Vorabskript zur Vorlesung

Analysis I und II

Sommersemester 2010/ Wintersemester 2010/ 11

Prof. Dr. Helmut Maier
Dipl.-Math. Hans- Peter Reck

Institut für Zahlentheorie und Wahrscheinlichkeitstheorie
Universität Ulm

Inhaltsverzeichnis

Kapitel 1

Einführung, reelle Zahlen

1.1 Allgemeines

Die Analysis ist neben der Linearen Algebra eine der zwei Grunddisziplinen der Mathematik. Fast alle weiterführenden mathematischen Theorien bauen auf ihnen auf. Somit gelten auch für die Analysis die Prinzipien für den Aufbau einer mathematischen Theorie. Eine mathematische Theorie besteht aus folgenden Bestandteilen:

1. Axiome
2. Definitionen
3. Lehrsätze

Wir kommen nun zur Beschreibung dieser Bestandteile:

1. Axiome:
   Dies sind Aussagen, die ohne Beweis als gültig angenommen werden. Die Aussagen werden über Objekte getroffen, über deren Natur nichts weiter ausgesagt wird. Als einer der ersten ist Euklid in seinem Werk "Elemente" auf diese Weise vorgegangen (um ca. 300 v. Chr.). Objekte, über die in den Euklidischen Axiomen Aussagen gemacht werden, sind unter anderem Punkte und Geraden.

   Ein Axiom (A) lautet:

   (A) Durch je zwei verschiedene Punkte geht genau eine Gerade.

   Euklids Auffassung war, daß unmittelbar einleuchtend ist, was unter Punkten und Geraden zu verstehen ist und daß auch die Aussage (A) unmittelbar einleuchtend ist. Punkte und Geraden wurden dabei als Gegenstände der Natur angesehen. In der modernen Mathematik herrscht diese Auffassung nicht mehr:

   Mathematik ist keine Naturwissenschaft!

   Ist in einer Theorie von Punkten und Geraden die Rede, so existieren sie unabhängig von der Natur.
   Macht man Aussagen über Gegenstände der Natur, wie zum Beispiel:

- Licht breitet sich geradlinig aus, oder
- die Bahn eines unbeschleunigten Körpers ist eine Gerade,

so besagt dies, daß die mathematische Theorie "Euklidische Geometrie" gut geeignet ist, die Ausbreitung des Lichts oder die Bewegung von unbeschleunigten Körpern zu beschreiben. Geraden sind gute Modelle für die Ausbreitung des Lichts oder die Bewegung eines unbeschleunigten Körpers. Jedoch ist eine Gerade nichts, was in der Natur vorkommt. Diese Unabhängigkeit der Axiome von der Natur hat zur Folge, daß zum Beweis von mathematischen Tatsachen nur die Axiome und was aus ihnen rein logisch abgeleitet wurde, benützt werden dürfen, nicht jedoch die sogenannte "Anschauung", die auf Naturerfahrung beruht.

2. Definitionen:
   Diese sind im Grunde nichts anderes als Vereinbarungen, welche Namen gewisse Objekte, Tatsachen oder Eigenschaften, die in der Theorie vorkommen, haben sollen.
   Wir werden uns zum Beispiel in dieser Vorlesung auf den Standpunkt stellen, daß wir nicht wissen, was der Begriff 2 ("zwei") bedeutet, bevor er nicht definiert wurde. Die Existenz des Objekts 1 ("eins") wird in den Axiomen gefordert werden, weiter auch die Existenz der Summe. Die Definition $2 := 1 + 1$ ist dann keine mathematische Aussage, sondern eine Vereinbarung, welchen Namen die Summe $1 + 1$ bekommen soll.

3. Lehrsätze:
   Ein Lehrsatz (kurz: Satz, manchmal auch Lemma (Hilfssatz)) besteht aus drei Teilen: Voraussetzung, Behauptung und Beweis.
   Die Behauptung macht Aussagen über gewisse Objekte der Theorie (z.B. Punkte, Geraden oder Zahlen). Diese Aussage gilt im allgemeinen nur, wenn die Objekte gewisse Voraussetzungen erfüllen. Die Bedeutung der Objekte muß klar sein, d.h. sofern sie nicht schon in den Axiomen vorkommen, muß schon eine Definition vorliegen. Im Beweis wird dann die Wahrheit der Aussage durch eine Kette von Schlüssen bewiesen. Dabei dürfen nur Tatsachen benützt werden, die entweder in den Axiomen festgestellt wurden oder deren Wahrheit schon früher bewiesen wurde. Berufung auf die Anschauung, etwa "Es ist doch klar, daß durch zwei verschiedene Punkte genau eine Gerade geht" sind nicht zulässig.
   Dies bedeutet jedoch nicht, daß die Anschauung wertlos ist. Sie gibt häufig Ideen, wie Beweise zu führen sind, kann als Erinnerungsstütze dienen oder Hinweise liefern, wie die Theorie aufzubauen ist. Wir werden in dieser Vorlesung oft Sachverhalte durch Skizzen veranschaulichen; häufig sind es Skizzen von Graphen von Funktionen. Die Ableitung einer Funktion kann man sich zum Beispiel als Steigung der Tangente an den Graph der Funktion vorstellen. Diese Skizzen werden jedoch niemals als Beweismittel verwendet werden.
   Wir werden beim Aufbau der Theorie nur von Dingen sprechen, die wir schon von einem früheren Teil der Vorlesung kennen. Bei der Wahl der Beispiele, mit denen wir die Theorie illustrieren, und auch bei den Übungsaufgaben werden wir gelegentlich anders verfahren. Um interessante Beispiele zu bekommen, werden wir dann wohlbekannte Dinge, wie etwa die Grundrechenarten, voraussetzen, auch wenn wir sie in der Vorlesung noch nicht besprochen hatten.

1.2 Mengen, Relationen, Abbildungen

Die Objekte einer mathematischen Theorie werden zu Mengen zusammengefaßt. Die Objekte sind dann Elemente dieser Menge. Der Begründer der Mengenlehre, Georg Cantor, gab folgende Beschreibung:

"Definition:"

(i) Unter einer Menge X verstehen wir jede Zusammenfassung von bestimmten wohlunterschiedenen Objekten $x, y, \ldots$ unserer Anschauung oder unseres Denkens, welche die Elemente von X genannt werden, zu einem Ganzen, einem neuen Objekt $X = \{x, y, \ldots\}$ unseres Denkens.

(ii) Für "x ist Element von X" schreiben wir $x \in X$ und für "x ist nicht Element von X" dann $x \notin X$.

Diese "Definition" werden wir nicht weiter benützen. Die in ihr vorkommenden Begriffe "Zusammenfassung", "Anschauung" oder "Denken" sind nicht klarer als der zu definierende Begriff "Menge". Wir nehmen an, wir wissen, was eine Menge ist. Sie wird in dem Axiomensystem der reellen Zahlen ein Grundbegriff sein, braucht also nicht definiert zu werden.
Mengen können auf verschiedene Weisen beschrieben werden: die Auflistung ihrer Elemente in geschweiften Klammern oder durch eine charakterisierende Eigenschaft.

Beispiel 1.2.1. Die Menge $X = \{12, 13, 14, 15\}$ kann auch als

$$X = \{x|\ x \text{ ist eine natürliche Zahl mit } 12 \leq x \leq 15\}$$

beschrieben werden.

Definition 1.2.1. Wir definieren:

(i) Die Menge, die kein Element enthält, heißt leere Menge und wird mit $\emptyset$ bezeichnet.

(ii) Eine Menge X heißt genau dann Teilmenge einer Menge Y oder in der Menge Y enthalten (Schreibweise: $X \subset Y$), wenn alle Elemente von X auch Elemente von Y sind, in Zeichen: $x \in X \Rightarrow x \in Y$.

Lemma 1.2.1. *Für jede beliebige Menge X gilt $\emptyset \subset X$.*

Beweis. Nach Definition 1.2.1 (ii) muß gezeigt werden: $x \in \emptyset \Rightarrow x \in X$.
Dieser Schluß ist richtig, weil die Voraussetzung $x \in \emptyset$ für alle x falsch ist. □

Mengen können verknüpft werden, ähnlich wie Zahlen durch Addition oder Multiplikation verknüpft werden können. Die wichtigsten Verknüpfungen sind Vereinigung und Durchschnitt.

Definition 1.2.2. Es seien X und Y Mengen.

(i) Unter der Vereinigung $X \cup Y$ der Mengen X und Y verstehen wir die Menge aller Elemente, die zu X oder zu Y (oder zu beiden) gehören:

$$X \cup Y := \{x|\ x \in X \ \underline{\text{oder}}\ x \in Y\}.$$

(ii) Unter dem Durchschnitt $X \cap Y$ der Mengen X und Y verstehen wir die Menge aller Elemente, die zu X und zu Y gehören.

Sind gewisse Mengen selbst Elemente einer Menge $\mathcal{M}$, so lassen sich die Vereinigung und der Durchschnitt all dieser Mengen erklären. Um eine kürzere Schreibweise zu erhalten, führen wir Abkürzungen für die sogenannten Quantoren "es gibt ein" und "für alle" ein:
Schreibweise:
Es ist $\exists$ eine Abkürzung für "es gibt" und $\forall$ eine Abkürzung für "für alle".

Definition 1.2.3. Es sei $\mathcal{M}$ eine Menge von Mengen.

(i) Unter der Vereinigung $\bigcup_{x\in\mathcal{M}} X$ verstehen wir:

$$\bigcup_{x\in\mathcal{M}} X := \{x|\ \exists x \in \mathcal{M}:\ x \in X\}.$$

(ii) Unter dem Durchschnitt $\bigcap_{x\in\mathcal{M}} X$ verstehen wir:

$$\bigcap_{x\in\mathcal{M}} X := \{x|\ \forall x \in \mathcal{M}:\ x \in X\}.$$

Definition 1.2.4. Es seien X und Y Mengen. Die Differenz von X und Y ist

$$X\backslash Y := \{x|\ x \in X \text{ und } x \notin Y\}.$$

Falls $Y \subset X$, dann heißt die Differenz $X\backslash Y$ auch Komplement von Y in X.
Schreibweise:

$$Y^C := X\backslash Y.$$

Definition 1.2.5. Mengen X und Y heißen disjunkt, wenn sie keine gemeinsamen Elemente besitzen, wenn also $X \cap Y = \emptyset$ gilt.

Für die Verknüpfung von Mengen gelten elementare Gesetze:

Satz 1.2.1 (Elementare Mengengesetze)**.** *Es seien X, Y, Z Mengen. Dann gelten die folgenden Gesetze:*

(i) $X \cup Y = Y \cup X$ *und* $X \cap Y = Y \cap X$ *(Kommutativgesetz)*

(ii) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ *und*
$(X \cap Y) \cap Z = X \cap (Y \cap Z)$ *(Assoziativgestz)*

(iii) $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$ *und*
$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ *(Distributivgesetz)*

(iv) $Z\backslash(X \cup Y) = (Z\backslash X) \cap (Z\backslash Y)$ *und*
$Z\backslash(X \cap Y) = (Z\backslash X) \cup (Z\backslash Y)$ *(de Morgansche Regeln)*

Beweis. Wir zeigen nur (i) und einen Teil von (iii):
(i) Es gilt:

$$\begin{aligned} x \in X \cup Y \quad &\underset{\text{Def. 1.2.2 (i)}}{\Leftrightarrow} \quad x \in X \ \underline{\text{oder}}\ x \in Y \\ &\Leftrightarrow \quad x \in Y \ \underline{\text{oder}}\ x \in X \underset{\text{Def. 1.2.2 (i)}}{\Leftrightarrow} x \in Y \cup X. \end{aligned}$$

Also ist $X \cup Y = Y \cup X$.

$$\begin{aligned} x \in X \cap Y \quad &\underset{\text{Def. 1.2.2 (ii)}}{\Leftrightarrow} \quad x \in X \ \underline{\text{und}}\ x \in Y \\ &\Leftrightarrow \quad x \in Y \ \underline{\text{und}}\ x \in X \underset{\text{Def. 1.2.2 (ii)}}{\Leftrightarrow} x \in Y \cap X. \end{aligned}$$

Damit ist auch $X \cap Y = Y \cap X$.

(iii) Wir zeigen lediglich $X \cap (Y \cup Z) \subset (X \cap Y) \cup (X \cap Z)$:

$$x \in X \cap (Y \cup Z) \Rightarrow x \in X \text{ und } x \in (Y \cup Z)$$

Insbesondere gilt $x \in Y$ oder $x \in Z$.
1. Fall: $x \in Y$

$$\begin{aligned} &\Rightarrow \quad x \in X \ \underline{\text{und}}\ x \in Y \underset{\text{Def. 1.2.2 (ii)}}{\Rightarrow} x \in X \cap Y. \\ &\Rightarrow \quad x \in (X \cap Y) \cup (X \cap Z). \end{aligned}$$

2. Fall: $x \in Z$

$$\begin{aligned} &\Rightarrow \quad x \in X \ \underline{\text{und}}\ x \in Z \underset{\text{Def. 1.2.2 (ii)}}{\Rightarrow} x \in X \cap Z \\ &\Rightarrow \quad x \in (X \cap Y) \cup (X \cap Z). \end{aligned}$$

In beiden Fällen gilt also die Behauptung. □

Definition 1.2.6. Es seien x, y Objekte. Unter dem Paar (x, y) verstehen wir die Menge

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

Bemerkung 1.2.1. Im Unterschied zu der Menge $\{x, y\}$, bei der es auf die Reihenfolge der x und y nicht ankommt, es ist nämlich $\{x, y\} = \{y, x\}$, kommt es beim Paar (x, y) sehr wohl auf die Reihenfolge an: x steht an erster Stelle, y an zweiter Stelle. Die Begriffe "Reihenfolge", "erste Stelle" bzw. "zweite Stelle" sind jedoch keine Grundbegriffe. Als Mathematiker wissen wir bisher noch nicht, was sie bedeuten. Es bleibt also nur die Möglichkeit, den Begriff "Paar" auf eine kompliziert anmutende Weise, wie z. B. durch Definition 1.2.6 zu definieren.

Definition 1.2.7. Es seien X und Y Mengen. Unter dem kartesischen Produkt $X \times Y$ verstehen wir die Menge aller Paare (x, y) mit $x \in X$ und $y \in Y$, also

$$X \times Y := \{(x, y) |\ x \in X, y \in Y\}.$$

Wir schreiben $X^2 := X \times X$.

Beispiel 1.2.2. Es sei $A = \{a, b\}$ und $B = \{0, 1, \omega\}$. Dann ist

$$A \times B = \{(a, 0), (a, 1), (a, \omega), (b, 0), (b, 1), (b, \omega)\}.$$

Definition 1.2.8. Es seien X und Y Mengen. Unter einer Relation von X zu Y versteht man eine Teilmenge $R \subset X \times Y$. Dabei steht $x \in X$ in R- Relation zu $y \in Y$, wenn $(x, y) \in R$ gilt, in Zeichen xRy. Ist $X = Y$, so heißt R eine Relation auf X.
Wie Mengen können auch Relationen durch charakterisierende Eigenschaften beschrieben werden.

Beispiel 1.2.3. Es sei $X = \{0, 2, 4\}$ und $Y = \{0, 1, 2, 3\}$. Wir definieren die Relation D:

$$xDy \Leftrightarrow x = 2y,$$

oder: x ist das Doppelte von y.
So ist

$$D = \{(0, 0), (2, 1), (4, 2)\}.$$

Definition 1.2.9. Ist R eine Relation von X zu Y, so ist die inverse Relation R^{-1} als Relation von Y zu X durch

$$R^{-1} \colon = \{(y, x) \in Y \times X |\ (x, y) \in R\}$$

erklärt.

Beispiel 1.2.4. Es sei $X = \{1, 2, 7, 8\}$ und $R :=$ "$<$" auf X.
Dann ist $R^{-1} =$"$>$". Also ist

$$R = \{(x, y) \in X^2 |\ x < y\} = \{(1, 2), (1, 7), (1, 8), (2, 7), (2, 8), (7, 8)\}$$

und

$$\begin{aligned} R^{-1} &= \{(y, x) \in X^2 |\ x < y\} = \{(2, 1), (7, 1), (8, 1), (7, 2), (8, 2), (8, 7)\} \\ &= \{(y, x) \in X^2 |\ (x, y) \in R\}. \end{aligned}$$

Definition 1.2.10. Eine Relation R auf einer Menge X heißt

- reflexiv, wenn $\forall x \in X$ gilt: xRx,
- symmetrisch, wenn $\forall x, y \in X$ gilt: $xRy \Rightarrow yRx$,
- transitiv, wenn $\forall x, y, z \in X$ gilt: xRy und $yRz \Rightarrow xRz$.

Eine Relation, die zugleich reflexiv, symmetrisch und transitiv ist, heißt Äquivalenzrelation.

Definition 1.2.11. Es seien X, Y Mengen. Eine Relation $f \subseteq X \times Y$ heißt Abbildung oder Funktion von X in Y, wenn es zu jedem $x \in X$ genau ein $y \in Y$ mit xfy gibt. In diesem Fall schreiben wir auch: $y = f(x)$.
Die Menge X heißt Definitionsbereich und Y heißt Wertebereich von f. Man nennt $y = f(x) \in Y$ das Bild von $x \in X$, und x heißt Urbild von $y = f(x)$.

Definition 1.2.12. Es seien X, Y Mengen und f eine Abbildung von X in Y.

(i) Das Bild einer Teilmenge $A \subset X$ ist die Menge der Bilder aller Elemente aus A:

$$f(A) := \{y \in Y |\ \exists x \in A \text{ mit } y = f(x)\}.$$

(ii) Das Bild des gesamten Definitionsbereichs, $f(X)$ heißt Bild von f und wird mit Imf bezeichnet.

(iii) Das Urbild von $B \subset Y$ ist die Menge aller Elemente $x \in X$, deren Bilder in B liegen:

$$f^{-1}(B) := \{x \in X |\ f(x) \in B\}.$$

Definition 1.2.13. Es seien $A \subset X$ und $B \subset Y$. Eine Abbildung $g \colon A \to B$ mit $g(x) = f(x)$ und $f(x) \in B$ für alle $x \in A$ heißt eine Restriktion oder Einschränkung von f. Umgekehrt ist f eine Erweiterung von g.
Die Abbildung

$$f_{|A} \colon A \to Y \quad \text{mit} \quad f_{|A}(x) = f(x) \quad \text{für} \quad x \in A$$

nennt man die Restriktion von f auf A.

Definition 1.2.14. Eine Funktion $f \colon X \to Y$ heißt

- injektiv (eineindeutig), falls aus $f(x_1) = f(x_2)$ die Aussage $x_1 = x_2$ folgt,

- surjektiv, falls $Imf = Y$ gilt,
- bijektiv, falls f sowohl injektiv als auch surjektiv ist.

Satz 1.2.2. *Es seien X, Y Mengen und f eine Abbildung von X in Y. Die zu f inverse Relation f^{-1} ist genau dann eine Abbildung, wenn f bijektiv ist.*

Beweis. f^{-1} Abbildung $\underset{Def.1.2.11}{\Leftrightarrow}$ $\forall y \in Y\ \exists$ genau ein $x \in X$ mit $yf^{-1}x$
$\underset{Def.1.2.9}{\Leftrightarrow}$ $\forall y \in Y\ \exists$ genau ein $x \in X$ mit xfy $\underset{Def.1.2.14}{\Leftrightarrow}$ f bijektiv. □

Beispiel 1.2.5. Wir werden uns später ausschließlich mit dem Fall $X, Y \subseteq \mathbb{R}$ befassen. Funktionen werden dann z. B. so notiert:

$$f\colon [0,3] \to \mathbb{R}, \quad x \to 7x.$$

Dabei bedeutet $[0,3]$ das Intervall $[0,3] = \{x|\ 0 \le x \le 3\}$.
Die Funktion f ist injektiv wegen

$$f(x_1) = f(x_2) \ \Leftrightarrow\ 7x_1 = 7x_2 \ \Leftrightarrow\ x_1 = x_2.$$

Aber f ist nicht surjektiv aufgrund

$$0 \le x \le 3 \ \Leftrightarrow\ 0 \le 7x \le 21.$$

Es ist $Imf = [0,21]$, aber nicht $Imf = \mathbb{R}$.
Die Abbildung $g\colon [0,3] \to [0,21]$, $x \to 7x$ ist bijektiv.

Beispiel 1.2.6. Die Abbildung

$$f\colon \mathbb{R} \to \mathbb{R},\ x \to x^4$$

ist weder injektiv noch surjektiv.
Es ist $f(x) = f(-x)$, z. B. $f^{-1}(\{16\}) = \{-2, 2\}$. Außerdem gilt stets $f(x) > 0$, also ist nicht $Imf = \mathbb{R}$.
Die Restriktion $f_{|[0,\infty]}$ ist injektiv.

Definition 1.2.15. Es seien X, Y_1, Y_2, Z Mengen sowie $g\colon X \to Y_1$ und $f\colon Y_2 \to Z$ Abbildungen mit $g(X) \subset Y_2$. Dann versteht man unter der Komposition von f und g die Abbildung

$$f \circ g\colon X \to Z \quad \text{mit} \quad (f \circ g)(x) := f(g(x)).$$

Beispiel 1.2.7. Es seien

$$\begin{aligned} g\colon &\quad \mathbb{R} \to \mathbb{R} \quad \text{mit} \quad g(x) = 2x + 1 \\ f\colon &\quad \mathbb{R} \to \mathbb{R} \quad \text{mit} \quad f(y) = y^2. \end{aligned}$$

Dann ist

$$(f \circ g)(x) = f(g(x)) = f(2x+1) = (2x+1)^2.$$

Definition 1.2.16. Es sei X eine Menge. Die Abbildung $id:\ X \to X, x \to x$ heißt die Identität auf X.

Satz 1.2.3. *Es seien X, Y Mengen und f eine bijektive Abbildung von X in Y mit inverser Abbildung $f^{-1}\colon Y \to X$. Dann ist*

$$f \circ f^{-1} = f^{-1} \circ f = id.$$

Also ist $f(f^{-1}(x)) = x$ und $f^{-1}(f(x)) = x$.

Beweis. Nach Definition 1.2.9 (inverse Abbildung) ist $y = f(x) \Leftrightarrow x = f^{-1}(y)$. Also gilt

$$y = f(f^{-1}(x)) \Leftrightarrow f^{-1}(x) = f^{-1}(y) \underset{f^{-1}\ \text{bijektiv}}{\Leftrightarrow} x = y \Leftrightarrow f(f^{-1}(x)) = x.$$

□

1.3 Die reellen Zahlen

Wir kommen nun zu den Axiomen (unbewiesenen Grundtatsachen) für die reellen Zahlen. Diese Axiome lassen sich in drei Gruppen gliedern. Die erste Gruppe, die Körperaxiome, beschreiben, wie mit reellen Zahlen gerechnet wird, d.h. die Addition und die Multiplikation reeller Zahlen. Dadurch sind die reellen Zahlen aber noch lange nicht charakterisiert. Es gibt Körper, das sind Strukturen, die ebenfalls die Körperaxiome erfüllen, die völlig anders aussehen als die reellen Zahlen. Die zweite Gruppe, die Anordnungsaxiome zeigen, daß die reellen Zahlen der Größe nach verglichen werden können. Sie bilden einen angeordneten Körper. Es gibt wiederum angeordnete Körper, die andere Eigenschaften aufweisen als die reellen Zahlen, etwa die rationalen Zahlen. Die dritte Gruppe, die nur aus einem einzigen Axiom, dem Vollständigkeitsaxiom, besteht, schließt die Charakterisierung ab.

Ein Körper ist eine Menge mit zwei inneren Verknüpfungen, Addition und Multiplikation genannt. Wir sollten daher zuerst klären, was unter einer Verknüpfung zu verstehen ist.

Definition 1.3.1. Eine Verknüpfung $\circ$ auf einer Menge X ist eine Abbildung $\circ\colon X \times X \to X$. Das Bild eines Paares (x, y) wird das Ergebnis der Verknüpfung genannt und mit $x \circ y$ bezeichnet. Verknüpfungen können mit jedem beliebigen Symbol (neben $\circ$) bezeichnet werden. Ist das Symbol $+$ (Pluszeichen), so heißt die Verknüpfung Addition, und das Ergebnis $x + y$ wird als Summe von x und y bezeichnet. Ist das Symbol $\cdot$ (Multiplikationspunkt), so heißt diese Multiplikation, und dementsprechend wird das Ergebnis $x \cdot y$ als Produkt von x und y bezeichnet.

Wichtige Strukturen mit einer Verknüpfung sind die Gruppen:

Definition 1.3.2. Eine Gruppe ist ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer Verknüpfung $\circ$ auf G, so daß gilt:

(G1) $a \circ (b \circ c) = (a \circ b) \circ c$ für alle $a, b, c \in G$ (Assoziativgesetz).

(G2) Es existiert ein Element e, neutrales Element (oder Einselement) genannt, so daß $e \circ a = a$ für alle $a \in G$.

(G3) Ist ein neutrales Element $e \in G$ gegeben, so gibt es zu jedem $a \in G$ ein inverses Element $a^{-1} \in G$, so daß $a^{-1} \circ a = e$.

Gilt zusätzlich noch, daß das Ergebnis der Verknüpfung von der Reihenfolge der Faktoren unabhängig ist, so heißt G eine kommutative oder abelsche Gruppe (nach N. H. Abel). Es gilt also

(G4) $a \circ b = b \circ a$ für alle $a, b \in G$ (Kommutativgesetz).

Bemerkung 1.3.1. Wird die Verknüpfung als Addition geschrieben, so wird das neutrale Element auch als Null bezeichnet (Schreibweise: 0). Das inverse Element eines Elements a wird dann mit $-a$ bezeichnet und heißt das Negative von a. Wird die Verknüpfung als Multiplikation geschrieben, so wird das neutrale Element auch als Eins bezeichnet (Schreibweise: 1).
Ein Körper ist ein Paar $((K, +), \cdot)$, wobei $(K, +)$ eine abelsche Gruppe, deren Verknüpfung $+$ Addition genannt wird, mit neutralem Element 0 ist, während $\cdot$ eine weitere Verknüpfung auf K ist, Multiplikation genannt, so daß $(K\backslash\{0\}, \cdot\}$ eine abelsche Gruppe ist.
Addition und Multiplikation sind durch das Distributivgesetz verbunden:

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

Die erste Gruppe von Axiomen für die Menge $\mathbb{R}$ der reellen Zahlen sind also folgende Körperaxiome:

Auf $\mathbb{R}$ existieren zwei Verknüpfungen $+$ (Addition) und $\cdot$ (Multiplikation), sowie Elemente 0 (Null) und 1 (Eins) mit $0 \neq 1$, so daß folgende Gesetze gelten:

(K1) $(a+b)+c = a+(b+c)$ und $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. (Assoziativgesetz).

(K2) $a+b = b+a$ und $a \cdot b = b \cdot a$. (Kommutativgesetz).

(K3) $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$ (Distributivgesetz).

(K4) $a+0 = a$ und $a \cdot 1 = a$ (Existenz neutraler Elemente).

(K5) Für alle $a \in \mathbb{R} \exists (-a) \in \mathbb{R}$ mit $a + (-a) = 0$ und
für alle $a \in \mathbb{R} \backslash \{0\} \ \exists \ a^{-1}$ mit $a \cdot a^{-1} = 1$ (Existenz inverser Elemente).

Definition 1.3.3. Das multiplikative Inverse a^{-1} wird auch Reziprokes oder Kehrwert von a genannt und "a hoch minus eins" gelesen.
Wir lassen künftig die Multiplikationspunkte weg und folgen der Regel "Punkt vor Strich", d.h. wenn keine Klammern das Gegenteil besagen, wird die Multiplikation vor der Addition ausgeführt, so kann z. B. das Distributivgesetz als $a(b+c) = ab + ac$ formuliert werden.

Die zweite Gruppe von Axiomen sind folgende Anordnungsaxiome:

Es gibt eine Relation "$<$" (sprich: kleiner) auf $\mathbb{R}$ zu $\mathbb{R}$, so daß folgende Gesetze gelten:

(A1) Für $a, b \in \mathbb{R}$ gilt stets eine und nur eine der Beziehungen $a < b$, $a = b$ und $b < a$. (Trichotomiegesetz).

(A2) $a < b$ und $b < c \Rightarrow a < c$. (Transitivitätsgesetz).

(A3) Ist $a < b$, so gilt $a + c < b + c \ \forall \ c \in \mathbb{R}$ und $ac < bc \ \forall \ c \in \mathbb{R}$ mit $0 < c$. (Monotoniegesetze).

Definition 1.3.4. Die zu der Relation "$<$" inverse Relation ist die Relation "$>$" (größer). Die Zeichen "$\leq$" bzw. "$\geq$" bedeuten kleiner oder gleich bzw. größer oder gleich.

Bemerkung 1.3.2. Jeder Körper, der außer den Körperaxiomen $(K1) - (K5)$ auch noch die Anordnungsaxiome $(A1) - (A3)$ erfüllt, heißt angeordneter Körper.

Zur Formulierung des letzten Axioms, des Vollständigkeitsaxioms, brauchen wir zunächst folgende

Definition 1.3.5. Es sei $X \subset \mathbb{R}$ und $s \in \mathbb{R}$. Dann heißt s eine obere Schranke von X, falls $x \leq s$ für alle $x \in X$ gilt. Falls eine obere Schranke von X existiert, heißt $X \subset \mathbb{R}$ nach oben beschränkt. Man nennt s kleinste obere Schranke oder Supremum von X, falls s eine obere Schranke von X ist und für alle oberen Schranken t von X gilt, daß $s \leq t$ ist.

Nun können wir noch unser letztes Axiom formulieren:

(V) Jede nichtleere nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum (Vollständigkeitsaxiom, Supremumsaxiom).

1.4 Ungleichungen, Rechenregeln, Betrag

Wir zeigen nun einige unmittelbare Folgerungen aus den Axiomen für die reellen Zahlen. Es werden sich vertraute Regeln über die Umformung von Ungleichungen und Regeln über das "Bruchrechnen" ergeben.

Satz 1.4.1. *(Kürzungsregel)*
Es seien $a, b, c \in \mathbb{R}$. Dann gilt

(i) $a + c = b + c \Rightarrow a = b$.

(ii) *Ist $c \neq 0$, so gilt $ac = bc \Rightarrow a = b$.*

Beweis. (i) $a+c = b+c \underset{(K5)}{\Rightarrow} (a+c)+(-c) = (b+c)+(-c) \underset{(K1)}{\Rightarrow} a+(c+(-c)) = b+(c+(-c)) \Rightarrow a+0 = b+0 \underset{(K4)}{\Rightarrow} a = b$.

(ii) $ac = bc \underset{(K5)}{\Rightarrow} (ac)c^{-1} = (bc)c^{-1} \underset{(K1)}{\Rightarrow} a(c \cdot c^{-1}) = b(c \cdot c^{-1}) \Rightarrow a \cdot 1 = b \cdot 1 \underset{(K4)}{\Rightarrow} a = b$.

□

Bevor wir fortfahren, müssen wir einen Punkt klären, der leicht zu übersehen ist. In den Axiomen $(K4)$ und $(K5)$ wurde zwar die Existenz der neutralen Elemente (0 und 1) sowie die Existenz der Inversen gefordert. Es wurde jedoch nicht die Eindeutigkeit dieser Objekte gefordert. Dies ergibt sich jedoch als Folgerung.

Satz 1.4.2. *(Eindeutigkeit der neutralen Elemente und der Inversen)*

(i) (Eindeutigkeit der Null):

$$a + 0' = a \quad \textit{für ein} \quad a \in \mathbb{R} \quad \Rightarrow \quad 0' = 0.$$

(ii) (Eindeutigkeit der Eins):

$$a \cdot 1' = a \quad \textit{für ein} \quad a \in \mathbb{R} \quad \Rightarrow \quad 1' = 1.$$

(iii) (Eindeutigkeit der Negativen):

$$\forall a \in \mathbb{R}: \quad a + x = 0 \quad \Rightarrow x = -a.$$

(iv) (Eindeutigkeit des Kehrwerts):

$$\forall a \in \mathbb{R}\backslash\{0\}: \quad a \cdot x = 1 \quad \Rightarrow x = a^{-1}.$$

Beweis. Die Beweise von Satz 1.4.2 (i), (ii) ergeben sich aus der Kürzungsregel (Satz 1.4.1):

(i) $a + 0' = a = a + 0 \underset{S.1.4.1(i)}{\Rightarrow} 0' = 0$.

(ii) $a \cdot 1' = a = a \cdot 1 \underset{S.1.4.1(ii)}{\Rightarrow} 1' = 1$

(iii) $a + x = 0 \underset{(K5)}{\Rightarrow} (-a) + (a + x) = 0 + (-a) \underset{(K1)}{\Rightarrow} (-a + a) + x = 0 + (-a) \underset{(K4),(K5)}{\Rightarrow} x = -a$.

(iv) Beweisskizze: Multiplikation mit a^{-1}. Sonst analog zu (iii).

□

Satz 1.4.3. *Für $a, b \in \mathbb{R}$ gilt:*

(i) $a \cdot 0 = 0 \cdot a = 0$

(ii) $(-a) \cdot b = a \cdot (-b) = -(ab)$

(iii) $(-a) \cdot (-b) = ab$

(iv) $a \cdot (b - c) = (b - c) \cdot a = ab - ac$

(v) $-(-a) = a$.

Beweis. Es ist

(i) $a \cdot 0 \underset{(K4)}{=} 0 + (a \cdot 0) = a \cdot (0 + 0) \underset{(K3)}{=} (a \cdot 0) + (a \cdot 0)$, also $0 + (a \cdot 0) = (a \cdot 0) + (a \cdot 0) \underset{S.1.4.1(i)}{\Leftrightarrow} a \cdot 0 = 0$.

(ii) $0 \underset{(i)}{=} 0 \cdot b = (a + (-a)) \cdot b = ab + (-a) \cdot b$. Aus der Eindeutigkeit des Negativen (Satz 1.4.2 (ii)) folgt: $(-a) \cdot b = -(ab)$. Es folgt $a \cdot (-b) = -(ab)$ wegen $a \cdot (-b) = (-b) \cdot a$ durch "Umbenennen" daraus.

(iii) $0 \underset{(i)}{=} 0 \cdot (-b) \underset{(K5)}{=} (a + (-a))(-b) \underset{(K3)}{=} a \cdot (-b) + (-a)(-b) \underset{(ii)}{=} -(ab) + (-a)(-b) \underset{S.1.4.2(ii)}{\Rightarrow} (-a)(-b) = ab$.

(iv) ohne Beweis

(v) ohne Beweis

□

Bemerkung 1.4.1. Wir vereinbaren auch für das Minuszeichen eine Punkt- vor- Strich- Regel:

$$ab - cd := (a \cdot b) - (c \cdot d).$$

Unter $a - b$ verstehen wir $a + (-b)$.

Definition 1.4.1. Es sei $a \in \mathbb{R}$ und $b \in \mathbb{R} \backslash \{0\}$. Wir schreiben

$$\frac{a}{b} := ab^{-1}.$$

Insbesondere ist $\frac{1}{b} := b^{-1}$.

Satz 1.4.4. *(Bruchrechnen) Es seien $a, b, c, d \in \mathbb{R}$, $b, d \neq 0$. Dann gilt:*

(i) $\dfrac{a}{b} = \dfrac{ad}{bd}$

(ii) $\dfrac{1}{(-b)} = -\dfrac{1}{b}$

(iii) $\dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad \pm bc}{bd}$

(iv) $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{ac}{bd}$

(v) $\left(\dfrac{a}{b}\right)^{-1} = \dfrac{b}{a}$, *falls auch* $a \neq 0$.

Beweis. Übungen $\square$

Definition 1.4.2. Man nennt $a \in \mathbb{R}$

- positiv, wenn $a > 0$ und
- negativ, wenn $a < 0$.

Satz 1.4.5. *Es seien* $a, b \in \mathbb{R}$. *Dann gilt*

(i) $0 < 1$

(ii) $a < b \Leftrightarrow b - a > 0$

(iii) $a < 0 \Leftrightarrow -a > 0$

(iv) $a < b \Leftrightarrow -b < -a$

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

Beweis. Wir beweisen nur (ii): $b - a > 0 \underset{(A3)}{\Leftrightarrow} (b - a) + a > 0 + a \Leftrightarrow b > a$. $\square$

Satz 1.4.6. *(Addition von gleichsinnigen Ungleichungen)*
Aus $a < b$ *und* $c < d$ *folgt* $a + c < b + d$.

Beweis. Aus der Monotonie $(A3)$ folgt $a + c < b + c$ und $b + c < b + d$. Mit der Transitivität $(A2)$ gilt dann $a + c < b + d$. $\square$

Satz 1.4.7. *(Durchmultiplikation von Ungleichungen)*
Es seien $a, b, c \in \mathbb{R}$, *und es sei* $a < b$.

(i) Ist $c > 0$, *so folgt* $ac < bc$.

(ii) Ist $c < 0$, *so folgt* $bc < ac$.

Beweis. (i) ist Axiom $(A3)$ (Monotoniegesetz)

(ii) Nach Satz 1.4.5 (ii) ist $-c > 0$. Dann gilt

$$\begin{aligned} a < b \quad & \underset{S.1.4.5(ii)}{\Leftrightarrow} \quad b - a > 0 \underset{\substack{-c>0 \\ (A3)}}{\Leftrightarrow} (b-a)(-c) > 0 \Leftrightarrow (a-b) \cdot c > 0 \\ & \underset{S.1.4.3(iv)}{\Leftrightarrow} \quad ac - bc > 0 \underset{S.1.4.5(ii)}{\Leftrightarrow} bc < ac. \end{aligned}$$

$\square$

Satz 1.4.8. *Es seien $a, b \in \mathbb{R}$ Dann folgt*

$$ab > 0 \quad \Leftrightarrow \quad (a > 0 \text{ und } b > 0) \quad \underline{\text{oder}} \quad (a < 0 \text{ und } b < 0).$$

Beweis. ohne Beweis □

Satz 1.4.9. *Es seien $a, b \in \mathbb{R}$ und $b \neq 0$. Dann gilt*

$$\frac{a}{b} > 0 \quad \Leftrightarrow \quad (a > 0 \text{ und } b > 0) \quad \underline{\text{oder}} \quad (a < 0 \text{ und } b < 0).$$

Beweis. Wir zeigen zunächst

(*) $b > 0 \Rightarrow \frac{1}{b} > 0$.

Nach $(A1)$ (Trichotomie) gilt genau einer der Fälle

$$\frac{1}{b} < 0, \quad \frac{1}{b} = 0, \quad \frac{1}{b} > 0.$$

Annahme:

- $\frac{1}{b} = 0 \Rightarrow 1 = b \cdot \frac{1}{b} = 0$ im Widerspruch zu $0 \neq 1$.
- $\frac{1}{b} < 0 \underset{(A3)}{\Rightarrow} b \cdot \frac{1}{b} < 0 \cdot b \underset{S.1.4.3(i)}{\Rightarrow} 1 < 0$ im Widerspruch zu Satz 1.4.5.

Also gilt $(*)$.
Wir kommen nun zum Beweis der Behauptung:
"$\Leftarrow$:"

Fall 1:

$$a > 0,\ b > 0 \underset{(*)}{\Rightarrow} a > 0,\ \frac{1}{b} > 0 \underset{(A3)}{\Rightarrow} \frac{a}{b} > 0.$$

Fall 2:

$$\begin{aligned} a < 0,\ b < 0 \quad &\underset{S.1.4.5(iii)}{\Rightarrow} \quad -a > 0,\ -b > 0 \underset{(*)}{\Rightarrow} -a > 0,\ -\frac{1}{b} > 0 \\ &\underset{(A3)}{\Rightarrow} \quad \frac{-a}{-b} > 0 \underset{S.1.4.4(ii)}{\Rightarrow} \frac{a}{b} > 0. \end{aligned}$$

"$\Rightarrow$:"
Zunächst folgt $a \neq 0$, denn $a = 0 \underset{S.1.4.3(i)}{\Rightarrow} a \cdot \frac{1}{b} = 0$.
Annahme:
Die Behauptung ($a > 0$ und $b > 0$) oder ($a < 0$ und $b < 0$) ist falsch.
Dann muß wegen $(A1)$ (Trichotomie) einer der folgenden Fälle zutreffen:

Fall 1: $a > 0$ und $b < 0$

Fall 2: $a < 0$ und $b > 0$

Diese beiden Fälle schließen wir der Reihe nach aus.

Fall 1: $a > 0,\ b < 0 \Rightarrow -a < 0,\ b < 0$. Nach der schon bewiesenen Richtung "$\Leftarrow$:" folgt dann $-\frac{a}{b} > 0 \Rightarrow \frac{a}{b} < 0$ im Widerspruch zur Voraussetzung.

Fall 2: $a < 0$ und $b > 0 \Rightarrow -a > 0,\ -b < 0 \underset{''\Leftarrow:''}{\Rightarrow} \frac{-a}{-b} < 0 \underset{S.1.4.4(iv)}{\Rightarrow} \frac{a}{b} < 0$, ebenfalls ein Widerspruch.

□

Definition 1.4.3. (Vorzeichen, Absolutbetrag)
Es sei $a \in \mathbb{R}$. Dann heißt

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

das Vorzeichen oder Signum von a.
Der Absolutbetrag (kurz: Betrag, Schreibweise: $|a|$) von $a \in \mathbb{R}$ ist durch

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

erklärt.

Satz 1.4.10. *(Eigenschaften des Absolutbetrages)*
Für alle $a, b \in \mathbb{R}$ ist

(i) $|a| \geq 0$ *und* $|a| = 0 \Leftrightarrow a = 0$ *(Definitheit)*

(ii) $|a \cdot b| = |a| \cdot |b|$ *(Multiplikativität)*

(iii) $|a + b| \leq |a| + |b|$ *(Dreiecksungleichung)*

Beweis. Wir beweisen nur (iii):
Es sei $\epsilon := \operatorname{sgn}(a + b)$. Wegen $a \leq |a|$ und $|\pm 1| = 1$ folgt

$$|a + b| = \epsilon(a + b) = \epsilon a + \epsilon b \leq |\epsilon a| + |\epsilon b| = |a| + |b|.$$

□

Satz 1.4.11. *Für alle $a, b, c \in \mathbb{R}$ gelten die folgenden Eigenschaften:*

(i) $a \leq |a|$

(ii) *Für* $a \neq 0$ *gilt* $a^2 = (-a)^2 = |a|^2 > 0$.

(iii) *Für* $a \neq 0$ *gilt* $|a^{-1}| = |a|^{-1}$.

(iv) $||a| - |b|| \leq |a - b|$.

Beweis. ohne Beweis □

Wichtige Teilmengen von $\mathbb{R}$ sind die Intervalle.

Definition 1.4.4. (Intervalle)
Es seien $a \leq b \in \mathbb{R}$. Die folgenden Mengen heißen Intervalle:

$$[a,b] := \{x \in \mathbb{R} |\ a \leq x \leq b\}$$

und falls $a < b$

$$\begin{aligned} [a,b) &:= \{x \in \mathbb{R} |\ a \leq x < b\} \\ (a,b] &:= \{x \in \mathbb{R} |\ a < x \leq b\} \\ (a,b) &:= \{x \in \mathbb{R} |\ a < x < b\} \end{aligned}$$

Dann heißt $[a,b]$ abgeschlossenes, auch kompaktes Intervall, (a,b) heißt offen und $[a,b)$ bzw. $(a,b]$ heißen halboffen. Die Eigenschaften abgeschlossen, kompakt und offen von Teilmengen von $\mathbb{R}$ werden später allgemein definiert werden.

Definition 1.4.5. (unendliche Intervalle)
Diese Intervalle werden mittels der Symbole ∞ und $-\infty$ (sprich: unendlich und minus unendlich) definiert. Es seien $a, b \in \mathbb{R}$. Wir definieren:

$$\begin{aligned} [a,\infty) &:= \{x \in \mathbb{R} |\ a \leq x\} \\ (a,\infty) &:= \{x \in \mathbb{R} |\ a < x\} \\ (-\infty,b] &:= \{x \in \mathbb{R} |\ x \leq b\} \\ (-\infty,b) &:= \{x \in \mathbb{R} |\ x < b\} \\ (-\infty,\infty) &:= \mathbb{R}. \end{aligned}$$

Definition 1.4.6. (Maximum, Minimum)
Es sei $X \subset \mathbb{R}$. Dann heißt m Minimum von X (Schreibweise: $\min X$), falls $m \in X$ und $m \leq x$ für alle $x \in X$ gilt, und M heißt Maximum von X (Schreibweise: $\max X$), falls $M \in X$ und $x \leq M$ für alle $x \in X$ ist.

Satz 1.4.12. *Für alle $a, b, c \in \mathbb{R}$ gilt*

(i) $|a| < c \Leftrightarrow a \in (-c, c)$

(ii) $|b - a| < c \Leftrightarrow b \in (a - c, a + c)$

Beweis. ohne Beweis □

1.5 Natürliche Zahlen, vollständige Induktion

Nun können wir die natürlichen Zahlen einführen, die wir bisher noch nicht kennen.

Definition 1.5.1. (i) Eine induktive Menge I reeller Zahlen ist eine Menge mit folgenden Eigenschaften:

(1) $1 \in I$

(2) $n \in I \Rightarrow n + 1 \in I$

(ii) Die Menge der natürlichen Zahlen ist der Durchschnitt aller induktiven Teilmengen von $\mathbb{R}$. Wir bezeichnen sie als $\mathbb{N}$.

Satz 1.5.1. *Es gilt:*

(i) $\mathbb{N} \neq \emptyset$, $\mathbb{N}$ *ist induktiv*

(ii) *Jede induktive Teilmenge von* $\mathbb{N}$ *ist* $\mathbb{N}$.

(iii) $\min \mathbb{N} = 1$

(iv) *Für* $n \in \mathbb{N}$ *gibt es kein* $m \in \mathbb{N}$ *mit* $n < m < n + 1$.

(v) $n, m \in \mathbb{N} \Rightarrow n + m \in \mathbb{N}$ *und* $n \cdot m \in \mathbb{N}$

(vi) *Es seien* $n \in \mathbb{N}$ *und* $m \in [1, \infty)$. *Dann ist* $n - m \in \mathbb{N} \Leftrightarrow m \in \mathbb{N}$ *und* $m < n$.

Zur Vorbereitung des Beweises zeigen wir zunächst:

Lemma 1.5.1. *Der Durchschnitt einer beliebigen Menge induktiver Teilmengen von* $\mathbb{R}$ *ist induktiv.*

Beweis. Es sei J eine Menge von induktiven Teilmengen von $\mathbb{R}$ und $M = \bigcap_{I \in J} I$. Dann ist

(1) $1 \in I$ für alle $I \in J \underset{Def.1.2.3(ii)}{\Rightarrow} 1 \in \bigcap_{I \in J} I = M$.

(2) $n \in M \underset{Def.1.2.3(ii)}{\Rightarrow} n \in I,\ \forall I \in J \underset{I\text{induktiv}}{\Rightarrow} n + 1 \in I,\ \forall I \in J \underset{Def.1.2.3(ii)}{\Rightarrow} n + 1 \in M$.

□

Beweis. (Beweis von Satz 1.5.1) Es sei J die Menge aller induktiven Teilmengen von $\mathbb{R}$.

(i) Nach Lemma 1.5.1 ist $\mathbb{N}$ induktiv und damit $1 \in \mathbb{N}$, also $\mathbb{N} \neq \emptyset$.

(ii) Es sei M eine induktive Teilmenge von $\mathbb{N}$. Dann gilt:

  (1) $M \subset \mathbb{N}$. Nun ist $\mathbb{N} = \bigcap_{I \in J} I$, also $\mathbb{N} \subset I,\ \forall I \in J$. Da $M \subset J$ ist, folgt

  (2) $\mathbb{N} \subset M$.

  Also ist $\mathbb{N} = M$.

(iii) Wegen $1 \in [1, \infty)$ und $n \geq 1 \underset{(A3)}{\Rightarrow} n + 1 \geq 1$ ist $[1, \infty)$ induktiv. Also ist $\mathbb{N} \subset [1, \infty) \Rightarrow n \geq 1$ für alle $n \in \mathbb{N}$.

(iv) Es sei
$$K := \{n \in \mathbb{N} |\ \nexists m \in \mathbb{N} \quad \text{mit} \quad n < m < n + 1\}.$$
Wir zeigen, daß K induktiv ist:
Es sei $M_1 := \mathbb{N} \backslash (1, 1 + 1)$. Dann ist M_1 induktiv, also $M_1 = \mathbb{N}$. Es ist $\mathbb{N} \cap (1, 1 + 1) = \emptyset$. Damit ist $1 \in K$.
Nun sei $n \in K$:
Setze $M_{n+1} := \mathbb{N} \backslash (n + 1, n + 1 + 1)$. Wiederum ist M_{n+1} induktiv, da $1 \in M_{n+1}$ ist und aus $m \in M_{n+1}$ folgt, da $n \in K$ ist, $m \notin (n, n + 1)$ ist. Also ist $m + 1 \notin (n + 1, n + 1 + 1)$ und folglich $m + 1 \in \mathbb{N} \backslash (n + 1, n + 1 + 1) = M_{n+1}$.
Damit ist $M_{n+1} = \mathbb{N}$. Also ist $\mathbb{N} \cap (n + 1, n + 1 + 1) = \emptyset$ und $n + 1 \in K$. Also ist $K = \mathbb{N}$.

(v) Für $n \in \mathbb{N}$ betrachten wir die Menge $L_n := \{m \in \mathbb{N} |\ n + m \in \mathbb{N}\}$. Diese Menge ist induktiv, denn $n \in L_n \underset{\mathbb{N} \in J}{\Rightarrow} n+1 \in \mathbb{N}$. Also ist $1 \in L_n$. Es sei $m \in L_n$, also $n+m \in \mathbb{N} \underset{\mathbb{N} \in J}{\Rightarrow} (n+m)+1 \underset{(K1)}{=} n + (m+1) \in L_n$. Also ist $m + 1 \in L_n$. Damit ist $L_n = \mathbb{N}$.

Wir verzichten auf den Beweis des zweiten Teils und auf den Beweis von Teil (vi). $\square$

Bemerkung 1.5.1.
Das Beweisverfahren in Satz 1.5.1, Teil (iv) und (v) ist der Beweis durch vollständige Induktion: Es ist eine Aussage $A(n)$ zu beweisen, die von der natürlichen Zahl n abhängt. Eigentlich handelt es sich um eine Folge von Aussagen $A(n)$, wie zum Beispiel in (iv) die Aussage $A(n)$: Es gibt kein $m \in \mathbb{N}$ mit $m \in (n, n + 1)$. Man definiert die Menge W aller n, für die die Aussage $A(n)$ wahr ist und zeigt, daß diese Menge induktiv ist, d.h.

(i) $1 \in W$: $A(1)$ ist wahr

(ii) $n \in W \Rightarrow (n + 1) \in W$: Wenn $A(n)$ wahr ist, dann ist auch $A(n + 1)$ wahr.

Nach Satz 1.5.1 (ii) folgt dann: $W = \mathbb{N}$, d.h. $A(n)$ ist für alle $n \in \mathbb{N}$ wahr.
Ein Beweis durch vollständige Induktion geht also nach folgendem Schema:

(i) $n = 1$: (Induktionsanfang): $A(1)$ ist wahr.

(ii) $n \to n + 1$ (Induktionsschritt): Wenn $A(n)$ wahr ist (Induktionshypothese), dann ist auch $A(n + 1)$ wahr.

Definition 1.5.2. (i) Mengen X und Y sind gleichmächtig (Schreibweise: $X \sim Y$), falls es eine injektive Abbildung $f : X \to Y$ und $g : Y \to X$ gibt.

(ii) Die Menge X heißt von kleinerer Mächtigkeit als die Menge Y (Schreibweise $X \prec Y$), falls es eine injektive Abbildung $f : X \to Y$, aber keine injektive Abbildung $g : Y \to X$ gibt.

(iii) Für $n \in \mathbb{N}$ sei der Abschnitt bis n (Schreibweise: $\mathbb{N}(n)$) durch $\mathbb{N}(n) := \{m \in \mathbb{N} |\ m \le n\}$ definiert.

Eine Menge X heißt endlich, falls es ein $n \in \mathbb{N}$ gibt, so daß $X \sim \mathbb{N}(n)$, andernfalls unendlich. Man nennt X abzählbar unendlich, falls $X \sim \mathbb{N}$ und abzählbar, falls X endlich oder abzählbar unendlich ist, andernfalls überabzählbar.

Satz 1.5.2. *Es seien X, Y, Z Mengen.*

(i) $X \sim Y \Leftrightarrow \exists f : X \to Y,\ f$ *bijektiv*

(ii) $X \prec Y$ *und* $Y \prec Z \Rightarrow X \prec Z$

(iii) $X \neq \emptyset$ *erfüllt genau eine der folgenden Möglichkeiten: X ist endlich, X ist abzählbar unenendlich oder X ist überabzählbar.*

(iv) Ist X endlich, so gibt es genau ein $n \in \mathbb{N}$ mit $X \sim \mathbb{N}(n)$.

Beweis. Übungen $\square$

Definition 1.5.3.

$$2 := 1 + 1,\ 3 := 2 + 1,\ 4 := 3 + 1, \ldots, 9 := 8 + 1,\ 10 := 9 + 1, \ldots$$

Definition 1.5.4. Es sei X eine endliche Menge. Ist $n \in \mathbb{N}$ die nach Satz 1.5.2 eindeutig bestimmte natürliche Zahl mit $X \sim \mathbb{N}(n)$, so schreiben wir $|X| = n$ und sagen, X hat $\underline{n \text{ Elemente}}$.

Beispiel 1.5.1. Es sei $X = \{a, b, c\}$. Für $n \geq 3$ haben wir eine bijektive Abbildung $f\colon X \to \mathbb{N}(3) = \{1, 2, 3\}$ mit $f(a) := 1$, $f(b) := 2$ und $f(c) := 3$. Also ist $|X| = 3$, d.h. X hat drei Elemente.

Satz 1.5.3. *(i) Jede nichtleere Teilmenge von $\mathbb{N}$ besitzt ein Minimum (Wohlordnungssatz).*

(ii) Jede endliche Teilmenge von $\mathbb{R}$ besitzt Maximum und Minimum.

Beweis. ohne Beweis. □

Definition 1.5.5. Wir setzen $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
Die Mengen $\mathbb{Z}$ der ganzen Zahlen und $\mathbb{Q}$ der rationalen Zahlen sind durch $\mathbb{Z} := \mathbb{N}_0 \cup \{-n | \ n \in \mathbb{N}\}$ und $\mathbb{Q} = \left\{\frac{r}{s} | \ r \in \mathbb{Z}, s \in \mathbb{Z}\backslash\{0\}\right\}$ definiert.

Man zeigt leicht:

Satz 1.5.4. *(i) $(\mathbb{Z}, +)$ ist eine kommutative Gruppe.*

(ii) $(\mathbb{Q}, +, \cdot)$ ist ein Körper.

Beweis. Durch Nachrechnen. □

Definition 1.5.6. Unter einer <u>Folge</u> a_n von Elementen einer Menge X verstehen wir eine Abbildung $f\colon \mathbb{N} \to X$, $n \to a_n$ oder $f\colon \mathbb{N}_0 \to X$, $n \to a_n$. Dabei heißt a_n auch das n- te <u>Glied der Folge</u> und n heißt <u>Folgenindex</u> (kurz: Index). Statt n kann auch jedes andere Symbol verwendet werden. Unter einer <u>endlichen Folge</u> verstehen wir eine Abbildung $f\colon \mathbb{N}(n) \to X$. Wir betrachten meist den Fall $X = \mathbb{R}$, also Folgen reeller Zahlen. Künfig soll unter einer Folge stets eine Folge reeller Zahlen gemeint sein, falls nichts anderes vereinbart ist. Eine Folge kann durch eine Gleichung, z. B. $a_n = 5n + 2$, beschrieben werden. Sie kann auch durch <u>vollständige Induktion</u> definiert werden.

Beispiel 1.5.2. (Folge der Fibonacci- Zahlen)
Es sei (a_n) durch

$$a_0 = 1,\ a_1 = 1,\ a_n = a_{n-1} + a_{n-2} \quad \text{für} \quad n \geq 2$$

gegeben. Die ersten Glieder der Folge berechnen sich zu

$$\begin{aligned} a_0 &= 1 \\ a_1 &= 1 \\ a_2 &= a_1 + a_0 = 1 + 1 = 2 \\ a_3 &= a_2 + a_1 = 2 + 1 = 3 \\ a_4 &= a_3 + a_2 = 3 + 2 = 5 \end{aligned}$$

Summen und Produkte einer beliebigen Anzahl von Gliedern einer Folge können induktiv definiert werden.

Definition 1.5.7. Es sei $(a_n)\colon \mathbb{N} \to \mathbb{R}$ eine Folge. Dann definieren wir die Folge $\left(\sum_{k=1}^{n} a_k\right)$ der <u>n- ten Partialsummen</u> induktiv durch

(i) $n = 1$:

$$\sum_{k=1}^{1} a_k := a_1$$

(ii) $n \to n+1$:

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

In $\sum\limits_{k=1}^{n} a_k$ heißt k die Summationsvariable, 1 (bzw. n) die untere (bzw. obere) Summationsgrenze, und $\{m \in \mathbb{N}|\ 1 \leq m \leq n\}$ heißt auch Summationsintervall. Die Sprechweise ist: Summe $k = 1$ bis n, a_k. Für die Summationsvariable kann auch jedes andere Symbol benützt werden. Allgemeiner können auch Summen $\sum\limits_{k=0}^{n} a_k$, falls $(a_n)\colon \mathbb{N}_0 \to \mathbb{R}$, oder $\sum\limits_{k=m}^{n} a_k$ induktiv definiert werden. Summen können auch für endliche Folgen definiert werden. Wir verzichten auf die Einzelheiten der Definition.

Durch vollständige Induktion beweist man leicht:

Satz 1.5.5. *Es seien $a, a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$. Dann gilt:*

(i) $\sum\limits_{k=1}^{n} a_k + \sum\limits_{k=1}^{n} b_k = \sum\limits_{k=1}^{n} (a_k + b_k)$

(ii) $a \sum\limits_{k=1}^{n} a_k = \sum\limits_{k=1}^{n} a a_k$

(iii) $\left|\sum\limits_{k=1}^{n} a_k\right| \leq \sum\limits_{k=1}^{n} |a_k|$ *(Dreiecksungleichung)*

Beweis. ohne Beweis. □

Definition 1.5.8. Es sei $(a_n)\colon \mathbb{N} \to \mathbb{R}$ eine Folge. Dann definieren wir die Folge $\left(\prod\limits_{k=1}^{n} a_k\right)$ der n- ten Partialprodukte induktiv durch

(i) $n = 1$:

$$\prod_{k=1}^{1} a_k := a_1$$

(ii) $n \to n+1$:

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

Es gelten die entsprechenden Bemerkungen wie bei der Definition der Partialsummen.

In Ausdrücken der Form $\sum\limits_{k=1}^{n} a$ bzw. $\prod\limits_{k=1}^{n} a$, bei denen der Folgenindex fehlt, werden die Partialsummen bzw. Partialprodukte von der konstanten Folge (a_k) mit $a_k = a$ für alle k gebildet.
Der Fall der Partialprodukte führt zur Definition der Potenzen:

Definition 1.5.9. (n- te Potenzen)
Es sei $a \in \mathbb{R}$ und $n \in \mathbb{N}$. Dann ist die <u>n- te Potenz</u> a^n (lies: a hoch n) durch $\prod_{k=1}^{n} a$ definiert. Es ist $a^0 = 1$ und für $a \neq 0$ sei $a^{-n} := \frac{1}{a^n}$.

Satz 1.5.6. *(Potenzgesetze)*
Es seien $a, b \in \mathbb{R}$ und $m, n \in \mathbb{Z}$. Dann gelten folgende Regeln (vorausgesetzt die Ausdrücke sind definiert):

(i) $(ab)^n = a^n \cdot b^n$

(ii) $a^{m+n} = a^m \cdot a^n$

(iii) $(a^m)^n = a^{mn}$

Beweis. ohne Beweis. □

Satz 1.5.7. *(Monotonie, Definitheit)*
Es seien $a < b \in \mathbb{R}^+$ und $n \in \mathbb{Z}$. Dann gilt

(i) $a^n < b^n$, *falls* $n > 0$

(ii) $a^n > b^n$, *falls* $n < 0$

(iii) $a^2 \geq 0$ *und* $a^2 = 0 \Leftrightarrow a = 0$.

Beweis. ohne Beweis. □

Definition 1.5.10. (Fakultät und Binomialkoeffizient)
Für $n \in \mathbb{N}$ ist $n!$ (lies: <u>n Fakultät</u>) durch

$$n! := \prod_{k=1}^{n} k$$

definiert. Weiter ist $0! := 1$.
Der Binomialkoeffizient $\binom{n}{k}$ (lies: n über k) ist für $k, n \in \mathbb{N}_0$ und $0 \leq k \leq n$ durch

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

definiert.

Eine wichtige Technik im Rechnen mit Summenzeichen ist die <u>Indexverschiebung</u>. Bevor wir diese Technik in einem Satz formulieren, geben wir zwei Beispiele:

Beispiel 1.5.3. Es ist

$$\sum_{m=2}^{n} (m-2) = (2-2) + \ldots + (n-2).$$

Wir führen eine neue Summationsvariable ein: $k = m - 2$. Diese <u>Substitution</u> ist mit einer bijektiven Abbildung f des ursprünglichen Summationsintervalls $\{m \in \mathbb{N} |\ 2 \leq m \leq n\}$ auf das neue Summationsintervall $\{n \in \mathbb{N} |\ 0 \leq k \leq n-2\}$ verbunden. Damit ist

$$\sum_{m=2}^{n} (m-2) = (2-2) + \ldots + (n-2) = 0 + \ldots + (n-2) = \sum_{k=0}^{n-2} k.$$

Eine der bekanntesten Anwendungen ist die Summenformel für die endliche geometrische Reihe.

Beispiel 1.5.4. Es sei $q \in \mathbb{R}\backslash\{1\}$ und $N \in \mathbb{N}$. Wir betrachten die endliche geometrische Reihe:

$$S := \sum_{n=0}^{N} q^n = 1 + q + \ldots + q^N.$$

Eine "skizzenhafte" Behandlung des Problems sieht wie folgt aus:
Differenzenbildung ergibt:

$$\begin{array}{rclllll} S & = & 1 & +q & +\ldots & +q^N & \\ qS & = & & q & +\ldots & +q^N & +q^{N+1} \\ \Rightarrow (1-q)S & = & 1 & & & & -q^{N+1} \end{array}$$

Somit ergibt sich: $S = \frac{1-q^{N+1}}{1-q}$.

Die mathematisch strenge Behandlung sieht wie folgt aus: es ist

$$S = \sum_{n=0}^{N} q^n. \tag{1}$$

Nach dem Distributivgesetz (Satz 1.5.5 (ii)) ist $qS = \sum_{n=0}^{N} q \cdot q^n = \sum_{n=0}^{N} q^{n+1}$. Wir führen die neue Summationsvariable $m = n + 1$ ein und erhalten

$$qS = \sum_{m=1}^{N+1} q^m = \sum_{n=1}^{N+1} q^n \tag{2}$$

mit Zurückbenennung $m \to n$. Aus (1) und (2) erhalten wir

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

Die Summationsintervalle in beiden Summen sind $\{n \in \mathbb{N}|\ 0 \leq n \leq N\}$ bzw. $\{n \in \mathbb{N}|\ 1 \leq n \leq N+1\}$. Wir spalten die beiden Indizes, die nicht im Durchschnitt liegen, ab und erhalten

$$\begin{array}{rclll} \sum_{n=0}^{N} q^n & = & 1+ & \sum_{n=1}^{N} q^n & \\ \sum_{n=0}^{N+1} q^n & = & & \sum_{n=1}^{N} q^n & +q^{N+1} \end{array}$$

Differenzenbildung ergibt:

$$(1-q)S = 1 + \sum_{n=1}^{N} (q^n - q^n) - q^{N+1} = 1 - q^{N+1},$$

also

$$S = \frac{1-q^{N+1}}{1-q}.$$

Satz 1.5.8. *Es sei* (a_k) *eine Folge (oder eine endliche Folge). Es sei* $m, n \in \mathbb{N}_0$ *mit* $m \leq n$ *und* $p, q \in \mathbb{N}$. *Dann ist*

$$\sum_{k=m}^{n} a_k = \sum_{k=m+p}^{n+p} a_{k-p} = \sum_{k=m-q}^{n-q} a_{k+q}.$$

Beweis. ohne Beweis. □

Beispiel 1.5.5. Wir wollen $T_n := \sum_{m=1}^{n} m$ bestimmen. Wir betrachten $S_n := \sum_{m=1}^{n} m^2$ und berechnen die Differenzen $S_{n+1} - S_n$ auf zwei verschiedene Weisen:
Es ist

$$S_{n+1} = \sum_{m=1}^{n} m^2 + (n+1)^2 = S_n + (n+1)^2,$$

also

$$S_{n+1} - S_n = (n+1)^2. \tag{1}$$

Die Berechnung mittels Indexverschiebung lautet:

$$\begin{aligned} S_{n+1} &= \sum_{m=1}^{n+1} m^2 = \sum_{k=0}^{n} (k+1)^2 = \sum_{m=0}^{n} (m+1)^2 = \sum_{m=0}^{n} (m^2 + 2m + 1) \\ &= \sum_{m=0}^{n} m^2 + 2 \sum_{m=0}^{n} m + \sum_{m=0}^{n} 1 = S_n + 2T_n + n + 1. \end{aligned}$$

Damit ist

$$S_{n+1} - S_n = 2T_n + n + 1. \tag{2}$$

Vergleich von (1) und (2) ergibt

$$(n+1)^2 = 2T_n + n + 1,$$

also

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

Die Behauptung $B(n)$ läßt sich auch durch vollständige Induktion beweisen:
Induktionsanfang $n = 1$:

$$\sum_{m=1}^{1} m = 1 \quad \text{und} \quad \frac{1}{2} \cdot 1 \cdot (1+1) = 1, \quad \text{also ist } B(1) \text{ wahr.}$$

Induktionsschritt $n \to n+1$:
Sei die Induktionshypothese für ein n richtig. Dann gilt

$$\begin{aligned} \sum_{m=1}^{n+1} m &= \sum_{m=1}^{n} m + (n+1) = T_n + n + 1 \underset{(IH)}{=} \frac{1}{2} n(n+1) + n + 1 \\ &= (n+1)\left(\frac{1}{2} n + 1\right) = \frac{1}{2}(n+1)(n+2) \end{aligned}$$

also ist $B(n)$ wahr.

Wir schließen mit Beispielen für Induktionsbeweise:

Satz 1.5.9. *(Bernoullische Ungleichung)*
Es sei $x > -1$ und $n \in \mathbb{N}$. Dann ist

$$(1+x)^n \geq 1 + nx.$$

Beweis. (Beweis durch vollständige Induktion)
Induktionsanfang $n = 1$:

$$(1+x)^1 = 1 + x.$$

Induktionsschritt $n \to n+1$:
Sei die Induktionshypothese für ein n richtig. Es gilt

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

nach Definition 1.5.8 und 1.5.9. Wegen $x > -1$ ist $(1+x) > 0$. Nach $(A3)$ (Monotonie) und der Induktionshypothese folgt:

$$(1+x)^n(1+x) \geq (1+nx)(1+x) = 1 + (n+1)x + nx^2 \geq 1 + (n+1)x$$

nach Satz 1.5.7 (Definitheit). □

Als letztes wollen wir eine Regel über die Berechnung von $(a+b)^n$, den Binomischen Lehrsatz beweisen.
Zur Vorbereitung zeigen wir eine Beziehung zwischen Binomialkoeffizienten.

