

**Skriptum zum Vorlesungszyklus
Analysis I/II
WS/SS 2004-2005**

Prof. Dr. Helmut Maier

Inhaltsverzeichnis

Einleitung	5
1. Mengen und Zahlen	6
1.1. Mengen und ihre Verknüpfungen	6
1.2. Die Axiome der reellen Zahlen	8
1.3. Folgerungen aus den Körperaxiomen	11
1.4. Folgerungen aus den Ordnungsaxiomen	14
1.5. Die natürlichen, ganzen und rationalen Zahlen	16
1.6. Folgerungen aus dem Supremumsprinzip	18
1.7. Rekursive Definitionen und induktive Beweise	19
2. Funktionsbegriff und Zahlenfolgen	23
2.1. Der Funktionsbegriff	23
2.2. Grenzwerte einer Zahlenfolge	26
2.3. Grenzwerte von Summen, Produkten, Quotienten	28
2.4. Wichtige Konvergenzkriterien	29
2.5. Häufungswerte, Limes Superior, Limes Inferior	30
2.6. Uneigentliche Grenzwerte	32
3. Unendliche Reihen	33
3.1. Grundbegriffe	33
3.2. Konvergenzkriterien	34
3.3. Die Exponentialfunktion	39
4. Stetigkeit	41
4.1. Grundbegriffe	41
4.2. Offene, abgeschlossene und kompakte Mengen	42
4.3. Grundeigenschaften stetiger Funktionen	44
4.4. Der Umkehrsatz für streng monotone Funktionen	45
4.5. Grenzwerte von Funktionen und einseitige Grenzwerte	46
4.6. Gleichmäßige Konvergenz und Stetigkeit	50
4.7. Stetigkeit der Exponentialfunktion, natürliche Logarithmusfunktion	51
5. Differenzierbarkeit	53
5.1. Definition der Ableitung	53
5.2. Differentiationsregeln	56
5.3. Differentiation elementarer Funktionen	57
5.4. Mittelwertsätze	58

5.5.	Grenzfunktionen und Potenzreihen	60
5.6.	Die Formel von Taylor, Taylorreihen	65
5.7.	Trigonometrische Funktionen und ihre Umkehrfunktionen	68
5.8.	Komplexe Folgen und Potenzreihen, komplexe Exponentialfunktion	71
5.9.	Regeln von l'Hôpital	74
5.10.	Konvexität, Kurvendiskussion	75
6.	Integration	79
6.1.	Das Riemannsches Integral	79
6.2.	Mittelwertsätze und Hauptsatz der Differential- und Integralrechnung	86
6.3.	Integration elementarer Funktionen	88
6.4.	Uneigentliche Integrale	91
6.5.	Vertauschung von Integration und Grenzwertbildung	96
	Einleitung zum zweiten Teil der Vorlesung	98
7.	Topologische Begriffe und Stetigkeit	98
7.1.	Der n -dimensionale euklidische Raum	98
7.2.	Konvergenz von Punktfolgen des \mathbb{R}^n	99
7.3.	Punktfolgen	101
7.4.	Stetigkeit von Funktionen mehrerer Variablen	102
8.	Differenzierbarkeit mehrwertiger Funktionen	105
8.1.	Partielle Differenzierbarkeit	105
8.2.	Totale Differenzierbarkeit	105
8.3.	Ableitungsregeln	109
8.4.	Partielle Ableitungen höherer Ordnung	110
8.5.	Kurven und Gebiete	112
8.6.	Mittelwertsatz und Taylorscher Satz	113
8.7.	Extrema bei Funktionen mehrerer Variablen	117
8.8.	Implizite Funktionen	119
8.9.	Extrema mit Nebenbedingungen	122
9.	Integration über mehreren Veränderlichen	123
9.1.	Bereichsintegrale	123
9.2.	Mehrfache Integrale	126
9.3.	Die Substitutionsformel	130
9.4.	Kurven- und Oberflächenintegrale	134
	Index	139

Einleitung

Zweck der Anfängervorlesungen ist neben dem Kennenlernen von Fakten (die zum Teil schon aus der Schule bekannt sind) die Gewöhnung an die Strenge in der Mathematik.

Eine mathematische Theorie umfasst zwei Arten von Aussagen:

- (1) Axiome: Dies sind (meist wenige) Grundaussagen, aus denen die Lehrsätze durch logische Schlüsse (Beweise genannt) abgeleitet werden.
- (2) Lehrsätze: Sie bestehen im Allgemeinen aus Voraussetzung, Behauptung und Beweis.

Eine Aussage, die zwar richtig erscheint, jedoch nicht bewiesen werden kann, kann nicht als Lehrsatz gelten, sondern hat nur den Status einer Vermutung. Ein berühmtes Beispiel ist die Goldbachsche Vermutung.

Beispiel eines Beweises:

Definitionen: $1 + 1 =: 2$, $2 + 1 =: 3$, $3 + 1 =: 4$, usw.

Axiome:

Es seien a, b, c natürliche Zahlen, dann gelte

- (i) das Assoziativgesetz der Addition: $(a + b) + c = a + (b + c)$,
- (ii) das Distributivgesetz (Verbindung von Addition und Multiplikation): $a \cdot (b + c) = a \cdot b + a \cdot c$,
- (iii) die Neutralität der Eins: $1 \cdot a = a$,

Behauptung: $2 \cdot 2 = 4$

Beweis:

$$\begin{aligned} (1 + 1) \cdot (1 + 1) &\stackrel{(ii)}{=} (1 + 1) \cdot 1 + (1 + 1) \cdot 1 \stackrel{(ii)}{=} (1 \cdot 1 + 1 \cdot 1) + (1 \cdot 1 + 1 \cdot 1) \\ &\stackrel{(iii)}{=} (1 + 1) + (1 + 1) \stackrel{\text{Def}}{=} 2 + (1 + 1) \stackrel{(i)}{=} (2 + 1) + 1 \stackrel{\text{Def}}{=} 3 + 1 \stackrel{\text{Def}}{=} 4. \end{aligned}$$

1. Mengen und Zahlen

1.1. Mengen und ihre Verknüpfungen

Georg Cantor begründete die aus der Schule bekannte Mengenlehre mit folgender Definition:

Eine Menge ist die Zusammenfassung von Objekten,
die Gegenstand unseres Denkens sein können, zu einer Gesamtheit.

BEISPIEL 1.1.1

$M = \{2, \text{Mond}, \text{Julius Cäsar}\}$, dann ist $\text{Mond} \in M$ aber $\text{NewYork} \notin M$. Eine Menge kann auch nur ein einziges Element enthalten: $M = \{m\}$, die Objekte M und m sind dann aber verschieden. Die leere Menge wird mit dem Symbol \emptyset bezeichnet.

Von besonderer Bedeutung in der Analysis sind die Mengen

- $\mathbb{N} := \{1, 2, 3, \dots\}$ (Menge der natürlichen Zahlen),
- $\mathbb{N}_0 := \{0, 1, 2, 3, \dots\}$,
- $\mathbb{Z} := \{0, 1, -1, 2, -2, \dots\}$ (Menge der ganzen Zahlen),
- $\mathbb{Q} :=$ Menge rationalen Zahlen, d. h. der Brüche ganzer Zahlen,
- $\mathbb{R} :=$ Menge der reellen Zahlen.

Wir nehmen in dieser Vorlesung an, dass wir wissen, was unter natürlichen, ganzen und rationalen Zahlen zu verstehen ist. Jedoch wird es in diesem Kapitel unsere Hauptaufgabe sein, ein Axiomensystem für die Menge der reellen Zahlen (und die darauf definierten Rechenoperationen Addition, Multiplikation, etc.) zu entwickeln.

\mathbb{N} ist ein Teil von \mathbb{Z} , d. h. $n \in \mathbb{N} \Rightarrow n \in \mathbb{Z}$.

DEFINITION 1.1.1

Es seien M und N Mengen. M heißt Teilmenge von N (bzw. N Obermenge von M), falls jedes Element von M auch ein Element von N ist. Schreibweise: $M \subseteq N$ (bzw. $N \supseteq M$). Beispielsweise gilt $\mathbb{N} \subseteq \mathbb{Z}$, $\mathbb{Z} \subseteq \mathbb{Q}$ und $\mathbb{Q} \subseteq \mathbb{R}$. Wir schreiben zusammenfassend $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$. Die Verneinung dieser Relation schreibt man $M \not\subseteq N$ (M ist keine Teilmenge von N , d. h. es gibt ein Element in M , das kein Element von N ist). Die Vereinigung ist $M \cup N = \{x \mid x \in M \text{ oder } x \in N\}$, beispielsweise $M = \{1, 2, 3\}$, $N = \{2, 3, 5\}$, dann ist $M \cup N = \{1, 2, 3, 5\}$. Der Durchschnitt ist $M \cap N = \{x \mid x \in M \text{ und } x \in N\}$, für $M = \{1, 2, 3\}$ und $N = \{2, 3, 5\}$ gilt dann $M \cap N = \{2, 3\}$.

DEFINITION 1.1.2

Sind M und N Mengen, so definieren wir deren Differenz durch

$$M \setminus N = \{x \in M \mid x \notin N\}.$$

Ist \mathcal{S} eine Menge von Mengen (ein Mengensystem), so definieren wir die Vereinigung aller Mengen aus \mathcal{S} als

$$\bigcup_{M \in \mathcal{S}} M = \{x \mid \text{es gibt ein } M \in \mathcal{S} \text{ mit } x \in M\}.$$

Sind n Mengen M_1, M_2, \dots, M_n gegeben, so schreiben wir $M_1 \cup M_2 \cup \dots \cup M_n$ auch als

$$\bigcup_{k=1}^n M_k.$$

Ist für jedes $k \in \mathbb{N}$ eine Menge M_k gegeben, so schreiben wir statt

$$\bigcup_{k \in \mathbb{N}} M_k \quad \text{auch} \quad \bigcup_{k=1}^{\infty} M_k.$$

Eine entsprechende Schreibweise verwenden wir für den Durchschnitt:

$$\bigcap_{M \in \mathcal{S}} M = \{x \mid x \text{ liegt in jedem } M \in \mathcal{S}\}, \quad \bigcap_{k=1}^n M_k = M_1 \cap M_2 \cap \dots \cap M_n, \quad \bigcap_{k=1}^{\infty} = \bigcap_{k \in \mathbb{N}} M_k.$$

DEFINITION 1.1.3 (Komplementmenge)

Es seien alle Mengen $M \in \mathcal{S}$ eines Mengensystems \mathcal{S} Teilmengen einer „Universalmenge“ U . Wir bezeichnen das Komplement $U \setminus M$ einer Menge $M \in \mathcal{S}$ mit M' .

BEISPIEL 1.1.2

Es sei $U = \mathbb{N} = \{1, 2, 3, \dots\}$ sowie $M = \{2, 4, 6, 8, \dots\}$, dann ist $U \setminus M = M' = \{1, 3, 5, 7, \dots\}$.

Es gelten die Komplementierungsregeln von de Morgan:

$$\left(\bigcup_{M \in \mathcal{S}} M \right)' = \bigcap_{M \in \mathcal{S}} (M') \quad \text{und} \quad \left(\bigcap_{M \in \mathcal{S}} M \right)' = \bigcup_{M \in \mathcal{S}} (M'),$$

d. h. das Komplement der Vereinigung ist gleich dem Schnitt der Komplemente, und das Komplement des Schnitts ist gleich der Vereinigung der Komplemente. Der Beweis dieser Aussage kann als Modell für die Beweise der Gleichheit von Mengen angesehen werden. Sind M und N zwei Mengen, so bezeichnen wir die Aussage $x \in M$ mit A und $x \in N$ mit B . Es muss dann gezeigt werden: $A \Rightarrow B$ und $B \Rightarrow A$.

BEWEIS

$$(K1) \quad A: x \in \left(\bigcup_{M \in \mathcal{S}} M \right)', \quad B: x \in \bigcap_{M \in \mathcal{S}} (M').$$

$A \Rightarrow B$:

$$\begin{aligned} x \in \left(\bigcup M \right)' &\Rightarrow (x \in U \text{ aber } x \notin \bigcup M) \Rightarrow (x \in U \text{ und } \forall M \in \mathcal{S}: x \notin M) \\ &\Rightarrow (\forall M \in \mathcal{S}: x \in M') \Rightarrow x \in \bigcap (M'). \end{aligned}$$

$B \Rightarrow A$:

$$\begin{aligned} x \in \bigcap (M') &\Rightarrow (\forall M \in \mathcal{S}: x \in M') \Rightarrow (x \in U \text{ und } \forall M \in \mathcal{S}: x \notin M) \\ &\Rightarrow (x \in U \text{ und } x \notin \bigcup M) \Rightarrow x \in \left(\bigcup M \right)'. \end{aligned}$$

Analog für die zweite Komplementierungsregel:

$$(K2) \quad A: x \in \left(\bigcap_{M \in \mathcal{S}} M \right)', \quad B: x \in \bigcup_{M \in \mathcal{S}} (M').$$

$A \Rightarrow B$:

$$\begin{aligned} x \in \left(\bigcap M \right)' &\Rightarrow (x \in U \text{ aber } x \notin \bigcap M) \Rightarrow (x \in U \text{ und } \exists M \in \mathcal{S}: x \notin M) \\ &\Rightarrow (\exists M \in \mathcal{S}: x \in M') \Rightarrow x \in \bigcup (M'). \end{aligned}$$

$B \Rightarrow A$:

$$\begin{aligned}x \in \bigcup(M') &\Rightarrow (\exists M \in \mathcal{S} : x \in M') \Rightarrow (x \in U \text{ und } \exists M \in \mathcal{S} : x \notin M) \\ &\Rightarrow (x \in U \text{ und } x \notin \bigcap M) \Rightarrow x \in (\bigcap M)'.\end{aligned}$$

□

DEFINITION 1.1.4

Zwei Mengen A und B heißen disjunkt, falls $A \cap B = \emptyset$ ist. Ein endliches oder unendliches Mengensystem \mathcal{P} von Teilmengen einer vorgegebenen Menge M heißt eine Partition von M , falls gilt:

$$M = \bigcup_{T \in \mathcal{P}} T \text{ und } S \cap T = \emptyset \text{ falls } S \neq T \text{ mit } S, T \in \mathcal{P}.$$

BEISPIEL 1.1.3

Eine mögliche Partition von $M = \mathbb{N}_0$ ist $\mathcal{P} = \{T_1, T_2, T_3\}$ mit

$$T_1 = \{0, 3, 6, 9, \dots\}, \quad T_2 = \{1, 4, 7, 10, \dots\}, \quad T_3 = \{2, 5, 8, 11, \dots\}.$$

DEFINITION 1.1.5

Eine Äquivalenzrelation ist eine zweistellige Relation \sim auf einer Menge M , die folgende Bedingungen erfüllt:

- (Ä1) Reflexivität: für alle $x \in M$ gilt $x \sim x$,
- (Ä2) Symmetrie: aus $x \sim y$ folgt $y \sim x$ für alle $x, y \in M$,
- (Ä3) Transitivität: aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ für alle $x, y, z \in M$.

Jede Partition \mathcal{P} einer Menge M erzeugt eine Äquivalenzrelation $\sim_{\mathcal{P}}$ auf M vermöge

$$x \sim_{\mathcal{P}} y \Leftrightarrow \exists T \in \mathcal{P} : x, y \in T.$$

Umgekehrt erzeugt jede Äquivalenzrelation \sim eine Partition \mathcal{P}_{\sim} von M wie folgt:

$$\mathcal{P}_{\sim} = \{T_x \mid x \in M \text{ und } T_x = \{y \in M \mid x \sim y\}\}.$$

In diesem Sinne entsprechen sich die Begriffe der Partition und der Äquivalenzrelation.

1.2. Die Axiome der reellen Zahlen

Die Körperaxiome

Ein Körper K ist eine Menge von Objekten, für die zwei Rechenoperationen erklärt sind, Addition und Multiplikation genannt, die jedem Paar (a, b) von Elementen $a, b \in K$ die Summe $a + b$ und das Produkt $a \cdot b$ zuordnen, so dass folgende Axiome erfüllt sind:

Axiome der Addition:

- (A1) Assoziativgesetz: für alle $a, b, c \in K$ gilt $(a + b) + c = a + (b + c)$,
- (A2) Neutrales Element: es gibt ein Element $0 \in K$, so dass $0 + a = a + 0 = a$ ist für alle $a \in K$,
- (A3) Existenz des Inversen: für jedes $a \in K$ gibt es ein Inverses $b \in K$, so dass $a + b = b + a = 0$ gilt (b wird dann mit $-a$ bezeichnet),
- (A4) Kommutativgesetz: es gilt $a + b = b + a$ für alle $a, b \in K$.

Axiome der Multiplikation:

- (M1) Assoziativgesetz: für alle $a, b, c \in K$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$,
- (M2) Neutrales Element: es gibt ein Element $1 \in K$ mit $1 \neq 0$, so dass $1 \cdot a = a \cdot 1 = a$ ist für alle $a \in K$,
- (M3) Existenz des Inversen: für jedes $a \in K \setminus \{0\}$ gibt es ein Inverses $b \in K$, so dass $a \cdot b = b \cdot a = 1$ gilt (b wird dann mit a^{-1} bezeichnet),
- (M4) Kommutativgesetz: es gilt $a \cdot b = b \cdot a$ für alle $a, b \in K$.

Verbindung von Addition und Multiplikation:

- (D1) Distributivgesetz 1: für alle $a, b, c \in K$ gilt $a \cdot (b + c) = a \cdot b + a \cdot c$,
- (D2) Distributivgesetz 2: für alle $a, b, c \in K$ gilt $(b + c) \cdot a = b \cdot a + c \cdot a$.

Die Begriffe der Gruppe und des Rings

Eine Menge G mit einer einzigen Rechenoperation \circ , welche die Axiome A1, A2 und A3 erfüllt, heißt Gruppe. Das neutrale Element wird gewöhnlich mit e bezeichnet. Die Axiome A1-A4 schreiben sich dann wie folgt:

- (A1) $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$,
- (A2) es gibt ein Element $e \in G$ mit $e \circ a = a \circ e = a$ für alle $a \in G$,
- (A3) zu jedem Element $a \in G$ gibt es ein $a^{-1} \in G$ mit $a \circ a^{-1} = a^{-1} \circ a = e$,
- (A4) es gilt $a \circ b = b \circ a$ für alle $a, b \in G$.

Eine Menge R mit zwei Operationen, genannt Addition bzw. Multiplikation (geschrieben $+$ bzw. \cdot) heißt Ring, wenn die Axiome A1-A4, M1 und M2 sowie D1 und D2 erfüllt sind. Gilt auch M4, so heißt R kommutativer Ring. Die ganzen Zahlen \mathbb{Z} bilden mit $+$ und \cdot beispielsweise einen Ring. Dort ist das Axiom M3 allerdings nicht erfüllt. Erfüllt die Menge R mit den Operationen $+$ und \cdot die Axiome A1-A4, M1-M4 und D1/D2, so heißt R ein Körper. Die gängigsten Beispiele für Körper sind natürlich $K = \mathbb{Q}$ sowie $K = \mathbb{R}$ mit der üblichen Addition/Multiplikation. Es gibt aber auch ganz andere Typen von Körpern, beispielsweise die Menge $M = \{0, 1\}$ mit den Rechenoperationen

$$\begin{array}{c|cc} \oplus & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}, \quad \begin{array}{c|cc} \boxminus & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Dieses Beispiel zeigt, dass die bisher aufgestellten Axiome nicht genügen, um die reellen Zahlen zu charakterisieren. Die Menge \mathbb{R} wird oft mit dem Zahlenstrahl identifiziert. Eine derartige Veranschaulichung ist für den obigen Körper unmöglich. Aus dem Konzept des Strahls werden weitere Axiome zur Eingrenzung von \mathbb{R} abgeleitet.

Die Ordnungsaxiome

Auf \mathbb{R} ist eine „Kleiner-Beziehung“ erklärt (geschrieben $a < b$ oder $b > a$), so dass die folgenden Axiome erfüllt sind:

- (O1) Trichotomie: für $a, b \in \mathbb{R}$ gilt stets genau eine der drei Beziehungen $a < b$, $a = b$ oder $a > b$,

- (O2) Transitivität: für $a, b, c \in \mathbb{R}$ folgt aus $a < b$ und $b < c$ stets $a < c$,
(O3) Monotoniegesetz: ist $a < b$, so gilt auch $a + c < b + c$ für jedes $c \in \mathbb{R}$ sowie $a \cdot c < b \cdot c$ für jedes $c > 0$.

Die Schreibweise $a \leq b$ (bzw. $b \geq a$) bedeutet, dass $a < b$ oder $a = b$ ist. Ein Körper K mit einer zweistelligen Relation $<$, für die O1-O3 gilt, heißt angeordneter Körper.

BEISPIEL 1.2.1

Die Körper $K = \mathbb{Q}$ und $K = \mathbb{R}$ bilden mit der üblichen kleiner-Relation $<$ jeweils angeordnete Körper, nicht jedoch $M = \{0, 1\}$, und zwar völlig unabhängig davon, wie $<$ auf M definiert wird. Im Folgenden kann man sich unter K daher stets \mathbb{Q} oder \mathbb{R} vorstellen.

DEFINITION 1.2.1

Eine Zahl $a \in \mathbb{R}$ heißt positiv, falls $a > 0$ ist, negativ, falls $a < 0$ ist, nichtnegativ, falls $a \geq 0$ und nichtpositiv, falls $a \leq 0$ ist.

Die wichtigsten Teilmengen von \mathbb{R} sind die Intervalle. Sie können für beliebige angeordnete Körper gebildet werden:

DEFINITION 1.2.2

Es seien $a, b \in K$ (K angeordneter Körper) mit $a < b$, dann nennen wir

- (i) $[a, b] := \{x \in K \mid a \leq x \leq b\}$ ein abgeschlossenes Intervall,
- (ii) $(a, b) := \{x \in K \mid a < x < b\}$ ein offenes Intervall,
- (iii) $[a, b) := \{x \in K \mid a \leq x < b\}$ bzw.
- (iv) $(a, b] := \{x \in K \mid a < x \leq b\}$ halboffene Intervalle.

DEFINITION 1.2.3

Schließlich definieren wir noch einseitig und zweiseitig unendliche Intervalle mittels der Symbole ∞ (unendlich) und $-\infty$ (minus unendlich):

- (i) $(-\infty, a) := \{x \in K \mid x < a\}$, $(-\infty, a] := \{x \in K \mid x \leq a\}$,
- (ii) $(a, \infty) := \{x \in K \mid x > a\}$, $[a, \infty) := \{x \in K \mid x \geq a\}$,
- (iii) $(-\infty, \infty) := K$.

BEMERKUNG 1.2.1

∞ und $-\infty$ sind reine *Symbole*, sie repräsentieren keine Elemente von K , und insbesondere *keine reellen Zahlen*.

DEFINITION 1.2.4

Wir sagen, dass die Zahl μ das kleinste Element (oder das Minimum) einer Menge $M \subseteq K$ (angeordneter Körper) ist, geschrieben: $\mu = \min M$, wenn μ selbst Element von M ist und $\mu \leq x$ für alle $x \in M$ gilt. Dem entsprechend definieren wir das größte Element (oder das Maximum).

BEISPIEL 1.2.2

Es seien $a, b \in K$ (angeordneter Körper) mit $a < b$. Das abgeschlossene Intervall $I_1 = [a, b]$ besitzt das Minimum a und das Maximum b . Das halboffene Intervall $I_2 = (a, b]$ hat das Maximum b , es besitzt aber kein Minimum, da a kein Element von I_2 ist.

Wir benötigen auch für solche Mengen das Konzept eines abschließenden Punktes:

DEFINITION 1.2.5

Ein $x \in K$ heißt obere Schranke einer Menge $M \subseteq K$, falls $y \leq x$ für jedes $y \in M$ gilt. Ein $x \in K$ heißt untere Schranke von M , falls $y \geq x$ für jedes $y \in M$ gilt.

DEFINITION 1.2.6

Eine Teilmenge M eines angeordneten Körpers K heißt nach oben beschränkt, falls es eine obere Schranke von M in K gibt. Analog heißt M nach unten beschränkt, falls es eine untere Schranke von M in K gibt. M heißt beschränkt schlechthin, wenn M nach oben und unten beschränkt ist, d. h. falls es $a < b$ in K gibt mit $M \subseteq [a, b]$.

Untere und obere Schranken sind im Gegensatz zu minimalen/maximalen Elementen nicht eindeutig.

BEISPIEL 1.2.3

Das Intervall $[0, \infty) \subset K$ ist nach unten, nicht aber nach oben beschränkt. Jedes nichtpositive $x \in K$ ist eine untere Schranke des Intervalls.

BEISPIEL 1.2.4

Ist $a < b$ für $a, b \in K$, so ist $[a, b] \subset K$ beschränkt mit oberer Schranke b (die dann auch das Maximum ist) und unterer Schranke a (die dann auch das Minimum ist).

Das Beispiel eines offenen Intervalls zeigt, dass es für gewisse Mengen einen abschließenden Punkt gibt, der jedoch nicht Teil der Menge ist.

DEFINITION 1.2.7

Ein Element s eines angeordneten Körpers K heißt kleinste obere Schranke oder Supremum einer Menge $M \subseteq K$ (geschrieben $s = \sup M$), wenn s eine obere Schranke von M ist, und überdies kein Element $s' < s$ in K obere Schranke von M sein kann, d. h. wenn es zu jedem $\varepsilon > 0$ in K mindestens ein $x \in M$ gibt mit $x > s - \varepsilon$. Analog heißt ein $i \in K$ größte untere Schranke oder Infimum von M ($i = \inf M$), falls i eine untere Schranke von M ist, und kein Element $i' > i$ in K existiert, das ebenfalls untere Schranke von M ist, d. h. für jedes $\varepsilon > 0$ in K gibt es ein $x \in M$ mit $x < i + \varepsilon$.

BEMERKUNG 1.2.2

Wenn eine nichtleere Menge M ein Supremum besitzt, so ist dieses eindeutig bestimmt.

Die bisher formulierten Axiome werden von $K = \mathbb{R}$ wie auch von $K = \mathbb{Q}$ erfüllt, zur vollständigen Charakterisierung von \mathbb{R} reichen sie daher nicht aus. Intuitiv kann man den Unterschied von \mathbb{Q} und \mathbb{R} so beschreiben: Die rationalen Zahlen \mathbb{Q} besitzen „Lücken“, die \mathbb{R} nicht besitzt. Man kann beispielsweise zeigen, dass es keine rationale Zahl $x = \frac{p}{q}$ gibt mit $x^2 = 2$ (Beweis später).

Zentrale Begriffe der Analysis wie Grenzwert, Stetigkeit und Differenzierbarkeit ergeben keinen Sinn, wenn der verwendete Zahlenbereich Lücken aufweist, beispielsweise können derartige Lücken die Existenz von Supremum und Infimum verhindern. Für einen angeordneten Körper K formulieren wir daher das als Supremumsprinzip bekannte Axiom

(S1) Jede nach oben beschränkte und nichtleere Teilmenge von K besitzt ein Supremum.

BEISPIEL 1.2.5

Es sei $M = \{x \in \mathbb{Q} \mid x^2 < 2 \text{ oder } x < 0\}$. Diese Menge ist nach oben beschränkt, zum Beispiel durch die obere Schranke 2, und offenbar nicht leer. Sie besitzt in \mathbb{R} ein Supremum, nämlich die positive Lösung der Gleichung $x^2 = 2$, die mit $\sqrt{2}$ bezeichnet wird. M ist auch eine Teilmenge von \mathbb{Q} , besitzt darin aber wegen $\sqrt{2} \notin \mathbb{Q}$ kein Supremum.

Das Axiom S1 unterscheidet also die Körper \mathbb{R} und \mathbb{Q} . Die reellen Zahlen sind durch die Axiome A1-A4, M1-M4, D1/D2, O1-O3 und S1 vollständig charakterisiert.

1.3. Folgerungen aus den Körperaxiomen

Die Körperaxiome beinhalten unter Anderem die Existenz

- (1) der Null $0 \in K$ mit $0 + a = a + 0 = a$ für alle $a \in K$,
- (2) des additiven Inversen $-a$ mit $a + (-a) = 0$ für alle $a \in K$,
- (3) der Eins $1 \in K$ mit $1 \cdot a = a \cdot 1 = a$ für alle $a \in K$,
- (4) des multiplikativen Inversen a^{-1} mit $a \cdot a^{-1} = a^{-1} \cdot a = 1$ für alle $a \in K \setminus \{0\}$.

Es wurde in den Körperaxiomen jedoch nicht die Eindeutigkeit gefordert, d. h. dass es jeweils genau ein neutrales bzw. inverses Element gibt.

In der Mathematik wird unter der Formulierung „es gibt ein ...“ verstanden, dass es mindestens ein, aber möglicherweise auch mehrere Elemente gibt, die eine bestimmte Bedingung erfüllen. Wir zeigen nun, dass es von jedem der in 1-4 beschriebenen Elemente genau eines gibt:

SATZ 1.3.1 (Eindeutigkeit der neutralen und inversen Elemente)

Voraussetzung: K ist ein Körper.

Behauptung: In K gibt es genau eine Null, genau eine Eins, zu jedem $a \in K$ genau ein additives Inverses $-a$ und zu jedem $a \in K \setminus \{0\}$ genau ein multiplikatives Inverses a^{-1} .

BEWEIS

Es besitze $\bar{0} \in \mathbb{R}$ die gleiche Neutralitätseigenschaft bzgl. der Addition wie das Element 0 aus dem Axiom A2, d. h. $\bar{0} + a = a + \bar{0} = a$ für alle $a \in K$. Dann ergibt sich für $a = 0$ die Gleichung $0 = \bar{0} + 0 = \bar{0}$, also $0 = \bar{0}$, d. h. die Elemente 0 und $\bar{0}$ sind tatsächlich gleich. Analog besitze $\bar{1}$ die Neutralitätseigenschaft der Eins 1 bzgl. der Multiplikation nach Axiom M2, d. h. $a \cdot \bar{1} = \bar{1} \cdot a = a$ für alle $a \in K$. Dann gilt $1 = \bar{1} \cdot 1 = \bar{1}$, und auch diese Elemente sind gleich. Nun sei $a \in K$ gegeben und $b, \bar{b} \in K$ seien additive Inverse (Negative) von a , also $a + b = a + \bar{b} = 0$. Addition von $-a$ ergibt:

$$(-a) + (a + b) = (-a) + (a + \bar{b}) \xrightarrow{A1} ((-a) + a) + b = ((-a) + a) + \bar{b} \xrightarrow{A3} 0 + b = 0 + \bar{b} \xrightarrow{A2} b = \bar{b}.$$

Nun sei $a \in K \setminus \{0\}$ gegeben und $c, \bar{c} \in K$ seien multiplikative Inverse von a , d. h. $a \cdot c = a \cdot \bar{c} = 1$. Multiplikation mit a^{-1} ergibt

$$a^{-1} \cdot (a \cdot c) = a^{-1} \cdot (a \cdot \bar{c}) \xrightarrow{M1} (a^{-1} \cdot a) \cdot c = (a^{-1} \cdot a) \cdot \bar{c} \xrightarrow{M3} 1 \cdot c = 1 \cdot \bar{c} \xrightarrow{M2} c = \bar{c}.$$

□

Als Nächstes führen wir gewisse Bezeichnungen ein und beweisen die von der Schule bekannten Vorzeichenregeln. Zur kurzen Schreibweise von Definitionen gehen wir von folgender Konvention aus: Ist A ein schon definierter Ausdruck, und soll B als A definiert werden, so schreiben wir $B := A$.

DEFINITION 1.3.1

Das Produkt $a \cdot b$ schreiben wir meist ohne den Multiplikationspunkt: $ab := a \cdot b$. Ebenso schreiben wir $a - b := a + (-b)$, $\frac{1}{a} := a^{-1}$ (für $a \neq 0$), $\frac{a}{b} := ab^{-1}$ (für $b \neq 0$), $a^2 := aa$, $a + b + c := a + (b + c) = (a + b) + c$, sowie $abc := a(bc) = (ab)c$.

SATZ 1.3.2 (Kürzungsregel)

Voraussetzung: K Körper und $a, b, c \in K$.

Behauptung: Es gilt

- (i) $a + b = a + c \Rightarrow b = c$,
- (ii) $b + a = c + a \Rightarrow b = c$, ist ferner $a \neq 0$, so gilt
- (iii) $ab = ac \Rightarrow b = c$,
- (iv) $ba = ca \Rightarrow b = c$.

BEWEIS

$$\begin{aligned} \text{Zu Teil (i): } \quad a + b = a + c &\Rightarrow -a + (a + b) = -a + (a + c) \\ &\xrightarrow{A1} (-a + a) + b = (-a + a) + c \Rightarrow 0 + b = 0 + c \xrightarrow{A2} b = c. \end{aligned}$$

Teil (ii) folgt aus (i) mittels des Kommutativgesetzes A4.

$$\text{Zu Teil (iii): } ab = ac \Rightarrow a^{-1}(ab) = a^{-1}(ac) \xrightarrow{\text{M1}} (a^{-1}a)b = (a^{-1}a)c \xrightarrow{\text{M3}} 1b = 1c \xrightarrow{\text{M2}} b = c.$$

Teil (iv) folgt daraus mit M4. □

SATZ 1.3.3 (Vorzeichenregeln)

Voraussetzung: K Körper und $a, b \in K$.

Behauptung: Es gilt

- (i) $0 \cdot a = a \cdot 0 = 0$,
- (ii) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$, $(-a)(-b) = ab$,
- (iii) $ab = 0 \Leftrightarrow a = 0$ oder $b = 0$.

BEWEIS

$$(i) : 0 \cdot a = (0 + 0) \cdot a \xrightarrow{\text{D2}} 0 \cdot a + 0 \cdot a \Rightarrow 0 \cdot a + 0 = 0 \cdot a + 0 \cdot a \xrightarrow{\text{Satz 1.3.2(i)}} 0 \cdot a = 0.$$

Daraus folgt dann $a \cdot 0 = 0$ mittels der Kommutativität. Zu Teil (ii): nach (i) ist

$$0 = 0 \cdot b \xrightarrow{\text{A3}} (a + (-a)) \cdot b \xrightarrow{\text{D2}} a \cdot b + (-a) \cdot b \quad \text{und}$$

$$a \cdot b + (-a) \cdot b = 0, \text{ also}$$

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

$$\xrightarrow{\text{1.3.2(i)}} (-a) \cdot b = -ab.$$

Damit folgt dann auch

$$a \cdot (-b) \xrightarrow{\text{M4}} (-b) \cdot a = -(ba) = -(ab), \text{ sowie}$$

$$0 = (-a)(b + (-b)) \xrightarrow{\text{D1}} (-a) \cdot b + (-a) \cdot (-b) \xrightarrow{\text{gezeigt}} -(ab) + (-a) \cdot (-b), \text{ also } 0 = -(ab) + ab \\ \Rightarrow -(ab) + (-a) \cdot (-b) = -(ab) + ab \xrightarrow{\text{1.3.2(i)}} (-a) \cdot (-b) = ab.$$

Zu Teil (iii): Falls $a = 0$ ist, folgt die Aussage trivialerweise. Für $a \neq 0$ gilt

$$a \cdot b = a \cdot 0 = 0 \xrightarrow[\text{a} \neq 0]{\text{1.3.2(iii)}} b = 0.$$

□

SATZ 1.3.4 (Bruchrechenregeln)

Voraussetzung: K Körper, $a, b, c, d \in K$ mit $b, d \neq 0$.

Behauptung: es gilt

$$(i) \frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \quad (ii) \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

BEWEIS

Teil (i), hier nur für das Pluszeichen:

$$bd \cdot \left(\frac{a}{b} + \frac{c}{d} \right) \xrightarrow{\text{D1}} (bd) \cdot \frac{a}{b} + (bd) \cdot \frac{c}{d} \xrightarrow{\text{Def.}} (bd) \cdot (b^{-1}a) + (bd) \cdot (d^{-1}c) = d(bb^{-1})a + b(dd^{-1})c = ad + bc. \\ \Rightarrow bd \cdot \left(\frac{a}{b} + \frac{c}{d} \right) = ad + bc$$

$$\Rightarrow \left(\frac{a}{b} + \frac{c}{d}\right) = \frac{ad + bc}{bd}.$$

Zu Teil (ii):

$$bd \cdot \left(\frac{a}{b} \cdot \frac{c}{d}\right) = d(bb^{-1}a)d^{-1}c = dad^{-1}c = a(dd^{-1})c = ac \Rightarrow \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

□

1.4. Folgerungen aus den Ordnungsaxiomen

Die folgenden Sätze werden für die reellen Zahlen formuliert. Da zu ihrem Beweis nur die bisher aufgestellten Axiome benutzt werden, gelten sie aber in jedem angeordneten Körper (z. B. \mathbb{Q}).

SATZ 1.4.1

Voraussetzung: $a, b \in \mathbb{R}$.

Behauptung: es gilt

- (i) $a < b \Leftrightarrow b - a > 0$,
- (ii) $a < 0 \Leftrightarrow -a > 0$,
- (iii) $a > 0 \Leftrightarrow -a < 0$,
- (iv) $a < b \Leftrightarrow -b < -a$.

BEWEIS

Wir zeigen zuerst die Richtung \Rightarrow von Teil (i):

$$a < b \stackrel{O_3}{\Rightarrow} 0 = a + (-a) < b + (-a) = b - a \Rightarrow b - a > 0.$$

Nun zur Richtung \Leftarrow : $b - a > 0 \Rightarrow b - a + a > 0 + a \Rightarrow b > a$.

Teil (ii) folgt aus (i) für $b = 0$. Teil (iii) folgt wiederum aus (ii), wenn man a durch $-a$ ersetzt und $-(-a) = a$ beachtet. Teil (iv) folgt aus (i):

$$a < b \stackrel{(i)}{\Leftrightarrow} b - a > 0 \Leftrightarrow (-a) - (-b) > 0 \Leftrightarrow -b < -a.$$

□

SATZ 1.4.2 (Addieren gleichsinniger Ungleichungen)

Voraussetzung: $a, b, c, d \in \mathbb{R}$ mit $a < b$ und $c < d$.

Behauptung: $a + c < b + d$.

BEWEIS

Wegen dem Monotoniegesetz gilt $a < b \Rightarrow a + c < b + c$ sowie $c < d \Rightarrow b + c < b + d$, also $a + c < b + c < b + d$, mit dem Transitivitätsgesetz folgt damit $a + c < b + d$. □

BEMERKUNG 1.4.1

Eine Ungleichung $a < b$ nennt man auch eine Abschätzung, und man sagt, man habe „ a nach oben durch b “ abgeschätzt (bzw. b von unten durch a abgeschätzt).

Im nächsten Satz lernen wir einen so genannten indirekten Beweis oder auch Widerspruchsbeweis kennen. Eine Aussage A wird bewiesen, indem das Gegenteil von A angenommen wird, um daraus einen Widerspruch, d. h. eine bekanntermaßen falsche Aussage, herzuleiten.

SATZ 1.4.3

Voraussetzung: $a, b \in \mathbb{R}$.

Behauptung: $ab > 0 \Leftrightarrow$ die Faktoren a, b sind entweder beide positiv oder beide negativ.

BEWEIS

Zunächst die Richtung \Leftarrow , es gilt

$$a, b > 0 \underset{O_3}{\Rightarrow} a \cdot b > 0 \cdot b = 0 \quad \text{aber auch} \quad a, b < 0 \underset{O_3}{\Rightarrow} ab = (-a) \cdot (-b) > 0$$

da für $a, b < 0$ die Negate $-a$ und $-b$ positiv sind. Nun zur Richtung \Rightarrow . Aus $ab > 0$ folgt $a \neq 0$ und $b \neq 0$ nach Satz 1.3.3(iii). Wir zeigen nun die Aussage

A: a und b sind entweder beide positiv, oder beide negativ

durch einen Widerspruchsbeweis. Annahme: A ist falsch, dann ist einer der Faktoren negativ und der andere positiv. O.B.d.A.¹ sei $a > 0$ und $b < 0$. Aus Satz 1.4.1(ii) folgt $-b > 0$, also $-(ab) = a(-b) > 0$, woraus wieder mit Satz 1.4.1 die Ungleichung $ab < 0$ folgt. Diese Ungleichung steht im Widerspruch zur Voraussetzung $ab > 0$, woraus folgt, dass die Annahme des Gegenteils von A falsch war, also muss A gelten. \square

SATZ 1.4.4

Voraussetzung: $a \in \mathbb{R}$.

Behauptung: Aus $a \neq 0$ folgt $a^2 > 0$, insbesondere ist $1 > 0$.

BEWEIS

Das folgt aus Satz 1.4.3, wenn man $a = b$ setzt. \square

SATZ 1.4.5

Voraussetzung: $a, b, c \in \mathbb{R}$ mit $a < b$ und $c < 0$.

Behauptung: $ac > bc$.

BEWEIS

Aus $c < 0$ folgt $-c > 0$ (Satz 1.4.1). Nach O_3 ist dann $-(ac) = a(-c) < b(-c) = -(bc)$, also $-(ac) < -(bc)$. Mit Satz 1.4.1(iv) erhalten wir: $bc < ac$ oder $ac > bc$. \square

Daraus ergibt sich die Merkregel, dass Ungleichungen mit positiven reellen Zahlen unter Beibehaltung des Ungleichheitszeichens multipliziert werden können, bei Multiplikation mit negativen reellen Zahlen muss das Ungleichheitszeichen dagegen umgedreht werden.

SATZ 1.4.6 (Multiplikation gleichsinniger Ungleichungen)

Voraussetzungen: $a, b, c, d \in \mathbb{R}$ mit $0 < a < b$ und $0 < c < d$.

Behauptung: $ac < bd$.

BEWEIS

Aus $a < b$ folgt $ac < bc$, aus $c < d$ folgt $bc < bd$, also gilt insgesamt $ac < bd$. \square

SATZ 1.4.7

Voraussetzung: $a, b \in \mathbb{R}$ mit $b \neq 0$.

Behauptung: $\frac{a}{b} > 0 \Leftrightarrow$ Zähler und Nenner sind entweder beide positiv oder beide negativ.

¹„Ohne Beschränkung der Allgemeinheit“, ein in der Mathematik übliches Kürzel um anzudeuten, dass eine Fallunterscheidung nur für einen Fall durchgeführt wird, weil der andere Fall in gleicher Weise bewiesen wird oder trivial ist.

BEWEIS

$$b \cdot \frac{1}{b} = 1 > 0 \quad \stackrel{\text{Satz 1.4.3}}{\Leftrightarrow} \quad \text{entweder } b, \frac{1}{b} > 0 \text{ oder } b, \frac{1}{b} < 0,$$

$$\frac{a}{b} = a \cdot \frac{1}{b} > 0 \quad \stackrel{\text{Satz 1.4.3}}{\Leftrightarrow} \quad \text{entweder } a, \frac{1}{b} > 0 \text{ oder } a, \frac{1}{b} < 0.$$

□

SATZ 1.4.8

Voraussetzung: $p_1, p_2, q \in \mathbb{R}$ mit $q > 0$ und $0 < p_1 < p_2$, sowie $q_1, q_2, p \in \mathbb{R}$ mit $p > 0$ und $0 < q_1 < q_2$.

Behauptung: $\frac{p_1}{q} < \frac{p_2}{q}$ sowie $\frac{p}{q_2} < \frac{p}{q_1}$. Insbesondere gilt $\frac{1}{q_2} < \frac{1}{q_1}$.

BEWEIS

Da $\frac{1}{q}$ wegen Satz 1.4.7 positiv ist, folgt die erste Behauptung aus O3. Zum Beweis der zweiten Ungleichung multiplizieren wir die Ungleichung $q_1 < q_2$ mit dem Bruch $\frac{p}{q_1 q_2}$. □

SATZ 1.4.9 (Ungleichung des arithmetischen Mittels)

Es seien $a, b \in \mathbb{R}$. Aus $a < b$ folgt: $a < \frac{a+b}{2} < b$.

BEWEIS

$$a < b \Rightarrow 2a = a + a < a + b < b + b = 2b.$$

□

Folgerung: zwischen zwei beliebigen reellen Zahlen liegt stets eine weitere Zahl. Das gilt ebenso für \mathbb{Q} .

1.5. Die natürlichen, ganzen und rationalen Zahlen

Es ist möglich, eine Definition für die Menge der natürlichen Zahlen zu geben, die sich auf das Axiomensystem für die reellen Zahlen gründet.

DEFINITION 1.5.1

Eine Menge M reeller Zahlen heißt induktive Menge, wenn sie die folgenden Eigenschaften erfüllt:

- (i) $1 \in M$,
- (ii) $a \in M \Rightarrow a + 1 \in M$.

Beispielsweise ist \mathbb{R} selbst eine induktive Menge.

DEFINITION 1.5.2

Die Menge \mathbb{N} der natürlichen Zahlen ist definiert als der Durchschnitt aller induktiven Teilmengen von \mathbb{R} , und \mathbb{N}_0 ist $\mathbb{N} \cup \{0\}$.

Damit können wir auch die Menge \mathbb{Z} der ganzen Zahlen sowie die Menge \mathbb{Q} der rationalen Zahlen definieren:

DEFINITION 1.5.3

$\mathbb{Z} := \{n \mid n \in \mathbb{N}\} \cup \{-n \mid n \in \mathbb{N}\} \cup \{0\}$, sowie $\mathbb{Q} := \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$.

SATZ 1.5.1 (Induktionsprinzip)

Die Menge \mathbb{N} ist induktiv. Jede induktive Menge natürlicher Zahlen stimmt mit \mathbb{N} überein.

BEWEIS

Da 1 allen induktiven Mengen angehört, liegt sie auch im Schnitt \mathbb{N} aller solchen Mengen. Ist $a \in \mathbb{N}$, so liegt a in allen induktiven Mengen, damit liegt $a + 1$ in allen induktiven Mengen und damit im Schnitt \mathbb{N} , also ist \mathbb{N} induktiv. Da \mathbb{N} der Durchschnitt aller induktiven Mengen ist, ist \mathbb{N} in jeder induktiven Menge enthalten. Ist daher M eine induktive Menge natürlicher Zahlen, so folgt $\mathbb{N} \subseteq M \subseteq \mathbb{N}$ und damit $M = \mathbb{N}$. \square

SATZ 1.5.2

Die Menge der natürlichen Zahlen ist additiv und multiplikativ abgeschlossen, d. h. Summe und Produkt natürlicher Zahlen sind wieder natürliche Zahlen.

BEWEIS

Es sei $m \in \mathbb{N}$ beliebig. Wir betrachten die Menge

$$A = A(m) := \{n \in \mathbb{N} \mid m + n \in \mathbb{N}\}.$$

Mit m liegt auch $m + 1$ in \mathbb{N} , da \mathbb{N} induktiv ist, also ist 1 ein Element von A . Nun sei $n \in A$ beliebig, d. h. $m + n \in \mathbb{N}$. Da \mathbb{N} induktiv ist folgt $m + n + 1 \in \mathbb{N}$, also $n + 1 \in A$. Insgesamt gilt also $1 \in A$ und $n \in A \Rightarrow n + 1 \in A$, d. h. A ist induktiv. Nach dem Induktionsprinzip ist $A = \mathbb{N}$, gleichbedeutend damit, dass für alle $n \in \mathbb{N}$ die Summe $m + n$ wieder in \mathbb{N} liegt. Da aber m beliebig gewählt war ist $m + n \in \mathbb{N}$ für alle $m, n \in \mathbb{N}$. Analog betrachten wir nun für ein beliebiges $m \in \mathbb{N}$ die Menge

$$M = M(m) := \{n \in \mathbb{N} \mid m \cdot n \in \mathbb{N}\}.$$

Das Produkt $m \cdot 1 = m$ liegt in \mathbb{N} , also $1 \in M$. Nun sei $n \in M$, d. h. $m \cdot n \in \mathbb{N}$. Dann gilt

$$m \cdot (n + 1) \stackrel{D1}{=} m \cdot n + m \cdot 1 \stackrel{M2}{=} m \cdot n + m.$$

Das Produkt $m \cdot n$ liegt wegen $n \in M$ in \mathbb{N} , ebenso m , aus der bereits gezeigten Abgeschlossenheit bzgl. der Addition folgt $m \cdot n + m \in \mathbb{N}$, und damit $n + 1 \in M$. Damit ist auch M induktiv, d. h. $M = \mathbb{N}$ und $m \cdot n$ liegt in \mathbb{N} für alle $m, n \in \mathbb{N}$. \square

SATZ 1.5.3

Für jedes $n \in \mathbb{N}$ gilt $n \geq 1$.

BEWEIS

Die Menge $M := \{x \in \mathbb{R} \mid x \geq 1\}$ ist induktiv. Da \mathbb{N} in jeder induktiven Menge enthalten ist gilt: $n \in \mathbb{N} \Rightarrow n \in M \Rightarrow n \geq 1$. \square

SATZ 1.5.4

Zwischen den natürlichen Zahlen n und $n + 1$ liegt keine weitere natürliche Zahl.

BEWEIS

Der Beweis gliedert sich in mehrere Schritte:

Schritt A:

Wir betrachten die Menge $M := \{1\} \cup \{x \in \mathbb{R} \mid x \geq 2\}$ und zeigen, dass M induktiv ist, d. h. $\mathbb{N} \subseteq M$. Zunächst ist nach Definition $1 \in M$. Nun sei $a \in M$ beliebig.

Fall 1: $a = 1$, dann ist $a + 1 = 2 \in \{x \in \mathbb{R} \mid x \geq 2\}$, also $a + 1 \in M$.

Fall 2: $a \in \{x \in \mathbb{R} \mid x \geq 2\}$, dann ist auch $a + 1 \in \{x \in \mathbb{R} \mid x \geq 2\} \subseteq M$.

Insgesamt folgt, dass $1 \in M$ ist und $a \in M \Rightarrow a + 1 \in M$ gilt, d. h. M ist induktiv.

Schritt B:

Wir zeigen: es gibt kein $m \in \mathbb{N}$ mit $1 < m < 2$. Dies zeigen wir durch einen Widerspruchsbeweis. Annahme: Es gibt $m \in \mathbb{N}$ mit $1 < m < 2$. Dann ist nach Teil A (wegen $\mathbb{N} \subseteq M$) aber $m \in M$. Für alle

$m \in M$ gilt jedoch entweder $m = 1$ oder $m \geq 2$, ein Widerspruch.

