf I ist.*
(b) *Genau dann ist f auf I konkav (bzw. streng konkav), wenn $f'(x)$ monoton (bzw. streng monoton) fallend auf I ist.*

Beweis. (a) "$\Longrightarrow$": Sei also f auf I konvex (bzw. streng konvex) und seien $x_1, x_2 \in I$ beliebig. Dann gilt für alle $u, v, x \in I$ mit $x_1 < u < x < v < x_2$:

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}, \quad \text{bzw.} \quad \frac{f(x) - f(x_1)}{x - x_1} < \frac{f(x_2) - f(x)}{x_2 - x} \tag{5.17}$$

sowie

$$\frac{f(u) - f(x_1)}{u - x_1} \leq \frac{f(x) - f(u)}{x - u} \quad \text{und} \quad \frac{f(v) - f(x)}{v - x} \leq \frac{f(x_2) - f(v)}{x_2 - v}. \tag{5.18}$$

Führen wir in (5.18) die Grenzübergänge für $u \to x_1$ und $v \to x_2$ durch, so erhalten wir

$$f'(x_1) \leq \frac{f(x) - f(x_1)}{x - x_1} \quad \text{und} \quad f'(x_2) \geq \frac{f(x_2) - f(x)}{x_2 - x},$$

woraus wegen (5.17) $f'(x_1) \leq f'(x_2)$ (bzw. $f'(x_1) < f'(x_2)$) folgt. f' ist also monoton (bzw. streng monoton) wachsend.

"$\Longleftarrow$": Sei nun f' monoton (bzw. streng monoton) wachsend auf I und seien $x_1, x_2, x \in I$ beliebig mit $x_1 < x < x_2$. Nach dem Mittelwertsatz der Differentialrechnung 5.20 gibt es Punkte $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$ mit

$$f'(\xi_1) = \frac{f(x) - f(x_1)}{x - x_1} \quad \text{und} \quad f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Nach Voraussetzung ist f' auf I monoton (bzw. streng monoton) wachsend. Es folgt daher:

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x} \quad \textit{bzw.} \quad \frac{f(x) - f(x_1)}{x - x_1} < \frac{f(x_2) - f(x)}{x_2 - x}.$$

Die Funktion f erfüllt also die Bedingung in Teil (c) von Satz 5.32 und ist daher nach diesem Satz konvex (bzw. streng konvex).

(b) folgt durch Anwenden von (a) auf die Funktion $-f$. □

5.34. Folgerung. *Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall und sei $f : I \to \mathbb{R}$ eine zweimal differenzierbare Funktion. Dann gilt:*

(a) *Genau dann ist f auf I konvex (bzw. konkav), wenn $f''(x) \geq 0$ (bzw. $f''(x) \leq 0$) für alle $x \in I$ gilt.*
(b) *Ist $f''(x) > 0$ (bzw. $f''(x) < 0$) für alle $x \in I$, so ist f auf I streng konvex (bzw. streng konkav).*

Beweis. Dies folgt unmittelbar aus Folgerung 5.28 und Satz 5.33 □

5.35. Definition. Sei I ein nicht zu einem Punkt ausgeartetes Intervall. Ein Punkt $x_0 \in I$ heißt *Wendepunkt* von f, falls es ein $\delta > 0$ gibt mit $(x_0 - \delta, x_0 + \delta) \subseteq I$, so daß f auf $(x_0 - \delta, x_0)$ konvex und auf $(x_0, x_0 + \delta)$ konkav ist oder dies für die Funktion $-f$ gilt.

5.36. Folgerung. *Hat eine differenzierbare Funktion $f : I \to \mathbb{R}$ (I wie vor) in einem Punkt $x_0 \in I$ einen Wendepunkt, so besitzt f' in x_0 eine lokale Extremstelle. Ist also f in x_0 sogar zweimal differenzierbar, so ist $f''(x_0) = 0$. Umgekehrt gilt: gibt es ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subseteq I$ und ist f' auf $(x_0 - \delta, x_0)$ monoton wachsend und auf $(x_0, x_0 + \delta)$ monoton fallend oder gilt dies für $-f'$, so besitzt f nach Definition 5.35 und Satz 5.33 in x_0 einen Wendepunkt. Dies ist insbesondere dann der Fall, wenn f in einer Umgebung von x_0 zweimal differenzierbar ist und f'' in x_0 einen Vorzeichenwechsel besitzt.*

KAPITEL 6

Integralrechnung in einer Veränderlichen

1. Das Riemann–Integral

In diesem Abschnitt sei $I = [a, b]$ stets ein beschränktes, abgeschlossenes Intervall mit $-\infty < a < b < \infty$.

6.1. Definition. (a) Eine endliche Teilmenge $\mathfrak{Z}$ von I mit $a, b \in \mathfrak{Z}$ nennt man eine *Zerlegung* oder eine *Partition* von I. Ist $\mathfrak{Z}$ eine Zerlegung von I mit $n+1$ paarweise verschiedenen Punkten, so können wir diese der Größe nach durchnumerieren und schreiben dann

$$\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\}\,.$$

Für $j = 1, \ldots, n$ nennen wir $[t_{j-1}, t_j]$ das *j–te Teilintervall* von $\mathfrak{Z}$. Die maximale Länge der Teilintervalle

$$\delta(\mathfrak{Z}) := \max\{t_j - t_{j-1}\,;\, 1 \le j \le n\}$$

heißt die *Feinheit der Zerlegung* $\mathfrak{Z}$ und $\mathfrak{Z}(I)$ bezeichne die Menge aller Zerlegungen des Intervalls I.

(b) Sind $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ mit $\mathfrak{Z}_1 \subseteq \mathfrak{Z}_2$, so nennen wir $\mathfrak{Z}_2$ eine *Verfeinerung von* $\mathfrak{Z}_1$ und sagen $\mathfrak{Z}_1$ ist *feiner als* $\mathfrak{Z}_2$. Offensichtlich gilt dann $\delta(\mathfrak{Z}_1) \ge \delta(\mathfrak{Z}_2)$.

(c) Sei $f : I \to \mathbb{R}$ eine beschränkte reellwertige Funktion auf I. Dann ist

$$\|f\|_I := \sup_{t \in I} |f(t)| < \infty\,.$$

Wir nennen $\|f\|_I$ die *Supremumsnorm* von f auf I. Ist

$$\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}(I),$$

so existieren für $j = 1, \ldots, n$ die Infima und Suprema von f auf dem j–ten Teilintervall,

$$m_j := \inf f([t_{j-1}, t_j])\,, \quad M_j := \sup f([t_{j-1}, t_j])\,. \tag{6.1}$$

Wir setzen

$$U(f, \mathfrak{Z}) := \sum_{j=1}^{n} m_j (t_j - t_{j-1}) \quad \text{und} \quad O(f, \mathfrak{Z}) := \sum_{j=1}^{n} M_j (t_j - t_{j-1}) \tag{6.2}$$

und nennen $U(f, \mathfrak{Z})$ die *Untersumme* und $O(f, \mathfrak{Z})$ die *Obersumme von f zur Zerlegung $\mathfrak{Z}$.*

6.2. Lemma. *Sei $f : I \to \mathbb{R}$ eine beschränkte Funktion und $m := \inf f(I)$, $M := \sup f(I)$.*

(a) *Für alle $\mathfrak{Z} \in \mathfrak{Z}(I)$ gilt $U(f, \mathfrak{Z}) \le O(f, \mathfrak{Z})$.*

(b) *Sind $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ zwei Zerlegungen mit $\mathfrak{Z}_1 \subset \mathfrak{Z}_2$ und ist $|\mathfrak{Z}_2 \setminus \mathfrak{Z}_1| = p$ (d.h. $\mathfrak{Z}_2$ hat genau p zusätzliche Unterteilungspunkte), so gilt*

$$\begin{aligned} 0 &\le U(f, \mathfrak{Z}_2) - U(f, \mathfrak{Z}_1) \le 2p\delta(\mathfrak{Z}_1)\|f\|_I\,, \\ 0 &\le O(f, \mathfrak{Z}_1) - O(f, \mathfrak{Z}_2) \le 2p\delta(\mathfrak{Z}_1)\|f\|_I\,. \end{aligned} \tag{6.3}$$

Wegen (a) gilt also insbesondere:

$$U(f, \mathfrak{Z}_1) \le U(f, \mathfrak{Z}_2) \le O(f, \mathfrak{Z}_2) \le O(f, \mathfrak{Z}_1).$$

(c) *Für alle* $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ *gilt* $U(f, \mathfrak{Z}_1) \leq O(f, \mathfrak{Z}_2)$.
(d) *Es existieren das Unterintegral*

$$U_a^b(f) := \sup\{U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\}$$

und das Oberintegral

$$O_a^b(f) := \inf\{O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\}$$

und es gilt: $m(b-a) \leq U_a^b(f) \leq O_a^b(f) \leq M(b-a)$.

Beweis. (a) Sei $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}(I)$ beliebig. Nach den Definitionen in (6.2) gilt (mit m_j, M_j wie in (6.1)):

$$O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) = \sum_{j=1}^{n} (M_j - m_j)(t_j - t_{j-1}) \geq 0.$$

(b) Wir beweisen nur die erste Ungleichung in (6.3), die zweite zeigt man analog oder durch Anwenden der ersten auf die Funktion $-f$ (unter Beachtung von $O(f, \mathfrak{Z}) = -U(-f, \mathfrak{Z})$ für alle $\mathfrak{Z}\in\mathfrak{Z}(I)$). Wir führen den Beweis durch vollständige Induktion nach p:

Induktionsanfang: Sei $p = 1$, also $\mathfrak{Z}_2 = \mathfrak{Z}_1 \cup \{u\}$ mit $u \notin \mathfrak{Z}_1$. Ist

$$\mathfrak{Z}_1 = \{a = t_0 < t_1 < \cdots < t_n = b\}\,,$$

so muß $u \in (t_{k-1}, t_k)$ für genau ein $k \in \{1, \ldots, n\}$ gelten. Seien m_j, M_j für $j = 1, \ldots, n$ zu f und $\mathfrak{Z}_1$ gemäß (6.1) gebildet. Wir setzen zusätzlich $m' := \inf f([t_{k-1}, u])$ und $m'' := \inf f([u, t_k])$. Dann ist $m_k \leq m'$ und $m_k \leq m''$ und es folgt

$$\begin{aligned} U(f, \mathfrak{Z}_2) - U(f, \mathfrak{Z}_1) = &\sum_{j=1}^{k-1} m_j(t_j - t_{j-1}) + m'(u - t_{k-1}) + m''(t_k - u) + \\ &+ \sum_{j=k+1}^{n} m_j(t_j - t_{j-1}) - \sum_{j=1}^{n} m_j(t_j - t_{j-1}) \\ =&m'(u - t_{k-1}) + m''(t_k - u) - m_k(t_k - t_{k-1}) = \\ =&(m' - m_k)(u - t_{k-1}) + (m'' - m_k)(t_k - u) \geq 0\,. \end{aligned} \tag{6.4}$$

Ferner gilt wegen $m_k \leq m'$ und $m_k \leq m''$:

$$\begin{aligned} 0 \leq m' - m_k = \inf f([t_{k-1}, u]) - &\inf f([t_{k-1}, t_k]) \leq \\ \leq \quad \|f\|_I \quad + &\quad \|f\|_I \quad = 2\|f\|_I\,. \end{aligned}$$

Ebenso erhält man

$$0 \leq m_k - m'' \leq 2\|f\|_I\,.$$

Setzen wir dies in (6.4) ein, so folgt

$$\begin{aligned} 0 \leq U(f, \mathfrak{Z}_2) - U(f, \mathfrak{Z}_1) &= (m' - m_k)(u - t_{k-1}) + (m'' - m_k)(t_k - u) \\ &\leq 2\|f\|_I(u - t_{k-1} + t_k - u) = 2\|f\|_I(t_k - t_{k-1}) \\ &\leq 2\|f\|_I\delta(\mathfrak{Z}_1)\,, \end{aligned}$$

womit die erste Ungleichung in (6.3) im Fall $p = 1$ gezeigt ist.

Sei nun schon die erste Ungleichung in (6.3) für ein $p \in \mathbb{N}$ bewiesen und seien $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ mit $\mathfrak{Z}_1 \subset \mathfrak{Z}_2$ und $|\mathfrak{Z}_2 \setminus \mathfrak{Z}_1| = p + 1$, also $\mathfrak{Z}_2 = \mathfrak{Z}_1 \cup \{u_1, \ldots, u_p, u_{p+1}\}$ mit paarweise verschiedenen $u_1, \ldots, u_p, u_{p+1} \in I \setminus \mathfrak{Z}_1$. Dann folgt mit $\mathfrak{Z}' := \mathfrak{Z}_1 \cup \{u_1, \ldots, u_p\}$ (und unter Beachtung von $\delta(\mathfrak{Z}') \le \delta(\mathfrak{Z}_1)$) nach Induktionsvoraussetzung

$$\begin{aligned} U(f,\mathfrak{Z}_2) - U(f,\mathfrak{Z}_1) &= \underbrace{U(f,\mathfrak{Z}_2) - U(f,\mathfrak{Z}')}_{\ge 0} + \underbrace{U(f,\mathfrak{Z}') - U(f,\mathfrak{Z}_1)}_{\ge 0} \\ &\le \|f\|_I \delta(\mathfrak{Z}') + 2p\|f\|_I \delta(\mathfrak{Z}_1) \le 2(p+1)\|f\|_I \delta(\mathfrak{Z}_1) \,. \end{aligned}$$

Damit ist gezeigt, daß die erste Ungleichung in (6.3) für alle $p \in \mathbb{N}$ gilt.

(c) Da $\mathfrak{Z}_3 := \mathfrak{Z}_1 \cup \mathfrak{Z}_2$ eine Verfeinerung von $\mathfrak{Z}_1$ und von $\mathfrak{Z}_2$ ist, erhalten wir mit (a) und (b):

$$U(f,\mathfrak{Z}_1) \le U(f,\mathfrak{Z}_3) \le O(f,\mathfrak{Z}_3) \le O(f,\mathfrak{Z}_2)\,.$$

(d) Für alle $\mathfrak{Z}, \mathfrak{Z}_1 \in \mathfrak{Z}(I)$ gilt nach (c): $U(f,\mathfrak{Z}) \le O(f,\mathfrak{Z}_1)$. Nach dem Supremumsaxiom existiert also

$$U_a^b(f) := \sup\{U(f,\mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\}$$

und es gilt für alle $\mathfrak{Z}_1 \in \mathfrak{Z}(I)$:

$$U_a^b(f) \le O(f,\mathfrak{Z}_1)\,.$$

Nochmals nach dem Supremumsaxiom existiert

$$O_a^b(f) := \inf\{O(f,\mathfrak{Z}_1)\,;\, \mathfrak{Z}_1 \in \mathfrak{Z}(I)\}$$

und es gilt $U_a^b(f) \le O_a^b(f)$. Insbesondere folgt für alle $\mathfrak{Z} \in \mathfrak{Z}(I)$:

$$U(f,\mathfrak{Z}) \le U_a^b(f) \le O_a^b(f) \le O(f,\mathfrak{Z})\,.$$

Mit der speziellen Wahl $\mathfrak{Z} := \{a = t_0 < t_1 = b\}$ erhalten wir die behauptete Ungleichung. □

Wir können nun das Riemann–Integral[1] definieren:

6.3. Definition. Eine beschränkte Funktion $f := I \to \mathbb{R}$ heißt auf $I = [a,b]$ *Riemann–integrierbar*, falls $U_a^b(f) = O_a^b(f)$. Wir nennen dann

$$\int_a^b f(t)dt := U_a^b(f) = O_a^b(f)$$

das *Riemann–Integral* von f über $[a,b]$. Die Menge der auf I reellwertigen, Riemann–integrierbaren Funktionen bezeichnen wir mit $\mathcal{R}(I)$.

Das Riemann–Integral können wir auch für $\mathbb{C}$–wertige Funktionen einführen: Wir nennen eine komplexwertige, beschränkte Funktion $f : I = [a,b] \to \mathbb{C}$ auf I Riemann–integrierbar, genau dann, wenn die reellwertigen Funktionen $\operatorname{Re} f, \operatorname{Im} f : I \to \mathbb{R}$ auf I Riemann–integrierbar sind. Ist dies der Fall, so sei

$$\int_a^b f(x)dx := \int_a^b \operatorname{Re} f(x)dx + i\int_a^b \operatorname{Im} f(x)dx\,.$$

Die Menge der auf I komplexwertigen, Riemann–integrierbaren Funktionen bezeichnen wir mit $\mathcal{R}(I,\mathbb{C})$.

Entsprechend gehen wir für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ im $\mathbb{K}^N$–wertigen Fall vor: Eine beschränkte, $\mathbb{K}^N$–wertige Funktion $f = (f_1, \ldots, f_N) : I = [a,b] \to \mathbb{K}^N$ heiße auf I Riemann–integrierbar genau dann, wenn die $\mathbb{K}$–wertigen Funktionen $f_j : I \to \mathbb{K}$ für $j = 1, \ldots, N$ auf I Riemann–integrierbar sind. Ist dies der Fall, so sei

$$\int_a^b f(x)dx := \left(\int_a^b f_1(x)dx, \ldots, \int_a^b f_N(x)dx\right)\,.$$

[1] Georg Friedrich Bernhard Riemann (17.9.1826–20.7.1866)

Die Menge der auf I Riemann–integrierbaren, $\mathbb{K}^N$–wertigen Funktionen bezeichnen wir mit $\mathcal{R}(I, \mathbb{K}^N)$.

Ist $f : I \to [0, \infty)$ eine nichtnegative Riemann–integrierbare Funktion, so nennen wir $\int_a^b f(x)dx$ auch den *Flächeninhalt* des Bereichs $B_f := \{(x, y)\,;\, a \leq x \leq b, 0 \leq y \leq f(x)\}$ zwischen dem Graphen von f und der x–Achse.

Wir geben nun einige äquivalente Beschreibungen der Riemann–Integrierbarkeit:

6.4. Satz. *Für eine beschränkte Funktion $f : I \to \mathbb{R}$ sind äquivalent:*

(a) *f ist auf I Riemann–integrierbar.*
(b) $\forall \varepsilon > 0\, \exists\, \mathfrak{Z} \in \mathfrak{Z}(I) : \quad O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) < \varepsilon$.
(c) *Es gibt eine Folge $(\mathfrak{Z}_n)_{n=1}^\infty$ aus $\mathfrak{Z}(I)$ mit*
$$\lim_{n\to\infty} U(f, \mathfrak{Z}_n) = \lim_{n\to\infty} O(f, \mathfrak{Z}_n)\,.$$
(d) *Es gibt eine Folge $(\mathfrak{Z}_n)_{n=1}^\infty$ aus $\mathfrak{Z}(I)$ mit*
$$\sup_{n\in\mathbb{N}} U(f, \mathfrak{Z}_n) = \inf_{n\in\mathbb{N}} O(f, \mathfrak{Z}_n)\,.$$

Ist (c) bzw. (d) erfüllt, so gilt
$$\int_a^b f(x)dx = \lim_{n\to\infty} U(f, \mathfrak{Z}_n) = \lim_{n\to\infty} O(f, \mathfrak{Z}_n),$$
bzw.
$$\int_a^b f(x)dx = \sup_{n\in\mathbb{N}} U(f, \mathfrak{Z}_n) = \inf_{n\in\mathbb{N}} O(f, \mathfrak{Z}_n).$$

Beweis. “(a)$\Longrightarrow$(b)”: Ist $f : I \to \mathbb{R}$ Riemann–integrierbar, so ist $U_a^b(f) = O_a^b(f)$. Nach Definition des Ober– und des Unterintegrals gibt es nach Lemma 1.16 zu beliebig vorgegebenem $\varepsilon > 0$ Zerlegungen $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ mit
$$U_a^b(f) - \frac{\varepsilon}{2} < U(f, \mathfrak{Z}_1) \leq U_a^b(f) \quad \text{und} \quad O_a^b(f) \leq O(f, \mathfrak{Z}_2) < O_a^b(f) + \frac{\varepsilon}{2}\,.$$
Für $\mathfrak{Z} := \mathfrak{Z}_1 \cup \mathfrak{Z}_2$ erhalten wir dann unter Verwendung von Lemma 6.2:
$$U_a^b(f) - \frac{\varepsilon}{2} < U(f, \mathfrak{Z}_1) \leq U(f, \mathfrak{Z}) \leq U_a^b(f) \leq O_a^b(f) \leq O(f, \mathfrak{Z}) \leq O(f, \mathfrak{Z}_2) < O_a^b(f) + \frac{\varepsilon}{2}$$
und hieraus wegen $O_a^b(f) = U_a^b(f)$
$$O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) < \varepsilon\,.$$

“(b)$\Longrightarrow$(c)”: Ist (b) erfüllt, so gibt es zu jedem $n \in \mathbb{N}$ eine Zerlegung $\mathfrak{Z}_n \in \mathfrak{Z}(I)$ mit $0 \leq O(f, \mathfrak{Z}_n) - U(f, \mathfrak{Z}_n) < \varepsilon_n := 1/n$. Insbesondere erfüllt die Folge $(\mathfrak{Z}_n)_{n=1}^\infty$ die Forderung (c).

“(c)$\Longrightarrow$(d)”: Ist $(\mathfrak{Z}_n)_{n=1}^\infty$ eine Folge in $\mathfrak{Z}(I)$, welche der Bedingung in (c) genügt, so gilt für alle $k \in \mathbb{N}$:
$$0 \leq \inf_{n\in\mathbb{N}} O(f, \mathfrak{Z}_n) - \sup_{n\in\mathbb{N}} U(f, \mathfrak{Z}_n) \leq O(f, \mathfrak{Z}_k) - U(f, \mathfrak{Z}_k) \to 0 \text{ für } k \to \infty\,.$$
Also erfüllt die Folge $(\mathfrak{Z}_n)_{n=1}^\infty$ auch (d).

“(d)$\Longrightarrow$(a)” und Beweis des Zusatzes: Sei nun $(\mathfrak{Z}_n)_{n=1}^\infty$ eine Folge in $\mathfrak{Z}(I)$, welche (d) erfüllt. Dann gilt mit $J := \sup_{n\in\mathbb{N}} U(f, \mathfrak{Z}_n) = \inf_{n\in\mathbb{N}} O(f, \mathfrak{Z}_n)$:
$$J \leq \sup\{U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = U_a^b(f) \leq O_a^b(f) = \inf\{O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} \leq J\,.$$
Also gilt $J = U_a^b(f) = O_a^b(f)$. f ist daher auf $I = [a, b]$ Riemann–integrierbar und es ist $J = \int_a^b f(t)dt$. $\square$

6.5. Beispiele. (a) Sei $c \in \mathbb{R}$ und sei $f_c : I \longrightarrow \mathbb{R}$ die konstante Funktion mit $f(t) = c$ für alle $t \in \mathbb{R}$. Dann ist f_c auf I Riemann-integrierbar und es gilt $\int_a^b f_c(t)dt = (b-a)c$.

(b) (Archimedes) Die Funktion $f : [0,1] \longrightarrow \mathbb{R}$ mit $f(x) := x^2$ für alle $x \in [0,1]$ ist auf I Riemann-integrierbar und es gilt $\int_0^1 f(x)dx = 1/3$.

(c) Für die Funktion $f : [0,1] \longrightarrow \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{für } x \in [0,1] \cap \mathbb{Q} \\ 0 & \text{für } x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

gilt $U_0^1(f) = 0$ und $O_0^1(f) = 1$ und daher $f \notin \mathcal{R}([0,1])$.

Beweis. (a) In diesem Fall gilt $m = M = c$ in Lemma 6.2 und es folgt die Riemann–Integrierbarkeit von f_c nach Definition 6.3 und Lemma 6.2 sowie $\int_a^b f_c(t)dt = (b-a)c$.

(b) Wir betrachten die Zerlegungsfolge $(\mathfrak{Z}_n)_{n=1}^\infty$ mit

$$\mathfrak{Z}_n := \left\{0 = \frac{0}{n} < \frac{1}{n} < \cdots < \frac{n}{n} = 1\right\} \quad \text{für alle } n \in \mathbb{N}\,.$$

Da f auf $[0,1]$ monoton wachsend ist, folgt hier für $j = 1, \ldots, n$:

$$m_j := \inf f\left(\left[\frac{j-1}{n}, \frac{j}{n}\right]\right) = \left(\frac{j-1}{n}\right)^2 \text{ und } M_j := \sup f\left(\left[\frac{j-1}{n}, \frac{j}{n}\right]\right) = \left(\frac{j}{n}\right)^2$$

und somit

$$O(f, \mathfrak{Z}_n) = \sum_{j=1}^n \left(\frac{j}{n}\right)^2 \cdot \frac{1}{n} = \frac{n(2n+1)(n+1)}{6n^3} \longrightarrow \frac{1}{3} \text{ für } n \to \infty$$

sowie

$$U(f, \mathfrak{Z}_n) = \sum_{j=1}^n \left(\frac{j-1}{n}\right)^2 \cdot \frac{1}{n} = O(f, \mathfrak{Z}_n) - \frac{1}{n} \longrightarrow \frac{1}{3} \text{ für } n \to \infty\,.$$

Nach Satz 6.4 ist f auf $[0,1]$ Riemann–integrierbar und es gilt $\int_0^1 f(x)dx = 1/3$.

(c) Sei $\mathfrak{Z} = \{0 = t_0 < t_1 < \cdots < t_n = 1\} \in \mathfrak{Z}([0,1])$ beliebig. Nach Definition von f und Satz 1.43 gilt hier für $k = 1, \ldots, n$:

$$m_j := \inf f([t_{j-1}, t_j]) = 0 \text{ und } M_j := \sup f([t_{j-1}, t_j]) = 1\,.$$

Es gilt also $U(f, \mathfrak{Z}) = 0$ und $O(f, \mathfrak{Z}) = 1$ für alle $\mathfrak{Z} \in \mathfrak{Z}([0,1])$. Hieraus folgt die Behauptung. □

6.6. Lemma. *Jede auf $I = [a,b]$ monotone Funktion $f : I \longrightarrow \mathbb{R}$ ist auf I Riemann–integrierbar.*

Beweis. Sei $f : I \longrightarrow \mathbb{R}$ monoton wachsend oder monoton fallend und sei

$$\mathfrak{Z}_n := \left\{a = t_0 := a + \frac{0(b-a)}{n} < t_1 := a + \frac{1(b-a)}{n} < \cdots < t_n := a + \frac{n(b-a)}{n} = b\right\}$$

eine *äquidistante* Zerlegung von I (mit also n Teilintervallen, die alle die Länge $(b-a)/n$ haben). Dann gilt für $j = 1, \ldots, n$:

$$m_j := \inf f([t_{j-1}, t_j]) = \begin{cases} f(t_{j-1}) & \text{im monoton wachsenden Fall} \\ f(t_j) & \text{im monoton fallenden Fall} \end{cases}$$

sowie

$$M_j := \sup f([t_{j-1}, t_j]) = \begin{cases} f(t_j) & \text{im monoton wachsenden Fall} \\ f(t_{j-1}) & \text{im monoton fallenden Fall.} \end{cases}$$

Wir erhalten im monoton wachsenden Fall

$$0 \leq O(f, \mathfrak{Z}_n) - U(f, \mathfrak{Z}_n) = \sum_{j=1}^{n} f(t_j)\frac{b-a}{n} - \sum_{j=1}^{n} f(t_{j-1})\frac{b-a}{n} =$$

$$= (f(b) - f(a))\frac{b-a}{n} \to 0 \text{ für } n \to \infty$$

und im monoton fallenden Fall

$$0 \leq O(f, \mathfrak{Z}_n) - U(f, \mathfrak{Z}_n) = \sum_{j=1}^{n} f(t_{j-1})\frac{b-a}{n} - \sum_{j=1}^{n} f(t_j)\frac{b-a}{n} =$$

$$= (f(a) - f(b))\frac{b-a}{n} \to 0 \text{ für } n \to \infty\,.$$

Mit Satz 6.4 folgt die Behauptung. □

Im folgenden Lemma fassen wir einige Rechenregeln für das Riemann–Integral zusammen:

6.7. Lemma. *Für alle $f, g \in \mathcal{R}(I)$ gilt:*

(a) *Für alle $\lambda \in \mathbb{R}$ ist die Funktion $\lambda f : x \mapsto \lambda f(x)$ auf I Riemann–integrierbar und es ist*

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx\,. \tag{6.5}$$

(b) *Die Funktion $f + g : x \mapsto f(x) + g(x)$ ist auf I Riemann–integrierbar und es gilt*

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx\,.$$

(c) *Ist $f \leq g$ auf I (d.h. ist $f(x) \leq g(x)$ für alle $x \in I$), so gilt auch*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx\,.$$

(d) *Mit $f_+(x) := \max\{f(x), 0\}$ und $f_-(x) := -\min\{f(x), 0\}$ für alle $x \in I$ gilt $f_+, f_-, |f| \in \mathcal{R}(I)$. Ferner gilt die Abschätzung*

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx \leq (b-a)\|f\|_I\,.$$

(e) *Für alle $1 \leq p \in \mathbb{R}$ ist $|f|^p \in \mathcal{R}(I)$.*

(f) *$fg \in \mathcal{R}(I)$.*

Beweis. (a) aus der Definition von Ober– und Untersummen folgt für alle $\mathfrak{Z} \in \mathfrak{Z}(I)$:

$$U(\lambda f, \mathfrak{Z}) = \begin{cases} \lambda U(f, \mathfrak{Z}) & \text{für } \lambda \geq 0 \\ \lambda O(f, \mathfrak{Z}) & \text{für } \lambda < 0 \end{cases} \quad \text{und} \quad O(\lambda f, \mathfrak{Z}) = \begin{cases} \lambda O(f, \mathfrak{Z}) & \text{für } \lambda \geq 0 \\ \lambda U(f, \mathfrak{Z}) & \text{für } \lambda < 0\,. \end{cases}$$

Unter Berücksichtigung von $\int_a^b f(x)dx = O_a^b(f) = U_a^b(f)$ und unter Verwendung von Lemma 1.17 im Fall $\lambda \geq 0$

$$U_a^b(\lambda f) = \sup\{U(\lambda f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \sup\{\lambda U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \lambda \sup\{U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} =$$

$$= \lambda U_a^b(f) = \lambda \int_a^b f(x)dx = \lambda O_a^b(f) = \lambda \inf\{O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} =$$

$$= \inf\{\lambda O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \inf\{O(\lambda f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = O_a^b(\lambda f)$$

sowie im Fall $\lambda < 0$:

$$U_a^b(\lambda f) = \sup\{U(\lambda f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \sup\{\lambda O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \lambda \inf\{O(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} =$$
$$= \lambda O_a^b(f) = \lambda \int_a^b f(x)dx = \lambda U_a^b(f) = \lambda \sup\{U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} =$$
$$= \inf\{\lambda U(f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = \inf\{O(\lambda f, \mathfrak{Z})\,;\, \mathfrak{Z} \in \mathfrak{Z}(I)\} = O_a^b(\lambda f)\,.$$

In beiden Fällen erhalten wir also die Riemann–Integrierbarkeit von λf und (6.5).

(b) Für eine Zerlegung $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}(I)$ und beschränkte Funktionen $h : I \to \mathbb{R}$ setzen wir:

$$m_j(h) := \inf h([t_{j-1}, t_j]) \quad \text{und} \quad M_j(h) := \sup h([t_{j-1}, t_j]) \qquad (j = 1, \ldots, n).$$

Mit dieser Bezeichnung gilt:

$$m_j(f+g) \geq m_j(f) + m_j(g) \quad \text{sowie} \quad M_j(f+g) \leq M_j(f) + M_j(g)$$

und daher

$$\begin{aligned} U(f+g, \mathfrak{Z}) &= \sum_{j=1}^n m_j(f+g)(t_j - t_{j-1}) \\ &\geq \sum_{j=1}^n m_j(f)(t_j - t_{j-1}) + \sum_{j=1}^n m_j(g)(t_j - t_{j-1}) = U(f, \mathfrak{Z}) + U(g, \mathfrak{Z}) \end{aligned}$$

und

$$\begin{aligned} O(f+g, \mathfrak{Z}) &= \sum_{j=1}^n M_j(f+g)(t_j - t_{j-1}) \\ &\leq \sum_{j=1}^n M_j(f)(t_j - t_{j-1}) + \sum_{j=1}^n M_j(g)(t_j - t_{j-1}) = O(f, \mathfrak{Z}) + O(g, \mathfrak{Z})\,. \end{aligned}$$

Für alle $\mathfrak{Z}_1, \mathfrak{Z}_2 \in \mathfrak{Z}(I)$ folgt mit $\mathfrak{Z} := \mathfrak{Z}_1 \cup \mathfrak{Z}_2$:

$$U_a^b(f+g) \geq U(f+g, \mathfrak{Z}) \geq U(f, \mathfrak{Z}) + U(g, \mathfrak{Z}) \geq U(f, \mathfrak{Z}_1) + U(g, \mathfrak{Z}_2)$$

und daher (durch Übergang zum Supremum bezüglich $\mathfrak{Z}_1$ und bezüglich $\mathfrak{Z}_2$):

$$U_a^b(f+g) \geq U_a^b(f) + U_a^b(g) = \int_a^b f(x)dx + \int_a^b g(x)dx\,.$$

Analog erhält man

$$O_a^b(f+g) \leq O_a^b(f) + O_a^b(g) = \int_a^b f(x)dx + \int_a^b g(x)dx\,.$$

Insgesamt folgt

$$\int_a^b f(x)dx + \int_a^b g(x)dx \leq U_a^b(f+g) \leq O_a^b(f+g) \leq \int_a^b f(x)dx + \int_a^b g(x)dx\,.$$

Nach Definition der Riemann–Integrierbarkeit und des Riemann–Integrals folgt die Behauptung.

(c) Ist $f \leq g$ auf I, so folgt für alle $\mathfrak{Z} \in \mathfrak{Z}(I)$:

$$\int_a^b f(x)dx = O_a^b(f) \leq O(f, \mathfrak{Z}) \leq O(g, \mathfrak{Z})\,.$$

Also folgt auch

$$\int_a^b f(x)dx \le O_a^b(g) = \int_a^b g(x)dx\,.$$

(d) Da f beschränkt ist sind auch die Funktionen f_+ und f_- auf I beschränkt. Um ihre Riemann–Integrierbarkeit zu beweisen, zeigen wir, daß die Bedingung (b) in Satz 6.4 für diese Funktionen erfüllt ist. Sei also $\varepsilon > 0$ beliebig. Wegen $f \in \mathcal{R}(I)$ gibt es nach Satz 6.4 eine Zerlegung $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}(I)$ mit $0 \le O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon$. Mit den Bezeichnung aus dem Beweis zu (b) gilt für alle $j = 1,\dots,n$ (wegen $f(x) \le f_+(x)$ für alle $x \in I$):

$$m_j(f) \le m_j(f_+) \quad \text{und} \quad M_j(f) \le M_j(f_+). \tag{6.6}$$

Aus der Definition von f_+ folgt weiter: Ist $M_j(f) < M_j(f_+)$, so ist schon $f(x) < 0$ und $f_+(x) = 0$ für alle $x \in [t_{j-1}, t_j]$. Es folgt in diesem Fall

$$M_j(f_+) - m_j(f_+) = 0 \le M_j(f) - m_j(f)\,.$$

Ist $M_j(f) = M_j(f_+)$, so erhalten wir wegen (6.6):

$$M_j(f_+) - m_j(f_+) = M_j(f) - m_j(f_+) \le M_j(f) - m_j(f)\,.$$

Für alle $j = 1,\dots,n$ gilt also $M_j(f_+) - m_j(f_+) \le M_j(f) - m_j(f)$ und daher

$$\begin{aligned} 0 \le O(f_+,\mathfrak{Z}) - U(f_+,\mathfrak{Z}) &= \sum_{j=1}^n (M_j(f_+) - m_j(f_+))(t_j - t_{j-1}) \\ &\le \sum_{j=1}^n (M_j(f) - m_j(f))(t_j - t_{j-1}) \\ &\le O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon\,. \end{aligned}$$

Nach Satz 6.4 sind also f_+ und damit nach (a) und (b) auch die Funktionen $f_- = f_+ - f$ und $|f| = f_+ + f_-$ Riemann–integrierbar auf I. Wegen $-f \le |f|$ und $f \le |f|$ auf I folgt nach (a) und (c)

$$-\int_a^b f(x)dx = \int_a^b -f(x)dx \le \int_a^b |f(x)|dx \quad \text{und} \quad \int_a^b f(x)dx \le \int_a^b |f(x)|dx\,.$$

Wegen $|f| \le \|f\|_I$ auf I folgt also

$$\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx \le \int_a^b \|f\|_I dx = (b-a)\|f\|_I\,.$$

(e) Ist $f \equiv 0$ auf I, so ist dies offensichtlich. Sei nun $f \not\equiv 0$ und somit $\|f\|_I > 0$. Wegen (d) ist $|f| \in \mathcal{R}(I)$. Nach Satz 6.4 gibt es also zu beliebig vorgegebenem $\varepsilon > 0$ eine Zerlegung $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}(I)$, so daß mit den Bezeichnungen aus (b) gilt:

$$0 \le O(|f|,\mathfrak{Z}) - U(|f|,\mathfrak{Z}) < \frac{\varepsilon}{p}\|f\|_I^{1-p}\,.$$

Wegen der Monotonie der Funktion $h : (0,\infty) \to (0,\infty)$, $h : t \mapsto t^p$ folgt

$$\begin{aligned} 0 \le O(|f|^p,\mathfrak{Z}) - U(|f|^p,\mathfrak{Z}) &= \sum_{j=1}^n (M_j(|f|^p) - m_j(|f|^p))(t_j - t_{j-1}) = \\ &= \sum_{j=1}^n (M_j(|f|)^p - m_j(|f|)^p)(t_j - t_{j-1}) \end{aligned}$$

Unter Verwendung des Mittelwertsatzes der Differentialrechnung (für h) und der Monotonie der Funktion $t \mapsto t^{p-1}$ auf I folgt also

$$\begin{aligned} 0 \le O(|f|^p, \mathfrak{Z}) - U(|f|^p, \mathfrak{Z}) &\le \sum_{j=1}^n pM_j(|f|)^{p-1}(M_j(|f|) - m_j(|f|))(t_j - t_{j-1}) \le \\ &\le p\|f\|_I^{p-1}(O(|f|, \mathfrak{Z}) - U(|f|, \mathfrak{Z})) < \varepsilon\,. \end{aligned}$$

Nach Satz 6.4 ist $|f|^p$ also auf I Riemann–integrierbar.

(f) Nach (a), (b) und (e) gilt für alle $u, v \in \mathcal{R}(I)$ mit $u \ge 0, v \ge 0$ auf I auch

$$uv = \frac{1}{2}((u+v)^2 - u^2 - v^2) \in \mathcal{R}(I)\,.$$

Zusammen mit (d) erhalten wir also

$$fg = (f_+ - f_-)(g_+ - g_-) = f_+g_+ - f_+g_- - f_-g_+ + f_-g_- \in \mathcal{R}(I)\,.$$

□

Es folgen einige Rechenregeln für den vektorwertigen Fall:

6.8. Lemma. *Seien $f = (f_1, \dots, f_N), g = (g_1, \dots, g_N) : I = [a,b] \to \mathbb{K}^N$ zwei auf I Riemann–integrierbare $\mathbb{K}^N$–wertige Funktionen.*

(a) *Für alle $\alpha, \beta \in \mathbb{K}$ ist auch $\alpha f + \beta g \in \mathcal{R}(I, \mathbb{K}^N)$ und*

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx\,.$$

$\mathcal{R}(I, \mathbb{K}^N)$ ist also ein Untervektorraum des $\mathbb{K}$–Vektorraums $\mathcal{F}(I, \mathbb{K}^N)$ aller $\mathbb{K}^N$–wertigen Funktionen auf I und das Riemann–Integral

$$\int_a^b \cdot\, dx : f \mapsto \int_a^b f(x)dx$$

definiert eine $\mathbb{K}$–lineare Abbildung von $\mathcal{R}(I, \mathbb{K}^N)$ nach $\mathbb{K}^N$.

(b) *Die Funktion*

$$|f| : x \mapsto |f(x)| = \left(\sum_{j=1}^N |f_j(x)|^2\right)^{1/2}$$

ist auf I Riemann–integrierbar und es gilt

$$\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx \le (b-a)\|f\|_I$$

mit $\|f\|_I := \sup_{x \in I} |f(x)| := \sup\{|f(x)|\,;\, x \in I\}$.

Beweis. (a) folgt leicht mit Hilfe der Rechenregeln aus Lemma 6.7 und der Definition des Riemann–Integrals für komplex– und vektorwertige Funktionen.

(b) Ohne Beschränkung der Allgemeinheit sei $\mathbb{K} = \mathbb{R}$ (da wir $\mathbb{C}$ mit $\mathbb{R}^2$ und $\mathbb{C}^N$ mit $\mathbb{R}^{2N}$ identifizieren können).

Wir zeigen zunächst, daß $|f|$ auf I Riemann–integrierbar ist. Da $f = (f_1, \dots, f_N) : I \to \mathbb{R}$ auf I Riemann–integrierbar ist, folgt $f_j \in \mathcal{R}(I)$ und somit (nach Lemma 6.7) auch $|f_j| \in \mathcal{R}(I)$ für $j = 1, \dots, N$. Sei nun $\varepsilon > 0$ beliebig. Wegen $|f_j| \in \mathcal{R}(I)$ gibt es nach Satz 6.4 eine Zerlegung $\mathfrak{Z}_j \in \mathfrak{Z}(I)$ mit

$$0 \le O(|f_j|, \mathfrak{Z}_j) - U(|f_j|, \mathfrak{Z}_j) < \frac{\varepsilon}{N} \qquad \text{für } j = 1, \dots, N.$$

Da $\mathfrak{Z} := \mathfrak{Z}_1 \cup \dots \cup \mathfrak{Z}_N$ eine Verfeinerung von $\mathfrak{Z}_j$ ist folgt auch

$$0 \le O(|f_j|, \mathfrak{Z}) - U(|f_j|, \mathfrak{Z}) < \frac{\varepsilon}{N} \qquad \text{für } j = 1, \dots, N.$$

$\mathfrak{Z}$ hat eine Darstellung der Form $\mathfrak{Z} = \{a = t_0 < t_1 < \dots < t_n = b\}$. Mit den Bezeichnungen

$$m_k(|f|) := \inf\{|f(x)|\,;\, t_{k-1} \le x \le t_k\},\ m_k(|f_j|) := \inf\{|f_j(x)|\,;\, t_{k-1} \le x \le t_k\},$$
$$M_k(|f|) := \sup\{|f(x)|\,;\, t_{k-1} \le x \le t_k\},\ M_k(|f_j|) := \sup\{|f_j(x)|\,;\, t_{k-1} \le x \le t_k\}$$

gilt

$$m_k(|f|) \ge \left(\sum_{j=1}^N m_k(|f_j|)^2\right)^{1/2}, \quad M_k(|f|) \le \left(\sum_{j=1}^N M_k(|f_j|)^2\right)^{1/2}$$

und daher (unter Verwendung von Teil (iv) von Satz 2.8 (b)):

$$\begin{aligned} 0 \le M_k(|f|) - m_k(|f|) &\le \left(\sum_{j=1}^N M_k(|f_j|)^2\right)^{1/2} - \left(\sum_{j=1}^N m_k(|f_j|)^2\right)^{1/2} \\ &\le \left(\sum_{j=1}^N (M_k(|f_j|) - m_k(|f_j|))^2\right)^{1/2} \\ &\le \sum_{j=1}^N (M_k(|f_j|) - m_k(|f_j|))\,. \end{aligned}$$

Hiermit folgt:

$$\begin{aligned} 0 \le O(|f|, \mathfrak{Z}) - U(|f|, \mathfrak{Z}) &= \sum_{k=1}^n (M_k(|f|) - m_k(|f|))(t_k - t_{k-1}) \\ &\le \sum_{k=1}^n \sum_{j=1}^N (M_k(|f_j|) - m_k(|f_j|))(t_k - t_{k-1}) \\ &\le \sum_{j=1}^N (O(|f_j|, \mathfrak{Z}) - U(|f_j|, \mathfrak{Z})) < \varepsilon\,. \end{aligned}$$

Nach Satz 6.4 ist $|f|$ also auf I Riemann–integrierbar.

Beim Nachweis der Ungleichungen in (b) genügt es wegen Lemma 6.7 (d), die erste zu zeigen. Sei $\varepsilon > 0$ beliebig vorgegeben. Wegen $|f|, f_j \in \mathcal{R}(I)$ gibt es $\mathfrak{Z}_0, \mathfrak{Z}_1, \dots, \mathfrak{Z}_N \in \mathfrak{Z}(I)$ mit

$$0 \le O(|f|, \mathfrak{Z}_0) - U(|f|, \mathfrak{Z}_0) < \frac{\varepsilon}{2} \text{ und } 0 \le O(f_j, \mathfrak{Z}_j) - U(f_j, \mathfrak{Z}_j) < \frac{\varepsilon}{2N} \text{ für } j = 1, \dots, N.$$

Für die gemeinsame Verfeinerung $\mathfrak{Z} := \bigcup_{j=0}^N \mathfrak{Z}_j$ gilt dann ebenfalls

$$0 \le O(|f|, \mathfrak{Z}) - U(|f|, \mathfrak{Z}) < \frac{\varepsilon}{2} \text{ und } 0 \le O(f_j, \mathfrak{Z}) - U(f_j, \mathfrak{Z}) < \frac{\varepsilon}{2\sqrt{N}} \text{ für } j = 1, \dots, N.$$

$\mathfrak{Z}$ hat eine Darstellung der Form $\mathfrak{Z} = \{a = t_0 < t_1 < \dots < t_n = b\}$. Wegen

$$\int_a^b f_j(x)dx \in [U(f_j, \mathfrak{Z}), O(f_j, \mathfrak{Z})]\,, \qquad S_j := \sum_{k=1}^n f_j(t_k)(t_k - t_{k-1}) \in [U(f_j, \mathfrak{Z}), O(f_j, \mathfrak{Z})]$$

folgt

$$\left|\int_a^b f_j(x)dx - S_j\right| < \frac{\varepsilon}{2\sqrt{N}} \qquad \text{für } j = 1, \dots, N.$$

Mit $S := (S_1, \dots, S_N) = \sum_{k=1}^n f(t_k)(t_k - t_{k-1})$ ist also

$$\Big|\int_a^b f(x)dx - S\Big| = \Big(\sum_{j=1}^N \Big(\int_a^b f_j(x)dx - S_j\Big)^2\Big)^{1/2} < \frac{\varepsilon}{2}$$

Unter Verwendung der Dreiecksungleichung für den euklidischen Betrag erhalten wir:

$$\begin{aligned}\Big|\int_a^b f(x)dx\Big| &\le \Big|\int_a^b f(x)dx - S\Big| + |S| < \Big|\sum_{k=1}^n f(t_k)(t_k - t_{k-1})\Big| + \frac{\varepsilon}{2} \\ &\le \sum_{k=1}^n |f(t_k)|(t_k - t_{k-1}) + \frac{\varepsilon}{2} \le O(|f|, \mathfrak{Z}) + \frac{\varepsilon}{2} < \int_a^b |f(x)|dx + \varepsilon\,.\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt die behauptete Ungleichung. □

6.9. Lemma (Additivität in den Integrationsgrenzen). *Sei $-\infty < a \le b \le c < +\infty$ und sei $f : [a, c] \to \mathbb{R}$ eine auf $[a, c]$ beschränkte Funktion. Wir vereinbaren für alle $y \in [a, c]$: $O_y^y(f) := U_y^y(f) := \int_y^y f(c)dx := 0$. Es gilt:*

(a) $U_a^c(f) = U_a^b(f) + U_b^c(f)$ *und* $O_a^c(f) = O_a^b(f) + O_b^c(f)$.
(b) *Genau dann ist f auf $[a, c]$ Riemann–integrierbar, wenn $f|_{[a,b]} \in \mathcal{R}([a, b])$ und $f|_{[b,c]} \in \mathcal{R}([b, c])$ gilt. In diesem Fall gilt*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx\,.$$

Beweis. (a) Ist $\mathfrak{Z} \in \mathfrak{Z}([a, c])$ beliebig vorgegeben, so gilt $\mathfrak{Z} \cup \{b\} = \mathfrak{Z}_1 \cup \mathfrak{Z}_2$ mit Zerlegungen $\mathfrak{Z}_1 \in \mathfrak{Z}([a, b])$ und $\mathfrak{Z}_2 \in \mathfrak{Z}([b, c])$. Es folgt

$$U(f, \mathfrak{Z}) \le U(f, \mathfrak{Z} \cup \{b\}) = U(f, \mathfrak{Z}_1) \cup U(f, \mathfrak{Z}_2) \le U_a^b(f) + U_b^c(f)\,.$$

Also gilt

$$U_a^c(f) \le U_a^b(f) + U_b^c(f)\,. \tag{6.7}$$

Sind umgekehrt Zerlegungen $\mathfrak{Z}_1 \in \mathfrak{Z}([a, b])$ und $\mathfrak{Z}_2 \in \mathfrak{Z}([b, c])$ beliebig vorgegeben, so ist $\mathfrak{Z} := \mathfrak{Z}_1 \cup \mathfrak{Z}_2 \in \mathfrak{Z}([a, c])$ und daher $U(f, \mathfrak{Z}_1) + U(f, \mathfrak{Z}_2) = U(f, \mathfrak{Z}) \le U_a^c(f)$. Durch Übergang zum Supremum bezüglich $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ folgt also

$$U_a^b(f) + U_b^c(f) \le U_a^c(f)\,.$$

Zusammen mit (6.7) ergibt dies die erste Behauptung in (a). Die zweite erhält man analog oder unter Beachtung von $O_a^c(f) = -U_a^c(-f)$ aus der ersten.

(b) Nach (a) und Lemma 6.2 (d) gilt

$$0 \le O_a^c(f) - U_a^c(f) = \underbrace{O_a^b(f) - U_a^b(f)}_{\ge 0} + \underbrace{O_b^c(f) - U_b^c(f)}_{\ge 0}\,.$$

Genau dann ist also $U_a^c(f) = O_a^c(f)$, wenn $U_a^b(f) = O_a^b(f)$ und $U_b^c(f) = O_b^c(f)$ gilt. Hiermit und mit Teil (a) folgen nun die Behauptungen in (b). □

Durch zweimalige Anwendung dieses Lemmas folgt also insbesondere, daß jede auf einem abgeschlossenen, beschränkten Intervall $I = [a, b]$ Riemann–integrierbare Funktion auch auf jedem abgeschlossenen Teilintervall $[\alpha, \beta] \subseteq [a, b]$ Riemann–integrierbar ist.

6.10. Definition. Sei $f : I \to \mathbb{R}$ auf dem Intervall $I = [a, b]$ Riemann–integrierbar und seien $\alpha, \beta \in I$ mit $\alpha \le \beta$. Wir definieren

$$\int_\beta^\alpha f(x)dx := -\int_\alpha^\beta f(x)dx,$$

wobei die rechte Seite nach Lemma 6.9 definiert ist. Damit ist $\int_\alpha^\beta f(x)dx$ für alle $\alpha, \beta \in I$ erklärt.

2. Der Hauptsatz der Differential– und Integralrechnung

Im folgenden sei I stets ein nicht zu einem Punkt ausgeartetes Intervall in $\mathbb{R}$, $N \in \mathbb{N}$ und $\mathbb{K}$ der Körper der reellen oder komplexen Zahlen.

6.11. Definition. Eine Funktion $F : I \to \mathbb{K}^N$ heißt *Stammfunktion* zu einer Funktion $f : I \to \mathbb{K}^N$, falls F auf I differenzierbar ist und $F'(x) = f(x)$ für alle $x \in I$ gilt.

6.12. Bemerkungen. (a) Sind $F, G : I \to \mathbb{K}^N$ Stammfunktionen zu einer Funktion $f : I \to \mathbb{K}^N$, so ist die Funktion $F - G$ auf I konstant,

(b) Ist $F : I \to \mathbb{K}^N$ Stammfunktion zu einer Funktion $f : I \to \mathbb{K}^N$, so ist für jedes $c \in \mathbb{K}^n$ auch $x \mapsto F(x)+c$ eine Stammfunktion von f. Man schreibt auch $\int f(t)dt = F+c$.

Beweis. (a) folgt aus Folgerung 5.23 und (b) ist unmittelbar klar. □

Der folgende Hauptsatz der Differential– und Integralrechnung zeigt, daß jede auf einem Intervall I stetige Funktion eine Stammfunktion besitzt.

6.13. Satz (Der Hauptsatz der Differential– und Integralrechnung). *Sei $f : I \to \mathbb{R}$ eine auf I stetige Funktion. Dann gilt:*

(a) *Für alle $a, b \in I$ mit $a < b$ ist f auf $[a, b]$ Riemann–integrierbar.*

(b) *Für alle $a \in I$ ist die durch*

$$F(x) := \int_a^x f(t)dt \qquad \text{für alle } x \in I$$

definierte Funktion $F : I \to \mathbb{R}$ eine Stammfunktion zu f auf I.

(c) *Für jede Stammfunktion $G : I \to \mathbb{R}$ und alle $a, b \in I$ gilt:*

$$\int_a^b f(t)dt = G(b) - G(a)\,.$$

Beweis. (a) Da f auf $[a, b]$ stetig ist, ist f auf $[a, b]$ nach Satz 4.38 auch gleichmäßig stetig und und nach Satz 4.17 beschränkt. Sei $\varepsilon > 0$ beliebig vorgegeben. Wegen der gleichmäßigen Stetigkeit von f gibt es ein $\delta > 0$ mit

$$\forall x, y \in I : \quad |x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b-a}\,.$$

Sei nun $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ eine beliebige Zerlegung von $[a, b]$ mit der Feinheit $\delta(\mathfrak{Z}) < \delta$. Nach dem Satz 4.19 von der Annahme des Maximums und Minimums stetiger Funktionen gibt es $x_j, y_j \in [t_{j-1}, t_j]$ mit

$$f(x_j) = m_j(f) := \min f([t_{j-1}, t_j]) \quad \text{und} \quad f(y_j) = M_j(f) := \max f([t_{j-1}, t_j])\,.$$

Wegen $|x_j - y_j| \le t_j - t_{j-1} \le \delta(\mathfrak{Z}) < \delta$ gilt

$$M_j(f) - m_j(f) < \frac{\varepsilon}{b-a} \qquad \text{für } j = 1, \ldots, n\,.$$

Es folgt

$$0 \leq O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) = \sum_{j=1}^{n}(M_j(f) - m_j(f))(t_j - t_{j-1}) < \frac{\varepsilon}{b-a}\sum_{j=1}^{n}(t_j - t_{j-1}) = \varepsilon\,.$$

Nach Satz 6.4 ist f also auf $[a,b]$ Riemann–integrierbar.

(b) Seien $a, x \in I$ beliebig. Wegen (a) existiert $F(y) := \int_a^y f(t)dt$ für alle $y \in I$. Da f in x stetig ist, gibt es zu beliebig vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ mit

$$\forall t \in I \cap U_\delta(x) : |f(t) - f(x)| < \frac{\varepsilon}{2}\,.$$

Für alle $y \in I$ mit $0 < |y - x| < \delta$ folgt mit Lemma 6.9 und Lemma 6.7 (d)

$$\begin{aligned}\left|\frac{F(y)-F(x)}{y-x} - f(x)\right| &= \frac{1}{|y-x|}\left|\int_x^y f(t)dt - f(x)(y-x)\right| = \frac{1}{|y-x|}\left|\int_x^y (f(t)-f(x))dt\right| \\ &\leq \frac{1}{|y-x|}|y-x|\sup\{|f(t)-f(x)|\,;\,|t-x|<\delta\} \leq \frac{\varepsilon}{2} < \varepsilon\,.\end{aligned}$$

Also existiert für alle $x \in I$ der Grenzwert $F'(x) := \lim_{y\to x}\frac{F(y)-F(x)}{y-x}$ und es gilt $F'(x) = f(x)$.

(c) Ist G eine beliebige Stammfunktion zu f auf I, so gibt es nach Bemerkung 6.12 eine Konstante $c \in \mathbb{R}$ mit

$$\forall x \in I : \quad G(x) = F(x) + c = \int_a^x f(t)dt + c\,.$$

Es gilt also nach Lemma 6.9: $G(b) - G(a) = F(b) - F(a) = \int_a^b f(t)dt$. □

Die Aussage (c) kann insbesondere zur Berechnung von Riemann–Integralen über stetige Funktionen verwendet werden, deren Stammfunktionen bekannt sind.

Statt $\int_a^b f(t)dt = G(b) - G(a)$ schreibt man auch $\int_a^b f(t)dt = G(x)|_a^b$.

6.14. Bemerkung. Durch komponentenweises Anwenden des Hauptsatzes der Differential– und Integralrechnung sieht man, daß dieser auch für $\mathbb{R}^N$–wertige, $\mathbb{C}$–wertige (wegen $\mathbb{C} = \mathbb{R}^2$) und $\mathbb{C}^N$–wertige ($\mathbb{C}^N = \mathbb{R}^{2N}$) stetige Funktionen gilt.

6.15. Beispiele. (a) Sei $\alpha \in \mathbb{R}\setminus\{-1\}$. Dann gilt

$$\int_a^b x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1}\Big|_a^b = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1} \quad \text{für alle } a,b \in \mathbb{R}$$

$$\text{mit } a < b \text{ und } \begin{cases}[a,b] \subset \mathbb{R} & \text{falls } \alpha \in \mathbb{N}_0 \\ [a,b] \subset [0,\infty) & \text{falls } 0 < \alpha \in \mathbb{R}\setminus\mathbb{N} \\ 0 \notin [a,b] \subset \mathbb{R} & \text{falls } \alpha \in -\mathbb{N} \\ [a,b] \subset (0,\infty) & \text{falls } \alpha \in (-\infty,0)\setminus\mathbb{Z}\,.\end{cases}$$

(b) Im Fall $\alpha = -1$ haben wir:

$$\int_a^b \frac{1}{x}dx = \begin{cases}\log(x)\big|_a^b = \log(b) - \log(a) = \log\left(\frac{b}{a}\right) & \text{für } 0 < a < b \\ \log(-x)\big|_a^b = \log(-b) - \log(-a) = \log\left(\frac{b}{a}\right) & \text{für } a < b < 0\,.\end{cases}$$

(c) Die Funktionen $e_n : \mathbb{R} \to \mathbb{C}$ mit $e_n(t) := \exp(int)$ für alle $t \in \mathbb{R}$, $n \in \mathbb{Z}$ sind nach Beispiel 5.4 (c) auf $\mathbb{R}$ differenzierbar und es gilt $e_n'(t) = in\exp(int) = ine_n(t)$ für alle

$t \in \mathbb{R}$. Für $n \neq 0$ ist also die Funktion $\frac{-i}{n} e_n$ eine Stammfunktion für e_n auf $\mathbb{R}$. Für alle $n \in \mathbb{Z} \setminus \{0\}$ und alle $c \in \mathbb{R}$ folgt mit Bemerkung 6.14:

$$\begin{aligned} \int_c^{c+2\pi} e_n(t)dt &= \frac{-i}{n} \exp(int)\Big|_c^{c+2\pi} = \frac{-i}{n}(\exp(2\pi in + inc) - \exp(inc)) \\ &= \frac{-i}{n} \exp(inc)(\exp(2\pi in) - 1) = 0\,. \end{aligned} \tag{6.8}$$

Im Fall $n = 0$ haben wir

$$\int_c^{c+2\pi} e_0(t)dt = \int_c^{c+2\pi} 1dt = 2\pi.$$

Wegen $e_n(t) = \cos(nt) + i\sin(nt)$ folgt aus (6.8) und der Definition des Riemann–Integrals im komplexwertigen Fall für alle $n \in \mathbb{Z} \setminus \{0\}$:

$$\begin{aligned} \int_c^{c+2\pi} \cos(nt)dt &= \int_c^{c+2\pi} \operatorname{Re} e_n(t)dt = \operatorname{Re} \int_c^{c+2\pi} e_n(t)dt = 0 \text{ und} \\ \int_c^{c+2\pi} \sin(nt)dt &= \int_c^{c+2\pi} \operatorname{Im} e_n(t)dt = \operatorname{Im} \int_c^{c+2\pi} e_n(t)dt = 0\,. \end{aligned}$$

6.16. Satz (Substitutionsregel). *Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{K}^N$ auf I stetig und $\varphi : [a,b] \to \mathbb{R}$ stetig differenzierbar mit $\varphi([a,b]) \subseteq I$. Dann gilt*

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx\,.$$

Beweis. Nach dem Hauptsatz der Differential– und Integralrechnung besitzt f eine Stammfunktion $F : I \to \mathbb{K}^N$ und nach der Kettenregel gilt für alle $t \in [a,b]$

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Nach dem Hauptsatz der Differential– und Integralrechnung folgt also:

$$\int_a^b f(\varphi(t))\varphi'(t)dt = (F \circ \varphi)(b) - (F \circ \varphi)(a) = \int_{\varphi(a)}^{\varphi(b)} f(x)dx\,.$$

□

6.17. Beispiele. (a) Sei $\varphi : [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion mit $0 \notin \varphi([a,b])$. Dann gilt unter Verwendung der Substitutionsregel 6.16 und Beispiel 6.15 (b):

$$\int_a^b \frac{\varphi'(t)}{\varphi(t)}dt = \int_{\varphi(a)}^{\varphi(b)} \frac{1}{x}dx = \log\left(\frac{\varphi(b)}{\varphi(a)}\right)\,.$$

(b) Als Spezialfall von (a) erhalten wir für $-\pi/2 < a < b < \pi/2$:

$$\int_a^b \tan(t)dt = \int_a^b \frac{\sin(t)}{\cos(t)}dt = -\int_a^b \frac{\cos'(t)}{\cos(t)}dt = \log\left(\frac{\cos(a)}{\cos(b)}\right)\,.$$

(c) Die Fläche des Halbkreises mit Radius 1 ist $\displaystyle\int_{-1}^1 \sqrt{1-x^2}dx = \frac{\pi}{2}$.