Lemma 1.5.2. *Es sei $k, n \in \mathbb{N}$ mit $0 \leq k \leq n-1$. Dann ist*

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

Beweis. Nach Definition 1.5.10 ist

$$(k+1)! = k!(k+1) \quad \text{und} \quad (n-k)! = (n-k-1)!(n-k)$$

Nach Definition 1.5.10 und Satz 1.4.4 (Bruchrechnen) folgt:

$$\begin{aligned}
\binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k-1)!(n-k)} + \frac{n!}{k!(k+1)(n-k-1)!} \\
&= \frac{n!(k+1)}{(k+1)!(n-k)!} + \frac{n!(n-k)}{(k+1)!(n-k)!} \\
&= \frac{n!(n+1)}{(k+1)!(n-k)!} = \frac{(n+1)!}{((n+1)-(k+1))!(k+1)!} \\
&= \binom{n+1}{k+1}.
\end{aligned}$$

□

Satz 1.5.10. *(Binomischer Lehrsatz)*
Es seien $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$. Dann ist

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

Beweis. Wir führen den Beweis durch vollständige Induktion:
Induktionsanfang $n = 1$:

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

Induktionsschritt $n \to n+1$:
Sei die Induktionshypothese für ein n richtig. Es gilt

$$\begin{aligned} (a+b)^{n+1} &\underset{\text{Def. 1.5.9}}{=} (a+b)^n(a+b) \underset{(IH)}{=} \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}(a+b) \\ &\underset{\substack{\text{Satz 1.5.5}(ii) \\ \text{Def. 1.5.9}}}{=} \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k+1}. \end{aligned}$$

Also ist

$$(a+b)^{n+1} = \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k+1}. \tag{1}$$

Wir spalten von der ersten Summe den Term $k = n$ und von der zweiten Summe den Term $k = 0$ ab und führen eine Indexverschiebung durch. Wir erhalten

$$\sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} = \sum_{k=0}^{n-1} \binom{n}{k} a^{k+1} b^{(n+1)-(k+1)} + a^{n+1} \tag{2}$$

und

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} = b^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} a^{k+1} b^{(n+1)-(k+1)}. \tag{3}$$

Aus (1), (2), (3) und Lemma 1.5.2 erhalten wir

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

□

1.6 Die komplexen Zahlen

Die komplexen Zahlen sind kein eigentlicher Bestandteil dieser Vorlesung. Viele Begriffe aus der reellen Analysis, wie Grenzwerte, Stetigkeit, usw., lassen sich auf die komplexe Analysis übertragen. Wir werden dies gelegentlich in Hinweisen andeuten. Jedoch ist die komplexe Analysis ein eigenständiges Gebiet und besitzt viele Züge, die in der reellen Analysis keine Paralellen haben. Sie ist Gegenstand der Vorlesung Analysis IV.

Definition 1.6.1. Die Menge $\mathbb{C}$ der komplexen Zahlen ist die Menge aller Paare reeller Zahlen:

$$\mathbb{C} := \{(x,y) |\ x,y \in \mathbb{R}\}.$$

Addition und Multiplikation sind wie folgt definiert:

(i) $(x_1,y_1) + (x_2,y_2) = (x_1 + x_2, y_1 + y_2)$

(ii) $(x_1,y_1) \cdot (x_2,y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$.

Diese Regeln werden durch folgende Definition übersichtlich:

Definition 1.6.2.

$$i = (0,1).$$

Aus Definition 1.6.1 (i), (ii) ergibt sich dann die folgende Regel:

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

Die Menge $\mathbb{R}$ der reellen Zahlen kann nun durch eine leichte Änderung der Definition 1.6.1 zu einer Teilmenge der komplexen Zahlen gemacht werden.

Definition 1.6.3. Es sei $\tilde{\mathbb{C}} = (\{\mathbb{C} \backslash \{(x,0) :\ x \in \mathbb{R}\}) \cup \mathbb{R}$.
(Wir "werfen die Elemente $(x,0)$ hinaus" und ersetzen sie durch die reellen Zahlen x. Die Regel $i^2 = (-1,0)$ wird zu $i^2 = -1$ und (x,y) kann als $(x,y) = x + iy$ geschrieben werden.)
Die Rechenregeln lassen sich wie folgt sehr leicht merken:
Es gelten die üblichen Regeln (Assoziativ-, Kommutativ- und Distributivgesetze), und es ist $i^2 = -1$.

Beispiel 1.6.1. Es ist

$$(4+3i)(7+5i) = 4 \cdot 7 + 4 \cdot 5i + 7 \cdot 3i + (3i) \cdot (5i) = 28 + 15i^2 + (4 \cdot 5 + 7 \cdot 3)i \underset{i^2=-1}{=} 13 + 41i.$$

Satz 1.6.1. *Die Struktur* $(\mathbb{C}, +, \cdot)$ *ist ein Körper mit der Null 0 und der Eins 1.*
Für $x + iy \neq 0$ *haben wir*

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

Kapitel 2

Folgen und Reihen

2.1 Folgen und Grenzwerte

Definition 2.1.1. (Grenzwert, Limes)
Eine Zahl a heißt Grenzwert (Limes) der Folge (a_n), (Schreibweise: $\lim_{n\to\infty} a_n = a$ oder $a_n \to a$ $(n \to \infty)$, wenn gilt:

$$\forall \epsilon > 0 \ \exists n_0 = n_0(\epsilon) \in \mathbb{N}, \quad \text{so daß} \quad \forall n \geq n_0 \quad \text{gilt} \quad |a_n - a| < \epsilon \ \forall n \geq n_0.$$

Man sagt dann auch:
Die Folge (a_n) konvergiert gegen a oder (a_n) ist eine konvergente Folge.
Hat (a_n) keinen Grenzwert, so heißt (a_n) divergent oder (a_n) divergiert.

Definition 2.1.2. Es sei $\epsilon > 0$ und $a \in \mathbb{R}$. Unter der ϵ- Umgebung $U_\epsilon(a)$ versteht man

$$U_\epsilon(a) := \{x| \ |x - a| < \epsilon\}.$$

Man sagt: Eine Eigenschaft gilt für fast alle Elemente einer Menge bzw. Glieder einer Folge, falls es höchstens endlich viele Elemente der Menge bzw. Glieder der Folge gibt, für die sie nicht gilt.

Bemerkung 2.1.1. Die Eigenschaft der Konvergenz läßt sich auch so ausdrücken:
Man nennt a den Grenzwert von (a_n), wenn in jeder ϵ- Umgebung $U_\epsilon(a)$ fast alle Glieder der Folge (a_n) liegen.

Satz 2.1.1. *Eine Folge (a_n) hat höchstens einen Grenzwert.*

Beweis. Annahme:
Es sei $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} a_n = b$ mit $a < b$. Wir setzen $\epsilon := 1/2(b - a)$. Es ist also

$$b - a = 2\epsilon \tag{1}$$

Nach Definition 2.1.1 gibt es ein $n_0 = n_0(\epsilon)$, so daß für $n \geq n_0$ gilt:

$$|a_n - a| < \epsilon \tag{2}$$

und ein $n_1 = n_1(\epsilon)$, so daß für $n \geq n_1$ gilt:

$$|b - a_n| < \epsilon. \tag{3}$$

Es sei $n \geq \max\{n_0, n_1\}$. Aus (2) und (3) folgt nach der Dreiecksungleichung (Satz 1.4.10 (iii)):

$$|b - a| = |(b - a_n) + (a_n - a)| < |b - a_n| + |a - a_n| < 2\epsilon$$

im Widerspruch zu (1). $\square$

Satz 2.1.2. *(Grenzwerte von konstanten Folgen)*
Für $c \in \mathbb{R}$ ist $\lim\limits_{n\to\infty} c = c$.

Beweis. Wir setzen $c_n := c$ für alle n. Dann ist $|c_n - c| = 0 < \epsilon$ für alle $\epsilon > 0$ und für alle n. □

Definition 2.1.3. Gilt $\lim\limits_{n\to\infty} a_n = 0$ für eine Folge (a_n), so heißt (a_n) eine <u>Nullfolge</u>.

Satz 2.1.3. *Die Folge $\left(\frac{1}{n}\right)_{n=1}^{\infty}$ ist eine Nullfolge, d.h. es ist $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.*

Beweis. Nach dem Vollständigkeitsaxiom (V) besitzt die Menge $X = \left\{-\frac{1}{n} \mid n \in \mathbb{N}\right\}$ ein Supremum S. Wegen $-\frac{1}{n} < 0$ für alle n ist 0 eine obere Schranke von $\left(-\frac{1}{n}\right)$, also ist $S \leq 0$.
<u>Annahme</u>: $S < 0$
Nach der Definition des Supremums existiert ein n mit $-\frac{1}{n} > \frac{4}{3}S$. Dann ist aber $-\frac{1}{2n} > \frac{2}{3}S > S$, ein Widerspruch.
Damit ist $\sup\left\{\left(-\frac{1}{n}\right) \mid n \in \mathbb{N}\right\} = 0$, d.h. $\forall \epsilon > 0\ \exists n_0 = n_0(\epsilon)$ mit $-\frac{1}{n_0} > -\epsilon$ und somit $0 < \frac{1}{n_0} < \epsilon$. Wegen $0 < \frac{1}{n} \leq \frac{1}{n_0}$ für $n \geq n_0$ folgt

$$\left|\frac{1}{n}\right| < \epsilon, \quad \forall n \geq n_0(\epsilon).$$

Nach Definition 2.1.1 ist $\lim\limits_{n\to\infty} \frac{1}{n} = 0$. □

Definition 2.1.4. (Beschränktheit)
Eine Folge (a_n) heißt <u>nach oben beschränkt</u>, wenn ein $A \in \mathbb{R}$ mit $a_n \leq A$ für alle $n \in \mathbb{N}$ existiert, bzw. <u>nach unten beschränkt</u>, wenn ein $B \in \mathbb{R}$ mit $a_n \geq B$ für alle $n \in \mathbb{N}$ existiert. Man nennt (a_n) <u>beschränkt</u>, falls es nach oben und nach unten beschränkt ist.

Satz 2.1.4. *(i) Die Folge $(n)_{n=1}^{\infty}$ ist <u>nach oben unbeschränkt</u>.*

(ii) Eine Folge (a_n) ist genau dann beschränkt, wenn $\exists S \in \mathbb{R}$, so daß $|a_n| \leq S\ \forall n \in \mathbb{N}$.

Beweis. (i) <u>Annahme</u>:
Die Folge (n) ist beschränkt, d.h. es gibt ein $S \in \mathbb{R}$ mit $n \leq S$ für alle $n \in \mathbb{N}$. Dann folgt $\frac{1}{n} \geq S^{-1} > 0$ für alle $n \in \mathbb{N}$ im Widerspruch zu $\lim_{n\to\infty} \frac{1}{n} = 0$ (Satz 2.1.3).

(ii) Nach Definition 2.1.3 ist (a_n) genau dann beschränkt, wenn $A, B \in \mathbb{R}$ mit

$$B \leq a_n \leq A \quad \forall n \in \mathbb{N} \tag{1}$$

existieren. Setze $S := \max\{|A|, |B|\}$. Aus (1) folgt

$$|a_n| \leq S \quad \forall n \in \mathbb{N}. \tag{2}$$

Umgekehrt folgt aus (2) mit $S > 0$, daß $-S \leq a_n \leq S$, also die Beschränktheit von (a_n).

□

Satz 2.1.5. *Eine konvergente Folge ist beschränkt.*

Beweis. Es sei $\lim_{n\to\infty} a_n = a$. Nach Definition 2.1.1 mit $\epsilon = 1$ existiert ein $n_0 = n_0(1)$, so daß

$$|a_n - a| < 1 \quad \text{für} \quad n \geq n_0. \tag{1}$$

Aus (1) folgt

$$|a_n| \leq |a| + 1. \tag{2}$$

Nach Satz 1.5.3 (*ii*) existiert $M := \max\{|a_n|,\ n < n_0\} \cup \{|a|+1\}$. Aus (2) folgt dann: $|a_n| \leq M$ für alle $n \in \mathbb{N}$. □

Es ist oft einfacher, Beweise über die Konvergenz von Folgen zu führen, indem man nicht direkt auf die Definition 2.1.1 zurückgeht, sondern den folgenden Satz verwendet:

Satz 2.1.6. *Für die Folge (a_n) sind folgende Aussagen äquivalent:*

(i) $\lim\limits_{n\to\infty} a_n = a.$

(ii) $\exists C > 0$, *so daß* $\forall \epsilon > 0\ \exists n_0 = n_0(\epsilon)$ *mit* $|a_n - a| < C\epsilon\ \forall n \geq n_0(\epsilon)$.

Beweis. "$\Rightarrow$":
Es sei $\epsilon > 0$ gegeben. Nach Definition 2.1.1 gibt es für alle $\epsilon' > 0$ ein $n_0 = n_0(\epsilon')$, so daß $|a_n - a| < \epsilon'$ für alle $n \geq n_0(\epsilon')$ ist. Dies gilt insbesondere auch für $\epsilon' = C\epsilon$. Es gilt also $|a_n - a| < \epsilon' = C\epsilon$ für $n \geq n_0(\epsilon')$.
"$\Leftarrow$":
Es sei $\epsilon > 0$ gegeben. Wir setzen $\epsilon' = C^{-1}\epsilon$. Dann gibt es nach Voraussetzung ein $n_0 = n_0(\epsilon')$, so daß $|a_n - a| < C\epsilon' = \epsilon$ für alle $\epsilon > 0$ ist. Nach Definition 2.1.1 ist $\lim_{n\to\infty} a_n = a$. □

Satz 2.1.7. *(Grenzwertsätze)*
Es seien (a_n) und (b_n) Folgen mit $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} b_n = b$. Dann ist

(i) $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n = a + b$

(ii) $\lim\limits_{n\to\infty} (a_n \cdot b_n) = \left(\lim\limits_{n\to\infty} a_n\right) \cdot \left(\lim\limits_{n\to\infty} b_n\right) = ab$

(iii) $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{a}{b}$, *falls* $b \neq 0$.

Beweis. (i) Es sei $\epsilon > 0$. Nach Definition 2.1.1 gibt es $n_0 = n_0(\epsilon)$ und $n_1 = n_1(\epsilon) \in \mathbb{N}$, so daß $|a_n - a| < \epsilon$ für alle $n \geq n_0(\epsilon)$ und $|b_n - b| < \epsilon$ für alle $n \geq n_1(\epsilon)$. Es sei $n_2 := \max\{n_0, n_1\}$. Dann ist für $n \geq n_2$ nach der Dreiecksungleichung (Satz 1.4.10 (ii))

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| < |a_n - a| + |b_n - b| < 2\epsilon.$$

Nach Satz 2.1.3 folgt $\lim\limits_{n\to\infty} (a_n + b_n) = a + b$.

(ii) Nach Satz 2.1.5 ist die Folge (b_n) beschränkt, d.h. es existiert ein $B \in \mathbb{R}$ mit

$$|b_n| \leq B \quad \forall n \in \mathbb{N}. \tag{1}$$

Es sei $\epsilon > 0$. Nach Definition 2.1.1 existiert ein $n_0 = n_0(\epsilon)$ und $n_1 = n_1(\epsilon) \in \mathbb{N}$ mit

$$|a_n - a| < \epsilon \quad \forall n \geq n_0, \tag{2}$$

$$|b_n - b| < \epsilon \quad \forall n \geq n_1. \tag{3}$$

Es sei $n_2 := \max\{n_0, n_1\}$. Dann ist für alle $n \geq n_2$

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - ab_n + ab_n - ab| \\ &\underset{\Delta-\text{Ugl.}}{\leq} |a_n - a||b_n| + |a||b_n - b| \\ &\underset{(1),(2),(3)}{<} (B + |a|)\epsilon. \end{aligned}$$

(iii) Nach Definition 2.1.1 gibt es $n_0 \in \mathbb{N}$, so daß $|b_n - b| \leq 1/2|b|$ für alle $n \geq n_0$ ist, und damit gilt nach Satz 1.4.11 *(iv)*

$$|b_n| \geq \frac{1}{2}|b| \quad \forall n \geq n_0. \tag{1}$$

Nach Definition 2.1.1 gibt es $n_1 = n_1(\epsilon)$ bzw. $n_2 = n_2(\epsilon)$, so daß für alle $n \geq n_1(\epsilon)$ bzw. für alle $n \geq n_2(\epsilon)$ gilt, daß $|a_n - a| < \epsilon$ bzw. $|b_n - b| < \epsilon$. Wir setzen $n_3 := \max\{n_1, n_2\}$ und erhalten

$$|a_n - a| < \epsilon, \quad |b_n - b| < \epsilon \quad \forall n \geq n_3. \tag{2}$$

Es folgt

$$\begin{aligned} \left|\frac{a_n}{b_n} - \frac{a}{b}\right| &= \frac{1}{|bb_n|}|a_n b - ab_n| = \frac{1}{|bb_n|}|a_n b - ab + ab - ab_n| \\ &\underset{\Delta-\text{Ugl.}}{\leq} \frac{1}{|bb_n|}\left(|b||a_n - a| + |a||b_n - b|\right) \\ &\underset{(1),(2)}{\leq} \frac{2}{|b|^2}\left(|a| + |b|\right)\epsilon. \end{aligned}$$

□

Satz 2.1.8. *(Erhaltung von Ungleichungen)*
Es seien (a_n) und (b_n) Folgen mit $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$ sowie $a_n \leq b_n$ für alle n. Dann ist $a \leq b$, also $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.

Beweis. Annahme: $a > b$
Wir setzen

$$2\epsilon := a - b. \tag{1}$$

Nach Definition 2.1.1 existiert $n_0 = n_0(\epsilon)$, so daß für alle $n \geq n_0(\epsilon)$ gilt:

$$|a_n - a| < \epsilon \quad \text{und} \quad |b_n - b| < \epsilon. \tag{2}$$

Aus (2) folgt

$$a_n > a - \epsilon \quad \text{und} \quad b_n < b + \epsilon \quad \forall n \geq n_0. \tag{3}$$

Aus (1) und (3) folgt $b_n < a_n$ für alle $n \geq n_0$, ein Widerspruch. Nach Axiom $(A1)$ (Trichotomiegesetz) folgt $a \leq b$. □

Bemerkung 2.1.2. Aus der scharfen Ungleichung $a_n < b_n$ kann nicht $\lim_{n\to\infty} a_n < \lim_{n\to\infty} b_n$ gefolgert werden, sondern auch nur $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$. Die sieht man am Beispiel $a_n = 0$ und $b_n = \frac{1}{n}$. Es ist $a_n < b_n$, aber $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Satz 2.1.9. *Es sei $q \in \mathbb{R}$ und $|q| < 1$. Dann ist $\lim_{n\to\infty} q^n = 0$.*

Beweis. Fall 1: $q = 0$
Dann ist $\lim_{n\to\infty} q^n = \lim_{n\to\infty} 0 = 0$.

Fall 2:
Es sei $Q := |q|^{-1}$. Dann ist $Q = 1 + \eta$ mit $\eta > 0$. Nach der Bernoullischen Ungleichung (Satz 1.5.9) ist $Q^n = (1+\eta)^n \geq 1 + n\eta$. Also ist $|q|^n = Q^{-n} \leq (1+n\eta)^{-1} \leq \eta^{-1}n^{-1}$ und damit $-\eta^{-1}n^{-1} \leq q^n \leq \eta^{-1}n^{-1}$. Nach den Sätzen 2.1.4, 2.1.7 und 2.1.8 ist

$$0 = \lim_{n\to\infty} \left(-\eta^{-1}n^{-1}\right) \leq \lim_{n\to\infty} q^n \leq \lim_{n\to\infty} \eta^{-1}n^{-1} = 0,$$

also $\lim_{n\to\infty} q^n = 0$. $\square$

Beispiel 2.1.1. Es sei $a_n = 1 + \frac{1}{n} + 10^{-6}(-1)^n$.
Wir untersuchen, ob das Konvergenzkriterium von Definition 2.1.1 für $a = 1$ erfüllt ist.

Fall 1:
Es sei $\epsilon > 10^{-6}$. Dann können wir schreiben: $\epsilon = 10^{-6} + \eta_1$ mit $\eta_1 > 0$. Wir wählen $n_0 = n_0(\epsilon) \in \mathbb{N}$, so daß $n_0 > \eta_1^{-1}$ ist. Für $n \geq n_0$ haben wir dann

$$|a_n - 1| = \left|\frac{1}{n} + 10^{-6}(-1)^n\right| < 10^{-6} + \eta_1 = \epsilon.$$

Fall 2:
Es sei $\epsilon \leq 10^{-6}$, z.B. $\epsilon = 10^{-6} - \eta_2$ mit $\eta_2 > 0$. Wir wählen $n_1 = n_1(\epsilon) \in \mathbb{N}$, so daß $n_1 > \eta_2^{-1}$. Für $n \geq n_1$ haben wir dann

$$|a_n - 1| = \left|\frac{1}{n} + 10^{-6}(-1)^n\right| > 10^{-6} - \eta_2 = \epsilon.$$

Damit ist das Kriterium von Definition 2.1.1 mit $a = 1$ zwar im Fall $\epsilon > 10^{-6}$ erfüllt, aber nicht im Falle $\epsilon \leq 10^{-6}$. Somit konvergiert a_n nicht gegen 1.

Es gibt auch keinen anderen Grenzwert a.
Es sei $a \in \mathbb{R}$. Nach der Dreieicksungleichung ist

$$2 \cdot 10^{-6} = \left|\left(1 + 10^{-6} - a\right) - \left(1 - 10^{-6} - a\right)\right| \leq \left|1 + 10^{-6} - a\right| + \left|1 - 10^{-6} - a\right|.$$

Daraus folgt: Für alle $a \in \mathbb{R}$ gilt einer der folgenden Fälle:
Fall a: $\left|a - \left(1 + 10^{-6}\right)\right| > 10^{-6}$ oder
Fall b: $\left|a - \left(1 - 10^{-6}\right)\right| > 10^{-6}$.

Wir zeigen, daß in Fall a) die Zahl a nicht der Grenzwert von (a_n) sein kann. Fall b) wird analog behandelt.
Es sei $n \geq 2 \cdot 10^6$. Dann ist für alle geraden n folglich $|a_n - a| \geq 10^{-6} - \frac{1}{n} \geq \frac{1}{2} \cdot 10^{-6}$.
Für $\epsilon = 10^{-6}$ ist das Kriterium von Definition 2.1.1 nicht erfüllt: es gibt kein $n_0(\epsilon)$, so daß für $n \geq n_0(\epsilon)$ dann $|a_n - a| < \epsilon$ gilt. Also ist a nicht Grenzwert von (a_n).

Es gibt jedoch Zahlen, nämlich $l_1 = 1 - 10^{-6}$ und $l_2 = 1 + 10^{-6}$, die schwächere Eigenschaften als das Konvergenzkriterium erfüllen. In beliebigen ϵ- Umgebungen $U_\epsilon(l_1)$ bzw. $U_\epsilon(l_2)$ liegen zwar nicht fast alle Glieder der Folge (a_n), aber doch unendlich viele. Die Zahlen l_1, l_2 sind Beispiele von Häufungswerten.

Definition 2.1.5. Es sei (a_n) eine Folge. Dann heißt a Häufungswert (HW) von (a_n), falls es für jedes $\epsilon > 0$ unendlich viele $n \in \mathbb{N}$ mit $a_n \in U_\epsilon(a)$ gibt.

Satz 2.1.10. *Es sei (a_n) eine Folge und $a \in \mathbb{R}$. Dann ist a genau dann Häufungswert von (a_n), wenn es eine Teilfolge $(a_{n_k})_{k=1}^{\infty}$ von (a_n) mit $\lim\limits_{k\to\infty} a_{n_k} = a$ gibt.*

Beweis. "$\Rightarrow$":
Es sei a Häufungswert von (a_n). Wir definieren durch vollständige Induktion eine Folge (n_k) mit $n_k \in \mathbb{N}$, $n_{k+1} > n_k$, so daß $|a_{n_k} - a| < \frac{1}{k}$:
$k = 1$:
Nach Definition 2.1.5 gibt es ein n_1 mit $|a_{n_1} - a| < 1$.
$k \to k+1$:
Es sei n_k schon definiert und $|\,|a_{n_k} - a| < \frac{1}{k}$. Nach Definition 2.1.5 gibt es $n_{k+1} > n_k$, so daß $\left|a_{n_{k+1}} - a\right| < \frac{1}{k+1}$ gilt. Nach Definition 2.1.1 ist $\lim_{n\to\infty} a_{n_k} = a$.

"$\Leftarrow$":
Es sei $\lim_{n\to\infty} a_{n_k} = a$. Es sei $\epsilon > 0$. Dann gibt es nach Definition 2.1.1 ein $k_0 = k_0(\epsilon)$, so daß $|a_{n_k} - a| < \epsilon$ für ∞- viele k. □

Beispiel 2.1.2. Es sei $a_n = 1 + \frac{1}{n} + 10^{-6}(-1)^n$ wie in Beispiel 2.1.1. Dann ist

$$\begin{aligned} \lim_{k\to\infty} a_{2k} &= 1 + 10^{-6} = l_2 \quad \text{und} \\ \lim_{k\to\infty} a_{2k+1} &= 1 - 10^{-6} = l_1. \end{aligned}$$

Definition 2.1.6. (Infimum, Supremum)
Es sei $X \subset \mathbb{R}$ und $u \in \mathbb{R}$. Dann heißt u untere Schranke von X, falls $u \le x$ für alle $x \in X$ gilt. Falls eine untere Schranke von X existiert, heißt X nach unten beschränkt. Man nennt u größte untere Schranke oder Infimum von X (Schreibweise: $\inf X$), falls u eine untere Schranke von X ist und für alle unteren Schranken t von X gilt, daß $t \le u$ ist. Man nennt X beschränkt, falls X nach oben und unten beschränkt ist. Für das Supremum von X (Definition 1.3.5) schreiben wir $\sup X$.

Satz 2.1.11. *(Existenz von Infimum und Supremum)*

(i) Eine nach oben beschränkte Menge reeller Zahlen hat stets ein Supremum.

(ii) Eine nach unten beschränkte Menge reeller Zahlen hat stets ein Infimum.

Beweis. (i) Dies ist das Vollständigkeitsaxiom (V).

(ii) Man betrachte die Menge $-X := \{-x|\ x \in X\}$. Dann gilt: s ist untere Schranke von $X \Leftrightarrow -s$ ist obere Schranke von $-X$. Also ist $\inf X = \sup(-X)$, und $\sup(-X)$ existiert nach (i).

□

Definition 2.1.7. (Monotonie)
Die Folge (a_n) heißt monoton wachsend (bzw. monoton fallend), wenn $a_{n+1} \ge a_n$ für alle n (bzw. $a_{n+1} \le a_n$ für alle n) gilt. Gilt die scharfe Ungleichung $a_{n+1} > a_n$ (bzw. $a_{n+1} < a_n$), so heißt (a_n) streng monoton wachsend (bzw. streng monoton fallend). Eine Folge heißt monoton, wenn sie monoton wachsend oder monoton fallend ist. Ist $\lim_{n\to\infty} a_n = a$ und (a_n) monoton wachsend (bzw. fallend), so schreiben wir auch $a_n \uparrow a$ (bzw. $a_n \downarrow a$).

Satz 2.1.12. *(Monotoniekriterium)*
Eine beschränkte monotone Folge ist konvergent.

Beweis. Fall 1: (a_n) ist monoton wachsend:
Nach Satz 2.1.11 hat $X := \{a_n |\ n \in \mathbb{N}\}$ ein Supremum s. Es sei $\epsilon > 0$. Nach der Definition des Supremums gibt es ein $n_0 = n_0(\epsilon)$, so daß $a_{n_0} \geq s - \epsilon$. Wegen der Monotonie ist dann $s - \epsilon \leq a_n \leq s$ für alle $n \geq n_0$. Nach Definition 2.1.1 folgt $\lim_{n\to\infty} a_n = s$.
Fall 2: (a_n) ist monoton fallend:
Dann ist $(-a_n)$ monoton wachsend, und es ist $\lim_{n\to\infty} a_n = -\lim_{n\to\infty}(-a_n)$. □

Definition 2.1.8. Unter der Länge eines Intervalls $I = [a, b]$ mit $a, b \in \mathbb{R}$ (Schreibweise: $|I|$), versteht man $|I| = b - a$.

Definition 2.1.9. (Intervallschachtelung)
Eine Folge (I_n) von kompakten Intervallen heißt Intervallschachtelung, wenn

(i) $I_{n+1} \subset I_n$

(ii) $\lim\limits_{n\to\infty} |I_n| = 0$.

Satz 2.1.13. *Es sei (I_n) eine Intervallschachtelung, und $I_n = [a_n, b_n]$. Dann existiert genau ein $a \in \mathbb{R}$ mit $a_n \uparrow a$ und $b_n \downarrow a$ mit $n \to \infty$.*

Beweis. Es sei $I_n = [a_n, b_n]$. Dann ist die Folge (a_n) monoton wachsend und die Folge (b_n) monoton fallend. Also ist $a_n \leq b_n$ für alle $n \in \mathbb{N}$, und somit ist (a_n) beschränkt. Nach Satz 2.1.12 ist (a_n) konvergent, d.h. $\lim_{n\to\infty} a_n = a$ für $a \in \mathbb{R}$. Wegen $a_{n_0} \leq a_n$ für alle $n \geq n_0$ folgt nach Satz 2.1.8, daß $a_{n_0} = \lim_{n\to\infty} a_{n_0} \leq \lim_{n\to\infty} a_n = a$ ist, also ist $a_{n_0} \leq a$ für alle n_0.
Analog zeigt man, daß die Folge b_n konvergiert. Es sei $\lim_{n\to\infty} b_n = b$. Dann ist $b \leq b_n$ für alle $n \in \mathbb{N}$. Also ist $0 \leq b - a \leq b_n - a_n$ für alle $n \in \mathbb{N}$. Aus $\lim_{n\to\infty} b_n - a_n = 0$ folgt $b - a = 0$ und daraus $a = b$. □

Satz 2.1.14. *(Bolzano- Weierstraß)*
Eine beschränkte Folge hat mindestens einen Häufungswert.

Beweis. Es sei
$$s \leq c_n \leq t \quad \text{für alle} \quad n \in \mathbb{N} \quad \text{und} \quad s, t \in \mathbb{R}. \tag{1}$$
Wir zeigen durch vollständige Induktion, daß eine Intervallschachtelung (I_m) mit $|I_m| = (t - s) \cdot 2^{-m}$ und $c_n \in I_m$ für unendlich viele n existiert:
$m = 0$:
Wir setzen $I_0 := [s, t]$. Wegen (1) gilt $c_n \in I_0$ für alle n.
$m \to m + 1$:
Es sei $|I_m| = (t - s) \cdot 2^{-m}$, $I_m = [a_m, b_m]$ und $c_n \in I_m$ für unendlich viele n. Dann ist $I_m = I_{m,1} \cup I_{m,2}$ mit $I_{m,1} = \left[a_m, \frac{a_m+b_m}{2}\right]$ und $I_{m,2} = \left[\frac{a_m+b_m}{2}, b_m\right]$. Für mindestens ein $j \in \{1, 2\}$ ist $a_n \in I_{m,j}$ für unendlich viele n. Wir setzen $I_{m+1} := I_{m,j}$. Es ist $|I_{m+1}| = \frac{1}{2}|I_m| = (t - s) \cdot 2^{-(m+1)}$. Nach Satz 2.1.9 ist $|I_m| \to 0$ für $m \to \infty$. Nach Satz 2.1.13 gibt es ein $a \in \mathbb{R}$ mit $a_n \uparrow a$ und $b_m \downarrow a$.
Es sei $\epsilon > 0$. Wegen $|I_m| \to 0$ für $m \to \infty$ existiert ein m mit $|I_m| < \epsilon$. Für die unendlich vielen n mit $c_n \in I_m$ gilt: $a_m < c_n < b_m$. Es folgt $|c_n - a| < \epsilon$. Damit ist a Häufungswert von (c_n). □

Definition 2.1.10. Eine Folge (a_n) heißt Cauchyfolge, falls für alle $\epsilon > 0$ ein $n_0 = n_0(\epsilon)$ existiert, so daß für alle Paare $(m, n) \in \mathbb{N}^2$ mit $m \geq n_0$ und $n \geq n_0$ gilt, daß $|a_m - a_n| < \epsilon$.

Satz 2.1.15. *(Cauchykriterium)*
Eine Folge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.

Beweis. "$\Rightarrow$":
Es sei $\lim\limits_{n\to\infty} a_n = a$.
Es sei $\epsilon > 0$. Nach Definition 2.1.1 existiert ein $n_0 = n_0(\epsilon/2)$, so daß $|a_n - a| < \epsilon/2$ für alle $n \geq n_0$ ist. Es sei $m \geq n_0$ und $n \geq n_0$ mit $m, n \in \mathbb{N}$. Aus $|a_m - a| < \epsilon/2$ und $|a_n - a| < \epsilon/2$ folgt

$$|a_m - a_n| = |(a_m - a) - (a_n - a)| \underset{\Delta-Ugl.}{\leq} |a_m - a| + |a_n - a| < \epsilon.$$

Damit ist (a_n) eine Cauchyfolge.
"$\Leftarrow$":
Es sei (a_n) eine Cauchyfolge.
Wenn wir in Definition 2.1.10 $\epsilon = 1$ setzen, erhalten wir ein $n_0 \in \mathbb{N}$, so daß $|a_m - a_n| < 1$ für alle (m, n) mit $m \geq n_0$ und $n \geq n_0$ ist. Es sei $M := \max\{a_n|\ n \leq n_0\}$. Dann ist für $n \geq n_0$ gerade $|a_n| \leq |a_{n_0}| + 1$. Also ist $|a_n| \leq \max\{M, |a_{n_0}| + 1\}$ für alle $n \in \mathbb{N}$, und damit ist (a_n) beschränkt. Nach Satz 2.1.14 (Bolzano- Weierstraß) hat (a_n) einen Häufungswert a.
Es sei $\epsilon > 0$. Da (a_n) eine Cauchyfolge ist, gibt es $n_0 = n_0(\epsilon)$, so daß für $m, n \geq n_0(\epsilon)$ gilt

$$|a_m - a_n| < \epsilon. \tag{1}$$

Da a Häufungswert von (a_n) ist, gibt es ein $m_0 \in \mathbb{N}$, $m_0 \geq n_0$, mit

$$|a_{m_0} - a| < \epsilon. \tag{2}$$

Da (1) auch für $m = m_0$ gilt, ist

$$|a_{m_0} - a_n| < \epsilon \tag{3}$$

für alle $n \geq n_0$. Aus (2) und (3) folgt

$$|a - a_n| = |a - a_{m_0} + a_{m_0} - a_n| \underset{\Delta-Ugl.}{\leq} |a - a_{m_0}| + |a_{m_0} - a_n| < 2\epsilon$$

für alle $n \geq n_0$. Nach Satz 2.1.6 ist $\lim\limits_{n\to\infty} a_n = a$. □

Satz 2.1.16. *Die Menge H der Häufungspunkte einer beschränkten Folge ist beschränkt.*

Beweis. Es sei s eine obere Schranke der Folge (a_n), d.h. $a_n \leq s$ für alle n. Weiter sei $a = s + \epsilon$ mit $\epsilon > 0$. Dann ist $a_n \notin U_\epsilon(a)$ für alle n. Also ist a kein Häufungswert von (a_n). Für jeden Häufungswert a von (a_n) gilt also $a \leq s$.
Analog zeigt man: $a \geq u$ für jede untere Schranke u von (a_n). □

Definition 2.1.11. Es sei (a_n) eine beschränkte Folge und H die Menge aller Häufungswerte von (a_n). Dann definiert man

$$\begin{aligned} \overline{\lim_{n\to\infty}}\, a_n &= \limsup_{n\to\infty} a_n := \sup H \quad \text{und} \\ \underline{\lim_{n\to\infty}}\, a_n &= \liminf_{n\to\infty} a_n := \inf H \end{aligned}$$

(Sprich: Limes Superior und Limes Inferior).

Satz 2.1.17. *Für eine beschränkte Folge (a_n) existieren stets $\limsup\limits_{n\to\infty} a_n$ und $\liminf\limits_{n\to\infty} a_n$ und sind eindeutig bestimmt.*
Insbesondere sind $\limsup\limits_{n\to\infty} a_n$ und $\liminf\limits_{n\to\infty} a_n$ der größte bzw. der kleinste Häufungswert von (a_n).

Beweis. Die Existenz folgt aus Satz 2.1.16 und die Eindeutigkeit aus der Eindeutigkeit des Supremums und des Infimums.
Es sei H die Menge der Häufungswerte von (a_n) und $l := \limsup_{n\to\infty} a_n$, also nach Definition 2.1.11 ist $l = \sup H$. Es sei $\epsilon > 0$. Da l nach Definition 1.3.5 die kleinste obere Schranke von H ist, gibt es einen Häufungswert w von (a_n) mit $l - \epsilon/2 < w \le l$. Nach Definition 2.1.5 gibt es unendlich viele n, so daß $a_n \in U_{\epsilon/2}(w)$. Für diese n gilt

$$|a_n - l| \le |(a_n - w) + (w - l)| \underset{\Delta - Ugl.}{\le} |a_n - w| + |w - l| < \epsilon.$$

Also ist $a_n \in U_\epsilon(l)$.Damit ist $l \in H$, also ist $l = \max H$.
Analog zeigt man, daß $\liminf_{n\to\infty} a_n \in H$. □

Satz 2.1.18. *Es sei (a_n) eine beschränkte Folge und $l \in \mathbb{R}$.*

(i) Es ist genau dann $l = \limsup_{n\to\infty} a_n$, wenn es für alle $\epsilon > 0$ unendlich viele n mit $a_n > l - \epsilon$, aber höchstens endlich viele n mit $a_n > l + \epsilon$ gibt. $(*)$

(ii) Es ist genau dann $l = \liminf_{n\to\infty} a_n$, wenn es für alle $\epsilon > 0$ unendlich viele n mit $a_n < l + \epsilon$, aber höchstens endlich viele n mit $a_n < l - \epsilon$ gibt.

Beweis. Wir zeigen nur (i).
"$\Rightarrow$":
Es sei $l = \limsup_{n\to\infty} a_n$ und $\epsilon > 0$. Nach Satz 2.1.17 ist l ein Häufungswert von (a_n). Nach Definition 2.1.5 gibt es unendlich viele n mit $a_n \in U_\epsilon(l) = (l - \epsilon, l + \epsilon)$. Für diese n gilt insbesondere $a_n > l - \epsilon$.
Annahme: es existieren unendlich viele n mit $a_n > l + \epsilon$.
Es sei $X = \{n| \ a_n > l + \epsilon\}$. Dann ist X eine unendliche Menge. Es sei $X = \{n_k| \ k \in \mathbb{N}\}$ mit $n_k < n_{k+1}$. Dann ist $(a_{n_k})_{k=1}^\infty$ eine unendliche beschränkte Teilfolge von (a_n). Diese hat nach Satz 2.1.14 (Bolzano- Weierstraß) einen Häufungswert l'. Wäre $l' \le l$, so gäbe es keine n_k mit $a_{n_k} \in U_\epsilon(l')$ im Widerspruch zur Definition von l' als Häufungswert von (a_{n_k}). Also muß $l' > l$ sein, was im Widerspruch zur Tatsache, daß l der größte Häufungswert von (a_n) ist, steht. Damit gilt $(*)$.

"$\Leftarrow$":
Wir nehmen die Gültigkeit von $(*)$ an. Dann ist l ein Häufungswert von (a_n). Es sei $l' = l + 2\epsilon$ mit $\epsilon > 0$. Dann gibt es höchstens endlich viele n mit $a_n \in U_\epsilon(l')$. Also ist l' kein Häufungswert. □

Bemerkung 2.1.3. Für $l = \limsup_{n\to\infty} a_n$ gibt es nach Satz 2.1.18 höchstens endlich viele n, welche $a_n > l + \epsilon$ erfüllen. Es kann jedoch unendlich viele n mit $a_n > l$ geben, wie das Beispiel $a_n = \frac{1}{n}$ und $l = 0$ zeigt.

Satz 2.1.19. *Es sei (a_n) eine beschränkte Folge. Die Folge ist genau dann konvergent, wenn sie nur einen einzigen Häufungswert besitzt.*
In diesem Fall ist dann

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n.$$

Beweis. "$\Leftarrow$":
Es sei $H = \{l\}$ mit $l \in \mathbb{R}$. Dann ist $\inf H = \sup H = l$, also

$$l = \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n.$$

Es sei $\epsilon > 0$. Nach Satz 2.1.18 gibt es höchstens endlich viele n mit $a_n > l + \epsilon$ (Teil (i)) und höchstens endlich viele n mit $a_n < l - \epsilon$ (Teil (ii)). Also gilt für fast alle n:

$$a_n \in (l - \epsilon, l + \epsilon) = U_\epsilon(l).$$

Nach Definition 2.1.1 bedeutet dies $\lim\limits_{n\to\infty} a_n = l$.

"$\Rightarrow$":
Es sei $a = \lim\limits_{n\to\infty} a_n$. Dann ist nach Satz 2.1.18 sowohl $a = \limsup\limits_{n\to\infty} a_n$, als auch $a = \liminf\limits_{n\to\infty} a_n$. Damit ist $H = \{a\}$. $\square$

Man kann nun die in diesem Abschnitt definierten Konzepte noch erweitern, indem man die Menge $\mathbb{R}$ zur Menge $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ mit den neuen Elementen ∞ (unendlich) und $-\infty$ (minus unendlich) erweitert. Diese Objekte ∞ und $-\infty$ können dann als uneigentliche Grenzwerte, Häufungswerte, etc. auftreten.

Definition 2.1.12. (Umgebungen von ∞ und $-\infty$)
Es sei $c \in \mathbb{R}$. Unter der c- Umgebung von ∞ ($U_c(\infty)$) bzw. der c- Umgebung von $-\infty$ ($U_c(-\infty)$) versteht man

$$\begin{aligned} U_c(\infty) &:= (c, \infty) = \{x \mid x > c\} \quad \text{bzw.} \\ U_c(-\infty) &:= (-\infty, c) = \{x \mid x < c\}. \end{aligned}$$

Definition 2.1.13. Es sei (a_n) eine Folge. Man sagt, (a_n) divergiert gegen ∞ bzw. gegen $-\infty$ (Schreibweise: $\lim\limits_{n\to\infty} a_n = \infty$, $a_n \to \infty$ für $n \to \infty$ bzw. $\lim\limits_{n\to\infty} a_n = -\infty$, $a_n \to -\infty$ für $n \to \infty$), wenn für alle $c > 0$ für fast alle n gilt, daß $a_n \in U_c(\infty)$ bzw. $a_n \in U_c(-\infty)$ ist.
Dann heißen ∞ bzw. $-\infty$ uneigentlicher Grenzwert von (a_n).

Beispiel 2.1.3. Nach Satz 2.1.4 ist $\lim\limits_{n\to\infty} n = \infty$, $n \to \infty$ für $n \to \infty$.

Definition 2.1.14. Man nennt ∞ bzw. $-\infty$ uneigentliche Häufungswerte der Folge (a_n), wenn für alle $c > 0$ es unendlich viele n mit $a_n \in U_c(\infty)$ bzw. $a_n \in U_c(-\infty)$ gibt.

Definition 2.1.15. Es sei $X \subset \mathbb{R}$ nach oben (bzw. nach unten) unbeschränkt. Dann schreiben wir $\sup X = \infty$ (bzw. $\inf X = -\infty$).

Definition 2.1.16. Es sei (a_n) eine Folge , H die Menge der eigentlichen und uneigentlichen Häufungswerte von (a_n). Dann setzen wir

$$\limsup_{n\to\infty} a_n = \sup H \quad \text{und} \quad \liminf_{n\to\infty} a_n = \inf H.$$

2.2 Die n- te Wurzel

Satz 2.2.1. *Es sei $n \in \mathbb{N}$ und $x \in [0, \infty)$. Dann gibt es genau ein $y \geq 0$, so daß $y^n = x$ ist.*

Beweis. Es sei $W := \{z \mid z \in [0, \infty), z^n \leq x\}$. Nach der Bernoullischen Ungleichung (Satz 1.5.9) ist für $z \geq 1 + \frac{x}{n}$:

$$z^n \geq \left(1 + \frac{x}{n}\right)^n \geq 1 + n \cdot \frac{x}{n} = x + 1 > x.$$

Damit ist W beschränkt und nach Satz 2.1.11 existiert $y_0 := \sup W$.

Wir zeigen im folgenden: $y_0^n = x$.
Annahme: $y_0^n < x$:
Dann gibt es ein $\delta > 0$, so daß $x = y_0^n + \delta$. Wir setzen

$$\begin{aligned} M &:= \sum_{k=0}^{n-1} \binom{n}{k} y_0^k, \\ \epsilon &:= \min\left\{\frac{1}{2}\delta M^{-1}, 1\right\} \quad \text{und} \\ z &:= y_0 + \epsilon. \end{aligned}$$

Wegen $\epsilon \leq 1$ ist $\epsilon^m \leq \epsilon$ für alle $m \in \mathbb{N}$. Nach dem Binomischen Lehrsatz (Satz 1.5.10) ist dann

$$z^n = (y_0 + \epsilon)^n = y_0^n + \sum_{k=0}^{n-1} \binom{n}{k} y_0^k \epsilon^{n-k} \leq y_0^n + M\epsilon \leq y_0^n + \frac{1}{2}\delta < x.$$

Damit ist $z \in W$ mit $z > y_0$, ein Widerspruch.

Annahme: $y_0^n > x$:
Dann gibt es ein $\delta > 0$ mit $x = y_0^n(1-\delta)$. Es sei $\epsilon := \min\left\{\frac{1}{2}n^{-1}\delta, \frac{1}{2}\right\}$ und $z = y_0(1-\delta)$. Nach der Bernoullischen Ungleichung ist

$$z^n = y_0^n(1-\epsilon)^n \geq y_0^n(1-n\epsilon) > x.$$

Damit ist z eine obere Schranke von W mit $z < y_0$, ein Widerspruch.
Es ist also $y_0^n = x$.
Weiter ist

$$\begin{aligned} 0 < z < y_0 &\Rightarrow z^n < y_0^n = x \\ z > y_0 &\Rightarrow z^n > y_0^n = x. \end{aligned}$$

Also ist

$$y^n = x, y \geq 0 \Leftrightarrow y = y_0.$$

□

2.3 Unendliche Reihen

Definition 2.3.1. Es sei (a_n) eine Zahlenfolge. Unter der unendlichen Reihe (kurz: Reihe) $\sum_{m=1}^{\infty} a_m$ (Sprechweise: Summe $m = 1$ bis unendlich a_m) versteht man die Folge $(S_n)_{n=1}^{\infty}$ der Partialsummen $S_n = \sum_{m=1}^{n} a_m$. Die Reihe $\sum_{m=1}^{\infty} a_m$ heißt konvergent, falls

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{m=1}^{n} a_m =: S$$

existiert. Man schreibt dann $\sum_{m=1}^{\infty} a_m = s$ und nennt S den Wert der unendlichen Reihe. Ist $\sum_{m=1}^{\infty} a_m$ nicht konvergent, so heißt es divergent.

Bemerkung 2.3.1. Man kann allgemeiner auch unendliche Reihen der Form $\sum_{m=k}^{\infty} a_m$ mit $k \in \mathbb{N}_0$ betrachten. Die Änderung in der Definition ist offensichtlich.

Definition 2.3.2. Es sei $q \in \mathbb{R}$. Unter der (unendlichen) geometrischen Reihe versteht man $\sum_{n=0}^{\infty} q^n$.

Satz 2.3.1. *Für* $|q| < 1$ *ist* $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$.
Für $|q| \geq 1$ *ist* $\sum_{n=0}^{\infty} q^n$ *divergent.*

Beweis. Für die Folge der Partialsummen $S_k := \sum_{n=0}^{k} q^n$ gilt: $S_{k+1} - S_k = q^{k+1}$. Nach dem Cauchykriterium (2.1.15) ist die unendliche Reihe höchstens dann konvergent, wenn $S_{k+1} - S_k \to 0$ für $k \to \infty$, also für $|q| < 1$.
Nach Beispiel 1.5.4 ist $S_k = \frac{1-q^{k+1}}{1-q}$. Nach Satz 2.1.9 ist $\lim_{k\to\infty} q^{k+1} = 0$ für $|q| < 1$. Es gilt also $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$ für $|q| < 1$.
Für $|q| \geq 1$ ist die unendliche Reihe divergent. □

Beispiel 2.3.1. Wir bestimmen $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$.
Es ist
$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$
Folglich ist
$$S_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} = 1 - \frac{1}{n+1}.$$
Also ist
$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} S_n = 1.$$

Satz 2.3.2. *(Grenzwertsätze)*
Es seien $\sum_{k=1}^{\infty} a_k$ *und* $\sum_{k=1}^{\infty} b_k$ *konvergente Reihen, und es sei* $a \in \mathbb{R}$. *Dann konvergieren die Reihen* $\sum_{k=1}^{\infty} (a_k + b_k)$ *und* $\sum_{k=1}^{\infty} a \cdot a_k$, *und es gilt*

(i) $\sum_{k=1}^{\infty} (a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$

(ii) $a \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a \cdot a_k$

(iii) Ist $a_k \leq b_k$ für alle $k \in \mathbb{N}$, dann gilt $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$.

Beweis. Wir beweisen nur (i):
Es seien $S_n := \sum_{k=1}^{n} a_k$ und $T_n := \sum_{k=1}^{n} b_k$ die Partialsummen der beiden Reihen. Nach Definition 2.3.1 ist

$$\sum_{k=1}^{\infty}(a_k + b_k) = \lim_{n\to\infty}\left(\sum_{k=1}^{n}(a_k + b_k)\right) = \lim_{n\to\infty}(S_n + T_n) \underset{S.2.1.7}{=} \lim_{n\to\infty} S_n + \lim_{n\to\infty} T_n = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k.$$

□

2.4 Konvergenzkriterien für unendliche Reihen

Satz 2.4.1. *Für eine konvergente Reihe* $\sum_{n=1}^{\infty} a_n$ *ist* $\lim_{n\to\infty} a_n = 0$.

Beweis. Mit $S_k = \sum_{n=1}^{k} a_n$ ist $S_{k+1} - S_k = a_{k+1}$.
Die Behauptung folgt nach Satz 2.1.15 (Cauchykriterium). □

Bemerkung 2.4.1. Aus $\lim_{n\to\infty} a_n = 0$ folgt nicht die Konvergenz von $\sum_{n=1}^{\infty} a_n$, wie wir bald in Beispielen sehen werden.

Definition 2.4.1. (absolute Konvergenz)
Eine unendliche Reihe $\sum_{n=1}^{\infty} a_n$ heißt absolut konvergent, wenn $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

Satz 2.4.2. *(i) Eine absolut konvergente Reihe* $\sum_{n=1}^{\infty} a_n$ *ist konvergent und es ist*

$$\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|.$$

(ii) Die Reihe $\sum_{n=1}^{\infty} a_n$ *ist genau dann absolut konvergent, wenn die Folge der Partialsummen* $\sum_{m=1}^{n} |a_m|$ *beschränkt ist.*

Beweis. (i) Wir wenden das Cauchykriterium (Satz 2.1.15) auf die konvergente Reihe $\sum_{k=1}^{\infty} |a_k|$ an.

Es sei $\epsilon > 0$ gegeben. Dann existiert ein $n_0 = n_0(\epsilon)$, so daß $\sum_{n_1<k\leq n_2} |a_k| < \epsilon$ für alle $n_1, n_2 \geq n_0$.

Aus der Dreiecksungleichung folgt

$$\left|\sum_{n_1<k\leq n_2} a_k\right| \leq \sum_{n_1<k\leq n_2} |a_k| < \epsilon.$$

Aus der anderen Richtung des Cauchykriteriums folgt die Konvergenz von $\sum_{k=1}^{\infty} a_k$. Es seien $S_n := \sum_{k=1}^{n} a_k$ und $T_n := \sum_{k=1}^{n} |a_k|$. Nach der Dreiecksungleichung ist $|S_n| \leq T_n$. Nach Satz 2.1.8 (Erhaltung von Ungleichungen) folgt

$$\left|\sum_{k=1}^{\infty} a_k\right| = \left|\lim_{n\to\infty} S_n\right| \leq \lim_{n\to\infty} T_n = \sum_{k=1}^{\infty} |a_k|.$$

(ii) Die Folge (T_n) ist monoton wachsend. Die Behauptung folgt aus Satz 2.1.5 und Satz 2.1.12 (Monotoniekriterium).

□

Satz 2.4.3. *(Leibnizkriterium)*
Es sei $a_n \downarrow 0$. Dann ist $\sum_{n=0}^{\infty} (-1)^n a_n$ konvergent.

Beweis. Es sei $S_k = \sum_{n=0}^{k} (-1)^n a_n$. Weiter ist

$$\begin{aligned} S_{2m+1} &= S_{2m} + (-1)^{2m+1} a_{2m+1} < S_{2m} \\ S_{2m+3} &= S_{2m+1} + (-1)^{2m+2}(a_{2m+2} - a_{2m+3}) > S_{2m+1} \\ S_{2m+2} &= S_{2m} + (-1)^{2m+1}(a_{2m+2} - a_{2m+1}) < S_{2m} \end{aligned}$$

Wir setzen $I_m := [S_{2m+1}, S_{2m}]$, für das wegen den obigen Ungleichungen $I_{m+1} \subset I_m$ folgt.
Es ist $|I_m| = |S_{2m} - S_{2m+1}| = |a_{2m}| \downarrow 0$. Nach Definition 2.1.9 ist die Folge (I_m) eine Intervallschachtelung. Nach Satz 2.1.13 existiert genau ein $a \in \mathbb{R}$ mit $S_{2m+1} \uparrow a$ und $S_{2m} \downarrow a$. □

Satz 2.4.4. *(Majorantenkriterium)*
Es seien (a_n) und (b_n) Folgen mit $|a_n| \leq b_n$. Es sei $\sum_{n=1}^{\infty} b_n$ konvergent. Dann ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent, und es ist $\sum_{n=1}^{\infty} |a_n| \leq \sum_{n=1}^{\infty} b_n$.

Beweis. Für die Folge der Partialsummen $S_n := \sum_{m=1}^{n} |a_m|$ gilt:

$$S_n \leq \sum_{m=1}^{n} b_m \leq \sum_{m=1}^{\infty} b_m.$$

Damit ist (S_n) beschränkt und konvergiert nach Satz 2.1.12 (Monotoniekriterium). □

Definition 2.4.2. Die Reihe $\sum_{n=1}^{\infty} b_n$ heißt <u>konvergente Majorante</u> der Reihe $\sum_{n=1}^{\infty} a_n$.

Beispiel 2.4.1. Wir betrachten $\sum_{k=1}^{\infty} \frac{1}{k^2}$.

Aus der Ungleichung $k+1 \leq 2k$ folgt $\frac{1}{k^2} \leq \frac{2}{k(k+1)}$. Nach Beispiel 2.3.1 ist $\sum_{k=1}^{\infty} \frac{2}{k(k+1)} = 2$. Damit ist die Reihe $\sum_{k=1}^{\infty} \frac{2}{k(k+1)}$ eine konvergente Majorante von $\sum_{k=1}^{\infty} \frac{1}{k^2}$. Nach Satz 2.4.4 ist $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergent.

Bemerkung 2.4.2. Mit dem Majorantenkriterium kann die Frage der Konvergenz einer unendlichen Reihe entschieden werden. Der Wert der Reihe kann jedoch nicht bestimmt werden. In Beispiel 2.4.1 ist der Wert der Majorante $\sum_{k=1}^{\infty} \frac{2}{k(k+1)} = 2$ sehr einfach zu bestimmen, nicht jedoch der Wert von $\sum_{k=1}^{\infty} \frac{1}{k^2}$.

Satz 2.4.5. *(Minorantenkriterium)*
Es seien (a_n) und (b_n) Folgen mit $0 \leq a_n \leq b_n$. Es sei $\sum_{n=1}^{\infty} a_n$ divergent. Dann ist $\sum_{n=1}^{\infty} b_n$ divergent.

Beweis. Annahme: $\sum_{n=1}^{\infty} b_n$ konvergiert. Dann ist nach dem Majorantenkriterium $\sum_{n=1}^{\infty} a_n$ konvergent, ein Widerspruch. □

Definition 2.4.3. Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt divergente Minorante der Reihe $\sum_{n=1}^{\infty} b_n$.

Definition 2.4.4. Es sei (a_n) eine Folge und $a_n \geq 0$. Ist $\sum_{n=1}^{\infty} a_n$ konvergent, so schreibt man auch $\sum_{n=1}^{\infty} a_n < \infty$.
Ist $\sum_{n=1}^{\infty} a_n$ divergent, so schreibt man auch $\sum_{n=1}^{\infty} a_n = \infty$.

Zwei wichtige Kriterien, das Quotientenkriterium und das Wurzelkriterium erhält man, wenn man eine Reihe mit der geometrischen Reihe als Majorante oder Minorante vergleicht.

Satz 2.4.6. *(Quotientenkriterium)*
Es sei $(a_n)_{n=1}^{\infty}$ eine Folge mit $a_n \neq 0$.

(i) Ist $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, so ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.

(ii) Ist $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$, so ist $\sum_{n=1}^{\infty} a_n$ divergent.

Beweis. (i) Es sei $l := \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$. Dann ist $l = 1 - 2\epsilon$ mit $\epsilon > 0$. Es sei $q := 1 - \epsilon$. Nach Satz 2.1.18 gibt es ein n_0, so daß $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für $n \geq n_0$ ist. Durch vollständige Induktion nach k zeigt man, daß $|a_{n_0+k}| \leq |a_{n_0}| q^k$ für alle $k \in \mathbb{N}$ ist. Somit ist

$$\sum_{n=1}^{n_0+k} |a_n| \leq \sum_{n=1}^{n_0} |a_n| + \sum_{k=1}^{\infty} |a_{n_0}| q^k < \infty.$$

Die Folge der Partialsummen $\sum_{n=1}^{N} |a_n|$ ist also beschränkt. Damit ist $\sum_{n=1}^{\infty} |a_n|$ konvergent.

(ii) Aus $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$ folgt $|a_n| \geq |a_{n_0}|$ für alle $n \geq n_0$. Damit ist das Konvergenzkriterium $\lim_{n\to\infty} a_n = 0$ von Satz 2.4.1 nicht erfüllt, und damit ist $\sum_{n=1}^{\infty} a_n$ divergent.

□

Satz 2.4.7. *(Wurzelkriterium)*
Es sei $(a_n)_{n=1}^{\infty}$ eine Folge. Es sei $l := \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Dann gilt

(i) Wenn $l < 1$ ist, dann konvergiert $\sum_{n=1}^{\infty} a_n$ absolut.

(ii) Wenn $l > 1$ ist, dann divergiert $\sum_{n=1}^{\infty} a_n$.

Beweis. (i) Aus $l < 1$ folgt $l = 1 - 2\epsilon$ mit $\epsilon > 0$. Es sei $q := 1 - \epsilon$. Nach Satz 2.1.18 gibt es ein n_0, so daß $\sqrt[n]{|a_n|} \leq q$ für $n \geq n_0$ ist. Also ist $|a_n| \leq q^n$ für $n \geq n_0$. Somit ist

$$\sum_{n=1}^{N} |a_n| \leq \sum_{n=1}^{n_0} |a_n| + \sum_{n=n_0+1}^{\infty} q^n < \infty.$$

Die Folge der Partialsummen $\sum_{n=1}^{N} |a_n|$ ist also beschränkt. Damit ist $\sum_{n=1}^{\infty} a_n$ konvergent.

(ii) Aus $l > 1$folgt $l = 1 + \epsilon$ mit $\epsilon > 0$. Nach Satz 2.1.18 gibt es ein n_0, so daß $\sqrt[n]{|a_n|} \geq 1$ für $n \geq n_0$ ist. Damit ist das Konvergenzkriterium $\lim_{n\to\infty} a_n = 0$ von Satz 2.4.1 nicht erfüllt, und damit ist $\sum_{n=1}^{\infty} a_n$ divergent.

□

Es gibt Fälle, in denen weder das Quotienten- noch das Wurzelkriterium geeignet sind, die Frage der Konvergenz oder Divergenz einer unendlichen Reihe zu entscheiden. In diesen Fällen führen oft andere Kriterien zur Antwort. Wir werden später die sogenannten Integralkriterien kennenlernen.

Wir begnügen uns zunächst mit

Satz 2.4.8. *(Cauchyscher Verdichtungssatz)*

Es sei $(a_n)_{n=1}^{\infty}$ *monoton fallend und* $a_n \geq 0$. *Dann ist die Reihe* $\sum_{n=1}^{\infty} a_n$ *genau dann konvergent, wenn die (verdichtete) Reihe* $\sum_{n=0}^{\infty} 2^n a_{2^n}$ *konvergiert.*

Beweis. Wir schreiben die Partialsumme $S_{2^{N+1}} = \sum_{n=1}^{2^{N+1}} a_n$ mit $N \in \mathbb{N}$ in der Form

$$S_{2^{N+1}} = \sum_{n=0}^{N+1} S_{2^n} - \sum_{n=0}^{N} S_{2^n} = \sum_{n=0}^{N} (S_{2^{n+1}} - S_{2^n}) + S_1.$$

Es ist

$$S_{2^{n+1}} - S_{2^n} = \sum_{k=2^n+1}^{2^{n+1}} a_k \quad \begin{cases} \geq 2^n a_{2^{n+1}} \\ \leq 2^n a_{2^n+1} \end{cases}$$

Deshalb folgt, daß

$$S_{2^{N+1}} \geq \frac{1}{2} \sum_{n=0}^{N} 2^{n+1} a_{2^{n+1}} + a_1 = \frac{1}{2} \sum_{n=0}^{N+1} 2^n a_{2^n} + \frac{a_1}{2}$$

und

$$S_{2^{N+1}} \leq \sum_{n=0}^{N} 2^n a_{2^n+1} + a_1 \leq \sum_{n=0}^{N} 2^n a_{2^n} + a_1.$$

Also ist die Folge $(S_{2^{N+1}})_{N=0}^{\infty}$ genau dann beschränkt, wenn die Reihe $\sum_{n=0}^{\infty} 2^n a_{2^n}$ beschränkt ist. Weil die Folge $(S_k)_{k=1}^{\infty}$ monoton wächst, ist $(S_k)_{k=1}^{\infty}$ genau dann beschränkt, wenn $(S_{s^{N+1}})_{N=0}^{\infty}$ beschränkt ist.
Die Behauptung folgt nach Satz 2.4.2 *(ii)*. □

Definition 2.4.5. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ heißt harmonische Reihe.

Satz 2.4.9. *Die harmonische Reihe ist divergent:* $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

Beweis. Nach Satz 2.4.8 (Cauchyscher Verdichtungssatz) ist $\sum_{n=1}^{\infty} \frac{1}{n}$ genau dann divergent, wenn die "verdichtete Reihe" $\sum_{n=1}^{\infty} 2^n \frac{1}{2^n}$ divergiert. Dies ist der Fall. □

Bemerkung 2.4.3. Die harmonische Reihe liefert ein Beispiel dafür, daß die Umkehrung von Satz 2.4.1 nicht gilt: Setzen wir $a_n = \frac{1}{n}$, so ist $\lim_{n\to\infty} a_n = 0$, aber $\sum_{n=1}^{\infty} a_n$ divergiert.
Weiter läßt sich aus der harmonischen Reihe ein Beispiel für eine Reihe angeben, die konvergiert, aber nicht absolut konvergiert. Die "alternierende harmonische Reihe" $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ konvergiert nach

Satz 2.4.3 (Leibnizkriterium). Durch Bildung der Absolutbeträge entsteht jedoch die harmonische Reihe, welche divergiert.
Die Divergenz der harmonischen Reihe kann weder mit dem Quotienten- noch mit dem Wurzelkriterium entschieden werden, wie wir gleich sehen werden.

Wir wollen im folgenden Wurzel- und Quotientenkriterium vergleichen:
Als Vorbereitung beweisen wir

Satz 2.4.10. *Für $a > 0$ ist $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1$.*

Beweis. (i) Wir nehmen zunächst $a > 1$ an:
Es sei $\sqrt[n]{a} = 1+\epsilon_n$ mit $\epsilon_n > 0$. Nach der Bernoullischen Ungleichung ist $a = (1+\epsilon_n)^n \geq 1+n\epsilon_n$, also $\epsilon_n \leq \frac{a-1}{n}$. Damit ist $\lim\limits_{n\to\infty} \epsilon_n = 0$ und $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1$.

(ii) Ist $a < 1$, so wenden wir (i) mit $\frac{1}{a}$ an.

(iii) Für $a = 1$ ist die Behauptung klar.

□

Satz 2.4.11. *Es sei (a_n) eine Folge mit $a_n \geq 0$ für alle $n \in \mathbb{N}$. Dann gilt:*

(i) $\limsup\limits_{n\to\infty} \sqrt[n]{a_n} \leq \limsup\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$

(ii) $\liminf\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} \leq \liminf\limits_{n\to\infty} \sqrt[n]{a_n}$

(iii) Wenn der Grenzwert $\lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$ existiert, dann existiert auch der Grenzwert $\lim\limits_{n\to\infty} \sqrt[n]{a_n}$, und es gilt $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = \lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$.

Beweis. Wir beweisen nur (i):
Es sei $l := \limsup\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$ und $\delta > 0$. Nach Satz 2.1.18 gibt es ein $n_0 = n_0(\delta)$, so daß $\frac{a_{n+1}}{a_n} \leq q$ für alle $n \geq n_0$ ist. Durch vollständige Induktion folgt $\frac{a_{n_0+l}}{a_{n_0}} \leq q^l$ für alle $l \in \mathbb{N}$. Wir setzen $n := n_0 + l$ und erhalten $a_n \leq q^{n-n_0} a_{n_0}$. Es folgt damit $\sqrt[n]{a_n} \leq q \sqrt[n]{q^{-n_0} a_{n_0}}$. Wegen $\lim\limits_{n\to\infty} \sqrt[n]{q^{-n_0} a_{n_0}} = 1$ gibt es zu jedem $\epsilon > 0$ ein $n_1 = n_1(\epsilon)$, so daß $\sqrt[n]{a_n} \leq q + \epsilon$ für alle $n \geq n_1$ ist. □

Beispiel 2.4.2. Die unendliche Reihe $\sum\limits_{n=1}^{\infty} \dfrac{n}{2^n}$ ist konvergent, denn

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)2^n}{2^{n+1}n} = \frac{1}{2} \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = \frac{1}{2} < 1.$$

Beispiel 2.4.3. Es sei $a_n = \frac{1}{n}$ und $b_n = \frac{1}{n^2}$.
Weder Quotienten- noch Wurzelkriterium sind geeignet, die Frage der Konvergenz der (divergenten) Reihe $\sum\limits_{n=1}^{\infty} a_n$ oder der (konvergenten) Reihe $\sum\limits_{n=1}^{\infty} b_n$ zu entscheiden. Es ist

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n}{n+1} = 1, \quad \lim_{n\to\infty} \frac{b_{n+1}}{b_n} = \lim_{n\to\infty} \frac{n^2}{(n+1)^2} = 1$$

und auch $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = 1$ und $\lim\limits_{n\to\infty} \sqrt[n]{b_n} = 1$ (nach Satz 2.4.11 (i)).
Nach Satz 2.4.11 (i) ist $\limsup\limits_{n\to\infty} \sqrt[n]{a_n} \leq \limsup\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$.

Kann daher die Konvergenz einer Reihe mit Hilfe des Quotientenkriteriums nachgewiesen werden ($\limsup\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} < 1$), so auch mit dem Wurzelkriterium. Die Umkehrung gilt nicht, wie das folgende Beispiel zeigt. Das Wurzelkriterium ist also stärker als das Quotientenkriterium.

Beispiel 2.4.4. Es sei $m \in \mathbb{N}$ und $a_n = \begin{cases} 2^{-n}, & \text{für} \quad n = 2m \\ n \cdot 2^{-n}, & \text{für} \quad n = 2m+1 \end{cases}$. es ist

$$\frac{a_{2m+1}}{a_{2m}} = \frac{1}{2}(2m+1) \to \infty, \quad \text{also } \limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty.$$

Es ist aber

$$\lim_{n\to\infty} \sqrt[2m]{2^{-2m}} = \frac{1}{2} \quad \text{bzw.} \quad \lim_{n\to\infty} \sqrt[2m+1]{(2m+1)\cdot 2^{-(2m+1)}} = \frac{1}{2},$$

und damit

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \frac{1}{2} < 1.$$

Das Wurzelkriterium zeigt also die Konvergenz der unendlichen Reihe, während das Quotientenkriterium keine Antwort liefert.