Schritt C:

Wir betrachten die Menge $N := \{n \in M \mid n - 1 \in \mathbb{N}_0\}$ und zeigen: N ist induktiv, d. h. $N = \mathbb{N}$. Zunächst ist $1 \in N$ wegen $1 - 1 = 0 \in \mathbb{N}_0$. Ist $a \in N$, so ist $a - 1 \in \mathbb{N}_0$. Dann ist auch $a + 1 \in N$, da $(a + 1) - 1 = a \in \mathbb{N}_0$ ist. Also: $a \in N \Rightarrow a + 1 \in N$. Die Menge N ist induktiv, d. h. $N = \mathbb{N}$.

Schritt D:

Wir betrachten die Menge $K := \{n \in \mathbb{N} \mid \nexists m \in \mathbb{N} \text{ mit } n < m < n + 1\}$ und zeigen, dass K induktiv ist. Nach Teil B gilt $1 \in K$, da es kein $m \in \mathbb{N}$ gibt mit $1 < m < 2$. Es sei nun $a \in K$, d. h. es gibt kein $m \in \mathbb{N}$ mit $a < m < a + 1$. Annahme: Es gibt $m \in \mathbb{N}$ mit $a + 1 < m < a + 2$. Nach Teil C ist $m - 1 \in \mathbb{N}_0$ und wegen $m > a + 1 \geq 2$ folgt $m - 1 \in \mathbb{N}$. Weiter ist $a < m - 1 < a + 1$, ein Widerspruch zu $a \in K$. Die Annahme ist also falsch, woraus $a + 1 \in K$ folgt. Damit ist K induktiv und $K = \mathbb{N}$.

Schritt E:

Ist $n \in \mathbb{N}$ beliebig, so ist nach Teil D auch $n \in K$, d. h. es gibt kein m mit $n < m < n + 1$, woraus die Behauptung des Satzes folgt. \square

SATZ 1.5.5 (Wohlordnungsprinzip)

Jede nichtleere Menge natürlicher Zahlen besitzt ein Minimum.

(ohne Beweis)

Die entsprechende Aussage für das Maximum gilt nicht:

SATZ 1.5.6

Jede reelle Zahl wird von einer natürlichen Zahl übertroffen. Die Menge der natürlichen Zahlen ist nach oben unbeschränkt.

BEWEIS

Annahme: \mathbb{N} ist beschränkt. Dann existiert nach dem Supremumsprinzip $S := \sup \mathbb{N}$. Nach der Definition des Supremums (mit der Wahl $\varepsilon = 1$) gibt es ein $n \in \mathbb{N}$ mit $n > S - 1$. Dies führt zu $n + 1 > S$, was wegen $n + 1 \in \mathbb{N}$ der Bedeutung von S widerspricht. \square

1.6. Folgerungen aus dem Supremumsprinzip

Der Unterschied von \mathbb{Q} und \mathbb{R} wurde durch das Supremumsprinzip S1 formalisiert. Die ursprüngliche Idee, dass \mathbb{Q} im Gegensatz zu \mathbb{R} Lücken aufweist, kann auch durch die folgende Konstruktion beschrieben werden:

DEFINITION 1.6.1

Ein Dedekindscher Schnitt von K ist ein Paar von Mengen $D = (U, O)$, der Unterklasse $U \subseteq K$ und der Oberklasse $O \subseteq K$, so dass gilt:

- (i) Beide Klassen sind nicht leer,
- (ii) für alle $u \in U$ und $o \in O$ ist $u < o$,
- (iii) O besitzt kein Minimum,
- (iv) $\{U, O\}$ ist eine Partition von K .

Der bisher benutzte Begriff der „Lücke“ wird nun wie folgt beschrieben:

DEFINITION 1.6.2

Ein Element $t \in K$ heißt Trennungszahl eines Dedekindschen Schnittes $D = (U, O)$, falls $u \leq t < o$ für alle $u \in U$ und $o \in O$ gilt.

BEISPIEL 1.6.1

Für ein $t \in K$ sei D_t der Dedekindsche Schnitt, der durch $U = (-\infty, t]$ und $O = (t, \infty)$ gegeben ist. Offenbar ist dann t gerade die Trennungszahl des Schnitts D_t .

Für die reellen Zahlen ist tatsächlich jeder Dedekindscher Schnitt von dieser Form:

SATZ 1.6.1 (Schnittaxiom)

Für jeden Dedekindschen Schnitt D von \mathbb{R} existiert eine Trennungszahl t .

BEWEIS

Es sei $D = (U, O)$ ein beliebiger Dedekindscher Schnitt von \mathbb{R} . Da O nicht leer ist, gibt es ein $o \in O$, so dass $u < o$ für alle $u \in U$ gilt. Also ist U nach oben beschränkt und nicht leer. Nach dem Supremumsprinzip S1 besitzt U ein Supremum $t \in \mathbb{R}$. Da $\{U, O\}$ eine Partition von \mathbb{R} ist, liegt t in O oder in U . Angenommen es liegt in O : da O kein Minimum besitzt gibt es $x \in O$ mit $x < t$. Nach Definition des Supremums ist $x < t$ dann keine obere Schranke von U , ein Widerspruch zu $x \in O$ und der Bedingung (ii) aus Definition 1.6.1. Damit liegt das Supremum s in der Unterklasse U , und ist damit sogar das Maximum von U . Wegen $t \in U$ gilt zudem $t < x$ für alle $x \in O$, d. h. t ist eine Trennungszahl von D . \square

BEMERKUNG 1.6.1

Man kann leicht nachprüfen, dass damit jeder Dedekindsche Schnitt D der reellen Zahlen tatsächlich von der Form $D = D_t$ für die existierende Trennungszahl t von D ist.

BEISPIEL 1.6.2

In \mathbb{Q} besitzt nicht jeder Dedekindscher Schnitt eine Trennungszahl, beispielsweise $D = (U, O)$ mit $O = \{x \in \mathbb{Q} \mid x^2 > 2 \text{ und } x \geq 0\}$ und $U = \mathbb{Q} \setminus O$.

Zur Erklärung dieses Beispiels sei noch nachgereicht:

SATZ 1.6.2

Es gibt keine rationale Zahl $x \in \mathbb{Q}$ mit $x^2 = 2$.

BEWEIS

Angenommen $x = \frac{a}{b}$ ist eine solche rationale Zahl. Wir nehmen an, dass Zähler und Nenner maximal gekürzt sind, insbesondere sind a und b nicht beide gerade. Aus $(ab^{-1})^2 = 2$ folgt durch Umformen $a^2 = 2 \cdot b^2$, insbesondere ist a^2 gerade. Das Quadrat einer ungeraden Zahl ist wieder ungerade, also muss a gerade sein, etwa $a = 2k$ für ein $k \in \mathbb{N}$. Daraus folgt $2b^2 = (2k)^2 = 4k^2$, also $b^2 = 2k^2$, mit dem gleichen Argument ist dann auch b gerade im Widerspruch zur Annahme. \square

BEMERKUNG 1.6.2

Das Schnittaxiom wurde hier aus dem Axiom S1 abgeleitet. Tatsächlich sind beide Axiome äquivalent: man kann auch S1 aus der Annahme des Schnittaxioms herleiten. Beide Axiome formulieren den Unterschied von \mathbb{Q} und \mathbb{R} .

1.7. Rekursive Definitionen und induktive Beweise

In den Körperaxiomen treten zunächst nur Summen von zwei Summanden auf. Um Summen mit einer beliebigen Anzahl von Summanden zu erklären, verwenden wir das Induktionsprinzip (Satz 1.5.1). Ist die Summe $a_1 + \dots + a_n$ aus n Summanden schon erklärt, so wird die $(n+1)$ -gliedrige Summe $a_1 + \dots + a_n + a_{n+1}$ durch

$$a_1 + \dots + a_n + a_{n+1} := (a_1 + \dots + a_n) + a_{n+1}$$

erklärt. Nach dem Induktionsprinzip folgt, dass damit die n -gliedrige Summe für alle $n \in \mathbb{N}$ erklärt ist.

Definitionen dieser Art, bei denen die Begriffe oder Größen $A(n)$, die von der natürlichen Zahl n

abhängen, mit Hilfe einiger oder aller schon erklärten $A(1), \dots, A(n-1)$ bestimmt werden, nennt man rekursive Definitionen oder auch Definitionen durch vollständige Induktion.

DEFINITION 1.7.1

Es sei $a \in \mathbb{R}$ und $n \in \mathbb{N}$. Wir definieren die n -te Potenz von a rekursiv durch $a^1 := a$ und $a^{n+1} := a^n \cdot a$. Wir erweitern diese Definition zu $a^0 := 1$ (unabhängig von a , insbesondere ist $0^0 = 1$) und $a^{-n} = \frac{1}{a^n}$ falls $a \neq 0$ ist und $n \in \mathbb{N}$. Einen Ausdruck der Form a^x nennen wir Potenz mit Basis a und Exponenten x .

Es gibt auch Beweise durch Induktion (oder durch vollständige Induktion). Um eine Eigenschaft $A(n)$ für alle $n \in \mathbb{N}$ zu zeigen, geht man wie folgt vor: Man zeigt zunächst $A(1)$ (Induktionsanfang). Dann zeigt man, dass aus der Wahrheit von $A(n)$ die Wahrheit von $A(n+1)$ folgt (Induktionsschluss). Wir betrachten die Menge $R := \{n \in \mathbb{N} \mid A(n) \text{ wahr}\}$. Der Induktionsanfang liefert $1 \in R$, der Induktionsschluss $n \in R \Rightarrow n+1 \in R$. Also ist R induktiv, mithin $R = \mathbb{N}$, d. h. die Aussage $A(n)$ ist dann für alle $n \in \mathbb{N}$ wahr (Satz 1.5.1).

Wir demonstrieren dieses Prinzip anhand von

SATZ 1.7.1

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

BEWEIS

Es sei $A(n)$ die Aussage, dass $1 + 2 + \dots + n$ gleich $\frac{n(n+1)}{2}$ ist. Induktionsanfang: $A(1)$ ist wahr, denn auf der linken Seite steht 1, auf der rechten Seite $\frac{1 \cdot 2}{2} = 1$. Induktionsschluss: $A(n)$ wird als wahr vorausgesetzt, also

$$(1+2+\dots+n)+(n+1) \underset{A(n)}{=} \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2} \Rightarrow A(n+1).$$

□

DEFINITION 1.7.2

Sind $a_1, \dots, a_n \in \mathbb{R}$, so schreiben wir die Summe $a_1 + \dots + a_n$ auch als

$$\sum_{k=1}^n a_k,$$

und das Produkt $a_1 \cdot \dots \cdot a_n$ als

$$\prod_{k=1}^n a_k.$$

Der Summationsindex k kann durch jedes andere Symbol ersetzt werden.

BEISPIEL 1.7.1

Satz 1.7.1 schreibt sich dann als

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

BEISPIEL 1.7.2

Unter

$$\sum_{n=1}^N n^2$$

versteht man die Summe $1^2 + 2^2 + \dots + N^2$.

Man kann auch Summen mit anderen Summationsgrenzen betrachten:

$$\sum_{m=r}^n a_m \text{ bedeutet } a_r + a_{r+1} + \cdots + a_n .$$

Dabei können r und n auch in \mathbb{Z} liegen. Hilfreich ist die Technik der Indexverschiebung. Durch Umbenennung der Terme können die Summationsgrenzen verändert werden.

BEISPIEL 1.7.3

Es ist

$$2^3 + 3^3 + \cdots + 100^3 = \sum_{k=2}^{100} k^3 .$$

Wir führen die Substitution $k \rightarrow j + 2$ durch und erhalten

$$\sum_{k=2}^{100} k^3 = \sum_{j=0}^{98} (j+2)^3 .$$

Dabei verwendet man meist wieder den ursprünglichen Bezeichner für den Summationsindex.

DEFINITION 1.7.3

Die Fakultät einer natürlichen Zahl ist definiert durch

$$0! := 1 \quad , \quad 1! := 1 \quad , \quad (n+1)! := (n!) \cdot (n+1) \quad , \quad n = 1, 2, \dots .$$

(gesprochen „ n Fakultät“)

Wir definieren ferner für $n, k \in \mathbb{N}_0$ mit $n \geq k$ den Binomialkoeffizient

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

(gesprochen „ n über k “)

BEMERKUNG 1.7.1

Man sieht leicht, dass $\binom{n}{0} = \binom{n}{n} = 1$ und $\binom{n}{k} = \binom{n}{n-k}$ gilt.

SATZ 1.7.2

Es seien $1 \leq k \leq n$ natürliche Zahlen, dann gilt

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

BEWEIS

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

□

Aus dem nächsten Satz wird der Name Binomialkoeffizient ersichtlich:

SATZ 1.7.3 (Binomischer Lehrsatz)

Es seien $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$, dann gilt

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

BEWEIS

Der Beweis wird durch vollständige Induktion geführt.

$n = 1$:

Die linke Seite ist $a + b$, für die rechte Seite ergibt sich

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

$n \rightarrow n + 1$:

Die Aussage

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

wird als richtig angenommen (Induktionshypothese). Daraus folgt

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

Um die Terme in beiden Summanden zusammenfassen zu können, müssen die Potenzprodukte $a^i b^j$ in beiden Summe gleiche Gestalt haben. Dies wird durch Indexverschiebung erreicht. Wir spalten in der ersten Summe den letzten Term ab, substituieren mit $k \rightarrow j + 1$, schreiben wieder k für j , und erhalten:

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

Von der zweiten Summe spalten wir den ersten Term ab:

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

Wir erhalten insgesamt

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

Wir beachten, dass

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

gilt, und erhalten mit Satz 1.7.2 den Induktionsschluss

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

□

2. Funktionsbegriff und Zahlenfolgen

2.1. Der Funktionsbegriff

DEFINITION 2.1.1

Es seien X und Y zwei nichtleere Mengen. Unter einer Funktion oder Abbildung f von X nach Y versteht man eine Vorschrift, die jedem $x \in X$ genau ein $y \in Y$ zuordnet. Dieses dem Element x zugeordnete Element y bezeichnen wir mit $f(x)$ und nennen es den Wert der Funktion f an der Stelle x , oder das Bild von x unter f , während x das Urbild von $y = f(x)$ heißt. X wird die Definitionsmenge (oder auch der Definitionsbereich), Y die Zielmenge von f genannt.

BEMERKUNG 2.1.1

Die detaillierteste Darstellung einer Funktion geschieht in der Form

$$f : \begin{cases} X & \rightarrow & Y \\ x & \mapsto & f(x) \end{cases}.$$

Es werden also zunächst die Mengen angegeben, zwischen denen die Funktion abbildet, dann die Zuordnung der einzelnen Elemente.

BEISPIEL 2.1.1

Es sei

$$f : \begin{cases} [1, 3] & \rightarrow & \mathbb{R} \\ x & \mapsto & 2x \end{cases}$$

diejenige Funktion, die jeder Zahl aus dem Intervall $[1, 3]$ ihr Doppeltes zuordnet. Dann ist beispielsweise $f(\frac{4}{3}) = \frac{8}{3}$.

BEMERKUNG 2.1.2

Anstelle von f kann natürlich jedes beliebige Symbol benutzt werden. Neben der Schreibweise des Arguments in Klammern sind auch andere Notationen üblich:

- (i) Das Argument im Index: a_n für $a(n)$ (gebräuchlich für Folgen),
- (ii) Darstellung ohne Klammern: Fg für $F(g)$ (gebräuchlich für höhere Operatoren),
- (iii) Funktionssymbol hinter dem Argument: A' für das Komplement von A .

Wir werden in dieser Vorlesung hauptsächlich zwei Typen von Definitionsbereichen betrachten: In diesem Kapitel untersuchen wir ausführlich den Fall $X \subseteq \mathbb{N}$ und $Y = \mathbb{R}$ (meist $X = \mathbb{N}$ oder $X = \mathbb{N}_0$). Ist beispielsweise $X = \mathbb{N}$, so ist jeder natürlichen Zahl n eine reelle Zahl a_n zugeordnet. Damit ist eine Zahlenfolge (a_n) definiert, aufgefasst als Funktion $a : \mathbb{N} \rightarrow \mathbb{R}$.

Im anderen Fall ist X ein Intervall der Menge der reellen Zahlen, oder eine Vereinigung von solchen Intervallen. Wiederum ist $Y \subseteq \mathbb{R}$. In diesem Fall sprechen wir von einer Funktion in einer reellen Veränderlichen (oder in einer reellen Variablen).

Zur Veranschaulichung der Funktion einer reellen Veränderlichen ist der Graph dieser Funktion von Wichtigkeit. Wir geben die allgemeine Definition eines Graphen, beginnen aber mit

DEFINITION 2.1.2

Es seien $k \in \mathbb{N}$, M_1, \dots, M_k beliebige Mengen. Das kartesische Produkt der Mengen M_1, \dots, M_k ist definiert als

$$M_1 \times M_2 \times \dots \times M_k := \{(m_1, \dots, m_k) \mid m_j \in M_j\},$$

d. h. als die Menge aller k -Tupel (m_1, \dots, m_k) , deren j -te Komponente jeweils in M_j liegen. Ist $M_1 = M_2 = \dots = M_k$, so schreiben wir kurz M^k .

DEFINITION 2.1.3

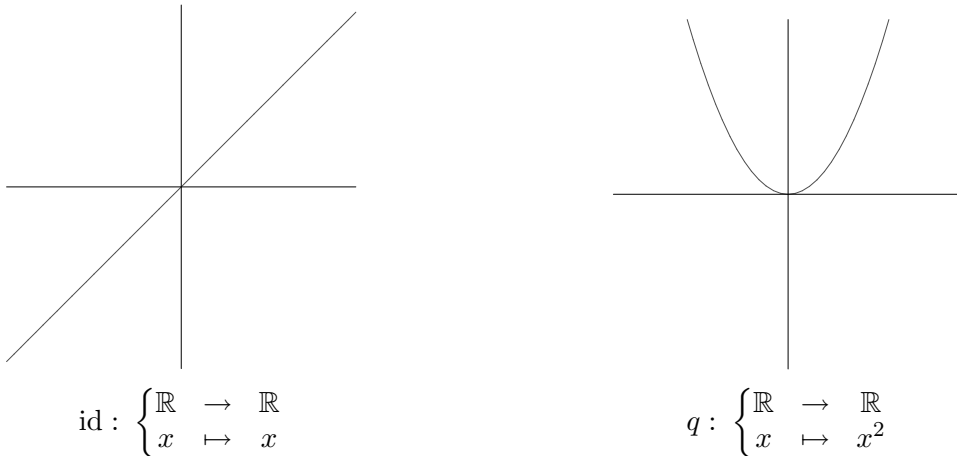
Es sei

$$f : \begin{cases} X & \rightarrow & Y \\ x & \mapsto & f(x) \end{cases}$$

eine Funktion von X nach Y . Der Graph von f ist die Menge $\{(x, f(x)) \mid x \in X\} \subseteq X \times Y$.

Es ist bekannt, dass durch Einführung eines kartesischen (d. h. rechtwinkligen) Koordinatensystems jedes $(x, y) \in \mathbb{R}^2$ mit einem Punkt der Zeichenebene identifiziert werden kann (wir werden dies nicht streng begründen). Sind $X \subseteq \mathbb{R}$ und $Y \subseteq \mathbb{R}$, so kann also der Graph einer Funktion durch eine Punktmenge der Ebene beschrieben werden.

BEISPIEL 2.1.2



Im Allgemeinen brauchen X und Y aber keine Zahlenmengen zu sein.

BEISPIEL 2.1.3

Ein Raum sei A Einheiten lang, B Einheiten breit, C Einheiten hoch. Durch die Einführung eines kartesischen Koordinatensystems, dessen Ursprung in der linken vorderen unteren Ecke des Raums liegt, kann jeder Punkt des Raums mit einem Tripel $(x, y, z) \in X = [0, A] \times [0, B] \times [0, C]$ identifiziert werden. Es sei T die Funktion, die jedem Punkt des Raums die dort herrschende Temperatur zuordnet. Dann kann T als Funktion von X nach \mathbb{R} betrachtet werden. T wird dann auch eine Funktion von mehreren Veränderlichen genannt, ein Hauptgegenstand der Analysis II.

DEFINITION 2.1.4

Die Funktionen $f_1 : X_1 \rightarrow Y_1$ und $f_2 : X_2 \rightarrow Y_2$ heißen gleich, wenn $X_1 = X_2$, $Y_1 = Y_2$ und $f_1(x) = f_2(x)$ für alle $x \in X_1 = X_2$ gilt.

BEISPIEL 2.1.4

Die Funktionen $f_R : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto x^2$ und $f_L : [-1, 2] \rightarrow \mathbb{R}$, $x \mapsto x^2$ sind nicht gleich, da die Definitionsmengen $[0, 1]$ und $[-1, 2]$ verschieden sind.

Eine Funktion $f : X \rightarrow Y$ ordnet nicht nur jedem Element $x \in X$ ein $y \in Y$ zu, sondern auch jeder Teilmenge A von X eine Teilmenge $f(A)$ von Y , sowie jeder Teilmenge B von Y eine Teilmenge $f^{-1}(B)$ von X :

DEFINITION 2.1.5

Es seien X und Y zwei nichtleere Mengen und $A \subseteq X$, $B \subseteq Y$, und $f : X \rightarrow Y$ eine Funktion von X nach Y . Das Bild $f(A)$ von A unter f ist die Menge $f(A) = \{f(x) \mid x \in A\}$. Das Urbild von B unter f ist die Menge $f^{-1}(B) = \{x \in X \mid f(x) \in B\}$. Ist $f(X) = Y$, so sagt man, f bildet X auf Y ab, oder f sei surjektiv. Haben verschiedene Urbilder immer verschiedene Bilder (folgt also aus $x_1 \neq x_2$ stets $f(x_1) \neq f(x_2)$), so nennt man f injektiv, oder auch umkehrbar eindeutig. In diesem (und nur in

diesem) Fall gibt es zu jedem $y \in f(X)$ genau ein Urbild $x \in X$ mit $f(x) = y$. Man kann dann die Umkehrabbildung (oder Umkehrfunktion) von f , bzw. die zu f inverse Funktion f^{-1} definieren durch

$$f^{-1} : \begin{cases} f(X) & \rightarrow X \\ f(x) & \mapsto x \end{cases} .$$

Offenbar ist $f^{-1}(f(x)) = x$ für alle $x \in X$ und $f(f^{-1}(y)) = y$ für alle $y \in f(X)$. Eine Funktion $f : X \rightarrow Y$, die sowohl injektiv als auch surjektiv ist, wird bijektiv genannt. Man sagt auch, f bilde X umkehrbar eindeutig nach Y ab, oder f sei eine Bijektion von X auf Y .

DEFINITION 2.1.6

Unter der Einschränkung $f|_A$ einer Funktion $f : X \rightarrow Y$ auf eine Menge $A \subseteq X$ versteht man diejenige Funktion, die jedem $x \in A$ den Wert $f(x)$ zuordnet:

$$f|_A : \begin{cases} A & \rightarrow Y \\ x & \mapsto f(x) \end{cases} .$$

Umgekehrt nennt man f die Fortsetzung von $f|_A$ auf X .

BEISPIEL 2.1.5

Es sei $X = [0, 1]$ und $Y_1 = [0, 3]$, sowie

$$f_1 : \begin{cases} X & \rightarrow Y_1 \\ x & \mapsto 2x \end{cases} .$$

Dann ist $f_1(X) = [0, 2] = Y_2$ (wegen $x \in [0, 1] \Leftrightarrow 2x \in [0, 2]$). Die Funktion f_1 ist injektiv (wegen $2x_1 = 2x_2 \Leftrightarrow x_1 = x_2$), aber nicht surjektiv, da $f_1(X) \neq Y_1$ ist. Die Funktion

$$f_2 : \begin{cases} X & \rightarrow Y_2 \\ x & \mapsto 2x \end{cases}$$

ist dagegen bijektiv. Die gemeinsame Umkehrfunktion $f_1^{-1} = f_2^{-1}$ von f_1 und f_2 ist gegeben durch

$$f_1^{-1} : \begin{cases} Y_2 & \rightarrow X \\ y & \mapsto \frac{1}{2}y \end{cases} .$$

BEISPIEL 2.1.6

Es sei $X_1 = [-1, 1]$ und $Y = [0, 1]$, sowie

$$g_1 : \begin{cases} X_1 & \rightarrow Y \\ x & \mapsto x^2 \end{cases} .$$

Aus einem späteren Teil der Vorlesung wird folgen, dass g_1 surjektiv ist. g_1 ist jedoch nicht injektiv, da $g_1(-x) = g_1(x)$ für alle $x \in [-1, 1]$ gilt. Hingegen ist die Einschränkung $g_2 = g_1|_{X_2}$ mit $X_2 = [0, 1]$ injektiv, da aus $0 \leq x_1 < x_2 \leq 1$ folgt: $0 \leq x_1^2 < x_2^2 \leq 1$ und damit $g_2(x_1) \neq g_2(x_2)$. Mit der Wurzelfunktion, die wir später definieren werden, können wir die Umkehrfunktion g_2^{-1} ausdrücken als

$$g_2^{-1} : \begin{cases} Y & \rightarrow X_2 \\ y & \mapsto \sqrt{y} \end{cases} .$$

Zum Schluss definieren wir noch den Begriff der Komposition (Hintereinanderausführung) zweier Funktionen:

DEFINITION 2.1.7

Sind zwei Funktionen $g : X \rightarrow Y_1$ und $f : Y_2 \rightarrow Z$ gegeben, und ist $g(X) \subseteq Y_2$, so bezeichnen wir die Funktion

$$f \circ g : \begin{cases} X & \rightarrow Z \\ x & \mapsto f(g(x)) \end{cases}$$

als Kompositum von f und g , die aus der inneren Funktion g und der äußeren Funktion f zusammengesetzt ist.

2.2. Grenzwerte einer Zahlenfolge

DEFINITION 2.2.1 (Betragfunktion)

Es sei $a \in \mathbb{R}$, dann definieren wir

$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{sonst} \end{cases} .$$

BEMERKUNG 2.2.1

Geometrisch gesehen gibt der Betrag einer reellen Zahl a den Abstand des der Zahl a entsprechenden Punkts vom Nullpunkt an.

SATZ 2.2.1

Es seien $a, b \in \mathbb{R}$. Der Betrag besitzt die folgenden Grundeigenschaften:

- (i) Definitheit: $|a| \geq 0$, wobei $|a| = 0$ nur für $a = 0$ gilt,
- (ii) Multiplikativität: $|ab| = |a| \cdot |b|$,
- (iii) Dreiecksungleichung: $|a + b| \leq |a| + |b|$,
- (iv) $|\frac{a}{b}| = \frac{|a|}{|b|}$ falls $b \neq 0$ sowie

$$\left| |a| - |b| \right| \leq \begin{cases} |a - b| \\ |a + b| \end{cases} .$$

BEWEIS

Alle Aussagen können durch Unterscheidung der vier Fälle

$$\alpha) a \geq 0, b \geq 0 \quad , \quad \beta) a \geq 0, b < 0 \quad , \quad \gamma) a < 0, b \geq 0 \quad , \quad \delta) a < 0, b < 0$$

bewiesen werden. Als Beispiel behandeln wir die Eigenschaft (iii) und überspringen die anderen Eigenschaften:

$$\alpha) \quad a \geq 0, b \geq 0 \quad \Rightarrow \quad a + b \geq 0 \quad \Rightarrow \quad |a + b| = a + b = |a| + |b| \quad ,$$

$$\beta) \quad a \geq 0, b < 0 \quad \Rightarrow \quad b \leq a + b = a - |b| \leq a \quad \Rightarrow \quad |a + b| \leq \max(|a|, |b|) \leq |a| + |b| \quad ,$$

$$\gamma) \quad \text{folgt aus } \beta) \text{ durch Umbenennung,}$$

$$\delta) \quad a < 0, b < 0 \quad \Rightarrow \quad a + b = -(|a| + |b|) \quad \Rightarrow \quad |a + b| \leq |a| + |b| \quad .$$

□

DEFINITION 2.2.2

Es sei $\varepsilon > 0$ und $x_0 \in \mathbb{R}$. Die ε -Umgebung $U_\varepsilon(x_0)$ von x_0 mit Radius ε definieren wir als $U_\varepsilon(x_0) := \{x \in \mathbb{R} \mid |x - x_0| < \varepsilon\}$.

DEFINITION 2.2.3

Es sei (a_n) eine Zahlenfolge und $a \in \mathbb{R}$. Die Zahlenfolge (a_n) konvergiert (oder strebt) gegen a , wenn es zu jeder positiven reellen Zahl $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ gibt, so dass für alle $n > n_0(\varepsilon)$ stets $|a - a_n| < \varepsilon$ ist. Dann heißt a der Grenzwert (oder Limes) der Folge (a_n) . Diese Eigenschaft drückt man folgendermaßen aus: „ $a_n \rightarrow a$ für $n \rightarrow \infty$ “, oder auch mit

$$a_n \rightarrow a \quad , \quad a_n \xrightarrow[n \rightarrow \infty]{} a \quad , \quad \lim_{n \rightarrow \infty} a_n = a .$$

Von einer Zahlenfolge, die gegen einen Grenzwert strebt, sagt man, sie konvergiere oder sie sei konvergent. Eine Zahlenfolge die nicht konvergiert, wird divergent genannt, sie divergiert.

BEMERKUNG 2.2.2

Oft ist es einfacher, anstelle der Ungleichung $|a_n - a| < \varepsilon$ für $n > n_0(\varepsilon)$ die Ungleichung $|a_n - a| < C\varepsilon$ für $n > n_0(\varepsilon)$ mit $C > 0$ zu zeigen. Dies reicht aus, da $\varepsilon > 0$ beliebig ist. Insbesondere kann statt ε auch $\varepsilon' = \frac{\varepsilon}{C}$ vorgeschrieben werden.

Die Konvergenzdefinition lässt sich mit der obigen der ε -Umgebung und der folgenden Definition anschaulicher und kürzer geben:

DEFINITION 2.2.4

Es sei (a_n) eine Zahlenfolge. Für jeden Index m heißt (a_m, a_{m+1}, \dots) ein Endstück der Folge (a_n) . Wir sagen, dass fast alle Glieder der Folge in einer Menge M liegen, wenn $a_n \in M$ für alle Indizes n bis auf höchstens endlich viele Ausnahmen gilt.

Dies ist offenbar gleichbedeutend mit: es gibt ein Endstück von (a_n) , das komplett in M liegt. Die obige Definition der Konvergenz können wir nun wie folgt fassen:

DEFINITION 2.2.5

Die Folge (a_n) konvergiert gegen a , wenn in jeder ε -Umgebung $U_\varepsilon(a)$ von a ein Endstück von (a_n) liegt, oder gleichbedeutend: wenn in jeder noch so kleinen Umgebung von a fast alle Folgenglieder von (a_n) enthalten sind.

BEISPIEL 2.2.1

Es gilt $\frac{1}{n} \rightarrow 0$ für $n \rightarrow \infty$: es sei $\varepsilon > 0$ beliebig vorgegeben, dann gilt für $n > \frac{1}{\varepsilon}$ stets $|\frac{1}{n} - 0| < \varepsilon$, woraus die Konvergenz gegen Null folgt.

BEISPIEL 2.2.2

Die Folge $a_n = (-1)^n$ ist divergent. Wir nehmen das Gegenteil an: es sei $a_n \rightarrow a$ für $n \rightarrow \infty$ und ein $a \in \mathbb{R}$. Wir betrachten die Umgebung $U_{\frac{1}{2}}(a)$. Offenbar ist $-1 \notin U_{\frac{1}{2}}(a)$ oder $1 \notin U_{\frac{1}{2}}(a)$. Da $a_n = -1$ für alle ungeraden n ist, sowie $a_n = 1$ für alle geraden n , gibt es unendlich viele n mit $a_n \notin U_{\frac{1}{2}}(a)$, im Widerspruch zur Konvergenzdefinition.

DEFINITION 2.2.6

Ist (a_n) eine Zahlenfolge und (n_j) eine Folge natürlicher Zahlen mit $n_j < n_{j+1}$ für alle j , so heißt die durch j indizierte Folge (a_{n_j}) eine Teilfolge von (a_n) .

SATZ 2.2.2

Jede Teilfolge einer konvergenten Folge ist konvergent und konvergiert gegen den gleichen Grenzwert.

BEWEIS

Es sei (a_n) eine Folge mit $a_n \rightarrow a$ für $n \rightarrow \infty$, und (a_{n_j}) eine durch j indizierte Teilfolge. Nun sei U eine beliebige ε -Umgebung von a . Wegen der Konvergenz von (a_n) gibt es ein N , so dass für alle $n > N = N(U)$ das Folgenglied a_n in U liegt. Es gibt ein J , so dass $n_J > N$ ist, da $n_j \geq j$ für alle j gilt, wegen $n_{j+1} \geq n_j + 1$. Also ist für alle $j > J$ auch $n_j > N$, d. h. das mit dem J -ten Folgenglied beginnende Endstück der Teilfolge (a_{n_j}) liegt komplett in U . Da die ε -Umgebung U von a beliebig war folgt, dass (a_{n_j}) gegen a strebt für $j \rightarrow \infty$. \square

DEFINITION 2.2.7

Eine Folge (a_n) heißt beschränkt, falls die Menge $\{a_n \mid n \in \mathbb{N}\}$ beschränkt ist. Im gleichen Sinne spricht man vom Supremum oder Infimum einer Folge, falls es existiert.

SATZ 2.2.3

Jede konvergente Folge ist beschränkt.

BEWEIS

Es sei a der Grenzwert von (a_n) . Nach Definition der Konvergenz gibt es ein n_0 , so dass $|a_n - a| < 1$ ist für alle $n \geq n_0$:

$$\begin{aligned} \Rightarrow |a_n| &\leq |a| + 1 \quad \text{falls } n \geq n_0 \\ \Rightarrow |a_n| &\leq \max(|a_1|, |a_2|, \dots, |a_{n_0}|, |a| + 1) \quad . \end{aligned}$$

□

2.3. Grenzwerte von Summen, Produkten, Quotienten

DEFINITION 2.3.1

Eine Folge (a_n) mit Grenzwert 0 heißt Nullfolge.

SATZ 2.3.1

Gilt mit einer Nullfolge (α_n) fast immer $|a_n - a| \leq \alpha_n$, so ist $a_n \rightarrow a$ gegen 0 für $n \rightarrow \infty$.

BEWEIS

Es sei $\varepsilon > 0$ gegeben. Dann gibt es $n_0 = n_0(\varepsilon)$, so dass für $n > n_0$ gilt: $\alpha_n < \varepsilon$. Dann gibt es höchstens endlich viele $n > n_0$ mit $|a_n - a| \geq \varepsilon$, also gilt fast immer $|a_n - a| < \varepsilon$, d. h. (a_n) konvergiert gegen a . □

SATZ 2.3.2

Strebt (a_n) gegen Null, und ist (b_n) beschränkt, so strebt auch die Folge der Produkte $(a_n b_n)$ gegen Null.

BEWEIS

Es sei $|b_n| < \beta$ für alle n und ein $\beta > 0$, und $\varepsilon > 0$ beliebig. Bestimme nun $n_0 = n_0(\varepsilon)$, so dass $|a_n| < \varepsilon$ für $n > n_0$ ist. Dann ist $|a_n b_n| < \beta \cdot \varepsilon$ für $n > n_0$. □

Aus Satz 2.2.1(iv) folgt

SATZ 2.3.3 (Betragssatz)

Strebt (a_n) gegen a , so strebt $(|a_n|)$ gegen $|a|$.

SATZ 2.3.4

Aus $a_n \rightarrow a$ und $b_n \rightarrow b$ für $n \rightarrow \infty$ folgt:

- (i) $a_n + b_n \rightarrow a + b$ (Summensatz),
- (ii) $a_n - b_n \rightarrow a - b$,
- (iii) $a_n \cdot b_n \rightarrow a \cdot b$ (Produktsatz),
- (iv) $\alpha \cdot a_n \rightarrow \alpha \cdot a$ (für jede Konstante $\alpha \in \mathbb{R}$),
- (v) ist überdies $b \neq 0$, so sind auch fast alle $b_n \neq 0$ und es gilt

$$\frac{a_n}{b_n} \rightarrow \frac{a}{b} \quad (\text{Quotientensatz}).$$

BEWEIS

Zu (i): Es sei $\varepsilon > 0$. Nach der Definition der Konvergenz gibt es Indizes n_0 und n_1 , so dass $|a_n - a| < \varepsilon$ für $n > n_0$ sowie $|b_n - b| < \varepsilon$ für $n > n_1$ gilt. Für alle $n > \max(n_0, n_1)$ gilt dann

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

Also $(a_n + b_n) \rightarrow a + b$ für $n \rightarrow \infty$. (ii) wird ebenso bewiesen. Zu (iii): Es ist $a_n b_n - ab = (a_n - a)b_n + a(b_n - b)$. Da $(a_n - a)$ und $(b_n - b)$ Nullfolgen sind, und (b_n) nach Satz 2.2.3 beschränkt ist, folgt mit (i), dass auch $(a_n b_n - ab)$ eine Nullfolge ist, also $a_n b_n \rightarrow ab$. Der Teil (iv) ist ein Spezialfall von (iii) für die konstante Folge $b_n = \alpha$. Zu (v): Es sei $b \neq 0$, dann ist nach Satz 2.3.3 $|b_n| \rightarrow |b| > 0$. Es gibt also zu $\varepsilon := \frac{1}{2}|b|$ ein n_0 , so dass $|b_n| > |b| - \varepsilon = \frac{1}{2}|b|$ ist für alle $n > n_0$. Für diese n ist somit $b_n \neq 0$ und

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \left| \frac{b - b_n}{b_n \cdot b} \right| \leq \frac{2}{|b|^2} \cdot |b - b_n|.$$

Nach (iii) strebt die rechte Seite gegen Null, womit aus Satz 2.2.3 die Konvergenz $\frac{1}{b_n} \rightarrow \frac{1}{b}$ für $n \rightarrow \infty$ folgt. Die Aussage (v) folgt somit aus (iii). □

DEFINITION 2.3.2

Es seien $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in \mathbb{R}$ mit $a_n \neq 0$ beliebig. Eine Funktion der Form

$$P: \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto P(x) := a_0 + a_1x + a_2x^2 + \dots + a_nx^n \end{cases}$$

heißt Polynom n -ten Grades. $\xi \in \mathbb{R}$ heißt Nullstelle von P , falls $P(\xi) = 0$ ist. Das Nullpolynom ist $P(x) = 0$ und besitzt keinen Grad.

DEFINITION 2.3.3

Es seien P und Q Polynome. N sei die Menge der Nullstellen von Q . Eine Funktion der Form

$$R: \begin{cases} \mathbb{R} \setminus N & \rightarrow \mathbb{R} \\ x & \mapsto \frac{P(x)}{Q(x)} \end{cases}$$

heißt rationale Funktion.

SATZ 2.3.5

Es seien P, Q Polynome, und $R(x) = \frac{P(x)}{Q(x)}$ eine rationale Funktion, sowie x_n eine Folge mit $x_n \rightarrow \xi$ für $n \rightarrow \infty$. Dann gilt:

$$\lim_{n \rightarrow \infty} P(x_n) = P(\xi) \quad \text{und} \quad \lim_{n \rightarrow \infty} R(x_n) = R(\xi) \quad \text{falls} \quad Q(\xi) \neq 0.$$

BEWEIS

Zu i): Der Beweis ergibt sich durch vollständige Induktion nach dem Grad von P aus Satz 2.3.4. Zu ii): Dies folgt aus i) und Satz 2.3.4(v). □

2.4. Wichtige Konvergenzkriterien

Um die Frage nach Konvergenz oder Divergenz einer Folge entscheiden zu können, ist es wichtig, über einfache Kriterien zu verfügen. Wir besprechen im folgenden einige Konvergenzkriterien. Diese erlauben es, die Konvergenz gewisser Folgen nachzuweisen, im Allgemeinen jedoch nicht, den Grenzwert zu berechnen.

DEFINITION 2.4.1

Eine Folge (a_n) heißt monoton wachsend, falls $a_n \leq a_{n+1}$ für alle n gilt, bzw. monoton fallend, falls $a_n \geq a_{n+1}$ für alle n gilt. Die Folge heißt monoton, falls sie monoton wachsend oder monoton fallend ist.

SATZ 2.4.1 (Monotonieprinzip)

Eine monotone Folge konvergiert genau dann, wenn sie beschränkt ist. In diesem Fall strebt sie gegen ihr Supremum, falls sie wächst, oder gegen ihr Infimum, falls sie fällt.

BEWEIS

Wir nehmen zunächst an, (a_n) sei wachsend und beschränkt mit $a = \sup(\{a_n \mid n \in \mathbb{N}\})$. Es sei $\varepsilon > 0$ gegeben. Dann gibt es ein $n_0 = n_0(\varepsilon)$ mit $a_{n_0} > a - \varepsilon$. Für alle $n > n_0$ ist dann $a_n \in U_\varepsilon(a)$, also $a_n \rightarrow a$ nach der Definition der Konvergenz. Die Behauptung für monoton fallende Folgen wird ähnlich bewiesen. Die Rückrichtung ist gerade Satz 2.2.3. □

SATZ 2.4.2 (Auswahlsatz von Bolzano-Weierstraß)

Jede beschränkte Folge enthält eine konvergente Teilfolge.

BEWEIS

Wir zeigen zunächst, dass jede Folge (a_n) eine monotone Teilfolge enthält. Wir nennen m eine Gipfelstelle von (a_n) , wenn für $n > m$ stets $a_n < a_m$ gilt. Besitzt (a_n) unendlich viele Gipfelstellen $m_1 < m_2 < \dots$, so ist $a_{m_1} > a_{m_2} > \dots$, d. h. (a_{m_n}) ist eine monotone Teilfolge, die nach Satz 2.4.1 konvergent ist. Gibt es nur endlich viele Gipfelstellen, so gibt es einen Index n_1 , der größer als alle Gipfelstellen und somit keine Gipfelstelle ist. Dann gibt es $n_2 > n_1$ mit $a_{n_2} \geq a_{n_1}$. Da auch n_2 keine

Gipfelstelle ist, gibt es ein $n_3 > n_2$ mit $a_{n_3} \geq a_{n_2}$, usw. (formaler Beweis durch vollständige Induktion). Es gibt also eine monoton wachsende Teilfolge (a_{n_k}) , die wiederum nach Satz 2.4.1 konvergent ist. \square

DEFINITION 2.4.2

Eine Folge (a_n) heißt Cauchyfolge, wenn es zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ gibt, so dass für alle $m, n > n_0$ stets $|a_m - a_n| < \varepsilon$ gilt.

SATZ 2.4.3 (Cauchy Kriterium)

Für eine Zahlenfolge (a_n) gilt: (a_n) konvergent $\Leftrightarrow (a_n)$ Cauchyfolge.

BEWEIS

\Rightarrow :

Es sei $a_n \rightarrow a$ konvergent, und $\varepsilon > 0$ gegeben. Nach der Definition der Konvergenz gibt es $n_0 = n_0(\frac{1}{2}\varepsilon)$, so dass $|a_n - a| < \frac{1}{2}\varepsilon$ für $n > n_0$ ist. Für $m, n > n_0$ gilt dann:

$$|a_m - a_n| \leq |a_m - a| + |a - a_n| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

Also ist (a_n) eine Cauchyfolge.

\Leftarrow :

Nun sei (a_n) eine Cauchyfolge. Wir zeigen zunächst, dass (a_n) beschränkt ist: zu $\varepsilon = 1$ gibt es ein n_0 , so dass für alle $m, n > n_0$ gilt: $|a_n - a_m| < 1 \Rightarrow |a_m| < |a_{n_0+1}| + 1$ für jedes $m > n_0$, und damit $|a_n| < \max(|a_1|, \dots, |a_{n_0}|, 1 + |a_{n_0+1}|)$. Also ist (a_n) beschränkt. Nach Satz 2.4.2 (Auswahlprinzip von Bolzano-Weierstraß) gibt es eine Teilfolge (a_{n_k}) , die gegen einen Grenzwert $a \in \mathbb{R}$ konvergiert. Wir zeigen, dass die Gesamtfolge gegen a konvergiert. Dazu sei $\varepsilon > 0$ gegeben. Wir bestimmen n_0 mit $|a_m - a_n| < \frac{1}{2}\varepsilon$ für alle $m, n > n_0$ sowie ein $N > n_0$, so dass $|a_N - a| < \frac{1}{2}\varepsilon$ ist. So ein N existiert, da (a_{n_k}) gegen a strebt. Für alle $n > n_0$ ist dann $|a_n - a| \leq |a_n - a_N| + |a_N - a| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$, also $a_n \rightarrow a$. \square

DEFINITION 2.4.3

Eine Folge abgeschlossener Intervalle $I_n = [a_n, b_n]$ heißt Intervallschachtelung, wenn $I_1 \supset I_2 \supset I_3 \supset \dots$ und $\lim(b_n - a_n) = 0$ gilt.

SATZ 2.4.4

Es sei $I_n = [a_n, b_n]$ eine Intervallschachtelung. Dann gibt es genau ein $a \in \mathbb{R}$, das allen I_n angehört.

BEWEIS

Die Folgen (a_n) und (b_n) sind offenbar Cauchyfolgen, und nach Satz 2.4.3 konvergent, etwa $a_n \rightarrow a$ und $b_n \rightarrow b$. Es ist $a_n \leq a$ und $b_n \geq b$ für alle n . Wegen $b_n - a_n \rightarrow 0$ folgt $a = b$, und diese Zahl liegt in allen I_n . \square

2.5. Häufungswerte, Limes Superior, Limes Inferior

Der Begriff des Häufungswerts stellt eine Abschwächung des Begriffs des Grenzwerts dar. Divergente Zahlenfolgen, also Folgen, die keinen Grenzwert besitzen, können immer noch einen Häufungswert besitzen.

DEFINITION 2.5.1

Eine Zahl α heißt Häufungswert der Folge (a_n) , wenn in jeder ε -Umgebung von α unendlich viele Folgenglieder liegen, wenn es also zu jedem $\varepsilon > 0$ unendlich viele Indizes n gibt mit $|\alpha - a_n| < \varepsilon$.

SATZ 2.5.1

Eine Zahl α ist genau dann Häufungswert von (a_n) , wenn sie der Grenzwert einer Teilfolge (a_{n_k}) ist.

BEWEIS

Es ist möglich, Indizes $n_1 < n_2 < \dots$ derart zu bestimmen, dass $a_{n_k} \in U_{\frac{1}{k}}(\alpha)$ gilt für alle k . Die Teilfolge (a_{n_k}) strebt dann gegen α . \square

BEISPIEL 2.5.1

Die Folge $a_n = (-1)^n + \frac{1}{n}$ besitzt die Häufungswerte -1 und 1 , aber keinen Grenzwert. Es gilt $(a_{2n}) \rightarrow 1$ und $(a_{2n+1}) \rightarrow -1$.

SATZ 2.5.2

Jede beschränkte Folge besitzt mindestens einen Häufungswert.

BEWEIS

Nach Satz 2.4.2 (Auswahlprinzip) hat (a_n) eine konvergente Teilfolge. Deren Grenzwert ist nach Satz 2.5.1 ein Häufungswert von (a_n) . \square

SATZ 2.5.3 (und Definition)

Jede beschränkte Folge (a_n) besitzt einen größten und einen kleinsten Häufungswert. Der erstere wird Limes superior genannt, der zweite Limes inferior. Schreibweisen:

$$\limsup_{n \rightarrow \infty} a_n, \quad \liminf_{n \rightarrow \infty} a_n.$$

Es sind auch die Symbole $\overline{\lim} a_n$ und $\underline{\lim} a_n$ gebräuchlich. Für jede konvergente Teilfolge (a'_n) von (a_n) gilt

$$\liminf_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} a'_n \leq \limsup_{n \rightarrow \infty} a_n,$$

ferner gibt es stets Teilfolgen von (a_n) , die jeweils gegen $\limsup a_n$ und $\liminf a_n$ konvergieren.

BEWEIS

Es sei \mathcal{H} die Menge der Häufungswerte von (a_n) , sowie $U \in \mathbb{R}$ eine untere Schranke und $V \in \mathbb{R}$ eine obere Schranke der Menge $\{a_n \mid n \in \mathbb{N}\}$. Dann sind U und V auch Schranken der Menge \mathcal{H} , also ist \mathcal{H} beschränkt und besitzt daher nach dem Supremumsprinzip (S1) ein Supremum S und ein Infimum I . Wir zeigen, dass auch $S \in \mathcal{H}$ gilt, und damit S der größte Häufungswert ist, und nennen S den Limes superior von (a_n) . Es sei $\varepsilon > 0$ gegeben. Dann gibt es wegen der Supremumseigenschaft von S einen Häufungswert $\alpha \in \mathcal{H}$ von (a_n) mit $\alpha \in (S - \frac{1}{2}\varepsilon, S)$. Nach der Definition des Häufungswerts gibt es unendlich viele n , so dass $|a_n - \alpha| < \frac{1}{2}\varepsilon$ ist. Für diese n folgt dann

$$|a_n - S| \leq |a_n - \alpha| + |\alpha - S| < \varepsilon.$$

Somit ist auch S ein Häufungswert von (a_n) . Analog zeigen wir, dass $I = \liminf a_n$ ein Häufungswert ist. Der Rest des Beweises folgt aus Satz 2.5.1. \square

SATZ 2.5.4

Die Zahl $\alpha \in \mathbb{R}$ ist genau dann der Limes inferior der beschränkten Folge (a_n) , wenn für jedes $\varepsilon > 0$ die Ungleichung $a_n < \alpha + \varepsilon$ für unendlich viele n , die Ungleichung $a_n < \alpha - \varepsilon$ aber nur für endlich viele n erfüllt ist. Entsprechend ist α' genau dann der Limes superior von (a_n) , wenn für jedes $\varepsilon > 0$ die Ungleichung $a_n > \alpha' - \varepsilon$ für unendlich viele n , die Ungleichung $a_n > \alpha' + \varepsilon$ aber nur für endlich viele n erfüllt ist.

BEWEIS

Wir beweisen nur die Aussage für den Limes superior.

\Rightarrow :

Da α' ein Häufungswert von (a_n) ist, gilt: für unendlich viele n ist $a_n \in (\alpha' - \varepsilon, \alpha' + \varepsilon)$, insbesondere $a_n > \alpha' - \varepsilon$. Annahme: Es gibt unendlich viele n mit $a_n > \alpha' + \varepsilon$. Wir betrachten die Teilfolge (a_{n_k}) von (a_n) mit $a_{n_k} > \alpha' + \varepsilon$. Da auch die Folge (a_{n_k}) beschränkt ist, besitzt sie einen Häufungswert $\beta \geq \alpha' + \varepsilon$, im Widerspruch dazu, dass α' der größte Häufungswert von (a_n) ist.