Beweis. Es ist einerseits (nach der Substitutionsregel)

$$\int_{-\pi/2}^{\pi/2} \cos(t)^2 dt = \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin(t)^2}\,\sin'(t)dt = \int_{-1}^{1} \sqrt{1-x^2}dx$$

und andererseits

$$\int_{-\pi/2}^{\pi/2} \cos(t)^2 dt = \int_{-\pi/2}^{\pi/2} \left(\frac{\exp(it)+\exp(-it)}{2}\right)^2 dt = \int_{-\pi/2}^{\pi/2} \frac{1}{4}(\exp(2it)+\exp(-2it)+2)dt$$

$$= \frac{1}{8i}(\exp(2it)-\exp(-2it)) + \frac{t}{2}\Big|_{-\pi/2}^{\pi/2} = \frac{1}{8i}(-1+1+1-1) + \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}\,.$$

Damit folgt die Behauptung. □

6.18. Satz (Partielle Integration). *Sei $I=[a,b]$ mit $a<b$ und seien $f,g : I \to \mathbb{K}$ zwei stetig differenzierbare Funktionen auf I. Dann gilt*

$$\int_a^b f(t)g'(t)dt = f(t)g(t)\Big|_a^b - \int_a^b f'(t)g(t)dt\,.$$

Beweis. Die Funktion $t \mapsto h(t) := f(t)g(t)$ ist auf I stetig differenzierbar mit $h'(t) = f(t)g'(t) + f'(t)g(t)$ für alle $t \in I$. Nach dem Hauptsatz der Differential– und Integralrechnung gilt also wegen der Linearität des Integrals (vergl. Lemma 6.7):

$$f(t)g(t)\Big|_a^b = h(t)\Big|_a^b = \int_a^b \big(f(t)g'(t)+f'(t)g(t)\big)dt = \int_a^b f(t)g'(t)dt + \int_a^b f'(t)g(t)dt\,.$$

Hieraus folgt die Behauptung. □

6.19. Beispiel. $\displaystyle\int_0^{2\pi} x\cdot\sin(x)dx = -x\cos(x)\Big|_0^{2\pi} + \int_0^{2\pi} 1\cdot\cos(x)dx = -2\pi,$

da $\displaystyle\int_0^{2\pi}\cos(x)dx = 0$ nach Beispiel 6.15 (c).

Bei Umkehrfunktionen f^{-1} kennt man wegen des Satzes 5.9 über die Differenzierbarkeit der Umkehrfunktion häufig deren Ableitung besser als f^{-1} selbst. Satz 6.18 legt daher folgenden Ansatz nahe:

$$\int_a^b f^{-1}(t)dt = t\cdot f^{-1}(t)\big|_a^b - \int_a^b t\cdot(f^{-1})'(t)dt\,.$$

Hierzu zwei Beispiele:

6.20. Beispiele. (a) Für $0<a<b$ gilt

$$\int_a^b \log(t)dt = t\cdot\log(t)\big|_a^b - \int_a^b t\cdot\frac{1}{t}dt = t(\log(t)-1)\big|_a^b\,.$$

Insbesondere ist $t\mapsto t(\log(t)-1)$ eine Stammfunktion zu log auf $(0,\infty)$.

(b) In den Übungen wurde gezeigt

$$\arctan'(x) = \frac{1}{1+x^2} \qquad \text{für alle } x\in\mathbb{R}\,.$$

Obiger Ansatz führt also zu

$$\int_a^b \arctan(x)dx = x\cdot\arctan(x)\big|_a^b - \int_a^b \frac{x}{1+x^2}dx\,.$$

Wenden wir hierauf die Substitutionsregel mit $\varphi(x) = x^2$ an, so erhalten wir

$$\int_a^b \arctan(x)dx = x \cdot \arctan(x)\Big|_a^b - \frac{1}{2}\int_{a^2}^{b^2} \frac{1}{1+s}ds = x \cdot \arctan(x)\Big|_a^b - \frac{1}{2}\log(1+s)\Big|_{a^2}^{b^2}$$
$$= \Big(x \cdot \arctan(x) - \frac{1}{2}\log(1+x^2)\Big)\Big|_a^b .$$

Insbesondere ist die Funktion

$$x \mapsto x \cdot \arctan(x) - \frac{1}{2}\log(1+x^2)$$

eine Stammfunktion zu arctan auf $\mathbb{R}$.

Wie der Mittelwertsatz der Differentialrechnung, so sind auch die beiden folgenden Mittelwertsätze der Integralrechnung wichtige Hilfsmittel bei Abschätzungen.

6.21. Satz (Erster Mittelwertsatz der Integralrechnung). *Sei $I = [a,b]$ mit $-\infty < a < b < \infty$ und seien $f, g : I \to \mathbb{R}$ zwei stetige Funktionen auf I mit $g \geq 0$ auf I. Dann gibt es ein $x_0 \in I$ mit*

$$\int_a^b f(x)g(x)dx = f(x_0)\int_a^b g(x)dx .$$

Speziell zu der Funktion $g \equiv 1$ gibt es also ein $x_0 \in I$ mit

$$\int_a^b f(x)dx = f(x_0)(b-a) .$$

Beweis. Nach Folgerung 4.20 gibt es Punkte $u, v \in I$ mit $f(I) = [f(u), f(v)]$. Wegen $g \geq 0$ auf I gilt also für alle $x \in I$:

$$f(u)g(x) \leq f(x)g(x) \leq f(v)g(x)$$

und daher nach Lemma 6.7 auch

$$f(u)\int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq f(v)\int_a^b g(x)dx .$$

Wegen der Stetigkeit der Funktion $t \mapsto f(t)\int_a^b g(x)dx$ gibt es also nach dem Zwischenwertsatz 4.12 ein x_0 zwischen u und v mit

$$f(x_0)\int_a^b g(x)dx = \int_a^b f(x)g(x)dx .$$

□

6.22. Satz (Zweiter Mittelwertsatz der Integralrechnung). *Sei $I = [a,b]$ wie vor, sei $f : I \to \mathbb{R}$ stetig und sei $g : I \to \mathbb{R}$ eine monotone stetig differenzierbare Funktion. Dann gibt es ein $x_0 \in I$ mit*

$$\int_a^b f(x)g(x)dx = g(a)\int_a^{x_0} f(x)dx + g(b)\int_{x_0}^b f(x)dx .$$

Beweis. Ohne Beschränkung der Allgemeinheit sei g als monoton wachsend vorausgesetzt (falls nicht, so ist $-g$ monoton wachsend). Nach dem Hauptsatz der Differential– und Integralrechnung besitzt f eine Stammfunktion $F : I \to \mathbb{R}$ auf I. Mittels partieller Integration folgt

$$\int_a^b f(x)g(x)dx = \int_a^b F'(x)g(x)dx = F(x)g(x)\Big|_a^b - \int_a^b F(x)g'(x)dx . \tag{6.9}$$

Nach dem ersten Mittelwertsatz der Integralrechnung gibt es ein $x_0 \in I$ mit

$$\int_a^b F(x)g'(x)dx = F(x_0)\int_a^b g'(x)dx = F(x_0)(g(b)-g(a)) .$$

Setzen wir dies in (6.9) ein, so folgt mit dem Hauptsatz der Differential– und Integralrechnung

$$\begin{aligned} \int_a^b f(x)g(x)dx &= F(b)g(b)-F(a)g(a)-F(x_0)(g(b)-g(a)) \\ &= (F(b)-F(x_0))g(b)+(F(x_0)-F(a))g(a) \\ &= g(a)\int_a^{x_0} f(x)dx + g(b)\int_{x_0}^b f(x)dx \end{aligned}$$

und damit die Behauptung. □

3. Der Satz von Taylor

6.23. Satz (Satz von Taylor[2]). *Sei $I \subseteq \mathbb{R}$ ein Intervall und sei $f \in C^{n+1}(I)$. Für alle $x_0, x \in I$ gilt*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(f,x) \tag{6.10}$$

mit dem Restglied

(6.11)

$$R_n(f,x) = \frac{1}{n!}\int_{x_0}^x (x-t)^n f^{(n+1)}(t)dt = \frac{(x-x_0)^{n+1}}{n!}\int_0^1 (1-s)^n f^{(n+1)}(x_0+s(x-x_0))ds .$$

Beweis. Für $x = x_0$ ist die Aussage trivial. Wir führen den Beweis für den Fall $x > x_0$. Im Fall $x < x_0$ verfährt man analog. Sei also $x > x_0$. Wir setzen für $0 \le k \le n$:

$$r_k(x) := \frac{1}{k!}\int_{x_0}^x (x-t)^k f^{(k+1)}(t)dt .$$

Durch partielle Integration erhalten wir für $1 \le k \le n$:

$$\begin{aligned} r_k(x) &= \frac{f^{(k)}(t)}{k!}(x-t)^k\Big|_{t=x_0}^{t=x} + \frac{1}{k!}\int_{x_0}^x k(x-t)^{k-1}f^{(k)}(t)dt \\ &= -\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + r_{k-1}(x) . \end{aligned}$$

Ferner

$$r_0(x) = \frac{1}{0!}\int_{x_0}^x 1\cdot f'(t)dt = f(x)-f(x_0) .$$

Es folgt

$$r_n(x)-f(x)+f(x_0) = r_n(x)-r_0(x) = \sum_{k=1}^n (r_k(x)-r_{k-1}(x)) = -\sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Einsetzen von $r_n(x)$ und Auflösung nach $f(x)$ ergibt die Behauptung mit der ersten Integralrestglieddarstellung. Die zweite erhält man aus der ersten mit der Substitution $s = \varphi(t) = \frac{t-x_0}{x-x_0}$. □

[2] Brook Taylor (18.8.1685–29.12.1731).

Wir nennen das in (6.10) auftretende Polynom

$$T_n(f,x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

das *n–te Taylorpolynom zu f*

Mit Hilfe des ersten Mittelwertsatzes der Integralrechnung erhalten wir weitere Darstellungen des Restglieds in (6.10):

6.24. Satz. *Unter der Voraussetzung und mit den Bezeichnungen des Satzes von Taylor 6.23 gilt für das Restglied $R_n(f,x)$ in (6.10): Es gibt Punkte Punkte y_0, z_0 zwischen x und x_0 (d.h. mit $x_0 \le y_0 \le x$ und $x_0 \le z_0 \le x$ im Fall $x_0 < x$ und $x \le y_0 \le x_0$ und $x \le z_0 \le x_0$ im Fall $x < x_0$), so daß*

$$R_n(f,x) = \frac{f^{(n+1)}(y_0)}{n!}(x-y_0)^n(x-x_0) \qquad (\textit{Cauchy–Darstellung des Restglieds}) \tag{6.12}$$

und

$$R_n(f,x) = \frac{f^{(n+1)}(z_0)}{(n+1)!}(x-x_0)^{n+1} \qquad (\textit{Lagrange–Darstellung des Restglieds})\,. \tag{6.13}$$

Beweis. Mit Hilfe des ersten Mittelwertsatzes der Integralrechnung folgt (6.12) unmittelbar aus der ersten Integraldarstellung des Restglieds in (6.11). Wenden wir den ersten Mittelwertsatzes der Integralrechnung auf die zweite Integraldarstellung in (6.11) an (unter Beachtung von $g(s) := (1-s)^n \ge 0$ auf $[0,1]$), so finden wir ein $s_0 \in [0,1]$ mit

$$\begin{aligned} R_n(f,x) =& \frac{(x-x_0)^{n+1}}{n!}\int_0^1 (1-s)^n f^{(n+1)}(x_0+s(x-x_0))ds \\ =& \frac{(x-x_0)^{n+1}}{n!} f^{(n+1)}(x_0+s_0(x-x_0))\int_0^1 (1-s)^n ds = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(z_0) \end{aligned}$$

mit $z_0 := x_0 + s_0(x-x_0)$. Damit ist auch (6.13) gezeigt. □

6.25. Beispiele. (a) Wir betrachten die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ mit dem Entwicklungspunkt $x_0 := 0$. Wegen $\exp' \equiv \exp$ ist exp beliebig oft differenzierbar und es gilt $\exp^{(n)}(x) = \exp(x)$ für alle $n \in \mathbb{N}_0, x \in \mathbb{R}$. Wir wenden den Satz von Taylor 6.23 mit der Lagrange–Darstellung des Fehlerrestglieds an und erhalten für alle $0 \ne x \in \mathbb{R}$:

$$\begin{aligned} \exp(x) &= \sum_{j=0}^{n} \frac{\exp(0)}{j!}(x-0)^j + R_n(\exp,x) \\ &= \sum_{j=0}^{n} \frac{x^j}{j!} + \frac{x^{n+1}}{(n+1)!}\exp(z_0) \quad \text{für ein } z_0 \text{ zwischen } x \text{ und } x_0. \end{aligned}$$

Wegen $\sum_{j=0}^{n} \frac{x^j}{j!} \to \exp(x)$ für $n \to \infty$ und alle $x \in \mathbb{R}$ folgt $R_n(\exp,x) \to 0$ für $n \to \infty$. Für $0 < |x| \le 1$ und $n = 10$ folgt

$$0 < \left| \exp(x) - \sum_{j=0}^{10} \frac{x^j}{j!} \right| = |R_{10}(\exp,x)| = \frac{|x|^{11}}{11!}\exp(z_0)$$

mit einem z_0 zwischen 0 und x. Insbesondere ist $0 < \exp(z_0) \le e$ und wir erhalten für alle $x \in [-1,1]$:

$$0 < \left| \exp(x) - \sum_{j=0}^{10} \frac{x^j}{j!} \right| \le \frac{e}{11!} < 6{,}81 \cdot 10^{-8}.$$

(b) Wir betrachten nun die Funktion $f : (-1, \infty) \to \mathbb{R}$ mit $f(x) := \log(1+x)$ für alle $x > -1$ und den Entwicklungspunkt $x_0 = 0$. Durch vollständige Induktion zeigt man: f ist beliebig oft differenzierbar und es gilt für alle $n \in \mathbb{N}$:

$$f^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{(1+x)^n}, \quad \text{insbesondere also } f^{(n)}(0) = (-1)^{n-1}(n-1)!\,.$$

Wir erhalten

$$\begin{aligned}\log(1+x) &= f(0) + \sum_{j=1}^{n}(-1)^{j-1}\frac{(j-1)!}{j!}x^j + R_n(f,x)\\ &= \sum_{j=1}^{n}(-1)^{j+1}\frac{x^j}{j} + R_n(f,x)\,.\end{aligned}$$

Wir untersuchen nun, für welche $x \in \mathbb{R}$ die Reihe $\sum_{j=1}^{\infty}(-1)^{j+1}\frac{x^j}{j}$ gegen $f(x) = \log(1+x)$ konvergiert.

1. Fall: Für $|x| > 1$ ist die Reihe $\sum_{j=1}^{\infty}(-1)^{j+1}\frac{x^j}{j}$ divergent, da $|x|^n/n \to \infty$ für $n \to \infty$.

2. Fall: Für $0 < x \le 1$ betrachten wir das Fehlerrestglied in der Lagrange–Darstellung:

$$R_n(f,x) = \frac{(-1)^n n!(x-0)^{n+1}}{(n+1)!(1+z_0)^{n+1}} \quad \text{für ein } z_0 \in (0,x)\,.$$

Es folgt

$$|R_n(f,x)| = \frac{x^{n+1}}{(n+1)(1+z_0)^{n+1}} < \frac{x^{n+1}}{(n+1)} \le \frac{1}{n+1} \to 0 \quad \text{für } n \to \infty\,.$$

Also gilt für $0 < x \le 1$ (und natürlich auch für $x = 0$):

$$\log(1+x) = \sum_{j=1}^{\infty}(-1)^{j+1}\frac{x^j}{j}\,.$$

Insbesondere haben wir für $x = 1$ gezeigt:

$$\log 2 = \sum_{j=1}^{\infty}(-1)^{j+1}\frac{1}{j}\,.$$

3. Fall: Für $-1 < x < 0$ betrachten wir das Fehlerrestglied in der Cauchy–Darstellung:

$$R_n(f,x) = \frac{(-1)^n n!(x-y_0)^n(x-0)}{(1+y_0)^{n+1}n!} \quad \text{für ein } y_0 \in [x,0]\,.$$

Wir erhalten für $n \to \infty$:

$$|R_n(f,x)| = \left|\frac{x-y_0}{1+y_0}\right|^n \cdot \frac{|x|}{1+y_0} \to 0\,,$$

denn wegen $-1 < x \le y_0 \le 0$ gilt:

$$\left|\frac{x-y_0}{1+y_0}\right| = \frac{y_0-x}{1+y_0} = -x + \underbrace{\frac{y_0(1+x)}{1+y_0}}_{\le 0} \le -x = |x| < 1\,.$$

Unter Berücksichtigung von Fall 2 sehen wir also, daß

$$\log(1+x) = \sum_{j=1}^{\infty}(-1)^{j+1}\frac{x^j}{j} \quad \text{für alle } x \in (-1,1]\,.$$

4. Fall: Für $x = -1$ ist $T_n(f,-1) = \sum_{j=1}^{n}\frac{1}{j} \to \infty$ für $n \to \infty$ nach Satz 3.9.

6.26. DEFINITION. Sei $-\infty \le a < x_0 < b \le \infty$ und sei $f : (a,b) \to \mathbb{R}$ eine auf (a,b) beliebig oft differenzierbare Funktion. Dann heißt die Reihe

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

die *Taylorreihe von* f *zum Entwicklungspunkt* x_0.

Wir haben in 6.25 also gezeigt, daß die Funktionen $x \mapsto \exp(x)$ auf $\mathbb{R}$ und $x \mapsto \log(1+x)$ auf $(-1,1]$ durch ihre Taylorreihen dargestellt werden.

6.27. WARNUNG. (a) Nicht alle beliebig oft differenzierbaren Funktionen werden auch in einer Umgebung eines Entwicklungspunktes durch ihre Taylorreihe dargestellt. So ist, wie in den Übungen gezeigt wird, die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) := \begin{cases} \exp(-x^{-2}) & \text{für } 0 \neq x \in \mathbb{R} \\ 0 & \text{für } x = 0 \end{cases}$$

auf $\mathbb{R}$ beliebig oft differenzierbar und erfüllt $f^{(n)} = 0$ für alle $n \in \mathbb{N}_0$. Es folgt $T_n(f,x) = 0$ für alle $n \in \mathbb{N}_0, x \in \mathbb{R}$, aber $f(x) \neq 0$ für alle $x \neq 0$ aus $\mathbb{R}$.

(b) Ohne Beweis vermerken wir, daß es beliebig oft differenzierbare Funktionen $f : (a,b) \to \mathbb{R}$ gibt, für die die Taylorreihe $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$ in keinem vom Entwicklungspunkt $x_0 \in (a,b)$ verschiedenen $x \in (a,b)$ konvergiert.

DEFINITION (Landau–Symbole[3] o und O). Sei $I \subseteq \mathbb{R}$ ein nicht zu einem Punkt ausgeartetes Intervall und sei $x_0 \in I$. Sind $f : I \setminus \{x_0\} \to \mathbb{R}^N$ und $h : I \setminus \{x_0\} \to (0,\infty)$ zwei Funktionen so schreiben wir

$$f(x) = \mathrm{O}(h(x)) \text{ für } x \to x_0 \;:\Longleftrightarrow\; \exists \delta > 0, C > 0 \;\forall x \in I \cap [x_0-\delta, x_0+\delta] \setminus \{x_0\} : \frac{|f(x)|}{h(x)} \le C$$

und

$$f(x) = \mathrm{o}(h(x)) \text{ für } x \to x_0 \;:\Longleftrightarrow\; \lim_{x \to x_0} \frac{|f(x)|}{h(x)} = 0\,.$$

Entsprechende Bezeichnungen führt man auch für das Verhalten bei $x \to \infty$ oder $x \to -\infty$ ein. Ist auch $g : I \setminus \{x_0\} \to \mathbb{R}^N$ eine Funktion, so schreiben wir

$$f(x) = g(x) + \mathrm{O}(h(x)) \text{ für } x \to x_0 \;:\Longleftrightarrow\; (f-g)(x) = \mathrm{O}(h(x))$$

und

$$f(x) = g(x) + \mathrm{o}(h(x)) \text{ für } x \to x_0 \;:\Longleftrightarrow\; (f-g)(x) = \mathrm{o}(h(x))$$

BEISPIEL. $\dfrac{1}{1-x} = 1 + x + \mathrm{O}(x^2)$ für $x \to 0$.

BEWEIS. Für $|x| < 1$ gilt:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 \sum_{n=0}^{\infty} x^n = 1 + x + \frac{x^2}{1-x}\,.$$

---

[3] von EDMUND GEORG HERMANN LANDAU (14.2.1877–19.2.1938) im Jahr 1909 eingeführt. Das Symbol O tritt auch schon vorher im Jahr 1894 bei PAUL GUSTAV HEINRICH BACHMANN (22.6.1837–31.3.1920) auf.

Hiermit folgt

$$\frac{\left|\dfrac{1}{1-x}-(1+x)\right|}{x^2} = \left|\frac{1}{1-x}\right| \leq C := 2 \quad \text{für } 0 < |x| \leq \delta := \frac{1}{2}$$

und damit die Behauptung. □

Wir zeigen nun eine Variante des Taylorschen Satzes für n mal stetig differenzierbare Funktionen:

6.28. Satz. *Sei* $-\infty \leq a < x_0 < b \leq \infty$, $n, N \in \mathbb{N}$ *und sei* $f \in C^n((a,b), \mathbb{R}^N)$. *Dann gilt*

$$\forall x \in (a,b) \setminus \{x_0\} : \quad f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x-x_0)^j + \eta(x)(x-x_0)^n$$

mit einer Funktion $\eta : (a,b) \to \mathbb{R}^N$ *für die* $\lim_{x\to x_0} \eta(x) = 0$ *ist. Es gilt also*

$$f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x-x_0)^j + \mathrm{o}(|x-x_0|^n) \quad \textit{für } x \to x_0\,.$$

Beweis. Wir beweisen die Aussage zunächst für den Spezialfall $N = 1$. Sei $x \in (a,b) \setminus \{x_0\}$ beliebig. Nach dem Taylorschen Satz 6.23 (angewendet auf $n-1$ statt n) gilt mit der Lagrange–Darstellung des Fehlerrestglieds: Es gibt ein $z_0(x)$ zwischen x_0 und x mit

$$\begin{aligned} f(x) &= \sum_{j=0}^{n-1} \frac{f^{(j)}(x_0)}{j!}(x-x_0)^j + \frac{f^{(n)}(z_0(x))}{n!}(x-x_0)^n \\ &= \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x-x_0)^j + \frac{(f^{(n)}(z_0(x)) - f^{(n)}(x_0))}{n!}(x-x_0)^n\,. \end{aligned}$$

Wegen $z_0(x) \to x_0$ für $x \to x_0$ folgt aus der Stetigkeit der n–ten Ableitung

$$f^{(n)}(z_0(x)) - f^{(n)}(x_0) \to f^{(n)}(x_0) - f^{(n)}(x_0) = 0 \qquad \text{für } x \to x_0$$

und damit

$$\eta(x) := \frac{f^{(n)}(z_0(x)) - f^{(n)}(x_0)}{n!} \to 0 \quad \text{für } x \to x_0\,.$$

Die Aussage des Satzes für den Fall $N > 1$ erhält man durch komponentenweises Anwenden der bereits gezeigten Aussage für den skalaren Fall. □

4. Integration rationaler Funktionen

Wir wollen uns nun überlegen, wie man Stammfunktionen und Integrale beliebiger rationaler Funktionen berechnen kann.

6.29. Integration rationaler Funktionen. Gegeben sei eine rationale Funktion $f = p/q$, wobei p, q teilerfremde Polynome mit Koeffizienten in $\mathbb{R}$ seien und q nicht identisch verschwinde (d.h. es sei $\deg(q) \geq 0$). $D_f := \{x \in \mathbb{R}\,;\, q(x) \neq 0\}$ bezeichne den natürlichen Definitionsbereich von f.

1. Reduktionsschritt: Durch die aus der Schule oder der Linearen Algebra bekannte *Polynomdivision mit Rest* findet man reelle Polynome p_0, h mit $\deg(h) < \deg(q)$ und

$$f(x) = p_0(x) + \frac{h(x)}{q(x)} \qquad \text{für alle } x \in D_f\,.$$

Zu p_0 können wir leicht eine Stammfunktion angeben. Ist etwa

$$p_0(x) = \sum_{j=0}^{d_0} u_{0,j} x^j \qquad \text{mit } d_0 = \deg(p_0) \text{ und } u_{0,0}, \dots, u_{0,d_0} \in \mathbb{R}\,,$$

so ist durch

$$\int p_0(x)dx = \sum_{j=0}^{d_0} u_{0,j} \frac{x^{j+1}}{j+1} + C$$

eine Stammfunktion zu p_0 gegeben.

2. Reduktionsschritt: Wir müssen noch für die rationale Funktion h/q eine Stammfunktion finden. Hierzu führen wir zunächst eine *Partialbruchzerlegung* durch, um zu einfacheren Ausdrücken zu gelangen. Sei etwa

$$q(x) = \sum_{j=0}^{d_q} v_j x^j \qquad \text{mit } v_{d_q} \neq 0, \text{ d.h. } d_q = \deg(q)\,,$$

Indem wir notfalls Zähler und Nenner von h/q durch v_{d_q} dividieren, können wir erreichen, daß $v_{d_q} = 1$ ist. Nach dem *Fundamentalsatz der Algebra* gilt dann

$$q(x) = \left(\prod_{j=1}^{r} (x - x_j)^{m_j} \right) \cdot \left(\prod_{k=1}^{s} (x - z_k)^{n_k} \right) .$$

Hierbei seien $x_1, \dots, x_r$ die reellen Nullstellen (mit den Vielfachheiten $m_1, \dots, m_r$) und $z_1 = \alpha_1 + i\beta_1, \dots, z_s = \alpha_s + i\beta_s$ die komplexen Nullstellen mit den Vielfachheiten $n_1, \dots, n_s$ und von Null verschiedenen Imaginärteilen $\beta_1, \dots, \beta_s$. Es ist dann

$$\sum_{j=1}^{r} m_j + \sum_{k=1}^{s} n_k = d_q = \deg(q)\,.$$

Ist $z = \alpha + i\beta$ eine komplexe Nullstelle von q mit Imaginärteil $\beta \neq 0$, so gilt auch

$$0 = \overline{q(z)} = \overline{\sum_{j=0}^{d_q} v_j (\alpha + i\beta)^j} = \sum_{j=0}^{d_q} v_j (\alpha - i\beta)^j\,.$$

Mit z ist somit auch $\overline{z}$ Nullstelle von q. s ist also eine gerade Zahl, d.h. $s = 2t$ für ein $t \in \mathbb{N}_0$, und wir erhalten

$$\begin{aligned} q(x) &= \left(\prod_{j=1}^{r} (x - x_j)^{m_j} \right) \cdot \left(\prod_{k=1}^{t} [(x - z_k)(x - \overline{z_k})]^{n_k} \right) \\ &= \left(\prod_{j=1}^{r} (x - x_j)^{m_j} \right) \cdot \left(\prod_{k=1}^{t} [(x - \alpha_k)^2 + \beta_k^2]^{n_k} \right) \end{aligned} \tag{6.14}$$

Wir zerlegen nun h/q in der Form

$$\frac{h(x)}{q(x)} = \sum_{j=1}^{r} \sum_{\mu=1}^{m_j} \frac{a_{j,\mu}}{(x - x_j)^\mu} + \sum_{k=1}^{t} \sum_{\nu=1}^{n_k} \frac{b_{k,\nu} x + c_{k,\nu}}{[(x - \alpha_k)^2 + \beta_k^2]^\nu}\,. \tag{6.15}$$

Hierbei lassen sich die Koeffizienten $a_{j,\mu}$, $b_{k,\nu}$, $c_{k,\nu}$ berechnen, indem man (6.15) mit q in der Form (6.14) durchmultipliziert, anschließend kürzt und dann Koeffizientenvergleich durchführt.

3. Reduktionsschritt: *Berechnung der Stammfunktionen zu den Summanden aus* (6.15). Folgende Typen von Stammfunktionen sind zu berechnen:

Typ 1: $\int (x-x_0)^{-\mu} dx$ mit $\mu \in \mathbb{N}, x_0 \in \mathbb{R}$.

Typ 2: $\int \frac{bx}{((x-\alpha)^2+\beta^2)^\nu} dx$ mit $\nu \in \mathbb{N}$ $\alpha, \beta, b \in \mathbb{R}$, $\beta \neq 0, b \neq 0$.

Typ 3: $\int \frac{c}{((x-\alpha)^2+\beta^2)^\nu} dx$ mit $\nu \in \mathbb{N}$ $\alpha, \beta, c \in \mathbb{R}$, $\beta \neq 0$.

Typ 1 bereitet keine Schwierigkeiten. Es ist

$$\int (x-x_0)^{-\mu} dx = \begin{cases} \frac{-1}{\mu-1}(x-x_0)^{1-\mu} + C & \text{für } \mu > 1\,, \\ \log|x-x_0| + C & \text{für } \mu = 1\,. \end{cases}$$

Zu Typ 2: Es ist

$$\frac{bx}{((x-\alpha)^2+\beta^2)^\nu} = \frac{b}{2} \cdot \frac{2(x-\alpha)}{((x-\alpha)^2+\beta^2)^\nu} + \frac{b\alpha}{((x-\alpha)^2+\beta^2)^\nu}\,,$$

wobei der zweite Summand vom Typ 3 ist. Für den ersten gilt mit der Substitution $\varphi(x) = (x-\alpha)^2 + \beta^2$, $\varphi'(x) = 2(x-\alpha)$:

$$\int \frac{b}{2} \cdot \frac{2(x-\alpha)}{((x-\alpha)^2+\beta^2)^\nu} = \frac{b}{2} \int \frac{\varphi'(x)}{\varphi(x)^\nu} dx = \begin{cases} \frac{-1}{\nu-1}((x-\alpha)^2+\beta^2)^{1-\nu} + C & \text{für } \nu > 1\,, \\ \log((x-\alpha)^2+\beta^2) + C & \text{für } \nu = 1\,. \end{cases}$$

Zu Typ 3: Für $\nu = 1$ hat man (mit einer naheliegenden Substitution und unter Verwendung von Beispiel 6.20 (b)):

$$\int \frac{c}{((x-\alpha)^2+\beta^2)} dx = \frac{1}{\beta} \arctan\left(\frac{x-\alpha}{\beta}\right) + C\,.$$

Für

$$I_\nu := \int \frac{c}{((x-\alpha)^2+\beta^2)^\nu} dx$$

gilt im Fall $\nu > 1$ (wie man durch Differentiation der folgenden Gleichung verifiziert):

$$I_\nu = \frac{x-\alpha}{2(\nu-1)\beta^2((x-\alpha)^2+\beta^2)^{\nu-1}} + \frac{2\nu-3}{2(\nu-1)\beta^2} I_{\nu-1}\,.$$

Damit kann I_ν induktiv berechnet werden.

6.30. Beispiel. Sei $f(x) := \frac{1-x}{x^2(x^2+1)} = \frac{h(x)}{q(x)}$. Polynomdivision ist hier nicht nötig, da $\deg(h) = 1 < 4 = \deg(q)$. Wir führen die Partialbruchzerlegung durch:

$$f(x) = \frac{1-x}{x^2(x^2+1)} = \frac{a_{1,1}}{x} + \frac{a_{1,2}}{x^2} + \frac{bx+c}{x^2+1}\,.$$

Zur Bestimmung der Koeffizienten $a_{1,1}, a_{1,2}, b, c$ multiplizieren wir mit dem Hauptnenner durch und erhalten

$$1-x = a_{1,1}x(x^2+1) + a_{1,2}(x^2+1) + bx^3 + cx^2 = (a_{1,1}+b)x^3 + (a_{1,2}+c)x^2 + a_{1,1}x + a_{1,2}\,.$$

Koeffizientenvergleich liefert die Gleichungen:

$$\begin{aligned} a_{1,1} + b &= 0 \\ a_{1,2} + c &= 0 \\ a_{1,1} &= -1 \\ a_{1,2} &= 1 \end{aligned}$$

Es folgt $a_{1,1} = c = -1$, $a_{1,2} = b = 1$ und somit

$$f(x) = -\frac{1}{x} + \frac{1}{x^2} + \frac{x-1}{x^2+1} = -\frac{1}{x} + \frac{1}{x^2} + \frac{1}{2} \cdot \frac{2x}{x^2+1} - \frac{1}{x^2+1}\,.$$

Als Stammfunktion erhalten wir also

$$\int f(x)dx = -\log|x| - \frac{1}{x} + \frac{1}{2}\log(x^2+1) - \arctan(x) + C\,.$$

5. Uneigentliche Integrale

Bisher wurden nur beschränkte Funktionen auf endlichen Intervallen integriert. Wir wollen nun untersuchen, wie man ein Integral auch für gewisse unbeschränkte Funktionen und über unendliche Intervalle erklären kann.

6.31. Definition. (a) Sei $-\infty < a < b \leq +\infty$. Eine Funktion $f : [a,b) \to \mathbb{K}^N$ heißt über $[a,b)$ *uneigentlich integrierbar*, falls $f|_{[a,c]}$ für alle $c \in (a,b)$ auf $[a,c]$ Riemann–integrierbar ist und der Grenzwert

$$\int_a^b f(x)dx := \lim_{c \to b^-} \int_a^c f(x)dx$$

in $\mathbb{K}^n$ existiert.

(b) Sei $-\infty \leq a < b < +\infty$. Eine Funktion $f : (a,b] \to \mathbb{K}^N$ heißt über $(a,b]$ *uneigentlich integrierbar*, falls $f|_{[c,b]}$ für alle $c \in (a,b)$ auf $[a,c]$ Riemann–integrierbar ist und der Grenzwert

$$\int_a^b f(x)dx := \lim_{c \to a^+} \int_c^b f(x)dx$$

in $\mathbb{K}^n$ existiert.

(c) Sei $-\infty \leq a < b \leq +\infty$. Eine Funktion $f : (a,b) \to \mathbb{K}^N$ heißt über (a,b) *uneigentlich integrierbar*, falls es ein $c \in (a,b)$ gibt, so daß $f|_{(a,c]}$ im Sinne von (b) und $f|_{[c,b)}$ im Sinne von (a) uneigentlich integrierbar sind. Man setzt dann

$$\int_a^b f(x)dx := \int_a^c f(x)dx + \int_c^b f(x)dx$$

und überlegt sich leicht, daß dieser Wert von der speziellen Wahl von c unabhängig ist.

Bei der Berechnung uneigentlicher Integrale hilft folgender Satz weiter:

6.32. Satz. *Sei $-\infty \leq a < b \leq +\infty$ und sei $f : (a,b) \to \mathbb{K}^N$ stetig. Für eine Stammfunktion $F : (a,b) \to \mathbb{K}^N$ von f gelte, daß die Grenzwerte*

$$F(a^+) := \lim_{x \to a^+} F(x) \qquad \text{und} \qquad F(b^-) := \lim_{x \to b^-} F(x)$$

existieren. Dann ist f über (a,b) uneigentlich integrierbar und es gilt

$$\int_a^b f(x)dx = F(b^-) - F(a^+)\,.$$

Beweis. Wir fixieren ein beliebiges $c \in (a,b)$. Nach dem Hauptsatz der Differential– und Integralrechnung ist für alle $s \in (a,c)$ die Funktion $f|_{[s,c]}$ auf $[s,c]$ und für alle $t \in (c,b)$ die Funktion $f|_{[c,t]}$ auf $[c,t]$ Riemann–integrierbar und es gilt

$$\lim_{s \to a^+} \int_s^c f(x)dx = \lim_{s \to a^+} (F(c) - F(s)) = F(c) - F(a^+)$$

$$\lim_{t \to b^-} \int_c^t f(x)dx = \lim_{t \to b^-} (F(t) - F(c)) = F(b^-) - F(c)\,.$$

f ist also über (a,b) uneigentlich integrierbar und es gilt $\int_a^b f(x)dx = F(b^-) - F(a^+)$. $\square$

6.33. Beispiele. (a) Wir betrachten für alle $\alpha \in \mathbb{R}$ das Integral $\displaystyle\int_1^\infty \frac{1}{x^\alpha} dx$.

Für $\alpha < 1$ gilt

$$\int_1^c \frac{1}{x^\alpha} dx = \frac{c^{1-\alpha} - 1}{1-\alpha} \to \infty \quad \text{für } c \to \infty$$

und das uneigentliche Integral existiert nicht.

Für $\alpha > 1$ gilt

$$\int_1^c \frac{1}{x^\alpha} dx = \frac{c^{1-\alpha} - 1}{1-\alpha} \to \frac{1}{\alpha - 1} \quad \text{für } c \to \infty$$

und wir erhalten

$$\int_1^\infty \frac{1}{x^\alpha} dx = \frac{1}{\alpha - 1}.$$

Im Grenzfall $\alpha = 1$ liegt Divergenz vor wegen

$$\int_1^c \frac{1}{x} dx = \log(c) \to \infty \quad \text{für } c \to \infty.$$

(b) Wir betrachten $\displaystyle\int_0^1 x^\alpha dx$ für alle $\alpha \in \mathbb{R}$.

Für $\alpha < -1$ gilt

$$\int_c^1 x^\alpha dx = \frac{1 - c^{1+\alpha}}{1+\alpha} \to \infty \quad \text{für } c \to 0^+$$

und das uneigentliche Integral existiert nicht.

Für $\alpha > -1$ gilt

$$\int_c^1 x^\alpha dx = \frac{1 - c^{1+\alpha}}{1+\alpha} \to \frac{1}{\alpha + 1} \quad \text{für } c \to 0^+$$

und wir erhalten

$$\int_0^1 x^\alpha dx = \frac{1}{\alpha + 1}.$$

Im Grenzfall $\alpha = -1$ liegt wegen

$$\int_c^1 \frac{1}{x} dx = -\log(c) \to \infty \quad \text{für } c \to 0^+$$

Divergenz vor.

6.34. Satz (Cauchy–Kriterium für uneigentliche Integrierbarkeit). *Sei $-\infty < a < b \leq +\infty$ und sei $f : [a,b) \to \mathbb{K}^N$ für alle $c \in (a,b)$ auf $[a,c]$ Riemann–integrierbar. Genau dann ist f auf $[a,b)$ uneigentlich integrierbar, wenn gilt:*

$$\forall \varepsilon > 0 \quad \exists x_0 = x_0(\varepsilon) \in (a,b) \quad \forall x_1, x_2 \in [x_0, b) : \qquad \left| \int_{x_1}^{x_2} f(t)dt \right| < \varepsilon. \tag{6.16}$$

Eine analoge Aussage gilt auch für Integrale, die an der unteren Integrationsgrenze uneigentlich sind.

Beweis. "$\Longrightarrow$": Existiert das uneigentliche Integral

$$J = \int_a^b f(t)dt = \lim_{c \to b^-} \int_a^c f(t)dt\,,$$

so gibt es zu beliebigem $\varepsilon > 0$ ein $x_0 \in (a,b)$ mit

$$\forall x \in [x_0, b) : \left| J - \int_a^x f(t)dt \right| < \frac{\varepsilon}{2}\,.$$

Für alle $x_1, x_2 \in [x_0, b)$ gilt dann:

$$\left| \int_{x_1}^{x_2} f(t)dt \right| \le \left| \int_a^{x_2} f(t)dt - J \right| + \left| J - \int_a^{x_1} f(t)dt \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also ist (6.16) erfüllt.

"$\Longleftarrow$": Sei nun (6.16) erfüllt. Sei $(x_n)_{n=1}^\infty$ eine beliebige Folge aus (a,b) mit $x_n \to b$ für $n \to \infty$. Dann ist wegen (6.16) das Cauchy–Kriterium für die Folge $(\int_a^{x_n} f(t)dt)_{n=1}^\infty$ erfüllt. Also existiert der Grenzwert

$$J := \lim_{n\to\infty} \int_a^{x_n} f(t)dt \quad \text{in } \mathbb{K}^N.$$

Wir zeigen $J = \lim_{x \to b^-} \int_a^x f(t)dt$. Sei also $\varepsilon > 0$ beliebig und sei x_0 zu $\varepsilon' := \varepsilon/2$ gemäß (6.16) gewählt. Wegen $x_n \to b$ für $n \to \infty$ und $\int_a^{x_n} f(t)dt \to J$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$\forall n \ge n_0 : \qquad \left| \int_a^{x_n} f(t)dt - J \right| < \frac{\varepsilon}{2} \quad \text{und } x_0 < x < b.$$

Für alle $x \in (x_0, b)$ gilt dann

$$\begin{aligned} \left| \int_a^x f(t)dt - J \right| &\le \left| \int_a^x f(t)dt - \int_a^{x_{n_0}} f(t)dt \right| + \left| \int_a^{x_{n_0}} f(t)dt - J \right| \\ &< \left| \int_{x_{n_0}}^x f(t)dt \right| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,. \end{aligned}$$

Also ist f auf $[a,b)$ uneigentlich integrierbar und $\int_a^b f(t)dt = J$. □

6.35. Beispiel. Für alle $\alpha > 0$ existiert das uneigentliche Integral

$$\int_1^\infty \frac{\sin(t)}{t^\alpha} dt\,.$$

Beweis. Für $1 < x_1 < x_2$ erhalten wir durch partielle Integration

$$\begin{aligned} \left| \int_{x_1}^{x_2} \frac{\sin(t)}{t^\alpha} dt \right| &= \left| -\frac{\cos(t)}{t^\alpha} \Big|_{x_1}^{x_2} - \alpha \int_{x_1}^{x_2} \frac{\cos(t)}{t^{\alpha+1}} dt \right| \le \frac{|\cos(x_2)|}{x_2^\alpha} + \frac{|\cos(x_1)|}{x_1^\alpha} + \alpha \int_{x_1}^{x_2} \frac{1}{t^{\alpha+1}} dt \\ &\le \frac{1}{x_2^\alpha} + \frac{1}{x_1^\alpha} + \frac{1}{x_1^\alpha} - \frac{1}{x_2^\alpha} = \frac{2}{x_1^\alpha}\,. \end{aligned}$$

Sei nun $\varepsilon > 0$ beliebig. Für alle $x_1, x_2 \in \mathbb{R}$ mit $x_1, x_2 \ge x_0 := \left(\frac{3}{\varepsilon}\right)^{1/\alpha}$ gilt dann (mit o.B.d.A. $x_1 < x_2$):

$$\left| \int_{x_1}^{x_2} \frac{\sin(t)}{t^\alpha} dt \right| \le \frac{2}{x_1^\alpha} \le \frac{2}{x_0^\alpha} = \frac{2}{3}\varepsilon\,.$$

Also ist (6.16) hier erfüllt und nach dem Cauchy–Kriterium 6.34 existiert das uneigentliche Integral. □

Auch zum Majorantenkriterium gibt es eine Variante für uneigentliche Integrale.

6.36. Satz (Majorantenkriterium für uneigentliche Integrale). *Sei* $-\infty < a < b \leq +\infty$ *und sei* $f : [a,b) \to \mathbb{K}^N$ *für alle* $c \in (a,b)$ *auf* $[a,c]$ *Riemann–integrierbar. Sei ferner* $h : [a,b) \to [0,\infty)$ *eine Funktion, die auf* $[a,b)$ *uneigentlich integrierbar ist und so, daß es ein* $d \in [a,b)$ *gibt mit* $|f(t)| \leq h(t)$ *für alle* $t \in [d,b)$. *Dann sind auch* f *und* $|f|$ *auf* $[a,b)$ *uneigentlich integrierbar.*

Beweis. Sei $\varepsilon > 0$ beliebig. Da h auf $[a,b)$ uneigentlich integrierbar ist, gibt es nach dem Cauchy–Kriterium 6.34 ein $x_* \in (a,b)$ mit

$$\forall x_1, x_2 \in [x_*, b) : \qquad \left| \int_{x_1}^{x_2} h(t)dt \right| < \varepsilon \,.$$

Mit $x_0 := \max\{d, x_*\}$ gilt dann für alle $x_1, x_2 \in [x_0, b)$ (mit o.B.d.A. $x_1 < x_2$):

$$\left| \int_{x_1}^{x_2} f(t)dt \right| \leq \int_{x_1}^{x_2} |f(t)|dt \leq \int_{x_1}^{x_2} h(t)dt < \varepsilon \,.$$

Mit dem Cauchy–Kriterium 6.34 folgt die Behauptung. □

Eine analoge Aussage gilt natürlich auch für am linken Intervallende oder an beiden Intervallenden uneigentliche Integrale.

Das folgende Integralkriterium stellt einen Zusammenhang zwischen der Konvergenz gewisser unendlicher Reihen und gewisser uneigentlicher Integrale her.

6.37. Satz (Integralkriterium). *Sei* $m \in \mathbb{Z}$, *sei* $(a_n)_{n=m}^{\infty}$ *eine monoton fallende Folge nicht negativer reeller Zahlen und sei* $f : [m,\infty) \to [0,\infty)$ *monoton fallende Funktion mit* $f(n) = a_n$ *für alle* $n \geq m$ *aus* $\mathbb{N}$. *Dann ist die Reihe* $\sum_{n=m}^{\infty} a_n$ *genau dann konvergent, wenn das uneigentliche Integral* $\int_m^{\infty} f(t)dt$ *existiert.*

Beweis. Da f monoton fallend und beschränkt ist, ist f nach Lemma 6.6 auf jedem abgeschlossenen Intervall $[a,b] \subset [m,\infty)$ Riemann–integrierbar und es gilt für alle $k \in \mathbb{Z}$ mit $k \geq m$:

$$\forall t \in [k, k+1] : \qquad a_k = f(k) \geq f(t) \geq f(k+1) = a_{k+1} \geq 0 \,.$$

Mit Lemma 6.7 erhalten wir durch Integration von k bis $k+1$:

$$\forall k \geq m, k \in \mathbb{N}_0 : \qquad a_k \geq \int_k^{k+1} f(t)dt \geq a_{k+1} \geq 0.$$

Hieraus folgt für alle $n_1, n_2 \in \mathbb{N}_0$ mit $m \leq n_1 < n_2$:

$$\sum_{k=n_1}^{n_2} a_k \geq \sum_{k=n_1}^{n_2} \int_k^{k+1} f(t)dt = \int_{n_1}^{n_2+1} f(t)dt \geq \sum_{k=n_1+1}^{n_2+1} a_k \geq 0. \tag{6.17}$$

Hieraus sieht man leicht, daß die unendliche Reihe $\sum_{k=m}^{\infty} a_k$ genau dann der Bedingung (b) des Cauchy–Kriteriums 3.41 genügt, wenn für das uneigentliche Integral die Bedingung (6.16) des Cauchy–Kriteriums 6.34 für uneigentliche Integrale erfüllt ist. Mit den beiden Cauchy–Kriterien folgt also die Behauptung. □

6.38. Beispiele. (a) Für $\alpha \in \mathbb{R}$ gilt: Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ konvergiert genau dann, wenn $\alpha > 1$. Dies folgt mit $f(t) := t^{-\alpha}$ unmittelbar aus dem Integralkriterium 6.37 und Beispiel 6.33 (a).

(b) Für $\alpha \in (0,\infty)$ gilt: Die Reihe $\sum_{n=2}^{\infty} \frac{1}{n \cdot \log(n)^\alpha}$ konvergiert genau dann, wenn $\alpha > 1$.

Beweis. Zu (b): Die durch

$$f(t) := \frac{1}{t\log(t)^\alpha} \qquad \text{für } t \geq 2$$

definierte Funktion $f : [2,\infty) \to \mathbb{R}$ nimmt ihre Werte in $(0,\infty)$ an und ist wegen

$$f'(t) = -\frac{\log(t)^\alpha + \alpha\log(t)^{\alpha-1}}{t^2\log(t)^{2\alpha}} = -\frac{\log(t)+\alpha}{t^2\log(t)^{\alpha+1}} < 0 \qquad \text{für } t \geq 2$$

streng monoton fallend. Mit der Substitution $u = \varphi(t) := \log(t)$ hat man

$$\int_2^x \frac{1}{t\log(t)^\alpha}dt = \int_{\log(2)}^{\log(x)} u^{-\alpha}du\,.$$

Nach Beispiel 6.33 (a) existiert also das uneigentliche Integral

$$\int_2^\infty \frac{1}{t\log(t)^\alpha}dt$$

genau dann, wenn $\alpha > 1$ gilt. Mit dem Integralkriterium folgt die Behauptung. $\square$

KAPITEL 7

Differentialrechnung in mehreren Veränderlichen

Falls nicht anders vermerkt, sei in diesem Kapitel Ω stets eine nicht leere, offene Teilmenge des $\mathbb{R}^N$. Wir betrachten nun Funktionen $f : \Omega \to \mathbb{R}^M$ (wobei $N, M \in \mathbb{N}$ natürliche Zahlen seien) und wollen untersuchen, welche der Ergebnisse aus Kapitel 5 sinnvolle Übertragungen auf den Fall mehrerer Veränderlicher zulassen.

1. Partielle Differentiation

7.1. Definition. Sei $f : \Omega \to \mathbb{R}^M$ eine $\mathbb{R}^M$–wertige Funktion auf Ω, sei $a \in \Omega$ und sei $u \in \mathbb{R}^N$ mit $|u| = 1$. Da Ω offen ist, gibt es ein $\delta > 0$, so daß für alle $t \in \mathbb{R}$ mit $|t| < \delta$ der Punkt $a + tu$ in Ω liegt. Wir definieren: *f ist im Punkt a in Richtung u differenzierbar*, falls der Grenzwert

$$(D_u f)(a) := \lim_{t\to 0} \frac{1}{t}(f(a + tu) - f(a))$$

existiert. Der Wert $(D_u f)(a)$ heißt dann die *Richtungsableitung von f an der Stelle a in Richtung u.*

7.2. Bemerkungen. Sei $f : \Omega \to \mathbb{R}^M$ eine $\mathbb{R}^M$–wertige Funktion auf Ω, sei $a \in \Omega$ und sei $u \in \mathbb{R}^N$ mit $|u| = 1$.

(a) Die Existenz der Richtungsableitung $(D_u f)(a)$ von f an der Stelle a in Richtung u ist also dazu äquivalent, daß die in $(-\delta, \delta)$ definierte Funktion g mit

$$g(t) := f(a + tu) \qquad \text{für alle } t \in (-\delta, \delta)$$

im Punkt 0 als Funktion in einer reellen Veränderlichen differenzierbar ist. Ist dies der Fall, so ist $g'(0) = (D_u f)(a)$.

(b) Existiert die Richtungsableitung nach u in allen Punkten $x \in \Omega$, so gilt

$$g'(t) = (D_u f)(a + tu) \quad \text{für alle } t \in \mathbb{R} \quad \text{mit } a + tu \in \Omega\,.$$

Beweis. (b) Ist $a+tu \in \Omega$, so gibt es ein $\delta > 0$ mit $a+(t+s)u \in \Omega$ für alle $s \in (-\delta, \delta)$. Es folgt

$$\frac{1}{s}(g(t + s) - g(t)) = \frac{1}{s}(f((a + tu) + su) - f(a + tu)) \to (D_u f)(a + tu) \quad \text{für } s \to 0$$

und damit die Behauptung. □

Im Fall $N = 1$ gibt es nur die Richtungen 1 und -1 und die Definition 7.1 stimmt mit der Definition der Differenzierbarkeit überein. Insbesondere folgt in diesem Fall aus der Existenz der Richtungsableitung von f an der Stelle a in Richtung 1 bzw. -1 die Stetigkeit von f im Punkt a. In mehreren Veränderlichen muß f nicht stetig sein, selbst, wenn alle Richtungsableitungen von f an der Stelle a existieren. Wir zeigen dies durch ein Beispiel:

7.3. Beispiel. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch

$$f(x_1, x_2) := \begin{cases} \dfrac{x_1 x_2^2}{x_1^2 + x_2^4} & \text{für alle } x_1, x_2 \in \mathbb{R} \text{ mit } x_1 \neq 0 \\ 0 & \text{für } x_1 = 0 \text{ und alle } x_2 \in \mathbb{R}\,. \end{cases}$$

Für alle $u = (u_1, u_2) \in \mathbb{R}^2$ mit $|u| = 1$ gilt dann

$$\lim_{t\to 0} \frac{1}{t}(f((0,0) + tu) - f(0,0)) = \lim_{t\to 0} \frac{1}{t} f(tu) = \begin{cases} 0 & \text{falls } u_1 = 0 \\ \dfrac{u_2^2}{u_1} & \text{für } u_1 \neq 0\,. \end{cases}$$

Die Richtungsableitung $(D_u f)(0,0)$ existiert also für alle $u \in \mathbb{R}^2$ mit $|u| = 1$. Die Funktion f ist jedoch in $(0,0)$ unstetig, denn für die gegen $(0,0)$ konvergente Folge $((n^{-2}, n^{-1}))_{n=1}^{\infty}$ gilt

$$\lim_{n\to\infty} f(n^{-2}, n^{-1}) = \frac{1}{2} \neq f(0,0) = 0\,.$$

7.4. Definition. Sei $f = (f_1, \ldots, f_M) : \Omega \to \mathbb{R}^M$ eine $\mathbb{R}^M$–wertige Funktion auf Ω, sei $a \in \Omega$ und sei $e_j := (0, \ldots, 0, 1, 0, \ldots, 0)$ der j–te Einheitsvektor der kanonischen Basis in $\mathbb{R}^N$, also mit der j–ten Komponente 1 und 0 in allen anderen Komponenten. Falls die Richtungsableitung $(D_{e_j} f)(a)$ existiert, so sagen wir f *ist in* a *partiell nach der* j*–ten Variablen differenzierbar* und schreiben auch

$$(D_j f)(a) \qquad \text{oder} \qquad \frac{\partial f}{\partial x_j}(a) \qquad \text{oder} \qquad f_{x_j}(a)$$

statt $(D_{e_j} f)(a)$. Der Wert $(D_j f)(a)$ heißt die *partielle Ableitung von* f *in* a *nach der* j*–ten Variablen.*

f heißt auf Ω *partiell differenzierbar*, falls für alle $a \in \Omega$ die partiellen Ableitungen $(D_1 f)(a), \ldots, (D_N f)(a)$ nach allen Variablen existieren.

7.5. Bemerkung. (a) Da Konvergenz in $\mathbb{R}^M$ mit der komponentenweisen Konvergenz übereinstimmt, existiert für eine Funktion $f = (f_1, \ldots, f_M) : \Omega \to \mathbb{R}^M$, ein $x \in \Omega$ und ein $j \in \{1, \ldots, N\}$ (bzw. ein $u \in \mathbb{R}^N$ mit $|u| = 1$) genau dann die partielle Ableitung $(D_j f)(x)$ (bzw. die Richtungsableitung $(D_u f)(x)$) in x, wenn alle $(D_j f_k)(x)$ (bzw. $(D_u f_k)(x)$) existieren für $k = 1, \ldots, M$. Es gilt dann

$$(D_j f)(x) = \big((D_j f_k)(x)\big)_{k=1}^{M} \qquad \text{bzw.} \quad (D_u f)(x) = \big((D_u f_k)(x)\big)_{k=1}^{M}.$$

(b) Nach Definition der Richtungsableitung und der partiellen Ableitung ist eine Funktion $f : \Omega \to \mathbb{R}^M$ genau dann in einem Punkt $a \in \Omega$ partiell nach der j–ten Variablen partiell differenzierbar, wenn für die in einer Umgebung von 0 durch

$$t \mapsto f(a_1, \ldots, a_{j-1}, a_j + t, a_{j+1}, \ldots, a_N)$$

definierte Funktion in 0 differenzierbar ist. Dies ist offensichtlich dazu äquivalent, daß die in einer Umgebung von a_j definierte Funktion

$$\xi \mapsto h(\xi) := f(a_1, \ldots, a_{j-1}, \xi, a_{j+1}, \ldots, a_N)$$

in a_j differenzierbar ist. Es gilt dann $(D_j f)(a) = h'(a_j)$. Die aus der Differentialrechnung einer Veränderlichen bekannten Summen–, Produkt– und Quotientenregeln übertragen sich also unmittelbar auf den Fall partieller Ableitungen.

7.6. Beispiele. Sei $r : \mathbb{R}^N \to \mathbb{R}$ die durch $r(x) := |x|$ für alle $x \in \mathbb{R}^N$ definierte Funktion.

(a) Für $0 \neq x \in \mathbb{R}^N$ und $j = 1, \ldots, N$ ist nach der Kettenregel in einer Veränderlichen die Funktion

$$\begin{aligned}\xi \mapsto h_j(\xi) :&= r(x_1, \ldots, x_{j-1}, \xi, x_{j+1}, \ldots, x_N)\\ &= \big(x_1^2 + \cdots + x_{j-1}^2 + \xi^2 + x_{j+1}^2 + \cdots + x_N^2\big)^{1/2}\end{aligned}$$

in dem Punkt x_j differenzierbar und es gilt: $h_j'(x_j) = \frac{x_j}{|x|}$. Nach der vorstehenden Bemerkung ist also r auf $\mathbb{R}^N \setminus \{0\}$ partiell differenzierbar und es gilt:

$$\frac{\partial r}{\partial x_j}(x) = \frac{x_j}{|x|} \qquad \text{für alle } 0 \neq x \in \mathbb{R}^N\,, j = 1, \ldots, N\,.$$

(b) Sei nun $f : (0, \infty) \to \mathbb{R}^M$ eine differenzierbare Funktion. Wenden wir für $j = 1, \ldots, N$ die Kettenregel in einer Veränderlichen an auf die Funktion

$$\xi \mapsto f\big(r(x_1, \ldots, x_{j-1}, \xi, x_{j+1}, \ldots, x_N)\big)\,,$$

so erhalten wir: $f \circ r$ ist auf $\mathbb{R}^N \setminus \{0\}$ partiell differenzierbar und es gilt

$$\frac{\partial (f \circ r)}{\partial x_j}(x) = \frac{x_j}{|x|} f'(r(x)) \qquad \text{für alle } 0 \neq x \in \mathbb{R}^N\,, j = 1, \ldots, N\,.$$

(c) Sei $g : \mathbb{R}^N \to \mathbb{R}$ definiert durch

$$g(x) := \begin{cases} \dfrac{x_1 \cdot \ldots \cdot x_N}{|x|^{2N}} & \text{für } 0 \neq x = (x_1, \ldots, x_N) \in \mathbb{R}^N \\ 0 & \text{für } x = 0\,. \end{cases}$$

Mit der Produktregel und (b) (angewendet auf die Funktion $f : t \mapsto f(t) := t^{-2N}$) folgt die partielle Differenzierbarkeit von g auf $\mathbb{R}^N \setminus \{0\}$ und wir erhalten für $0 \neq x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ und $j = 1, \ldots, N$:

$$\begin{aligned}\frac{\partial g}{\partial x_j}(x) &= \frac{x_1 \cdot \ldots \cdot x_{j-1} \cdot x_{j+1} \cdot \ldots \cdot x_N}{r(x)^{2N}} + x_1 \cdot \ldots \cdot x_N \cdot \frac{\partial r^{-2N}}{\partial x_j}(x)\\ &= \frac{x_1 \cdot \ldots \cdot x_{j-1} \cdot x_{j+1} \cdot \ldots \cdot x_N}{|x|^{2N}} - 2N \cdot x_1 \cdot \ldots \cdot x_N \cdot \frac{x_j}{|x|} \cdot \frac{1}{r(x)^{2N+1}}\\ &= \frac{x_1 \cdot \ldots \cdot x_{j-1} \cdot x_{j+1} \cdot \ldots \cdot x_N}{|x|^{2N}} \Big(1 - 2N \cdot \frac{x_j^2}{|x|^2}\Big)\,.\end{aligned}$$

Auch in $x = 0$ ist g partiell differenzierbar. Es gilt

$$\frac{\partial g}{\partial x_j}(0) = \lim_{t \to 0} \frac{g(te_j) - g(0)}{t} = 0\,.$$

Ist $u \in \mathbb{R}^N$ mit $|u| = 1$, so gilt

$$\begin{aligned}\frac{g(tu) - g(0)}{t} &= \frac{t^N \cdot u_1 \cdot \ldots \cdot u_N}{t^{2N+1}|u|^{2N}} = \frac{u_1 \cdot \ldots \cdot u_N}{t^{N+1}}\\ &\to \begin{cases} 0 & \text{falls } u_j = 0 \quad \text{für wenigstens ein } j \in \{1, \ldots, N\} \\ \pm\infty & \text{sonst.} \end{cases}\end{aligned}$$

Aus der partiellen Differenzierbarkeit folgt also *nicht* die Existenz aller Richtungsableitungen.

Es folgt ein hinreichendes Kriterium für die Vertauschbarkeit von Integration und Bildung der Richtungsableitungen:

7.7. Satz. *Sei* $-\infty < a < b < \infty$*, sei* $\Omega \subseteq \mathbb{R}^N$ *offen und sei* $u \in \mathbb{R}^N$ *mit* $|u| = 1$. $f : [a,b] \times \Omega \to \mathbb{R}^M$ *sei eine Funktion mit den beiden folgenden Eigenschaften:*

(a) *Für alle* $x \in \Omega$ *ist die Funktion* $s \mapsto f(s,x)$ *stetig auf* $[a,b]$.
(b) *Für alle* $(s,x) \in [a,b] \times \Omega$ *existiert die Richtungsableitung*

$$(D_u f)(s,x) = \lim_{t\to 0} \frac{1}{t}(f(s,x+tu) - f(s,x))$$

und die hierdurch definierte Funktion

$$D_u f : [a,b] \times \Omega \to \mathbb{R}^M, \qquad (s,x) \mapsto (D_u f)(s,x),$$

ist stetig auf $[a,b] \times \Omega$.