2.5 Bedingte und unbedingte Konvergenz, Produktreihen

Definition 2.5.1. Unter der Umordnung einer unendlichen Reihe $\sum\limits_{n=1}^{\infty} a_n$ versteht man eine Reihe $\sum\limits_{k=1}^{\infty} a_{n_k}$, wobei n_k durch eine bijektive Abbildung $\tau\colon \mathbb{N} \to \mathbb{N}$, $k \to \tau(k) = n_k$ gegeben ist.

Beispiel 2.5.1. Es sei die alternierende harmonische Reihe $\sum\limits_{n=1}^{\infty} \dfrac{(-1)^{n+1}}{n} = \sum\limits_{n=1}^{\infty} a_n$ mit $a_n = \frac{(-1)^{n+1}}{n}$ gegeben. Die bijektive Abbildung $\tau\colon \mathbb{N} \to \mathbb{N}$, $k \to \tau(k) = n_k$ sei durch

$$\begin{aligned} \tau(3m) &= 2m \\ \tau(3m-2) &= 4m-3 \quad \text{und} \\ \tau(3m-1) &= 4m-1 \end{aligned}$$

für alle $m \in \mathbb{N}$ definiert. Dann haben wir die Umordnung

$$\sum_{k=1}^{\infty} a_{n_k} = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} \dots$$

Diese umgeordnete Reihe kann auch folgendermaßen erhalten werden: Es sei $s = \sum\limits_{n=1}^{\infty} a_n$, also

$$\begin{array}{rclllllllll} s & = & 1 & -\frac{1}{2} & +\frac{1}{3} & -\frac{1}{4} & +\frac{1}{5} & -\frac{1}{6} & +\frac{1}{7} & -\frac{1}{8} & \pm\dots \\ \frac{1}{2}s & = & & \frac{1}{2} & & -\frac{1}{4} & & +\frac{1}{6} & & -\frac{1}{8} & \pm\dots \end{array}$$

Dann ist

$$\frac{3}{2}s = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \pm \ldots = \sum_{n=1}^{\infty} a_{n_k}.$$

Die umgeordnete Reihe $\sum_{n=1}^{\infty} a_{n_k}$ hat also einen <u>anderen Wert</u> als die ursprüngliche Reihe $\sum_{n=1}^{\infty} a_n$.

Definition 2.5.2. Eine konvergente Reihe heißt <u>unbedingt konvergent</u>, wenn jede Umordnung von ihr gegen denselben Wert konvergiert, sonst <u>bedingt konvergent</u>.

Satz 2.5.1. *(Umordnungssatz)*
Eine Reihe ist genau dann unbedingt konvergent, wenn sie absolut konvergiert.

Beweis. ohne Beweis. □

Definition 2.5.3. (Cauchyprodukt)
Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ unendliche Reihen: das Cauchyprodukt ist dann durch $\sum_{n=0}^{\infty} c_n$, wobei $c_n = \sum_{k+l=n} a_k b_l = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{l=0}^{n} a_{n-l} b_l$ definiert ist.

Satz 2.5.2. *Sind $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergent, so ist dann das Cauchyprodukt $\sum_{n=0}^{\infty} c_n$ mit $c_n = \sum_{k+l=n} a_k b_l$ absolut konvergent, und es gilt die Cauchysche Produktformel*

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty} c_n.$$

Beweis. ohne Beweis. □

2.6 Dezimalbruchentwicklung

Definition 2.6.1. Es sei $g \in \mathbb{N}$ und $g \geq 2$. Unter der g- Bruchentwicklung von $a \in [0, \infty)$ versteht man eine Darstellung der Form

$$a = \sum_{m=0}^{n} a_m g^m + \sum_{n=1}^{\infty} b_n g^{-n}, \qquad (*)$$

wobei $a_m \in \{0, 1, \ldots, g-1\}$ und $b_n \in \{0, 1, \ldots, g-1\}$ mit $a_m \neq 0$ für $m > 0$ ist und folgende Form <u>ausgeschlossen ist</u>: $b_n = g - 1$ für alle $n \geq n_0$.

Bemerkung 2.6.1. Für $g = 10$ erhält man die vertraute Dezimalbruchentwicklung:
Für $a = \frac{1}{3}$ erhalten wir $a = \sum_{n=1}^{\infty} 3 \cdot 10^{-n} = 0{,}33\ldots = 0{,}\overline{3}$.
Es ist also in $(*)$ dann $g = 10$, $a_m = 0$ für alle $m \geq 0$ und $b_n = 3$ für alle $n \in \mathbb{N}$.

Satz 2.6.1. *Es sei $g \in \mathbb{N}$ und $g \geq 2$. Jedes $a \in [0, \infty)$ besitzt genau eine g- Bruchentwicklung der Form $(*)$.*

Kapitel 3

Stetigkeit, Differenzierbarkeit

3.1 Grenzwerte von Funktionen

Definition 3.1.1. (Häufungspunkt)
Es sei $\mathcal{D} \subseteq \mathbb{R}$. Dann heißt $\xi \in \mathbb{R}$ Häufungspunkt (HP) von $\mathcal{D}$, wenn in jeder Umgebung $U_\epsilon(\xi)$ ein $x \in \mathcal{D} \backslash \{\xi\}$ liegt.
Ein $\xi \in \mathcal{D}$, das nicht Häufungspunkt von $\mathcal{D}$ ist, heißt isolierter Punkt.

Bemerkung 3.1.1. Ein Häufungspunkt von $\mathcal{D}$ braucht nicht zu $\mathcal{D}$ gehören. Ist $\mathcal{D} = (a, b)$ mit $a < b \in \mathbb{R}$, so sind a und b Häufungspunkte von $\mathcal{D}$, gehören aber nicht zu $\mathcal{D}$.

Definition 3.1.2. Es sei $\mathcal{D} \subseteq \mathbb{R}$, $f \colon \mathcal{D} \to \mathbb{R}$, $x \to f(x)$ eine Funktion und x_0 ein Häufungspunkt von $\mathcal{D}$.
Dann heißt $a \in \mathbb{R}$ Grenzwert der Funktion f an der Stelle x_0 (Schreibweise: $\lim\limits_{x \to x_0} f(x) = a$ oder $f(x) \to a$, $(x \to x_0)$), falls gilt, daß für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ existiert, so daß $|f(x) - a| < \epsilon$ für alle x mit $0 < |x - x_0| < \delta$ gilt.

Ein sehr wichtiges Hilfsmittel zum Beweis von Aussagen über Grenzwerte von Funktionen ist das Folgenkriterium. Dadurch können die aus Kapitel 2 bekannten Tatsachen über die Grenzwerte von Folgen zur Anwendung gebracht werden.

Satz 3.1.1. *(Folgenkriterium)*
Es sei $\mathcal{D} \subseteq \mathbb{R}$, $f \colon \mathcal{D} \to \mathbb{R}$ und x_0 ein HP von $\mathcal{D}$. Dann sind folgende Aussagen äquivalent:

(i) $\lim\limits_{x \to x_0} f(x) = a$

(ii) Für jede Folge $(z_n)_{n=1}^{\infty}$ mit $z_n \in \mathcal{D} \backslash \{x_0\}$ und $\lim\limits_{n \to \infty} z_n = x_0$ gilt:

$$\lim_{n \to \infty} f(z_n) = a.$$

Beweis. $(i) \Rightarrow (ii)$:
Es sei $(z_n)_{n=1}^{\infty}$ eine Folge mit $z_n \in \mathcal{D} \backslash \{x_0\}$ und $\lim\limits_{n \to \infty} z_n = x_0$.
Es sei $\epsilon > 0$ gegeben. Nach Definition 3.1.2 existiert ein $\delta = \delta(\epsilon)$, so daß für alle $x \in \mathcal{D}$ mit $0 < |x - x_0| < \delta$ gilt:

$$|f(x) - a| < \epsilon. \qquad (*)$$

Nach Definition 2.1.1 existiert ein $n_0 = n_0(\delta)$, so daß für alle $n \geq n_0$ gilt: $|z_n - x_0| < \delta$. Wegen $(*)$ ist dann auch $|f(z_n) - a| < \epsilon$ für alle $n \geq n_0$. Nach Definition 2.1.1 bedeutet dies

$$\lim_{n\to\infty} f(z_n) = a.$$

$(ii) \Rightarrow (i)$:
Annahme: Die Aussage $\lim\limits_{x\to x_0} f(x) = a$ ist falsch.
Dann existiert ein $\epsilon > 0$, so daß für alle $n \in \mathbb{N}$ ein z_n mit $|z_n - x_0| < \frac{1}{n}$ mit $z_n \in \mathcal{D}\backslash\{x_0\}$ und $|f(z_n) - a| > \epsilon$ existiert. Dann ist nach Definition 2.1.1 die Aussage $\lim\limits_{n\to\infty} f(z_n) = a$ falsch, ein Widerspruch. □

Satz 3.1.2. *(Eindeutigkeit des Grenzwerts)*
Es sei $f\colon \mathcal{D} \to \mathbb{R}$ und x_0 ein HP von $\mathcal{D}$. Dann hat die Funktion f höchstens einen Grenzwert an der Stelle a.

Beweis. Dies folgt aus dem Folgenkriterium (Satz 3.1.1) und der Eindeutigkeit des Grenzwerts für Folgen (Satz 2.1.1). □

Satz 3.1.3. *(Grenzwertsätze)*
Es sei $\mathcal{D} \subseteq \mathbb{R}$, x_0 ein HP von $\mathcal{D}$ und $f, g\colon \mathcal{D} \to \mathbb{R}$ Funktionen. Die Grenzwerte $\lim\limits_{x\to x_0} f(x)$ und $\lim\limits_{x\to x_0} g(x)$ mögen existieren. Dann existieren auch die Grenzwerte

$$\lim_{x\to x_0} (f(x) + g(x)) \quad \textit{und} \quad \lim_{x\to x_0} (f(x) \cdot g(x)),$$

und es ist

$$\begin{aligned} \lim_{x\to x_0} (f(x) + g(x)) &= \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x) \\ \lim_{x\to x_0} (f(x) \cdot g(x)) &= \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x). \end{aligned}$$

Ist $\lim\limits_{x\to x_0} g(x) \neq 0$, so gibt es eine Umgebung $U_\delta(x_0)$, so daß $g(x) \neq 0$ für alle $x \in \mathcal{D} \cap U_\delta(x_0)$. Es sei $h = \frac{f}{g}|_{\{x|\ g(x)\neq 0\}\cap\mathcal{D}}$. Dann ist

$$\lim_{x\to x_0} h(x) = \frac{\lim\limits_{x\to x_0} f(x)}{\lim\limits_{x\to x_0} g(x)}.$$

Beweis. Es sei $b := \lim\limits_{x\to x_0} g(x) \neq 0$. Wir wenden Definition 3.1.2 mit $\epsilon = \frac{|b|}{2}$ an. Dann existiert ein $\delta > 0$, so daß für alle $x \in U_\delta(x_0)$ gilt: $|g(x) - b| < \frac{|b|}{2}$ und damit $g(x) \neq 0$.
Alle anderen Aussagen werden bewiesen, indem man das Folgenkriterium und die entsprechenden Grenzwertsätze für Folgen (Satz 2.1.7) anwendet. □

3.2 Stetigkeit

Definition 3.2.1. Es sei $\mathcal{D} \subseteq \mathbb{R}$, $f\colon \mathcal{D} \to \mathbb{R}$ und $x_0 \in \mathcal{D}$. Dann heißt f stetig im Punkt x_0, wenn für alle $\epsilon > 0$ ein $\delta = \delta(x_0, \epsilon) > 0$ gibt, so daß $|f(x) - f(x_0)| < \epsilon$ für alle $x \in \mathcal{D}$ mit $|x - x_0| < \delta$.

Bemerkung 3.2.1. Ist x_0 ein isolierter Punkt von $\mathcal{D}$, so ist jede auf $\mathcal{D}$ definierte Funktion f stetig in x_0.

Satz 3.2.1. *Es sei $f\colon \mathcal{D} \to \mathbb{R}$ und $x_0 \in \mathcal{D}$ ein HP von $\mathcal{D}$. Dann ist f im Punkt x_0 genau dann stetig, wenn $\lim\limits_{x\to x_0} f(x) = f(x_0)$ ist.*

Beweis. Das folgt unmittelbar aus Definition 3.1.2 für den Grenzwert und aus Definition 3.2.1 für die Stetigkeit. □

Bemerkung 3.2.2. Der Wert der Funktion im Punkt x_0, nämlich $f(x_0)$, spielt bei der Bestimmung des Grenzwerts $\lim\limits_{x\to x_0} f(x)$ keine Rolle, wohl aber bei der Frage, ob f im Punkt x_0 stetig ist. Der Grenzwert $\lim\limits_{x\to x_0} f(x)$ kann selbst dann existieren, wenn $f(x_0)$ gar nicht definiert ist. Aber f ist stetig in x_0, wenn $\lim\limits_{x\to x_0} f(x)$ existiert und mit dem Wert der Funktion $f(x_0)$ übereinstimmt.

Beispiel 3.2.1. Wir betrachten drei Beispiele einander sehr ähnlicher Funktionen:

a) Es sei $f\colon \mathcal{D}_1 \to \mathbb{R}$, $x \to 2x$ mit $\mathcal{D}_1 = \mathbb{R}$. Aus dem Folgenkriterium für Grenzwerte folgt, daß $\lim\limits_{x\to 0} f(x) = 0$ ist. Daher ist f stetig im Punkt 0, da $\lim\limits_{x\to 0} f(x) = f(0)$ ist.

b) Nun sei $g\colon \mathcal{D}_2 \to \mathbb{R}$, $x \to 2x$ mit $\mathcal{D}_2 = \mathbb{R}\backslash\{0\}$. Es ist $0 \notin \mathcal{D}_2$, aber 0 ist Häufungspunkt von $\mathcal{D}_2$. Da $f|_{\mathcal{D}_2} = g$, folgt $\lim\limits_{x\to 0} f(x) = \lim\limits_{x\to 0} g(x) = 0$. Aber g ist <u>nicht stetig</u> in 0, da 0 nicht zum Definitionsbereich von g gehört. Allerdings kann g durch die Definition $g(0) = 0$ <u>stetig im Punkt 0 fortgesetzt</u> werden. Damit wird der Definitionsbereich $\mathcal{D}_2$ zu $\mathcal{D}_1$ erweitert und g mit f identifiziert.

c) Schließlich sei $h\colon \mathcal{D}_1 \to \mathbb{R}$, $x \to h(x)$ mit $h(x) = \begin{cases} 2x & \text{für} \quad x \neq 0 \\ 1 & \text{für} \quad x = 0. \end{cases}$
Es ist $\lim\limits_{x\to 0} h(x) = 0 \neq h(0)$. Also ist h nicht stetig in $x = 0$.

Aus dem Folgenkriterium für Grenzwerte läßt sich leicht das Folgenkriterium für Stetigkeit gewinnen:

Satz 3.2.2. *(Folgenkriterium für Stetigkeit)*
Es sei $\mathcal{D} \subset \mathbb{R}$ und x_0 ein HP von $\mathcal{D}$. Dann sind folgende Aussagen äquivalent:

(i) Die Funktion f ist stetig in x_0.

(ii) Für jede Folge $(z_n)_{n=1}^{\infty}$ mit $z_n \in \mathcal{D}\backslash\{x_0\}$ und $\lim\limits_{n\to\infty} z_n = x_0$ gilt: $\lim\limits_{n\to\infty} f(z_n) = f(x_0)$.

Beweis. Das folgt unmittelbar aus dem Folgenkriterium für Grenzwerte (Satz 3.1.1) und der Definition der Stetigkeit (Definition 3.2.1). □

Bemerkung 3.2.3. Satz 3.2.2 gibt ein Beispiel für die <u>Vertauschbarkeit zweier Operationen</u>: Die eine Operation ist die <u>Grenzwertoperation</u>. Von einer Folge (a_n) wird der Grenzwert $\lim_{n\to\infty} a_n$ gebildet. Die zweite Operation ist die Bildung des Funktionswertes $x \to f(x)$. Satz 3.2.2 sagt, daß bei <u>stetigen Funktionen f</u> diese Operationen vertauscht werden können. Es ist $\lim\limits_{n\to\infty} f(z_n) = f\left(\lim\limits_{n\to\infty} z_n\right)$. Bei unstetigen Funktionen ist eine Vertauschung i.a. nicht möglich. Für die Funktion aus Beispiel 3.2.1 c) $h(x) = \begin{cases} 2x & \text{für} \quad x \neq 0 \\ 1 & \text{für} \quad x = 0 \end{cases}$ und die Folge (z_n) mit $z_n = \frac{1}{n}$ haben wir

$$\lim_{n\to\infty} h(z_n) = \lim_{n\to\infty} 2 \cdot \frac{1}{n} = 0 \quad \text{aber } h\left(\lim_{n\to\infty} z_n\right) = h(0) = 1.$$

Satz 3.2.3. *Es sei $\mathcal{D} \subset \mathbb{R}$ und $x_0 \in \mathcal{D}$. Die Funktionen $f, g\colon \mathcal{D} \to \mathbb{R}$ seien im Punkt x_0 stetig. Dann sind auch die Funktionen $f+g$ und $f \cdot g$ in x_0 stetig. Ist $g(x_0) \neq 0$, so gibt es eine Umgebung $U_\delta(x_0)$, so daß $g(x) \neq 0$ für alle $x \in \mathcal{D} \cap U_\delta(x_0)$. Es sei $h = \frac{f}{g}|_{\{x|\ g(x)\neq 0\}\cap\mathcal{D}}$. Dann ist h im Punkt x_0 stetig.*

Beweis. Mann kann zunächst annehmen, daß x_0 ein HP von $\mathcal{D}$ ist, da Funktionen in isolierten Punkten ihres Definitionsbereichs immer stetig sind. Nach Voraussetzung ist $\lim\limits_{x\to x_0} f(x) = f(x_0)$ und $\lim\limits_{x\to x_0} g(x) = g(x_0)$. Nach Satz 3.1.3 (Grenzwertsätze) folgt $\lim\limits_{x\to x_0}(f(x)+g(x)) = f(x_0)+g(x_0)$ und damit die Stetigkeit von $f+g$ in x_0.
Die anderen Teile der Behauptung folgen ebenfalls aus Satz 3.1.3. □

Wir kommen nun zur Komposition stetiger Funktionen:

Satz 3.2.4. *(Kettenregel für stetige Funktionen)*
Es seien $\mathcal{D}, \mathcal{E} \subset \mathbb{R}$ und $f\colon \mathcal{D} \to \mathcal{E}$ sowie $g\colon \mathcal{E} \to \mathbb{R}$. Es sei f in $x_0 \in \mathcal{D}$ stetig und g stetig in $y_0 = f(x_0) \in \mathcal{E}$. Dann ist die Funktion $g \circ f\colon \mathcal{D} \to \mathbb{R}$, $x \to (g \circ f)(x) = g(f(x))$ in x_0 stetig.

Beweis. Wir können annehmen, daß x_0 ein HP von f ist. Es sei (z_n) eine Folge mit $\lim\limits_{n\to\infty} z_n = x_0$. Nach dem Folgenkriterium für Stetigkeit (Satz 3.2.2) ist $\lim\limits_{n\to\infty} f(z_n) = f(x_0) = y_0$. Wiederum nach dem Folgenkriterium, angewandt auf g, gilt für die Folge $(f(z_n))$ dann $\lim\limits_{n\to\infty} g(f(z_n)) = g(y_0) = g(f(x_0))$. Also gilt für jede Folge (z_n) mit $\lim\limits_{n\to\infty} z_n = x_0$ dann

$$\lim_{n\to\infty}(g\circ f)(z_n) = (g\circ f)\left(\lim_{n\to\infty} z_n\right) = (g\circ f)(x_0).$$

Nach dem Folgenkriterium ist $g \circ f$ in x_0 stetig. □

3.3 Einseitige und uneigentliche Grenzwerte, einseitige Stetigkeit

Definition 3.3.1. Es sei $\mathcal{D} \subset \mathbb{R}$, $f\colon \mathcal{D} \to \mathbb{R}$ und $x_0 \in \mathbb{R}$ ein HP von $\mathcal{D}$.

(i) So heißt $a \in \mathbb{R}$ rechtsseitiger Limes von f an der Stelle x_0 (Schreibweise: $\lim\limits_{x\to x_0^+} f(x) = a$), falls x_0 HP von $\mathcal{D} \cap (x_0, \infty)$ ist, und wenn die Restiktion

$$g := f|_{\mathcal{D}\cap(x_0,\infty)}\colon \mathcal{D} \cap (x_0, \infty) \to \mathbb{R}$$

den Limes a besitzt, d.h. wenn $\lim\limits_{x\to x_0} g(x) = a$ gilt.

(ii) Analog wird der linksseitige Limes von f an der Stelle x_0 definiert.
(Schreibweise: $\lim\limits_{x\to x_0^-} f(x) = a$).

(iii) Um Grenzwerte von einseitigen Grenzwerten zu unterscheiden, nennt man Grenzwerte im Sinne von Definition 3.1.2 auch beidseitige Grenzwerte.

Satz 3.3.1. *Es sei $f\colon \mathcal{D} \to \mathbb{R}$ und x_0 sei HP sowohl von $\mathcal{D} \cap (x_0, \infty)$, als auch von $\mathcal{D} \cap (-\infty, x_0)$. Dann existiert $\lim\limits_{x\to x_0} f(x)$ genau dann, wenn der rechtsseitige und der linksseitige Limes von f an der Stelle x_0 existieren und übereinstimmen. In diesem Fall ist*

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x).$$

Beweis. ohne Beweis. □

Definition 3.3.2. (i) Eine Funktion $f\colon \mathcal{D} \to \mathbb{R}$ heißt im Punkt $x_0 \in \mathcal{D}$ rechtsseitig stetig, wenn die Restiktion $g := f|_{\mathcal{D}\cap[x_0,\infty)}$ in x_0 stetig ist.

(ii) Analog wird linksseitige Stetigkeit definiert.

Satz 3.3.2. *Es sei $f\colon \mathcal{D} \to \mathbb{R}$ und $x_0 \in \mathcal{D}$. Dann ist f in x_0 genau dann stetig, wenn es in x_0 rechts- und linksseitig stetig ist.*

Beweis. ohne Beweis. □

Es besteht nun noch die Möglichkeit, (ein- oder beidseitige) Grenzwerte für $x \to \infty$ oder $x \to -\infty$ zu definieren, bzw. ∞ oder $-\infty$ als Grenzwert zuzulassen.

Definition 3.3.3. (i) Es sei $X \subset \mathbb{R}$. Dann heißt ∞ (bzw $-\infty$) uneigentlicher HP von X, wenn in jeder Umgebung $U_c(\infty) = (c, \infty)$ (bzw. $U_c(-\infty) = (-\infty, c)$) ein $x \in X$ liegt.

(ii) Es sei $f\colon \mathcal{D} \to \mathbb{R}$ und ∞ ein HP von $\mathcal{D}$. Dann ist $\lim\limits_{x\to\infty} f(x) = a$ für $a \in \mathbb{R}$, falls für alle $\epsilon > 0$ ein $c = c(\epsilon)$ existiert, so daß $|f(x) - a| < \epsilon$ für alle x mit $x \in U_c(\infty)$ ist.

(iii) Analog wird $\lim\limits_{x\to-\infty} f(x)$ definiert.

Definition 3.3.4. (i) Es sei $\mathcal{D} \subset \mathbb{R}$, $f\colon \mathcal{D} \to \mathbb{R}$ und x_0 ein HP von $\mathcal{D}$. Man sagt $\lim\limits_{x\to x_0} f(x) = \infty$, falls für alle $c > 0$ ein $\delta = \delta(c) > 0$ existiert, so daß $f(x) > c$, d.h. $f(x) \in U_c(\infty)$ für alle x mit $|x - x_0| < \delta$ gilt.

(ii) Es sei $\mathcal{D} \subset \mathbb{R}$ und x_0 ein HP von $\mathcal{D}\cap(x_0, \infty)$. Man sagt $\lim\limits_{x\to x_0^+} f(x) = \infty$, falls für $g := f|_{\mathcal{D}\cap(x_0,\infty)}$ gilt: $\lim\limits_{x\to x_0} g(x) = \infty$.

(iii) Analog werden die Beziehungen

$$\lim_{x\to x_0} f(x) = -\infty, \quad \lim_{x\to x_0^+} f(x) = -\infty, \quad \lim_{x\to x_0^-} f(x) = \pm\infty \quad \text{bzw.} \quad \lim_{x\to\pm\infty} f(x) = \pm\infty$$

definiert.

Beispiel 3.3.1. Es sei $\mathcal{D} = \mathbb{R}\backslash\{0\}$ und $f\colon \mathcal{D} \to \mathbb{R}$, $x \to f(x) = \frac{1}{x}$. Für alle $c > 0$ gilt dann $f(x) = \frac{1}{x} > c \Leftrightarrow 0 < x < c$ und $f(x) = \frac{1}{x} < -c \Leftrightarrow -c^{-1} < x < 0$. Also ist

$$\lim_{x\to 0^+} \frac{1}{x} = \infty \quad \text{und} \quad \lim_{x\to 0^-} \frac{1}{x} = -\infty.$$

3.4 Polynome und rationale Funktionen

Definition 3.4.1. Es sei $\mathcal{D} \subset \mathbb{R}$.

(i) Eine Funktion $P\colon \mathcal{D} \to \mathbb{R}$, $x \to P(x)$ mit $P(x) = \sum\limits_{k=0}^{n} a_k x^k$ mit $a_k \in \mathbb{R}$ und $0 \le k \le n$ heißt Polynom. Die a_k heißen Koeffizienten des Polynoms. Ist $a_n \neq 0$, so heißt n der Grad des Polynoms, $n = gradP$. Ist $a_k = 0$ für $0 \le k \le n$, so heißt P das Nullpolynom und man setzt $gradP := -\infty$.

(ii) Es seien $P, Q\colon \mathcal{D} \to \mathbb{R}$ Polynome. Es sei $E \subset \{x \in \mathcal{D} |\ Q(x) \neq 0\} \neq \emptyset$. Dann heißt

$$R\colon E \to \mathbb{R}, \quad x \to R(x) = \frac{P(x)}{Q(x)}$$

rationale Funktion.

Satz 3.4.1. *Polynome und rationale Funktionen sind in jedem Punkt ihres Definitionsbereiches stetige Funktionen.*

Beweis. (i) Zunächst zeigen wir die Aussage für Polynome:
Wir können annehmen, daß $x_0 \in \mathcal{D}$ ein HP von $\mathcal{D}$ ist. Konstante Funktionen c mit $c(x) = c$ für alle $x \in \mathcal{D}$ sind nach dem Folgenkriterium (Satz 3.2.2) in $x_0 \in \mathcal{D}$ stetig. Für jede Folge (z_n) mit $z_n \in \mathcal{D}$ und $\lim\limits_{n\to\infty} z_n = x_0$ ist $c(z_n) = c$ für alle n und damit $\lim\limits_{n\to\infty} c(z_n) = c = c(x_0)$.
Ebenso folgt die Stetigkeit der Identität: $id\colon x \to id(x) = x$.
Mit vollständiger Induktion folgt nach Satz 3.2.3, daß die Monome $M\colon x \to x^k$ stetig sind. Nach Satz 3.2.3 sind dann auch die konstanten Vielfachen $x \to c_k x^k$ stetig. Durch vollständige Induktion (nach der Anzahl der Summanden) folgt schließlich die Stetigkeit von $P(x)$.

(ii) Die Aussage für rationale Funktionen folgt aus Satz 3.2.3 (ii).

□

Beispiel 3.4.1. Es sei f durch $f(x) = \begin{cases} x^2 & \text{für} \quad x \in (-\infty, 1] \\ 2x - 1 & \text{für} \quad x \in (1, \infty] \end{cases}$ definiert. In welchen $x_0 \in \mathbb{R}$ ist f stetig?

Lösung:
Nach Satz 3.4.1 sind $f_{|(-\infty,1)}$ bzw. $f_{|(1,\infty)}$ in allen $x_0 \in (-\infty, 1)$ bzw. $x_0 \in (1, \infty)$ stetig. Also ist f in $\mathbb{R}\backslash\{1\}$ stetig. In allen $x_0 \neq 1$ existieren also die beidseitigen Grenzwerte und stimmen mit dem Funktionswert $f(x_0)$ überein. Also ist f stetig für $x_0 \neq 1$. Weiter ist

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

Somit ist $\lim\limits_{x\to 1^-} f(x) = \lim\limits_{x\to 1^+} f(x) = \lim\limits_{x\to 1} f(x) = f(1)$.
Also ist f auch im Punkt $x_0 = 1$ stetig. Damit ist f in allen $x_0 \in \mathbb{R}$ stetig.

Beispiel 3.4.2. Es sei g durch $g(x) = \begin{cases} 7x & \text{für} \quad x \in (-\infty, 2) \\ x^3 - 4 & \text{für} \quad x \in [2, \infty] \end{cases}$ definiert. In welchen $x_0 \in \mathbb{R}$ ist g stetig?

Lösung:
Nach Satz 3.4.1 sind $g_{|(-\infty,2)}$ bzw. $g_{|(2,\infty)}$ in allen $x_0 \in (-\infty, 2)$ bzw. $x_0 \in (2, \infty)$ stetig. Also ist g in $\mathbb{R}\backslash\{2\}$ stetig. Es ist

$$\begin{aligned} \lim_{x\to 2^-} g(x) &= \lim_{x\to 2^-} 7x = 14 \\ \lim_{x\to 2^+} g(x) &= \lim_{x\to 2^+} (x^3 - 4) = 8 - 4 = 4. \end{aligned}$$

Also existiert $\lim\limits_{x\to 2} g(x)$ nicht. Damit ist g in $x_0 = 2$ nicht stetig.

Wir kommen nun zur Frage, inwieweit die Koeffizienten und der Grad eines Polynoms eindeutig bestimmt sind.

Beispiel 3.4.3. Es sei $\mathcal{D} = \{0, 1\}$, $f : \mathcal{D} \to \mathbb{R}$, $x \to f(x)$ mit $f(x) = \begin{cases} 0 & \text{für} \quad x = 0 \\ 1 & \text{für} \quad x = 1. \end{cases}$ So kann f auf unendlich viele Arten als Polynom geschrieben werden. Für jedes $n \in \mathbb{N}$ stimmt das Polynom $P_n : \mathcal{D} \to \mathbb{R}$, $x \to x^n$ mit f überein. Die Differenz $P_{m,n} := g_n - g_m : x \to x^n - x^m$ stellen stets das Nullpolynom auf $\mathcal{D}$ dar, 0 und 1 sind Nullstellen des Polynoms $P_{m,n}$.

Diese Vieldeutigkeit hat ihre Ursache darin, daß der Definitionsbereich $\mathcal{D}$ sehr klein ist. Wir wollen als erstes prüfen, wieviele Nullstellen ein Polynom haben kann.

Definition 3.4.2. Es sei $f : \mathcal{D} \to \mathbb{R}$. Ein $x_0 \in \mathcal{D}$ mit $f(x_0) = 0$ heißt Nullstelle von f.
Ist $f(x) = c$ für alle $x \in \mathcal{D}$, so schreiben wir auch $f(x) \equiv c$ (auf $\mathcal{D}$).
Das Polynom $P : \mathcal{D} \to \mathbb{R}$ mit $P(x) \equiv 0$ (auf $\mathcal{D}$) heißt Nullpolynom (auf $\mathcal{D}$).

Satz 3.4.2. *Es sei $n \in \mathbb{N}$ und $P : \mathcal{D} \to \mathbb{R}$, $x \to P(x)$ ein Polynom mit $P(x) = \sum_{k=0}^{n} a_k x^k$ mit $a_k \in \mathbb{R}$ und $a_n \neq 0$. Für $x_0 \in \mathcal{D}$ gibt es dann ein Polynom $Q(x) = \sum_{k=0}^{n-1} b_k x^k$ mit $b_k \in \mathbb{R}$ und $b_{n-1} = a_n$, so daß*

$$P(x) = (x - x_0)Q(x) + P(x_0) \quad \textit{für alle } x \in \mathcal{D} \quad \textit{gilt.} \tag{$*$}$$

Ist insbesondere x_0 eine Nullstelle von P, so ist $P(x) = (x - x_0)Q(x)$ für alle $x \in \mathcal{D}$.

Beweis. Wir führen vollständige Induktion nach n durch:
Induktionsanfang $n = 1$:
Dann haben wir das Polynom $P(x) = a_1 x + a_0$ mit $a_1 \neq 0$ gegeben.
Es ist $P(x) = a_1(x - x_0) + a_1 x_0 + a_0$. Es gilt also $(*)$ mit $Q(x) = a_1$.
Induktionsschritt $n \to n + 1$:
Es gelte die Induktionshypothese, und es sei $P(x) = \sum_{k=0}^{n+1} a_k x^k$ mit $a_k \in \mathbb{R}$. Dann ist mit $R(x) = \sum_{k=0}^{n} a_{k+1} x^k$

$$P(x) = (x - x_0)R(x) + x_0 R(x) + a_0. \tag{1}$$

Nach Induktionshypothese gibt es $c_i \in \mathbb{R}$, so daß mit $S(x) = \sum_{k=0}^{n-1} c_k x^k$ gilt:

$$R(x) = (x - x_0)S(x) + R(x_0). \tag{2}$$

Aus (1) und (2) folgt

$$P(x) = (x - x_0)Q(x) + P(x_0) \quad \text{mit } Q(x) := R(x) + x_0 S(x).$$

□

Satz 3.4.3. *Es sei $n \in \mathbb{N}$. Ein Polynom vom Grad n hat höchstens n Nullstellen.*

Beweis. Es seien $x_1, \ldots, x_n \in \mathbb{R}$ verschiedene Nullstellen von $P(x)$. Wir zeigen zunächst durch vollständige Induktion nach k:

$$\text{Es ist } P(x) = (x - x_1) \cdots (x - x_k) \cdot Q_{n-k}(x) \text{ mit einem Polynom } Q_{n-k} \text{ vom Grad } n - k. \tag{1}$$

Induktionsanfang $k = 1$:
Dies ist Satz 3.4.2.
Induktionsschritt $k \to k + 1$:
Nach Induktionshypothese ist

$$P(x) = (x - x_1) \cdots (x - x_k) \cdot Q_{n-k}(x) \text{ mit einem Polynom } Q_{n-k} \text{ vom Grad } n - k. \tag{2}$$

Nach Satz 1.4.8 ist

$$\prod_{l=1}^{k} (x_{k+1} - x_l) \neq 0$$

und damit $Q_{n-k}(x_{k+1}) = 0$ ebenfalls nach Satz 1.4.8. Nach Satz 3.4.2 ist

$$Q_{n-k}(x) = (x - x_{k+1}) Q_{n-(k+1)}(x) \tag{3}$$

mit $grad Q_{n-(k+1)}(x) = n - (k + 1)$. Aus (2) und (3) folgt

$$P(x) = \prod_{l=1}^{k+1} (x - x_l) Q_{n-(k+1)}(x),$$

womit (1) gezeigt ist.
Für $k = n$ ergibt sich

$$P(x) = a_n \prod_{l=1}^{n} (x - x_l)$$

mit $a_n \in \mathbb{R} \backslash \{0\}$. Wiederum nach Satz 1.4.8 folgt

$$P(x) = 0 \Leftrightarrow x \in \{x_1, \ldots, x_n\}.$$

□

Satz 3.4.4. *(Identitätssatz für Polynome)*
Es sei $\mathcal{D} \subset \mathbb{R}$ und $P, Q \colon \mathcal{D} \to \mathbb{R}$ Polynome. Mit $n \in \mathbb{N}$ und $a_k, b_k \in \mathbb{R}$ seien $P(x) = \sum_{k=0}^{n} a_k x^k$ und $Q(x) = \sum_{k=0}^{n} b_k x^k$ gegeben. Weiterhin gebe es $(n + 1)$ Elemente $x_1, \ldots, x_{n+1} \in \mathcal{D}$ mit $P(x_j) = Q(x_j)$ für $1 \leq j \leq n + 1$. Dann ist $a_k = b_k$ für $0 \leq k \leq n$. Ist $P(x_j) = 0$ für $1 \leq j \leq n + 1$, so ist P das Nullpolynom.
Insbesondere sind der Grad und die Koeffizienten eines Polynoms durch seine Werte auf einer unendlichen Menge eindeutig bestimmt. Ist $P(x) = 0$ für unendlich viele Werte von x, so ist P das Nullpolynom.

Beweis. Man wendet Satz 3.4.3 auf das Polynom $P - Q$ an. □

Satz 3.4.2 ist ein Spezialfall des Euklidischen Algorithmus:

Satz 3.4.5. *(Euklidischer Algorithmus)*
Es sei $\mathcal{D} \subset \mathbb{R}$, $|\mathcal{D}| = \infty$ und $P, Q \not\equiv 0$ zwei Polynome auf $\mathcal{D}$. Dann gibt es zwei eindeutig bestimmte Polynome S, R mit $grad R < grad Q$, so daß $P(x) = S(x)Q(x) + R(x)$ für alle $x \in \mathcal{D}$ ist.

Beweis. ohne Beweis. □

Beispiel 3.4.4. Es sei $P(x) = 3x^5 + x^3 + 1$ und $Q(x) = x^3 + 2$. Division ergibt

$$3x^5 + x^3 + 1 = (x^3 + 2) \cdot (3x^2 + 1) + (-6x^2 - 1).$$

Es gelten der Faktorisierungssatz und der Satz von der Partialbruchzerlegung.

Satz 3.4.6. *(Faktorisierungssatz)*
Es sei P ein Polynom vom Grad $n \geq 1$ mit paarweise verschiedenen Nullstellen $x_1, \ldots, x_k \in \mathbb{R}$ und den Vielfachheiten $\nu_1, \ldots, \nu_k \in \mathbb{N}$. Ist $\nu_1 + \ldots + \nu_k < n$, so gibt es eindeutig bestimmte, paarweise verschiedene normierte Polynome $P_1(x) = x^2 + b_1 x + c_1, \ldots, P_l(x) = x^2 + b_l x + c_l$ vom Grad 2, welche keine reelle Nullstelle besitzen, d.h. es gilt $4c_j - b_j^2 > 0$ für $1 \leq j \leq l$, und es gibt eindeutig bestimmte Zahlen $\mu_1, \ldots, \mu_l \in \mathbb{N}$, so daß P die <u>Produktdarstellung</u>

$$P(x) = a_n (x - x_1)^{\nu_1} \ldots (x - x_k)^{\nu_k} (P_1(x))^{\mu_1} \ldots (P_l(x))^{\mu_l}$$

besitzt. Es gilt $\nu_1 + \ldots + \nu_k + 2(\mu_1 + \ldots + \mu_l) = n$.

Beweis. ohne Beweis. □

Satz 3.4.7. *(Partialbruchzerlegung)*
Es seien P, Q Polynome mit grad$P <$ grad$Q$. Außerdem habe Q die (nach Satz 3.4.6 existierende) Produktdarstellung

$$Q(x) = (x - x_1)^{\nu_1} \cdots (x - x_k)^{\nu_k} \cdot (x^2 + b_1 x + c_1)^{\mu_1} \cdots (x^2 + b_l x + c_l)^{\mu_l}.$$

Dann besitzt R eine <u>Partialbruchdarstellung</u> der Form

$$R(x) = \sum_{i=1}^{k} \left(\frac{A_i^{(1)}}{x - x_i} + \ldots + \frac{A_i^{(\nu_i)}}{(x - x_i)^{\nu_i}} \right) + \sum_{j=1}^{l} \left(\frac{B_j^{(1)} x + C_j^{(1)}}{x^2 + b_j x + c_j} + \ldots + \frac{B_j^{(\mu_j)} x + C_j^{(\mu_j)}}{(x^2 + b_j x + c_j)^{\mu_j}} \right)$$

mit $A_i^{(1)}, \ldots, A_i^{(\nu_i)}$, $i = 1, \ldots, k$ und $B_j^{(1)}, C_j^{(1)}, \ldots, B_j^{(\mu_j)}, C_j^{(\mu_j)}$, $j = 1, \ldots, l$.

Wir schließen mit der Diskussion von Grenzwerten von Polynomen und rationalen Funktionen:

Satz 3.4.8. *(i) Es sei $P\colon \mathbb{R} \to \mathbb{R}$ ein Polynom vom Grad n, $P(x) := \sum_{k=0}^{n} a_k x^k$ mit $a_k \in \mathbb{R}$ und $a_n = 1$. Dann ist*

$$\lim_{x \to \infty} P(x) = \infty \quad \text{und} \quad \lim_{x \to -\infty} P(x) = \begin{cases} \infty, & \text{falls } n \text{ gerade} \\ -\infty, & \text{falls } n \text{ ungerade ist.} \end{cases}$$

*(ii) Es sei $R(x) := \dfrac{P(x)}{Q(x)}$ mit Polynomen P, Q mit grad$P <$ grad$Q$.
Dann ist $\lim_{x \to \pm\infty} R(x) = 0$.*

(iii) Es sei $x_0 \in \mathbb{R}$ und $n \in \mathbb{N}$. Dann ist für gerades n

$$\lim_{x \to x_0} (x - x_0)^{-n} = \infty$$

und für ungerades n

$$\lim_{x \to x_0^+} (x - x_0)^{-n} = \infty \quad \text{und} \quad \lim_{x \to x_0^-} (x - x_0)^{-n} = -\infty$$

(iv) Es seien R, S rationale Funktionen und $x_0 \in \mathbb{R} \cup \{-\infty, \infty\}$. Es sei $\lim\limits_{x \to x_0} R(x) = \pm\infty$ und $\lim\limits_{x \to x_0} S(x) = c \in \mathbb{R} \backslash \{0\}$. Es sei $T(x) := R(x) + S(x)$ und $U(x) := R(x) \cdot S(x)$. Dann ist

$$\lim_{x \to x_0} T(x) = \lim_{x \to x_0} R(x) \quad \textit{und} \; \lim_{x \to x_0} U(x) = \begin{cases} \lim_{x \to x_0} R(x), & \textit{falls} \quad c > 0 \\ -\lim_{x \to x_0} R(x), & \textit{falls} \quad c < 0. \end{cases}$$

Dabei ist $-(-\infty) = \infty$ gesetzt.
Für einseitige Grenzwerte gelten entsprechende Ausagen.

Beweis. ohne Beweis. □

Die Behandlung von uneigentlichen (ein- oder beidseitigen) Grenzwerten rationaler Funktionen oder Grenzwerte für $x \to \pm\infty$ kann zusammen mit dem Euklidischen Algorithmus, Faktorisierung und Partialbruchzerlegung auf Satz 3.4.8 zurückgeführt werden.

3.5 Stetige Funktionen auf kompakten Intervallen

Definition 3.5.1. Es sei $f : \mathcal{D} \to \mathbb{R}$ eine Funktion, $\mathcal{E} \subset \mathcal{D}$. Dann heißt f stetig auf $\mathcal{E}$, falls f in jedem Punkt $x \in \mathcal{E}$ stetig ist. Weiter heißt f nach oben beschränkt auf $\mathcal{E}$, falls es eine obere Schranke $s \in \mathbb{R}$ gibt, so daß $f(x) \leq s$ für alle $x \in \mathcal{E}$ gilt. Entsprechend wird "nach unten beschränkt" definiert. Man nennt f dann beschränkt auf $\mathcal{E}$, falls es nach oben und nach unten beschränkt ist.

Satz 3.5.1. *(Stetigkeit auf kompakten Intervallen)*
Es sei $a \leq b \in \mathbb{R}$, $I = [a, b]$ und $f : I \to \mathbb{R}$ stetig auf I. Dann gilt:

(i) Die Funktion f ist beschränkt auf I.

(ii) Die Funktion f nimmt auf I Minimum und Maximum an, d.h. es existieren $m, M \in \mathbb{R}$ und $x_{\min}, x_{\max} \in I$, so daß $f(x_{\min}) = m$, $f(x_{\max}) = M$ und $m \leq f(x) \leq M$ für alle $x \in [a, b]$ gilt.

Beweis. Es sei $M \in \mathbb{R} \cup \{\infty\}$ das Supremum des Bildes $f(I) = \{f(x)|\ x \in I\}$. Dann gibt es eine Folge $(x_n)_{n=1}^{\infty}$, so daß $\lim\limits_{n \to \infty} f(x_n) = M$ ist. Ist $M = \infty$, so ist der Limes uneigentlich. Nach dem Satz von Bolzano- Weiterstraß (Satz 2.1.14) hat die Folge $(x_n)_{n=1}^{\infty}$ mindestens einen Häufungswert z_0. Damit ist z_0 Häufungspunkt von I. Also enthält I alle seine Häufungspunkte, somit ist $z_0 \in I$. Nach Satz 2.1.10 gibt es eine Teilfolge $(x_{n_k})_{n=1}^{\infty}$ von (x_n) mit $\lim\limits_{n \to \infty} x_{n_k} = z_0$. Nach dem Folgenkriterium für Stetigkeit (Satz 3.2.2) ist $f(z_0) = M$. Damit folgt $M \in \mathbb{R}$, d.h. $M \neq \infty$. Hiermit ist die Beschränktheit nach oben und die Existenz des Maximums gezeigt. Benützt man diese Tatsachen für $-f$ anstelle von f, so folgt die Existenz des Minimums und die Beschränktheit nach unten. □

Definition 3.5.2. Es sei $f : \mathcal{D} \to \mathbb{R}$ eine Funktion, $\mathcal{E} \subset \mathcal{D}$. Dann heißt f gleichmäßig stetig auf $\mathcal{E}$, falls für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon)$ existiert, so daß $|f(x) - f(x_0)| < \epsilon$ für alle $x, x_0 \in \mathcal{D}$ mit $|x - x_0| < \delta$ gilt.

Bemerkung 3.5.1. Jede auf $\mathcal{E}$ gleichmäßig stetige Funktion ist dort auch stetig. Dies folgt aus Definition 3.2.1. Die Zahl $\delta = \delta(\epsilon) > 0$ wird dann im allgemeinen jedoch nicht nur von ϵ, sondern auch von x_0 abhängen. Ist $\mathcal{E}$ kein kompaktes Intervall, so braucht eine auf $\mathcal{E}$ stetige Funktion nicht gleichmäßig stetig zu sein, wie das folgende Beispiel zeigt:

Beispiel 3.5.1. Es sei $\mathcal{D} = \mathcal{E} = (0,1)$ und $f\colon \mathcal{D} \to \mathbb{R}$, $x \to f(x) = \frac{1}{x}$. Dann ist f auf $\mathcal{E}$ stetig, aber nicht gleichmäßig stetig.
Es sei $\epsilon > 0$ gegeben. Es ist

$$|f(x) - f(x_0)| = \left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x - x_0|}{|x||x_0|}$$

und damit

$$|f(x) - f(x_0)| \geq \epsilon \Leftrightarrow |x - x_0| \geq \epsilon|x||x_0|.$$

Dies zeigt, daß für beliebige $\delta > 0$ die Aussage $|f(x) - f(x_0)| < \epsilon$ höchstens dann gilt, wenn $|x - x_0| < \epsilon|x_0|$ ist. Also ist $|f(x) - f(x_0)| < \epsilon$ für $|x - x_0| < \delta(\epsilon, x_0)$ mit $\delta(\epsilon, x_0) < \epsilon|x_0|$ erfüllt, also für kein $\delta > 0$, das nur von ϵ abhängt.

Auf kompakten Intervallen sind stetige Funktionen hingegen gleichmäßig stetig. Dieses Ergebnis ist eine Folgerung des Überdeckungssatzes von Heine- Borel.
Zunächst definieren wir den Begriff der Überdeckung:

Definition 3.5.3. Es seien $I, X \subset \mathbb{R}$. Für alle $x \in X$ sei $U(x) := U_{\epsilon(x)}(x)$ mit $\epsilon(x) > 0$ eine $\epsilon(x)$-Umgebung von x. Die Menge $\mathcal{U} = \{U_\epsilon(x) |\ x \in X\}$ heißt eine Überdeckung von I, wenn $I \subseteq \bigcup\limits_{x \in X} U(x)$.
Man nennt $\mathcal{V}$ eine Teilüberdeckung von $\mathcal{U}$, falls $\mathcal{V} = \{U(x) |\ x \in Y\}$ mit $Y \subset X$ und $I \subseteq \bigcup\limits_{x \in Y} U(x)$ ist.
Ist dazu Y endlich, so heißt $\mathcal{V}$ eine endliche Teilüberdeckung von $\mathcal{U}$.

Satz 3.5.2. *(Überdeckungssatz von Heine- Borel)*
Es sei $\mathcal{U}$ eine Überdeckung des kompakten Intervalls I. Dann besitzt $\mathcal{U}$ eine endliche Teilüberdeckung.

Beweis. Es sei $I = [a, b]$.
Annahme: Es gibt eine Überdeckung $\mathcal{U}$ von I, die keine endliche Teilüberdeckung besitzt.
Es sei $\mathcal{U} := \{U(x) |\ x \in X\}$ mit $X \subset \mathbb{R}$. Wir definieren nun durch vollständige Induktion eine Folge (I_n) von Intervallen mit folgenden Eigenschaften:

(i) Es bildet (I_n) eine Intervallschachtelung.

(ii) Es gilt: $|I_n| = |I| \cdot 2^{-n}$.

(iii) Die Überdeckung $\mathcal{U}$ von I besitzt keine endliche Teilüberdeckung.

Induktionsanfang: $n = 1$:
Dann ist $I_1 = I$.

Induktionsschritt: $n \to n + 1$:
Nach dem dritten Teil der Induktionshypothese besitzt die Überdeckung $\mathcal{U}$ von $I_n := [a_n, b_n]$ keine endliche Teilüberdeckung, und es ist $|I_n| = |I| \cdot 2^{-n}$. Dann ist $\mathcal{U}$ auch eine Überdeckung für die zwei Hälften $I_{n,1} := \left[a_n, \frac{a_n+b_n}{2}\right]$ und $I_{n,2} := \left[\frac{a_n+b_n}{2}, b_n\right]$. Für mindestens eine der zwei Hälften $I_{n,j}$ mit $j \in \{1, 2\}$ existiert dann ebenfalls keine endliche Teilüberdeckung von $\mathcal{U}$. Dann setzen wir $I_{n+1} := I_{n,j}$. Damit erfüllt die Folge (I_n) die Bedingungen (i), (ii) und (iii). Nach Satz 2.1.3 existiert ein $z_0 \in \mathbb{R}$, das allen Intervallen angehört. Da $\mathcal{U}$ eine Überdeckung von I ist, gibt es $x \in X$ mit $z_0 \in U(x)$. Dann gibt es ein $\epsilon > 0$, so daß $U_\epsilon(z_0) \subset U(x)$ ist. Weiter existiert ein n_0, so daß für alle $n \geq n_0$ gilt: $I_n \subset U_\epsilon(z_0)$. Diese I_n werden aber von der einzigen Umgebung $U_\epsilon(z_0)$ überdeckt, im Widerspruch dazu, daß für die keine endliche Teilüberdeckung von $\mathcal{U}$ existiert. □

Satz 3.5.3. *Es sei I ein kompaktes Intervall und $f\colon I \to \mathbb{R}$ stetig auf I. Dann ist f auf I gleichmäßig stetig.*

Beweis. Es sei $\epsilon > 0$. Zu jedem $x_0 \in I$ gibt es ein $\delta = \delta(\epsilon, x_0) > 0$, so daß für alle $x \in U_{3\delta}$ gilt:

$$|f(x) - f(x_0)| < \frac{\epsilon}{3}. \tag{1}$$

Wir setzen

$$U(x_0) = U_\delta(x_0). \tag{2}$$

Es bildet $\mathcal{U} = \{U(x_0)|\ x_0 \in I\}$ eine Überdeckung von I. Nach Satz 3.5.2 besitzt $\mathcal{U}$ eine endliche Teilüberdeckung. Es gilt also $x_1, x_2, \ldots, x_m \in I$, so daß es für alle $x \in I$ ein k mit $1 \le k \le m$ und $x \in U(x_k) = U_\delta(\epsilon, x_k)$ gibt. Wir setzen $\delta := \min\{\delta(\epsilon, x_1), \ldots, \delta(\epsilon, x_m\}$. Es seien nun $x, x' \in I$ mit $|x - x'| < \delta$. Dann gibt es k, l mit $1 \le k, l \le m$ mit $x \in U(x_k)$ und $x' \in U(x_l)$. Es ist (nach der Dreiecksungleichung, Satz 1.4.10) dann

$$|x_k - x_l| \le |x_k - x| + |x - x'| + |x' - x_l| < 3\delta.$$

Wegen (1) und (2) folgt $|f(x_k) - f(x_l)| < \frac{\epsilon}{3}$. Schließlich ist

$$|f(x) - f(x')| \le |f(x) - f(x_k)| + |f(x_k) - f(x_l)| + |f(x_l) - f(x')| < \epsilon.$$

Also ist $|f(x) - f(x')| < \epsilon$ für alle $x, x' \in I$ mit $|x - x'| < \delta$, die gleichmäßige Stetigkeit. □

Satz 3.5.4. *(Zwischenwertsatz)*
Es sei $a < b \in \mathbb{R}$, $I = [a, b]$, und $f\colon I \to \mathbb{R}$ sei stetig auf I. Es sei $f(a) < c < f(b)$. Dann gibt es ein $\xi \in (a, b)$ mit $f(\xi) = c$, d.h. es gilt $[f(a), f(b)] \subset f(I)$.

Beweis. Wir betrachten die Menge

$$X := \{z|\ a \le z \le b : f(x) \le c,\ \forall x \in [a, z]\}.$$

Es ist $X \ne \emptyset$, da $a \in X$. Wegen $X \subset [a, b]$ ist X beschränkt. Also existiert $s = \sup X$.

(i) Wir zeigen zunächst: $f(s) \le c$.
Es sei $x_n = s - \frac{s-a}{n}$. Wegen $a \le x_n \le s$ ist $f(x_n) \le c$. Wegen $\lim\limits_{n\to\infty} x_n = s$ ist nach dem Folgenkriterium $\lim\limits_{n\to\infty} f(x_n) = f(s)$ und wegen $f(x_n) \le c$ ist $f(s) = \lim\limits_{n\to\infty} f(x_n) \le c$.
Damit ist (i) gezeigt.

(ii) Annahme: $f(s) < c$.
Dann gibt es ein $\epsilon > 0$, so daß $f(s) = c - 2\epsilon$ gilt. Wegen der Stetigkeit von f existiert ein $\delta = \delta(a)$, so daß für alle $x \in U_\delta(s) = (s - \delta, s + \delta)$ gilt: $|f(x) - f(s)| < \epsilon$ und somit $|f(x)| < c - \epsilon$. Damit ist aber $[a, s + \delta) \cap [a, b] \subset X$ im Widerspruch zu $\sup X = s$. Also ist $f(s) \ge c$.

Zusammengefaßt folgt $f(s) = c$. □

3.6 Monotone Funktionen, Umkehrfunktion

Definition 3.6.1. Es sei $\mathcal{D} \subset \mathbb{R}$ und $f\colon \mathcal{D} \to \mathbb{R}$.
Dann heißt f monoton wachsend, wenn $f(x_1) \leq f(x_2)$ für alle $x_1, x_2 \in \mathcal{D}$ mit $x_1 \leq x_2$ ist, und f heißt streng monoton wachsend, wenn $f(x_1) < f(x_2)$ gilt. Entsprechend wird monoton fallend bzw. streng monoton fallend definiert.

Satz 3.6.1. *Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Es sei $f\colon I \to \mathbb{R}$ stetig und $f(a) < f(b)$. Dann sind folgende Aussagen äquivalent:*

(i) Die Funktion f besitzt eine auf $[f(a), f(b)]$ definierte Inverse f^{-1}, d.h. die Gleichung $f(x) = c$ besitzt für alle $c \in [f(a), f(b)]$ eine eindeutig bestimmte Lösung $x = f^{-1}(c)$.

(ii) Die Funktion f ist injektiv.

(iii) Die Funktion f ist streng monoton wachsend.

Beweis. ohne Beweis. □

Satz 3.6.2. *Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Es sei $f\colon I \to \mathbb{R}$ eine stetige, streng monoton wachsende Funktion auf dem kompakten Intervall I.*
Dann ist die Inverse $f^{-1}\colon [f(a), f(b)] \to [a, b]$ eine stetige, streng monoton wachsende Funktion auf $[f(a), f(b)]$.

Beweis. ohne Beweis. □

Satz 3.6.3. *Es sei $n \in \mathbb{N}$. Die n- te Wurzelfunktion $g\colon [0, \infty) \to \mathbb{R}$, $x \to g(x) = \sqrt[n]{x}$ ist auf $[0, \infty)$ stetig.*

Beweis. Dies folgt aus Satz 3.6.2, wenn wir anstatt f dann die streng monoton wachsende Funktion $f\colon [0, b] \to \mathbb{R}$, $x \to x^n$ wählen. Sie hat die Inverse $f^{-1}\colon [0, b^n] \to \mathbb{R}$, $x \to \sqrt[n]{x}$. Die Behauptung folgt mit $g := f^{-1}$, da b beliebig groß gewählt werden kann. □

3.7 Differenzierbarkeit

Definition 3.7.1. (i) Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall. Dann heißt $f\colon I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar, falls der Grenzwert

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Er heißt Ableitung oder Differentialquotient von f an der Stelle x_0.
Schreibweise: $f'(x_0) = \dfrac{df}{dx}(x_0) = \dfrac{dy}{dx}$.

(ii) Ist f für jedes $x_0 \in I$ differenzierbar, so heißt f differenzierbar auf I.
Die Funktion $f'\colon I \to \mathbb{R}$, $x \to f'(x)$ heißt die Ableitung von f.

Satz 3.7.1. *Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, $f\colon I \to \mathbb{R}$ und $x_0 \in I$. Folgende Aussagen sind äquivalent:*

(i) Die Funktion f besitzt an der Stelle x_0 die Ableitung $f'(x_0)$.

(ii) Es existiert eine Funktion $r\colon I \to \mathbb{R}$, $x \to r(x)$, wobei r in x_0 stetig und $r(x_0) = 0$ ist, so daß

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)(x - x_0). \qquad (*)$$

Beweis. $(i) \to (ii)$:
Wir setzen

$$\tilde{r}(x) := \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0).$$

Dann ist $(*)$ für $x \neq x_0$ erfüllt, und es ist

$$\lim_{x \to x_0} \tilde{r}(x) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) = 0.$$

Wir definieren

$$r(x) = \begin{cases} \tilde{r}(x) & \text{für} \quad x \neq x_0 \\ 0 & \text{für} \quad x = x_0. \end{cases}$$

$(ii) \to (i)$:
Es gilt

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \big(f'(x_0) + r(x)\big) = f'(x_0).$$

□

Bemerkung 3.7.1. Satz 3.7.1 besagt, daß für kleine Werte von $|x - x_0|$ die Funktion f sehr gut durch die lineare Funktion (Polynom 1. Grades) $L_f\colon x \to L_f(x)$ approximiert wird. Dabei ist $L_f(x)$ die lineare Approximation von f. Wir werden später Approximationen durch Polynome höheren Grades, sogenannte Taylorpolynome kennenlernen. Der Graph von L_f ist eine Gerade , die Tangente an der Kurve $y = f(x)$ im Punkt $(x_0, f(x_0))$. Weiter ist $f'(x_0)$ die Steigung der Tangente. Sie ist der Grenzwert der Differenzenquotienten

$$\frac{f(x) - f(x_0)}{x - x_0},$$

der Steigungen der Sekanten der Kurve $y = f(x)$ durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$.
Eine gröbere Approximation ist durch die konstante Funktion $C_f(x) = f(x_0)$ (Polynom nullten Grades) gegeben. Ihr Graph ist die horizontale Gerade durch $(x_0, f(x_0))$. Es ist $f(x) = C_f(x) + s(x)$ mit $\lim\limits_{x \to x_0} s(x) = 0$. Diese existiert auch für stetige Funktionen.

Satz 3.7.2. *(Eindeutigkeit der linearen Approximation)*
Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, $f\colon I \to \mathbb{R}$ und $x_0 \in I$. Gilt $f(x) = f(x_0) + c(x - x_0) + r(x)(x - x_0)$ mit $c \in \mathbb{R}$, r stetig in x_0 und $r(x_0) = 0$, so ist f differenzierbar in x_0, und es ist $c = f'(x_0)$.

Beweis. Aus $f(x) = f(x_0) + c(x - x_0) + r(x)(x - x_0)$ folgt

$$\frac{f(x) - f(x_0)}{x - x_0} = c + r(x).$$

Also ist

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = c + \lim_{x \to x_0} r(x) \equiv c.$$

Nach Definition 3.7.1 ist $c = f'(x_0)$. □

Satz 3.7.3. *(Differenzierbarkeit impliziert Stetigkeit)*
Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, $f\colon I \to \mathbb{R}$ und $x_0 \in I$. Ist f in x_0 differenzierbar, so ist es dort auch stetig.

Beweis. Aus $(*)$ aus Satz 3.7.1 folgt

$$\lim_{x \to x_0} f(x) = f(x_0) + \lim_{x \to x_0} f'(x_0)(x - x_0) + \lim_{x \to x_0} r(x)(x - x_0) = f(x_0)$$

nach Satz 3.1.3. Nach Satz 3.2.1 folgt die Stetigkeit von f in x_0. $\square$

Bemerkung 3.7.2. Die Umkehrung von Satz 3.7.3 gilt nicht. Die Betragsfunktion $f\colon \mathbb{R} \to \mathbb{R}, x \to |x|$ ist in $x_0 = 0$ stetig, aber nicht differenzierbar. Es ist

$$\begin{aligned} \lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} &= \lim_{x \to 0^+} \frac{x}{x} = 1, \quad \text{aber} \\ \lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} &= \lim_{x \to 0^-} \frac{-x}{x} = -1. \end{aligned}$$

Also existiert $\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0}$ nicht.

3.8 Ableitungsregeln

Satz 3.8.1. *Es sei $I \subset \mathbb{R}$ ein Intervall. Die Funktionen $f, g\colon I \to \mathbb{R}$ seien in $x_0 \in \mathbb{R}$ differenzierbar. Es seien $a, b \in \mathbb{R}$. Dann sind die Funktionen $af + bg$ und $f \cdot g$ in x_0 differenzierbar, und es gilt*

(i) $(af + bg)'(x_0) = af'(x_0) + bg'(x_0)$ *(Linearität)*

(ii) $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$ *(Produktregel)*

Beweis. Wir beweisen nur (ii):
Für $x \in I$ mit $x \neq x_0$ gilt

$$\begin{aligned} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x)g(x) - f(x_0)g(x)}{x - x_0} + \frac{f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0) \frac{g(x) - g(x_0)}{x - x_0}. \end{aligned}$$

Also ist

$$\begin{aligned} \lim_{x \to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \lim_{x \to x_0} g(x) + \lim_{x \to x_0} f(x_0) \lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

$\square$

Satz 3.8.2. *(Ableitung von Polynomen)*
Es sei $n \in \mathbb{N}$, und $P\colon \mathbb{R} \to \mathbb{R}$, $x \to P(x) = \sum_{k=0}^{n} a_k x^k$ mit $a_k \in \mathbb{R}$ sei ein Polynom.
Dann ist P für alle $x \in \mathbb{R}$ differenzierbar, und es ist $P'(x) = \sum_{k=1}^{n} k a_k x^{k-1}$.
Spezialfall: Es ist $\frac{d}{dx} x^n = n \cdot x^{n-1}$.

Beweis. Wir betrachten zunächst die Ableitung der Identität $id\colon x \to x$. Es ist

$$id'(x_0) = \lim_{x\to x_0} \frac{id(x) - id(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1.$$

Der Spezialfall ergibt sich dann durch vollständige Induktion nach n.
Der allgemeine Fall folgt aus der Linearität. □

Satz 3.8.3. *Es sei $I \subset \mathbb{R}$ ein Intervall, und es seien $f, g\colon I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar. Außerdem sei $g(x) \neq 0$ für $x \in I$. Dann ist $\frac{f}{g}$ in x_0 differenzierbar, und es gilt die Quotientenregel*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

Beweis. Für $x, x_0 \in I$ und $x \neq x_0$ gilt

$$\frac{\frac{f}{g}(x) - \frac{f}{g}(x_0)}{x - x_0} = \frac{f(x)g(x_0) - f(x_0)g(x)}{(x - x_0)g(x)g(x_0)} = \left(\frac{f(x) - f(x_0)}{x - x_0} \cdot g(x_0) - f(x_0) \cdot \frac{g(x) - g(x_0)}{x - x_0}\right) \frac{1}{g(x)g(x_0)}.$$

Der Grenzübergang $x \to x_0$ liefert die Behauptung. □

Satz 3.8.4. *Es seien $I, J \subset \mathbb{R}$ Intervalle, $f\colon I \to J$ sei im Punkt $x_0 \in I$ differenzierbar und $g\colon J \to \mathbb{R}$ sei im Punkt $y_0 = f(x_0) \in J$ differenzierbar. Dann ist $g \circ f\colon I \to \mathbb{R}$ im Punkt x_0 differenzierbar, und es gilt die Kettenregel*

$$(g \circ f)'(x_0) = g'(y_0) \cdot f'(x_0).$$

Beweis. Nach Satz 3.7.1 ist für alle $x \in I$

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)(x - x_0) \tag{1}$$

mit r stetig in x_0 sowie $r(x_0) = 0$ und

$$g(y) = g(y_0) + g'(y_0)(y - y_0) + s(y)(y - y_0). \tag{2}$$

für alle $y \in J$ mit s stetig in y_0 und $s(y_0) = 0$.
Indem wir in (2) dann $y = f(x)$ setzen, erhalten wir aus (1)

$$g(f(x)) = g(f(x_0)) + g'(y_0)(f'(x_0)(x - x_0) + r(x)(x - x_0)) + s(f(x))(f'(x_0)(x - x_0) + r(x)(x - x_0)). \tag{3}$$

Es sei

$$t(x) := g'(y_0)r(x) + s(f(x))(f'(x_0) + r(x)).$$

Die Funktion f ist stetig in $x = x_0$ und wegen $s(f(x_0)) = s(y_0) = 0$ folgt $t(x_0) = 0$. Also folgt aus (3)

$$(g \circ f)(x) = (g \circ f)(x_0) + g'(y_0)f'(x_0)(x - x_0) + t(x)(x - x_0)$$

mit t stetig in $x = x_0$ und $t(x_0) = 0$. Aus Satz 3.7.1 und 3.7.2 folgt $g \circ f$ ist differenzierbar in x_0, und es ist $(g \circ f)'(x_0) = g'(y)f'(x_0)$. □

Satz 3.8.5. *(Differenzierbarkeit der Umkehrfunktion)*
Es seien I, J Intervalle. Es sei $f\colon I \to J$ bijektiv und $x_0 \in I$. Es sei f im Punkt x_0 differenzierbar, und es sei $f'(x_0) \neq 0$. Dann ist die inverse Funktion $f^{-1}\colon J \to I$ in $y_0 = f(x_0)$ differenzierbar, und es ist

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Beweis. Es sei $(y_n)_{n=1}^{\infty}$ mit $y_n \in J \backslash \{y_0\}$ und $\lim_{n\to\infty} y_n = y_0$. Die Folge $(x_n)_{n=1}^{\infty}$ sei durch

$$y_n = f(x_n) \Leftrightarrow x_n = f^{-1}(y_n)$$

definiert. Wegen der Stetigkeit von f^{-1} ist dann

$$\lim_{n\to\infty} x_n = f^{-1}\left(\lim_{n\to\infty} y_n\right) = f^{-1}(y_0) = x_0.$$

nach dem Folgenkriterium ist

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \left(\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}\right)^{-1} = \frac{1}{f'(x_0)}.$$

Wieder ist nach dem Folgenkriterium

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)}.$$

□

Beispiel 3.8.1. Die Funktion $f\colon (0,\infty) \to (0,\infty)$, $x \to x^2$ besitzt die Umkehrfunktion

$$f^{-1}\colon (0,\infty) \to (0,\infty), \quad y \to \sqrt{y}.$$

Nach Satz 3.8.5 ist für $y_0 \in (0,\infty)$ dann

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(f^{-1}(y_0))} = \frac{1}{2\sqrt{y}}.$$

Also ist, wenn wir wieder x statt y schreiben:

$$\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}.$$

3.9 Mittelwertsatz, Monotonie

Definition 3.9.1. Es sei I ein Intervall. Für das Innere von I schreiben wir $\overset{\circ}{I}$.
Es sei $f\colon I \to \mathbb{R}$ und $\xi \in I$. Man sagt: f besitzt in ξ ein relatives Maximum (bzw. relatives Minimum), falls es ein $\delta > 0$ gibt, so daß $f(x) \le f(\xi)$ (bzw. $f(x) \ge f(\xi)$) für alle $x \in U_\delta(\xi)$, (d.h. für alle $x \in I$ mit $|x - \xi| < \delta$), gibt. Die Funktion f besitzt in ξ ein relatives Extremum, falls es dort ein relatives Maximum oder Minimum besitzt.