⇐:

Es sei $\alpha' \in \mathbb{R}$ beliebig, so dass für jedes $\varepsilon > 0$ die Ungleichung $a_n > \alpha' - \varepsilon$ für unendlich viele n , die Ungleichung $a_n > \alpha' + \varepsilon$ aber nur für endlich viele n erfüllt ist. Es liegen also unendlich viele Folgenglieder im Intervall $(\alpha' - \varepsilon, \infty)$, aber nur endlich viele in $(\alpha' + \varepsilon, \infty)$. Also liegen unendlich viele Folgenglieder in $(\alpha' - \varepsilon, \alpha' + \varepsilon]$, d. h. es liegen für jedes $\varepsilon > 0$ unendlich viele Folgenglieder in $U_\varepsilon(\alpha')$. Damit ist α' ein Häufungswert der Folge (a_n) . Es bleibt zu zeigen, dass α' der größtmögliche Häufungswert ist. Sei nun $\alpha > \alpha'$ beliebig und $\delta = \frac{1}{2}(\alpha - \alpha') > 0$. Dann liegt $U_\delta(\alpha)$ vollständig im Intervall $(\alpha' + \delta, \infty)$. Nach Voraussetzung liegen darin nur endlich viele Folgenglieder, d. h. α ist kein Häufungswert von (a_n) . Damit ist α' der größte Häufungswert, nach Satz 2.5.3 also $\alpha' = \limsup a_n$. \square

SATZ 2.5.5

Eine Folge (a_n) konvergiert genau dann, wenn sie beschränkt ist, und nur einen Häufungswert besitzt. In diesem Fall ist $\lim a_n = \limsup a_n = \liminf a_n$.

BEWEIS

Dies folgt aus den Sätzen 2.5.1 und 2.5.3. \square

2.6. Uneigentliche Grenzwerte

DEFINITION 2.6.1

Wir sagen, die Folge (a_n) divergiert gegen $+\infty$ (bzw. gegen $-\infty$), in Zeichen $a_n \rightarrow \infty$ ($a_n \rightarrow -\infty$), wenn es zu jedem $C > 0$ ein n_0 gibt, so dass für alle $n > n_0$ stets $a_n > C$ (bzw. $a_n < -C$) gilt.

Für $a_n \rightarrow \infty$ sagt man, $+\infty$ sei der uneigentliche Grenzwert von (a_n) , und schreibt auch $\lim a_n = \infty$ (entsprechend für $-\infty$).

SATZ 2.6.1

Es sei $a_n \rightarrow \infty$, $b_n \rightarrow \infty$ und $c_n \rightarrow c$. Dann gilt:

- (i) $a_n + b_n \rightarrow \infty$,
- (ii) $a_n + c_n \rightarrow \infty$,
- (iii) $\alpha a_n \rightarrow \infty$ für $\alpha > 0$ bzw. $\alpha a_n \rightarrow -\infty$ für $\alpha < 0$,
- (iv) $a_n b_n \rightarrow \infty$,
- (v) $\frac{\alpha}{a_n} \rightarrow 0$.

BEWEIS

(Übungsaufgabe) \square

3. Unendliche Reihen

3.1. Grundbegriffe

DEFINITION 3.1.1

Es sei (a_n) eine Zahlenfolge. Wir setzen $S_n := a_0 + a_1 + \dots + a_n$ für $n = 0, 1, 2, \dots$. Unter der unendlichen Reihe (oder kurz Reihe)

$$\sum_{k=0}^{\infty} a_k$$

versteht man die Folge (S_n) . Die S_n heißen Teil- oder Partialsummen der unendlichen Reihe.

Die Indizierung braucht nicht immer mit $k = 0$ zu beginnen. So versteht man unter

$$\sum_{k=r}^{\infty} a_k$$

die Folge der Partialsummen $S_n = a_r + a_{r+1} + \dots + a_n$. Auch auf die Bezeichnung des Summationsindex kommt es nicht an.

DEFINITION 3.1.2

Die Reihe

$$\sum_{k=0}^{\infty} a_k$$

heißt konvergent, wenn die Folge der Partialsummen (S_n) konvergiert. Strebt $S_n \rightarrow S$, so sagt man, die Reihe konvergiere gegen S , geschrieben

$$\sum_{k=0}^{\infty} a_k = S,$$

und nennt S den Reihenwert oder die Summe der Reihe. Eine nicht konvergente Reihe wird divergent genannt. Man sagt auch, die Reihe existiere bzw. existiere nicht, um auszudrücken, dass sie konvergiert oder divergiert.

BEISPIEL 3.1.1

Es gilt

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

Beweis: wir zeigen zunächst durch vollständige Induktion

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

Induktionsanfang $n = 1$: $S_1 = \frac{1}{2} = 1 - \frac{1}{1+1}$. Induktionsschritt $n \rightarrow n+1$:

$$S_{n+1} = S_n + \frac{1}{(n+1)(n+2)} = 1 - \frac{1}{n+1} + \left(\frac{1}{n+1} - \frac{1}{n+2} \right) = 1 - \frac{1}{n+2}.$$

Nun gilt

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n \rightarrow \infty} S_n = 1.$$

Eine wichtige Ungleichung im Zusammenhang mit Reihen ist die

SATZ 3.1.1 (Bernoulli-Ungleichung)

Für alle $x \in \mathbb{R}$ mit $x > -1$ und $x \neq 0$ sowie alle $n \in \mathbb{N}$ gilt $(1+x)^n \geq 1+nx$, Gleichheit gilt nur für $n = 1$.

BEWEIS
(Übungsaufgabe)

□

3.2. Konvergenzkriterien

SATZ 3.2.1 (Cauchy Kriterium, Formulierung für Reihen)

Die Reihe

$$\sum_{k=0}^{\infty} a_k$$

konvergiert genau dann, wenn sie eine Cauchyfolge ist, d. h. wenn die Partialsummen eine Cauchyfolge bilden, wenn also zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ existiert, so dass für alle $n > n_0$ und alle natürlichen Zahlen p gilt: $|a_{n+1} + a_{n+2} + \dots + a_{n+p}| < \varepsilon$.

SATZ 3.2.2 (Monotoniekriterium, Formulierung für Reihen)

Eine Reihe mit nicht negativen Gliedern konvergiert genau dann, wenn die Folge ihrer Teilsummen beschränkt ist.

Offenbar wird die Eigenschaft, eine Cauchyreihe zu sein, durch Abänderung von endlich vielen Gliedern nicht beeinflusst. Bei Konvergenzuntersuchungen kommt es daher nicht auf den Index an, mit dem die Summation beginnt. Wir dürfen deshalb eine vorgelegte Reihe kurz mit $\sum a_k$ bezeichnen. Der Wert einer unendlichen Reihe wird im Falle seiner Existenz jedoch durch Änderung von endlich vielen Gliedern sehr wohl verändert. Es gilt

$$\sum_{k=p+1}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - (a_0 + a_1 + \dots + a_p).$$

SATZ 3.2.3

Bei einer konvergenten Reihe $\sum a_k$ bilden sowohl die Glieder a_k als auch die Reste

$$r_n = \sum_{k=n+1}^{\infty} a_k$$

stets eine Nullfolge.

BEWEIS

Das folgt aus dem Cauchy Kriterium (Satz 3.2.1).

□

BEMERKUNG 3.2.1

Die Bedingung $a_n \rightarrow 0$ ist also eine notwendige Bedingung für die Konvergenz von $\sum a_n$. Dass die Bedingung nicht hinreichend ist, zeigt der folgende Satz.

SATZ 3.2.4

Die harmonische Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

divergiert.

BEWEIS

Wir betrachten die Teilsummen

$$S_{2^n} = \sum_{k=1}^{2^n} \frac{1}{k}.$$

Wir zeigen durch vollständige Induktion $S_{2^n} \geq \frac{1}{2}n$. Induktionsanfang: $S_{2^1} = S_2 = 1 + \frac{1}{2} > \frac{1}{2}$. Induktionsschritt:

$$S_{2^{n+1}} = S_{2^n} + \frac{1}{2^n+1} + \frac{1}{2^n+2} + \cdots + \frac{1}{2^{n+1}} = S_{2^n} + \frac{1}{2^n+1} + \frac{1}{2^n+2} + \cdots + \frac{1}{2^n+(2^{n+1}-2^n)}.$$

$S_{2^{n+1}}$ lässt sich also als Summe aus S_{2^n} und $2^{n+1} - 2^n$ Summanden s_j schreiben, die jeweils nicht kleiner als $\frac{1}{2^{n+1}}$ sind. Also gilt

$$S_{2^{n+1}} = S_{2^n} + \sum_{j=1}^{2^{n+1}-2^n} s_j \geq S_{2^n} + (2^{n+1} - 2^n) \cdot \frac{1}{2^{n+1}} = S_{2^n} + (1 - \frac{1}{2}) \geq \frac{1}{2}n + \frac{1}{2} = \frac{1}{2}(n+1).$$

Damit divergiert $S_n \rightarrow \infty$, denn die 2^n -ten Partialsummen werden beliebig groß, und es gilt $S_1 < S_2 < S_3 < \cdots$ wegen $\frac{1}{n} > 0$. \square

SATZ 3.2.5 (Rechenregeln für konvergente Reihen)

Sind die Reihen

$$\sum_{k=0}^{\infty} a_k, \quad \sum_{k=0}^{\infty} b_k$$

jeweils konvergent, so ist

$$\sum_{k=0}^{\infty} (a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k, \quad \sum_{k=0}^{\infty} (a_k - b_k) = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{\infty} b_k, \quad \sum_{k=0}^{\infty} (\alpha a_k) = \alpha \sum_{k=0}^{\infty} a_k.$$

BEWEIS

Dieser Satz folgt durch Anwendung von Satz 2.3.4 (Rechenregeln für Folgen) auf die Folge der Partialsummen. \square

BEISPIEL 3.2.1

Die unendliche Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

ist konvergent: Aus Beispiel 3.1.1 folgt durch Indexverschiebung

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

Die Partialsummen können nun wie folgt abgeschätzt werden:

$$S_n = \sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \frac{1}{k(k-1)} \xrightarrow{n \rightarrow \infty} 2.$$

Also ist die Folge (S_n) einerseits streng monoton wachsend, andererseits durch 2 beschränkt. Damit ist $\sum \frac{1}{k^2}$ konvergent.

DEFINITION 3.2.1

Eine Reihe $\sum a_k$ heißt absolut konvergent, wenn die Reihe $\sum |a_k|$ konvergiert.

SATZ 3.2.6

Eine absolut konvergente Reihe ist auch gewöhnlich konvergent.

BEWEIS

Ist $\sum |a_k|$ konvergent, so ist $\sum |a_k|$ eine Cauchyreihe (Satz 3.2.1). Es sei $\varepsilon > 0$ gegeben. Dann gibt es $n_0(\varepsilon)$, so dass für $n > n_0$ und jedes $p \in \mathbb{N}$ gilt: $|a_{n+1}| + \cdots + |a_{n+p}| < \varepsilon$. Daraus folgt mit der Dreiecksungleichung

$$|a_{n+1} + \cdots + a_{n+p}| \leq |a_{n+1}| + \cdots + |a_{n+p}| < \varepsilon.$$

Damit ist auch $\sum a_k$ eine Cauchyreihe und nach Satz 3.2.1 konvergent. \square

Beispiel 3.2.1 gibt Anlass zur folgenden Definition:

DEFINITION 3.2.2

Die Reihe $\sum b_k$ heißt (konvergente) Majorante der Reihe $\sum a_k$, wenn $|a_k| \leq b_k$ für alle k gilt, und $\sum b_k$ konvergiert.

SATZ 3.2.7 (Majorantenkriterium)

Ist $\sum b_k$ eine konvergente Majorante von $\sum a_k$, so ist $\sum a_k$ konvergent.

BEWEIS

Es sei $\varepsilon > 0$ gegeben: $\sum b_k$ konvergiert $\Rightarrow \sum b_k$ ist eine Cauchyreihe $\Rightarrow \exists n_0 = n_0(\varepsilon)$, so dass für $n > n_0$ und $p \in \mathbb{N}$ gilt:

$$|b_{n+1} + \dots + b_{n+p}| < \varepsilon \Rightarrow |a_{n+1} + \dots + a_{n+p}| \underset{\text{DrÜngl.}}{\leq} |a_{n+1}| + \dots + |a_{n+p}| \leq |b_{n+1} + \dots + b_{n+p}| < \varepsilon.$$

Also ist auch $\sum a_k$ eine Cauchyreihe und damit konvergent. \square

BEMERKUNG 3.2.2

Ist $\sum b_k$ eine (konvergente) Majorante der Reihe $\sum a_k$, so ist sie nach Definition auch eine Majorante von $\sum |a_k|$. Nach Satz 3.2.7 ist daher $\sum a_k$ sogar absolut konvergent.

DEFINITION 3.2.3

Die Reihe $\sum c_k$ heißt (divergente) Minorante von $\sum d_k$, falls $0 \leq c_k \leq d_k$ für alle k gilt, und $\sum c_k$ divergiert.

SATZ 3.2.8 (Minorantenkriterium)

Ist $\sum c_k$ eine divergente Minorante von $\sum d_k$, so ist $\sum d_k$ divergent.

BEWEIS

Annahme: $\sum d_k$ konvergiert. Dann ist $\sum d_k$ eine konvergente Majorante von $\sum c_k$ und damit nach Satz 3.2.7 auch $\sum c_k$ konvergent, ein Widerspruch. \square

DEFINITION 3.2.4

Es sei $q \in \mathbb{R}$. Die Reihe

$$\sum_{k=0}^{\infty} q^k$$

heißt (unendliche) geometrische Reihe mit Quotient q .

SATZ 3.2.9

Die unendliche geometrische Reihe ist konvergent genau für $|q| < 1$, in diesem Fall ist

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

BEWEIS

Es sei zunächst $q \neq 1$ und S_n die n -te Partialsumme $S_n = q^0 + q^1 + \dots + q^n$. Dann gilt

$$q \cdot S_n = \sum_{k=1}^{n+1} q^k, \text{ Subtraktion ergibt } S_n = \frac{1 - q^{n+1}}{1 - q}.$$

Es ist S_n konvergent $\Leftrightarrow (q^{n+1})$ konvergiert $\Leftrightarrow |q| < 1$. In diesem Fall strebt $q^{n+1} \rightarrow 0$ für $n \rightarrow \infty$ und somit $S_n \rightarrow \frac{1}{1-q}$. Ist $q = 1$, so ist $S_n = 1 + \dots + 1 = n$ und somit $\sum q^k$ divergent. \square

Majoranten- bzw. Minorantenkriterium führen in Verbindung mit der geometrischen Reihe zu einem wichtigen Konvergenzkriterium:

SATZ 3.2.10 (Quotientenkriterium)

Es sei (a_k) eine Zahlenfolge mit $a_k \neq 0$ für $k > k_0$. Dann gilt:

- (i) Ist $\limsup \left| \frac{a_{k+1}}{a_k} \right| < 1$, so ist $\sum a_k$ absolut konvergent.
(ii) Ist $\liminf \left| \frac{a_{k+1}}{a_k} \right| > 1$, so ist $\sum a_k$ divergent.

BEWEIS

Zu i): Es sei $\limsup \left| \frac{a_{k+1}}{a_k} \right| = 1 - 2\delta_0 < 1$ für ein $\delta_0 > 0$. Nach Satz 2.5.4 gilt dann $\left| \frac{a_{k+1}}{a_k} \right| > 1 - \delta_0 =: q_0$ für höchstens endlich viele k . Es gibt also ein k_1 (das ohne Einschränkung größer als k_0 ist), so dass $a_k \neq 0$ und $\left| \frac{a_{k+1}}{a_k} \right| \leq q_0$ für $k > k_1$ gilt. Damit folgt für $k > k_1$:

$$\left| \frac{a_k}{a_{k_1}} \right| = \left| \frac{a_k}{a_{k-1}} \right| \cdot \left| \frac{a_{k-1}}{a_{k-2}} \right| \cdots \left| \frac{a_{k_1+1}}{a_{k_1}} \right| \leq q_0^{k-k_1},$$

also $|a_k| \leq M_0 \cdot q_0^k$ mit $M_0 := |a_{k_1}| q_0^{-k_1}$. Damit ist $\sum M_0 q_0^k$ nach Satz 3.2.9 eine konvergente Majorante für $\sum |a_k|$ und somit $\sum a_k$ nach dem Majorantenkriterium absolut konvergent.

Zu ii): Es sei $\liminf \left| \frac{a_{k+1}}{a_k} \right| = 1 + \delta_1$ für ein $\delta_1 > 0$. Nach Satz 2.5.4 ist $\left| \frac{a_{k+1}}{a_k} \right| < 1 + \delta_1$ für höchstens endlich viele k . Es gibt also wieder ein $k_1 > k_0$, so dass $a_k \neq 0$ und $\left| \frac{a_{k+1}}{a_k} \right| \geq 1 + \delta_1 =: q_1$ ist für $k > k_1$. Damit gilt für alle $k > k_1$:

$$\left| \frac{a_k}{a_{k_1}} \right| = \left| \frac{a_k}{a_{k-1}} \right| \cdot \left| \frac{a_{k-1}}{a_{k-2}} \right| \cdots \left| \frac{a_{k_1+1}}{a_{k_1}} \right| \geq q_1^{k-k_1},$$

woraus $|a_k| \geq M_1 q_1^k$ mit $M_1 := |a_{k_1}| q_1^{-k_1}$ folgt. Damit gilt $|a_k| \rightarrow \infty$ für $k \rightarrow \infty$, und $\sum a_k$ ist divergent. \square

BEISPIEL 3.2.2

Die Reihe

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}$$

ist konvergent. Wir wenden dazu das Quotientenkriterium an mit $a_n = \frac{n!}{n^n}$:

$$\left| \frac{a_n}{a_{n+1}} \right| = \frac{(n+1)^{n+1}}{(n+1)!} \cdot \frac{n!}{n^n} = \frac{1}{n+1} \cdot \frac{(n+1)^{n+1}}{n^n} = \left(1 + \frac{1}{n}\right)^n \underset{3.1.1}{\geq} 1 + n \cdot \frac{1}{n} = 2.$$

Somit ist $\left| \frac{a_{n+1}}{a_n} \right| \leq \frac{1}{2}$ für alle n , insbesondere auch $\limsup \left| \frac{a_{n+1}}{a_n} \right| \leq \frac{1}{2}$. Nach dem Quotientenkriterium 3.2.10(i) ist die Reihe $\sum \frac{n!}{n^n}$ damit konvergent.

SATZ 3.2.11 (Leibnizsche Regel)

Es sei (a_k) eine monoton fallende Zahlenfolge mit $\lim a_n = 0$. Dann ist die alternierende Reihe

$$\sum_{k=0}^{\infty} (-1)^k a_k = a_0 - a_1 + a_2 - a_3 \cdots$$

konvergent.

BEWEIS

Strebt eine monoton fallende Folge gegen Null, so sind alle Glieder nichtnegativ. Es ist $S_{2n+1} = S_{2n} - a_{2n+1} \leq S_{2n}$ und $\lim(S_{2n+1} - S_{2n}) = \lim a_n = 0$. Ferner gilt

$$\begin{aligned} S_{2n+2} - S_{2n} &= -a_{2n+1} + a_{2n+2} \leq 0 \quad \text{und} \\ S_{2n+1} - S_{2n+3} &= a_{2n+3} - a_{2n+2} \leq 0 \quad . \end{aligned}$$

Somit bildet die Folge der Intervalle $I_n := [S_{2n+1}, S_{2n}]$ eine Intervallschachtelung. Nach Satz 2.4.4 gibt es genau ein S , das allen I_n angehört. Es folgt

$$S = \lim_{n \rightarrow \infty} S_{2n} = \sum_{k=0}^{\infty} (-1)^k a_k .$$

\square

BEISPIEL 3.2.3

Die alternierende harmonische Reihe

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}$$

konvergiert nach der Leibnizschen Regel. Sie ist nach Satz 3.2.4 damit ein Beispiel für eine Reihe, die konvergiert, aber nicht absolut konvergiert. Man kann auch Umordnungen dieser Reihe betrachten, beispielsweise

$$1 - \frac{1}{2} + \underbrace{\frac{1}{3}}_{1 \text{ Summand}} - \frac{1}{4} + \underbrace{\frac{1}{5} + \frac{1}{7}}_{2^1 \text{ Summanden}} - \frac{1}{6} + \underbrace{\frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15}}_{2^2 \text{ Summanden}} - \frac{1}{8} \pm \dots$$

$$= 1 - \frac{1}{2} + \sum_{k=0}^{\infty} \left(\sum_{l=2^k+1}^{2^{k+1}-1} \frac{1}{2l-1} - \frac{1}{2(k+1)} \right),$$

welche, wie man leicht zeigt, divergiert.

Man kann nun fragen: welche Eigenschaft muss eine konvergente Reihe haben, damit auch jede ihrer Umordnungen konvergiert?

DEFINITION 3.2.5

Es sei n_1, n_2, \dots eine Folge natürlicher Zahlen, in der jede natürliche Zahl einmal, aber auch nur einmal auftritt (d. h. die Abbildung $j \mapsto n_j$ ist eine Bijektion $\mathbb{N} \rightarrow \mathbb{N}$). Dann heißt die Reihe

$$\sum_{k=1}^{\infty} a_{n_k}$$

eine Umordnung der Reihe $\sum a_n$. Entsprechend erklärt man Umordnungen für Reihen mit beliebigem Startindex. Eine konvergente Reihe mit dem Reihenwert S heißt unbedingt konvergent, wenn jede ihrer Umordnungen gegen S konvergiert. Andernfalls heißt die Reihe bedingt konvergent.

SATZ 3.2.12

Es gilt für jede Reihe: $\sum a_k$ unbedingt konvergent $\Leftrightarrow \sum a_k$ absolut konvergent.

(OHNE BEWEIS)

□

SATZ 3.2.13 (Riemannscher Umordnungssatz)

Eine bedingt konvergente Reihe besitzt immer eine Umordnung, die gegen eine willkürlich vorgegebene Zahl konvergiert.

(OHNE BEWEIS)

□

Wir betrachten jetzt die Multiplikation unendlicher Reihen.

DEFINITION 3.2.6

Sind

$$\sum_{j=0}^{\infty} a_j, \sum_{k=0}^{\infty} b_k$$

unendliche Reihen, so versteht man unter einer Produktreihe der beiden Reihen eine unendliche Reihe, in der jedes Produkt $a_j b_k$ genau einmal vorkommt (die Reihenfolge ist beliebig).

Eine spezielle Produktreihe gibt

DEFINITION 3.2.7

Sind

$$\sum_{j=0}^{\infty} a_j, \sum_{k=0}^{\infty} b_k$$

unendliche Reihe, so verstehen wir unter ihrem Cauchyprodukt die unendliche Reihe

$$\sum_{n=0}^{\infty} \left(\sum_{j=0}^n a_j b_{n-j} \right).$$

SATZ 3.2.14

Sind $\sum a_j$ und $\sum b_k$ absolut konvergente unendliche Reihen, so ist jede ihrer Produktreihen absolut konvergent (insb. auch ihr Cauchyprodukt).

(OHNE BEWEIS)

□

3.3. Die Exponentialfunktion

SATZ 3.3.1

Die unendliche Reihe

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ist für alle $x \in \mathbb{R}$ absolut konvergent.

BEWEIS

Wir wenden das Quotientenkriterium an:

$$a_n = \frac{x^n}{n!} \Rightarrow \left| \frac{a_{n+1}}{a_n} \right| = \frac{|x|^{n+1}}{(n+1)!} \cdot \frac{n!}{|x|^n} = \frac{|x|}{n+1},$$

und damit $\limsup \left| \frac{a_{n+1}}{a_n} \right| = 0$. Nach dem Quotientenkriterium ist die Reihe damit für alle $x \in \mathbb{R}$ absolut konvergent. □

DEFINITION 3.3.1

Wir definieren die natürliche Exponentialfunktion durch

$$\exp : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & \exp(x) \end{cases} \text{ mit } \exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

(Aus Satz 3.3.1 folgt, dass diese Definition für alle $x \in \mathbb{R}$ Sinn macht)

Der Name Exponentialfunktion ist nicht zufällig. Als Anwendung des Konzepts des Cauchyprodukts ergibt sich vielmehr, dass die Exponentialfunktion der vertrauten Funktionalgleichung genügt.

SATZ 3.3.2

Für $x, y \in \mathbb{R}$ gilt $\exp(x + y) = \exp(x) \exp(y)$.

BEWEIS

Es ist nach Definition

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!}.$$

Dann ist nach dem Produktreihensatz 3.2.14 auch das Cauchyprodukt der beiden unendlichen Reihen absolut konvergent. Wir erhalten

$$\begin{aligned} \exp(x) \exp(y) &= \sum_{n=0}^{\infty} \sum_{j=0}^n \frac{x^j}{j!} \cdot \frac{y^{n-j}}{(n-j)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^n \frac{n!}{j!(n-j)!} x^j y^{n-j} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^n \binom{n}{j} x^j y^{n-j} = \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n = \exp(x+y). \end{aligned}$$

□

DEFINITION 3.3.2 (Eulersche Zahl)

Wir setzen

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Der numerische Wert von e ergibt sich zu $e = 2,71828\dots$

SATZ 3.3.3

Es gilt:

- (i) Für $m \in \mathbb{Z}$ ist $\exp(m) = e^m$.
- (ii) Es ist $\exp(x) > 0$ für alle $x \in \mathbb{R}$.

BEWEIS

Man beweist (i) zunächst durch vollständige Induktion nach $m \in \mathbb{N}$. Für $m = 0$ folgt die Aussage aus der Definition von e . Ist $m < 0$, so ist nach dem eben Bewiesenen $\exp(-m) = e^{-m}$, also

$$\exp(m) \exp(-m) = \exp(m + (-m)) = \exp(0) = 1 \Rightarrow \exp(m) = (\exp(-m))^{-1} = (e^{-m})^{-1} = e^m.$$

Zu (ii): Für $x > 0$ ist offenbar $\exp(x) > 0$. Die Aussage gilt dann wegen $\exp(0) = \exp(x) \exp(-x) = 1$ auch für $x < 0$. □

DEFINITION 3.3.3

Wir schreiben auch $\exp(x) =: e^x$. Nach Satz 3.3.3(i) liefert für $x \in \mathbb{Z}$ die Definition dasselbe Resultat wie die Definition als Potenz.

4. Stetigkeit

4.1. Grundbegriffe

Im letzten Kapitel haben wir uns mit Zahlenfolgen beschäftigt, d. h. mit Funktionen, deren Definitionsbereich eine Teilmenge der ganzen Zahlen ist. Wir beschäftigen uns nun mit Funktionen, die typischerweise auf Intervallen oder Vereinigungen von solchen definiert sind. Versucht man den Graphen einer auf einem Intervall definierten Funktion mit einem Stift zu zeichnen, so kann dies möglich sein, ohne den Stift abzusetzen. Andererseits, wenn der Graph Sprünge macht, wird es nötig sein, den Stift abzusetzen. Diese unpräzise Idee wird durch den Begriff der Stetigkeit erfasst.

DEFINITION 4.1.1

Es sei $X \subseteq \mathbb{R}$ und $f : X \rightarrow \mathbb{R}$ eine Funktion. f heißt stetig in einem Punkt $\xi \in X$, wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon)$ gibt, so dass für alle $x \in X$ mit $|x - \xi| < \delta$ immer $|f(x) - f(\xi)| < \varepsilon$ gilt. Anders formuliert: für $\eta = f(\xi)$ existiert zu jeder Umgebung $U_\varepsilon(\eta)$ stets eine genügend kleine Umgebung $U_\delta(\xi)$ mit $f(U_\delta(\xi)) \subseteq U_\varepsilon(\eta)$.

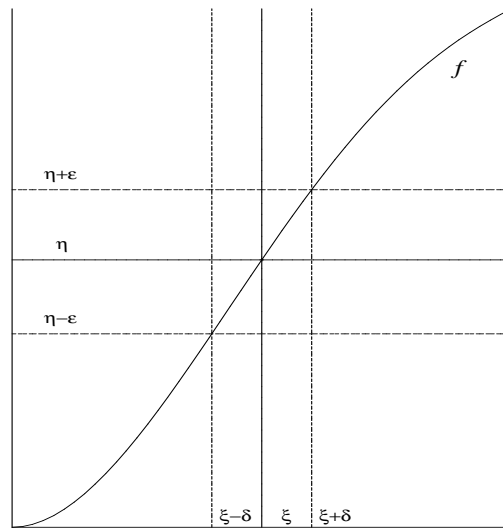


ABBILDUNG 4.1. Stetigkeit in ξ

SATZ 4.1.1

Es sei $X \subseteq \mathbb{R}$ und $f : X \rightarrow \mathbb{R}$ eine Funktion. Dann sind folgende Bedingungen äquivalent:

- (i) f stetig in $\xi \in X$.
- (ii) Für jede Folge (x_n) in X mit $x_n \rightarrow \xi$ gilt $f(x_n) \rightarrow f(\xi)$.

BEWEIS

\Rightarrow :

Es sei f stetig in $\xi \in X$. Es sei $\varepsilon > 0$ gegeben. Wir bestimmen $\delta = \delta(\varepsilon) > 0$, so dass für alle $x \in X$ mit $|x - \xi| < \delta$ immer $|f(x) - f(\xi)| < \varepsilon$ gilt. Es gibt dann ein $n_0 = n_0(\delta)$, so dass für alle $n > n_0$ stets $|x_n - \xi| < \delta$ gilt. Daraus folgt $|f(x_n) - f(\xi)| < \varepsilon$ für alle $n > n_0$. Also $f(x_n) \rightarrow f(\xi)$ für $n \rightarrow \infty$.

\Leftarrow :

Für jede Folge (x_n) mit $x_n \rightarrow \xi$ und $x_n \in X$ gelte $f(x_n) \rightarrow f(\xi)$. Annahme: f ist nicht stetig in $\xi \in X$. Dann gibt es ein Ausnahme- ε , zu dem es kein δ gibt, so dass aus $|x - \xi| < \delta$ für $x \in X$ immer $|f(x) - f(\xi)| < \varepsilon$ folgt, d. h. zu jedem $n \in \mathbb{N}$ gibt es ein $x_n \in X$ mit $|x_n - \xi| < \frac{1}{n}$ aber $|f(x_n) - f(\xi)| \geq \varepsilon$ für das Ausnahme- ε . Damit gilt $x_n \rightarrow \xi$ aber nicht $f(x_n) \rightarrow f(\xi)$. Widerspruch. \square

SATZ 4.1.2

Es seien f, g auf $X \subseteq \mathbb{R}$ definiert und stetig in $\xi \in X$. Dann sind auch $f + g$ und $f \cdot g$ in ξ stetig. Ist überdies $g(\xi) \neq 0$, so ist die auf $\{x \in X \mid g(x) \neq 0\}$ definierte Funktion $\frac{f}{g}$ ebenfalls in ξ stetig.

BEWEIS

Die Aussage folgt aus Satz 4.1.1 und der entsprechenden Aussage für Folgen in Satz 2.3.4. □

SATZ 4.1.3

Es seien P, Q Polynome und $R(x) = \frac{P(x)}{Q(x)}$ eine rationale Funktion. Es gilt: Alle Polynome sind stetig in allen $\xi \in \mathbb{R}$. Rationale Funktionen R sind stetig in allen ξ mit $Q(\xi) \neq 0$.

BEWEIS

Die konstanten Funktionen

$$\text{const} : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & c \text{ fest} \end{cases}$$

sind offenbar stetig, ebenso die Identität

$$\text{id} : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & x \end{cases} .$$

Wiederholte Anwendung von Satz 4.1.2 ergibt die Stetigkeit der Polynome P, Q sowie von R . □

SATZ 4.1.4 (Komposition zweier Funktionen)

Das Kompositum $f \circ g$ möge existieren. g sei in ξ und f in $g(\xi)$ stetig. Dann ist auch $f \circ g$ in ξ stetig.

BEWEIS

Aus $x_n \rightarrow \xi$ folgt zunächst $g(x_n) \rightarrow g(\xi)$, daraus dann $(f \circ g)(x_n) \rightarrow (f \circ g)(\xi)$. Die Behauptung folgt aus Satz 4.1.1. □

SATZ 4.1.5

Sind die Funktionen f und g auf X definiert und in ξ stetig, so sind auch die Funktionen $|f|$, $\max(f, g)$ und $\min(f, g)$ in ξ stetig.

BEWEIS

(Übungsaufgabe) □

Man kann auch von Stetigkeit einer Funktion auf einer Menge X sprechen.

DEFINITION 4.1.2

Eine Funktion $f : X \rightarrow \mathbb{R}$ heißt stetig auf Y für ein $Y \subseteq X$, wenn f stetig in jedem $\xi \in Y$ ist.

4.2. Offene, abgeschlossene und kompakte Mengen

Viele Eigenschaften stetiger Funktionen f und die Definition der Stetigkeit selbst lassen sich sehr kurz beschreiben, wenn man gewisse Typen von Mengen und deren Bilder oder Urbilder unter der Funktion f betrachtet. Wir werden in der Folge die Typen der offenen, abgeschlossenen bzw. kompakten Mengen einführen.

DEFINITION 4.2.1

Eine Teilmenge $O \subseteq \mathbb{R}$ heißt offen, wenn es zu jedem $\xi \in O$ eine ε -Umgebung $U_\varepsilon(\xi)$ gibt mit $U_\varepsilon(\xi) \subseteq O$. Ist $X \subseteq \mathbb{R}$, so heißt ein $O \subseteq X$ X -offen (bzw. relativ offen bzgl. X), wenn es zu jedem $\xi \in O$ ein $U_\varepsilon(\xi)$ gibt, so dass $U_\varepsilon(\xi) \cap X \subseteq O$ gilt. Eine Teilmenge $A \subseteq \mathbb{R}$ heißt abgeschlossen, wenn der Grenzwert jeder konvergenten Folge (x_n) mit $x_n \in A$ selbst in A liegt. Eine Teilmenge $K \subseteq \mathbb{R}$ heißt kompakt, wenn jede Folge in K eine konvergente Teilfolge besitzt, deren Grenzwert wieder zu K gehört.

BEISPIEL 4.2.1

Wie man leicht sieht, sind offene Intervalle (a, b) oder $(-\infty, b)$ bzw. (a, ∞) offen, und abgeschlossene Intervalle $[a, b]$ oder $[a, \infty)$ bzw. $(-\infty, b]$ abgeschlossen.

SATZ 4.2.1

Eine Teilmenge $A \subseteq \mathbb{R}$ ist genau dann abgeschlossen, wenn das Komplement $A' = \mathbb{R} \setminus A$ offen ist.

BEWEIS

\Rightarrow :

Es sei $A \subseteq \mathbb{R}$ abgeschlossen und $\xi \in A'$. Wir haben zu zeigen: es gibt ein $\varepsilon > 0$, so dass $U_\varepsilon(\xi) \subseteq A'$ gilt. Annahme: es gibt kein solches $\varepsilon > 0$. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in A \cap U_{\frac{1}{n}}(\xi)$. Offenbar gilt $x_n \rightarrow \xi$ für $n \rightarrow \infty$ und $\xi \notin A$ im Widerspruch zur Abgeschlossenheit von A .

\Leftarrow :

Es sei A' offen und (x_n) eine Folge mit $x_n \in A$ und $x_n \rightarrow \xi$ für ein $\xi \notin A$. Dann gibt es $\varepsilon > 0$ mit $U_\varepsilon(\xi) \in A' \Rightarrow U_\varepsilon(\xi) \cap A = \emptyset \Rightarrow x_n \notin U_\varepsilon(\xi) \Rightarrow x_n \not\rightarrow \xi$, ein Widerspruch. \square

SATZ 4.2.2

Eine Teilmenge $K \subseteq \mathbb{R}$ ist kompakt genau dann, wenn sie abgeschlossen und beschränkt ist.

BEWEIS

\Rightarrow :

Es sei $K \subseteq \mathbb{R}$ kompakt. Dann liegt der Limes ξ jeder konvergenten Folge aus K selbst in K . Also ist K abgeschlossen. Annahme: K ist unbeschränkt. Dann gibt es eine Zahlenfolge (x_n) in K , die gegen ∞ divergiert, und daher keine konvergente Teilfolge enthält, Widerspruch.

\Leftarrow :

Es sei $K \subseteq \mathbb{R}$ abgeschlossen und beschränkt, sowie (x_n) eine Folge in K . Dann ist auch (x_n) beschränkt, und hat nach Satz 2.4.2 (Bolzano-Weierstraß) eine konvergente Teilfolge $x_{n_k} \rightarrow \xi$ für $k \rightarrow \infty$. Da K abgeschlossen ist folgt $\xi \in K$. Damit ist K kompakt. \square

Wir können nun die Intervalle folgendermaßen klassifizieren:

SATZ 4.2.3

Es gilt:

- (i) Ein Intervall I ist offen $\Leftrightarrow I = (a, b)$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$.
- (ii) Ein Intervall I ist abgeschlossen $\Leftrightarrow I = [a, b]$, $I = (-\infty, b]$, $I = [a, \infty)$, $I = \mathbb{R}$.
- (iii) Ein Intervall I ist kompakt $\Leftrightarrow I = [a, b]$ mit $a, b \in \mathbb{R}$.

OHNE BEWEIS

\square

Nun kommen wir zum Zusammenhang zwischen stetigen Funktionen und offenen Mengen.

SATZ 4.2.4

Es sei $f: X \rightarrow \mathbb{R}$ eine Funktion mit $X \subseteq \mathbb{R}$. Dann gilt:

f stetig auf $X \Leftrightarrow$ das Urbild jeder offenen Menge ist X -offen.

BEWEIS

Wir beweisen nur die Hinrichtung. Es sei $O \subseteq \mathbb{R}$ eine offene Menge und $\xi \in f^{-1}(O)$. Dann gibt es ein $y \in O$ mit $f(\xi) = y$. Da O offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(y) \subseteq O$. Wegen der Stetigkeit von f in ξ gibt es ein $\delta > 0$, so dass aus $x \in U_\delta(\xi) \cap X$ stets $f(x) \in U_\varepsilon(y) \subseteq O$ folgt. Also gilt $U_\delta(\xi) \cap X \subseteq f^{-1}(O)$. Damit ist $f^{-1}(O)$ eine X -offene Menge. \square

4.3. Grundeigenschaften stetiger Funktionen

SATZ 4.3.1 (Zwischenwertsatz)

Eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.

BEWEIS

Für $f(a) = f(b)$ ist nichts zu beweisen. Es sei $f(a) < f(b)$ und $\eta \in (f(a), f(b))$. Dazu sei $A = \{x \in [a, b] \mid f(x) \leq \eta\}$. Zunächst ist $A \neq \emptyset$, da $a \in A$ ist. Da A beschränkt ist, existiert $\xi = \sup(A)$, und es ist $\xi \in [a, b]$. Es sei $x_n \rightarrow \xi$ mit $x_n \in A$. Wegen der Stetigkeit von f muss $f(x_n) \rightarrow f(\xi)$ gelten. Da $f(x_n) \leq \eta$ gilt folgt $f(\xi) \leq \eta$. Annahme: $f(\xi) < \eta$. Dann gibt es $U_\varepsilon(f(\xi)) \subseteq (-\infty, \eta)$. Wegen der Stetigkeit von f in ξ gibt es $\delta > 0$, so dass $x \in U_\delta(\xi) \Rightarrow f(x) \in U_\varepsilon(f(\xi))$ gilt, also $f(x) \leq \eta$ für alle $x \in U_\delta(\xi)$. Damit ist $U_\delta(\xi) \subseteq A$ im Widerspruch zu $\sup(A) = \xi$. Also ist $f(\xi) = \eta$, und der Zwischenwert η wird angenommen. \square

SATZ 4.3.2

Es sei $X \subseteq \mathbb{R}$, $K \subseteq X$ und K kompakt, sowie f stetig auf K . Dann ist auch das Bild $f(K)$ kompakt (kurz: das stetige Bild kompakter Mengen ist wieder kompakt).

BEWEIS

Es sei (y_n) eine Zahlenfolge mit $y_n \in f(K)$ für alle n . Zu zeigen: (y_n) hat eine konvergente Teilfolge. Es gibt $x_n \in K$ mit $f(x_n) = y_n$. Da K kompakt ist, besitzt (x_n) eine Teilfolge (x_{n_k}) mit $x_{n_k} \rightarrow \xi$ für $k \rightarrow \infty$. Wegen der Stetigkeit von f folgt $y_{n_k} \rightarrow f(\xi)$. \square

SATZ 4.3.3

Eine stetige Funktion f mit kompaktem Definitionsbereich $X \subseteq \mathbb{R}$ ist beschränkt, und besitzt Minimum sowie Maximum, d. h. es gibt in X eine Minimalstelle x_1 und eine Maximalstelle x_2 , so dass $f(x_1) \leq f(x) \leq f(x_2)$ für alle $x \in X$ gilt.

BEWEIS

Nach Satz 4.3.2 ist $f(X)$ kompakt und damit nach Satz 4.2.2 abgeschlossen und beschränkt. Damit ist $s = \sup(f(X)) \in f(X)$, d. h. es gibt ein $x_2 \in X$ mit $f(x_2) = s$. Ebenso gibt es $i = \inf(f(X)) \in f(X)$ und $x_1 \in X$ mit $f(x_1) = i$. \square

DEFINITION 4.3.1

Die Funktion f heißt gleichmäßig stetig auf X , wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x, y \in X$ gilt: $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

SATZ 4.3.4

Jede stetige Funktion mit kompaktem Definitionsbereich X ist darauf sogar gleichmäßig stetig.

Wir geben diesen Beweis nicht, weisen jedoch darauf hin, dass er mittels der im Folgenden zu erklärenden Überdeckungseigenschaft einer kompakten Menge gegeben werden kann.

DEFINITION 4.3.2

Es sei $M \subseteq \mathbb{R}$ eine Teilmenge von \mathbb{R} , und jedes $x \in M$ möge in einer gewissen offenen Menge O_x liegen, so dass

$$M \subseteq \bigcup_{x \in M} O_x$$

gilt. In diesem Fall nennen wir das Mengensystem $\mathcal{O} = \{O_x \mid x \in M\}$ eine offene Überdeckung von M . Gibt es ein endliches Teilsystem $\mathcal{E} = \{O_{x_1}, \dots, O_{x_n}\} \subseteq \mathcal{O}$, das bereits ganz M überdeckt, so sagt man, \mathcal{O} enthalte eine endliche Überdeckung (von M).

SATZ 4.3.5 (Überdeckungssatz von Heine-Borel)

Eine Teilmenge von \mathbb{R} ist genau dann kompakt, wenn jede ihrer offenen Überdeckungen eine endliche Überdeckung enthält.

OHNE BEWEIS

□

4.4. Der Umkehrsatz für streng monotone Funktionen

DEFINITION 4.4.1

Es sei $f : X \rightarrow \mathbb{R}$, $X \subseteq \mathbb{R}$ eine Funktion.

- (i) f heißt wachsend (bzw. streng wachsend) auf X , falls $f(x_1) \leq f(x_2)$ (bzw. $f(x_1) < f(x_2)$) gilt für alle $x_1, x_2 \in X$ mit $x_1 < x_2$.
- (ii) f heißt abnehmend (bzw. streng abnehmend) auf X , falls $f(x_1) \geq f(x_2)$ (bzw. $f(x_1) > f(x_2)$) gilt für alle $x_1, x_2 \in X$ mit $x_1 < x_2$.
- (iii) f heißt monoton (bzw. streng monoton), falls f (streng) wachsend oder (streng) abnehmend ist.

SATZ 4.4.1 (Umkehrsatz)

Es sei I ein beliebiges Intervall, und $f : I \rightarrow \mathbb{R}$ auf I streng wachsend. Dann ist die Umkehrfunktion $f^{-1} : f(I) \rightarrow I$ vorhanden, streng wachsend, und stetig. Ist f selbst stetig, so ist $f(I)$ ein Intervall. Eine entsprechende Aussage gilt für streng abnehmende Funktionen.

OHNE BEWEIS

□

SATZ 4.4.2

Es sei $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ ein Intervall und

$$f : \begin{cases} I & \rightarrow \mathbb{R} \\ x & \mapsto x^n \end{cases}.$$

- (i) Ist n ungerade, so besitzt f für jedes $I \subseteq \mathbb{R}$ eine stetige Umkehrfunktion auf $f(I)$. Es ist $f(\mathbb{R}) = \mathbb{R}$ und $f([0, \infty)) = [0, \infty)$.
- (ii) Ist n gerade, so besitzt f für jedes $I \subseteq [0, \infty)$ eine stetige Umkehrfunktion auf $f(I)$. Es ist $f(\mathbb{R}) = f([0, \infty)) = [0, \infty)$.

BEWEIS

Für $n \in \mathbb{N}$ folgt durch wiederholte Multiplikation der Ungleichung $0 \leq x_1 < x_2$ mit sich selbst (formaler Beweis durch vollständige Induktion) $0 \leq x_1^n < x_2^n$. Für $1 \leq x$ gilt $x \leq x^n$. Damit ist f nach oben unbeschränkt. Nach dem Zwischenwertsatz 4.3.1 nimmt f auf $[0, \infty)$ jeden Wert aus $[0, \infty)$ an. Der Rest folgt aus Satz 4.4.1. □

DEFINITION 4.4.2 (Wurzelfunktion)

Es sei $n \in \mathbb{N}$. Die nach Satz 4.4.2 definierte Umkehrfunktion von

$$f : \begin{cases} [0, \infty) & \rightarrow \mathbb{R} \\ x & \mapsto x^n \end{cases}$$

heißt n -te Wurzel. Schreibweise: $f^{-1}(y) = \sqrt[n]{y}$ oder $f^{-1}(y) = y^{\frac{1}{n}}$. Kurz: Für $x, y \geq 0$ gilt $x = \sqrt[n]{y} \Leftrightarrow x^n = y$. Im Fall $n = 2$ schreibt man kurz \sqrt{y} (Quadratwurzel).

BEMERKUNG 4.4.1

Aus den Sätzen von Abschnitt 4.1 folgt die Stetigkeit aller aus $\text{id} : x \mapsto x$ mittels der vier Grundrechenarten und der Wurzelfunktion gebildeten Funktionen in jedem Punkt, in dem Sie definiert sind. Beispielsweise

$$f(x) = \frac{\sqrt[4]{3 + \sqrt{\frac{x+1}{x+2}}}}{x^3 + \sqrt[5]{x^2 + 1}}.$$

4.5. Grenzwerte von Funktionen und einseitige Grenzwerte

Der Grenzwertbegriff ist, wie bei Folgen, von großer Wichtigkeit auch für Funktionen. Er ermöglicht es, in manchen Fällen, Funktionen mit „lückenhaftem“ Definitionsbereich „stetig zu ergänzen“.

BEISPIEL 4.5.1

Es sei $X = \mathbb{R} \setminus \{1\}$ („Punktierter Zahlenstrahl“) und

$$f : \begin{cases} X & \rightarrow \mathbb{R} \\ x & \mapsto \frac{x^2-1}{x-1} \end{cases} .$$

f ist nicht stetig in $\xi = 1$ wegen $\xi \notin X$. Jedoch ist $f(x) = x + 1$ für $x \neq 1$. Man sagt, f ist im Punkt $\xi = 1$ stetig ergänzbar: wir können f zu einer auf ganz \mathbb{R} stetigen Funktion $g : x \mapsto x + 1$ fortsetzen (d. h. es ist $g|_X = f$). Der Wert der Fortsetzung g in der „Lücke“ ξ wird sich als Grenzwert der Funktion f in diesem Punkt erweisen. Zur Existenz dieses Grenzwerts ist es zunächst notwendig, dass 1 ein Häufungspunkt von X ist.

DEFINITION 4.5.1

Ein $\xi \in \mathbb{R}$ heißt Häufungspunkt (HP) der Menge $M \subseteq \mathbb{R}$, wenn es eine Folge (x_n) aus M gibt, die gegen ξ konvergiert, und deren Glieder alle von ξ verschieden sind.

BEISPIEL 4.5.2

Ist $\xi \in \mathbb{R}$ und $X = \mathbb{R} \setminus \{\xi\}$, so ist ξ ein Häufungspunkt von X .

SATZ 4.5.1

Es sei $M \subseteq \mathbb{R}$. Ein $\xi \in \mathbb{R}$ ist genau dann Häufungspunkt von M , wenn in jeder ε -Umgebung von ξ ein von ξ verschiedener Punkt aus M liegt.

BEWEIS

Folgt aus der Definition der Folgenkonvergenz. □

DEFINITION 4.5.2

Die Funktion f sei auf X definiert und ξ sei ein Häufungspunkt von X . Wir sagen: $f(x) \rightarrow \eta$ für $x \rightarrow \xi$ ($f(x)$ konvergiert oder strebt gegen η , wenn x gegen ξ strebt), auch geschrieben

$$\lim_{x \rightarrow \xi} f(x) = \eta ,$$

wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in X$ mit $0 < |x - \xi| < \delta$ immer $|f(x) - \eta| < \varepsilon$ ist.

SATZ 4.5.2

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion, ξ ein Häufungspunkt von X . Es ist

$$\lim_{x \rightarrow \xi} f(x) = \eta$$

genau dann, wenn für jede Folge (x_n) aus X , die gegen ξ konvergiert und deren Glieder alle verschieden von ξ sind, stets $f(x_n) \rightarrow \eta$ konvergiert.

BEWEIS

Der Beweis ist ganz analog dem Beweis von Satz 4.1.1. Wir überspringen die Einzelheiten. □

DEFINITION 4.5.3

Es sei $X \subseteq \mathbb{R}$. Ein Punkt $\xi \in X$ heißt isolierter Punkt von X , falls ξ kein Häufungspunkt von $X \setminus \{\xi\}$ ist.

SATZ 4.5.3

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion, $\xi \in X$. f ist genau dann stetig in ξ , wenn ξ ein isolierter Punkt von X ist oder $\lim_{x \rightarrow \xi} f(x)$ existiert und $\lim_{x \rightarrow \xi} f(x) = f(\xi)$ gilt.

OHNE BEWEIS

□

Gegeben sei eine Funktion $f : X \rightarrow \mathbb{R}$ und $\xi \in X$. Ist man nur am Verhalten von f auf einer Seite des Punkts ξ interessiert, so wird man die Begriffe der einseitigen Grenzwerte bzw. der einseitigen Stetigkeit verwenden.