Dann existiert für die durch

$$F(x) := \int_a^b f(s,x)ds \qquad \text{für alle } x \in \Omega$$

definierte Funktion $F : \Omega \to \mathbb{R}^M$ *in alle Punkten* $x \in \Omega$ *die Richtungsableitung* $(D_u F)(x)$ *und es gilt*

$$(D_u F)(x) = \int_a^b (D_u f)(s,x)ds\,. \tag{7.1}$$

Die hierdurch definierte Funktion $D_u F : \Omega \to \mathbb{R}^M$ *ist auf* Ω *stetig.*

Beweis. Da die Bildung der Richtungsableitung komponentenweise erfolgt, genügt es die Behauptung für den Fall $M = 1$ zu beweisen. Sei also $f : [a,b] \times \Omega \to \mathbb{R}$ eine Funktion, die (a) und (b) erfüllt und seien $x \in \Omega$ und $\varepsilon > 0$ beliebig. Da Ω offen ist gibt es ein $\delta_0 > 0$ mit $B_{\delta_0}(x) := \{\xi \in \mathbb{R}^N ; |\xi - x| \leq \delta_0\} \subset \Omega$. Wegen (a) ist $s \mapsto f(s,x)$ für alle $x \in B_{\delta_0}(x)$ auf $[a,b]$ stetig und daher Riemann–integrierbar. Da $[a,b] \times B_{\delta_0}(x)$ eine beschränkte und abgeschlossene Teilmenge von $\mathbb{R}^{N+1}$ ist, ist die Funktion $D_u f$ schon gleichmäßig stetig auf $[a,b] \times B_{\delta_0}(x)$ (vergl. Satz 4.38). Es gibt also ein $\delta \in (0,\delta_0)$, so daß für alle $(s_1,\xi_1),(s_2,\xi_2) \in [a,b] \times B_{\delta_0}(x)$ gilt:

$$|(s_1,\xi_1) - (s_2,\xi_2)| < \delta \quad \Longrightarrow \quad |(D_u f)(s_1,\xi_1) - (D_u f)(s_2,\xi_2)| < \frac{\varepsilon}{b-a+1}\,. \tag{7.2}$$

Sei nun $t \in (-\delta,\delta)$ beliebig. Nach dem Mittelwertsatz der Differentialrechnung, angewendet auf die Funktion $g : t \mapsto f(s,x+tu)$ gibt es zu jedem $s \in [a,b]$ ein $\tau(s,t)$ zwischen 0 und t mit (vergl. Bemerkung 7.2):

$$\frac{1}{t}(f(s,x+tu) - f(s,x)) = g'(\tau(s,t)) = D_u f(s,x+\tau(s,t)u)\,.$$

Für alle $s \in [a,b]$ hat man für $0 < |t| < \delta$

$$|(s,x+\tau(s,t)u) - (s,x)| = |x+\tau(s,t)u - x| = |\tau(s,t)| \leq |t| < \delta$$

und daher wegen (7.2):

$$\begin{aligned}\Big|\frac{1}{t}(F(x+tu)-F(x))-&\int_a^b (D_uf)(s,x)ds\Big| =\\ &=\Big|\int_a^b\Big(\frac{1}{t}(f(s,x+tu)-f(s,x))-(D_uf)(s,x)\Big)ds\Big|\\ &\le \int_a^b\Big|\frac{1}{t}(f(s,x+tu)-f(s,x))-(D_uf)(s,x)\Big|ds\\ &\le \int_a^b|(D_u)f(s,x+\tau(s,t)u)-(D_uf)(s,x)|ds\\ &\le \int_a^b \frac{\varepsilon}{b-a+1}ds=\frac{(b-a)\varepsilon}{b-a+1}<\varepsilon\,.\end{aligned}$$

Also gilt

$$\lim_{t\to 0}\frac{1}{t}(F(x+tu)-F(x))=\int_a^b (D_uf)(s,x)ds$$

und damit die Behauptung. Die Stetigkeitsaussage folgt mit Satz 7.9. □

Da partielle Ableitungen nach der j-ten Variablen insbesondere Richtungsableitungen nach dem j–ten Einheitsvektor sind, erhalten wir als unmittelbare Folgerung:

7.8. Folgerung. *Sei $-\infty<a<b<\infty$, sei $\Omega\subseteq\mathbb{R}^N$ offen und sei $j\in\{1,\dots,N\}$. $f:[a,b]\times\Omega\to\mathbb{R}^M$ sei eine Funktion mit den beiden folgenden Eigenschaften:*

(a) *Für alle $x\in\Omega$ ist die Funktion $s\mapsto f(s,x)$ stetig auf $[a,b]$.*
(b) *Für alle $(s,x)\in[a,b]\times\Omega$ existiert die partielle Ableitung*

$$\frac{\partial f}{\partial x_j}(s,x)=\lim_{t\to 0}\frac{1}{t}(f(s,x+te_j)-f(s,x))$$

und die hierdurch definierte Funktion

$$\frac{\partial f}{\partial x_j}:[a,b]\times\Omega\to\mathbb{R}^M\,,\qquad (s,x)\mapsto\frac{\partial f}{\partial x_j}(s,x)\,,$$

ist stetig auf $[a,b]\times\Omega$.

Dann existiert für die durch

$$F(x):=\int_a^b f(s,x)ds\qquad \textit{für alle } x\in\Omega$$

definierte Funktion $F:\Omega\to\mathbb{R}^M$ in alle Punkten $x\in\Omega$ die partielle Ableitung $(\frac{\partial F}{\partial x_j})(x)$ und es gilt

$$\Big(\frac{\partial F}{\partial x_j}\Big)(x)=\int_a^b\frac{\partial f}{\partial x_j}(s,x)ds\,. \tag{7.3}$$

Die hierdurch definierte Funktion $\frac{\partial F}{\partial x_j}:\Omega\to\mathbb{R}^M$ ist auf Ω stetig.

Für Integrale, deren Integrand stetig von einem Parameter abhängt, erhalten wir folgende Stetigkeitsaussage (aus der man auch leicht die Stetigkeitsaussage in Satz 7.7 erhält).

7.9. Satz. *Sei $K \subset \mathbb{R}^M$ beschränkt und abgeschlossen. Sei $-\infty < a < b < \infty$. Für alle stetigen Funktionen $f : [a,b] \times K \to \mathbb{K}^N$ gilt: Die durch*

$$h(u) := \int_a^b f(x,u)dx \qquad \text{für alle } u \in K$$

definierte Funktion $h : K \to \mathbb{K}^N$ ist gleichmäßig stetig.

Beweis. Da auch $H := [a,b] \times K$ beschränkt und abgeschlossen ist, muß f schon gleichmäßig stetig auf H sein (vergl. Satz 4.38). Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es also ein $\delta > 0$, so daß für alle $x, y \in [a,b]$, $u, v \in K$ mit $|(x,u) - (y,v)| < \delta$ gilt

$$|f(x,u) - f(y,v)| < \frac{\varepsilon}{b-a+1}$$

Für alle $u, v \in K$ mit $|u - v| < \delta$ folgt also:

$$\begin{aligned} |h(u) - h(v)| &= \Big| \int_a^b f(x,u) - f(x,v)dx \Big| \leq \int_a^b |f(x,u) - f(x,v)|dx \\ &\leq \frac{\varepsilon}{b-a+1} L(b-a) < \varepsilon \end{aligned}$$

und damit die gleichmäßige Stetigkeit von h auf K. □

7.10. Definition. (a) Sei $f : \Omega \to \mathbb{R}$ eine partiell differenzierbare Funktion. Dann heißt die $\mathbb{R}^N$–wertige Funktion

$$\mathrm{grad} f : x \mapsto (\mathrm{grad} f)(x) := \Big(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_N}(x) \Big)$$

der *Gradient von f*. Statt grad$f$ schreibt man auch ∇f mit $\nabla := \big(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_N}\big)$.

(b) Unter einem *Vektorfeld* F auf einer offenen Menge $\Omega \subseteq \mathbb{R}^N$ versteht man eine Abbildung $F : \Omega \to \mathbb{R}^N$. So ist zum Beispiel der Gradient ∇f einer auf Ω partiell differenzierbaren Funktion $f : \Omega \to \mathbb{R}$ stets ein Vektorfeld. Ist $F = (f_1, \ldots, f_N) : \Omega \to \mathbb{R}^N$ ein partiell differenzierbares Vektorfeld auf Ω, so heißt die Funktion

$$\mathrm{div} F := \langle \nabla, F \rangle := \sum_{j=1}^N \frac{\partial f_j}{\partial x_j}$$

die *Divergenz* von F.

(c) Ist $\Omega \subseteq \mathbb{R}^3$ offen und $F : \Omega \to \mathbb{R}^3$ ein auf Ω partiell differenzierbares Vektorfeld, so heißt das durch

$$\mathrm{rot} F := \nabla \times F := \Big(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \Big)$$

definierte Vektorfeld rot$F : \Omega \to \mathbb{R}^3$ die *Rotation* des Vektorfeldes F.

Aus den Produktregeln für die partiellen Ableitungen erhält man unmittelbar die folgenden Produktregeln für ∇ und div:

7.11. Lemma. *Seien $f, g : \Omega \to \mathbb{R}$ partiell differenzierbare Funktionen und sei $F = (f_1, \ldots, f_N) : \Omega \to \mathbb{R}^N$ ein partiell differenzierbares Vektorfeld. Dann gilt:*

$$\nabla(fg) = f \cdot \nabla g + g \cdot \nabla f \tag{7.4}$$

und

$$
\begin{aligned}
\operatorname{div}(fF) &= \sum_{j=1}^{N} \frac{\partial (f\cdot f_j)}{\partial x_j} = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}\cdot f_j + f\cdot \sum_{j=1}^{N} \frac{\partial f_j}{\partial x_j} \\
&= \langle \nabla f, F\rangle + f\cdot \operatorname{div} F \\
&= \langle \nabla f, F\rangle + f\cdot \langle \nabla, F\rangle \,.
\end{aligned} \tag{7.5}
$$

7.12. Beispiele. (a) Die Identität $\mathrm{id} : \mathbb{R}^N \to \mathbb{R}^N$ mit $\mathrm{id}(x) = x$ für alle $x \in \mathbb{R}^N$ ist ein partiell differenzierbares Vektorfeld mit $\operatorname{div}(\mathrm{id})(x) = N$ für alle $x \in \mathbb{R}^N$.

(b) Ist $r : \mathbb{R}^N \to \mathbb{R}$, $r : x \mapsto |x|$, wie in Beispiel 7.6 und ist $f : (0,\infty) \to \mathbb{R}$ eine differenzierbare Funktion, so folgt für alle $0 \neq x \in \mathbb{R}^N \setminus \{0\}$:

$$(\nabla r)(x) = \frac{1}{r(x)} x = \frac{1}{|x|} x$$

und

$$(\nabla(f\circ r))(x) = \frac{f'(r(x))}{r(x)} x = \frac{f'(|x|)}{|x|} x \,.$$

Insbesondere ist $\nabla r : \mathbb{R}^N \setminus \{0\} \to \mathbb{R}^N$ ein partiell differenzierbares Vektorfeld und wir erhalten mit der Produktregel 7.11 und (a) für alle $0 \neq x \in \mathbb{R}^N$:

$$
\begin{aligned}
(\operatorname{div}(\nabla r))(x) &= \Big(\operatorname{div}\Big(\frac{1}{r}\mathrm{id}\Big)\Big)(x) = \Big\langle \Big(\nabla \frac{1}{r}\Big)(x), x \Big\rangle + \frac{(\operatorname{div}\mathrm{id})(x)}{r(x)} = -\frac{\langle x, x\rangle}{|x|^3} + \frac{N}{|x|} \\
&= \frac{N-1}{|x|} \,.
\end{aligned}
$$

7.13. Definition. Ist $f : \Omega \to \mathbb{R}^M$ eine auf einer offenen Menge $\Omega \subseteq \mathbb{R}^N$ partiell differenzierbare Funktion und sind die partiellen Ableitungen $D_1 f, \ldots, D_N f$ ebenfalls partiell differenzierbar auf Ω, so heißt f *zweimal partiell differenzierbar* auf Ω. Statt $(D_j(D_k f))(x)$ schreibt man auch

$$(D_j D_k f)(x) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x)\,, \qquad j,k = 1,\ldots,N\,.$$

Induktiv definieren wir: $f : \Omega \to \mathbb{R}^M$ heißt auf Ω *$(k+1)$–mal partiell differenzierbar*, falls f auf Ω k–mal partiell differenzierbar ist und falls für alle $j_1, \ldots, j_k \in \{1, \ldots, k\}$ auch die Funktionen $D_{j_k} \ldots D_{j_1} f$ partiell differenzierbar sind. Man definiert dann für alle $j_{k+1} \in \{1, \ldots, k\}$ und alle $x \in \Omega$:

$$\frac{\partial^{k+1} f}{\partial x_{j_{k+1}} \partial x_{j_k} \ldots \partial x_{j_1}} := (D_{j_{k+1}} D_{j_k} \ldots D_{j_1} f)(x) := (D_{j_{k+1}}(D_{j_k} \ldots D_{j_1} f))(x)\,.$$

Für die k–fache partielle Ableitung nach der j–ten Variablen schreiben wir auch

$$D_j^k f \qquad \text{oder} \qquad \frac{\partial^k f}{\partial x_j^k}\,.$$

Für alle $k \in \mathbb{N}_0$ definieren wir: $f : \Omega \to \mathbb{R}^M$ heißt auf Ω *k–mal stetig partiell differenzierbar*, falls f auf Ω k–mal partiell differenzierbar ist und sowohl f als auch alle partiellen Ableitungen der Ordnung $\leq k$ auf Ω stetig sind. Mit $C^k(\Omega, \mathbb{R}^M)$ bezeichnen wir die Menge aller auf Ω k–mal stetig partiell differenzierbaren Funktionen mit Werten in

$\mathbb{R}^M$. Insbesondere hat man $C^0(\Omega, \mathbb{R}^M) = C(\Omega, \mathbb{R}^M)$. Durch vollständige Induktion zeigt man, daß $C^k(\Omega, \mathbb{R}^M)$ ein Untervektorraum von $C(\Omega, \mathbb{R}^M)$ ist. Auch

$$C^\infty(\Omega, \mathbb{R}^M) := \bigcap_{k=0}^{\infty} C^k(\Omega, \mathbb{R}^M)$$

ist ein Untervektorraum von $C(\Omega, \mathbb{R}^M)$. Für $k \in \mathbb{N}_0 \cup \{\infty\}$ schreiben wir im Fall $M = 1$ auch $C^k(\Omega)$ statt $C^k(\Omega, \mathbb{R})$

Wie man an einem Beispiel in den Übungen sieht, kann es vorkommen, daß $D_i D_j f$ und $D_j D_i f$ verschieden sind. Der folgende Satz von H. A. Schwarz zeigt, daß dies bei zweimal stetig partiell differenzierbaren Funktionen nicht passieren kann.

7.14. Satz. *Seien $i, j \in \{1, \ldots, N\}$ und sei $f \in C^1(\Omega, \mathbb{R}^M)$ eine Funktion für die $D_i D_j f$ auf ganz Ω existiert und stetig ist. Dann existiert auch $D_j D_i f$ auf ganz Ω und es gilt für alle $x \in \Omega$:*

$$(D_i D_j f)(x) = (D_j D_i f)(x)\,.$$

Insbesondere ist auch $D_j D_i f$ auf Ω stetig.

Beweis. Da die partiellen Ableitungen komponentenweise gebildet werden, genügt es die Behauptung für den Fall $M = 1$ zu beweisen. Sei wieder $\{e_1, \ldots, e_N\}$ die kanonische Basis des $\mathbb{R}^N$. Sei $x \in \Omega$ beliebig. Da Ω offen ist, gibt es ein $\delta > 0$, so daß $x + se_i + te_j \in \Omega$ für alle $s, t \in \mathbb{R}$ mit $|s| < \delta$ und $|t| < \delta$. Nach dem Hauptsatz der Differential– und Integralrechnung (angewendet auf die Funktion $t \mapsto f(x + se_i + te_j)$) und Bemerkung 7.2 gilt für alle $(s, t) \in (-\delta, \delta) \times (-\delta, \delta)$:

$$f(x + se_i + te_j) = f(x + se_i) + \int_0^t (D_j f)(x + se_i + \tau e_j) d\tau\,.$$

Partielle Differentiation nach s ergibt wegen der Stetigkeit von $D_i D_j f$ nach Folgerung 7.8 für alle $(s, t) \in (-\delta, \delta) \times (-\delta, \delta)$:

$$(D_i f)(x + se_i + te_j) = (D_i f)(x + se_i) + \int_0^t (D_i D_j f)(x + se_i + \tau e_j) d\tau\,.$$

Für alle $s \in (-\delta, \delta)$ ist nach dem Hauptsatz der Differential– und Integralrechnung die rechte Seite (und damit auch die linke Seite) differenzierbar nach t. $(D_j D_i f)(x + se_i + te_j)$ existiert also und wir erhalten somit für alle $(s, t) \in (-\delta, \delta) \times (-\delta, \delta)$ unter Verwendung von Bemerkung 7.2:

$$(D_j D_i f)(x + se_i + te_j) = (D_i D_j f)(x + se_i + te_j)\,.$$

Speziell für $s = t = 0$ ergibt dies die Behauptung. □

7.15. Folgerung. *Sei $f \in C^\infty(\Omega, \mathbb{R}^M)$. Dann gilt für alle Permutationen π von $1, \ldots k$, alle $i_1, \ldots, i_k \in \{1, \ldots, N\}$ und alle $x \in \Omega$:*

$$\big(D_{i_{\pi(1)}} D_{i_{\pi(2)}} \ldots D_{i_{\pi(k)}} f\big)(x) = \big(D_{i_1} D_{i_2} \ldots D_{i_k} f\big)(x)\,.$$

Dies zeigt man durch vollständige Induktion nach k unter Verwendung von Satz 7.14 und der Tatsache, daß sich jede Permutation als eine Hintereinanderausführung von Transpositionen schreiben läßt.

7.16. Folgerung. *Ist $\Omega \subseteq \mathbb{R}^3$ offen und $f \in C^2(\Omega)$, so gilt*

$$\operatorname{rot}(\operatorname{grad} f) \equiv (0, 0, 0)\,.$$

Beweis. Es ist gilt nach Satz 7.14:

$$\operatorname{rot}(\operatorname{grad} f) = \Big(\frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2}, \frac{\partial^2 f}{\partial x_3 \partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_3}, \frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1}\Big) \equiv (0,0,0)\,.$$

□

7.17. Definition. Ist $\Omega \subseteq \mathbb{R}^N$ offen und $f \in C^2(\Omega, \mathbb{R}^M)$, so definiert man

$$\Delta f := \sum_{j=1}^{N} \frac{\partial^2 f}{\partial x_j^2}\,.$$

Im Fall $M = 1$ hat man insbesondere $\Delta f = \operatorname{div} \operatorname{grad} f$. Der Differentialoperator

$$\Delta := \sum_{j=1}^{N} \frac{\partial^2}{\partial x_j^2}$$

heißt *Laplace–Operator* und die partielle Differentialgleichung

$$\Delta f \equiv 0$$

die *Potentialgleichung*. Ihre Lösungen heißen *harmonische Funktionen*.

Man kann zeigen, daß die harmonischen Funktionen schon beliebig oft stetig partiell differenzierbar sind.

7.18. Beispiel. Sei $f \in C^2((0,\infty))$. Wir betrachten die Funktion $f \circ r$ mit $r(x) = |x|$ für alle $x \in \mathbb{R}^N$. Nach Beispiel 7.12 (b) ist $f \circ r$ auf $\mathbb{R}^N \setminus \{0\}$ partiell differenzierbar und es gilt $(\nabla(f \circ r))(x) = \frac{f'(x)}{r(x)} \cdot x$. Mit der Produktregel 7.11 für den Divergenz–Operator folgt: $\nabla(f \circ r)$ ist partiell differenzierbar und

$$\begin{aligned}
\Delta(f \circ r)(x) =& (\operatorname{div} \nabla(f \circ r))(x) \\
=& \Big\langle \nabla(f' \circ r)(x), \frac{1}{r(x)} x \Big\rangle + f'(r(x)) \Big(\operatorname{div} \frac{\mathrm{id}}{r}\Big)(x) \\
=& \Big\langle \frac{f''(r(x))}{r(x)} x, \frac{1}{r(x)} x \Big\rangle + f'(|x|) \frac{N-1}{|x|} \\
=& f''(r(x)) + f'(|x|) \frac{N-1}{|x|}
\end{aligned}$$

Ist $N \geq 3$ und $f(t) := t^{2-N}$ für alle $t > 0$, so folgt für die Funktion $h_N : \mathbb{R}^N \setminus \{0\} \to \mathbb{R}$ mit $h_N(x) := |x|^{2-N}$ für alle $0 \neq x \in \mathbb{R}^N$:

$$(\Delta h_N)(x) = (\Delta r^{2-N})(x) = (2-N)(1-N)|x|^{-N} + (2-N)|x|^{1-N} \frac{N-1}{|x|} = 0\,.$$

h_N ist also eine auf $\mathbb{R}^N \setminus \{0\}$ harmonische Funktion.

Im Fall $N = 2$ gilt mit $f = \log$ für die Funktion $h_2 : x \mapsto \log(|x|)$:

$$(\Delta h_2)(x) = (\Delta(\log \circ r))(x) = -|x|^{-2} + |x|^{-2} = 0.$$

Also ist auch h_2 eine auf $\mathbb{R}^2 \setminus \{0\}$ harmonische Funktion.

2. Totale Differenzierbarkeit und totales Differential

7.19. Definition. Eine auf einer offenen Menge $\Omega \subseteq \mathbb{R}^N$ definierte $\mathbb{R}^M$–wertige Funktion heißt in einem Punkt $x_0 \in \Omega$ *total differenzierbar*, falls es eine lineare Abbildung $T : \mathbb{R}^N \to \mathbb{R}^M$ gibt mit:

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \big| f(x) - f(x_0) - T(x - x_0) \big| = 0\,. \tag{7.6}$$

f heißt auf Ω total differenzierbar, falls f in allen Punkten von Ω total differenzierbar ist.

Im Spezialfall $N = 1$ ist jede lineare Abbildung $T : \mathbb{R} \to \mathbb{R}^N$ von der Gestalt $T : u \mapsto u y_T$ mit einem Vektor $y_T \in \mathbb{R}^M$. Die Aussage (7.6) ist in diesem Fall äquivalent zu:

$$\lim_{x \to x_0} \left| \frac{1}{|x - x_0|} (f(x) - f(x_0)) - y_T \right| = 0$$

d.h. zur Differenzierbarkeit von f in x_0 und zu $f'(x_0) = y_T$. Die Abbildung $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ ist dann die Tangente in $(x_0, f(x_0))$ an den Graphen von f.

Entsprechend werden wir auch im Fall $N > 1$ die Abbildung $x \mapsto f(x_0) + T(x - x_0)$ mit T wie in (7.6) die *Tangente* in $(x_0, f(x_0))$ an den Graphen von f nennen. Im Fall $M = 1$ ist hierdurch eine Hyperebene in $\mathbb{R}^{N+1}$ erklärt.

Wir schreiben im folgenden die Vektoren des $\mathbb{R}^N$ und $\mathbb{R}^M$ als Spaltenvektoren. Aus der linearen Algebra ist bekannt: Ist $T : \mathbb{R}^N \to \mathbb{R}^M$ eine lineare Abbildung, so gibt es genau eine Matrix A_T aus der Menge $M_{M \times N}(\mathbb{R})$ aller $M \times N$ Matrizen über $\mathbb{R}$,

$$A_T = \begin{pmatrix} a_{1,1} & \dots & a_{1,N} \\ \vdots & & \vdots \\ a_{M,1} & \dots & a_{M,N} \end{pmatrix}$$

mit

$$Tu = A_T u = \begin{pmatrix} a_{1,1} & \dots & a_{1,N} \\ \vdots & & \vdots \\ a_{M,1} & \dots & a_{M,N} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^N a_{1,k} u_k \\ \vdots \\ \sum_{k=1}^N a_{M,k} u_k \end{pmatrix} \quad \text{für alle } u = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \in \mathbb{R}^N.$$

Ist also $f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix} : \Omega \to \mathbb{R}^M$ eine in einem Punkt $x_0 \in \Omega$ total differenzierbare Funktion und ist T wie in (7.6), so gilt für die zugehörige Matrix $A_T = \big(a_{j,k}\big)_{\substack{j=1,\dots,M \\ k=1,\dots,N}} \in M_{M \times N}(\mathbb{R})$ nach (7.6)

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \left| f(x) - f(x_0) - \Big(\sum_{k=1}^N a_{M,k}(x_k - x_{0,k}) \Big)_{j=1}^M \right| = 0$$

und daher auch

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \left| f_j(x) - f_j(x_0) - \sum_{k=1}^N a_{j,k}(x_k - x_{0,k}) \right| = 0$$

für $j = 1, \dots, M$. Aus der totalen Differenzierbarkeit von f in x_0 folgt also die totale Differenzierbarkeit von $f_1, \dots, f_M$ in x_0. Sind umgekehrt die Funktionen $f_1, \dots, f_M$ in x_0

total differenzierbar, so gibt es lineare Abbildungen $T_j : \mathbb{R}^N \to \mathbb{R}$ mit

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \big| f_j(x) - f_j(x_0) - T_j(x - x_0) \big| = 0$$

für $j = 1, \ldots, M$. Zu T_j gibt es eindeutig bestimmte $a_{j,1}, \ldots, a_{j,N} \in \mathbb{R}$ mit

$$T_j u = (a_{j,1}, \ldots, a_{j,N}) \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \quad \text{für alle} \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \in \mathbb{R}^N.$$

Es folgt also für $j = 1, \ldots, M$

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \Big| f_j(x) - f_j(x_0) - \sum_{k=1}^{N} a_{j,k}(x_k - x_{0,k}) \Big| = 0$$

und somit auch

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} \Big| f(x) - f(x_0) - \Big(\sum_{k=1}^{N} a_{j,k}(x_k - x_{0,k}) \Big)_{j=1}^{M} \Big| =$$

$$= \lim_{x \to x_0} \frac{1}{|x - x_0|} |f(x) - f(x_0) - A(x - x_0)| = 0,$$

wobei $A = \big(a_{j,k}\big)_{\substack{j=1,\ldots,M \\ k=1,\ldots,N}} \in M_{M \times N}(\mathbb{R})$. Die durch A

$$T : \mathbb{R}^N \to \mathbb{R}^M, \qquad u \mapsto T(u) := Au$$

definierte lineare Abbildung $\mathbb{R}^N$ nach $\mathbb{R}^M$ erfüllt also (7.6) und es folgt die totale Differenzierbarkeit von f in x_0. Wir halten also fest:

7.20. Lemma. *Eine Funktion* $f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix} : \Omega \to \mathbb{R}^M$ *ist genau dann in einem Punkt x_0 der offenen Menge Ω total differenzierbar, wenn die $\mathbb{R}$–wertigen Funktionen $f_1, \ldots, f_M$ in x_0 total differenzierbar sind.*

7.21. Satz. *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei* $f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix} : \Omega \to \mathbb{R}^M$ *in $x_0 \in \Omega$ total differenzierbar, so daß also*

$$\lim_{x \to x_0} \frac{1}{|x - x_0|} |f(x) - f(x_0) - A(x - x_0)| = 0 \tag{7.7}$$

gilt mit einer Matrix $A = \big(a_{j,k}\big)_{\substack{j=1,\ldots,M \\ k=1,\ldots,N}} \in M_{M \times N}(\mathbb{R})$. *Dann gilt:*

(a) *f ist in x_0 stetig.*
(b) *f ist in x_0 partiell differenzierbar und man hat*

$$a_{j,k} = \frac{\partial f_j}{\partial x_k} \quad \textit{für alle } j = 1, \ldots, M,\ k = 1, \ldots, N.$$

Insbesondere ist also die Matrix A aus (7.7) und somit auch die lineare Abbildung T aus (7.6) eindeutig bestimmt. Man nennt T dann das *totale Differential von f in x_0* und schreibt

$$T = (Df)(x_0).$$

Die zu T gehörige Matrix heißt die *Jacobi–Matrix*[1] und wird mit $J_f(x_0)$ bezeichnet. Es ist also

$$J_f(x_0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dots & \dfrac{\partial f_1}{\partial x_N} \\ \vdots & & \vdots \\ \dfrac{\partial f_M}{\partial x_1} & \dots & \dfrac{\partial f_M}{\partial x_N} \end{pmatrix}.$$

Beweis. (a) Aus (7.7) folgt insbesondere auch

$$f(x) - f(x_0) = f(x) - f(x_0) - A(x - x_0) + A(x - x_0) \to 0 \quad \text{für } x \to x_0,$$

da

$$A(x - x_0) = \Big(\sum_{k=1}^{N} a_{j,k}(x_k - x_{0,k}) \Big)_{j=1}^{M} \to 0 \quad \text{für } x \to x_0.$$

f ist also in x_0 stetig.

(b) Für $j = 1, \dots, M$, $k = 1, \dots, N$ gilt (mit $x = x_0 + te_k$, e_k der k–te Einheitsvektor)

$$\lim_{t\to 0} \frac{1}{|t|} \Big| f_j(x_0 + te_k) - f_j(x_0) - t \sum_{l=1}^{N} a_{j,l}\delta_{l,k} \Big| = 0$$

wobei

$$\delta_{l,k} := \begin{cases} 0 & \text{für } l \neq k \\ 1 & \text{für } l = k \end{cases}$$

das *Kroneckersymbol* bezeichnet. Es folgt also

$$\lim_{t\to 0} \Big| \frac{f_j(x_0 + te_k) - f_j(x_0)}{t} - a_{j,k} \Big| = 0\,.$$

f_j ist daher in x_0 nach der k–ten Variablen partiell differenzierbar und es gilt

$$\frac{\partial f_j}{\partial x_k}(x_0) = a_{j,k} \quad \text{für alle } j = 1, \dots, M, k = 1, \dots, N\,.$$

□

Offensichtlich ist (mit $x = x_0 + \xi$) die Aussage (7.6) äquivalent zu

(7.8) $$\lim_{\xi\to 0} \frac{1}{|\xi|} |f(x_0 + \xi) - f(x_0) - T\xi| = 0$$

und (7.7) ist äquivalent zu

(7.9) $$\lim_{\xi\to 0} \frac{1}{|\xi|} |f(x_0 + \xi) - f(x_0) - A\xi| = 0\,.$$

Wir werden dies im folgenden häufig verwenden.

Natürlich kann die Jacobimatrix für alle in x_0 partiell differenzierbaren Funktionen $f : \Omega \to \mathbb{R}^M$ gebildet werden. Das Beispiel 7.3 einer in 0 unstetigen aber in 0 partiell differenzierbaren Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ zeigt jedoch wegen Satz 7.22, daß aus der partiellen Differenzierbarkeit in x_0 nicht die totale Differenzierbarkeit in x_0 folgt. Es gilt jedoch:

[1] Carl Gustav Jacob Jacobi (10.12.1804–18.2.1851).

7.22. SATZ. *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei $f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix} : \Omega \to \mathbb{R}^M$ auf Ω partiell differenzierbar. Sind die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}$ für alle $j = 1, \ldots, M$ und alle $k = 1, \ldots, N$ in x_0 stetig, so ist f in x_0 total differenzierbar.*

BEWEIS. Da f genau dann in x_0 total differenzierbar ist, wenn die Komponentenfunktionen $f_1, \ldots, f_n$ in x_0 total differenzierbar sind, genügt es nach Lemma 7.20, den Fall $M = 1$ zu betrachten. Sei also nun f als reellwertig vorausgesetzt.

Da Ω offen ist gibt es ein $\delta > 0$ mit $U_\delta(x_0) = \{x \in \mathbb{R}^N\,;\, |x - x_0| < \delta\} \subseteq \Omega$. Wir definieren für $j = 0, 1, \ldots, N$ und alle $\xi \in \mathbb{R}^N$ mit $|\xi| < \delta$:

$$u_j(\xi) := x_0 + \sum_{\nu=1}^{j} \xi_\nu e_\nu \,,$$

wobei e_ν den ν–ten Einheitsvektor in $\mathbb{R}^N$ bezeichne. Dann liegen die Punkte $u_0(\xi) = x_0, u_1(\xi), \ldots, u_N(\xi) = x_0 + \xi$ alle in $U_\delta(x_0)$ und $u_j(\xi)$ unterscheidet sich von $u_{j-1}(\xi)$ nur in der j–ten Komponente, $u_j(\xi) = u_{j-1}(\xi) + \xi_j e_j$. Nach dem Mittelwertsatz der Differentialrechnung, angewendet auf die Funktion $t \mapsto f(u_{j-1}(\xi) + t\xi_j e_j)$ gibt es ein $\theta_j(\xi) \in (0, 1)$ mit

$$f(u_j(\xi)) - f(u_{j-1}(\xi)) = \xi_j (D_j f)(v_j(\xi)) \quad \text{mit } v_j(\xi) := u_{j-1}(\xi) + \theta_j(\xi)\xi_j e_j \,.$$

Wegen $u_0(\xi) = x_0$ und $u_N(\xi) = x_0 + \xi$ folgt also

$$f(x_0 + \xi) - f(x_0) = \sum_{j=1}^{N} f(u_j(\xi)) - f(u_{j-1}(\xi)) = \sum_{j=1}^{N} \xi_j (D_j f)(v_j(\xi)) \,. \tag{7.10}$$

Für $\xi \to 0$ gilt $u_j(\xi) \to x_0$ und $v_j(\xi) \to x_0$ und daher (wegen der Stetigkeit von $D_j f$ in x_0):

$$\lim_{\xi \to 0} (D_j f)(v_j(\xi)) = (D_j f)(x_0) \quad \text{für } j = 1, \ldots, N \,.$$

Wir erhalten

$$\begin{aligned}
0 \leq & \frac{1}{|\xi|} \Big| f(x_0 + \xi) - f(x_0) - \sum_{j=1}^{N} (D_j f)(x_0)\xi_j \Big| \leq \\
\leq & \frac{1}{|\xi|} \Big| f(x_0 + \xi) - f(x_0) - \sum_{j=1}^{N} (D_j f)(v_j(\xi))\xi_j \Big| + \frac{1}{|\xi|} \Big| \sum_{j=1}^{N} \big((D_j f)(v_j(\xi)) - (D_j f)(x_0) \big) \xi_j \Big| \leq \\
\leq & 0 + \Big(\sum_{j=1}^{N} \big((D_j f)(v_j(\xi)) - (D_j f)(x_0) \big)^2 \Big)^{1/2} \to 0 \quad \text{for } \xi \to 0 \,,
\end{aligned}$$

wobei bei der letzten Ungleichung (7.10) und die Cauchy–Schwarzsche Ungleichung verwendet wurden. Also ist (7.8) mit der linearen Abbildung

$$T : \mathbb{R}^N \to \mathbb{R} \,, \qquad \xi \mapsto T(\xi) := \sum_{j=1}^{N} (D_j f)(x_0)\xi_j$$

erfüllt und es folgt die totale Differenzierbarkeit von f in x_0. □

7.23. FOLGERUNG. *Jede auf einer offenen Menge $\Omega \subseteq \mathbb{R}^N$ stetig partiell differenzierbare Funktion $f : \Omega \to \mathbb{R}^M$ ist auf Ω total differenzierbar.*

Der Beweis der folgenden Rechenregeln ist inzwischen Routine und muß daher hier nicht ausgeführt werden.

7.24. Lemma. *Sei $\Omega \subseteq \mathbb{R}^N$ offen und seinen $f, g : \Omega \to \mathbb{R}^M$ zwei in dem Punkt $x_0 \in \Omega$ total differenzierbare Funktionen.*

(a) *Für alle $\alpha, \beta \in \mathbb{R}$ ist die Funktion $\alpha f + \beta g$ in x_0 total differenzierbar und es gilt*

$$(D(\alpha f + \beta g))(x_0) = \alpha(Df)(x_0) + \beta(Dg)(x_0).$$

(b) *Ist $M = 1$, so ist auch die Funktion fg in x_0 total differenzierbar und es gilt*

$$(D(fg))(x_0) = f(x_0)(Dg)(x_0) + g(x_0)(Df)(x_0).$$

7.25. Beispiele. (a) Sei $T : \mathbb{R}^N \to \mathbb{R}^M$ eine lineare Abbildung. Dann ist T auf $\mathbb{R}^N$ total differenzierbar und es gilt $(DT)(a) = T$ für alle $a \in \mathbb{R}^N$.

(b) Wir können die Menge $M_{N\times N}(\mathbb{R})$ der $(N \times N)$–Matrizen identifizieren mit dem $\mathbb{R}^{N^2}$. Der euklidische Betrag stimmt mit der in den Übungen eingeführten Frobeniusnorm $\|\cdot\|_2$ überein. Die Abbildung

$$f : M_{N\times N}(\mathbb{R}) \to M_{N\times N}(\mathbb{R}), \qquad A \mapsto f(A) := A^2,$$

ist auf $M_{N\times N}(\mathbb{R})$ total differenzierbar und es gilt für alle $A, C \in M_{N\times N}(\mathbb{R})$:

$$((Df)(A))(C) = AC + CA.$$

Beweis. (a) Dies ist unmittelbar klar, denn für alle $a \in \mathbb{R}^N$ gilt:

$$\lim_{\xi\to 0} \frac{1}{|\xi|}\big|T(a+\xi) - T(a) - T(\xi)\big| = 0.$$

(b) Für alle $A \in M_{N\times N}(\mathbb{R})$ und $0 \neq C \in M_{N\times N}(\mathbb{R})$ gilt:

$$\begin{aligned} 0 \leq \frac{1}{\|C\|_2}\big\|f(A+C) - f(A) - AC - CA\big\|_2 &= \frac{1}{\|C\|_2}\big\|(A+C)^2 - A^2 - AC - CA\big\|_2 \\ &= \frac{1}{\|C\|_2}\big\|C^2\big\|_2 \leq \frac{\|C\|_2^2}{\|C\|_2} = \|C\|_2 \to 0 \end{aligned}$$

für $C \to 0$. □

7.26. Satz (Kettenregel). *Seien $\Omega_1 \subseteq \mathbb{R}^N$ und $\Omega_2 \subseteq \mathbb{R}^M$ offen und seien $g : \Omega_1 \to \mathbb{R}^M$ und $f : \Omega_2 \to \mathbb{R}^L$ zwei Funktionen mit $g(\Omega_1) \subseteq \Omega_2$. Ist g in einem Punkt $x_0 \in \Omega_1$ und f in dem Punkt $g(x_0) \in \Omega_2$ total differenzierbar, so ist auch die Hintereinanderausführung $f \circ g : \Omega_1 \to \mathbb{R}^L$, $x \mapsto (f \circ g)(x) := f(g(x))$, in x_0 total differenzierbar und es gilt*

$$(D(f \circ g))(x_0) = (Df)(g(x_0)) \circ (Dg)(x_0).$$

Für die Jacobimatrix $J_{f\circ g}(x_0)$ erhält man also

$$J_{f\circ g}(x_0) = J_f(g(x_0))J_g(x_0).$$

Beweis. Wir setzen $A := (Dg)(x_0)$, $B := (Df)(g(x_0))$ sowie für $\xi \in \mathbb{R}^N$

$$\varphi(\xi) := g(x_0 + \xi) - g(x_0) - A(\xi) \quad \text{für } \xi \in \mathbb{R}^N$$

und

$$\psi(\eta) := f(g(x_0) + \eta) - f(g(x_0)) - B(\eta) \quad \text{für } \eta \in \mathbb{R}^M.$$

Wegen der totalen Differenzierbarkeit von g in x_0 und von f in $g(x_0)$ gilt

$$\frac{|\varphi(\xi)|}{|\xi|} \to 0 \quad \text{für } \xi \to 0 \qquad \text{sowie} \qquad \frac{|\psi(\eta)|}{|\eta|} \to 0 \quad \text{für } \eta \to 0.$$

Also ist die durch

$$\varphi_0(\xi) := \begin{cases} \dfrac{1}{|\xi|}\varphi(\xi) & \text{für } 0 \neq \xi \in \mathbb{R}^N \\ 0 & \text{für } \xi = 0 \end{cases}$$

definierte Funktion $\varphi_0 : \mathbb{R}^N \to \mathbb{R}^M$ stetig in 0 und auch $\psi_0 : \mathbb{R}^M \to \mathbb{R}^L$ mit

$$\psi_0(\eta) := \begin{cases} \dfrac{1}{|\eta|}\psi(\eta) & \text{für } 0 \neq \eta \in \mathbb{R}^M \\ 0 & \text{für } \eta = 0 \end{cases}$$

ist stetig in 0. Die Kettenregel folgt, wenn wir zeigen können, daß mit

$$\chi(\xi) := f(g(x_0 + \xi)) - f(g(x_0)) - B(A(\xi)) \qquad \text{für } \xi \in \mathbb{R}^N$$

gilt

$$\frac{|\chi(\xi)|}{|\xi|} \to 0 \qquad \text{für } \xi \to 0. \tag{7.11}$$

Wegen $g(x_0 + \xi) - g(x_0) = A(\xi) + \varphi(\xi)$ und

$$\begin{aligned} f\big(g(x_0 + \xi)\big) =& f\big(g(x_0) + (g(x_0 + \xi) - g(x_0))\big) \\ =& f(g(x_0)) + B\big(g(x_0 + \xi) - g(x_0)\big) + \psi\big(g(x_0 + \xi) - g(x_0)\big) \\ =& f\big(g(x_0)\big) + B\big(A(\xi) + \varphi(\xi)\big) + \psi\big(g(x_0 + \xi) - g(x_0)\big) \\ =& f\big(g(x_0)\big) + B(A(\xi)) + B\big(\varphi(\xi)\big) + \psi\big(g(x_0 + \xi) - g(x_0)\big) \end{aligned}$$

folgt

$$\begin{aligned} \chi(\xi) =& B\big(\varphi(\xi)\big) + \psi\big(g(x_0 + \xi) - g(x_0)\big) \\ =& |\xi| \cdot B\big(\varphi_0(\xi)\big) + \big|A(\xi) + |\xi|\varphi_0(\xi)\big|\psi_0\big(g(x_0 + \xi) - g(x_0)\big) \end{aligned}$$

und hieraus mit der Dreiecksungleichung für $|\xi| > 0$ unter Verwendung der in den Übungen eingeführten Matrixnorm $\|\cdot\|$:

$$\begin{aligned} 0 \le \frac{|\chi(\xi)|}{|\xi|} \le& \big|B\big(\varphi_0(\xi)\big)\big| + \big|A(|\xi|^{-1}\xi) + \varphi_0(\xi)\big| \cdot \big|\psi_0\big(g(x_0 + \xi) - g(x_0)\big)\big| \\ \le& \|B\| \cdot |\varphi_0(\xi)| + (\|A\| + |\varphi_0(\xi)|) \cdot \big|\psi_0\big(g(x_0 + \xi) - g(x_0)\big)\big| \,. \end{aligned}$$

Wegen der Stetigkeit von g in x_0 und der Funktionen φ_0 und ψ_0 im Nullpunkt folgt (7.11) und damit die Behauptung. □

7.27. Folgerung. *In der Situation von Satz 7.26 ist die Funktion $f \circ g$ in x_0 partiell differenzierbar und es gilt*

$$\frac{\partial(f \circ g)}{\partial x_j}(x_0) = \sum_{k=1}^{M} \frac{\partial f}{\partial y_k}(g(x_0))\frac{\partial g_k}{\partial x_j}(x_0) \qquad \textit{für } j = 1, \dots, N\,.$$

Beweis. Dies folgt unmittelbar aus Satz 7.26 durch Berechnung der j–ten Spalte von $J_{f\circ g}(x_0) = J_f(g(x_0))J_g(x_0)$. □

7.28. Folgerung. *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei $f : \Omega \to \mathbb{R}$ eine in dem Punkt $x_0 \in \Omega$ total differenzierbare Funktion. Dann existiert für alle $u \in \mathbb{R}^N$ mit $|u| = 1$ die Richtungsableitung $(D_u f)(x_0)$ von f in x_0 und es gilt*

$$(D_u f)(x_0) = \langle u, (\nabla f)(x_0)\rangle\,.$$

Beweis. Für alle $t \in \mathbb{R}$ mit $x + tu \in \Omega$ definieren wir $g(t) := x_0 + tu$. Da Ω offen ist, ist g in einer Umgebung von 0 definiert und offensichtlich differenzierbar. Nach Definition der Richtungsableitung gilt

$$(D_u f)(x_0) = \lim_{t\to 0} \frac{f(x_0 + tu) - f(x_0)}{t} = \lim_{t\to 0} \frac{f(g(t)) - f(g(0))}{t} = (f \circ g)'(0)\,.$$

Der letzte Grenzwert existiert nach der Kettenregel und es folgt

$$(D_u f)(x_0) = J_{f\circ g}(0) = J_f(g(0))J_g(0) = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(x_0)u_j = \langle u, (\nabla f)(x_0)\rangle\,.$$

□

Ist $(\nabla f)(x_0) \neq 0$, so ist der Winkel φ zwischen u und $(\nabla f)(x_0)$ gegeben durch

$$\cos(\varphi) = \frac{\langle u, (\nabla f)(x_0)\rangle}{|u| \cdot |(\nabla f)(x_0)|}\,.$$

Die Richtungsableitung wird also am größten für $\varphi = 0$. Der Gradient gibt somit die Richtung des steilsten Anstiegs an.

3. Der Mittelwertsatz und der Taylorsche Satz in mehreren Veränderlichen

In Analogie zum Fall einer reellen Veränderlichen (vergl. Def. 5.15) definieren wir: Sei $\emptyset \neq \Omega \subseteq \mathbb{R}^N$ und sei $f : \Omega \to \mathbb{R}$ eine reellwertige Funktion auf Ω. Ein Punkt $x_0 \in \Omega$ heißt *lokale Maximumstelle* (bzw. *lokale Minimumstelle*) von f, falls es ein $\varepsilon > 0$ gibt mit

$$\forall x \in \Omega \cap U_\varepsilon(x_0) \setminus \{x_0\} : f(x) \leq f(x_0) \quad (\text{bzw. } f(x) \geq f(x_0)). \tag{7.12}$$

$f(x_0)$ heißt dann ein *lokales Maximum* (bzw. *lokales Minimum*) von f. Ist x_0 lokale Maximums– oder Minimumsstelle von f, so nennen wir x_0 auch eine *lokale Extremstelle* und $f(x_0)$ ein *lokales Extremum* von f. Gilt in (7.12) das $<$–Zeichen bzw. das $>$–Zeichen, so nennen wir x_0 eine *isolierte* lokale Maximum– bzw. Minimumstelle von f.

7.29. Satz. *Sei $f : \Omega \to \mathbb{R}$ eine auf einer offenen Menge $\Omega \subseteq \mathbb{R}^N$ definierte reellwertige Funktion, die in dem Punkt $x_0 \in \Omega$ eine lokale Extremstelle besitze. Ist f in x_0 partiell differenzierbar, so ist $(\nabla F)(x_0) = 0$.*

Beweis. Da Ω offen ist gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq \Omega$. Für $j = 1, \ldots, N$ betrachten wir die Funktion

$$\varphi_j : (-\varepsilon, \varepsilon) \to \mathbb{R}\,, \qquad t \mapsto \varphi_j(t) := f(x_0 + te_j)\,,$$

wobei e_j wieder den j–ten Einheitsvektor in $\mathbb{R}^N$ bezeichne. Da f nach Voraussetzung in x_0 partiell differenzierbar ist, ist φ_j in 0 differenzierbar mit $\varphi_j'(0) = (D_j f)(x_0)$. Aus der Tatsache, daß f in x_0 eine lokale Extremstelle hat, folgt daß 0 eine lokale Extremstelle von φ_j ist. Nach Satz 5.16 muß also für $j = 1, \ldots, N$ gelten: $0 = \varphi_j(0) = (D_j f)(x_0)$. □

7.30. Satz (Mittelwertsatz der Differentialrechnung in mehrereren Veränderlichen). *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei $f \in C^1(\Omega, \mathbb{R}^M)$. Sind $x \in \Omega$ und $\xi \in \mathbb{R}^N$ so gegeben, daß die Strecke $\{x + t\xi\,;\, 0 \leq t \leq 1\}$ ganz in Ω enthalten ist, so gilt*

$$f(x + \xi) - f(x) = \int_0^1 J_f(x + t\xi)\xi\, dt \tag{7.13}$$

und daher auch[2]

$$|f(x+\xi)-f(x)| \le \sup_{t\in[0,1]} \|J_f(x+t\xi)\| \cdot |\xi| \,. \tag{7.14}$$

Beweis. Die Funktion

$$g:[0,1]\to\mathbb{R}^M\,, \qquad t\mapsto g(t):=f(x+t\xi)$$

ist nach der Kettenregel auf $[0,1]$ differenzierbar und mit $h(t):=x+t\xi$ gilt wegen $J_h(t)=\xi$:

$$g'(t)=J_g(t)=J_f(x+t\xi)J_h(t)=J_f(x+t\xi)\xi \qquad \text{für alle } t\in[0,1]\,.$$

Es folgt daher

$$f(x+\xi)-f(x)=g(1)-g(0)=\int_0^1 g'(t)dt=\int_0^1 J_f(x+t\xi)\xi dt$$

und hieraus

$$|f(x+\xi)-f(x)| \le \int_0^1 |J_f(x+t\xi)\xi|dt \le \int_0^1 \|J_f(x+t\xi)\|\cdot|\xi|dt \le \sup_{t\in[0,1]} \|J_f(x+t\xi)\|\cdot|\xi|\,.$$

□

Im Spezialfall $M=1$ gibt es nach dem Mittelwertsatz in einer Veränderlichen ein $\theta\in(0,1)$, so daß mit den Bezeichnungen aus dem obigen Beweis gilt:

$$\begin{aligned} f(x+\xi)-f(x) =& g(1)-g(0)=g'(\theta)=J_f(x+\theta\xi)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_N\end{pmatrix} \\ =& \sum_{j=1}^N \frac{\partial f}{\partial x_j}(x+\theta\xi)\xi_j = \langle \xi, (\nabla f)(x+\theta\xi)\rangle\,. \end{aligned}$$

Damit erhalten wir:

7.31. Satz (Mittelwertsatz der Differentialrechnung in mehrereren Veränderlichen für reellwertige Funktionen). *Sei $\Omega\subseteq\mathbb{R}^N$ offen und sei $f\in C^1(\Omega,\mathbb{R}^M)$. Sind $x\in\Omega$ und $\xi\in\mathbb{R}^N$ so gegeben, daß die Strecke $\{x+t\xi\,;\,0\le t\le 1\}$ ganz in Ω enthalten ist, so gibt es ein $\theta\in(0,1)$ mit*

$$f(x+\xi)-f(x)=\langle \xi, (\nabla f)(x+\theta\xi)\rangle\,. \tag{7.15}$$

Eine Menge $\Omega\subseteq\mathbb{R}^n$ heißt *konvex*, falls für alle Punkte $x,y\in\Omega$ die Verbindungstrecke $\{x+t(y-x)\,;\,0\le t\le 1\}$ in Ω enthalten ist.

7.32. Folgerung. *Sei Ω eine offene, konvexe Teilmenge des $\mathbb{R}^N$. Ist $f\in C^1(\Omega,\mathbb{R}^M)$ eine Funktion mit $(Df)(x)=0$ für alle $x\in\Omega$, so ist f auf Ω konstant.*

[2] Die Matrixnorm $\|\cdot\|$ wurde in den Übungen eingeführt. Wegen der stetigen partiellen Differenzierbarkeit von f auf Ω ist die Abbildung $t\mapsto J_f(x+t\xi)$ auf $[0,1]$ stetig. Es gilt also für alle $t\in[0,1]$ für $s\to t$ unter Verwendung der Ungleichung zwischen Matrixnorm und Frobeniusnorm:

$$0\le \big|\|J_f(x+t\xi)\|-\|J_f(x+s\xi)\|\big| \le \|J_f(x+t\xi)-J_f(x+s\xi)\| \le \sqrt{N}\|J_f(x+t\xi)-J_f(x+s\xi)\|_2 \to 0\,.$$

Insbesondere ist also auch die Funktion $t\mapsto\|J_f(x+t\xi)\|$ auf $[0,1]$ stetig und somit

$$\sup_{t\in[0,1]}\|J_f(x+t\xi)\|<\infty\,.$$

Beweis. $(Df)(x) = 0$ für alle $x \in \Omega$ ist äquivalent zu $J_f(x) \equiv 0$ auf Ω. Sind $u, v \in \Omega$ beliebig, so gilt wegen der Konvexität von Ω auch $\{u + t(v-u)\,;\, 0 \le t \le 1\} \subset \Omega$ und wir erhalten aus Satz 7.30 (mit $x = u$ und $\xi = v - u$):

$$f(v) - f(u) = \int_0^1 J_f(u + t(v-u))(v-u)dt = 0\,.$$

f muß also auf Ω konstant sein. □

Allgemeiner gilt:

7.33. Folgerung. *Sei $\Omega \subseteq \mathbb{R}^N$ offen mit der Eigenschaft, daß sich je zwei Punkte aus Ω durch einen ganz in Ω verlaufenden Polygonzug verbinden lassen. Ist $f \in C^1(\Omega, \mathbb{R}^M)$ eine Funktion mit $(Df)(x) = 0$ für alle $x \in \Omega$, so ist f auf Ω konstant.*

Beweis. Seien $u, v \in \Omega$ beliebig. Nach Voraussetzung gibt es einen ganz in Ω verlaufenden Polygonzug der u und v verbindet. Seien $u_0 = u, u_1, \dots, u_{n-1}, u_n = v$ die Ecken dieses Polygonzuges. Da die Strecke von u_{j-1} bis u_j ganz in Ω verläuft können wir Satz 7.30 (mit $x = u_{j-1}$ und $\xi = u_j - u_{j-1}$) anwenden und erhalten wie im Beweis zu Folgerung 7.32:

$$f(u_j) - f(u_{j-1}) = 0 \qquad \text{für } j = 1, \dots, n$$

und somit $f(u) = f(u_0) = f(u_1) = \cdots = f(u_{n-1}) = f(u_n) = f(v)$. Also ist f auf Ω konstant. □

Um bei der Formulierung des Taylorschen Satzes den Schreibaufwand in Grenzen zu halten und die Übersichtlichkeit zu verbessern, führen wir die abkürzende *Multiindexschreibweise* ein:

7.34. Bezeichnungen (Multiindexschreibweise). Die Elemente von $\mathbb{N}_0^N$ nennen wir *Multiindizes*. Für einen Multiindex $\alpha \in \mathbb{N}_0^N$ definieren wir die Ordnung von α durch

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_N$$

(*Vorsicht: Nicht mit dem euklidischen Betrag verwechseln!*) und

$$\alpha! := \prod_{j=1}^N \alpha_j!$$

sowie

$$D^\alpha := D_1^{\alpha_1} D_2^{\alpha_2} \dots D_N^{\alpha_N} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_N^{\alpha_N}}\,.$$

Ist $\mathcal{R}$ ein kommutativer Ring mit Einselement 1 und $h = (h_1, \dots, h_N) \in \mathcal{R}^N$, so sei

$$h^\alpha := \prod_{j=1}^N h_j^{\alpha_j} = h_1^{\alpha_1} h_2^{\alpha_2} \dots h_N^{\alpha_N}\,.$$

Hierbei vereinbaren wir $h_j^0 := 1$ im Fall $\alpha_j = 0$.

7.35. Satz (Polynomischer Satz). *Sei $\mathcal{R}$ ein kommutativer Ring mit Einselement 1 und sei $h = (h_1, \dots, h_N) \in \mathcal{R}^N$. Dann gilt für alle $k \in \mathbb{N}$:*

$$(h_1 + \cdots + h_N)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} h^\alpha. \tag{7.16}$$

Hierbei bedeutet $\sum_{|\alpha|=k}$, daß über alle Multiindizes $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}_0^N$ mit $|\alpha| = \alpha_1 + \ldots + \alpha_N = k$ summiert werden soll.

Beweis. Wir führen den Beweis durch vollständige Induktion nach k. Für $k=1$ ist die Aussage (7.16) offensichtlich richtig. Sei nun $p \in \mathbb{N}$ so, daß (7.16) für $1 \leq k \leq p$ erfüllt ist. Es ist zu zeigen, daß (7.16) dann auch für $k = p+1$ gilt.

Durch Anwendung der Induktionsvoraussetzung und Ausmultiplizieren und Ordnen nach den Potenzen h^β erhalten wir eine Darstellung der Form:

$$(h_1 + \cdots + h_N)^{p+1} = (h_1 + \cdots + h_N) \sum_{|\alpha|=p} \frac{p!}{\alpha!} h^\alpha = \sum_{|\beta|=p+1} C_\beta h^\beta. \tag{7.17}$$

Hierbei ist für alle $\beta \in \mathbb{N}_0^N$ mit $|\beta| = p+1$:

$$C_\beta h^\beta = \sum_{\substack{1 \leq j \leq N \\ \beta_j \geq 1}} h_j \frac{p!}{\beta_1! \dots \beta_{j-1}!(\beta_j - 1)!\beta_{j+1}! \dots \beta_N!} h^{\beta - e_j},$$

wobei e_j der j–te Einheitsvektor sei. Es folgt

$$\begin{aligned} C_\beta h^\beta &= \sum_{\substack{1 \leq j \leq N \\ \beta_j \geq 1}} \frac{p!\beta_j}{\beta!} h^\beta = \sum_{j=1}^N \frac{p!\beta_j}{\beta!} h^\beta = |\beta| \cdot \frac{p!}{\beta!} h^\beta \\ &= \frac{(p+1)!}{\beta!} h^\beta \end{aligned}$$

Setzen wir dies in (7.17) ein so folgt die Gültigkeit von (7.16) für $k = p+1$ und damit die Behauptung. □

7.36. Bemerkung. Sei $\Omega \subseteq \mathbb{R}^N$ offen und $\xi \in \mathbb{R}^N$. Der Differentialoperator

$$\langle \xi, \nabla \rangle = \sum_{j=1}^N \xi_j D_j = \sum_{j=1}^N \xi_j \frac{\partial}{\partial x_j}$$

sei definiert durch

$$\langle \xi, \nabla \rangle f := \sum_{j=1}^N \xi_j D_j f = \sum_{j=1}^N \xi_j \frac{\partial f}{\partial x_j}$$

d.h. durch

$$(\langle \xi, \nabla \rangle f)(x) := \sum_{j=1}^N \xi_j (D_j f)(x) = \sum_{j=1}^N \xi_j \frac{\partial f}{\partial x_j}(x)$$

für alle $f \in C^1(\Omega, \mathbb{R})$, $x \in \Omega$. Wegen der Vertauschbarkeit der Differentiationsreihenfolge (vergl. die Folgerungen 7.15 und 7.16) erhalten wir mit einem zu dem Beweis von Satz 7.16 analogen Beweis durch vollständige Induktion nach p:

$$\langle \xi, \nabla \rangle^k f := \Big(\sum_{j=1}^N \xi_j D_j \Big)^k f = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \xi^\alpha D^\alpha f \tag{7.18}$$

für $1 \leq k \leq p \in \mathbb{N}$, $f \in C^k(\Omega, \mathbb{R})$. Man ersetze im Beweis zu Satz 7.16 einfach h_j durch $\xi_j D_j$. Wir definieren noch $\langle \xi, \nabla \rangle^0 f := f$.