Satz 3.9.1. *Es sei I ein Intervall, $f\colon I \to \mathbb{R}$ und $\xi \in \overset{\circ}{I}$. Zudem sei f in ξ differenzierbar. Besitzt f in ξ ein relatives Extremum, so ist $f'(\xi) = 0$.*

Beweis. Wir können, falls wir nötigenfalls f durch $-f$ ersetzen, annehmen, daß f in ξ ein relatives Maximum besitzt. Es ist dann

$$f'(\xi) = \lim_{x\to\xi^+} \frac{f(x) - f(\xi)}{x - \xi} \le 0, \tag{1}$$

da $\dfrac{f(x) - f(\xi)}{x - \xi} \le 0$ für $x > \xi$ ist, und

$$f'(\xi) = \lim_{x\to\xi^-} \frac{f(x) - f(\xi)}{x - \xi} \ge 0, \tag{2}$$

da $\dfrac{f(x) - f(\xi)}{x - \xi} \ge 0$ für $x < \xi$ ist.
Aus (1) und (2) folgt $f'(\xi) = 0$. □

Satz 3.9.2. *(Satz von Rolle)*
Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Die Funktion $f\colon I \to \mathbb{R}$ sei auf I stetig und auf $\mathring{I} = (a, b)$ differenzierbar. Es sei $f(a) = f(b) = 0$. Dann gibt es ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.

Beweis. Nach Satz 3.5.1 nimmt f auf I sein Maximum M und sein Minimum m an. Ist $f \equiv 0$, so ist $f'(\xi) = 0$ für alle $\xi \in (a, b)$. Andernfalls ist $M > 0$ oder $m < 0$. Wir nehmen $f(\xi) = M > 0$ an. Wegen $f(a) = f(b) = 0$ ist $\xi \notin \{a, b\}$, also ist $\xi \in \mathring{I} = (a, b)$. Nach Satz 3.9.1 ist $f'(\xi) = 0$. □

Satz 3.9.3. *(1. Mittelwertsatz)*
Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Weiter sei $f\colon I \to \mathbb{R}$ auf I stetig und auf $\mathring{I}$ differenzierbar. Dann gibt es $\xi \in (a, b)$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Es sei

$$g(x) := f(x) - f(a) - (x - a)\frac{f(b) - f(a)}{b - a}.$$

Dann ist $g(a) = g(b) = 0$. Nach dem Satz von Rolle (Satz 3.9.2) gibt es ein $\xi \in (a, b)$ mit

$$g'(\xi) = 0 \Leftrightarrow f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

□

Bemerkung 3.9.1. Der Satz von Rolle und der 1. Mittelwertsatz lassen sich geometrisch so deuten, daß es mindestens einen Punkt ξ im Inneren des Intervalls $[a, b]$ gibt, so daß die Tangente an die Kurve $y = f(x)$ in $(\xi, f(\xi))$ parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist.

Satz 3.9.4. *(2. Mittelwertsatz)*
Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Es seien $f, g\colon I \to \mathbb{R}$ auf $[a, b]$ stetig und auf (a, b) differenzierbar. Es sei $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist $g(a) \neq g(b)$, und es gibt ein $\xi \in (a, b)$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Beweis. Wegen des Satzes von Rolle (Satz 3.9.2) ist $g(b) \neq g(a)$. Die Funktion $h\colon I \to \mathbb{R}$ sei durch

$$h(x) := f(x) - \left(f(a) + \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a))\right)$$

definiert. Dann ist $h(a) = h(b) = 0$. Nach dem Satz von Rolle gibt es ein $\xi \in (a, b)$ mit $h'(\xi) = 0$, d.h.

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(\xi).$$

Wegen $g'(\xi) \neq 0$ folgt die Behauptung. □

Satz 3.9.5. *Es sei I ein beliebiges Intervall und $f\colon I \to \mathbb{R}$ stetig in I und differenzierbar in $\mathring{I}$. Dann gilt*

$$f'(x) = 0 \quad \text{für alle } x \in \mathring{I} \Leftrightarrow \ f \equiv const. \text{ auf } I.$$

Beweis. "$\Rightarrow$":
Es sei $f'(x) = 0$ für alle $x \in \overset{\circ}{I}$. es seien $x_1, x_2 \in I$ und $x_1 < x_2$. Nach dem Mittelwertsatz gilt dann

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) = 0$$

für ein $\xi \in (x_1, x_2)$. Deshalb ist $f(x) \equiv \text{const}$.
"$\Leftarrow$":
Klar. □

Satz 3.9.6. *(Monotonietest)*
Es sei I ein beliebiges Intervall und $f\colon I \to \mathbb{R}$ in I stetig und in $\overset{\circ}{I}$ differenzierbar. Dann gilt

(i) Es gilt

$$f'(x) \geq 0 \quad \text{für alle } x \in \overset{\circ}{I} \Leftrightarrow \ f \text{ ist monoton wachsend.}$$

(ii) Es gilt

$$f'(x) > 0 \quad \text{für alle } x \in \overset{\circ}{I} \Rightarrow \ f \text{ ist streng monoton wachsend.}$$

Entsprechende Aussagen gelten für monoton fallenden Funktionen.

Beweis. (i) "$\Rightarrow$":
Es seien $x_1, x_2 \in I$ mit $x_1 < x_2$. Nach dem Mittelwertsatz (Satz 3.9.3) gibt es $\xi \in (x_1, x_2)$, so daß

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \underset{f'(\xi) \geq 0}{\Rightarrow} f(x_1) \leq f(x_2).$$

"$\Leftarrow$":
Es sei $x_0 \in \overset{\circ}{I}$. Es ist

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Wegen $f(x) \geq f(x_0)$ für $x \geq x_0$ ist $f'(x_0) \geq 0$.

(ii) ohne Beweis.

□

Bemerkung 3.9.2. Die Rückrichtung in *(ii)* gilt nicht. Dies zeigt das Beispiel $f\colon \mathbb{R} \to \mathbb{R}$, $x \to x^3$. Obwohl f auf $\mathbb{R}$ streng monoton wachsend ist, ist $f'(0) = 0$.

Beispiel 3.9.1. Es sei $f\colon \mathbb{R} \to \mathbb{R}$, $x \to f(x) := x^3 - 3x$. Man finde maximale Intervalle, auf denen f streng monoton wächst bzw. streng monoton fällt.
Lösung:
Es ist $f'(x) = 3x^2 - 3 = 3(x-1)(x+1)$. Es ist $f'(x) = 0$ für $x_1 = -1$ und $x_2 = 1$. Somit ist $f'(x) > 0$ für $x \in (-\infty, -1) \cup (1, \infty)$. Es ist $f'(x) < 0$ für $x \in (-1, 1)$. Nach Satz 3.9.6 ist f streng monoton wachsend auf $(-\infty, -1)$ und auf $(1, \infty)$ und streng monoton fallend auf $(-1, 1)$.

3.10 Höhere Ableitungen, Taylorpolynome

Definition 3.10.1. Es sei $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$, dessen Ableitung f' auf I existiere.

(i) Ist f' auf I stetig, so heißt f stetig differenzierbar auf I (Schreibweise: $f \in C^1(I)$).

(ii) Die Ableitung $f'(x_0)$ existiere in $x_0 \in I$. Dann heißt

$$f''(x_0) = \frac{d^2 f}{dx^2}(x_0) := (f')'(x_0) = \left(\frac{d}{dx}\left(\frac{df}{dx}\right)\right)(x_0)$$

die zweite Ableitung (oder Ableitung 2. Ordnung) von f an der Stelle x_0.

(iii) Falls $f''(x)$ für alle $x \in I$ existiert, dann heißt f zweimal differenzierbar auf I.

(iv) Ist zusätzlich f'' auf I stetig, dann heißt f zweimal stetig differenzierbar auf I (Schreibweise: $f \in C^2(I)$).

(v) Allgemein wird durch vollständige Induktion nach n die n- te Ableitung (oder Ableitung n- ter Ordnung) $f^{(n)}(x_0) = \dfrac{d^n f}{dx^n}(x_0)$ definiert und die Klasse $C^n(I)$ der n- mal stetig differenzierbaren Funktionen auf I.
Außerdem setzen wir $f^{(0)}(x_0) := f(x_0)$.

(vi) Existiert $f^{(n)}(x_0)$ für alle $n \in \mathbb{N}$, so heißt f unendlich oft differenzierbar in x_0.

(vii) Ist $f^{(n)}(x_0)$ für alle $n \in \mathbb{N}$ und für alle $x \in I$ erklärt, so heißt f unendlich oft differenzierbar in I (Schreibweise: $f \in C^\infty(I)$). In diesem Fall sind alle Ableitungen automatisch stetig, also ist f unendlich oft stetig differenzierbar.

Wir haben in Satz 3.7.1 gesehen, daß die erste Ableitung einer Funktion f für die lineare Approximation von f von Bedeutung ist. Es ist

$$f(x) = L_f(x) + r(x)(x - x_0)$$

mit $L_f(x) = f(x_0) + f'(x_0)(x - x_0)$.
Es ist $L_f(x_0) = f(x_0)$ und $L'_f(x_0) = f'(x_0)$. So ist dann L_f das eindeutig bestimmte Polynom vom Grad ≤ 1, das in x_0 dieselbe Ableitung bis zur 1. Ordnung besitzt wie f. Es ist nun zu erwarten, daß das Polynom höchstens n- ten Grades, das in x_0 dieselbe Ableitungen bis zur n- ten Ordnung besitzt wie f, eine besonders gute Approximation von f liefert.

Definition 3.10.2. Es sei $I \subset \mathbb{R}$ ein Intervall. Es sei $f\colon I \to \mathbb{R}$ in $x_0 \in I$ dann n- mal differenzierbar. Dann heißt

$$T^{(n)} f(x_0, x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

das n- te Taylorpolynom von f mit Entwicklungspunkt x_0.

Satz 3.10.1. *Es sei $I \subset \mathbb{R}$ ein Intervall. Es sei $f\colon I \to \mathbb{R}$ in $x_0 \in I$ genau n- mal differenzierbar und P ein Polynom vom Grad höchstens n. Dann sind folgende Aussagen äquivalent:*

(i) $P^{(k)}(x_0) = f^{(k)}(x_0)$ *mit* $0 \leq k \leq n$

(ii) $P(x) = T^{(n)} f(x_0, x)$.

Satz 3.10.2. *(Satz von Taylor)*
Es sei $n \in \mathbb{N}$ und $I \subset \mathbb{R}$ ein kompaktes Intervall. Gegeben sei die Funktion $f\colon I \to \mathbb{R}$, welche $(n+1)$-mal stetig differenzierbar auf $\overset{\circ}{I}$ sei und n- mal stetig differenzierbar auf I. Dann gilt die Taylorsche Formel

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} = T^{(n)}f(x_0,x) + R_{n+1}(x_0,x)$$

für alle $x \in I$ und $x \neq x_0$. Dabei ist $\xi = x_0 + t(x-x_0)$ für ein $t \in (0,1)$ und

$$R_{n+1}(x_0,x) := \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

das Restglied von Lagrange.

Beweis. ohne Beweis. □

3.11 de L'Hopitalsche Regeln

Die Bestimmung des Grenzwertes eines Quotientens $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ kann nach den Grenzwertsätzen (Satz 3.1.3) mittels

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$$

erfolgen, falls der Grenzwert des Nenners $\lim_{x\to a} g(x) \neq 0$ ist.
Gilt sowohl $\lim_{x\to a} f(x) = 0$ als auch $\lim_{x\to a} g(x) = 0$, so führen oft die de L'Hopitalschen Regeln zum Ziel. Diese können auch in Fällen angewandt werden, in denen $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$ ist.

Satz 3.11.1. *Die Funktionen f und g seien auf dem offenen Intervall (a,b) mit $-\infty \leq a < b \leq \infty$ differenzierbar. Weiter sei $g'(x) \neq 0$ für $x \in (a,b)$. Es sei*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \qquad \left(\textit{der Fall } \frac{0}{0}\right)$$

oder

$$\lim_{x\to a} g(x) = \pm\infty \qquad \left(\textit{der Fall } \frac{\infty}{\infty} \textit{ bzw. } \frac{c}{\infty}\right).$$

Außerdem existiere der Grenzwert $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$.
Dann ist $g(x) \neq 0$ für $x \in (a,b)$, der Limes $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ existiert, und es gilt die de L'Hopitalsche Regel

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Beweis. Wir beschränken uns auf den Fall $\frac{0}{0}$.

(i) Zunächst zeigen wir, daß

$$g(x) \neq 0 \text{ für alle } x \in (a,b) \tag{1}$$

gilt. Wir führen diesen Beweis durch Widerspruch:
<u>Annahme:</u> $\exists x_0 \in (a,b)$ mit $g(x_0) = 0$.

(a) Fall 1: $a \in \mathbb{R}$ (d.h. $a \neq -\infty$):
Nach dem Satz von Rolle (Satz 3.9.2) gibt es dann ein ξ mit $g'(\xi) = 0$, ein Widerspruch.

(b) Fall 2: $a = -\infty$:
Es ist $g(x_0 - 1) \neq 0$. Andernfalls würde der Satz von Rolle wieder ein $\xi \in (x_0 - 1, x_0)$ mit $g'(\xi) = 0$ liefern, wiederum ein Widerspruch.
Es folgt die Existenz eines $x_2 < x_1 := x_0 - 1$ mit $|g(x_2)| < |g(x_1)|$. Es gibt also ein relatives Extremum im Intervall (x_2, x_0). Der Satz von Rolle liefert wieder ein $\xi \in (a, b)$ mit $g'(\xi) = 0$, erneut ein Widerspruch.

(ii) Es sei $l := \lim\limits_{x \to a} \dfrac{f'(x)}{g'(x)}$.
Falls $l \in \mathbb{R} \cup \{-\infty\}$, dann sei $l' > l$ fest vorgegeben. Dann gibt es ein $a' > a$ mit

$$\frac{f'(x)}{g'(x)} < l'$$

für alle x mit $a < x < a'$. Für $a < x < y < a'$ folgt aus dem zweiten Mittelwertsatz

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(\xi)}{g'(\xi)} < l'$$

für ein $\xi \in (x, y)$. Für $x \to a$ folgt

$$\frac{f(y)}{g(y)} \leq l' \tag{2}$$

für $a < y < a'$.

(iii) Analog folgt für $l'' \leq l$: es existiert ein $a'' > a$ mit

$$\frac{f(y)}{g(y)} \geq l''$$

für alle y mit $a < y < a''$.

Aus (ii) und (iii) folgt die Behauptung. □

3.12 Konvexität und relative Extrema, Kurvendiskussion

Die Aufgabe der Kurvendiskussion ist es, die wesentlichen Eigenschaften des Graphen einer Funktion f zu ermitteln. Dazu gehört die Lage von absoluten und relativen Maxima und Minima. Wir haben mit Satz 3.9.1 zwar eine notwendige Bedingung für ein relatives Extremum in $x = \xi$ gefunden ($f'(\xi) = 0$), kennen aber noch keine hinreichende Bedingung. Eine solche kann durch Betrachtung der zweiten Ableitung f'' erhalten werden. Wir beginnen mit einem notwendigen Kriterium, das die zweite Ableitung benützt.

Satz 3.12.1. *(Notwendiges zweites Ableitungskriterium)*
Die Funktionen $f\colon I \to \mathbb{R}$ sei auf I zweimal stetig differenzierbar und besitze in einem Punkt $x_0 \in \mathring{I}$ ein relatives Maximum bzw. Minimum. Dann gilt

$$f''(x_0) \leq 0 \quad \textit{bzw.} \quad f''(x_0) \geq 0.$$

Beweis. O.B.d.A. besitze f ein relatives Maximum. Nach Satz 3.9.1 ist $f'(x_0) = 0$. Nach dem Satz von Taylor (Satz 3.10.2) haben wir

$$f(x) = f(x_0) + \frac{1}{2} f''(\xi)(x - x_0)^2 \tag{1}$$

mit ξ zwischen x und x_0.
<u>Annahme:</u> $f''(x_0) > 0$.
Wegen der Stetigkeit von f'' in x_0 gibt es ein $\delta > 0$, so daß $f''(\xi) > 0$ für alle $\xi \in U_\delta(x_0)$ ist. Aus (1) folgt dann $f(x) > f(x_0)$ für alle $x \in U_\delta(x_0)$, ein Widerspruch.
Also folgt $f''(x_0) \leq 0$.
Ebenso führt die Annahme $f''(x_0) < 0$ zu einem Widerspruch dazu, daß f in x_0 ein relatives Minimum besitzt. □

Satz 3.12.2. *(Hinreichendes zweites Ableitungskriterium)*
Es sei $f\colon I \to \mathbb{R}$ auf I zweimal stetig differenzierbar. In einem inneren Punkt $x_0 \in \overset{\circ}{I}$ sei $f'(x_0) = 0$ und $f''(x_0) < 0$ bzw. $f''(x_0) > 0$. Dann besitzt f an der Stelle x_0 ein isoliertes relatives Maximum bzw. Minimum.

Beweis. Wir betrachten nur den Fall $f''(x_0) < 0$. Der Fall $f''(x_0) > 0$ verläuft analog. Wie im Beweis von Satz 3.12.1 haben wir

$$f(x) = f(x_0) + \frac{1}{2} f''(\xi)(x - x_0)^2 \tag{1}$$

mit ξ zwischen x und x_0. Wegen der Stetigkeit von f'' gibt es ein $\delta = \delta(\epsilon)$, so daß $f''(\xi) \leq \frac{f''(x_0)}{2}$ für alle $\xi \in U_\delta(x_0)$ gilt. Daraus und aus (1) folgt $f(x) < f(x_0)$ für alle $x \in U_\delta(x_0)$.
Damit besitzt f ein isoliertes Maximum in $x = x_0$. □

Definition 3.12.1. Es sei $f\colon I \to \mathbb{R}$ auf I einmal differenzierbar. Dann besitzt f in $x_0 \in \overset{\circ}{I}$ einen <u>Wendepunkt</u> $(x_0, f(x_0))$, wenn die Ableitung f' in x_0 ein isoliertes relatives Extremum besitzt.

Definition 3.12.2. Es sei I ein (endliches oder unendliches) Intervall, das aus mehr als einem Punkt besteht. Dann heißt f <u>konvex</u> auf I, wenn die Ungleichung

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2) \tag{1}$$

für alle $x_1, x_2 \in I$ und alle $t \in (0,1)$ gilt.
Zudem heißt f <u>streng konvex</u>, wenn die strikte Ungleichung

$$f((1-t)x_1 + tx_2) < (1-t)f(x_1) + tf(x_2) \tag{2}$$

für alle $x_1 \neq x_2$ und alle $t \in (0,1)$ gilt.
Gelten (1) bzw. (2) mit dem $\geq$- bzw. $>$- Zeichen, so heißt f <u>konkav</u> bzw. <u>streng konkav</u>.

Bemerkung 3.12.1. Die Ungleichungen (1) bzw. (2) haben folgende geometrische Bedeutung:
Durchläuft t das Intervall $(0,1)$, so durchläuft der Punkt $((1-t)x_1 + tx_2, (1-t)f(x_1) + tf(x_2))$ die Verbindungsstrecke der Punkte $(x_1, f(x_1))$ und $(x_2, f(x_2))$, d.h. die <u>Sekante</u>, während $((1-t)x_1 + tx_2, f((1-t)x_1 + tx_2))$ die Punkte des Graphen $(x, f(x))$ mit den gleichen Abszissen durchläuft. Die Ungleichungen (1) bzw. (2) besagen also, daß das Segment des Graphen von $y = f(x)$ zwischen zweien seiner Punkte nicht unter bzw. über der <u>Sekante</u> ist.

Satz 3.12.3. *Es sei I ein beliebiges Intervall, $f\colon I \to \mathbb{R}$ sei auf I stetig und auf $\overset{\circ}{I}$ differenzierbar. Dann gilt:*

(i) *Die Funktion f ist auf I genau dann konvex, wenn die Ableitung f' auf $\overset{\circ}{I}$ monoton wächst.*

(ii) *Weiter ist f auf I genau dann streng konvex, wenn die Ableitung f' auf $\overset{\circ}{I}$ streng monoton wächst.*

Beweis. ohne Beweis. $\square$

Satz 3.12.4. *(Konvexität, zweites Ableitungskriterium)*
Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall, und $f\colon I \to \mathbb{R}$ sei auf I stetig und auf $\overset{\circ}{I}$ differenzierbar. Dann gilt:

(i) *Die Funktion f ist auf I genau dann konvex, wenn $f''(x) \geq 0$ für alle $x \in \overset{\circ}{I}$ gilt.*

(ii) *Ist $f''(x) > 0$ für alle $x \in \overset{\circ}{I}$, so ist f streng konvex.*

Beweis. ohne Beweis. $\square$

Kapitel 4

Stetigkeit und Differenzierbarkeit von Grenzfunktionen, Potenzreihen

4.1 Gleichmäßige Konvergenz und Stetigkeit

Die Typen der uns bisher bekannten Funktionen sind sehr begrenzt. Sie umfassen lediglich Ausdrücke, die durch Kombination der vier Grundrechenarten mit der Wurzelbildung erhalten werden. Wichtige andere "elementare" Funktionen werden als Grenzwerte von Funktionenfolgen erhalten, z.B. die Exponentialfunktion $E(x)$ (oder e^x) durch

$$E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}.$$

Somit ist $E(x)$ der Grenzwert der Folge der Funktionen $(T_n(x))_{n=0}^{\infty}$ mit $T_n(x) := \sum_{k=0}^{n} \frac{x^k}{k!}$.

Jede Funktion $T_n(x)$ dieser Folge ist als Polynom stetig und differenzierbar. Gilt dies auch für die Grenzfunktion $E(x)$? Wir werden im folgenden die Grundlagen zur Behandlung dieser Fragen bereitstellen.

Das folgende Beispiel zeigt, daß die Grenzfunktion einer Folge stetiger Funktionen nicht stetig zu sein braucht.

Beispiel 4.1.1. Für $n \in \mathbb{N}$ sei

$$f_n\colon [0,1] \to \mathbb{R}, \quad x \to f_n(x) = x^n.$$

Es ist

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{für} \quad 0 \le x < 1 \\ 1 & \text{für} \quad x = 1. \end{cases}$$

Die Grenzfunktion $f(x)$ ist also unstetig in $x = 1$, obwohl alle Funktionen f_n der Folge stetig sind.

Die Stetigkeit der Grenzfunktion folgt jedoch, wenn man die Forderung der Konvergenz durch die stärkere Forderung der gleichmäßigen Konvergenz ersetzt.

Definition 4.1.1. Es sei $D \subseteq \mathbb{R}$. Für $n \in \mathbb{N}$ seien $f_n\colon D \to \mathbb{R}$ Funktionen. Außerdem sei $f\colon D \to \mathbb{R}$. Man sagt $(f_n)_{n=1}^{\infty}$ konvergiert gleichmäßig gegen f (auf D), Schreibweise: $f_n \overset{glm.}{\longrightarrow} f$, falls gilt:

$$\forall \epsilon > 0 \; \exists n_0 = n_0(\epsilon), \text{ so daß } \forall n \ge n_0 : |f(x) - f_n(x)| < \epsilon.$$

Bemerkung 4.1.1. Der Unterschied zur punktweisen Konvergenz besteht darin, daß für gegebenes $\epsilon > 0$ ein n_0 existieren muß, das von x unabhängig ist. In Beispiel 4.1.1 läßt sich zu jedem Paar (ϵ, x) mit $\epsilon > 0$ und $x_0 \in [0,1]$ ein $n_0 = n_0(\epsilon, x_0)$ finden, so daß $|f(x_0) - f_n(x_0)| < \epsilon$ für alle $n \geq n_0$ ist. Es läßt sich jedoch kein solches n_0 finden, das für alle x_0 funktioniert.

Satz 4.1.1. *Es sei $D \subseteq \mathbb{R}$, und $f_n\colon D \to \mathbb{R}$ mit $n \in \mathbb{N}$ sei auf D stetig. Die Folge $(f_n)_{n=1}^{\infty}$ konvergiere gleichmäßig gegen $f\colon D \to \mathbb{R}$. Dann ist auch f auf D stetig.*

Beweis. Es sei $x_0 \in D$. Ist x_0 isolierter Punkt von D, so ist f nach Bemerkung 3.2.1 in x_0 stetig. Wir können also annehmen, daß x_0 Häufungspunkt von D ist. Es sei $\epsilon > 0$ gegeben. Dann gibt es nach der Definition der gleichmäßigen Konvergenz ein $n_0 = n_0(\epsilon/3) \in \mathbb{N}$, so daß

$$|f(x) - f_n(x)| < \epsilon/3 \quad \text{für alle } n \geq n_0(\epsilon/3) \tag{1}$$

ist. Wegen der Stetigkeit von f_n gibt es nach Definition 3.2.1 ein $\delta = \delta(\epsilon/3) > 0$, so daß

$$|f_n(x) - f_n(x_0)| < \epsilon/3 \quad \text{für alle } x \in U_\delta(x_0) \tag{2}$$

ist. Aus (1) und (2) folgt nun für alle $x \in U_\delta(x_0)$

$$\begin{aligned}
|f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\
&\underset{\Delta-Ugl.}{\leq} |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\
&< \epsilon.
\end{aligned}$$

Dies zeigt nach Definition 3.2.1 die Stetigkeit von f in x_0. □

Satz 4.1.2. *(Cauchykriterium für gleichmäßige Konvergenz)*
Eine Folge von Funktionen $f_n\colon D \to \mathbb{R}$ ist genau dann gleichmäßig konvergent, wenn

$$\forall \epsilon > 0\ \exists n_0 = n_0(\epsilon),\ \textit{so daß für } n_1, n_2 \geq n_0 \textit{ gilt } |f_{n_1}(x) - f_{n_2}(x)| < \epsilon.$$

Beweis. Nach dem Cauchykriterium (Satz 2.1.5) konvergiert die Folge $(f_n(x))$ für alle $x \in D$. Es gibt also eine Funktion $f\colon D \to \mathbb{R}$ mit $\lim\limits_{n\to\infty} f_n(x) = f(x)$.
Es sei $\epsilon > 0$ gegeben. Dann gibt es $n_0 = n_0(\epsilon/2)$, so daß für $n_1, n_2 \geq n_0$ gilt:

$$|f_{n_1}(x) - f_{n_2}(x)| < \epsilon/2.$$

Satz 2.1.8 (Erhaltung von Ungleichungen) ergibt mit dem Grenzübergang $n_2 \to \infty$ schließlich

$$|f_{n_1}(x) - f(x)| \leq \epsilon/2 < \epsilon.$$

□

4.2 Differenzierbarkeit der Grenzfunktion

Satz 4.2.1. *Es sei $I = [a,b]$ mit $a < b \in \mathbb{R}$. Die Funktionen $f_n\colon I \to \mathbb{R}$ mit $n \in \mathbb{N}$ seien differenzierbar mit den Ableitungen $f_n'\colon I \to \mathbb{R}$. Für $x_0 \in I$ konvergiere (f_n) und (f_n') konvergiere gleichmäßig auf I. Dann konvergiert (f_n) auf I gleichmäßig gegen eine auf I differenzierbare Funktion f, und es gilt*

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

für alle $x \in I$.

Beweis. i) Es sei $\epsilon > 0$ gegeben. Nach dem Cauchykriterium (Satz 2.1.5) gibt es ein $n_0 \in \mathbb{N}$, so daß

$$|f_k(x_0) - f_l(x_0)| < \epsilon/2 \quad \text{für alle } k, l \geq n_0. \tag{1}$$

Weiter folgt wegen der gleichmäßigen Konvergenz von $(f'_n(x))$ auf I die Existenz von $n_1 \in \mathbb{N}$, so daß für $k, l \geq n_1$

$$|f'_k(\xi) - f'_l(\xi)| < \frac{\epsilon}{2|I|} \quad \text{für alle } \xi \in I \tag{2}$$

gilt. Es sei $n_2 = \max\{n_0, n_1\}$ mit $k, l \geq n_2$ und $x \in I$. Nach dem Mittelwertsatz (Satz 3.9.3) gibt es ein ξ zwischen x_0 und x, so daß

$$(f_k(x) - f_l(x)) - (f_k(x_0) - f_l(x_0)) = (f'_k(\xi) - f'_l(\xi))(x - x_0). \tag{3}$$

Aus (1), (2) und (3) folgt

$$|f_k(x) - f_l(x)| < \epsilon \quad \text{für alle } k, l \geq n_2.$$

Damit erfüllt die Folge (f_n) das Cauchykriterium für gleichmäßige Konvergenz und konvergiert nach Satz 4.1.2 gleichmäßig gegen eine Grenzfunktion f:

$$\lim_{n\to\infty} f_n(x) = f(x) \tag{4}$$

für alle $x \in I$. Nach Satz 4.1.1 ist f stetig.

ii) Für $n \in \mathbb{N}$ sei

$$g_n(x) = \begin{cases} \dfrac{f_n(x) - f_n(x_0)}{x - x_0}, & \text{falls} \quad x \neq x_0 \\ f'_n(x_0), & \text{falls} \quad x = x_0. \end{cases}$$

Nach der Definition der Ableitung (Definition 3.7.1) ist g_n auf I stetig. Nach dem Mittelwertsatz (Satz 3.9.3) folgt für alle $x \neq x_0$:

$$g_k(x) - g_l(x) = f'_k(\xi) - f'_l(\xi) \quad \text{für ein } \xi \in (x_0, x).$$

Wegen der gleichmäßigen Konvergenz von (f'_n) auf I erfüllt auch die Folge (g_n) das Cauchykriterium der gleichmäßigen Konvergenz und konvergiert nach Satz 4.1.2 gegen eine Grenzfunktion g. Nach Satz 4.1.1 ist g stetig und damit

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} g(x) = g(x_0) = \lim_{n\to\infty} f'_n(x_0).$$

Nach Definition 3.7.1 ist damit f in x_0 differenzierbar und $f'(x_0) = \lim_{n\to\infty} f'_n(x_0)$. Da nach (4) die Voraussetzung $\lim_{n\to\infty} f_n(x_0) = f(x_0)$ für alle $x \in I$ erfüllt ist, ist f für alle $x \in I$ differenzierbar, und es gilt $f'(x) = \lim_{n\to\infty} f'_n(x)$.

□

4.3 Stetigkeit und Differenzierbarkeit durch unendliche Reihen definierter Funktionen

Wie schon in Abschnitt 4.1 ausgeführt wurde, sind unendliche Reihen von Funktionen Spezialfälle von Grenzfunktionen von Funktionenfolgen. Betrachtet man die Sätze 4.1.1 und 4.1.2 für diesen Spezialfall, so erhält man sofort

Satz 4.3.1. *i) Es sei* $D \subseteq \mathbb{R}$. *Die Funktionen* $g_n\colon D \to \mathbb{R}$ *mit* $n \in \mathbb{N}$ *seien auf* D *stetig. Die unendliche Reihe* $\sum_{n=1}^{\infty} g_n(x)$ *konvergiere auf* D *gleichmäßig gegen* $f(x)$. *Dann ist auch* f *auf* D *stetig.*

ii) Es sei $I = [a, b]$ *mit* $a < b \in \mathbb{R}$. *Die Funktionen* $h_n\colon I \to \mathbb{R}$ *mit* $n \in \mathbb{N}$ *seien differenzierbar mit den Ableitungen* $h_n'\colon I \to \mathbb{R}$. *Für* $x_0 \in I$ *konvergiere* $\sum_{n=1}^{\infty} h_n(x_0)$ *und* $\sum_{n=1}^{\infty} h_n'(x_0)$ *konvergiere gleichmäßig auf* I. *Dann konvergiert* $h(x) = \sum_{n=1}^{\infty} h_n(x)$ *auf* I *gleichmäßig. Weiter ist* $h(x)$ *auf* I *differenzierbar, und es gilt* $h'(x) = \sum_{n=1}^{\infty} h_n'(x)$ *für alle* $x \in I$.

Wir geben noch ein Kriterium für die gleichmäßige Konvergenz von unendlichen Reihen:

Satz 4.3.2. *(Weierstraßscher M- Test)*
Es sei $D \subset \mathbb{R}$ *und* $f_n\colon D \to \mathbb{R}$ *mit* $n \in \mathbb{N}$ *Funktionen. Es gebe Zahlen* $M_n \in \mathbb{R}$, *so daß* $|f_n(x)| \leq M_n$ *für alle* $x \in D$ *und* $n \in \mathbb{N}_0$ *mit* $\sum_{n=1}^{\infty} M_n < \infty$.
Dann konvergiert die Reihe $\sum_{n=1}^{\infty} f_n$ *auf* D *gleichmäßig.*

Beweis. Es sei $\epsilon > 0$. Nach dem Cauchykriterium (Satz 2.1.5) existiert ein n_0, so daß $\sum_{k=m}^{n} M_k < \epsilon$ für alle m, n mit $n_0 \leq m \leq n$. Dann ist auch $\left|\sum_{k=m}^{n} f_k(x)\right| < \epsilon$. Die gleichmäßige Konvergenz der Reihe folgt aus Satz 4.1.2. $\square$

4.4 Potenzreihen

Definition 4.4.1. Es sei $(a_n)_{n=0}^{\infty}$ eine Folge reeller Zahlen und $x, x_0 \in \mathbb{R}$. Die unendliche Reihe

$$p(x) := \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

heißt Potenzreihe (in x) mit Entwicklungspunkt x_0 und Koeffizienten a_n.

Bemerkung 4.4.1. Zur Bezeichnung der Folge (a_n) können auch andere Symbole verwendet werden, ebenso an Stelle von x, z. B. ist $\sum_{n=0}^{\infty} b_n (w - w_0)^n$ eine Potenzreihe in w mit Entwicklungspunkt w_0 und Koeffizienten b_n.

Die grundlegende Frage betrifft den Konvergenzbereich der Potenzreihe, die Menge aller x, für die $p(x)$ konvergiert.

Beispiel 4.4.1. Die geometrische Reihe $\sum_{n=0}^{\infty} x^n$ konvergiert nach Satz 2.3.1 für $|x| < 1$ und divergiert für $|x| \geq 1$. Der Konvergenzbereich ist also das Intervall $(-1, 1)$.

Wir werden sofort sehen, daß der Konvergenzbereich einer Potenzreihe stets ein Intervall ist.
Zur Vorbereitung zeigen wir

Satz 4.4.1. *Es sei* $p(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ *eine Potenzreihe, und* $p(x)$ *konvergiere für* $x = x_1$. *Es sei* $0 < \epsilon < |x_1 - x_0|$.
Dann konvergiert $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *gleichmäßig auf der Menge* $\{x|\ |x-x_0| \leq |x_1-x_0| - \epsilon\}$.

Beweis. Nach Satz 2.4.1 folgt aus der Konvergenz für $x = x_1$

$$|a_n||x_1 - x_0|^n \to 0 \quad (n \to \infty).$$

Insbesondere ist $|a_n||x_1 - x_0|^n \leq M$ für ein (von n unabhängiges) M. Es ist

$$|a_n||x-x_0|^n \leq |a_n|(|x_1-x_0| - \epsilon)^n = |a_n||x_1-x_0|^n \left(1 - \frac{\epsilon}{|x_1-x_0|}\right)^n \leq Mq^n$$

mit $q = 1 - \dfrac{\epsilon}{|x_1 - x_0|} < 1$.
Die Behauptung folgt nach dem Weierstraßschen M- Test (Satz 4.3.2). □

Satz 4.4.2. *Es sei* $p(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ *eine Potenzreihe. Dann tritt genau einer der folgenden drei Fälle ein:*

i) Die Potenzreihe $p(x)$ *konvergiert nur für* $x = x_0$.

ii) Es gibt eine Zahl $r > 0$, *so daß* $p(x)$ *für* $|x-x_0| < r$ *konvergiert und für* $|x-x_0| > r$ *divergiert.*

iii) Die Potenzreihe $p(x)$ *konvergiert für alle* x.

Beweis. Die drei Fälle schließen sich offenbar gegenseitig aus. Wir nehmen an, daß die Fälle (i) und (iii) nicht eintreten. Zu zeigen ist, daß dann der Fall (ii) eintritt. Es sei

$$\mathcal{K} = \{x' \colon p(x) \text{ konvergiert für } x \in (x_0 - |x'-x_0|, x_0 + |x'-x_0|)\} \tag{1}$$

und

$$r = \inf\{|x'-x_0| \colon x' \in \mathcal{K}^C\}. \tag{2}$$

Da Fall (i) nicht eintritt, existiert ein $x_1 \neq x_0$, so daß $p(x_1)$ konvergiert. Nach Satz 4.4.1 ist dann $(x_0 - |x_1-x_0|, x_0 + |x_1-x_0|) \subseteq \mathcal{K}$ und damit $r > 0$.
Da Fall (iii) nicht eintritt ist $\mathcal{K}^C \neq \infty$ und damit $r < \infty$.

i) Es sei $x \in (x_0 - r, x_0 + r)$.
Nach den Definitionen (1) und (2) existiert ein $x' \in (x_0 - r, x_0 + r)$ so daß $p(x')$ konvergiert.
Nach Satz 4.4.1 konvergiert auch $p(x)$.

ii) Es sei $x \notin [x_0 - r, x_0 + r]$.
Nach den Definitionen (1) und (2) gibt es ein x'' mit $|x_0 - x''| < |x_0 - x|$, so daß $p(x'')$ divergiert.
Annahme: $p(x)$ ist konvergent.
Dann ist nach Satz 4.4.1 auch $p(x'')$ konvergent, ein Widerspruch.

$\square$

Definition 4.4.2. Trifft auf eine Potenzreihe der Fall (ii) aus Satz 4.4.2 zu, so nennen wir den Wert r den Konverenzradius der Potenzreihe.
Im Fall (i) sagen wir: p hat den Konvergenzradius $r = 0$ und im Fall (iii): p hat den Konvergenzradius $r = \infty$.

Satz 4.4.3. *(Formel für den Konvergenzradius)*
Es gilt

$$r = \left(\limsup_{n\to\infty} \sqrt[n]{|a_n|}\right)^{-1}.$$

Hierbei ist $0^{-1} := \infty$ *und* $\infty^{-1} := 0$ *zu setzen.*

Beweis. Es sei $l := \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.
Nach dem Wurzelkriterium (Satz 2.4.7) ist $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ konvergent, wenn

$$l \cdot |x - x_0| < 1, \quad \text{also für } |x - x_0| < l^{-1},$$

und divergent für

$$l \cdot |x - x_0| > 1, \quad \text{also für } |x - x_0| > l^{-1}.$$

Die Behauptung folgt nach Definition 4.4.2. $\square$

Bei der Behandlung einiger wichtiger Beispiele hiflt

Satz 4.4.4. *Es ist*

$$\lim_{n\to\infty} \sqrt[n]{n} = 1 \quad \textit{und} \quad \lim_{n\to\infty} \sqrt[n]{n!} = \infty.$$

Beweis. Nach Satz 2.4.11 ist

$$\limsup_{n\to\infty} \sqrt[n]{n} \leq \limsup_{n\to\infty} \frac{n}{n+1} = 1$$

und

$$\liminf_{n\to\infty} \sqrt[n]{n!} \geq \liminf_{n\to\infty} \frac{(n+1)!}{n!} = \infty.$$

$\square$

Beispiel 4.4.2. Es sei

$$p(x) = \sum_{n=1}^{\infty} \frac{2^n}{n} x^n.$$

Nach Satz 4.4.4 ist

$$\lim_{n\to\infty} \sqrt[n]{\frac{2^n}{n}} = 2 \lim_{n\to\infty} \sqrt[n]{n^{-1}} = 2.$$

Nach Satz 4.4.3 hat $p(x)$ den Konvergenzradius $r = 1/2$.

Bemerkung 4.4.2. Satz 4.4.2 (ii) macht keine Aussage über die Konvergenz der Potenzreihe in den Endpunkten des Intervalls $(x_0 - r, x_0 + r)$. Es können alle vier Fälle auftreten: Konvergenz in keinem der beiden Endpunkte, nur in dem rechten, nur in dem linken oder in beiden Endpunkten.

Beispiel 4.4.3. Es sei

$$p(x) = \sum_{n=1}^{\infty} \frac{x^n}{n}.$$

Nach Satz 4.4.4 ist $r = \lim_{n\to\infty} \left(\sqrt[n]{\frac{1}{n}}\right)^{-1} = 1$. Weiter ergibt Satz 4.4.2 Konvergenz in $(-1, 1)$ und Divergenz außerhalb von $[-1, 1]$.

Für den Endpunkt $x_1 = 1$ erhalten wir $p(x_1) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$, die harmonische Reihe, also Divergenz nach Satz 2.4.9. Für den Endpunkt $x_2 = -1$ erhalten wir $p(x_2) = \displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$, also Konvergenz nach dem Leibnizkriterium (Satz 2.4.3).

Satz 4.4.5. *(Stetigkeit und Differenzierbarkeit von Potenzreihen)*
Es sei $p(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *eine Potenzreihe mit Konvergenzradius* $r > 0$. *Dann ist* $p(x)$ *stetig und in* $(x_0 - r, x_0 + r)$ *unendlich oft differenzierbar, falls* $r < \infty$ *bzw. auf* $\mathbb{R}$ *unendlich oft differenzierbar, falls* $r = \infty$ *ist. Die Ableitungen können durch gliedweise Differentiation erhalten werden, d.h.*

$$p'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Die Potenzreihen für alle Ableitungen $p^{(k)}(x)$ *mit* $k \in \mathbb{N}$ *haben denselben Konvergenzradius wie* $p(x)$.

Beweis. Wir behandeln nur den Fall $r < \infty$. Den Fall $r = \infty$ erhält man durch kleine Änderungen. Wegen

$$\lim_{n\to\infty} \sqrt[n]{n} = 1 \quad \text{ist} \quad \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} \sqrt[n]{n|a_n|} = r^{-1}.$$

Damit haben $p(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x - x_0)^n$ und $p'(x) = \displaystyle\sum_{n=1}^{\infty} n a_n (x - x_0)^n$ denselben Konvergenzradius r. Nach Satz 4.4.1 konvergieren sie beide gleichmäßig in $[x_0 - r + \epsilon, x_0 + r - \epsilon]$ für jedes $\epsilon > 0$. Nach Satz 4.3.1 ist $p(x)$ in $(x_0 - r, x_0 + r)$ stetig und differenzierbar, und es ist $p'(x) = \displaystyle\sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}$. □

Satz 4.4.6. *(Identitätssatz für Polynome)*
Es sei $x \in \mathbb{R}$. *Es seien* $p_1(x) = \displaystyle\sum_{n=0}^{\infty} a_n(x - x_0)^n$ *bzw.* $p_2(x) = \displaystyle\sum_{n=0}^{\infty} b_n(x - x_0)^n$ *Potenzreihen mit Entwicklungspunkt* x_0 *und Konvergenzradius* $r_1 > 0$ *bzw.* $r_2 > 0$. *Es existiere ein* $r_3 > 0$, *so daß* $p_1(x) = p_2(x)$ *für alle* $x \in (x_0 - r_3, x_0 + r_3)$ *ist. Dann ist* $a_n = b_n$ *für alle* $n \in \mathbb{N}_0$.

Beweis. Wegen $p_1(x) = p_2(x)$ für alle $x \in (x_0 - r_3, x_0 + r_3)$ gilt auch $p_1^{(k)}(x) = p_2^{(k)}(x)$ für alle $x \in (x_0 - r_3, x_0 + r_3)$ und für alle $k \in \mathbb{N}$, insbesondere auch $p_1^{(k)}(x_0) = p_2^{(k)}(x_0)$ für alle $k \in \mathbb{N}$. Durch k- fache gliedweise Differentiation nach Satz 4.4.5 erhalten wir

$$p_1^{(k)}(x_0) = k! \cdot a_n \quad \text{und} \quad p_2^{(k)}(x_0) = k! \cdot b_n.$$

Also ist $a_n = b_n$. □

4.5 Taylorreihen

Wir erinnern an den Satz von Taylor (Satz 3.10.2):
Es sei $n \in \mathbb{N}$ und $I \subset \mathbb{R}$ ein kompaktes Intervall. Gegeben sei die Funktion $f\colon I \to \mathbb{R}$, welche $(n+1)$-mal stetig differenzierbar auf $\overset{\circ}{I}$ sei und n- mal stetig differenzierbar auf I. Dann gilt die Taylorsche Formel

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} = T^{(n)}f(x_0,x) + R_{n+1}(x_0,x) \tag{1}$$

für alle $x \in I$ und $x \neq x_0$. Dabei ist $\xi = x_0 + t(x-x_0)$ für ein $t \in (0,1)$ und

$$R_{n+1}(x_0,x) := \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

das Restglied von Lagrange.
Ist f auf I unendlich oft differenzierbar, so kann (1) für jeden Wert von n berechnet werden.
Die Folge der Taylorpolynome $(T^{(n)}f(x_0,x))$ bildet dann eine Potenzreihe, die Taylorreihe

$$T^{(n)}f(x_0,x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

Definition 4.5.1. Es sei f auf I unendlich oft differenzierbar und $x_0 \in I$. Unter der Taylorreihe $Tf(x_0,x)$ von f mit Entwicklungspunkt x_0 versteht man

$$Tf(x_0,x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Es stellt sich die Frage: Für welche x gilt $Tf(x_0,x) = f(x)$?

Satz 4.5.1. *Es gilt genau dann*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,$$

wenn $\lim\limits_{n\to\infty} R_{n+1}(x_0,x) = 0$ *ist.*

Beweis. Dies folgt sofort aus der Taylorschen Formel (1). □

Bemerkung 4.5.1. Eine Potenzreihe $p(x) = \sum\limits_{k=0}^{\infty} a_k(x-x_0)^k$ mit positivem Konvergenzradius ist ihre eigene Taylorreihe. Denn k- fache gliedweise Differenzierung ergibt, wie im Beweis von Satz 4.4.6

$$a_k = \frac{p^{(k)}(x_0)}{k!}.$$

Kapitel 5

Die elementaren transzendenten Funktionen

5.1 Die Exponentialfunktion

Satz 5.1.1. *Die Potenzreihe* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ *ist für alle* $x \in \mathbb{R}$ *konvergent.*

Beweis. Nach Satz 4.4.4 ist $\lim_{n\to\infty} \left(\frac{1}{n!}\right)^{1/n} = 0$. Nach Satz 4.4.3 hat die Potenzreihe den Konvergenzradius ∞. □

Definition 5.1.1. Die nach Satz 5.1.1 für alle $x \in \mathbb{R}$ definierte Funktion

$$E(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

heißt Exponentialfunktion. Andere Schreibweisen sind e^x oder $\exp(x)$.

Satz 5.1.2. *Die Exponentialfunktion hat folgende Eigenschaften:*

i) (Differenzierbarkeit)
Die Funktion $E(x)$ *ist unendlich oft differenzierbar, und es gilt* $E^{(n)}(x) = E(x)$ *für alle* $n \in \mathbb{N}$ *und für alle* $x \in \mathbb{R}$*, insbesondere also* $E'(x) = E(x)$ *für alle* $x \in \mathbb{R}$*.*

ii) (Funktionalgleichung)
Es gilt

$$\begin{aligned} E(x_1 + x_2) &= E(x_1) \cdot E(x_2) \quad \textit{für alle } x_1, x_2 \in \mathbb{R}, \\ E(-x) &= E(x)^{-1} \quad \textit{und } E(0) = 1, \\ E(x) &> 0 \quad \textit{für alle } x. \end{aligned}$$

iii) (Monotonie, Konvexität)
$E(x)$ *ist streng monoton wachsend und streng konvex.*

iv) (Asymptotik)
Es gilt

$$\lim_{x\to\infty} E(x) = \infty \quad \textit{und} \quad \lim_{x\to-\infty} E(x) = 0$$

v) (Wachstum)
Es gilt

$$\lim_{x\to\infty} \frac{x^n}{E(x)} = 0 \quad \textit{für alle } n \in \mathbb{N}.$$

vi) (Wertebereich)
E hat den Wertebereich $E(\mathbb{R}) = (0,\infty)$.

Beweis. i) Nach Satz 4.4.5 ist die Potenzreihe $E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ differenzierbar, und die Ableitung $E'(x)$ kann durch "gliedweise Differentiation' erhalten werden:

$$E'(x) = \sum_{n=0}^{\infty} \left(\frac{d}{dx}\frac{x^n}{n!}\right) = \sum_{n=1}^{\infty} \frac{1}{n!}\left(n \cdot x^{n-1}\right) = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \underset{\substack{Index-\\verschiebung}}{=} \sum_{n=0}^{\infty} \frac{x^n}{n!} = E(x).$$

Durch vollständige Induktion nach n folgt $E^{(n)}(x) = E(x)$ für alle $n \in \mathbb{N}$.

ii) Es seien $x_1, x_2 \in \mathbb{R}$. Dann ist

$$E(x_1) \cdot E(x_2) = \sum_{m=0}^{\infty} \frac{x_1^m}{m!} \sum_{n=0}^{\infty} \frac{x_2^n}{n!}.$$

Nach Satz 2.5.2 (Cauchyprodukt) ist

$$\begin{aligned} E(x_1) \cdot E(x_2) &= \sum_{l=0}^{\infty}\sum_{m=0}^{l} \frac{x_1^m}{m!} \cdot \frac{x_2^{l-m}}{(l-m)!} = \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{m=0}^{l} \frac{l!}{m!(l-m)!} x_1^m x_2^{l-m} \\ &= \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{m=0}^{l} \binom{l}{m} x_1^m x_2^{l-m} \underset{\substack{S.\ 1.5.10\\ Binom.Lehrsatz}}{=} \sum_{l=0}^{\infty} \frac{1}{l!}(x_1+x_2)^l = E(x_1+x_2). \end{aligned}$$

Es ist weiter

$$E(0) = \sum_{n=0}^{\infty} \frac{0^n}{0!} = 0^0 + \frac{0^1}{1!} + \ldots = 0^0 = 1.$$

Mit $x_1 = x$ und $x_2 = -x$ folgt $E(x) \cdot E(-x) = E(x + (-x)) = E(0) = 1$ und damit ist $E(-x) = E(x)^{-1}$.
Für $x \geq 0$ ist

$$E(x) = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} \geq 1,$$

und für $x < 0$ ist $E(x) = E(-x)^{-1} > 0$.

iii) Es ist $E'(x) = E(x) > 0$ und $E''(x) = E(x) > 0$.

iv) und v) Für $x > 0$ ist

$$E(x) = 1 + x + \ldots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \ldots \geq \frac{x^{n+1}}{(n+1)!}$$

für alle n. Damit ist

$$\frac{x^n}{E(x)} \leq \frac{x^n}{x^{n+1}/(n+1)!} = \frac{(n+1)!}{x}, \quad \text{also } \lim_{x\to\infty} \frac{x^n}{E(x)} = 0.$$

Es folgt $E(x) \geq x^n$ für alle $x \geq x_0(n)$ und damit

$$\lim_{x\to\infty} E(x) = \infty \quad \text{und} \quad \lim_{x\to-\infty} E(x) = \lim_{x\to-\infty} E(-x)^{-1} = 0.$$

vi) Zu jedem $y > 0$ gibt es nach iv) $x_1, x_2 \in \mathbb{R}$ mit $E(x_1) < y < E(x_2)$. Nach dem Zwischenwertsatz (Satz 3.5.4) existiert ein x zwischen x_1 und x_2 mit $E(x) = y$.

□

Definition 5.1.2. (Hyperbelfunktionen)
Die Funktionen sinh und cosh (Sprechweise: Sinus hyperbolicus und Kosinus hyperbolicus) sind durch

$$\begin{aligned}\sinh x &:= \frac{e^x - e^{-x}}{2} \ (= \frac{1}{2}(E(x) - E(-x)))\\ \cosh x &:= \frac{e^x + e^{-x}}{2} \ (= \frac{1}{2}(E(x) + E(-x)))\end{aligned}$$

definiert.

Satz 5.1.3. *Es ist*

i) $\cosh(-x) = \cosh x$ *und* $\sinh(-x) = -\sinh x$ *für alle* $x \in \mathbb{R}$

ii) $\sinh' x = \cosh x$ *und* $\cosh' x = \sinh x$.

iii) $\cosh^2 x - \sinh^2 x = 1$

iv) $\lim_{x\to\infty} \cosh x = \lim_{x\to-\infty} \cosh x = \infty$ *und*
$\lim_{x\to\infty} \sinh x = \infty$ *bzw.* $\lim_{x\to-\infty} \sinh x = -\infty$.

Beweis. Durch Nachrechnen. □

5.2 Der Logarithmus

Satz 5.2.1. *Die Exponentialfunktion* $E\colon \mathbb{R} \to \mathbb{R}$ *besitzt eine Inverse* $E^{-1}\colon \mathbb{R}^+ \to \mathbb{R}$.

Beweis. Dies folgt aus den Eigenschaften (iii) und (vi) der Exponentialfunktion (nach Satz 5.1.2) und Satz 3.6.2. □

Definition 5.2.1. Die Funktion E^{-1} von Satz 5.2.1 wird mit "log" bezeichnet und heißt der (natürliche) Logarithmus von x.
Sie hat die folgenden Eigenschaften:

i) $E(\log x) = x$ und $\log(E(x)) = x$ für alle $x > 0$
$\log 1 = 0$

ii) log ist differenzierbar, und es ist $\frac{d}{dx}(\log x) = \frac{1}{x}$ für alle $x > 0$

iii) log ist streng monoton und streng konkav

iv) Es ist $\lim_{x\to\infty} \log x = \infty$ und $\lim_{x\to 0^+} \log x = -\infty$

Beweis. ohne Beweis. $\square$

Satz 5.2.2. *Für $|x| < 1$ gilt die Taylorentwicklung:*

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^n}{n}. \tag{$*$}$$

Beweis. Wir berechnen zuerst das n- te Taylorpolynom für $f(x) := \log(1+x)$ mit Entwicklungspunkt 0. Die Folge der Ableitungen $f^{(n)}$ ergibt sich wie folgt:

$$f'(x) = (1+x)^{-1}.$$

Durch vollständige Induktion nach k zeigt man:

$$f^{(k)}(x) = (-1)^{k+1}(k-1)!(1+x)^{-k}.$$

Also ist

$$f^{(k)}(0) = \begin{cases} 0 & \text{für} \quad k = 0 \\ (-1)^{k+1}(k-1)! & \text{für} \quad k \geq 1. \end{cases}$$

Damit haben wir nach Definition 3.10.2

$$T^{(n)}f(0,x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k.$$

Nach Satz 4.5.1 gilt genau dann

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k,$$

wenn

$$R_{n+1}(0,x) = \frac{f^{(n+1)}(\xi_{n+1})}{(n+1)!} = \frac{(1+\xi_{n+1})^{-(n+1)}}{n+1} x^{n+1} \longrightarrow 0 \quad (n \to \infty).$$

Dabei liegen die ξ_{n+1} zwischen 0 und x.
Es sei zunächst $-1/2 < x < 1$. Dann folgt $|1+\xi_{n+1}| \geq |x|$ und

$$|R_{n+1}(0,x)| = \left| \frac{x}{1+\xi_{n+1}} \right|^{n+1} \longrightarrow 0 \quad (n \to \infty).$$

Damit folgt $(*)$ für $-1/2 < x < 1$.
Nach Satz 4.4.3 und 4.4.4 hat die Potenzreihe in $(*)$ den Konvergenzradius $r = 1$.
Man kann nun auf anderem Wege zeigen, daß $(*)$ für den weiteren Bereich $|x| < 1$ gilt.
Es sei

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

Es ist

$$\frac{d}{dx} \log(1+x) = \frac{1}{1+x}$$

und nach Satz 4.4.5

$$\frac{d}{dx}p(x) = \sum_{n=0}^{\infty}(-x)^n = \frac{1}{1+x}.$$

Also ist

$$\frac{d}{dx}\left(\log(1+x) - p(x)\right) = 0$$

und nach Satz 3.9.5 ist dann $\log(1+x)-p(x) = const$ auf I. Für $-1/2 < x < 1$ ist $\log(1+x)-p(x) = 0$, also ist $const. = 0$. □

5.3 Allgemeine Exponentialfunktionen, Logarithmus- und Potenzfunktionen

In Definition 1.5.9 hatten wir für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ die Potenzen a^n rekursiv definiert. Für festes a ist dann die Abbildung $\mathbb{N} \to \mathbb{R}$, $n \to a^n$ eine Funktion mit Definitionsbereich $\mathbb{N}$, d.h. eine Zahlenfolge. Es stellt sich die Frage, wie dieser Definitionsbereich erweitert werden kann, wenn alle Potenzgesetze nach wie vor gelten sollen. Eine erste Erweiterung zu einer Abbildung $\mathbb{Z} \to \mathbb{R}$ erhält man durch die Definition $a^{-n} := \frac{1}{a^n}$ für $n > 0$ und $a \neq 0$ und $a^0 = 1$ für $a \in \mathbb{R}$. Für $a > 0$ und rationale Exponenten $x = \frac{r}{s}$ läßt sich $a^{r/s}$ durch $a^{r/s} := \sqrt[s]{a^r}$ definieren. Ein Weg, den Definitionsbereich für $a > 0$ auf ganz $\mathbb{R}$ zu erweitern, also a^x für beliebiges $x \in \mathbb{R}$ zu definieren, ist, eine Folge $x_n = \frac{r_n}{s_n}$ rationaler Zahlen mit $\lim_{n\to\infty} x_n = x$ zu wählen und a^x durch

$$a^x := \lim_{n\to\infty} a^{r_n/s_n}$$

zu definieren.
Es ist dann zu zeigen, daß dieser Grenzwert existiert und nur von x und nicht von der gewählten Folge (r_n/s_n) abhängt.
Die Definition der Exponentialfunktion $x \to a^x$ kann tatsächlich auf diese Weise durchgeführt.
Wir werden jedoch einen einfacheren Weg wählen, der auf der bereits definierten Exponentialfunktion $x \to E(x)$ von Abschnitt 5.1 beruht.

Definition 5.3.1. (Allgemeine Exponential-, Logarithmus- und Potenzfunktion)

i) Für $a > 0$ und $x \in \mathbb{R}$ ist $a^x := \exp(x \log a)$.

ii) $e := E(1)$

iii) Für $b > 0$, $b \neq 1$ und $x > 0$ ist $\log_b x := \frac{\log x}{\log b}$.

iv) Für $x > 0$ und $s \in \mathbb{R}$ ist $x^s := \exp(s \log x)$.

Satz 5.3.1. *Für alle $a, b > 0$ und alle $x, y \in \mathbb{R}$ gelten die Potenzgesetze:*

i) $a^x \cdot a^y = a^{x+y}$

ii) $a^x \cdot b^x = (ab)^x$

iii) $(a^x)^y = a^{xy}$

Beweis. ohne Beweis. □

Satz 5.3.2. *Es sei $s \in \mathbb{R}$. Die allgemeine Potenzfunktion $(0,\infty) \to \mathbb{R}$, $x \to x^s$ ist unendlich oft differenzierbar, für $s > 0$ streng monoton wachsend, $\lim_{x\to\infty} x^s = \infty$, für $s < 0$ streng monoton fallend, $\lim_{x\to 0^+} x^s = \infty$, $\lim_{x\to-\infty} x^s = 0$, für $x > 1$ und $x < 0$ streng konvex und für $0 < x < 1$ streng konkav. Außerdem gilt*

$$\frac{d}{dx}x^s = s \cdot x^{s-1}.$$

Beweis. ohne Beweis. □

Definition 5.3.2. (allgemeiner Binomialkoeffizient)
Für $s \in \mathbb{R}$ und $k \in \mathbb{N}$ sei

$$\binom{s}{k} := \frac{s \cdot (s-1) \cdots (s-k+1)}{k!} \quad \text{und} \quad \binom{s}{0} = 1.$$

Satz 5.3.3. *(Binomialreihe)*
Für $s \in \mathbb{R}$ und $|x| < 1$ gilt

$$(1+x)^s = \sum_{k=0}^{\infty} \binom{s}{k} x^k.$$

Beweis. ohne Beweis. □

5.4 Die trigonometrischen Funktionen

Satz 5.4.1. *Die Potenzreihen*

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad \textit{und} \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

konvergieren für alle $x \in \mathbb{R}$.

Beweis. Dies folgt unmittelbar aus Satz 4.4.3 und Satz 4.4.4. □

Definition 5.4.1. Die nach Satz 5.4.1 für alle $x \in \mathbb{R}$ definierte Funktion

$$x \to \sin x := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

(Sprechweise: sinus x) heißt Sinusfunktion.
Die nach Satz 5.4.1 für alle $x \in \mathbb{R}$ definierte Funktion

$$x \to \cos x := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

(Sprechweise: cosinus x) heißt Kosinusfunktion.

Satz 5.4.2. *Die Funktionen $\sin x$ und $\cos x$ sind unendlich oft differenzierbar, und es gilt*

$$\sin' x = \cos x \quad \textit{und} \quad \cos' x = -\sin x.$$

Beweis. Dies folgt aus Satz 4.4.5 mittels gliedweiser Differentiation. □

Lemma 5.4.1. *Es ist*

i) $\cos x > 0$ *für* $x \in (0, \sqrt{2})$

ii) $\cos 2 < 0$

Beweis. i) Es ist

$$\cos x = \sum_{l=0}^{\infty} \left(\frac{x^{4l}}{(4l)!} - \frac{x^{4l+2}}{(4l+2)!} \right) = \sum_{l=0}^{\infty} \frac{x^{4l}}{(4l)!} \left(1 - \frac{x^2}{(4l+1) \cdot (4l+2)} \right)$$

Für $0 < x < \sqrt{2}$ ist $1 - \frac{x^2}{(4l+1)\cdot(4l+2)} > 0$ für alle $l \in \mathbb{N}$.

ii) Es gilt

$$\cos 2 = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \left(\frac{2^6}{6!} - \frac{2^8}{8!} \right) - \ldots \leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = 1 - 2 + \frac{16}{24} = -\frac{1}{3} < 0.$$

□

Definition 5.4.2. Die Zahl π (sprich: pi) ist durch

$$\frac{\pi}{2} := \inf\{x_0 > 0\colon\ \cos x_0 = 0\}$$

definiert.

Satz 5.4.3. *Es gilt* $\sin^2 x + \cos^2 x = 1$ *für alle* $x \in \mathbb{R}$.

Beweis. Es sei $f(x) = \sin^2 x + \cos^2 x$. Nach Satz 5.4.2 und Satz 3.8.4 (Kettenregel) erhalten wir

$$f'(x) = 2 \sin x \cos x - 2 \sin x \cos x = 0.$$

Nach Satz 3.9.5 folgt $f(x) \equiv const.$. Einsetzen von $x = 0$ ergibt $f(x) \equiv 1$. □

Satz 5.4.4. *Es gilt*

i) $2\sqrt{2} < \pi < 4$

ii) $\cos \frac{\pi}{2} = 0$ *und* $\sin \frac{\pi}{2} = 1$.

iii) Auf $\left[0, \frac{\pi}{2}\right]$ *ist* $\sin x$ *streng monoton wachsend und* $\cos x$ *streng monoton fallend.*

Beweis. i) Nach Definition 3.9.2 und Lemma 5.4.1 folgt $\sqrt{2} < \frac{\pi}{2} < 2$.

ii) und (iii) Nach Definition 5.4.2 folgt wegen der Stetigkeit von $\cos x$, daß $\cos \frac{\pi}{2} = 0$ ist. Es ist $\cos x > 0$ für $x \in \left[0, \frac{\pi}{2}\right]$. Nach Satz 3.9.6 (Monotonietest) folgt, daß $\sin x$ auf $\left[0, \frac{\pi}{2}\right]$ streng monoton wachsend ist. Wegen $\sin 0 = 0$ folgt $\sin x > 0$ für $x \in \left[0, \frac{\pi}{2}\right]$. Aus $(\cos x)' = -\sin x$ und Satz 3.9.6 folgt, daß $\cos x$ auf $\left[0, \frac{\pi}{2}\right]$ streng monoton fallend ist. Aus $\cos^2 \frac{\pi}{2} + \sin^2 \frac{\pi}{2} = 1$ und $\cos \frac{\pi}{2} = 0$ folgt dann $\sin \frac{\pi}{2} = 1$.

□

Satz 5.4.5. *(Additionstheoreme)*
Für $x', x \in \mathbb{R}$ *gilt:*

i) $\sin(x'+x) = \cos x' \sin x + \sin x' \cos x$

ii) $\cos(x'+x) = \cos x' \cos x - \sin x' \sin x$

Beweis. i) Für festes $x' \in \mathbb{R}$ setzen wir: $f(x) = \sin(x'+x)$ und bestimmen die Taylorreihe von f um $x_0 = 0$. Die Folge der Ableitungen $f^{(k)}(0)$ ergibt sich als

$$\begin{aligned} f(0) &= &= \sin x' \\ f'(0) &= &= \cos x' \\ f''(0) &= &= -\sin x' \\ f^{(3)}(0) &= &= -\cos x' \end{aligned}$$

und allgemein für alle $m \in \mathbb{N}_0$

$$\begin{aligned} f^{(4m)}(0) &= &= \sin x' \\ f^{(4m+1)}(0) &= &= \cos x' \\ f^{(4m+2)}(0) &= &= -\sin x' \\ f^{(4m+3)}(0) &= &= -\cos x' \end{aligned}$$

Die Taylorreihe ergibt sich als

$$\begin{aligned} Tf(0,x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \\ &= \cos x' \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots \right) + \sin x' \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots \right) \\ &= \cos x' \sin x + \sin x' \cos x. \end{aligned}$$

Für das Restglied gilt

$$R_{n+1}(0,x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1}$$

mit ξ zwischen 0 und x.
Es ist $f^{(n+1)}(\xi) = \pm \sin(x'+\xi)$ oder $f^{(n+1)}(\xi) = \pm \cos(x'+\xi)$, also $\left| f^{(n+1)}(\xi) \right| \leq 1$. Damit gilt $R_{n+1}(0,x) \to 0$ für $n \to 0$. Es folgt

$$\sin(x+x') = f(x) = Tf(0,x) = \cos x' \sin x + \sin x' \cos x.$$

ii) folgt aus (i) durch Differentiation beider Seiten nach x.

□

Satz 5.4.6. *(Periodizität)*

i) $\sin\left(x + \frac{\pi}{2}\right) = \cos x$ *und* $\cos\left(x + \frac{\pi}{2}\right) = -\sin x$

ii) $\sin(x+\pi) = -\sin x$ *und* $\cos(x+\pi) = -\cos x$

iii) $\sin(x+2k\pi) = \sin x$ *und* $\cos(x+2k\pi) = \cos x$ *für alle* $k \in \mathbb{Z}$.

Beweis. i) Dies folgt aus Satz 5.4.5 mit $x' = \frac{\pi}{2}$.

ii) Dies folgt durch zweimalige Anwendung von (i).

iii) Durch zweimalige Anwendung von (ii) folgt zunächst

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

Durch vollständige Induktion nach k folgt die Behauptung für alle $k \in \mathbb{N}$. Ersetzt man x durch $x - 2k\pi$, ergibt sich die Behauptung für alle $k \in \mathbb{Z}$.

□

Satz 5.4.7. *i) Die Funktion* $\sin x$ *ist in* $(0, \pi)$ *positiv und streng konkav; in* $(\pi, 2\pi)$ *negativ und streng konvex, jeweils zwischen den Nullstellen 0 und* π *bzw.* π *und* 2π*, welche auch Wendepunkte sind.*

ii) Die Funktion $\sin x$ *ist im Intervall* $(-\frac{\pi}{2}, \frac{\pi}{2})$ *vom Minimum* $\sin\left(-\frac{\pi}{2}\right) = -1$ *bis zum Maximum* $\sin \frac{\pi}{2} = 1$ *streng monoton wachsend und im Intervall* $(\frac{\pi}{2}, \frac{3\pi}{2})$ *vom Maximum* $\sin \frac{\pi}{2} = 1$ *bis zum Minimum* $\sin \frac{3\pi}{2} = -1$ *streng monoton fallend.*

iii) Die Funktion $\sin x$ *hat die Periode* 2π*. Die Nullstellen sind durch*

$$\sin x = 0 \Leftrightarrow x = k\pi, \quad k \in \mathbb{Z}$$

gegeben.