DEFINITION 4.5.4

Es sei f auf X erklärt und ξ sei ein Häufungspunkt von

$$X_l := \{x \in X \mid x < \xi\} \text{ bzw. } X_r := \{x \in X \mid x > \xi\}.$$

Im ersten Fall sei $f_l := f|_{X_l}$, im zweiten Fall $f_r := f|_{X_r}$. Die links- bzw. rechtsseitigen Grenzwerte von f werden als die Grenzwerte der Funktionen f_l bzw. f_r erklärt. Strebt $f_l(x) \rightarrow \eta_l$ für $x \rightarrow \xi$ bzw. $f_r(x) \rightarrow \eta_r$ für $x \rightarrow \xi$, so sagen wir, f besitze für $x \rightarrow \xi^-$ bzw. für $x \rightarrow \xi^+$ den linksseitigen Grenzwert η_l bzw. den rechtsseitigen Grenzwert η_r . Schreibweise:

$$\begin{aligned} \lim_{x \rightarrow \xi^-} f(x) = \eta_l, \quad \lim_{x \rightarrow \xi^+} f(x) = \eta_r \text{ oder} \\ f(x) \rightarrow \eta_l \quad x \rightarrow \xi^-, \quad f(x) \rightarrow \eta_r \quad x \rightarrow \xi^+ \text{ oder kurz} \\ f(\xi^-) = \eta_l, \quad f(\xi^+) = \eta_r. \end{aligned}$$

SATZ 4.5.4

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und ξ ein Häufungspunkt von $X_l = \{x \in X \mid x < \xi\}$. Es gilt

$$\lim_{x \rightarrow \xi^-} f(x) = \eta$$

genau dann, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in U_\delta(\xi) \cap X_l$ gilt: $|f(x) - \eta| < \varepsilon$. Eine entsprechende Aussage gilt für rechtsseitige Grenzwerte.

OHNE BEWEIS

□

SATZ 4.5.5

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und ξ sei Häufungswert von $X_l = \{x \in X \mid x < \xi\}$. Es gilt

$$\lim_{x \rightarrow \xi^-} f(x) = \eta,$$

wenn für jede Folge (x_n) aus X_l , die gegen ξ strebt, auch $f(x_n)$ gegen η strebt.

OHNE BEWEIS

□

BEISPIEL 4.5.3

Das Signum (Vorzeichen) von $x \in \mathbb{R}$ ist definiert als die Funktion $\text{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ mit $\text{sgn}(x) = 1$ für $x > 0$, $\text{sgn}(x) = -1$ für $x < 0$, und $\text{sgn}(0) = 0$.

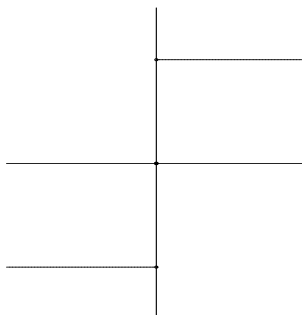


ABBILDUNG 4.2. Das Signum

Dann gilt: $\lim_{x \rightarrow 0^-} \operatorname{sgn}(x) = -1$, $\lim_{x \rightarrow 0^+} \operatorname{sgn}(x) = +1$ und $\operatorname{sgn}(0) = 0$.

SATZ 4.5.6

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und $\xi \in X$ sei Häufungspunkt sowohl von X_l als auch von X_r . Genau dann existiert der Grenzwert η von f an der Stelle ξ , wenn $f(\xi^-)$ und $f(\xi^+)$ existieren und mit η übereinstimmen.

OHNE BEWEIS

□

BEISPIEL 4.5.4

Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 2x & \text{falls } x < 2 \\ x^2 & \text{falls } x \geq 2 \end{cases} .$$

Dann ist

$$\lim_{x \rightarrow 2^-} f(x) = \lim_{x \rightarrow 2^-} (2x) = 4 \text{ und}$$

$$\lim_{x \rightarrow 2^+} f(x) = \lim_{x \rightarrow 2^+} (x^2) = 4 .$$

Also ist $\lim_{x \rightarrow 2} f(x) = 4$.

BEISPIEL 4.5.5

Es sei $g : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$g(x) = \begin{cases} x & \text{falls } x < 1 \\ x^2 - 1 & \text{falls } x \geq 1 \end{cases} .$$

Dann ist

$$\lim_{x \rightarrow 1^-} g(x) = \lim_{x \rightarrow 1^-} (x) = 1 \text{ aber}$$

$$\lim_{x \rightarrow 1^+} g(x) = \lim_{x \rightarrow 1^+} (x^2 - 1) = 0 .$$

Also besitzt $g(x)$ an der Stelle 1 keinen Grenzwert.

SATZ 4.5.7

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und $\xi \in X$ ein Häufungspunkt sowohl von X_l als auch von X_r . Genau dann ist f in ξ stetig, wenn die einseitigen Grenzwerte $f(\xi^-)$ und $f(\xi^+)$ existieren und mit $f(\xi)$ übereinstimmen.

OHNE BEWEIS

□

BEISPIEL 4.5.6

Es sei $h : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$h(x) = \begin{cases} x & \text{falls } x < 0 \\ 1 & \text{falls } x = 0 \\ x^2 & \text{falls } x > 0 \end{cases} .$$

Dann ist $f(0^-) = f(0^+) = 0 \neq f(0)$. Also ist f in $\xi = 0$ nicht stetig, obwohl beide einseitigen Grenzwerte existieren und übereinstimmen, da der Grenzwert an der Stelle 0 nicht mit dem Wert von f an der Stelle 0 übereinstimmt.

DEFINITION 4.5.5

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und $\xi \in X$. Die Funktion f heißt linksseitig stetig in ξ , falls $f|_{X_l \cup \{\xi\}}$ stetig in ξ ist, bzw. rechtsseitig stetig in ξ , falls $f|_{X_r \cup \{\xi\}}$ stetig in ξ ist.

SATZ 4.5.8

Es gilt für jedes $f : X \rightarrow \mathbb{R}$ und jedes $\xi \in X$:

$$f \text{ linksseitig stetig in } \xi \Leftrightarrow \xi \text{ isolierter Punkt in } X \text{ oder } f(\xi) = \lim_{x \rightarrow \xi^-} f(x).$$

OHNE BEWEIS □

Grenzwerte von Funktionen können schließlich auch für die uneigentlichen Grenzen $\pm\infty$ erklärt werden.

DEFINITION 4.5.6

Es sei $f : X \rightarrow \mathbb{R}$ eine auf einer nach oben unbeschränkten Menge X erklärte Funktion. Wir sagen: $f(x) \rightarrow \eta$ für $x \rightarrow \infty$ ($f(x)$ konvergiert/strebt gegen η für $x \rightarrow \infty$) genau dann, wenn es zu jedem $\varepsilon > 0$ ein $x_0 = x_0(\varepsilon)$ gibt, so dass für alle $x > x_0$ aus X stets $|f(x) - \eta| < \varepsilon$ gilt. Eine entsprechende Definition gilt für $x \rightarrow -\infty$.

Wir haben wieder ein Folgenkriterium für Grenzwerte $x \rightarrow \pm\infty$:

SATZ 4.5.9

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion, und X nach oben (bzw. unten) unbeschränkt. Es ist

$$\lim_{x \rightarrow \infty} f(x) = \eta \text{ bzw. } \lim_{x \rightarrow -\infty} f(x) = \eta$$

genau dann, wenn für jede Folge (x_n) aus X , die gegen ∞ (bzw. $-\infty$) divergiert, stets $f(x_n) \rightarrow \eta$ konvergiert.

OHNE BEWEIS □

SATZ 4.5.10 (Rechenregeln für Grenzwerte)

Es seien $f, g : X \rightarrow \mathbb{R}$ Funktionen mit $f(x) \rightarrow \eta$ und $g(x) \rightarrow \zeta$ für eine der Bedingungen $x \rightarrow \xi$, $x \rightarrow \xi^-$, $x \rightarrow \xi^+$, $x \rightarrow \infty$ oder $x \rightarrow -\infty$. Dann gilt:

- (i) $f(x) + g(x) \rightarrow \eta + \zeta$, $f(x) - g(x) \rightarrow \eta - \zeta$,
- (ii) $f(x) \cdot g(x) \rightarrow \eta \cdot \zeta$, $a \cdot f(x) \rightarrow a \cdot \eta \forall a \in \mathbb{R}$,
- (iii) $f(x)/g(x) \rightarrow \eta/\zeta$ falls $\zeta \neq 0$,
- (iv) $|f(x)| \rightarrow |\eta|$.

BEWEIS

Dieser Satz folgt aus den entsprechenden Sätzen 2.3.3/2.3.4 für Folgen und den Folgenkriterien für die Grenzwerte von Funktionen 4.5.2/4.5.5/4.5.9. □

SATZ 4.5.11

f und g mögen die gleichen Voraussetzungen erfüllen wie im vorigen Satz. Strebt $f(x)$ gegen Null und ist $g(x)$ beschränkt, so strebt auch $f(x)g(x)$ gegen Null.

BEWEIS

Das folgt aus Satz 2.3.2 und den Folgenkriterien für Grenzwerte. □

BEISPIEL 4.5.7

Es sei $b_p \neq 0$. Dann ist

$$\lim_{x \rightarrow \infty} \frac{a_0 + a_1x + \dots + a_px^p}{b_0 + b_1x + \dots + b_px^p} = \lim_{x \rightarrow \infty} \frac{\frac{a_0}{x^p} + \frac{a_1}{x^{p-1}} + \dots + a_p}{\frac{b_0}{x^p} + \frac{b_1}{x^{p-1}} + \dots + b_p} = \frac{a_p}{b_p},$$

d. h. genau wie bei Folgen ist der Grenzwert rationaler Funktionen durch die Koeffizienten der höchsten Potenzen festgelegt.

Wenn eine Funktion bei einer der Bewegungen $x \rightarrow \xi$, $x \rightarrow \xi-$, $x \rightarrow \xi+$ oder $x \rightarrow \pm\infty$ nicht konvergiert, so sagt man, sie divergiere. Bei gewissen Formen der Divergenz definieren wir schließlich wie für Folgen auch die uneigentlichen Grenzwerte über eine (ε, δ) -Forderung mit zugehörigem Folgenkriterium:

DEFINITION 4.5.7

Es sei $f : X \rightarrow \mathbb{R}$ eine Funktion und ξ ein Häufungspunkt von X . Wir sagen, f divergiert gegen ∞ (bzw. gegen $-\infty$) für $x \rightarrow \xi$, wenn es zu jedem $C > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in U_\delta(\xi) \setminus \{\xi\}$ stets $f(x) > C$ (bzw. $f(x) < -C$) gilt.

BEMERKUNG 4.5.1

Das zugehörige Folgenkriterium lautet: f divergiert gegen ∞ (bzw. $-\infty$) für $x \rightarrow \xi$ genau dann, wenn für jede Folge $x_n \rightarrow \xi$ mit $x_n \neq \xi$ gilt: $f(x_n) \rightarrow \infty$ (bzw. $f(x_n) \rightarrow -\infty$) für $n \rightarrow \infty$.

Entsprechende Definitionen/Folgenkriterien gelten für einseitige uneigentliche Grenzwerte.

SATZ 4.5.12

Es sei $n \in \mathbb{N}$, $a_j \in \mathbb{R}$ und $a_n \neq 0$. Dann gilt

$$\lim_{x \rightarrow \infty} (a_0 + a_1x + \dots + a_nx^n) = \begin{cases} \infty & \text{falls } a_n > 0 \\ -\infty & \text{falls } a_n < 0 \end{cases},$$

$$\lim_{x \rightarrow -\infty} (a_0 + a_1x + \dots + a_nx^n) = \begin{cases} \infty & \text{falls } a_n > 0 \text{ und } n \text{ gerade} \\ \infty & \text{falls } a_n < 0 \text{ und } n \text{ ungerade} \\ -\infty & \text{falls } a_n > 0 \text{ und } n \text{ ungerade} \\ -\infty & \text{falls } a_n < 0 \text{ und } n \text{ gerade} \end{cases}.$$

OHNE BEWEIS

□

4.6. Gleichmäßige Konvergenz und Stetigkeit

Bisher wissen wir noch nicht, ob die Exponentialfunktion stetig ist. Da wir diese durch eine unendliche Reihe und somit als Grenzwert einer Folge von Funktionen definiert haben, brauchen wir zur Behandlung dieser - und vieler anderer - Fragen ein Kriterium über die Stetigkeit von Grenzfunktionen von Funktionenfolgen.

DEFINITION 4.6.1

Es sei $f_n : X \rightarrow \mathbb{R}$ für $n \in \mathbb{N}$ eine Folge von Funktionen. Wir sagen, die Funktionenfolge (f_n) konvergiert/strebt gleichmäßig auf X gegen eine Grenzfunktion f , wenn es zu jedem $\varepsilon > 0$ einen Index n_0 gibt, so dass für alle $n > n_0$ und alle $x \in X$ stets $|f_n(x) - f(x)| < \varepsilon$ gilt.

BEMERKUNG 4.6.1

Aus der Definition der gleichmäßigen Stetigkeit folgt insbesondere, dass für jedes feste x

$$(*) \quad \lim_{n \rightarrow \infty} f_n(x) = f(x),$$

also punktweise Konvergenz vorliegt. (*) ist jedoch schwächer als die gleichmäßige Konvergenz. (*) besagt: zu jedem $\varepsilon > 0$ gibt es ein $n_0 = n_0(\varepsilon, x)$, das von ε und x abhängt, so dass für alle $n > n_0$ immer $|f_n(x) - f(x)| < \varepsilon$ gilt. Entscheidend bei der gleichmäßigen Konvergenz ist, dass zu $\varepsilon > 0$ ein n_0 gefunden werden kann, dass nur von ε und nicht von x abhängig ist.

BEISPIEL 4.6.1

Es sei $X = [0, 1]$ und

$$f_n : \begin{cases} X & \rightarrow \mathbb{R} \\ x & \mapsto x^n \end{cases} \quad \text{und} \quad f : \begin{cases} X & \rightarrow \mathbb{R} \\ x & \mapsto f(x) \end{cases} \quad \text{mit} \quad f(x) = \begin{cases} 0 & \text{falls } x \in [0, 1) \\ 1 & \text{falls } x = 1 \end{cases}.$$

Es ist offensichtlich $\lim f_n(x) = f(x)$, d. h. die Folge (f_n) konvergiert punktweise auf $[0, 1]$ gegen f . Die Konvergenz ist jedoch nicht gleichmäßig. Es sei $x_m = 1 - \frac{1}{2m}$ nach der Bernoullischen Ungleichung ist $(1+x)^n \geq 1+nx$ für $x > -1$ und $n \in \mathbb{N}$, also $f_n(x_m) \geq \frac{1}{2}$. Für $\varepsilon = \frac{1}{2}$ ist also $|f_n(x_m) - f(x_m)| \geq \varepsilon$ für alle $n \leq m$ und damit $n_0(\frac{1}{2}, x_m) \geq m$. Für $\varepsilon = \frac{1}{2}$ ist n_0 damit zwangsläufig auch von x abhängig. Wir beobachten zudem, dass alle Funktionen f_n stetig auf X sind, nicht aber die Grenzfunktion f , die unstetig in $\xi = 1$ ist.

Eine derartige Situation kann bei gleichmäßiger Konvergenz nicht auftreten:

SATZ 4.6.1

Es sei $X \subseteq \mathbb{R}$ und $f_n : X \rightarrow \mathbb{R}$ stetig auf X für alle n . Die Funktionenfolge (f_n) strebe gleichmäßig auf X gegen eine Grenzfunktion f , dann ist auch f stetig auf X .

BEWEIS

Es sei $\xi \in X$ und $\varepsilon > 0$ gegeben. Dann gibt es $n_0 = n_0(\frac{1}{3}\varepsilon)$, so dass $|f_n(x) - f(x)| < \frac{1}{3}\varepsilon$ gilt für alle $x \in X$ und $n > n_0$. Es sei $n_1 > n_0$. Da f_{n_1} stetig in ξ ist, gibt es $\delta = \delta(\frac{1}{3}\varepsilon)$, so dass $|f_{n_1}(x) - f_{n_1}(\xi)| < \frac{1}{3}\varepsilon$ ist für alle $x \in X \cap U_\delta(\xi)$. Für $x \in U_\delta(\xi)$ ist dann

$$|f(\xi) - f(x)| \leq |f(\xi) - f_{n_1}(\xi)| + |f_{n_1}(\xi) - f_{n_1}(x)| + |f_{n_1}(x) - f(x)| < \varepsilon.$$

Dies zeigt, dass f in ξ stetig ist. □

4.7. Stetigkeit der Exponentialfunktion, natürliche Logarithmusfunktion

SATZ 4.7.1

Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$ ist auf ganz \mathbb{R} stetig. Sie ist streng wachsend auf \mathbb{R} mit Wertebereich $\exp(\mathbb{R}) = (0, \infty)$. \exp besitzt eine streng wachsende Umkehrfunktion mit Definitionsbereich $(0, \infty)$. Es ist $\lim_{x \rightarrow -\infty} \exp(x) = 0$ sowie $\lim_{x \rightarrow \infty} \exp(x) = \infty$.

BEWEIS

Es ist

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Wir zeigen zunächst, dass die Folge der Partialsummen S_n auf jedem Intervall $[-M, M]$ für $M \in \mathbb{N}$ gleichmäßig konvergiert. Es sei $\varepsilon > 0$ gegeben und $N = 2M$. Für $k \geq N$ und $x \in [-M, M]$ haben wir

$$\frac{|x|^{k+1}}{(k+1)!} \cdot \frac{k!}{|x|^k} = \frac{|x|}{k+1} \leq \frac{1}{2} \Rightarrow \frac{|x|^{N+j}}{(N+j)!} \leq \frac{|x|^N}{N!} 2^{-j} \leq \frac{M^N}{N!} 2^{-j}.$$

Wegen der Konvergenz der unendlichen Reihe

$$\frac{M^N}{N!} \cdot \sum_{j=0}^{\infty} 2^{-j}$$

gibt es $j_0 = j_0(\varepsilon)$, so dass

$$\frac{M^N}{N!} \cdot \sum_{j=j_0+1}^{\infty} 2^{-j} < \varepsilon$$

gilt. Setze $n_0 = n_0(\varepsilon) := N + j_0$. Dann ist für $n > n_0$ und $x \in [-M, M]$

$$\left| \exp(x) - \sum_{k=0}^n \frac{x^k}{k!} \right| < \varepsilon.$$

Nach Satz 4.6.1 ist \exp damit stetig. Für $x_0 < x_1$ ist

$$\exp(x_1) = \exp(x_0) \cdot \sum_{k=0}^{\infty} \frac{(x_1 - x_0)^k}{k!} = \exp(x_0) \cdot \left(1 + \sum_{k=1}^{\infty} \frac{(x_1 - x_0)^k}{k!} \right) > \exp(x_0).$$

Wegen $\lim e^x = \infty$ und $\lim e^{-x} = 0$ nimmt $\exp(x)$ nach dem Zwischenwertsatz jeden Wert aus $(0, \infty)$ an. Nach Satz 4.4.1 besitzt \exp eine stetige Umkehrfunktion. \square

DEFINITION 4.7.1

Die Umkehrfunktion der Exponentialfunktion \exp heißt natürlicher Logarithmus, Schreibweise $\log x$.

SATZ 4.7.2

Die Funktion $\log : (0, \infty) \rightarrow \mathbb{R}$ ist stetig und streng wachsend. Es ist

$$\lim_{x \rightarrow 0^+} \log x = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \log x = \infty.$$

Für $x, y > 0$ ist $\log(xy) = \log(x) + \log(y)$.

BEWEIS

Das folgt aus den Eigenschaften der Exponentialfunktion. \square

Wir können jetzt auch allgemeine Exponentialfunktionen und Logarithmusfunktionen definieren:

DEFINITION 4.7.2

Für $a > 0$ aus \mathbb{R} definieren wir die verallgemeinerte Exponentialfunktion a^x durch $a^x := \exp(x \log a)$.

SATZ 4.7.3

Die allgemeine Exponentialfunktion a^x ist auf \mathbb{R} stetig. Für $a \neq 1$ ist ihr Wertebereich $(0, \infty)$. Es ist $1^x = 1$ für alle $x \in \mathbb{R}$. Für $a > 1$ ist a^x streng wachsend auf \mathbb{R} und $\lim a^x = \infty$ bzw. $\lim a^{-x} = 0$. Für $0 < a < 1$ ist a^x streng fallend auf \mathbb{R} und $\lim a^x = 0$ bzw. $\lim a^{-x} = \infty$. Für $a > 0$ und $a \neq 1$ besitzt a^x auf $(0, \infty)$ eine stetige Umkehrfunktion.

OHNE BEWEIS \square

DEFINITION 4.7.3

Die nach Satz 4.7.3 existierende Umkehrfunktion von a^x für $a > 0$ und $a \neq 1$ bezeichnen wir mit $\log_a x : (0, \infty) \rightarrow \mathbb{R}$. Es gilt also

$$y = \log_a x \quad \Leftrightarrow \quad x = a^y.$$

SATZ 4.7.4

Es sei $a > 0$ und $a \neq 1$. Der Definitionsbereich von $\log_a x$ ist $(0, \infty)$. Für $a > 1$ ist $\log_a x$ streng wachsend und stetig mit

$$\lim_{x \rightarrow 0^+} \log_a x = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \log_a x = \infty.$$

Für $a < 1$ ist $\log_a x$ streng fallend und stetig mit

$$\lim_{x \rightarrow 0^+} \log_a x = \infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \log_a x = -\infty.$$

5. Differenzierbarkeit

5.1. Definition der Ableitung

DEFINITION 5.1.1

Die reelle Funktion f sei auf dem (völlig beliebigen) Intervall I definiert. Wir sagen, f sei differenzierbar im Punkt $\xi \in I$, wenn

$$(*) \quad \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi}$$

oder (was dasselbe ist)

$$\lim_{h \rightarrow 0} \frac{f(\xi + h) - f(\xi)}{h}$$

existiert und endlich ist. Dieser Limes wird mit $f'(\xi)$ bezeichnet und die Ableitung von f an der Stelle ξ genannt. Besitzt der Differenzenquotient (*) nur einen links- bzw. rechtsseitigen Limes, so spricht man auch von links- bzw. rechtsseitiger Ableitung.

SATZ 5.1.1

Die Funktion f ist in jedem Punkt stetig, in dem sie differenzierbar ist.

BEWEIS

Es sei f in ξ differenzierbar. Wir haben zu zeigen, dass für jede Folge (x_n) , die gegen ξ konvergiert, und deren Glieder verschieden von ξ sind, $f(x_n) \rightarrow f(\xi)$ gilt. Es ist

$$f(x_n) - f(\xi) = \frac{f(x_n) - f(\xi)}{x_n - \xi} \cdot (x_n - \xi) \rightarrow f'(\xi) \cdot 0 = 0$$

und damit $f(x_n) \rightarrow f(\xi)$. □

Die Umkehrung gilt nicht, wie das folgende Beispiel zeigt.

BEISPIEL 5.1.1

Es sei $f(x) = |x|$ die Betragsfunktion. Sie ist stetig in $\xi = 0$, aber es ist

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{f(0+h) - f(0)}{h} &= \lim_{h \rightarrow 0^+} \frac{h}{h} = 1, \\ \lim_{h \rightarrow 0^-} \frac{f(0-h) - f(0)}{h} &= \lim_{h \rightarrow 0^-} \frac{-h}{h} = -1. \end{aligned}$$

Es existieren also die rechts- bzw. linksseitigen Ableitungen in $\xi = 0$. Da diese jedoch verschieden sind, existiert die Ableitung nicht.

DEFINITION 5.1.2

Wir sagen, die Funktion f sei auf einem Intervall I differenzierbar, wenn f in jedem Punkt von I differenzierbar ist. Ist f auf I differenzierbar, so wird durch $x \mapsto f'(x)$ eine Funktion auf I definiert, die wir mit f' bezeichnen und die Ableitung von f auf I nennen. Ist die Funktion f' in $\xi \in I$ selbst differenzierbar, so bezeichnen wir ihre Ableitung an dieser Stelle mit $f''(\xi)$, und nennen sie die zweite Ableitung von f an der Stelle ξ , und sagen, dass f zweimal Differenzierbar ist in ξ . Existiert $f''(\xi)$ für alle $\xi \in I$, so heißt f auf I zweimal Differenzierbar und $f'' : x \mapsto f''(x)$ die zweite Ableitung von f auf I . Allgemein wird durch Fortsetzung dieses Prozesses die n -te Ableitung $f^{(n)}$ von f definiert, falls sie existiert. Die nullte Ableitung $f^{(0)}$ ist f selbst.

Die Abbildung D , die jeder auf I differenzierbaren Funktion f ihre Ableitung f' zuordnet, heißt Differentiationsoperator. Mit seiner Hilfe können wir f' in der Form $D(f)$ oder kürzer Df schreiben. Die Ableitung $f'(\xi)$ von f an der Stelle ξ ist dann auch durch $(Df)(\xi)$ oder $Df(\xi)$ darstellbar.

Die höheren Ableitungen f'', f''', \dots bzw. ihre Werte $f''(\xi), f'''(\xi), \dots$ an der Stelle ξ schreiben wir gelegentlich in der Form

$$D^2 f, D^3 f, \dots \text{ bzw. } D^2 f(\xi), D^3 f(\xi), \dots$$

Schließlich ist noch die Verwendung des Differenzenquotienten $\frac{df}{dx}$ (oder auch $\frac{df(x)}{dx}$ bzw. $\frac{d}{dx}f(x)$) für die Ableitung f' üblich. Die höheren Ableitungen f'', f''', \dots werden dann mit

$$\frac{d^2 f}{dx^2}, \frac{d^3 f}{dx^3}, \dots \text{ oder } \frac{d^2 f(x)}{dx^2}, \frac{d^3 f(x)}{dx^3}, \dots$$

bezeichnet. Den Quotienten $\frac{f(x)-f(\xi)}{x-\xi}$ in ξ bezeichnet man auch als $\frac{\Delta f}{\Delta x}$. Die Idee von Leibniz war, Δx „unendlich klein“ zu wählen und mit dx zu bezeichnen. Es wurde später erkannt, dass diesem Gedankengang keine strenge logische Begründung gegeben werden kann. Aus dem Verhalten der Ableitung einer Funktion können wichtige „geometrische“ Schlüsse gezogen werden. Wir werden dies später bei Kurvendiskussionen im Detail erörtern und wollen hier nur die wichtigsten Gesichtspunkte besprechen.

SATZ 5.1.2

Es sei $f : I \rightarrow \mathbb{R}$ eine Funktion und f sei in $\xi \in I$ differenzierbar. Ist $f'(\xi) > 0$, so gibt es eine Umgebung U von ξ , so dass für alle $x \in U \cap I$ gilt:

$$f(x) < f(\xi) \text{ falls } x < \xi \quad \text{und} \quad f(x) > f(\xi) \text{ falls } x > \xi.$$

Die umgekehrte Aussage gilt für $f'(\xi) < 0$.

BEWEIS

Da

$$f'(\xi) = \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi}$$

ist gibt es eine Umgebung U von ξ , so dass für alle $x \in (U \cap I) \setminus \{\xi\}$ gilt:

$$\frac{f(x) - f(\xi)}{x - \xi} > 0.$$

Daraus folgt die Behauptung. □

DEFINITION 5.1.3

Wir sagen, die reelle Funktion f besitzt auf $X \subseteq \mathbb{R}$ an der Stelle $\xi \in X$ ein lokales Maximum bzw. ein lokales Minimum, wenn es eine ε -Umgebung U von ξ gibt, so dass für alle $x \in U \cap X$ stets $f(x) \leq f(\xi)$ gilt bzw. $f(x) \geq f(\xi)$. Lokale Maxima oder lokale Minima heißen auch lokale Extrema. Ein lokales Maximum kann sehr wohl kleiner sein als das globale Maximum $\max_{x \in X} f(x)$.

DEFINITION 5.1.4

Es sei $X \subseteq \mathbb{R}$ beliebig. Ein $\xi \in X$ heißt innerer Punkt von X , falls es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(\xi) \subseteq X$.

SATZ 5.1.3

Es sei $\xi \in X$ ein innerer Punkt und $f : X \rightarrow \mathbb{R}$ besitze in ξ ein lokales Extremum und sei überdies in ξ differenzierbar. Dann ist notwendig $f'(\xi) = 0$.

BEWEIS

f besitze in ξ ein lokales Maximum. Dann ist

$$\frac{f(x) - f(\xi)}{x - \xi} \geq 0 \text{ oder } \leq 0,$$

je nachdem $x < \xi$ oder $x > \xi$ ist. Wegen

$$f'(\xi) = \lim_{x \rightarrow \xi} \frac{f(x) - f(\xi)}{x - \xi}$$

ist $f'(\xi) = 0$. □

Schließlich ist auch die Interpretation der Differenzierbarkeit als lineare Approximation/Näherung von Bedeutung. Ihre volle Bedeutung wird erst in der Analysis mehrerer Variablen sichtbar werden.

SATZ 5.1.4

Genau dann ist die Funktion $f : I \rightarrow \mathbb{R}$ im Punkt $\xi \in I$ differenzierbar, wenn f in der Form

$$(*) \quad f(\xi + h) = f(\xi) + ah + r(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0$$

dargestellt werden kann. In diesem Fall ist $a = f'(\xi)$.

BEWEIS

Es sei $a \in \mathbb{R}$ gegeben, so dass (*) gilt. Dann ist

$$\lim_{h \rightarrow 0} \frac{f(\xi + h) - f(\xi)}{h} = a + \lim_{h \rightarrow 0} \frac{r(h)}{h} = a.$$

Also existiert $f'(\xi)$ und ist gleich a . Es sei umgekehrt f differenzierbar in ξ . Wir definieren $r(h)$ durch $f(\xi + h) - f(\xi) = ah + r(h)$, wobei $a = f'(\xi)$ ist. Dann gilt

$$\lim_{h \rightarrow 0} \frac{r(h)}{h} = \lim_{h \rightarrow 0} \frac{f(\xi + h) - f(\xi)}{h} - a = 0.$$

□

DEFINITION 5.1.5

Ist $f : I \rightarrow \mathbb{R}$ differenzierbar in $\xi \in I$, so nennt man den Graph der Funktion

$$L_\xi(x) = f(\xi) + f'(\xi) \cdot (x - \xi)$$

die Tangente der Kurve $y = f(x)$ im Punkt $x = \xi$ (offenbar ist die Tangente eine Gerade).

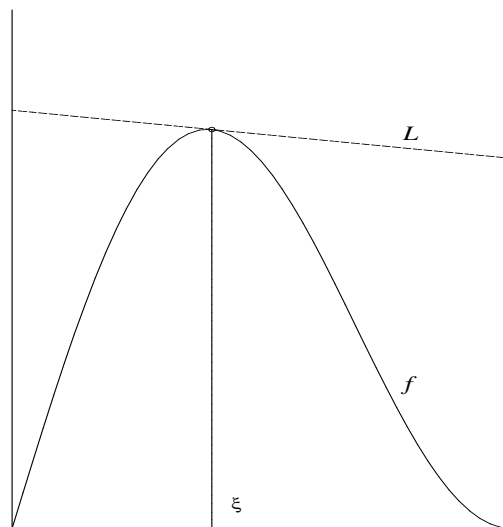


ABBILDUNG 5.1. Tangente an einer Funktion

BEMERKUNG 5.1.1

Satz 5.1.4 sagt also, dass der Graph einer in ξ differenzierbaren Funktion sehr gut durch eine Gerade durch ξ angenähert werden kann.

5.2. Differentiationsregeln

SATZ 5.2.1 (Rechenregeln für die Ableitung)

Die Funktionen f und g seien auf dem Intervall I definiert und in $\xi \in I$ differenzierbar. Dann sind auch $f + g$, $f - g$, αf und f/g (für $g(\xi) \neq 0$) in ξ differenzierbar, und die Ableitungen werden durch die folgenden Formeln gegeben, in denen als Argument ξ einzutragen ist:

$$(f + g)' = f' + g' \quad , \quad (f - g)' = f' - g' \quad , \quad (\alpha f)' = \alpha(f') \quad , \quad (fg)' = f'g + fg' \quad , \quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} .$$

BEWEIS

Aus Satz 4.5.10 (Rechenregeln für Grenzwerte) und dem Stetigkeitssatz 5.1.1 folgt für $x \rightarrow \xi$:

$$\frac{(f(x) \pm g(x)) - (f(\xi) \pm g(\xi))}{x - \xi} = \frac{f(x) - f(\xi)}{x - \xi} \pm \frac{g(x) - g(\xi)}{x - \xi} \longrightarrow f'(\xi) \pm g'(\xi) .$$

Ebenso ist

$$\frac{\alpha f(x) - \alpha f(\xi)}{x - \xi} = \alpha \cdot \frac{f(x) - f(\xi)}{x - \xi} \longrightarrow \alpha \cdot f'(\xi)$$

sowie

$$\frac{f(x)g(x) - f(\xi)g(\xi)}{x - \xi} = f(x)\frac{g(x) - g(\xi)}{x - \xi} + g(\xi)\frac{f(x) - f(\xi)}{x - \xi} \longrightarrow f(\xi)g'(\xi) + g(\xi)f'(\xi) ,$$

$$\frac{\frac{f(x)}{g(x)} - \frac{f(\xi)}{g(\xi)}}{x - \xi} = \frac{1}{g(x)g(\xi)} \cdot \left[g(\xi)\frac{f(x) - f(\xi)}{x - \xi} - f(\xi)\frac{g(x) - g(\xi)}{x - \xi} \right] \longrightarrow \frac{g(\xi)f'(\xi) - f(\xi)g'(\xi)}{g(\xi)^2} .$$

□

BEMERKUNG 5.2.1

Da die Ableitung der konstanten Funktionen $f(x) = c$ überall verschwindet, folgt aus der Quotientenregel von Satz 5.2.1 die Reziprokenregel

$$\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$$

falls g in ξ differenzierbar und $g(\xi) \neq 0$ ist.

SATZ 5.2.2 (Kettenregel)

Es seien $u : I_u \rightarrow \mathbb{R}$ sowie $f : I_f \rightarrow \mathbb{R}$ Funktionen, wobei I_u und I_f Intervalle sind und $I_f \supseteq u(I_u)$ gilt. Die Funktion u sei in ξ , und f in $u(\xi)$ differenzierbar. Dann ist $f \circ u$ in ξ differenzierbar, und es gilt

$$(f \circ u)'(\xi) = f'(u(\xi)) \cdot u'(\xi) .$$

BEWEIS

Wir setzen $\eta = u(\xi)$ und schreiben gemäß Satz 5.1.4:

$$u(\xi + h) - u(\xi) = u'(\xi)h + r_1(h) \quad , \quad \frac{r_1(h)}{h} =: \varphi_1(h) \rightarrow 0 \quad (h \rightarrow 0) ,$$

$$f(\eta + k) - f(\eta) = f'(\eta)k + r_2(k) \quad , \quad \frac{r_2(k)}{k} =: \varphi_2(k) \rightarrow 0 \quad (k \rightarrow 0) .$$

Aus diesen Darstellungen folgt

$$(f \circ u)(\xi + h) - (f \circ u)(\xi) = f(u(\xi + h)) - f(u(\xi)) = f'(\eta)[u(\xi + h) - u(\xi)] + r_2(u(\xi + h) - u(\xi))$$

$$= f'(\eta)[u'(\xi)h + r_1(h)] + r_2(u'(\xi)h + r_1(h)) = f'(\eta)u'(\xi)h + r(h) \text{ mit}$$

$$r(h) := f'(\eta)r_1(h) + r_2(u'(\xi)h + r_1(h)) = f'(\eta)h\varrho_1(h) + [u'(\xi)h + r_1(h)]\varrho_2(u'(\xi)h + r_1(h)).$$

Offenbar strebt $\frac{r(h)}{h} \rightarrow 0$ für $h \rightarrow 0$. Nach Satz 5.1.4 folgt die Behauptung. \square

SATZ 5.2.3 (Satz über die Umkehrfunktion)

Die Funktion f sei monoton und stetig auf dem Intervall I , so dass die Umkehrfunktion f^{-1} auf dem Intervall $f(I)$ existiert. Ist in $\xi \in I$ die Ableitung $f'(\xi)$ vorhanden und nicht Null, so kann f^{-1} in $\eta = f(\xi)$ differenziert werden, und es gilt:

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{f'(f^{-1}(\eta))}.$$

BEWEIS

Ist (y_n) eine beliebige Folge aus $f(I)$, die gegen η strebt, und deren Glieder alle von η verschieden sind, so liegen die $x_n = f^{-1}(y_n)$ alle in I und sind von ξ verschieden. Wegen der Stetigkeit von f^{-1} strebt $x_n \rightarrow f^{-1}(\eta) = \xi$. Also konvergiert

$$\frac{f^{-1}(y_n) - f^{-1}(\eta)}{y_n - \eta} = \frac{x_n - \xi}{f(x_n) - f(\xi)} = \frac{1}{\frac{f(x_n) - f(\xi)}{x_n - \xi}} \rightarrow \frac{1}{f'(\xi)}.$$

\square

5.3. Differentiation elementarer Funktionen

SATZ 5.3.1

Es ist

$$(x^k)' = k \cdot x^{k-1} \begin{cases} \forall x & \text{falls } k = 1, 2, \dots \\ \forall x \neq 0 & \text{falls } k = -1, -2, \dots \end{cases}.$$

BEWEIS

Es sei $f(x) = x$. Dann ist

$$f'(\xi) = \lim_{h \rightarrow 0} \frac{\xi + h - \xi}{h} = 1.$$

Damit ist die Behauptung für $k = 1$ gezeigt. Mittels der Produktregel von Satz 5.2.1 folgt sie für alle $k \in \mathbb{N}$ durch Induktion nach k :

$$(x^{k+1})' = (x \cdot x^k)' = x(x^k)' + x'(x^k) \stackrel{\text{IH}}{=} x \cdot kx^{k-1} + 1 \cdot x^k = (k+1) \cdot x^k.$$

Für negative k folgt die Behauptung durch die Quotientenregel. \square

Aus Satz 5.3.1 ergibt sich sofort die Differentiation für Polynome:

SATZ 5.3.2

Es sei $P: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto P(x) = a_0 + a_1x + \dots + a_nx^n$ ein Polynom, dann ist $P'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1}$.

SATZ 5.3.3 (Wurzelfunktion)

Es sei $k \in \mathbb{N}$. Dann ist

$$\left(x^{\frac{1}{k}}\right)' = \frac{1}{k}x^{\frac{1}{k}-1} \quad \forall x \in (0, \infty).$$

BEWEIS

Wir setzen $f : (0, \infty) \rightarrow (0, \infty)$, $y \mapsto y^k$. Dann ist

$$f^{-1} : \begin{cases} (0, \infty) & \rightarrow & (0, \infty) \\ x & \mapsto & \sqrt[k]{x} \end{cases}.$$

Nach Satz 5.2.3 gilt mit $\eta = f(\xi) = \xi^k$ die Gleichung

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{k\xi^{k-1}} = \frac{1}{k\eta^{1-\frac{1}{k}}} = \frac{1}{k} \cdot \eta^{\frac{1}{k}-1}.$$

□

5.4. Mittelwertsätze

SATZ 5.4.1

Es seien die beiden Funktionen $f, g : I \rightarrow \mathbb{R}$ stetig auf dem kompakten Intervall $I = [a, b]$ und differenzierbar auf dem offenen Intervall (a, b) . Dann gilt:

(i) **Satz von Rolle**

Ist $f(a) = f(b)$, so existiert ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.

(ii) **1. Mittelwertsatz der Differentialrechnung**

Es existiert ein $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

(iii) **2. Mittelwertsatz der Differentialrechnung**

Ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so existiert $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

BEMERKUNG 5.4.1

Der 1. Mittelwertsatz hat folgende geometrische Bedeutung: Betrachtet man den Graphen der Funktion f , so ist die Tangente an diesen Graphen im Punkt $(\xi, f(\xi))$ parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$.

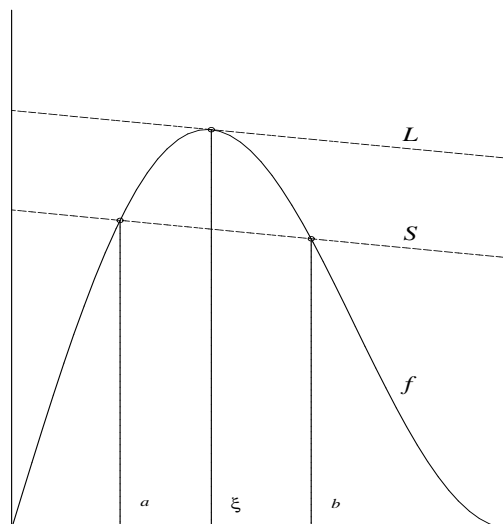


ABBILDUNG 5.2. Der Mittelwertsatz

BEWEIS

Zu (i): Fall 1: $f(x) = f(a)$ für alle $x \in [a, b]$. Dann ist $f'(x) = 0$ für alle $x \in [a, b]$ und jedes $\xi \in (a, b)$ funktioniert. Fall 2: $f(x) \neq f(a)$ für mindestens ein $x \in [a, b]$. Da f stetig ist, besitzt f auf dem kompakten Intervall I ein Maximum und ein Minimum. Mindestens eine dieser Extremstellen ξ liegt im Inneren (a, b) von $[a, b]$. Nach Satz 5.1.3 ist $f'(\xi) = 0$.

Zu (ii): Wir wenden den Satz von Rolle an auf die Hilfsfunktion

$$h : [a, b] \rightarrow \mathbb{R}, x \mapsto (b - a) \cdot (f(x) - f(a)) - (x - a)(f(b) - f(a)).$$

Es folgt die Existenz eines $\xi \in (a, b)$ mit $0 = (b - a)f'(\xi) - (f(b) - f(a))$.

Zu (iii): Wir wenden wieder den Satz von Rolle an auf

$$h : [a, b] \rightarrow \mathbb{R}, x \mapsto (g(b) - g(a)) \cdot (f(x) - f(a)) - (g(x) - g(a)) \cdot (f(b) - f(a)).$$

Es folgt die Existenz eines $\xi \in (a, b)$ mit $0 = (g(b) - g(a))f'(\xi) - (f(b) - f(a))g'(\xi)$. Hieraus folgt die Behauptung durch Division, da wegen $g'(x) \neq 0$ auf (a, b) der erste Mittelwertsatz $g(a) \neq g(b)$ liefert. \square

SATZ 5.4.2

Es sei $f : I \rightarrow \mathbb{R}$ stetig auf $I = [a, b]$ und differenzierbar in (a, b) . Dann gilt:

- (i) $f'(x) = 0$ für alle $x \in (a, b) \Leftrightarrow f$ ist konstant auf I .
- (ii) $f'(x) \geq 0$ für alle $x \in (a, b) \Leftrightarrow f$ ist monoton wachsend auf I .
- (iii) $f'(x) > 0$ für alle $x \in (a, b) \Rightarrow f$ ist streng monoton wachsend auf I .

BEWEIS

Zu (i): Wir fixieren $x \in (a, b)$ und wenden auf die Funktion f und auf das Intervall $[a, x]$ den 1. Mittelwertsatz an. Er liefert ein ξ mit

$$\frac{f(x) - f(a)}{x - a} = f'(\xi) = 0 \text{ also } f(x) = f(a).$$

Die Umkehrung der Aussage folgt aus den Ableitungsregeln.

Zu (ii): \Rightarrow : Für $x < y$ gilt

$$f(y) - f(x) = \frac{f(y) - f(x)}{y - x}(y - x) = f'(\xi)(y - x) \geq 0.$$

\Leftarrow : Da f wächst, ist der Differenzenquotient $\frac{f(x+h) - f(x)}{h}$ stets nichtnegativ, damit auch dessen Grenzwert $f'(x)$.

Zu (iii): Aus $x < y$ folgt

$$f(y) - f(x) = \frac{f(y) - f(x)}{y - x}(y - x) = f'(\xi)(y - x) > 0.$$

\square

BEMERKUNG 5.4.2

Die Umkehrung von (iii) ist nicht richtig: die Funktion $f(x) = x^3$ ist auf \mathbb{R} streng monoton wachsend, aber es gilt $f'(0) = 0$.

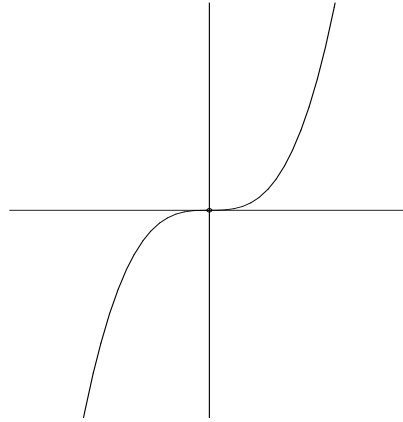


ABBILDUNG 5.3. $f(x) = x^3$

5.5. Grenzfunktionen und Potenzreihen

Wichtige Funktionen der Analysis (wie die Exponentialfunktion) können als Grenzfunktionen definiert werden. Um Aussagen über Stetigkeit und Ableitungen solcher Funktionen zu gewinnen, braucht man daher Sätze über Stetigkeit und Ableitungen von Grenzfunktionen. Den Fall der Stetigkeit haben wir in Satz 4.6.1 behandelt. Entscheidend für die Stetigkeit der Grenzfunktion f einer Funktionenfolge (f_n) war die gleichmäßige Konvergenz von (f_n) . Für die Existenz und Berechnung der Ableitung einer Grenzfunktion ist die gleichmäßige Konvergenz der Ableitungen (f'_n) entscheidend. Als Vorbereitung benötigen wir

SATZ 5.5.1 (Vertauschung von Grenzwertübergängen)

Es sei $f_n : I \rightarrow \mathbb{R}$ eine auf einem Intervall $I \subseteq \mathbb{R}$ definierte Folge von Funktionen (nicht notwendig stetig), die gleichmäßig gegen $f : I \rightarrow \mathbb{R}$ strebe. Zudem sei $x_0 \in I$ fest, so dass für jedes n der Grenzwert $f_n(x)$ für $x \rightarrow x_0$ existiert. Dann gilt

$$\lim_{x \rightarrow x_0} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = \lim_{n \rightarrow \infty} \left(\lim_{x \rightarrow x_0} f_n(x) \right).$$

BEWEIS

Es sei a_n jeweils der Grenzwert von $f_n(x)$ für $x \rightarrow x_0$ und $\varepsilon > 0$ beliebig. Es gibt nach dem Cauchy-kriterium ein $n_0 = n_0(\varepsilon)$, so dass $|f_m(x) - f_n(x)| < \varepsilon$ gilt für alle $m, n > n_0$ und alle $x \in I$ (da die f_n gleichmäßig gegen f konvergieren ist n_0 nicht von x abhängig). Wird nun $x \in I \setminus \{x_0\}$ nahe genug bei x_0 gewählt, so folgt $|a_m - a_n| < \varepsilon$ für alle $m, n > n_0$, also ist (a_n) eine Cauchyfolge mit einem Grenzwert $a = \lim a_n$. Aus der Dreiecksungleichung folgt

$$|f(x) - a| \leq |f(x) - f_m(x)| + |f_m(x) - a_m| + |a_m - a|.$$

Wir wählen m so groß, dass $|f(x) - f_m(x)| < \frac{1}{3}\varepsilon$ ist für alle $x \in I$ und gleichzeitig $|a_m - a| < \frac{1}{3}\varepsilon$. Dazu sei $\delta > 0$ genügend klein, so dass $|f_m(x) - a_m| < \frac{1}{3}\varepsilon$ ist für alle $x \in I$ mit $0 < |x - x_0| < \delta$. Dann ist für diese x nach der obigen Ungleichung $|f(x) - a| \leq \varepsilon$. Das bedeutet $f(x) \rightarrow a$ für $x \rightarrow x_0$, und daraus folgt die Behauptung:

$$\lim_{x \rightarrow x_0} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = \lim_{x \rightarrow x_0} f(x) = a = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(\lim_{x \rightarrow x_0} f_n(x) \right).$$

□

SATZ 5.5.2

Es sei $I = [a, b]$ ein reelles Intervall und (f_n) für $n = 0 \dots \infty$ eine Folge von auf I differenzierbaren Funktionen $f_n : I \rightarrow \mathbb{R}$. Die Folge der Ableitungen (f'_n) konvergiere auf I gleichmäßig und (f_n) konvergiere in wenigstens einem Punkt $x^* \in I$. Dann konvergiert (f_n) auf I gleichmäßig und $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ ist auf I differenzierbar:

$$f'(x) = \left(\lim_{n \rightarrow \infty} f'_n \right) (x).$$

Also gilt unter diesen Voraussetzungen kurz ausgedrückt $(\lim f_n)' = \lim(f'_n)$.

BEWEIS

Da (f_n) in x^* konvergiert, gibt es nach dem Cauchy Kriterium zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$, so dass $|f_m(x^*) - f_n(x^*)| < \frac{1}{2}\varepsilon$ ist für alle $m, n > n_0$. Da die Folge der Ableitungen sogar gleichmäßig konvergiert, kann n_0 groß genug gewählt werden, so dass zusätzlich $|f'_m(x) - f'_n(x)| < \frac{\varepsilon}{2(b-a)}$ ist für alle $x \in I$. Daraus folgt durch Anwendung des 1. Mittelwertsatzes auf $f_m - f_n$

$$(*) \quad |(f_m(x) - f_n(x)) - (f_m(y) - f_n(y))| < \frac{\varepsilon}{2(b-a)} \cdot |x - y| \quad \forall m, n > n_0.$$

Durch Einsetzen von x^* für y kann der Abstand $|f_m(x^*) - f_n(x^*)|$ auf alle $x \in I$ übertragen werden:

$$|f_m(x) - f_n(x)| \leq |(f_m(x) - f_n(x)) - (f_m(x^*) - f_n(x^*))| + |f_m(x^*) - f_n(x^*)| < \frac{\varepsilon \cdot |x - x^*|}{2(b-a)} + \frac{1}{2}\varepsilon \leq \varepsilon.$$

Daraus folgt die gleichmäßige Konvergenz der Folge (f_n) auf I gegen eine Grenzfunktion $f : I \rightarrow \mathbb{R}$. Für $x_0 \in I$ und $x \neq x_0$ sei

$$Q_n(x) := \frac{f_n(x) - f_n(x_0)}{x - x_0} \quad \text{und} \quad Q(x) := \frac{f(x) - f(x_0)}{x - x_0}.$$

Wegen den Rechenregeln für Folgen strebt Q_n punktweise gegen Q für $n \rightarrow \infty$. Einsetzen von x_0 für y in $(*)$ ergibt

$$\begin{aligned} |(f_m(x) - f_n(x)) - (f_m(x_0) - f_n(x_0))| &< \frac{\varepsilon}{2(b-a)} \cdot |x - y| \\ \Rightarrow |Q_m(x) - Q_n(x)| &< \frac{\varepsilon}{2(b-a)} \quad \forall m, n > n_0. \end{aligned}$$

Offenbar ist die Wahl von x beliebig, daher ist die Konvergenz $Q_n \rightarrow Q$ sogar gleichmäßig auf I . Damit gilt

$$f'(x_0) = \lim_{x \rightarrow x_0} Q(x) = \lim_{x \rightarrow x_0} \left(\lim_{n \rightarrow \infty} Q_n(x) \right) \stackrel{\text{Satz 5.5.1}}{=} \lim_{n \rightarrow \infty} \left(\lim_{x \rightarrow x_0} Q_n(x) \right) = \lim_{n \rightarrow \infty} f'_n(x_0).$$

□

DEFINITION 5.5.1

Es sei (a_n) für $n \geq 0$ eine Folge reeller Zahlen und $x_0 \in \mathbb{R}$. Unter einer Potenzreihe um den Punkt x_0 versteht man die unendliche Reihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

Die Menge

$$\left\{ x \in \mathbb{R} \mid \sum_{n=0}^{\infty} a_n (x - x_0)^n \text{ konvergiert} \right\}$$

heißt der Konvergenzbereich der Potenzreihe.