7.37. Satz (Satz von Taylor in mehreren Veränderlichen für skalarwertige Funktionen). *Sei $\Omega \subseteq \mathbb{R}^N$ offen, $p \in \mathbb{N}_0$ und $f \in C^{p+1}(\Omega, \mathbb{R})$. Sind $x_0 \in \Omega$ und $\xi \in \mathbb{R}^N$ gegeben, so daß die Strecke $\{x_0 + t\xi\,;\, t \in [0,1]\}$ in Ω enthalten ist, so gilt*

$$\begin{aligned} f(x_0+\xi) &= \sum_{k=0}^{p} \frac{1}{k!}(\langle \xi, \nabla\rangle^k f)(x_0) + R_p(x_0,\xi) \\ &= \sum_{k=0}^{p} \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!}(D^\alpha f)(x_0) + R_p(x_0,\xi)\,. \end{aligned} \tag{7.19}$$

Für das Restglied $R_p(x_0,\xi)$ gilt mit einem $\theta \in (0,1)$:

$$\begin{aligned} R_p(x_0,\xi) &= \frac{1}{(p+1)!}(\langle \xi, \nabla\rangle^{p+1} f)(x_0+\theta\xi) \\ &= \sum_{|\alpha|=p+1} \frac{\xi^\alpha}{\alpha!}(D^\alpha f)(x_0+\theta\xi) \end{aligned} \tag{7.20}$$

und die Integraldarstellung

$$\begin{aligned} R_p(x_0,\xi) &= \frac{1}{p!}\int_0^1 (1-s)^p(\langle \xi, \nabla\rangle^{p+1} f)(x_0+s\xi)ds \\ &= (p+1) \sum_{|\alpha|=p+1} \frac{\xi^\alpha}{\alpha!}\int_0^1 (1-s)^p(D^\alpha f)(x_0+s\xi)ds\,. \end{aligned} \tag{7.21}$$

Beweis. Wir definieren die Hilfsfunktion $\varphi : [0,1] \to \mathbb{R}$ durch $\varphi(t) := f(x_0+t\xi)$ für alle $t \in [0.1]$. Durch vollständige Induktion zeigt man mit der Kettenregel: $\varphi \in C^{p+1}([0,1],\mathbb{R})$ und

$$\varphi^{(k)}(t) = (\langle \xi, \nabla\rangle^k f)(x_0+t\xi) \tag{7.22}$$

für alle $t \in [0,1]$, $0 \leq k \leq p+1$. Wenden wir nun den Satz von Taylor in einer Veränderlichen mit der Lagrange–Restglieddarstellung bzw. der Integralrestglieddarstellung auf φ an, so erhalten wir für ein $\theta \in (0,1)$ (unter Verwendung von (7.22)):

$$\begin{aligned} f(x_0+\xi) &= \varphi(0+1) = \sum_{k=0}^{p} \frac{1}{k!}\varphi^{(k)}(0)1^k + \frac{1^{p+1}}{(p+1)!}\varphi^{(p+1)}(\theta) \\ &= \sum_{k=0}^{p} \frac{1}{k!}(\langle \xi, \nabla\rangle^k f)(x_0) + \frac{1}{(p+1)!}(\langle \xi, \nabla\rangle^{p+1} f)(x_0+\theta\xi) \end{aligned}$$

sowie

$$\begin{aligned} f(x_0+\xi) &= \varphi(0+1) = \sum_{k=0}^{p} \frac{1}{k!}\varphi^{(k)}(0)1^k + \frac{1}{p!}\int_0^1 (1-s)^p\varphi^{(p+1)}(s)ds \\ &= \sum_{k=0}^{p} \frac{1}{k!}(\langle \xi, \nabla\rangle^k f)(x_0) + \frac{1}{p!}\int_0^1 (1-s)^p(\langle \xi, \nabla\rangle^{p+1} f)(x_0+s\xi)ds\,. \end{aligned}$$

Unter Verwendung von Bemerkung 7.36 folgt die Darstellung (7.19) mit den Restglieddarstellungen (7.20) und (7.21). □

Die Restglieddarstellung (7.20) hat den Nachteil, daß sie im $\mathbb{R}^M$–wertigen Fall bei komponentenweiser Anwendung von Satz 7.37 für jede Komponente f_j von $f = (f_1, \ldots, f_M)$ ein womöglich anderes θ_j liefert. Wenden wir jedoch (7.19) mit der Integralrestglieddarstellung (7.21) komponentenweise an, so folgt:

7.38. Satz (Satz von Taylor in mehreren Veränderlichen für $\mathbb{R}^M$–wertige Funktionen). *Sei $\Omega \subseteq \mathbb{R}^N$ offen, $p \in \mathbb{N}_0$ und $f \in C^{p+1}(\Omega, \mathbb{R}^M)$. Sind $x_0 \in \Omega$ und $\xi \in \mathbb{R}^N$ gegeben, so daß die Strecke $\{x_0 + t\xi\,;\, t \in [0,1]\}$ in Ω enthalten ist, so gilt*

$$f(x_0+\xi) = \sum_{k=0}^{p} \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!}(D^\alpha f)(x_0) + R_p(x_0,\xi)\,. \tag{7.23}$$

mit der Integralrestglieddarstellung

$$R_p(x_0,\xi) =(p+1) \sum_{|\alpha|=p+1} \frac{\xi^\alpha}{\alpha!} \int_0^1 (1-s)^p (D^\alpha f)(x_0+s\xi)ds\,. \tag{7.24}$$

Als Folgerung hieraus erhalten wir die folgende qualitative Variante des Taylorschen Satzes (vergl. auch Satz 6.28 für den Fall einer Veränderlichen):

7.39. Folgerung. *Sei $\Omega \subseteq \mathbb{R}^N$ offen, $p \in \mathbb{N}_0$, $x_0 \in \Omega$ und $f \in C^p(\Omega, \mathbb{R}^M)$. Dann gilt für alle $\xi \in \mathbb{R}^N$ mit $x_0 + \xi \in \Omega$:*

$$f(x_0+\xi) = \sum_{k=0}^{p} \sum_{|\alpha|=k} (D^\alpha f)(x_0)\frac{\xi^\alpha}{\alpha!} + \varphi(\xi)$$

mit $\varphi(\xi) = \mathrm{o}(|\xi|^p)$ für $\xi \to 0$, d.h. mit

$$\frac{\varphi(\xi)}{|\xi|^p} \to 0 \qquad \text{für } \xi \to 0.$$

Beweis. Wir vermerken zunächst

$$\frac{|\xi^\alpha|}{|\xi|^p} \le 1 \qquad \text{für alle } \alpha \in \mathbb{N}_0^N \text{ mit } |\alpha| = p \text{ und alle } \xi \in \mathbb{R}^N \setminus \{0\}\,. \tag{7.25}$$

In der Tat gilt wegen $|\alpha| = p$:

$$\frac{|\xi^\alpha|}{|\xi|^p} = \prod_{j=1}^{N} \Big(\frac{|\xi_j|}{|\xi|}\Big)^{\alpha_j} \le 1\,.$$

Da Ω offen ist, gibt es ein $\varepsilon > 0$, so daß $x_0 + \xi \in \Omega$ für alle $\xi \in \mathbb{R}^N$ mit $|\xi| \le \varepsilon$. Wenden wir den Taylorschen Satz 7.38 an für $p-1$ statt p, so folgt für alle $\xi \in \mathbb{R}^N$ mit $|\xi| \le \varepsilon$:

$$\begin{aligned}\varphi(\xi) =& f(x_0+\xi) - \sum_{k=0}^{p} \sum_{|\alpha|=k} (D^\alpha f)(x_0)\frac{\xi^\alpha}{\alpha!}\\ =& p \sum_{|\alpha|=p} \frac{\xi^\alpha}{\alpha!} \int_0^1 (1-s)^{p-1}(D^\alpha f)(x_0+s\xi)ds - \sum_{|\alpha|=p} (D^\alpha f)(x_0)\frac{\xi^\alpha}{\alpha!}\end{aligned}$$

Wegen

$$p\int_0^1 (1-s)^{p-1} ds = 1$$

folgt für $0 < |\xi| \leq 1$ unter Verwendung von (7.25):

$$\begin{aligned}\frac{|\varphi(\xi)|}{|\xi|^p} &= \Big| \sum_{|\alpha|=p} \frac{\xi^\alpha}{\alpha!|\xi|^p} p \int_0^1 (1-s)^{p-1}((D^\alpha f)(x_0+s\xi) - (D^\alpha)(x_0)) f ds \Big| \\ &\leq \sum_{|\alpha|=p} \frac{|\xi^\alpha|}{\alpha!|\xi|^p} p \Big| \int_0^1 (1-s)^{p-1}\big((D^\alpha f)(x_0+s\xi) - (D^\alpha f)(x_0)\big) ds \Big| \\ &\leq \sum_{|\alpha|=p} \frac{p}{\alpha!} \Big| \int_0^1 (1-s)^{p-1}\big((D^\alpha f)(x_0+s\xi) - (D^\alpha f)(x_0)\big) ds \Big|\end{aligned}$$

Da nach Voraussetzung für alle $\alpha \in \mathbb{N}_0^N$ mit $|\alpha| = p$ die Funktion

$$(s, \xi) \mapsto (D^\alpha f)(x_0 + s\xi) - (D^\alpha f)(x_0)$$

auf $[0,1] \times \{\xi \in \mathbb{R}^N\, ;\, |\xi| \leq \varepsilon\}$ stetig ist, folgt unter Verwendung von Satz 7.9:

$$\lim_{\xi \to 0} \frac{|\varphi(\xi)|}{|\xi|^p} = 0$$

und damit die Behauptung. □

In dem Spezialfall $p = 2$ und $M = 1$ erhalten wir aus Folgerung 7.39:

$$\begin{aligned}f(x_0+\xi) &= f(x_0) + \sum_{j=1}^N \xi_j \frac{\partial f}{\partial x_j}(x_0) + \frac{1}{2!} \sum_{j=1}^N \sum_{k=1}^N \xi_j \xi_k \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0) + \varphi(\xi) \\ &= f(x_0) + \langle \xi, (\nabla f)(x_0) \rangle + \langle H_f(x_0)\xi, \xi \rangle + \varphi(\xi)\end{aligned}$$

mit

$$\frac{\varphi(\xi)}{|\xi|^p} \to 0 \qquad \text{für } \xi \to 0.$$

Hierbei sei $H_f(x_0)$ die Matrix

$$H_f(x_0) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_N}(x_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_N}(x_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_N \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_N \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_N^2}(x_0) \end{pmatrix}.$$

$H_f(x_0)$ heißt die *Hesse–Matrix*[3] von f in x_0. Sie spielt bei der Untersuchung auf lokale Extrema eine wichtige Rolle. Wegen der Vertauschbarkeit der partiellen Ableitungen (nach dem Satz 7.14 von H. A. Schwarz) ist $H_f(x_0)$ für $f \in C^2(\Omega, \mathbb{R})$ eine *symmetrische*[4] Matrix mit reellen Einträgen.

Wir benötigen einige Bezeichnungen und Hilfsmittel aus der Linearen Algebra.

7.40. Definition. Eine symmetrische Matrix $A \in M_{N\times N}(\mathbb{R})$ über $\mathbb{R}$ heißt

(a) *positiv definit*, falls $\langle A\xi, \xi\rangle > 0$ für alle $0 \neq \xi \in \mathbb{R}^N$ gilt.
(b) *negativ definit*, falls $\langle A\xi, \xi\rangle < 0$ für alle $0 \neq \xi \in \mathbb{R}^N$ gilt.
(c) *positiv semidefinit*, falls $\langle A\xi, \xi\rangle \geq 0$ für alle $\xi \in \mathbb{R}^N$ gilt.
(d) *negativ semidefinit*, falls $\langle A\xi, \xi\rangle \leq 0$ für alle $\xi \in \mathbb{R}^N$ gilt.

---

[3] Ludwig Otto Hesse (22.4.1811–4.8.1874)

[4] Eine quadratische Matrix $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{K})$ über einem Körper $\mathbb{K}$ heißt bekanntlich *symmetrisch*, wenn für alle $j,k \in \{1,\dots,N\}$ gilt $a_{j,k} = a_{k,j}$.

(e) *indefinit*, falls es $\xi, \eta \in \mathbb{R}^N$ gibt mit $\langle A\xi, \xi\rangle < 0 < \langle A\eta, \eta\rangle$.

Für eine Matrix $A = (a_{j,k})_{\substack{j=1,\dots,M\\k=1,\dots,N}} \in M_{M\times N}(\mathbb{K})$ über einem Körper $\mathbb{K}$ bezeichnen wir mit A^t die zu A transponierte Matrix $A^t := (a_{j,k})_{\substack{k=1,\dots,N\\j=1,\dots,M}} \in M_{N\times M}(\mathbb{K})$. Eine quadratische Matrix $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{K})$ über einem Körper $\mathbb{K}$ ist also genau dann symmetrisch, wenn $A^t = A$ gilt. Eine Matrix $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{K})$ heißt *orthogonal*, falls $A^tA = E_N$ gilt, wobei E_N die Einheitsmatrix in $M_{N\times N}(\mathbb{K})$ bezeichnet. Eine Matrix $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{K})$ mit $a_{j,k} = 0$ für alle $j,k \in \{1,\dots,N\}$ mit $j \neq k$ nennen wir eine *Diagonalmatrix*. Sind $\lambda_1,\dots,\lambda_N \in \mathbb{K}$ gegeben, so bezeichnen wir mit $\operatorname{diag}(\lambda_1,\dots,\lambda_N)$ die Diagonalmatrix, die auf der Diagonalen die Einträge $\lambda_1,\dots,\lambda_N$ hat. Eine Zahl $\lambda \in \mathbb{K}$ heißt ein *Eigenwert* der quadratischen Matrix $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{K})$, falls die Gleichung

$$A\xi = \lambda\xi \tag{7.26}$$

eine Lösung $0 \neq \xi \in \mathbb{K}^N$ besitzt. Die Eigenwerte sind gerade die Nullstellen des charakteristischen Polynoms $q_A(\lambda) = \det(A - \lambda E_N)$. Ist ξ eine vom Nullvektor verschiedene Lösung von (7.26), so nennen wir ξ einen *Eigenvektor* von A zum Eigenwert λ. Aus der Linearen Algebra ist bekannt (siehe z.B. in [**16**]):

7.41. Satz. *Zu jeder symmetrischen reellen Matrix $A \in M_{N\times N}(\mathbb{R})$ gibt es eine orthogonale Matrix $U \in M_{N\times N}(\mathbb{R})$ und eine Diagonalmatrix $D = \operatorname{diag}(\lambda_1,\dots,\lambda_N)$ mit*

$$A = U^tDU\,.$$

Hierbei sind $\lambda_1,\dots,\lambda_N$ die Eigenwerte von A die entsprechend ihrer algebraischen Vielfachheit (als Nullstellen des charakteristischen Polynoms von A) aufgeführt sind.

Die beiden folgenden Kriterien sind nützlich bei der Untersuchung von quadratischen, reellen Matrizen auf ihre Definitheitseigenschaften:

7.42. Lemma. *Für eine symmetrische, reelle Matrix $A \in M_{N\times N}(\mathbb{R})$ gilt:*

(a) *A ist positiv (bzw. negativ) definit genau dann, wenn alle Eigenwerte von A positiv (bzw. negativ) sind.*
(b) *A ist positiv (bzw. negativ) semidefinit genau dann, wenn alle Eigenwerte von A nicht negativ (bzw. nicht positiv) sind.*
(c) *A ist genau dann indefinit, wenn A sowohl positive als auch negative Eigenwerte besitzt.*

Beweis. (a) Wir beweisen die Aussage nur für den positiv definiten Fall. Die andere Äquivalenz zeigt man analog.

"$\Longrightarrow$": Sei also A positiv definit und sei λ ein Eigenvektor von A. Dann gibt es einen Eigenvektor $\xi \neq 0$ von A zu diesem Eigenvektor. Es folgt

$$0 < \langle A\xi, \xi\rangle = \lambda|\xi|^2$$

und daher $\lambda > 0$. Also sind alle Eigenwerte von A positiv.

"$\Longleftarrow$": Sind alle Eigenwerte von A positiv, so gilt nach Satz 7.41

$$A = U^tDU$$

mit einer orthogonalen Matrix U und einer Diagonalmatrix $D = \operatorname{diag}(\lambda_1,\dots,\lambda_N)$, deren Einträge auf der Diagonalen alle positiv sind. Sei nun $0 \neq \xi \in \mathbb{R}^N$. Es folgt

$$\langle A\xi, \xi\rangle = \langle U^tDU\xi, \xi\rangle = \langle DU\xi, U\xi\rangle = \sum_{j=1}^N \lambda_j b_j^2$$

mit positiven $\lambda_1, \ldots, \lambda_N$ und

$$\begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} := U\xi \neq 0 \quad \text{wegen } 0 \neq \xi = U^t U \xi.$$

Also folgt $\langle A\xi, \xi \rangle > 0$.

(b), (c) als Übung. □

Ist $A = (a_{j,k})_{j,k=1,\ldots,N}$ eine quadratische Matrix über einem Körper $\mathbb{K}$, so heißen für $k = 1, \ldots, N$ die Zahlen

$$\Delta_k(A) := \det \begin{pmatrix} a_{1,1} & \ldots & a_{1,k} \\ \vdots & & \vdots \\ a_{k,1} & \ldots & a_{k,k} \end{pmatrix}$$

die *Hauptminoren* von A. Den Beweis des folgenden Kriteriums für positive Definitheit einer symmetrischen Matrix $A \in M_{N\times N}(\mathbb{R})$ findet man in Lehrbüchern zur Linearen Algebra, z.B. in [10], p. 328.

7.43. Satz. *Eine symmetrische Matrix $A \in M_{N\times N}(\mathbb{R})$ über $\mathbb{R}$ ist genau dann positiv definit, wenn alle Ihre Hauptminoren positiv sind.*

Speziell für $N = 2$ hat man also: $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ ist genau dann positiv definit, wenn $a > 0$ und $ac - b^2 > 0$ erfüllt ist.

Eine notwendige Bedingung für das Vorliegen einer lokalen Extremstelle wird gegeben durch:

7.44. Satz. *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei $f \in C^2(\Omega, \mathbb{R})$. f habe in dem Punkt $x_0 \in \Omega$ eine lokale Extremstelle. Dann ist $(\nabla f)(x_0) = 0$ und es gilt:*

(a) *Ist x_0 eine lokale Maximumstelle, so ist $H_f(x_0)$ negativ semidefinit.*
(b) *Ist x_0 eine lokale Minimumstelle, so ist $H_f(x_0)$ positiv semidefinit.*

Beweis. Nach Satz 7.29 ist $(\nabla f)(x_0) = 0$. Sei nun x_0 eine lokale Maximumstelle.

Annahme: $H_f(x_0)$ ist nicht negativ semidefinit. Dann gibt es ein $\xi \in \mathbb{R}^N$ mit

$$\langle H_f(x_0)\xi, \xi \rangle > 0.$$

Insbesondere ist $\xi \neq 0$. Wegen $f \in C^2(\Omega, \mathbb{R})$ ist die durch

$$h(x) := \langle H_f(x)\xi, \xi \rangle \qquad \text{für alle } x \in \Omega$$

definierte Funktion $h : \Omega \to \mathbb{R}$ stetig. Es gibt also ein $\delta > 0$, so daß für alle $u \in \mathbb{R}^N$ mit $|u| \leq \delta$ gilt

$$x_0 + u \in \Omega \quad \text{und} \quad \langle H_f(x_0 + u)\xi, \xi \rangle > 0.$$

Sei $t \in (0, \delta)$ beliebig. Für $\eta := \frac{t}{|\xi|}\xi$ gilt nach dem Taylorschen Satz 7.37 mit dort $p = 1$:

$$f(x_0 + \eta) = f(x_0) + \langle \eta, (\nabla f)(x_0) \rangle + R_1(x_0, \eta) = f(x_0) + R_1(x_0, \eta)$$

mit der Restglieddarstellung

$$\begin{aligned} R_1(x_0,\eta) = \sum_{|\alpha|=2} \frac{\eta^\alpha}{\alpha!}(D^\alpha f)(x_0+\theta\eta) = & \quad \frac{1}{2}\sum_{j=1}^N\sum_{k=1}^N \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0+\theta\eta)\eta_j\eta_k \\ = & \frac{t^2}{2|\xi|^2}\langle H_f(x_0+\theta\eta)\xi,\xi\rangle \end{aligned}$$

mit einem $\theta \in (0,1)$. Wegen $|\theta\eta| = \theta t < \delta$ folgt $R_1(x_0,\eta) > 0$ und somit

$$f\Big(x_0 + \frac{t}{|\xi|}\xi\Big) > f(x_0) \qquad \text{für alle } t \in (0,\delta)$$

im Widerspruch dazu, daß x_0 nach Voraussetzung eine lokale Maximumstelle ist. Also war die oben gemachte Annahme falsch. $H_f(x_0)$ muß negativ semidefinit sein.

Die Behauptung in (b) beweist man analog oder durch Betrachten von $-f$ statt f. □

7.45. Satz. *Sei $\Omega \subseteq \mathbb{R}^N$ offen, $f \in C^2(\Omega,\mathbb{R})$, $x_0 \in \Omega$ und $(\nabla f)(x_0) = 0$.*

(a) *Ist $H_f(x_0)$ positiv definit, so ist x_0 eine isolierte lokale Minimumstelle.*
(b) *Ist $H_f(x_0)$ negativ definit, so ist x_0 eine isolierte lokale Maximumstelle.*
(c) *Ist $H_f(x_0)$ indefinit, so ist x_0 keine lokale Extremstelle.*

Beweis. (a) Sei also $H_f(x_0)$ positiv definit. Wir zeigen zunächst

(i) $\varepsilon := \inf\{\langle H_f(x_0)\xi,\xi\rangle\,;\, \xi \in \mathbb{R}^N, |\xi| = 1\} > 0$.

Die Menge $S := \{\xi \in \mathbb{R}^N, |\xi| = 1\}$ ist eine beschränkte und abgeschlossene Teilmenge des $\mathbb{R}^N$. Da die Funktion

$$g : S \to \mathbb{R} \qquad \xi \mapsto g(\xi) := \langle H_f(x_0)\xi,\xi\rangle,$$

auf S stetig ist, nimmt sie in einem Punkt $\xi_0 \in S$ ihr Minimum an. Wegen $|\xi_0| = 1$ ist $\xi \neq 0$ und es folgt nach Voraussetzung

$$0 < \varepsilon := \langle H_f(x_0)\xi_0,\xi_0\rangle = \min_{\xi\in S}\langle H_f(x_0)\xi,\xi\rangle\,.$$

(ii) Wegen $f \in C^2(\Omega,\mathbb{R})$ ist die Abbildung $x \mapsto H_f(x)$ von Ω nach $M_{N\times N}(\mathbb{R})$ stetig. Es gibt also ein $\delta > 0$, so daß für alle $u \in \mathbb{R}^N$ mit $|u| < \delta$ gilt:

$$\forall u \in \mathbb{R}^N : |u| < \delta \quad \Longrightarrow \quad x_0 + u \in \Omega \quad \text{und} \quad \|H_f(x_0+u) - H_f(x_0)\|_2 < \varepsilon\,,$$

wobei für $A = (a_{j,k})_{j,k=1,\dots,N} \in M_{N\times N}(\mathbb{R})$

$$\|A\|_2 := \Big(\sum_{j=1}^N\sum_{k=1}^N a_{j,k}^2\Big)^{1/2}$$

wieder die Frobeniusnorm von A sei. Dann gilt für alle $u \in \mathbb{R}^N$ mit $|u| < \delta$ und alle $\xi \in S$ (unter Verwendung der in den Übungen hergeleiteten Beziehungen zwischen der Matrixnorm $\|\cdot\|$ und der Frobeniusnorm $\|\cdot\|_2$):

$$\begin{aligned} \langle H_f(x_0+u)\xi,\xi\rangle = & \langle H_f(x_0)\xi,\xi\rangle + \langle (H_f(x_0+u) - H_f(x_0))\xi,\xi\rangle \\ \geq & \varepsilon - |\xi|\cdot\big|(H_f(x_0+u) - H_f(x_0))\xi\big| \\ \geq & \varepsilon - \|H_f(x_0+u) - H_f(x_0)\| \geq \varepsilon - \|H_f(x_0+u) - H_f(x_0)\|_2 > 0. \end{aligned}$$

(iii) Sei nun $u \in \mathbb{R}^N$ beliebig mit $0 < |u| < \delta$. Dann gilt nach dem Taylorschen Satz 7.37 (wie im Beweis zu Satz 7.44) für ein $\theta \in (0,1)$:

$$f(x_0+u) = f(x_0) + \langle u, (\nabla f)(x_0)u\rangle + \frac{1}{2}\langle H_f(x_0+\theta u)u,u\rangle = f(x_0) + \frac{1}{2}\langle H_f(x_0+\theta u)u,u\rangle$$

mit

$$\frac{1}{2}\langle H_f(x_0+\theta u)u,u\rangle = \frac{|u|^2}{2}\left\langle H_f(x_0+\theta u)\frac{1}{|u|}u, \frac{1}{|u|}u\right\rangle > 0$$

nach (ii) (wegen $|\theta u| = \theta|u| < \delta$). x_0 ist also eine isolierte lokale Minimumstelle.

(b) behandelt man analog oder durch Anwenden von (a) auf die Funktion $-f$.

(c) folgt unmittelbar aus Satz 7.44. □

7.46. Beispiele. (a) $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch $f(x,y) := 1+x^2+xy+2y^2+19y^3$ für alle $(x,y) \in \mathbb{R}^2$. Man rechnet nach:

$$(\nabla f)(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{und} \quad H_f(0,0) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

Mit Satz 7.43 sieht man, daß $H_f(0,0)$ positiv definit ist. Es liegt also in $(0,0)$ eine isolierte lokale Minimumstelle vor.

(b) Sei nun $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) := 1+x^2+3xy-y^2+y^3$ für alle $(x,y) \in \mathbb{R}^2$. Man rechnet nach:

$$(\nabla f)(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{und} \quad H_f(0,0) = \begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix}.$$

Es folgt

$$\left\langle H_f(0,0)\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\rangle = -2 < 0 \quad \text{und} \quad \left\langle H_f(0,0)\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\rangle = 2 > 0.$$

$H_f(0,0)$ ist also indefinit. Nach Satz 7.45 liegt daher in $(0,0)$ keine lokale Extremstelle vor.

7.47. Bemerkung. Man beachte, daß Satz 7.45 im semidefiniten Fall *keine* Aussage liefert.

Hierzu einige Beispiele:

7.48. Beispiele. (a) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) := x^4+y^4$ für alle $x,y \in \mathbb{R}$ hat in $(0,0)$ eine isolierte lokale Minmumstelle mit

$$(\nabla f)(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{und} \quad H_f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.27}$$

$H_f(0,0)$ ist also positivsemidefinit und negativ semidefinit.

(b) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) := -x^4-y^4$ für alle $x,y \in \mathbb{R}$ hat in $(0,0)$ eine isolierte lokale Maximumstelle und erfüllt ebenfalls (7.27).

(c) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) := x^4-y^4$ für alle $x,y \in \mathbb{R}$ erfüllt ebenfalls (7.27), besitzt aber in $(0,0)$ keine lokale Extremstelle, denn für alle $t > 0$ gilt

$$f(t,0) = t^4 > 0 = f(0,0) > -t^4 = f(0,t).$$

KAPITEL 8

Metrische Räume

1. Definition und erste Eigenschaften, Stetigkeit

Sei E ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $\|\cdot\| : E \to [0, \infty)$ heißt eine *Norm* auf E, falls sie den folgenden Bedingungen genügt:

(N1) $\forall x \in X : \qquad \|x\| = 0 \iff x = 0.$
(N2) $\forall x \in X \, \forall \lambda \in \mathbb{K} : \qquad \|\lambda x\| = |\lambda| \cdot \|x\|.$
(N3) $\forall x, y \in X : \qquad \|x + y\| \leq \|x\| + \|y\|.$ (*Dreiecksungleichung*)

Wir nennen E versehen mit dieser Norm dann einen *normierten Raum* und schreiben $(E, \|\cdot\|)$ um anzudeuten, daß E mit der Norm $\|\cdot\|$ versehen ist.

Beispiele. (a) $(\mathbb{R}^N, |\cdot|)$ ist ein normierter $\mathbb{R}$–Vektorraum.

(b) Sei $\emptyset \neq K \in \mathbb{R}^N$ abgeschlossen und beschränkt. Dann ist der Raum der stetigen, $\mathbb{K}$–wertigen Funktionen versehen mit der Supremumsnorm auf K

$$f \mapsto \|f\|_K := \sup_{x \in K} |f(x)|$$

ein normierter $\mathbb{K}$–Vektorraum.

(Beweis als Übung).

In einem normierten Raum $(E, \|\cdot\|)$ kann für $x, y \in E$ der Wert $\|x - y\|$ als Abstand von x zu y angesehen werden. Wir führen einen noch allgemeineren Abstandsbegriff ein:

8.1. Definition. Sei X eine nicht leere Menge. Eine Abbildung $d : X \times X \to [0, \infty)$ heißt *Metrik* auf X, falls die folgenden drei Bedingungen erfüllt sind:

(M1) $\forall x, y \in X : \qquad d(x, y) = 0 \iff x = y.$
(M2) $\forall x, y \in X : \qquad d(x, y) = d(y, x).$ (*Symmetrie*)
(M3) $\forall x, y, z \in X : \qquad d(x, z) \leq d(x, y) + d(y, z).$ (*Dreiecksungleichung*)

(X,d) heißt dann ein *metrischer Raum*.

8.2. Beispiele. (a) Sei $(E, \|\cdot\|)$ ein normierter Raum. Dann ist $(E, d_{\|\cdot\|})$ mit $d_{\|\cdot\|}(x, y) := \|x - y\|$ für alle $x, y \in X$ ein metrischer Raum. Wir nennen $d_{\|\cdot\|}$ die von der Norm $\|\cdot\|$ induzierte Metrik auf E. Insbesondere ist also $\mathbb{R}^N$ versehen mit der durch den euklidischen Betrag induzierten Metrik $d_{|\cdot|}$ ein metrischer Raum.

(b) Sei $\emptyset \neq F \subseteq \mathbb{R}^2$ und $p \in F$. Wir definieren die *Zentralismus–Metrik* $d_p : F \times F \to [0, \infty)$ durch

$$d_p(x, y) := \begin{cases} |x - p| + |p - y| & \text{falls } x \neq y \\ 0 & \text{falls } x = y \end{cases}$$

für alle $x, y \in F$. Man rechnet leicht nach, daß die Bedingungen (M1)–(M3) erfüllt sind.

(c) Sei X eine nicht leere Menge. Wir definieren die *diskrete Metrik* $d : X \times X \to [0, \infty)$ durch

$$d(x, y) := \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{falls } x \neq y \end{cases}$$

für alle $x, y \in X$. Auch hier rechnet man unmittelbar nach, daß die Bedingungen (M1)–(M3) erfüllt sind.

8.3. Bemerkung. Sei (X, d) ein metrischer Raum und sei $\emptyset \neq Y \subseteq X$. Man rechnet unmittelbar nach, daß dann $(Y, d|_{Y\times Y})$ ebenfalls ein metrischer Raum ist. Wir nennen $d|_{Y\times Y}$ die von d auf Y induzierte Metrik. Insbesondere wird also jede nicht leere Teilmenge Y eines normierten Raums, versehen mit der durch $d_{\|\cdot\|}$ auf Y induzierten Metrik zu einem metrischen Raum.

Auch die aus dem $\mathbb{R}^N$ bekannten topologischen Begriffe wie der Umgebungsbegriff und die Begriffe der offenen und der abgeschlossenen Mengen lassen sich mühelos auf allgemeine metrische Räume übertragen.

8.4. Definition. Sei (X, d) ein metrischer Raum und $x_0 \in X$.

(a) Für $\varepsilon > 0$ heißt die Menge $U_\varepsilon(x_0) := \{x \in X\,;\, d(x, x_0) < \varepsilon\}$ die *ε–Umgebung* von x_0 in (X, d).
(b) $U \subseteq X$ heißt *Umgebung* von x_0, falls es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq U$ gibt.
(c) $\Omega \subseteq X$ heißt *offen* in (X, d), falls Ω Umgebung eines jeden Punktes von Ω ist, d.h. falls es zu jedem $x \in \Omega$ ein $\varepsilon = \varepsilon(x) > 0$ gibt mit $U_\varepsilon(x) \subseteq \Omega$.
(d) x_0 heißt *innerer Punkt* einer Teilmenge M von X, wenn M eine Umgebung von x_0 ist, d.h. wenn es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(x_0) \subseteq M$.
(e) Für eine Teilmenge M von X heißt die Menge

$$\mathring{M} := \operatorname{int}(M) := \{x \in M\,;\, x \text{ ist innerer Punkt von } M\}$$

der inneren Punkte von M das *Innere* oder der *offene Kern* von M in (X, d).
(f) Eine Teilmenge A von X heißt *abgeschlossen*0 in (X, d), falls ihr Komplement $X \setminus A$ offen in (X, d) ist.
(g) x_0 heißt *Berührungspunkt* einer Menge $M \subseteq X$, falls für jede Umgebung U von x_0 in (X, d) gilt: $U \cap M \neq \emptyset$.
(h) Für eine Teilmenge M von X heißt die Menge

$$\overline{M} := \{x \in M\,;\, x \text{ Berührungspunkt von } M \text{ in } (X, d)\}$$

der Berührungspunkte von M in (X, d) die *abgeschlossene Hülle* oder die *Abschließung* von M in (X, d).

Bemerkungen. Sei (X, d) ein metrischer Raum und $x_0 \in X$.

(a) Die Mengen X und $\emptyset$ sind sowohl offen als auch abgeschlossen in (X, d).
(b) Für jedes $\varepsilon > 0$ ist die ε–Umgebung $U_\varepsilon(x_0)$ von x_0 in (X, d) offen.
(c) Für jedes $\varepsilon > 0$ ist die Menge

$$B_\varepsilon(x_0) := \{x \in X\,;\, d(x, x_0) \leq \varepsilon\}$$

abgeschlossen in (X, d).

Beweis. (a) ist klar, da sowohl X als auch $\emptyset$ die Bedingung in Definition 8.4 (c) erfüllen.

(b) Sei $\varepsilon > 0$ beliebig und $y \in U_\varepsilon(x_0)$ beliebig. Dann ist $\delta := \varepsilon - d(y, x_0) > 0$ und für alle $x \in U_\delta(y)$ gilt nach der Dreiecksungleichung:

$$d(x, x_0) \leq d(x, y) + d(y, x_0) < \delta + d(y, x_0) = \varepsilon\,.$$

Es folgt $U_\delta(y) \subseteq U_\varepsilon(x_0)$. $U_\varepsilon(x_0)$ ist also Umgebung für alle in $U_\varepsilon(x_0)$ enthaltenen Punkte und ist somit offen nach Definition 8.4 (c).

(c) Wir müssen zeigen, daß $X \setminus B_\varepsilon(x_0)$ offen in (X,d) ist, daß es also zu jedem Punkt $y \in X \setminus B_\varepsilon(x_0)$ ein $\delta > 0$ gibt mit $U_\delta(y) \subseteq X \setminus B_\varepsilon(x_0)$. Sei also $y \in X \setminus B_\varepsilon(x_0)$ beliebig. Dann gilt $\delta := d(y,x_0) - \varepsilon > 0$. Nach der Dreiecksungleichung gilt für alle $x \in U_\delta(y)$

$$d(x,x_0) + d(y,x) \geq d(y,x_0)$$

und somit

$$d(x,x_0) \geq d(y,x_0) - d(y,x) > d(y,x_0) - \delta = \varepsilon\,.$$

Also ist $U_\delta(y) \subseteq X \setminus B_\varepsilon(x_0)$ und wir haben gezeigt, daß $B_\varepsilon(x_0)$ abgeschlossen ist. □

Beispiele. Wir betrachten in $(\mathbb{R}, d_{|\cdot|})$ die Menge $M := (0,1]$. Es ist $1 \in M$ aber für alle $\varepsilon > 0$ ist $U_\varepsilon(1) \not\subseteq M$. Also ist M nicht offen. Es ist $0 \in \mathbb{R} \setminus M$ aber für alle $\varepsilon > 0$ ist $U_\varepsilon(0) \not\subseteq \mathbb{R} \setminus M$. Also ist $\mathbb{R} \setminus M$ nicht offen und M daher auch nicht abgeschlossen. Man rechnet leicht nach, daß $\mathrm{int}(M) = (0,1)$ und $\overline{M} = [0,1]$ gilt.

(b) Sei X eine nicht leere Menge und $d : X \times X \to [0,\infty)$ die diskrete Metrik auf X (vergl. Beispiel 8.2 (c)). Dann gilt für alle $\varepsilon \in (0,1]$ und alle $x \in X$: $U_\varepsilon(x) = \{x\}$ und für alle $\varepsilon > 1$ ist $U_\varepsilon(x) = X$. Sei $M \subseteq X$ beliebig. Es folgt für alle $x \in M$: $U_1(x) = \{x\} \subseteq M$. Also ist M offen in (X,d). Da dies auch für $X \setminus M$ gilt, ist M auch abgeschlossen in (X,d). Alle Teilmengen von X sind somit bezüglich der diskreten Metrik auf X sowohl offen als auch abgeschlossen.

Das folgende Lemma enthält Aussagen über Durchschnitte bzw. Vereinigungen von offenen und von abgeschlossenen Teilmengen eines metrischen Raums.

8.5. Lemma. *Sei (X,d) ein metrischer Raum.*

(a) *Der Durchschnitt von endlich vielen in (X,d) offenen Teilmengen von X ist offen in (X,d).*
(b) *Die Vereinigung von beliebig vielen in (X,d) offenen Teilmengen von X ist offen in (X,d).*
(c) *Die Vereinigung von endlich vielen in (X,d) abgeschlossenen Teilmengen von X ist abgeschlossen in (X,d).*
(d) *Der Durchschnitt von beliebig vielen in (X,d) abgeschlossenen Teilmengen von X ist abgeschlossen in (X,d).*

Beweis. (a) Seien $G_1, \ldots, G_n$ endlich viele offene Mengen in (X,d) und sei $x \in \bigcap_{j=1}^n G_j$ beliebig. Zu jedem $j \in \{1,\ldots,n\}$ gibt es ein $\varepsilon_j > 0$ mit $U_{\varepsilon_j}(x) \subseteq G_j$. Mit $\varepsilon := \min\{\varepsilon_1, \ldots, \varepsilon_n\}$ gilt dann: $\varepsilon > 0$ und $U_\varepsilon(x) \subseteq \bigcap_{j=1}^n G_j$. Nach Definition 8.4 (c) ist $\bigcap_{j=1}^n G_j$ also offen.

(b) Sei $(G_i)_{i\in I}$ eine beliebige Familie von in (X,d) offenen Teilmengen von X und sei $x \in \bigcup_{i\in I} G_i$ beliebig. Dann gibt es ein $i_0 \in I$, so daß $x \in G_{i_0}$. Da G_{i_0} offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq G_{i_0} \subseteq \bigcup_{i\in I} G_i$. Also ist $\bigcup_{i\in I} G_i$ offen in (X,d).

(c) und (d) erhält man aus (a) bzw. (b) durch Übergang zum Komplement. □

Dieses Lemma legt eine noch allgemeinere Begriffsbildung nahe: Sei $X \neq \emptyset$ eine Menge und $\mathcal{O}$ eine Menge von Teilmengen von X mit den folgenden Eigenschaften:

(T1) $\emptyset \in \mathcal{O}$ und $X \in \mathcal{O}$.
(T2) Endliche Durchschnitte von Elementen aus $\mathcal{O}$ liegen in $\mathcal{O}$.
(T3) Vereinigungen von beliebig vielen Elementen aus $\mathcal{O}$ liegen in $\mathcal{O}$.

Dann heißt $\mathcal{O}$ eine *Topologie* auf X und $(X, \mathcal{O})$ ein *topologischer Raum*. Die Elemente von $\mathcal{O}$ heißen die in $(X, \mathcal{O})$ offenen Mengen.

Die Bezeichnungen *offener Kern* und *abgeschlossene Hülle* in Definition 8.4 werden gerechtfertigt durch das folgende Lemma:

8.6. Lemma. *Sei (X, d) ein metrischer Raum und $M \subseteq X$.*

(a) $\operatorname{int}(M)$ *ist die größte in M enthaltene offene Teilmenge von X. Es gilt*

$$\operatorname{int}(M) = \bigcup \{G \subseteq M\,;\, G \text{ ist offen in } (X, d)\}\,. \tag{8.1}$$

(b) *Es gilt:* $M = \operatorname{int}(M) \iff M$ *ist offen in* (X, d).

(c) $\operatorname{int}(M) = \operatorname{int}(\operatorname{int}(M))$.

(d) $\overline{M}$ *ist die kleinste abgeschlossene Teilmenge von X, die M enthält. Es gilt*

$$\overline{M} = \bigcap \{A \subseteq Y\,;\, M \subseteq A \text{ und } A \text{ ist abgeschlossen}\}\,. \tag{8.2}$$

(e) *Es gilt:* $\overline{M} = M \iff M$ *ist abgeschlossen.*

(f) $\overline{M} = \overline{\overline{M}}$.

Beweis. (a) Wir zeigen zunächst, daß $\operatorname{int}(M)$ offen in (X, d) ist. Sei also $x_0 \in \operatorname{int}(M)$ beliebig. Nach Definition von $\operatorname{int}(M)$ gibt es dann ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq M$. Sei $y \in U_\varepsilon(x_0)$ beliebig. Dann ist $\delta := \varepsilon - d(x_0, y) > 0$ und $U_\delta(y) \subseteq U_\varepsilon(x_0) \subseteq M$ (vergl. den Beweis zu Teil (b) der Bemerkungen nach Definition 8.4). Daher ist auch y ein innerer Punkt von M. Es folgt $U_\varepsilon(x_0) \subseteq \operatorname{int}(M)$. $\operatorname{int}(M)$ ist also offen.

Ist nun G eine beliebige in (X, d) offene Teilmenge von M, so gibt es zu jedem $x \in G$ ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq G \subseteq M$. Also sind alle Punkte aus G innere Punkte von M und es folgt (8.1).

(b) "$\Longrightarrow$" folgt unmittelbar aus (a) und "$\Longleftarrow$" ergibt sich unmittelbar aus der Definition der offenen Mengen und des offenen Kerns.

(c) folgt aus (a).

(d) Wir zeigen zunächst $M \subseteq \overline{M}$. Ist $x \in M$ beliebig, so gilt für alle Umgebungen U von x in (X, d): $x \in U \cap M$ und somit $U \cap M \neq \emptyset$. Jeder Punkt aus M ist also Berührungspunkt von M.

Wir zeigen nun, daß $\overline{M}$ abgeschlossen ist in (X, d), indem wir zeigen, daß $X \setminus \overline{M}$ offen ist. Sei also $x_0 \in X \setminus \overline{M}$ beliebig. Da x_0 kein Berührungspunkt von M ist, gibt es eine Umgebung U von x mit $U \cap M = \emptyset$. Nach Definition der Umgebung gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq U$. Wir zeigen, daß $U_\varepsilon(x_0)$ keine Berührungspunkte von M enthält und somit in $X \setminus \overline{M}$ enthalten ist. Sei hierzu $y \in U_\varepsilon(x_0)$ beliebig. Es folgt mit $\delta := \varepsilon - d(x_0, y) > 0$, daß $U_\delta(y) \subseteq U_\varepsilon(x_0) \subseteq U$ gilt (vergl. den Beweis zu Teil (b) der Bemerkungen nach Definition 8.4). Es folgt $U_\delta(y) \cap M \subseteq U \cap M = \emptyset$. y ist somit kein Berührungspunkt von M und es ist $U_\varepsilon(x_0) \subseteq X \setminus \overline{M}$. Also ist $X \setminus \overline{M}$ offen und somit $\overline{M}$ abgeschlossen.

Sei nun A eine beliebige abgeschlossene Teilmenge von X mit $M \subseteq A$. Da $X \setminus A$ offen ist gibt es zu jedem $x \in X \setminus A$ ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq X \setminus A$, also auch $U_\varepsilon(x) \subseteq X \setminus M$. Die Punkte aus $X \setminus A$ sind daher keine Berührungspunkte von M. Also ist $\overline{M} \subseteq A$ und es folgt die Behauptung.

(e) folgt unmittelbar aus (d) und (f) ist eine Konsequenz von (d) und (e). $\square$

Auch die Begriffe Konvergenz und Stetigkeit lassen sich ohne Schwierigkeiten im Rahmen der Theorie metrischer Räume formulieren.

8.7. Definition. Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_{n=1}^\infty$ heißt

(a) *Cauchy–Folge* in (X, d), falls gilt:

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \qquad d(x_n, x_m) < \varepsilon \,.$$

(b) *konvergent gegen ein* $y \in X$ *bzgl.* d, falls gilt:

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \qquad d(x_n, y) < \varepsilon \,.$$

Wir schreiben dann auch $x_n \to y$ für $n \to \infty$. Dies ist also genau dann der Fall, wenn $d(x_n, y) \to 0$ in $\mathbb{R}$ für $n \to \infty$.

Bemerkung. Das Element y ist durch die Bedingung in (b) eindeutig bestimmt; sind $y, z \in X$ mit $x_n \to y$ und $x_n \to z$ für $n \to \infty$, so folgt mit der Dreiecksungleichung

$$0 \leq d(y, z) \leq d(y, x_n) + d(x_n, z) \to 0 \quad \text{für } n \to \infty \,.$$

Wir nennen daher y wie in (b) den *Grenzwert* oder den *Limes der Folge* $(x_n)_{n=1}^\infty$ und schreiben $y = \lim_{n\to\infty} x_n$.

Jede konvergente Folge $(x_n)_{n=1}^\infty$ in einem metrischen Raum (X, d) ist eine Cauchy–Folge. Ist etwa $x_n \to y$ für $n \to \infty$, so gibt es zu beliebigem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt $d(x_n, y) < \varepsilon/2$. Dann folgt mit der Dreiecksungleichung für alle $n, m \geq n_0$:

$$d(x_n, x_m) \leq d(x_n, y) + d(y, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \,.$$

Also ist $(x_n)_{n=1}^\infty$ in der Tat eine Cauchy–Folge.

Ein metrischer Raum (X, d) heißt *vollständig*, falls jede Cauchy–Folge in (X, d) konvergent ist. Eine Teilmenge M eines metrischen Raumes (X, d) heißt vollständig, falls der metrische Raum $(M, d|_{M\times M})$ vollständig ist.

Nach dem Cauchy–Kriterium 3.23 ist $(\mathbb{R}^N, d_{|\cdot|})$ ein vollständiger metrischer Raum.

Bemerkung. Jede vollständige Teilmenge A eines metrischen Raums (X, d) ist abgeschlossen.

Beweis. Sei x ein beliebiger Berührungspunkt von A. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in A \cap U_{1/n}(x)$. Wegen $0 \leq d(x, x_n) < 1/n \to 0$ für $n \to \infty$ folgt $x_n \to x$ für $n \to \infty$. Insbesondere ist $(x_n)_{n=1}^\infty$ eine Cauchy–Folge in X. Für alle $\varepsilon > 0$ existiert also ein $n_0 \in \mathbb{N}$ mit

$$\forall n, m \geq n_0 : \qquad d|_{A\times A}(x_n, x_m) = d(x_n, x_m) < \varepsilon \,.$$

$(x_n)_{n=1}^\infty$ ist daher auch eine Cauchy–Folge in $(A, d|_{A\times A})$ konvergiert also wegen der Vollständigkeit in $(A, d|_{A\times A})$ gegen ein $a \in A$ für $n \to \infty$. Wegen

$$0 \leq d(x, a) \leq d(x, x_n) + d(x_n, a) \to 0 \qquad \text{für } n \to \infty$$

ist $d(x, a) = 0$ und somit $x = a \in A$. Damit ist gezeigt, daß A mit der Menge seiner Berührungspunkte übereinstimmt. Nach Lemma 8.6 ist A daher abgeschlossen. □

Bemerkung. Sei (X, d) ein vollständiger Raum. Dann ist jede abgeschlossene Teilmenge A von X vollständig.

Beweis. Ist $(x_n)_{n=1}^\infty$ eine Cauchy–Folge in $(A, d|_{A\times A})$, so ist $(x_n)_{n=1}^\infty$ offensichtlich auch eine Cauchy–Folge in (X, d). Da (X, d) nach Voraussetzung vollständig ist, gilt also $x_n \to x_0 \in X$ für $n \to \infty$ für ein $x_0 \in X$. Sei U eine beliebige Umgebung von x_0. Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq U$. Wegen $d(x_n, x_0) \to 0$ für $n \to \infty$ gibt es ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt, $d(x_n, x_0) < \varepsilon$ und daher $x_n \in U_\varepsilon(x_0) \cap A \subseteq U \cap A$. Wegen $U \cap A \neq \emptyset$ für alle Umgebungen U von x_0 ist also $x_0 \in \overline{A} = A$. Ferner gilt

$$d|_{A\times A}(x_n, x_0) = d(x_n, x_0) \to 0 \qquad \text{für } n \to \infty$$

und somit $x_n \to x_0$ in $(A, d|_{A\times A})$. Jede Cauchy–Folge in $(A, d|_{A\times A})$ ist also in $(A, d|_{A\times A})$ konvergent und die Vollständigkeit von A ist bewiesen. □

Insbesondere ist also jede abgeschlossene Teilmenge des $(\mathbb{R}^N, d_{|\cdot|})$ vollständig.

8.8. Definition. Seien (X, d) und (Y, d') zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt

(a) *stetig in einem Punkt* $x_0 \in X$, falls für jede Umgebung U von $f(x_0)$ in (Y, d') die Menge $f^{-1}(U)$ eine Umgebung von x_0 in (X, d) ist.
(b) *stetig auf einer Menge* $M \subseteq X$, falls f in allen Punkten von M stetig ist.
(c) eine *Homöomorphie*, falls f bijektiv und stetig ist und auch die Umkehrabbildung stetig ist.

Wie im $\mathbb{R}^N$ gibt es eine Reihe von äquivalenten Bedingungen für die Stetigkeit in einem Punkt:

8.9. Lemma. *Seien (X, d) und (Y, d') zwei metrische Räume. Für eine Abbildung $f : X \to Y$ und einen Punkt $x_0 \in X$ sind die folgenden vier Aussagen äquivalent:*

(a) *f ist in x_0 stetig.*
(b) $\forall \varepsilon > 0\, \exists \delta > 0 \qquad f(U_\delta(x_0)) \subseteq U_\varepsilon(f(x_0))$.
(c) $\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in X : \quad d(x, x_0) < \delta \Longrightarrow d'(f(x), f(x_0)) < \varepsilon$.
(d) *f ist folgenstetig in x_0, d.h. für jede gegen x_0 konvergente Folge $(x_n)_{n=1}^\infty$ in (X, d) gilt $f(x_n) \to f(x_0)$ für $n \to \infty$.*

Beweis. "(a)$\Longrightarrow$(b)": Sei $\varepsilon > 0$ beliebig. Da $U_\varepsilon(f(x_0))$ eine Umgebung von $f(x_0)$ in (Y, d') ist, ist $f^{-1}(U_\varepsilon(f(x_0)))$ eine Umgebung von x_0 in (X, d). Es gibt also ein $\delta > 0$ mit $U_\delta(x_0) \subseteq f^{-1}\big(U_\varepsilon(f(x_0))\big)$. Hieraus folgt $f(U_\delta(x_0)) \subseteq U_\varepsilon(f(x_0))$.

"(b)$\Longrightarrow$(c)" ist offensichtlich.

"(c)$\Longrightarrow$(d)": Sei $(x_n)_{n=1}^\infty$ eine beliebige in (X, d) gegen x_0 konvergente Folge aus X und sei $\varepsilon > 0$ beliebig. Nach Voraussetzung gibt es ein $\delta > 0$ mit

$$\forall\, x \in X : \qquad d(x, x_0) < \delta \Longrightarrow d'(f(x), f(x_0)) < \varepsilon. \tag{8.3}$$

Wegen $x_n \to x_0$ für $n \to \infty$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$\forall\, n \geq n_0 : d(x_n, x_0) < \delta\,.$$

Zusammen mit (8.3) folgt für alle $n \geq n_0$: $d'(f(x_n), f(x_0)) < \varepsilon$. Also konvergiert die Folge $(f(x_n))_{n=1}^\infty$ in (Y, d') gegen $f(x_0)$ für $n \to \infty$ und (d) ist erfüllt.

"(d)$\Longrightarrow$(a)": Sei nun (d) erfüllt und sei V eine beliebige Umgebung von $f(x_0)$ in (Y, d'). Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(f(x_0)) \subseteq V$.

Annahme: $f^{-1}(V)$ ist keine Umgebung von x_0 in (X, d). Dann gilt für alle $\delta > 0$: $U_\delta(x_0) \not\subseteq f^{-1}(V)$. Zu jedem $n \in \mathbb{N}$ gibt es also ein $x_n \in U_{1/n}(x_0)$ mit $f(x_n) \notin V$ und daher mit $d'(f(x_n), f(x_0)) \geq \varepsilon$. Wegen $d(x_n, x_0) < 1/n \to 0$ für alle $n \in \mathbb{N}$ folgt $x_n \to x_0$ für $n \to \infty$ und somit nach Voraussetzung $f(x_n) \to f(x_0)$ im Widerspruch zu $d'(f(x_n), f(x_0)) \geq \varepsilon$ für alle $n \in \mathbb{N}$. Da die obige Annahme zu einem Widerspruch geführt hat, muß $f^{-1}(V)$ doch eine Umgebung von x_0 in (X, d) sein. □

Auch den Begriff der gleichmäßigen Stetigkeit können wir für Abbildungen zwischen metrischen Räumen einführen:

Definition. Seien (X,d) und (Y,d') zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt *gleichmäßig stetig auf* X, falls gilt:

$$\forall\,\varepsilon > 0\,\exists\,\delta > 0\,\forall x,y \in X : \qquad d(x,y) < \delta \quad \Longrightarrow \quad d'(f(x),f(y)) < \varepsilon\,.$$

Seien (X,d) und (Y,d') zwei metrische Räume. Eine Abbildung $f : X \to Y$ heißt *Lipschitz–stetig auf* X, falls es ein $L > 0$ gibt mit

$$\forall u,v \in X : \qquad d'(f(u),f(v)) \le Ld(u,v)\,. \tag{8.4}$$

L heißt dann eine *Lipschitz–Konstante* zu f.

Bemerkung. Seien (X,d) und (Y,d') zwei metrische Räume.

(a) Jede Lipschitz–stetige Abbildung $f : X \to Y$ ist gleichmäßig stetig auf X.

(b) Für alle $u,x,y \in X$ gilt

$$|d(u,x) - d(u,y)| \le d(x,y)\,.$$

Insbesondere ist die Abbildung $x \mapsto d(u,x)$ Lipschitz–stetig mit Lipschitz–Konstante $L = 1$ und daher nach (a) gleichmäßig stetig auf X.

(c) Wie man leicht nachrechnet, wird $X \times X$ zu einem metrischen Raum vermöge der durch

$$d_2 : X \times X \to [0,\infty), \qquad ((x,y),(u,v)) \mapsto d_2((x,y),(u,v)) := d(x,u) + d(y,v),$$

definierten Metrik. Die Abbildung $d : X \times X \to \mathbb{R}$ ist Lipschitz–stetig mit der Lipschitz–Konstanten $L = 1$ und daher nach (a) gleichmäßig stetig auf $X \times X$.

Beweis. (a) Sei also (8.4) erfüllt mit einer Lipschitz–Konstanten $L > 0$ und sei $\varepsilon > 0$ beliebig. Dann gilt für alle $u,v \in X$ mit $d(u,v) < \delta := \varepsilon/L$: $d'(f(u),f(v)) \ge Ld(u,v) < \varepsilon$. f ist also gleichmäßig stetig auf X.

(b) Für alle u,x,y gilt nach der Dreiecksungleichung

$$d(u,x) - d(u,y) \le d(y,x) \quad \text{und} \quad d(u,y) - d(u,x) \le d(x,y).$$

Hieraus folgt die behauptete Ungleichung.

(c) Für alle $(x,y),(u,v) \in X \times X$ gilt unter Verwendung der Ungleichung aus (b):

$$\begin{aligned} |d(x,y) - d(u,v)| \le& |d(x,y) - d(x,v)| + |d(x,v) - d(u,v)| \\ \le& d(y,v) + d(x,u) = d_2((x,y),(u,v))\,. \end{aligned}$$

Hieraus folgt die Behauptung. □

Im folgenden Lemma werden einige äquivalente Bedingungen für die Stetigkeit von Abbildungen zwischen zwei metrischen Räumen angegeben.

8.10. Lemma. *Seien (X,d) und (Y,d') zwei metrische Räume. Für eine Abbildung $f : X \to Y$ sind die folgenden vier Aussagen äquivalent:*

(a) *f ist stetig auf X.*

(b) *Für jede Teilmenge M von X ist $f(\overline{M}) \subseteq \overline{f(M)}$.*

(c) *Für jede in (Y,d') abgeschlossene Menge $A \subseteq Y$ ist $f^{-1}(A)$ abgeschlossen in (X,d).*

(d) *Für jede in (Y,d') offene Menge $G \subseteq Y$ ist $f^{-1}(G)$ offen in (X,d).*

Beweis. "(a)$\Longrightarrow$(b)": Sei f auf X stetig und sei $M \subseteq X$ beliebig. Ist x ein beliebiger Punkt aus $\overline{M}$ und $V \subseteq Y$ eine beliebige Umgebung von $f(x)$ in (Y,d'), so ist wegen der Stetigkeit von f in x die Menge $f^{-1}(V)$ eine Umgebung von x in (X,d). Da x ein Berührungspunkt von M ist, gibt es daher ein $u \in f^{-1}(V) \cap M$. Es folgt $f(u) \in V \cap f(M)$

und somit $V \cap f(M) \neq \emptyset$ für alle Umgebungen V von $f(x)$. Also ist $f(x) \in \overline{f(M)}$ für alle $x \in \overline{M}$.

"(b)$\Longrightarrow$(c)": Ist (*b*) erfüllt und ist A eine beliebige abgeschlossene Teilmenge von Y, so folgt mit $M := f^{-1}(A)$ (wegen $A = \overline{A}$ nach Lemma 8.6) aus (b)

$$f(\overline{M}) \subseteq \overline{f(M)} \subseteq \overline{A} = A$$

und hieraus

$$f^{-1}(A) = M \subseteq \overline{M} = \overline{f^{-1}(A)} \subseteq f^{-1}(A)\,.$$

Nach Lemma 8.6 ist $f^{-1}(A)$ also abgeschlossen in (X, d).

"(c)$\Longrightarrow$(d)": Sei (c) erfüllt und sei $G \subseteq Y$ eine beliebige in (Y, d') offene Menge. Dann ist $Y \setminus G$ abgeschlossen in (Y, d') und daher nach Voraussetzung auch $f^{-1}(Y \setminus G) = X \setminus f^{-1}(G)$ abgeschlossen in (X, d). Also ist $f^{-1}(G)$ offen in (X, d).

"(d)$\Longrightarrow$(a)": Ist schließlich (d) erfüllt und $x \in X$ sowie $\varepsilon > 0$ beliebig vorgegeben, so ist $U_\varepsilon(f(x))$ offen in (Y, d') und daher $f^{-1}\big(U_\varepsilon(f(x))\big)$ offen in (X, d). Es gibt also ein $\delta > 0$ mit $U_\delta(x) \subseteq f^{-1}\big(U_\varepsilon(f(x))\big)$. Es folgt $f\big(U_\delta(x)\big) \subseteq U_\varepsilon(f(x))$. Nach Lemma 8.9 ist f in x stetig. Da dies für alle $x \in X$ gilt, ist f auf ganz X stetig. □

Der folgende Satz ist eine Verallgemeinerung des Intervallschachtelungsprinzips 3.18.

8.11. Cantorscher Durchschnittssatz. *Sei (X, d) ein vollständiger metrischer Raum und sei $(A_n)_{n=1}^\infty$ eine Folge abgeschlossener Teilmengen von X mit den folgenden Eigenschaften:*

(i) $\forall n \in \mathbb{N}: \qquad \emptyset \neq A_{n+1} \subseteq A_n$.

(ii) *Es gibt ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ der Durchmesser*

$$d_n := \operatorname{diam}(A_n) := \sup_{x,y \in A_n} d(x, y)$$

endlich ist, und es gilt $d_n \to 0$ für $n \to \infty$.

Dann gibt es genau einen Punkt $x \in \bigcap\limits_{n=1}^{\infty} A_n$ und für jede Folge $(a_n)_{n=1}^\infty$ mit $a_n \in A_n$ für alle $n \in \mathbb{N}$ gilt $a_n \to x$ für $n \to \infty$.

Beweis. Sei also $(a_n)_{n=1}^\infty$ eine Folge mit $a_n \in A_n$ für alle $n \in \mathbb{N}$ und sei $\varepsilon > 0$ beliebig. Wegen $d_n \to 0$ für $n \to \infty$ gibt es ein $n_1 \in \mathbb{N}$, so daß für alle $n \geq n_1$ gilt $d_n < \varepsilon$. Für alle $n, m \geq n_1$ folgt dann wegen $a_n \in A_n \subseteq A_{n_1}$ und $a_m \in A_m \subseteq A_{n_1}$: $d(x_n, x_m) \leq d_{n_1} < \varepsilon$. Die Folge $(a_n)_{n=1}^\infty$ ist also eine Cauchy–Folge in (X, d) und daher wegen der Vollständigkeit von (X, d) konvergent gegen ein $x \in X$.

Wir zeigen $x \in \bigcap\limits_{n=1}^{\infty} A_n$: Sei $n \in \mathbb{N}$ beliebig und U eine beliebige Umgebung von x. Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U$. Zu ε existiert ein $k_0 \in \mathbb{N}$ mit $d(x, a_k) < \varepsilon$ für alle $k \geq k_0$. Für $k := \max\{k_0, n\}$ gilt also $a_k \in A_k \cap U_\varepsilon(x) \subseteq A_n \cap U$. Es folgt $x \in \overline{A_n} = A_n$ für alle $n \in \mathbb{N}$ und damit $x \in \bigcap\limits_{n=1}^{\infty} A_n$.

Zur Eindeutigkeit: Sind $x, y \in \bigcap\limits_{n=1}^{\infty} A_n$, so folgt für alle $n \in \mathbb{N}$: $x, y \in A_n$ und daher $0 \leq d(x, y) \leq d_n \to 0$ für $n \to \infty$. Daher gilt $d(x, y) = 0$ und somit $x = y$. □

8.12. Lemma. *Sei (X, d) ein metrischer Raum und $\emptyset \neq M \subseteq X$. Für $G \subseteq M$ sind äquivalent:*

(a) *G ist offen in $(M, d|_{M\times M})$.*
(b) *Es gibt eine in (X, d) offene Menge $G' \subseteq X$ mit $G = G' \cap M$.*

Beweis. Für alle $x \in M$ und alle $\varepsilon > 0$ bezeichne

$$U_\varepsilon(x) = \{y \in X\,;\, d(x,y) < \varepsilon\} \text{ die } \varepsilon\text{–Umgebung von } x \text{ in } (X,d) \text{ und}$$
$$V_\varepsilon(x) = \{y \in M\,;\, d(x,y) < \varepsilon\} \text{ die } \varepsilon\text{–Umgebung von } x \text{ in } (M, d|_{M\times M}).$$

Offensichtlich gilt $V_\varepsilon(x) = U_\varepsilon(x) \cap M$.

"(a)$\Longrightarrow$(b)": Ist $G \subseteq M$ offen in $(M, d|_{M\times M})$, so gibt es zu jedem $x \in G$ ein $\varepsilon(x) > 0$ mit $V_{\varepsilon(x)}(x) \subseteq G$. Es folgt

$$G = \bigcup_{x\in G} V_{\varepsilon(x)}(x) = \bigcup_{x\in G} (U_{\varepsilon(x)}(x) \cap M) = \Big(\bigcup_{x\in G} U_{\varepsilon(x)}(x)\Big) \cap M = G' \cap M\,,$$

wobei $G' := \bigcup_{x\in G} U_{\varepsilon(x)}(x)$ als Vereinigung von offenen Mengen offen ist in (X, d).