Beweis. Dies folgt aus den Sätzen 5.4.4 und 5.4.6. □

Satz 5.4.8. *i) Die Funktion* $\cos x$ *ist im Intervall* $(-\frac{\pi}{2}, \frac{\pi}{2})$ *positiv und streng konkav und in* $(\frac{\pi}{2}, \frac{3\pi}{2})$ *negativ und streng konvex, jeweils zwischen den Nullstellen* $-\frac{\pi}{2}$ *und* $\frac{\pi}{2}$ *bzw.* $\frac{\pi}{2}$ *und* $\frac{3\pi}{2}$*, welche auch Wendepunkte sind.*

ii) Die Funktion $\cos x$ *ist im Intervall* $(0, \pi)$ *streng monoton fallend vom Maximum* $\cos 0 = 1$ *bis zum Minimum* $\cos \pi = -1$ *und in* $(\pi, 2\pi)$ *streng monoton wachsend vom Minimum* $\cos \pi = -1$ *bis zum Maximum* $\cos 2\pi = 1$*.*

iii) Die Funktion $\cos x$ *hat die Periode* 2π*. Die Nullstellen sind durch*

$$\cos x = 0 \Leftrightarrow x = \left(k + \frac{1}{2}\right)\pi, \quad k \in \mathbb{Z}$$

gegeben.

Definition 5.4.3. (Tangens, Cotangens)

i) Die Funktion $\tan\colon x \to \tan x$ (Sprechweise: Tangens x) ist durch

$$\tan x := \frac{\sin x}{\cos x} \quad \text{für} \quad x \neq \left(k + \frac{1}{2}\right)\pi,\ k \in \mathbb{Z}$$

definiert.

ii) Die Funktion $\cot\colon x \to \cot x$ (Sprechweise: Kotangens x) ist durch

$$\cot x := \frac{\cos x}{\sin x} \quad \text{für} \quad x \neq k\pi,\ k \in \mathbb{Z}$$

definiert.

Satz 5.4.9. *Die Funktionen* $\tan x$ *und* $\cot x$ *sind in ihren Definitionsbereichen unendlich oft differenzierbar, und es gilt*

$$\begin{aligned}
\tan' x &= 1+\tan^2 x = \frac{1}{\cos^2 x} \quad \textit{für} \quad x \neq \left(k+\frac{1}{2}\right)\pi,\ k\in\mathbb{Z} \quad \textit{und} \\
\cot' x &= -(1+\cot^2 x) = -\frac{1}{\sin^2 x} \quad \textit{für} \quad x\neq\ k\pi,\ k\in\mathbb{Z}.
\end{aligned}$$

Beweis. Dies folgt aus der Quotientenregel und den Differentiationsregeln für $\sin x$ und $\cos x$ (Satz 5.4.2). □

Satz 5.4.10. *(Additionstheoreme)*
Es gilt

i) $\tan(x+x') = \dfrac{\tan x + \tan x'}{1-\tan x\cdot\tan x'}$ *für* $x, x', x+x' \neq \left(k+\frac{1}{2}\right)\pi,\ k\in\mathbb{Z}$

ii) $\cot(x+x') = \dfrac{\cot x\cdot\cot x' - 1}{\cot x + \cot x'}$ *für* $x, x', x+x' \neq k\pi,\ k\in\mathbb{Z}$

Beweis. Wir beweisen nur (i):
Aus Satz 5.4.5 folgt:

$$\tan(x+x') = \frac{\sin(x+x')}{\cos(x+x')} = \frac{\sin x\cos x' + \cos x \sin x'}{\cos x\cos x' - \sin x\sin x'} = \frac{\tan x+\tan x'}{1-\tan x\tan x'}.$$

□

Satz 5.4.11. *(Periodizität)*
Für $x\neq\left(k+\frac{1}{2}\right)\pi$ *und* $x\neq k\pi$ *mit* $k\in\mathbb{Z}$ *gilt:*

$$\begin{aligned}
\tan\left(x+\frac{\pi}{2}\right) &= -\cot x, &\qquad \cot\left(x+\frac{\pi}{2}\right) &= -\tan x \\
\tan(x+\pi) &= \tan x, & \cot(x+\pi) &= \cot x.
\end{aligned}$$

Beweis. Dies folgt aus den entsprechenden Gleichungen für $\sin x$ und $\cos x$ (Satz 5.4.6). □

Satz 5.4.12. *i) Die Funktion* $\tan x$ *ist in* $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ *streng monoton wachsend von* $-\infty$ *bis* ∞, *in* $\left(-\frac{\pi}{2},0\right)$ *streng konkav und in* $\left[0,\frac{\pi}{2}\right)$ *streng konvex mit* $\tan 0 = 0$, *d.h. 0 ist Nullstelle und Wendepunkt.*

ii) Die Funktion $\cot x$ *ist in* $(0,\pi)$ *streng monoton fallend von* $-\infty$ *bis* ∞, *in* $\left(0,\frac{\pi}{2}\right]$ *streng konvex und in* $\left[\frac{\pi}{2},\pi\right]$ *streng konkav mit* $\cot\frac{\pi}{2}=0$, *d.h.* $\frac{\pi}{2}$ *ist Nullstelle und Wendepunkt.*

Beweis. Dies folgt aus den Sätzen 5.4.7 bis 5.4.9. □

5.5 Die Arcusfunktionen

Satz 5.5.1. *i) Die Funktion* $\sin\colon \left[-\frac{\pi}{2},\frac{\pi}{2}\right] \to [-1,1]$, $x \to \sin x$ *besitzt eine Inverse:*

$$\sin^{-1}\colon [-1,1] \to \left[-\frac{\pi}{2},\frac{\pi}{2}\right].$$

ii) Die Funktion $\cos\colon [0,\pi] \to [-1,1]$, $x \to \cos x$ *besitzt eine Inverse:*

$$\cos^{-1}\colon [-1,1] \to [0,\pi].$$

iii) Die Funktion $\tan\colon \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \to \mathbb{R}$, $x \to \tan x$ *besitzt eine Inverse:*

$$\tan^{-1}\colon \mathbb{R} \to \left(-\frac{\pi}{2},\frac{\pi}{2}\right).$$

iv) Die Funktion $\cot\colon (0,\pi) \to \mathbb{R}$, $x \to \cot x$ *besitzt eine Inverse:*

$$\cot^{-1}\colon \mathbb{R} \to (0,\pi).$$

Beweis. Dies folgt aus der strengen Monotonie der Funktionen in den angegebenen Intervallen (Sätze 5.4.7, 5.4.8, 5.4.12). □

Definition 5.5.1. (Arcusfunktionen)

i) Die Inverse des Sinus auf dem Intervall $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ ist der arcsin (Sprechweise: Arcussinus), d.h. für alle $x \in [-1,1]$ und $y \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ gilt

$$\arcsin x = y \Leftrightarrow x = \sin y.$$

ii) Die Inverse des Kosinus auf dem Intervall $[0,\pi]$ ist der arccos (Sprechweise: Arcuskosinus), d.h. für alle $x \in [-1,1]$ und $y \in [0,\pi]$ gilt

$$\arccos x = y \Leftrightarrow x = \cos y.$$

iii) Die Inverse des Tangens auf $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ ist der arctan (Sprechweise: Arcustangens), d.h. für alle $x \in \mathbb{R}$ und $y \in \left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ gilt

$$\arctan x = y \Leftrightarrow x = \tan y.$$

iv) Die Inverse des Kotangens auf $(0,\pi)$ ist der arccot (Sprechweise: Arcuskotangens), d.h. für alle $x \in \mathbb{R}$ und $y \in (0,\pi)$ gilt

$$\operatorname{arccot} x = y \Leftrightarrow x = \cot y.$$

Satz 5.5.2. *Die Arcusfunktionen sind in ihren Definitionsbereichen unendlich oft differenzierbar, und es gilt*

i) $\arcsin' x = \dfrac{1}{\sqrt{1-x^2}}$ *für* $x \in (-1,1)$

ii) $\arccos' x = -\dfrac{1}{\sqrt{1-x^2}}$ *für* $x \in (-1,1)$

iii) $\arctan' x = \dfrac{1}{1+x^2}$ *für* $x \in \mathbb{R}$

iv) $\operatorname{arccot}' x = -\dfrac{1}{1+x^2}$ *für* $x \in \mathbb{R}$.

Beweis. Wir beweisen nur (i). Der Beweis der anderen Teile verläuft ähnlich.

i) Nach Satz 3.8.5 (Differenzierbarkeit der Umkehrfunktion) ist die Arcussinusfunktion als Umkehrfunktion der Sinusfunktion im Intervall $(-1, 1)$ differenzierbar, und es gilt

$$\arcsin' x = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}.$$

□

Satz 5.5.3. *(Taylorreihe)*

i) Für $|x| < 1$ *ist*

$$\arcsin x = \sum_{k=0}^{\infty} (-1)^k \binom{1/2}{k} \frac{x^{2k+1}}{2k+1}.$$

ii) Für $|x| < 1$ *ist*

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

Beweis. Wir gehen analog zur Herleitung der Taylorentwicklung für $\log(1+x)$ vor (Satz 5.2.2): Für die Ableitungen $\arcsin' x$ und $\arctan' x$ sind die Taylorreihen $p_1(x)$ und $p_2(x)$ leicht zu finden. Für die ursprüngliche Funktionen ergeben sich die Taylorreihen durch "Integration" (mehr darüber im nächsten Kapitel). Man findet Potenzreihen, deren Ableitungen $p_1(x)$ bzw. $p_2(x)$ sind. Wir beschreiben die Details:

i) Es ist

$$\arcsin' x = \frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-1/2}.$$

Nach Satz 5.3.3 (Binomialreihe) ist

$$(1-u)^{-1/2} = \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} u^k \quad \text{für } |u| < 1.$$

Die Substitution $u = x^2$ ergibt

$$\arcsin' x = \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} x^{2k} \quad \text{für } |x| < 1.$$

Es sei

$$p_1(x) := \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} \frac{x^{2k+1}}{2k+1}.$$

Man sieht leicht, daß $p_1(x)$ den Konvergenzradius 1 hat. Nach Satz 4.4.5 kann $p_1(x)$ gliedweise differenziert werden:

$$p_1'(x) = \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} x^{2k} \quad \text{für } |x| < 1,$$

also
$$\frac{d}{dx}\left(\arcsin x - p_1(x)\right) = 0.$$
Nach Satz 3.9.5 ist $\arcsin x - p_1(x) = const.$ auf $(-1,1)$. Einsetzen von $x = 0$ ergibt schließlich $p_1(x) = \arcsin x$.

ii) Es ist
$$\arctan' x = \frac{1}{1+x^2}.$$
Nach Satz 2.3.1 ist $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ für $|q| < 1$ (geometrische Reihe). Die Substitution $q = -x^2$ ergibt
$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k} \quad \text{für } |x| < 1.$$
Es sei
$$p_2(x) := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$
Nach Satz 4.4.3 und Satz 4.4.4 hat $p_2(x)$ den Konvergenzradius 1. Nach Satz 4.4.5 kann $p_2(x)$ gliedweise differenziert werden:
$$p_2'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} \quad \text{für } |x| < 1,$$
also
$$\frac{d}{dx}\left(\arctan x - p_2(x)\right) = 0.$$
Nach Satz 3.9.5 ist $\arctan x - p_2(x) = const.$ auf $(-1,1)$. Einsetzen von $x = 0$ ergibt wiederum $p_2(x) = \arctan x$.

□

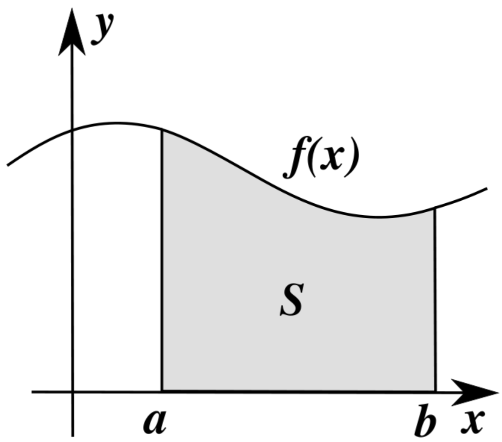
Kapitel 6

Integralrechnung

6.1 Das Flächenproblem

So wie sich die Differentialrechnung historisch aus dem Tangentenproblem entwickelt hat, hat sich die Integralrechnung aus dem Flächenproblem entwickelt. Newtons Entdeckung, daß die Probleme Umkehrungen voneinander sind, kann als Geburtsstunde der modernen Analysis angesehen werden.

Wir beschreiben das Flächenproblem:
Es sei $f\colon [a,b] \to \mathbb{R}$ und $f(x) \geq 0$ für alle $x \in [a,b]$. Wir wollen die Fläche $S = \mathcal{F}$ bestimmen, die von der x- Achse, den Vertikalen $x = a$ und $x = b$ sowie der Kurve $y = f(x)$ begrenzt wird.



Diese Aufgabe besteht nicht im Beweis eines mathematischen Lehrsatzes, sondern in der Erstellung einer Definition. Flächen sind zunächst Eigenschaften von in der Natur vorkommenden Objekten: Diese Anschauung führt uns zu Forderungen an die Eigenschaften, welche der Begriff "Fläche" erfüllen muß:

i) Die Fläche soll additiv sein: sind zwei Teilmengen $\mathcal{M}_1, \mathcal{M}_2 \subseteq \mathbb{R}^2$ disjunkt, d.h. $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$, so sollte die Fläche der Vereinigung $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2$ die Summe der Flächen von $\mathcal{M}_1$ und $\mathcal{M}_2$ sein:
$$\mathcal{F}(\mathcal{M}) = \mathcal{F}(\mathcal{M}_1) + \mathcal{F}(\mathcal{M}_2).$$

ii) Die Fläche eines Rechtecks ist gleich dem Produkt seiner Seitenlängen.

iii) Die Fläche einer Menge $\mathcal{M}$, die in einer Geraden enthalten ist, ist 0.

Aus den Eigenschaften (i)- (iii) ergibt sich folgendes:

1. Ist eine Vereinigung V_1 von nicht überlappenden vertikalen Rechtecken in $\mathcal{F}$ enthalten, so ist die Fläche von $\mathcal{F}$ nicht kleiner als die Fläche von V_1.

2. Ist die Fläche $\mathcal{F}$ in einer Vereinigung V_2 von nicht überlappenden vertikalen Rechtecken enthalten, so ist die Fläche von $\mathcal{F}$ nicht größer als die Fläche von V_2.

Beispiel 6.1.1. Es sei $f(x) = x^2$ sowie $a = 0$ und $b = 1$. Wir zerlegen das Intervall $[0, 1]$ in n Teilintervalle der gleichen Länge $\frac{1}{n}$. Das j- te Teilintervall ist $I_j = \left[\frac{j}{n}, \frac{j+1}{n}\right]$. Dadurch wird auch das Flächenstück $\mathcal{F}$ in n Teilflächen zerlegt, wobei $\mathcal{F}_j$ diejenige Menge an Punkten ist, deren x-Koordinate im Teilintervall I_j liegt:

$$\mathcal{F}_j := \{(x, y) \colon x \in I_j, 0 \leq y \leq f(x)\}.$$

Wir betrachten nun über jedem Teilintervall I_j das größte Rechteck, das in $\mathcal{F}_j$ enthalten ist und das kleinste Rechteck, das $\mathcal{F}_j$ enthält.
Das größte Rechteck über $I_j = \left[\frac{j}{n}, \frac{j+1}{n}\right]$, das in $\mathcal{F}_j$ enthalten ist, ist

$$\underline{R_j} = \left[\frac{j}{n}, \frac{j+1}{n}\right] \times \left[0, \left(\frac{j}{n}\right)^2\right]$$

mit Fläche $\frac{j^2}{n^3}$.
Das kleinste Rechteck über $I_j = \left[\frac{j}{n}, \frac{j+1}{n}\right]$, das $\mathcal{F}_j$ enthält, ist

$$\overline{R_j} = \left[\frac{j}{n}, \frac{j+1}{n}\right] \times \left[0, \left(\frac{j+1}{n}\right)^2\right]$$

mit Fläche $\frac{(j+1)^2}{n^3}$.
Wir können nun die Fläche $\mathcal{F}$ zwischen den Summen der $\underline{R_j}$, der Untersumme

$$\underline{S_n} := \sum_{j=0}^{n-1} \frac{j^2}{n^3}$$

und der Summe der $\overline{R_j}$, der Obersumme

$$\overline{S_n} := \sum_{j=1}^{n} \frac{j^2}{n^3}$$

eingrenzen.
Wie man durch vollständige Induktion leicht beweist, ist

$$\sum_{j=1}^{n} j^2 = \frac{1}{6} n \cdot (n+1) \cdot (2n+1),$$

und damit ist

$$\begin{aligned}\underline{S_n} &= \frac{1}{6n^3}(n-1)\cdot n\cdot(2n-1)\\ \overline{S_n} &= \frac{1}{6n^3}n\cdot(n+1)\cdot(2n+1).\end{aligned}$$

Für $n \to \infty$ haben $\underline{S_n}$ und $\overline{S_n}$ denselben Grenzwert:

$$\lim_{n\to\infty}\underline{S_n} = \lim_{n\to\infty}\overline{S_n} = \frac{1}{3}.$$

Damit sollte die gesuchte Fläche $\frac{1}{3}$ sein, oder $\displaystyle\int_0^1 x^2\,dx = \frac{1}{3}$ (Sprechweise: Integral $x^2\,dx$ von 0 bis 1).

Hiermit haben wir die Grundideen für den Integralbegriff beschrieben. Die Definition soll im nächsten Abschnitt präzise entwickelt werden.

Es werden dabei ein paar Verallgemeinerungen auftreten:

1. Wir verzichten auf die Forderung, daß f positiv sein soll. Das Integral kann somit auch negative Werte annehmen. Dies ist geometrisch so zu verstehen, daß Flächenstücke unter der x- Achse negativ gewertet werden.
2. Die Teilintervalle, die bei der Zerlegung des Intervalls $[a,b]$ auftreten, brauchen nicht gleich lang zu sein.

6.2 Das Riemannsche Integral

Definition 6.2.1. Es sei $I=[a,b]$ mit $a<b\in\mathbb{R}$ ein kompaktes Intervall. Unter einer Zerlegung (oder Partition) $\mathcal{Z}$ von I versteht man ein Tupel $\mathcal{Z}:=(x_0,x_1,\ldots,x_n)$ mit $a=x_0<x_1<\ldots<x_n=b$. Die x_j heißen Teilungspunkte der Zerlegung $\mathcal{Z}$, und $I_j:=[x_j,x_{j+1}]$ mit $j\in\{0,\ldots,n-1\}$ heißt das j- te Teilintervall von $\mathcal{Z}$.

Definition 6.2.2. Es sei $I=[a,b]$ und $a<b\in\mathbb{R}$. Weiter sei $f\colon I\to\mathbb{R}$ eine auf I beschränkte Funktion und $\mathcal{Z}=(x_0,\ldots,x_n)$ eine Zerlegung von I. Es sei

$$M_j := \sup\{f(\xi) \mid \xi\in[x_j,x_{j+1}]\} \quad \text{und} \quad m_j := \inf\{f(\xi) \mid \xi\in[x_j,x_{j+1}]\}$$

mit $j\in\{0,\ldots,n-1\}$.

Unter der Riemannschen Obersumme $\overline{S}(f,\mathcal{Z})$ von f bzgl. der Zerlegung $\mathcal{Z}$ versteht man

$$\overline{S}(f,\mathcal{Z}) := \sum_{j=0}^{n-1} M_j(x_{j+1}-x_j).$$

Unter der Riemannschen Untersumme $\underline{S}(f,\mathcal{Z})$ von f bzgl. der Zerlegung $\mathcal{Z}$ vesteht man

$$\underline{S}(f,\mathcal{Z}) := \sum_{j=0}^{n-1} m_j(x_{j+1}-x_j).$$

Definition 6.2.3. Es sei $I = [a, b]$ mit $a < b \in \mathbb{R}$ ein kompaktes Intervall und $f\colon I \to \mathbb{R}$ eine auf I beschränkte Funktion. Dann heißt

$$\overline{\int_a^b} f(x)\,dx := \inf\{\overline{S}(f, \mathcal{Z}) |\ \mathcal{Z} \text{ Zerlegung von } I\}$$

das Oberintegral von f über I und

$$\underline{\int_a^b} f(x)\,dx := \sup\{\underline{S}(f, \mathcal{Z}) |\ \mathcal{Z} \text{ Zerlegung von } I\}$$

das Unterintegral von f über I.

Falls $\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx$, so schreiben wir für den gemeinsamen Wert

$$\int_a^b f(x)\,dx$$

(Sprechweise: Integral $f(x)\,dx$ von a bis b) und nennen ihn das (Riemann-) Integral von f über I. Die Funktion f heißt dann (Riemann-) integrierbar über I und x heißt Integrationsvariable. Statt "x" kann jedes beliebige Symbol verwendet werden.

Im Rest dieses Abschnitts sei stets $I = [a, b]$ mit $a < b \in \mathbb{R}$ und $f\colon I \to \mathbb{R}$. Jede auf I definierte Funktion wird als beschränkt vorausgesetzt.

Definition 6.2.4. Es seien $\mathcal{Z}, \mathcal{Z}', \mathcal{Z}_1, \mathcal{Z}_2$ Zerlegungen von I.

i) Es sei $\mathcal{Z} = (x_0, \ldots, x_n)$, $I_j = [x_j, x_{j+1}]$ mit $j \in \{0, \ldots, n-1\}$, und $\eta(\mathcal{Z}) := \max_j\{x_{j+1} - x_j\} = \max |I_j|$, die Länge des längsten Teilintervalls von $\mathcal{Z}$, heißt die Feinheit der Zerlegung $\mathcal{Z}$.

ii) Es heißt $\mathcal{Z}'$ Verfeinerung von $\mathcal{Z}$, wenn jeder Teilungspunkt von $\mathcal{Z}$ auch Teilungspunkt von $\mathcal{Z}'$ ist.

iii) Unter der Superposition $\mathcal{Z}_1 \cup \mathcal{Z}_2$ verstehen wir diejenige Zerlegung von I, deren Menge von Teilungspunkten die Vereinigung der Mengen von Teilungspunkten von $\mathcal{Z}_1$ und $\mathcal{Z}_2$ ist.

Satz 6.2.1. *Es sei $|f(x)| \leq M$ für alle $x \in I$, $\mathcal{Z}$ eine Zerlegung von I mit Feinheit η und $\mathcal{Z}'$ eine Verfeinerung von $\mathcal{Z}$.*

i) Es ist stets $\underline{S}(f, \mathcal{Z}) \leq \overline{S}(f, \mathcal{Z})$.

ii) Es ist $\overline{S}(f, \mathcal{Z}') \leq \overline{S}(f, \mathcal{Z})$ und $\underline{S}(f, \mathcal{Z}') \geq \underline{S}(f, \mathcal{Z})$.

iii) Hat $\mathcal{Z}'$ genau l Teilungspunkte mehr als $\mathcal{Z}$, so gilt

$$\overline{S}(f, \mathcal{Z}) - \overline{S}(f, \mathcal{Z}') \leq 2lM\eta \quad \textit{und} \quad \underline{S}(f, \mathcal{Z}') - \underline{S}(f, \mathcal{Z}) \leq 2lM\eta.$$

Beweis. Es genügt, die Behauptungen nur für die Obersummen $\overline{S}$ zu beweisen. Daraus folgen sie wegen $\underline{S}(f, \mathcal{Z}) = -\overline{S}(-f, \mathcal{Z})$ auch für die Untersummen.

i) Es sei $\mathcal{Z} = (x_0, \dots, x_n)$ mit $a = x_0 < \dots < x_n = b$. Nach Definition 6.2.2 ist

$$\begin{aligned}\overline{S}(f,\mathcal{Z}) &= \sum_{j=0}^{n-1} M_j|I_j| \quad \text{mit} \quad M_j = \sup\{f(\xi)|\ \xi \in I_j\}\\ \underline{S}(f,\mathcal{Z}) &= \sum_{j=0}^{n-1} m_j|I_j| \quad \text{mit} \quad m_j = \inf\{f(\xi)|\ \xi \in I_j\}.\end{aligned}$$

Wegen $m_j \leq M_j$ folgt $\underline{S}(f,\mathcal{Z}) \leq \overline{S}(f,\mathcal{Z})$.

ii) Die Teilungspunkte von $\mathcal{Z}'$ seien $a = x_0 < \dots < x_n = b$ mit den Teilintervallen $I'_j = [x_j, x_{j+1}]$. Da $\mathcal{Z}'$ eine Verfeinerung von $\mathcal{Z}$ ist, ist $\mathcal{Z} := (x_{j_0}, \dots, x_{j_\nu})$ mit $j_0, \dots, j_\nu \in \{0, \dots, n\}$ und $a = x_{j_0} < \dots < x_{j_\nu} = b$. Die Teilintervalle I_μ von $\mathcal{Z}$ sind Vereinigungen der Teilintervalle I'_j von $\mathcal{Z}'$. Es ist

$$I_\mu = [x_{j,\mu}, x_{j,\mu+1}] = \bigcup_{j_\mu \leq j \leq j_{\mu+1}-1} [x_j, x_{j+1}] = \bigcup_{j_\mu \leq j \leq j_{\mu+1}-1} I'_j$$

und damit

$$|I_\mu| = \sum_{j_\mu \leq j \leq j_{\mu+1}-1} |I'_j| \quad \text{mit} \quad \mu \in \{0, \dots, \nu-1\}. \tag{1}$$

Außerdem ist

$$M'_j := \sup\{f(\xi)|\ \xi \in I'_j\} \leq M_\mu := \sup\{f(\xi)|\ \xi \in I_\mu\} \tag{2}$$

für $j_\mu \leq j \leq j_{\mu+1} - 1$. Damit ist

$$\overline{S}(f,\mathcal{Z}') = \sum_{j=0}^{n-1} M'_j|I'_j| = \sum_{\mu=0}^{\nu-1} \sum_{j_\mu \leq j \leq j_{\mu+1}-1} M'_j|I'_j| \underset{(2)}{\leq} \sum_{\mu=0}^{\nu-1} M_\mu \sum_{j_\mu \leq j \leq j_{\mu+1}-1} |I'_j| \tag{3}$$

$$\underset{(1)}{=} \sum_{\mu=0}^{\nu-1} M_\mu |I_\mu| = \overline{S}(f,\mathcal{Z}).$$

iii) Die Bezeichnungen seien wie in (ii). Die Anzahl der Teilungspunkte von I_μ, die nicht Teilungspunkte von $\mathcal{Z}$ sind, ist $j_{\mu+1} - j_\mu - 1$. Somit ist

$$\sum_{\mu=0}^{\nu-1} (j_{\mu+1} - j_\mu - 1) = l.$$

Es ist nach (3)

$$\begin{aligned}\overline{S}(f,\mathcal{Z}) - \overline{S}(f,\mathcal{Z}') &= \sum_{\mu=0}^{\nu-1} \sum_{j_\mu \leq j \leq j_{\mu+1}-1} (M_\mu - M'_j)|I'_j|\\ &\leq 2M\eta \sum_{\mu=0}^{\nu-1} (j_{\mu+1} - j_\mu - 1) = 2lM\eta.\end{aligned}$$

□

Satz 6.2.2. *(Ober- und Unterintegral)*
Das Ober- und das Unterintegral existieren, und es ist

$$\underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx.$$

Beweis. Nach Satz 6.2.1 gilt für beliebige Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$

$$\underline{S}(f,\mathcal{Z}_2) \le \underline{S}(f,\mathcal{Z}_1 \cup \mathcal{Z}_2) \le \overline{S}(f,\mathcal{Z}_1 \cup \mathcal{Z}_2) \le \overline{S}(f,\mathcal{Z}_1).$$

Damit existieren

$$\begin{aligned}\overline{\int_a^b} f(x)\,dx &:= \inf\{\overline{S}(f,\mathcal{Z})\colon \mathcal{Z} \text{ Zerlegung von } I\} \quad \text{bzw.}\\ \underline{\int_a^b} f(x)\,dx &:= \sup\{\underline{S}(f,\mathcal{Z})\colon \mathcal{Z} \text{ Zerlegung von } I\},\end{aligned}$$

da die Mengen nach unten bzw. nach oben beschränkt und nichtleer sind.

$$\begin{aligned}&\underline{S}(f,\mathcal{Z}_2) \le \overline{S}(f,\mathcal{Z}_1) \quad \forall \mathcal{Z}_1,\mathcal{Z}_2 \Rightarrow \sup\{\underline{S}(f,\mathcal{Z})\} \le \overline{S}(f,\mathcal{Z}_1) \quad \forall \mathcal{Z}_1\\ \Rightarrow \quad &\underline{\int_a^b} f(x)\,dx = \sup\{\underline{S}(f,\mathcal{Z})\} \le \inf\{\overline{S}(f,\mathcal{Z})\} = \overline{\int_a^b} f(x)\,dx.\end{aligned}$$

□

Satz 6.2.3. *Es ist*

i) $\forall \epsilon > 0\ \exists \delta = \delta(\epsilon) > 0$, *so daß*

$$\left| \overline{S}(f,\mathcal{Z}) - \overline{\int_a^b} f(x)\,dx \right| < \epsilon$$

für alle Zerlegungen $\mathcal{Z}$ *mit* $\eta(\mathcal{Z}) < \delta$.

ii) $\forall \epsilon > 0\ \exists \delta = \delta(\epsilon) > 0$, *so daß*

$$\left| \underline{S}(f,\mathcal{Z}) - \underline{\int_a^b} f(x)\,dx \right| < \epsilon$$

für alle Zerlegungen $\mathcal{Z}$ *mit* $\eta(\mathcal{Z}) < \delta$.

Beweis. Wieder genügt es, nur (i) zu beweisen.
Nach Definition 6.2.3 gibt es eine Zerlegung $\mathcal{Z}_0$, so daß $\overline{S}(f,\mathcal{Z}_0) \le \overline{\int_a^b} f(x)\,dx + \frac{\epsilon}{2}$, wobei $\mathcal{Z}_0$ genau l Teilungspunkte habe. Wir wählen δ so, daß $2Ml\delta < \frac{\epsilon}{2}$ mit $M := \sup\{|f(\xi)|\colon \xi \in I\}$. Es sei $\mathcal{Z}$ eine Zerlegung mit $\eta(\mathcal{Z}) < \delta$. Nach Satz 6.2.1 ist

$$\overline{S}(f,\mathcal{Z} \cup \mathcal{Z}_0) \le \overline{S}(f,\mathcal{Z}_0) \le \overline{\int_a^b} f(x)\,dx + \frac{\epsilon}{2} \tag{1}$$

und

$$\left|\overline{S}(f,\mathcal{Z} \cup \mathcal{Z}_0) - \overline{S}(f,\mathcal{Z})\right| < 2Ml\delta < \frac{\epsilon}{2}. \tag{2}$$

Aus (1) und (2) folgt die Behauptung. □

Definition 6.2.5. Es sei $(\mathcal{Z}_k)_{k=1}^{\infty}$ eine Folge von Zerlegungen von I.
Dann heißt $(\mathcal{Z}_k)$ eine <u>ausgezeichnete Zerlegungsfolge</u>, falls $\lim\limits_{k\to\infty} \eta(\mathcal{Z}_k) = 0$ ist.

Satz 6.2.4. *Es sei* $(\mathcal{Z}_k)$ *eine ausgezeichnete Zerlegungsfolge. Dann ist*

$$\begin{aligned} \lim_{k\to\infty} \overline{S}(f,\mathcal{Z}_k) &= \overline{\int_a^b} f(x)\,dx \quad \textit{und} \\ \lim_{k\to\infty} \underline{S}(f,\mathcal{Z}_k) &= \underline{\int_a^b} f(x)\,dx. \end{aligned}$$

Beweis. Dies folgt sofort aus Satz 6.2.3 und Definition 6.2.5. □

Satz 6.2.5. *(Riemannsches Integrabilitätskriterium)*
Eine Funktion f *ist genau dann über* I *integrierbar, wenn es zu jedem* $\epsilon > 0$ *eine Zerlegung* $\mathcal{Z}$ *von* I *gibt, so daß* $\overline{S}(f,\mathcal{Z}) - \underline{S}(f,\mathcal{Z}) < \epsilon$ *gilt.*

Beweis. "$\Rightarrow$":
Es sei f über I integrierbar. Weiter sei $\epsilon > 0$ und $(\mathcal{Z}_k)_{k=1}^{\infty}$ eine ausgezeichnete Zerlegungsfolge. Nach Satz 6.2.4 ist

$$\lim_{k\to\infty} \overline{S}(f,\mathcal{Z}_k) = \lim_{k\to\infty} \underline{S}(f,\mathcal{Z}_k) = \int_a^b f(x)\,dx.$$

Damit existiert ein $k_0 = k_0(\epsilon)$, so daß für alle $k \geq k_0$ gilt:

$$\left|\overline{S}(f,\mathcal{Z}_k) - \underline{S}(f,\mathcal{Z}_k)\right| < \epsilon.$$

"$\Leftarrow$":
Es seien $\epsilon > 0$ und $\mathcal{Z}$ so gewählt, daß

$$\overline{S}(f,\mathcal{Z}_k) - \underline{S}(f,\mathcal{Z}_k) < \epsilon.$$

Wegen

$$\overline{\int_a^b} f(x)\,dx \leq \overline{S}(f,\mathcal{Z}) \quad \text{und} \quad \underline{\int_a^b} f(x)\,dx \geq \underline{S}(f,\mathcal{Z})$$

folgt

$$\overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx < \epsilon.$$

Da ϵ beliebig ist, folgt

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx.$$

□

Definition 6.2.6. Es sei $\mathcal{Z} = (x_0,\ldots,x_n)$ eine Zerlegung von I. Unter einer Besetzung $\mathcal{B}$ von I (bzgl. $\mathcal{Z}$) versteht man ein Tupel $(\xi_0,\ldots,\xi_{n-1})$ mit $x_j \leq \xi_j \leq x_{j+1}$. Unter der Riemannschen Summe $S(f,\mathcal{Z},\mathcal{B})$ versteht man

$$S(f,\mathcal{Z},\mathcal{B}) := \sum_{j=0}^{n-1} f(\xi_j)(x_{j+1} - x_j).$$

Satz 6.2.6. *Es sei* f *über* I *integrierbar. Weiter sei* $(\mathcal{Z}_k)$ *eine ausgezeichnete Zerlegungsfolge und* $(\mathcal{B}_k)$ *eine Folge zugehöriger Besetzungen. Dann gilt:*

$$\lim_{k\to\infty} S(f,\mathcal{Z}_k,\mathcal{B}_k) = \int_a^b f(x)\,dx.$$

Beweis. Es ist $\underline{S}(f,\mathcal{Z}_k) \leq S(f,\mathcal{Z}_k,\mathcal{B}_k) \leq \overline{S}(f,\mathcal{Z}_k)$. Die Behauptung folgt aus Satz 6.2.4. □

6.3 Integrierbare Funktionen

In diesem Abschnitt sei stets $I = [a, b]$ mit $a < b \in \mathbb{R}$.

Definition 6.3.1. i) Es sei $J := [c, d]$ mit $c < d \in \mathbb{R}$ und $f\colon J \to \mathbb{R}$ eine auf J beschränkte Funktion. Unter der Schwankung $\sigma(f, J)$ der Funktion f auf dem Intervall J versteht man

$$\sigma(f, J) = \sup\{f(\xi)|\ \xi \in J\} - \inf\{f(\xi)|\ \xi \in J\}.$$

ii) Es sei $f\colon I \to \mathbb{R}$ eine auf I beschränkte Funktion und $\mathcal{Z} = (x_0, \ldots, x_n)$ eine Zerlegung von I mit dem j- ten Teilintervall $I_j := [x_j, x_{j+1}]$ mit $j \in \{0, \ldots, n-1\}$. Unter der Schwankungssumme $\tau(f, \mathcal{Z})$ (auf I bzgl. der Zerlegung $\mathcal{Z}$) versteht man

$$\tau(f, \mathcal{Z}) := \sum_{j=0}^{n-1} \sigma(f, I_j)|I_j|.$$

Satz 6.3.1. *Es sei $f\colon I \to \mathbb{R}$ eine auf I beschränkte Funktion und $\mathcal{Z}$ eine Zerlegung von I. Dann ist $\overline{S}(f, \mathcal{Z}) - \underline{S}(f, \mathcal{Z}) = \tau(f, \mathcal{Z})$.*

Beweis. Dies folgt sofort aus der Definition 6.2.2. □

Satz 6.3.2. *Es sei $f\colon I \to \mathbb{R}$ eine auf I beschränkte Funktion. Dann ist f genau dann auf I integrierbar, wenn es für alle $\epsilon > 0$ eine Zerlegung $\mathcal{Z}$ von I mit $\tau(f, \mathcal{Z}) < \epsilon$ gibt.*

Beweis. Dies folgt sofort aus dem Riemannschen Integrabilitätskriterium (Satz 6.2.5) und Satz 6.3.1. □

Satz 6.3.3. *(Eigenschaften der Schwankung)*
Es sei $J = [c, d]$ mit $c < d \in \mathbb{R}$ und $f, g\colon J \to \mathbb{R}$ beschränkte Funktionen, also $|f|, |g| \le M$ auf J. Weiter sei $\alpha \in \mathbb{R}$. Dann gilt:

i) $\sigma(\alpha f, J) = |\alpha|\sigma(f, J)$

ii) $\sigma(f + g, J) \le \sigma(f, J) + \sigma(g, J)$

iii) $\sigma(fg, J) \le M(\sigma(f, J) + \sigma(g, J))$

iv) $\sigma(\max\{f, g\}, J) \le \sigma(f, J) + \sigma(g, J)$.

Beweis. Es sei $M(f) := \sup\{f(\xi)|\ \xi \in J\}$ und $m(f) := \inf\{f(\xi)|\ \xi \in J\}$.

i) Für $\alpha \ge 0$ ist $M(\alpha f) = \alpha M(f)$ und $m(\alpha f) = \alpha m(f)$. Für $\alpha < 0$ ist $M(\alpha f) = \alpha m(f)$ und $m(\alpha f) = \alpha M(f)$.

ii) Es ist $M(f + g) \le M(f) + M(g)$ und $m(f + g) \le m(f) + m(g)$, also $M(f + g) - m(f + g) \le (M(f) - m(f)) + (M(g) - m(g))$.

iii) Für $x_1, x_2 \in J$ ist

$$\begin{aligned} |f(x_1)g(x_1) - f(x_2)g(x_2)| &\le |f(x_1)| \cdot |g(x_1) - g(x_2)| + |g(x_2)| \cdot |f(x_1) - f(x_2)| \\ &\le M(\sigma(f, J) + \sigma(g, J)). \end{aligned}$$

Damit ist auch

$$\sigma(fg, J) \le M(\sigma(f, J) + \sigma(g, J)).$$

iv) Es ist $M(\max\{f,g\}) \in (M(f), M(g))$ und $m(\max\{f,g\}) \in (m(f), m(g))$.

□

Satz 6.3.4. *(Schwankungssumme)*
Es seien $f, g\colon I \to \mathbb{R}$ auf I beschränkt, also $|f|, |g| \le M$ für alle $x \in I$. Weiter sei $\alpha \in \mathbb{R}$ und $\mathcal{Z}$ eine Zerlegung von I. Dann gilt

i) $\tau(\alpha f, \mathcal{Z}) = |\alpha| \tau(f, \mathcal{Z})$

ii) $\tau(f+g, \mathcal{Z}) \le \tau(f, \mathcal{Z}) + \tau(g, \mathcal{Z})$

iii) $\tau(fg, \mathcal{Z}) \le M(\tau(f, \mathcal{Z}) + \tau(g, \mathcal{Z}))$

iv) $\tau(\max\{f,g\}, \mathcal{Z}) \le \tau(f, \mathcal{Z}) + \tau(g, \mathcal{Z})$

Beweis. Die Aussagen (i)- (iv) folgen, wenn man die entsprechenden Gleichungen oder Ungleichungen von Satz 6.3.3 mit $J = I_j$ mit I_j als die Teilintervalle von $\mathcal{Z}$ für jedes $j \in \{0, \dots, n-1\}$ anwendet, dann mit $|I_j|$ multipliziert und über j summiert. □

Satz 6.3.5. *Es seien $f, g\colon I \to \mathbb{R}$ über I integrierbare Funktionen und $\alpha, \beta \in \mathbb{R}$. Dann sind folgende Funktionen über I ebenfalls integrierbar:*

i) $\alpha f + \beta g$

ii) $f \cdot g$

iii) $\max\{f, g\}$

iv) $|f|$

v) Falls $\inf\{g(\xi) |\ \xi \in I\} > 0$ *ist, ist auch fg^{-1} über I integrierbar.*

Beweis. Es sei $(\mathcal{Z}_k)$ eine ausgezeichnete Zerlegungsfolge. Nach Satz 6.2.4 und Satz 6.3.1 ist

$$\lim_{k\to\infty} \tau(f, \mathcal{Z}_k) = \lim_{k\to\infty} \tau(g, \mathcal{Z}_k) = 0.$$

Nach Satz 6.3.4 sind dann auch

$$\lim_{k\to\infty} \tau(\alpha f + \beta g, \mathcal{Z}_k) = 0, \quad \lim_{k\to\infty} \tau(fg, \mathcal{Z}_k) = 0 \quad \text{und} \quad \lim_{k\to\infty} \tau(\max\{f,g\}, \mathcal{Z}_k) = 0.$$

Damit folgen die Behauptungen (i), (ii) und (iii). Aussage (iv) folgt wegen $|f| = \max\{-f, f\}$ aus (iii). Wir verzeichten auf den Beweis von (v). □

Satz 6.3.6. *(Gleichungen und Ungleichungen für Integrale)*
Es seien f, g über I integrierbar und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

i) (Linearität)

$$\int_a^b (\alpha f(x) + \beta g(x))\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx$$

ii) (Erhaltung von Ungleichungen)
Ist $f(x) \leq g(x)$ für alle $x \in I$, so ist

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

iii) (konstantes Integral)

$$\int_a^b \alpha\,dx = \alpha \cdot (b-a).$$

iv) (Mittelwertsatz der Integralrechung)
Gilt $m \leq f(x) \leq M$ für alle $x \in I$, so ist

$$m \cdot (b-a) \leq \int_a^b f(x)\,dx \leq M \cdot (b-a).$$

v) (Dreiecksungleichung)

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$$

vi) (Cauchy- Schwarz)

$$\left(\int_a^b f(x) \cdot g(x)\,dx\right)^2 \leq \int_a^b f(x)^2\,dx \cdot \int_a^b g(x)^2\,dx.$$

Beweis. Es sei $(\mathcal{Z}_k)$ eine ausgezeichnete Zerlegungsfolge mit zugehörigen Besetzungen $\mathcal{B}_k$. Die Teilintervalle von $\mathcal{Z}_k$ seien $I_j^{(k)}$, die Punkte der Besetzung $\mathcal{B}_k$ seien $\xi_j^{(k)}$ mit $0 \leq j \leq n_k - 1$. Dann gilt nach Satz 6.2.6:

$$\lim_{k\to\infty} S(f, \mathcal{Z}_k, \mathcal{B}_k) = \int_a^b f(x)\,dx \quad \text{und} \lim_{k\to\infty} S(g, \mathcal{Z}_k, \mathcal{B}_k) = \int_a^b g(x)\,dx. \qquad (*)$$

i) Es ist

$$S(\alpha f + \beta g, \mathcal{Z}_k, \mathcal{B}_k) = \alpha S(f, \mathcal{Z}_k, \mathcal{B}_k) + \beta S(g, \mathcal{Z}_k, \mathcal{B}_k).$$

Wegen $(*)$ folgt

$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx.$$

ii) Aus $f(x) \leq g(x)$ folgt

$$S(f, \mathcal{Z}_k, \mathcal{B}_k) \leq S(g, \mathcal{Z}_k, \mathcal{B}_k)$$

für alle k. Daher ist

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} S(f, \mathcal{Z}_k, \mathcal{B}_k) \leq \lim_{k\to\infty} S(g, \mathcal{Z}_k, \mathcal{B}_k) = \int_a^b g(x)\,dx.$$

iii) Es ist

$$S(\alpha, \mathcal{Z}_k, \mathcal{B}_k) = \alpha \sum_{j=0}^{n_k-1} |I_j| = \alpha \cdot (b-a).$$

iv) Dies folgt aus (ii) und (iii).

v) Nach der (gewöhnlichen) Dreiecksungleichung (Satz 1.4.10 (iii)) folgt

$$|S(f,\mathcal{Z}_k,\mathcal{B}_k)| = \left|\sum_{j=0}^{n_k-1} f(\xi_j^{(k)})|I_j^{(k)}|\right| \underset{\Delta-Ugl.}{\leq} \sum_{j=0}^{n_k-1} \left|f(\xi_j^{(k)})\right| \left|I_j^{(k)}\right| = S(|f|,\mathcal{Z}_k,\mathcal{B}_k).$$

Damit ist

$$\left|\int_a^b f(x)\,dx\right| = \lim_{k\to\infty} |S(f,\mathcal{Z}_k,\mathcal{B}_k)| \leq \lim_{k\to\infty} S(|f|,\mathcal{Z}_k,\mathcal{B}_k) = \int_a^b |f(x)|\,dx.$$

vi) Nach der (gewöhnlichen) Cauchy- Schwarzschen Ungleichung (siehe Vorlesung in Linearer Algebra) folgt

$$\begin{aligned} S(fg,\mathcal{Z}_k,\mathcal{B}_k)^2 &= \left(\sum_{j=0}^{n_k-1} f(\xi_j^{(k)}) \left|I_j^{(k)}\right|^{1/2} g(\xi_j^{(k)}) \left|I_j^{(k)}\right|^{1/2}\right)^2 \\ &\leq \left(\sum_{j=0}^{n_k-1} f(\xi_j^{(k)})^2 \left|I_j^{(k)}\right|\right) \cdot \left(\sum_{j=0}^{n_k-1} g(\xi_j^{(k)})^2 \left|I_j^{(k)}\right|\right) \\ &= S(f^2,\mathcal{Z}_k,\mathcal{B}_k) \cdot S(g^2,\mathcal{Z}_k,\mathcal{B}_k). \end{aligned}$$

Damit gilt

$$\left(\int_a^b f(x)\cdot g(x)\,dx\right)^2 \leq \int_a^b f(x)^2\,dx \cdot \int_a^b g(x)^2\,dx.$$

□

Wir wollen nun untersuchen, welche der bisher betrachteten Eigenschaften von Funktionen, wie z. B. Stetigkeit, die Integrierbarkeit zur Folge haben.

Satz 6.3.7. *(Integrierbarkeit stetiger Funktionen)*
Es sei $f\colon I \to \mathbb{R}$ auf I stetig. Dann ist f integrierbar.

Beweis. Es sei $\epsilon > 0$. Nach Satz 3.5.3 ist f auf I gleichmäßig stetig. Daher existiert ein $\delta = \delta(\epsilon) > 0$, so daß

$$|f(x_1) - f(x_2)| < \epsilon(b-a)^{-1}$$

für alle x_1, x_2 mit $|x_1 - x_2| < \delta$. Es sei $\mathcal{Z}$ eine Zerlegung mit $\eta(\mathcal{Z}) < \delta$ mit den Teilintervallen I_j mit $j \in \{0,\ldots,n-1\}$. Dann gilt für jedes Teilintervall I_j von $\mathcal{Z}$, daß $\sigma(f,\mathcal{Z}_j) < \epsilon$ ist. So gilt also für die Schwankungssumme

$$\tau(f,\mathcal{Z}) = \sum_{j=0}^{n-1} \sigma(f,I_j)|I_j| < \epsilon(b-a)^{-1} \sum_{j=0}^{n-1} |I_j| = \epsilon.$$

Nach dem Riemannschen Integrabilitätskriterium (Satz 6.2.5) folgt also die Integrierbarkeit von f.
□

Definition 6.3.2. Es sei $\mathcal{Z} = (x_0, \ldots, x_n)$ eine Zerlegung von I. Dann heißt

$$V(f, \mathcal{Z}) = \sum_{j=0}^{n-1} |f(x_{j+1}) - f(x_j)|$$

die Variation von f bzgl. $\mathcal{Z}$ und $V(f) := \sup_{\mathcal{Z}} V(f, \mathcal{Z})$ die Totalvariation von f auf I.
Ist $V(f) < \infty$, so heißt f von beschränkter Variation.

Satz 6.3.8. *Ist f von beschränkter Variation auf I, so ist f integrierbar.*

Beweis. Es sei $\mathcal{Z} = (x_0, \ldots, x_n)$ eine Zerlegung von I. Für die Schwankungssumme $\tau(f, \mathcal{Z})$ gilt:

$$\begin{aligned} \tau(f, \mathcal{Z}) &= \sum_{j=0}^{n-1} \sigma(f, I_j)|I_j| = \sum_{j=0}^{n-1} \left(\sup\{f(\xi) |\ \xi \in I_j\} - \inf\{f(\xi) |\ \xi \in I_j\}\right) |I_j| \\ &\leq \sup \sum_{j=0}^{n-1} \left(\left| f(\xi_j^{(1)}) \right| - \left| f(\xi_j^{(2)}) \right| \right) |I_j| \leq \eta(\mathcal{Z}) V(f). \end{aligned}$$

Damit gilt für jede ausgezeichnete Zerlegungsfolge $(\mathcal{Z}_k)$:

$$\lim_{k \to \infty} \tau(f, \mathcal{Z}_k) = 0.$$

Damit ist f nach dem Riemannschen Integrabilitätskriterium integrierbar. □

Satz 6.3.9. *Es sei $f\colon I \to \mathbb{R}$ auf I monoton. Dann ist $V(f) = |f(b) - f(a)|$. Insbesondere ist f von beschränkter Variation.*

Beweis. Ist f monoton wachsend, so ist für jede Zerlegung $\mathcal{Z} = (x_0, \ldots, x_n)$

$$|f(x_{j+1}) - f(x_j)| = f(x_{j+1}) - f(x_j)$$

und damit

$$V(f, \mathcal{Z}) = \sum_{j=0}^{n-1} (f(x_{j+1}) - f(x_j)) = f(b) - f(a).$$

Ist f monoton fallend, so folgt die Behauptung analog für $-f$. □

Satz 6.3.10. *(Integrierbarkeit monotoner Funktionen)*
Es sei $f : I \to \mathbb{R}$ auf I monoton. Dann ist f auf I integrierbar.

Beweis. Dies folgt aus Satz 6.3.8 und Satz 6.3.9. □

Bemerkung 6.3.1. Satz 6.3.10 zeigt, daß die Klasse der integrierbaren Funktionen größer als die Klasse der stetigen Funktionen ist, da es nichtstetige monotone Funktionen gibt. Weitere Beispiele liefert der folgende Satz:

Satz 6.3.11. *Es seien $f, g\colon I \to \mathbb{R}$, f auf I integrierbar, und es gelte $f(x) = g(x)$ bis auf endlich viele $x_j \in I$. Dann ist auch g integrierbar, und es gilt*

$$\int_a^b f(x)\, dx = \int_a^b g(x)\, dx.$$

Beweis. Es genügt zu zeigen, daß $g - f$ integrierbar ist und

$$\int_a^b g(x) - f(x)\,dx = 0.$$

Wir können also annehmen, daß $f(x)$ auf $I = [a, b]$ die identische Nullfunktion und $g(x) = 0$ für alle $x \in I - \{\xi_1, \ldots, \xi_m\}$ ist. Es sei $\mathcal{Z}$ eine Zerlegung von I mit Teilintervallen $I_1, \ldots, I_{n-1}$. Da jedes ξ_j höchstens zwei Teilintervallen angehört, ist

$$\sup\{f(\xi)|\ \xi \in I_j\} = \inf\{f(\xi)|\ \xi \in I_j\}$$

bis auf höchstens $2m$ Werte von j. Für diese ist

$$\sup\{f(\xi)|\ \xi \in I_j\} \le \max_{1\le j\le m} |g(\xi_j)| \quad \text{und } \inf\{f(\xi)|\ \xi \in I_j\} \ge -\max_{1\le j\le m} |g(\xi_j)|\,.$$

Damit ist

$$\begin{aligned}\overline{S}(g, \mathcal{Z}) &\le 2m \max_{1\le j\le m} |g(\xi_j)|\,\eta(\mathcal{Z})\\ \underline{S}(g, \mathcal{Z}) &\ge -2m \max_{1\le j\le m} |g(\xi_j)|\,\eta(\mathcal{Z}).\end{aligned}$$

Für eine ausgezeichnete Zerlegungsfolge $(\mathcal{Z}_k)_{k=1}^{\infty}$ ist somit

$$\lim_{k\to\infty} \overline{S}(g, \mathcal{Z}_k) = 0 = \lim_{k\to\infty} \underline{S}(g, \mathcal{Z}_k).$$

Nach Satz 6.2.4 ist $\int_a^b g(x)\,dx = 0$. □

Satz 6.3.12. *(Integration über Teilintervalle, Additivität)*

i) Es sei $f\colon I \to \mathbb{R}$ über I integrierbar, und $J \subseteq I$ sei ein Teilintervall von I. Dann ist f auch über J integrierbar.

ii) Es sei $a < c < b$. Dann ist

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

Beweis. i) Es sei $(\mathcal{Z}_k)$ eine ausgezeichnete Zerlegungsfolge, so daß die Endpunkte von J Teilungspunkte von jedem $\mathcal{Z}_k$ sind. Dann ist

$$\tau(f, J, \mathcal{Z}_k) \le \tau(f, I, \mathcal{Z}_k),$$

also nach Satz 6.2.4 und Satz 6.3.1

$$\lim_{k\to\infty} \tau(f, J, \mathcal{Z}_k) = \lim_{k\to\infty} \tau(f, I, \mathcal{Z}_k) = 0.$$

ii) Es sei $(\mathcal{Z}_k)$ eine ausgezeichnete Zerlegungsfolge mit Besetzungen $\mathcal{B}_k$, so daß c Teilungspunkt für jedes I ist. Für jedes $\mathcal{Z}_k$ bilden die Teilungspunkte, die zu $[a, c]$ gehören, eine Zerlegung $\mathcal{Z}_k^{(1)}$ von $[a, c]$, und diejenigen, die zu $[c, b]$ gehören, eine Zerlegung $\mathcal{Z}_k^{(2)}$ von $[c, b]$. Eben so erhält man Besetzungen $\mathcal{B}_k^{(1)}$ und $\mathcal{B}_k^{(2)}$. Also ist

$$\begin{aligned}\int_a^b f(x)\,dx &= \lim_{k\to\infty} S(f, \mathcal{Z}_k, \mathcal{B}_k) = \lim_{k\to\infty} S(f, \mathcal{Z}_k^{(1)}, \mathcal{B}_k^{(1)}) + \lim_{k\to\infty} S(f, \mathcal{Z}_k^{(2)}, \mathcal{B}_k^{(2)})\\ &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.\end{aligned}$$

□

Schließlich wollen wir noch den Integralbegriff auch auf Fälle ausdehnen, in denen die Bedingung $a < b$ nicht erfüllt ist.

Definition 6.3.3. Es sei

i) $\displaystyle\int_a^a f(x)\,dx = 0$, falls $f(a)$ definiert ist.

ii) $\displaystyle\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$, falls $\int_a^b f(x)\,dx$ existiert.

Satz 6.3.13. *Es gilt für $a < c < b$*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

falls alle Ausdrücke definiert sind.

Beweis. ohne Beweis □

6.4 Der Hauptsatz der Differenzial- und Integralrechnung

Der Hauptsatz der Differenzial- und Integralrechnung zeigt, daß die Operationen der Differentiation und Integration Umkehrungen voneinander sind.

Satz 6.4.1. *(Hauptsatz der Differential- und Integralrechnung)*
Es sei $f\colon [a,b] \to \mathbb{R}$ Riemann- integrierbar über $[a,b]$. Weiter sei $F\colon [a,b] \to \mathbb{R}$, $x \to F(x)$ durch

$$F(x) := \int_a^x f(t)\,dt$$

definiert.
Ist f in $x_0 \in [a,b]$ stetig, so ist F in x_0 differenzierbar, und es ist $F'(x_0) = f(x_0)$.

Beweis. Es sei $\epsilon > 0$. Wegen der Stetigkeit von f in x_0 gibt es ein $\delta = \delta(\epsilon) > 0$ mit $|f(t) - f(x_0)| < \epsilon$ für alle t mit $|t - x_0| < \delta$. Es folgt

$$\left|\int_{x_0}^x f(t)\,dt - \int_{x_0}^x f(x_0)\,dt\right| = \left|\int_{x_0}^x f(t)\,dt - |x - x_0| f(x_0)\right| \leq \epsilon |x - x_0|.$$

Damit ist

$$\left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt - f(x_0)\right| < \epsilon \qquad (1)$$

für alle $x \in U_\delta(x_0) \cap [a,b] - \{x_0\}$. Mit (1) folgt

$$\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0).$$

□

Definition 6.4.1. Es sei $I = [a,b]$ mit $a < b \in \mathbb{R}$ und $f\colon I \to \mathbb{R}$. Dann heißt $F\colon I \to \mathbb{R}$ Stammfunktion oder unbestimmtes Integral von f auf I (Schreibweise: $F(x) = \displaystyle\int f(x)\,dx$), falls $F'(x) = f(x)$ für alle $x \in I$.

Satz 6.4.2. *Es sei $a < b \in \mathbb{R}$ und $f\colon [a,b] \to \mathbb{R}$. Weiter seien F_1 und F_2 Stammfunktionen von f. Dann ist $F_1(x) - F_2(x) = const.$ auf $[a,b]$.*

Beweis. Es ist $(F_1(x) - F_2(x))' = 0$. Nach Satz 3.9.5 ist $F_1(x) - F_2(x) = const.$ □

Satz 6.4.3. *Es sei $a < b \in \mathbb{R}$, $f\colon [a,b] \to \mathbb{R}$ und F eine Stammfunktion von f auf $[a,b]$. Dann ist*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Beweis. Nach Satz 6.4.1 ist $G(x) := \displaystyle\int_a^x f(t)\,dt$ eine Stammfunktion von f. Nach Satz 6.4.2 ist $F(x) = G(x) + C$ für eine Konstante $C \in \mathbb{R}$.
Es ist

$$\int_a^b f(t)\,dt = G(b) - G(a) = (G(b) + C) - (G(a) + C) = F(b) - F(a).$$

□

Definition 6.4.2. Für den Ausdruck $F(b) - F(a)$ verwenden wir die Abkürzung $[F(x)]_a^b$ oder $[F(x)]_{x=a}^{x=b}$.
Die Aussage von Satz 6.4.3 ist also

$$\int_a^b f(x)\,dx = [F(x)]_a^b.$$

Satz 6.4.3 ermöglicht es, viele Integrale einfach zu berechnen.

Beispiel 6.4.1. Wir kommen auf das schon in Beispiel 6.1.1 berechnete Integral $\displaystyle\int_0^1 x^2\,dx$ zurück.
Es sei $F(x) = \frac{x^3}{3}$. Dann ist $F'(x) = f(x) = x^2$.
Also gilt nach Satz 6.4.3 und Definition 6.4.2

$$\int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

Im nächsten Abschnitt geben wir eine Übersicht über die wichtigsten unbestimmten Integrale, die Grundintegrale.
In den darauffolgenden Abschnitten beschreiben wir, wie daraus durch Integrationstechniken weitere Integrale gewonnen werden können.

6.5 Grundintegrale

Ist $F(x)$ eine Stammfunktion von $f(x)$, so ist nach Satz 6.4.2 jede Stammfunktion von $f(x)$ von der Form $F(x) + C$. Dieses C heißt auch Integrationskonstante. Der Einfachheit halber wird sie in der folgenden Liste weggelassen. Jede Diffferentiationsregel aus den vorigen Abschnitten führt zu einer Integrationsregel:

i) $\displaystyle\int x^n\,dx = \frac{x^{n+1}}{n+1}$ für $n \in \mathbb{N}$ oder $x \neq 0$, $n \in \mathbb{Z}$, $n \neq -1$

ii) $\displaystyle\int \frac{dx}{x} = \log|x|$ für $x \neq 0$

iii) $\displaystyle\int x^\mu \, dx = \frac{x^{\mu+1}}{\mu+1}$ für $x > 0$, $\mu \in \mathbb{R}$, $\mu \neq -1$

iv) $\displaystyle\int e^x \, dx = e^x$

v) $\displaystyle\int a^x \, dx = \frac{a^x}{\log a}$ für $a > 0$

vi) $\displaystyle\int \sin x \, dx = -\cos x$ und $\displaystyle\int \cos x \, dx = \sin x$

vii) $\displaystyle\int \frac{dx}{\cos^2 x} = \tan x$ für $x \neq \left(k + \frac{1}{2}\right)\pi$ mit $k \in \mathbb{Z}$

viii) $\displaystyle\int \frac{dx}{\sin^2 x} = -\cot x$ für $x \neq k\pi$ mit $k \in \mathbb{Z}$

ix) $\displaystyle\int \cot x \, dx = \log \sin x$ für $\sin x > 0$

x) $\displaystyle\int \tan x \, dx = -\log \cos x$ für $\cos x > 0$

xi) $\displaystyle\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x$ für $|x| < 1$

xii) $\displaystyle\int \frac{dx}{1+x^2} = \arctan x$

6.6 Partielle Integration und Substitution, Integrationstechniken

Satz 6.6.1. *Es sei I ein nicht ausgeartetes Intervall, und $f, g\colon I \to \mathbb{R}$ seien in I differenzierbar. Die Funktion fg' besitze in I eine Stammfunktion. Dann hat auch $f'g$ eine Stammfunktion in I, und es gilt die partielle Integrationsregel*

$$\int f(x) \cdot g'(x) \, dx = f(x) \cdot g(x) - \int f'(x) \cdot g(x) \, dx.$$

Beweis. Es sei

$$F(x) := f(x) \cdot g(x) - \int f(x) \cdot g'(x) \, dx.$$

Dann ist $F(x)$ in I differenzierbar, und es gilt

$$F'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x) - f(x) \cdot g'(x) = f'(x) \cdot g(x).$$

Also ist F eine Stammfunktion von $f'g$ mit $F(x) = \int f'(x) \cdot g(x) \, dx$, und es gilt

$$\int f(x) \cdot g'(x) \, dx = f(x) \cdot g(x) - \int f'(x) \cdot g(x) \, dx.$$

□

Satz 6.6.2. *(Partielle Integration für bestimmte Integrale)*
Es sei $I = [a, b]$ *mit* $a < b \in \mathbb{R}$ *und* $\int_a^b f(x) \cdot g'(x)\,dx$ *existiere.*
Dann existiert auch $\int_a^b f'(x) \cdot g(x)\,dx$*, und es gilt die* *partielle Integrationsregel*

$$\int_a^b f(x) \cdot g'(x)\,dx = [f(x) \cdot g(x)]_a^b - \int_a^b f'(x) \cdot g(x)\,dx.$$

Beweis. Dies folgt aus Satz 6.6.1 und Satz 6.4.3. □

Beispiel 6.6.1. Wir berechnen $\int xe^x\,dx$.
Wir wenden Satz 6.6.1 mit $f(x) = x$ und $g'(x) = e^x$ an. Dann ist

$$\int xe^x\,dx = \int f(x) \cdot g'(x)\,dx = f(x) \cdot g(x) - \int f'(x) \cdot g(x)\,dx = xe^x - \int e^x\,dx = xe^x - e^x.$$

Beispiel 6.6.2. Wir bestimmen $\int e^x \sin x\,dx$.
Wir wenden erneut Satz 6.6.1 an, diesmal mit $f(x) = e^x$ und $g(x) = -\cos x$ an. Dann ist

$$\int e^x \sin x\,dx = \int f(x) \cdot g'(x)\,dx = -e^x \cos x + \int e^x \cos x\,dx.$$

Eine zweite Anwendung von Satz 6.6.1, dieses Mal mit $f(x) = e^x$ und $g(x) = \sin x$ ergibt

$$\int e^x \sin x\,dx = -e^x \cos x + e^x \sin x - \int e^x \sin x\,dx.$$

Also gilt

$$\int e^x \sin x\,dx = \frac{1}{2}\left(e^x \sin x - e^x \cos x\right).$$

Beispiel 6.6.3. Wir bestimmen $\int \log x\,dx$.
Wir setzen $f(x) = \log x$ und $g(x) = x$. Somit ist

$$\int \log x\,dx = \int f(x) \cdot g'(x)\,dx = x \log x - \int 1\,dx = x \log x - x.$$

Satz 6.6.3. *Es seien* $a < b \in \mathbb{R}$ *und* $c < d \in \mathbb{R}$*. Weiter sei* $f\colon [a, b] \to [c, d]$ *in* $[a, b]$ *differenzierbar, und es gelte* $f'(t) \neq 0$ *für* $t \in [a, b]$ *und* $g\colon [c, d] \to \mathbb{R}$*. Besitzt die Funktion* $(g \circ f) \cdot f'\colon [a, b] \to \mathbb{R}$ *in* $[a, b]$ *eine Stammfunktion, so besitzt* g *eine Stammfunktion in* $[c, d]$*, und es gilt die* *Substitutionsregel*

$$\int g(x)\,dx = \int g(f(t)) \cdot f'(t)\,dt \Big|_{t=f^{-1}(x)}.$$

Beweis. Es sei $F(t) = \int g(f(t)) \cdot f'(t)\,dt$. Dann ist $F'(t) = g(f(t)) \cdot f'(t)$. Weiter gilt nach der Kettenregel (Satz 3.8.4)

$$(F \circ f^{-1})'(x) = F'(f^{-1}(x)) \cdot (f^{-1})'(x) = g(f(f^{-1}(x))) \cdot f'(f^{-1}(x)) \frac{1}{f'(f^{-1}(x))} = g(x)$$

für $x \in [c,d]$. Also ist $F \circ f^{-1}$ eine Stammfunktion von g auf $[c,d]$, und es gilt

$$\int g(x)\,dx = F(f^{-1}(x)) = \int g(f(t)) \cdot f'(t)\,dt\bigg|_{t=f^{-1}(x)}.$$

□

Satz 6.6.4. *(Substitution bei bestimmten Integralen)*
Es seien $a < b \in \mathbb{R}$ und $c < d \in \mathbb{R}$. Weiter sei $f\colon [a,b] \to [c,d]$ in $[a,b]$ differenzierbar, und es gelte $f'(t) \neq 0$ für $t \in [a,b]$ und $g\colon [c,d] \to \mathbb{R}$. Besitzt die Funktion $(g \circ f) \cdot f'\colon [a,b] \to \mathbb{R}$ in $[a,b]$ eine Stammfunktion, so besitzt g eine Stammfunktion in $[c,d]$, und es gilt die Substitutionsregel

$$\int_a^b g(f(t)) \cdot f'(t)\,dt = \int_{f(a)}^{f(b)} g(x)\,dx$$

Beweis. Dies folgt aus Satz 6.6.3 und Satz 6.4.3. □

Beispiel 6.6.4. Es ist

$$\int e^{\sin t} \cos t\,dt = \int g(f(t)) \cdot f'(t)\,dt$$

mit $f(t) = \sin t$ und $g(x) = e^x$. Also ist

$$\int e^{\sin t} \cos t\,dt = \int e^x\,dx\bigg|_{x=f(t)} = e^x\bigg|_{x=f(t)} = e^{\sin t}.$$

Definition 6.6.1. (Trigonometrische Substitution)
Unter einer trigonometrischen Substitution versteht man eine Substitution der Art

$$t = \arcsin x. \tag{$*$}$$

Bemerkung 6.6.1. Löst man die Gleichung $(*)$ nach x auf, so erhält man

$$t = \arcsin x \Leftrightarrow x = \sin t.$$

Wegen der trigonometrischen Identität $\sin^2 x + \cos^2 x = 1$ (Satz 5.4.3) erhält man

$$\sqrt{1-x^2} = \cos t.$$

Die trigonometrische Substitution ist also für die Behandlung von Integralen geeignet, in denen der Ausdruck $\sqrt{1-x^2}$ vorkommt.