Bei der Bestimmung des Konvergenzbereichs dürfen wir o. B. d. A. $x_0 = 0$ annehmen. Um das einzusehen verende man die Substitution $w := x - x_0$ und schreibe anschließend wieder x an Stelle von w . Natürlich konvergiert die Reihe

$$(*) \quad \sum_{n=0}^{\infty} a_n w^n$$

trivialerweise für $w = 0$. Es kann sein, dass die Reihe $(*)$ nur für $w = 0$ konvergiert:

BEISPIEL 5.5.1

Die Potenzreihe

$$\sum_{n=0}^{\infty} n! w^n$$

divergiert nach dem Quotientenkriterium für alle $w \neq 0$, denn es ist

$$\left| \frac{(n+1)! w^{k+1}}{n! w^n} \right| = |(n+1)w| \longrightarrow \infty.$$

Andererseits zeigt die Reihe für die Exponentialfunktion, dass eine Reihe der Form $(*)$ auch für alle $w \in \mathbb{R}$ konvergieren kann. Um den Konvergenzbereich einer Reihe genauer zu studieren, benötigen wir das folgende

LEMMA 5.5.3

Es sei (a_n) eine Zahlenfolge:

- (i) Konvergiert die Reihe $\sum a_n w^n$ für ein $w^* \neq 0$, so gilt: Für jedes feste $0 < r < |w^*|$ konvergiert die Reihe $\sum a_n w^n$ gleichmäßig auf $\{w \in \mathbb{R} \mid |w| < r\}$. Dies gilt auch dann, wenn man in der Reihe die Absolutbeträge der Koeffizienten nimmt, also für

$$\sum_{n=0}^{\infty} |a_n| w^n$$

und sogar für die Reihe

$$\sum_{n=0}^{\infty} |a_n w^n|.$$

Diese Reihen konvergieren für jedes w mit $|w| < |w^*|$.

- (ii) Divergiert die Reihe $\sum a_n w^n$ für ein \tilde{w} , so divergiert sie auch für jedes $w \in \mathbb{R}$ mit $|w| > |\tilde{w}|$.

BEWEIS

Zu (i): Da $(*)$ für w^* konvergiert, handelt es sich bei der Folge $a_n (w^*)^n$ um eine Nullfolge, die insbesondere beschränkt ist. Es gibt also ein reelles K mit $\forall n \in \mathbb{N}_0 : |a_n (w^*)^n| \leq K$. Für alle w mit $|w| < r$ gilt dann (man beachte $0 < q := \frac{r}{|w^*|} < 1$):

$$|a_n w^n| \leq |a_n r^n| = |a_n (w^*)^n| \cdot \left| \frac{r}{w^*} \right|^n \leq K q^n.$$

Deshalb folgt die gleichmäßige Konvergenz unmittelbar. Wendet man dies für $r := |w|$ mit $|w| < |w^*|$ an, so folgt die Konvergenz für w . Zu (ii): Würde die Reihe $(*)$ für ein w mit $|w| > |\tilde{w}|$ konvergieren, so nach (i) auch für \tilde{w} , ein Widerspruch. \square

SATZ 5.5.4

Konvergiert die Reihe

$$(*) \quad \sum_{n=0}^{\infty} a_n w^n$$

für ein $w^* \neq 0$ und divergiert sie für ein $\tilde{w} \in \mathbb{R}$, so existiert (genau) ein $R > 0$ derart, dass die Reihe $(*)$ für alle $|w| < R$ konvergiert, und für $|w| > R$ divergiert. Außerdem ist die Potenzreihe für jedes

beliebige aber feste r mit $0 < r < R$ gleichmäßig konvergent auf $\{w \in \mathbb{R} \mid |w| < r\}$. Dies gilt auch, wenn wir in der Reihe die Absolutbeträge nehmen.

BEWEIS

Nach dem vorigen Lemma kann man $R := \sup\{r > 0 \mid \sum a_n r^n \text{ konvergiert}\}$ setzen. □

DEFINITION 5.5.2

Wir nennen R den Konvergenzradius der Reihe (*). Wir setzen zusätzlich

$$R := \begin{cases} 0 & \text{falls (*) nur in } w = 0 \text{ konvergiert} \\ \infty & \text{falls (*) auf ganz } \mathbb{R} \text{ konvergiert} \end{cases} .$$

SATZ 5.5.5 (Cauchy-Hadamard)

Für den Konvergenzradius R der Potenzreihe (*) gilt

$$R = \left(\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \right)^{-1} \text{ wenn wir noch } \frac{1}{0} = \infty \text{ und } \frac{1}{\infty} = 0 \text{ setzen.}$$

OHNE BEWEIS □

LEMMA 5.5.6

Es sei $r \in \mathbb{R}$ beliebig, aber fest. Die Potenzreihe $\sum n^r x^n$ hat den Konvergenzradius $R = 1$.

BEWEIS

Dies folgt aus dem Quotientenkriterium wegen

$$\lim_{n \rightarrow \infty} \left| \frac{(n+1)^r x^{n+1}}{n^r x^n} \right| = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^r |x| = |x| .$$

Teil (i) des Kriteriums trifft also gerade für $|x| < 1$ zu, Teil (ii) für $|x| > 1$. □

SATZ 5.5.7 (Stetigkeit und Differenzierbarkeit von Potenzreihen)

Hat die Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

einen positiven Konvergenzradius R (oder ist $R = \infty$), so ist $f(x)$ für alle $|x| < R$ differenzierbar zum Wert

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1} .$$

Damit ist $f(x)$ auch stetig für $|x| < R$. Zudem hat die Potenzreihe

$$\sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}$$

ebenfalls den Konvergenzradius R .

BEWEIS

Wie oben nehmen wir $x_0 = 0$ an. Wir zeigen zunächst, dass die Reihe

$$\sum_{n=1}^{\infty} n a_n x^{n-1}$$

den Konvergenzradius R besitzt, dass sie also für jedes x mit $|x| < R$ konvergiert und für jedes x mit $|x| > R$ divergiert. Es sei $x \in \mathbb{R}$ beliebig (aber fest), und $|x| < R$. Wir wählen ein festes ϱ mit

$|x| < \varrho < R$. Dann konvergiert die Reihe $\sum a_n \varrho^n$, d. h. es gibt ein $K > 0$ derart, dass für alle $n \in \mathbb{N}_0$ gilt: $|a_n| \varrho^n \leq K$. Es folgt

$$|na_n x^{n-1}| \leq n \cdot |a_n| \cdot \varrho^{n-1} \cdot \left| \frac{x}{\varrho} \right|^{n-1} \leq K^* n \varrho^{n-1}$$

mit $q := \left| \frac{x}{\varrho} \right| < 1$ und $K^* = K \varrho^{-1}$. Da nach dem vorigen Lemma die Reihe

$$\sum_{n=1}^{\infty} n q^{n-1}$$

konvergiert, ist auf Grund des Majorantenkriteriums die Reihe

$$\sum_{n=1}^{\infty} na_n x^{n-1}$$

konvergent. Nun sei x^* beliebig mit $|x^*| > R$. Würde die Reihe

$$\sum_{n=1}^{\infty} na_n (x^*)^{n-1}$$

konvergieren, so wäre die Folge $|na_n (x^*)^{n-1}|$ beschränkt, also auch die Folge $|a_n (x^*)^n|$. Dann würde aber

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

für alle x mit $R < |x| < |x^*|$ konvergieren, im Widerspruch dazu, dass der Konvergenzradius von

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

gerade $R < |x|$ ist. Es sei nun x beliebig aber fest mit $|x| < R$. Wir wählen ein festes η mit $|x| < \eta < R$. Nach dem Obigen und Satz 5.5.4 sind

$$\sum_{n=0}^{\infty} a_n x^n \text{ und } \sum_{n=1}^{\infty} na_n x^{n-1}$$

gleichmäßig konvergent auf $\{x \in \mathbb{R} \mid |x| < \eta\}$. Nach Satz 5.5.2 ist dann

$$f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1}.$$

□

SATZ 5.5.8 (Ableitung von Exponentialfunktion und Logarithmus)

Es gilt:

- (i) $(e^x)' = e^x$ für $x \in \mathbb{R}$ sowie
- (ii) $(\log x)' = \frac{1}{x}$ für $x \in (0, \infty)$.

BEWEIS

Es ist

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \forall x \in \mathbb{R}.$$

Nach Satz 5.5.7 ist

$$(e^x)' = \sum_{n=0}^{\infty} \left(\frac{d}{dx} \frac{x^n}{n!} \right) = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{m=0}^{\infty} \frac{x^m}{m!} = e^x.$$

Nun setzen wir

$$f : \begin{cases} \mathbb{R} & \rightarrow (0, \infty) \\ y & \mapsto e^y \end{cases} .$$

Dann ist

$$f^{-1} : \begin{cases} (0, \infty) & \rightarrow \mathbb{R} \\ x & \mapsto \log x \end{cases} .$$

Es sei $\eta = e^\xi$. Nach Satz 5.2.3 gilt

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{e^\xi} = \frac{1}{\eta} .$$

□

SATZ 5.5.9 (Rechnen mit Potenzreihen)

Haben die Potenzreihen $\sum a_n(x-x_0)^n$ bzw. $\sum b_n(x-x_0)^n$ die positiven Konvergenzradien r_a bzw. r_b , so gilt für alle x mit $|x-x_0| < \min(r_a, r_b)$

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n \pm \sum_{n=0}^{\infty} b_n(x-x_0)^n = \sum_{n=0}^{\infty} (a_n \pm b_n)(x-x_0)^n$$

und

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n \cdot \sum_{n=0}^{\infty} b_n(x-x_0)^n = \sum_{n=0}^{\infty} (a_0b_n + a_1b_{n-1} + \dots + a_nb_0)(x-x_0)^n .$$

OHNE BEWEIS

□

Man kann fragen, ob eine Funktion auf mehr als eine Weise als Potenzreihe um einen Punkt dargestellt werden kann. Dies ist nicht der Fall, wie der folgende Satz zeigt:

SATZ 5.5.10 (Identitätssatz für Potenzreihen)

Es sei

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n \quad , \quad g(x) = \sum_{n=0}^{\infty} b_n(x-x_0)^n$$

und beide Reihen mögen das gemeinsame Konvergenzintervall K haben. Stimmen dann die Funktionen f und g auf einer Folge (x_n) überein, deren Glieder verschieden von x_0 sind, die aber gegen x_0 strebt, gilt also $f(x_k) = g(x_k)$ für $k = 1, 2, \dots$, so müssen beide Funktionen und beide Reihen vollständig identisch sein, also $f(x) = g(x)$ für alle $x \in K$ und $a_n = b_n$ für alle $n \in \mathbb{N}_0$.

OHNE BEWEIS

□

5.6. Die Formel von Taylor, Taylorreihen

SATZ 5.6.1

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ auf I n -mal differenzierbar. Es sei $x_0 \in I$. Dann gibt es genau ein Polynom T_n vom Grad kleiner oder gleich n , das die $(n+1)$ Bedingungen

$$T_n^{(k)}(x_0) = f^{(k)}(x_0) \quad k = 0 \dots n$$

erfüllt, nämlich

$$T_n(x) := T_n(x; f, x_0) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k .$$

BEWEIS

Man rechnet nach, dass T_n die angegebenen Bedingungen erfüllt:

$$T_n(x_0) = f(x_0) \quad , \quad T'_n(x_0) = \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} \cdot k \cdot (x_0 - x_0)^{k-1} = f'(x_0) \text{ usw. .}$$

Die Eindeutigkeit ist eine Übungsaufgabe. □

DEFINITION 5.6.1

Das Polynom des vorigen Satzes heißt Taylorpolynom der Ordnung n von f (um x_0).

SATZ 5.6.2 (Satz von Taylor, Restglied von Lagrange)

Die Funktion $f : I \rightarrow \mathbb{R}$ sei n -mal stetig differenzierbar auf $I = [a, b]$ und $(n+1)$ -mal differenzierbar auf (a, b) . Es sei $x_0 \in I$ beliebig aber fest vorgegeben. Dann gilt: zu jedem $x \in I$ gibt es (mindestens) ein $\xi = \xi(f, x, n, x_0) \in (x, x_0)$ (bzw. (x_0, x) falls $x_0 < x$) mit

$$f(x) = T_n(x) + R_n(x) \quad , \quad R_n(x) := \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} \text{ , also}$$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} .$$

BEWEIS

Wir dürfen $x \neq x_0$ voraussetzen, da die Behauptung für $x = x_0$ trivialerweise erfüllt ist. Für ein festes $x \neq x_0$ betrachten wir die Funktion

$$g : \begin{cases} I & \rightarrow \mathbb{R} \\ t & \mapsto g(t) \end{cases} \quad , \quad g(t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x - t)^k .$$

Es folgt $g(x) = f(x)$ und $g(x_0) = T_n(x)$, sowie

$$g'(t) = \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!} (x - t)^k - \sum_{k=1}^n \frac{f^{(k)}(t)}{(k-1)!} (x - t)^{k-1} .$$

In der 2. Summe ersetzen wir den Summationsindex k durch $p = k - 1$ und erhalten

$$g'(t) = \frac{f^{(n+1)}(t)}{n!} (x - t)^n .$$

Wir wenden den 2. Mittelwertsatz auf das Intervall von x_0 bis x und auf die beiden Funktionen $g(t)$ und $(x - t)^{n+1}$ an und erhalten

$$\frac{T_n(x) - f(x)}{(x - x_0)^{n+1}} = \frac{g(x_0) - g(x)}{(x - x_0)^{n+1} - (x - x)^{n+1}} = \frac{g'(\xi)}{(n+1)(x - \xi)^n} = \frac{f^{(n+1)}(\xi)}{(n+1)!} .$$

□

BEMERKUNG 5.6.1

Der Satz von Taylor ermöglicht in vielen Fällen gute Approximationen einer Funktion f durch Polynome. Die Gestalt des Restglieds

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

legt die Vermutung nahe, dass in vielen Fällen die Qualität der Approximation mit wachsendem n wächst, da (für x genügend nahe bei x_0) gilt

$$\lim_{x \rightarrow x_0} \frac{(x - x_0)^{n+1}}{(x - x_0)^{m+1}} = 0 \text{ falls } n > m .$$

Es ergibt sich die Frage: Wann ist $\lim_{n \rightarrow \infty} T_n(x) = f(x)$?

DEFINITION 5.6.2

Die Funktion $f : I \rightarrow \mathbb{R}$ sei beliebig oft in $I = [a, b]$ differenzierbar. Für $x_0 \in [a, b]$ heißt

$$T(x) := T(x; f, x_0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

die Taylorreihe von f um den Punkt x_0 .

Unmittelbar aus Satz 5.6.2 folgt

SATZ 5.6.3

Die Funktion $f : I \rightarrow \mathbb{R}$ sei in $I = [a, b]$ beliebig oft differenzierbar und $x_0 \in [a, b]$. Dann gilt:

$$f(x) = T(x; f, x_0) \Leftrightarrow \lim_{n \rightarrow \infty} R_n(x) = 0 \text{ (punktweise) .}$$

BEISPIEL 5.6.1

Die Funktion $f(x) = \log(1 + x)$ ist auf $I = (-1, \infty)$ beliebig oft differenzierbar, und es gilt für alle $k \in \mathbb{N}$:

$$f^{(k)}(x) = \frac{(-1)^{k+1}(k-1)!}{(1+x)^k},$$

wie man leicht durch vollständige Induktion beweist. Der Satz von Taylor besagt (für $x_0 = 0$):

$$\log(1+x) = \sum_{k=1}^n \frac{(-1)^{k+1}}{k} x^k + \frac{(-1)^n}{n+1} \left(\frac{x}{1+\delta x} \right)^{n+1}$$

mit einem geeigneten $\delta = \delta(n, x) \in (0, 1)$. Das Restglied $R_n(x)$ ist gegeben durch

$$R_n(x) = \frac{(-1)^n}{n+1} \cdot \left(\frac{x}{1+\delta x} \right)^{n+1}.$$

Für $-\frac{1}{2} \leq x \leq 1$ gilt

$$\left| \frac{x}{1+\delta x} \right| \leq 1, \text{ also } |R_n(x)| \leq \frac{1}{n+1} \xrightarrow{n \rightarrow \infty} 0 \text{ und somit}$$

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k.$$

Dieses Resultat gilt sogar für $-1 < x \leq 1$. Wir wollen dies auf andere Weise zeigen: Nach dem Quotientenkriterium hat die Potenzreihe

$$f(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$$

den Konvergenzradius 1. Also ist die Funktion

$$g(x) = \log(1+x) - \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$$

nach Satz 5.5.7 differenzierbar auf $I = (-1, 1)$ und

$$g'(x) = \frac{1}{1+x} - \sum_{k=1}^{\infty} (-1)^{k-1} x^{k-1} = \frac{1}{1+x} + \frac{-1}{1+x} = 0.$$

Somit ist nach Satz 5.5.10(i) die Funktion $g(x)$ tatsächlich eine Konstante, diese besitzt den Wert $g(0) = 1 - 1 = 0$.

5.7. Trigonometrische Funktionen und ihre Umkehrfunktionen

SATZ 5.7.1

Die Potenzreihen

$$\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \quad \text{und} \quad \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

sind für alle $x \in \mathbb{R}$ konvergent.

BEWEIS

Dies folgt aus dem Quotientenkriterium. □

DEFINITION 5.7.1

Wir definieren die trigonometrischen Funktionen Sinus und Kosinus durch

$$\sin : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & \sin(x) \end{cases} \quad \sin(x) := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \quad \text{und}$$

$$\cos : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & \cos(x) \end{cases} \quad \cos(x) := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} .$$

SATZ 5.7.2

Es gilt $(\sin x)' = \cos x$ und $(\cos x)' = -\sin x$.

BEWEIS

Dies folgt aus Satz 5.5.7 (Differentiation von Potenzreihen). □

SATZ 5.7.3

Es gilt $\sin^2 x + \cos^2 x = 1$ für alle $x \in \mathbb{R}$. Insbesondere ist $|\sin x| \leq 1$ und $|\cos x| \leq 1$.

BEWEIS

Es sei $f(x) = \sin^2 x + \cos^2 x$. Aus Satz 5.7.2 folgt nach der Kettenregel $f'(x) = 2 \sin x \cos x - 2 \sin x \cos x = 0$. Also ist $f(x)$ konstant, und damit $f(x) = f(0) = 1$. □

SATZ 5.7.4

Es gilt für alle $x \in \mathbb{R}$

- (i) $\sin(x+y) = \sin x \cos y + \cos x \sin y$ und
- (ii) $\cos(x+y) = \cos x \cos y - \sin x \sin y$.

BEWEIS

Für festes $y \in \mathbb{R}$ setzen wir $f(x) = \sin(x+y)$. Es gilt

$$f'(x) = \cos(x+y) \quad , \quad f''(x) = -\sin(x+y) \quad , \quad f^{(3)}(x) = -\cos(x+y) \quad , \quad f^{(4)}(x) = \sin(x+y) = f(x)$$

und daraus allgemein per Induktion für alle $m \in \mathbb{N}_0$

$$\begin{array}{ll} f^{(4m+0)}(x) = \sin(x+y) & f^{(4m+0)}(0) = \sin y \\ f^{(4m+1)}(x) = \cos(x+y) & f^{(4m+1)}(0) = \cos y \\ f^{(4m+2)}(x) = -\sin(x+y) & f^{(4m+2)}(0) = -\sin y \\ f^{(4m+3)}(x) = -\cos(x+y) & f^{(4m+3)}(0) = -\cos y \end{array} .$$

Damit erhalten wir für $f(x)$ die Taylorreihe (um den Nullpunkt)

$$T(x) = \sin y \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots \right) + \cos y \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots \right) = (\cos x)(\sin y) + (\sin x)(\cos y)$$

mit dem Restglied

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1},$$

das wegen

$$f^{(n+1)}(\xi) \in \{\pm \cos(\xi + y), \pm \sin(\xi + y)\}$$

und $\sin^2(\xi + y) + \cos^2(\xi + y) = 1$ abgeschätzt werden kann:

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Somit ist $R_n(x) \rightarrow 0$ für $n \rightarrow \infty$. Nach Satz 5.6.3 ist damit $f(x) = T(x)$, womit (i) gezeigt ist. Der Teil (ii) folgt durch eine ähnliche Überlegung. \square

SATZ 5.7.5

Es gilt

- (i) $\cos x > 0$ für $x \in [0, 1]$ und
- (ii) es gibt ein $\xi \in (1, 2)$ mit $\cos(\xi) = 0$.

BEWEIS

Zu (i): es ist $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots \geq 1 - \frac{x^2}{2}$ für $x \in [0, 1]$, also $\cos x > 0$ für $x \in [0, 1]$. Zu (ii): Es gilt

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \left(1 - \frac{x^2}{7 \cdot 8}\right) - \frac{x^{10}}{10!} \left(1 - \frac{x^2}{11 \cdot 12}\right) - \dots$$

Für $0 \leq x \leq 2$ sind die großen Klammern positiv, deshalb gilt dann

$$\cos x < 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \Rightarrow \cos 2 < 1 - \frac{4}{2} + \frac{16}{24} = -\frac{1}{3} < 0.$$

Nach dem Zwischenwertsatz 4.3.1 gibt es ein $\xi \in (1, 2)$ mit $\cos(\xi) = 0$. \square

DEFINITION 5.7.2 (Die Kreiszahl)

Wir setzen $\pi := 2 \cdot \min\{\xi > 0 \mid \cos \xi = 0\}$.

BEMERKUNG 5.7.1

Die Zahl $\xi = \frac{\pi}{2}$ ist also die kleinste positive Nullstelle der Kosinusfunktion. Nach Satz 5.7.5 ist $2 < \pi < 4$. Eine numerische Rechnung liefert $\pi = 3,14159265\dots$

SATZ 5.7.6

Es gilt

- (i) \sin und \cos sind 2π -periodisch: $\sin(x + 2\pi) = \sin x$ und $\cos(x + 2\pi) = \cos x$ für alle $x \in \mathbb{R}$.
- (ii) $\sin(-x) = -\sin x$ und $\cos(-x) = \cos x$ für alle $x \in \mathbb{R}$.
- (iii) $\sin(x + \frac{\pi}{2}) = \cos x$, $\sin(x + \pi) = -\sin x$, $\cos(x + \frac{\pi}{2}) = -\sin x$, $\cos(x + \pi) = -\cos x$.
- (iv)

$$\sin x = 0 \Leftrightarrow x = k\pi \quad (k \in \mathbb{Z})$$

$$\sin x = 1 \Leftrightarrow x = \frac{1}{2}\pi + 2k\pi \quad (k \in \mathbb{Z})$$

$$\sin x = -1 \Leftrightarrow x = \frac{3}{2}\pi + 2k\pi \quad (k \in \mathbb{Z})$$

$$\cos x = 0 \Leftrightarrow x = \frac{1}{2}\pi + k\pi \quad (k \in \mathbb{Z})$$

$$\cos x = 1 \Leftrightarrow x = 2k\pi \quad (k \in \mathbb{Z})$$

$$\cos x = -1 \Leftrightarrow x = (2k + 1)\pi \quad (k \in \mathbb{Z})$$

- (v) \sin ist streng wachsend auf $[0, \frac{\pi}{2}]$, streng abnehmend auf $[\frac{\pi}{2}, \frac{3}{2}\pi]$, streng wachsend auf $[\frac{3}{2}\pi, 2\pi]$.

BEWEIS

Für $x \in (0, \frac{\pi}{2})$ ist $(\sin x)' = \cos x > 0$ nach Definition von $\frac{\pi}{2}$ als kleinste positive Nullstelle des Kosinus. Nach Satz 5.4.2(iii) ist der Sinus damit streng wachsend auf $[0, \frac{\pi}{2}]$, also $\sin \frac{\pi}{2} > 0$, da $\sin 0 = 0$. Aus $\sin^2 \frac{\pi}{2} + \cos^2 \frac{\pi}{2} = 1$ folgt $\sin \frac{\pi}{2} = 1$. Nach Satz 5.7.4 ist $\sin(x + \frac{\pi}{2}) = \sin x \cos \frac{\pi}{2} + \cos x \sin \frac{\pi}{2} = \cos x$, also gilt

$$\sin \pi = \cos \frac{\pi}{2} = 0, \quad \cos \pi = (\cos \frac{\pi}{2})^2 - (\sin \frac{\pi}{2})^2 = -1.$$

Ferner gilt $\sin(x + \pi) = \sin x \cos \pi - \cos x \sin \pi = -\sin x$ sowie $\sin(2\pi) = -\sin \pi = 0$ und $\cos(2\pi) = \cos^2 \pi - \sin^2 \pi = 1$. Zudem ist $\sin(x + 2\pi) = \sin x \cos(2\pi) + \cos x \sin(2\pi) = \sin x$ und $\cos(x + 2\pi) = \cos x \cos(2\pi) - \sin x \sin(2\pi) = \cos x$. Die übrigen Tatsachen lassen sich aus dem bisher Bewiesenen einfach herleiten. \square

Mittels der Funktionen \sin und \cos lassen sich noch weitere trigonometrische Funktionen definieren:

DEFINITION 5.7.3

Für $x \neq (k + \frac{1}{2})\pi$ ($k \in \mathbb{Z}$) setzen wir $\tan x := \frac{\sin x}{\cos x}$ (Tangens) und für $x \neq k\pi$ ($k \in \mathbb{Z}$) sei $\cot x := \frac{\cos x}{\sin x} = \tan^{-1} x$ (Kotangens).

SATZ 5.7.7

Es gilt:

- (i) Die Funktionen \tan und \cot haben die Periode π .
- (ii) \tan ist streng wachsend auf $(-\frac{\pi}{2}, \frac{\pi}{2})$.
- (iii) $\tan(-x) = -\tan x$ und $\cot(-x) = -\cot x$.
- (iv) $\tan(x + \frac{\pi}{2}) = -\cot x$ und $\cot(x + \frac{\pi}{2}) = -\tan x$.
- (v) $(\tan x)' = \frac{1}{\cos^2 x} = 1 + \tan^2 x$, ebenso $(\cot x)' = -\frac{1}{\sin^2 x} = -(1 + \cot^2)x$.
- (vi) $\lim_{x \rightarrow \frac{\pi}{2}^+} \tan x = -\infty$ und $\lim_{x \rightarrow \frac{\pi}{2}^-} \tan x = \infty$.

BEWEIS

Diese Eigenschaften folgen leicht aus der Definition und aus den Eigenschaften von \sin und \cos . \square

Wir behandeln nun die Umkehrfunktionen der trigonometrischen Funktionen.

DEFINITION 5.7.4

Die Umkehrfunktion von

$$\sin : \begin{cases} [-\frac{\pi}{2}, \frac{\pi}{2}] & \rightarrow & [-1, 1] \\ x & \mapsto & \sin x \end{cases} \text{ wird mit } \arcsin : \begin{cases} [-1, 1] & \rightarrow & [-\frac{\pi}{2}, \frac{\pi}{2}] \\ x & \mapsto & \arcsin x \end{cases} \quad (\underline{\text{Arcussinus}})$$

bezeichnet. Die Umkehrfunktion von

$$\cos : \begin{cases} [0, \pi] & \rightarrow & [-1, 1] \\ x & \mapsto & \cos x \end{cases} \text{ wird mit } \arccos : \begin{cases} [-1, 1] & \rightarrow & [0, \pi] \\ x & \mapsto & \arccos x \end{cases} \quad (\underline{\text{Arcuskosinus}})$$

bezeichnet. Die Umkehrfunktion von

$$\tan : \begin{cases} (-\frac{\pi}{2}, \frac{\pi}{2}) & \rightarrow & \mathbb{R} \\ x & \mapsto & \tan x \end{cases} \text{ wird mit } \arctan : \begin{cases} \mathbb{R} & \rightarrow & (-\frac{\pi}{2}, \frac{\pi}{2}) \\ x & \mapsto & \arctan x \end{cases} \quad (\underline{\text{Arcustangens}})$$

bezeichnet. Die Umkehrfunktion von

$$\cot : \begin{cases} (0, \pi) & \rightarrow & \mathbb{R} \\ x & \mapsto & \cot x \end{cases} \text{ wird mit } \operatorname{arccot} : \begin{cases} \mathbb{R} & \rightarrow & (0, \pi) \\ x & \mapsto & \operatorname{arccot} x \end{cases} \quad (\underline{\text{Arcuskotangens}})$$

bezeichnet.

SATZ 5.7.8

Die Funktionen \arcsin , \arccos , \arctan und arccot sind auf ihrem Definitionsbereich stetig und im Inneren differenzierbar mit

- (i) $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$, $x \in (-1, 1)$
- (ii) $(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$, $x \in (-1, 1)$
- (iii) $(\arctan x)' = \frac{1}{1+x^2}$, $x \in \mathbb{R}$
- (iv) $(\operatorname{arccot} x)' = -\frac{1}{1+x^2}$, $x \in \mathbb{R}$.

BEWEIS

Die Existenz und Stetigkeit der Umkehrfunktionen folgt aus Satz 4.4.1 und der strengen Monotonie von \sin , \cos , \tan und \cot auf den betrachteten Intervallen. Zu (i): wir setzen

$$f : \begin{cases} (-\frac{\pi}{2}, \frac{\pi}{2}) & \rightarrow (-1, 1) \\ y & \mapsto \sin y \end{cases} .$$

Dann ist

$$f^{-1} : \begin{cases} (-1, 1) & \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2}) \\ x & \mapsto \arcsin x \end{cases} .$$

Es sei $\eta = \sin \xi$, dann ist nach Satz 5.2.3

$$(f^{-1})'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{\cos \xi} = \frac{1}{\sqrt{1-\sin^2 \xi}} = \frac{1}{\sqrt{1-\eta^2}} .$$

Die übrigen Differentiationsregeln werden analog bewiesen. □

5.8. Komplexe Folgen und Potenzreihen, komplexe Exponentialfunktion

Manche Tatsachen über die trigonometrischen Funktionen Sinus und Kosinus, insbesondere die Additionstheoreme, sind besser zu verstehen, wenn sie im Zusammenhang mit der Exponentialfunktion betrachtet werden. Dieser Zusammenhang erschließt sich, wenn diese Funktionen für komplexwertige Argumente betrachtet werden. Dies erfordert die Betrachtung von Potenzreihen mit komplexen Argumenten und damit die Betrachtung komplexer Zahlenfolgen. Zunächst führen wir den Körper \mathbb{C} der komplexen Zahlen ein.

DEFINITION 5.8.1

Die Menge \mathbb{C} ist die Menge aller Paare reeller Zahlen: $\mathbb{C} = \{(x, y) \mid x, y \in \mathbb{R}\}$. Addition und Multiplikation sind wie folgt erklärt:

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

SATZ 5.8.1

Mit diesen Rechenoperationen bildet \mathbb{C} einen Körper mit Nullelement $(0, 0)$ und Einselement $(1, 0)$. Das multiplikative Inverse von $z = (x, y) \in \mathbb{C} \setminus \{(0, 0)\}$ ist gegeben durch

$$z^{-1} = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right) .$$

BEWEIS

(Übungsaufgabe) □

Die Darstellung von \mathbb{C} wird übersichtlicher durch die Einführung der imaginären Einheit $i := (0, 1)$ und durch die Identifizierung von $(x, 0)$ mit $x \in \mathbb{R}$, deren strenge Definition wir nicht geben werden. Durch diese Identifizierung wird \mathbb{R} zu einem Unterkörper von \mathbb{C} . Wir erhalten $\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}$. Addition und Multiplikation ergeben sich auf die folgende übersichtliche Weise: es gelten die Körperaxiome und die Regel $i^2 = -1$, damit gilt

$$(x_1 + iy_1) \cdot (x_2 + iy_2) \stackrel{D1/D2}{=} x_1x_2 + ix_1y_2 + ix_2y_1 + i^2y_1y_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1).$$

DEFINITION 5.8.2

Für $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ heißt x der Realteil von z und y der Imaginärteil (Bezeichnung: $x = \operatorname{Re}(z)$ und $y = \operatorname{Im}(z)$). Die Zahl $\bar{z} = x - iy$ heißt die zu z konjugierte komplexe Zahl. Die (reelle) Zahl $|z| := \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$ heißt der Betrag von z . Es gilt also $z \in \mathbb{R} \Leftrightarrow \operatorname{Im}(z) = 0$. Ein $z \in \mathbb{C}$ mit $\operatorname{Re}(z) = 0$ heißt rein-imaginär.

SATZ 5.8.2 (Betragseigenschaften)

Für $z_1, z_2, z \in \mathbb{C}$ gilt:

- (i) $|z_1 + z_2| \leq |z_1| + |z_2|$,
- (ii) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$,
- (iii) $|z| = |\bar{z}|$,
- (iv) $|z| = 0 \Leftrightarrow z = 0$.

OHNE BEWEIS

□

So wie der Körper \mathbb{R} durch die Zahlengerade, lässt sich der Körper \mathbb{C} durch die komplexe Zahlenebene geometrisch veranschaulichen. Sie enthält als Teilmengen die reelle Achse \mathbb{R} und die imaginäre Achse $\{ix \mid x \in \mathbb{R}\}$.

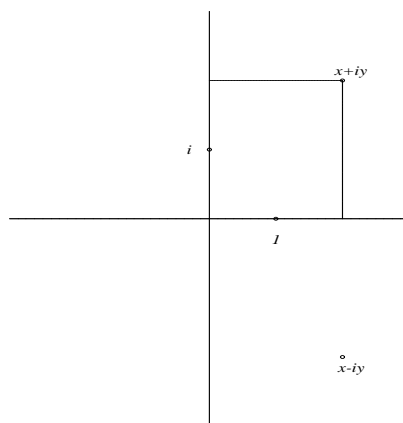


ABBILDUNG 5.4. Der Punkt $z = x + iy$ in der Zahlenebene

Die geometrische Bedeutung des Betrags z ergibt sich als Abstand des Punkts z vom Nullpunkt. \bar{z} ist das Spiegelbild von z bzgl. der reellen Achse.

Wir geben nun einen kurzen Überblick über die Theorie der komplexen Zahlenfolgen. Diese ist weitgehend analog der Theorie der reellen Zahlenfolgen. Der Grenzwert einer Folge komplexer Zahlen (a_n) wird wie im Reellen definiert: $\lim a_n = a \Leftrightarrow$ zu jedem $\varepsilon > 0$ gibt es ein $n_0(\varepsilon)$, so dass für alle $n > n_0$ stets $|a_n - a| < \varepsilon$ ist. Die Konvergenz lässt sich durch die Konvergenz reeller Zahlenfolgen charakterisieren:

$$\lim_{n \rightarrow \infty} (x_n + iy_n) = x + iy \Leftrightarrow \lim_{n \rightarrow \infty} x_n = x \text{ und } \lim_{n \rightarrow \infty} y_n = y.$$

Konzepte für die reellen Zahlen lassen sich übertragen, sofern sie nicht mit den Ordnungsaxiomen in \mathbb{R} verbunden sind, welche für \mathbb{C} nicht existieren. So lässt sich auf komplexe Zahlenfolgen der Begriff des Häufungswerts übertragen, und es gilt das Cauchy Kriterium. Hingegen lassen sich die Begriffe Limes superior und Limes inferior und das Monotonieprinzip nicht übertragen.

Mit jeder komplexen Zahlenfolge (a_n) ist die unendliche Reihe $\sum a_n$ verbunden, die wiederum als Folge ihrer Partialsummen definiert wird. Wichtige Begriffe und Kriterien, wie absolute Konvergenz, Quotientenkriterium, Majorantenkriterium, lassen sich auf komplexe unendliche Reihen übertragen. Für jede Folge (a_n) mit $a_n \in \mathbb{C}$ und $z_0 \in \mathbb{C}$ definieren wir die Potenzreihe

$$(*) \quad \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

und fragen, für welche $z \in \mathbb{C}$ sie konvergiert. An die Stelle des Konvergenzintervalls für reelle Potenzreihen tritt jetzt der Konvergenzkreis. Es gilt folgende Erweiterung von Satz 5.5.4:

SATZ 5.8.3

Konvergiert die Reihe () für ein $z^* \in \mathbb{C}$, und divergiert sie für ein $\tilde{z} \in \mathbb{C}$, so existiert (genau) ein $R > 0$, so dass die Reihe (*) für alle $z \in \mathbb{C}$ mit $|z - z_0| < R$ konvergiert, und für alle $z \in \mathbb{C}$ mit $|z - z_0| > R$ divergiert.*

Wiederum nennt man R den Konvergenzradius mit entsprechender Interpretation der Fälle $R = 0$ (nur in z_0 konvergent) und $R = \infty$ (auf ganz \mathbb{C} konvergent). Komplexe Potenzreihen bilden einen Hauptgegenstand der Funktionentheorie (komplexe Analysis). Wir wenden uns nun den Spezialfällen der Exponentialfunktion und der trigonometrischen Funktionen zu:

DEFINITION 5.8.3

Für $z \in \mathbb{C}$ definieren wir

$$\exp(z) := e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad \sin(z) = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}, \quad \cos(z) = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.$$

Für $z \in \mathbb{R}$ stimmt das mit der ursprünglichen Definition überein. Wie im Reellen beweist man die Funktionalgleichung $e^{z_1+z_2} = e^{z_1} e^{z_2}$ für alle $z_1, z_2 \in \mathbb{C}$. Grundlegend ist der folgende Zusammenhang zwischen der Exponentialfunktion und den trigonometrischen Funktionen:

SATZ 5.8.4 (Formel von Euler)

Für alle $z \in \mathbb{C}$ gilt $e^{iz} = \cos z + i \cdot \sin z$.

BEWEIS

Wir benutzen die Periodizität der Folge $i^k = 1, i, -1, -i, 1, i, -1, -i, \dots$ und erhalten

$$\sum_{k=0}^{\infty} \frac{(iz)^k}{k!} = 1 + iz + \frac{i^2 z^2}{2!} + \frac{i^3 z^3}{3!} + \dots = \underbrace{\left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} \pm \dots\right)}_{=\cos z} + i \cdot \underbrace{\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} \pm \dots\right)}_{=\sin z}$$

□

Aus diesem Satz lassen sich einige Eigenschaften von \sin und \cos auf durchsichtigere Weise ableiten, als es durch Beschränkung auf reelle z möglich ist.

SATZ 5.8.5

Es seien $x, y \in \mathbb{R}$ und $z = x + iy$, dann gilt:

- (i) $|e^{iy}| = |\cos y + i \sin y| = 1 \quad (\Leftrightarrow \sin^2 y + \cos^2 y = 1).$
- (ii) $e^z = e^{x+iy} = e^x (\cos y + i \sin y).$
- (iii) $|e^z| = e^x = e^{\operatorname{Re}(z)}.$

BEWEIS

Zu (i): es sei $w = e^{iy}$, dann ist $|w|^2 = w \cdot \bar{w} = e^{iy} e^{-iy} = 1$. Der Rest folgt aus Satz 5.8.4. (ii) und (iii) folgen aus (i) und der Funktionalgleichung für \exp . \square

Auch die Additionstheoreme für \cos und \sin ergeben sich leicht aus dem Zusammenhang mit der Exponentialfunktion:

$$\begin{aligned} \cos(x+y) + i \sin(x+y) &= e^{i(x+y)} = e^{ix} e^{iy} = (\cos x + i \sin x) \cdot (\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\cos x \sin y + \sin x \cos y). \end{aligned}$$

Separate Betrachtung von Real- und Imaginärteil ergibt die beiden Teilaussagen von Satz 5.7.4. Schließlich erwähnen wir noch die Verwendung der Exponentialfunktion in der Polardarstellung komplexer Zahlen. Zu $z \in \mathbb{C}$ gibt es ein eindeutig bestimmtes $\phi \in [0, 2\pi)$ mit $z = |z| \cdot (\cos \phi + i \sin \phi) = |z| \cdot e^{i\phi}$.

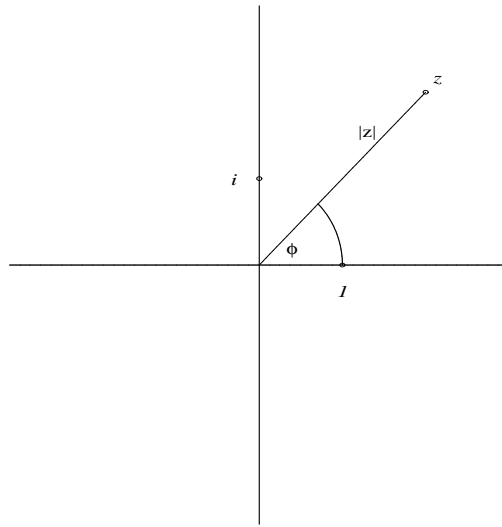


ABBILDUNG 5.5. Polardarstellung $z = x + iy = |z| \cdot e^{i\phi}$

5.9. Regeln von l'Hôpital

SATZ 5.9.1

Es sei $I = (a, b)$ ein reelles Intervall mit $-\infty \leq a < b \leq \infty$. Auf I seien die beiden Funktionen $f, g : I \rightarrow \mathbb{R}$ differenzierbar, und es habe $g'(x)$ keine Nullstelle auf I . Ferner gelte

$$\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R} \cup \{\pm\infty\} \text{ mit } c = a \text{ oder } c = b.$$

Dann ist

$$\begin{aligned} &\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \alpha, \text{ falls} \\ \text{a) } &\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0 \text{ oder b) } \lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = \pm\infty. \end{aligned}$$

BEWEIS

Wir führen den Beweis lediglich für den Fall a) und setzen noch $c \neq \pm\infty$ voraus. Da c endlich ist, können f und g durch die Festsetzung $f(c) := 0, g(c) := 0$ stetig fortgesetzt werden. Aus dem 2. Mittelwertsatz folgt

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f'(\xi)}{g'(\xi)}$$

mit $\xi \in (c, x)$. Für $x \rightarrow c$ geht auch $\xi \rightarrow c$ und damit

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{\xi \rightarrow c} \frac{f'(\xi)}{g'(\xi)} = \alpha.$$

□

BEISPIEL 5.9.1

Es ist

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} \stackrel{5.9.1}{=} \lim_{x \rightarrow 0} \frac{\sin x}{2x} \stackrel{5.9.1}{=} \lim_{x \rightarrow 0} \frac{\cos x}{2} = \frac{1}{2}.$$

Man kann auch mit der Taylorentwicklung arbeiten:

$$\frac{1 - \cos x}{x^2} = \frac{1 - \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} \pm \dots\right)}{x^2} = \frac{x^2 \left(\frac{1}{2!} + \frac{x^2}{4!} \mp \dots\right)}{x^2} = \frac{1}{2} + x^2(\dots) \rightarrow \frac{1}{2}.$$

5.10. Konvexität, Kurvendiskussion

DEFINITION 5.10.1

Es sei I ein Intervall. Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt konvex bzw. konkav auf I , wenn gilt:

$$f(\lambda x_1 + (1 - \lambda)x_2) \underset{\text{bzw. } \geq}{\leq} \lambda f(x_1) + (1 - \lambda)f(x_2) \quad \forall x_1, x_2 \in I, \forall \lambda \in [0, 1].$$

BEMERKUNG 5.10.1

Die Punkte der Sekante zwischen den Kurvenpunkten $(x_1, f(x_1))$ und $(x_2, f(x_2))$ werden durch $\lambda \in [0, 1]$ wie folgt parametrisiert: $(\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2))$. Geometrisch bedeutet die Konvexität (bzw. Konkavität) also, dass die Sekante zwischen $(x_1, f(x_1))$ und $(x_2, f(x_2))$ nirgends unterhalb (bzw. oberhalb) des Graphen von f verläuft.

SATZ 5.10.1

Ist I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion, so sind folgende Aussagen äquivalent:

- (i) f ist konvex auf I .
- (ii) $-f$ ist konkav auf I .
- (iii) Sind $x_1, \dots, x_n \in I$ und $\lambda_1, \dots, \lambda_n \geq 0$ mit $\sum \lambda_k = 1$, so gilt die Jensensche Ungleichung

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \sum_{k=1}^n \lambda_k f(x_k).$$

- (iv) Für $x_1, x_2, x_3 \in I$ mit $x_1 < x_2 < x_3$ gilt stets

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

BEWEIS

Die Äquivalenz (i) \Leftrightarrow (ii) und (iii) \Rightarrow (i) sind klar. Die Richtung (i) \Rightarrow (iii) wird durch Induktion nach n bewiesen. Fall $n = 1$: Es ist $\lambda_1 = 1$ und die Behauptung ist klar. $n \rightarrow n + 1$: Es seien $x_1, \dots, x_{n+1} \in I$ und $\lambda_1, \dots, \lambda_{n+1} \geq 0$ mit $\lambda_1 + \dots + \lambda_{n+1} = 1$ gegeben. Ohne Einschränkung kann $\lambda_1 > 0$ angenommen werden. Wir setzen

$$\lambda := \sum_{k=1}^n \lambda_k, \quad \mu_k := \frac{\lambda_k}{\lambda} \quad (k = 1 \dots n), \quad x := \sum_{k=1}^n \mu_k x_k.$$

Dann ist $0 < \lambda \leq 1$, $\mu_1 + \dots + \mu_n = 1$, $1 - \lambda = \lambda_{n+1}$ und $x \in I$, denn

$$\min\{x_1, \dots, x_n\} \leq \sum_{k=1}^n \mu_k x_k \leq \max\{x_1, \dots, x_n\}$$

$$\begin{aligned} \Rightarrow f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &= f(\lambda x + (1-\lambda)x_{n+1}) \stackrel{(i)}{\leq} \lambda f\left(\sum_{k=1}^n \mu_k x_k\right) + \lambda_{n+1} f(x_{n+1}) \\ &\stackrel{\text{IH}}{\leq} \sum_{k=1}^n \lambda \mu_k f(x_k) + \lambda_{n+1} f(x_{n+1}) = \sum_{k=1}^{n+1} \lambda_k f(x_k). \end{aligned}$$

Nun zu (i) \Leftrightarrow (iv): Wir schreiben $x_2 = \lambda x_1 + (1-\lambda)x_3$ mit $\lambda \in (0, 1)$. Dann gilt

$$f(\lambda x_1 + (1-\lambda)x_3) \leq \lambda f(x_1) + (1-\lambda)f(x_3) \Leftrightarrow f(x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_3)$$

$$\Leftrightarrow \lambda(f(x_2) - f(x_1)) \leq (1-\lambda)(f(x_3) - f(x_2))$$

$$\Leftrightarrow f(x_2) - f(x_1) \leq \frac{1-\lambda}{\lambda}(f(x_3) - f(x_2)) = \frac{x_2 - x_1}{x_3 - x_2}(f(x_3) - f(x_2))$$

$$\Leftrightarrow \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Das bedeutet gerade (i) \Leftrightarrow (iv). □

SATZ 5.10.2

Es sei I ein Intervall mit den Endpunkten a, b für $-\infty \leq a < b \leq \infty$ und $f : I \rightarrow \mathbb{R}$ stetig:

- (i) Ist f auf (a, b) differenzierbar, so gilt
 - (a) f konvex auf $I \Leftrightarrow f'$ wachsend auf (a, b) .
 - (b) f konkav auf $I \Leftrightarrow f'$ abnehmend auf (a, b) .
- (ii) Ist f auf (a, b) zweimal differenzierbar, so gilt
 - (a) f konvex auf $I \Leftrightarrow f'' \geq 0$ in (a, b) .
 - (b) f konkav auf $I \Leftrightarrow f'' \leq 0$ in (a, b) .

BEWEIS

Richtung \Rightarrow aus (i)a): Es sei f konvex in I und $a < x_1 < x_2 \leq b$. Für $x_1 < t < u < x_2$ gilt dann

$$\frac{f(t) - f(x_1)}{t - x_1} \leq \frac{f(u) - f(t)}{u - t} \leq \frac{f(x_2) - f(u)}{x_2 - u}.$$

Für $t \rightarrow x_1+$ strebt die linke Seite gegen $f'(x_1)$ und für $u \rightarrow x_2-$ strebt die rechte Seite gegen $f'(x_2)$. Also ist $f'(x_1) \leq f'(x_2)$. (i)b) folgt aus a) durch Betrachtung von $-f$. Rückrichtung \Leftarrow : Es sei f' wachsend in (a, b) und $x_1, x_2, x_3 \in I$ mit $x_1 < x_2 < x_3$. Dann gibt es nach dem 1. Mittelwertsatz $\xi \in (x_1, x_2)$ und $\vartheta \in (x_2, x_3)$ mit

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) \leq f'(\vartheta) = \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Nach Satz 5.10.1(iv) ist f dann konvex auf I . Teil b) ergibt sich aus a) wieder durch Betrachtung von $-f$. Teil (ii) folgt aus Satz 5.4.2 und (i). □

DEFINITION 5.10.2

Es sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ beliebig. Ein innerer Punkt $x_0 \in I$ heißt Wendepunkt von f , wenn es ein $\varepsilon > 0$ gibt, so dass eine der folgenden Aussagen zutrifft:

- (i) f ist konvex auf $(x_0 - \varepsilon, x_0]$ und konkav auf $[x_0, x_0 + \varepsilon)$.
- (ii) f ist konkav auf $(x_0 - \varepsilon, x_0]$ und konvex auf $[x_0, x_0 + \varepsilon)$.