"(b)$\Longrightarrow$(a)": Ist $G = G' \cap M$ mit einer in (X, d) offenen Menge $G' \subseteq X$, so gibt es zu jedem $x \in G$ ein $\varepsilon > 0$ mit $x \in U_\varepsilon(x) \subseteq G'$. Also ist $V_\varepsilon(x) \subseteq G' \cap M = G$. Dies zeigt, daß G offen in $(M, d|_{M\times M})$ ist. $\square$

2. Kompakte metrische Räume

8.13. Definition. Sei (X, d) ein metrischer Raum und $M \subseteq X$.

(a) Eine Familie $(G_i)_{i\in I}$ von Teilmengen von X heißt eine *Überdeckung von M*, falls $M \subseteq \bigcup_{i\in I} G_i$. Wir sprechen von einer *offenen Überdeckung*, falls alle G_i, $i \in I$, offen in (X, d) sind. Eine Überdeckung $(G_i)_{i\in I}$ von M heißt *endlich*, falls I eine endliche Menge ist. Ist $(G_i)_{i\in I}$ eine Überdeckung von M und $J \subseteq I$ mit $M \subseteq \bigcup_{j\in J} G_j$, so nennen wir $(G_j)_{j\in J}$ eine *Teilüberdeckung* von $(G_i)_{i\in I}$.
(b) (X, d) heißt *kompakt*, falls jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt. Eine nicht leere Teilmenge M eines metrischen Raums (X, d) heißt *kompakt*, falls $(M, d|_{M\times M})$ ein kompakter metrischer Raum ist.

8.14. Lemma. *Eine nicht leere Teilmenge K eines metrischen Raums (X, d) ist genau dann kompakt, wenn jede in (X, d) offene Überdeckung von K eine endliche Teilüberdeckung besitzt.*

Beweis. "$\Longrightarrow$": Sei $\emptyset \neq K \subseteq X$ und K kompakt, d.h. $(K, d|_{K\times K})$ ist ein kompakter metrischer Raum. Sei $(G_i)_{i\in I}$ eine beliebige in (X, d) offene Überdeckung von K. Dann ist $(G_i \cap K)_{i\in I}$ nach Lemma 8.12 eine in $(K, d|_{K\times K})$ offene Überdeckung von K, besitzt also nach Voraussetzung eine endliche Teilüberdeckung $(G_j \cap K)_{j\in J}$ ($J \subseteq I$ eine endliche Teilmenge von I). $(G_j)_{j\in J}$ ist dann eine endliche Teilüberdeckung der Überdeckung $(G_i)_{i\in I}$ von K.

"$\Longleftarrow$": Sei nun (b) vorausgesetzt und sei $(G_i)_{i\in I}$ eine beliebige in $(K, d|_{K\times K})$ offene Überdeckung von K. Nach Lemma 8.12 gibt es in (X, d) offene Mengen G_i', $i \in I$, mit $G_i' \cap K = G_i$ für alle $i \in I$. $(G_i')_{i\in I}$ ist dann eine in (X, d) offene Überdeckung von K, aus der wir nach Voraussetzung eine endliche Teilüberdeckung $(G_j')_{j\in J}$ auswählen können. $(G_j)_{j\in J}$ ist dann eine endliche Teilüberdeckung der Überdeckung $(G_i)_{i\in I}$ von K. Damit ist gezeigt, daß $(K, d|_{K\times K})$ ein kompakter metrischer Raum ist. $\square$

8.15. Definition. Ein metrischer Raum (X, d) heißt *präkompakt*, falls es zu jedem $\varepsilon > 0$ endlich viele $x_1, \ldots, x_n \in X$ gibt mit

$$X \subseteq \bigcup_{j=1}^{n} U_\varepsilon(x_j) .$$

$\{x_1, \ldots, x_n\}$ heißt dann ein endliches ε–Netz für X. Eine nicht leere Teilmenge M eines metrischen Raum (X, d) heißt *präkompakt*, falls $(M, d|_{M\times M})$ ein präkompakter metrischer Raum ist.

Beispiel. $((0, 1), d)$ mit $d(x, y) := |x - y|$ für alle $x, y \in (0, 1)$ ist präkompakt aber nicht kompakt.

Beweis. Zu beliebigem $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $1/n < \varepsilon$. $\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\}$ ist dann ein endliches ε–Netz. Der metrische Raum $((0, 1), d)$ ist jedoch nicht kompakt, denn $\big((\frac{1}{k}, \frac{k-1}{k})\big)_{k=2}^{\infty}$ ist eine offene Überdeckung von $(0, 1)$, die keine endliche Teilüberdeckung zuläßt. $\square$

Im folgenden Satz geben wir einige äquivalente Beschreibungen der kompakten metrischen Räume an.

8.16. Satz. *Für einen metrischen Raum (X, d) sind die folgenden Aussagen äquivalent:*

(a) *(X, d) ist ein kompakter metrischer Raum.*
(b) *(X, d) hat die endliche Durchschnittseigenschaft, d.h. für jede Familie $(A_i)_{i\in I}$ von in (X, d) abgeschlossenen Mengen mit $\bigcap_{i\in I} A_i = \emptyset$ gibt es bereits eine endliche Teilmenge J von I mit $\bigcap_{j\in J} A_j = \emptyset$.*
(c) *Jede Folge aus (X, d) besitzt eine konvergente Teilfolge. (Man sagt auch: (X, d) ist folgenkompakt.)*
(d) *(X, d) ist präkompakt und vollständig.*

Beweis. "(a)$\Longrightarrow$(b)": Sei (X, d) kompakt und sei $(A_i)_{i\in I}$ eine Familie von in (X, d) abgeschlossenen Mengen mit $\bigcap_{i\in I} A_i = \emptyset$. Dann gilt

$$\bigcup_{i\in I}(X \setminus A_i) = X \setminus \bigcap_{i\in I} A_i = X .$$

Durch $\big(X \setminus A_i\big)_{i\in I}$ ist also eine offene Überdeckung von X gegeben, aus der wir wegen der Kompaktheit von (X, d) eine endliche Teilüberdeckung $\{X \setminus A_{i_1}, \ldots, X \setminus A_{i_n}\}$ auswählen können. Es ist dann

$$X = \bigcup_{k=1}^{n}(X \setminus A_{i_k}) = X \setminus \bigcap_{k=1}^{n} A_{i_k} .$$

Also ist $\bigcap_{k=1}^{n} A_{i_k} = \emptyset$. Damit ist gezeigt, daß (X, d) die endliche Durchschnittseigenschaft hat.

"(b)$\Longrightarrow$(c)": Sei nun (b) erfüllt und sei $(x_n)_{n=1}^{\infty}$ eine beliebige Folge in (X, d). Dann ist für alle $n \in \mathbb{N}$ die Menge

$$A_n := \overline{\{x_k\,;\, k \geq n\}}$$

eine in (X, d) abgeschlossene Menge mit $A_{n+1} \subseteq A_n$ für alle $n \in \mathbb{N}$. Wäre $\bigcap_{n=1}^{\infty} A_n = \emptyset$, so gäbe es nach Voraussetzung endlich viele $n_1, \ldots, n_r \in \mathbb{N}$ mit $\bigcap_{j=1}^{r} A_{n_j} = \emptyset$ im Widerspruch dazu, daß $x_n \in A_n \subseteq \bigcap_{j=1}^{r} A_{n_j}$ für alle $n \geq m := \max\{n_1, \ldots, n_r\}$. Also gilt $\bigcap_{n=1}^{\infty} A_n \neq \emptyset$.

Sei nun $x \in \bigcap_{n=1}^{\infty} A_n$ beliebig. Wir konstruieren induktiv eine gegen x konvergente Teilfolge $(x_{n_k})_{k=1}^{\infty}$ von $(x_n)_{n=1}^{\infty}$. Zunächst gibt es wegen $x \in A_1$ ein $n_1 \in \mathbb{N}$ mit $x_{n_1} \in U_1(x) \cap A_1$. Seien nun schon $n_1, \ldots, n_k \in \mathbb{N}$ konstruiert mit

(i) $n_j < n_{j+1}$ für alle $j \in \{i \in \mathbb{N}\,;\, 1 \leq i < k\}$.
(ii) $d(x, x_{n_j}) < \frac{1}{j}$ für alle $j \in \{i \in \mathbb{N}\,;\, 1 \leq i \leq k\}$.

Wegen $x \in A_{n_k+1}$ gibt es dann ein $n_{k+1} \in \mathbb{N}$ mit $n_{k+1} \geq n_k + 1 > n_k$ und $x_{n_{j+1}} \in U_{1/(k+1)}(x) \cap A_{n_k+1}$ also mit $d(x, x_{n_{j+1}}) < \frac{1}{j+1}$. Dann sind die Bedingungen (i) und (ii) auch für $k+1$ erfüllt. Wegen $d(x, x_{n_j}) < \frac{1}{j} \to 0$ für $j \to \infty$ konvergiert die so konstruierte Teilfolge gegen x. Also ist (c) erfüllt.

"(c)$\Longrightarrow$(d)": Sei (c) erfüllt. Wir zeigen zunächst, daß (X, d) dann vollständig ist. Sei also $(x_n)_{n=1}^{\infty}$ eine beliebige Cauchy–Folge in (X, d). Nach Voraussetzung besitzt diese eine gegen ein $x \in X$ konvergente Teilfolge $(x_{n_k})_{k=1}^{\infty}$. Wir zeigen, daß dann auch die Folge $(x_n)_{n=1}^{\infty}$ gegen x konvergiert. Sei hierzu $\varepsilon > 0$ beliebig. Wegen $x_{n_k} \to x$ für $k \to \infty$ gibt es ein $k_0 \in \mathbb{N}$ mit

$$\forall k \geq k_0 : \qquad d(x, x_{n_k}) < \frac{\varepsilon}{2}\,.$$

Da $(x_n)_{n=1}^{\infty}$ eine Cauchy–Folge in (X, d) ist, gibt es auch ein $n_0 \in \mathbb{N}$ mit

$$\forall n, m \geq n_0 : \qquad d(x_n, x_m) < \frac{\varepsilon}{2}\,.$$

Für alle $m \geq m_0 := \max\{k_0, n_0\}$ gilt dann nach der Dreiecksungleichung:

$$d(x, x_m) \leq d(x, x_{n_m}) + d(x_{n_m}, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also folgt $x_n \to x$ für $n \to \infty$. Damit ist die Vollständigkeit von (X, d) gezeigt.

Annahme: (X, d) ist nicht präkompakt, d.h.

$$\exists \varepsilon > 0 \quad \forall m \in \mathbb{N} \quad \forall x_1, \ldots, x_m \in X : \qquad \bigcup_{j=1}^{m} U_\varepsilon(x_j) \neq X\,. \tag{8.5}$$

Wir konstruieren induktiv eine Folge $(u_n)_{n=1}^{\infty}$ in X mit

$$d(x_n, x_m) \geq \varepsilon \qquad \text{für alle } n, m \in \mathbb{N} \text{ mit } n \neq m\,. \tag{8.6}$$

Sei hierzu $u_1 \in X$ ein beliebiger Punkt. Sind $u_1, \ldots, u_k \in X$ schon konstruiert, so daß (8.6) für alle $n, m \leq k$ aus $\mathbb{N}$ mit $n \neq m$ erfüllt ist, so gibt es wegen (8.5) ein $u_{k+1} \in X$ mit

$$u_{k+1} \notin \bigcup_{n=1}^{k} U_\varepsilon(u_n)\,.$$

Damit ist (8.6) auch für alle $n, m \leq k+1$ aus $\mathbb{N}$ mit $n \neq m$ erfüllt. Nach Voraussetzung besitzt die so induktiv konstruierte Folge $(u_n)_{n=1}^{\infty}$ eine in (X, d) konvergente Teilfolge $(u_{n_j})_{j=1}^{\infty}$. Diese ist insbesondere eine Cauchy–Folge im Widerspruch dazu, daß wegen (8.6) für alle $i, j \in \mathbb{N}$ mit $i \neq j$ gilt: $d(x_{n_i}, x_{n_j}) \geq \varepsilon$. Also hat die obige Annahme zu einem Widerspruch geführt. Der Raum (X, d) muß daher doch präkompakt sein.

"(d)$\Longrightarrow$(a)": Sei nun (X, d) ein vollständiger und präkompakter metrischer Raum. Dann gibt es zu jedem $n \in \mathbb{N}$ endlich viele $x_{1,n}, \ldots, x_{k(n),n}$ mit

$$X \subseteq \bigcup_{j=1}^{k(n)} U_{1/n}(x_{j,n}) \subseteq \bigcup_{j=1}^{k(n)} B_{1/n}(x_{j,n})$$

mit $B_{1/n}(x_{j,n}) := \overline{U_{1/n}(x_{j,n})}$.

Annahme: Es gibt eine offene Überdeckung $(U_i)_{i \in I}$ von X, die keine endliche Teilüberdeckung besitzt. Sei $A_0 := X$. Wir konstruieren induktiv eine Folge $(A_n)_{n=1}^{\infty}$ von abgeschlossenen Teilmengen von X, so daß für alle $j \in \mathbb{N}$ gilt:

(i) $A_{j-1} \supseteq A_j = \overline{A_j} \neq \emptyset$.

(ii) $d_j := \operatorname{diam} A_j \leq \frac{2}{j}$.

(iii) Die Überdeckung $(U_i)_{i\in I}$ von A_j besitzt keine endliche Teilüberdeckung.

Wegen $X = \bigcup_{j=1}^{k(1)} B_1(x_{j,1})$ gibt es ein $j_1 \in \{1,\dots,k(1)\}$, so daß die Überdeckung $(U_i)_{i\in I}$ von $B_1(x_{j_1,1})$ keine endliche Teilüberdeckung besitzt. Mit $A_1 := B_1(x_{j_1,1})$ sind dann die Bedingungen (i)–(iii) für $j = 1$ erfüllt. Seien nun schon die Mengen $A_1,\dots,A_n$ so konstruiert, daß die Bedingungen (i)–(iii) für alle $j = 1,\dots,n$ erfüllt sind. Wegen

$$X = \bigcup_{j=1}^{k(n+1)} B_{1/n+1}(x_{j,n+1})$$

folgt

$$A_n = \bigcup_{j=1}^{k(n+1)} (B_{1/n+1}(x_{j,n+1}) \cap A_n)\,.$$

Da (iii) für $j = n$ erfüllt ist, gibt es ein $j_{n+1} \in \{1,\dots,k(n+1)\}$, so daß die Überdeckung $(U_i)_{i\in I}$ von $A_{n+1} := B_{1/(n+1)}(x_{j_{n+1},n+1}) \cap A_n$ keine endliche Teilüberdeckung besitzt. A_{n+1} erfüllt nach Konstruktion offensichtlich die Bedingungen (i) und (iii) für $j = n+1$. Mit der Dreiecksungleichung rechnet man nach, daß auch (ii) für $j = n+1$ gilt.

Die so definierte Folge $(A_n)_{n=1}^\infty$ erfüllt die Voraussetzungen zum Cantorschen Durchschnittssatz. Da (X,d) nach Voraussetzung vollständig ist, gibt es also genau ein Element $x \in \bigcap_{n=1}^\infty A_n$. Wegen $X = \bigcup_{i\in I} U_i$ gibt es ein $i_0 \in I$ mit $x \in U_{i_0}$. Da U_{i_0} offen ist gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U_{i_0}$. Sei nun $n \in \mathbb{N}$ beliebig mit $n > 2/\varepsilon$. Dann folgt für alle $y \in A_n$: $d(x,y) \leq d_n \leq 2/n < \varepsilon$. Es folgt daher $A_n \subseteq U_\varepsilon(x) \subseteq U_{i_0}$ im Widerspruch dazu, daß die Überdeckung $(U_i)_{i\in I}$ von A_n keine endliche Teilüberdeckung besitzt. Die Annahme war also falsch. Es folgt, daß jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt, d.h. daß (X,d) ein kompakter metrischer Raum ist. □

8.17. Folgerung. *Jede kompakte Teilmenge K eines metrischen Raums (X,d) ist vollständig und abgeschlossen.*

Beweis. Nach Satz 8.16 ist $(K, d|_{K\times K})$ vollständig und nach der zweiten Bemerkung nach Definition 8.7 somit abgeschlossen in (X,d). □

8.18. Folgerung. *Jede abgeschlossene Teilmenge A eines kompakten metrischen Raums (X,d) ist kompakt.*

Beweis. Sei $(x_n)_{n=1}^\infty$ eine beliebige Folge aus A. Da (X,d) kompakt ist, besitzt diese eine in (X,d) konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Insbesondere ist $(x_{n_k})_{k=1}^\infty$ eine Cauchy–Folge in (X,d), deren Glieder in A liegen, also auch eine Cauchy–Folge in $(A, d|_{A\times A})$. Nach Satz 8.16 ist (X,d) vollständig. Nach der Bemerkung unmittelbar vor der Definition 8.8 ist daher auch $(A, d|_{A\times A})$ vollständig und somit die Folge $(x_{n_k})_{k=1}^\infty$ konvergent in $(A, d|_{A\times A})$. Da jede Folge aus A eine in $(A, d|_{A\times A})$ konvergente Teilfolge besitzt, ist $(A, d|_{A\times A})$ nach Satz 8.16 kompakt. □

8.19. Satz. *Seien (X,d) und (Y,d') zwei metrische Räume und sei $f : X \to Y$ eine stetige Abbildung. Dann ist für jede kompakte Teilmenge K von (X,d) auch die Menge $f(K)$ kompakt in (Y,d').*

Beweis. Sei $(U_i)_{i\in I}$ eine beliebige offene Überdeckung von $f(K)$. Nach Lemma 8.10 ist dann $f^{-1}(U_i)$ offen in (X,d) für alle $i \in I$. Wegen

$$K \subseteq \bigcup_{i\in I} f^{-1}(U_i)$$

ist $(f^{-1}(U_i))_{i\in I}$ eine in (X,d) offene Überdeckung von K, aus der wir wegen der Kompaktheit von K eine endliche Teilüberdeckung $f^{-1}(U_{i_1}),\dots,f^{-1}(U_{i_n})$ auswählen können. Es folgt

$$f(K)\subseteq\bigcup_{j=1}^{n}U_{i_j}\,.$$

$(U_{i_j})_{j=1}^n$ ist also eine endliche Teilüberdeckung der Überdeckung $(U_i)_{i\in I}$ von $f(K)$. Nach Lemma 8.14 ist $f(K)$ kompakt. □

Sei $(E,\|\cdot\|)$ ein normierter Raum. Eine Menge $B\subseteq E$ heißt *beschränkt*, falls

$$\sup_{b\in B}\|b\|<\infty\,.$$

8.20. Lemma. *Jede kompakte Teilmenge K eines normierten Raums $(E,\|\cdot\|)$ ist abgeschlossen und beschränkt.*

Beweis. Nach Folgerung 8.17 ist K abgeschlossen. Durch $(U_n(0)\cap K)_{n=1}^\infty$ ist eine in $(K,d_{\|\cdot\|}|_{K\times K})$ offene Überdeckung gegeben, aus der wir wegen der Kompaktheit von K eine endliche Teilüberdeckung $(U_n(0)\cap K)_{n=1}^m$ auswählen können. Es folgt $K\subseteq\bigcup_{n=1}^m U_n(0)=U_m(0)$ und damit

$$\sup_{x\in K}\|x\|\le m<\infty\,.$$

K ist also beschränkt. □

In $\mathbb{R}^N$ gilt auch die Umkehrung:

Bemerkung. Eine Teilmenge $K\neq\emptyset$ des $\mathbb{R}^N$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Beweis. Wegen Lemma 8.20 ist nur die Rückrichtung zu zeigen. Sei also $\emptyset\neq K\subset\mathbb{R}^N$ beschränkt und abgeschlossen. Ist $(x_n)_{n=1}^\infty$ eine beliebige Folge aus K, so ist diese beschränkt, besitzt also nach dem Satz 3.34 von Bolzano und Weierstraß eine in $\mathbb{R}^N$ konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Diese ist Cauchy–Folge in $\mathbb{R}^N$ also auch in $(K,d_{|\cdot|}|_{K\times K})$. Da K als abgeschlossene Teilmenge des $\mathbb{R}^N$ vollständig ist (vergl. die Bemerkung unmittelbar vor Definition 8.8), ist $(x_{n_k})_{k=1}^\infty$ auch in $(K,d_{|\cdot|}|_{K\times K})$ konvergent. K erfüllt also die Bedingung (c) in Satz 8.16 und ist daher nach diesem Satz kompakt. □

Da wir $\mathbb{C}^N$ mit $\mathbb{R}^{2N}$ identifizieren können, gilt auch:

Folgerung. *Eine nicht leere Teilmenge des $\mathbb{C}^N$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Wir wollen dieses Ergebnis auf alle endlich dimensionalen, normierten Räume übertragen. Hierzu zeigen wir zunächst:

8.21. Satz. *Sei $(E,\|\cdot\|)$ ein normierter Vektorraum über $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$ mit $N:=\dim E<\infty$. Dann ist $(E,\|\cdot\|)$ topologisch isomorph zu $(\mathbb{K}^N,|\cdot|)$, d.h. es gibt eine lineare Homöomorphie $T:\mathbb{K}^N\to E$.*

Beweis. Sei $e_1,\dots,e_N$ eine Basis von E. Dann ist die Abbildung

$$T:\mathbb{K}^N\to E\,,\qquad (z_1,\dots,z_N)\mapsto T(z_1,\dots,z_N):=\sum_{j=1}^N z_je_j\,,$$

ein $\mathbb{K}$–Vektorraumisomorphismus. Wie man unmittelbar nachrechnet, ist durch

$$|||\cdot|||:\mathbb{K}^N\to[0,\infty)\,,\qquad (z_1,\dots,z_N)\mapsto|||(z_1,\dots,z_N)|||:=\|T(z_1,\dots,z_N)\|$$

eine Norm auf $\mathbb{K}^N$ gegeben. In den Übungen (Aufgabe 46) wird gezeigt, daß diese zum euklidischen Betrag äquivalent ist. Es gibt also Konstanten $c > 0$ und $C > 0$ mit

$$\forall z = (z_1, \dots, z_N) \in \mathbb{K}^N : \qquad c|z| \le |||z||| = \|T(z)\| \le C|z| . \tag{8.7}$$

Aus der rechten Ungleichung folgt für alle $u, v \in \mathbb{K}^N$:

$$\|T(u) - T(v)\| = \|T(u - v)\| \le C|u - v|$$

T ist also Lipschitz–stetig und somit stetig (vergl. die Bemerkung vor Lemma 8.10). Die Abbildung $T^{-1} : E \to \mathbb{K}^N$ ist ebenfalls linear und erfüllt wegen (8.7)

$$\forall x \in E : |T^{-1}(x)| \le \frac{1}{c}\|T(T^{-1}(x))\| = \frac{1}{c}\|x\| . \tag{8.8}$$

Daher ist auch T^{-1} Lipschitz–stetig und somit stetig, denn für alle $x, y \in E$ gilt

$$|T^{-1}(x) - T^{-1}(y)| = |T^{-1}(x - y)| \le \frac{1}{c}\|x - y\| .$$

T ist also eine lineare Homöomorphie. □

8.22. Folgerung. *Sei $(E, \|\cdot\|)$ ein endlich dimensionaler normierter Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für $\emptyset \ne K \subseteq E$ sind die beiden folgenden Aussagen äquivalent:*

(a) *K ist beschränkt und abgeschlossen.*
(b) *K ist kompakt.*

Beweis. "(a)$\Longrightarrow$(b)" Sei K beschränkt und abgeschlossen und sei T wie im Beweis zu Satz 8.21. Da T stetig ist, ist $\tilde{K} := T^{-1}(K)$ abgeschlossen und wegen (8.8) folgt aus der Beschränktheit von K die von $\tilde{K}$. Wie vor Satz 8.21 gezeigt wurde, ist $\tilde{K}$ daher kompakt. Wegen der Stetigkeit von T ist nach Satz 8.19 auch $K = T(\tilde{K})$ kompakt in $(E, \|\cdot\|)$. □

Im unendlichdimensionalen Fall ist diese Äquivalenzaussage nicht mehr gültig. Bevor wir dies zeigen stellen wir ein Lemma von F. Riesz bereit:

8.23. Lemma (von Riesz). *Sei $(E, \|\cdot\|)$ ein normierter Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und sei $F \subsetneq E$ ein echter abgeschlossener Untervektorraum von E. Dann gibt es zu jedem $\varepsilon \in (0, 1)$ ein $x \in E \setminus F$ mit*

$$\|x\| = 1 \quad \textit{und} \quad \inf_{y \in F} \|x - y\| \ge 1 - \varepsilon .$$

Beweis. Wegen $F \ne E$ gibt es ein $x_0 \in E \setminus F$. Da $E \setminus F$ offen ist, folgt

$$d := \inf_{y \in F} \|x_0 - y\| > 0 .$$

Wegen $d(1+\varepsilon) > d$ gibt es nach Definition des Infimums ein $y_0 \in F$ mit $\|x_0 - y_0\| < d(1+\varepsilon)$. Der Vektor

$$x := \frac{1}{\|x_0 - y_0\|}(x_0 - y_0)$$

erfüllt dann $\|x\| = 1$. Da F ein Untervektorraum ist und $x_0 \notin F$, folgt wegen

$$x_0 = y_0 + \|x_0 - y_0\| x ,$$

daß x nicht in F liegt. Ist nun $u \in F$ beliebig, so folgt

$$\begin{aligned}\|x - u\| &= \left\| \frac{1}{\|x_0 - y_0\|}(x_0 - y_0) - u \right\| = \frac{1}{\|x_0 - y_0\|} \|x_0 - \underbrace{(y_0 + \|x_0 - y_0\| u)}_{\in F}\| \ge \\ &\ge \frac{d}{\|x_0 - y_0\|} > \frac{1}{1+\varepsilon} > 1 - \varepsilon\end{aligned}$$

Also gilt $\inf_{u\in F}\|x-u\| \geq 1-\varepsilon$. □

8.24. Lemma. *Sei $(E,\|\cdot\|)$ ein endlich dimensionaler $\mathbb{K}$–Vektorraum ($\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$). Dann ist $(E,\|\cdot\|)$ vollständig.*

Beweis. Sei also $\dim E = N < \infty$. Nach Satz 8.21 und dem zugehörigen Beweis gibt es einen stetigen $\mathbb{K}$–Vektorraumisomorphismus $T:\mathbb{K}^N \to E$ und Konstanten $c>0$ und $C>0$ mit

$$c|z| \leq \|T(z)\| \leq C|z| \qquad \text{für alle } z\in\mathbb{K}^N.$$

Sei nun $(x_n)_{n=1}^\infty$ eine beliebige Cauchy–Folge in $(E,\|\cdot\|)$ und sei $\varepsilon>0$ beliebig. Dann gibt es ein $n_0\in\mathbb{N}$, so daß für alle $n,m\geq n_0$ gilt: $\|x_n-x_m\|<c\varepsilon$. Hieraus folgt $|T^{-1}(x_n)-T^{-1}(x_m)|<\varepsilon$ für alle $n,m\geq n_0$. Die Folge $(T^{-1}(x_n))_{n=1}^\infty$ ist also eine Cauchy–Folge in $\mathbb{K}^N$ und daher konvergent in $\mathbb{K}^N$ gegen ein $z_0\in\mathbb{K}^N$. Wegen der Stetigkeit von T gilt dann $x_n = T(T^{-1}(x_n)) \to T(z_0)$ für $n\to\infty$. Damit ist gezeigt, daß jede Cauchy–Folge aus $(E,\|\cdot\|)$ in $(E,\|\cdot\|)$ konvergent ist. $(E,\|\cdot\|)$ ist daher vollständig. □

8.25. Satz. *Für einen normierten $\mathbb{K}$–Vektorraum $(E,\|\cdot\|)$ sind die beiden folgenden Aussagen äquivalent:*

(a) $\dim E<\infty$.
(b) *Die Einheitskugel $B_1(0) := \{x\in E\,;\,\|x\|<1\}$ ist kompakt.*

Beweis. Da $B_1(0)$ beschränkt und abgeschlossen ist, ergibt sich die Aussage “(a)$\Longrightarrow$(b)” unmittelbar aus Folgerung 8.22.

“(b)$\Longrightarrow$(a)”: Sei nun E unendlichdimensional. Wir konstruieren induktiv eine Folge $(x_n)_{n=1}^\infty$ in $B_1(0)$ mit $\|x_n\|=1$ und $\|x_n-x_m\|\geq 1/2$ für alle $n,m\in\mathbb{N}$ mit $m\neq n$. Wir wählen zunächst einen beliebigen Vektor $x_1\in E$ mit $\|x_1\|=1$. Seien $x_1,\dots,x_n$ schon gefunden mit

$$\|x_j\|=1 \quad \text{und} \quad \|x_i-x_j\|\geq\frac{1}{2} \quad \text{für alle } i,j\in\{1,\dots,n\} \text{ mit } i\neq j\,. \tag{8.9}$$

Dann hat der von $x_1,\dots,x_n$ erzeugte Untervektorraum F die Dimension $\leq n<\infty$ ist also nach Lemma 8.24 vollständig und daher nach der Bemerkung vor Definition 8.8 abgeschlossen. Nach dem Lemma 8.23 von Riesz gibt es also ein $x_{n+1}\in E\backslash F$ mit $\|x_{n+1}\|=1$ und $\inf_{u\in F}\|x_{n+1}-u\|\geq 1/2$. Insbesondere gilt also $\|x_{n+1}-x_j\|\geq 1/2$ für alle $j=1,\dots,n$. Daher ist nun (8.9) für $n+1$ statt n erfüllt.

Wäre nun $B_1(0)$ kompakt, so hätte die so induktiv konstruierte Folge $(x_n)_{n=1}^\infty$ nach Satz 8.16 eine konvergente Teilfolge $(x_{n_k})_{k=1}^\infty$. Insbesondere würde folgen $\|x_{n_k}-x_{n_{k+1}}\|\to 0$ im Widerspruch zu $\|x_{n_{k+1}}-x_{n_k}\|\geq 1/2$ für alle $k\in\mathbb{N}$. Also kann $B_1(0)$ nicht kompakt sein. □

8.26. Definition. Eine Teilmenge $M\neq\emptyset$ eines metrischen Raums (X,d) heißt *relativkompakt*, falls $\overline{M}$ kompakt ist.

8.27. Definition. Sei (X,d) ein metrischer Raum und sei $(x_n)_{n=1}^\infty$ eine Folge aus X. Ein Punkt $u\in X$ heißt *Häufungswert* der Folge $(x_n)_{n=1}^\infty$, falls diese eine gegen u konvergente Teilfolge besitzt.

8.28. Lemma. *Sei (X,d) ein metrischer Raum und $\emptyset\neq M\subseteq X$.*

(a) *Ist M relativkompakt, so ist M auch präkompakt.*
(b) *M ist genau dann relativkompakt, wenn jede Folge aus M wenigstens einen Häufungswert in X besitzt.*

Beweis. (a) Nach Voraussetzung ist $\overline{M}$ kompakt, also auch präkompakt. Es gibt also zu beliebig vorgegebenem $\varepsilon > 0$ endlich viele $x_1, \ldots, x_n \in \overline{M}$ mit

$$M \subseteq \overline{M} \subseteq \bigcup_{j=1}^{n} U_{\varepsilon/2}(x_j)\,.$$

Da die Punkte x_j Berührungspunkte von M sind, gibt es einen Punkt $u_j \in U_{\varepsilon/2}(x_j) \cap M$ für $j = 1, \ldots, n$. Dann folgt mit der Dreiecksungleichung

$$U_{\varepsilon/2}(x_j) \subseteq U_{\varepsilon}(u_j) \qquad \text{für } j = 1, \ldots, n$$

und damit

$$M \subseteq \bigcup_{j=1}^{n} U_{\varepsilon}(u_j)\,.$$

$u_1, \ldots, u_n$ ist also ein endliches ε–Netz für M.

(b) Sei M relativkompakt und sei $(x_n)_{n=1}^{\infty}$ eine beliebige Folge aus M. Da $\overline{M}$ kompakt ist besitzt diese eine in $\overline{M}$ und damit auch in (X, d) konvergente Teilfolge und somit auch einen Häufungswert in X.

Hat umgekehrt jede Folge aus M einen Häufungswert in X und ist $(x_n)_{n=1}^{\infty}$ eine beliebige Folge aus $\overline{M}$, so ist für alle $n \in \mathbb{N}$ der Punkt x_n Berührungspunkt von M. Es gibt also zu jedem $n \in \mathbb{N}$ ein $u_n \in U_{1/n}(x_n) \cap M$. Die Folge $(u_n)_{n=1}^{\infty}$ hat nach Voraussetzung eine in (X, d) gegen ein $v \in X$ konvergente Teilfolge $(u_{n_k})_{k=1}^{\infty}$. Es ist $v \in \overline{M}$ und

$$d(x_{n_k}, v) \leq d(x_{n_k}, u_{n_k}) + d(u_{n_k}, v) \leq \frac{1}{n_k} + d(u_{n_k}, v) \to 0 \quad \text{für } k \to \infty\,.$$

Also besitzt $(x_n)_{n=1}^{\infty}$ eine in $\overline{M}$ konvergente Teilfolge. Nach Satz 8.16 ist $\overline{M}$ also kompakt und M somit relativkompakt. $\square$

In vollständigen metrischen Räumen gilt in (a) auch die Umkehrung. (Beweis als Übung).

KAPITEL 9

Folgen und Reihen stetiger Funktionen

1. Punktweise und gleichmäßige Konvergenz

9.1. Definition. Sei X eine nicht leere Menge und sei (Y, d) ein metrischer Raum. Eine Folge $(f_n)_{n=1}^{\infty}$ von Abbildungen $f_n : X \to Y$ heißt

(a) *punktweise konvergent* auf X gegen eine Abbildung $f : X \to Y$, falls für alle $x \in X$ gilt: $f(x) = \lim_{n\to\infty} f_n(x)$, d.h.

$$\forall x \in X \,\forall \varepsilon > 0 \,\exists n_0 = n_0(x, \varepsilon) \in \mathbb{N} \,\forall n \geq n_0 : \quad d(f(x), f_n(x)) < \varepsilon\,. \tag{9.1}$$

(b) eine *punktweise Cauchy–Folge* auf X, falls für alle $x \in X$ die Folge $(f_n(x))_{n=1}^{\infty}$ eine Cauchy–Folge in (Y, d) ist, d.h. falls gilt:

$$\forall x \in X \,\forall \varepsilon > 0 \,\exists n_0 = n_0(x, \varepsilon) \in \mathbb{N} \,\forall n, m \geq n_0 : \quad d(f_n(x), f_m(x)) < \varepsilon\,. \tag{9.2}$$

(c) *gleichmäßig konvergent* auf X gegen eine Abbildung $f : X \to Y$, falls gilt:

$$\forall \varepsilon > 0 \,\exists n_0 = n_0(\varepsilon) \in \mathbb{N} \,\forall x \in X \,\forall n \geq n_0 : \quad d(f(x), f_n(x)) < \varepsilon\,. \tag{9.3}$$

(d) eine *gleichmäßige Cauchy–Folge*, falls gilt:

$$\forall \varepsilon > 0 \,\exists n_0 = n_0(\varepsilon) \in \mathbb{N} \,\forall x \in X \,\forall n, m \geq n_0 : \quad d(f_n(x), f_m(x)) < \varepsilon\,. \tag{9.4}$$

9.2. Bemerkungen. (a) Offensichtlich folgt (9.1) aus (9.3) und (9.2) aus (9.4). Jede auf X gleichmäßig konvergente Folge von Abbildungen ist also insbesondere punktweise konvergent auf X und jede gleichmäßige Cauchy–Folge von Abbildungen ist insbesondere eine punktweise Cauchy–Folge.

(b) Es gibt höchstens eine Abbildung $f : X \to Y$ mit (9.1) und daher auch höchstens eine Abbildung $f : X \to Y$ mit (9.3).

(c) Die Bedingung (9.3) für die gleichmäßige Konvergenz ist äquivalent zu

$$\forall \varepsilon > 0 \,\exists n_0 \in \mathbb{N} \,\forall n \geq n_0 : \quad \sup_{x\in X} d(f(x), f_n(x)) < \varepsilon \tag{9.5}$$

und die Bedingung (9.4) für das Vorliegen einer gleichmäßigen Cauchy–Folge ist äquivalent zu

$$\forall \varepsilon > 0 \,\exists n_0 \in \mathbb{N} \,\forall n, m \geq n_0 : \quad \sup_{x\in X} d(f_n(x), f_m(x)) < \varepsilon\,. \tag{9.6}$$

9.3. Beispiele. (a) Die Funktionen $f_n : \mathbb{R} \to \mathbb{R}$ seien definiert durch:

$$f_n(x) := \max\left\{0, n - n^2 \left| x - \frac{1}{n} \right| \right\} \qquad \text{für alle } x \in \mathbb{R}\,,\ n \in \mathbb{N}\,.$$

Offensichtlich gilt $f_n(x) \to 0$ für alle $x \in \mathbb{R}$, d.h. die Folge $(f_n)_{n=1}^{\infty}$ ist punktweise auf $\mathbb{R}$ konvergent gegen die Funktion $f \equiv 0$. $(f_n)_{n=1}^{\infty}$ ist aber nicht gleichmäßig gegen 0 konvergent, denn

$$\left| f_n\left(\frac{1}{n}\right) - 0 \right| = n \to \infty\,.$$

(b) Die Funktionen $g_n : [0,1] \to \mathbb{R}$ seien definiert durch $g_n(x) := x^n$ für alle $x \in [0,1], n \in \mathbb{N}$. Für alle $x \in [0,1]$ gilt dann für $n \to \infty$

$$g_n(x) \to g(x) := \begin{cases} 0 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x = 1\,. \end{cases}$$

Die Folge $(g_n)_{n=1}^{\infty}$ ist also punktweise auf $[0,1]$ konvergent gegen die Funktion g. Die Grenzfunktion g ist in 1 unstetig.

Bei gleichmäßiger Konvergenz ist die Situation besser. Gleichmäßige Limites Folgen stetiger Abbildungen sind stetig:

9.4. Satz. *Seien (X,d) und (Y,d') zwei metrische Räume und sei $(f_n)_{n=1}^{\infty}$ eine Folge von auf X gleichmäßig gegen eine Abbildung $f : X \to Y$ konvergenten stetigen Abbildungen $f_n : X \to Y$. Dann ist auch f auf X stetig.*

Beweis. Wir müssen zeigen, daß f in allen Punkten von X stetig ist. Sei also $x \in X$ beliebig und $\varepsilon > 0$ beliebig. Da die Folge $(f_n)_{n=1}^{\infty}$ gleichmäßig auf X gegen f konvergiert, gibt es zu $\varepsilon' := \varepsilon/3 > 0$ ein $n_0 \in \mathbb{N}$, so daß

$$\forall n \geq n_0 \, \forall u \in X : \qquad d'(f_n(u), f(u)) < \varepsilon' = \frac{\varepsilon}{3}\,.$$

Da die Abbildung f_{n_0} in x stetig ist, gibt es zu $\varepsilon' := \varepsilon/3 > 0$ ein $\delta > 0$, so daß

$$\forall u \in U_\delta(x) : \qquad d'(f_{n_0}(x), f_{n_0}(u)) < \varepsilon' = \frac{\varepsilon}{3}\,.$$

Für alle $u \in U_\delta(x)$ gilt dann nach der Dreiecksungleichung:

$$\begin{aligned} d'(f(x), f(u)) \leq & d'(f(x), f_{n_0}(x)) + d'(f_{n_0}(x), f_{n_0}(u)) + d'(f_{n_0}(u), f(u)) \\ < & \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,. \end{aligned}$$

f ist also in x stetig für alle $x \in X$. □

9.5. Satz. *Sei X eine nicht leere Menge, (Y,d) ein metrischer Raum und $(f_n)_{n=1}^{\infty}$ eine auf X gleichmäßig gegen eine Abbildung $f : X \to Y$ konvergente Folge von Abbildungen $f_n : X \to Y$. Dann ist $(f_n)_{n=1}^{\infty}$ eine gleichmäßige Cauchy–Folge auf X.*

Beweis. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ aus $\mathbb{N}$ gilt:

$$\forall x \in X : d(f(x), f_n(x)) < \frac{\varepsilon}{2}\,.$$

Für alle $n, m \geq n_0$ aus $\mathbb{N}$ gilt dann mit der Dreiecksungleichung:

$$\forall x \in X : d(f_m(x), f_n(x)) \leq d(f_m(x), f(x)) + d(f(x), f_n(x)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also ist $(f_n)_{n=1}^{\infty}$ eine gleichmäßige Cauchy–Folge auf X. □

Ist (Y,d) vollständig, so gilt auch die Umkehrung:

9.6. Satz (Cauchy–Kriterium für die gleichmäßige Konvergenz). *Sei X eine nicht leere Menge und (Y,d) ein vollständiger metrischer Raum. Ist $(f_n)_{n=1}^{\infty}$ eine gleichmäßige Cauchy–Folge von Abbildungen $f_n : X \to Y$, so gibt es genau eine Abbildung $f : X \to Y$, so daß $f_n \to f$ gleichmäßig auf X für $n \to \infty$.*

Beweis. Da $(f_n)_{n=1}^\infty$ insbesondere auch eine punktweise Cauchy–Folge ist, ist für alle $x \in X$ die Folge $(f_n(x))_{n=1}^\infty$ eine Cauchy–Folge in dem vollständigen metrischen Raum (Y, d) und daher in (Y, d) konvergent gegen einen Wert $f(x) \in Y$. Hierdurch ist eine Abbildung $f : X \to Y$ definiert. Zu zeigen ist noch die gleichmäßige Konvergenz der Folge $(f_n)_{n=1}^\infty$ gegen f. Sei also $\varepsilon > 0$ beliebig. Da $(f_n)_{n=1}^\infty$ eine gleichmäßige Cauchy–Folge ist, gibt es ein $n_0 \in \mathbb{N}$, so daß für alle $n, m \geq n_0$ gilt

$$\forall x \in X : \qquad d(f_n(x), f_m(x)) < \frac{\varepsilon}{2}\,.$$

Wegen $f_m(x) \to f(x)$ in (Y, d) für $m \to \infty$ und der Stetigkeit der Abbildung

$$d(f_n(x), \cdot) : Y \to \mathbb{R}\,, \qquad y \mapsto d(f_n(x), y),$$

erhalten wir mit Lemma 8.9 für $m \to \infty$ für alle $n \geq n_0$:

$$\forall x \in X : \qquad d(f_n(x), f(x)) \leq \frac{\varepsilon}{2} < \varepsilon\,.$$

Also ist $(f_n)_{n=1}^\infty$ gleichmäßig auf X konvergent gegen f für $n \to \infty$. □

Aus Satz 9.6 und Satz 9.4 erhalten wir unmittelbar:

9.7. Folgerung. *Seien (X, d) und (Y, d') zwei metrische Räume, sei (Y, d') vollständig und sei $(f_n)_{n=1}^\infty$ eine gleichmäßige Cauchy–Folge stetiger Abbildungen von X nach Y. Dann ist $(f_n)_{n=1}^\infty$ gleichmäßig auf X konvergent gegen eine eine stetige Abbildung $f : X \to Y$.*

Sind (X, d) und (Y, d') zwei metrische Räume, so bezeichnen wir mit $C(X, Y)$ die Menge aller stetigen Abbildungen von X nach Y. Ist X kompakt, so können wir $C(X, Y)$ mit einer natürlichen Metrik versehen:

9.8. Satz. *Sei (Y, d') ein metrischer Raum und sei (K, d) ein kompakter metrischer Raum. Dann ist durch*

$$d_K : C(K, Y) \times C(K, Y) \to [0, \infty)$$

mit

$$d_K(f, g) := \sup_{x \in K} d'(f(x), g(x)) \qquad \textit{für alle } f, g \in C(K, Y).$$

eine Metrik auf $C(K, Y)$ gegeben. Eine Folge $(f_n)_{n=1}^\infty$ aus $C(K, Y)$ ist genau dann eine Cauchy–Folge bzgl. d_K (bzw. genau dann bzgl. d_K gegen eine Abbildung $f \in C(K, Y)$ konvergent), wenn sie auf K eine gleichmäßige Cauchy–Folge ist (bzw. wenn sie auf K gleichmäßig gegen f konvergiert.)

Ist (Y, d') vollständig, so ist auch $(C(K, Y), d_K)$ vollständig.

Beweis. Sind $f, g \in C(K, Y)$, so ist, wie man leicht verifiziert, auch die Abbildung $x \mapsto (f(x), g(x))$ stetig von K nach $Y \times Y$, wobei wir $Y \times Y$ versehen mit der Metrik

$$\rho : (Y \times Y) \times (Y \times Y) \to [0, \infty)$$

die gegeben ist durch

$$\rho((u_1, u_2), (v_1, v_2)) := d'(u_1, v_1) + d'(u_2, v_2) \quad \text{für alle } u_1, u_2, v_1, v_2 \in Y.$$

Da $d' : Y \times Y \to [0, \infty)$ stetig ist (sogar gleichmäßig stetig, vergl. Teil (c) der Bemerkung nach der Definition der gleichmäßigen Stetigkeit und Lemma 8.9), ist die Abbildung $x \mapsto d'(f(x), g(x))$ als Hintereinanderausführung stetiger Abbildungen wieder eine stetige Abbildung und somit wegen der Kompaktheit von K auf K beschränkt. Also folgt $d_K(f, g) \in [0, \infty)$ für alle $f, g \in C(K, Y)$. Daß d_K eine Metrik ist, rechnet man nach. Nach Bemerkung 9.2 ist eine Folge $(f_n)_{n=1}^\infty$ aus $C(K, Y)$ ist genau dann eine Cauchy–Folge bzgl.

d_K (bzw. genau dann bzgl. d_K gegen eine Abbildung $f \in C(K,Y)$ konvergent), wenn sie auf K eine gleichmäßige Cauchy–Folge ist (bzw. wenn sie auf K gleichmäßig gegen f konvergiert.)

Die Vollständigkeitsaussage ergibt sich aus Folgerung 9.7. □

Einen vollständigen normierten Raum $(Y, \|\cdot\|)$ nennt man auch einen *Banachraum*.

9.9. Satz. *Sei (X,d) ein metrischer Raum und $(Y,\|\cdot\|)$ ein normierter $\mathbb{K}$–Vektorraum ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$). Dann ist $C(X,Y)$ bezüglich der punktweisen Operationen der Addition und der Multiplikation mit Skalaren ein $\mathbb{K}$–Vektorraum. Ist (X,d) kompakt, so wird $C(X,Y)$ durch*

$$\|\cdot\|_X : C(X,Y) \to [0,\infty)\,, \qquad f \mapsto \|f\|_X := \sup_{x\in X} \|f(x)\|$$

zu einem normierten Raum. Dieser ist vollständig, falls $(Y,\|\cdot\|)$ vollständig ist.

Beweis. Sind $f, g \in C(X,Y)$ und $\lambda \in \mathbb{K}$, so sind auch die Funktionen $f+g$ und λf stetig: Sei hierzu $\varepsilon > 0$ und $x \in X$ beliebig. Da f, g stetig in x sind, gibt es ein $\delta > 0$ mit

$$\forall u \in U_\delta(x): \qquad \|f(x) - f(u)\| < \frac{\varepsilon}{2} \text{ und } \|g(x) - g(u)\| < \frac{\varepsilon}{2}$$

bzw.

$$\forall u \in U_\delta(x): \qquad \|f(x) - f(u)\| < \frac{\varepsilon}{|\lambda|+1}\,.$$

Dann folgt für alle $u \in U_\delta(x)$:

$$\|(f+g)(x) - (f+g)(u)\| \le \|f(x) - f(u)\| + \|g(x) - g(u)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

bzw.

$$\|(\lambda f)(x) - (\lambda f)(u)\| = |\lambda| \cdot \|f(x) - f(u)\| \le |\lambda| \frac{\varepsilon}{|\lambda|+1} < \varepsilon\,.$$

Dies zeigt die Stetigkeit von $f+g$ und λf. $C(X,Y)$ ist also ein Untervektorraum des $\mathbb{K}$–Vektorraums aller Abbildungen von X nach Y.

Sei nun (X,d)ein kompakter metrischer Raum. Daß $\|\cdot\|_X$ eine Norm auf $C(X,Y)$ definiert rechnet man nach. Der Rest folgt durch Anwenden von Satz 9.8 auf $(Y, d_{\|\cdot\|})$. □

9.10. Definition. Sei (X,d) ein metrischer Raum und $(Y,\|\cdot\|)$ ein Banachraum. Eine Reihe $\sum_{n=0}^\infty f_n$ von Abbildungen $f_n : X \to Y$ heißt *gleichmäßig konvergent auf X* gegen eine Abbildung $f : X \to Y$, falls die Folgen der Partialsummen der Reihe gleichmäßig auf X gegen f konvergiert. Die Reihe $\sum_{n=0}^\infty f_n$ heißt *gleichmäßig absolut konvergent*, falls

$$\sum_{n=0}^\infty \sup_{x\in X} \|f_n(x)\| < \infty\,.$$

9.11. Satz. *Sei (X,d) ein metrischer Raum und sei $(Y,\|\cdot\|)$ ein Banachraum.*

(a) *Sind die Abbildungen $f_n : X \to Y$ stetig und ist $\sum_{n=0}^\infty f_n$ gleichmäßig konvergent auf X gegen eine Abbildung $f : X \to Y$, so ist f nach Satz 9.4 stetig.*
(b) *Sind die Abbildungen $f_n : X \to Y$ stetig und ist $\sum_{n=0}^\infty f_n$ absolut gleichmäßig konvergent auf X, so ist $\sum_{n=0}^\infty f_n$ gleichmäßig konvergent auf X gegen eine (nach (a) dann) stetige und beschränkte Abbildung $f : X \to Y$.*

Beweis. Zu (b): Nach Folgerung 9.7 genügt es zu zeigen, daß die Folge $(s_n)_{n=0}^\infty$ der Partialsummen $s_n = \sum_{k=0}^n f_n$, $n \in \mathbb{N}_0$, eine gleichmäßige Cauchy–Folge ist. Sei also $\varepsilon > 0$ beliebig. Wegen der Konvergenz der Reihe $\sum_{n=0}^\infty \sup_{x\in X} \|f_n(x)\|$ gibt es ein n_0, so daß für alle $n, m \geq n_0$ (mit O.B.d.A. $m \leq n$) gilt

$$\sum_{k=m}^n \sup_{x\in X} \|f_k(x)\| < \varepsilon\,.$$

Dann gilt für alle $n, m \geq n_0$ (mit $m \leq n$) nach der Dreiecksungleichung:

$$\forall u \in X: \quad \|s_n(u) - s_m(u)\| = \left\| \sum_{k=m+1}^n f_k(x) \right\| \leq \sum_{k=m+1}^n \sup_{x\in X} \|f_k(x)\| < \varepsilon\,.$$

Also ist $(s_n)_{n=0}^\infty$ eine gleichmäßige Cauchy–Folge. Nach Folgerung 9.7 konvergiert $(s_n)_{n=0}\infty$ also gleichmäßig auf X gegen eine stetige Abbildung $f : X \to Y$. Für alle $u \in X$ und alle $n \in \mathbb{N}_0$ gilt

$$\|s_n(u)\| = \left\| \sum_{k=0}^n f_k(u) \right\| \leq \sum_{k=0}^n \sup_{x\in X} \|f_k(x)\| \leq \sum_{k=0}^\infty \sup_{x\in X} \|f_k(x)\|$$

Also auch

$$\|f(u)\| = \lim_{n\to\infty} \|s_n(u)\| \leq \sum_{k=0}^\infty \sup_{x\in X} \|f_k(x)\|$$

wegen der Stetigkeit der Norm $\|\cdot\| : Y \to \mathbb{R}$. □

Als Folgerung erhalten wir:

9.12. Folgerung (Majorantenkriterium). *Sei (X, d) ein metrischer Raum und sei $(Y, \|\cdot\|)$ ein Banachraum. Ist $(f_n)_{n=0}^\infty$ eine Folge stetiger Abbildungen von X nach Y, für die es eine Folge $(c_n)_{n=0}^\infty$ in $[0,\infty)$ und ein $n_0 \in \mathbb{N}_0$ gibt, so daß die Reihe $\sum_{n=0}^\infty c_n$ in $\mathbb{R}$ konvergiert und so daß für alle $n \geq n_0$ gilt*

$$\sup_{x\in X} \|f_n(x)\| \leq c_n\,,$$

so ist die Reihe $\sum_{n=0}^\infty f_n$ gleichmäßig konvergent auf X gegen eine stetige Abbildung $f : X \to Y$.

Beweis. Die Reihe $\sum_{n=n_0}^\infty f_n$ ist offensichtlich gleichmäßig absolut konvergent, konvergiert also nach Satz 9.11 gleichmäßig auf X gegen eine beschränkte stetige Abbildung $g : X \to Y$. Die Gesamtreihe $\sum_{n=0}^\infty f_n$ ist daher gleichmäßig auf X konvergent gegen die stetige Abbildung $f := g + \sum_{n=0}^{n_0-1} f_k$. □

Eine *beidseitig* unendliche Reihe $\sum_{n=-\infty}^\infty a_n$ in einem normierten Raum $(Y, \|\cdot\|)$ heißt *konvergent*, falls die Reihen $\sum_{n=0}^\infty a_n$ und $\sum_{n=1}^\infty a_{-n}$ in $(Y, \|\cdot\|)$ konvergieren. Konvergiert die Reihe $\sum_{n=-\infty}^\infty \|a_n\|$ in $\mathbb{R}$, so nennen wir die Reihe $\sum_{n=-\infty}^\infty a_n$ *absolut konvergent.* Ist $(Y, \|\cdot\|)$ vollständig, so ist jede absolut konvergente Reihe in $(Y, \|\cdot\|)$ auch konvergent. Entsprechend definiert man punktweise Konvergenz, gleichmäßige Konvergenz und gleichmäßig absolute Konvergenz für beidseitig unendliche Reihen $\sum_{n=-\infty}^\infty f_n$ von Abbildungen $f_n : X \to Y$ zwischen einem metrischen Raum (X, d) und einem normierten Raum $(Y, \|\cdot\|)$.

Mit Satz 9.11 erhalten wir:

9.13. Folgerung. *Sei $\sum_{n=-\infty}^{\infty} a_n$ eine absolut konvergente Reihe in $\mathbb{C}$. Dann ist die durch*

$$t \mapsto f(t) := \sum_{n=-\infty}^{\infty} a_n \exp(int)$$

definierte Funktion $f : \mathbb{R} \to \mathbb{C}$ stetig und beschränkt auf $\mathbb{R}$. Die Reihe ist offensichtlich eine auf $\mathbb{R}$ gleichmäßig absolut konvergente Reihe stetiger Funktionen.

Funktionenreihen der Form $\sum_{n=-\infty}^{\infty} a_n \exp(int)$ nennt man auch *Fourierreihen.*

2. Vertauschung von Differentiation und Grenzwertbildung

9.14. Beispiel. Für alle $n \in \mathbb{N}$ und $x \in \mathbb{R}$ definieren wir

$$f_n(x) := \frac{1}{n} \sin(nx)\,.$$

In den Übungen wird gezeigt, daß die Folge $(f_n)_{n=1}^{\infty}$ gleichmäßig auf $\mathbb{R}$ gegen die Funktion $f \equiv 0$ konvergiert, daß aber die Folge $(f_n')_{n=1}^{\infty}$ der Ableitungen mit

$$f_n'(x) = \cos(nx) \qquad \text{für alle } x \in \mathbb{R}, n \in \mathbb{N}$$

noch nicht einmal punktweise gegen die Ableitung $f' \equiv 0$ der Grenzfunktion konvergiert.

Die Vertauschung von Grenzwertbildung und Differentiation wird also nur unter besonderen Voraussetzungen möglich sein.

9.15. Satz. *Sei $I = [a, b]$ ein nicht ausgeartetes, kompaktes Intervall und sei $(f_n)_{n=0}^{\infty}$ eine Folge von auf I differenzierbaren reellwertigen Funktionen, für die gilt*

(a) *Die Folge $(f_n')_{n=0}^{\infty}$ der Ableitungsfunktionen konvergiert gleichmäßig auf I gegen eine Funktion $g : I \to \mathbb{R}$.*

(b) *Es gibt ein $x_0 \in I$, so daß die reelle Folge $(f_n(x_0))_{n=0}^{\infty}$ in $\mathbb{R}$ konvergiert.*

Dann konvergiert $(f_n)_{n=0}^{\infty}$ gleichmäßig auf I gegen eine differenzierbare Funktion $f : I \to \mathbb{R}$ und es gilt $f' \equiv g$.

Beweis. Als differenzierbare Funktionen sind die Funktionen f_n insbesondere auf I stetig. Wir zeigen zunächst, daß sie gleichmäßig auf I gegen eine stetige Funktion $f : I \to \mathbb{R}$ konvergieren. Sei $\varepsilon > 0$ beliebig vorgegeben. Da die Folge $(f_n')_{n=0}^{\infty}$ der Ableitungsfunktionen gleichmäßig konvergent ist, ist sie insbesondere eine gleichmäßige Cauchy–Folge (vergl. Satz 9.5). Es gibt also ein $n_0 \in \mathbb{N}_0$, so daß gilt

$$\forall n \geq n_0 \quad \forall p \in \mathbb{N}_0 : \qquad \sup_{x \in I} |f_{n+p}'(x) - f_n'(x)| < \frac{\varepsilon}{2(b-a)}\,. \tag{9.7}$$

Da die Folge $(f_n(x_0))_{n=0}^{\infty}$ in $\mathbb{R}$ konvergiert, gibt es ein $n_1 \in \mathbb{N}_0$ mit

$$\forall n \geq n_1 \quad \forall p \in \mathbb{N}_0 : \qquad |f_{n+p}(x_0) - f_n(x_0)| < \frac{\varepsilon}{2}\,. \tag{9.8}$$

Für alle $n \geq m_0 := \max\{n_0, n_1\}$, alle $p \in \mathbb{N}_0$ und alle $x \in I$ erhalten wir unter Verwendung des Mittelwertsatzes der Differentialrechnung und unter Verwendung von (9.8) und (9.7):

$$\begin{aligned} |f_{n+p}(x) - f_n(x)| \leq& |f_{n+p}(x_0) - f_n(x_0)| + |f_{n+p}(x) - f_n(x) - (f_{n+p}(x_0) - f_n(x_0))| \\ <& \frac{\varepsilon}{2} + |x - x_0| \cdot |(f_{n+p} - f_n)'(\xi)| \\ \leq& \frac{\varepsilon}{2} + |x - x_0| \sup_{t \in I} |(f_{n+p} - f_n)'(t)| \\ \leq& \frac{\varepsilon}{2} + |x - x_0| \frac{\varepsilon}{2(b-a)} \leq \varepsilon\,. \end{aligned}$$

mit einem ξ zwischen x und x_0. Die Folge $(f_n)_{n=0}^\infty$ ist also eine gleichmäßige Cauchy–Folge auf I und konvergiert daher nach Folgerung 9.7 gleichmäßig auf I gegen eine stetige Funktion $f : I \to \mathbb{R}$.

Wir müssen noch zeigen, daß f auf I differenzierbar ist und daß $f'(u) = g(u)$ für alle $u \in I$ gilt. Sei also $u \in I$ beliebig und im folgenden fest. Für alle $n \in \mathbb{N}_0$ definieren wir die Funktion $h_n : I \to \mathbb{R}$ durch

$$h_n(x) := \begin{cases} \dfrac{f_n(u) - f_n(x)}{u - x} & \text{für } x \neq u, x \in I, \\ f_n'(u) & \text{für } x = u\,. \end{cases}$$

Nach Definition der Ableitung und des Grenzwertes ist h_n eine auf I stetige Funktion für alle $n \in \mathbb{N}_0$. Wir zeigen, daß die Folge $(h_n)_{n=0}^\infty$ eine gleichmäßige Cauchy–Folge ist. Sei hierzu $\varepsilon > 0$ beliebig. Nach Voraussetzung gibt es ein $n_2 \in \mathbb{N}_0$ mit

$$\forall n \geq n_2 \quad \forall p \in \mathbb{N}_0 \quad \forall x \in I : \qquad |f_{n+p}'(x) - f_n'(x)| < \varepsilon\,.$$

Für alle $n \geq n_2$, $p \in \mathbb{N}_0$ und alle $x \in I \setminus \{u\}$ gilt dann mit einem ξ zwischen x und u nach dem Mittelwertsatz der Differentialrechnung:

$$|h_{n+p}(x) - h_n(x)| = \left| \frac{(f_{n+p} - f_n)(u) - (f_{n+p} - f_n)(x)}{u - x} \right| = |f_{n+p}'(\xi) - f_n'(\xi)| < \varepsilon$$

und für $x = u$ erhalten wir ebenfalls

$$|h_{n+p}(u) - h_n(u)| = |f_{n+p}'(u) - f_n'(u)| < \varepsilon\,.$$

Die Folge $(h_n)_{n=0}^\infty$ ist also eine gleichmäßige Cauchy–Folge auf I und konvergiert daher gleichmäßig auf I gegen eine stetige Funktion $h : I \to \mathbb{R}$. Nach den Grenzwertrechenregeln gilt für alle $x \in I \setminus \{u\}$:

$$h(x) = \lim_{n\to\infty} h_n(x) = \lim_{n\to\infty} \frac{f_n(u) - f_n(x)}{u - x} = \frac{f(u) - f(x)}{u - x}\,.$$

Für $x = u$ erhalten wir

$$h(u) = \lim_{n\to\infty} h_n(u) = \lim_{n\to\infty} f_n'(u) = g(u)\,.$$

Da h in u stetig ist, folgt

$$\lim_{x\to u} \frac{f(u) - f(x)}{u - x} = \lim_{x\to u} h(x) = h(u) = g(u)\,.$$

f ist also in u differenzierbar und erfüllt $f'(u) = g(u)$. □

Auch in mehreren Veränderlichen kann man eine entsprechende Aussage beweisen:

9.16. Satz. *Sei $\emptyset \neq G \subset \mathbb{R}^N$ offen, beschränkt und konvex und sei $(f_n)_{n=0}^\infty$ eine Folge von Funktionen aus $C^1(G, \mathbb{R})$ mit den beiden folgenden Eigenschaften:*

(a) *Die Folge $(\nabla f_n)_{n=0}^\infty$ konvergiert gleichmäßig auf G gegen ein Vektorfeld $g : G \to \mathbb{R}^N$.*
(b) *Es gibt ein $x_0 \in G$, für das die Folge $(f_n(x_0))_{n=0}^\infty$ in $\mathbb{R}$ konvergent ist.*

Dann konvergiert die Folge $(f_n)_{n=0}^\infty$ gleichmäßig auf G gegen eine Funktion $f \in C^1(G, \mathbb{R})$ und es gilt $\nabla f \equiv g$ auf G.