Beispiel 6.6.5. Was ist $\int \sqrt{1-x^2}\,dx$ für $|x| < 1$?
Mit der Substitution $x = \sin t$ und $\frac{dx}{dt} = \cos t$ folgt

$$\begin{aligned}
\int \sqrt{1-x^2}\,dx &= \int \cos^2 t\,dt \underset{part.Int.}{=} \cos t \sin t + \int \sin^2 t\,dt = \cos t \sin t + \int (1 - \cos^2 t)\,dt \\
&= t + \cos t \sin t - \int \cos^2 t\,dt.
\end{aligned}$$

Also ist

$$\int \cos^2 t\,dt = \frac{1}{2}(t + \cos t \sin t)$$

und damit

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}(t + \cos t \sin t)\bigg|_{t=\arcsin x} = \frac{1}{2}\left(\arcsin x + x\sqrt{1-x^2}\right).$$

Bemerkung 6.6.2. Wenn das Integral den Ausdruck $\sqrt{1+x^2}$ enthält, führt oft die Substitution $x = \sinh t$ zum Ziel.

6.7 Integration rationaler Funktionen, Partialbruchzerlegung

Integrale rationaler Funktionen, also

$$\int R(x)\,dx \quad \text{mit} \quad R(x) = \frac{P(x)}{Q(x)},$$

wobei P und Q Polynome sind, können stets durch elementare Funktionen ausgedrückt werden. Dies beruht auf der Partialbruchzerlegung. Nach Satz 3.4.7 gibt es eine Darstellung $R(x) = P_1(x) + S(x)$, wobei $P_1(x)$ ein Polynom und $S(x)$ eine Summe von Ausdrücken der Form

$$\frac{A_i^{(1)}}{x-x_i} + \frac{A_i^{(2)}}{(x-x_i)^2} + \ldots + \frac{A_i^{(\nu_i)}}{(x-x_i)^{\nu_i}} \quad \text{und} \quad \frac{B_j^{(1)}x + C_j^{(1)}}{x^2+b_jx+c_j} + \frac{B_j^{(2)}x + C_j^{(2)}}{(x^2+b_jx+c_j)^2} + \ldots + \frac{B_j^{(\mu_j)}x + C_j^{(\mu_j)}}{(x^2+b_jx+c_j)^{\mu_j}}$$

ist. Dabei sind $(x-x_i)^{\nu_i}$ und $(x^2+b_jx+c_j)^{\mu_j}$ Faktoren des Nennerpolynoms $Q(x)$, wobei $x^2+b_jx+c_j$ keine reellen Nullstellen hat. Nachdem die Partialbruchzerlegung von $R(x)$ durchgeführt ist, besteht die restliche Aufgabe also darin, Ausdrücke der Form

$$\frac{A}{(x-x_i)^\nu} \quad \text{und} \quad \frac{Bx+C}{(x^2+bx+c)^\mu}$$

zu integrieren.
Wir besprechen im folgenden diese Fälle:

i) $$\int \frac{dx}{(x-x_i)^\nu} = \begin{cases} \frac{1}{1-\nu}(x-x_i)^{-\nu+1}, & \text{falls} \quad \nu \neq 1 \\ \log|x-x_i|, & \text{falls} \quad \nu = 1. \end{cases}$$

ii) $$\int \frac{x}{(x^2+bx+c)^\mu}\,dx = \frac{1}{2}\int \frac{2x+b}{(x^2+bx+c)^\mu}\,dx - \frac{b}{2}\int \frac{dx}{(x^2+bx+c)^\mu}$$

Die Substitution $u = x^2+bx+c$ ergibt

$$\int \frac{2x+b}{(x^2+bx+c)^\mu} = \begin{cases} \frac{1}{1-\mu}(x^2+bx+c)^{-\mu+1}, & \text{falls} \quad \mu \neq 1 \\ \log|x^2+bx+c|, & \text{falls} \quad \mu = 1. \end{cases}$$

Fall (ii) läßt sich somit auf die Behandlung von $\displaystyle\int \frac{dx}{(x^2+bx+c)^\mu}$ zurückführen.

iii) $$\int \frac{dx}{(x^2+bx+c)^\mu}$$

Für $\mu = 1$ haben wir

$$\begin{aligned} \int \frac{dx}{x^2+bx+c} &= \int \frac{dx}{\left(x+\frac{b}{2}\right)^2 + \frac{4c-b^2}{4}} = \frac{4}{4c-b^2}\int \frac{dx}{\left(\frac{2x+b}{\sqrt{4c-b^2}}\right)^2+1} \\ &= \frac{2}{\sqrt{4c-b^2}} \cdot \arctan\left(\frac{2x+b}{\sqrt{4c-b^2}}\right) \end{aligned}$$

Für $\mu > 1$ haben wir

$$\begin{aligned}
\int \frac{dx}{(x^2+bx+c)^{\mu-1}} &= \int \frac{dx}{\left(\left(x+\frac{b}{2}\right)^2+\frac{4c-b^2}{4}\right)^{\mu-1}} \\
&= \left(\frac{4}{4c-b^2}\right)^{\mu-1} \int \frac{dx}{\left(\left(\frac{2x+b}{\sqrt{4c-b^2}}\right)^2+1\right)^{\mu-1}} \\
&= \left(\frac{4}{4c-b^2}\right)^{\mu-3/2} \int \frac{dt}{(t^2+1)^{\mu-1}}\Bigg|_{t=\frac{2x+b}{\sqrt{4c-b^2}}}
\end{aligned}$$

Durch partielle Integration ergibt sich

$$\int \frac{dt}{(t^2+1)^{\mu-1}} = \frac{t}{(t^2+1)^{\mu-1}} + 2\cdot(\mu-1)\int \frac{t^2}{(t^2+1)^{\mu}}\,dt.$$

Wegen

$$\frac{t^2}{(t^2+1)^{\mu}} = \frac{1}{(t^2+1)^{\mu-1}} - \frac{1}{(t^2+1)^{\mu}}$$

gilt

$$\begin{aligned}
\int \frac{dt}{(t^2+1)^{\mu-1}} &= \frac{t}{(t^2+1)^{\mu-1}} + 2\cdot(\mu-1)\int \frac{1}{(t^2+1)^{\mu-1}} - \frac{1}{(t^2+1)^{\mu}}\,dt \\
&= \frac{t}{(t^2+1)^{\mu-1}} + 2\cdot(\mu-1)\int \frac{1}{(t^2+1)^{\mu-1}}\,dt - 2\cdot(\mu-1)\int \frac{1}{(t^2+1)^{\mu}}\,dt,
\end{aligned}$$

also

$$\begin{aligned}
(1+2\cdot(1-\mu))\int \frac{dt}{(t^2+1)^{\mu-1}} &= \frac{t}{(t^2+1)^{\mu-1}} - 2\cdot(\mu-1)\int \frac{dt}{(t^2+1)^{\mu}} \quad \text{bzw.} \\
(2\nu-3)\int \frac{dt}{(t^2+1)^{\mu-1}} &= -\frac{t}{(t^2+1)^{\mu-1}} + 2\cdot(\mu-1)\int \frac{dt}{(t^2+1)^{\mu}},
\end{aligned}$$

und deshalb folgt

$$\begin{aligned}
&\int \frac{dx}{(x^2+bx+c)^{\mu-1}} \\
&= \frac{1}{2\mu-3}\left(\frac{4}{4c-b^2}\right)^{\mu-3/2}\cdot\left(-\frac{2x+b}{\sqrt{4c-b^2}}\cdot\frac{1}{\left(\left(\frac{2x+b}{\sqrt{4c-b^2}}\right)^2+1\right)^{\mu-1}} + 2\cdot(\mu-1)\int \frac{dt}{(t^2+1)^{\mu}}\Bigg|_{t=\frac{2x+b}{\sqrt{4c-b^2}}}\right) \\
&= \frac{1}{2\mu-3}\left(\frac{4}{4c-b^2}\right)^{\mu-3/2}\cdot\left(-\frac{1}{\sqrt{4c-b^2}}\left(\frac{4c-b^2}{4}\right)^{\mu-1}\frac{2x+b}{\left(\frac{2x+b}{2}\right)^2+\frac{4c-b^2}{4}}\right. \\
&\quad \left. +2\cdot(\mu-1)\left(\frac{4c-b^2}{4}\right)^{\mu-1/2}\int \frac{dx}{(x^2+bx+c)^{\mu}}\right) \\
&= \frac{1}{2\cdot(2\mu-3)}\left(-\frac{2x+b}{(x^2+bx+c)^{\mu-1}} + (\mu-1)\cdot(4c-b^2)\int \frac{dx}{(x^2+bx+c)^{\mu}}\right).
\end{aligned}$$

Daraus ergibt sich die Rekursionsformel

$$\int \frac{dx}{(x^2+bx+c)^{\mu}} = \frac{1}{(\mu-1)\cdot(4c-b^2)}\left(\frac{2x+b}{(x^2+bx+c)^{\mu-1}} + 2\cdot(2\mu-3)\int \frac{dx}{(x^2+bx+c)^{\mu-1}}\right).$$

Beispiel 6.7.1. Man bestimme

$$\int \frac{dx}{(x-1)^2 \cdot (x^2+1)}, \quad x \neq 1.$$

Lösung:
Partialbruchzerlegung:
Der Integrand $S(x) = \dfrac{1}{(x-1)^2 \cdot (x^2+1)}$ hat die Form

$$S(x) = \frac{A^{(1)}}{x-1} + \frac{A^{(2)}}{(x-1)^2} + \frac{Bx+C}{x^2+1} = \frac{A^{(1)}(x-1)(x^2+1) + A^{(2)}(x^2+1) + (Bx+C)(x-1)^2}{(x-1)^2 \cdot (x^2+1)} \tag{1}$$

Wir müssen nun die unbekannten Koeffizienten A, B, C bestimmen. "Koeffizientenvergleich" ergibt

$$A^{(1)}(x-1)(x^2+1) + A^{(2)}(x^2+1) + (Bx+C)(x-1)^2 = 1.$$

Durch Einsetzen geeigneter Werte für x können oft einige Terme "zum Verschwinden gebracht" werden. So ergibt Einsetzen von $x = 1$

$$2A^{(2)} = 1 \quad \text{oder } A^{(2)} = \frac{1}{2}. \tag{2}$$

Einsetzen von $x = 0, -1, 2$ führen auf das lineare Gleichungssystem

$$\begin{array}{llll} 2A^{(1)} & & +\,C & = \frac{1}{2} \\ 4A^{(1)} & -\,4B & +\,4C & = 0 \\ 5A^{(1)} & +\,2B & +\,C & = -\frac{3}{2}, \end{array}$$

welches mit Methoden der Linearen Algebra (z. B. dem Gaußalgorithmus) gelöst werden kann. Es ergibt sich

$$A^{(1)} = -3, \quad B = \frac{7}{2} \quad \text{bzw. } C = \frac{13}{2}. \tag{3}$$

Aus (1), (2) und (3) erhalten wir

$$\frac{1}{(x-1)^2 \cdot (x^2+1)} = \frac{-3}{x-1} + \frac{1}{2 \cdot (x-1)^2} + \frac{7x+13}{2 \cdot (x^2+1)}.$$

Damit ist für $x \neq 1$

$$\int \frac{dx}{(x-1)^2 \cdot (x^2+1)} = -3 \log|x-1| - \frac{1}{2 \cdot (x-1)} + \frac{7}{4} \log(x^2+1) + \frac{13}{2} \arctan x.$$

6.8 Uneigentliche Integrale

Der Ausdruck $\displaystyle\int_0^1 \frac{dx}{\sqrt{x}}$ erfüllt nicht die Voraussetzungen von Definition 6.2.3, da der Integrand $\dfrac{1}{\sqrt{x}}$ wegen $\displaystyle\lim_{x \to 0^+} \frac{1}{\sqrt{x}} = \infty$ auf dem Integrationsintervall $[0, 1]$ nicht beschränkt ist. Verkleinert man das Integrationsintervall jedoch zu $[\epsilon, 1]$ mit $\epsilon > 0$ beliebig klein, so ist

$$\int_\epsilon^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_{x=\epsilon}^{x=1} = 2 - 2\sqrt{\epsilon}$$

definiert. Außerdem existiert $\lim\limits_{\epsilon \to 0^+} \int_\epsilon^1 \frac{dx}{\sqrt{x}} = 2$. Wir sagen: Das uneigentliche Integral $\int_0^1 \frac{dx}{\sqrt{x}}$ hat den Wert 2 oder $\int_0^1 \frac{dx}{\sqrt{x}} = 2$. Wir werden außerdem uneigentliche Integrale der Form $\int_a^\infty f(x)\,dx$, $\int_{-\infty}^b f(x)\,dx$ bzw. $\int_{-\infty}^\infty f(x)\,dx$ betrachten.

Definition 6.8.1. (uneigentliche Integrale)

i) Es sei $I = [a, b]$ mit $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{\infty\}$ und $a < b$. Weiter sei $f : I \to \mathbb{R}$ auf jedem kompakten Teilintervall $J = [a, c]$ mit $a < c < b$ beschränkt und Riemann- integrierbar. Dann ist

$$\int_a^b f(x)\,dx := \lim_{c \to b^-} \int_a^c f(x)\,dx,$$

falls der Grenzwert existiert.

ii) Es sei $I = [a, b]$ mit $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R}$ und $a < b$. Dann ist

$$\int_a^b f(x)\,dx := \lim_{c \to a^+} \int_c^b f(x)\,dx,$$

falls der Grenzwert existiert.

iii) Es sei $I = [a, b]$ mit $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$. Dann ist

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

falls die beiden Summanden im Sinne von (i) und (ii) existieren.

In den Fällen (i) bis (iii) heißt jeweils $\int_a^b f(x)\,dx$ das uneigentliche Integral $\int_a^b f(x)\,dx$, falls es nicht im Sinne von Definition 6.2.3 existiert. Dann heißt $\int_a^b f(x)\,dx$ konvergent. Existiert in (i) oder (ii) der Grenzwert nicht, oder existiert in (iii) mindestens einer der beiden Summanden nicht, so heißt das uneigentliche Integral divergent.

Beispiel 6.8.1. Wir berechnen

$$\int_0^\infty e^{-x}\,dx.$$

Es ist

$$\int_0^\infty e^{-x}\,dx = \lim_{b \to \infty} \int_0^b e^{-x}\,dx = \lim_{b \to \infty} \left(1 - e^{-b}\right) = 1.$$

Beispiel 6.8.2. Wir berechnen

$$\int_{-\infty}^\infty e^{-|x|}\,dx.$$

Es ist

$$\begin{aligned} \int_{-\infty}^\infty e^{-|x|}\,dx &= \int_{-\infty}^0 e^x\,dx + \int_0^\infty e^{-x}\,dx = \lim_{a \to -\infty} \int_a^0 e^x\,dx + \lim_{b \to \infty} \int_0^b e^{-x}\,dx \\ &= \lim_{a \to -\infty} (1 - e^a) + \lim_{b \to \infty} \left(1 - e^{-b}\right) = 2. \end{aligned}$$

Beispiel 6.8.3. Es sei

$$\int_0^\infty \frac{1}{x^2}\,dx$$

zu berechnen.
Wir spalten das Integral in die beiden uneigentlichen Teilintegrale

$$\int_0^1 \frac{1}{x^2}\,dx \quad \text{und} \quad \int_1^\infty \frac{1}{x^2}\,dx$$

auf. Es ist

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{b\to\infty} \int_1^b \frac{1}{x^2}\,dx = \lim_{b\to\infty} \left(1 - \frac{1}{b}\right) = 1$$

und

$$\int_0^1 \frac{1}{x^2}\,dx = \lim_{a\to 0^+} \int_a^1 \frac{1}{x^2}\,dx = \lim_{a\to 0^+} \left(\frac{1}{a} - 1\right) \qquad (= \infty).$$

Damit ist das uneigentliche Integral $\displaystyle\int_0^\infty \frac{1}{x^2}\,dx$ divergent.

Einen wichtigen Spezialfall uneigentlicher Integrale behandelt

Satz 6.8.1. *i) Das Integral* $\displaystyle\int_0^1 \frac{1}{x^\alpha}\,dx$ *ist für* $\alpha < 1$ *konvergent und für* $\alpha \geq 1$ *divergent.*

ii) Das Integral $\displaystyle\int_1^\infty \frac{1}{x^\alpha}\,dx$ *ist für* $\alpha > 1$ *konvergent und für* $\alpha \leq 1$ *divergent.*

Beweis. Dies folgt aus Definition 6.8.1 durch Nachrechnen. □

Den Wert eines uneigentlichen Integrals zu berechnen ist oft unmöglich. Es läßt sich jedoch im allgemeinen entscheiden, ob ein Integral konvergiert oder divergiert.
Grundlage dafür sind das Majoranten- und Minorantenkriterium.

Definition 6.8.2. i) Ein konvergentes Integral $\displaystyle\int_a^b g(x)\,dx$ mit $g(x) \geq 0$ in $[a,b]$ heißt eine (konvergente) Majorante von $\displaystyle\int_a^b f(x)\,dx$, falls $f(x) \leq g(x)$ für alle $x \in [a,b]$ gilt.

ii) Ein divergentes Integral $\displaystyle\int_a^b g(x)\,dx$ mit $g(x) \geq 0$ in $[a,b]$ heißt eine (divergente) Minorante von $\displaystyle\int_a^b f(x)\,dx$, falls $f(x) \geq g(x)$ für alle $x \in [a,b]$ gilt.

Satz 6.8.2. *i) Majorantenkriterium:*
Ein uneigentliches Integral mit einer konvergenten Majorante konvergiert.

ii) Minorantenkriterium:
Ein uneigentliches Integral mit einer divergenten Minorante divergiert.

Beweis. ohne Beweis. □

Beispiel 6.8.4. Wir betrachten
$$\int_{-\infty}^{\infty} e^{-x^2}\,dx.$$
Es sei $h\colon x \to e^{-x}$ und $f\colon x \to e^{-x^2}$. Für $x \geq 1$ ist $x^2 \geq x$ und daher ist nach Satz 5.1.2 (Monotonie von $E(x)$)
$$e^{-x^2} \leq e^{-x} \quad \text{für alle } x \in [1,\infty). \tag{1}$$
Die Funktionen f und h sind stetig und daher beschränkt auf $[0,1]$ und $\min\{f(x)|\ x \in [0,1]\} > 0$. Daher gibt es eine Konstante $C > 0$, so daß
$$e^{-x^2} \leq Ce^{-x} \quad \text{für alle } x \in [0,1]. \tag{2}$$
Aus (1) und (2) folgt, daß die Funktion $g\colon [0,\infty) \to \mathbb{R}$, $x \to Ce^{-x}$ eine konvergente Majorante von $f(x) = e^{-x^2}$ ist.
Nach Satz 6.8.2 konvergiert
$$\int_0^{\infty} e^{-x^2}\,dx, \quad \text{d.h. } \lim_{b\to\infty} \int_0^b e^{-x^2}\,dx$$
existiert. Die Substitution $u = -x$ zeigt, daß
$$\int_0^b e^{-x^2}\,dx = \int_{-b}^0 e^{-u^2}\,du.$$
Damit konvergiert auch
$$\int_{-\infty}^0 e^{-x^2}\,dx = \lim_{b\to\infty} \int_{-b}^0 e^{-u^2}\,du.$$
Also konvergiert
$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \int_{-\infty}^0 e^{-x^2}\,dx + \int_0^{\infty} e^{-x^2}\,dx.$$

Beispiel 6.8.5. Es sei
$$\int_0^{\infty} \frac{1-e^{-x}}{x^{3/2}}\,dx.$$
Wir spalten dies in die beiden Teilintegrale
$$\int_0^1 \frac{1-e^{-x}}{x^{3/2}}\,dx \quad \text{und} \quad \int_1^{\infty} \frac{1-e^{-x}}{x^{3/2}}\,dx$$
auf und untersuchen diese auf Konvergenz.

i) $\displaystyle\int_0^1 \frac{1-e^{-x}}{x^{3/2}}\,dx = \int_0^1 \frac{1-e^{-x}}{x} \cdot x^{-1/2}\,dx$

Nach dem Satz von de L'Hospital (Satz 3.11.1) oder nach der Potenzreihenentwicklung für e^x ist
$$\lim_{x\to 0^+} \frac{1-e^{-x}}{x} = \lim_{x\to 0^+} \frac{e^{-x}}{1} = 1.$$
Damit ist die Funktion
$$h(x) := \begin{cases} \dfrac{1-e^{-x}}{x}, & \text{falls} \quad x \in (0,1] \\ 1, & \text{falls} \quad x = 0 \end{cases}$$

auf $[0,1]$ stetig und damit beschränkt. Also gibt es ein $C > 0$ mit $\left|\frac{1-e^{-x}}{x}\right| \leq C$ und somit

$$\frac{1-e^{-x}}{x^{3/2}} \leq \frac{C}{x^{1/2}}.$$

Also besitzt $\int_0^1 \frac{1-e^{-x}}{x^{3/2}}\,dx$ die konvergente Majorante $\int_0^1 \frac{C}{x^{1/2}}\,dx$ und ist damit nach Satz 6.8.2 konvergent.

ii) $\int_1^\infty \frac{1-e^{-x}}{x^{3/2}}\,dx$

Dieses Integral hat die konvergente Majorante $\int_1^\infty \frac{1}{x^{3/2}}\,dx$. Damit ist $\int_1^\infty \frac{1-e^{-x}}{x^{3/2}}\,dx$ konvergent.

iii) Damit ist auch

$$\int_0^\infty \frac{1-e^{-x}}{x^{3/2}}\,dx = \int_0^1 \frac{1-e^{-x}}{x^{3/2}}\,dx + \int_1^\infty \frac{1-e^{-x}}{x^{3/2}}\,dx$$

konvergent.

Uneigentliche Integrale können auch benützt werden, um die Konvergenz oder Divergenz unendlicher Reihen nachzuweisen. Uneigentliche Integrale sind leichter zu berechnen als unendliche Reihen.

Satz 6.8.3. *(Integralkriterium) Es sei $f\colon [0,\infty) \to [0,\infty)$ eine nicht- negative monoton fallende Funktion, und es sei $a_k := f(k)$ für alle $k \in \mathbb{N}_0$. Dann gelten die Ungleichungen*

$$0 \leq \sum_{k=1}^{n} a_k \leq \int_0^n f(x)\,dx \leq \sum_{k=0}^{n-1} a_k$$

für alle $n \in \mathbb{N}$. Insbesondere konvergiert die Reihe $\sum_{k=0}^{\infty} a_k$ genau dann, wenn das uneigentliche Integral $\int_0^\infty f(x)\,dx$ konvergiert.

Beweis. Da f monoton fallend ist, gilt für $k \leq x \leq k+1$

$$a_{k+1} = f(k+1) \leq f(x) \leq f(k) = a_k$$

und damit

$$a_{k+1} = \int_k^{k+1} f(k+1)\,dx \leq \int_k^{k+1} f(x)\,dx \leq \int_k^{k+1} f(k)\,dx = a_k.$$

Summation über $k = 0 \ldots n-1$ ergibt

$$0 \leq \sum_{k=1}^{n} a_k \leq \int_0^n f(x)\,dx \leq \sum_{k=0}^{n-1} a_k$$

für alle $n \in \mathbb{N}$. □

Beispiel 6.8.6. Wir betrachten

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}.$$

Nach (einer leicht variierten Form) von Satz 6.8.3 ist $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ genau dann konvergent, wenn $\int_1^\infty \frac{dx}{x^\alpha}$ konvergiert. Satz 6.8.1 (ii) ergibt: $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ konvergiert für $\alpha > 1$ und divergiert für $\alpha \leq 1$.

Beispiel 6.8.7. Nun ist

$$\sum_{n=2}^{\infty} \frac{1}{n \cdot (\log n)^\alpha}.$$

Die unendliche Reihe konvergiert genau dann, wenn $\int_2^\infty \frac{dx}{x \cdot (\log x)^\alpha}$ konvergiert. Die Substitution $u = \log x$ ergibt

$$\int \frac{dx}{x \cdot (\log x)^\alpha} = \int \frac{du}{u^\alpha}\bigg|_{u=\log x} = \begin{cases} \frac{1}{1-\alpha}(\log x)^{1-\alpha}, & \text{falls} \quad \alpha \neq 1 \\ \log\log x, & \text{falls} \quad \alpha = 1. \end{cases}$$

Damit ist $\sum_{n=2}^{\infty} \frac{1}{n \cdot (\log n)^\alpha}$ für $\alpha > 1$ konvergent und für $\alpha \leq 1$ divergent.

6.9 Vertauschung von Integration und Grenzwertübergang

Satz 6.9.1. *Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Weiter sei $(f_k)_{k=1}^\infty$ eine Folge von Funktionen mit $f_k \colon I \to \mathbb{R}$, wobei die f_k auf I Riemann- integrierbar sind, und $f_k \underset{glm.}{\to} f$ $(k \to \infty)$ auf I. Dann ist f auf I Riemann- integrierbar, und es gilt*

$$\int_a^b f(x)\, dx = \lim_{k\to\infty} \int_a^b f_k(x)\, dx.$$

Beweis. ohne Beweis. □

Als Spezialfall von Satz 6.9.1 erhalten wir

Satz 6.9.2. *Es sei $a < b \in \mathbb{R}$ und $I = [a, b]$. Es sei $\sum_{k=0}^{\infty} f_k(x)$ eine gleichmäßig konvergente Reihe auf I Riemann- integrierbarer Funktionen. Dann definiert $\sum_{k=0}^{\infty} f_k(x)$ eine auf I Riemann- integrierbare Funktion, und es gilt*

$$\int_a^b \sum_{k=0}^{\infty} f_k(x)\, dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)\, dx.$$

Kapitel 7

Der n- dimensionale Raum, Stetigkeit

7.1 Der n- dimensionale Raum, lineare Struktur

Definition 7.1.1. Es sei $n \in \mathbb{N}$. Unter dem n- dimensionalen Raum $\mathbb{R}^n$ versteht man

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n-\text{mal}} = \{\vec{x} = (x_1, \ldots, x_n) |\ x_j \in \mathbb{R},\ 1 \leq j \leq n\}.$$

Wir nennen $\vec{x}$ auch Punkt des $\mathbb{R}^n$, welcher dabei als Zeilenvektor geschrieben wird. Es wird auch der zugehörige Spaltenvektor

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

eine Rolle spielen. Um eine kurze Schreibweise dafür zu haben, machen wir von dem aus der Linearen Algebra bekannten Begriff der Transponierten Gebrauch.
Es sei

$$\mathcal{A} = (a_{jk})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq q}}$$

eine Matrix vom Typ (p, q). Unter der Transponierten $\mathcal{A}^T$ von $\mathcal{A}$ versteht man

$$\mathcal{A}^T = (a_{kj})_{\substack{1 \leq k \leq q \\ 1 \leq j \leq p}}$$

eine Matrix (q, p), die aus $\mathcal{A}$ dadurch hervorgeht, daß die Rolle von Zeilen und Spalten vertauscht wird. Fassen wir den Zeilenvektor $\vec{x} = (x_1, \ldots, x_n)$ als einzeilige Matrix und den dazugehörigen Spaltenvektor als einspaltige Matrix auf, so führt dies zur Schreibweise

$$\vec{x}^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

von der wir in Zukunft Gebrauch machen wollen.

In Analysis I haben wir Eigenschaften von Abbildungen (Funktionen, z. B. Stetigkeit und Differenzierbarkeit, $f \colon X \to Y$, $x \to f(x)$ mit X und Y Teilmengen von $\mathbb{R}$ betrachtet. Man bezeichnet f daher auch als Funktion einer Variablen (bei unserer Wahl ist "x" diese Variable).

In Analysis II betrachten wir nun die allgemeine Situation, daß X und Y Teilmengen von (möglicherweise) höherdimensionalen Räumen sind: $X \subseteq \mathbb{R}^p$ und $Y \subseteq \mathbb{R}^q$ mit $p, q \in \mathbb{N}$. Analysis I ergibt sich dann als Spezialfall $p = q = 1$.

Definition 7.1.2. Die Menge sämtlicher Abbildungen $\vec{f}\colon X \to Y$ mit $X \subseteq \mathbb{R}^p$ und $Y \subseteq \mathbb{R}^q$ bezeichnen wir mit $\mathcal{F}_p^q$. Für den Wert von $\vec{f}$ an der Stelle $\vec{x} \in X$ schreiben wir gewöhnlich $\vec{f}(\vec{x})$, benützen manchmal jedoch auch eine andere Bezeichnungweise, z. B. $\vec{x} = (x, y, z)$ und dann statt $\vec{f}(\vec{x})$ auch $\vec{f}(x, y, z)$, wobei eine Klammer weggelassen wird. Ist $q = 1$, d.h. $Y \subseteq \mathbb{R}$, so schreiben wir die Funktion meist ohne Vektorpfeil: f statt $\vec{f}$.

Beispiel 7.1.1. Es sei $p = 3$ und $q = 1$ sowie $X = \mathbb{R}^3$ und $Y = [0, \infty)$. Es sei $f(\vec{x})$ die Entfernung des Punktes (x, y, z) vom Ursprung $(0, 0, 0)$, also die Länge des Vektors $\vec{x}$. Dann können folgende Schreibweisen benützt werden:

$$f(\vec{x}) = (x^2 + y^2 + z^2)^{1/2} \quad \text{oder} \quad f(x, y, z) = (x^2 + y^2 + z^2)^{1/2}.$$

Es geht nun darum, grundlegende Eigenschaften und Folgerungen daraus, wie z. B. Stetigkeit und Differenzierbarkeit, die in Analysis I für Funktionen einer Variablen definiert wurden, auf Abbildungen in mehreren Variablen zu übertragen. Der Definition der Differenzierbarkeit (Definition 3.7.1) lag der Vergleich mit Funktionen eines besonders einfachen Typs, den linearen Funktionen zugrunde (Satz 3.7.1).
Auch bei der Differenzierbarkeit von Funktionen mehrerer Variablen spielen lineare Abbildungen eine Rolle. Diese basieren auf den linearen Strukturen des $\mathbb{R}^n$, insbesondere der Struktur des Vektorraums, der Gegenstand der Vorlesung "Lineare Algebra" ist.
Diese werden wir in diesem Abschnitt betrachten.

Definition 7.1.3. Es sei $n \in \mathbb{N}$ und $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$.

i) Unter der Summe $\vec{x} + \vec{y}$ verstehen wir $\vec{x} + \vec{y} := (x_1 + y_1, \ldots, x_n + y_n)$.

ii) Es sei $\lambda \in \mathbb{R}$. Unter dem (skalaren) Vielfachen $\lambda\vec{x}$ verstehen wir $\lambda\vec{x} = (\lambda x_1, \ldots, \lambda x_n)$.

iii) Die Standardbasis $\mathcal{B}$ ist $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ mit $\vec{e}_j = (0, \ldots, 1, \ldots, 0)$, das in der j- ten Spalte einen Eintrag besitzt.

Satz 7.1.1. *Es sei $n \in \mathbb{N}$. Es bildet $\mathbb{R}^n$ mit den in Definition 7.1.3 definierten Rechenoperationen einen n- dimensionalen Vektorraum über $\mathbb{R}$ mit Nullvektor $\vec{0} = (0, \ldots, 0)$. Die Standardbasis ist eine Basis des $\mathbb{R}^n$.*

Beweis. Für die Definition der in Satz 7.1.1 vorkommenden Begriffe und seinen Beweis verweisen wir auf die Vorlesung "Lineare Algebra". □

Definition 7.1.4. Es seien $p, q \in \mathbb{N}$.

i) Eine Abbildung $L \in \mathcal{F}_p^q$ ist genau dann linear, wenn für alle $\vec{x}, \vec{y} \in \mathbb{R}^p$ und $\lambda \in \mathbb{R}$

$$\begin{aligned} L(\vec{x} + \vec{y}) &= L(\vec{x}) + L(\vec{y}) \\ L(\lambda\vec{x}) &= \lambda L(\vec{x}) \end{aligned}$$

gilt.

ii) Man nennt $\vec{f} \in \mathcal{F}_p^q$ affin, wenn $\vec{f}(\vec{x}) = \vec{x}_0 + L(\vec{x})$ mit festem $\vec{x}_0 \in \mathbb{R}^q$ gilt.

Aus der Vorlesung "Lineare Algebra" ist folgender Satz bekannt:

Satz 7.1.2. *Eine Abbildung $L \in \mathcal{F}_p^q$ ist genau dann linear, wenn $L(\vec{x}) = \mathcal{A}\vec{x}^T$ mit einer Matrix $\mathcal{A}$ vom Typ (q, p) ist.*

Definition 7.1.5. Es heißt $\mathcal{A}$ Matrix der linearen Abbildung L.

Weiter besitzt $\mathbb{R}^n$ die zusätzliche Struktur eines Euklidischen Vektorraums:

Definition 7.1.6. Es sei $n \in \mathbb{N}$ und $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$.

i) Unter dem inneren Produkt $\vec{x} \cdot \vec{y}$ versteht man
$$\vec{x} \cdot \vec{y} = \sum_{j=1}^{n} x_j y_j.$$

ii) Unter der Norm (Länge) $||\vec{x}||$ von $\vec{x}$ versteht man
$$||\vec{x}|| = \sqrt{\vec{x} \cdot \vec{x}}.$$

Satz 7.1.3. *Es seien $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Die Norm hat folgende Eigenschaften:*

i) $||\vec{x}|| \geq 0$ *und* $||\vec{x}|| = 0 \Leftrightarrow \vec{x} = \vec{0}$ *(Definitheit)*

ii) $||\lambda\vec{x}|| = |\lambda| \cdot ||\vec{x}||$ *(Homogenität)*

iii) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ *für alle* $\vec{x}, \vec{y} \in \mathbb{R}^n$ *(Dreiecksungleichung)*

Beweis. siehe Vorlesung "Lineare Algebra" □

7.2 Der n- dimensionale Raum, topologische Struktur

In der Analysis einer Variablen spielte der Begriff der Umgebung eines Punktes, der wiederum auf dem Abstandsbegriff aufgebaut war, eine wichtige Rolle. Diese Begriffe wollen wir zunächst auf den $\mathbb{R}^n$ übertragen.

Definition 7.2.1. Es seien $\vec{x}_0, \vec{x}, \vec{y} \in \mathbb{R}^n$. Dann heißt $d(\vec{x}, \vec{y}) := ||\vec{x} - \vec{y}||$ Abstand zwischen $\vec{x}$ und $\vec{y}$. Es sei $\delta > 0$. Die δ- Umgebung $U_\delta(\vec{x}_0)$ von $\vec{x}_0$ ist durch $U_\delta(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n | \; d(\vec{x}, \vec{x}_0) < \delta\}$ definiert.

Bemerkung 7.2.1. Für $n = 1$ ist das äquivalent zu Definition 2.1.2.
Die in Definition 7.2.1 eingeführte Abstandsfunktion $d(\cdot)$ ist der Spezialfall einer Metrik. Weiter ist $(\mathbb{R}^n, d)$ ein metrischer Raum.

Definition 7.2.2. Es sei $X \neq \emptyset$ und $d \colon X \times X \to \mathbb{R}$. Dann heißt (X, d) genau dann ein metrischer Raum, wenn d die folgenden Eigenschaften besitzt:
Es seien $x, y, z \in X$.

i) $d(x, y) = d(y, x)$ (Symmetrie)

ii) $d(x, y) \geq 0$ und $d(x, y) = 0 \Leftrightarrow x = y$ (Definitheit)

iii) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Beispiel 7.2.1. Es sei $X = \{f\colon [0,1] \to \mathbb{R} |\ f \text{ stetig}\}$ und

$$d(f,g) := \int_0^1 |f(x) - g(x)|\, dx.$$

Dann ist (X, d) ein metrischer Raum.

Von den Eigenschaften (i)-(iii) in Definition 7.2.2 sind (i) und (iii), die Symmetrie und die Dreiecksungleichung, unmittelbar klar, während (ii) etwas Überlegung erfordert:
Es sei $f(x_0) \neq g(x_0)$. Wegen der Stetigkeit von f und g gibt es dann ein $\epsilon > 0$ und ein δ mit $0 < \delta < \frac{1}{2}$, so daß $|f(x) - g(x)| > \epsilon$ für alle $x \in U_\delta(x_0) \cap [0,1]$. Es folgt

$$\int_0^1 |f(x) - g(x)|\, dx \geq \int_{\max\{0, x_0-\delta\}}^{\min\{x_0+\delta, 1\}} \epsilon\, dx \geq \frac{1}{2}\delta\epsilon > 0.$$

Die Elemente von X werden auch Punkte des metrischen Raumes genannt. Wie Beispiel 7.2.1 zeigt, ist der Begriff des metrischen Raumes allgemeiner als der Begriff des $\mathbb{R}^n$ und umfaßt auch Räume von Funktionen. Dadurch wird es möglich, Ideen und Begriffe, die für den $\mathbb{R}^n$ entwickelt werden, zum Studium von Mengen von Funktionen nutzbar zu machen. Diese Vorgehensweise ist zentral in der mathematischen Disziplin Funktionalanalysis. Wir werden als Beispiel den Banachschen Fixpunktsatz kennenlernen.
Der Begriff der Konvergenz einer Punktfolge kann für jeden metrischen Raum erklärt werden:

Definition 7.2.3. (Konvergenz in metrischen Räumen)
Es sei (X, d) ein metrischer Raum.

i) Es sei $x_0 \in X$ und $\delta > 0$.
   Die δ- Umgebung $U_\delta(x_0)$ des Punktes x_0 ist durch $U_\delta(x_0) := \{x \in X |\ d(x, x_0) < \delta\}$ definiert.

ii) Es sei $(a_k)_{k=1}^\infty$ mit $a_k \in X$ eine Folge von Punkten von X, und es sei $a \in X$.
   Wir sagen:
   Die Folge (a_k) hat den Grenzwert $a \in X$ oder konvergiert gegen a (Schreibweise: $\lim\limits_{k\to\infty} a_k = a$), falls gilt: $\forall \epsilon > 0$, $\exists k_0 = k_0(\epsilon)$, so daß $a_k \in U_\epsilon(a)$, $\forall k \geq k_0$.

Satz 7.2.1. *(Eindeutigkeit des Grenzwertes)*
Es sei (X, d) ein metrischer Raum und $(a_k)_{k=1}^\infty$ mit $a_k \in X$ eine Punktfolge. Dann hat (a_k) höchstens einen Grenzwert.

Beweis. Es sei $\lim\limits_{k\to\infty} a_k = a^{(1)}$ und $\lim\limits_{k\to\infty} a_k = a^{(2)}$.
Annahme: $a^{(1)} \neq a^{(2)}$.
Nach Definition 7.2.2 (ii) (Definitheit) ist dann $d(a^{(1)}, a^{(2)}) > 0$. Es sei $d(a^{(1)}, a^{(2)}) = 2\epsilon$ mit $\epsilon > 0$.
Nach Definition 7.2.3 gibt es ein k_0, so daß für alle $k \geq k_0$ dann $d(a_k, a^{(1)}) < \epsilon$ und $d(a_k, a^{(2)}) < \epsilon$ gilt.
Nach Definition 7.2.2 (ii) (Dreiecksungleichung) folgt: $d(a^{(1)}, a^{(2)}) \leq d(a^{(1)}, a_k) + d(a_k, a^{(2)}) < 2\epsilon$, ein Widerspruch. $\square$

Viele Konvergenzkriterien in $\mathbb{R}$, wie z. B. das Monotoniekriterium ergeben für den allgemeinen Fall eines metrischen Raumes keinen Sinn. Der Begriff der Monotonie beruht auf der Anordnungseigenschaft der reellen Zahlen (Axiome (A1), (A2), (A3) von Abschnitt 1.3). Diese ist im $\mathbb{R}^n$ und für allgemeine metrische Räume nicht gegeben.

Jedoch können die Begriffe "Beschränktheit" und "Cauchyfolge" für beliebige metrische Räume eingeführt werden.

Definition 7.2.4. Es sei (X,d) ein metrischer Raum.

i) Eine Menge $B \subset X$ heißt beschränkt, wenn es $x_0 \in X$ und $M \in \mathbb{R}$ mit $d(x,x_0) < M$ für alle $x \in B$ gibt.

ii) Eine Folge $(a_k)_{k=1}^{\infty}$ mit $a_k \in X$ heißt beschränkt, wenn die Menge $\{a_k |\ k \in \mathbb{N}\}$ beschränkt ist.

Definition 7.2.5. (Cauchyfolgen in metrischen Räumen)
Es sei (X,d) ein metrischer Raum. Eine Folge $(a_k)_{k=1}^{\infty}$ mit $a_k \in X$ heißt Cauchyfolge, falls $\forall \epsilon > 0$, $\exists k_0 = k_0(\epsilon)$, so daß $d(a_{k_1}, a_{k_2}) < \epsilon\ \forall k_1, k_2 \geq k_0(\epsilon)$.

Satz 7.2.2. *(Konvergente Folgen sind Cauchyfolgen)*
Es sei (X,d) ein metrischer Raum. Eine konvergente Folge von Punkten in X ist stets eine Cauchyfolge.

Beweis. Es sei $\lim\limits_{k\to\infty} a_k = a$ und $\epsilon > 0$. Dann gibt es $k_0 = k_0(\epsilon/2)$, so daß für $k \geq k_0$ dann $d(a_k, a) < \frac{\epsilon}{2}$ gilt. Für $k_1, k_2 \geq k_0$ gilt nach Definition 7.2.2 (iii) (Dreiecksungleichung)

$$d(a_{k_1}, a_{k_2}) \leq d(a_{k_1}, a) + d(a, a_{k_2}) < \epsilon.$$

Also ist $(a_k)_{k=1}^{\infty}$ eine Cauchyfolge. □

In metrischen Räumen gilt also der "notwendige Teil" des Cauchykriteriums (Satz 2.1.15). Notwendig für die Konvergenz einer Folge ist, daß sie eine Cauchyfolge ist. Im allgemeinen ist diese Bedingung in metrischen Räumen nicht hinreichend. Metrische Räume, in denen das Cauchykriterium gilt, heißen vollständig.

Definition 7.2.6. Ein metrischer Raum (X,d) heißt vollständig, wenn in ihm jede Cauchyfolge konvergiert.

Da wir uns fast ausschließlich mit dem Spezialfall $X = \mathbb{R}^n$ und $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$ beschäftigen werden, formulieren wir die Konvergenzeigenschaft noch für diesen Fall gesondert (Definition 7.2.3'). Im Grunde ist Definition 7.2.3' überflüssig, da sie sich aus Definition 7.2.3 durch die Spezialisierung $X = \mathbb{R}^n$ und $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$ ergibt.

Definition 7.2.3.' (Konvergenz im $\mathbb{R}^n$)
Es sei $(\vec{a}_k)_{k=1}^{\infty}$ mit $\vec{a}_k \in \mathbb{R}^n$ eine Folge von Punkten im $\mathbb{R}^n$, und es sei $\vec{a} \in \mathbb{R}^n$. Wir sagen: Die Folge $(\vec{a}_k)$ hat den Grenzwert $\vec{a} \in \mathbb{R}^n$ oder konvergiert gegen $\vec{a}$ (Schreibweise: $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$), falls gilt: $\forall \epsilon > 0$, $\exists k_0 = k_0(\epsilon)$, so daß $||\vec{a}_k - \vec{a}|| < \epsilon$, $\forall k \geq k_0(\epsilon)$.

Die Konvergenz im $\mathbb{R}^n$ ist streng mit der in Analysis I behandelten Konvergenz in $\mathbb{R}$ verknüpft: Die Vektoren $\vec{a}_k$ haben n Komponenten $a_{k,j}$ derart, daß $\vec{a}_k = (a_{k,1}, \ldots, a_{k,n})$ ist. Mit der Folge $(\vec{a}_k)_{k=1}^{\infty}$ sind somit n "Komponentenfolgen" verbunden: $(a_{k,1})_{k=1}^{\infty}$, $(a_{k,2})_{k=1}^{\infty}, \ldots, (a_{k,n})_{k=1}^{\infty}$. Die Konvergenz der Punktfolge $(\vec{a}_k)$ ist äquivalent zur Konvergenz sämtlicher n Komponentenfolgen.

Satz 7.2.3. *Es sei $(\vec{a}_k)_{k=1}^{\infty}$ eine Folge von Punkten des $\mathbb{R}^n$ mit $\vec{a}_k = (a_{k,1}, \ldots, a_{k,n})$. Weiter sei $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Es ist genau dann $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$, wenn $\lim\limits_{k\to\infty} a_{k,j} = a_j$ für alle $j \in \{1, \ldots, n\}$.*

Beweis. "$\Rightarrow$":
Es sei $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$. Es ist nun zu zeigen, daß $\lim\limits_{k\to\infty} a_{k,j} = a_j$ für alle $j \in \{1,\dots,n\}$ ist.
Es sei $\epsilon > 0$ gegeben. Dann gibt es nach Definition 7.2.3' ein $k_0 = k_0(\epsilon)$, so daß $||\vec{a}_k - \vec{a}|| < \epsilon$ für alle $k \geq k_0(\epsilon)$ gilt. Daraus folgt für $j \in \{1,2,\dots,n\}$ dann

$$|a_{k,j} - a_j| \leq \left(\sum_{j=1}^{n} (a_{k,j} - a_j)^2\right)^{1/2} = ||\vec{a}_k - \vec{a}|| < \epsilon$$

für alle $k \geq k_0(\epsilon)$. Also ist $\lim\limits_{k\to\infty} a_{k,j} = a_j$.

"$\Leftarrow$":
Nun sei $\lim\limits_{k\to\infty} a_{k,j} = a_j$ für alle $j \in \{1,2,\dots,n\}$, und es ist $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$ zu zeigen.
Wiederum sei $\epsilon > 0$. Dann gibt es nach Definition 2.1.1 ein k_0, so daß $|a_{k,j} - a_j| < \frac{\epsilon}{\sqrt{n}}$ für alle $k \geq k_0(\epsilon)$ ist. Damit gilt

$$||\vec{a}_k - \vec{a}|| = \left|\sum_{j=1}^{n} (a_{k,j} - a_j)^2\right|^{1/2} < \left|n\left(\frac{\epsilon}{\sqrt{n}}\right)^2\right|^{1/2} = \epsilon$$

für alle $k \geq k_0(\epsilon)$. Also ist $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$. $\square$

Aus Satz 7.2.3 ergibt sich sofort

Satz 7.2.4. *Es seien $(\vec{a}_k)_{k=1}^{\infty}$ und $(\vec{b}_k)_{k=1}^{\infty}$ Folgen von Punkten des $\mathbb{R}^n$ und $(\lambda_k)_{k=1}^{\infty}$ eine Folge reeller Zahlen. Weiter sei*

$$\lim_{k\to\infty} \vec{a}_k = \vec{a}, \quad \lim_{k\to\infty} \vec{b}_k = \vec{b} \quad \text{und} \quad \lim_{k\to\infty} \lambda_k = \lambda.$$

Dann gilt

i) $\lim\limits_{k\to\infty} (\vec{a}_k + \vec{b}_k) = \vec{a} + \vec{b}$

ii) $\lim\limits_{k\to\infty} \lambda_k \vec{a}_k = \lambda \vec{a}$

iii) $\lim\limits_{k\to\infty} \vec{a}_k \vec{b}_k = \vec{a}\vec{b}$

iv) $\lim\limits_{k\to\infty} ||\vec{a}_k|| = ||\vec{a}||$

Satz 7.2.5. *(Cauchykriterium im $\mathbb{R}^n$)*
Es sei $(\vec{a}_k)_{k=1}^{\infty}$ eine Folge von Punkten des $\mathbb{R}^n$. Dann konvergiert $(\vec{a}_k)$ genau dann, wenn es eine Cauchyfolge ist, d.h. wenn gilt: $\forall \epsilon > 0\ \exists k_0 = k_0(\epsilon)$, so daß $\forall k_1, k_2 \geq k_0(\epsilon)$ gilt: $||\vec{a}_{k_1} - \vec{a}_{k_2}|| < \epsilon$.

Bemerkung 7.2.2. Nach Definition 7.2.6 ist $(\mathbb{R}^n, d)$ mit $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$ also ein vollständiger metrischer Raum.

Beweis. (Beweis von Satz 7.2.5)
Es sei $\vec{a}_k = (a_{k,1}, \dots, a_{k,n})$. Nach Satz 7.2.4 ist $(\vec{a}_k)$ genau dann konvergent, wenn die Komponentenfolgen $(a_{k,j})_{k=1}^{n}$ für $j \in \{1,\dots,n\}$ konvergieren, also nach Satz 2.1.15 (Cauchykriterium) genau dann, wenn sie Cauchyfolgen sind. Man zeigt leicht, daß dies dazu äquivalent ist, daß $(\vec{a}_k)_{k=1}^{\infty}$ eine Cauchyfolge ist. $\square$

Definition 7.2.7. Es sei (X, d) ein metrischer Raum und $(\vec{a}_k)_{k=1}^{\infty}$ eine Folge von Punkten in X. Dann heißt a Häufungswert (HW) von (a_k), wenn es für alle $\epsilon > 0$ unendlich viele $k \in \mathbb{N}$ mit $a_k \in U_\epsilon(a)$ gibt.

Satz 7.2.6. *Es sei (X, d) ein metrischer Raum, und $a \in X$ sei Häufungswert der Folge (a_k). Dann gibt es eine Teilfolge $(a_{k_j})_{j=1}^{\infty}$ mit $\lim\limits_{j\to\infty} a_{k_j} = a$.*

Beweis. Die Folge (k_j) wird durch vollständige Induktion definiert.
$j = 1$:
Zuerst wird k_1 so gewählt, daß $d(a_{k_1}, a) < 1$ ist.
$j \to j + 1$:
Dann wird k_{j+1} so gewählt, daß $k_{j+1} > k_j$ und $d(a_{k_{j+1}}, a) < \frac{1}{j+1}$ gilt. □

Satz 7.2.7. *(Bolzano- Weierstraß im $\mathbb{R}^n$)*
Eine beschränkte Folge im $\mathbb{R}^n$ hat mindestens einen Häufungswert.

Beweis. Es sei $(\vec{a}_k)_{k=1}^{\infty}$ mit $\vec{a}_k = (a_{k,1}, \ldots a_{k,n})$. Nach Satz 7.2.3 genügt es zu zeigen, daß es eine Folge $(k_l^{(n)})_{n=1}^{\infty}$ gibt, so daß sämtliche Komponentenfolgen $(a_{k_l^{(n)},j})$ mit $j \in \{1, \ldots, n\}$ konvergieren. Wir konstruieren die Folge (k_m) in n Schritten:
Schritt 1:
Da $(a_{k,1})_{k=1}^{\infty}$ beschränkt ist, besitzt diese Folge nach dem Satz von Bolzano- Weierstraß (Satz 2.1.14) eine konvergente Teilfolge $(a_{k_l}^{(n)})_{l=1}^{\infty}$ mit $\lim\limits_{l\to\infty} a_{k_l^{(1)}} = a^{(1)}$.
Schritt r:
Es sei schon die Folge $(k_l^{r-1})_{l,r-1=1}^{\infty}$ so konstruiert, daß $\lim\limits_{l\to\infty} a_{k_{l,r-1}}^{(j)} = a^{(j)}$ für $1 \le j \le r - 1$ ist.
Da die Folge $(a_{k_{l,r-1}}^{(r-1)})$ beschränkt ist, hat sie eine konvergente Teilfolge. Es gibt also eine konvergente Teilfolge $(k_l^{(r)}$ von $(k_l^{(r-1)})$, so daß $\lim\limits_{l\to\infty} a_{k_l^{(r)}} = a^{(r)}$ gilt.
In Schritt n erhalten wir eine Folge $(k_l^{(n)})_{l=1}^{\infty}$, so daß $\lim\limits_{l_n\to\infty} a_{k_l^{(j)}} = a^{(j)}$ für $j \in \{1, \ldots, n\}$ ist.
Nach Satz 7.2.3 folgt $\lim\limits_{l\to\infty} \vec{a}_{k_l^{(n)}} = (a^{(1)}, \ldots, a^{(n)})$. □

In Analysis I haben wir gesehen, daß viele Eigenschaften von stetigen Funktionen, wie Beschränktheit, Existenz von Maximum und Minimum, etc., von Eigenschaften ihres Definitionsbereiches abhängen: es handelte sich meistens um ein kompaktes Intervall.
Für $n > 1$ werden die betrachteten Mengen im allgemeinen komplizierter sein, jedoch spielen auch hier die Eigenschaften abgeschlossen, offen, kompakt und beschränkt eine entscheidende Rolle. Diese topologischen Eigenschaften einer Menge wollen wir im folgenden definieren:
Wir geben zunächst die Definition der ersten beiden Eigenschaften für den allgemeinen Fall eines metrischen Raumes:

Definition 7.2.8. Es sei (X, d) ein metrischer Raum.

i) Die Menge $O \subseteq X$ heißt offen, wenn es für alle $x \in O$ ein $\delta > 0$ gibt, so daß $U_\delta(x) \subset O$ ist.

ii) Die Menge $A \subset X$ heißt abgeschlossen, wenn $X \backslash A$ offen ist.

Definition 7.2.9. Es sei (X, d) ein metrischer Raum, $M \subset X$ und $a \in X$.

i) Man nennt a einen Häufungspunkt (HP) von M, wenn $U_\delta(a) \cap (M \backslash \{a\}) \neq \emptyset$ für alle $\delta > 0$ ist.

ii) Es heißt a ein Berührpunkt (BP) von M, wenn für alle $\delta > 0$ dann $M \cap U_\delta(a) \neq \emptyset$ gilt.

iii) Weiter heißt a ein isolierter Punkt von M, wenn $a \in M$ ist und ein $\delta > 0$ existiert, so daß $U_\delta(a) \cap (M \backslash \{a\}) = \emptyset$ gilt.

iv) Ein äußerer Punkt von M ist a, wenn ein $\delta > 0$ existiert, so daß $U_\delta(a) \cap M = \emptyset$ gilt.

v) Es ist a ein innerer Punkt von M, wenn ein $\delta > 0$ existiert, so daß $U_\delta(a) \subset M$ gilt.

vi) Schließlich heißt a Randpunkt von M, wenn für alle $\delta > 0$ gilt, daß $U_\delta(a) \cap M \neq \emptyset$ und $U_\delta(a) \cap (X \backslash M) \neq \emptyset$ gilt.

Definition 7.2.10. Es sei (X, d) ein metrischer Raum und $M \subset X$.

i) Der Durchschnitt aller abgeschlossener Mengen, die M enthalten, heißt die abgeschlossene Hülle $\overline{M}$ von M.

ii) Die Vereinigung aller offenen Mengen O mit $O \subset M$ heißt offener Kern $\overset{\circ}{M}$ von M.

Zwischen den Begriffen in den Definitionen 7.2.8, 7.2.9 und 7.2.10 bestehen zahlreiche Beziehungen, von denen wir im folgenden einige beschreiben.

Satz 7.2.8. *Es sei (X, d) ein metrischer Raum, $M \subset X$ und $a \in X$. Dann ist a genau dann ein Berührpunkt von M, wenn a Häufungspunkt oder isolierter Punkt von M ist.*

Beweis. Übungen. □

Satz 7.2.9. *Es sei (X, d) ein metrischer Raum und $O \subset X$.*

i) Die Menge O ist genau dann offen, wenn jeder Punkt von O ein innerer Punkt von O ist.

ii) Die Vereinigung beliebig vieler offener Mengen ist offen.

iii) Der Durchschnitt zweier offener Mengen ist offen.

Beweis. i) Dies folgt aus Definition 7.2.4 (i) und der Definition 7.2.5 (v) des inneren Punktes.

ii) Es sei $\{O_j, j \in J\}$ eine Menge offener Mengen und $O = \bigcup_{j \in J} O_j$. Ist $a \in O$, so existiert ein $j \in J$, so daß a innerer Punkt von O_j ist. Damit ist a innerer Punkt von O. Also ist O offen.

□

Satz 7.2.10. *Es sei (X, d) ein metrischer Raum und $M \subset X$.*

i) Der Durchschnitt einer beliebigen Menge von abgeschlossenen Mengen ist abgeschlossen.

ii) Die abgeschlossene Hülle $\overline{M}$ ist abgeschlossen.

Beweis. i) Es sei $\{A_j, j \in J\}$ eine Menge abgeschlossener Mengen. Dann ist $X \backslash A_j$ nach Definition 7.2.8 für alle $j \in J$ offen. Nach den Regeln von de Morgan ist $O = X \backslash (\bigcap_{j \in J} A_j) = \bigcup_{j \in J} (X \backslash A_j)$. Nach Satz 7.2.9 ist O offen. Also ist $\bigcap_{j \in J} A_j$ abgeschlossen.

ii) Dies folgt unmittelbar aus Definition 7.2.10 (i).

□

Satz 7.2.11. *Es sei (X,d) ein metrischer Raum und $M \subset X$. Dann gilt*

i) Ein Punkt $a \in X$ ist genau dann Berührpunkt von M, wenn es Berührpunkt von $\overline{M}$ ist.

ii) Die Menge M ist genau dann abgeschlossen, wenn sie sämtliche Berührpunkte enthält.

iii) Die Menge $\overline{M}$ ist die Menge aller Berührpunkt von M.

iv) Die Menge M ist genau dann abgeschlossen, wenn $M = \overline{M}$ ist.

Beweis. Es sei $\mathcal{B}(M)$ die Menge aller Berührpunkte von M. Wir zeigen zunächst:

$$\text{Der Punkt } a \text{ ist Berührpunkt von } \mathcal{B}(M) \Rightarrow a \text{ ist Berührpunkt von } M. \tag{1}$$

Es sei a ein Berührpunkt von $\mathcal{B}(M)$ und $\delta > 0$. Dann gibt es ein $b \in \mathcal{B}(M)$ mit $b \in U_\delta(a)$. Dann ist $d(a,b) < \delta$, und somit existiert ein $\delta_1 > 0$, so daß $d(a,b) + \delta_1 < \delta$ ist. Da b Berührpunkt von M ist, existiert ein $c \in U_{\delta_1}(b) \cap M$. Dann ist nach der Dreiecksungleichung $d(a,c) < \delta$, also ist $c \in U_\delta(a)$. Also ist a Berührpunkt von M, und damit ist (1) gezeigt.

Wir beginnen nun mit (ii):
"$\Rightarrow$":
Es sei M abgeschlossen. Dann ist $X \backslash M$ offen. Ist $a \in X \backslash M$, so ist es nach Satz 7.2.9 innerer Punkt von $X \backslash M$. Also existiert ein $\delta > 0$, so daß $U_\delta(a) \cap M = \emptyset$ gilt. Nach Definition 7.2.9 (ii) ist a kein Berührpunkt von M.
"$\Leftarrow$":
Die Menge M enthalte alle ihre Berührpunkte. Weiter sei $a \in X \backslash M$. Dann ist a kein Berührpunkt von M. Nach Definition 7.2.9 existiert ein $\delta > 0$ mit $M \cap U_\delta(a) = \emptyset$, d.h. a ist innerer Punkt von $X \backslash M$. Nach Satz 7.2.9 ist $X \backslash M$ offen, also ist nach Definition 7.2.8 die Menge M abgeschlossen.

(iii) Da $\overline{M}$ abgeschlossen ist, folgt nach (ii) sogleich $\overline{M} = \mathcal{B}(\overline{M})$. Daraus und aus $M \subset \overline{M}$ folgt

$$\mathcal{B}(M) \subset \mathcal{B}(\overline{M}) = \overline{M}. \tag{2}$$

Nach (1) enthält $\mathcal{B}(M)$ alle seine Berührpunkte und ist daher nach (ii) abgeschlossen. Wegen

$$\overline{M} = \bigcap_{\substack{M \subset A \\ A \text{ abgeschlossen}}} A$$

folgt

$$\overline{M} \subset \mathcal{B}(M). \tag{3}$$

Aus (2) und (3) folgt $\mathcal{B}(M) = \overline{M}$, also die Behauptung.

Aus (iii) und (1) folgt die Behauptung (i).

Schließlich folgt (iv) aus (ii) und (iii). □

Eine noch allgemeinere Struktur als die des metrischen Raumes ist die des topologischen Raumes. Topologische Räume sind Gegenstand der mathematischen Disziplin Topologie.

Definition 7.2.11. Ein topologischer Raum ist ein Paar (X,T), bestehend aus einer Menge $X \neq \emptyset$ und einer Menge T offener Teilmengen von X, die folgende Bedingungen erfüllen:

i) $X, \emptyset \in T$

ii) $O_1, O_2 \in T \Rightarrow O_1 \cap O_2 \in T$

iii) Die Vereinigung beliebig vieler offener Mengen ist offen.

Satz 7.2.12. *Ein metrischer Raum (X, d) ist stets ein topologischer Raum, wenn die offenen Mengen wie in Definition 7.2.8 (i) definiert wurden.*

Beweis. Es sei T die Menge der Teilmengen O, die im Sinne von Definition 7.2.8 (i) offen sind. Wir überprüfen nun die Eigenschaften von Definition 7.2.11:

i) Es ist klar, daß X und $\emptyset$ offen sind.

ii) Sind O_1 und O_2 offen, so ist nach Satz 7.2.9 (iii) auch $O_1 \cap O_2$ offen.

iii) Es seien O_j für $j \in J$ offen. Nach Satz 7.2.9 (ii) ist auch $O = \bigcup_{j \in J} O_j$ offen.

□

Definition 7.2.12. Es sei (X, T) ein topologischer Raum und $a \in X$. Unter einer offenen Umgebung von a versteht man eine Menge $U \in T$ mit $a \in U$.

Bemerkung 7.2.3. Für metrische Räume (X, d) führt Definition 7.2.12 auf einen Umgebungsbegriff, der weiter gefaßt ist als in Definition 7.2.3 der Begriff der δ- Umgebung $U_\delta(a)$. Nach Definition 7.2.12 ist jede offene Obermenge eines $U_\delta(a)$ eine Umgebung von a.

Auch der Begriff der Konvergenz kann für allgemeine topologische Räume erklärt werden.

Definition 7.2.13. Es sei (X, T) ein topologischer Raum. Weiter sei $(x_k)_{k=1}^{\infty}$ eine Folge mit $x_k \in X$ und $a \in X$. Dann heißt a Grenzwert der Folge $(x_k)_{k=1}^{\infty}$ (Schreibweise: $\lim\limits_{k\to\infty} x_k = a$), wenn es für jede offene Umgebung U von a ein $k_0 = k_0(U)$ gibt, so daß $x_k \in U$ für alle $k \geq k_0$ gilt.

Definition 7.2.14. Es sei (X, T) ein topologischer Raum und $M \subset X$. Eine Menge $\mathcal{U}$ von offenen Mengen von X heißt eine Überdeckung von M, wenn für alle $x \in M$ ein $U = U(x) \in \mathcal{U}$ mit $x \in U$ existiert. Weiter heißt $\mathcal{V} \subset \mathcal{U}$ Teilüberdeckung von $\mathcal{U}$, wenn auch $\mathcal{V}$ eine Überdeckung von M ist. Zudem heißt $\mathcal{V}$ endlich, wenn die Menge $\mathcal{V}$ endlich ist.

Definition 7.2.15. Es sei (X, T) ein topologischer Raum und $K \subset X$.

i) Die Menge K heißt kompakt (oder überdeckungskompakt), falls jede Überdeckung von K eine endliche Teilüberdeckung besitzt.

ii) Weiter heißt K folgenkompakt, falls jede Folge $(x_k)_{k=1}^{\infty}$ mit $x_k \in K$ eine konvergente Teilfolge besitzt, deren Grenzwert in K liegt.

Ist der topologische Raum ein metrischer Raum, so sind die Eigenschaften "kompakt" und "folgenkompakt" äquivalent.

Satz 7.2.13. *Es sei (X, d) ein metrischer Raum und $K \subset X$. Dann ist K genau dann kompakt, wenn es folgenkompakt ist.*

Zur Vorbereitung des Beweises von Satz 7.2.13 beweisen wir zunächst zwei Lemmata:

Lemma 7.2.1. *Es sei (X,d) ein metrischer Raum, $K \subset X$ folgenkompakt und $\epsilon > 0$. Dann gibt es endlich viele Punkte $y_1, \ldots, y_m \in K$, so daß die Umgebungen $U_\epsilon(y_j)$ mit $1 \le j \le m$ die Menge K überdecken, d.h. $K \subset \bigcup_{j=1}^{m} U_\epsilon(y_j)$.*

Beweis. Annahme: Es gibt keine solche endliche Menge $M = \{y_1, \ldots, y_m\}$, so daß $K \subset \bigcup_{j=1}^{m} U_\epsilon(y_j)$.
Wir konstruieren durch vollständige Induktion eine unendliche Folge $(x_k)_{k=1}^\infty$ mit $d(x_i, x_j) \ge \epsilon$ für $i \ne j$.
$k = 1$:
Es sei $x_1 \in K$ beliebig.
$k \to k+1$:
Es seien $x_1, \ldots, x_k \in K$, so daß $d(x_i, x_j) \ge \epsilon$ für $1 \le i, j \le k$ und $i \ne j$. Da die Umgebungen $U_\epsilon(x_i)$ für $1 \le i \le k$ nach der Annahme K nicht überdecken, gibt es ein $x_{k+1} \notin \bigcup_{j=1}^{k} U_\epsilon(x_j)$. Damit ist auch $d(x_{k+1}, x_j) \ge \epsilon$ für alle $j \in \{1, \ldots, k\}$. Somit gilt für jede Teilfolge von (x_k), daß sie keine Cauchyfolge ist und damit nach Satz 7.2.2 nicht konvergiert. Damit enthält (x_k) also keine konvergente Teilfolge, im Widerspruch zur Folgenkompaktheit von K. □

Lemma 7.2.2. *Es sei (X,d) ein metrischer Raum, $K \subset X$ folgenkompakt und $L \subset K$ abgeschlossen. Dann ist auch L folgenkompakt.*

Beweis. Es sei $(x_k)_{k=1}^\infty$ mit $x_k \in L$. Wegen der Folgenkompaktheit von K enthält (x_k) eine konvergente Teilfolge $(x_{k_j})_{j=1}^\infty$ mit $\lim_{j\to\infty} x_{k_j} = x^*$. Dann ist x^* Berührpunkt von $\{x_{k_j} |\ j \in \mathbb{N}\}$ und damit auch Berührpunkt von L. Nach Satz 7.2.11 ist $x^* \in L$. □

Beweis. (Beweis von Satz 7.2.13):
"$\Rightarrow$":
Es sei $K \subset X$ kompakt, und $(x_k)_{k=1}^\infty$ sei eine Folge mit $x_k \in K$.
Annahme: Die Folge (x_k) hat keinen Häufungswert in K.
Dann gibt es für alle $y \in K$ ein $\delta = \delta(y) > 0$, so daß $x_k \in U_{\delta(y)}(y)$ für höchstens endlich viele $k \in \mathbb{N}$ gilt. Wegen $K \subset \bigcup_{y \in K} U_{\delta(y)}(y)$ ist $\mathcal{U} = \{U_{\delta(y)}(y) |\ y \in K\}$ eine Überdeckung von K. Wegen der Kompaktheit von K hat $\mathcal{U}$ eine endliche Teilüberdeckung

$$\mathcal{V} = \{U_{\delta(y_j)}(y_j) |\ j \in \{1, \ldots, n\}\}.$$

Nach Konstruktion liegen höchstens endlich viele x_k in jedem $U_{\delta(y)}(y)$ und daher auch höchstens endlich viele in

$$\bigcup_{j=1}^{n} U_{\delta(y_j)}(y_j) \supset K,$$

ein Widerspruch, da $x_k \in K$ für alle k ist.
Damit hat (x_k) einen Häufungswert $x^* \in K$ und deshalb auch eine konvergente Teilfolge mit Grenzwert $x^* \in K$. Somit ist K folgenkompakt.

"$\Leftarrow$":
Es sei $K \subset X$ folgenkompakt.
Annahme: Es gibt eine Überdeckung $\mathcal{U}$ von K, die keine endliche Teilüberdeckung besitzt.

Wir konstruieren durch vollständige Induktion eine Folge $(x_k)_{k=1}^{\infty}$ mit $x_k \in K$, so daß $\overline{U_{2^{-k}}(x_k)}$ durch keine endliche Teilüberdeckung von $\mathcal{U}$ überdeckt wird.