SATZ 5.10.3 (Notwendige Bedingung für Wendepunkte)

Es sei x_0 Wendepunkt von f und f in x_0 zweimal differenzierbar (d. h. f' existiert in einer Umgebung von x_0 und ist in x_0 differenzierbar), dann ist $f''(x_0) = 0$.

BEWEIS

Aus obiger Definition und aus Satz 5.10.2 folgt, dass es ein $\varepsilon > 0$ gibt mit: f' ist wachsend (bzw. fallend) auf $(x_0 - \varepsilon, x_0]$ und fallend (bzw. wachsend) auf $[x_0, x_0 + \varepsilon)$. f' besitzt also ein lokales Extremum in x_0 . Nach Satz 5.1.3 ist $f''(x_0) = 0$. \square

SATZ 5.10.4 (Hinreichende Bedingung für Wendepunkte)

Es sei I ein Intervall, $f : I \rightarrow \mathbb{R}$ und x_0 ein innerer Punkt von I . f sei $(2n+1)$ -mal differenzierbar in x_0 für ein $n \in \mathbb{N}$ (d. h. $f', f'', \dots, f^{(2n)}$ existieren in einer Umgebung von x_0 und $f^{(2n)}$ ist differenzierbar in x_0). Es sei $f^{(k)}(x_0) = 0$ für $k = 2 \dots 2n$ und $f^{(2n+1)}(x_0) \neq 0$, dann ist x_0 ein Wendepunkt von f .

BEWEIS

Es sei ohne Einschränkung $f^{(2n+1)}(x_0) > 0$. Es folgt

$$\lim_{x \rightarrow x_0} \frac{f^{(2n)}(x) - f^{(2n)}(x_0)}{x - x_0} = f^{(2n+1)}(x_0) > 0.$$

Also gibt es ein $\varepsilon > 0$, so dass $f^{(2n)}$ existiert und

$$\frac{f^{(2n)}(x)}{x - x_0} > 0$$

ist für $0 < |x - x_0| < \varepsilon$. Es sei also $0 < |x - x_0| < \varepsilon$. Die Anwendung des Satzes von Taylor auf f'' liefert die Existenz eines ξ mit $x_0 < \xi < x$ oder $x < \xi < x_0$ und

$$\begin{aligned} f''(x) &= \sum_{k=0}^{2n-3} \frac{f^{(k+2)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(2n)}(\xi)}{(2n-2)!} (x - x_0)^{2n-2} \\ &= \frac{1}{(2n-2)!} \cdot \frac{f^{(2n)}(\xi)}{\xi - x_0} \cdot \frac{\xi - x_0}{x - x_0} \cdot (x - x_0)^{2n-1} \begin{cases} > 0 & \text{falls } x > x_0 \\ < 0 & \text{falls } x < x_0 \end{cases}. \end{aligned}$$

Aus Satz 5.10.2 folgt damit, dass f konkav in $(x_0 - \varepsilon, x_0]$ und konvex in $[x_0, x_0 + \varepsilon)$ ist. Damit ist x_0 ein Wendepunkt von f . \square

Wir verfügen jetzt über fast alle Voraussetzungen für die so genannte Kurvendiskussion. Dabei handelt es sich um die Aufgabe, einen Überblick über die wesentlichen Charakteristiken einer Funktion (bzw. ihres Graphen) zu erhalten. Solche Charakteristiken umfassen Eigenschaften der Monotonie (Angabe der Intervalle, in denen die Funktion wächst bzw. abnimmt), der Konvexität, globale bzw. lokale Extrema, Wendepunkte sowie Nullstellen. Wir haben in Satz 5.1.3 schon das Verschwinden der Ableitung als notwendige Bedingung für das Vorliegen eines lokalen Extremums erkannt. Wir ergänzen sie noch durch eine hinreichende Bedingung:

SATZ 5.10.5 (Hinreichende Bedingung für lokale Extrema)

Es sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ $(2n)$ -mal differenzierbar bei x_0 für ein $n \in \mathbb{N}$. Es sei $f^{(k)}(x_0) = 0$ für $k = 1, \dots, 2n-1$ und $f^{(2n)}(x_0) \neq 0$. Dann besitzt f an der Stelle x_0 ein lokales Extremum, und zwar ein lokales Maximum falls $f^{(2n)}(x_0) < 0$ ist, bzw. ein lokales Minimum falls $f^{(2n)}(x_0) > 0$ ist.

BEWEIS

Ohne Einschränkung sei $f^{(2n)}(x_0) > 0$, ansonsten betrachte $-f$. Es gilt

$$\lim_{x \rightarrow x_0} \frac{f^{(2n-1)}(x) - f^{(2n-1)}(x_0)}{x - x_0} = f^{(2n)}(x_0) > 0.$$

Also gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x_0) \subseteq I$ so dass $f^{(2n-1)}$ dort existiert und

$$\frac{f^{(2n-1)}(x)}{x - x_0} > 0$$

in $U_\varepsilon(x_0) \setminus \{x_0\}$ ist. Nach dem Satz von Taylor gibt es ein $\xi \in U_\varepsilon(x_0)$ mit $x_0 < \xi < x$ oder $x < \xi < x_0$ und

$$\begin{aligned} f(x) &= \sum_{k=0}^{2n-2} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(2n-1)}(\xi)}{(2n-1)!} (x - x_0)^{2n-1} \\ &= f(x_0) + \frac{1}{(2n-1)!} \cdot \frac{f^{(2n-1)}(\xi)}{\xi - x_0} \cdot \frac{\xi - x_0}{x - x_0} \cdot (x - x_0)^{2n} > f(x_0). \end{aligned}$$

Daraus folgt $f(x) > f(x_0)$ für alle $x \in U_\varepsilon(x_0) \setminus \{x_0\}$, also besitzt f in x_0 ein lokales Minimum. \square

Wir schließen mit dem Beispiel einer Kurvendiskussion:

BEISPIEL 5.10.1

Es sei

$$f : \begin{cases} [-3, 2] & \rightarrow & \mathbb{R} \\ x & \mapsto & x^3 - 3x + 1 \end{cases}.$$

Lokale Extrema und Monotonie:

Es ist $f'(x) = 3x^2 - 3 = 3(x-1)(x+1)$, also $f'(x) = 0$ für $x = \pm 1$. Dabei ist $f'(x) > 0$ in $(-\infty, -1)$ und $(1, \infty)$ sowie $f'(x) < 0$ in $(-1, 1)$. Also haben wir in $x_1 = -1$ das lokale Maximum $f(-1) = 3$. Analog besitzt f in 1 das lokale Minimum $f(1) = -1$. f ist wachsend in $(-\infty, -1)$ und $(1, \infty)$, aber fallend in $(-1, 1)$.

Globale Extrema:

Die Randwerte von f sind $f(-3) = -17$ und $f(2) = 3$. Vergleich mit den lokalen Extrema ergibt, dass das globale Maximum $\max f(x) = 3$ in den Punkten -1 und 2 angenommen wird, das globale Minimum $\min f(x) = -17$ im Punkt -3 .

Konvexität und Wendepunkte:

Es ist

$$f''(x) = 6x \begin{cases} < 0 & \text{in } (-\infty, 0) \\ > 0 & \text{in } (0, \infty) \end{cases}.$$

Also ist f konkav in $(-\infty, 0)$ und konvex in $(0, \infty)$. Damit ist $x_3 = 0$ der einzige Wendepunkt von f .

6. Integration

6.1. Das Riemannsches Integral

Historisch hat sich die Analysis aus zwei geometrischen Grundproblemen entwickelt. Das erste ist das Tangentenproblem: bestimme die Tangente an einer Kurve mit der Gleichung $y = f(x)$. Die Lösung dieses Grundproblems führt zur Differentialrechnung. Das andere Grundproblem ist das Flächenproblem: bestimme die Fläche unter der Kurve $y = f(x)$ (zunächst für $f(x) \geq 0$) zwischen $x = a$ und $x = b$.

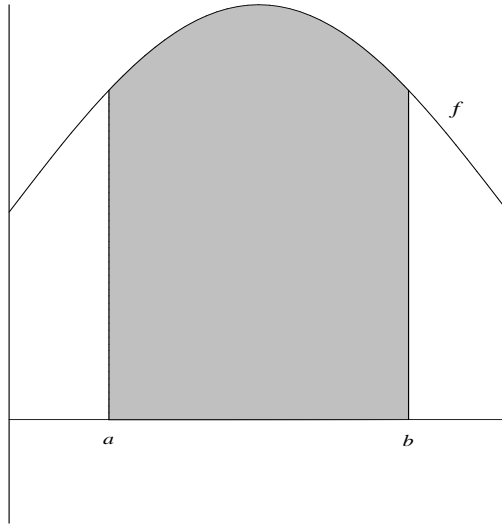


ABBILDUNG 6.1. Flächeninhalt von $f(x)$ über $[a, b]$

Die Lösung dieses Grundproblems führt zur Definition des Integrals und zur Integralrechnung. Es war die grundlegende Entdeckung von Newton und Leibniz, dass die beiden Grundprobleme Umkehrungen von einander sind.

DEFINITION 6.1.1

Es sei f eine reelle, auf dem kompakten Intervall $[a, b]$ definierte und beschränkte Funktion. Eine Zerlegung \mathcal{Z} von $[a, b]$ ist eine endliche Folge (x_0, \dots, x_n) mit $a = x_0 < x_1 < \dots < x_n = b$. Die x_j heißen Teilungspunkte von \mathcal{Z} . Einer Zerlegung \mathcal{Z} werden die Obersumme und die Untersumme wie folgt zugeordnet:

$$\begin{aligned}\overline{S}(\mathcal{Z}) &:= \sum_{\nu=1}^n \overline{M}_{\nu}(x_{\nu} - x_{\nu-1}) \\ \underline{S}(\mathcal{Z}) &:= \sum_{\nu=1}^n \underline{M}_{\nu}(x_{\nu} - x_{\nu-1}) \\ \overline{M}_{\nu} &:= \sup\{f(x) \mid x_{\nu-1} \leq x \leq x_{\nu}\} \\ \underline{M}_{\nu} &:= \inf\{f(x) \mid x_{\nu-1} \leq x \leq x_{\nu}\}\end{aligned}$$

Im Fall $f(x) \geq 0$ können Ober- und Untersummen als obere bzw. untere Approximationen der Fläche unter $f(x)$ gedeutet werden.

DEFINITION 6.1.2

Eine Zerlegung $\mathcal{Z} = (x_0, \dots, x_n)$ heißt eine Verfeinerung einer Zerlegung $\mathcal{Z}^* = (y_0, \dots, y_m)$, wenn die y_i eine Teilfolge der x_i bilden.

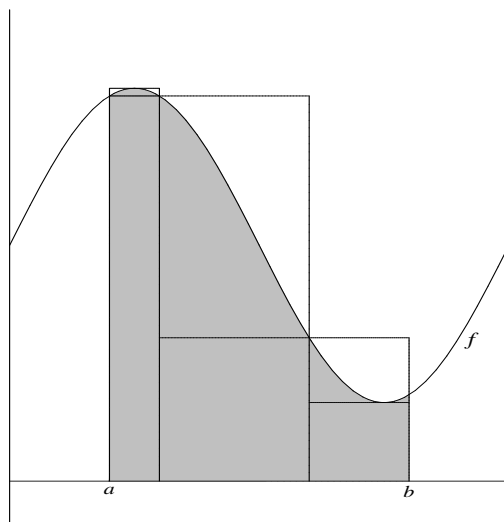


ABBILDUNG 6.2. Ober- und Untersummen

SATZ 6.1.1

Es sei f wie in Definition 6.1.1 und \mathcal{Z} bzw. \mathcal{Z}^* Zerlegungen von $[a, b]$, so dass \mathcal{Z} eine Verfeinerung von \mathcal{Z}^* ist. Dann gilt:

$$\overline{S}(\mathcal{Z}) \leq \overline{S}(\mathcal{Z}^*) \quad \text{und} \quad \underline{S}(\mathcal{Z}) \geq \underline{S}(\mathcal{Z}^*).$$

BEWEIS

Es sei $\mathcal{Z} = (x_0, \dots, x_n)$. Die Teilungspunkte von \mathcal{Z}^* lassen sich, weil sie auch Teilungspunkte von \mathcal{Z} sind, wie folgt niederschreiben

$$\mathcal{Z}^* = (x_{j_0}, x_{j_1}, \dots, x_{j_m}) \quad \text{mit} \quad 0 = j_0 < j_1 < \dots < j_m = n.$$

Wir setzen

$$\overline{M}_\mu^* = \sup\{f(x) \mid x_{j_{\mu-1}} \leq x \leq x_{j_\mu}\}.$$

Offenbar gilt

$$x_{j_\mu} - x_{j_{\mu-1}} = \sum_{\nu=j_{\mu-1}+1}^{j_\mu} (x_\nu - x_{\nu-1}), \quad \text{und} \quad \overline{M}_\mu^* \geq \overline{M}_\nu \quad \forall \nu \text{ mit } j_{\mu-1} \leq \nu \leq j_\mu, \text{ also}$$

$$\sum_{\nu=j_{\mu-1}+1}^{j_\mu} \overline{M}_\nu \cdot (x_\nu - x_{\nu-1}) \leq \overline{M}_\mu^* \cdot \sum_{\nu=j_{\mu-1}+1}^{j_\mu} (x_\nu - x_{\nu-1}) = \overline{M}_\mu^* (x_{j_\mu} - x_{j_{\mu-1}}).$$

Addition dieser Ungleichungen für $\mu = 1, 2, \dots, m$ führt zur ersten Ungleichung des Satzes. Die zweite Ungleichung wird analog bewiesen. \square

DEFINITION 6.1.3

Unter der Feinheit $\eta(\mathcal{Z})$ einer Zerlegung $\mathcal{Z} = (x_0, \dots, x_n)$ versteht man $\eta(\mathcal{Z}) := \max_{0 \leq i \leq n-1} (x_{i+1} - x_i)$.

Eine Folge von Zerlegungen (\mathcal{Z}_j) heißt ausgezeichnet, falls $\lim_{k \rightarrow \infty} \eta(\mathcal{Z}_k) = 0$ ist.

SATZ 6.1.2

Es sei $|f(x)| \leq M$ für ein $x \in [a, b]$ und \mathcal{Z} eine Verfeinerung der Zerlegung \mathcal{Z}^* von $[a, b]$. \mathcal{Z} habe λ Teilungspunkte mehr als \mathcal{Z}^* . Dann gelten

$$\overline{S}(\mathcal{Z}^*) - \overline{S}(\mathcal{Z}) \leq 2\lambda M \cdot \eta(\mathcal{Z}) \quad \text{sowie} \quad \underline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}^*) \leq 2\lambda M \cdot \eta(\mathcal{Z}).$$

BEWEIS

Die Teilungspunkte von \mathcal{Z} seien x_0, x_1, \dots, x_n , die von \mathcal{Z}^* seien x_{j_0}, \dots, x_{j_m} (mit $0 = j_0 < j_1 < \dots < j_m = n$), so dass $\lambda = n - m$ ist. \overline{M}_μ^* sei die größte unter den Zahlen $\overline{M}_{j_{\mu-1}+1}, \dots, \overline{M}_{j_\mu}$ und es sei $\overline{M}_\mu^* = \overline{M}_{\nu_\mu}^*$ für $j_{\mu-1} < \nu_\mu < j_\mu$. Dann ist

$$\begin{aligned} \overline{M}_\mu^*(x_{j_\mu} - x_{j_{\mu-1}}) &= \sum_{\nu=j_{\mu-1}+1}^{j_\mu} \overline{M}_\nu(x_\nu - x_{\nu-1}) + \sum_{\nu=j_{\mu-1}+1}^{j_\mu} (\overline{M}_\mu^* - \overline{M}_\nu)(x_\nu - x_{\nu-1}) \\ &= 2M \sum_{\substack{\nu=j_{\mu-1}+1 \\ \nu \neq \nu_\mu}}^{j_\mu} (x_\nu - x_{\nu-1}) \leq 2M(j_\mu - j_{\mu-1} - 1)\eta(\mathcal{Z}). \end{aligned}$$

Summation dieser Ungleichungen für $\mu = 1, 2, \dots, m$ liefert

$$\begin{aligned} \sum_{\mu=1}^m \overline{M}_\mu^*(x_{j_\mu} - x_{j_{\mu-1}}) - \sum_{\mu=1}^m \sum_{\nu=j_{\mu-1}}^{j_\mu} \overline{M}_\nu(x_\nu - x_{\nu-1}) &= \overline{S}(\mathcal{Z}^*) - \overline{S}(\mathcal{Z}) \\ &\leq 2M \sum_{\mu=1}^m (j_\mu - j_{\mu-1} - 1)\eta(\mathcal{Z}) = 2M(j_m - j_0 - m)\eta(\mathcal{Z}) = 2\lambda M\eta(\mathcal{Z}). \end{aligned}$$

Die zweite Ungleichung wird analog bewiesen. □

SATZ 6.1.3

Für je zwei Zerlegungen \mathcal{Z}_1 und \mathcal{Z}_2 ist stets $\underline{S}(\mathcal{Z}_1) \leq \overline{S}(\mathcal{Z}_2)$.

BEWEIS

Ist \mathcal{Z} eine gemeinsame Verfeinerung von \mathcal{Z}_1 und \mathcal{Z}_2 , so ist $\underline{S}(\mathcal{Z}_1) \leq \underline{S}(\mathcal{Z}) \leq \overline{S}(\mathcal{Z}) \leq \overline{S}(\mathcal{Z}_2)$. □

Nach Satz 6.1.3 ist die Menge der Obersummen nach unten und die Menge der Untersummen nach oben beschränkt. Nach dem Infimums- bzw. Supremumsprinzip besitzt die Menge Ober- bzw. Untersummen ein Infimum bzw. ein Supremum.

DEFINITION 6.1.4

Wir nennen

$$\int_a^b f(x)dx := \inf_{\mathcal{Z}} \overline{S}(\mathcal{Z}) \quad \text{bzw.} \quad \int_a^b f(x)dx := \sup_{\mathcal{Z}} \underline{S}(\mathcal{Z})$$

das Oberintegral bzw. das Unterintegral der Funktion $f(x)$ über $[a, b]$.

Aus Satz 6.1.3 erhalten wir unmittelbar

SATZ 6.1.4

Es ist stets

$$\int_a^b f(x)dx \leq \int_a^b f(x)dx.$$

DEFINITION 6.1.5

Unter der Superposition $\mathcal{Z}_1 \vee \mathcal{Z}_2$ zweier Zerlegungen versteht man die Zerlegung, deren Menge von Teilungspunkten die Vereinigung der Mengen der Teilungspunkte von \mathcal{Z}_1 und \mathcal{Z}_2 ist.

SATZ 6.1.5

Es seien die Voraussetzungen wie oben. Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle Zerlegungen \mathcal{Z} von $[a, b]$ mit $\eta(\mathcal{Z}) < \delta$ stets gilt:

$$0 \leq \overline{S}(\mathcal{Z}) - \int_a^b f(x)dx < \varepsilon \quad , \quad 0 \leq \int_a^b f(x)dx - \underline{S}(\mathcal{Z}) < \varepsilon .$$

BEWEIS

Nach Definition des Oberintegrals gibt es eine Zerlegung \mathcal{Z}^* mit

$$\overline{S}(\mathcal{Z}^*) < \int_a^b f(x)dx + \frac{\varepsilon}{2} .$$

Die Anzahl der Teilungspunkte von \mathcal{Z}^* sei α . Die Superposition $\mathcal{Z} \vee \mathcal{Z}^*$ hat dann höchstens α Teilungspunkte mehr als \mathcal{Z} . Nach Satz 6.1.2 ist demnach

$$\overline{S}(\mathcal{Z}) - \overline{S}(\mathcal{Z} \vee \mathcal{Z}^*) \leq 2\alpha M\eta(\mathcal{Z} \vee \mathcal{Z}^*) .$$

Da nach Satz 6.1.1 $\overline{S}(\mathcal{Z} \vee \mathcal{Z}^*) \leq \overline{S}(\mathcal{Z}^*)$ ist und da $\eta(\mathcal{Z} \vee \mathcal{Z}^*) \leq \eta(\mathcal{Z})$ ist erhalten wir

$$\overline{S}(\mathcal{Z} \vee \mathcal{Z}^*) - \int_a^b f(x)dx < \frac{\varepsilon}{2} + 2\alpha M\eta(\mathcal{Z}) .$$

Mit $\delta = \frac{\varepsilon}{4\alpha M}$ erhalten wir die Behauptung. □

KOROLLAR 6.1.6

Für jede ausgezeichnete Zerlegungsfolge (\mathcal{Z}_n) von $[a, b]$ gilt

$$\lim_{n \rightarrow \infty} \overline{S}(\mathcal{Z}_n) = \int_a^b f(x)dx \quad , \quad \lim_{n \rightarrow \infty} \underline{S}(\mathcal{Z}_n) = \int_a^b f(x)dx .$$

DEFINITION 6.1.6

Die Funktion f sei auf $[a, b]$ definiert und dort beschränkt. Sie heißt integrierbar auf $[a, b]$, wenn Ober- und Unterintegral von f übereinstimmen. Der gemeinsame Wert von Ober- und Unterintegral heißt dann das (Riemannsches) Integral von f über $I = [a, b]$. Geschrieben

$$\int_I f(x)dx = \int_a^b f(x)dx = \int_a^b f(x)dx = \int_a^b f(x)dx .$$

Die Intervallgrenze a heißt untere Schranke, b obere Schranke des Integrals, I der Integrationsbereich und x die Integrationsvariable.

Die Bezeichnung der Integrationsvariable kann beliebig gewählt werden:

$$\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(u)du .$$

SATZ 6.1.7 (Riemannsches Integritätskriterium)

Die Funktion f ist genau dann über $[a, b]$ integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{Z} von $[a, b]$ gibt, so dass $\overline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}) < \varepsilon$ gilt.

BEWEIS

\Rightarrow : Es sei f integrierbar über $[a, b]$. Nach Satz 6.1.5 gibt es zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{Z} , so dass

$$\overline{S}(\mathcal{Z}) - \int_a^b f(x)dx < \frac{\varepsilon}{2} \quad \text{und} \quad \int_a^b f(x)dx - \underline{S}(\mathcal{Z}) < \frac{\varepsilon}{2}$$

gilt. Addition der Ungleichungen ergibt $\overline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}) < \varepsilon$.

\Leftarrow : Es existiere nun umgekehrt zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{Z} mit $\overline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}) < \varepsilon$. Aus der Definition von Ober- und Unterintegral folgt

$$\int_a^b f(x)dx - \int_a^b f(x)dx \leq \overline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}) < \varepsilon,$$

und da ε beliebig klein gewählt werden kann folgt

$$\int_a^b f(x)dx - \int_a^b f(x)dx \leq 0 \quad \Rightarrow \quad \int_a^b f(x)dx = \int_a^b f(x)dx,$$

also die Integrierbarkeit von f . □

SATZ 6.1.8

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann ist f integrierbar über $[a, b]$.

BEWEIS

Nach Satz 4.3.4 ist f auf $[a, b]$ sogar gleichmäßig stetig. Zu jedem $\varepsilon > 0$ gibt es also ein $\delta = \delta(\varepsilon) > 0$, so dass aus $|x - y| < \delta$ und $x, y \in [a, b]$ stets $|f(x) - f(y)| < \varepsilon$ folgt. Es sei nun $\mathcal{Z} = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$ mit Feinheit $\eta(\mathcal{Z}) < \delta$. Es gilt

$$\overline{S}(\mathcal{Z}) - \underline{S}(\mathcal{Z}) = \sum_{\nu=1}^n \left(\sup_{x \in [x_{\nu-1}, x_{\nu}]} f(x) - \inf_{x \in [x_{\nu-1}, x_{\nu}]} f(x) \right) \cdot (x_{\nu} - x_{\nu-1}) \leq \varepsilon \sum_{\nu=1}^n (x_{\nu} - x_{\nu-1}) = \varepsilon(b-a).$$

Da ε beliebig klein gewählt werden kann, folgt nach Satz 6.1.7 die Behauptung. □

Wir kommen nun zu einer Modifikation der Ober- und Untersummen:

DEFINITION 6.1.7

Es sei $\mathcal{Z} = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$. Eine endliche Folge von Zahlen (ξ_1, \dots, ξ_n) mit $x_{\nu-1} \leq \xi_{\nu} \leq x_{\nu}$ heißt eine Besetzung B von \mathcal{Z} . Man bezeichnet

$$S(\mathcal{Z}, B) = \sum_{\nu=1}^n f(\xi_{\nu})(x_{\nu} - x_{\nu-1})$$

als Riemannsche Summe.

Für stetige Funktionen f , die auf kompakten Intervallen ihre Suprema/Infima annehmen, sind die Ober- und Untersummen spezielle Riemannsche Summen. Das gilt auch für monotone Funktionen.

SATZ 6.1.9

Es sei (\mathcal{Z}_n) eine ausgezeichnete Zerlegungsfolge von $[a, b]$ und (B_n) eine Folge von zugehörigen Besetzungen. Die Funktion f sei über $[a, b]$ integrierbar. Dann gilt

$$\lim_{n \rightarrow \infty} S(\mathcal{Z}_n, B_n) = \int_a^b f(x)dx.$$

BEWEIS

Das folgt unmittelbar aus Korollar 6.1.6, da stets $\underline{S}(\mathcal{Z}_n) \leq S(\mathcal{Z}_n, B_n) \leq \overline{S}(\mathcal{Z}_n)$ gilt. \square

SATZ 6.1.10

Wenn f über $[a, b]$ integrierbar ist, so ist f auch über jedem Teilintervall $[a_1, b_1] \subseteq [a, b]$ integrierbar.

OHNE BEWEIS \square

SATZ 6.1.11

Wenn f sowohl über $[a, b]$ als auch über $[b, c]$ integrierbar ist, so ist f über $[a, c]$ integrierbar mit

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

OHNE BEWEIS \square

In der Definition des Integrals setzten wir bisher $a < b$ voraus. Wir ergänzen den Integralbegriff durch

DEFINITION 6.1.8

Für $a < b$ und f integrierbar über $[a, b]$ ist

$$\int_a^a f(x) dx := 0 \quad \text{sowie} \quad \int_b^a f(x) dx := - \int_a^b f(x) dx.$$

Mit dieser Definition gilt Satz 6.1.11 für beliebige $a, b, c \in I$ in einem kompakten Intervall I , über das f integrierbar ist. Mehrfache Anwendung dieses Satzes ergibt

SATZ 6.1.12

Falls $a_1, \dots, a_n \in I$ (I kompaktes Intervall) und f über I integrierbar ist, so gilt

$$\int_{a_1}^{a_2} f(x) dx + \int_{a_2}^{a_3} f(x) dx + \dots + \int_{a_{n-1}}^{a_n} f(x) dx = \int_{a_1}^{a_n} f(x) dx.$$

SATZ 6.1.13 (Linearität des Integrals)

Wenn die Integrale

$$\int_a^b f_1(x) dx \quad \text{und} \quad \int_a^b f_2(x) dx$$

existieren, so existiert für beliebige $c_1, c_2 \in \mathbb{R}$ auch das Integral über $c_1 f_1(x) + c_2 f_2(x)$, und es gilt

$$\int_a^b (c_1 f_1(x) + c_2 f_2(x)) dx = c_1 \int_a^b f_1(x) dx + c_2 \int_a^b f_2(x) dx.$$

OHNE BEWEIS \square

SATZ 6.1.14

Existiert das Integral $\int_a^b f(x) dx$, so existiert auch $\int_a^b |f(x)| dx$, und es gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

OHNE BEWEIS

□

SATZ 6.1.15

Produkt und Quotient zweier integrierbarer Funktionen sind integrierbar. Letzteres unter der Voraussetzung, dass der Absolutbetrag der Nennerfunktion ein positives Infimum auf dem Integrationsbereich besitzt.

OHNE BEWEIS

□

SATZ 6.1.16

Jede auf dem kompakten Intervall $[a, b]$ monotone Funktion f ist über $[a, b]$ integrierbar.

OHNE BEWEIS

□

SATZ 6.1.17

Wenn f und g über $[a, b]$ integrierbar sind ($a < b$) und $f(x) \leq g(x)$ für alle $x \in [a, b]$ ist, so gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

BEWEIS

Die Behauptung folgt aus den entsprechenden Ungleichungen für Ober- bzw. Untersummen.

□

SATZ 6.1.18 (Cauchy-Schwarzsche Ungleichung für Integrale)

Es seien f und g integrierbar über $[a, b]$, dann gilt

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \left(\int_a^b f(x)^2 dx \right) \cdot \left(\int_a^b g(x)^2 dx \right).$$

BEWEIS

Es sei $a < b$ und (\mathcal{Z}_n) eine ausgezeichnete Zerlegungsfolge von $[a, b]$ mit jeweils $n + 1$ Teilungspunkten. Für $n \in \mathbb{N}$ sei $B_n = (\xi_1^{(n)}, \dots, \xi_n^{(n)})$ eine Besetzung von \mathcal{Z}_n . Für die Riemannschen Summen $S_1(\mathcal{Z}_n, B_n)$ von f^2 , $S_2(\mathcal{Z}_n, B_n)$ von g^2 und $S_3(\mathcal{Z}_n, B_n)$ von $f \cdot g$ gilt nach der gewöhnlichen Cauchy-Schwarzschen Ungleichung (siehe lineare Algebra):

$$\begin{aligned} S_3^2(\mathcal{Z}_n, B_n) &= \left(\sum_{\nu=1}^n f(\xi_\nu)g(\xi_\nu)(x_\nu - x_{\nu-1}) \right)^2 \\ &\leq \left(\sum_{\nu=1}^n f(\xi_\nu)^2(x_\nu - x_{\nu-1}) \right) \cdot \left(\sum_{\nu=1}^n g(\xi_\nu)^2(x_\nu - x_{\nu-1}) \right) = S_1(\mathcal{Z}_n, B_n) \cdot S_2(\mathcal{Z}_n, B_n). \end{aligned}$$

Die Behauptung folgt aus Korollar 6.1.6 durch den Übergang $n \rightarrow \infty$.

□

6.2. Mittelwertsätze und Hauptsatz der Differential- und Integralrechnung

SATZ 6.2.1 (1. Mittelwertsatz der Integralrechnung)

Wenn f in einem kompakten Intervall $[a, b]$ stetig ist, so gibt es ein $\xi \in [a, b]$ mit

$$\int_a^b f(x) dx = f(\xi) \cdot (b - a).$$

BEWEIS

Es seien \overline{M} und \underline{M} das Supremum bzw. Infimum von f auf $[a, b]$, dann gilt

$$\int_a^b \underline{M} dx \leq \int_a^b f(x) dx \leq \int_a^b \overline{M} dx, \text{ also}$$

$$\underline{M} \cdot (b - a) \leq \int_a^b f(x) dx \leq \overline{M} \cdot (b - a).$$

Es gibt also ein $M \in \mathbb{R}$ mit $\underline{M} \leq M \leq \overline{M}$ mit

$$\int_a^b f(x) dx = M \cdot (b - a).$$

Nach dem Zwischenwertsatz 4.3.1 gibt es $\xi \in [a, b]$ mit $f(\xi) = M$, woraus die Behauptung folgt. \square

SATZ 6.2.2

Das Integral ist eine stetige Funktion der oberen (und auch der unteren) Grenze.

BEWEIS

Es sei f integrierbar über $[a, b]$ und

$$F(u) := \int_a^u f(t) dt, \quad u, x_0, x_0 + h \in [a, b],$$

dann ist

$$F(x_0 + h) - F(x_0) = \int_a^{x_0+h} f(t) dt - \int_a^{x_0} f(t) dt = \int_{x_0}^{x_0+h} f(t) dt.$$

Es sei $|f(x)| \leq M$ für $x \in [a, b]$, nach den Sätzen 6.1.14 und 6.1.17 folgt

$$|F(x_0 + h) - F(x_0)| \leq \int_{x_0}^{x_0+h} M dt = M \cdot |h|.$$

Das bedeutet gerade $\lim_{h \rightarrow 0} F(x_0 + h) = F(x_0)$. \square

DEFINITION 6.2.1

Eine differenzierbare Funktion F heißt Stammfunktion oder unbestimmtes Integral zu f auf dem Intervall I , wenn $F'(x) = f(x)$ für alle $x \in I$ ist. Bezeichnung: $F(x) = \int f(x) dx$.

SATZ 6.2.3

Sind F und G Stammfunktionen von f auf einem Intervall I , so gibt es eine Konstante $c \in \mathbb{R}$ mit $G(x) = F(x) + c$ für alle $x \in I$.

BEWEIS

Es ist $(G(x) - F(x))' = 0$. Nach Satz 5.4.2 ist $G(x) - F(x)$ eine Konstante. □

SATZ 6.2.4

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann besitzt f in $[a, b]$ die Stammfunktion

$$F(x) = \int_a^x f(t) dt.$$

Jede weitere Stammfunktion von f unterscheidet sich von F nur um eine Konstante.

BEWEIS

Es seien $x_0, x_0 + h \in [a, b]$. Aus Satz 6.2.1 folgt

$$F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t) dt = f(x_0 + \vartheta h) \cdot h \quad \text{mit } 0 \leq \vartheta \leq 1.$$

Es folgt

$$F'(x_0) = \lim_{h \rightarrow 0} \frac{F(x_0 + h) - F(x_0)}{h} = \lim_{h \rightarrow 0} (f(x_0 + \vartheta h)) = f(x_0).$$

□

BEMERKUNG 6.2.1

Satz 6.2.4 besagt, dass die Operationen Integration und Differentiation Umkehrungen voneinander sind.

SATZ 6.2.5 (Hauptsatz der Differential- und Integralrechnung)

Es sei f stetig auf einem Intervall I . Für jede Stammfunktion F von f auf I und beliebige $a, b \in I$ gilt

$$\int_a^b f(x) dx = F(b) - F(a).$$

BEWEIS

Nach Satz 6.2.4 ist für $c \in I$ die Funktion

$$G(x) := \int_c^x f(t) dt$$

eine Stammfunktion von f . Es gilt

$$(*) \quad \int_a^b f(x) dx = G(b) - G(a).$$

Es sei nun F eine beliebige Stammfunktion von f auf I . Nach Satz 6.2.3 gibt es $c \in \mathbb{R}$ mit $F(x) = G(x) + c$, woraus mit (*) die Behauptung folgt. □

BEMERKUNG 6.2.2

Für $F(b) - F(a)$ schreiben wir auch $[F(x)]_a^b$ oder $[F(x)]_{x=a}^{x=b}$. Also

$$\int_a^b f(x) dx = [F(x)]_a^b.$$

BEISPIEL 6.2.1

Es gilt $\int x^2 dx = \frac{1}{3}x^3 + c$. Die einfachste Wahl ist $c = 0$. Aus Satz 6.2.5 ergibt sich

$$\int_0^1 x^2 dx = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3}.$$

Der Hauptsatz ermöglicht die Berechnung vieler bestimmter Integrale durch das Auffinden von unbestimmten Integralen oder Stammfunktionen. Dieser Vorgang wird Integration genannt. Allerdings besitzen die „meisten“ etwas komplizierteren Funktionen keine elementaren Stammfunktionen (d. h. sie lassen sich nicht durch bekannte Funktionen ausdrücken). Im nächsten Abschnitt besprechen wir die wichtigsten Integrationsverfahren. Die Kenntnis dieser Techniken hat etwas an Bedeutung verloren, seit es Computeralgebrasysteme gibt, die diese Aufgabe ausführen können.

6.3. Integration elementarer Funktionen

Wir beginnen mit einer Liste der unbestimmten Integrale gewisser Grundfunktionen:

$$1. \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + c \begin{cases} \text{auf } \mathbb{R} & \text{falls } \alpha \in \mathbb{N}_0 \\ \text{auf } \mathbb{R} \setminus \{0\} & \text{falls } \alpha \in \mathbb{Z} \setminus \{-1\} \\ \text{auf } (0, \infty) & \text{falls } \alpha \in \mathbb{R} \setminus \{-1\} \end{cases},$$

$$2. \int \frac{dx}{x} = \log|x| + c \text{ auf } \mathbb{R} \setminus \{0\} \quad , \quad 3. \int e^x dx = e^x + c \text{ auf } \mathbb{R} ,$$

$$4. \int \sin dx = -\cos x + c \text{ auf } \mathbb{R} \quad , \quad 5. \int \cos dx = \sin x + c \text{ auf } \mathbb{R} \quad ,$$

$$6. \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + c \text{ auf } (-1, 1) \quad , \quad 7. \int \frac{dx}{1+x^2} = \arctan x + c \text{ auf } \mathbb{R} .$$

Die Techniken der Integration durch Substitution bzw. partielle Integration ergeben sich aus der Umkehrung der Kettenregel bzw. der Produktregel für Ableitungen:

SATZ 6.3.1 (Integration durch Substitution)

Es sei $g : I \rightarrow \mathbb{R}$ stetig differenzierbar und f stetig auf $g(I)$. Dann gilt auf I

$$\int f(g(x))g'(x)dx = \int f(u)du .$$

BEMERKUNG 6.3.1

Formal lässt sich Integration durch Substitution auch so ausdrücken: Die Substitution $u = g(x)$, $du = g'(x)dx$ führt $\int f(g(x))g'(x)dx$ über in $\int f(u)du$.

SATZ 6.3.2 (Partielle Integration)

Sind $f, g : I \rightarrow \mathbb{R}$ stetig differenzierbar, so gilt auf I

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx .$$

BEMERKUNG 6.3.2

Mit $u = f(x)$ und $v = g(x)$ lässt sich die partielle Integration formal so ausdrücken: $\int u dv = uv - \int v du$.

BEISPIEL 6.3.1

Es gilt $\int e^{\sin x} \cos x dx = e^{\sin x} + c$. Substitution mit $u = \sin x$ ergibt

$$\int e^{\sin x} \cos dx = \int e^u du = e^u + c = e^{\sin x} + c .$$

BEISPIEL 6.3.2

Es gilt $\int x e^x dx = (x-1)e^x + c$. Partielle Integration mit $u = x$ und $v = e^x$ ergibt

$$\int x e^x dx = x e^x - \int e^x dx = x e^x - e^x + c.$$

Dieses Beispiel ist ein Spezialfall der allgemeinen Problemstellungen

$$\int P(x) e^x dx, \quad \int P(x) \sin(kx) dx, \quad \int P(x) \cos(kx) dx,$$

die durch (für Polynome P höheren Grades wiederholte) partielle Integration gelöst werden:

BEISPIEL 6.3.3

Es gilt

$$\begin{aligned} \int x^2 \sin(3x) &= -\frac{1}{3} x^2 \cos(3x) - \int 2x \left(-\frac{1}{3} \cos(3x)\right) = -\frac{1}{3} x^2 \cos(3x) + \frac{2}{3} \int x \cos(3x) \\ &= -\frac{1}{3} x^2 \cos(3x) + \frac{2}{3} \left(\frac{1}{3} x \sin(3x) + \frac{1}{9} \cos(3x) \right) = \frac{2}{9} x \sin(3x) - \frac{1}{3} x^2 \cos(3x) + \frac{2}{27} \cos(3x). \end{aligned}$$

BEISPIEL 6.3.4

Es gilt

$$\begin{aligned} \int e^x \sin x dx &= e^x \sin x - \int e^x \cos x dx = e^x \sin x - e^x \cos x - \int e^x \sin x dx. \\ \Rightarrow 2 \int e^x \sin x dx &= e^x (\sin x - \cos x) + c \quad \Rightarrow \quad \int e^x \sin x dx = \frac{1}{2} e^x (\sin x - \cos x) + c'. \end{aligned}$$

BEISPIEL 6.3.5

Es gilt

$$\int x^2 \log x dx \stackrel{u=\log x, v'=x^2}{=} \frac{1}{3} x^3 \log x - \int \frac{1}{3} x^3 \frac{1}{x} dx = \frac{1}{3} x^3 \log x - \frac{1}{3} \int x^2 dx = \frac{1}{3} x^3 \log x - \frac{1}{9} x^3 + c.$$

Wir diskutieren noch die Integration rationaler Funktionen. Rationale Funktionen können stets durch Partialbruchzerlegung integriert werden. Wir führen die grundlegenden Techniken ohne Beweis ein:

LEMMA 6.3.3

Jedes reelle Polynom P besitzt eine Zerlegung der Form

$$P(x) = A \cdot \prod_{i=1}^n (x^2 + b_i x + c_i)^{\nu_i} \cdot \prod_{j=1}^l (x - a_j)^{\mu_j}$$

mit geeigneten Konstanten $A, a_j, b_i, c_i \in \mathbb{R}$, wobei die quadratischen Polynome $x^2 + b_i x + c_i$ keine reellen Nullstellen besitzen.

LEMMA 6.3.4

Es sei R ein reelles Polynom vom Grad m , Q ein reelles Polynom vom Grad $\deg(Q) > \deg(R)$ mit

$$Q(x) = A \cdot \prod_{i=1}^n (x^2 + b_i x + c_i)^{\nu_i} \cdot \prod_{j=1}^l (x - a_j)^{\mu_j},$$

dann gibt es eindeutig bestimmte reelle Zahlen $d_r^{(i)}$ und $e_r^{(i)}$ ($1 \leq r \leq \nu_i, 1 \leq i \leq n$) und $f_s^{(j)}$ ($1 \leq s \leq \mu_j, 1 \leq j \leq l$) mit

$$(*) \quad \frac{R(x)}{Q(x)} = \sum_{i=1}^n \sum_{r=1}^{\nu_i} \frac{d_r^{(i)} x + e_r^{(i)}}{(x^2 + b_i x + c_i)^r} + \sum_{j=1}^l \sum_{s=1}^{\mu_j} \frac{f_s^{(j)}}{(x - a_j)^s}.$$

Wichtig ist die praktische Bestimmung der Partialbruchzerlegung. Wir erwähnen mehrere Techniken:

1) Koeffizientenvergleich: Man bringt die rechte Seite von (*) auf den Hauptnenner und vergleicht die Zähler der linken und rechten Seite. Man erhält ein lineares Gleichungssystem für die Konstanten $d_r^{(i)}, e_r^{(i)}, f_s^{(j)}$ aus dem sich diese ermitteln lassen:

BEISPIEL 6.3.6

$$\begin{aligned} \frac{x^3 + 2x + 1}{(x^2 + 1)(x^2 + 2x + 3)} &\stackrel{!}{=} \frac{d^{(1)}x + e^{(1)}}{x^2 + 1} + \frac{d^{(2)}x + e^{(2)}}{x^2 + 2x + 3} \\ &= \frac{(d^{(1)} + d^{(2)})x^3 + (2d^{(1)} + e^{(1)} + e^{(2)})x^2 + (3d^{(1)} + 2e^{(1)} + d^{(2)})x + (3e^{(1)} + e^{(2)})}{(x^2 + 1)(x^2 + 2x + 3)}. \end{aligned}$$

Vergleich der Koeffizienten der Zähler führt zum Gleichungssystem

$$\begin{aligned} d^{(1)} + d^{(2)} &= 1 \\ 2d^{(1)} + e^{(1)} + e^{(2)} &= 0 \\ 3d^{(1)} + d^{(2)} + 2e^{(1)} &= 2 \\ 3e^{(1)} + e^{(2)} &= 1 \end{aligned}$$

mit der Lösung $d^{(1)} = 0, d^{(2)} = 1, e^{(1)} = \frac{1}{2}, e^{(2)} = -\frac{1}{2}$. Wir erhalten die Partialbruchzerlegung

$$\frac{x^3 + 2x + 1}{(x^2 + 1)(x^2 + 2x + 3)} = \frac{\frac{1}{2}}{x^2 + 1} + \frac{x - \frac{1}{2}}{x^2 + 2x + 3}.$$

2) Ist das Nennerpolynom ein Produkt aus verschiedenen linearen Polynomen, so können auch spezielle Werte für die Variable x substituiert werden:

$$\begin{aligned} \frac{x^2 + x + 7}{(x - 1)(x + 1)(x + 2)} &\stackrel{!}{=} \frac{f_1}{x - 1} + \frac{f_2}{x + 1} + \frac{f_3}{x + 2} \\ &= \frac{f_1(x + 1)(x + 2) + f_2(x - 1)(x + 2) + f_3(x - 1)(x + 1)}{(x - 1)(x + 1)(x + 2)}. \end{aligned}$$

$$x = -1 \Rightarrow -2f_2 = 7 \Rightarrow f_2 = -\frac{7}{2}$$

$$x = 1 \Rightarrow 6f_1 = 9 \Rightarrow f_1 = \frac{3}{2}$$

$$x = -2 \Rightarrow 3f_3 = 9 \Rightarrow f_3 = 3$$

Also

$$\frac{x^2 + x + 7}{(x - 1)(x + 1)(x + 2)} = \frac{\frac{3}{2}}{x - 1} + \frac{-\frac{7}{2}}{x + 1} + \frac{3}{x + 2}.$$

3) Durch Verwendung komplexer Koeffizienten kann das Nennerpolynom $Q(x)$ stets vollständig in Linearfaktoren zerlegt werden:

BEISPIEL 6.3.7

Es gilt

$$\frac{1}{x^4 - 1} = \frac{1}{(x^2 + 1)(x + 1)(x - 1)} = \frac{1}{(x + i)(x - i)(x + 1)(x - 1)} \stackrel{!}{=} \frac{f_1}{x + i} + \frac{f_2}{x - i} + \frac{f_3}{x + 1} + \frac{f_4}{x - 1}$$

$$= \frac{f_1(x-i)(x+1)(x-1) + f_2(x+i)(x+1)(x-1) + f_3(x+i)(x-i)(x-1) + f_3(x+i)(x-i)(x+1)}{(x+i)(x-i)(x+1)(x-1)}$$

mit den zugehörigen Lösungen

$$x = -i \Rightarrow 4if_1 = 1 \Rightarrow f_1 = -\frac{1}{4}i$$

$$x = i \Rightarrow -4if_2 = 1 \Rightarrow f_2 = \frac{1}{4}i$$

$$x = -1 \Rightarrow -4f_3 = 1 \Rightarrow f_3 = -\frac{1}{4}$$

$$x = 1 \Rightarrow 4f_4 = 1 \Rightarrow f_4 = \frac{1}{4}$$

Wir erhalten

$$\frac{1}{x^4 - 1} = \frac{-\frac{1}{4}i}{x+i} + \frac{\frac{1}{4}i}{x-i} + \frac{-\frac{1}{4}}{x+1} + \frac{\frac{1}{4}}{x-1}.$$

Indem wir nun die beiden konjugiert komplexen Ausdrücke kombinieren, erhalten wir

$$\frac{-\frac{1}{4}i}{x+i} + \frac{\frac{1}{4}i}{x-i} = \frac{-\frac{1}{2}}{x^2+1}$$

und damit die reelle Partialbruchzerlegung

$$\frac{1}{x^4 - 1} = \frac{-\frac{1}{2}}{x^2+1} + \frac{-\frac{1}{4}}{x+1} + \frac{\frac{1}{4}}{x-1}.$$

Durch die Partialbruchzerlegung lässt sich die Integration jeder rationalen Funktion auf die Grundaufgaben

$$\int \frac{dx}{(x-a)^n}, \quad \int \frac{cx+d}{(x^2+ax+b)^n} dx$$

reduzieren.

6.4. Uneigentliche Integrale

Bei der Definition des Riemann-Integrals wurde das Integrationsintervall $I = [a, b]$ als kompakt und der Integrand als beschränkt vorausgesetzt. Der Integrationsbegriff kann nun auf die Fälle erweitert werden, in denen eine oder mehrere dieser Voraussetzungen nicht erfüllt sind.

DEFINITION 6.4.1

Im Folgenden setzen wir stets voraus, dass f über $[a, b]$ *nicht* integrierbar ist:

- (i) Ist $-\infty < a \leq b \leq \infty$ und f integrierbar über jedes kompakte Teilintervall von $[a, b)$, so nennt man den Grenzwert

$$\int_a^b f(x) dx := \lim_{\xi \rightarrow b^-} \int_a^\xi f(x) dx$$

das uneigentliche Integral von f über $[a, b)$. Es heißt konvergent, falls dieser Grenzwert existiert, andernfalls divergent. Im Fall $b = \infty$ ist unter $\lim_{\xi \rightarrow b^-}$ der Übergang $\lim_{\xi \rightarrow \infty}$ zu verstehen.

- (ii) Es sei $-\infty \leq a \leq b < \infty$ und f integrierbar über jedem kompakten Teilintervall von $(a, b]$. Dann nennt man entsprechend den Grenzwert

$$\int_a^b f(x) dx := \lim_{\xi \rightarrow a^+} \int_\xi^b f(x) dx$$

das uneigentliche Integral von f über $(a, b]$. Es heißt konvergent, falls dieser Grenzwert existiert, andernfalls divergent. Im Fall $a = -\infty$ ist unter $\lim_{\xi \rightarrow a+}$ der Übergang $\lim_{\xi \rightarrow -\infty}$ zu verstehen.

- (iii) Es sei $-\infty \leq a \leq b \leq \infty$ und f integrierbar über jedes kompakte Teilintervall von (a, b) und $c \in (a, b)$. Dann nennt man die Summe der uneigentlichen Integrale

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx$$

das uneigentliche Integral von f über (a, b) . Diese Definition ist unabhängig von c . Das uneigentliche Integral heißt konvergent, wenn *beide* Summanden konvergent sind, andernfalls divergent.

BEISPIEL 6.4.1

Es gilt:

$$(i) \int_0^1 \frac{dx}{\sqrt{x}} = \lim_{\xi \rightarrow 0+} \int_{\xi}^1 \frac{dx}{\sqrt{x}} = \lim_{\xi \rightarrow 0+} [2\sqrt{x}]_{\xi}^1 = 2,$$

$$(ii) \int_{-\infty}^{\infty} e^{-|x|} dx = \int_{-\infty}^0 e^{-|x|} dx + \int_0^{\infty} e^{-|x|} dx = \int_{-\infty}^0 e^x dx + \int_0^{\infty} e^{-x} dx$$

$$= \lim_{\xi \rightarrow -\infty} \int_{\xi}^0 e^x dx + \lim_{\eta \rightarrow \infty} \int_0^{\eta} e^{-x} dx = \lim_{\xi \rightarrow -\infty} [e^x]_{\xi}^0 + \lim_{\eta \rightarrow \infty} [-e^{-x}]_0^{\eta} = 2.$$

$$(iii) \int_{-\infty}^{\infty} e^{-x} dx \text{ ist divergent, da } \lim_{\xi \rightarrow -\infty} \int_{\xi}^0 e^{-x} dx = \lim_{\xi \rightarrow -\infty} (e^{-\xi} - 1) \text{ nicht endlich ist.}$$

$$(iv) \int_1^{\infty} \frac{1}{x} dx \text{ ist divergent, da } \lim_{\xi \rightarrow \infty} \int_1^{\xi} \frac{1}{x} dx = \lim_{\xi \rightarrow \infty} \log(\xi) \text{ nicht endlich ist.}$$

Wir definieren das uneigentliche Integral zunächst für einen noch allgemeineren Fall:

DEFINITION 6.4.2

Es sei $a = a_0 < a_1 < \dots < a_n = b$ eine Zerlegung von $I = [a, b]$. Die Summe uneigentlicher Integrale

$$\int_a^b f(x) dx := \sum_{k=1}^n \int_{a_{k-1}}^{a_k} f(x) dx$$

nennt man das uneigentliche Integral von f über I .