Beweis. Nach Voraussetzung ist

$$0 < C := \sup_{x\in G} |x| < \infty\,.$$

Da die Folge $(\nabla f_n)_{n=0}^{\infty}$ gleichmäßig auf G konvergent ist, ist sie auch eine gleichmäßige Cauchy–Folge. Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es also ein $n_0 \in \mathbb{N}_0$ mit

$$\forall n \geq n_0 \quad \forall p \in \mathbb{N} : \qquad \sup_{x\in G} |(\nabla f_{n+p})(x) - (\nabla f_n)(x)| < \frac{\varepsilon}{4C}\,. \tag{9.9}$$

Auch $(f_n(x_0))_{n=0}^{\infty}$ ist konvergent und somit eine Cauchy–Folge in $\mathbb{R}$. Es gibt also ein $n_1 \in \mathbb{N}_0$ mit

$$\forall n \geq n_1 \quad \forall p \in \mathbb{N} : \qquad |f_{n+p}(x_0) - f_n(x_0)| < \frac{\varepsilon}{2}\,. \tag{9.10}$$

Aus (9.9) und (9.10) erhalten wir unter Verwendung des Mittelwertsatzes 7.31 (mit einem $\theta \in (0,1)$) für alle $n \geq m_0 := \max\{n_0, n_1\}$, alle $p \in \mathbb{N}_0$ und alle $x \in G$:

$$\begin{aligned}
|f_{n+p}(x) - f_n(x)| \leq& |f_{n+p}(x_0) - f_n(x_0)| + |f_{n+p}(x) - f_n(x) - (f_{n+p}(x_0) - f_n(x_0))| \\
<& \frac{\varepsilon}{2} + |\langle x - x_0, (\nabla(f_{n+p} - f_n))(x_0 + \theta(x - x_0))\rangle| \\
\leq& \frac{\varepsilon}{2} + |x - x_0| \cdot |(\nabla(f_{n+p} - f_n))(x_0 + \theta(x - x_0))| \\
\leq& \frac{\varepsilon}{2} + 2C \sup_{x\in G} |(\nabla f_{n+p})(x) - (\nabla f_n)(x)| \\
<& \frac{\varepsilon}{2} + 2C\frac{\varepsilon}{4C} = \varepsilon\,.
\end{aligned}$$

Die Folge $(f_n)_{n=0}^{\infty}$ ist also eine gleichmäßige Cauchy–Folge und daher nach Folgerung 9.7 gleichmäßig auf G gegen eine stetige Funktion $f : G \to \mathbb{R}$ konvergent.

Zu zeigen bleibt noch, daß f auf G stetig partiell differenzierbar ist und daß für alle $x \in G$ und alle $j = 1, \ldots, N$ gilt $(D_j f)(x) = g_j(x)$, wobei g_j die j–te Komponente des Vektorfelds g bezeichnet. Da g nach Satz 9.4 stetig ist, folgt dann auch $f \in C^1(G, \mathbb{R})$.

Seien also $x \in G$ und $j \in \{1, \ldots, N\}$ beliebig und bezeichne e_j den j–Einheitsvektor in $\mathbb{R}^N$. Da G offen ist, gibt es ein $\delta > 0$, so daß $x + te_j \in G$ für alle $t \in I := [-\delta, \delta]$. Wir definieren nun Funktionen $\varphi_n, \varphi, \psi : [-\delta, \delta] \to \mathbb{R}$ durch

$$\varphi_n(t) := f_n(x + te_j)\,, \quad \varphi(t) := f(x + te_j)\,, \quad \psi(t) := g_j(x + te_j) \qquad \text{für alle } t \in [0, 1]\,.$$

Dann folgt $\varphi_n \in C^1([-\delta, \delta], \mathbb{R})$ mit

$$\varphi_n'(t) = \frac{\partial f}{\partial x_j}(x + te_j) \to g_j(x + te_j) = \psi(t)$$

gleichmäßig auf $[-\delta, \delta]$ für $n \to \infty$. Damit sind für φ_n, φ und ψ die Voraussetzungen an f_n, f und g in Satz 9.15 erfüllt und es folgt: φ ist auf $[-\delta, \delta]$ differenzierbar mit $\varphi' \equiv \psi$ auf $[-\delta, \delta]$. Insbesondere folgt

$$\lim_{t\to 0} \frac{f(x + te_j) - f(x)}{t} = \varphi'(0) = \psi(0) = g_j(x)$$

und damit die Behauptung. □

9.17. Definition. Sei Ω eine nicht leere Teilmenge des $\mathbb{R}^N$. Eine Folge $(f_n)_{n=0}^{\infty}$ von Funktionen $f_n : \Omega \to \mathbb{R}^M$ heißt *lokal gleichmäßig konvergent auf* Ω gegen eine Funktion $f : \Omega \to \mathbb{R}^N$, falls es zu jedem $x \in \Omega$ ein $\delta > 0$ gibt, so daß $(f_n)_{n=0}^{\infty}$ auf $U_\delta(x) \cap \Omega$ gleichmäßig gegen f konvergiert.

In den Übungen wird gezeigt:

9.18. Bemerkung. Ist $(f_n)_{n=0}^{\infty}$ eine auf $\emptyset \neq \Omega \subseteq \mathbb{R}^N$ lokal gleichmäßig gegen eine Funktion $f : \Omega \to \mathbb{R}^M$ konvergente Folge von auf Ω stetigen $\mathbb{R}^M$–wertigen Funktionen, so ist auch die Grenzfunktion stetig.

9.19. Folgerung. *Sei $(f_n)_{n=0}^{\infty}$ eine punktweise auf einer offenen Menge $\emptyset \neq \Omega \subseteq \mathbb{R}^N$ gegen eine Funktion $f : \Omega \to \mathbb{R}^M$ konvergente Folge von Funktionen aus $C^1(\Omega, \mathbb{R}^M)$. Für alle $j = 1, \ldots, N$ sei die Folge $\left(\frac{\partial f}{\partial x_j}\right)_{n=0}^{\infty}$ lokal gleichmäßig auf Ω gegen eine Funktion $g_j : \Omega \to \mathbb{R}^M$ konvergent. Dann ist $f \in C^1(\Omega, \mathbb{R}^M)$ und es gilt $\frac{\partial f}{\partial x_j} \equiv g_j$ auf Ω für alle $j = 1, \ldots, N$.*

Beweis. Nach Bemerkung 9.18 sind die Funktionen $g_1, \ldots, g_N : \Omega \to \mathbb{R}^M$ stetig auf Ω. Sei $x \in \Omega$ beliebig. Dann gibt es ein $\delta > 0$, so daß die Folge $\left(\frac{\partial f}{\partial x_j}\right)_{n=0}^{\infty}$ gleichmäßig auf $U_\delta(x) \cap \Omega$ gegen g_j konvergiert für $j = 1, \ldots, N$. Da Ω offen ist, gibt es ein $\eta > 0$ mit $\eta \leq \delta$, so daß $U_\eta(x) \subset \Omega \cap U_\delta(x)$. Wir wenden nun Satz 9.16 an auf die Komponenten von f und die konvexe Menge $G := U_\eta(x)$. Es folgt die stetige partielle Differenzierbarkeit von f auf $U_\eta(x)$. Da f in einer Umgebung eines jeden Punktes aus Ω stetig partiell differenzierbar ist, folgt $f \in C^1(\Omega, \mathbb{R}^N)$. $\square$

3. Potenzreihen

Sei im folgenden $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Ist $(a_n)_{n=0}^{\infty}$ eine Folge aus $\mathbb{K}$ und ist $z_0 \in \mathbb{K}$, so nennen wir die Reihe $\sum_{n=0}^{\infty} f_n$ der Funktionen $f_n : z \mapsto a_n(z - z_0)^n$, $n \in \mathbb{N}_0$, die *Potenzreihe* mit den Koeffizienten a_n, $n \in \mathbb{N}_0$, und dem *Entwicklungspunkt* z_0. Statt $\sum_{n=0}^{\infty} f_n$ schreiben wir auch $\sum_{n=0}^{\infty} a_n(z - z_0)^n$.

9.20. Satz. *Sei $(a_n)_{n=0}^{\infty}$ eine Folge aus $\mathbb{K}$ und sei $z_0 \in \mathbb{K}$. Ist $z_1 \in \mathbb{K} \setminus \{z_0\}$ ein Punkt für den die Reihe $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ konvergiert, so konvergiert die Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ mit dem Entwicklungspunkt z_0 gleichmäßig absolut auf $B_\rho(z_0) = \{z \in \mathbb{K}\,; |z - z_0| \leq \rho\}$ für jedes $\rho \in (0, |z_1 - z_0|)$. Die durch*

$$f(z) := \sum_{n=0}^{\infty} a_n(z - z_0)^n \qquad \text{für alle } z \text{ mit } |z - z_0| < |z_1 - z_0| \tag{9.11}$$

definierte Funktion $f : U_{|z_1 - z_0|}(z_0) \to \mathbb{K}$ ist auf $U_{|z_1 - z_0|}(z_0)$ stetig.

Beweis. Da die Reihe $\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ konvergiert, muß die Folge $(a_n(z_1 - z_0)^n)_{n=0}^{\infty}$ insbesondere beschränkt sein. Es gibt also ein $M > 0$ mit

$$\forall n \in \mathbb{N}_0 : \qquad |a_n(z_1 - z_0)^n| \leq M\,.$$

Für die auf ganz $\mathbb{K}$ definierten stetigen Funktionen $f_n : z \mapsto a_n(z - z_0)^n$ und $0 < \rho < |z_1 - z_0|$ gilt dann:

$$\sup_{z \in U_\rho(z_0)} |f_n(z)| = |a_n|\rho^n = |a_n(z_1 - z_0)^n| \left(\frac{\rho}{|z_1 - z_0|}\right)^n \leq M \left(\frac{\rho}{|z_1 - z_0|}\right)^n.$$

Da die Reihe $\sum_{n=0}^{\infty} M\left(\frac{\rho}{|z_1 - z_0|}\right)^n$ wegen $\frac{\rho}{|z_1 - z_0|} < 1$ konvergent ist, folgt nach dem Majorantenkriterium 9.12 die gleichmäßig absolute Konvergenz der Reihe $\sum_{n=0}^{\infty} f_n$ gegen eine stetige Grenzfunktion auf $B_\rho(z_0)$. Da dies für alle $0 < \rho < |z_1 - z_0|$ gilt, ist die durch (9.11) definierte Funktion f auf $U_{|z_1 - z_0|}(z_0)$ stetig. $\square$

9.21. Definition. Für eine Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ mit Entwicklungspunkt $z_0 \in \mathbb{K}$ und Koeffizienten $a_n \in \mathbb{K}$, $n \in \mathbb{N}_0$, heißt

$$R := \sup \left\{ |z - z_0|\,;\, z \in \mathbb{K}, \sum_{n=0}^{\infty} a_n(z - z_0)^n \text{ ist konvergent} \right\}$$

der *Konvergenzradius* der Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$.

9.22. Folgerung. *Gegeben sei eine Potenzreihe* $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ *mit Entwicklungspunkt* $z_0 \in \mathbb{K}$*, Koeffizienten* $a_n \in \mathbb{K}$*,* $n \in \mathbb{N}_0$*, und Konvergenzradius* R*. Dann gilt:*

(a) *Für alle* $\rho \in [0, R)$ *ist* $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ *gleichmäßig absolut konvergent auf* $B_\rho(z_0)$*. Insbesondere ist die Potenzreihe* $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ *auf* $U_r(z_0)$ *lokal gleichmäßig konvergent.*

(b) *Die durch* $z \mapsto \sum_{n=0}^{\infty} a_n(z-z_0)^n$ *auf* $U_R(z_0)$ *definierte Funktion ist auf* $U_R(z_0)$ *stetig.*

Dies erhält man in nahe liegender Weise aus der Definition des Konvergenzradius und Satz 9.20.

9.23. Satz (von Cauchy und Hadamard[1]). *Für den Konvergenzradius einer Potenzreihe* $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ *mit dem Entwicklungspunkt* $z_0 \in \mathbb{K}$ *und den Koeffizienten* $a_n \in \mathbb{K}$*,* $n \in \mathbb{N}_0$*, gilt:*

$$R = \frac{1}{\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|}} \tag{9.12}$$

mit der Vereinbarung $\frac{1}{0} := \infty$ *und* $\frac{1}{\infty} := 0$.

Beweis. Sei R der Konvergenzradius von $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ und bezeichne $\tilde{R}$ die rechte Seite von (9.12). Wir zeigen zunächst

(a) $\tilde{R} \leq R$. Ist $R = \infty$ oder $\tilde{R} = 0$ so ist dies offensichtlich. Sei nun $0 < \tilde{R} < \infty$ und sei $z \in \mathbb{K}$ beliebig mit $0 < |z-z_0| < \tilde{R}$. Dann gilt:

$$\limsup_{k\to\infty} \sqrt[k]{|a_k(z-z_0)^k|} = |z-z_0| \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \frac{|z-z_0|}{\tilde{R}} < 1\,.$$

Nach dem Wurzelkriterium 3.57 und der anschließenden Bemerkung 3.58 ist also die Reihe $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ absolut konvergent und daher $|z-z_0| \leq R$. Da dies für alle $z \in \mathbb{K}$ mit $0 < |z-z_0| < \tilde{R}$ gilt folgt $\tilde{R} \leq R$.

(b) Zu zeigen ist noch: $R \leq \tilde{R}$. Für $\tilde{R} = \infty$ ist dies offensichtlich. Sei nun $\tilde{R} < \infty$ und $z \in \mathbb{K}$ beliebig mit $|z-z_0| > \tilde{R}$. Dann gilt

$$\limsup_{k\to\infty} \sqrt[k]{|a_k(z-z_0)^k|} = |z-z_0| \limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1\,.$$

Für unendlich viele $n \in \mathbb{N}_0$ gilt also $\sqrt[k]{|a_k(z-z_0)^k|} > 1$ und damit auch $|a_k(z-z_0)^k| > 1$. Die Reihe $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ ist also divergent, da ihre Glieder keine Nullfolge bilden. Da dies für alle $z \in \mathbb{K}$ mit $|z-z_0| > \tilde{R}$ gilt folgt $R \leq \tilde{R}$. $\square$

9.24. Lemma (Cauchy–Produkt von Potenzreihen). *Gegeben seien zwei Potenzreihen* $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ *und* $\sum_{n=0}^{\infty} b_n(z-z_0)^n$ *mit positiven Konvergenzradien* R_a *und* R_b*, mit Koeffizienten in* $\mathbb{K}$ *und mit dem gleichen Entwicklungspunkt* z_0*. Mit*

$$c_n := \sum_{k=0}^{n} a_k b_{n-k} \qquad \text{für alle } n \in \mathbb{N}_0$$

gilt für den Konvergenzradius R *der Potenzreihe* $\sum_{n=0}^{\infty} c_n(z-z_0)^n$*: Es ist*

$$R \geq \min\{R_a, R_b\}$$

[1]Jacques Salomon Hadamard (8.12.1865–17.10.1963).

und für alle $z \in \mathbb{K}$ mit $|z-z_0| < \min\{R_a, R_b\}$ hat man

$$\sum_{n=0}^{\infty} c_n (z-z_0)^n = \Big(\sum_{n=0}^{\infty} a_n (z-z_0)^n\Big)\Big(\sum_{n=0}^{\infty} b_n (z-z_0)^n\Big).$$

Beweis. Sei also $z \in \mathbb{K}$ beliebig mit $|z-z_0| < \min\{R_a, R_b\}$. Nach Folgerung 9.22 sind dann die Reihen $\sum_{n=0}^{\infty} a_n (z-z_0)^n$ und $\sum_{n=0}^{\infty} b_n (z-z_0)^n$ absolutkonvergent. Nach dem Cauchyschen Produktreihensatz 3.67 ist dann auch die Reihe

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k (z-z_0)^k b_{n-k} (z-z_0)^{n-k} = \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)(z-z_0)^n$$

absolutkonvergent gegen $\Big(\sum_{n=0}^{\infty} a_n (z-z_0)^n\Big)\Big(\sum_{n=0}^{\infty} b_n (z-z_0)^n\Big)$. Es ist also $|z-z_0| \leq R$ für alle $z \in \mathbb{K}$ mit $|z-z_0| < \min\{R_a, R_b\}$ und wir erhalten $R \geq \min\{R_a, R_b\}$. □

9.25. Satz. *Sei $\sum_{n=0}^{\infty} a_n (z-z_0)^n$ eine Potenzreihe mit positivem Konvergenzradius R, mit Koeffizienten in $\mathbb{K}$ und mit dem Entwicklungspunkt $z_0 \in \mathbb{K}$. Sei $f : U_R(z_0) \to \mathbb{K}$ die durch*

$$f(z) := \sum_{n=0}^{\infty} a_n (z-z_0)^n \qquad \text{für alle } z \in U_R(z_0)$$

definierte, nach Folgerung 9.22 stetige Funktion. Ist z_0 ein Häufungspunkt der Nullstellenmenge von f, d.h. gibt es eine Folge $(z_n)_{n=0}^{\infty}$ in $U_R(z_0) \setminus \{z_0\}$ mit $f(z_n) = 0$ für alle n und mit $z_n \to z_0$ für $n \to \infty$, so gilt schon $a_n = 0$ für alle $n \in \mathbb{N}_0$ und somit $f \equiv 0$.

Beweis. Sei also $(z_n)_{n=0}^{\infty}$ eine Folge in $U_R(z_0) \setminus \{z_0\}$ mit $f(z_n) = 0$ für alle n und mit $z_n \to z_0$ für $n \to \infty$. Wir verfahren nach der *Methode des kleinsten Verbrechers*:

Annahme: Es gibt ein $n \in \mathbb{N}_0$ mit $a_n \neq 0$. Dann existiert

$$m := \min\{n \in \mathbb{N}_0\,;\, a_n \neq 0\}\,.$$

Es ist also

$$f(z) := \sum_{n=m}^{\infty} a_n (z-z_0)^n \qquad \text{für alle } z \in U_R(z_0)\,.$$

Nach den Grenzwertrechenregeln folgt für $0 < |z-z_0| < R$:

$$\frac{f(z)}{(z-z_0)^m} = \sum_{n=0}^{\infty} a_{m+n} (z-z_0)^n.$$

Insbesondere ist der Konvergenzradius der rechten Seite ≥ 0. Die durch

$$g(z) := \begin{cases} \frac{f(z)}{(z-z_0)^m} & \text{für alle } z \in U_R(z_0) \setminus \{z_0\} \\ a_m & \text{für } z = z_0 \end{cases}$$

definierte Funktion $g : U_R(z_0) \to \mathbb{K}$ ist also nach Folgerung 9.22 stetig auf $U_R(z_0)$. Insbesondere folgt

$$a_m = g(z_0) = \lim_{n\to\infty} \frac{f(z_n)}{(z_n - z_0)^m} = 0$$

im Widerspruch zur Wahl von m. Es gibt also keinen "kleinsten Verbrecher" m, d.h. es ist $a_m = 0$ für alle $m \in \mathbb{N}_0$. □

9.26. Satz. *Sei* $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *eine Potenzreihe mit positivem Konvergenzradius* R*, mit reellen Koeffizienten und mit dem Entwicklungspunkt* $x_0 \in \mathbb{R}$*. Dann ist die durch*

$$f(x) := \sum_{n=0}^{\infty} a_n(x-x_0)^n \qquad \textit{für alle } x \in (x_0-R, x_0+R)$$

definierte Funktion $f : (x_0-R, x_0+R) \to \mathbb{R}$ *auf* (x_0-R, x_0+R) *beliebig oft differenzierbar und es gilt für alle* $k \in \mathbb{N}_0$*:*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n(x-x_0)^{n-k} = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} a_{n+k}(x-x_0)^n.$$

Insbesondere ist also

$$f^{(k)}(x_0) = k!a_k \qquad \textit{für alle } k \in \mathbb{N}_0$$

und daher

$$f(x) := \sum_{n=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^n \qquad \textit{für alle } x \in (x_0-R, x_0+R)\,.$$

Die Funktion f *wird also auf* $U_R(x_0)$ *durch ihre Taylorreihe um* x_0 *dargestellt.*

Beweis. Nach dem Satz 9.23 von Cauchy–Hadamard gilt für die Konvergenzradien R von $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ und R_1 der gliedweise differenzierten Potenzreihe

$$\sum_{n=1}^{\infty} n a_n(x-x_0)^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}(x-x_0)^n,$$

daß $R = 1/\mu$, $R_1 = 1/\mu_1$ mit

$$\begin{aligned} \mu &= \limsup_{k\to\infty} \sqrt[k]{|a_k|}\,, \\ \mu_1 &= \limsup_{k\to\infty} \sqrt[k]{(k+1)|a_{k+1}|}\,. \end{aligned}$$

Nach Definition des Limes superior gibt es eine eine streng monoton wachsende Folge $(k_j)_{j=1}^{\infty}$ mit

$$\mu_1 = \lim_{j\to\infty} \sqrt[k_j]{(k_j+1)|a_{k_j+1}|} = \lim_{j\to\infty} \sqrt[k_j]{|a_{k_j+1}|} \tag{9.13}$$

wegen

$$\lim_{j\to\infty} \sqrt[k_j]{k_j+1} = 1\,.$$

Nach den Grenzwertrechenregeln gilt für $x \neq x_0$: Die Reihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ ist genau dann konvergent,wenn die Reihe $\sum_{n=1}^{\infty} a_n(x-x_0)^n$ konvergiert und dies ist genau dann der Fall, wenn $\sum_{n=1}^{\infty} a_n(x-x_0)^{n-1} = \sum_{n=0}^{\infty} a_{n+1}(x-x_0)^n$ konvergiert. Also gilt auch

$$\mu = \limsup_{k\to\infty} \sqrt[k]{|a_{k+1}|}\,.$$

Wegen (9.13) folgt also $\mu_1 \le \mu$ und somit $R_1 \ge R$. Insbesondere sind die gliedweise differenzierte Reihe und die Ausgangsreihe auf $U_R(x_0)$ lokal gleichmäßig konvergent. Mit Folgerung 9.19 erhalten wir für alle $x \in U_R(x_0)$:

$$f'(x) = \sum_{n=1}^{\infty} n a_n(x-x_0)^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}(x-x_0)^n\,.$$

Denn Rest beweist man durch vollständige Induktion nach der Ableitungsordnung. □

9.27. Folgerung. *Für alle $x \in (-1,1)$ und alle $n \in \mathbb{N}$ gilt:*

$$\frac{1}{(1-x)^{n+1}} = \sum_{j=n}^{\infty} \binom{j}{n} x^{j-n}.$$

Beweis. Für die Funktion $h : (-1,1) \to \mathbb{R}$ mit

$$h(x) := \sum_{j=0}^{\infty} x^j = \frac{1}{1-x} \qquad \text{für alle } x \in (-1,1)$$

zeigt man durch vollständige Induktion:

$$h^{(n)}(x) = \frac{n!}{(1-x)^{n+1}} \qquad \text{für alle } x \in (-1,1), n \in \mathbb{N}_0\,.$$

Mit Satz 9.26 folgt

$$\frac{n!}{(1-x)^{n+1}} = h^{(n)}(x) = \sum_{j=n}^{\infty} \frac{j!}{(j-n)!} x^{j-n}.$$

Division durch $n!$ ergibt die Behauptung. □

9.28. Definition. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und sei $\Omega \subseteq \mathbb{K}$ offen in $\mathbb{K}$. Eine Funktion $f : \Omega \to \mathbb{K}$ heißt auf Ω *lokal in eine Potenzreihe entwickelbar* oder $\mathbb{K}$–*analytisch* , falls es zu jedem $z_0 \in \Omega$ ein $r > 0$ mit $U_r(z_0) \subseteq \Omega$ und eine in $U_r(z_0)$ konvergente Potenzreihe $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ gibt, so daß für alle $z \in U_r(z_0)$ gilt:

$$f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n.$$

9.29. Satz. *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall. Für eine Funktion $f : I \to \mathbb{R}$ sind die beiden folgenden Aussagen äquivalent:*

(a) *f ist auf I reell analytisch.*
(b) *$f \in C^\infty(I,\mathbb{R})$ und für alle $x_0 \in I$ gibt es positive reelle Zahlen δ, M und q mit $U_\delta(x_0) \subseteq I$ und*

$$\forall n \in \mathbb{N}_0 \quad \forall x \in U_\delta(x_0) : \qquad |f^{(n)}(x)| < n! \cdot M \cdot q^n. \tag{9.14}$$

Ist eine dieser beiden Aussagen erfüllt, so gibt es zu jedem $x_0 \in I$ ein $r > 0$ mit $U_r(x_0) \subseteq I$ und:

$$\forall x \in U_r(x_0) : \qquad f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n. \tag{9.15}$$

Zu jedem Punkt aus $x_0 \in I$ gibt es also eine Umgebung $U_r(x_0)$ von x_0, so daß die Funktion f in $U_r(x_0)$ durch ihre Taylorreihe dargestellt wird.

Beweis. “ (b)$\Longrightarrow$(a):” Sei (b) erfüllt und sei $x_0 \in I$ beliebig. Seien δ, M, q gemäß (a) mit (9.14) gewählt. Sei $x \in \mathbb{R}$ mit

$$|x-x_0| < \min\{\delta, 1/q\}\,.$$

Mit dem Satz 6.23 von Taylor folgt für alle $n \in \mathbb{N}$:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x,x_0)$$

mit

$$R_n(x,x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \qquad \text{für ein } \xi \text{ zwischen } x_0 \text{ und } x\,.$$

Unter Verwendung von (9.14) erhalten wir für $n \to \infty$ wegen $|x-x_0|q < 1$:

$$|R_n(x,x_0)| \leq \frac{(n+1)! \cdot M \cdot q^{n+1}}{(n+1)!}|x-x_0|^{n+1} = M(|x-x_0|q)^{n+1} \to 0\,.$$

Damit folgt (b) und die Darstellung (9.15).

„ (a)$\Longrightarrow$(b):" Sei nun f reell analytisch auf I und sei $x_0 \in I$ beliebig. Dann gibt es ein $\delta_0 > 0$ mit $U_{\delta_0}(x_0) \subseteq I$ und eine in $U_{\delta_0}(x_0)$ konvergente Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ mit

$$\forall x \in U_{\delta_0}(x_0): \qquad f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n.$$

Nach dem Satz 9.23 von Cauchy–Hadamard gilt also

$$\delta_0 \leq \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}$$

und damit

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \frac{1}{\delta_0} < \infty\,.$$

Dann gibt es ein $n_0 \in \mathbb{N}_0$, so daß für alle $n \geq n_0$ gilt:

$$\sqrt[n]{|a_n|} < \frac{1}{\delta_0} + \frac{1}{\delta_0} =: q_0\,.$$

Setzen wir also

$$M_0 := 1 + \max\left\{\frac{|a_j|}{q_0^j}\,;\, 0 \leq j \leq n_0\right\},$$

so folgt für alle $n \in \mathbb{N}_0$:

$$|a_n| \leq M_0\Big(\frac{2}{\delta_0}\Big)^n.$$

Für alle $x \in \mathbb{R}$ mit $|x-x_0| < \delta_0/4$ folgt dann mit Satz 9.26 und Folgerung 9.27:

$$\begin{aligned}
|f^{(n)}(x)| =& \Big|\sum_{j=n}^{\infty} \frac{j!}{(j-n)!} a_j(x-x_0)^{j-n}\Big| \leq \sum_{j=n}^{\infty} \frac{j!}{(j-n)!}|a_j| \cdot |x-x_0|^{j-n} \\
\leq& \sum_{j=n}^{\infty} \frac{j!}{(j-n)!} M_0\Big(\frac{2}{\delta_0}\Big)^j \Big(\frac{\delta_0}{4}\Big)^{j-n} = n! M_0 \Big(\frac{2}{\delta_0}\Big)^n \sum_{j=n}^{\infty} \binom{j}{n}\Big(\frac{1}{2}\Big)^{j-n} \\
=& n! M_0 \Big(\frac{2}{\delta_0}\Big)^n \frac{1}{(1-1/2)^{n+1}} = n! 2M_0\Big(\frac{4}{\delta_0}\Big)^n.
\end{aligned}$$

Mit $\delta := \delta_0/4$, $q := 4/\delta_0$ und $M := 2M_0$ ist also (9.14) erfüllt. $\square$

9.30. Satz. *Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit reellen Koeffizienten und positivem Konvergenzradius R. Dann ist die durch*

$$f(x) := \sum_{n=0}^{\infty} a_n(x-x_0)^n \qquad \text{für alle } x \in (x_0-R, x_0+R)$$

definierte Funktion $f : (x_0-R, x_0+R) \to \mathbb{R}$ auf (x_0-R, x_0+R) reell analytisch.

Beweis. Wir zeigen, daß die Bedingung (b) in Satz 9.29 erfüllt ist. Sei also $u_0 \in (x_0-R, x_0+R)$ beliebig vorgegeben und sei $r \in (|u_0-x_0|, R)$ fest gewählt. Mit den Bezeichnungen

$$\delta := \frac{1}{2}(r - |x_0-u_0|) \quad \text{und} \quad \tilde{r} := \frac{1}{2}(r + |x_0-u_0|) = |x_0-u_0| + \delta$$

gilt $\delta > 0$ und $\tilde{r} < r < R$. Nach dem Satz 9.23 von Cauchy–Hadamard gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{1}{R} < \frac{1}{r} < \infty\,.$$

Es gibt also ein $n_0 \in \mathbb{N}_0$, so daß

$$\forall n \geq n_0 : \qquad \sqrt[n]{|a_n|} < \frac{1}{r}$$

also auch

$$\forall n \geq n_0 : \qquad |a_n| < \frac{1}{r^n}\,.$$

Für alle $x \in U_\delta(u_0)$ ist $|x - x_0| \leq |x - u_0| + |u_0 - x_0| < \tilde{r} < r$ und es folgt mit Satz 9.26 und Folgerung 9.27 für alle $n \geq n_0$:

$$\begin{aligned}
|f^{(n)}(x)| =& \Big| \sum_{j=n}^{\infty} \frac{j!}{(j-n)!} a_j (x - x_0)^{j-n} \Big| \leq \sum_{j=n}^{\infty} \frac{j!}{(j-n)!} |a_j| \cdot |x - x_0|^{j-n} \\
\leq& n! \sum_{j=n}^{\infty} \binom{j}{n} \frac{\tilde{r}^{j-n}}{r^j} = \frac{n!}{r^n} \sum_{j=n}^{\infty} \binom{j}{n} \Big(\frac{\tilde{r}}{r}\Big)^{j-n} \\
=& \frac{n!}{r^n} \cdot \frac{1}{\big(1 - \frac{\tilde{r}}{r}\big)^{n+1}} = n! \tilde{M} \cdot \Big(\frac{1}{r - \tilde{r}}\Big)^n
\end{aligned}$$

mit $\tilde{M} := \dfrac{r}{r - \tilde{r}}$.

Für $0 \leq n \leq n_0$ ist $f^{(n)}$ auf $[u_0 - \delta, u_0 + \delta]$ stetig und damit auch beschränkt. Es gibt daher ein $M_n > 0$ mit

$$\forall x \in [u_0 - \delta, u_0 + \delta] : \qquad |f^{(n)}(x)| \leq n! M_n \cdot \Big(\frac{1}{r - \tilde{r}}\Big)^n.$$

Mit $M := \max\{M_0, \ldots, M_{n_0}, \tilde{M}\}$ folgt

$$\forall n \in \mathbb{N}_0 \quad \forall x \in U_\delta(u_0) : \qquad |f^{(n)}(x)| \leq n! M \cdot \Big(\frac{1}{r - \tilde{r}}\Big)^n.$$

Also ist (9.14) erfüllt mit $q := \dfrac{1}{r - \tilde{r}}$. Nach Satz 9.29 ist f daher reell analytisch auf $(x_0 - R, x_0 + R)$. $\square$

9.31. Folgerung. *Die Funktionen* exp, sin, cos, sinh, cosh *sind auf* $\mathbb{R}$ *reell analytisch.*

9.32. Satz. *Sei $I = (a, b)$ mit $-\infty \leq a < b \leq \infty$ ein offenes Intervall und sei $f : I \to \mathbb{R}$ eine nicht identisch verschwindende reell analytische Funktion auf I. Dann hat die Menge*

$$N_f := \{x \in I\,;\, f(x) = 0\}$$

der Nullstellen von I keinen Häufungspunkt in I.

Beweis. (a) Wir zeigen zunächst, daß die Menge H_f der Häufungspunkte von N_f offen ist.

Sei $x_0 \in H_f$ beliebig. Da f auf I reell analytisch ist, gibt es ein $\varepsilon > 0$, so daß $U_\varepsilon(x_0) \subseteq I$ und so daß f auf $U_\varepsilon(x_0)$ eine Darstellung durch eine in $U_\varepsilon(x_0)$ konvergente Potenzreihe mit Entwicklungspunkt x_0 hat. Nun ist x_0 Häufungspunkt von N_f, also auch Häufungspunkt von $N_{f|_{U_\varepsilon(x_0)}}$. Nach Satz 9.25 folgt somit $f \equiv 0$ auf $U_\varepsilon(x_0)$. Insbesondere ist $U_\varepsilon(x_0) \subseteq H_f \subseteq N_f$.

(b) *Annahme*: Es gibt einen Häufungspunkt t_0 von N_f und $H_f \subseteq N_f$.

Da f nicht identisch verschwindet, gibt es ein $u \in I$ mit $f(u) \neq 0$ Wegen der Stetigkeit von f gibt es dann noch ein $\delta > 0$ mit $U_\delta(u) \subseteq I$ und $U_\delta(u) \cap N_f = \emptyset$. Ohne Beschränkung der Allgemeinheit können wir $u < t_0$ annehmen (den Fall $u > t_0$ behandelt man analog). Für

$$v_0 := \inf((u, b) \cap H_f) \tag{9.16}$$

gilt dann: $v_0 > u$ und $v_0 \notin H_f$, da sonst nach (a) ein $\varepsilon > 0$ existiert mit $U_\varepsilon(v_0) \subseteq H_f$ im Widerspruch zur Definition von v_0. Wegen $v_0 \notin H_f$ und (9.16) gibt eine Folge $(s_n)_{n=1}^\infty$ aus $(v_0, b) \cap H_f \subseteq N_f$ mit $s_n \to v_0$ für $n \to \infty$. Dies steht im Widerspruch dazu, daß v_0 kein Häufungspunkt von N_f ist.

N_f kann also keine Häufungspunkte haben. □

KAPITEL 10

Nichtlineare Gleichungen

1. Der Fixpunktsatz von Banach

Ist $T : X \to X$ eine Abbildung von einer nicht leeren Menge X in sich, so nennen wir $x_0 \in X$ einen *Fixpunkt von* T, falls

$$T(x_0) = x_0$$

gilt.

10.1. Beispiel. In den Übungen zur Analysis 1 zeigt man mit Hilfe des Zwischenwertsatzes, daß jede stetige Abbildung $f : [0,1] \to [0,1]$ wenigstens einen Fixpunkt besitzt.

10.2. Definition. Sei (X,d) ein metrischer Raum. Eine Abbildung $T : X \to X$ heißt *stark kontrahierend*, falls es eine Konstante $K \in [0,1)$ gibt mit

$$\forall x, y \in X : \qquad d(T(x), T(y)) \leq K d(x,y)\,.$$

Wir nennen K dann auch eine *Kontraktionskonstante* der starken Kontraktion T.

Eine Abbildung $T : X \to X$ ist also genau dann stark kontrahierend, wenn sie Lipschitz–stetig mit einer Lipschitz–Konstanten $K \in [0,1)$ ist.

10.3. Beispiel. Die Abbildung

$$f : (0,1] \to (0,1]\,, \qquad x \mapsto \frac{x}{2}\,,$$

ist stark kontrahierend mit Kontraktionskonstante $1/2$ und besitzt keinen Fixpunkt in $(0,1]$.

Ein wichtiges konstruktives Kriterium für die Existenz eines Fixpunktes ist der folgende Fixpunktsatz von Banach.

10.4. Banachscher Fixpunktsatz. *Sei (X,d) ein vollständiger metrischer Raum und sei $T : X \to X$ eine stark kontrahierende Abbildung mit Kontraktionskonstante $K \in [0,1)$. Dann besitzt T genau einen Fixpunkt $a \in X$. Für diesen gilt: Ist $x_0 \in X$ beliebig, so ist*

$$a = \lim_{n\to\infty} x_n \quad \textit{mit } x_{n+1} := T(x_n) \textit{ für alle } n \in \mathbb{N}_0\,. \tag{10.1}$$

Man hat die a priori Fehlerabschätzung

$$\forall n \in \mathbb{N} : \qquad d(x_n, a) \leq \frac{K^n}{1-K} d(x_1, x_0) \tag{10.2}$$

und die a posteriori Fehlerabschätzungen

$$\forall n \in \mathbb{N} : \qquad d(x_{n+1}, a) \leq \frac{K}{1-K} d(x_{n+1}, x_n) \tag{10.3}$$

und

$$\forall x \in X : \qquad d(T(x), a) \leq \frac{K}{1-K} d(T(x), x)\,. \tag{10.4}$$

Beweis. (a) *Zur Eindeutigkeit*: Sind $a, b \in X$ mit $T(a) = a$ und $T(b) = b$, so folgt nach Voraussetzung

$$d(a,b) = d(T(a), T(b)) \le K d(a,b)\,.$$

Wegen $0 \le K < 1$ ist dies nur möglich, wenn $d(a,b) = 0$ und damit $a = b$ gilt. Zu zeigen bleibt also noch die Existenz eines Fixpunktes.

(b) Sei $x_0 \in X$ beliebig und im folgenden fest und sei $(x_n)_{n=0}^{\infty}$ die gemäß (10.1) induktiv konstruierte Folge aus X. Durch vollständige Induktion zeigen wir:

$$\forall n \in \mathbb{N}_0 : \qquad d(x_{n+1}, x_n) \le K^n d(x_1, x_0)\,. \tag{10.5}$$

Für $n = 0$ ist dies trivial. Ist $n \in \mathbb{N}_0$ mit $d(x_{n+1}, x_n) \le K^n d(x_1, x_0)$, so folgt nach Voraussetzung

$$d(x_{n+2}, x_{n+1}) = d(T(x_{n+1}), T(x_n)) \le K d(x_{n+1}, x_n) \le K^{n+1} d(x_1, x_0)\,.$$

(c) Unter Verwendung von (10.5) und der Dreiecksungleichung erhalten wir für alle $n, p \in \mathbb{N}_0$:

$$d(x_{n+p}, x_n) \le \sum_{k=0}^{p-1} d(x_{n+k+1}, x_{n+k}) \le \sum_{k=0}^{p-1} K^{n+k} d(x_1, x_0) \le \frac{K^n}{1-K} d(x_1, x_0)\,. \tag{10.6}$$

Die Folge $(x_n)_{n=0}^{\infty}$ ist also eine Cauchy–Folge in (X, d). Da (X, d) ein vollständiger metrischer Raum ist, existiert der Grenzwert

$$a := \lim_{n\to\infty} x_n \qquad \text{in } (X, d).$$

Wegen der Stetigkeit der Abbildung $d(\cdot, x_n) : X \to \mathbb{R}$ folgt aus (10.6) für $p \to \infty$ die a priori Abschätzung (10.2).

(d) Wegen $T(x_n) = x_{n+1}$ gilt nach Voraussetzung

$$0 \le d(T(a), a) \le d(T(a), x_{n+1}) + d(x_{n+1}, a) \le d(a, x_n) + d(x_{n+1}, a) \to 0 \quad \text{für } n \to \infty$$

und damit $d(T(a), a) = 0$, d.h. $T(a) = a$. Der Punkt a ist also der nach (a) einzige Fixpunkt von T.

(e) Nimmt man x_n als neuen Startwert für die Iteration, so erhält man aus der a priori Abschätzung (10.2)

$$\forall p \in \mathbb{N}_0 : \qquad d(x_{n+p}, a) \le \frac{K^p}{1-K} d(x_{n+1}, x_n)\,.$$

Der Spezialfall $p = 1$ liefert die a posteriori Abschätzung (10.3).

Sei nun $x \in X$ beliebig. Nimmt man als neuen Startwert der Iteration $x_0 := x$, so folgt die Abschätzung (10.4) aus (10.2) mit speziell $n = 1$. □

10.5. Beispiele. (a) Die Tangensfunktion besitzt in dem Intervall $(\pi, \frac{3}{2}\pi)$ offensichtlich einen Fixpunkt x_0. Für alle $x, y \in \mathbb{R}$ gilt jedoch nach dem Mittelwertsatz 5.20 der Differentialrechnung mit einem ξ zwischen x und y:

$$|\tan(x) - \tan(y)| = |\tan'(\xi)| \cdot |x - y| = \frac{|x-y|}{\cos(\xi)^2} \ge |x - y|\,.$$

Die Tangensfunktion ist daher auf keiner Teilmenge von $\mathbb{R}$ mit mehr als einem Punkt stark kontrahierend. Es gilt jedoch für $x \in [\pi, \frac{3}{2}\pi]$:

$$\tan(x) = x \iff x = \tan(x - \pi) \iff \arctan(x) = x - \pi \iff g(x) = x$$

mit $g(t) := \arctan(t) + \pi$. Wegen $\arctan([0, \infty)) \subset [0, \frac{1}{2}\pi]$ gilt

$$g([\pi, \frac{3}{2}\pi]) \subseteq [\pi, \frac{3}{2}\pi]\,.$$

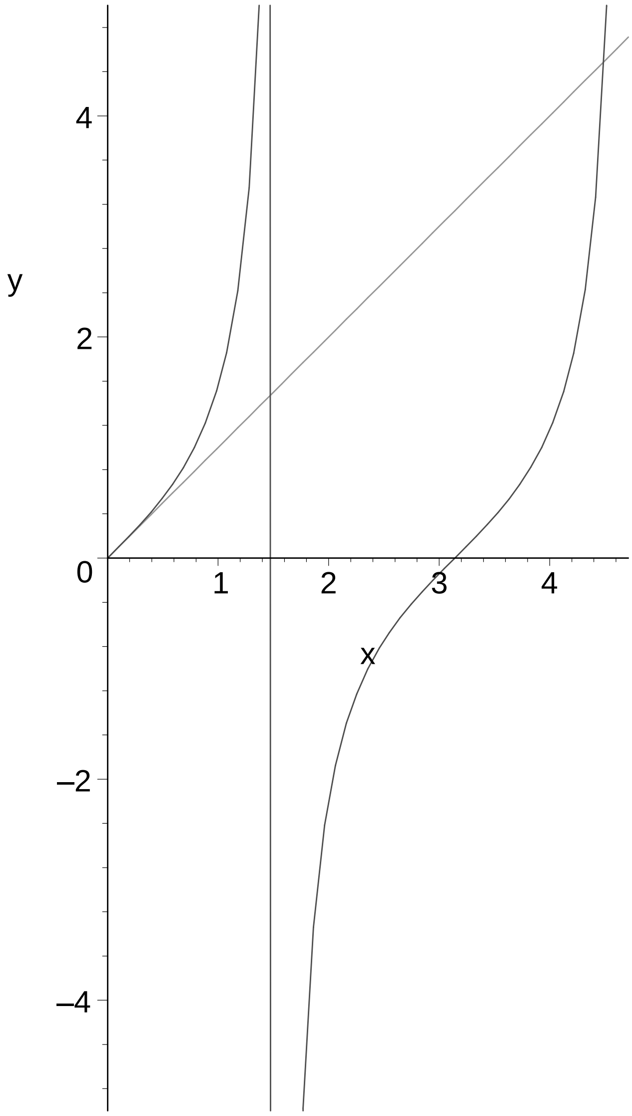


Abbildung 1. Die Tangensfunktion im Bereich $[0, 3\pi/2]$.

Da $[\pi, \frac{3}{2}\pi]$ eine abgeschlossene Teilmenge des vollständigen metrischen Raums $(\mathbb{R}, d_{|\cdot|})$ ist, ist $([\pi, \frac{3}{2}\pi], d)$ ein vollständiger metrischer Raum, wenn wir mit d die von der euklidischen Metrik $d_{|\cdot|}$ auf $[\pi, \frac{3}{2}\pi] \times [\pi, \frac{3}{2}\pi]$ eingeschränkte Metrik bezeichnen. Für alle $x, y \in [\pi, \frac{3}{2}\pi]$ gibt es nach dem Mittelwertsatz ein ξ zwischen x und y mit

$$|g(x) - g(y)| = |g'(\xi)| \cdot |x - y| = \frac{1}{1 + \xi^2}|x - y| \leq \frac{1}{1 + \pi^2}|x - y| \, .$$

g ist also eine stark kontrahierende Abbildung von $[\pi, \frac{3}{2}\pi]$ in sich mit einer Kontraktionskonstante $K_0 \leq \frac{1}{1+\pi^2} < K := 0.092 < 1$ und hat nach dem Banachschen Fixpunktsatz genau einen Fixpunkt a in $[\pi, \frac{3}{2}\pi]$, der dann auch der einzige Fixpunkt der Tangensfunktion in $[\pi, \frac{3}{2}\pi]$ sein muß. Mit dem im Banachschen Fixpunktsatz angegebenen iterativen Verfahren können wir a näherungsweise berechnen, wobei wir für die Näherungen Fehlerschranken angeben können. Wir verwenden als Startwert den Wert $x_0 := 4$ und erhalten in den ersten Iterationsschritten folgende Näherungen und Fehlerabschätzungen:

Fixpunktiteration zur Berechnung des Fixpunktes a von g			
n	x_n	$\frac{K^n}{1-K}\lvert x_1 - x_0\rvert$	$\frac{K}{1-K}\lvert x_n - x_{n-1}\rvert$
1	4.467410318	0.04735875469	0.04735875469
2	4.492175746	0.004357005431	0.002509272440
3	4.493351224	0.0004008444997	0.0001191012952
4	4.493406710	0.00003687769397	0.000005621929515

(b) Sei $I := [0,1]$. Dann ist $(C(I,\mathbb{R}), d_I)$ mit

$$d_I(f,g) := \|f-g\|_I = \sup_{t\in I}|f(t)-g(t)| \qquad \text{für alle } f,g \in C(I,\mathbb{R})$$

nach Satz 9.8 ein vollständiger metrischer Raum. Für alle $f \in C(I,\mathbb{R})$ definieren wir

$$(T(f))(t) := \int_0^1 \cos(f(s)st)ds \qquad \text{für } 0 \le t \le 1\,.$$

Nach Satz 7.9 ist $T(f)$ wieder eine stetige Funktion auf I. Seien nun $f,g \in C(I,\mathbb{R})$ beliebig. Für alle $s,t \in I$ gibt es nach dem Mittelwertsatz der Differentialrechnung ein ξ zwischen $f(s)$ und $g(s)$ mit

$$|\cos(f(s)st) - \cos(g(s)st)| = |st\sin(\xi st)| \cdot |f(s)-g(s)| \le s d_I(f,g)\,.$$

Es folgt also

$$d_I(T(f),T(g)) = \sup_{t\in I}|(T(f))(t) - (T(g))(t)| \le \int_0^1 s d_I(f,g)ds = \frac{1}{2}d_I(f,g)\,.$$

T ist also eine stark kontrahierende Abbildung von $(C(I,\mathbb{R}), d_I)$ in sich und besitzt daher nach dem Banachschen Fixpunktsatz genau einen Fixpunkt, den man näherungsweise mit Hilfe des im Banachschen Fixpunktsatz angegebenen Iterationsverfahrens berechnen kann.

Wir geben noch einige Varianten des Banachschen Fixpunktsatzes an:

10.6. Folgerung. *Sei (X,d) ein vollständiger metrischer Raum und sei $T : X \to X$ eine Abbildung, so daß für ein $m \in \mathbb{N}$ die m–malige Hintereinanderausführung*

$$T^m := \underbrace{T \circ T \circ \cdots \circ T}_{m\text{–mal}}$$

eine stark kontrahierende Abbildung von X in sich ist. Dann besitzt T genau einen Fixpunkt.

Beweis. Nach dem Banachschen Fixpunktsatz hat T^m genau einen Fixpunkt a in X. Da jeder Fixpunkt von T auch einer von T^m ist, kann T höchstens einen Fixpunkt besitzen. Wegen $T^m(T(a)) = T^{m+1}(a) = T(T^m(a)) = T(a)$ ist auch $T(a)$ ein Fixpunkt von T^m. Wegen der Eindeutigkeitsaussage im Banachschen Fixpunktsatz muß $a = T(a)$ gelten. $\square$

10.7. Satz. *Sei (X,d) ein vollständiger metrischer Raum, sei $\emptyset \neq A \subseteq X$ und sei $T : A \to X$ eine Abbildung, für die eine Konstante $K \in [0,1)$ existiert mit*

$$\forall x, y \in A : \qquad d(T(x), T(y)) \leq K d(x,y).$$

Gibt es ein $x_0 \in A$ und ein $r > 0$ mit $B_r(x_0) = \{x \in X\,;\, d(x, x_0) \leq r\} \subseteq A$ und

$$d(T(x_0), x_0) \leq (1-K)r\,, \tag{10.7}$$

so gilt: Die durch $x_{n+1} := T(x_n)$ für alle $n \in \mathbb{N}_0$ induktiv definierte Folge konvergiert gegen einen Fixpunkt $a = T(a) \in B_r(x_0)$. a ist der einzige Fixpunkt von T in $B_r(x_0)$ und es gelten die Fehlerabschätzungen (10.2) und (10.3) aus dem Banachschen Fixpunktsatz.

Beweis. (a) Die Eindeutigkeitsaussage folgt wie im Beweis zum Banachschen Fixpunktsatz.

(b) Durch vollständige Induktion nach n zeigen wir zunächst, daß für alle $n \in \mathbb{N}_0$ gilt:

$$x_n \in B_r(x_0) \quad \text{und} \quad d(x_{n+1}, x_n) \leq K^n d(x_1, x_0)\,. \tag{10.8}$$

Im Fall $n = 0$ ist dies offensichtlich erfüllt. Sei nun $k \in \mathbb{N}_0$, so daß (10.8) für alle $k = 0, \ldots, n$ erfüllt ist. Mit dieser Induktionsvoraussetzung und der Dreiecksungleichung erhalten wir unter Verwendung von (10.7):

$$d(x_{n+1}, x_0) \leq \sum_{j=0}^{n} d(x_{j+1}, x_j) \leq \sum_{j=0}^{n} K^j d(x_1, x_0) \leq \frac{d(x_1, x_0)}{1-K} = \frac{d(T(x_0), x_0)}{1-K} \leq r$$

Also ist $x_{n+1} \in B_r(x_0) \subseteq A$ und es folgt

$$d(x_{n+2}, x_{n+1}) = d(T(x_{n+1}), T(x_n)) \leq K d(x_{n+1}, x_n) \leq K^{n+1} d(x_1, x_0)\,.$$

(c) Wie in Teil (c) des Beweises zum Banachschen Fixpunktsatz erhält man die Konvergenz der Folge gegen einen Grenzwert $a \in B_r(x_0)$ (man beachte, daß $B_r(x_0)$ eine vollständige Teilmenge von (X,d) ist und in A liegt) und die a priori Abschätzung (10.2). Auch die Tatsache, daß a ein Fixpunkt von T ist, zeigt man analog wie im Beweis zum Banachschen Fixpunktsatz.

(d) Durch vollständige Induktion nach j zeigt man

$$\forall n, j \in \mathbb{N}_0 : \qquad d(x_{n+2+j}, x_{n+1+j}) \leq K^{j+1} d(x_{n+1}, x_n)\,.$$

Für alle $n, p \in \mathbb{N}_0$ erhalten wir also nach (b) und der Dreiecksungleichung

$$d(x_{n+2+p}, x_{n+1}) \leq \sum_{j=0}^{p} d(x_{n+2+j}, x_{n+1+j}) \leq \sum_{j=0}^{p} K^{j+1} d(x_{n+1}, x_n) \leq \frac{K}{1-K} d(x_{n+1}, x_n)\,.$$

Für $k \to \infty$ folgt hieraus die a posteriori Abschätzung (10.3). □

10.8. Satz (Lokaler Konvergenzsatz). *Sei (X,d) ein vollständiger metrischer Raum und sei $T : X \to X$ eine Abbildung, die einen Fixpunkt $a \in X$ besitze. Es gebe ein $\varepsilon > 0$ und eine Konstante $K \in [0,1)$ mit*

$$\forall x, y \in U_\varepsilon(a) : \qquad d(T(x), T(y)) \leq K d(x,y)\,.$$

Dann gilt für $0 < r < \varepsilon$: Es ist $T(B_r(a)) \subseteq B_r(a)$ und für jedes $x_0 \in B_r(a)$ konvergiert die durch $x_{n+1} := T(x_n)$ für alle $n \in \mathbb{N}_0$ induktiv definierte Folge gegen a und es gelten die Fehlerabschätzungen (10.2) und (10.3) aus dem Banachschen Fixpunktsatz.

Beweis. Für $0 < r < \varepsilon$ gilt $B_r(a) \subset U_\varepsilon(a)$ und für alle $x \in B_r(a)$ ist auch $T(x) \in B_r(a)$ wegen

$$d(T(x), a) = d(T(x), T(a)) \leq Kd(x, a) \leq Kr < r\,.$$

Als abgeschlossene Teilmenge des vollständigen metrischen Raums (X, d) ist $B_r(a)$ vollständig. Der Banachsche Fixpunktsatz ist also auf $T|_{B_r(a)} : B_r(a) \to B_r(a)$ und den vollständigen metrischen Raum $(B_r(a), d|_{B_r(a)\times B_r(a)})$ anwendbar und es folgen die Behauptungen □

10.9. Definition. Sei $(x_n)_{n=0}^\infty$ eine gegen einen Punkt $a \in X$ konvergente Folge in einem metrischen Raum (X, d) und sei $p \geq 1$. Die Folge $(x_n)_{n=0}^\infty$ heißt *konvergent von der Ordnung* p, falls eine Konstante $C > 0$ existiert mit

$$\forall n \in \mathbb{N}_0 : \qquad d(x_{n+1}, a) \leq Cd(x_n, a)^p.$$

Im Fall $p = 1$ fordern wir zusätzlich $C < 1$ und sprechen von *linearer Konvergenz.*

Die Konvergenz der Iteration des Banachschen Fixpunktsatzes ist also zumindest linear.

2. Das Newton–Verfahren

Wir betrachten die folgende *Problemstellung*: Gegeben sei eine auf einer offenen Menge $\emptyset \neq \Omega \subseteq \mathbb{R}^N$ definierte Abbildung $f : \Omega \to \Omega$. Es gebe einen Punkt $a \in \Omega$ mit $f(a) = 0$. Wir suchen nach einem geeigneten Verfahren zur näherungsweisen Berechnung von a unter geeigneten, noch zu präzisierenden Voraussetzungen an f.

Zur Motivation des Ansatzes im folgenden Satz betrachten wir zunächst den Fall $N = 1$ und $\Omega = I$, $I \neq \emptyset$ ein offenes Intervall in $\mathbb{R}$. Ist f auf I stetig differenzierbar, so wird sich für x_0 in der "Nähe" von a die Tangente t an den Graphen von f im Punkt $(x_0, f(x_0))$ nicht "wesentlich" von f unterscheiden. Die Nullstelle x_1 von t sollte also eine verbesserte Näherung für a sein. Die Tangente

$$t : \mathbb{R} \to \mathbb{R}\,, \qquad x \mapsto t(x) = f(x_0) + f'(x_0)(x - x_0)$$

schneidet die x–Achse natürlich nur, wenn $f'(x_0) \neq 0$ ist. Dies trifft wegen der Stetigkeit von f' in einer Umgebung von a zu, falls $f'(a) \neq 0$ erfüllt ist. Man erhält für x_1:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}\,.$$

Man fährt auf diese Weise fort, definiert also induktiv:

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{für alle } n \in \mathbb{N}_0\,.$$

a ist Fixpunkt von g mit

$$g(x) := x - \frac{f(x)}{f'(x)} \qquad \text{für alle } x \in I \text{ mit } f'(x) \neq 0.$$

Unter der zusätzlichen Voraussetzung, daß f' Lipschitz–stetig ist, kann man mit Hilfe des lokalen Konvergenzsatzes 10.8 lineare Konvergenz dieses Verfahrens gegen a zeigen. Setzen wir $f \in C^2(I, \mathbb{R})$ voraus, so können wir sogar quadratische Konvergenz (also Konvergenz der Ordnung $p = 2$) beweisen.

10.10. Satz (Newton[1]–Verfahren). *Sei* $\Omega \subseteq \mathbb{R}^N$ *offen und* $f \in C^2(\Omega, \mathbb{R}^N)$. *Es gebe ein* $a \in \Omega$ *mit* $f(a) = 0$. *Ferner sei die Jacobi–Matrix* $J_f(a)$ *von* f *in* a *invertierbar. Dann gibt es Konstanten* $C > 0$ *und* $r \in (0, 1/C)$, *so daß für alle* $x_0 \in \Omega$ *mit* $|x_0 - a| \leq r$ *gilt: Die durch*

$$x_{n+1} := x_n - J_f(x_n)^{-1} f(x_n) \tag{10.9}$$

induktiv definierte Folge $(x_n)_{n=0}^{\infty}$ *konvergiert gegen* a *und erfüllt*

$$\forall n \in \mathbb{N}_0 : \qquad |x_{n+1} - a| \leq C|x_n - a|^2.$$

Beweis. Wie in den Übungen gezeigt wurde ist die Menge G der invertierbaren Matrizen aus $M_{N\times N}(\mathbb{R})$ offen in $M_{N\times N}(\mathbb{R})$ und die Matrixinversion ist als Abbildung von G in sich stetig. Es gibt daher Konstanten $\rho, m_1, m_2 > 0$ mit

$$B_\rho(a) = \{x \in \mathbb{R}^N \,;\, |x - a| \leq \rho\} \subset \Omega \quad \text{und} \quad J_f(B_\rho(a)) \subset G$$

sowie

$$\sup_{x \in B_\rho(a)} \|J_f(x)\| \leq m_1 \quad \text{und} \quad \sup_{x \in B_\rho(a)} \|J_f(x)^{-1}\| \leq m_2 \,.$$

Man beachte hierzu, daß $B_\rho(a)$ eine kompakte Teilmenge des $\mathbb{R}^N$ ist und stetige Funktionen auf kompakten Teilmengen des $\mathbb{R}^N$ beschränkt sind. Da f zweimal stetig partiell differenzierbar ist, ist $J_f \in C^1(\Omega, M_{N\times N}(\mathbb{R}))$ und daher nach dem Mittelwertsatz der Differentialrechnung Lipschitz–stetig auf $B_\rho(a)$. Es gibt also ein $L > 0$ mit

$$\forall x, y \in B_\rho(x) : \quad \|J_f(x) - J_f(y)\| \leq \|J_f(x) - J_f(y)\|_2 \leq L|x - y| \,.$$

Nach dem Satz von Taylor 7.38 gilt für alle $x \in B_\rho(x_0)$:

$$f(x) = f(a) + J_f(a)(x - a) + R_1(x, a) = J_f(a)(x - a) + R_1(x, a)$$

mit

$$\begin{aligned} |R_1(x,a)| &= \Big| 2 \sum_{|\alpha|=2} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 (1-s)(D^\alpha f)(a + s(x-a)) ds \Big| \\ &\leq 2 \sum_{|\alpha|=2} \frac{|x-a|^2}{\alpha!} \int_0^1 (1-s) \big|(D^\alpha f)(a + s(x-a))\big| ds \\ &\leq \max_{|\alpha|=2} \max_{u \in B_\rho(a)} \big|(D^\alpha f)(u)\big| |x-a|^2 \sum_{|\alpha|=2} \frac{2!}{\alpha!} \int_0^1 (1-s) ds \\ &\leq C|x-a|^2, \end{aligned} \tag{10.10}$$

wobei

$$C := \frac{N^2}{2} \max_{|\alpha|=2} \max_{u \in B_\rho(a)} \big|(D^\alpha f)(u)\big| < \infty \,,$$

da die zweiten partiellen Ableitungen von f stetig und somit auf $B_\rho(a)$ beschränkt sind. Offensichtlich ist a Fixpunkt der in einer Umgebung U von a definierten Abbildung

$$g : U \to \mathbb{R}^N \,, \qquad x \mapsto g(x) := x - J_f(x)^{-1} f(x) \,.$$

Wir setzen $K := m_2 C + m_1 m_2^2 L$ und fixieren $r < \min\{\rho, 1/K\}$. Für alle $x \in B_r(a) \subset B_\rho(a)$ können wir nun abschätzen:

$$\begin{aligned} |g(x) - a| &= |g(x) - g(a)| = |x - J_f(x)^{-1} f(x) - a + J_f(a)^{-1} f(a)| \\ &\leq \underbrace{|J_f(a)^{-1}(J_f(a)(x-a) + f(a) - f(x))|}_{=:A(x)} + \underbrace{|(J_f(a)^{-1} - J_f(x)^{-1}) f(x)|}_{=:B(x)} \end{aligned}$$

[1] Sir Isaac Newton (4.1.1643–31.3.1727).