<u>$k = 0$</u>:
Nach Lemma 7.2.1 gibt es endlich viele Punkte $y_{1,0}, \dots, y_{m_0,0}$, so daß die Umgebungen $U_{2^{-l}}(y_{j,l})$ mit $j \in \{1, \dots, m_0\}$ die Menge K überdecken. Dann gibt es mindestens ein $j_0 \in \{1, \dots, m_0\}$, so daß $\overline{U_{2^{-l}}(y_{j,l})}$ durch keine endliche Teilüberdeckung von $\mathcal{U}$ überdeckt wird. Wir setzen $x_0 := y_{j_0,0}$.

<u>$k \to k+1$</u>:
Es werde nun $\overline{U_{2^{-k}}(x_k)}$ durch keine endliche Teilüberdeckung von $\mathcal{U}$ überdeckt. Nach Lemma 7.2.1 und 7.2.2 gibt es endlich viele Punkte $y_{1,k+1}, \dots, y_{m_{k+1},k+1}$, so daß die Umgebungen $U_{2^{-(k+1)}}(y_{j,k+1})$ mit $j \in \{1, \dots, m_{k+1}\}$ die Menge $\overline{U_{2^{-k}}(x_k)}$ überdecken. Für mindestens ein $j_0 \in \{1, \dots, m_{k+1}\}$ wird $\overline{U_{2^{-(k+1)}}(y_{j_0,k+1})}$ durch keine endliche Teilüberdeckung von $\mathcal{U}$ überdeckt. Setze $x_{k+1} := y_{j_0,k+1}$.

Wegen der Folgenkompaktheit von K gibt es eine Teilfolge $(x_{k_j})_{j=1}^{\infty}$ von (x_k) mit

$$\lim_{j\to\infty} x_{k_j} = x^* \in K.$$

Da (x_k) wegen $d(x_k, x_{k+1}) \le 2^{-k}$ eine Cauchyfolge ist, ist auch

$$\lim_{k\to\infty} x_k = x^*.$$

Es gibt ein $U^* \in \mathcal{U}$ mit $x^* \in U^*$. Da U^* offen ist, gibt es ein $\epsilon > 0$ mit $U_\epsilon(x^*) \subset U^*$. Wir wählen k_0 derart, daß $x_k \in U_{\epsilon/2}(x^*)$ für $k \ge k_0$ und wählen $k_1 \ge k_0$ so, daß $2^{-k_1} \le \frac{\epsilon}{4}$ ist. Nach der Dreiecksungleichung ist dann $\overline{U_{2^{-k_1}}(x_{k_1})} \subset U_\epsilon(x^*)$. Also wird $\overline{U_{2^{-k_1}}(x_{k_1})}$ von der <u>einzigen</u> Umgebung $U^* \in \mathcal{U}$ überdeckt, im Widerspruch dazu, daß $U_{2^{-k_1}}(x_{k_1})$ durch keine endliche Teilüberdeckung von $\mathcal{U}$ überdeckt wird. □

Zum Schluß dieses Abschnitts betrachten wir den Spezialfall $(\mathbb{R}^n, d)$ mit $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$.

Satz 7.2.14. *Die Menge $K \subset \mathbb{R}^n$ ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.*

Beweis. "$\Rightarrow$:"
Es sei $K \subset \mathbb{R}^n$ kompakt.

i) <u>Annahme:</u> K ist nicht beschränkt.
Dann können wir mittels einer vollständigen Induktion eine Folge $(\vec{x}_k)_{k=0}^{\infty}$ mit $||\vec{x}_0|| \ge 1$ und $||\vec{x}_{k+1}|| \ge ||\vec{x}_j|| + 1$ für alle $j \le k$ konstruieren. Es gilt dann $||\vec{x}_i - \vec{x}_j|| \ge 1$ für alle i, j mit $i \ne j$. Für jede Teilfolge von $(\vec{x}_k)$ gilt dann, daß sie keine Cauchyfolge und daher nach Satz 7.2.2 nicht konvergent ist. Also ist K nicht folgenkompakt und daher nach Satz 7.2.13 nicht kompakt, ein Widerspruch.
Also ist K beschränkt.

ii) Es sei $\vec{b}$ ein Berührpunkt von K. Für alle $k \in \mathbb{N}$ existiert dann ein $\vec{x}_k \in U_{1/k}(\vec{b}) \cap K$. Also gilt $\lim\limits_{k\to\infty} \vec{x}_k = \vec{b}$, und wegen der Folgenkompaktheit ist daher $\vec{b} \in K$. Damit enthält K alle seine Berührpunkte und ist daher nach Satz 7.2.11 (ii) abgeschlossen.

"$\Leftarrow$:"
Es sei $K \subset \mathbb{R}^n$ abgeschlossen und beschränkt.
Weiter sei $(\vec{x}_k)_{k=1}^{\infty}$ eine Folge mit $\vec{x}_k \in K$. Nach Satz 7.2.7 (Bolzano- Weierstraß im $\mathbb{R}^n$) hat $(\vec{x}_k)$

mindestens einen Häufungswert und damit eine konvergente Teilfolge $(\vec{x}_{k_j})_{j=1}^{\infty}$ mit $\lim_{j\to\infty} \vec{x}_{k_j} = \vec{x}^*$. Damit ist $\vec{x}^*$ Berührpunkt von K, und daher ist nach Satz 7.2.11 (ii) auch $\vec{x}^* \in K$. Damit ist K folgenkompakt, also nach Satz 7.2.13 auch kompakt. □

7.3 Stetigkeit

Definition 7.3.1. Es seien $p, q \in \mathbb{N}$. Mit $\mathcal{F}_p^q$ bezeichnen wir die Menge der Abbildungen $f\colon X \to Y$ mit $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$.

Wir verallgemeinern zunächst die Grenzwertdefinition für den Fall $p = q = 1$ (Definition 3.1.2).

Definition 7.3.2. Es sei $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$, x_0 ein Häufungspunkt von X und $\vec{f}\colon X \to \mathbb{R}^q$. Dann heißt $\vec{a} \in \mathbb{R}^q$ Grenzwert von $\vec{f}(\vec{x})$ für $\vec{x}$ gegen $\vec{x}_0$ (Schreibweise: $\lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) = \vec{a}$), wenn für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon)$ existiert, so daß für alle $\vec{x} \in X$ mit $0 < ||\vec{x} - \vec{x}_0|| < \delta$ dann $||\vec{f}(\vec{x}) - \vec{a}|| < \epsilon$ gilt.

Bemerkung 7.3.1. Wie im Fall $p = q = 1$ läßt sich die Bedingung in Definition 7.3.2 wieder mit dem Umgebungsbegriff ausdrücken:

$$\forall \epsilon > 0\ \exists \delta = \delta(\epsilon), \text{ so daß } \quad \forall \vec{x} \in (X \backslash \{\vec{x}_0\}) \cap U_\delta(\vec{x}) \text{ gilt: } \quad \vec{f}(\vec{x}) \in U_\epsilon(\vec{a}).$$

Satz 7.3.1. *(Folgenkriterium)*
Es seien $p, q \in \mathbb{N}$, $X \subseteq \mathbb{R}^p$, $\vec{f}\colon X \to \mathbb{R}^q$ und x_0 ein Häufungspunkt von X. Dann sind folgende Aussagen äquivalent:

i) $\lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) = \vec{a}$.

ii) Für jede Folge $(z_n)_{n=1}^{\infty}$ mit $\vec{z}_n \in X \backslash \{\vec{x}_0\}$ und $\lim_{n\to\infty} \vec{z}_n = \vec{x}_0$ gilt $\lim_{n\to\infty} \vec{f}(\vec{z}_n) = \vec{a}$.

Beweis. Der Bewies verläuft fast wörtlich exakt wie der Beweis von Satz 3.1.1, wobei lediglich der Betrag $|\cdot|$ durch die Norm $||\cdot||$ zu ersetzen ist. □

Aus dem Folgenkriterium und Satz 7.2.3 folgt sofort, daß die Existenz des Grenzwerts einer Funktion zur Existenz der Grenzwerte für die Komponentenfunktionen äquivalent ist.

Satz 7.3.2. *Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$, $\vec{f}\colon X \to \mathbb{R}^q$ und x_0 ein Häufungspunkt von X. Weiter sei $\vec{f}(\vec{x}) = (f_1(\vec{x}), \ldots, f_q(\vec{x}))$ und $\vec{a} = (a_1, \ldots, a_q)$. Dann gilt genau dann $\lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) = \vec{a}$, wenn $\lim_{\vec{x}\to\vec{x}_0} f_j(\vec{x}) = a_j$ für alle $j \in \{1, \ldots, q\}$ gilt.*

Aus dem Folgenkriterium und aus Satz 7.2.4 ergibt sich

Satz 7.3.3. *Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ und x_0 ein Häufungspunkt von X. Weiter seien $\vec{f}, \vec{g}\colon X \to \mathbb{R}^q$ und $\lambda\colon X \to \mathbb{R}$. Dann gilt*

i) $\lim_{\vec{x}\to\vec{x}_0} (\vec{f}(\vec{x}) + \vec{g}(\vec{x})) = \lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) + \lim_{\vec{x}\to\vec{x}_0} \vec{g}(\vec{x})$

ii) $\lim_{\vec{x}\to\vec{x}_0} \lambda(\vec{x}) \cdot \vec{f}(\vec{x}) = \left(\lim_{\vec{x}\to\vec{x}_0} \lambda(\vec{x})\right) \cdot \left(\lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x})\right)$

iii) $\lim\limits_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) \cdot \vec{g}(\vec{x}) = \left(\lim\limits_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x})\right) \cdot \left(\lim\limits_{\vec{x}\to\vec{x}_0} \vec{g}(\vec{x})\right)$

iv) $\lim\limits_{\vec{x}\to\vec{x}_0} ||\vec{f}(\vec{x})|| = ||\lim\limits_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x})||$

Definition 7.3.3. (Stetigkeit)
Es seien $p, q \in \mathbb{N}$.

i) Es seien $X \subset \mathbb{R}^p$, $\vec{f}\colon X \to \mathbb{R}^q$ und $\vec{x}_0 \in X$. Dann heißt $\vec{f}$ in x_0 stetig, wenn für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon, x_0)$ existiert, so daß für alle $\vec{x} \in X$ mit $||\vec{x} - \vec{x}_0|| < \delta$ dann $||\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0)|| < \epsilon$ gilt.

ii) Weiter heißt $\vec{f}$ auf X stetig, wenn $\vec{f}$ in jedem $\vec{x}_0 \in X$ stetig ist.

Bemerkung 7.3.2. Die Stetigkeit läßt sich auch mit dem Umgebungsbegriff wie folgt ausdrücken:

$$\forall \epsilon > 0\ \exists \delta = \delta(\epsilon, x_0), \text{ so daß } \quad \forall \vec{x} \in X \cap U_\delta(\vec{x}) \text{ gilt: } \quad \vec{f}(\vec{x}) \in U_\epsilon(\vec{f}(\vec{x}_0)).$$

Bemerkung 7.3.3. Ist $\vec{x}_0$ ein isolierter Punkt von X, so ist $\vec{f}$ in $\vec{x}_0$ immer stetig.

Aus den Definitionen 7.3.2 und 7.3.3 ergibt sich sofort

Satz 7.3.4. *Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$, $\vec{f}\colon X \to \mathbb{R}^q$ und x_0 ein Häufungspunkt von X. Dann ist $\vec{f}$ in $\vec{x}_0$ genau dann stetig, wenn $\lim\limits_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) = \vec{f}(\vec{x}_0)$.*

Daraus ergibt sich, daß auch die Stetigkeit "komponentenweise" überprüft werden kann.

Satz 7.3.5. *Es seien $p, q \in \mathbb{N}$, $X \subseteq \mathbb{R}^p$, $\vec{f}\colon X \to \mathbb{R}^q$ und $x_0 \in X$. Zudem gilt $\vec{f}(\vec{x}) = (f_1(\vec{x}), \ldots, f_q(\vec{x}))$. Dann ist $\vec{f}$ genau dann in $\vec{x}_0$ stetig, wenn f_j in $\vec{x}_0$ stetig ist (für alle $j = 1, \ldots, q$).*

Beweis. Ist $\vec{x}_0$ ein isolierter Punkt von X, so sind nach Bemerkung 7.3.3 $\vec{f}$ und die f_j in $\vec{x}_0$ stetig. Es sei dann also $\vec{x}_0$ ein Häufungspunkt von X. Dann gilt

$$\begin{aligned}
\vec{f} \text{ stetig in } \vec{x}_0 \quad &\underset{S.7.3.4}{\Leftrightarrow} \quad \lim_{\vec{x}\to\vec{x}_0} \vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) \\
&\underset{S.7.3.2}{\Leftrightarrow} \quad \lim_{\vec{x}\to\vec{x}_0} f_j(\vec{x}) = f_j(\vec{x}_0) \quad \text{für alle } j \in \{1, \ldots, n\} \\
&\underset{S.7.3.4}{\Leftrightarrow} \quad f_j \text{ in } \vec{x}_0 \text{ stetig für alle } j \in \{1, \ldots, n\}.
\end{aligned}$$

□

Satz 7.3.6. *Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ und $x_0 \in X$. Weiter seien $\vec{f}, \vec{g}\colon X \to \mathbb{R}^q$, $\lambda\colon X \to \mathbb{R}$ und $\vec{f}, \vec{g}, \lambda$ in $\vec{x}_0$ stetig. Dann sind auch folgende Abbildungen in $\vec{x}_0$ stetig:*

i) $\vec{f} + \vec{g}$

ii) $\lambda \cdot \vec{f}$

iii) $\vec{f} \cdot \vec{g}$

iv) $||\vec{f}||$

Beweis. Ist $\vec{x}_0$ ein isolierter Punkt von X, so sind alle erwähnten Funktionen dort stetig. Ist $\vec{x}_0$ ein Häufungspunkt von X, so folgt die Behauptung aus Satz 7.3.3 und Satz 7.3.4. □

Satz 7.3.7. *(Stetigkeit der Komposition)*
Es seien $p, q, r \in \mathbb{N}$, $X \subset \mathbb{R}^p$, $Y \subset \mathbb{R}^q$, $\vec{x}_0 \in X$, $\vec{y}_0 \in Y$, $\vec{g}\colon X \to Y$, $\vec{f}\colon Y \to \mathbb{R}^r$ und $\vec{g}(\vec{x}_0) = \vec{y}_0$. Weiter sei $\vec{g}$ in $\vec{x}_0$ und $\vec{f}$ in $\vec{y}_0$ stetig. Dann ist auch $\vec{f} \circ \vec{g}$ in $\vec{x}_0$ stetig.

Beweis. Wir können wieder annehmen, daß $\vec{x}_0$ ein Häufungspunkt von X ist. Es sei $(\vec{w}_k)_{k=1}^{\infty}$ mit $\lim_{k\to\infty} \vec{w}_k = \vec{x}_0$. Nach Satz 7.3.4 ist

$$\lim_{k\to\infty} \vec{g}(\vec{w}_k) = \vec{g}(\vec{x}_0) \quad \text{und} \quad \lim_{k\to\infty} \vec{f}(\vec{g}(\vec{w}_k)) = \vec{f}(\vec{g}(\vec{x}_0)).$$

Wiederum nach Satz 7.3.4 folgt die Stetigkeit von $\vec{f} \circ \vec{g}$. □

Die Stetigkeit einer Abbildung läßt sich wiederum mit Begriffen der Topologie formulieren. Wir beginnen mit dem Spezialfall der topologischen Räume $\mathbb{R}^p$ und $\mathbb{R}^q$ sowie einerAbbildung $\vec{f}\colon \mathbb{R}^p \to \mathbb{R}^q$, die auf dem gesamten Raum $\mathbb{R}^p$ definiert ist.

Satz 7.3.8. *Es sei $\vec{f}\colon \mathbb{R}^p \to \mathbb{R}^q$. Dann ist $\vec{f}$ genau dann auf $\mathbb{R}^p$ stetig, wenn das Urbild jeder offenen Menge von $\mathbb{R}^q$ wieder offen ist, d.h. wenn $\vec{f}^{-1}(O)$ für alle offenen Mengen $O \subset \mathbb{R}^q$ eine offene Menge von $\mathbb{R}^p$ ist.*

Beweis. Es sei $O \subset \mathbb{R}^q$ eine offene Menge und $\vec{x}_0 \in \vec{f}^{-1}(O)$. Dann ist $\vec{y}_0 := \vec{f}(\vec{x}_0) \in O$. Da O offen ist, existiert ein $\epsilon > 0$, so daß $U_\epsilon(\vec{y}_0) \subset O$ ist. Wegen der Stetigkeit von $\vec{f}$ in $\vec{y}_0$ existiert ein $\delta > 0$, so daß $\vec{f}(\vec{x}) \in U_\epsilon(\vec{y}_0)$ und damit $\vec{f}(\vec{x}) \in O$ für alle $\vec{x} \in U_\delta(\vec{x}_0)$ ist. Damit ist $U_\delta(\vec{x}_0) \subset \vec{f}^{-1}(O)$. Also ist jedes $\vec{x}_0 \in \vec{f}^{-1}(O)$ innerer Punkt von $\vec{f}^{-1}(O)$. Nach Satz 7.2.9 ist $\vec{f}^{-1}(O)$ offen. □

Die in Satz 7.3.7 beschriebene Eigenschaft stetiger Abbildungen dient zur Definition der Stetigkeit von Abbildungen zwischen topologischen Räumen.

Definition 7.3.4. Es seien (X, T) und (Y, S) topologische Räume. Eine Abbildung $f\colon X \to Y$ heißt stetig, falls für alle $O \in S$ dann $f^{-1}(O) \subset T$ gilt.

Ist f nicht auf dem gesamten Raum, sondern nur auf einer Teilmenge $W \subset X$ definiert, so kann die Stetigkeit mittels der sogenannten relativen Topologie definiert werden.

Definition 7.3.5. Es sei (X, T) ein topologischer Raum und $W \subset X$ mit $W \neq \emptyset$. Eine Teilmenge $O_W \subset W$ heißt relativ offen (bzgl. W), falls es eine Menge $O \subset T$ mit $O_W = O \cap W$ gibt.

Satz 7.3.9. *Es sei (X, T) ein topologischer Raum und $W \subset X$ mit $W \neq \emptyset$. Weiter sei $T\big|_W$ die Menge der relativ offenen Teilmengen von W. Dann ist $(W, T\big|_W)$ ein topologischer Raum.*

Beweis.

i) Die Mengen $W = X \cap W$ und $\emptyset = \emptyset \cap W$ sind nach den Definitionen 7.2.11 und 7.3.5 relativ offen.

ii) Es seien $O_W^{(1)}, O_W^{(2)}$ relativ offen. Dann existieren $O^{(1)}, O^{(2)} \in T$ mit $O_W^{(1)} = O^{(1)} \cap W$ und $O_W^{(2)} = O^{(2)} \cap W$. Dann ist $O_W^{(1)} \cap O_W^{(2)} = (O^{(1)} \cap O^{(2)}) \cap W \in T\big|_W$.

iii) Es seien $O_W^{(j)}$ relativ offen für $j \in J$. Dann existiert ein $O^{(j)} \in T$ mit $O_W^{(j)} = O^{(j)} \cap W$. Dann ist

$$\bigcup_{j\in J} O_W^{(j)} = \left(\bigcup_{j\in J} O^{(j)}\right) \cap W \in T\big|_W.$$

Nach Definition 7.2.11 ist $(W, T\big|_W)$ ein topologischer Raum.

□

Definition 7.3.6. Es seien (X, T) und (Y, S) topologische Räume. Weiter sei $W \subset X$ mit $W \neq \emptyset$. Eine Funktion $f: W \to Y$ heißt stetig, falls für jede Menge O, die bzgl. $f(W)$ relativ offen ist, die Menge $f^{-1}(O)$ relativ offen bzgl. W ist.

Bemerkung 7.3.4. Für den Fall $W = X$ stimmt Definition 7.3.6 mit Definition 7.3.4 überein.

Satz 7.3.10. *Es seien $p, q \in \mathbb{N}$ und $W \subset \mathbb{R}^p$ mit $W \neq \emptyset$. Dann heißt $\vec{f}: W \to \mathbb{R}^q$ stetig auf W (im Sinne von Definition 7.3.3 (ii)), wenn es stetig im Sinne von Definition 7.3.6 ist, wobei $\mathbb{R}^p$ bzw. $\mathbb{R}^q$ mit der Topologie der metrischen Räume $(\mathbb{R}^p, d)$ bzw. $(\mathbb{R}^q, d)$ versehen sind.*

Beweis. ohne Beweis. □

Satz 7.3.11. *Es sei (X, T) ein topologischer Raum und $K \subset X$ mit $K \neq \emptyset$. Dann ist K genau dann kompakt, wenn K bzgl. der Topologie $(K, T\big|_K)$ kompakt ist.*

Beweis. "$\Rightarrow$":
Es sei K kompakt. Es sei $\tilde{\mathcal{U}}$ eine Menge von (bzgl. K) relativ offenen Teilmengen $\tilde{U} \subset K$, die K überdecken, d.h.

$$K = \bigcup_{\tilde{U} \in \tilde{\mathcal{U}}} \tilde{U}.$$

Nach Definition 7.3.5 gibt es zu jedem $\tilde{U}$ ein offenes $U(\tilde{U}) \subset X$ mit $U(\tilde{U}) \in T$, so daß $\tilde{U} = K \cap U(\tilde{U})$. Somit ist $\mathcal{U} = \{U(\tilde{U}| \ \tilde{U} \in \tilde{\mathcal{U}}\}$ eine Überdeckung von K. Nach Definition 7.2.10 (Kompaktheit) ist $\mathcal{U}$ eine endliche Teilüberdeckung $\mathcal{V} = \{\mathcal{U}(\tilde{U_j})| \ j \in \{1, \ldots, n\}\}$. Dann ist $\tilde{\mathcal{V}} = \{\tilde{U_j}| \ j \in \{1, \ldots, n\}\}$ eine endliche Teilüberdeckung von $\tilde{\mathcal{U}}$. Nach Definition 7.2.10 ist K kopmpakt (bzgl. der Topologie $(K, T\big|_K)$).
"$\Leftarrow$":
Wir verzichten auf den Beweis der Rückrichtung. □

Satz 7.3.12. *(Stetige Bilder kompakter Mengen sind kompakt)*
Es seien (X, T) und (Y, S) topologische Räume. Weiter sei $K \subset X$ kompakt, und $f: K \to Y$ sei auf K stetig. Dann ist auch $f(K)$ kompakt.

Beweis. Es sei $\mathcal{V}$ eine Überdeckung von $f(K)$. Dann sind alle $V \in \mathcal{V}$ relativ offen (bzgl. $f(K)$). Dann ist

$$K = \bigcup_{V \in \mathcal{V}} f^{-1}(V).$$

Nach Definition 7.3.6 ist $f^{-1}(V)$ relativ offen bzgl. K für alle $V \in \mathcal{V}$. Also bildet $\mathcal{U} = \{f^{-1}(V)| \ V \in \mathcal{V}\}$ eine Überdeckung (bzgl. der Topologie $(K, T\big|_K)$ von K). Damit ist K nach Satz 7.3.11 auch bzgl. der Topologie $(K, T\big|_K)$ kompakt.
Daher hat $\mathcal{U}$ eine endliche Teilüberdeckung $\mathcal{U}_{end} = \{f^{-1}(V_j)| \ j \in \{1, \ldots, m\}\}$ mit $V_j \in \mathcal{V}$. Dann ist $\mathcal{V}_{end} = \{V_j| \ j \in \{1, \ldots, m\}\}$ eine endliche Teilmenge von $\mathcal{V}$. Also ist $f(K)$ kompakt bzgl $f(X)$ und nach Satz 7.3.11 auch kompakt bzgl. der Topologie (Y, S). □

Wir kommen nun zu einer Anwendung auf den Fall $(X, T) = (\mathbb{R}^p, d)$, $(Y, S) = (\mathbb{R}^q, d)$ und $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$.

Satz 7.3.13. *Es seien $p, q \in \mathbb{N}$, $K \subset \mathbb{R}^p$ kompakt und $\vec{f}: K \to \mathbb{R}^q$ sei auf K stetig. Dann ist auch $\vec{f}(K)$ kompakt.*

Beweis. Dies folgt sofort aus Satz 7.3.12. □

Wir kommen nun zum wichtigen Spezialfall $q = 1$.

Satz 7.3.14. *(Annahme der Extremwerte)*
Es seien $p \in \mathbb{N}$, $K \subset \mathbb{R}^p$ kompakt und $f\colon K \to \mathbb{R}$ sei auf K stetig. Dann nimmt f auf K Maximum und Minimum an, d.h. es gibt $\vec{x}_{max}, \vec{x}_{min} \in K$ mit $f(\vec{x}_{min}) \le f(\vec{x}) \le f(\vec{x}_{max})$ für alle $\vec{x} \in K$.

Beweis. Nach Satz 7.3.12 ist $f(K)$ kompakt, also nach Satz 7.2.14 abgeschlossen und beschränkt. Damit besitzt die Menge $f(K)$ ein endliches Infimum $s = \inf\{f(K)\}$ und ein endliches Supremum $S = \sup\{f(K)\}$. Da $f(K)$ abgeschlossen ist, gilt $s, S \in f(K)$. □

Definition 7.3.7. (gleichmäßige Stetigkeit)
Es seien $p, q \in \mathbb{N}$, $M \subset \mathbb{R}^p$ und $\vec{f}\colon M \to \mathbb{R}^q$. Dann heißt $\vec{f}$ gleichmäßig stetig auf M, wenn es für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ gibt, so daß $||\vec{f}(\vec{x}_1) - \vec{f}(\vec{x}_2)|| < \epsilon$ für alle $\vec{x}_1, \vec{x}_2$ mit $||\vec{x}_1 - \vec{x}_2|| < \delta$ gilt.

Satz 7.3.15. *(gleichmäßige Stetigkeit auf kompakten Mengen)*
Es seien $p, q \in \mathbb{N}$, $K \subset \mathbb{R}^p$ kompakt und $\vec{f}\colon K \to \mathbb{R}^q$ stetig. Dann ist $\vec{f}$ auf K gleichmäßig stetig.

Beweis. Es sei $\epsilon > 0$ gegeben. Nach der Definition der Stetigkeit (Definition 7.3.3) gibt es zu jedem $\vec{z} \in K$ ein $\delta(\vec{z}) > 0$, so daß für alle $x \in U_{3\delta(\vec{z})}(\vec{z})$ gilt

$$||f(\vec{x}) - f(\vec{z})|| < \frac{\epsilon}{3}. \tag{1}$$

Es bildet $\mathcal{U} = \{U_{\delta(\vec{z})}(\vec{z}) |\ \vec{z} \in K\}$ eine Überdeckung von K. Wegen der Kompaktheit von K hat $\mathcal{U}$ eine endliche Teilüberdeckung, d.h. es gibt $\{\vec{z}_1, \ldots, \vec{z}_m\}$, so daß $K \subset \mathcal{V} = \{U_{\delta(\vec{z}_j)}(\vec{z}_j) |\ j \in \{1, \ldots, m\}\}$ ist. Es sei $\delta = \delta(\epsilon) := \min\{\delta(\vec{z}_j) |\ j \in \{1, \ldots, m\}\}$. Weiter seien $\vec{x}_1, \vec{x}_2 \in K$ mit $||\vec{x}_1 - \vec{x}_2|| < \delta(\epsilon)$. Dann gibt es $j_1, j_2 \in \{1, \ldots, m\}$ mit $||\vec{x}_1 - \vec{z}_{j_1}|| < \delta$ und $||\vec{x}_2 - \vec{z}_{j_2}|| < \delta$. Nach der Dreiecksungleichung ist dann

$$||\vec{z}_{j_1} - \vec{z}_{j_2}|| \le ||\vec{z}_{j_1} - \vec{x}_1|| + ||\vec{x}_1 - \vec{x}_2|| + ||\vec{x}_2 - \vec{z}_{j_2}|| < 3\delta.$$

Wegen (1) folgt

$$||f(\vec{x}_1) - f(\vec{x}_2)|| \underset{\Delta - Ugl.}{\le} ||f(\vec{x}_1) - f(\vec{z}_{j_1})|| + ||f(\vec{z}_{j_1}) - f(\vec{z}_{j_2})|| + ||f(\vec{z}_{j_2}) - f(\vec{x}_2)|| < \epsilon.$$

□

7.4 Partielle Differenzierbarkeit

Definition 7.4.1. Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $\vec{x}_0$ ein innerer Punkt von X. Dann heißt eine Funktion $\vec{f}\colon X \to \mathbb{R}^q$ in $\vec{x}_0$ partiell differenzierbar nach x_j, wenn für alle $j \in \{1, \ldots, p\}$

$$\lim_{h \to 0} \frac{\vec{f}(\vec{x}_0 + h\vec{e}_j) - \vec{f}(\vec{x}_0)}{h} =: \frac{\partial \vec{f}}{\partial x_j}(\vec{x}_0)$$

existiert. Dann heißt $\dfrac{\partial \vec{f}}{\partial x_j}$ partielle Ableitung von $\vec{f}$ nach x_j (in $\vec{x}_0$). Andere Schreibweisen sind $\dfrac{\partial}{\partial x_j}\vec{f}(\vec{x}_0)$ oder $\vec{f}_{x_j}(\vec{x}_0)$.

Bemerkung 7.4.1. Nach Satz 7.2.3 können die partiellen Ableitungen "komponentenweise" ermittelt werden. Ist $\vec{f} = (f_1, \ldots, f_q)$, so ist

$$\frac{\partial \vec{f}}{\partial x_j} = \left(\frac{\partial f_1}{\partial x_1}, \ldots, \frac{\partial f_q}{\partial x_p} \right)$$

mit $j \in \{1, \ldots, p\}$.
Die Bestimmung der partiellen Ableitungen kann also auf den Fall $q = 1$ zurückgeführt werden.
Für $q = 1$ hängt der Begriff der partiellen Ableitung eng mit dem Begriff der Ableitung aus Definition 3.7.1 zusammen:
Es sei $\vec{x}_0 = (x_{1,0}, \ldots, x_{p,0})$. Setzen wir $g_j(x_j) = f(x_{1,0}, \ldots, x_j, \ldots, x_{p,0})$, so ergibt sich aus Definition 7.4.1

$$\frac{\partial f}{\partial x_j}(x_0) = \lim_{x_j \to x_{j,0}} \frac{g_j(x_j) - g_j(x_{j,0})}{x_j - x_{j,0}} = g_j'(x_{j,0}).$$

Wir sehen schon, daß die einzige "wirkliche Variable" die Variable x_j ist. Für $i \neq j$ nimmt x_j jeweils den festen Wert $x_{i,0}$ an und wird als Konstante behandelt. Die Regeln für die Bestimmung der partiellen Ableitung ergeben sich daher aus den alten Ableitungregeln.

Beispiel 7.4.1. Es sei $f\colon \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \to f(x, y, z) = xye^{yz}$. Bestimme die partiellen Ableitungen von f.

Lösung:
Es werden jeweils die zwei anderen Variablen als Konstante behandelt. Man erhält

$$\begin{aligned}
\frac{\partial}{\partial x}\left(xye^{yz}\right) &= ye^{yz} \\
\frac{\partial}{\partial y}\left(xye^{yz}\right) &= xe^{yz}\frac{\partial}{\partial y}(y) + xy\frac{\partial}{\partial y}e^{yz} = xe^{yz} + xyze^{yz} \\
\frac{\partial}{\partial z}\left(xye^{yz}\right) &= xy^2 e^{yz}
\end{aligned}$$

unter Verwendung der Produkt- bzw. Kettenregel.

Es stellt sich nun heraus, daß der Begriff der partiellen Ableitung noch nicht geeignet ist, den Begriff der gewöhnlichen Ableitung zu verallgemeinern. Nach Satz 3.7.3 folgt aus der Differenzierbarkeit einer Funktion in einem Punkt auch deren Stetigkeit. Hingegen folgt aus der partiellen Differenzierbarkeit einer Funktion mehrerer Variablen nicht deren Stetigkeit, wie das folgende Beispiel zeigt:

Beispiel 7.4.2. Die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ sei durch

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{für} \quad (x, y) \neq (0, 0), \\ 0 & \text{für} \quad (x, y) = (0, 0) \end{cases}$$

definiert. Die Stetigkeit und die Existenz der partiellen Ableitungen für $(x, y) \neq (0, 0)$ sind klar. Existenz der partiellen Ableitungen und Stetigkeit sind also nur noch im Ursprung $(x, y) = (0, 0)$ zu untersuchen. Aus Definition 7.4.1 ergibt sich

$$\begin{aligned}
f_x(0, 0) &= \lim_{x \to 0} \frac{x \cdot 0}{x^2 + 0^2} = \lim_{x \to 0} 0 = 0 \\
f_y(0, 0) &= \lim_{y \to 0} \frac{0 \cdot y}{0^2 + y^2} = \lim_{y \to 0} 0 = 0.
\end{aligned}$$

Die partiellen Ableitungen f_x und f_y existieren also auch für $(x, y) = (0, 0)$. Wir untersuchen nun die Stetigkeit von f in $(0, 0)$ mit dem Folgenkriterium:

Für $t \in \mathbb{R}$ betrachten wir die Folge $\vec{x}_{n,t} = \left(\frac{1}{n}, \frac{t}{n}\right)$. Die Punkte der Folge nähern sich dem Ursprung $(0,0)$ also auf der Geraden $y = tx$. Es ist

$$\lim_{n\to\infty} f(\vec{x}_{n,t}) = \lim_{n\to\infty} \frac{t/n^2}{(1+t^2)/n^2} = \frac{t}{1+t^2}.$$

Der Grenzwert hängt also von t ab. Nach dem Folgenkriterium müßte er im Falle der Stetigkeit für alle t derselbe sein. Damit ist f in $(0,0)$ nicht stetig.

Eine einfache Rechnung zeigt, daß die partiellen Ableitungen der Funktion f in Beispiel 7.4.2 im Nullpunkt nicht beschränkt sind. Aus der zusätzlichen Forderung der Beschränktheit der partiellen Ableitungen folgt nun tatsächlich die Stetigkeit der Funktion f.

Satz 7.4.1. *Es seien $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $f\colon X \to \mathbb{R}$. Sämtliche partielle Ableitungen $\frac{\partial f}{\partial x_j}$ von f mit $1 \le j \le p$ mögen auf X existieren und seien dort beschränkt. Dann ist f auf X stetig.*

Beweis. Es sei

$$\left|\frac{\partial f}{\partial x_j}\right| \le M \tag{1}$$

für alle $\vec{x} \in X$ und $\vec{x}_0 \in X$. Da $\vec{x}_0$ ein innerer Punkt von X ist, gibt es ein $\delta_0 > 0$, so daß für alle $\vec{x} \in \mathbb{R}^p$ ein $\vec{x}_0 \in X$ mit

$$||\vec{x} - \vec{x}_0|| < \delta_0 \tag{2}$$

gibt. Für festes $\vec{x}^* \in U_{\delta_0}(\vec{x}_0)$ definieren wir die endliche Folge $(\vec{x}_k)$ mit $1 \le k \le p$ wie folgt: Es sei

$$\vec{x}^* = \vec{x}_0 + \sum_{m=1}^{p} h_m \vec{e}_m.$$

Dann sei

$$\vec{x}_k = \vec{x}_0 + \sum_{m=1}^{k} h_m \vec{e}_m.$$

Es ist

$$f(\vec{x}^*) = f(\vec{x}_0) + \sum_{k=1}^{p} (f(\vec{x}_k) - f(\vec{x}_{k-1})). \tag{3}$$

Wir wenden nun den Mittelwertsatz (Satz 3.9.3) auf die Funktionen

$$\varphi_k\colon [0, h_k] \to \mathbb{R} \text{ (bzw. } [h_k, 0], \text{ falls } h_k < 0), \; t \to \varphi_k(t) = f(\vec{x}_{k-1} + t\vec{e}_k)$$

an. Wir haben $\varphi_k(0) = f(\vec{x}_{k-1})$, $\varphi_k(h_k) = f(\vec{x}_k)$ und

$$f(\vec{x}_k) - f(\vec{x}_{k-1}) = \varphi_k(h_k) - \varphi_k(0) = \varphi_k'(\xi_k) h_k$$

für $\xi_k \in (0, h_k)$. Es ist

$$\varphi_k'(\xi_k) = \frac{\partial f}{\partial x_k}(\vec{x}_{k-1} + \xi_k \vec{e}_k)$$

und damit

$$|f(\vec{x}_k) - f(\vec{x}_{k-1})| \le h_k M. \tag{4}$$

Aus (3) und (4) folgt

$$|f(\vec{x}^*) - f(\vec{x}_0)| \le \left(\max_{1\le k\le p} h_k\right) pM. \tag{5}$$

Es sei $\epsilon > 0$. Wir setzen $\delta := \min\{\delta_0, \frac{\epsilon}{pM}\}$, und es sei $||\vec{x} - \vec{x}_0|| < \delta$. Wegen $\max_{1 \le k \le p} h_k < \delta$ folgt aus (5)

$$|f(\vec{x}^*) - f(\vec{x}_0)| < \epsilon.$$

□

Die gesamte Information über die partiellen Ableitungen einer Funktion kann in der Funktionalmatrix gesammelt werden.

Definition 7.4.2. i) Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $\vec{f}\colon X \to \mathbb{R}^q$ mit $\vec{f} = (f_1, \ldots, f_q)$. Die partiellen Ableitungen $\frac{\partial \vec{f}}{\partial x_j}(\vec{x}_0)$ mögen für $j \in \{1, \ldots, q\}$ existieren. Dann versteht man unter der Funktionalmatrix oder Jacobimatrix $\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)$ (von $\vec{f}$ in $\vec{x}_0$) die Matrix

$$\frac{\partial \vec{f}}{\partial \vec{x}} := \begin{pmatrix} \dfrac{\partial f_1(\vec{x}_0)}{\partial x_1} & \cdots & \dfrac{\partial f_1(\vec{x}_0)}{\partial x_p} \\ \vdots & & \vdots \\ \dfrac{\partial f_q(\vec{x}_0)}{\partial x_1} & \cdots & \dfrac{\partial f_q(\vec{x}_0)}{\partial x_p} \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(\vec{x}_0) \right)_{\substack{1 \le i \le q \\ 1 \le j \le p}}$$

In der i- ten Zeile von $\frac{\partial \vec{f}}{\partial \vec{x}}$ stehen also die partiellen Ableitungen der j- ten Komponentenfunktion, in der j- ten Spalte die partiellen Ableitungen sämtlicher Komponentenfunktionen nach der i- ten Variable. Die Matrix $\frac{\partial \vec{f}}{\partial \vec{x}}$ ist vom Typ (q, p).

ii) Im Spezialfall $q = 1$ heißt

$$\operatorname{grad} f(\vec{x}_0) := \left(\frac{\partial f(\vec{x}_0)}{\partial x_1}, \ldots, \frac{\partial f(\vec{x}_0)}{\partial x_p} \right)$$

der Gradient von f in $\vec{x}_0$.
Der Gradient ist also die einzige Zeile der einzeiligen Funktionalmatrix.

7.5 Totale Differenzierbarkeit

Wir haben in Beispiel 7.4.2 gesehen, daß aus der partiellen Differenzierbarkeit einer Funktion nicht deren Stetigkeit folgt. Die partielle Differenzierbarkeit einer Funktion mehrerer Variablen ist daher als Verallgemeinerung der Differenzierbarkeit einer Funktion einer Variablen nicht gut geeignet. Bei der Suche nach einer guten Verallgemeinerung lassen wir uns daher von Satz 3.7.1 leiten. Die Differenzierbarkeit einer Funktion f einer Variablen ist zur linearen Approximierbarkeit dieser Funktion äquivalent:
Es gibt $c \in \mathbb{R}$, so daß

$$f(x) = f(x_0) + c(x - x_0) + r(x) \qquad (*)$$

mit

$$\lim_{x \to x_0} \frac{|r(x)|}{|x - x_0|} = 0.$$

Bei mehreren Variablen tritt an die Stelle der linearen Funktion $L(x) = f(x_0) + c(x - x_0)$ die affine Abbildung $\vec{L}(\vec{x}) = \vec{f}(\vec{x}_0) + (C(\vec{x} - \vec{x}_0)^T)^T$ mit einer Matrix C.

Definition 7.5.1. (totale Differenzierbarkeit)
Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $\vec{f}\colon X \to \mathbb{R}^q$. Dann heißt $\vec{f}$ in $\vec{x}_0$ <u>(total) differenzierbar</u>, falls eine Matrix C vom Typ (q, p) und eine Funktion $\vec{r}\colon X \to \mathbb{R}^q$ mit

$$\lim_{\vec{x}\to\vec{x}_0} \frac{\vec{r}(\vec{x})}{||\vec{x}-\vec{x}_0||} = \vec{0}$$

existieren, so daß

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + (C(\vec{x}-\vec{x}_0)^T)^T + \vec{r}(\vec{x})$$

gilt.

Bemerkung 7.5.1. Der Begriff <u>total differenzierbar</u> wird verwendet, wenn der Gegensatz zur bloßen <u>partiellen Differenzierbarkeit</u> betont werden soll.

Definition 7.5.2. Es seien $p, q \in \mathbb{N}$ und C eine Matrix vom Typ (q, p). Unter der <u>Norm</u> von C (Schreibweise $||C||$) versteht man

$$||C|| := \sup\{||(C\vec{x}^T)^T||\colon \vec{x} \in \mathbb{R}^p,\ ||\vec{x}|| = 1\}.$$

Satz 7.5.1. *Es seien $p, q \in \mathbb{N}$ und C eine Matrix vom Typ (q, p). Dann gilt*

i) Es ist $||C|| < \infty$.

ii) Für $\vec{x} \in \mathbb{R}^p$ ist $||(C\vec{x}^T)^T|| \leq ||C|| \cdot ||\vec{x}||$.

iii) Die Abbildung $\vec{L}_C\colon \mathbb{R}^p \to \mathbb{R}^q$, $\vec{x} \to (C\vec{x}^T)^T$ ist auf $\mathbb{R}^p$ stetig.

Beweis. i) Es sei $C = (c_{ij})_{\substack{1\leq i\leq q\\ 1\leq j\leq p}}$ mit den <u>Zeilenvektoren</u> $\vec{c}_i = (c_{i1}, \ldots, c_{ip})$. Es sei $\vec{x} \in \mathbb{R}^p$.
Dann ist

$$C\vec{x}^T = \begin{pmatrix} \vec{c}_1 \cdot \vec{x} \\ \vdots \\ \vec{c}_q \cdot \vec{x} \end{pmatrix}.$$

Nach der Cauchy- Schwarzschen Ungleichung ist $|\vec{c}_i \cdot \vec{x}| \leq ||\vec{c}_i|| \cdot ||\vec{x}||$, also

$$||(C\vec{x}^T)^T|| \leq q \cdot \left(\max_{1\leq i\leq q} ||\vec{c}_i||\right),$$

falls $||\vec{x}|| = 1$ ist. Damit ist $||C|| < \infty$.

ii) Für $\vec{x} = \vec{0}$ ist die Behauptung klar. Es sei nun also $\vec{x} \neq \vec{0}$. Dann ist $\vec{v} = \frac{\vec{x}}{||\vec{x}||}$ ein <u>Einheitsvektor</u>, d.h. $||\vec{v}|| = 1$. Nach Definition 7.5.2 ist $||(C\vec{v}^T)^T|| \leq ||C|| \cdot ||\vec{v}||$, und damit

$$||(C\vec{x}^T)^T|| \leq ||C|| \cdot ||\vec{x}|| \cdot ||\vec{v}|| = ||C|| \cdot ||\vec{x}||.$$

iii) Es seien $\vec{x}_0, \vec{x} \in \mathbb{R}^p$. Nach (ii) ist

$$||(C\vec{x}^T)^T - (C\vec{x}_0^T)^T|| \leq ||C|| \cdot ||\vec{x} - \vec{x}_0||.$$

Damit ist

$$\lim_{\vec{x}\to\vec{x}_0} C\vec{x} = C\vec{x}_0.$$

Nach Satz 7.3.4 ist $\vec{L}_C$ in $\vec{x}_0$ stetig.

□

Satz 7.5.2. *(Aus totaler Differenzierbarkeit folgt Stetigkeit)*
Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $\vec{f}\colon X \to \mathbb{R}^q$. Weiter sei $\vec{f}$ in $\vec{x}_0$ total differenzierbar. Dann ist $\vec{f}$ in $\vec{x}_0$ stetig.

Beweis. Nach Definition 7.5.1 ist

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + (C(\vec{x} - \vec{x}_0)^T)^T + \vec{r}(\vec{x}).$$

Nach Satz 7.5.1 ist die Abbildung $\vec{x} \to (C(\vec{x} - \vec{x}_0)^T)^T$ auf $\mathbb{R}^p$ stetig. Die Abbildung $\vec{x} \to \vec{r}(\vec{x})$ ist in $\vec{x} = \vec{x}_0$ nach Satz 7.3.4 wegen $\lim\limits_{\vec{x}\to\vec{x}_0} \vec{r}(\vec{x}) = \vec{0}$ stetig. □

Satz 7.5.3. *Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und C eine Matrix vom Typ (q, p). Weiter sei $\vec{r}\colon X \to \mathbb{R}^q$ mit $\lim\limits_{\vec{x}\to\vec{x}_0} \dfrac{\vec{r}(\vec{x})}{||\vec{x} - \vec{x}_0||} = 0$ und $\vec{f}\colon X \to \mathbb{R}^q$ mit*

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + (C(\vec{x} - \vec{x}_0)^T)^T + \vec{r}(\vec{x}), \tag{$*$}$$

d.h. $\vec{f}$ ist nach Definition 7.5.1 in $\vec{x}_0$ total differenzierbar. Dann ist $\vec{f}$ auch partiell nach allen Variablen differenzierbar. Die Matrix C in $(*)$ ist eindeutig bestimmt. Es ist

$$C = \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0),$$

die Funktionalmatrix von $\vec{f}$ in $\vec{x}_0$.

Beweis. Es sei

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{q1} & \cdots & c_{qj} & \cdots & c_{qp} \end{pmatrix}$$

und $\vec{f} = (f_1, \ldots, f_q)$. Für $k \in \{1, \ldots, p\}$ bilden wir für $h \neq 0$ den Differenzenquotienten

$$\frac{\vec{f}(\vec{x}_0 + h\vec{e}_j) - \vec{f}(\vec{x}_0)}{h} = \frac{h(C\vec{e}_j^T)^T}{h} + \frac{\vec{r}(\vec{x}_0 + h\vec{e}_j)}{h} \tag{1}$$

Dabei ist $C\vec{e}_j^T$ die j- te Spalte der Matrix C, also

$$(C\vec{e}_j^T)^T = (c_{1j}, \ldots, c_{qj}). \tag{2}$$

Weiter ist

$$\lim_{h\to 0} \frac{\vec{r}(\vec{x}_0 + h\vec{e}_j)}{h} = \vec{0}. \tag{3}$$

Aus (1), (2) und (3) und Definition 7.4.1 folgt

$$\lim_{h\to 0} \frac{\vec{f}(\vec{x}_0 + h\vec{e}_j) - \vec{f}(\vec{x}_0)}{h} = (c_{1j}, \ldots, c_{qj}) = \left(\frac{\partial f_1}{\partial x_j}, \ldots, \frac{\partial f_q}{\partial x_j}\right)$$

Nach Definition 7.4.2 ist $C = \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)$. □

Satz 7.5.4. *Es seien $p, q \in \mathbb{N}$, $X \subseteq \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $\vec{f}\colon X \to \mathbb{R}^q$. Weiter sei $\vec{f}$ in X nach allen Variablen partiell differenzierbar, und die partiellen Ableitungen seien in $\vec{x}_0$ stetig. Dann ist $\vec{f}$ in $\vec{x}_0$ total differenzierbar.*

Beweis. Wir beschränken uns auf den Fall $q = 1$. Der allgemeine Fall kann darauf zurückgeführt werden, indem man die Komponentenfunktionen betrachtet.
Es sei also $f\colon X \to \mathbb{R}$ und $\vec{x}_0 \in X$. Da $\vec{x}_0$ ein innerer Punkt von X ist, gibt es $\delta_0 > 0$, so daß $U_{\delta_0}(\vec{x}_0) \subset X$ ist. Es sei $\epsilon > 0$ gegeben. Wegen der Stetigkeit von $\frac{\partial f}{\partial x_j}$ in $\vec{x}_0$ gibt es δ mit $0 < \delta < \delta_0$, so daß

$$\left|\frac{\partial f}{\partial x_j}(\vec{x}) - \frac{\partial \vec{f}}{\partial x_j}(\vec{x}_0)\right| < \epsilon \tag{1}$$

für alle $j \in \{1, \ldots, p\}$, falls $||\vec{x} - \vec{x}_0|| < \delta$ ist.
Wie im Beweis von Satz 7.4.1 definieren wir für festes $\vec{x}^* \in U_\delta(\vec{x}_0)$ die Folge $(\vec{x}_k)$: es sei

$$\vec{x}^* = \vec{x}_0 + \sum_{m=1}^{p} h_m \vec{e}_m$$

mit $|h_m| < \delta$. Dann sei

$$\vec{x}_k = \vec{x}_0 + \sum_{m=1}^{k} h_m \vec{e}_m.$$

Nach dem Mittelwertsatz ist

$$f(\vec{x}_k) - f(\vec{x}_{k-1}) = h_k \cdot \frac{\partial f}{\partial x_k}(\vec{x}_{k-1} + \xi_k \vec{e}_k)$$

mit einem ξ_k zwischen 0 und h_k. Wegen (1) folgt

$$\left|f(\vec{x}_k) - f(\vec{x}_{k-1}) - h_k \cdot \frac{\partial f}{\partial x_k}(\vec{x}_0)\right| \leq |h_k| \cdot \left|\frac{\partial f}{\partial x_k}(\vec{x}_{k-1} + \xi_k \vec{e}_k) - \frac{\partial f}{\partial x_k}(\vec{x}_0)\right| \leq \epsilon |h_k|.$$

Also gilt

$$\left|f(\vec{x}^*) - f(\vec{x}_0) + \left(\frac{\partial f}{\partial \vec{x}}(\vec{x}_0) \cdot (\vec{x}^* - \vec{x}_0)^T\right)^T\right| \leq \epsilon \sum_{k=1}^{p} |h_k|.$$

Nach Definition 7.5.1 ist f in $\vec{x}_0$ total differenzierbar. □

7.6 Differentiationsregeln

Satz 7.6.1. *(Ableitungen von Summen, Produkten und inneren Produkten)*
Es seien $p, q \in \mathbb{N}$, $\vec{f}, \vec{g} \in \mathcal{F}_p^q$ und $\lambda \in \mathcal{F}_p^1$. Falls für $\vec{x}_0 \in \mathbb{R}^p$ die partiellen Ableitungen $\frac{\partial \vec{f}}{\partial x_j}$, $\frac{\partial \vec{g}}{\partial x_j}$ bzw. $\frac{\partial \lambda}{\partial x_j}$ existieren, dann existieren sie auch für $\vec{f} + \vec{g}$, $\vec{f} \cdot \vec{g}$ und $\lambda \cdot \vec{f}$, und es gilt

$$\begin{aligned}
\frac{\partial}{\partial x_j}(\vec{f} + \vec{g}) &= \frac{\partial}{\partial x_j}\vec{f} + \frac{\partial}{\partial x_j}\vec{g} \\
\frac{\partial}{\partial x_j}(\vec{f} \cdot \vec{g}) &= \left(\frac{\partial \vec{f}}{\partial x_j}\right) \cdot \vec{g} + \vec{f} \cdot \left(\frac{\partial \vec{g}}{\partial x_j}\right) \\
\frac{\partial}{\partial x_j}(\lambda \cdot \vec{f}) &= \left(\frac{\partial \lambda}{\partial x_j}\right) \cdot \vec{f} + \lambda \cdot \frac{\partial \vec{f}}{\partial x_j}.
\end{aligned}$$

Beweis. Diese Regeln folgen unmittelbar aus Definition 7.4.1 und aus den Regeln für die Ableitung von Funktionen einer Variablen. □

Satz 7.6.2. *(Kettenregel)*
Es seien $p, q, r \in \mathbb{N}$, $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$ offen sowie $\vec{f}\colon X \to Y$ und $\vec{g}\colon Y \to \mathbb{R}^r$. Es sei $\vec{x}_0 \in X$ und $\vec{f}(\vec{x}_0) = \vec{y}_0$. Die Funktion $\vec{f}$ sei in $\vec{x}_0$ und $\vec{g}$ in $\vec{y}_0$ total differenzierbar und haben die Funktionalmatrizen $\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)$ und $\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0)$. Dann ist auch die Komposition $\vec{g} \circ \vec{f}$ in $\vec{x}_0$ total differenzierbar und habe die Funktionalmatrix

$$\frac{\partial(\vec{g} \circ \vec{f})}{\partial \vec{x}}(\vec{x}_0) = \frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0).$$

Beweis. Nach Definition 7.5.1 und Satz 7.5.3 ist für alle $\vec{x} \in X$ und für alle $\vec{y} \in Y$

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + \left(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)^T\right)^T + \vec{r}(\vec{x}) \tag{1}$$

und

$$\vec{g}(\vec{y}) = \vec{g}(\vec{y}_0) + \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot (\vec{y} - \vec{y}_0)^T\right)^T + \vec{s}(\vec{y}) \tag{2}$$

mit Abbildungen $\vec{r}$, $\vec{s}$, für die

$$\lim_{\vec{x} \to \vec{x}_0} \frac{\vec{r}(\vec{x})}{||\vec{x} - \vec{x}_0||} = \vec{0} \quad \text{und} \quad \lim_{\vec{y} \to \vec{y}_0} \frac{\vec{s}(\vec{y})}{||\vec{y} - \vec{y}_0||}$$

gilt. Wir wenden (1) und (2) mit $\vec{y}_0 = \vec{f}(\vec{x}_0)$ und $\vec{y} = \vec{f}(\vec{x})$ an und erhalten

$$\begin{aligned}
(\vec{g} \circ \vec{f})(\vec{x}) &= \vec{g}\left(\vec{f}(\vec{x}_0) + \left(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)^T\right)^T + \vec{r}(\vec{x})\right) \\
&= \vec{g}(\vec{y}_0) + \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot (\vec{y} - \vec{y}_0)^T\right)^T + \vec{s}(\vec{y}) \\
&= \vec{g}(\vec{f}(\vec{x}_0)) + \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \left(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)^T + \vec{r}(\vec{x})\right)^T\right)^T + \vec{s}(\vec{f}(\vec{x})) \\
&= \vec{g}(\vec{f}(\vec{x}_0)) + \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)^T\right)^T + \vec{u}(\vec{x})
\end{aligned}$$

mit

$$\vec{u}(\vec{x}) = \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \vec{r}(\vec{x})^T\right)^T + \vec{s}(\vec{f}(\vec{x})).$$

Nach Satz 7.5.1 ist

$$\left|\left|\left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \vec{r}(\vec{x})^T\right)^T\right|\right| \le \left|\left|\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0)\right|\right| \cdot ||\vec{r}(\vec{x})||\,.$$

Also ist

$$\lim_{\vec{x} - \vec{x}_0} \frac{1}{||\vec{x} - \vec{x}_0||} \cdot \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \vec{r}(\vec{x})^T\right)^T = \vec{0}. \tag{3}$$

Wir setzen

$$\vec{v}(\vec{y}) := \begin{cases} \frac{\vec{s}(\vec{y})}{||\vec{y}-\vec{y}_0||} & \text{für} \quad \vec{y} \neq \vec{y}_0, \\ 0 & \text{für} \quad \vec{y} = \vec{y}_0. \end{cases}$$

Dann ist $\lim\limits_{\vec{y}\to\vec{y}_0} \vec{v}(\vec{y}) = \vec{0}$ für $\vec{x} \neq \vec{x}_0$ und

$$\frac{\vec{s}(\vec{f}(\vec{x}))}{||\vec{x}-\vec{x}_0||} = \vec{v}(\vec{f}(\vec{x})) \cdot \vec{w}(\vec{x}),$$

wobei

$$\vec{w}(\vec{x}) = \frac{||\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0)||}{||\vec{x}-\vec{x}_0||}$$

auf $U_\delta(\vec{x}_0)$ für hinreichend kleine $\delta > 0$ beschränkt ist. Es folgt

$$\lim_{\vec{x}-\vec{x}_0} \frac{\vec{s}(\vec{f}(\vec{x}))}{||\vec{x}-\vec{x}_0||} = \vec{0} \tag{4}$$

Aus (1), (2), (3) und (4) folgt

$$(\vec{g} \circ \vec{f})(\vec{x}) = (\vec{g} \circ \vec{f})(\vec{x}_0) + \left(\left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0) \right) \cdot (\vec{x}-\vec{x}_0)^T \right)^T + \vec{u}(\vec{x})$$

mit

$$\lim_{\vec{x}-\vec{x}_0} \frac{\vec{u}(\vec{x})}{||\vec{x}-\vec{x}_0||} = \vec{0}.$$

Nach Definition 7.5.1 und Satz 7.5.3 folgt die totale Differenzierbarkeit von $g \circ f$ in $\vec{x}_0$ mit

$$\frac{\partial(\vec{g} \circ \vec{f})}{\partial \vec{x}}(\vec{x}_0) = \frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) \cdot \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0).$$

□

7.7 Richtungsableitung

Bei der Bildung der partiellen Ableitung

$$\frac{\partial f}{\partial x_j}(\vec{x}_0) = \lim_{h\to 0} \frac{f(\vec{x}_0 + h\vec{e}_j) - f(\vec{x}_0)}{h}$$

spielen nur die Werte von f auf der Geraden $\vec{x} = \vec{x}_0 + h\vec{e}_j$, deren Richtung durch den j- ten Einheitsvektor $\vec{e}_j$ gegeben ist, eine Rolle. Betrachtet man stattdessen eine beliebige Gerade durch $\vec{x}_0$, so kommt man auf das Konzept der Richtungsableitung.

Definition 7.7.1. Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $\vec{x}_0 \in X$. Es sei $f\colon X \to \mathbb{R}$ und $\vec{v} \in \mathbb{R}^p$ mit $||\vec{v}|| = 1$. Dann heißt f in Richtung $\vec{v}$ differenzierbar, falls der Grenzwert

$$D_{\vec{v}}(f)(\vec{x}_0) = \lim_{h\to 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h}$$

existiert. Dann heißt $D_{\vec{v}}(f)(x_0)$ die Richtungsableitung von f in Richtung $\vec{v}$ im Punkt $\vec{x}_0$.

Bemerkung 7.7.1. Wie schon in der Einleitung ausgeführt, ergibt sich für $\vec{v} = \vec{e}_j$ gerade die partielle Ableitung $\frac{\partial f}{\partial x_j}$.

Satz 7.7.1. *Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $\vec{x}_0 \in X$. Zudem sei f in $\vec{x}_0$ total differenzierbar. Dann existieren im Punkt $\vec{x}_0$ für alle $v \in \mathbb{R}^p$ mit $||\vec{v}|| = 1$ die Richtungsableitungen von f in Richtung $\vec{v}$, und es ist $D_{\vec{v}}(f)(\vec{x}_0) = \vec{v} \cdot \operatorname{grad} f$.*

Beweis. Nach der Definition der totalen Differenzierbarkeit (Definition 7.5.1) ist für alle h mit $\vec{x}_0 + h\vec{v} \in X$

$$f(\vec{x}_0 + h\vec{v}) = f(\vec{x}_0) + h\vec{v}(\operatorname{grad} f) + r(\vec{x}_0 + h\vec{v}),$$

mit $\lim\limits_{h \to 0} \dfrac{r(\vec{x}_0 + h\vec{v})}{|h|} = 0$. Es folgt

$$\lim_{h \to 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h} = \vec{v} \cdot \operatorname{grad} f$$

□

Satz 7.7.2. *Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $||\vec{v}|| = 1$. Zudem sei $f\colon X \to \mathbb{R}$ in $\vec{x}_0$ total differenzierbar. Dann ist $|D_{\vec{v}}(f)(\vec{x}_0)| \leq ||\operatorname{grad} f||$.*
Es ist genau dann $D_{\vec{v}}(f)(\vec{x}_0) = \operatorname{grad} f$, wenn $\operatorname{grad} f = \vec{0}$ oder $\vec{v} = \lambda \cdot \operatorname{grad} f$ mit $\lambda > 0$ gilt.

Beweis. Dies folgt aus Satz 7.7.1 und der Cauchy- Schwarzschen Ungleichung. □

Bemerkung 7.7.2. Stellt man die Frage, in welche Richtung die Funktion f am stärksten wächst, d.h. für welchen Einheitsvektor $\vec{v}$ die Richtungsableitung $D_{\vec{v}}(f)(x_0)$ am größten ist, so liefert Satz 7.7.2 die Antwort: in Richtung des Gradienten.
Der Gradient gibt somit die Richtung des stärksten Anstiegs der Funktion f an.

Beispiel 7.7.1. Es beschreibe $f(x, y, z)$ das Quadrat der Entfernung von $\vec{x} = (x, y, z)$ vom Ursprung $\vec{0} = (0, 0, 0)$. Es ist also

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad (x, y, z) \to f(x, y, z) = x^2 + y^2 + z^2.$$

Dann ist $\operatorname{grad} f = 2(x, y, z)$. Befindet sich ein Objekt im Punkt $\vec{x} = (x, y, z)$, so kann es sich am schnellsten vom Ursprung entfernen, indem es sich in Richtung $\vec{x}$ bewegt.

7.8 Ableitungen höherer Ordnung

Definition 7.8.1. Es seien $p, q \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $\vec{x}_0 \in X$. Weiter seien $j, k \in \{1, \ldots, p\}$.

i) Die partielle Ableitung $\frac{\partial \vec{f}}{\partial x_k}$ möge in allen $\vec{x} \in X$ existieren. Existiert die partielle Ableitung

$$\frac{\partial}{\partial x_j}\left(\frac{\partial \vec{f}}{\partial x_k}\right)$$

in $\vec{x} = \vec{x}_0$, so schreibt man dafür

$$\frac{\partial^2 \vec{f}}{\partial x_k \partial x_j}(\vec{x}_0) \quad (\text{oder } \vec{f}_{x_k x_j}).$$

Ableitungen der Form $\frac{\partial^2 \vec{f}}{\partial x_k \partial x_j}$ heißen partielle Ableitungen zweiter Ordnung.

ii) Partielle Ableitungen höherer Ordnung werden induktiv definiert:
Die Ableitung m- ter Ordnung $\frac{\partial^m \vec{f}}{\partial x_{j_1} \dots \partial x_{j_m}}$ möge in $\vec{x} \in X$ und in $\vec{x} = \vec{x}_0$ möge

$$\frac{\partial}{\partial x_{j_{m+1}}} \left(\frac{\partial^m \vec{f}}{\partial x_{j_1} \dots \partial x_{j_m}} \right)$$

existieren. Für $\frac{\partial}{\partial x_{j_{m+1}}} \left(\frac{\partial^m \vec{f}}{\partial x_{j_1} \dots \partial x_{j_m}} \right) (\vec{x}_0)$ schreibt man dann auch

$$\frac{\partial^{m+1} \vec{f}}{\partial x_{j_1} \dots \partial x_{j_m} \partial x_{j_{m+1}}} (\vec{x}_0)$$

und nennt es eine partielle Ableitung $(m+1)$- ter Ordnung.

Beispiel 7.8.1. Es sei $f\colon \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \to f(x, y, z)$. Dann gibt es sechs "gemischte" Ableitungen dritter Ordnung, die Ableitungen nach allen drei Variablen umfassen:

$$\frac{\partial^3 f}{\partial x \partial y \partial z}, \quad \frac{\partial^3 f}{\partial x \partial z \partial y}, \quad \frac{\partial^3 f}{\partial y \partial x \partial z}, \quad \frac{\partial^3 f}{\partial y \partial z \partial x}, \quad \frac{\partial^3 f}{\partial z \partial x \partial y} \quad \text{und} \quad \frac{\partial^3 f}{\partial z \partial y \partial x}.$$

Haben diese Ableitungen alle denselben Wert, d.h. ist der Wert unabhängig von der Reihenfolge der Differentiationen?
Dies würde folgen, wenn wir zeigen könnten, daß die Ableitungen zweiter Ordnung von f und deren drei Ableitungen erster Ordnung unabhängig von der Differentiationsreihenfolge sind. Es folgt dann

$$\begin{aligned} \frac{\partial^3 f}{\partial x \partial y \partial z} &= \frac{\partial^2}{\partial x \partial y}(f_z) = \frac{\partial^2}{\partial y \partial x}(f_z) = \frac{\partial}{\partial y} \left(\frac{\partial^2 f}{\partial x \partial z} \right) = \frac{\partial}{\partial y} \left(\frac{\partial^2 f}{\partial z \partial x} \right) = \frac{\partial^3 f}{\partial y \partial z \partial x} \\ &= \frac{\partial^2}{\partial y \partial z}(f_x) = \frac{\partial^2}{\partial z \partial y}(f_x) = \frac{\partial^3 f}{\partial z \partial y \partial x} = \frac{\partial}{\partial z} \left(\frac{\partial^2 f}{\partial y \partial x} \right) = \frac{\partial}{\partial z} \left(\frac{\partial^2 f}{\partial x \partial y} \right) \\ &= \frac{\partial^3 f}{\partial z \partial x \partial y} = \frac{\partial^2}{\partial z \partial x}(f_y) = \frac{\partial^2}{\partial x \partial z}(f_y) = \frac{\partial^3 f}{\partial x \partial z \partial y}. \end{aligned}$$

Das Beispiel legt jedenfalls nahe- und es ist auch beweisbar- daß sich die Fragestellung der Vertauschung der Differentiation auf den Fall einer skalaren Funktion zweier Variablen reduzieren läßt:
Es sei $f\colon X \to \mathbb{R}$ mit $X \subset \mathbb{R}^2$ zweimal partiell differenzierbar. Wann gilt

$$f_{xy} = f_{yx}? \tag{1}$$

Das folgende Beispiel zeigt, daß dies nicht immer gelten braucht.

Beispiel 7.8.2. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x, y) = \begin{cases} xy \cdot \dfrac{x^2 - y^2}{x^2 + y^2} & \text{für} \quad (x_1, x_2) \neq (0, 0), \\ 0 & \text{für} \quad (x_1, x_2) = (0, 0) \end{cases}$$

gegeben. Für $(x, y) \neq (0, 0)$ ist

$$f_x(x, y) = y \cdot \frac{x^2 - y^2}{x^2 + y^2} + xy \cdot \frac{2x(x^2 + y^2) - 2x(x^2 - y^2)}{(x^2 + y^2)^2} = y \cdot \frac{x^2 - y^2}{x^2 + y^2} + 4 \cdot \frac{xy^2}{(x^2 + y^2)^2}$$

und

$$f_y(x, y) = x \cdot \frac{x^2 - y^2}{x^2 + y^2} - 4 \cdot \frac{x^2 y}{(x^2 + y^2)^2}.$$

Daraus ergibt sich $f_x(0, 0) = f_y(0, 0) = 0$. Es ist $f_x(0, y) = -y$ und $f_y = (x, 0) = x$ und damit $f_{xy}(0, 0) = -1$ und $f_{yx}(0, 0) = 1$.

Um die Vertauschbarkeit (1) zu gewährleisten, müssen also zusätzliche Forderungen an die Funktion f gestellt werden. Es stellt sich heraus, daß die Stetigkeit der zweiten Ableitung hinreichend ist.

Satz 7.8.1. *(Schwarz)*
Es sei $X \subset \mathbb{R}^2$, X offen und nichtleer, sowie $f\colon X \to \mathbb{R}$. Die Ableitungen zweiter Ordnung f_{xy} und f_{yx} mögen in X existieren und seien stetig in $\vec{x}_0 = (x_0, y_0) \in X$. Dann ist $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$.