Mit dieser Definition braucht f nur auf $I \setminus \{a_0, \dots, a_n\}$ definiert zu sein:

BEISPIEL 6.4.2

$$\int_{-1}^1 \log |x| dx = \int_{-1}^0 \log(-x) dx + \int_0^1 \log x dx$$

$$\begin{aligned}
&= \lim_{\xi \rightarrow 0^-} \int_{-1}^{\xi} \log(-x) dx + \lim_{\eta \rightarrow 0^+} \int_{\eta}^1 \log x dx = 2 \lim_{\xi \rightarrow 0^+} \left(\int_{\xi}^1 \log x dx \right) = 2 \lim_{\xi \rightarrow 0^+} [x \log(x) - x]_{\xi}^1 \\
&= 2 \lim_{\xi \rightarrow 0^+} (1 \cdot \log(1) - 1 - \xi \log(\xi) + \xi) = -2 + 0 + 2 \lim_{\xi \rightarrow 0^+} (-\xi \log(\xi)) = -2.
\end{aligned}$$

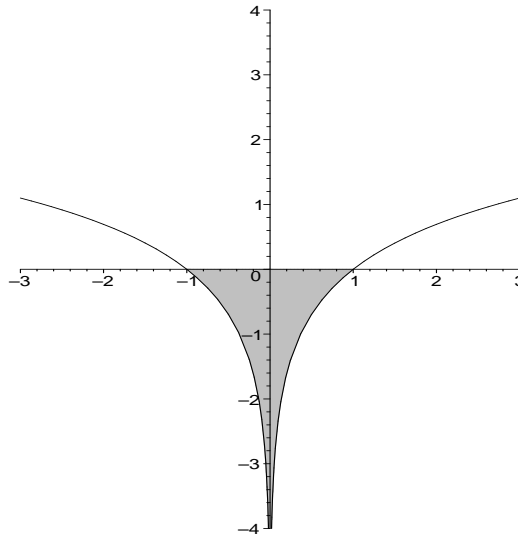


ABBILDUNG 6.3. Der Flächeninhalt von $\log|x|$ über $[-1, 1]$ ist endlich.

In vielen Fällen ist es schwierig, die Werte von uneigentlichen Integralen zu berechnen. Es ist jedoch im Allgemeinen einfach, die Frage der Konvergenz oder Divergenz zu entscheiden. Wir geben die wichtigsten Kriterien an:

SATZ 6.4.1 (Cauchy-Kriterium für Integrale)

Das an der oberen Grenze uneigentliche Integral

$$\int_a^b f(x) dx$$

konvergiert genau dann, wenn zu jedem $\varepsilon > 0$ eine Umgebung U von b existiert, so dass

$$\left| \int_c^d f(x) dx \right| < \varepsilon$$

ist für alle $c, d \in [a, b] \cap U$.

OHNE BEWEIS

□

DEFINITION 6.4.3

Ein uneigentliches Integral $\int_a^b f(x) dx$ heißt absolut konvergent, wenn $\int_a^b |f(x)| dx$ konvergent ist.

SATZ 6.4.2

Ein absolut konvergentes uneigentliches Integral ist auch konvergent.

OHNE BEWEIS

□

SATZ 6.4.3 (Majoranten-/Minorantenkriterium für Integrale)

Es gilt:

- (i) Ist $|f(x)| \leq g(x)$ auf (a, b) und konvergiert $\int_a^b g(x)dx$, so konvergiert $\int_a^b f(x)dx$ absolut.
- (ii) Ist $f(x) \geq g(x) \geq 0$ auf (a, b) und divergiert $\int_a^b g(x)dx$, so divergiert auch $\int_a^b f(x)dx$.

OHNE BEWEIS

□

DEFINITION 6.4.4

Gilt (i), so heißt $\int_a^b g(x)dx$ eine konvergente Majorante von $\int_a^b f(x)dx$, bzw. divergente Minorante in (ii).

BEISPIEL 6.4.3

Das Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

ist konvergent, denn es gilt $-x^2 \leq 1 - |x|$ für alle $x \in \mathbb{R}$. Also besitzt das Integral die konvergente Majorante

$$\int_{-\infty}^{\infty} e^{1-|x|} dx = 2e.$$

BEISPIEL 6.4.4 (Eulersche Γ -Funktion)

Das Integral

$$\Gamma(x) := \int_0^{\infty} t^{x-1} e^{-t} dt$$

ist konvergent für alle $x > 0$. Das Integral

$$\int_0^1 t^{x-1} e^{-t} dt$$

ist an der unteren Grenze uneigentlich nur für $x < 1$. Es besitzt dort aber die konvergente Majorante

$$\int_0^1 t^{x-1} dt = \lim_{\xi \rightarrow 0+} \int_{\xi}^1 t^{x-1} dt = \lim_{\xi \rightarrow 0+} \left[\frac{t^x}{x} \right]_{t=\xi}^{t=1} = \frac{1}{x}.$$

Also ist auch

$$\int_0^1 t^{1-x} e^{-t} dt$$

für $0 < x < 1$ konvergent. Andererseits ist die Funktion $f(t) = t^{x-1}e^{-\frac{t}{2}}$ für festes x in $t \geq 1$ beschränkt, d. h. $f(t) \leq M$ für ein $M \in \mathbb{R}$. Also besitzt

$$\int_1^{\infty} t^{x-1} e^{-t} dt$$

die konvergente Majorante

$$\int_1^{\infty} M e^{-\frac{t}{2}} dt .$$

Es sind also die beiden Teilintegrale und damit $\Gamma(x)$ konvergent für $x > 0$.

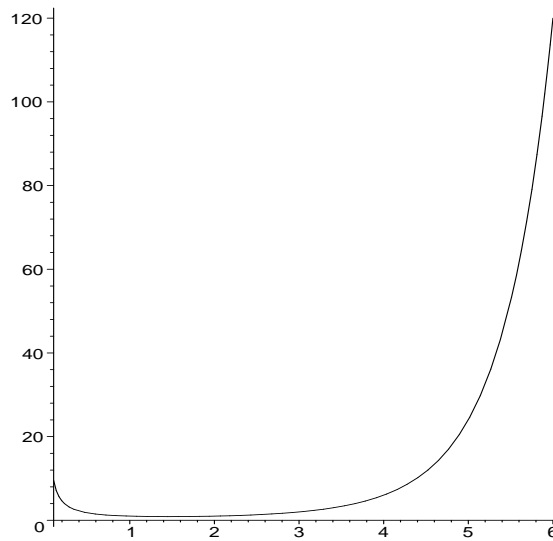


ABBILDUNG 6.4. Die Eulersche Γ -Funktion.

Ohne Beweis sei angemerkt, dass $\Gamma(n) = (n - 1)!$ ist für alle $n \in \mathbb{N}$.

SATZ 6.4.4 (Integralkriterium für unendliche Reihen)

Es sei $f : [1, \infty) \rightarrow \mathbb{R}$ monoton, dann gilt

$$\sum_{n=1}^{\infty} f(n) \text{ konvergent} \Leftrightarrow \int_1^{\infty} f(x) dx \text{ konvergent} .$$

BEWEIS

Es sei ohne Einschränkung f monoton fallend, dann gilt

$$\int_k^{k+1} f(x) dx \leq f(k) \leq \int_{k-1}^k f(x) dx \quad \forall k \geq 2 .$$

Summation über k ergibt

$$\int_2^{n+1} f(x) dx \leq \sum_{k=2}^n f(k) \leq \int_1^n f(x) dx .$$

Damit ist $\sum_{k=1}^n f(k) = f(1) + \sum_{k=2}^n f(k)$ genau dann konvergent für $n \rightarrow \infty$, wenn $\int_1^{\infty} f(x) dx$ konvergiert. \square

BEISPIEL 6.4.5

Die Reihen

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}$$

divergieren für $\alpha \leq 1$ und konvergieren für $\alpha > 1$. Denn für $\alpha < 1$ ist

$$\int_1^{\infty} x^{-\alpha} dx$$

divergent, denn es gilt

$$\lim_{\xi \rightarrow \infty} \int_1^{\xi} x^{-\alpha} dx = \lim_{\xi \rightarrow \infty} \left[\frac{x^{1-\alpha}}{1-\alpha} \right]_1^{\xi} = \lim_{\xi \rightarrow \infty} \left(\frac{\xi^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha} \right) = \infty.$$

Für $\alpha = 1$ ist

$$\int_1^{\infty} \frac{1}{x} dx$$

divergent, da

$$\lim_{\xi \rightarrow \infty} \int_1^{\xi} \frac{dx}{x} = \lim_{\xi \rightarrow \infty} [\log(x)]_1^{\xi} = \lim_{\xi \rightarrow \infty} (\log(\xi) - \log(1)) = \infty$$

ist. Für $\alpha > 1$ konvergiert dagegen

$$\begin{aligned} \int_1^{\infty} x^{-\alpha} dx &= \lim_{\xi \rightarrow \infty} \int_1^{\xi} x^{-\alpha} dx = \lim_{\xi \rightarrow \infty} \left[\frac{x^{1-\alpha}}{1-\alpha} \right]_1^{\xi} = \frac{1}{1-\alpha} \left(\lim_{\xi \rightarrow \infty} [x^{1-\alpha}]_1^{\xi} \right) \\ &= \frac{1}{1-\alpha} \left(\lim_{\xi \rightarrow \infty} (\overbrace{\xi^{1-\alpha}}^{\text{negativ}} - 1) \right) = \frac{1}{1-\alpha} (0 - 1) = \frac{1}{\alpha - 1}. \end{aligned}$$

6.5. Vertauschung von Integration und Grenzwertbildung

SATZ 6.5.1

Es sei $f_n : [a, b] \rightarrow \mathbb{R}$ jeweils integrierbar für $n \in \mathbb{N}$ und $f_n \rightarrow f$ gleichmäßig konvergent auf $[a, b]$. Dann ist die Grenzfunktion f integrierbar auf $[a, b]$ mit

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

BEWEIS

Das folgt aus Satz 5.5.1. □

Satz 6.5.1 kann unter Anderem dazu verwendet werden, Potenzreihenentwicklungen, die wir früher schon auf anderem Weg gewonnen haben, auf direkte Weise herzuleiten:

BEISPIEL 6.5.1

Es sei $\varepsilon > 0$ und

$$f_n(t) = \sum_{k=0}^n (-t)^k.$$

Auf $[-1 + \varepsilon, 1 - \varepsilon]$ konvergiert die Folge (f_n) gleichmäßig gegen $f(t) = \frac{1}{1+t}$. Nach Satz 6.5.1 gilt für $x \in [-1 + \varepsilon, 1 - \varepsilon]$:

$$\lim_{n \rightarrow \infty} \int_0^x f_n(t) dt = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(-1)^k x^{k+1}}{k+1} = \int_0^x \frac{dt}{1+t} = \log(1+x).$$

Also folgt

$$\log(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$$

für $|x| < 1$.

Einleitung zum zweiten Teil der Vorlesung

Auch in diesem Teil ist der Funktionsbegriff zentral. Während jedoch bisher die betrachteten Funktionen von einer Variable abhingen, beispielsweise $f(x) = \sqrt{1 + e^x}$, betrachten wir künftig Funktionen, die von einer beliebigen Anzahl von Variablen abhängen, beispielsweise

$$(*) \quad T(x, y, z, t) = (x^2 + y^2 + z^2)^{-\frac{1}{2}} \cdot e^{-t}.$$

Außerdem können die Funktionen mehrere Komponenten besitzen. Indem wir sowohl die unabhängigen Variablen als auch die Komponenten zu Vektoren zusammenfassen, können wir solche Funktionen als Abbildungen zwischen höherdimensionalen Räumen auffassen, die Funktion (*) zum Beispiel als

$$T : \mathbb{R}^4 \setminus \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\} \longrightarrow \mathbb{R}, \quad \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \mapsto (x^2 + y^2 + z^2)^{-\frac{1}{2}} \cdot e^{-t}.$$

Wir beginnen daher mit einer Betrachtung von Mengen und Folgen im euklidischen Raum \mathbb{R}^n .

7. Topologische Begriffe und Stetigkeit

7.1. Der n -dimensionale euklidische Raum

DEFINITION 7.1.1

Für $n \in \mathbb{N}$ sei $\mathbb{R}^n = \{\vec{x} = (x_1, \dots, x_n)^T \mid x_i \in \mathbb{R}, 1 \leq i \leq n\}$.

Wir stellen uns also die Elemente des \mathbb{R}^n als Spaltenvektoren vor, schreiben sie jedoch aus der Gründen der Raumersparnis meist als Transponierte von Zeilenvektoren. Wir bezeichnen die Elemente des \mathbb{R}^n auch als Punkte.

DEFINITION 7.1.2

Die Menge \mathbb{R}^n wird durch die folgenden beiden Verknüpfungen zu einem Vektorraum: Die Addition: es seien $\vec{x} = (x_1, \dots, x_n)^T$ und $\vec{y} = (y_1, \dots, y_n)^T$ Elemente des \mathbb{R}^n . Wir setzen $\vec{x} + \vec{y} := (x_1 + y_1, \dots, x_n + y_n)^T$. Die Skalarmultiplikation: es seien $\lambda \in \mathbb{R}$ und $\vec{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, wir setzen $\lambda\vec{x} := (\lambda x_1, \dots, \lambda x_n)^T$.

DEFINITION 7.1.3 (Skalarprodukt)

Für $\vec{x} = (x_1, \dots, x_n)^T$ und $\vec{y} = (y_1, \dots, y_n)^T$ aus dem \mathbb{R}^n setzen wir

$$\vec{x} \cdot \vec{y} := \sum_{k=1}^n x_k y_k,$$

bzw. als Matrixprodukt geschrieben $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$. Der Betrag eines $\vec{x} \in \mathbb{R}^n$ wird über

$$\|\vec{x}\| := \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\sum_{k=1}^n x_k^2}$$

erklärt.

SATZ 7.1.1

Die Funktion

$$\|\cdot\| : \begin{cases} \mathbb{R}^n & \rightarrow [0, \infty) \\ \vec{x} & \mapsto \|\vec{x}\| \end{cases}$$

erfüllt die folgenden Eigenschaften:

- (i) Sie ist positiv definit: $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0} = (0, \dots, 0)^T$, und $\|\vec{x}\| \geq 0$ für alle \vec{x} ,
- (ii) für alle $\vec{x} \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt $\|\lambda\vec{x}\| = |\lambda| \cdot \|\vec{x}\|$,
- (iii) es gilt die Dreiecksungleichung: $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$.

BEWEIS

Siehe Lineare Algebra. □

DEFINITION 7.1.4

Eine Menge X zusammen mit einer Funktion $d : X \times X \rightarrow [0, \infty)$ heißt ein metrischer Raum, falls gilt:

- (i) $d(x, y) = 0 \Leftrightarrow x = y$,
- (ii) d ist symmetrisch: $d(x, y) = d(y, x)$ für alle $x, y \in X$,
- (iii) d erfüllt die Dreiecksungleichung: $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in X$.

Die Funktion d heißt dann Metrik auf X , $d(x, y)$ heißt Abstand zwischen x und y .

DEFINITION 7.1.5

Auf \mathbb{R}^n erklären wir den Abstand oder die euklidische Metrik wie folgt: es seien $\vec{x} = (x_1, \dots, x_n)^T, \vec{y} = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, wir setzen

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}.$$

SATZ 7.1.2

Die Menge \mathbb{R}^n bildet mit der euklidischen Metrik einen metrischen Raum.

BEWEIS

Folgt aus den Eigenschaften des euklidischen Betrags. □

7.2. Konvergenz von Punktfolgen des \mathbb{R}^n

Die zentralen Begriffe über Folgen und Konvergenz in der Analysis einer Variablen lassen sich weitgehend auf die Analysis mehrerer Variablen übertragen. Eine Ausnahme bilden die Konzepte, die auf der Anordnung der reellen Zahlen beruhen, wie Supremum/Infimum oder Limes inferior/superior.

Als allgemeine Regel lässt sich feststellen, dass sich Definitionen und Lehrsätze im Fall einer Variablen dadurch auf den Fall mehrerer Variablen verallgemeinern lassen, dass der Abstand $|x - y|$ zwischen zwei reellen Zahlen $x, y \in \mathbb{R}$ durch den Abstand $\|\vec{x} - \vec{y}\|$ zwischen zwei Punkten des \mathbb{R}^n ersetzt wird.

DEFINITION 7.2.1

Es sei (\vec{a}_k) eine Folge von Punkten des \mathbb{R}^n . Die Folge (\vec{a}_k) konvergiert oder strebt gegen ein $\vec{a} \in \mathbb{R}^n$, wenn es zu jeder positiven Zahl $\varepsilon > 0$ ein $k_0(\varepsilon)$ gibt, so dass für alle $k > k_0$ stets $\|\vec{a}_k - \vec{a}\| < \varepsilon$ ist. \vec{a} heißt dann der Grenzwert oder Limes der Folge (\vec{a}_k) . Wir schreiben $\vec{a}_k \rightarrow \vec{a}$ für $k \rightarrow \infty$ oder $\lim_{k \rightarrow \infty} \vec{a}_k = \vec{a}$.

Für jede Punktfolge $(\vec{a}_k) = ((a_{k,1}, \dots, a_{k,n})^T)$ können die Zahlenfolgen der ν -ten Komponente $(a_{k,\nu})$ betrachtet werden. Die Konvergenz der Punktfolge (\vec{a}_k) ist dann äquivalent zur Konvergenz der n Komponentenfolgen:

SATZ 7.2.1

Gegeben sei die Punktfolge (\vec{a}_k) mit $\vec{a}_k = (a_{k,1}, \dots, a_{k,n})^T$ und $\vec{a} = (a_1, \dots, a_n)^T \in \mathbb{R}^n$. Es ist $\vec{a}_k \rightarrow \vec{a}$ für $k \rightarrow \infty$ genau dann, wenn $a_{k,\nu} \rightarrow a_\nu$ für jedes $\nu = 1, \dots, n$ gilt.

BEWEIS

Das folgt unmittelbar aus

$$|a_{k,\nu} - a_\nu| \leq \|\vec{a}_k - \vec{a}\| = \sqrt{\sum_{\nu=1}^n (a_{k,\nu} - a_\nu)^2}.$$

□

SATZ 7.2.2

Jede Teilfolge einer konvergenten Folge ist konvergent, und zwar zum selben Grenzwert wie die Ausgangsfolge.

BEWEIS

Das folgt aus Satz 7.2.1 und Satz 2.2.2. □

DEFINITION 7.2.2

Eine Punktfolge (\vec{a}_k) heißt beschränkt, wenn es ein $K \in \mathbb{R}$ gibt mit $\|\vec{a}_k\| \leq K$ für alle k .

SATZ 7.2.3

Jede konvergente Punktfolge ist beschränkt.

BEWEIS

Das folgt aus Satz 7.2.1 und Satz 2.2.3. □

SATZ 7.2.4

Es seien (\vec{a}_k) und (\vec{b}_k) Punktfolgen und (γ_k) eine Zahlenfolge mit $\vec{a}_k \rightarrow \vec{a}$, $\vec{b}_k \rightarrow \vec{b}$ und $\gamma_k \rightarrow \gamma$ für $k \rightarrow \infty$. Dann gilt

$$\lim_{k \rightarrow \infty} (\vec{a}_k \pm \vec{b}_k) = \vec{a} \pm \vec{b} \quad , \quad \lim_{k \rightarrow \infty} (\gamma_k \vec{a}_k) = \gamma \vec{a} \quad , \quad \lim_{k \rightarrow \infty} (\vec{a}_k \cdot \vec{b}_k) = \vec{a} \cdot \vec{b} .$$

BEWEIS

Folgt direkt aus Satz 7.2.1. □

SATZ 7.2.5

Es gilt $\lim_{k \rightarrow \infty} \vec{a}_k = \vec{a} \Rightarrow \lim_{k \rightarrow \infty} \|\vec{a}_k\| = \|\vec{a}\|$.

BEWEIS

Folgt direkt aus Satz 7.2.1. □

SATZ 7.2.6 (Cauchy-Kriterium)

Die Punktfolge (\vec{a}_k) ist genau dann konvergent gegen ein $\vec{a} \in \mathbb{R}^n$, wenn es zu jedem $\varepsilon > 0$ ein $k_0 = k_0(\varepsilon)$ gibt mit $\|\vec{a}_k - \vec{a}_m\| < \varepsilon$ für alle $k, m > k_0$.

BEWEIS

Dies folgt aus Satz 7.2.1 und Satz 2.4.3. □

SATZ 7.2.7 (Auswahlprinzip von Bolzano-Weierstraß)

Jede beschränkte Folge enthält eine konvergente Teilfolge.

BEWEIS

(Übungsaufgabe) □

DEFINITION 7.2.3

Es sei $\varepsilon > 0$. Unter der (offenen) ε -Umgebung eines Punktes $\vec{a} \in \mathbb{R}^n$ versteht man die Menge $U_\varepsilon(\vec{a}) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{a}\| < \varepsilon\}$.

DEFINITION 7.2.4

Ein $\vec{a} \in \mathbb{R}^n$ heißt Häufungswert der Folge (\vec{a}_k) , wenn es für jedes $\varepsilon > 0$ unendlich viele Indizes k gibt mit $\vec{a}_k \in U_\varepsilon(\vec{a})$.

SATZ 7.2.8

Jede beschränkte Punktfolge besitzt mindestens einen Häufungswert.

BEWEIS

Das folgt aus Satz 7.2.7. □

7.3. Punktmengen

Die Begriffe der offenen, abgeschlossenen und kompakten Teilmengen von \mathbb{R} , die im Abschnitt 4.2 eingeführt wurden, können leicht auf Teilmengen des \mathbb{R}^n verallgemeinert werden.

DEFINITION 7.3.1

Eine Teilmenge $O \subseteq \mathbb{R}^n$ heißt offen, wenn es zu jedem $\vec{\xi} \in O$ ein $\varepsilon > 0$ gibt mit $U_\varepsilon(\vec{\xi}) \subseteq O$. Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt abgeschlossen, wenn der Grenzwert jeder konvergenten Folge aus A in A liegt. Eine Teilmenge $K \subseteq \mathbb{R}^n$ heißt kompakt, wenn jede Folge aus K eine konvergente Teilfolge besitzt, die gegen einen Grenzwert aus K konvergiert. Ist $X \subseteq \mathbb{R}^n$, so heißt $O \subseteq \mathbb{R}^n$ X -offen (oder relativ offen bzgl. X), wenn es zu jedem $\vec{\xi} \in O$ ein $\varepsilon > 0$ gibt mit $U_\varepsilon(\vec{\xi}) \cap X \subseteq O$.

DEFINITION 7.3.2

Es sei $M \subseteq \mathbb{R}^n$. Ein $\vec{\xi} \in M$ heißt innerer Punkt von M , falls es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(\vec{\xi}) \subseteq M$.

BEMERKUNG 7.3.1

Eine Teilmenge $O \subseteq \mathbb{R}^n$ ist offen genau dann, wenn sie nur aus inneren Punkten besteht.

SATZ 7.3.1

Eine Teilmenge $A \subseteq \mathbb{R}^n$ ist genau dann abgeschlossen, wenn das Komplement $A' = \mathbb{R}^n \setminus A$ offen ist.

BEWEIS

Wird genau wie Satz 4.2.1 bewiesen. □

DEFINITION 7.3.3

Ein $\vec{\xi} \in \mathbb{R}^n$ heißt Häufungspunkt der Menge $M \subseteq \mathbb{R}^n$, wenn es eine Folge (\vec{x}_k) aus M gibt, die gegen $\vec{\xi}$ konvergiert, und deren Glieder alle von $\vec{\xi}$ verschieden sind.

SATZ 7.3.2

Es sei $M \subseteq \mathbb{R}^n$. Ein $\vec{\xi} \in \mathbb{R}^n$ ist genau dann Häufungspunkt von M , wenn in jeder ε -Umgebung von $\vec{\xi}$ ein von $\vec{\xi}$ verschiedener Punkt von M liegt.

BEWEIS

Das ist trivial. □

SATZ 7.3.3

Vereinigungen endlich vieler und Durchschnitte beliebig vieler abgeschlossener Mengen sind wieder abgeschlossen. Durchschnitte endlich vieler und Vereinigungen beliebig vieler offener Mengen sind wieder offen.

BEWEIS

Es seien endlich viele offene Mengen $O_1, \dots, O_m \subseteq \mathbb{R}^n$ gegeben und $\vec{\xi} \in O_1 \cap \dots \cap O_m$. Für jedes k mit $1 \leq k \leq m$ gibt es dann eine Umgebung $U_{\varepsilon_k}(\vec{\xi})$, die komplett in O_k liegt. Dann gilt für $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_m\}$

$$U_\varepsilon(\vec{\xi}) = \bigcap_{k=1}^m U_{\varepsilon_k}(\vec{\xi}) \subseteq \bigcap_{k=1}^m O_k,$$

d. h. das beliebige $\vec{\xi}$ ist innerer Punkt und der Schnitt der O_k damit offen. Der Fall der Vereinigung beliebiger offener Mengen ist trivial. Die Behauptung für Schnitte und Vereinigungen abgeschlossener Mengen folgt durch Komplementbildung mit Satz 7.3.1. □

DEFINITION 7.3.4

Der Durchschnitt aller abgeschlossenen Obermengen von $M \subseteq \mathbb{R}$ heißt die abgeschlossene Hülle oder der Abschluss \overline{M} von M . Die Vereinigung aller offenen Teilmengen von M heißt der innere Kern oder das Innere M^0 von M . Die Punktmenge $\overline{M} \setminus M^0$ heißt der Rand $\text{Rd}(M)$ von M , jeder ihrer Punkte ein Randpunkt von M . Der Rand ist offenbar eine abgeschlossene Punktmenge.

DEFINITION 7.3.5

Ein $\vec{\xi} \in M$ heißt isolierter Punkt von M , falls $\vec{\xi}$ kein Häufungspunkt von M ist.

DEFINITION 7.3.6

Eine Teilmenge $M \subseteq \mathbb{R}^n$ heißt beschränkt, falls es ein $C \in \mathbb{R}$ gibt mit $\|\vec{x}\| \leq C$ für alle $\vec{x} \in M$.

SATZ 7.3.4

Eine Teilmenge $K \subseteq \mathbb{R}^n$ ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.

BEWEIS

Wie in Satz 4.2.2. □

Wie im Fall $n = 1$ können auch für beliebige Dimension $n \in \mathbb{N}$ kompakte Teilmengen durch eine Überdeckungseigenschaft charakterisiert werden:

DEFINITION 7.3.7

Es sei $M \subseteq \mathbb{R}^n$ und jedes $\vec{x} \in M$ möge in einer gewissen offenen Menge $O_{\vec{x}} \subseteq \mathbb{R}^n$ liegen, so dass $M \subseteq \bigcup O_{\vec{x}}$ ist. In diesem Fall nennen wir das Mengensystem $\mathcal{O} = \{O_{\vec{x}} \mid \vec{x} \in M\}$ eine offene Überdeckung von M . Gibt es ein endliches Teilsystem $\mathcal{O}' = \{O_{\vec{x}_1}, \dots, O_{\vec{x}_m}\}$, das bereits ganz M überdeckt, so sagt man, \mathcal{O} enthalte eine endliche Überdeckung von M .

SATZ 7.3.5 (Überdeckungssatz von Heine-Borel)

Eine Teilmenge des \mathbb{R}^n ist genau dann kompakt, wenn jede ihrer offenen Überdeckungen eine endliche Überdeckung enthält.

OHNE BEWEIS □

7.4. Stetigkeit von Funktionen mehrerer Variablen

DEFINITION 7.4.1

Es seien $l, n \in \mathbb{N}$ und $M \subseteq \mathbb{R}^n$. Unter $\mathcal{F}_n^l(M)$ verstehen wir die Menge aller Abbildungen (Funktionen) $\vec{f}: M \rightarrow \mathbb{R}^l$.

DEFINITION 7.4.2

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ und $\vec{\xi} \in M \subseteq \mathbb{R}^n$. Die Funktion \vec{f} heißt stetig in $\vec{\xi}$, falls es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt, so dass $\|\vec{f}(\vec{\xi}) - \vec{f}(\vec{x})\| < \varepsilon$ ist für alle $\vec{x} \in M$ mit $\|\vec{\xi} - \vec{x}\| < \delta$. \vec{f} heißt stetig auf M , falls \vec{f} in jedem $\vec{\xi} \in M$ stetig ist. \vec{f} heißt gleichmäßig stetig auf M , wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt, so dass für $\vec{x}, \vec{y} \in M$ mit $\|\vec{x} - \vec{y}\| < \delta$ immer $\|\vec{f}(\vec{x}) - \vec{f}(\vec{y})\| < \varepsilon$ gilt.

Wie im Fall $n = 1$ steht der Stetigkeitsbegriff in engem Zusammenhang zum Grenzwertbegriff für Funktionen. Dieser ist wiederum mit der Konvergenz von Punktfolgen verbunden.

DEFINITION 7.4.3 (Funktionsgrenzwerte)

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ für $M \subseteq \mathbb{R}^n$ und $\vec{a} \in \mathbb{R}^n$ ein Häufungspunkt von M . Wir sagen: $\vec{f}(\vec{x})$ strebt oder konvergiert gegen ein $\vec{b} \in \mathbb{R}^l$ für $\vec{x} \rightarrow \vec{a}$, Schreibweise

$$\vec{f}(\vec{x}) \longrightarrow \vec{b} \quad \text{oder} \quad \lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{b},$$

falls gilt: zu jedem $\varepsilon > 0$ gibt es ein $\delta = \delta(\varepsilon) > 0$, so dass für alle $\vec{x} \in M$ mit $0 < \|\vec{x} - \vec{a}\| < \delta$ immer $\|\vec{f}(\vec{x}) - \vec{b}\| < \varepsilon$ gilt.

SATZ 7.4.1

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ mit $M \subseteq \mathbb{R}^n$ und einem Häufungspunkt $\vec{a} \in \mathbb{R}^n$ von M der in M liegt. Dann ist \vec{f} genau dann stetig in \vec{a} , wenn $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a})$ gilt.

BEWEIS

Das folgt direkt aus der Definition der Stetigkeit und des Grenzwerts. □

SATZ 7.4.2

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ und $\vec{a} \in \mathbb{R}^n$ ein Häufungspunkt von M . Dann gilt $\vec{f}(\vec{x}) \rightarrow \vec{f}(\vec{a})$ für $\vec{x} \rightarrow \vec{a}$ genau dann, wenn für jede in $M \setminus \{\vec{a}\}$ gelegene Folge (\vec{x}_k) mit $\vec{x}_k \rightarrow \vec{a}$ auch $\vec{f}(\vec{x}_k) \rightarrow \vec{f}(\vec{a})$ gilt.

BEWEIS

Genau wie der Beweis zu Satz 4.5.2. □

SATZ 7.4.3

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ und $\vec{a} \in M$ beliebig. Dann ist \vec{f} in \vec{a} genau dann stetig, wenn \vec{a} ein isolierter Punkt von M ist oder $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a})$ ist.

OHNE BEWEIS □

Die Stetigkeit vektorwertiger Funktionen lässt sich auf die Stetigkeit ihrer reellwertigen Komponenten zurückführen:

SATZ 7.4.4

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ mit $\vec{f}(\vec{x}) = (f_1(\vec{x}), \dots, f_l(\vec{x}))^T$. \vec{f} ist genau dann stetig in einem $\vec{a} \in M$, wenn jede Komponente $f_k(\vec{x})$ für $k = 1 \dots l$ stetig in \vec{a} ist.

BEWEIS

Ist \vec{a} ein isolierter Punkt von M , so ist nach Satz 7.4.3 nichts zu beweisen. Andernfalls ist nach den Sätzen 7.4.2 und 7.4.3 die Stetigkeit äquivalent zur Konvergenz von Folgen. Aus Satz 7.2.1 folgt dann die Behauptung. □

SATZ 7.4.5

Summe, Differenz und Produkte (sowohl Skalarprodukte als auch Produkte mit skalaren Funktionen) in \vec{a} stetiger Funktionen sind in \vec{a} stetig. Für Funktionen aus $\mathcal{F}_n^1(M)$ gilt Entsprechendes auch für den Quotienten, wenn die Nennerfunktion in \vec{a} nicht verschwindet.

BEWEIS

Mit Satz 7.4.4 kann man alles auf Funktionen aus \mathcal{F}_n^1 zurückgeführt werden. Für diese verläuft der Beweis analog zum Beweis von Satz 4.1.2. □

SATZ 7.4.6

Es sei $\vec{g} \in \mathcal{F}_n^l(M)$ mit $M \subseteq \mathbb{R}^n$ mit Werten in $N \subseteq \mathbb{R}^l$ und $\vec{f} \in \mathcal{F}_l^k(N)$. Ist \vec{g} stetig in $\vec{\xi} \in M$ und \vec{f} stetig in $\vec{\eta} = \vec{g}(\vec{\xi})$, so ist $\vec{h}(\vec{x}) = \vec{f}(\vec{g}(\vec{x}))$ stetig in $\vec{\xi} \in M$.

BEWEIS

Ist wie Satz 4.1.4 zu beweisen. □

SATZ 7.4.7

Die Wertemenge $\vec{f}(K)$ einer auf einer kompakten Punktmenge $K \subseteq \mathbb{R}^n$ definierten stetigen Funktion ist kompakt. Insbesondere ist $\vec{f}(K)$ beschränkt.

OHNE BEWEIS □

Stetige Funktionen $\vec{f}: K \rightarrow \mathbb{R}$ in den $\mathbb{R}^1 = \mathbb{R}$ nehmen auf kompakten Mengen damit Maxima und Minima an.

SATZ 7.4.8

Jede auf einer kompakten Punktmenge stetige Funktion ist dort gleichmäßig stetig.

OHNE BEWEIS

□

DEFINITION 7.4.4

Eine Funktionenfolge $(\vec{f}_k)_{k=k_0}^{\infty}$ mit $\vec{f}_k \in \mathcal{F}_n^l(M)$ für ein $M \subseteq \mathbb{R}^n$ heißt gleichmäßig konvergent gegen eine Grenzfunktion $\vec{f} \in \mathcal{F}_n^l(M)$, falls es zu jedem $\varepsilon > 0$ ein $k_0 = k_0(\varepsilon) > 0$ gibt, so dass aus $k > k_0$ und $\vec{x} \in M$ stets $\|\vec{f}_k(\vec{x}) - \vec{f}(\vec{x})\| < \varepsilon$ folgt.

SATZ 7.4.9

Sind die Funktionen $\vec{f}_k \in \mathcal{F}_n^l(M)$ stetig in $\vec{\xi} \in M$, und konvergiert (\vec{f}_k) gleichmäßig auf M gegen $\vec{f} \in \mathcal{F}_n^l(M)$, so ist auch \vec{f} stetig in $\vec{\xi}$.

BEWEIS

Genau wie Satz 4.6.1.

□

Wir schließen das Kapitel mit Beispielen stetiger Funktionen. Nach Satz 7.4.5 ist jedes Polynom P in n Variablen

$$P : \mathbb{R}^n \rightarrow \mathbb{R}, P(\vec{x}) = \sum_{0 \leq k_1, \dots, k_n \leq d} a_{k_1, \dots, k_n} x_1^{k_1} \cdots x_n^{k_n}$$

stetig auf ganz \mathbb{R}^n . Weiterhin sind die Quotienten solcher Polynome $R(\vec{x}) = \frac{P(\vec{x})}{Q(\vec{x})}$ überall stetig, wo $Q(\vec{x}) \neq 0$ ist. Nach Satz 7.4.6 sind auch aus stetigen Funktionen zusammengesetzte Funktionen wie beispielsweise

$$\vec{f}(x_1, x_2, x_3) = \sin \left(\frac{e^{x_1 \cdot x_2} + x_3^2}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \right)$$

auf ihrem Definitionsbereich stetig.

8. Differenzierbarkeit mehrwertiger Funktionen

8.1. Partielle Differenzierbarkeit

DEFINITION 8.1.1

Es sei $\vec{f} = \vec{f}(x_1, \dots, x_n) \in \mathcal{F}_n^l(M)$ mit $M \subseteq \mathbb{R}^n$ und $\vec{x}_0 = (\xi_1, \dots, \xi_n) \in M^0$ ein innerer Punkt. \vec{f} heißt partiell differenzierbar nach x_k in \vec{x}_0 für $1 \leq k \leq n$, falls der Grenzwert

$$\frac{\partial \vec{f}}{\partial x_k}(\vec{x}_0) := \lim_{h \rightarrow 0} \frac{\vec{f}(\xi_1, \dots, \xi_k + h, \dots, \xi_n) - \vec{f}(\xi_1, \dots, \xi_k, \dots, \xi_n)}{h}$$

existiert und endlich ist. Der Vektor

$$\frac{\partial \vec{f}}{\partial x_k}(\vec{x}_0) = \vec{f}_{x_k}(\vec{x}_0)$$

wird als partielle Ableitung von \vec{f} nach x_k an der Stelle \vec{x}_0 bezeichnet.

BEMERKUNG 8.1.1

Die partielle Differenzierbarkeit ist also äquivalent zur Differenzierbarkeit der Funktion

$$g(t) = \vec{f}(\xi_1, \dots, \xi_{k-1}, t, \xi_{k+1}, \dots, \xi_n)$$

in einer reellen Variablen im Punkt $t = \xi_k$. Wie man leicht sieht, ist $\vec{f} = (f_1, \dots, f_l)^T$ partiell differenzierbar nach x_k genau dann, wenn alle Komponentenfunktionen f_j nach x_k partiell differenzierbar sind.

BEISPIEL 8.1.1

Es sei $f(\vec{x}) = e^{x_1 x_2} \sin(x_3)$ aus $\mathcal{F}_3^1(\mathbb{R})$. Dann ist

$$\frac{\partial f}{\partial x_1}(\vec{x}) = x_2 e^{x_1 x_2} \sin(x_3) \quad , \quad \frac{\partial f}{\partial x_2}(\vec{x}) = x_1 e^{x_1 x_2} \sin(x_3) \quad , \quad \frac{\partial f}{\partial x_3}(\vec{x}) = e^{x_1 x_2} \cos(x_3) .$$

DEFINITION 8.1.2

Es sei $f \in \mathcal{F}_n^1(M)$ für $M \subseteq \mathbb{R}$ und $\vec{x}_0 \in M^0$. Es sei $\vec{v} \in \mathbb{R}^n$ irgend ein Vektor der Länge $\|\vec{v}\| = 1$. Dann heißt (falls existent und endlich)

$$f_{\vec{v}}(\vec{x}_0) := \lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h}$$

die Richtungsableitung von f in Richtung \vec{v} .

BEMERKUNG 8.1.2

Die partielle Ableitung f_{x_k} ist eine spezielle Richtungsableitung für $\vec{v} = \vec{e}_k = (0, \dots, \underbrace{1}_k, \dots, 0)^T$.

8.2. Totale Differenzierbarkeit

Nach Satz 5.1.1 ist eine Funktion in einer Variablen überall wo sie differenzierbar ist auch stetig. Hingegen folgt für Funktionen in mehreren Variablen aus der partiellen Differenzierbarkeit nicht die Stetigkeit, wie folgendes Beispiel zeigt:

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

Die Funktion f ist partiell differenzierbar in $(0, 0)$ mit

$$\frac{\partial f}{\partial x_1}(0, 0) = \frac{d}{dx_1} f(x_1, 0)|_{x_1=0} = 0 \quad \text{und} \quad \frac{\partial f}{\partial x_2}(0, 0) = \frac{d}{dx_2} f(0, x_2)|_{x_2=0} = 0 .$$

Für die Punktfolgen (\vec{x}_k) mit $\vec{x}_k = (\frac{1}{k}, 0)^T$ und (\vec{y}_k) mit $\vec{y}_k = (\frac{1}{k}, \frac{1}{k})^T$ gilt aber

$$\lim_{k \rightarrow \infty} \vec{x}_k = \lim_{k \rightarrow \infty} \vec{y}_k = 0 ,$$

jedoch ist

$$\lim_{k \rightarrow \infty} f(\vec{x}_k) = \vec{0} \text{ und } \lim_{k \rightarrow \infty} f(\vec{y}_k) = \lim_{k \rightarrow \infty} \frac{\frac{1}{k^2}}{\frac{2}{k^2}} = \frac{1}{2}.$$

Nach Satz 7.4.2 ist f damit unstetig in $\vec{0}$.

Die partielle Differenzierbarkeit für Funktionen mehrerer Variablen ist also eine wesentlich schwächere Eigenschaft als die Differenzierbarkeit für eine Funktion in einer Variablen. Um die richtige Verallgemeinerung des Differenzierbarkeitsbegriffs auf Funktionen in mehreren Variablen zu finden, lassen wir uns von der Idee der linearen Approximierbarkeit leiten. Nach Satz 5.1.4 ist eine Funktion $f \in \mathcal{F}_1^1$ differenzierbar in x_0 genau dann, wenn f durch eine lineare Funktion $f(x) + a(x - x_0)$ approximiert werden kann, d. h. wenn

$$f(x) = f(x_0) + a(x - x_0) + r(x) \text{ mit } \lim_{x \rightarrow x_0} \frac{r(x)}{|x - x_0|} = 0$$

ist. Die allgemeine lineare Funktion in mehreren Variablen ist durch

$$\vec{f}(\vec{x}_0) + A(\vec{x} - \vec{x}_0)$$

mit einer Matrix $A \in \mathbb{R}^{l \times n}$ gegeben. Die Definition für (totale) Differenzierbarkeit von Funktionen mehrerer Variablen lautet somit:

DEFINITION 8.2.1

Eine Funktion $\vec{f} \in \mathcal{F}_n^l(M)$ heißt in einem inneren Punkt $\vec{x}_0 \in M^0$ (total) differenzierbar, wenn eine Matrix $A \in \mathbb{R}^{l \times n}$ existiert, sowie eine auf einer Umgebung $U \subseteq M$ von \vec{x}_0 definierte Funktion $\vec{r} \in \mathcal{F}_n^l$ mit

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{\vec{r}(\vec{x})}{\|\vec{x} - \vec{x}_0\|} = \vec{0},$$

so dass gilt:

$$\forall \vec{x} \in U : \vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + A(\vec{x} - \vec{x}_0) + \vec{r}(\vec{x}).$$

SATZ 8.2.1

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$, $\vec{x}_0 \in M^0$ und \vec{f} in \vec{x}_0 differenzierbar. Dann ist

$$\frac{\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0)}{\|\vec{x} - \vec{x}_0\|}$$

beschränkt in $U_\varepsilon(\vec{x}_0) \setminus \{\vec{x}_0\}$ für ein genügend kleines $\varepsilon > 0$.

BEWEIS

Es gilt mit A und \vec{r} aus der vorigen Definition

$$\frac{\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0)}{\|\vec{x} - \vec{x}_0\|} = A \frac{\vec{x} - \vec{x}_0}{\|\vec{x} - \vec{x}_0\|} + \frac{\vec{r}(\vec{x})}{\|\vec{x} - \vec{x}_0\|}.$$

Der 2. Summand ist wegen der Konvergenzbedingung an \vec{r} beschränkt. Der 1. Summand ist ebenfalls beschränkt, da

$$\vec{y} = \frac{\vec{x} - \vec{x}_0}{\|\vec{x} - \vec{x}_0\|}$$

vom Betrag 1 ist und $A\vec{y}$ als stetige Funktion auf einer kompakten Menge beschränkt ist. □

SATZ 8.2.2

Wenn $\vec{f}(\vec{x})$ in \vec{x}_0 differenzierbar ist, so ist \vec{f} in \vec{x}_0 stetig.

BEWEIS

Das folgt direkt aus Satz 8.2.1. □

DEFINITION 8.2.2

Die Funktion $\vec{f} \in \mathcal{F}_n^l$ sei nach allen Variablen x_1, \dots, x_n partiell differenzierbar. Die Matrix dieser partiellen Ableitungen

$$\frac{d\vec{f}}{d\vec{x}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_l}{\partial x_1} & \frac{\partial f_l}{\partial x_2} & \cdots & \frac{\partial f_l}{\partial x_n} \end{pmatrix}$$

heißt die Funktionalmatrix von \vec{f} . Im Fall $l = 1$ ($\vec{f} = f$) nennt man die transponierte Funktionalmatrix, also den Spaltenvektor

$$\text{grad}(f) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)^T,$$

den Gradienten von f (auch mit ∇f bezeichnet, sprich „nabla f “). Die Determinante einer quadratischen Funktionalmatrix

$$\det \frac{d\vec{f}}{d\vec{x}}$$

heißt Funktionaldeterminante oder Jacobische Determinante von \vec{f} . Die Spur einer quadratischen Funktionalmatrix wird als Divergenz bezeichnet:

$$\text{div} \vec{f} = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \cdots + \frac{\partial f_n}{\partial x_n}.$$

Funktionalmatrizen, Gradienten, Funktionaldeterminanten und Divergenzen an einer Stelle \vec{x}_0 werden durch Anbringen von (\vec{x}_0) bezeichnet:

$$\frac{d\vec{f}}{d\vec{x}}(\vec{x}_0), \text{grad} f(\vec{x}_0), \det \frac{d\vec{f}}{d\vec{x}}(\vec{x}_0), \text{div} \vec{f}(\vec{x}_0).$$

SATZ 8.2.3

Es sei $\vec{f} \in \mathcal{F}_n^l(M)$ in einem $\vec{x}_0 \in M^0$ differenzierbar. Dann existieren in \vec{x}_0 alle partiellen Ableitungen $\vec{f}_{x_1}, \dots, \vec{f}_{x_n}$, und die Matrix A aus Definition 8.2.1 ist gerade die Funktionalmatrix von \vec{f} in \vec{x}_0 .

BEWEIS

Es gilt

$$(*) \quad \lim_{\vec{x} \rightarrow \vec{x}_0} \frac{\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0) - A(\vec{x} - \vec{x}_0)}{\|\vec{x} - \vec{x}_0\|} = \vec{0}.$$

Wir setzen $\vec{x} = \vec{x}_0 + h\vec{e}_k$, dann folgt aus (*)

$$\lim_{h \rightarrow 0} \frac{\vec{f}(\vec{x}_0 + h\vec{e}_k) - \vec{f}(\vec{x}_0) - hA\vec{e}_k}{h} \cdot \frac{h}{|h|} = \vec{0} \text{ also}$$

$$\lim_{h \rightarrow 0} \frac{\vec{f}(\vec{x}_0 + h\vec{e}_k) - \vec{f}(\vec{x}_0)}{h} = A\vec{e}_k.$$

Daraus folgt die Existenz von $\frac{d\vec{f}}{d\vec{x}}$ im Punkt \vec{x}_0 , und

$$A\vec{e}_k = \left(\frac{\partial f_1}{\partial x_k}, \dots, \frac{\partial f_l}{\partial x_k} \right)^T.$$

Jedoch ist $A\vec{e}_k$ gerade die k -te Spalte von A , woraus die Behauptung folgt. □

BEMERKUNG 8.2.1

Es folgt unmittelbar, dass die Matrix einer linearen Funktion ihre Funktionalmatrix ist:

$$\frac{d}{d\vec{x}}(A\vec{x} + \vec{a}) = A.$$

SATZ 8.2.4

Es sei $f \in \mathcal{F}_n^1$ und $\vec{x}_0 \in M$ ein innerer Punkt, sowie f differenzierbar in \vec{x}_0 . Dann existiert die Richtungsableitung $f_{\vec{v}}(\vec{x}_0)$ für jeden Vektor $\vec{v} \in \mathbb{R}^n$ der Länge Eins, und es gilt

$$f_{\vec{v}}(\vec{x}_0) = \vec{v} \cdot \text{grad}f(\vec{x}_0).$$

BEWEIS

Wegen der Differenzierbarkeit gilt

$$\lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0) - h\vec{v} \cdot \text{grad}f(\vec{x}_0)}{\|h\vec{v}\|} = 0,$$

und da $\|h\vec{v}\| = h$ ist

$$\lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0) - h\vec{v} \cdot \text{grad}f(\vec{x}_0)}{h} = 0.$$

Also auch

$$\lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h} = \vec{v} \cdot \text{grad}f(\vec{x}_0).$$

□

SATZ 8.2.5

Es sei $f \in \mathcal{F}_n^1(M)$, $\vec{x}_0 \in M^0$ und f in \vec{x}_0 differenzierbar. Ist $\text{grad}f(\vec{x}_0) = \vec{0}$, so verschwinden alle Richtungsableitungen in \vec{x}_0 . Ist $\text{grad}f(\vec{x}_0) \neq \vec{0}$, so gibt es unter allen Richtungsableitungen $f_{\vec{v}}(\vec{x}_0)$ eine größte, und zwar die Ableitung in Richtung des Gradienten $\text{grad}f(\vec{x}_0)$. Ihr Wert ist $\|\text{grad}f(\vec{x}_0)\|$.

BEWEIS

Nach der Cauchy-Schwarzschen Ungleichung gilt

$$|f_{\vec{v}}(\vec{x}_0)| \leq \|\vec{v}\| \cdot \|\text{grad}f(\vec{x}_0)\| = \|\text{grad}f(\vec{x}_0)\|.$$

Für $\text{grad}f(\vec{x}_0) \neq \vec{0}$ haben wir für

$$\vec{v}_0 = \frac{\text{grad}f(\vec{x}_0)}{\|\text{grad}f(\vec{x}_0)\|}$$

die Abschätzung

$$0 \leq f_{\vec{v}_0}(\vec{x}_0) = \frac{\text{grad}f(\vec{x}_0) \cdot \text{grad}f(\vec{x}_0)}{\|\text{grad}f(\vec{x}_0)\|} = \|\text{grad}f(\vec{x}_0)\|.$$

□

BEMERKUNG 8.2.2

Man sagt daher auch, der Gradient gibt die Richtung des „stärksten Anstiegs“ von f an, und sein Betrag ist gerade dieser stärkste Anstieg. Für die zu \vec{v} entgegengesetzte Richtung $-\vec{v}$ gilt

$$f_{-\vec{v}}(\vec{x}_0) = (-\vec{v}) \cdot \text{grad}f(\vec{x}_0) = -f_{\vec{v}}(\vec{x}_0),$$

also ist die Gegenrichtung des Gradienten in \vec{x}_0 die Richtung des „stärksten Abstiegs“ von f , und dieser stärkste Abstieg wird durch $-\|\text{grad}f(\vec{x}_0)\|$ gegeben.