Für den ersten Summanden $A(x)$ erhalten wir nach (10.10):

$$A(x) \leq \|J_f(a)^{-1}\| \cdot |J_f(a)(x-a) + f(a) - f(x)| \leq m_2 C|x-a|^2$$

und für den zweiten Summanden folgt mit Hilfe des Mittelwertsatzes 7.30 der Differentialrechnung in mehreren Veränderlichen wegen $f(a) = 0$ und der Lipschitz–Stetigkeit von $\xi \mapsto J_f(\xi)$ auf $B_\rho(a)$:

$$\begin{aligned} B(x) \leq& \|J_f(a)^{-1} - J_f(x)^{-1}\| \cdot |f(x)| = \\ =& \|J_f(x)^{-1} J_f(x) J_f(a)^{-1} - J_f(x)^{-1} J_f(a) J_f(a)^{-1}\| \cdot |f(x) - f(a)| \leq \\ \leq& \|J_f(x)^{-1}\| \cdot \|J_f(x) - J_f(a)\| \cdot \|J_f(a)^{-1}\| \cdot |x-a| \sup_{0 \leq t \leq 1} \|J_f(a + t(x-a))\| \leq \\ \leq& m_2^2 L m_1 |x-a|^2 . \end{aligned}$$

Also gilt:

$$|g(x) - a| \leq K|x-a|^2 \leq Kr|x-a| \leq qr$$

mit $q := rK < 1$ Insbesondere ist $g(x) \in B_r(a)$ für alle $x \in B_r(a)$. Durch vollständige Induktion sieht man nun: Die durch (10.9) induktiv definierte Folge nimmt für $x_0 \in B_r(a)$ ihre Werte in $B_r(a)$ an und erfüllt

$$|x_{n+1} - a| = |g(x_n) - a| \leq K|x_n - a|^2 \leq q|x_n - a| . \tag{10.11}$$

Durch vollständige Induktion folgt

$$|x_n - a| \leq q^n |x_0 - a| \to 0 \qquad \text{für } n \to \infty .$$

Wegen (10.11) ist die Konvergenz quadratisch. □

10.11. Bemerkung. Damit man nicht in jedem Schritt eine aufwendige Matrixinversion durchführen muß, geht man in der Praxis wie folgt vor: Multiplizieren wir (10.9) von links mit $J_f(x)$, so erhalten wir

$$J_f(x_n) x_{n+1} = J_f(x_n) x_n - f(x_n)$$

also

$$J_f(x_n) z_n = -f(x_n) \tag{10.12}$$

mit $z_n = x_{n+1} - x_n$. Man berechnet nun z_n als Lösung des linearen Gleichungssystems (10.12) und erhält x_{n+1} als $x_{n+1} = z_n + x_n$.

3. Implizite Funktionen und Umkehrfunktionen

10.12. Beispiele. (a) Lösen wir die Gleichung

$$2x^2 + 3y = 0 \tag{10.13}$$

nach y auf, so erhalten wir $y = -\frac{2}{3}x^2$ für alle $x \in \mathbb{R}$. Man sagt, die hierdurch gegebene Funktion

$$g : \mathbb{R} \to \mathbb{R} , \qquad x \mapsto -\frac{2}{3}x^2$$

ist implizit durch (10.13) definiert. Wollen wir hingegen (10.13) nach x auflösen, so erhalten wir nur dann reelle Lösungen, wenn $y \leq 0$ ist. In diesem Fall erhalten wir zwei Lösungen

$$x_1(y) = \sqrt{-\frac{3}{2}y} \quad \text{und} \quad x_2(y) = -\sqrt{-\frac{3}{2}y} .$$

(b) Die Gleichung $x^2 + y^2 + 1 = 0$ hat für keine Werte von x eine reelle Lösung $y(x)$ und für keine Werte von y eine reelle Lösung $x(y)$.

Wir wollen in diesem Abschnitt das folgende Problem behandeln: Gegeben sei eine Funktion f in zwei Veränderlichen x und y. Unter welchen Bedingungen läßt sich die Gleichung

$$f(x,y) = 0 \tag{10.14}$$

nach y auflösen, d.h. läßt sich eine Funktion $x \mapsto g(x)$ finden mit

$$f(x, g(x)) \equiv 0\,?$$

Im Falle der Auflösbarkeit nennt man g die durch (10.14) bestimmte *implizite Funktion*. Was läßt sich hierbei über die Differenzierbarkeitseigenschaften von g sagen, wenn f stetig partiell differenzierbar ist?

Allgemeiner werden wir $\mathbb{R}^N$–wertige Funktionen in Veränderlichen $x \in \mathbb{R}^M$, $y \in \mathbb{R}^N$ betrachten.

10.13. Satz (über implizite Funktionen). *Sei $\Omega \subseteq \mathbb{R}^{M+N} = \mathbb{R}^M \times \mathbb{R}^N$ offen und sei $f \in C^1(\Omega, \mathbb{R}^N)$. Für einen Punkt $(x_0, y_0) \in \Omega$ gelte*

$$f(x_0, y_0) = 0.$$

Ferner sei die Matrix

$$A := \frac{\partial f}{\partial y}(x_0, y_0) := \Big(\frac{\partial f_j}{\partial y_k}(x_0, y_0)\Big)_{\substack{j=1,\dots,N\\k=1,\dots,N}}$$

invertierbar. Dann gilt:

(a) *Es gibt ein $\delta_1 > 0$ und ein $\delta_2 > 0$, so daß zu jedem $x \in U_1 := U_{\delta_1}(x_0)$ genau ein $g(x) \in U_2 := U_{\delta_2}(y_0)$ existiert mit*

$$f(x, g(x)) = 0\,.$$

(b) *Die durch (a) definierte Funktion $g : U_1 \to U_2 \subset \mathbb{R}^N$ ist stetig partiell differenzierbar und für alle $x \in U_1$ gilt*

$$J_g(x) = -\Big(\frac{\partial f}{\partial y}(x, g(x))\Big)^{-1}\frac{\partial f}{\partial x}(x, g(x)). \tag{10.15}$$

(c) *Ist $f \in C^k(\Omega, \mathbb{R}^N)$ für ein $k \in \mathbb{N} \cup \{\infty\}$, so gilt auch $g \in C^k(U_1, \mathbb{R}^N)$.*

Beweis. Da f nach Voraussetzung stetig partiell differenzierbar ist, gibt es $\delta_1', \delta_2 > 0$ mit $B_{\delta_1'}(x_0) \times B_{\delta_2}(y_0) \subset \Omega$ und

$$\sup\Big\{\Big\|\frac{\partial f}{\partial y}(x_0, y_0) - \frac{\partial f}{\partial y}(x, y)\Big\|\,;\, |x - x_0| \le \delta_1'\,,\, |y - y_0| \le \delta_2\Big\} < \frac{1}{2\|A^{-1}\|}\,. \tag{10.16}$$

Da $A = \frac{\partial f}{\partial y}(x_0, y_0)$ invertierbar ist und die Menge der invertierbaren Matrizen in $M_{N\times N}(\mathbb{R})$ offen ist, können wir wegen der Stetigkeit von $\frac{\partial f}{\partial y}$ durch eventuelles Verkleinern von δ_1' und δ_2 erreichen, daß gilt:

$$\forall u \in B_{\delta_1'}(x_0)\,,\, v \in B_{\delta_2}(y_0) : \qquad \frac{\partial f}{\partial y}(u, v) \quad \text{ist in } M_{N\times N}(\mathbb{R}) \text{ invertierbar.} \tag{10.17}$$

Wegen der Stetigkeit von f in (x_0, y_0) und wegen $f(x_0, y_0) = 0$ gibt es ein $\delta_1 \in (0, \delta_1')$ mit

$$\forall x \in B_{\delta_1}(x_0) : \qquad |f(x, y_0)| < \frac{\delta_2}{3\|A^{-1}\|}\,. \tag{10.18}$$

Angeregt durch das Newtonverfahren betrachten wir für alle $x \in B_{\delta_1}(x_0)$ die Abbildung

$$T_x : B_{\delta_2}(y_0) \to \mathbb{R}^N\,, \qquad y \mapsto T_x(y) := y - A^{-1}f(x, y)\,.$$

Für die Jacobi–Matrix $J_{T_x}(y)$ von T_x in $y \in B_{\delta_2}(y_0)$ erhält man unter Verwendung von (10.16) die Abschätzung (wobei E_N die Einheitsmatrix in $M_{N\times N}(\mathbb{R})$ bezeichne):

$$\begin{aligned}\|J_{T_x}(y)\| =&\Big\|E_N - A^{-1}\frac{\partial f}{\partial y}(x,y)\Big\| = \Big\|A^{-1}\Big(\frac{\partial f}{\partial y}(x_0,y_0) - \frac{\partial f}{\partial y}(x,y)\Big)\Big\|\\ \le&\|A^{-1}\|\cdot\Big\|\frac{\partial f}{\partial y}(x_0,y_0) - \frac{\partial f}{\partial y}(x,y)\Big\| < \frac{1}{2}\end{aligned}$$

Unter Verwendung dieser Ungleichung folgt für alle $x \in B_{\delta_1}(x_0)$ und alle $y_1, y_2 \in B_{\delta_2}(y_0)$ folgt nach dem Mittelwertsatz 7.30 der Differentialrechnung:

$$|T_x(y_1) - T_x(y_2)| \le \sup_{t\in[0,1]}\big\|J_{T_x}(y_1 + t(y_2-y_1))\big\|\cdot|y_1-y_2| \le \frac{1}{2}|y_1-y_2|\,. \tag{10.19}$$

Für alle $y \in B_{\delta_2}(y_0)$ und alle $x \in B_{\delta_1}(x_0)$ erhalten wir mit (10.18) und (10.19)

$$\begin{aligned}|T_x(y) - y_0| =&|T_x(y) - T_x(y_0) - A^{-1}f(x,y_0)|\\ \le&|T_x(y) - T_x(y_0)| + \|A^{-1}\|\cdot|f(x,y_0)|\\ \le&\frac{1}{2}|y-y_0| + \frac{\delta_2}{3} \le \frac{5}{6}\delta_2\,.\end{aligned} \tag{10.20}$$

Wir halten fest: Für alle $x \in B_{\delta_1}(x_0)$ ist T_x eine stark kontrahierende Abbildung der abgeschlossenen und daher vollständigen Teilmenge $B_{\delta_2}(y_0)$ des $\mathbb{R}^N$ in sich mit Kontraktionskonstante $K \le 1/2$. Nach dem Banachschen Fixpunktsatz 10.4 gibt es daher genau ein $g(x) \in B_{\delta_2}(y_0)$ mit $T_x(g(x)) = g(x)$, d.h. mit $A^{-1}f(x,g(x)) = 0$, d.h. mit $f(x,g(x)) = 0$. Insbesondere ist $g(x_0) = y_0$. Wegen $T_x(g(x)) = g(x)$ fokgt aus (10.20) auch $g(x) \in B_{\frac{5}{6}\delta_2}(y_0) \subset U_{\delta_2}(y_0)$ für alle $x \in B_{\delta_1}(x_0)$. Damit ist (a) bewiesen.

Definieren wir induktiv die Funktionen $g_n : B_{\delta_1}(x_0) \to \mathbb{R}^N$, $n \in \mathbb{N}_0$, durch

$$g_0 \equiv y_0 \quad \text{und } g_{n+1}(x) := T_x(g_n(x)) = g_n(x) + A^{-1}f(x,g_n(x)) \quad \text{für alle } x \in B_{\delta_1}(x_0), n \in \mathbb{N}_0\,,$$

so folgt durch vollständige Induktion die Stetigkeit von g_n für alle $n \in \mathbb{N}_0$. Aus der a priori Abschätzung (10.2) im Banachschen Fixpunktsatz folgt

$$\forall x \in B_{\delta_1}(x_0): \qquad |g_n(x) - g(x)| \le \frac{\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}}|g_1(x) - g_0(x)|$$

und somit

$$\|g_n - g\|_{B_{\delta_1}(x_0)} \le 2^{1-n}\|g_1 - g_0\|_{B_{\delta_1}(x_0)} \to 0 \quad \text{für } n \to \infty\,.$$

Die Folge $(g_n)_{n=0}^\infty$ ist also eine auf $B_{\delta_1}(x_0)$ gleichmäßig gegen g konvergente Folge von auf $B_{\delta_1}(x_0)$ stetigen Funktionen. Nach Satz 9.4 ist also auch die Funktion g auf $B_{\delta_1}(x_0)$ stetig. Ferner gilt $g(x) \in B_{\delta_2}(y_0)$ für alle $x \in U_{\delta_1}(x_0)$ und daher

$$\forall x \in U_{\delta_1}(x_0): \qquad g(x) = T_x(g(x)) \in B_{\frac{5}{6}\delta_2}(y_0) \subset U_{\delta_2}(g(x_0)) = U_{\delta_2}(y_0)\,. \tag{10.21}$$

Für alle $(x,y) \in \Omega$ und alle $\xi \in \mathbb{R}^M, \eta \in \mathbb{R}^N$ gilt

$$J_f(x,y)\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x,y) & \dots & \frac{\partial f_1}{\partial x_M}(x,y) & \frac{\partial f_1}{\partial y_1}(x,y) & \dots & \frac{\partial f_1}{\partial y_N}(x,y) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_N}{\partial x_1}(x,y) & \dots & \frac{\partial f_N}{\partial x_M}(x,y) & \frac{\partial f_N}{\partial y_1}(x,y) & \dots & \frac{\partial f_N}{\partial y_N}(x,y) \end{pmatrix} \begin{pmatrix}\xi_1\\ \vdots \\ \xi_M \\ \eta_1 \\ \vdots \\ \eta_N\end{pmatrix}$$

$$= \frac{\partial f}{\partial x}(x,y)\xi + \frac{\partial f}{\partial y}(x,y)\eta\,.$$

Da f auf Ω stetig partiell und somit auch total differenzierbar ist, folgt für alle $x \in U_{\delta_1}(x_0)$, $y \in U_{\delta_2}(y_0)$ und alle $\xi \in \mathbb{R}^M$, $\eta \in \mathbb{R}^N$ mit $x+\xi \in U_{\delta_1}(x_0)$ $y+\eta \in U_{\delta_2}(y_0)$:

$$f(x+\xi, y+\eta) = f(x,y) + \frac{\partial f}{\partial x}(x,y)\xi + \frac{\partial f}{\partial y}(x,y)\eta + \varphi(\xi,\eta) \tag{10.22}$$

mit

$$\frac{|\varphi(\xi,\eta)|}{|(\xi,\eta)|} \to 0 \quad \text{für } |(\xi,\eta)| \to 0\,. \tag{10.23}$$

Speziell mit $y = g(x), \eta = g(x+\xi) - g(x)$ erhalten wir für alle $x \in B_{\delta_1}(x_0)$ wegen $f(x,g(x)) = 0 = f(x+\xi, g(x+\xi))$

$$0 = \frac{\partial f}{\partial x}(x,g(x))\xi + \frac{\partial f}{\partial y}(x,g(x))(g(x+\xi)-g(x)) + \varphi(\xi, g(x+\xi)-g(x))\,.$$

Da $\frac{\partial f}{\partial y}(x,g(x))$ invertierbar ist, folgt hieraus

$$g(x+\xi) = g(x) - \Big(\frac{\partial f}{\partial y}(x,g(x))\Big)^{-1}\frac{\partial f}{\partial x}(x,g(x))\xi + \psi(\xi) \tag{10.24}$$

mit

$$\psi(\xi) = -\Big(\frac{\partial f}{\partial y}(x,g(x))\Big)^{-1}\varphi(\xi, g(x+\xi)-g(x))\,. \tag{10.25}$$

Für alle $0 \neq \xi \in \mathbb{R}^M$ mit $x+\xi \in U_{\delta_1}(x_0)$ hat man dann

$$\frac{|\psi(\xi)|}{|\xi|} \leq \Big\|\Big(\frac{\partial f}{\partial y}(x,g(x))\Big)^{-1}\Big\| \cdot \frac{|\varphi(\xi, g(x+\xi)-g(x))|}{|(\xi, g(x+\xi)-g(x))|} \cdot \frac{|(\xi, g(x+\xi)-g(x))|}{|\xi|} \tag{10.26}$$

Wegen der Stetigkeit von g gilt $g(x+\xi) \to g(x)$ für $|\xi| \to 0$ und somit auch (nach (10.23))

$$\frac{|\varphi(\xi, g(x+\xi)-g(x))|}{|(\xi, g(x+\xi)-g(x))|} \to 0 \quad \text{für } |\xi| \to 0.$$

Da der erste Faktor in (10.26) von ξ unabhängig ist, folgt

$$\frac{|\psi(\xi)|}{|\xi|} \to 0 \quad \text{für } |\xi| \to 0, \tag{10.27}$$

wenn wir zeigen können, daß es Konstanten $\varepsilon > 0$ und $C > 0$ gibt, so daß $U_\varepsilon(x) \subset U_{\delta_1}(x_0)$ gilt und

$$\text{(10.28)}\qquad \begin{aligned}\forall \xi \in U_\varepsilon(0) \setminus \{0\} : \frac{|(\xi, g(x+\xi) - g(x))|}{|\xi|} &\le \frac{|\xi|}{|\xi|} + \frac{|g(x+\xi) - g(x)|}{|\xi|} = \\ &= 1 + \frac{|g(x+\xi) - g(x)|}{|\xi|} \le 1 + C\,.\end{aligned}$$

Wegen (10.23) gibt es zu $K := \left\| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \right\|^{-1}$ ein $\varepsilon_1 > 0$ mit $U_{\varepsilon_1}(x) \subset U_{\delta_1}(x_0)$, $U_{\varepsilon_1}(g(x)) \subset U_{\delta_2}(y_0)$ und

$$\text{(10.29)}\qquad \forall \xi \in U_{\varepsilon_1}(0), \eta \in U_{\varepsilon_1}(0) : \quad |\varphi(\xi, \eta)| \le \frac{K}{2} |(\xi, \eta)| \le \frac{K}{2}(|\xi| + |\eta|)\,.$$

Wegen der Stetigkeit von g in x gibt es ein $\varepsilon > 0$ mit $\varepsilon < \varepsilon_1$, so daß für alle $\xi \in U_\varepsilon(0)$ gilt $|g(x+\xi) - g(x)| < \varepsilon_1$. Mit (10.24), (10.25) und (10.29) folgt für $0 < |\xi| < \varepsilon$:

$$\begin{aligned} |g(x+\xi) - g(x)| \le& \left| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial f}{\partial x}(x, g(x))\xi \right| + \\ &+ \left| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \varphi(\xi, g(x+\xi) - g(x)) \right| \\ \le& \left\| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \right\| \cdot |\xi| + \\ &+ \left\| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \right\| \cdot \frac{K}{2} \cdot (|\xi| + |g(x+\xi) - g(x)|) \\ \le& \left\| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \right\| \cdot |\xi| + \frac{|\xi|}{2} + \frac{1}{2} |g(x+\xi) - g(x)| \end{aligned}$$

Hieraus erhalten wir für $0 < |\xi| < \varepsilon$:

$$\frac{|g(x+\xi) - g(x)|}{|\xi|} \le C := 2 \left\| \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \right\| + 1\,.$$

Also gilt (10.28) und somit auch (10.27). Zusammen mit (10.24) folgt die totale Differenzierbarkeit von g in allen $x \in U_{\delta_1}(x_0)$ und (10.15). Mit Hilfe der Cramerschen Regel folgt hieraus wegen der Stetigkeit von g die Stetigkeit von J_g auf $U_{\delta_1}(x_0)$. Also ist g auf $U_{\delta_1}(x_0)$ stetig partiell differenzierbar. Durch vollständige Induktion nach k schließt man mit Hilfe der Cramerschen Regel und (10.15): Ist $f \in C^k(\Omega, \mathbb{R}^N)$, so ist g auf $U_{\delta_1}(x_0)$ ebenfalls k mal stetig partiell differenzierbar. □

10.14. Satz (von der Umkehrfunktion). *Sei $G \subseteq \mathbb{R}^N$ offen, $k \in \mathbb{N} \cup \{\infty\}$ und sei $f \in C^k(G, \mathbb{R}^N)$. Ist $a \in G$ und ist $J_f(a)$ invertierbar, so gibt es offene Mengen $U, V \subseteq \mathbb{R}^N$ mit $a \in U \subseteq G$ und $f(a) \in V$, so daß $f|_U : U \to V$ bijektiv ist und $(f|_U)^{-1} : V \to U$ auf V k–mal stetig partiell differenzierbar ist. Es gilt dann*

$$\text{(10.30)}\qquad J_{(f|_U)^{-1}}(f(a)) = (J_f(a))^{-1}.$$

Beweis. Die Menge $\Omega := \mathbb{R}^N \times G$ ist offen in $\mathbb{R}^{2N} = \mathbb{R}^N \times \mathbb{R}^N$ und die Abbildung

$$F : \Omega \to \mathbb{R}^N, \qquad (x, y) \mapsto F(x, y) := x - f(y),$$

ist auf Ω k–mal stetig partiell differenzierbar. Es ist $F(f(a), a) = 0$ und $\frac{\partial F}{\partial y}(f(a), a) = -J_f(a)$ nach Voraussetzung invertierbar. Nach dem Satz über implizite Funktionen gibt es also $\delta_1 > 0$ und $\delta_2 > 0$, so daß es zu jedem $x \in U_{\delta_1}(f(a))$ genau ein $g(x) \in U_{\delta_2}(a)$ gibt mit $F(x, g(x)) = 0$, d.h. mit $x = f(g(x))$ und die hierdurch definierte Funktion

$g : U_{\delta_1}(f(a)) \to \mathbb{R}^N$ ist k–mal stetig partiell differenzierbar. Wegen der Stetigkeit von f ist die Menge $U := U_{\delta_2}(a) \cap f^{-1}\big(U_{\delta_1}(f(a))\big)$ eine offene Umgebung von a und $f|_U : U \to U_{\delta_1}(f(a))$ ist bijektiv mit $(f|_U)^{-1} = g$.

(10.30) folgt mit Hilfe von (10.15) oder auch durch Anwendung der Kettenregel auf $x = f(g(x))$ für alle $x \in U_{\delta_1}(a)$ □

10.15. Beispiel. Die Abbildung

$$f : [0,\infty) \times \mathbb{R} \to \mathbb{R}^2, \quad (r,\varphi) \mapsto f(r,\varphi) := (r\cos(\varphi), r\sin(\varphi))$$

ist surjektiv und beliebig oft stetig partiell differenzierbar mit

$$J_f(r,\varphi) = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} \qquad \text{und} \quad \det(J_f(r,\varphi)) = r\,.$$

Wegen $f(r,\varphi) = f(r,\varphi+2\pi)$ ist f nicht injektiv, hat aber nach dem Satz von der Umkehrfunktion für alle $r_0 > 0$ und $\varphi_0 \in \mathbb{R}$ in einer Umgebung V eines jeden Punktes eine Umkehrfunktion $g : V \to U$ von V auf eine Umgebung U von (r_0,φ_0). Mit $(x,y) := f(r,\varphi)$ folgt $r = \sqrt{x^2+y^2}$ und

$$J_g(x,y) = J_f(r,\varphi)^{-1} = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\frac{\sin(\varphi)}{r} & \frac{\cos(\varphi)}{r} \end{pmatrix} = \frac{1}{x^2+y^2} \begin{pmatrix} x\sqrt{x^2+y^2} & y\sqrt{x^2+y^2} \\ -y & x \end{pmatrix}.$$

Im folgenden Abschnitt geben wir eine wichtige Anwendung des Satzes über implizite Funktionen auf die Untersuchung von lokalen Extremstellen bei Vorliegen von Nebenbedingungen.

4. Extrema mit Nebenbedingungen

10.16. Satz (Multiplikatorenregel von Lagrange). *Sei $\Omega \subseteq \mathbb{R}^N$ offen und seien $f \in C^1(\Omega,\mathbb{R})$ und $g = (g_1,\dots,g_L) \in C^1(\Omega,\mathbb{R}^L)$ mit $L < N$. Sei $x_0 \in \Omega$ mit $g(x_0) = 0$ und* $\operatorname{rang} J_g(x_0) = L$ *(also mit maximal möglichem Rang). Ist x_0 lokale Extremstelle von f auf $\{x \in \Omega\,;\, g(x) = 0\}$, so gibt es $\lambda_1,\dots,\lambda_L \in \mathbb{R}$ mit*

$$\frac{\partial f}{\partial x_j}(x_0) + \sum_{k=1}^{L} \lambda_k \frac{\partial g_k}{\partial x_j}(x_0) = 0 \quad \text{für } j = 1,\dots N, \tag{10.31}$$

d.h. mit

$$(\nabla f)(x_0) + J_g(x_0)\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_L \end{pmatrix} = 0\,.$$

Beweis. Wegen $\operatorname{rang} J_g(x_0) = L$ gibt es L der Variablen $x_1,\dots,x_N$ etwa (nach eventueller Umnumerierung) $x_1,\dots,x_L$, mit

$$\det\left(\frac{\partial g_j}{\partial x_k}(x_0)\right)_{\substack{j=1,\dots,L\\ k=1,\dots L}} \neq 0\,. \tag{10.32}$$

Schreiben wir $x = (x_j)_{j=1}^N$ in der Form $x = \binom{u}{v}$ mit $u = (x_j)_{j=1}^L \in \mathbb{R}^L$ und $v = (x_j)_{j=L+1}^N \in \mathbb{R}^{N-L}$ und entsprechend $x_0 = \binom{u_0}{v_0}$, so gibt es nach dem Satz über implizite Funktionen

Umgebungen U von u_0 und V von v_0 sowie eine Funktion $\varphi \in C^1(V, \mathbb{R}^L)$ mit $\varphi(v_0) = u_0$ und $g(\varphi(v), v) \equiv 0$ auf V sowie mit

$$J_\varphi(v) = -\Big(\frac{\partial g}{\partial u}(x_0)\Big)^{-1}\frac{\partial g}{\partial v}(x_0) \tag{10.33}$$

Die durch $F(v) := f(\varphi(v), v)$ für alle $v \in V$ definierte Funktion $F : V \to \mathbb{R}$ hat nach Voraussetzung in v_0 eine lokale Extremstelle. Es folgt daher mit der Kettenregel und (10.33):

$$\begin{aligned} 0 =& (\nabla F)(v_0)^t = J_F(v_0) = \frac{\partial f}{\partial u}(u_0, v_0) J_\varphi(v_0) + \frac{\partial f}{\partial v}(u_0, v_0) \\ =& -\frac{\partial f}{\partial u}(x_0)\Big(\frac{\partial g}{\partial u}(x_0)\Big)^{-1}\frac{\partial g}{\partial v}(x_0) + \frac{\partial f}{\partial v}(x_0) \end{aligned} \tag{10.34}$$

Da wegen (10.32) die Zeilen von $\frac{\partial g}{\partial u}(x_0)$ linear unabhängig sind, gibt es $\lambda_1, \dots, \lambda_L \in \mathbb{R}$ mit

$$-\frac{\partial f}{\partial u}(x_0) = \sum_{k=1}^{L} \lambda_k \frac{\partial g_k}{\partial u}(x_0) = (\lambda_1, \dots, \lambda_L)\frac{\partial g}{\partial u}(x_0)\,, \tag{10.35}$$

also mit

$$\frac{\partial f}{\partial u}(x_0) + \sum_{k=1}^{L} \lambda_k \frac{\partial g_k}{\partial u}(x_0) = 0\,.$$

Dies liefert die ersten L Gleichungen in (10.31). Setzen wir nun noch (10.35) in (10.34) ein, so folgt

$$0 = (\lambda_1, \dots, \lambda_L)\frac{\partial g}{\partial u}(x_0)\Big(\frac{\partial g}{\partial u}(x_0)\Big)^{-1}\frac{\partial g}{\partial v}(x_0) + \frac{\partial f}{\partial v}(x_0)$$

und damit

$$\frac{\partial f}{\partial v}(x_0) + (\lambda_1, \dots, \lambda_L)\frac{\partial g}{\partial v}(x_0) = 0\,,$$

d.h. es gilt auch für $j = L+1, \dots, N$

$$\frac{\partial f}{\partial x_j}(x_0) + \sum_{k=1}^{L} \lambda_k \frac{\partial g_k}{\partial x_j}(x_0) = 0\,,$$

Damit ist gezeigt, daß (10.31) erfüllt ist. □

Es gilt also: Ist $g : \Omega \to \mathbb{R}^L$ stetig differenzierbar mit $\operatorname{rang} J_g(x) = L$ für alle $x \in \Omega$ mit $g(x) = 0$, so können höchstens solche $x_0 \in \Omega$ als lokale Extremstellen von $f \in C^1(\Omega, \mathbb{R})$ unter der Nebenbedingung $g(x_0) = 0$ d.h. mit

$$g_j(x_0) = 0 \qquad \text{für } j = 1, \dots, L, \tag{10.36}$$

auftreten, für die $\lambda_1, \dots, \lambda_L \in \mathbb{R}$ existieren mit (10.31). Mit den Gleichungssystemen (10.31) und (10.36) stehen uns $N+L$ Gleichungen zur Berechnung der $N+L$ Unbekannten $\lambda_1, \dots, \lambda_N$ und $x_{0,1}, \dots, x_{0,N}$ zur Verfügung.

10.17. Beispiel. Gesucht sind alle Extremstellen der offensichtlich stetig differenzierbaren Funktion

$$f : \mathbb{R}^3 \to \mathbb{R}\,, \qquad (x, y, z) \mapsto f(x, y, z) := xyz$$

unter den Nebenbedingungen

$$g_1(x, y, z) := x + y + z - 5 = 0 \tag{10.37}$$

und

$$g_2(x, y, z) := xy + yz + zx - 8 = 0\,. \tag{10.38}$$

Mit $g := (g_1, g_2)$ gilt $g \in C^1(\mathbb{R}^3, \mathbb{R}^2)$ und

$$J_g(x) = \begin{pmatrix} 1 & 1 & 1 \\ y+z & x+z & x+y \end{pmatrix}.$$

Es folgt

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x}(x,y,z) & \frac{\partial g_1}{\partial y}(x,y,z) \\ \frac{\partial g_2}{\partial x}(x,y,z) & \frac{\partial g_2}{\partial y}(x,y,z) \end{pmatrix} = x + z - (y+z) = x - y$$

und

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x}(x,y,z) & \frac{\partial g_1}{\partial z}(x,y,z) \\ \frac{\partial g_2}{\partial x}(x,y,z) & \frac{\partial g_2}{\partial z}(x,y,z) \end{pmatrix} = x + y - (y+z) = x - z\,.$$

$J_g(x,y,z)$ hat also höchstens in solchen Punkten (x,y,z) einen Rang < 1, für die $x = y = z$ gilt. Für solche Punkte gilt aber

$$g_1(x,y,z) = 3x - 5 \neq 0 \quad \text{oder} \quad g_2(x,y,z) = 3x^2 - 8 \neq 0.$$

Für alle (x,y,z), die die Nebenbedingungen (10.37) und (10.38) erfüllen ist also $\operatorname{rang} J_g(x,y,z) = 2$. Damit ist die Rangbedingung in Satz 10.16 erfüllt. Zu lösen sind nun nach Satz 10.16 die Gleichungen (10.37) und (10.38) und

$$(10.39)\quad 0 = \frac{\partial f}{\partial x}(x,y,z) + \lambda_1 \frac{\partial g_1}{\partial x}(x,y,z) + \lambda_2 \frac{\partial g_2}{\partial x}(x,y,z) = yz + \lambda_1 + \lambda_2(y+z)$$

$$(10.40)\quad 0 = \frac{\partial f}{\partial y}(x,y,z) + \lambda_1 \frac{\partial g_1}{\partial y}(x,y,z) + \lambda_2 \frac{\partial g_2}{\partial y}(x,y,z) = xz + \lambda_1 + \lambda_2(x+z)$$

$$(10.41)\quad 0 = \frac{\partial f}{\partial z}(x,y,z) + \lambda_1 \frac{\partial g_1}{\partial z}(x,y,z) + \lambda_2 \frac{\partial g_2}{\partial z}(x,y,z) = xy + \lambda_1 + \lambda_2(x+y)$$

Addition der drei Gleichungen (10.39)–(10.41) liefert unter Verwendung von (10.37) und (10.38):

$$(10.42)\qquad 8 + 3\lambda_1 + 10\lambda_2 = 0$$

Subtrahieren wir (10.40) von (10.39) und (10.41), so erhalten wir

$$(10.43)\qquad (y-x)(z+\lambda_2) = 0 \quad \text{und} \quad (y-z)(x+\lambda_2) = 0\,.$$

Hieraus folgt

Fall (a): $x = y$ oder

Fall (b): $z = -\lambda_2$.

In Fall (a) erhalten wir mit (10.37)

$$z = 5 - 2x$$

und hieraus mit (10.38)

$$x^2 + 2x(5-2x) - 8 = 0$$

Es ergeben sich die Lösungen (unter Verwendung der rechten Gleichung in (10.43) und von (10.42))

$$x_1 = y_1 = 2, \quad z_1 = 1, \quad \lambda_2^{(1)} = -x_1 = -2 \neq -z_1, \quad \lambda_1^{(1)} = 4\,.$$

und

$$x_2 = y_2 = \frac{4}{3}, \quad z_2 = \frac{7}{3}, \quad \lambda_2^{(2)} = -x_2 = -\frac{4}{3} \neq -z_1, \quad \lambda_1^{(2)} = \frac{16}{9}\,.$$

Insbesondere sieht man hieraus, daß die Fälle (a) und (b) nicht beide gemeinsam eintreffen können.

In Fall (b) muß also $x \neq y$ gelten. Einsetzen von $\lambda_2 = -z$ in der rechten Gleichung von (10.43) liefert

$$(y-z)(x-z) = 0\,.$$

Im Fall $y = z$ erhält man wie in Fall (a):

$$y_3 = z_3 = 2 \quad \text{und} \quad x_3 = 1$$

und

$$y_4 = z_4 = \frac{4}{3}, \quad \text{und} \quad x_4 = \frac{7}{3}\,.$$

Im verbleibenden Fall $x = z$ erhält man analog

$$x_5 = z_5 = 2 \quad \text{und} \quad y_5 = 1$$

und

$$x_6 = z_6 = \frac{4}{3}, \quad \text{und} \quad y_6 = \frac{7}{3}\,.$$

Insbesondere folgt

$$(x_j, y_j, z_j) \in N_g := \{(x, y, z) \in \mathbb{R}^3\,;\, g(x, y, z) = 0\}\,.$$

Subtrahieren wir zweimal die Gleichung (10.38) von der ins Quadrat erhobenen Gleichung (10.37), so folgt für alle $(x, y, z) \in N_g$: $x^2 + y^2 + z^2 = 3$. Die Menge N_g ist als nicht leerer Schnitt einer Ebene mit einer Sphäre eine Kreislinie in $\mathbb{R}^3$ und damit kompakt. Die Funktion f nimmt also auf N_g sowohl ihr Maximum als auch ihr Minimum an. Wegen

$$\begin{aligned} f(x_j, y_j, z_j) =& 4 \quad \text{für } j = 1, 3, 5 \quad \text{und} \\ f(x_j, y_j, z_j) =& \frac{112}{27} = 4 + \frac{4}{27} \quad \text{für } j = 2, 4, 6 \end{aligned}$$

folgt: x_j ist Maximumstelle für $j = 2, 4, 6$ und Minimumstelle für $j = 1, 3, 5$.

KAPITEL 11

Weiterer Ausbau der Integralrechnung

1. Weglänge und Wegintegral

Im folgenden sei $[a,b]$ stets ein abgeschlossenes und beschränktes Intervall in $\mathbb{R}$ mit $-\infty < a < b < \infty$.

11.1. Definition. Ein *Weg* im $\mathbb{R}^N$ ist eine stetige Abbildung $\gamma : [a,b] \to \mathbb{R}^N$. Die Menge $\gamma([a,b]) = \{\gamma(t)\,;\, a \le t \le b\}$ heißt die *Spur des Weges* γ. Nach Satz 4.18 ist dies eine abgeschlossene und beschränkte Teilmenge von $\mathbb{R}^N$.

11.2. Beispiele. (a) Ist $f : [a,b] \to \mathbb{R}$ eine stetige, reellwertige Funktion, so ist durch $\gamma_f(t) := (t, f(t))$ für $a \le t \le b$ ein Weg in $\mathbb{R}^2$ definiert, der den Graphen von f durchläuft.

(b) Für $n \in \mathbb{N}$ sei die Abbildung $\gamma_n : [0,1] \to \mathbb{R}^2$ (bzw. $\gamma_n : [0,1] \to \mathbb{C}$) definiert durch $\gamma_n(t) := (\cos(2n\pi t), \sin(2n\pi t))$ (bzw. $\gamma_n(t) := \exp(i2n\pi t)$) für $0 \le t \le 1$. Der Weg γ_n durchläuft die Einheitskreislinie genau n mal im mathematischen Drehsinn (gegen den Uhrzeigersinn).

(c) Sind $x_0, x_1, \ldots, x_n$ Punkte in $\mathbb{R}^N$, so ist durch $\pi : [0,n] \to \mathbb{R}^N$ mit

$$\pi(t) := (t-j+1)x_j + (j-t)x_{j-1} \qquad \text{für } j-1 \le t \le j \quad \text{und } j = 1, \ldots, n$$

ein Weg gegeben, der der Reihe nach die Strecken von x_{j-1} nach x_j für $j = 1, \ldots, n$ durchläuft. Wir nennen π den *Polygonzug*, der die Punkte $x_0, x_1, \ldots, x_n$ durchläuft.

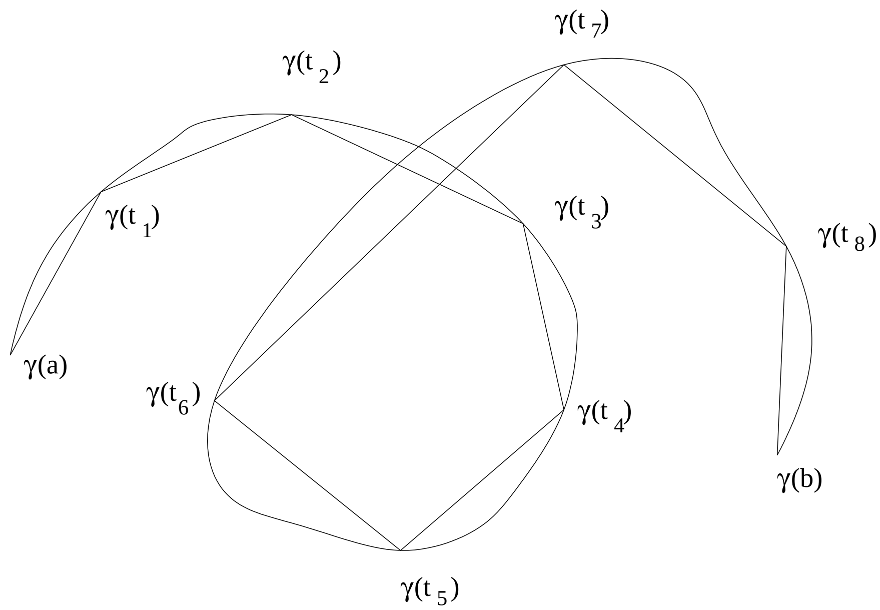


Abbildung 1. Einbeschriebener Polygonzug zu einem Weg

11.3. Definition. Ist $\gamma : [a,b] \to \mathbb{R}^N$ ein Weg und sei

$$\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\} \in \mathfrak{Z}([a,b])$$

eine Zerlegung des Intervalls $[a, b]$. Dann ist

$$L_{\mathfrak{Z}}(\gamma) := \sum_{j=1}^{n} |\gamma(t_j) - \gamma(t_{j-1})|$$

die elementargeometrische Länge des einbeschriebenen Polygonzuges, der die Punkte

$$\gamma(t_0) = \gamma(a), \gamma(t_1), \dots, \gamma(t_{n-1}), \gamma(t_n) = \gamma(b)$$

in dieser Reihenfolge durchläuft. Der Weg γ heißt *rektifizierbar*, falls

$$L(\gamma) := \sup_{\mathfrak{Z} \in \mathfrak{Z}([a,b])} L_{\mathfrak{Z}}(\gamma) < \infty\,.$$

$L(\gamma)$ heißt dann die *Länge des Weges* γ.

Nicht jeder Weg ist rektifizierbar wie das folgende Beispiel zeigt.

11.4. Beispiel. Die Abbildung $\gamma : [0, 1] \to \mathbb{R}$ mit

$$\gamma(t) := \begin{cases} t\cos\left(\frac{\pi}{t}\right) & \text{für } 0 < t \leq 1 \\ 0 & \text{für } t = 0 \end{cases}$$

ist stetig, aber der hierdurch gegebene Weg in $\mathbb{R}$ ist nicht rektifizierbar.

Zum Beweis betrachten wir für alle $n \in \mathbb{N}$ betrachten wir die Zerlegung

$$\mathfrak{Z}_n := \left\{0, \frac{1}{2n}, \frac{1}{2n-1}, \frac{1}{2n-2}, \dots, \frac{1}{2}, 1\right\}$$

des Intervalls $[0, 1]$. Man erhält wegen $\cos(j\pi) = (-1)^j$ und der Divergenz der harmonischen Reihe (Satz 3.9) für $n \to \infty$:

$$\begin{aligned} L_{\mathfrak{Z}_n}(\gamma) :=& \left|\gamma\left(\frac{1}{2n}\right) - \gamma(0)\right| + \sum_{j=1}^{2n-1} \left|\gamma\left(\frac{1}{j}\right) - \gamma\left(\frac{1}{j+1}\right)\right| \\ =& \frac{1}{2n} + \sum_{j=1}^{2n-1}\left(\frac{1}{j} + \frac{1}{j+1}\right) > \sum_{j=1}^{2n} \frac{1}{j} \to \infty\,. \end{aligned}$$

11.5. Lemma. *Sei $\gamma : [a, b] \to \mathbb{R}^N$ ein Weg in $\mathbb{R}^N$ und sei $\mathfrak{Z}_0 \in \mathfrak{Z}([a, b])$ eine beliebige Zerlegung von $[a, b]$. Dann gilt für jede Verfeinerung $\mathfrak{Z}$ von $\mathfrak{Z}_0$:*

$$L_{\mathfrak{Z}_0}(\gamma) \leq L_{\mathfrak{Z}}(\gamma)\,.$$

Beweis. Ist $\mathfrak{Z}_0 = \{a = t_0 < t_1 < \cdots < t_n = b\}$ und zunächst $\mathfrak{Z} = \mathfrak{Z}_0 \cup \{s\}$ für ein $s \in [a, b]$ mit $s \notin \{t_0, t_1, \dots, t_n\}$, so ist $t_{k-1} < s < t_k$ für ein $k \in \{1, \dots, n\}$ und wir erhalten mit der Dreiecksungleichung

(11.1)

$$\begin{aligned} L_{\mathfrak{Z}_0}(\gamma) =& \sum_{j=1}^{k-1} |\gamma(t_j) - \gamma(t_{j-1})| + |\gamma(t_k) - \gamma(t_{k-1})| + \sum_{j=k+1}^{n} |\gamma(t_j) - \gamma(t_{j-1})| \\ \leq& \sum_{j=1}^{k-1} |\gamma(t_j) - \gamma(t_{j-1})| + |\gamma(t_k) - \gamma(s)| + |\gamma(s) - \gamma(t_{k-1})| + \sum_{j=k+1}^{n} |\gamma(t_j) - \gamma(t_{j-1})| \\ \leq& L_{\mathfrak{Z}}(\gamma) \end{aligned}$$

Durch vollständige Induktion nach $p := |\mathfrak{Z} \setminus \mathfrak{Z}_0|$ zeigt man nun leicht die Behauptung für beliebige Verfeinerungen $\mathfrak{Z}$ von $\mathfrak{Z}_0$. $\square$

11.6. Satz. *Ein Weg $\gamma : [a,b] \to \mathbb{R}^N$ ist genau dann rektifizierbar, wenn gilt:*

$$\exists L \geq 0 \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall \mathfrak{Z} \in \mathfrak{Z}([a,b]) : \quad \delta(\mathfrak{Z}) < \delta \Longrightarrow |L_{\mathfrak{Z}}(\gamma) - L| < \varepsilon\,. \tag{11.2}$$

Hierbei bezeichnet

$$\delta(\mathfrak{Z}) := \max_{1 \leq j \leq n} (t_j - t_{j-1})$$

die Feinheit der Zerlegung $\mathfrak{Z} = \{a = t_0, t_1, \dots, t_n = b\} \in \mathfrak{Z}([a,b])$. Es gilt dann $L = L(\gamma)$.

Beweis. „$\Longleftarrow$“: Sei also (11.2) erfüllt und sei L wie in (11.2). Sei $\varepsilon > 0$ beliebig und sei $\delta = \delta(\varepsilon) > 0$ wie in (11.2). Ist $\mathfrak{Z}$ eine beliebige Zerlegung von $[a,b]$, so gibt es eine Verfeinerung $\mathfrak{Z}_0$ von $\mathfrak{Z}$ mit $\delta(\mathfrak{Z}_0) < \delta$. Wegen Lemma 11.5 und (11.2) folgt:

$$L_{\mathfrak{Z}}(\gamma) \leq L_{\mathfrak{Z}_0}(\gamma) < L + \varepsilon\,.$$

Also gilt für alle $\varepsilon > 0$:

$$\sup_{\mathfrak{Z} \in \mathfrak{Z}([a,b])} L_{\mathfrak{Z}}(\gamma) < L + \varepsilon\,.$$

γ ist also rektifizierbar und es ist $L(\gamma) \leq L$.

„$\Longrightarrow$“: Sei nun $\gamma : [a,b] \to \mathbb{R}^N$ rektifizierbar. Wir zeigen, daß (11.2) erfüllt ist mit $L = L(\gamma)$.

Sei also $\varepsilon > 0$ beliebig vorgegeben. Wegen $L = \sup\{L_{\mathfrak{Z}}(\gamma)\,;\, \mathfrak{Z} \in \mathfrak{Z}([a,b])\}$ gibt es eine Zerlegung $\mathfrak{Z}_0 = \{a = s_0 < s_1, \dots < s_m = b\} \in \mathfrak{Z}([a,b])$ mit

$$L - \frac{\varepsilon}{2} < L_{\mathfrak{Z}_0}(\gamma) \leq L\,. \tag{11.3}$$

Als stetige Funktion auf dem abgeschlossenen und beschränkten Intervall $[a,b]$ ist γ nach Satz 4.38 schon gleichmäßig stetig. Es gibt also ein $\delta > 0$ mit

$$\forall s,t \in [a,b] : \qquad |s - t| < \delta \implies |\gamma(s) - \gamma(t)| < \frac{\varepsilon}{4m}\,. \tag{11.4}$$

Sei nun $\mathfrak{Z} = \{a = t_0 < t_1 < \dots < t_n = b\} \in \mathfrak{Z}([a,b])$ beliebig mit $\delta(\mathfrak{Z}) < \delta$. Wie in (11.1) im Beweis zu Lemma 11.5 erhält man, wenn $k \in \{1, \dots, n\}$ so gewählt ist, daß $t_{k-1} \leq s_1 \leq t_k$ gilt:

$$\begin{aligned} 0 \leq L_{\mathfrak{Z} \cup \{s_1\}}(\gamma) - L_{\mathfrak{Z}}(\gamma) \leq& |\gamma(s_1) - \gamma(t_{k-1})| + |\gamma(t_k) - \gamma(s_1)| \\ <& \frac{\varepsilon}{4m} + \frac{\varepsilon}{4m} = \frac{\varepsilon}{2m} \qquad \text{(unter Verwendung von (11.4))}. \end{aligned}$$

Induktiv zeigt man durch sukzessive Hinzunahme der Punkte $s_1, s_2, \dots, s_{m-1}$:

$$0 \leq L_{\mathfrak{Z} \cup \mathfrak{Z}_0}(\gamma) - L_{\mathfrak{Z}}(\gamma) < (m-1)\frac{\varepsilon}{2m} < \frac{\varepsilon}{2}\,. \tag{11.5}$$

Da $\mathfrak{Z} \cup \mathfrak{Z}_0$ eine Verfeinerung von $\mathfrak{Z}_0$ ist, folgt nach (11.3) mit Lemma 11.5 und der Definition von L

$$L - \frac{\varepsilon}{2} < L_{\mathfrak{Z}_0}(\gamma) \leq L_{\mathfrak{Z} \cup \mathfrak{Z}_0}(\gamma) \leq L$$

und hieraus mit (11.5)

$$|L - L_{\mathfrak{Z}}(\gamma)| \leq |L - L_{\mathfrak{Z} \cup \mathfrak{Z}_0}(\gamma)| + |L_{\mathfrak{Z} \cup \mathfrak{Z}_0}(\gamma) - L_{\mathfrak{Z}}(\gamma)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Also ist (11.2) erfüllt. □

Wir wollen nun zeigen, daß jeder stetig differenzierbare Weg schon rektifizierbar ist. Hierzu benötigen wir:

11.7. Lemma. *Sei $\gamma : [a,b] \to \mathbb{R}^N$ stetig differenzierbar. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß für alle $s,t \in [a,b]$ mit $0 < |s-t| < \delta$ gilt:*

$$\left| \frac{1}{s-t}\,(\gamma(s) - \gamma(t)) - \gamma'(t) \right| < \varepsilon\,.$$

Beweis. (a) Wir führen den Beweis zunächst für den Spezialfall $N = 1$. Da die Ableitung $\gamma' : [a,b] \to \mathbb{R}$ stetig ist, ist γ' nach Satz 4.38 schon gleichmäßig stetig auf $[a,b]$. Es gibt also ein $\delta > 0$ mit

$$\forall\, \sigma, \tau \in [a,b] : \qquad |\sigma - \tau| < \delta \quad \Longrightarrow \quad |\gamma'(\sigma) - \gamma'(\tau)| < \varepsilon\,.$$

Sind nun $s,t \in [a,b]$ beliebig vorgegeben mit $0 < |s-t| < \delta$, so gibt es nach dem Mittelwertsatz der Differentialrechnung 5.20 ein σ zwischen s und t mit

$$\gamma'(\sigma) = \frac{1}{s-t}(\gamma(s) - \gamma(t))\,.$$

Damit folgt

$$\left| \frac{1}{s-t}(\gamma(s) - \gamma(t)) - \gamma'(t) \right| = |\gamma'(\sigma) - \gamma'(t)| < \varepsilon\,.$$

(b) Sei nun $N \in \mathbb{N}$ beliebig, also $\gamma(t) = (\gamma_1(t), \ldots, \gamma_N(t))$ mit stetig differenzierbaren Funktionen $\gamma_1, \ldots, \gamma_N : [a,b] \to \mathbb{R}$ und sei $\varepsilon > 0$ beliebig. Indem wir (a) auf die Wege $\gamma_1, \ldots, \gamma_N : [a,b] \to \mathbb{R}$ anwenden, finden wir zu $\varepsilon' := \varepsilon/\sqrt{N}$ positive Zahlen $\delta_1, \ldots, \delta_N$ mit

$$\forall s,t \in [a,b] : \qquad 0 < |s-t| < \delta_j \quad \Longrightarrow \left| \frac{1}{s-t}\,(\gamma_j(s) - \gamma_j(t)) - \gamma_j'(t) \right| < \frac{\varepsilon}{\sqrt{N}}\,.$$

für $j = 1, \ldots, N$. Damit folgt für alle $s,t \in [a,b]$ mit $|s-t| < \delta := \min\limits_{1 \le j \le N} \delta_j$:

$$\begin{aligned} \left| \frac{1}{s-t}\,(\gamma(s) - \gamma(t)) - \gamma'(t) \right| &= \left(\sum_{j=1}^N \left| \frac{1}{s-t}\,(\gamma_j(s) - \gamma_j(t)) - \gamma_j'(t) \right|^2 \right)^{1/2} < \\ &< \left(\sum_{j=1}^N \left(\frac{\varepsilon}{\sqrt{N}} \right)^2 \right)^{1/2} = \varepsilon\,. \end{aligned}$$

□

11.8. Satz. *Jeder stetig differenzierbare Weg $\gamma : [a,b] \to \mathbb{R}^N$ rektifizierbar und hat die Weglänge*

$$L(\gamma) = \int_a^b |\gamma'(t)|\, dt\,.$$

Beweis. Sei $\varepsilon > 0$ beliebig. Wegen Satz 11.6 genügt es zu zeigen, daß es ein $\delta > 0$ gibt, so daß für alle $\mathfrak{z} \in \mathfrak{Z}([a,b])$ mit $\delta(\mathfrak{z}) < \delta$ gilt:

$$\left| L_{\mathfrak{z}}(\gamma) - \int_a^b |\gamma'(t)|\, dt \right| < \varepsilon\,. \tag{11.6}$$

Nach Lemma 11.7 gibt es ein $\delta > 0$ mit

$$\forall s,t \in [a,b] : \quad 0 < |s-t| < \delta \quad \Longrightarrow \quad \left| \frac{1}{s-t}\,(\gamma(s) - \gamma(t)) - \gamma'(t) \right| < \frac{\varepsilon}{2(b-a)}\,. \tag{11.7}$$

Durch eventuelles Verkleinern von δ können wir wegen der gleichmäßigen Stetigkeit von γ' auf $[a,b]$ erreichen, daß zusätzlich gilt:

$$\forall s,t \in [a,b] : \quad 0 < |s-t| < \delta \quad \Longrightarrow \quad |\gamma'(s) - \gamma'(t)| < \frac{\varepsilon}{3(b-a)}\,. \tag{11.8}$$

Sei nun $\mathfrak{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ eine beliebige Zerlegung des Intervalls $[a, b]$ mit $\delta(\mathfrak{Z}) < \delta$. Wir erhalten unter Verwendung von (11.7) und (11.8) die folgende Abschätzung:

$$\begin{aligned}
\left|L_{\mathfrak{Z}}(\gamma) - \int_a^b |\gamma'(t)| dt\right| &= \left|\sum_{j=1}^n |\gamma(t_j) - \gamma(t_{j-1})| - \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |\gamma'(t)|\, dt\right| \\
&\leq \sum_{j=1}^n \left|\int_{t_{j-1}}^{t_j} \left|\frac{1}{t_j - t_{j-1}}(\gamma(t_j) - \gamma(t_{j-1}))\right| dt - \int_{t_{j-1}}^{t_j} |\gamma'(t)|\, dt\right| \\
&\leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left|\left|\frac{1}{t_j - t_{j-1}}(\gamma(t_j) - \gamma(t_{j-1}))\right| - |\gamma'(t)|\right| dt \\
&\leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left|\frac{1}{t_j - t_{j-1}}(\gamma(t_j) - \gamma(t_{j-1})) - \gamma'(t)\right| dt \\
&\leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \Bigg(\underbrace{\left|\frac{1}{t_j - t_{j-1}}(\gamma(t_j) - \gamma(t_{j-1})) - \gamma'(t_{j-1})\right|}_{< \frac{\varepsilon}{2(b-a)}} + \underbrace{|\gamma'(t_{j-1}) - \gamma'(t)|}_{< \frac{\varepsilon}{3(b-a)}}\Bigg) dt \\
&\leq \frac{5\varepsilon}{6(b-a)} \sum_{j=1}^n (t_j - t_{j-1}) = \frac{5}{6}\varepsilon < \varepsilon\,.
\end{aligned}$$

Also gilt (11.6) und damit die Behauptung. $\square$

11.9. Lemma. *Sei $a < c < b$ und sei $\gamma : [a, b] \to \mathbb{R}^N$ stetig. Genau dann ist γ rektifizierbar, wenn die Teilwege $\gamma|_{[a,c]}$ und $\gamma|_{[c,b]}$ rektifizierbar sind. Es gilt dann:*

$$L(\gamma) = L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]}).$$

Beweis. "$\Longrightarrow$": Sei γ rektifizierbar. Dann gilt für alle $\mathfrak{Z}_1 \in \mathfrak{Z}([a, c])$, $\mathfrak{Z}_2 \in \mathfrak{Z}([c, b])$:

$$L_{\mathfrak{Z}_1}(\gamma|_{[a,c]}) + L_{\mathfrak{Z}_2}(\gamma|_{[c,b]}) = L_{\mathfrak{Z}_1 \cup \mathfrak{Z}_2}(\gamma) \leq L(\gamma)\,.$$

Durch Übergang zum Supremum bezüglich $\mathfrak{Z}_1$ und $\mathfrak{Z}_2$ sieht man, daß $\gamma|_{[a,c]}$ und $\gamma|_{[c,b]}$ rektifizierbar sind und daß $L(\gamma) \geq L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]})$ gilt.

"$\Longleftarrow$": Seien nun die Wege $\gamma|_{[a,c]}$ und $\gamma|_{[c,b]}$ rektifizierbar und sei $\mathfrak{Z} \in \mathfrak{Z}([a, b])$ beliebig. Dann gilt mit $\mathfrak{Z}_1 := (\mathfrak{Z} \cap [a, c]) \cup \{c\}$ und $\mathfrak{Z}_2 := (\mathfrak{Z} \cap [c, b]) \cup \{c\}$:

$$L_{\mathfrak{Z}}(\gamma) \leq L_{\mathfrak{Z} \cup \{c\}}(\gamma) = L_{\mathfrak{Z}_1}(\gamma) + L_{\mathfrak{Z}_2}(\gamma) \leq L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]}).$$

Also ist γ rektifizierbar und es gilt $L(\gamma) \leq L(\gamma|_{[a,c]}) + L(\gamma|_{[c,b]})$. $\square$

11.10. Definition. Eine stetige Funktion $f : [a, b] \to \mathbb{R}^N$ heiße *stückweise stetig differenzierbar* , falls es eine Zerlegung $\{a = t_0 < t_1 < \cdots < t_n = b\}$ von $[a, b]$ gibt, so daß für alle $j = 1, \ldots, 1$ die Funktion $f|_{[t_{j-1},t_j]} : [t_{j-1}, t_j] \to \mathbb{R}^N$ auf $[t_{j-1}, t_j]$ stetig differenzierbar ist.

Aus Satz 11.8 und Lemma 11.9 erhalten wir unmittelbar:

11.11. Folgerung. *Sei $\gamma : [a, b] \to \mathbb{R}^N$ ein stückweise stetig differenzierbarer Weg. Dann ist γ rektifizierbar und es gilt:*

$$L(\gamma) = \int_a^b |\gamma'(t)|\, dt\,.$$

11.12. Folgerung. *Sei* $f : [a,b] \to \mathbb{R}$ *stückweise stetig differenzierbar. Dann ist der durch* $\gamma_f(t) := (t, f(t))$ *definierte Weg* $\gamma_f : [a,b] \to \mathbb{R}^2$ *rektifizierbar und es gilt:*

$$L(\gamma_f) = \int_a^b \sqrt{1 + f'(t)^2}\, dt\,.$$

11.13. Beispiele. (a) $\gamma : [0,1] \to \mathbb{R}^3$ sei definiert durch

$$\gamma(t) := (\cos(3\pi t), \sin(3\pi t), 4\pi t) \qquad \text{für alle } t \in [0,1].$$

Dieser Weg ist stetig differenzierbar mit

$$\gamma'(t) = (-3\pi \sin(3\pi t), 3\pi \cos(3\pi t), 4\pi).$$

Nach Satz 11.8 ist γ also rektifizierbar und es gilt:

$$L(\gamma) = \int_0^1 |\gamma'(t)| dt = \int_0^1 \sqrt{9\pi^2 + 16\pi^2} dt = 5\pi\,.$$

(b) Für $n \in \mathbb{N}$ sei $\gamma_n : [0,1] \to \mathbb{C}\,(= \mathbb{R}^2)$ definiert durch $\gamma_n(t) := \exp(2n\pi i t)$ für $0 \le t \le 1$. Auch dieser Weg ist stetig differenzierbar und somit rektifizierbar. Wegen $\gamma_n'(t) = 2n\pi i \exp(2n\pi i t)$ folgt für die Weglänge nach Satz 11.8:

$$L(\gamma_n) = \int_a^b |\gamma_n'(t)| dt = \int_a^b 2n\pi dt = 2n\pi\,.$$

11.14. Satz. *Sei* $\gamma : [a,b] \to \mathbb{R}^N$ *ein rektifizierbarer Weg und sein* $\varphi : [c,d] \to [a,b]$ *eine bijektive, stetige Abbildung mit* $\varphi(c) = a$ *und* $\varphi(d) = b$. *Dann ist der Weg* $\gamma \circ \varphi : [c,d] \to \mathbb{R}^N$ *rektifizierbar mit* $\operatorname{Spur}(\gamma \circ \varphi) = \operatorname{Spur}(\gamma)$ *und es gilt* $L(\gamma \circ \varphi) = L(\gamma)$.