Beweis. Es seien $h, k \in \mathbb{R}$, so daß $(x_0 + \delta_1 h, y_0 + \delta_2 k) \in X$ für alle δ_1, δ_2 mit $0 < \delta_1, \delta_2 < 1$ gilt. Wir wenden nun auf die Funktionen

$$g(y) := f(x_0 + h, y) - f(x_0, y) \quad \text{bzw.} \quad l(x) := f(x, y_0 + k) - f(x, y_0)$$

den Mittelwertsatz an und erhalten

$$f(x_0+h, y_0+k) - f(x_0+h, y_0) - f(x_0, y_0+k) + f(x_0, y_0) = g(y_0+k) - g(y_0) = k\left(\frac{d}{dy} g(y_0 + \vartheta_1 k)\right) \quad (1)$$
$$= k(f_y(x_0 + h, y_0 + \vartheta_1 k) - f_y(x_0, y_0 + \vartheta_1 k)$$

mit $0 < \vartheta_1 < 1$ bzw.

$$f(x_0+h, y_0+k) - f(x_0+h, y_0) - f(x_0, y_0+k) + f(x_0, y_0) = l(x_0+h) - l(x_0) = h\left(\frac{d}{dx} l(x_0 + \vartheta_2 h)\right) \quad (2)$$
$$= h(f_x(x_0 + \vartheta_2 h, y_0 + k) - f_x(x_0, y_0)$$

mit $0 < \vartheta_2 < 1$. Wir wenden auf die rechten Seiten von (1) und (2) nochmals den Mittelwertsatz an und erhalten nach Division durch hk

$$\begin{aligned} \frac{1}{hk}\left(f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)\right) &= f_{yx}(x_0 + \vartheta_3 h, y_0 + \vartheta_1 k) \\ &= f_{xy}(x_0 + \vartheta_4 h, y_0 + \vartheta_2 k) \end{aligned}$$

mit $0 < \vartheta_3, \vartheta_4 < 1$. □

7.9 Taylorpolynome, Satz von Taylor

Wir kommen nun zum Satz von Taylor in p Variablen. Wie im Falle $p = 1$ ist der einfachste Spezialfall der Mittelwertsatz.

Satz 7.9.1. *(Mittelwertsatz)*
Es sei $X \subset \mathbb{R}^p$ offen, $f\colon X \to \mathbb{R}$ sei auf X differenzierbar, und es seien $\vec{x}_0$, $\vec{x}_0 + \vec{h} \in X$, so daß auch die Verbindungsstrecke von $\vec{x}_0$ und $\vec{x}_0 + \vec{h}$ in X liegt. Dann gibt es ein ϑ mit $0 < \vartheta < 1$, so daß

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \operatorname{grad}(f(\vec{x}_0 + \vartheta\vec{h})) \cdot \vec{h}$$

gilt.

Beweis. Wir setzen $\vec{g}(t) := \vec{x}_0 + t\vec{h}$ und $\varphi = f \circ g$. Nach der Kettenregel (Satz 7.6.2) ist φ auf $(0,1)$ differenzierbar, und es ist

$$\varphi'(t) = \operatorname{grad} f(\vec{g}(t)) \cdot g'(t) = \operatorname{grad} f(\vec{x}_0 + t\vec{h}) \cdot \vec{h}.$$

Nach Satz 7.3.7 (Stetigkeit der Komposition) ist φ auf $[0,1]$ stetig. Nach dem Mittelwertsatz (Satz 3.9.3) gibt es ϑ mit $0 < \vartheta < 1$, so daß

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \varphi(1) - \varphi(0) = \varphi'(\vartheta) = f'(\vec{x}_0 + \vartheta\vec{h}) \cdot \vec{h}$$

gilt. □

Bei der Formulierung des allgemeinen Satzes von Taylor in mehreren Variablen sind "symbolische Potenzen" des "Operators" $h \cdot \nabla$ von Nutzen.

Definition 7.9.1. Es sei $p \in \mathbb{N}$ und $f\colon \mathbb{R}^p \to \mathbb{R}$. Alle partiellen Ableitungen bis zur Ordnung m mögen existieren und stetig sein. Es sei $\vec{h} \in \mathbb{R}^p$. Dann definieren wir induktiv:
$k = 1$:

$$(\vec{h} \cdot \nabla) f = h_1 \frac{\partial}{\partial x_1} f + \ldots + h_p \frac{\partial}{\partial x_p} f = \vec{h} \cdot \operatorname{grad} f.$$

$k \to k+1$:

$$(\vec{h} \cdot \nabla)^{k+1} \cdot f = (\vec{h} \cdot \nabla) \cdot \Big((\vec{h} \cdot \nabla)^k (f) \Big).$$

Beispiel 7.9.1. Es sei $p = 2$, $m = 3$ und $\vec{h} \in \mathbb{R}^2$. Damit ist $(\vec{h} \cdot \nabla) \cdot f = h_1 \frac{\partial}{\partial x} f + h_2 \frac{\partial}{\partial y} f$ und

$$\begin{aligned}
(\vec{h} \cdot \nabla)^2 f &= \vec{h} \cdot \nabla \left(h_1 \frac{\partial}{\partial x} f + h_2 \frac{\partial}{\partial y} f \right) \\
&= h_1 \frac{\partial}{\partial x} \left(h_1 \frac{\partial}{\partial x} f + h_2 \frac{\partial}{\partial y} f \right) + h_2 \frac{\partial}{\partial y} \left(h_1 \frac{\partial}{\partial x} f + h_2 \frac{\partial}{\partial y} f \right) \\
&= h_1^2 \frac{\partial^2 f}{\partial x^2} + h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2 h_1 \frac{\partial^2 f}{\partial y \partial x} + h_2^2 \frac{\partial^2 f}{\partial y^2}.
\end{aligned}$$

Wegen der Stetigkeit von f_{xy} und f_{yx} ist nach Satz 7.8.1 dann $f_{xy} = f_{yx}$ und damit

$$\begin{aligned}
(\vec{h} \cdot \nabla)^2 f &= h_1^2 \frac{\partial^2 f}{\partial x^2} + 2 h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2^2 \frac{\partial^2 f}{\partial y^2} \\
(\vec{h} \cdot \nabla)^3 f &= h_1 \frac{\partial}{\partial x} \left(h_1^2 \frac{\partial^2 f}{\partial x^2} + 2 h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2^2 \frac{\partial^2 f}{\partial y^2} \right) + h_2 \frac{\partial}{\partial y} \left(h_1^2 \frac{\partial^2 f}{\partial x^2} + 2 h_1 h_2 \frac{\partial^2 f}{\partial x \partial y} + h_2^2 \frac{\partial^2 f}{\partial y^2} \right) \\
&= h_1^3 \frac{\partial^3 f}{\partial x^3} + 3 h_1^2 h_2 \frac{\partial^3 f}{\partial x^2 \partial y} + 3 h_1 h_2^2 \frac{\partial^3 f}{\partial x \partial y^2} + h_2^3 \frac{\partial^3 f}{\partial y^3}.
\end{aligned}$$

Definition 7.9.2. Es seien $p, m \in \mathbb{N}$ und $X \subset \mathbb{R}^p$ offen. Dann ist $C^m(X)$ die Menge aller Funktionen $f\colon X \to \mathbb{R}$, deren partielle Ableitungen bis zur m- ten Ordnung auf X existieren und stetig sind.

Definition 7.9.3. Es seien $p, m \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$ und $f \in C^m(X)$.
Das Taylorpolynom m- ter Ordnung von f in $\vec{x}_0$ ist

$$T(f, \vec{x}_0, m)(\vec{h}) = \sum_{j=0}^{m} \frac{1}{j!} (\vec{h} \cdot \nabla)^j f(\vec{x}_0).$$

Satz 7.9.2. *(Taylor)*
Es seien $p, m \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $f \in C^{m+1}(X)$. Liegen die Punkte $\vec{x}_0$ und $\vec{x}_0 + \vec{h}$ samt ihrer Verbindungsstrecke in X, so gibt es ein ϑ mit $0 < \vartheta < 1$, so daß

$$f(\vec{x}_0 + \vec{h}) = T(f, \vec{x}_0, m)(\vec{h}) + \frac{1}{(m+1)!} (\vec{h} \cdot \nabla)^{m+1} f(\vec{x}_0 + \vartheta \vec{h}).$$

Beweis. ohne Beweis. □

7.10 Extremwerte

In diesem Abschnitt betrachten wir nur Funktionen $f \in \mathcal{F}_p^1$. Notwendige Bedingungen für das Vorliegen eines Extremwerts ergeben sich aus der Approximation durch das Taylorpolynom erster Ordnung, hinreichende aus der Approximation durch das Taylorpolynom zweiter Ordnung.

Satz 7.10.1. *(Notwendige Bedingung für Extrema)*
Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$ offen und $\vec{x}_0 \in X$. Die Funktion f sei in $\vec{x}_0$ nach allen Variablen partiell differenzierbar. Besitzt f in $\vec{x}_0$ ein lokales Extremum, so ist $\operatorname{grad} f = \vec{0}$.

Beweis. Hat f im Punkt $\vec{x}_0$ ein lokales Extremum, so haben auch die Funktionen einer Variablen $\varphi_k(h) := f(\vec{x}_0 + h\vec{e}_k)$ mit $k \in \{1, \ldots, p\}$ lokale Extrema in $h = 0$. Nach Satz 3.9.1 folgt daraus

$$\varphi_k'(o) = \frac{\partial}{\partial x_k} f(\vec{x}_0) = 0.$$

Also ist $\operatorname{grad} f = \vec{0}$. □

Zur Formulierung von hinreichenden Bedingungen erinnern wir an den Begriff der <u>Definitheit</u> einer quadratischen Matrix aus der Linearen Algebra:

Definition 7.10.1. Es sei $p \in \mathbb{N}$ und $\mathcal{A} = (a_{jk})_{\substack{1 \le j \le p \\ 1 \le k \le p}}$ eine <u>symmetrische Matrix</u>, d.h. $a_{jk} = a_{kj}$ für alle $j, k \in \{1, \ldots, p\}$. Die Matrix $\mathcal{A}$ bzw. die zu $\mathcal{A}$ gehörende <u>quadratische Form</u> $Q\colon \vec{x} \to Q(\vec{x}) = \vec{x} \cdot (\mathcal{A}\vec{x}^T)$ heißen <u>positiv</u> (bzw. <u>negativ</u>) <u>definit</u>, falls $Q(\vec{x}) > 0$ (bzw. $Q(\vec{x}) < 0$) für alle $\vec{x} \neq \vec{0}$ ist. Es heißen $\mathcal{A}$ bzw. Q <u>positiv</u> (bzw. <u>negativ</u>) <u>semidefinit</u>, falls $Q(\vec{x}) \geq 0$ (bzw. $Q(\vec{x}) \leq 0$) für alle $\vec{x} \neq \vec{0}$ ist. Man nennt $\mathcal{A}$ bzw. Q <u>indefinit</u>, falls $\vec{x}_1, \vec{x}_2$ mit $Q(\vec{x}_1) > 0$ und $Q(\vec{x}_2) < 0$ existieren.

Definition 7.10.2. Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$, $\vec{x}_0 \in X$ und $f \in C^2(X)$. Dann versteht man unter der <u>Hesseschen Matrix</u> $H_f(\vec{x}_0)$ von f in $\vec{x}_0$

$$H_f(\vec{x}_0) = \left(\frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) \right)_{\substack{1 \le j \le p \\ 1 \le k \le p}} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}_0) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_p}(\vec{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\vec{x}_0) & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_p}(\vec{x}_0) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_p \partial x_1}(\vec{x}_0) & & \ldots & \frac{\partial^2 f}{\partial x_p^2}(\vec{x}_0) \end{pmatrix}.$$

Bemerkung 7.10.1. Nach dem Satz von Schwarz (Satz 7.8.1) ist die Matrix $H_f(\vec{x}_0)$ symmetrisch, da $\frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) = \frac{\partial^2 f}{\partial x_k \partial x_j}(\vec{x}_0)$ für alle $1 \leq j, k \leq p$ gilt.

Satz 7.10.2. *Es sei $p \in \mathbb{N}$, $X \subset \mathbb{R}^p$, $\vec{x}_0 \in X$ und $f \in C^2(X)$. Weiter sei* $\operatorname{grad} f(\vec{x}_0) = \vec{0}$.

i) Ist $H_f(\vec{x}_0)$ <u>positiv</u> (bzw. negativ) <u>definit</u>, so hat f in $\vec{x}_0$ ein strenges <u>Minimum</u> (bzw. Maximum).

ii) Ist $H_f(\vec{x}_0)$ <u>indefinit</u>, so hat f in $\vec{x}_0$ <u>kein Extremum</u>.

Beweis. Wir führen den Beweis nur für den Fall, daß $H_f(\vec{x}_0)$ positiv definit ist. Die anderen Fälle werden ähnlich behandelt.
Für $\vec{h} \in \mathbb{R}^p$ und $|\vartheta| < 1$ sei die quadratische Form $Q(\vec{y}) = Q(\vec{y}, \vec{x}_0, \vartheta, \vec{h})$ durch

$$Q(\vec{y}) := \vec{y} \cdot H_f(\vec{x}_0 + \vartheta\vec{h}) \cdot \vec{y}^T$$

definiert. Ist $H_f(\vec{x}_0)$ positiv definit, so nimmt $Q(\vec{y}, \vec{x}_0, 0, h) = \vec{y} \cdot H_f(\vec{x}_0) \cdot \vec{y}^T$ auf der kompakten Menge $S_1 = \{\vec{y}\colon ||\vec{y}|| = 1\}$ ein positives Minimum an:

$$Q(\vec{y}, \vec{x}_0, 0, \vec{h}) \geq m > 0$$

für $||\vec{y}|| = 1$.
Wegen der Stetigkeit der partiellen Ableitungen zweiter Ordnung gibt es ein $\delta > 0$, so daß für alle $\vec{h}$ mit $||\vec{h}|| < \delta$ und für alle $\vec{y}$ mit $||\vec{y}|| = 1$

$$Q(\vec{y}, \vec{x}_0, \vartheta, \vec{h}) \geq \frac{m}{2} > 0 \tag{1}$$

gilt. Nach dem Satz von Taylor (Satz 3.9.2) mit $m = 1$ haben wir wegen (1)

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + ||\vec{h}||^2 \cdot Q\left(\frac{\vec{h}}{||\vec{h}||}, \vec{x}_0, \vartheta, \vec{h}\right) \geq f(\vec{x}_0) + \frac{m}{2} \cdot ||\vec{h}||^2.$$

Damit hat f in $\vec{x}_0$ ein lokales Minimum. □

7.11 Banachscher Fixpunktsatz

Der Banachsche Fixpunktsatz hat zahlreiche Anwendungen, insbesondere in Funktionenräumen, z.B. bei Existenzsätzen in der Theorie der gewöhnlichen Differentialgleichungen. Wir werden eine andersartige Anwendung in Abschnitt 7.12 kennenlernen.

Definition 7.11.1. Es sei (X, d) ein vollständiger metrischer Raum. Eine Abbildung $f\colon X \to X$ heißt kontrahierend, wenn es ein $0 < q < 1$ gibt, so daß

$$d(f(x), f(y)) < q \cdot d(x, y)$$

für alle $x, y \in X$ gilt.

Satz 7.11.1. *(Banachscher Fixpunktsatz)*
Es sei (X, d) ein vollständiger metrischer Raum und $f\colon X \to X$ sei kontrahierend, d.h. es gibt ein $q < 1$ gibt, so daß

$$d(f(x), f(y)) < q \cdot d(x, y) \tag{1}$$

für alle $x, y \in X$ gilt. Dann hat f genau einen Fixpunkt $x^* \in X$, welcher iterativ gewonnen werden kann. Es sei $x_0 \in X$ beliebig und

$$x_{n+1} = f(x_n) \tag{2}$$

mit $n \in \mathbb{N}_0$. Dann ist $\lim\limits_{n\to\infty} x_n = x^*$, und es ist

$$d(x_n, x^*) \leq \frac{q^n}{1-q} \cdot d(x_n, x_0)$$

für alle $n \in \mathbb{N}_0$.

Beweis. Aus (1) und (2) folgt durch vollständige Induktion nach n

$$d(x_{n+1}, x_n) < q^n \cdot d(x_1, x_0). \tag{3}$$

Nach der Dreiecksungleichung (Definition 7.2.2 (iii)) folgt für $k \in \mathbb{N}$

$$d(x_{n+k}, x_n) \leq \sum_{j=0}^{k-1} d(x_{n+j+1}, x_{n+j}) \leq \sum_{j=0}^{\infty} q^{n+j} = \frac{q^n}{1-q}. \tag{4}$$

Also ist $\lim\limits_{n\to\infty} d(x_{n+k}, x_n) = 0$ für alle $k \in \mathbb{N}$. Damit ist (x_n) eine Cauchyfolge. Nach Definition 7.2.6 (vollständiger metrischer Raum) ist (x_n) konvergent, d.h. es existiert ein $x^* \in X$ mit $\lim\limits_{n\to\infty} x_n = x^*$.
Nach der Dreiecksungleichung gilt für alle $n \in \mathbb{N}$

$$d(x^*, f(x^*)) \leq d(x^*, x_n) + d(x_n, x_{n+1}) + d(x_{n+1}, f(x^*)). \tag{5}$$

Nach (1) ist $d(f(x^*), x_{n+1}) \leq q \cdot d(x^*, x_n)$, also ist nach (3)

$$d(x^*, f(x^*)) \leq D_n := (1+q) \cdot d(x_n, x^*) + q^n \cdot d(x_1, x_0).$$

Wegen $\lim\limits_{n\to\infty} d(x_n, x^*) = 0$ und $\lim\limits_{n\to\infty} q^n = 0$ folgt $\lim\limits_{n\to\infty} D_n = 0$.
Also ist $d(x^*, f(x^*)) = 0$ und nach Definition 7.2.2 (ii) ist $f(x^*) = x^*$.
Weiter ist x^* vom "Startwert" x_0 unabhängig. Es sei $y_0 \in X$ und $y_{n+1} = f(y_n)$ für alle $n \in \mathbb{N}_0$ sowie $y^* = \lim\limits_{n\to\infty} y_n$. Nach der Dreiecksungleichung gilt für alle $n \in \mathbb{N}$

$$d(x^*, y^*) \leq d(x^*, x_n) + d(x_n, y_n) + d(y_n, y^*).$$

Aus (1) folgt $d(x_n, y_n) \leq q^n \cdot d(x_0, y_0)$, also wiederum $d(x^*, y^*) = 0$. Damit ist $x^* = y^*$. □

7.12 Inverse Funktionen im $\mathbb{R}^p$, implizite Funktionen

Für die Definition der Funktionen einer Variablen gibt es die Möglichkeit der impliziten und der expliziten Definition.

Beispiel 7.12.1. Die Punkte (x, y) des Einheitskreises im $\mathbb{R}^2$ sind durch

$$x^2 + y^2 = 1 \tag{1}$$

bestimmt. Dies ist die implizite Definition zweier Funktionen $f_1, f_2 \colon [-1, 1] \to \mathbb{R}$, die durch

$$y = f_1(x) = \sqrt{1-x^2} \quad \text{und} \quad y = f_2(x) = -\sqrt{1-x^2}$$

explizit definiert sind. Die Gleichung (1) wurde "nach y aufgelöst". Falls die Stetigkeit der Lösung verlangt wird, ist diese Auflösung in hinreichend kleinen Umgebungen der Punkte (x_0, y_0) mit $x_0 \neq 0$ eindeutig. Ist $y_0 > 0$, so ist für genügend kleines $\epsilon > 0$ für $(x, y) \in U_\epsilon(x_0, y_0)$ die Aussage (1) zu $y = f_1(x)$ äquivalent (bei $y_0 < 0$ zu $y = f_2(x)$).

Die Gleichung (1) ist ein Spezialfall von $F(x, y) = 0$. Wir betrachten nun die Verallgemeinerung

$$\vec{F}(\vec{x}, \vec{y}) = \vec{0} \tag{2}$$

für den Fall, daß $\vec{x} \in X \subset \mathbb{R}^p$, $\vec{y} \in Y \subset \mathbb{R}^q$ und $\vec{F} \colon X \times Y \to \mathbb{R}^q$. Der Definitionsbereich der Funktion $\vec{F}$ ist somit $X \times Y = \{(\vec{x}, \vec{y}) |\ \vec{x} \in X, \vec{y} \in Y\}$. Der Einfachheit halber schreiben wir nun für $(\vec{x}, \vec{y}) = ((x_1, \ldots, x_p), (y_1, \ldots, y_q))$ kurz $(\vec{x}, \vec{y}) = (x_1, \ldots, x_p, y_1, \ldots, y_q)$. Damit wird $(\vec{x}, \vec{y})$ zu einem Element des $\mathbb{R}^{p+q}$ und $X \times Y$ zu einer offenen Teilmenge des $\mathbb{R}^{p+q}$.
Im folgenden wird stets vorausgesetzt, daß sämtliche Komponentenfunktionen von $\vec{F}$ aus $C^1(X \times Y)$

sind, also stetige partielle Ableitungen nach allen Variablen besitzen.
Ist $\vec{F} = (F_1, \ldots, F_q)$, so gilt für die Funktionalmatrix von F

$$\frac{\partial \vec{F}}{\partial(\vec{x}, \vec{y})}(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_1}{\partial x_p}(\vec{x}_0, \vec{y}_0) & \frac{\partial F_1}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_1}{\partial y_q}(\vec{x}_0, \vec{y}_0) \\ \vdots & & & & & \vdots \\ \frac{\partial F_q}{\partial x_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_q}{\partial x_p}(\vec{x}_0, \vec{y}_0) & \frac{\partial F_q}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_q}{\partial y_q}(\vec{x}_0, \vec{y}_0) \end{pmatrix}.$$

Diese Matrix zerfällt in zwei Teilmatrizen $\frac{\partial \vec{F}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0)$ und $\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)$. Es ist

$$\frac{\partial \vec{F}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_1}{\partial x_p}(\vec{x}_0, \vec{y}_0) \\ \vdots & & \vdots \\ \frac{\partial F_q}{\partial x_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_q}{\partial x_p}(\vec{x}_0, \vec{y}_0) \end{pmatrix}$$

und

$$\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_1}{\partial y_q}(\vec{x}_0, \vec{y}_0) \\ \vdots & & \vdots \\ \frac{\partial F_q}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \ldots & \frac{\partial F_q}{\partial y_q}(\vec{x}_0, \vec{y}_0) \end{pmatrix},$$

welche mit q Zeilen und q Spalten quadratisch ist.

Bevor wir den Hauptsatz über implizite Funktionen formulieren, behandeln wir zum besseren Verständnis einen einfachen Spezialfall:

Es sei $\vec{F}$ eine affine Abbildung

$$\vec{F}(\vec{x}, \vec{y}) = D((\vec{x}, \vec{y}) - (\vec{x}_0, \vec{y}_0)). \tag{3}$$

Dann ist nach Satz 7.5.3

$$\frac{\partial \vec{F}}{\partial(\vec{x}, \vec{y})} = D.$$

Die Funktionalmatrix ist also konstant, d.h. unabhängig von $\vec{x}, \vec{y}$.
Wir suchen nun eine Auflösung $\vec{f}$ der Gleichung (2) nach $\vec{y}$, also eine Abbildung $\vec{f}\colon \mathbb{R}^p \to \mathbb{R}^q$ mit

$$\vec{F}(\vec{x}, \vec{y}) = \vec{0} \Leftrightarrow \vec{y} = \vec{f}(\vec{x}).$$

Bei festem $\vec{x}$ ist (2) ein lineares Gleichungssystem für die Variable $\vec{y}$. Da $\left(\frac{\partial \vec{F}}{\partial \vec{y}}\right)^{-1} = D^{-1}$ existiert, ist es für beliebiges $\frac{\partial \vec{F}}{\partial \vec{x}}$ eindeutig lösbar und besitzt

$$\vec{y} = \vec{f}(\vec{x}) = \vec{y}_0 - D^{-1}\left(\frac{\partial \vec{F}}{\partial \vec{x}}\right)(\vec{x} - \vec{x}_0)^T \tag{4}$$

als Lösung. Diese Lösung kann aus einer beliebigen "Näherungslösung" durch folgenden Operator Φ erhalten werden.

Definition 7.12.1. Es seien $\vec{F}, X, Y$ wie oben definiert. Weiter sei $D := \frac{\partial \vec{F}}{\partial(\vec{x}, \vec{y})}(\vec{x}_0, \vec{y}_0)$ und $\vec{g}\colon X \to Y$. Dann versteht man unter der Funktion $\Phi(\vec{g})$

$$\Phi(\vec{g})(\vec{x}) := \vec{g}(\vec{x}) - D^{-1}\vec{F}(\vec{x}, \vec{g}(\vec{x})).$$

Satz 7.12.1. *Es seien $\vec{F}, X, Y$ wie oben definiert und $\vec{f}\colon X \to Y$. Dann verschwindet $\vec{F}(\vec{x}, \vec{f}(\vec{x}))$ genau dann identisch, wenn $\vec{f}$ eine Lösung der Fixpunktgleichung $\vec{g} = \Phi(\vec{g})$ ist.*

Beweis. Durch Nachrechnen. □

Satz 7.12.2. *Es sei $\vec{F}\colon \mathbb{R}^{p+q} \to \mathbb{R}^q$ eine affine Abbildung mit $\vec{F}(\vec{x}_0, \vec{y}_0) = \vec{0}$. Weiter sei $\vec{g}\colon \mathbb{R}^p \to \mathbb{R}^q$. Dann gilt für $\vec{f} = \Phi(\vec{g})$*

$$\vec{F}(\vec{x}, \vec{y}) = \vec{0} \Leftrightarrow \vec{y} = \vec{f}(\vec{x}).$$

Beweis. Es ist

$$D^{-1}\frac{\partial \vec{F}}{\partial \vec{y}} = I_q,$$

die Einheitsmatrix. Nach (4) hat $\vec{F}(\vec{x}, \vec{y}) = \vec{0}$ die eindeutige Lösung

$$\vec{f}(\vec{x}) = \vec{y}_0 - D^{-1}\left(\frac{\partial \vec{F}}{\partial \vec{x}}\right)(\vec{x} - \vec{x}_0)^T.$$

Es sei $\vec{g}(\vec{x}) = \vec{f}(\vec{x}) + \vec{r}(\vec{x})$. Dann ist wegen $\vec{F}(\vec{x}, \vec{f}(\vec{x})) = \vec{0}$

$$\Phi(\vec{g})(\vec{x}) = \vec{f}(\vec{x}) + \vec{r}(\vec{x}) - D^{-1}(\vec{F}(\vec{x}, \vec{f}(\vec{x})) - D^{-1}\frac{\partial \vec{F}}{\partial \vec{y}}\vec{r}(\vec{x}) = \vec{f}(\vec{x}).$$

□

Bemerkung 7.12.1. Bei einer affinen Abbildung $\vec{F}$ führt dann also der Operator Φ eine beliebige "Näherungslösung" $\vec{y} = \vec{g}(\vec{x})$ von $\vec{F}(\vec{x}, \vec{y}) = \vec{0}$ in eine exakte Lösung $\Phi(\vec{g}) = \vec{f}$ über. Nach Satz 7.12.1 ist $\vec{g} = \vec{f}$ eine Lösung der Fixpunktgleichung $\vec{g} = \Phi(\vec{g})$.
Im allgemeinen Fall ist $\vec{F}$ nicht mehr affin, besitzt aber nach Definition 7.5.1 (totale Differenzierbarkeit) und Satz 7.5.4 eine affine Approximation

$$\vec{F}(\vec{x}, \vec{y}) = \frac{\partial \vec{F}}{\partial(\vec{x}, \vec{y})}\left((\vec{x}, \vec{y}) - (\vec{x}_0, \vec{y}_0)\right)^T + \vec{r}(\vec{x}, \vec{y})$$

mit

$$\lim_{(\vec{x},\vec{y}) \to (\vec{x}_0, \vec{y}_0)} \frac{\vec{r}(\vec{x}, \vec{y})}{||(\vec{x} - \vec{x}_0, \vec{y} - \vec{y}_0)||} = \vec{0}.$$

Dies legt die Anwendung des Banachschen Fixpunktsatzes (Satz 7.11.1) nahe:
Ist $\vec{g}_0$ ein geeigneter Startwert, so wird die Folge $(\vec{g}_n)$, die durch $\vec{g}_{n+1} = \Phi(\vec{g}_n)$ definiert ist, gegen die Lösung $\vec{g} = \vec{f}$ der Fixpunktgleichung $\vec{g} = \Phi(\vec{g})$ konvergieren.

Satz 7.12.3. *(Hauptsatz über implizite Funktionen)*
Die Mengen $X \subset \mathbb{R}^p$, $Y \subset \mathbb{R}^q$ und die Funktion $\vec{F}\colon X \times Y \to \mathbb{R}^q$ mögen die beschriebenen Voraussetzungen erfüllen. Ferner seien $\vec{x}_0 \in X$ und $\vec{y}_0 \in Y$ mit $\vec{F}(\vec{x}_0, \vec{y}_0) = \vec{0}$, und die Funktionalmatrix $D = \frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)$ sei invertierbar mit inverser Matrix D^{-1}. Dann gibt es $\delta > 0$, $\epsilon > 0$ und genau eine stetige Funktion $\vec{f}\colon U_\delta(\vec{x}_0) \to U_\epsilon(\vec{y}_0)$, so daß für alle $\vec{x} \in U_\delta(\vec{x}_0)$ und für alle $\vec{y} \in U_\epsilon(\vec{y}_0)$

$$\vec{F}(\vec{x}, \vec{y}) = \vec{0} \Leftrightarrow \vec{y} = \vec{f}(\vec{x})$$

gilt.
Es sei $\vec{g}_0(\vec{x}) = \vec{y}_0$ und $\vec{g}_{n+1} = \Phi(\vec{g}_n)$. Dann ist $\vec{g}_n(\vec{x}) \underset{glm.}{\to} \vec{f}(\vec{x})$ auf $U_\delta(\vec{x}_0)$.

Beweis. Für $\vec{g}\colon X \to Y$ ist $\Phi(\vec{g})(\vec{x}) = \varphi(\vec{x}, \vec{g}(\vec{x}))$ mit

$$\varphi(\vec{x}, \vec{y}) := \vec{y} - D^{-1}\vec{F}(\vec{x}, \vec{y}). \tag{1}$$

Wir benötigen eine Aussage über $\varphi(\vec{x}, \vec{y}_2) - \varphi(\vec{x}, \vec{y}_1)$, wobei $\vec{x}$, $\vec{y}_1$ und $\vec{y}_2$ folgende Forderungen erfüllen sollen:
Es sei

$$\eta := \frac{1}{4q}||D^{-1}||^{-1}. \tag{2}$$

Es sei $F = (F_1, \ldots, F_q)$, und $\delta > 0$ und $\epsilon > 0$ seien so gewählt, daß für alle $j \in \{1, \ldots, q\}$, für alle $\vec{x} \in U_\delta(\vec{x}_0)$ und für alle $\vec{y} \in \overline{U_\epsilon(\vec{y}_0)}$

$$||\operatorname{grad} F_j(\vec{x}, \vec{y}) - \operatorname{grad} F_j(\vec{x}_0, \vec{y}_0)|| \leq \eta \tag{3}$$

gelte. Wegen der Stetigkeit der partiellen Ableitungen von F_j existieren δ und ϵ.
Wir setzen nun $U = U_\delta(\vec{x}_0)$ sowie $V = U_\epsilon(\vec{y}_0)$, und es sei

$$\vec{x} \in U \quad \text{und} \quad \vec{y}_1, \vec{y}_2 \in V. \tag{4}$$

Es sei

$$\vec{y}_2 = \vec{y}_1 + \vec{k}. \tag{5}$$

Wir finden nun eine lineare Approximation für $\vec{F}(\vec{x}^*, \vec{y}_2) - \vec{F}(\vec{x}^*, \vec{y}_1)$ für ein festes $\vec{x}^* \in U$. Wir beginnen mit den Komponentenfunktionen F_j. Da $\vec{x}^*$ fest ist, ist F_j eine Funktion von $\vec{y}$ allein. Der Gradient $\operatorname{grad} F_j$ soll nur die partiellen Ableitungen nach $x_1, \ldots x_q$ beinhalten. Nach dem Mittelwertsatz (7.9.1) ist

$$F_j(\vec{x}^*, \vec{y}_2) - F_j(\vec{x}^*, \vec{y}_1) = \operatorname{grad} F_j(\vec{x}^*, \vec{y}_1 + \vartheta_j \vec{h}) \cdot (\vec{y}_2 - \vec{y}_1).$$

Damit ergibt sich aus (2) und (3)

$$F_j(\vec{x}^*, \vec{y}_2) - F_j(\vec{x}^*, \vec{y}_1) - \operatorname{grad} F_j(\vec{x}_0, \vec{y}_0) \cdot (\vec{y}_2 - \vec{y}_1) = (\operatorname{grad} F_j(\vec{x}^*, \vec{y}_1 + \vartheta_j \vec{h}) - \operatorname{grad} F_j(\vec{x}_0, \vec{y}_0)) \cdot (\vec{y}_2 - \vec{y}_1).$$

Aus (3), (4) und (5) sowie der Cauchy- Schwarzschen Ungleichung folgt

$$||F_j(\vec{x}^*, \vec{y}_2) - F_j(\vec{x}^*, \vec{y}_1) - \operatorname{grad} F_j(\vec{x}_0, \vec{y}_0) \cdot (\vec{y}_2 - \vec{y}_1)|| \leq \eta||\vec{y}_2 - \vec{y}_1||$$

für alle $j \in \{1, \ldots, q\}$.
Wir fassen dies für alle j zusammen und erhalten für alle $\vec{x} \in U$ und für alle $\vec{y}_1, \vec{y}_2 \in \overline{V}$

$$||\vec{F}(\vec{x}, \vec{y}_2) - \vec{F}(\vec{x}, \vec{y}_1) - \left(\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot (\vec{y}_2 - \vec{y}_1)^T\right)^T \leq q\eta||\vec{y}_2 - \vec{y}_1||. \tag{6}$$

Es ist nach (1)

$$\begin{aligned}\varphi(\vec{x}, \vec{y}_2) - \varphi(\vec{x}, \vec{y}_1) &= (\vec{y}_2 - \vec{y}_1) - \left(D^{-1}\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot (\vec{y}_2 - \vec{y}_1)^T\right)^T \\ &\quad + \left(D^{-1}(\vec{F}(\vec{x}, \vec{y}_2)^T - \vec{F}(\vec{x}, \vec{y}_1)^T - \frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot (\vec{y}_2 - \vec{y}_1)^T\right)^T.\end{aligned}$$

Wegen

$$D^{-1}\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) = I_q,$$

die Einheitsmatrix, verschwindet der erste Summand, und wir erhalten aus (6)

$$||\varphi(\vec{x},\vec{y}_2)-\varphi(\vec{x},\vec{y}_1)|| \le ||D^{-1}||\cdot q\eta \cdot ||\vec{y}_2-\vec{y}_1|| = \frac{1}{4}||\vec{y}_2-\vec{y}_1||. \tag{7}$$

Wir kommen nun zur Definition des metrischen Raumes (M,d). Es sei M die Menge aller stetigen Abbildungen $\vec{g}\colon U \to \overline{V}$ mit $\vec{g}(\vec{x}_0)=\vec{y}_0$. Für $\vec{g}_1,\vec{g}_2 \in M$ sei

$$d(\vec{g}_1,\vec{g}_2) = \sup\{||\vec{g}_2(\vec{x})-\vec{g}_1(\vec{x})||\colon \vec{x}\in U\}.$$

Die Menge M ist wegen $\vec{g}_0 \in M$ nichtleer. Die Cauchyfolgen von (M,d) erfüllen das Cauchykriterium für gleichmäßige Konvergenz und konvergieren daher gegen stetige Grenzfunktionen $h\colon U \to \overline{V}$. Damit ist M ein vollständiger metrischer Raum.
Es sei $\vec{g}\in M$, $\vec{g}(\vec{x}_0)=\vec{y}_0$ und $\vec{g}(\vec{x})\in\overline{V}$. Wir wenden (7) mit $\vec{y}_1=\vec{y}_0$ und $\vec{y}_2=\vec{g}(\vec{x})$ an und erhalten

$$\sup\{||\varphi(\vec{x},\vec{g}(\vec{x}))-\vec{y}_0||\colon \vec{x}\in U\} \le \frac{1}{4}\epsilon.$$

Alos ist $\Phi(\vec{g})\in M$ und insbesondere

$$d(\Phi(\vec{g}_0),\vec{g}_0) \le \frac{1}{4}\epsilon. \tag{8}$$

Aus (7) folgt, daß Φ mit $q=\frac{1}{4}$ kontrahierend ist.
Damit sind alle Voraussetzungen für die Anwendung des Banachschen Fixpunktsatzes (Satz 7.11.1) erfüllt: Die Folge (g_n) konvergiert gegen einen Fixpunkt $\vec{f}$.
Es ist $\vec{f}\in M$, d.h. $\vec{f}(\vec{x})\in\overline{V}$ für alle $\vec{x}\in U$, $\vec{f}$ ist also stetig. Es ist

$$d(\vec{f},\vec{g}_0) \le d(\Phi(\vec{g}_0,\vec{g}_0) + d(\Phi(\vec{g}_0),\vec{f}). \tag{9}$$

Nach Satz 7.12.1 ist

$$d(\Phi(\vec{g}_0),\vec{f}) \le \frac{1/4}{1-1/4} d(\Phi(\vec{g}_0,\vec{g}_0) \le \frac{1}{12}\epsilon.$$

Aus (8) und (9) folgt

$$d(\vec{f},\vec{g}_0) \le \frac{\epsilon}{3}.$$

Damit ist $\vec{f}(\vec{x})\in V$ für alle $x\in U$. Da $\vec{f}$ eindeutig bestimmt ist, ist der Beweis von Satz 7.12.3 beendet. □

Satz 7.12.4. *(Differenzierbarkeit der implizit definierten Funktion)*
Die Voraussetzungen und Bezeichnungen seien wie in Satz 7.12.3. Dann ist $\vec{f}$ in $\vec{x}_0$ total differenzierbar, und es ist

$$\frac{\partial\vec{f}}{\partial\vec{x}}(\vec{x}_0) = -\left(\frac{\partial\vec{F}}{\partial\vec{y}}(\vec{x}_0,\vec{y}_0)\right)^{-1}\cdot\frac{\partial\vec{F}}{\partial\vec{x}}(\vec{x}_0,\vec{y}_0).$$

Beweis. Es seien $||\vec{h}||<\delta$ und $||\vec{k}||<\epsilon$.
Nach der Definition der totalen Differenzierbarkeit (Definition 7.5.1) ist

$$\vec{F}(\vec{x}_0+\vec{h},\vec{y}_0+\vec{k}) = \frac{\partial\vec{F}}{\partial\vec{x}}(\vec{x}_0,\vec{y}_0)\cdot\vec{h} + \frac{\partial\vec{F}}{\partial\vec{y}}(\vec{x}_0,\vec{y}_0)\cdot\vec{k} + \vec{r}(\vec{h},\vec{k}) \tag{1}$$

mit

$$\lim_{(\vec{h},\vec{k})\to\vec{0}} \frac{\vec{r}(\vec{h},\vec{k})}{||(\vec{h},\vec{k})||} = \vec{0}.$$

Wir setzen

$$\vec{k}(\vec{h}) = \vec{f}(\vec{x}_0 + \vec{h}) - \vec{f}(\vec{x}_0). \tag{2}$$

Wegen der Stetigkeit von $\vec{f}$ ist

$$\lim_{\vec{h}\to\vec{0}} ||\vec{k}(\vec{h})|| = \vec{0}. \tag{3}$$

Dann folgt aus (1) und nach den Aussagen und Bezeichnungen von Satz 7.12.3

$$\vec{0} = \vec{F}(\vec{x}_0 + \vec{h}, \vec{f}(\vec{x}_0 + \vec{h})) = \vec{F}(\vec{x}_0 + \vec{h}, \vec{y}_0 + \vec{k}(\vec{h})) = \frac{\partial \vec{F}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) \cdot \vec{h} + \frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot \vec{k}(\vec{h}) + \vec{r}(h, \vec{k}(\vec{h})). \tag{4}$$

Wegen (3) ist

$$\lim_{\vec{h}\to\vec{0}} \frac{\vec{r}(\vec{h}, \vec{k}(\vec{h}))}{||(\vec{h}, \vec{k}(\vec{h}))||} = \vec{0}. \tag{5}$$

Es folgt aus (4)

$$\vec{k}(\vec{h}) = \vec{f}(\vec{x}_0 + \vec{h}) - \vec{f}(\vec{x}_0) = -\left(\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right)^{-1} \frac{\partial \vec{F}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) \cdot \vec{h} = -\left(\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right)^{-1} \vec{r}(\vec{h}, \vec{k}(\vec{h})). \tag{6}$$

Die Behauptung ist bewiesen, wenn statt (5) die schärfere Aussage

$$\lim_{\vec{h}\to\vec{0}} \frac{\vec{r}(\vec{h}, \vec{k}(\vec{h}))}{||\vec{h}||} = \vec{0} \tag{7}$$

gezeigt werden kann.
Es ist

$$||(\vec{h}, \vec{k}(\vec{h}))|| \leq ||\vec{h}|| + ||\vec{k}(\vec{h})||,$$

also wegen (5)

$$\lim_{\vec{h}\to\vec{0}} \frac{\vec{r}(\vec{h}, \vec{k}(\vec{h}))}{||\vec{h}|| + ||\vec{k}(\vec{h})||} = \vec{0}.$$

Damit gibt es δ_0 mit $0 < \delta_0 < \delta$, so daß für $||\vec{h}|| < \delta_0$

$$||\vec{r}(\vec{h}, \vec{k}(\vec{h}))|| \leq \frac{1}{2} ||\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)||^{-1} \left(||\vec{h}|| + ||\vec{k}(\vec{h})||\right) \tag{8}$$

gilt. Aus (6) und (8) folgt

$$\frac{1}{2}||\vec{k}(\vec{h})|| \leq ||\vec{k}(\vec{h})|| - \frac{1}{2}||\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot \vec{k}|| \cdot ||\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)||^{-1} \leq C||\vec{h}|| \tag{9}$$

mit

$$C = || \left(\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right)^{-1} \frac{\partial \vec{F}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0)|| + \frac{1}{2}||\frac{\partial \vec{F}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)||^{-1}.$$

Daraus folgt (7) und die Behauptung.

□

Wir kommen nun zu einem Spezialfall von Satz 7.12.3:
Wir untersuchen die Umkehrbarkeit der Abbildung $\vec{f}\colon X \to \mathbb{R}^p$, $x \to \vec{y} = \vec{f}(\vec{x})$. Dies läuft auf die Auflösung der Gleichung $\vec{F}(\vec{x}, \vec{y}) = \vec{y} - \vec{f}(\vec{x}) = \vec{0}$ nach $\vec{x}$ hinaus.

Satz 7.12.5. *(Inverse Funktionen im $\mathbb{R}^p$)*

i) Es sei $X \subset \mathbb{R}^p$ offen, $\vec{x}_0 \in X$, $\vec{f}\colon X \to \mathbb{R}^p$ und $\vec{f} \in C^1(X)$. Es sei $\vec{y}_0 = \vec{f}(\vec{x}_0)$ und $\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)$ invertierbar. Dann gibt es $\delta > 0$, $\epsilon > 0$ und genau eine stetige Funktion $\vec{g}\colon U_\epsilon(\vec{y}_0) \to \mathbb{R}^p$, $\vec{y} \to \vec{g}(\vec{y})$, so daß für alle $\vec{x} \in U_\delta(\vec{x}_0)$ und für alle $\vec{y} \in U_\epsilon(\vec{y}_0)$

$$\vec{y} = \vec{f}(\vec{x}) \Leftrightarrow \vec{x} = \vec{g}(\vec{y})$$

gilt.
Weiter ist $\vec{g}$ in $\vec{y}_0$ total differenzierbar, und es ist

$$\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{y}_0) = \left(\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)\right)^{-1}.$$

ii) Das Bild $\vec{f}(\vec{x})$ ist offen.

Beweis. i) Wir wenden Satz 7.12.3 und Satz 7.12.4 mit $\vec{F}(\vec{y},\vec{x}) = \vec{y} = \vec{f}(\vec{x})$ an. Es ist

$$\frac{\partial \vec{F}(\vec{y},\vec{x})}{\partial(\vec{y},\vec{x})} = \left(\frac{\partial \vec{F}(\vec{y},\vec{x})}{\partial \vec{y}}(\vec{y}_0,\vec{x}_0), \frac{\partial \vec{F}(\vec{y},\vec{x})}{\partial \vec{x}}(\vec{y}_0,\vec{x}_0)\right)$$

mit

$$\frac{\partial \vec{F}(\vec{y},\vec{x})}{\partial \vec{y}}(\vec{y}_0,\vec{x}_0) = I_p \quad \text{und} \quad \frac{\partial \vec{F}(\vec{y},\vec{x})}{\partial \vec{x}}(\vec{y}_0,\vec{x}_0) = -\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{y}_0,\vec{x}_0).$$

ii) Es sei $\vec{y}_0 \in \vec{f}(X)$, d.h. es existiert ein $\vec{x}_0 \in X$ mit $\vec{f}(\vec{x}_0) = \vec{x}_0$.
Nach (i) ist dann auch $U_\epsilon(\vec{y}_0) \in \vec{f}(X)$, d.h. $\vec{x}_0$ ist innerer Punkt von $\vec{f}(X)$.

□

Bemerkung 7.12.2. In Satz 7.12.5 werden die Bedingungen für die lokale Invertierbarkeit einer Abbildung $\vec{f}$ gegeben. Es gibt Umgebungen $U_\delta(\vec{x}_0)$ bzw. $U_\epsilon(\vec{y}_0)$, die $\vec{f}$ bijektiv aufeinander abbildet. Aus der lokalen Invertierbarkeit folgt nicht die globale, wie das folgende Beispiel zeigt.

Beispiel 7.12.2. Es sei $\vec{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \to (e^x \cos y, e^x \sin y)$. Es ist

$$\frac{\partial \vec{f}}{\partial \vec{x}} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Wegen

$$\det\left(\frac{\partial \vec{f}}{\partial \vec{x}}\right) = e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} \neq 0$$

ist $\left(\frac{\partial \vec{f}}{\partial \vec{x}}\right)$ für alle $(x,y) \in \mathbb{R}^2$ invertierbar. Nach Satz 7.12.5 ist $\vec{f}$ lokal invertierbar. Allerdings ist $\vec{f}$ nicht global invertierbar, da etwa $\vec{f}(x, y+2\pi) = \vec{f}(x,y)$ gilt.

7.13 Kurven

Definition 7.13.1. i) Es sei $p \in \mathbb{N}$, $a < b \in \mathbb{R}$ und $I = [a, b]$. Eine stetige Abbildung $\vec{f}\colon I \to \mathbb{R}^p$ heißt Kurve. Dabei heißt $\vec{f}(a)$ Anfangspunkt und $\vec{f}(b)$ Endpunkt von $\vec{f}$.
Die Menge $\{\vec{f}(t)\colon t \in [a, b]\}$ heißt der Träger von $\vec{f}$.

ii) Es sei $\vec{f} = (f_1, \ldots, f_p)$. Dann heißt $\vec{f}$ auf $[a, b]$ stetig differenzierbar, falls alle f_j auf $[a, b]$ stetig differenzierbar sind.

Definition 7.13.2. Es sei $\vec{f}\colon [a, b] \to \mathbb{R}^p$ eine Kurve und $\mathcal{Z} = (t_0, t_1, \ldots, t_n)$ mit $a = t_0 < t_1 < \ldots < t_n = b$ eine Zerlegung von $[a, b]$. Es sei $l(\mathcal{Z}) := \sum_{k=0}^{n-1} ||\vec{f}(t_{k+1}) - \vec{f}(t_k)||$. Ist $l := \sup\{l(\mathcal{Z})\colon \mathcal{Z}\} < \infty$, so heißt $\vec{f}$ rektifizierbar, und l heißt die Länge (Bogenlänge) der Kurve $\vec{f}$.

Satz 7.13.1. *Es sei $p \in \mathbb{N}$ und $a < b \in \mathbb{R}$. Weiter sei $\vec{f}\colon [a, b] \to \mathbb{R}^p$ stetig differenzierbar. Dann ist $\vec{f}$ rektifizierbar mit der Länge $l = \int_a^b \left\|\frac{d\vec{f}}{dt}(t)\right\| dt.$*

Beweis. ohne Beweis. □

Beispiel 7.13.1. Es sei $\vec{f}\colon [0, 2\pi] \to \mathbb{R}^2$, $\varphi \to (\cos\varphi, \sin\varphi)$. Der Träger von $\vec{f}$ ist der Einheitskreis. Die Kurve ist geschlossen, da Anfangs- und Endpunkt übereinstimmen: $\vec{f}(0) = \vec{f}(2\pi) = (1, 0)$. Man nennt φ dann auch Winkel zwischen $\vec{\varphi}_1 = (1, 0)$ und $\vec{v} = (\cos\varphi, \sin\varphi)$.
Es sei $(x_0, y_0) = (\cos\varphi_0, \sin\varphi_0)$, und wir betrachten die Teilkurve $\vec{f}(\cdot, \varphi_0)\colon [0, \varphi_0] \to (\cos\varphi, \sin\varphi)$. Nach Satz 7.13.1 ist $\vec{f}(\cdot, \varphi_0)$ rektifizierbar mit der Länge

$$l(\varphi_0) = \int_0^{\varphi_0} \left(\left(\frac{d}{dt}\cos t\right)^2 + \left(\frac{d}{dt}\sin t\right)^2\right)^{1/2} dt = \int_0^{\varphi_0} 1\, dt = \varphi_0.$$

Hiermit ergibt sich der Zusammenhang mit der seit dem Altertum üblichen Interpretation des Winkels φ zwischen zwei Einheitsvektoren als der Bogenlänge der von ihnen begrenzten Teilkurven des Einheitskreises und der Winkelfunktionen als Quotienten von Seiten des rechtwinkligen Dreiecks mit Winkel φ.

Kapitel 8

Integralrechnung im $\mathbb{R}^p$

8.1 Riemannsche Summen und Riemannsches Integral

Definition 8.1.1. Es sei $p \in \mathbb{N}$.

i) Unter einem (kompakten) Intervall des R^p (auch p- dimensionales Intervall) versteht man eine Menge I der Form
$$I = \{(x_1, \ldots, x_p) |\; a_j \leq x_j \leq b_j,\; 1 \leq j \leq p\}$$
mit $a_j \leq b_j \in \mathbb{R}$ für $1 \leq j \leq p$.

ii) Unter dem Maß (Inhalt, Volumen) $|I|$ des Intervalls I verstehen wir
$$|I| := \prod_{j=1}^{p} (b_j - a_j).$$

Bemerkung 8.1.1. Ein kompkates p- dimensionales Intervall I ist das kartesische Produkt von p (eindimensionalen) Intervallen
$$I = [a_1, b_1] \times \ldots \times [a_p, b_p]. \qquad (*)$$
Für $p = 1$ ist das Maß von I gerade seine Länge. Für $p = 2$ ist das Maß auch als Fläche bekannt, und es ergibt sich die schon in Abschnitt 6.1 erwähnte Formel für die Fläche eines Rechtecks. Allgemein ist das Maß des Intervalls $(*)$ gleich dem Produkt der Längen der Intervalle $[a_j, b_j]$.

Definition 8.1.2. Es sei $p \in \mathbb{N}$ und I ein p- dimensionales Intervall. Unter einer Zerlegung $\mathcal{Z}$ von I versteht man eine Menge von Teilintervallen $I_\nu \subset I$ mit $\nu \in \{1, \ldots, n\}$ mit $I = \bigcup_{\nu=1}^{n} I_\nu$, so daß die offenen K von I_ν disjunkt sind.

Definition 8.1.3. Es sei $p \in \mathbb{N}$ und I ein p- dimensionales Intervall. Es sei $f\colon I \to \mathbb{R}$ beschränkt, und $\mathcal{Z} = \{I_\nu\colon 1 \leq \nu \leq n\}$ sei eine Zerlegung von I. Weiter seien $M_\nu := \sup\{f(\vec{x}) |\; \vec{x} \in I_\nu\}$ und $m_\nu := \inf\{f(\vec{x}) |\; \vec{x} \in I_\nu\}$. Unter der Riemannschen Obersumme $\overline{S}(f, \mathcal{Z})$ (bzw. Untersumme $\underline{S}(f, \mathcal{Z})$) versteht man
$$\overline{S}(f, \mathcal{Z}) := \sum_{\nu=1}^{n} M_\nu |I_\nu| \quad \text{bzw.} \quad \underline{S}(f, \mathcal{Z}) := \sum_{\nu=1}^{n} m_\nu |I_\nu|.$$

Definition 8.1.4. Es sei $p \in \mathbb{N}$, I ein p- dimensionales Intervall und $f\colon I \to \mathbb{R}$. Unter dem Riemannschen Oberintegral $\overline{\int_I} f(\vec{x})\,d\vec{x}$ (bzw. Unterintegral $\underline{\int_I} f(\vec{x})\,d\vec{x}$) versteht man

$$
\begin{aligned}
\overline{\int_I} f(\vec{x})\,d\vec{x} &= \inf\{\overline{S}(f,\mathcal{Z}) \mid \mathcal{Z} \quad \text{Zerlegung von } I\} \quad \text{bzw.} \\
\underline{\int_I} f(\vec{x})\,d\vec{x} &= \sup\{\underline{S}(f,\mathcal{Z}) \mid \mathcal{Z} \quad \text{Zerlegung von } I\}.
\end{aligned}
$$

Satz 8.1.1. *Die Voraussetzungen seien wie in Definition 8.1.4. Dann ist*

$$\underline{\int_I} f(\vec{x})\,d\vec{x} \le \overline{\int_I} f(\vec{x})\,d\vec{x}.$$

Beweis. ohne Beweis. □

Definition 8.1.5. Die Voraussetzungen seien wie in Definition 8.1.4.
Es heißt f integrierbar über I, falls Ober- und Unterintegral übereinstimmen. Der gemeinsame Wert heißt das Integral von f über I (Schreibweise: $\int_I f(\vec{x})\,\vec{x}\,d\vec{x}$).

Riemannsche Integrale können auch über anderen Mengen als Intervallen definiert werden.

Lemma 8.1.1. *Es sei $p \in \mathbb{N}$, $\mathcal{M} \subset \mathbb{R}^p$ beschränkt und nichtleer. Es sei I^* der Durchschnitt aller Intervalle, die $\mathcal{M}$ enthalten. Dann ist auch I^* ein Intervall.*

Beweis. ohne Beweis. □

Definition 8.1.6. Es sei $p \in \mathbb{N}$, $\mathcal{M} \subset \mathbb{R}^p$ beschränkt und nichtleer. Weiter sei I^* der Durchschnitt aller Intervalle, die $\mathcal{M}$ enthalten. Es sei $f\colon \mathcal{M} \to \mathbb{R}$ beschränkt. Wir setzen

$$\tilde{f}(\vec{x}) = \begin{cases} f(\vec{x}) & \text{für} \quad \vec{x} \in I^*, \\ 0 & \text{für} \quad \vec{x} \in I^* \setminus \mathcal{M} \end{cases}$$

und

$$\overline{\int_{\mathcal{M}}} f(\vec{x})\,d\vec{x} := \overline{\int_{I^*}} f(\vec{x})\,d\vec{x} \quad \text{und} \quad \underline{\int_{\mathcal{M}}} f(\vec{x})\,d\vec{x} := \underline{\int_{I^*}} f(\vec{x})\,d\vec{x}.$$

Man nennt f integrierbar über $\mathcal{M}$, falls

$$\overline{\int_{\mathcal{M}}} f(\vec{x})\,d\vec{x} = \underline{\int_{\mathcal{M}}} f(\vec{x})\,d\vec{x}$$

gilt, und der gemeinsame Wert von Ober- und Unterintegral heißt dann das Integral von f über $\mathcal{M}$ (Schreibweise: $\int_{\mathcal{M}} f(\vec{x})\,d\vec{x}$).

Definition 8.1.7. Es sei $p \in \mathbb{N}$, $\mathcal{M} \subset \mathbb{R}^p$ beschränkt und nichtleer. Es heißt $\mathcal{M}$ meßbar, falls das Integral $\int_{\mathcal{M}} 1\,d\vec{x}$ existiert. Weiter heißt $|\mathcal{M}| = \int_{\mathcal{M}} 1\,d\vec{x}$ das Maß (Inhalt oder Volumen) von $\mathcal{M}$.
Für $p = 1$ spricht man auch von der Länge, für $p = 2$ vom Flächeninhalt und für $p = 3$ vom Rauminhalt von $\mathcal{M}$.

Satz 8.1.2. *Es sei $p \in \mathbb{N}$ und $\mathcal{M} \subset \mathbb{R}^p$ beschränkt. Es seien $f, g\colon \mathcal{M} \to \mathbb{R}$ über $\mathcal{M}$ integrierbare Funktionen und $\alpha, \beta \in \mathbb{R}$. Dann sind auch*

a) $\alpha f + \beta g$,

b) $f \cdot g$,

c) $\max\{f, g\}$,

d) $|f|$,

e) falls $\inf\{g(\xi)|\ \xi \in \mathcal{M}\} > 0$ *ist, auch* fg^{-1}

über $\mathcal{M}$ *integrierbar.*
Es gelten folgende Gleichungen und Ungleichungen

i) (Linearität)

$$\int_{\mathcal{M}} (\alpha f(\vec{x}) + \beta g(\vec{x}))\, d\vec{x} = \alpha \int_{\mathcal{M}} f(\vec{x})\, d\vec{x} + \beta \int_{\mathcal{M}} g(\vec{x})\, d\vec{x}$$

ii) (Erhaltung von Ungleichungen)
Ist $f(\vec{x}) \leq g(\vec{x})$ *für alle* $\vec{x} \in \mathcal{M}$, *so ist auch*

$$\int_{\mathcal{M}} f(\vec{x})\, d\vec{x} \leq \int_{\mathcal{M}} g(\vec{x})\, d\vec{x}$$

iii) (Mittelwertsatz)
Ist $\mathcal{M}$ *meßbar und* $m \leq f(\vec{x}) \leq M$ *für alle* $\vec{x} \in \mathcal{M}$, *so ist*

$$m|\mathcal{M}| \leq \int_{\mathcal{M}} f(\vec{x})\, d\vec{x} \leq M|\mathcal{M}|.$$

iv) (Dreiecksungleichung)

$$\left|\int_{\mathcal{M}} f(\vec{x})\, d\vec{x}\right| \leq \int_{\mathcal{M}} |f(\vec{x})|\, d\vec{x}$$

v) (Cauchy- Schwarz)

$$\left(\int_{\mathcal{M}} f(\vec{x})g(\vec{x})\, d\vec{x}\right)^2 \leq \left(\int_{\mathcal{M}} f(\vec{x})^2\, d\vec{x}\right) \cdot \left(\int_{\mathcal{M}} g(\vec{x})^2\, d\vec{x}\right)$$

Beweis. ohne Beweis. □

Satz 8.1.3. *(Integrierbarkeit stetiger Funktionen)*
Es sei $p \in \mathbb{N}$ *und* $\mathcal{M} \subset \mathbb{R}^p$ *meßbar. Weiter sei* $f\colon \overline{\mathcal{M}} \to \mathbb{R}$ *auf* $\overline{\mathcal{M}}$ *stetig. Dann ist* f *über* $\overline{\mathcal{M}}$ *integrierbar.*

Beweis. ohne Beweis. □

8.2 Mehrfache Integrale

Wir kommen nun zur Frage, wie Integrale im $\mathbb{R}^p$ berechnet werden. Der Einfachheit halber beschränken wir uns auf den Fall $p = 2$, in dem schon die wesentlichen Ideen sichtbar werden.

Definition 8.2.1. Es sei $\mathcal{M} \subset \mathbb{R}^2$ beschränkt und nichtleer.
Für $y \in \mathbb{R}$ setzen wir $\mathcal{M}(y) := \{x \in \mathbb{R} |\ (x,y) \in \mathcal{M}\}$. Es sei $Y(\mathcal{M}) := \{y \in \mathbb{R} \colon \mathcal{M}(y) \neq \emptyset\}$. Falls für jedes $y \in Y(\mathcal{M})$ das Integral $F(y) := \int_{\mathcal{M}(y)} f(x,y)\, dx$ existiert und falls $\int_{Y(\mathcal{M})} F(y)\, dy$ existiert, so heißt

$$\int_{\mathcal{M}}\int f(x,y)\, dx\, dy := \int_{Y(\mathcal{M})}\int_{\mathcal{M}(y)} f(x,y)\, dx\, dy$$

das Doppelintegral von f über $\mathcal{M}$.

Beispiel 8.2.1. Es sei $\mathcal{M} = \{(x,y) \in \mathbb{R}^2 |\ 0 \leq y \leq 1,\ -y \leq x \leq y^2\}$ und $f(x,y) = xy$. Dann ist $Y(\mathcal{M}) = [0,1]$, $\mathcal{M}(y) = \{x |\ -y \leq x \leq y^2\}$ und

$$\int_{\mathcal{M}}\int f(x,y)\, dx\, dy = \int_0^1 \int_{-y}^{y^2} xy\, dx\, dy = \int_0^1 y \cdot \left(\frac{y^4}{2} - \frac{y^2}{2}\right) dy = -\frac{1}{24}.$$

Satz 8.2.1. *Es sei $\mathcal{M} \subset \mathbb{R}^2$ beschränkt, $f \colon \mathcal{M} \to \mathbb{R}$ sei über $\mathcal{M}$ integrierbar, und das Doppelintegral $\int_{\mathcal{M}}\int f(\vec{x})\, dx\, dy$ existiere. Dann ist*

$$\int_{\mathcal{M}}\int f(\vec{x})\, d\vec{x} = \int_{\mathcal{M}}\int f(x,y)\, dx\, dy.$$

Beweis. ohne Beweis. □

8.3 Substitutionsregel

Definition 8.3.1. (Funktionaldeterminante)
Es sei $p \in \mathbb{N}$ und $X \subset \mathbb{R}^p$ offen. Weiter sei $\vec{g} \colon X \to \mathbb{R}^p$ in $\vec{x}_0 \in X$ total differenzierbar.
Dann heißt die Determinante der Funktionalmatrix $\frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0)$ die Funktionaldeterminante von $\vec{g}$ in $\vec{x}_0$ (Schreibweise: $\det \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0)$).

Satz 8.3.1. *(Substitutionsregel)*
Es sei $p \in \mathbb{N}$ und $\mathcal{M}$ eine meßbare Teilmenge des $\mathbb{R}^p$. Es sei $\overline{\mathcal{M}} \subset X$, X offen, $\vec{g} \colon X \to \mathbb{R}^p$ sei auf X stetig differenzierbar, und es sei $\vec{g}\big|_{\mathcal{M}}$ bijektiv. Es sei $\det \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0)) \neq 0$ für $\vec{x} \in \mathcal{M}$, und $f \colon \overline{\vec{g}(\mathcal{M})} \to \mathbb{R}$ sei stetig auf $\overline{\vec{g}(\mathcal{M})}$. Dann ist auch $\vec{g}(\mathcal{M})$ meßbar, und es ist

$$\int_{\vec{g}(\mathcal{M})} f(\vec{u})\, d\vec{u} = \int_{\mathcal{M}} f(\vec{g}(\vec{x})) \cdot \left|\det \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0)\right| d\vec{x}.$$

Beweis. ohne Beweis. □

Einer der wichtigsten Spezialfälle ist die Substitutionsformel für Polarkoordinaten.

Satz 8.3.2. *Es sei $(x,y) \in \mathbb{R}^2 \backslash \{(0,0)\}$. Dann gibt es genau ein Paar (r,φ) mit $r > 0$ und $\varphi \in [0, 2\pi]$, so daß*

$$x = r\cos\varphi \quad \textit{und} \quad y = r\sin\varphi$$

gilt.

Beweis. ohne Beweis. □

Definition 8.3.2. Es heißen (r, φ) die Polarkoordianten von $\vec{x} = (x, y)$.

Satz 8.3.3. *Es sei $\mathcal{M}$ meßbar, $\overline{\mathcal{M}} \subset P \subset \overline{(0, \infty) \times (0, 2\pi)}$ und P offen. Es sei $\vec{g}\colon P \to \mathbb{R}^2$ und $\vec{g}(r, \varphi) = (r \cos \varphi, r \sin \varphi)$ sowie $\vec{f}\colon \overline{\vec{g}(\mathcal{M})} \to \mathbb{R}$ auf $\overline{\vec{g}(\mathcal{M})}$ stetig. Dann ist*

$$\int_{\vec{g}(\mathcal{M})} f(\vec{x})\, d\vec{x} = \int_{\mathcal{M}} f(r \cos \varphi, r \sin \varphi) r\, d(r, \varphi).$$

Beweis. Satz 8.3.3 ergibt sich aus Satz 8.3.1 wegen

$$\det \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r(\cos^2 \varphi + \sin^2 \varphi) = r.$$

□

Lemma 8.3.1. *Für $R > 0$ sei*

$$K(R) := \{(x, y)|\ x^2 + y^2 \leq R\} \quad \textit{und} \quad Q(R) := \{(x, y)|\ -R \leq x \leq R,\ -R \leq y \leq R\}.$$

Weiter sei $f\colon K(R) \to \mathbb{R}$ auf $K(R)$ stetig. Dann sind $K(R)$ und $Q(R)$ meßbar, und es gilt

$$\int_{K(R)} f(\vec{x})\, d\vec{x} = \int_0^{2\pi} \int_0^R f(r \cos \varphi, r \sin \varphi) r\, dr\, d\varphi.$$

Lemma 8.3.2. *Es sei $p \in \mathbb{N}$ und $\mathcal{M}_1, \mathcal{M}_2 \subset \mathbb{R}^p$, welche meßbar seien mit $\mathcal{M}_1 \subset \mathcal{M}_2$. Dann ist $|\mathcal{M}_1| \leq |\mathcal{M}_2|$, und $\mathcal{M}_2 \backslash \mathcal{M}_1$ ist meßbar.*

Bemerkung 8.3.1. Lemma 8.3.1 und 8.3.2 können aus der bisher behandelten Theorie leicht abgeleitet werden.

Beispiel 8.3.1. Man bestimme

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx.$$

Lösung:
Es ist

$$\int_0^{\infty} e^{-x^2}\, dx = \int_{-\infty}^{0} e^{-x^2}\, dx$$

und

$$\int_1^A e^{-x^2}\, dx \leq \int_1^A x \cdot e^{-x^2}\, dx \leq \int_1^{\infty} x \cdot e^{-x^2}\, dx = (2e)^{-1}.$$

Die Konvergenz des Integrals folgt also aus dem Majorantenkriterium.
Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \to f(x, y) = e^{-(x^2+y^2)}$. Wir betrachten für $R > 0$ die beiden Integrale

$$\int_{K(R)} f(\vec{x})\, d\vec{x} \quad \text{und} \quad \int_{Q(R)} f(\vec{x})\, d\vec{x}.$$

Es ist

$$f(r \cos \varphi, r \sin \varphi) = \exp(-r^2(\cos^2 \varphi + \sin^2 \varphi)) = e^{-r^2}.$$

Nach Lemma 8.3.1 ist

$$\begin{aligned} \int_{K(R)} f(\vec{x})\, d\vec{x} &= \int_0^{2\pi} \int_0^R f(r \cos \varphi, r \sin \varphi) \cdot r\, dr\, d\varphi = \int_0^{2\pi} \int_0^R r \cdot e^{-r^2}\, dr\, d\varphi \\ &= \frac{1}{2} \int_0^{2\pi} \left(1 - e^{-R^2}\right) d\varphi = \pi \left(1 - e^{-R^2}\right). \end{aligned}$$

Es ist $K(R) \subset Q(R) \subset K(\sqrt{2}R)$. Daher ist nach Satz 8.1.2 und Lemma 8.3.2

$$\pi\left(1-e^{-R^2}\right) \leq \int_{Q(R)} e^{-(x^2+y^2)}\,d\vec{x} \leq \pi\left(1-e^{-2R^2}\right) \tag{$*$}$$

Nach Satz 8.2.1 ist

$$\int_{Q(R)} e^{-(x^2+y^2)}\,d\vec{x} = \int_{-R}^{R}\int_{-R}^{R} e^{-(x^2+y^2)}\,dx\,dy = \left(\int_{-R}^{R} e^{-x^2}\,dx\right)^2.$$

Nach $(*)$ ist

$$\pi^{1/2}\left(1-e^{-R^2}\right)^{1/2} \leq \int_{-R}^{R} e^{-x^2}\,dx \leq \pi^{1/2}\left(1-e^{-2R^2}\right)^{1/2}.$$

Also ist

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$