Beweis. Daß γ und $\gamma \circ \varphi$ die gleiche Spur haben, ist offensichtlich. Ist $\mathfrak{Z} = \{c = t_0 < t_1 < \cdots < t_n = d\} \in \mathfrak{Z}([c,d])$ beliebig, so folgt wegen der strengen Monotonie von φ (vergl. Satz 4.32) für $j = 1, \ldots, n$, daß $\varphi(t_{j-1}) < \varphi(t_j)$ gilt. $\varphi(\mathfrak{Z}) = \{a = \varphi(t_0) < \varphi(t_1) < \cdots < \varphi(t_n) = b\}$ ist also eine Zerlegung von $[a,b]$ und es folgt

$$L_{\mathfrak{Z}}(\gamma \circ \varphi) = \sum_{j=1}^{n} |\gamma(\varphi(t_j)) - \gamma(\varphi(t_{j-1}))| = L_{\varphi(\mathfrak{Z})}(\gamma) \le L(\gamma)\,.$$

$\gamma \circ \varphi$ ist also rektifizierbar und $L(\gamma \circ \varphi) \le L(\gamma)$. Wenden wir dies auf $\gamma \circ \varphi$ statt γ und φ^{-1} statt φ an, so folgt auch $L(\gamma) = L((\gamma \circ \varphi) \circ \varphi^{-1}) \le L(\gamma \circ \varphi)$ und somit $L(\gamma \circ \varphi) = L(\gamma)$. □

11.15. Definition. Zwei Wege $\gamma_1 : [c,d] \to \mathbb{R}^N$ und $\gamma_2 : [a,b] \to \mathbb{R}^N$ heißen *äquivalent* (in Kurzform: $\gamma_1 \sim \gamma_2$), falls es eine bijektive, stetige Abbildung $\varphi : [c,d] \to [a,b]$ gibt mit $\varphi(c) = a$, $\varphi(d) = b$ und $\gamma_1 = \gamma_2 \circ \varphi$. Die Abbildung $\varphi : [c,d] \to [a,b]$ heißt dann auch eine *Parametertransformation*. Man rechnet unmittelbar nach, daß "$\sim$" tatsächlich eine Äquivalenzrelation auf der Menge aller Wege in $\mathbb{R}^N$ definiert.

Eine *Kurve* Γ in $\mathbb{R}^N$ ist eine Äquivalenzklasse von Wegen in $\mathbb{R}^N$. Die Elemente dieser Äquivalenzklassen nennen wir auch *Parametrisierungen* der Kurve Γ. Eine Kurve Γ in $\mathbb{R}^N$ heißt rektifizierbar, falls sie eine rektifizierbare Parametrisierung $\gamma : [a,b] \to \mathbb{R}^N$ besitzt. Wir definieren dann die *Kurvenlänge* von Γ durch $L(\Gamma) := L(\gamma)$. Nach Satz 11.14 ist diese unabhängig von der speziellen Parametrisierung. Die *Spur* einer Kurve Γ ist per Definition die Spur einer beliebigen Parametrisierung von Γ. Nach Satz 11.14 ist auch diese unabhängig von der speziellen Parametrisierung. Eine Kurve Γ heißt *stückweise stetig differenzierbar*, falls sie eine stückweise stetige Parametrisierung besitzt.

Durch ein Beispiel zeigen wir nun, daß nicht alle Parametrisierungen einer (stückweise) stetig differenzierbaren Kurve (stückweise) stetig differenzierbar sein müssen.

11.16. Beispiel. Die Wege $\gamma_1 : [0,\pi] \to \mathbb{R}^2$, $t \mapsto \gamma_1(t) := (\cos(t), \sin(t))$, und $\gamma_2 : [-1,1] \to \mathbb{R}^2$, $t \mapsto \gamma_2(t) := (-t, \sqrt{1-t^2})$, sind äquivalent denn $\varphi : [0,\pi] \to [0,1]$ mit $\varphi(t) := -\cos(t)$ für $0 \le t \le \pi$ ist stetig mit $\varphi(0) = -1$, $\varphi(\pi) = 1$, ist wegen $\varphi'(t) = \sin(t) > 0$ auf $(0,\pi)$ streng monoton wachsend, auf $[0,\pi]$ also bijektiv und erfüllt $\gamma_1 = \gamma_2 \circ \varphi$. γ_1 ist sogar stetig differenzierbar, aber γ_2 ist in den Randpunkten nicht differenzierbar.

Die folgende Definition von Wegintegralen ist zwar nicht die allgemeinst mögliche, reicht aber für viele Anwendungen aus.

11.17. Definition. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und sei $\gamma : [a,b] \to \mathbb{K}^N$ ein stückweise stetig differenzierbarer Weg.

(a) Für stetige Funktionen $f : \mathrm{Spur}(\gamma) \to \mathbb{K}$ definieren wir das *Wegintegral erster Art* über f längs des Weges γ durch

$$\int_\gamma f(x)dx := \int_a^b f(\gamma(t))\gamma'(t)dt\,.$$

(b) Ist $F : \mathrm{Spur}(\gamma) \to \mathbb{R}^N$ eine stetige $\mathbb{R}^N$–wertige Funktion, so ist das *Wegintegral zweiter Art* über F längs des Weges γ definiert durch:

$$\int_\gamma \langle F(x), dx\rangle := \int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle dt\,.$$

Physikalisch kann man dies interpretieren als die Arbeit, die entlang des Weges γ bei Einwirkung der äußeren Kraft F geleistet wird.

11.18. Lemma. *Die in Definition 11.17 definierten Wegintegrale ändern sich nicht bei stetig differenzierbaren Parametertransformationen. Ist $\varphi : [c,d] \to [a,b]$ bijektiv und stetig differenzierbar mit $\varphi(c) = a$ und $\varphi(d) = b$, so gilt*

$$\int_\gamma f(x)dx = \int_{\gamma\circ\varphi} f(x)dx\,; \qquad \int_\gamma \langle F(x), dx\rangle = \int_{\gamma\circ\varphi} \langle F(x), dx\rangle\,.$$

Beweis. Nach Definition der Wegintegrale gilt:

$$\begin{aligned}\int_{\gamma\circ\varphi} f(x)dx &= \int_c^d f(\gamma(\varphi(t)))(\gamma\circ\varphi)'(t)dt = \int_c^d f(\gamma(\varphi(t)))\gamma'(\varphi(t))\varphi'(t)dt \\ &= \int_{\varphi(c)}^{\varphi(d)} f(\gamma(s))\gamma'(s)ds = \int_\gamma f(x)dx\end{aligned}$$

sowie

$$\begin{aligned}\int_{\gamma\circ\varphi} \langle F(x), dx\rangle &= \int_c^d \langle F(\gamma(\varphi(t))), (\gamma\circ\varphi)'(t)\rangle dt = \int_c^d \langle F(\gamma(\varphi(t))), \varphi'(t)\gamma'(\varphi(t))\rangle dt \\ &= \int_{\varphi(c)}^{\varphi(d)} \langle F(\gamma(s)), \gamma'(s)\rangle ds = \int_\gamma \langle F(x), dx\rangle\,.\end{aligned}$$

□

11.19. Lemma (Rechenregeln). (a) *Sei $\gamma : [a,b] \to \mathbb{K}^N$ ein stückweise stetig differenzierbarer Weg. Dann gilt für alle stetigen Funktionen $f, g : \gamma([a,b]) \to \mathbb{K}$:*

(i) $\forall\, \alpha, \beta \in \mathbb{K}$: $\displaystyle\int_\gamma (\alpha f + \beta g)(x)dx = \alpha\int_\gamma f(x)dx + \beta\int_\gamma g(x)dx.$

(ii) $\left|\int_\gamma f(x)dx\right| \le L(\gamma)\|f\|_{\gamma([a,b])}$.

(b) *Sei* $\gamma : [a,b] \to \mathbb{R}^N$ *ein stückweise stetig differenzierbarer Weg. Dann gilt für alle stetigen Funktionen* $F, G : \gamma([a,b]) \to \mathbb{R}^N$*:*

(i) $\forall\, \alpha, \beta \in \mathbb{R}: \quad \int_\gamma \langle(\alpha F + \beta G)(x), dx\rangle = \alpha \int_\gamma \langle F(x), dx\rangle + \beta \int_\gamma \langle G(x), dx\rangle$.

(ii) $\left|\int_\gamma \langle F(x), dx\rangle\right| \le L(\gamma)\|F\|_{\gamma([a,b])}$.

Beweis. (i) folgt in beiden Fällen aus den entsprechenden Rechenregeln für das Riemann–Integral.

Zu (ii) Nach Definition der Wegintegrale und unter Verwendung von Lemma 6.8 und Folgerung 11.11 erhalten wir:

$$\begin{aligned}\left|\int_\gamma f(x)dx\right| &= \left|\int_a^b f(\gamma(t))\gamma'(t)dt\right| \le \int_a^b |f(\gamma(t))| \cdot |\gamma'(t)|dt \\ &\le \|f\|_{\gamma([a,b])} \int_a^b |\gamma'(t)|dt = L(\gamma)\|f\|_{\gamma([a,b])}\end{aligned}$$

sowie

$$\begin{aligned}\left|\int_\gamma \langle F(x), dx\rangle\right| &= \left|\int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle dt\right| \le \int_a^b |\langle F(\gamma(t))\gamma'(t)\rangle|dt \le \int_a^b |F(\gamma(t))| \cdot |\gamma'(t)|dt \\ &\le \|F\|_{\gamma([a,b])} \int_a^b |\gamma'(t)|dt = L(\gamma)\|F\|_{\gamma([a,b])}\,.\end{aligned}$$

□

Analog wie in Satz 7.9 erhalten wir auch für Wegintegrale, deren Integrand stetig von einem Parameter abhängt, eine Stetigkeitsaussage:

11.20. Satz. *Sei* $K \subset \mathbb{R}^M$ *beschränkt und abgeschlossen.*

(a) *Sei* $\gamma : [a,b] \to \mathbb{K}^N$ *ein stückweise stetig differenzierbarer Weg. Dann gilt für alle stetigen Funktionen* $f : \gamma([a,b]) \times K \to \mathbb{K}$*: Die durch*

$$h(u) := \int_\gamma f(x,u)dx \qquad \text{für alle } u \in K$$

definierte Funktion $h : K \to \mathbb{K}^N$ *ist gleichmäßig stetig.*

(b) *Sei* $\gamma : [a,b] \to \mathbb{R}^N$ *ein stückweise stetig differenzierbarer Weg. Dann gilt für alle stetigen Funktionen* $F : \gamma([a,b]) \times K \to \mathbb{R}^N$*: Die durch*

$$h(u) := \int_\gamma \langle F(x,u), dx\rangle \qquad \text{für alle } u \in K$$

definierte Funktion $h : K \to \mathbb{R}$ *ist gleichmäßig stetig.*

Beweis. Da auch $H := \gamma([a,b]) \times K$ beschränkt und abgeschlossen ist, muß f bzw. F schon gleichmäßig stetig auf H sein (vergl. Satz 4.38). Zu beliebig vorgegebenem $\varepsilon > 0$ gibt es also ein $\delta > 0$, so daß für alle $x, y \in \gamma([a,b])$, $u, v \in K$ mit $|(x,u) - (y,v)| < \delta$ gilt

$$|f(x,u) - f(y,v)| < \frac{\varepsilon}{L(\gamma)+1} \quad \text{bzw.} \quad |F(x,u) - F(y,v)| < \frac{\varepsilon}{L(\gamma)+1}\,.$$

Für alle $u, v \in K$ mit $|u - v| < \delta$ folgt mit Lemma 11.19 (und dem zugehörigen Beweis):

$$\begin{aligned}|h(u) - h(v)| &= \left| \int_\gamma f(x,u) - f(x,v)dx \right| \leq \int_a^b |f(\gamma(t),u) - f(\gamma(t),v)| \cdot |\gamma'(t)|dt \\ &\leq \frac{\varepsilon}{L(\gamma)+1} L(\gamma) < \varepsilon\end{aligned}$$

bzw.

$$\begin{aligned}|h(u) - h(v)| &= \left| \int_\gamma \langle F(x,u) - F(x,v), dx\rangle \right| \leq \int_a^b |F(\gamma(t),u) - F(\gamma(t),v)| \cdot |\gamma'(t)|dt \\ &\leq \frac{\varepsilon}{L(\gamma)+1} L(\gamma) < \varepsilon\end{aligned}$$

In beiden Fällen ist also h auf K gleichmäßig stetig. □

2. Integration in mehreren Veränderlichen

Seien $I_j := [a_j, b_j]$, $j = 1, \ldots, N$, kompakte Intervalle, $2 \leq N \in \mathbb{N}$ und sei

$$Q := I_1 \times \cdots \times I_N$$

der von diesen Intervallen aufgespannte achsenparallele Quader in $\mathbb{R}^N$. Ist $f : Q \to \mathbb{R}^M$ eine stetige Funktion und π eine Permutation von $1, 2, \ldots, N$, so existiert das Riemann-Integral

$$h_1(\xi) := \int_{a_{\pi(1)}}^{\xi_{\pi(1)}} f(\xi_1, \ldots, \xi_{\pi(1)-1}, x_{\pi(1)}, \xi_{\pi(1)+1}, \ldots, \xi_N) dx_{\pi(1)}\,, \qquad \xi = (\xi_1, \ldots, \xi_N) \in Q.$$

Die hierdurch definierte Funktion $h_1 : Q \to \mathbb{R}^M$ ist nach dem Hauptsatz der Differential– und Integralrechnung bezüglich der Variablen $\xi_{\pi(1)}$stetig partiell differenzierbar und es gilt

$$\frac{\partial h_1}{\partial \xi_{\pi(1)}}(\xi) = h_0(\xi) := f(\xi)\,, \qquad \text{für alle } \xi \in Q.$$

Wir zeigen:
h_1 ist stetig auf Q.

Beweis. Seien $\xi \in Q$ und $\varepsilon > 0$ beliebig vorgegeben. Nach dem Satz 7.9 über die stetige Parameterabhängigkeit bei Riemann–Integralen gibt es ein $\delta_1 > 0$, so daß für alle Punkte $(\eta_1, \ldots, \eta_{\pi(1)-1}, \eta_{\pi(1)+1}, \ldots, \eta_N) \in \mathbb{R}^{N-1}$ mit $\eta_j \in I_j$, und $|\xi_j - \eta_j| < \delta_1$ für alle $j = 1, \ldots, \pi(1) - 1, \pi(1) + 1, \ldots, N$ gilt:

$$|h_1(\xi) - h_1(\eta_1, \ldots, \eta_{\pi(1)-1}, \xi_{\pi(1)}, \eta_{\pi(1)+1}, \ldots, \eta_N)| < \frac{\varepsilon}{2}.$$

Mit

$$\delta := \min\left\{\delta_1, \frac{\varepsilon}{2\|f\|_Q + 1}\right\}$$

folgt also für alle $\eta \in Q$ mit $|\xi - \eta| < \delta$:

$$\begin{aligned}|h_1(\xi) - h_1(\eta)| \leq &|h_1(\xi) - h_1(\eta_1, \ldots, \eta_{\pi(1)-1}, \xi_{\pi(1)}, \eta_{\pi(1)+1}, \ldots, \eta_N)| + \\ &|h_1(\eta_1, \ldots, \eta_{\pi(1)-1}, \xi_{\pi(1)}, \eta_{\pi(1)+1}, \ldots, \eta_N) - h_1(\eta)| \\ < &\frac{\varepsilon}{2} + \delta\|f\|_Q < \varepsilon,\end{aligned}$$

womit die Stetigkeit von h_1 gezeigt ist. □

Die Funktion h_1 kann nun bezüglich der $\pi(2)$–ten Variablen integriert werden und durch

$$h_2(\xi) := \int_{a_{\pi(2)}}^{\xi_{\pi(2)}} h_1(\xi_1, \ldots, \xi_{\pi(2)-1}, x_{\pi(2)}, \xi_{\pi(2)+1}, \ldots, \xi_N) dx_{\pi(2)}\,, \qquad \xi = (\xi_1, \ldots, \xi_N) \in Q.$$

ist eine stetige Funktion $h_2 : Q \to \mathbb{R}^M$ gegeben, die bezüglich $\xi_{\pi(2)}$ stetig partiell differenzierbar ist mit

$$\frac{\partial h_2}{\partial \xi_{\pi(2)}}(\xi) = h_1(\xi)\,, \qquad \text{für alle } \xi \in Q.$$

Nach dem Satz über die Vertauschbarkeit von Differentiation und Integration (Folgerung 7.8) ist h_2 auch weiterhin bezüglich $\xi_{\pi(1)}$ stetig partiell differenzierbar und es gilt

$$\begin{aligned}\frac{\partial h_2}{\partial \xi_{\pi(1)}}(\xi) &= \int_{a_{\pi(2)}}^{\xi_{\pi(2)}} \frac{\partial h_1}{\partial \xi_{\pi(1)}}(\xi_1, \ldots, \xi_{\pi(2)-1}, x_{\pi(2)}, \xi_{\pi(2)+1}, \ldots, \xi_N) dx_{\pi(2)} \\ &= \int_{a_{\pi(2)}}^{\xi_{\pi(2)}} f(\xi_1, \ldots, \xi_{\pi(2)-1}, x_{\pi(2)}, \xi_{\pi(2)+1}, \ldots, \xi_N) dx_{\pi(2)}\end{aligned}$$

für alle $\xi \in Q$. Insbesondere folgt mit $\xi = (\xi_1, \ldots, \xi_N) \in Q$, $\xi_{\pi(1)} = b_{\pi(1)}$, $\xi_{\pi(2)} = b_{\pi(2)}$ sowie in den inneren Integralen $x = (x_1, \ldots, x_N)$ mit $x_j := \xi_j$ für alle $j \in \{1, \ldots, N\} \setminus \{\pi(1), \pi(2)\}$, daß gilt:

$$\begin{aligned}\int_{a_{\pi(2)}}^{b_{\pi(2)}} \Big(\int_{a_{\pi(1)}}^{b_{\pi(1)}} f(x) dx_{\pi(1)} \Big) dx_{\pi(2)} &= h_2(\xi) \\ &= \int_{a_{\pi(1)}}^{b_{\pi(1)}} \frac{\partial h_2}{\partial x_{\pi(1)}}(x_1, \ldots, x_{\pi(2)-1}, b_{\pi(2)}, x_{\pi(1)+1}, \ldots, x_N) dx_{\pi(1)} \\ &= \int_{a_{\pi(1)}}^{b_{\pi(1)}} \Big(\int_{a_{\pi(2)}}^{b_{\pi(2)}} f(x)\, dx_{\pi(2)} \Big) dx_{\pi(1)}\,.\end{aligned}$$

Vertauschen der Integrationsreihenfolge führt also zum gleichen Ergebnis. Da sich jede Permutation von $1, 2, \ldots, N$ als Hintereinanderausführung von endlich vielen Transpositionen schreiben läßt, erhalten wir:

11.21. Satz (Elementare Form des Satzes von Fubini). *Seien $I_j := [a_j, b_j]$, $j = 1, \ldots, N$, kompakte Intervalle und sei $Q := I_1 \times \cdots \times I_N$. Ist $f : Q \to \mathbb{R}^M$ stetig, so existiert für alle Permutationen π von $1, 2, \ldots, N$ das iterierte Riemann–Integral*

$$I_\pi := \int_{a_{\pi(N)}}^{b_{\pi(N)}} \Big(\int_{a_{\pi(N-1)}}^{b_{\pi(N-1)}} \Big(\ldots \int_{a_{\pi(2)}}^{b_{\pi(2)}} \Big(\int_{a_{\pi(1)}}^{b_{\pi(1)}} f(x)\, dx_{\pi(1)} \Big) dx_{\pi(2)} \Big) \ldots dx_{\pi(N-1)} \Big) dx_{\pi(N)}$$

und es gilt

$$I_\pi = \int_{a_N}^{b_N} \Big(\int_{a_{N-1}}^{b_{N-1}} \Big(\ldots \int_{a_2}^{b_2} \Big(\int_{a_1}^{b_1} f(x)\, dx_1 \Big) dx_2 \Big) \ldots dx_{N-1} \Big) dx_N$$

Der Wert des Integrals ist also von der Reihenfolge der Integrationen unabhängig.

Aufgrund dieses Satzes definieren wir für stetige Funktionen $f : Q \to \mathbb{R}^M$:

$$\int_Q f(x)\, dx := \int_{a_{\pi(n)}}^{b_n} \Big(\int_{a_{n-1}}^{b_{n-1}} \Big(\ldots \int_{a_2}^{b_2} \Big(\int_{a_1}^{b_1} f(x)\, dx_1 \Big) dx_2 \Big) \ldots dx_{n-1} \Big) dx_n.$$

Eine sehr viel allgemeinere und besser anwendbare Fassung des Satzes von Fubini werden wir in Analysis 3 nach der Einführung des Lebesgues-Integrals kennenlernen.

Ist $f : \mathbb{R}^N \to \mathbb{R}^M$ eine Funktion, so ist der *Träger* $\operatorname{supp}(f)$ (englisch bzw. französisch: *support*) definiert als die Abschließung der Menge aller $x \in \mathbb{R}^N$ mit $f(x) \neq 0$, also:

$$\operatorname{supp}(f) := \overline{\{x \in \mathbb{R}^N \,;\, f(x) \neq 0\}}^{\mathbb{R}^N}.$$

Ist $f : \mathbb{R}^N \to \mathbb{R}^M$ eine stetige Funktion mit kompaktem Träger und ist $Q \subset \mathbb{R}^N$ ein achsenparalleler Quader mit $\operatorname{supp}(f) \subseteq Q$, so definieren wir

$$\int_{\mathbb{R}^N} f(x)dx := \int_Q f(x)\, dx\,.$$

Diese Definition ist offensichtlich unabhängig von der speziellen Wahl des achsenparallelen Quaders Q, solange nur Q den Träger von f enthält.

Das folgende Beispiel in zwei Veränderlichen zeigt, daß der Satz von Fubini im allgemeinen nicht für beliebige iterierte Riemann–Integrale gilt selbst, wenn beide iterierten Integrale existieren.

11.22. Beispiel. (Vergl. [17] Chapter 10, Exercise 2.) Zu jedem beschränkten, offenen Intervall (a, b) mit $-\infty < a < b < \infty$ gibt es eine stetige Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ mit $\operatorname{supp} \varphi \subset (a, b)$ und $\int_{-\infty}^{\infty} \varphi(t)\, dt = \int_a^b \varphi(t)\, dt = 1$, z.B. die durch

$$\varphi(t) := \max\Big\{0, \frac{4}{b-a}\Big(1 - \frac{4}{b-a}\Big|\frac{a+b}{2} - t\Big|\Big)\Big\}$$

für $t \in \mathbb{R}$ definierte Zackenfunktion. Für alle $k \in \mathbb{N}$ gibt es also eine stetige Funktion $\varphi_k : \mathbb{R} \to \mathbb{R}$ mit Träger in $(2^{-k}, 2^{1-k})$ und $\int_{-\infty}^{\infty} \varphi_k(t)\, dt = \int_{2^{-k}}^{2^{1-k}} \varphi_k(t)\, dt = 1$. Wir definieren $f : \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x, y) := \sum_{k=1}^{\infty} (\varphi_k(x) - \varphi_{k+1}(x))\varphi_k(y) = \varphi_1(x)\varphi_1(y) + \sum_{k=2}^{\infty} \varphi_k(x)(\varphi_k(y) - \varphi_{k-1}(y))$$

für alle $x, y \in \mathbb{R}$. Offensichtlich gilt $\operatorname{supp} f \subseteq [0, 1] \times [0, 1]$ und f ist auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ stetig. Man beachte hierzu, daß die Träger der Funktionen φ_k, $k \in \mathbb{N}$, paarweise disjunkt sind. Für $y \in \mathbb{R}$ und alle $x \in \mathbb{R}$ gilt

$$f(x, y) = \begin{cases} 0 & \text{falls } y \leq 0 \\ 0 & \text{falls } y > 0 \\ (\varphi_k(x) - \varphi_{k+1}(x))\varphi_k(y) & \text{falls } 0 < y \leq 1 \text{ mit } 2^{-k} < x \leq 2^{1-k} \text{ für ein } k \in \mathbb{N} \end{cases}$$

und somit $\int_{-\infty}^{\infty} f(x, y)\, dx = 0$.

Für $x \in \mathbb{R}$ und alle $y \in \mathbb{R}$ gilt

$$f(x, y) = \begin{cases} 0 & \text{falls } x \leq 0 \\ 0 & \text{falls } x > 0 \\ \varphi_1(x)\varphi_1(y) & \text{falls } 2^{-1} < x \leq 1 \\ \varphi_k(x)(\varphi_k(y) - \varphi_{k-1}(y)) & \text{falls } 0 < x \leq 2^{-1} \text{ mit } 2^{-k} < x \leq 2^{1-k} \\ & \text{für ein } k \in \mathbb{N} \text{ mit } k \geq 2. \end{cases}$$

und damit $\int_{-\infty}^{\infty} f(x, y)\, dy = \varphi_1(x)$.

Es folgt: Für alle $y \in \mathbb{R}$ ist die Funktion $x \mapsto f(x, y)$ stetig auf $\mathbb{R}$ mit Träger in $[0, 1]$ und für alle $x \in \mathbb{R}$ ist die Funktion $y \mapsto f(x, y)$ stetig auf $\mathbb{R}$ mit Träger in $[0, 1]$. Für die iterierten Riemann–Integrale gilt jedoch

$$\int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x, y)\, dx\Big)\, dy = 0 \neq 1 = \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} f(x, y)\, dy\Big)\, dx\,.$$

Es liegt nahe die Definition des mehrdimensionalen Riemann–Integrals für stetige Funktionen mit kompaktem Träger zu erweitern auf Funktionen, die sich in geeigneter Weise durch Stetige Funktionen mit kompakten Trägern approximieren lassen. Der richtige Rahmen hierfür ist die Theorie des Lebesgue–Integrals, die wir in der Analysis 3 behandeln werden. Wir betrachten hier nur ein einfaches Beispiel, in dem die Approximation naheliegend ist.

Mit S^N bezeichnen wir den *Standard-N-Simplex* im $\mathbb{R}^N$, der definiert ist durch

$$S^N := \left\{x = (x_1, \dots, x_N) \in \mathbb{R}^N\,;\, x_1 + \cdots + x_N \leq 1,\, x_k \geq 0 \text{ für } k = 1, \dots, N\right\}.$$

Ferner bezeichne $W^N := [0,1]^N$ den *Standardwürfel der Kantenlänge* 1 im $\mathbb{R}^N$. Mit diesen Bezeichnungen gilt:

11.23. Satz. *Sei $f \in C(S^N, \mathbb{R}^M)$ und sei $F : W^N \to \mathbb{R}^M$ die durch*

$$F(x) := \begin{cases} f(x) & \text{für } x \in S^N \\ 0 & \text{für } x \in W^N \setminus S^N \end{cases}$$

definierte, im allgemeinen unstetige Fortsetzung von f auf W^N. Dann existiert das Integral

$$\int_{S^N} f(x)\,dx := \int_{W^N} F(x)\,dx$$

als iteriertes Riemann–Integral und ist von der Integrationsreihenfolge unabhängig.

Beweis. Für $0 < \delta < 1$ definieren wir

$$\varphi_\delta(t) := \begin{cases} 1 & \text{für } t \leq 1 - \delta \\ \frac{1-t}{\delta} & \text{für } 1 - \delta < t \leq \\ 0 & \text{für } t > 1 \end{cases}$$

und

$$F_\delta(x) := \varphi_\delta(x_1 + \cdots + x_N)F(x) \quad \text{für alle } x = (x_1, \dots, x_N) \in W^N.$$

Die so definierte Funktion F_δ ist nun stetig auf W^N. Jeder Punkt $x \in W^N$ hat eine eindeutige Darstellung der Form $x = (x', x_N)$ mit $x' \in W^{N-1}$, $x_N \in [0,1]$. Ist $x' \in W^{N-1}$ beliebig, so ist die Menge der $x_N \in [0,1]$ mit $F(x', x_N) \neq F_\delta(x', x_n)$ entweder leer oder ein Intervall der Länge $\leq \delta$. daher folgt mit

$$h_N(x') := \int_0^1 F(x', x_N)\,dx_N\,, \qquad h_{N,\delta}(x') := \int_0^1 F_\delta(x', x_N)\,dx_N \qquad (x' \in W^{N-1}),$$

es ist

$$\begin{aligned} |h_N(x') - h_{N,\delta}(x')| &\leq \int_0^1 |F(x', x_N) - F_\delta(x', x_N)|\,dx_N \\ &= \int_0^1 |(1 - \varphi(x_1 + \cdots + x_N))F(x', x_N)|\,dx_N \\ &\leq \delta \|F\|_{W^N} = \delta \|f\|_{S^N}\,, \end{aligned} \tag{11.9}$$

wobei $\|\cdot\|_{W^N}$ bzw. $\|\cdot\|_{S^N}$ die Supremumsnorm auf W^N bzw. S^N sei. Insbesondere ist h_N stetig auf W^{N-1}, da die Folge $h_{N,1/n}$ gleichmäßig auf W^{N-1} gegen h_N konvergiert für $n \to \infty$. Die weiteren Integrationen bereiten also keine Schwierigkeiten. Aus (11.9) folgt für alle $0 < \delta < 1$:

$$\left|\int_{W^N} F(x)\,dx - \int_{W^N} F_\delta(x)\,dx\right| \leq \delta \|f\|_{S^N}$$

Da nach Satz 11.21 das Integral $\int_{W^N} F_\delta(x)\,dx$ von der Integrationsreihenfolge unabhängig ist, muß dies auch für $\int_{W^N} F(x)\,dx$ gelten. □

Wir wollen nun einen Transformationssatz für iterierte Riemann-Integrale bei Variablenwechsel herleiten. Wir beginnen mit einfachen Variablensubstitutionen.

11.24. Definition. Sei $\Omega \subseteq \mathbb{R}^N$ offen. Eine *einfache Abbildung* $\Phi : \Omega \to \mathbb{R}^N$ ist eine Abbildung der Form

$$\Phi(x) = \sum_{j=1,\, j\neq k}^{N} x_j \mathbf{e}_j + g(x)\mathbf{e}_k = x + (g(x) - x_k)\mathbf{e}_k = \begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ g(x) \\ x_{k+1} \\ \vdots \\ x_N \end{pmatrix} .$$

Hierbei sei $1 \leq k \leq N$ und $g : \Omega \to \mathbb{R}$ eine reellwertige Funktion. $\mathbf{e}_1, \dots \mathbf{e}_N$ seien die kanonischen Einheitsvektoren. Eine einfache Abbildung ist also eine Abbildung, die nur eine Koordinate verändert. Sei nun g in einem Punkt $a \in \Omega$ total differenzierbar. Dann unterscheidet sich die Jacobimatrix $J_\Phi(a)$ von Φ Im Punkt a von der Einheitsmatrix nur in der k–ten Zeile. Diese hat die Einträge

$$\frac{\partial g}{\partial x_1}(a), \dots, \frac{\partial g}{\partial x_N}(a) .$$

Für die Determinante von $J_\Phi(a)$ erhält man also

$$\det J_\phi(a) = \frac{\partial g}{\partial x_k}(a) .$$

$J_\Phi(a)$ ist also genau dann invertierbar, wenn $\frac{\partial g}{\partial x_k}(a) \neq 0$.

Ein weiterer einfacher Typ sind lineare Abbildungen $V : \mathbb{R}^N \to \mathbb{R}^n$, die zwei Variablen vertauschen, für die es also $j, k \in \{1, 2, \dots, N\}$ gibt mit.

$$V(x) = x + (x_j - x_k)\mathbf{e}_k + (x_k - x_j)\mathbf{e}_j \qquad \text{für alle } x \in \mathbb{R}^N.$$

Im Fall $j = k$ ist dies die Einheitsmatrix. Da es sich um lineare Abbildungen handelt, stimmen sie mit ihrem totalen Differential überein. Ihre Matrixdarstellungen bezüglich der kanonischen Basis im $\mathbb{R}^N$ sind Permutationsmatrizen und haben (im Fall $j \neq k$) die Determinante -1, da sie aus der Einheitsmatrix durch Vertauschen zweier Zeilen entstehen.

Ferner führen wir die Projektionen $P_k : \mathbb{R}^N \to \mathbb{R}^N$ ein mit

$$P_0(x) := 0, \quad P_j(x) := \sum_{j=1}^{k} x_j \mathbf{e}_j$$

für $x = (x_1, \dots, x_N) \in \mathbb{R}^N$, $1 \leq k \leq N$. Für diese Projektionen gilt nach Definition:

$$\ker P_k = \mathrm{LH}(\mathbf{e}_{k+1}, \dots, \mathbf{e}_N) , \quad \operatorname{ran} P_k = \mathrm{LH}(\mathbf{e}_1, \dots, \mathbf{e}_k).$$

(Hierbei bezeichnet LH die lineare Hülle).

11.25. Satz. *Sei $\Omega \subseteq \mathbb{R}^N$ offen mit $0 \in \Omega$ und sei $F \in C^1(\Omega, \mathbb{R}^N$ mit $F(0) = 0$ und invertierbarer Jacobimatrix $J_F(0)$ in 0 gibt es eine Umgebung U von 0, in der F eine Darstellung der Form*

$$F(x) = (V_1 \circ \cdots \circ V_{N-1} \circ \Phi_N \circ \cdots \circ \Phi_1)(x) \qquad (x \in U) \tag{11.10}$$

besitzt mit einfachen $C^1(U,\mathbb{R}^N)$–Abbildungen, für die gilt

$$\Phi_j(0) = 0, \quad \det J_{\Phi_j}(0) \neq 0, \qquad (x \in U)$$

und linearen Abbildungen $V_1, \dots, V_{N-1}$ die entweder zwei Variablen vertauschen oder gleich der Identität sind.

Beweis. Wir konstruieren induktiv für $k = 1, \dots, N$ Nullumgebungen W_k in $\mathbb{R}^N$, Abbildungen $F_k \in C^1(W_k,\mathbb{R}^N)$ mit $F_k(0) = 0$, invertierbarer Jacobimatrix $J_{F_j}(0)$ und

$$P_{k-1}F_k(x) = P_{k-1}x \qquad \text{für alle } x \in W_k. \tag{11.11}$$

Für $k = 1$ sind diese Bedingen erfüllt mit $W_1 := \Omega$, $F_1 := F$.

Sei nun $1 \leq k < N$ und sei W_k eine Nullumgebung in $\mathbb{R}^N$ und sei $F_k \in C^1(W_k,\mathbb{R}^N)$ mit $F_k(0) = 0$ und invertierbarer Jacobimatrix $J_{F_j}(0)$, so daß (11.11) erfüllt ist. Dann gilt nach der Induktionsvoraussetzung für alle $x \in W_k$:

$$F_k(x) = P_{k-1}(x) + \sum_{j=k}^{N} \alpha_j(x)\mathbf{e_j} \tag{11.12}$$

mit $\alpha_k, \dots, \alpha_N \in C^1(W_k,\mathbb{R})$. Es folgt

$$(DF_k)(0)\mathbf{e}_k = J_{F_k}(0)\mathbf{e}_k = \sum_{j=k}^{N} \frac{\partial \alpha_j}{\partial x_k}(0)\mathbf{e}_j.$$

Da $J_{F_k}(0)$ nach Induktionsvoraussetzung invertierbar ist, kann dieser Ausdruck nicht Null sein. Es gibt also ein $m \in \{k+1, \dots, N\}$ mit $\frac{\partial \alpha_m}{\partial x_k}(0) \neq 0$.

Sei V_k die lineare Abbildung die genau die Variablen x_k und x_m vertauscht. Im Fall $k = m$ ist dies die Einheitsmatrix, sonst ist es eine echte Vertauschungsmatrix. Wir setzen nun

$$\Phi_k(x) := x + (\alpha_m(x) - x_k)\mathbf{e}_k \qquad (x \in W_k).$$

Dann ist Φ_k eine elementare Abbildung mit $\Phi_k(0) = 0$ und $\det J_{\Phi_k}(0) = \frac{\partial \alpha_m}{\partial x_k}(0) \neq 0$. Nach dem Satz 10.14 von der Umkehrfunktion gibt es eine offene Umgebung $U_k \subseteq W_k$ von 0, so daß Φ_k die Umgebung U_k bijektiv auf eine Umgebung W_{k+1} von 0 abbildet und Φ_k^{-1} stetig partiell differenzierbar auf W_{k+1} ist. Wir definieren nun

$$F_{k+1}(\xi) := V_k(F_k \circ \Phi_k^{-1})(\xi)) \qquad (\xi \in W_{k+1}).$$

Dann ist $F_{k+1} \in C^1(W_{k+1},\mathbb{R}^N)$ mit $F_{k+1}(0) = 0$ und mit der Kettenregel und dem Multiplikationssatz für die Determinanten erhalten wir

$$\det J_{F_{k+1}}(0) = (\det V_k)(\det J_{F_k}(0))(\det J_{\Phi_k}(0))^{-1} \neq 0.$$

Für alle $x \in U_k$ gilt

$$\begin{aligned} P_kF_{k+1}(\Phi_k(x)) =& P_kV_kF_k(x) = P_k(P_{k-1}x + \alpha_m(x)\mathbf{e}_k + \beta_{k+1}\mathbf{e}_{k+1} + \dots + \beta_N\mathbf{e}_N) \\ =& P_{k-1}x + \alpha_m(x)\mathbf{e}_k = P_k\Phi_k(x) \end{aligned}$$

mit $\beta_j(x) = \alpha_j(x)$ für $m \neq j \in \{k+1, \dots, N\}$ und $\beta_m = \alpha_k$ falls $m \neq k$. Also folgt

$$P_kF_{k+1}(y) = P_ky \qquad \text{für alle } y \in W_{k+1}.$$

Nachdem wir diesen Schritt für $k = 1, \dots, N-1$ durchgeführt haben erhalten wir schließlich

$$\begin{aligned} F(x) =& F_1(x) = V_1F_2(\Phi_1(x)) = V_1V_2(F_3 \circ \Phi_2 \circ \Phi_1)(x) = \dots \\ =& V_1 \dots V_{N-1}(F_N \circ \Phi_{N-1} \circ \dots \circ \Phi_1)(x) \end{aligned}$$

für alle x in einer Umgebung U von 0. Da nach Konstruktion $P_{N-1}F_n(x) = P_{N-1}x$ für alle $x \in U$ gilt, ist $\Phi_N := F_N$ eine einfache $C^1(U,\mathbb{R}^N)$–Abbildung und es folgt die Behauptung. $\square$

Mit Hilfe des folgenden Satzes werden wir stetige Funktionen f mit kompaktem Träger zerlegen können in eine Summe von Funktionen mit “kleinen” Trägern.

11.26. Satz (Zerlegung der Eins). *Sei $(U_\alpha)_{\alpha\in A}$ eine offene Überdeckung einer kompakten Menge $K \subset \mathbb{R}^N$. Dann gibt es endlich viele stetige Funktionen $\varphi_1, \dots, \varphi_m \in C(\mathbb{R}^N, \mathbb{R})$ mit den folgenden Eigenschaften:*

(a) $\varphi_j(\mathbb{R}^N) \subseteq [0,1]$ *für alle* $j = 1, \dots, N$.
(b) *Zu jedem $j \in \{1, \dots, m\}$ gibt es ein $\alpha \in A$, so daß* $\operatorname{supp}\varphi_j$ *eine kompakte Teilmenge von U_α ist.*
(c) *Für alle $\xi \in \mathbb{R}^N$ gilt $\varphi_1(x) + \dots + \varphi_m(x) = 1$.*

Ist also $f : \mathbb{R}^N \to \mathbb{R}^M$ eine stetige Funktion mit $\operatorname{supp}(f) \subseteq K$, *so gilt*

$$f = \varphi_1 f + \dots + \varphi_m f$$

mit $\operatorname{supp}(\varphi_j f) \in U_\alpha$ *für ein $\alpha = \alpha(j) \in A$ für $j = 1, \dots, m$.*

Wir nennen $\varphi_1, \dots, \varphi_m$ dann auch eine endliche *Zerlegung* oder*Partition der Eins* zur Überdeckung $(U_\alpha)_{\alpha\in A}$.

Beweis. Zu jedem $x \in K$ fixieren wir ein $\alpha(x) \in A$ mit $x \in U_{\alpha(x)}$. Dann gibt es ein $\varepsilon(x) > 0$ mit $U_{3\varepsilon(x)}(x) \subseteq U_{\alpha(x)}$. Die Funktion ψ_x mit

$$\psi_x(\xi) := \begin{cases} 1 & \text{für } |\xi - x| \le \varepsilon(x) \\ 1 - \frac{|\xi - x| - \varepsilon(x)}{\varepsilon(x)} & \text{für } \varepsilon < |\xi - x| \le 2\varepsilon(x) \\ 0 & \text{für } |\xi - x| > 2\varepsilon(x) \end{cases}$$

hat dann kompakten Träger in $U_{3\varepsilon(x)}(x) \subseteq U_{\alpha(x)}$ und nimmt nur Werte zwischen 0 und 1 an. Da K kompakt ist, gibt es endlich viele $x_1, \dots, x_m \in K$ mit

$$K \subset U_{\varepsilon(x_1)}(x_1) \cup \ldots \cup U_{\varepsilon(x_m)}(x_m).$$

Wir definieren nun für alle $x \in \mathbb{R}^N$: $\varphi_1(x) := \psi_{x_1}(x)$ und für $2 \le j \le m$:

$$\varphi_j(x) := \psi_{x_j}(x) \prod_{k=1}^{j-1} (1 - \psi_{x_k}(x)).$$

Dann sind (a) und (b) für $\varphi_1, \dots, \varphi_m$ erfüllt und es gilt für $1 \le k \le m$

$$\sum_{j=1}^{k} \varphi_j(x) = 1 - \prod_{j=1}^{k} (1 - \psi_{x_j}(x)) \qquad (x \in \mathbb{R}^N). \tag{11.13}$$

Für $k = 1$ ist dies offensichtlich. Ist (11.13) für ein $k \in \{1, \dots, m-1\}$ erfüllt, so folgt wegen

$$\sum_{j=1}^{k+1} \varphi_j(x) = 1 - \prod_{j=1}^{k} (1 - \psi_{x_j}(x)) + \psi_{x_{k+1}}(x) \prod_{j=1}^{k} (1 - \psi_{x_j}(x)) \qquad (x \in \mathbb{R}^N)$$

die Gültigkeit von (11.13) auch für $k+1$. Ist $x \in K$ so folgt $x \in U_{\varepsilon(x_j)}(x_j)$ für wenigstens ein $j \in \{1, \dots, m\}$, also auch $1 - \psi_{x_j}(x) = 0$ und somit nach (11.13) auch die Aussage (c). $\square$

Wir können nun zumindest in einem einfachen Fall den Transformationssatz beweisen:

11.27. Satz (Transformationssatz). *Sei $\Omega \subseteq \mathbb{R}^N$ offen und sei $\Psi \in C^1(\Omega, \mathbb{R}^N)$ eine injektive Abbildung mit* $\det J_\Psi(x) \neq 0$ *für alle $x \in \Omega$. Dann gilt für alle stetigen Funktionen $f : \Omega \to \mathbb{R}^M$ mit kompaktem Träger und mit* $\operatorname{supp} f \subset \Psi(\Omega)$*:*

$$\int_{\mathbb{R}^N} f(y)\,dy = \int_{\mathbb{R}^N} f(\Psi(x))\big|\det J_\Psi(x)\big|\,dx \tag{11.14}$$

Beweis. Wegen $\det J_\Psi(x) \neq 0$ ist nach dem Satz von der Umkehrfunktion die Menge $\Psi(\Omega)$ wieder offen und die Umkehrabbildung $\Psi^{-1} : \Psi(\Omega) \to \Omega$ stetig differenzierbar. Insbesondere ist daher der Integrand der rechten Seite von (11.14) stetig mit kompaktem Träger in Ω.

Ist $\Phi : Q \to \mathbb{R}^N$ eine einfache Abbildung im Sinne von Definition 11.24 auf einem offenen Quader $Q = (a_1, b_1) \times \cdots \times (a_n, b_n) \subset \mathbb{R}^N$,

$$\Phi(x) = \sum_{j=1,\, j\neq k}^{N} x_j \mathbf{e}_j + g(x)\mathbf{e}_k \qquad (x \in Q)$$

für ein $k \in \{1, \ldots, N\}$ mit $\frac{\partial g}{\partial x_k}(x) \neq 0$ für alle $x \in Q$, so hat man für jede stetige Funktion $h : Q \to \mathbb{R}^N$ mit kompaktem in $\Phi(Q)$ enthaltenen Träger mit der Variablensubstitution $y_k = g(x)$ und $y_j := x_j$ für $k \neq j \in \{1, \ldots, N\}$:

$$\begin{aligned}\int_{[a_k,b_k]} h(\Phi(x))\big|\det J_\Phi(x)\big|\,dx_k &= \int_{[a_k,b_k]} h(x)\Big|\frac{\partial g}{\partial x_k}(x)\Big|\,dx_k\\ &= \int_{-\infty}^{\infty} h(y)\,dy_k\end{aligned}$$

und daher nach Satz 11.21

$$\int_{\mathbb{R}^N} h(y)\,dy = \int_Q h(\Phi(x))\big|\det J_\Phi(x)\big|\,dx.$$

Ist V eine Abbildung, die genau zwei Variablen vertauscht, so ist $|\det J_V(x)| = |\det V| = 1$ für alle $x \in \mathbb{R}^N$ und nach Satz 11.21 folgt für alle stetigen Funktionen $h : \mathbb{R}^N \to \mathbb{R}^M$ mit kompaktem Träger

$$\int_{\mathbb{R}^N} h(y)\,dy = \int_{\mathbb{R}^N} h(V(x))|\det J_V(x)| = |\det V|\,dx.$$

Also gilt der Satz für alle einfachen Abbildungen und alle Variablenvertauschungen. Ist $T : \mathbb{R}^N \to \mathbb{R}^N$ eine Verschiebung $x \mapsto T(x) := x + b$ für ein $b \in \mathbb{R}^N$, so folgt durch direkte Berechnung des iterierten Riemann–Integrals für alle stetigen Funktionen $h : \mathbb{R}^N \to \mathbb{R}^M$ mit kompaktem Träger:

$$\int_{\mathbb{R}^N} h(y)\,dy = \int_{\mathbb{R}^N} h(x+b)\,dx = \int_{\mathbb{R}^N} h(T(x))\big|\det J_T(x)\big|\,dx$$

Ist der Satz richtig für zwei Abbildungen Ψ_1, Ψ_2 und ist $\Psi_0 = \Psi_1 \circ \Psi_2$, so folgt für alle stetigen Funktionen h mit kompaktem Träger im Bild von Ψ_1:

$$\begin{aligned}\int_{\mathbb{R}^N} h(y)\,dy &= \int_{\mathbb{R}^N} h(\Psi_1(\xi))\big|\det J_{\Psi_1}(\xi)\big|\,d\xi\\ &= \int_{\mathbb{R}^N} h(\Psi_1(\Psi_2(x)))\big|\det J_{\Psi_1}(\Psi_2(x))\big| \cdot \big|\det J_{\Psi_2}(x)\big|\,dx\\ &= \int_{\mathbb{R}^N} h(\Psi_0(x))\big|\det J_{\Psi_0}(x)\big|\,dx\,,\end{aligned}$$

denn nach der Kettenregel und dem Multiplikationssatz für Determinanten gilt

$$\det(J_{\Psi_1}(\Psi_2(x))) \cdot \det(J_{\Psi_2}(x)) = \det(J_{\Psi_1}(\Psi_2(x))J_{\Psi_2}(x)) = \det J_{\Psi_0}(x)\,.$$

Nun hat nach Satz 11.25 jeder Punkt $a \in \Omega$ eine Umgebung U, in der gilt

$$\Psi(x) = \Psi(a) + (V_1 \circ \cdots \circ V_{k-1} \circ \Phi_k \circ \cdots \circ \Phi_1)(x-a)$$

mit Variablenvertauschungen $V_1, \ldots, V_{k-1}$ und elementaren Abbildungen $\Phi_1, \ldots, \Phi_k$. Mit $W := \Psi(U)$ folgt also nach dem, was bisher gezeigt wurde, für alle stetigen Funktionen $f : \mathbb{R}^N \to \mathbb{R}^M$ mit kompaktem in W enthaltenen Träger die Gültigkeit von (11.14). Es folgt also: Zu jedem Punkt $y \in \Psi(\Omega)$ gibt es eine Umgebung W_y von y, so daß (11.14) richtig ist für alle stetigen Funktionen mit kompaktem Träger in W_y.

Sei nun $f : \mathbb{R}^N \to \mathbb{R}^M$ eine beliebige stetige Funktion mit kompaktem Träger $K := \operatorname{supp} f \subset \Psi(\Omega)$. Nach dem Satz 11.26 über die Zerlegung der Eins gibt es dann endlich viele stetige Funktionen $\varphi_1, \ldots, \varphi_m$ mit kompakten Trägern, so daß $\operatorname{supp} \varphi_j(x) \subset W_{y_j}$ für ein $y_j \in \Psi(\Omega)$, $j = 1, \ldots, m$, und

$$f(x) = \varphi_1(x)f(x) + \cdots + \varphi_m(x)f(x) \qquad (x \in \mathbb{R}^N).$$

Also gilt (11.14) für alle $\varphi_j f$, $j = 1, \ldots, m$ und somit (wegen der Linearität des iterierten Riemann–Integrals) auch für f. □

Die bewiesene Fassung des Transformationssatzes ist noch zu restriktiv. Wir werden später eine besser anwendbare Fassung (im Rahmen der Theorie des Lebesgues–Integrals) beweisen. Durch approximatives Vorgehen wie in Satz 11.23 kann man jedoch auch schon interessante Aussagen erhalten.

11.28. Beispiel. Sei $\Psi : [0,R]\times[0,2\pi] \to \mathbb{R}^2$ definiert durch $\Psi(r,\varphi) := (r\cos\varphi, r\sin\varphi)$ für alle $r \in [0,R]$, $\varphi \in [0,2\pi]$. Dann bildet Ψ das Rechteck $Q := [0,R] \times [0,2\pi]$ surjektiv auf die abgeschlossene Kreisscheibe $D := \overline{U_R(0)}$ in $\mathbb{R}^2$ ab, ist injektiv auf dem Inneren von Q und erfüllt dort $\det J_\Psi(r,\varphi) = r \neq 0$. Indem man ähnlich wie in Satz 11.23 verfährt, zeigt man, daß für alle $f \in C(D,\mathbb{R}^M)$ gilt:

$$\int_D f(y)\,dy = \int_0^R \int_0^{2\pi} f(\Psi(r,\varphi)) r\,dr\,d\varphi\,.$$

KAPITEL 12

Der Approximationssatz von Weierstraß

Ziel dieses kurzen Abschnittes ist es, zu zeigen, daß sich jede auf einem kompakten Intervall stetige Funktion gleichmäßig durch Polynomfunktionen approximieren läßt. Wir beweisen zunächst einen abstrakten Approximationssatz.

Sei $I := [a, b]$ mit $-\infty < a < b < \infty$ ein kompaktes Intervall, $\mathbb{K} := \mathbb{R}$ oder $\mathbb{K} := \mathbb{C}$. Eine lineare Abbildung $K : C(I, \mathbb{K}) \to C(I, \mathbb{K})$ heißt *positiv*, falls für alle reellwertigen $f \in C(I, \mathbb{K})$ mit $f \geq 0$ auf I auch gilt $(Kf)(x) \geq 0$ für alle $x \in I$. Sind dann $f, g \in C(I, \mathbb{K})$ mit $f(x) - g(x) \geq 0$ für alle $x \in I$, so folgt dann $(Kg)(x) \leq (Kf)(x)$ für alle $x \in I$. Sei ferner $\pi_k : I \to \mathbb{K}$ die Funktion $x \mapsto x^k$ auf I $(k \in \mathbb{N}_0)$.

12.1. Satz (Korovkin, 1953). *Sei $(K_n)_{n=1}^{\infty}$ eine Folge von positiven Operatoren, so daß für $k = 0, 1, 2$ gilt: $K_n : C(I, \mathbb{K}) \to C(I, \mathbb{K})$ mit $K_n \pi_k \to \pi_k$ gleichmäßig auf I für $n \to \infty$. Dann gilt schon für alle $f \in C(I, \mathbb{K})$:*

$$K_n f \to f \qquad \text{gleichmäßig für } n \to \infty.$$

Beweis. (a) Sei zunächst $f \in C(I, \mathbb{R})$ und sei $\varepsilon > 0$ beliebig vorgegeben. Wegen der Kompaktheit von I ist f schon gleichmäßig stetig auf I. Es gibt also ein $\delta > 0$ mit

$$\forall x, y \in I \quad |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}.$$

Dann gilt für alle $x, t \in I$:

$$|f(x) - f(t)| \leq \begin{cases} \frac{\varepsilon}{2} & \text{falls } |x - t| < \delta \\ 2\|f\|_I \left(\frac{x-t}{\delta}\right)^2 & \text{falls } |x - t| \geq \delta \end{cases}$$

und daher

$$|f(x) - f(t)| \leq \frac{\varepsilon}{2} + 2\|f\|_I \left(\frac{x-t}{\delta}\right)^2$$

also auch

$$f(t) - \frac{\varepsilon}{2} - 2\|f\|_I \left(\frac{x-t}{\delta}\right)^2 \leq f(x) \leq f(t) + \frac{\varepsilon}{2} + 2\|f\|_I \left(\frac{x-t}{\delta}\right)^2.$$

Wenden wir hierauf bezüglich der Variablen x bei festem $t \in I$ den Operator K_n an, so folgt

$$K_n\Big(f(t) - \frac{\varepsilon}{2} - 2\|f\|_I \Big(\frac{\pi_1 - t}{\delta}\Big)^2\Big)(x) \leq (K_n f)(x) \leq K_n\Big(f(t) + \frac{\varepsilon}{2} + 2\|f\|_I \Big(\frac{\pi_1 - t}{\delta}\Big)^2\Big)(x).$$

Da die Argumente der äußeren Ausdrücke als Polynome vom Grad < 2 Linearkombinationen von π_0, π_1, π_2 sind, gibt es nach Voraussetzung ein $n_0 \in \mathbb{N}$, so daß für alle $n \geq n_0$ gilt

$$\Big\| f(t) \pm \frac{\varepsilon}{2} \pm +2\|f\|_I \Big(\frac{\pi_1 - t}{\delta}\Big)^2 - K_n\Big(f(t) \pm \frac{\varepsilon}{2} \pm 2\|f\|_I \Big(\frac{\pi_1 - t}{\delta}\Big)^2\Big)\Big\|_I < \frac{\varepsilon}{2}.$$

Wir erhalten also für alle $n \geq n_0$, $x, t \in I$:

$$f(t) - \frac{\varepsilon}{2} - 2\|f\|_I \Big(\frac{x-t}{\delta}\Big)^2 - \frac{\varepsilon}{2} < (K_n f)(x) < f(t) + \frac{\varepsilon}{2} + 2\|f\|_I \Big(\frac{x-t}{\delta}\Big)^2 + \frac{\varepsilon}{2}.$$

Speziell mit $t = x$ folgt für alle $n \geq n_0$, $x \in I$:

$$f(x) - \varepsilon < (K_n f)(x) < f(x) + \varepsilon$$

und somit $\|f - K_n f\|_I < \varepsilon$. Also konvergiert $K_n f$ gleichmäßig auf I gegen f für $n \to \infty$.

(b) Ist f komplexwertig, so wenden wir (a) auf $\operatorname{Re} f$ und $\operatorname{Im} f$ an und erhalten ebenfalls die Behauptung. □

Sei nun $I = [0,1]$. Für $f \in C(I, \mathbb{K})$ und $n \in \mathbb{N}$ definieren wir, das n–te *Bernsteinpolynom*[1] $B_n(f, \cdot)$ zu f durch

$$B_n(f,x) := \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-k)^{n-k} \qquad (x \in I).$$

Die Abbildung $B_n : f \mapsto B_n(f, \cdot)$ ist offensichtlich linear. Ist $f \geq 0$ auf I, so ist auch $B_n(f,x) \geq 0$ für alle $x \in I$. B_n ist also eine positive lineare Abbildung für alle $n \in \mathbb{N}$.

12.2. Lemma. *Für alle $x \in I$, $n \in \mathbb{N}$ gilt:*

(a) $B_n(\pi_0, x) = 1 = \pi_0(x)$.
(b) $B_n(\pi_1, x) = x = \pi_1(x)$.
(c) $B_n(\pi_2, x) = x^2 + \frac{x(1-x)}{n} = \pi_2(x) + \frac{x(1-x)}{n}$.

Insbesondere gilt $B_n(\pi_j, \cdot) \to \pi_j$ gleichmäßig auf I für $n \to \infty$.

Beweis. (a) $B_n(\pi_0, x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-k)^{n-k} = (x + (1-x))^n = 1$.

(b) Für alle $x, y \in I$, $n \in \mathbb{N}$, gilt:

$$nx(x+y)^{n-1} = x\frac{\partial (x+y)^n}{\partial x} = x\frac{\partial}{\partial x}\sum_{k=0}^{n}\binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{n} k \binom{n}{k} x^k y^{n-k}.$$

Speziell für $y = 1 - x$ folgt also (nach Division durch n):

$$x = \sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1-x)^{n-k}$$

und damit die Behauptung.

(c) Für $n = 1$ rechnet man dies unmittelbar nach. Für $n \geq 2$ schließen wir, wie folgt:

$$x^2 n(n-1)(x+y)^{n-2} = x^2 \frac{\partial^2 (x+y)^n}{\partial x^2} = x^2 \frac{\partial^2}{\partial x^2} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{n} k(k-1)\binom{n}{k} x^k y^{n-k}.$$

Speziell für $y = 1 - x$ folgt also (nach Division durch $n(n-1)$):

$$x^2 = \sum_{k=0}^{n} \frac{k(k-1)}{n(n-1)} \binom{n}{k} x^k (1-x)^{n-k}.$$

[1] Sergei Natanovich Bernstein (5.3.1880–26.10.1968)

Nach Definition von $B_n(\pi_2, x)$ gilt also für alle $n \geq 2$, $x \in I$:

$$\begin{aligned} B_n(\pi_2, x) &= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} x^k (1-x)^{n-k} \\ &= \sum_{k=0}^{n} \frac{k(k-1)}{n^2} \binom{n}{k} x^k (1-x)^{n-k} + \sum_{k=0}^{n} \frac{k}{n^2} \binom{n}{k} x^k (1-x)^{n-k} \\ &= \frac{n-1}{n} \sum_{k=0}^{n} \frac{k(k-1)}{n(n-1)} \binom{n}{k} x^k (1-x)^{n-k} + \frac{1}{n} B_n(\pi_1, x) \\ &= \frac{n-1}{n} x^2 + \frac{1}{n} x = x^2 + \frac{x(1-x)}{n}. \end{aligned}$$

und damit die Behauptung. □

Mit diesem Lemma und dem Satz von Korovkin erhalten wir unmittelbar:

12.3. Satz (Approximationssatz von Weierstraß). *Jede auf* $[0, 1]$ *stetige Funktion kann gleichmässig auf* $[0, 1]$ *durch eine Folge von Polynomfunktionen approximiert werden.*

Beweis. Nach Lemma 12.2 und dem Satz von Korovkin gilt für alle $f \in C([0, 1], \mathbb{K})$: $B_n(f, x) \to f(x)$ gleichmäßig auf $[0, 1]$ für $n \to \infty$.

□

Literaturverzeichnis

[1] Bergren, Lennart, Borwein, Jonathan, and Borwein, Peter, *Pi: A Source Book*, Springer–Verlag, New York–Berlin–Heidelberg: 1997.

[2] Forster, Otto, *Analysis 1, Differential– und Integralrechnung einer Veränderlichen*, 6. Auflage, Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden: 2001.

[3] Forster, Otto, *Analysis 2, Differentialrechnung im* $\mathbb{R}^n$ *– Gewöhnliche Differentialgleichungen*, 5. Auflage, Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden: 1984.

[4] Forster, Otto, *Analysis 3, Integralrechnung im* $\mathbb{R}^n$ *mit Anwendungen*, 3. Auflage, Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden: 1984.

[5] Grauert, Hans, und Lieb, Ingo, *Differential– und Integralrechnung* I, 4. Auflage, Springer–Verlag, Berlin–Heidelberg–New York: 1976.

[6] Heuser, Harro, *Lehrbuch der Analysis*, Teil 1, 15. Auflage, Teubner-Verlag, Stuttgart: 2003.

[7] Heuser, Harro *Lehrbuch der Analysis, Teil 2*. 12. Auflage, Teubner-Verlag, Stuttgart: 2002.

[8] Hildebrandt, Stefan *Analysis 1*, Springer–Verlag, Berlin–Heidelberg–New York: 2002.

[9] Hildebrandt, Stefan *Analysis 2*, Springer–Verlag, Berlin–Heidelberg–New York: 2003.

[10] Hoffman, Kenneth and Kunze, Ray *Linear Algebra*, 2. edition, Prentice–Hall, Inc., Englewood Cliffs, New Jersey: 1971.

[11] Kaballo, Winfried, *Einführung in die Analysis I*, Spektrum Akademischer Verlag, Heidelberg-Berlin-Oxford: 1996.

[12] Kaballo, Winfried, *Einführung in die Analysis II*, Spektrum Akademischer Verlag, Heidelberg-Berlin-Oxford: 1997.

[13] Kamke, Erich, *Mengenlehre*, 4. Auflage, Walter de Gruyter & Co., Berlin:1962.

[14] Königsberger, Konrad, *Analysis 1*, 6. Auflage, Springer–Verlag, Berlin . . . : 2004.

[15] Königsberger, Konrad, *Analysis 2*, 4. Auflage, Springer–Verlag, Berlin . . . : 2002.

[16] Lamprecht, Erich *Einführung in die Algebra*, UTB 739, Birkhäuser Verlag, Basel und Stuttgart: 1978.

[17] Rudin, Walter, *Analysis*, Oldenbourg Verlag, München–Wien: 1998. (Deutsche Übersetzung von *Principles of Mathematical Analysis*, McGraw–Hill, third edition, 1976).

[18] Schulz, Friedmar, *Analysis 1*, Oldenbourg Verlag, München–Wien: 2002.

[19] Walter, Wolfgang, *Analysis 1*, 6. Auflage, Springer–Verlag, Berlin–Heidelberg–New York: 2001.

[20] Walter, Wolfgang, *Analysis 2*, 5. Auflage, Springer–Verlag, Berlin–Heidelberg–New York: 2002.

Index