8.3. Ableitungsregeln

In diesem Abschnitt verallgemeinern wir die wichtigsten Ableitungsregeln auf vektorwertige Funktionen:

SATZ 8.3.1 (Mehrdimensionale Umkehrregel)

Es sei $\vec{f} \in \mathcal{F}_n^n(M)$ auf der offenen Menge $M \subseteq \mathbb{R}^n$ stetig partiell differenzierbar mit $\det\left(\frac{d\vec{f}}{d\vec{x}}\right)(\vec{x}_0) \neq 0$ in einem Punkt $\vec{x}_0 \in M$, dann gibt es eine Umgebung U von \vec{x}_0 in M , so dass $V = \vec{f}(U)$ eine Umgebung von $\vec{y}_0 = \vec{f}(\vec{x}_0)$ ist, und \vec{f} auf U eindeutig umkehrbar ist. Die Umkehrung ist auf V stetig differenzierbar mit

$$\frac{d\vec{f}^{-1}}{d\vec{y}}(\vec{y}) = \left(\frac{d\vec{f}}{d\vec{x}}(\vec{f}^{-1}(\vec{y})) \right)^{-1} \in \mathbb{R}^{n \times n}$$

für alle $\vec{y} \in V$.

OHNE BEWEIS

□

SATZ 8.3.2 (Mehrdimensionale Kettenregel)

Es sei $\vec{g} \in \mathcal{F}_n^l(M)$ für $M \subseteq \mathbb{R}^n$ mit Werten in $N \subseteq \mathbb{R}^l$, sowie $\vec{f} \in \mathcal{F}_l^k(N)$. Für $\vec{x}_0 \in M^0$ sei $\vec{y}_0 = \vec{g}(\vec{x}_0)$ in N^0 . Die Funktion \vec{f} sei in \vec{y}_0 , und die Funktion \vec{g} in \vec{x}_0 differenzierbar. Dann gilt: $\vec{s} = \vec{f} \circ \vec{g}$ ist im Punkt \vec{x}_0 differenzierbar mit

$$\frac{d\vec{s}}{d\vec{x}}(\vec{x}_0) = \frac{d\vec{f}}{d\vec{y}}(\vec{y}_0) \cdot \frac{d\vec{g}}{d\vec{x}}(\vec{x}_0).$$

BEWEIS

Nach Definition der totalen Differenzierbarkeit gilt

$$\vec{f}(\vec{y}_0 + \vec{t}) = \vec{f}(\vec{y}_0) + B\vec{y}_0 + \vec{\varrho}(\vec{t}) \quad \text{mit} \quad \lim_{\vec{t} \rightarrow \vec{0}} \frac{\vec{\varrho}(\vec{t})}{\|\vec{t}\|} = \vec{0} \quad \text{und}$$

$$\vec{g}(\vec{x}_0 + \vec{h}) = \vec{g}(\vec{x}_0) + A\vec{x}_0 + \vec{\vartheta}(\vec{h}) \quad \text{mit} \quad \lim_{\vec{h} \rightarrow \vec{0}} \frac{\vec{\vartheta}(\vec{h})}{\|\vec{h}\|} = \vec{0},$$

wobei B und A jeweils die Funktionalmatrizen von \vec{f} in \vec{y}_0 bzw. von \vec{g} in \vec{x}_0 sind. Wir setzen $\vec{y}_0 = \vec{g}(\vec{x}_0)$, $\vec{t} = \vec{g}(\vec{x}_0 + \vec{h}) - \vec{g}(\vec{x}_0)$. Dann folgt aus den obigen Gleichungen $\vec{t} = A\vec{h} + \vec{\vartheta}(\vec{h})$ und somit

$$\vec{s}(\vec{x}_0 + \vec{h}) = \vec{f}(\vec{g}(\vec{x}_0 + \vec{h})) = \vec{f}(\vec{y}_0 + \vec{t}) = \vec{f}(\vec{y}_0) + B\vec{t} + \vec{\varrho}(\vec{t}) = \vec{s}(\vec{x}_0) + BA\vec{h} + B\vec{\vartheta}(\vec{h}) + \vec{\varrho}(A\vec{h} + \vec{\vartheta}(\vec{h})).$$

Es gibt nun ein $r > 0$ und ein $c > 0$, so dass für $\|\vec{h}\| < r$ gilt:

$$\|A\vec{h} + \vec{\vartheta}(\vec{h})\| \leq c\|\vec{h}\|.$$

Daraus folgt

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{B\vec{\vartheta}(\vec{h}) + \vec{\varrho}(A\vec{h} + \vec{\vartheta}(\vec{h}))}{\|\vec{h}\|} = \vec{0}.$$

Also gilt

$$\frac{d\vec{s}}{d\vec{x}}(\vec{x}_0) = B \cdot A = \frac{d\vec{f}}{d\vec{y}}(\vec{y}_0) \cdot \frac{d\vec{g}}{d\vec{x}}(\vec{x}_0).$$

□

BEMERKUNG 8.3.1

Wir erhalten folgenden Spezialfall: ist $\vec{g} \in \mathcal{F}_1^n$ und $f \in \mathcal{F}_n^1$, also

$$s(x) = f(\vec{g}(x)) = f(g_1(x), \dots, g_n(x)),$$

so gilt

$$\frac{ds}{dx} = \frac{df}{d\vec{x}} \cdot \frac{d\vec{g}}{dx} = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \frac{\partial g_j}{\partial x}.$$

8.4. Partielle Ableitungen höherer Ordnung

DEFINITION 8.4.1

Die partielle Ableitung \vec{f}_{x_m} einer Funktion \vec{f} sei in einer offenen Menge definiert. Existiert die partielle Ableitung von \vec{f}_{x_m} nach x_n , so wird diese mit $\vec{f}_{x_m x_n}$ bezeichnet. Weitere Schreibweisen:

$$\frac{\partial^2 \vec{f}}{\partial x_m \partial x_n} \text{ oder } \frac{\partial^2}{\partial x_m \partial x_n} \vec{f},$$

im Falle $n = m$ auch

$$\frac{\partial^2 \vec{f}}{\partial x_m^2} \text{ oder } \frac{\partial^2}{\partial x_m^2} \vec{f}.$$

Dieses Verfahren kann fortgesetzt werden. Im Falle der Existenz ist

$$\vec{f}_{x_{m_1} x_{m_2} \dots x_{m_k}} = \frac{\partial^k \vec{f}}{\partial x_{m_1} \partial x_{m_2} \dots \partial x_{m_k}}$$

als eine durch k -malige partielle Ableitung aus \vec{f} hervorgehende Funktion definiert. Man spricht von einer partiellen Ableitung k -ter Ordnung.

DEFINITION 8.4.2

Wenn für die Funktion \vec{f} in einer offenen Menge M alle partiellen Ableitungen bis zur k -ten Ordnung existieren und stetig sind, so sagt man, \vec{f} sei in M k -mal stetig differenzierbar, oder \vec{f} gehöre der Differentiationsklasse $\mathcal{C}^k(M)$ an. Zusätzlich definiert man $\mathcal{C}^0(M)$ als die Klasse der auf M stetigen Funktionen und $\mathcal{C}^\infty(M)$ als die Klasse der Funktionen auf M , zu denen partielle Ableitungen beliebiger Ordnung existieren und stetig sind.

SATZ 8.4.1

Wenn $f \in \mathcal{F}_n^1$ der Differentiationsklasse \mathcal{C}^k angehört für $k \geq 2$, so ist jede partielle Ableitung k -ter Ordnung von f unabhängig von der Reihenfolge der Einzelableitungen, d. h. es ist

$$f_{x_{m_{\pi(1)}} \dots x_{m_{\pi(k)}}} = f_{x_{m_1} \dots x_{m_k}}$$

für jede Permutation π von $1, \dots, k$.

OHNE BEWEIS

□

BEMERKUNG 8.4.1

Im Spezialfall $k = n = 2$ schreibt sich das als $f_{xy} = f_{yx}$.

DEFINITION 8.4.3

Wird eine Funktion $f \in \mathcal{C}^k$ ν_1 -mal nach x_1 , ν_2 -mal nach x_2 usw. bis ν_n -mal nach x_n partiell abgeleitet (mit $\nu_1 + \dots + \nu_n = k$), so schreibt man auch

$$\frac{\partial^k f}{\partial x_1^{\nu_1} \dots \partial x_n^{\nu_n}} = \frac{\partial^k}{\partial x_1^{\nu_1} \dots \partial x_n^{\nu_n}} f.$$

Wir setzen

$$\Delta f = \frac{\partial^2}{\partial x_1^2} f + \cdots + \frac{\partial^2}{\partial x_n^2} f .$$

Δ heißt Laplace-Operator. Offenbar gilt $\Delta f = \operatorname{div}(\operatorname{grad} f)$.

SATZ 8.4.2

Es sei $M \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{F}_n^l(M)$. Die partiellen Ableitungen f_{x_1}, \dots, f_{x_n} mögen in M existieren und stetig sein. Dann ist f in M (total) differenzierbar.

BEWEIS

Es sei ohne Einschränkung $f \in \mathcal{F}_n^1(M)$ und $\vec{x}_0 \in M = M^0$. Wir haben zu zeigen, dass

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - \vec{h} \cdot \operatorname{grad} f(\vec{x}_0)|}{\|\vec{h}\|} = 0$$

ist. Dazu wählen wir $r > 0$ so klein, dass alle $\vec{x} \in \mathbb{R}^n$ mit $\|\vec{x} - \vec{x}_0\| < r$ zu M gehören. Es sei $\vec{h} = (h_1, \dots, h_n)^T$ und $\|\vec{h}\| < r$. Wir setzen

$$\vec{h}_m = (h_1, \dots, h_m, 0, \dots, 0)^T = \sum_{k=1}^m h_k \vec{e}_k, \quad 0 \leq m \leq n,$$

also $\vec{h}_0 = \vec{0}$ und $\vec{h}_n = \vec{h}$. Wegen $\|\vec{h}_m\| \leq \|\vec{h}\|$ ist auch $\|\vec{h}_m\| < r$. Wir haben

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \sum_{m=1}^n (f(\vec{x}_0 + \vec{h}_m) - f(\vec{x}_0 + \vec{h}_{m-1}))$$

und setzen

$$g_m(h) = f(\vec{x}_0 + \vec{h}_{m-1} + h \vec{e}_m), \quad 0 \leq h \leq h_m.$$

Nach dem Mittelwertsatz existiert ein $t \in (0, h_m)$ mit

$$g_m(h_m) - g_m(0) = h_m g'(t).$$

Wir setzen $\vec{x}_m = \vec{x}_0 + \vec{h}_{m-1} + t \vec{e}_m$ und erhalten

$$f(\vec{x}_0 + \vec{h}_m) - f(\vec{x}_0 + \vec{h}_{m-1}) = g_m(h_m) - g_m(0) = h_m f_{x_m}(\vec{x}_m).$$

Es ist demnach

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \sum_{m=1}^n h_m f_{x_m}(\vec{x}_m).$$

Weiter gilt:

$$\vec{h} \cdot \operatorname{grad} f(\vec{x}_0) = \sum_{m=1}^n h_m f_{x_m}(\vec{x}_0),$$

also auch

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - \vec{h} \cdot \operatorname{grad} f(\vec{x}_0) = \sum_{m=1}^n h_m (f_{x_m}(\vec{x}_m) - f_{x_m}(\vec{x}_0)) \text{ und}$$

$$\frac{|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - \vec{h} \cdot \operatorname{grad} f(\vec{x}_0)|}{\|\vec{h}\|} \leq \sum_{m=1}^n \frac{|h_m|}{\|\vec{h}\|} |f_{x_m}(\vec{x}_m) - f_{x_m}(\vec{x}_0)| \leq \sum_{m=1}^n |f_{x_m}(\vec{x}_m) - f_{x_m}(\vec{x}_0)|.$$

Wegen der Stetigkeit der f_{x_m} gilt

$$\lim_{\vec{h} \rightarrow \vec{0}} (f_{x_m}(\vec{x}_m) - f_{x_m}(\vec{x}_0)) = 0$$

und damit

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - \vec{h} \cdot \operatorname{grad} f(\vec{x}_0)|}{\|\vec{h}\|} = 0,$$

d. h. die (totale) Differenzierbarkeit. □

8.5. Kurven und Gebiete

DEFINITION 8.5.1

Der in der Schule benutzte Begriff der Kurve wird wie folgt präzisiert:

- (i) Jede auf I stetige Funktion $\gamma \in \mathcal{F}_1^n$, $\gamma : I \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}(t)$, bei welcher der Definitionsbereich $I \subseteq \mathbb{R}$ ein kompaktes Intervall ist, heißt eine Kurve im \mathbb{R}^n . Die Menge der Punkte $\Gamma = \{\vec{x}(t) \mid t \in [a, b]\}$ heißt Bogen oder Träger der Kurve γ . Man sagt auch, $\gamma : t \mapsto \vec{x}(t)$ sei eine Parameterdarstellung von Γ mit Parameter $t \in I$, oder γ sei eine orientierte Kurve.
- (ii) Ist $I = [a, b]$, so heißen die Punkte $\vec{x}(a)$ und $\vec{x}(b)$ Anfangspunkt und Endpunkt von γ .
- (iii) Die Kurve $\gamma : t \mapsto \vec{x}(t) = (x_1(t), \dots, x_n(t))^T$ heißt differenzierbar, falls die Komponentenfunktionen $x_k(t)$ differenzierbar sind für $1 \leq k \leq n$. Für eine differenzierbare Kurve γ heißt der Vektor $\frac{d}{dt}\vec{x}(t) = (x'_1(t), \dots, x'_n(t))^T$ der Tangentenvektor an die Kurve γ im Punkt $\vec{x}(t)$ (es ist auch die Schreibweise $\vec{x}' = (\dot{x}_1(t), \dots, \dot{x}_n(t))^T$ üblich).
- (iv) Eine stetig differenzierbare Kurve heißt glatt, falls der Tangentenvektor nirgends verschwindet, d. h. falls $\frac{d}{dt}\vec{x}(t) \neq \vec{0}$ ist für alle $t \in I$.
- (v) Die Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt stückweise differenzierbar, wenn das Intervall $I = [a, b]$ eine Zerlegung $a = t_0 < t_1 < \dots < t_k = b$ besitzt, so dass für $1 \leq j \leq k$ die Teilkurven

$$\gamma|_{[t_{j-1}, t_j]} : [t_{j-1}, t_j] \rightarrow \mathbb{R}^n, t \mapsto \vec{x}(t)$$

differenzierbar sind. Entsprechend erklärt man stückweise glatt.

BEISPIEL 8.5.1

Für $\vec{a}, \vec{b} \in \mathbb{R}^n$ mit $\vec{a} \neq \vec{b}$ ist die orientierte Verbindungsstrecke

$$\gamma : [0, 1] \rightarrow \mathbb{R}^n, t \mapsto \vec{x}(t) = \vec{a} + t(\vec{b} - \vec{a})$$

glatt wegen $\vec{x}'(t) = \vec{b} - \vec{a} \neq \vec{0}$.

DEFINITION 8.5.2

Es seien $\gamma_j : I_j \rightarrow \mathbb{R}^n$ jeweils Kurven.

- (i) Die Kurve $\gamma_1 : I_1 \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}_1(t)$ heißt äquivalent zur Kurve $\gamma_2 : I_2 \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}_2(t)$, wenn es eine stetige und streng wachsende Bijektion $\varphi : I_1 \rightarrow I_2$ gibt mit $\vec{x}_2(\varphi(t)) = \vec{x}_1(t)$. Man sagt auch, dass γ_2 eine Umparametrisierung von γ_1 ist. Wie man leicht sieht, ist die Äquivalenz von Wegen eine Äquivalenzrelation.
- (ii) Wir sagen, dass die Kurve $\gamma_1 : I_1 \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}_1(t)$ die zur Kurve $\gamma_2 : I_1 \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}_2(t)$ entgegengesetzt orientierte Kurve ist, wenn es eine stetige und streng wachsende Bijektion $\varphi : I_1 \rightarrow I_2$ gibt mit $\vec{x}_2(\varphi(t)) = \vec{x}_1(a + b - t)$ für alle $t \in [a, b]$.

BEISPIEL 8.5.2

Es seien $\vec{a}, \vec{b} \in \mathbb{R}^n$, dann sind die Kurven

$$\begin{aligned} \gamma_1 : [0, 1] &\rightarrow \mathbb{R}^n, t \mapsto \vec{x}_1(t) = \vec{a} + t(\vec{b} - \vec{a}) \\ \gamma_2 : [0, \frac{1}{2}] &\rightarrow \mathbb{R}^n, t \mapsto \vec{x}_2(t) = \vec{a} + 2t(\vec{b} - \vec{a}) \end{aligned}$$

äquivalent mit $\vec{x}_2(\varphi(t)) = \vec{x}_1(t)$ für $\varphi(t) = \frac{1}{2}t$. Der Träger ist jeweils die Verbindungsstrecke von \vec{a} und \vec{b} .

BEISPIEL 8.5.3

Die Kurven

$$\begin{aligned} \gamma_1 : [0, \pi] &\rightarrow \mathbb{R}^2, t \mapsto \vec{x}_1(t) = (\cos t, \sin t)^T \\ \gamma_2 : [0, \pi] &\rightarrow \mathbb{R}^2, t \mapsto \vec{x}_2(t) = (\cos(\pi - t), \sin(\pi - t))^T \end{aligned}$$

sind entgegengesetzt orientiert, denn es ist $\vec{x}_2(\varphi(t)) = \vec{x}_1(\pi - t)$ mit $\varphi(t) = t$. Der Träger ist jeweils der obere Einheitskreis $\{(x, y)^T \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \geq 0\}$.

BEMERKUNG 8.5.1

Wir können Kurven offenbar auf drei verschiedene Arten betrachten:

- (i) Kurven als Parameterdarstellungen $t \mapsto \vec{x}(t)$,
- (ii) Kurven als Äquivalenzklassen von Parameterdarstellungen,
- (iii) Kurven als Bögen, d. h. als Punktmengen von Kurven nach (i).

Trivialerweise besitzen äquivalente Kurven gleichen Träger. Dasselbe gilt für entgegengesetzt orientierte Kurven. Man prüft leicht nach, dass die wichtigsten Merkmale einer Kurve, insb. der Träger und die (noch zu definierende) Länge, sich nicht ändern, wenn man die Kurve umparametrisiert. Daher ist (ii) die gängigste Betrachtungsweise. Sie erlaubt es beispielsweise, als Parameterintervall stets $[0, 1]$ zu benutzen. Wir verwenden im Folgenden die Darstellung (i) und überlassen es dem Leser nachzuprüfen, dass die Aussagen tatsächlich unabhängig von der Parametrisierung sind.

DEFINITION 8.5.3

Unter der Länge einer stetig differenzierbaren Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ verstehen wir die Zahl

$$L(\gamma) := \int_a^b \|\dot{\vec{x}}(t)\| dt.$$

BEISPIEL 8.5.4

Der Längenbegriff entspricht nicht der anschaulichen Länge der Trägermenge: es seien $\vec{a} \neq \vec{b}$ Punkte im \mathbb{R}^n . Die Länge der Verbindungsstrecke

$$\gamma_1 : [0, 1] \rightarrow \mathbb{R}^n, t \mapsto \vec{x}(t) = \vec{a} + t(\vec{b} - \vec{a})$$

ist $\|\vec{b} - \vec{a}\|$, dagegen besitzt die Kurve

$$\gamma_2 : [0, 1] \rightarrow \mathbb{R}^n, t \mapsto \vec{x}(t) = \begin{cases} \vec{a} + 2t(\vec{b} - \vec{a}) & \text{falls } t \in [0, \frac{1}{2}] \\ \vec{a} + (2 - 2t)(\vec{b} - \vec{a}) & \text{falls } t \in [\frac{1}{2}, 1] \end{cases}$$

die doppelte Länge $2\|\vec{b} - \vec{a}\|$, aber den gleichen Träger.

DEFINITION 8.5.4

Es sei $M \subseteq \mathbb{R}^n$ irgend eine Menge:

- (i) M heißt sternförmig, falls es ein $\vec{x} \in M$ gibt, so dass für jedes $\vec{y} \in M$ die Verbindungsstrecke von \vec{x} nach \vec{y} komplett in M liegt.
- (ii) M heißt konvex, wenn je zwei Punkte aus M sich durch eine ganz in M verlaufende Strecke verbinden lassen.
- (iii) M heißt zusammenhängend, wenn je zwei Punkte von M sich durch eine ganz in M verlaufende Kurve verbinden lassen.
- (iv) M heißt Gebiet, falls M offen und zusammenhängend ist.

8.6. Mittelwertsatz und Taylorscher Satz

SATZ 8.6.1 (Mittelwertsatz der Differentialrechnung im \mathbb{R}^n)

Die Funktion $f \in \mathcal{F}_n^1(M)$ sei auf einer offenen Menge $M \subseteq \mathbb{R}^n$ definiert und dort differenzierbar. Zu jedem Paar von Punkten $\vec{x}_0, \vec{x} \in M$, deren Verbindungsstrecke in M liegt, gibt es ein $d \in (0, 1)$, so dass gilt:

$$f(\vec{x}) = f(\vec{x}_0) + (\vec{x} - \vec{x}_0) \cdot \text{grad} f(\vec{x}_0 + d(\vec{x} - \vec{x}_0)).$$

BEWEIS

Wir betrachten die Funktion $h(t) = f(\vec{x}_0 + t(\vec{x} - \vec{x}_0))$ für $t \in [0, 1]$. Es ist $h(t) = f(\vec{g}(t))$ mit $\vec{g}(t) = \vec{x}_0 + t(\vec{x} - \vec{x}_0)$. Nach der Kettenregel (Satz 8.3.2) ist

$$h'(t) = \frac{df}{d\vec{y}}(\vec{g}(t)) \cdot \frac{d\vec{g}}{dt}(t) = \text{grad}f(\vec{x}_0 + t(\vec{x} - \vec{x}_0)) \cdot (\vec{x} - \vec{x}_0).$$

Nach dem (eindimensionalen) Mittelwertsatz 5.4.1(ii) ist

$$f(\vec{x}) - f(\vec{x}_0) = h(1) - h(0) = h'(d) = (\vec{x} - \vec{x}_0) \cdot \text{grad}f(\vec{x}_0 + d(\vec{x} - \vec{x}_0))$$

für ein $d \in (0, 1)$, woraus die Behauptung folgt. \square

SATZ 8.6.2

Eine in einem Gebiet \mathcal{G} differenzierbare Funktion $f \in \mathcal{F}_n^1(\mathcal{G})$, deren sämtliche partiellen Ableitungen identisch verschwinden, ist eine Konstante.

BEWEIS

Es seien $\vec{a}, \vec{b} \in \mathcal{G}$. Da \mathcal{G} zusammenhängend ist, lassen sich \vec{a} und \vec{b} durch eine in \mathcal{G} verlaufende Kurve $\gamma: [\alpha, \beta] \rightarrow \mathbb{R}^n$, $t \mapsto \vec{x}(t)$ mit $\vec{x}(\alpha) = \vec{a}$ und $\vec{x}(\beta) = \vec{b}$ verbinden. Wir nehmen an, es gäbe ein t , so dass $f(\vec{x}(t)) \neq f(\vec{a})$ wäre und setzen $t_0 = \inf\{t \in [\alpha, \beta] \mid f(\vec{x}(t)) \neq f(\vec{a})\}$. Da $f(\vec{x}(t)) = f(\vec{a}) = f(\vec{x}(\alpha))$ für alle $t < t_0$ ist, folgt wegen der Stetigkeit von γ und $f: f(\vec{x}(t_0)) = f(\vec{a})$. Da \mathcal{G} offen ist gibt es ein $\varepsilon > 0$, so dass $\vec{x} \in \mathcal{G}$ ist für $\|\vec{x} - \vec{x}(t_0)\| < \varepsilon$. Wegen der Stetigkeit von γ gibt es ein $\delta = \delta(\varepsilon) > 0$, so dass $\|\vec{x}(t) - \vec{x}(t_0)\| < \varepsilon$ ist für alle t mit $|t - t_0| < \delta$. Aus der Definition von t_0 folgt die Existenz von t_1 mit $|t_1 - t_0| < \delta$, so dass $f(\vec{x}(t_1)) \neq f(\vec{a}) = f(\vec{x}(\alpha))$ ist. Da die Verbindungsstrecke von $\vec{x}(t_0)$ und $\vec{x}(t_1)$ ganz zu \mathcal{G} gehört, gilt nach Satz 8.6.1 für ein $\vartheta \in (0, 1)$ die Gleichung

$$f(\vec{x}(t_1)) = f(\vec{x}(t_0)) + (\vec{x}(t_1) - \vec{x}(t_0)) \cdot \text{grad}f(\vec{x}(t_0) + \vartheta(\vec{x}(t_1) - \vec{x}(t_0))) = f(\vec{x}(t_0)),$$

ein Widerspruch. Daher gibt es kein $t \in [\alpha, \beta]$ mit $f(\vec{x}(t)) \neq f(\vec{a})$, insbesondere gilt $f(\vec{b}) = f(\vec{x}(\beta)) = f(\vec{a})$. Da aber $\vec{a}, \vec{b} \in \mathcal{G}$ beliebig waren folgt, dass f auf \mathcal{G} konstant ist. \square

Wir kommen zur Formulierung des Satzes von Taylor für Funktionen mehrerer Variablen. Es sei $M \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{F}_n^1(M)$ gehöre in M der Differentiationsklasse \mathcal{C}^{m+1} an. Die Punkte \vec{x}_0 und $\vec{x} = \vec{x}_0 + \vec{h}$ samt ihrer Verbindungsstrecke mögen zu M gehören. Wir verallgemeinern die Methode, die zum Beweis des Mittelwertsatzes 8.6.1 führte. Dort betrachteten wir die Hilfsfunktion $g(t) = f(\vec{x}_0 + t\vec{h})$ für $t \in [0, 1]$, und wendeten auf sie den Mittelwertsatz an. Nun wenden wir auf sie den Satz 5.6.2 von Taylor in einer Variablen an. Wir erhalten

$$f(\vec{x}_0 + \vec{h}) = g(1) = \sum_{k=0}^m \frac{g^{(k)}(0)}{k!} + \frac{g^{(m+1)}(\vartheta)}{(m+1)!}$$

für ein $\vartheta \in (0, 1)$. Das Hauptproblem besteht darin, die Ableitungen $g^{(k)}(\vartheta)$ zu bestimmen. Dies geschieht durch wiederholte Anwendung der Kettenregel 8.3.2: es sei $\vec{h} = (h_1, \dots, h_n)^T$, wir haben

$$g'(\vartheta) = \frac{df}{d\vec{x}}(\vec{x}_0 + \vartheta\vec{h}) \cdot \frac{d(\vec{x}_0 + t\vec{h})}{dt}(\vartheta) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\vec{x}_0 + \vartheta\vec{h}) h_j$$

oder symbolisch

$$g'(\vartheta) = (\vec{h}\nabla)f(\vec{x}_0 + \vartheta\vec{h}).$$

Weiter erhalten wir

$$g''(\vartheta) = (\vec{h}\nabla)^{(2)}f(\vec{x}_0 + \vartheta\vec{h}) = \sum_{j_1=1}^n \frac{\partial}{\partial x_{j_1}} \left(\sum_{j_2=1}^n \frac{\partial f}{\partial x_{j_2}}(\vec{x}_0 + \vartheta\vec{h}) h_{j_2} \right) h_{j_1}$$

$$= \sum_{j_1, j_2=1}^n \frac{\partial^2 f}{\partial x_{j_1} \partial x_{j_2}}(\vec{x}_0 + \vartheta \vec{h}) h_{j_1} h_{j_2}$$

und allgemein

$$g^{(k)}(\vartheta) = (\vec{h} \nabla)^{(k)} f(\vec{x}_0 + \vartheta \vec{h}) = \sum_{j_1, \dots, j_k=1}^n \frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}}(\vec{x}_0 + \vartheta \vec{h}) h_{j_1} \cdots h_{j_k}.$$

Wir erhalten somit

SATZ 8.6.3 (Satz von Taylor, 1. Version)

Es sei $M \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{F}_n^1(M)$ gehöre der Differentiationsklasse \mathcal{C}^{m+1} an. Es sei $\vec{h} = (h_1, \dots, h_n)^T$. Liegen die Punkte \vec{x}_0 und $\vec{x}_0 + \vec{h}$ samt ihrer Verbindungsstrecke in M , so gibt es ein $\vartheta \in (0, 1)$, so dass gilt:

$$\begin{aligned} f(\vec{x}_0 + \vec{h}) &= \sum_{k=0}^m \frac{1}{k!} (\vec{h} \nabla)^{(k)} f(\vec{x}_0) + \frac{1}{(m+1)!} (\vec{h} \nabla)^{(m+1)} f(\vec{x}_0 + \vartheta \vec{h}) \\ &= \sum_{k=0}^m \frac{1}{k!} \sum_{j_1, \dots, j_k=1}^n \frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}}(\vec{x}_0) h_{j_1} \cdots h_{j_k} \\ &\quad + \frac{1}{(m+1)!} \sum_{j_1, \dots, j_{m+1}=1}^n \frac{\partial^{m+1} f}{\partial x_{j_1} \cdots \partial x_{j_{m+1}}}(\vec{x}_0 + \vartheta \vec{h}) h_{j_1} \cdots h_{j_{m+1}}. \end{aligned}$$

BEISPIEL 8.6.1

Es sei

$$f : \begin{cases} M & \rightarrow \mathbb{R} \\ (x, y, z)^T & \mapsto f((x, y, z)^T) \end{cases}$$

in $M \subseteq \mathbb{R}^3$ dreimal stetig differenzierbar. Im Folgenden identifizieren wir der Einfachheit halber die Schreibweisen $f(x, y, z)$ und $f((x, y, z)^T)$. Die Verbindungsstrecke von $(x_0, y_0, z_0)^T$ und $(x_0 + h, y_0 + j, z_0 + k)^T$ liege in M . Dann gilt:

$$\begin{aligned} f(x_0 + h, y_0 + j, z_0 + k) &= f(x_0, y_0, z_0) + \frac{\partial f}{\partial x}(x_0, y_0, z_0)h + \frac{\partial f}{\partial y}(x_0, y_0, z_0)j + \frac{\partial f}{\partial z}(x_0, y_0, z_0)k \\ &\quad + \frac{1}{2!} \left(\frac{\partial^2 f}{\partial x \partial x}(x_0, y_0, z_0)h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0, z_0)hj + \frac{\partial^2 f}{\partial x \partial z}(x_0, y_0, z_0)hk \right. \\ &\quad + \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0, z_0)hj + \frac{\partial^2 f}{\partial y \partial y}(x_0, y_0, z_0)j^2 + \frac{\partial^2 f}{\partial y \partial z}(x_0, y_0, z_0)jk \\ &\quad \left. + \frac{\partial^2 f}{\partial z \partial x}(x_0, y_0, z_0)hk + \frac{\partial^2 f}{\partial z \partial y}(x_0, y_0, z_0)jk + \frac{\partial^2 f}{\partial z \partial z}(x_0, y_0, z_0)k^2 \right) \\ &\quad + \frac{1}{3!} \left(\frac{\partial^3 f}{\partial x \partial x \partial x}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h^3 + \frac{\partial^3 f}{\partial x \partial x \partial y}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h^2 j + \cdots \right). \end{aligned}$$

Wegen der Stetigkeit der partiellen Ableitungen kann nach Satz 8.4.1 die Reihenfolge der Differentiationen vertauscht werden. Durch Zusammenfassen erhalten wir

$$\begin{aligned}
f(x_0 + h, y_0 + j, z_0 + k) &= f(x_0, y_0, z_0) + \frac{\partial f}{\partial x}(x_0, y_0, z_0)h + \frac{\partial f}{\partial y}(x_0, y_0, z_0)j + \frac{\partial f}{\partial z}(x_0, y_0, z_0)k \\
&+ \frac{1}{2!} \left(\frac{\partial^2 f}{\partial x \partial x}(x_0, y_0, z_0)h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0, z_0)hj + 2 \frac{\partial^2 f}{\partial x \partial z}(x_0, y_0, z_0)hk \right. \\
&+ \left. \frac{\partial^2 f}{\partial y \partial y}(x_0, y_0, z_0)j^2 + 2 \frac{\partial^2 f}{\partial y \partial z}(x_0, y_0, z_0)jk + \frac{\partial^2 f}{\partial z \partial z}(x_0, y_0, z_0)k^2 \right) \\
&+ \frac{1}{3!} \left(\frac{\partial^3 f}{\partial x^3}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h^2 j \right. \\
&+ 3 \frac{\partial^3 f}{\partial x \partial y^2}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h j^2 + 3 \frac{\partial^3 f}{\partial x^2 \partial z}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h^2 k \\
&+ 6 \frac{\partial^3 f}{\partial x \partial y \partial z}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h j k + \frac{\partial^3 f}{\partial y^3}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)j^3 \\
&+ 3 \frac{\partial^3 f}{\partial y^2 \partial z}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)j^2 k + 3 \frac{\partial^3 f}{\partial y \partial z^2}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)j k^2 \\
&+ \left. 3 \frac{\partial^3 f}{\partial x \partial z^2}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)h k^2 + \frac{\partial^3 f}{\partial z^3}(x_0 + \vartheta h, y_0 + \vartheta j, z_0 + \vartheta k)k^3 \right).
\end{aligned}$$

Führt man die Zusammenfassung von Termen im allgemeinen Fall aus, so erhält man

SATZ 8.6.4 (Satz von Taylor, 2. Version)

Es sei $M \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{F}_n^1(M)$ gehöre der Differentiationsklasse \mathcal{C}^{m+1} an. Es sei $\vec{h} = (h_1, \dots, h_n)^T$. Liegen die Punkte \vec{x}_0 und $\vec{x}_0 + \vec{h}$ samt ihrer Verbindungsstrecke in M , so gibt es ein $\vartheta \in (0, 1)$, so dass gilt:

$$\begin{aligned}
f(\vec{x}_0 + \vec{h}) &= \sum_{k=0}^m \frac{1}{k!} \sum_{\substack{\nu_1, \dots, \nu_n=0 \dots k \\ \nu_1 + \dots + \nu_n = k}} \frac{k!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_n!} \cdot \frac{\partial^k}{\partial x_1^{\nu_1} \dots \partial x_n^{\nu_n}} f(\vec{x}_0) \cdot h_1^{\nu_1} \dots h_n^{\nu_n} \\
&+ \frac{1}{(m+1)!} \sum_{\substack{\nu_1, \dots, \nu_n=0 \dots m+1 \\ \nu_1 + \dots + \nu_n = m+1}} \frac{(m+1)!}{\nu_1! \cdot \nu_2! \cdot \dots \cdot \nu_n!} \cdot \frac{\partial^{m+1}}{\partial x_1^{\nu_1} \dots \partial x_n^{\nu_n}} f(\vec{x}_0 + \vartheta \vec{h}) \cdot h_1^{\nu_1} \dots h_n^{\nu_n}.
\end{aligned}$$

BEWEIS

In Satz 8.6.3 fassen wir alle Terme zusammen, bei denen ν_1 Indizes j_l den Wert 1, ν_2 Indizes den Wert 2, \dots , ν_n Indizes den Wert n haben. Die Anzahl dieser Terme ist $\frac{k!}{\nu_1! \dots \nu_n!}$. \square

8.7. Extrema bei Funktionen mehrerer Variablen

DEFINITION 8.7.1

Es sei $M \subseteq \mathbb{R}^n$ und $f \in \mathcal{F}_n^1(M)$ sowie $\vec{x}_0 \in M^0$ ein innerer Punkt. Man sagt, f besitze in \vec{x}_0 ein relatives Maximum bzw. ein relatives Minimum (gemeinsame Bezeichnung: relatives Extremum), wenn es eine Umgebung $U \subseteq M$ von \vec{x}_0 gibt, so dass $f(\vec{x}) \leq f(\vec{x}_0)$ bzw. $f(\vec{x}) \geq f(\vec{x}_0)$ ist für alle $\vec{x} \in U$. Schließt man in der Ungleichung das Gleichheitszeichen aus, so spricht man von einem strengen relativen Extremum.

SATZ 8.7.1 (Notwendige Bedingung für Extrema)

Wenn die Funktion f in \vec{x}_0 partiell differenzierbar nach allen Variablen ist und in \vec{x}_0 ein relatives Extremum besitzt, so ist $\text{grad}f(\vec{x}_0) = \vec{0}$.

BEWEIS

Dies folgt sofort, wenn Satz 5.1.3 auf f als Funktion der Variablen x_j bei festgehaltenen übrigen Variablen angewandt wird. \square

Um ein hinreichendes Kriterium zu erhalten, benötigen wir den Begriff der Definitheit aus der Linearen Algebra:

DEFINITION 8.7.2

Eine symmetrische Matrix $B \in \mathbb{R}^{n \times n}$ heißt (positiv bzw. negativ) definit, wenn für die zugehörige quadratische Form

$$Q : \mathbb{R}^n \rightarrow \mathbb{R}, \quad Q(\vec{x}) = \vec{x}^T \cdot B \cdot \vec{x}$$

gilt, dass $Q(\vec{x}) > 0$ bzw. $Q(\vec{x}) < 0$ ist für alle $\vec{x} \neq 0$. Die Matrix heißt (positiv bzw. negativ) semidefinit, falls $Q(\vec{x}) \geq 0$ bzw. $Q(\vec{x}) \leq 0$ für alle $\vec{x} \neq 0$ gilt. Sie heißt indefinit, falls sie nicht semidefinit ist.

SATZ 8.7.2 (Charakterisierung der Definitheit)

Die folgenden Aussagen sind für eine symmetrische Matrix B äquivalent:

- (i) B ist positiv definit,
- (ii) alle Eigenwerte von B sind positiv,
- (iii) alle Hauptunterdeterminanten von B sind positiv.

BEWEIS

Siehe Lineare Algebra. \square

DEFINITION 8.7.3

Es sei $f \in \mathcal{F}_n^1$ in einer Umgebung von $\vec{x} \in \mathbb{R}^n$ definiert und gehöre dort der Differentiationsklasse \mathcal{C}^2 an. Die Hessesche Matrix von f ist

$$H_f(\vec{x}) := \begin{pmatrix} \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\vec{x})}{\partial x_2 \partial x_1} & \dots & \frac{\partial^2 f(\vec{x})}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\vec{x})}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f(\vec{x})}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_n} & \frac{\partial^2 f(\vec{x})}{\partial x_2 \partial x_n} & \dots & \frac{\partial^2 f(\vec{x})}{\partial x_n \partial x_n} \end{pmatrix}.$$

Satz 8.4.1 besagt, dass die Hessesche Matrix symmetrisch ist.

SATZ 8.7.3 (Hinreichende Bedingung für Extrema)

Es sei $f \in \mathcal{F}_n^1$ in einer Umgebung U von $\vec{x}_0 \in \mathbb{R}^n$ definiert, und gehöre dort der Differentiationsklasse \mathcal{C}^2 an. Es sei $\text{grad}f(\vec{x}_0) = \vec{0}$. Wenn $H_f(\vec{x}_0)$ definit ist, besitzt f in \vec{x}_0 ein strenges relatives Extremum. Es ist ein Maximum, falls $H_f(\vec{x}_0)$ negativ definit ist, bzw. ein Minimum, falls $H_f(\vec{x}_0)$ positiv definit ist. Ist $H_f(\vec{x}_0)$ indefinit, so besitzt f in \vec{x}_0 kein Extremum.

BEWEIS

Es sei $\vec{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$ so gewählt, dass die Verbindungsstrecke von \vec{x}_0 und $\vec{x}_0 + \vec{h}$ in U liegt. Nach Satz 8.6.3 (Taylor) mit $m = 1$ gibt es ein $\vartheta \in (0, 1)$ mit

$$\begin{aligned} f(\vec{x}_0 + \vec{h}) &= f(\vec{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\vec{x}_0) h_j + \frac{1}{2!} \sum_{j_1, j_2=1}^n \frac{\partial^2 f}{\partial x_{j_1} \partial x_{j_2}}(\vec{x}_0 + \vartheta \vec{h}) h_{j_1} h_{j_2} \\ &= f(\vec{x}_0) + \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) h_j h_k + r(\vec{h}) \end{aligned}$$

$$\text{mit } r(\vec{h}) = \frac{1}{2} \sum_{j,k=1}^n \left(\frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0 + \vartheta \vec{h}) - \frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) \right) h_j h_k.$$

Es ist

$$\frac{|r(\vec{h})|}{\|\vec{h}\|^2} \leq \frac{1}{2} \sum_{j,k=1}^n \left| \frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0 + \vartheta \vec{h}) - \frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) \right|,$$

wegen der Stetigkeit der zweiten partiellen Ableitungen folgt

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{r(\vec{h})}{\|\vec{h}\|^2} = 0.$$

Damit haben wir

$$(*) \quad \frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0)}{\|\vec{h}\|^2} = \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k} \cdot \frac{h_j}{\|\vec{h}\|} \cdot \frac{h_k}{\|\vec{h}\|} + \frac{r(\vec{h})}{\|\vec{h}\|^2} = \frac{1}{2} \left(\frac{\vec{h}}{\|\vec{h}\|} \right)^T \cdot H_f(\vec{x}_0) \cdot \left(\frac{\vec{h}}{\|\vec{h}\|} \right) + \frac{r(\vec{h})}{\|\vec{h}\|^2}.$$

Ist $H_f(\vec{x}_0)$ positiv definit, so folgt

$$\frac{1}{2} \left(\frac{\vec{h}}{\|\vec{h}\|} \right)^T \cdot H_f(\vec{x}_0) \cdot \left(\frac{\vec{h}}{\|\vec{h}\|} \right) \geq \min_{\|\vec{y}\|=1} (\vec{y}^T \cdot H_f(\vec{x}_0) \cdot \vec{y}) > 0,$$

da $\{\vec{y} \in \mathbb{R}^n \mid \|\vec{y}\| = 1\}$ eine kompakte Menge bildet. Wegen

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{r(\vec{h})}{\|\vec{h}\|^2} = 0$$

ist $f(\vec{x}_0 + \vec{h}) > f(\vec{x}_0)$ für alle \vec{h} in einer Umgebung von $\vec{0}$, d. h. f besitzt ein strenges relatives Minimum in \vec{x}_0 . Die Behauptung für negativ definites $H_f(\vec{x}_0)$ folgt analog. Ist $H_f(\vec{x}_0)$ indefinit, so gibt es \vec{a}, \vec{b} mit $\|\vec{a}\| = \|\vec{b}\| = 1$, so dass $\vec{a}^T H_f(\vec{x}_0) \vec{a} > 0$ und $\vec{b}^T H_f(\vec{x}_0) \vec{b} < 0$ ist. Setzen wir $\vec{h} = c\vec{a}$, so folgt aus (*) für hinreichend kleines $c > 0$, dass $f(\vec{x}_0 + c\vec{a}) > f(\vec{x}_0)$ ist, ebenso aber auch $f(\vec{x}_0 + c\vec{b}) < f(\vec{x}_0)$. Daher besitzt f dann in \vec{x}_0 kein Extremum. \square

Wir formulieren diesen Satz noch im Spezialfall $n = 2$:

SATZ 8.7.4

Es sei $f \in \mathcal{F}_2^1$ in einer Umgebung von $\vec{x}_0 \in \mathbb{R}^2$ definiert und gehöre dort dem Differentiationsbereich \mathcal{C}^2 an. Ist $f_{xx}(\vec{x}_0) \cdot f_{yy}(\vec{x}_0) - f_{xy}(\vec{x}_0)^2 > 0$, so besitzt $f(x, y)$ in \vec{x}_0 ein strenges relatives Extremum, und zwar ein Maximum bzw. Minimum, wenn $f_{xx}(\vec{x}_0) < 0$ bzw. $f_{xx}(\vec{x}_0) > 0$ ist. Wenn $f_{xx}(\vec{x}_0) \cdot f_{yy}(\vec{x}_0) - f_{xy}(\vec{x}_0)^2 < 0$ ist, so besitzt f in \vec{x}_0 kein relatives Extremum.

8.8. Implizite Funktionen

Wir betrachten eine Funktion $\vec{g}(\vec{x}, \vec{y})$ zweier vektorwertiger Variablen $\vec{x} \in \mathbb{R}^n$ und $\vec{y} \in \mathbb{R}^l$ und fragen, ob die Gleichung $\vec{g}(\vec{x}, \vec{y}) = \vec{0}$ nach \vec{y} auflösbar ist, d. h. gibt es eine Funktion $\vec{f} \in \mathcal{F}_n^l$, so dass $\vec{g}(\vec{x}, \vec{f}(\vec{x})) = \vec{0}$ ist für alle \vec{x} aus einer offenen Menge? Ist \vec{f} in diesem Fall eindeutig bestimmt, stetig, oder sogar differenzierbar? Im Falle der Existenz von \vec{f} sagt man, \vec{f} sei implizit durch die Gleichung $\vec{g}(\vec{x}, \vec{y}) = \vec{0}$ definiert.

BEISPIEL 8.8.1

Die Funktion $f(x) = \sqrt{1 - x^2}$ ist implizit definiert durch $g(x, y) = x^2 + y^2 - 1 = 0$.

DEFINITION 8.8.1

Es sei $\vec{g} \in \mathcal{F}_{n+l}^l$, also $\vec{g} : (x_1, \dots, x_n, y_1, \dots, y_l)^T \mapsto g((x_1, \dots, x_n, y_1, \dots, y_l)^T)$. Wir betrachten $\vec{g} = \vec{g}(\vec{x}, \vec{y})$ als eine Funktion der zwei vektorwertigen Variablen $\vec{x} = (x_1, \dots, x_n)^T$ und $\vec{y} = (y_1, \dots, y_l)^T$. Die Funktion \vec{g} sei nach allen $n + l$ Variablen partiell differenzierbar. Funktionalmatrizen, die sich ergeben, wenn der Differentiationsprozess nur auf einen Teil der Variablen von \vec{g} angewendet wird, werden mit ∂ statt d bezeichnet, insbesondere ist

$$\frac{\partial \vec{g}}{\partial \vec{x}} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_l}{\partial x_1} & \dots & \frac{\partial g_l}{\partial x_n} \end{pmatrix} \quad \text{und} \quad \frac{\partial \vec{g}}{\partial \vec{y}} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \dots & \frac{\partial g_1}{\partial y_l} \\ \vdots & & \vdots \\ \frac{\partial g_l}{\partial y_1} & \dots & \frac{\partial g_l}{\partial y_l} \end{pmatrix}.$$

SATZ 8.8.1 (Hauptsatz der Theorie der impliziten Funktionen)

Es sei $M \subseteq \mathbb{R}^{n+l}$ eine offene Menge und $(\vec{x}_0, \vec{y}_0) \in M$ ein Punkt mit $\vec{x}_0 \in \mathbb{R}^n$ bzw. $\vec{y}_0 \in \mathbb{R}^l$, und $\vec{g} \in \mathcal{F}_{n+l}^l(M)$ gehöre der Differentiationsklasse \mathcal{C}^1 an. Es sei $\det \frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0$ und $\vec{g}(\vec{x}_0, \vec{y}_0) = \vec{0}$. Dann gibt es eine Umgebung $U \subseteq \mathbb{R}^n$ von \vec{x}_0 sowie genau eine Funktion $\vec{f} \in \mathcal{F}_n^l(U)$ mit $\vec{f}(\vec{x}_0) = \vec{y}_0$ und $\vec{g}(\vec{x}, \vec{f}(\vec{x})) = \vec{0}$ für alle $\vec{x} \in U$. Zudem ist die Matrix $\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}, \vec{y})$ invertierbar für alle $\vec{x} \in U$ und $\vec{y} = \vec{f}(\vec{x})$. Die Funktion \vec{f} gehört in U der Differentiationsklasse \mathcal{C}^1 an, und es ist

$$\frac{d\vec{f}}{d\vec{x}}(\vec{x}) = - \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}, \vec{f}(\vec{x})) \right)^{-1} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}, \vec{f}(\vec{x})).$$

BEWEIS

Wir ergänzen \vec{g} zu einer Funktion $\vec{G} : \mathbb{R}^{n+l} \rightarrow \mathbb{R}^{n+l}$ durch

$$\vec{G}(\vec{x}, \vec{y}) := (\vec{x}, \vec{g}(\vec{x}, \vec{y}))^T = (x_1, \dots, x_n, g_1(\vec{x}, \vec{y}), \dots, g_l(\vec{x}, \vec{y}))^T.$$

Nach der Kettenregel ist mit \vec{g} auch $\vec{G} \in \mathcal{F}_{n+l}^{n+l}$ stetig differenzierbar. Es gilt

$$\det \left(\frac{d\vec{G}}{d(\vec{x}, \vec{y})} \right) = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} & \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_l} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} & \frac{\partial g_2}{\partial y_1} & \cdots & \frac{\partial g_2}{\partial y_l} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_l}{\partial x_1} & \frac{\partial g_l}{\partial x_2} & \cdots & \frac{\partial g_l}{\partial x_n} & \frac{\partial g_l}{\partial y_1} & \cdots & \frac{\partial g_l}{\partial y_l} \end{vmatrix} = \begin{vmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_l} \\ \vdots & & \vdots \\ \frac{\partial g_l}{\partial y_1} & \cdots & \frac{\partial g_l}{\partial y_l} \end{vmatrix} = \det \left(\frac{\partial \vec{g}}{\partial \vec{y}} \right) \neq 0$$

im Punkt (\vec{x}_0, \vec{y}_0) nach Voraussetzung, nach Satz 8.3.1 gibt es also eine Umgebung $\tilde{U} \subseteq M$ von (\vec{x}_0, \vec{y}_0) , auf der \vec{G} umkehrbar ist zu einer Funktion $\vec{G}^{-1} : \tilde{V} \rightarrow \tilde{U}$ mit $\tilde{V} = \vec{G}(\tilde{U})$. Wir teilen die $n+l$ Komponenten von \vec{G}^{-1} auf in $\vec{G}^{-1}(\vec{x}, \vec{y}) = (\vec{a}(\vec{x}, \vec{y}), \vec{b}(\vec{x}, \vec{y}))$ mit $\vec{a} \in \mathcal{F}_{l+n}^n(\tilde{V})$ und $\vec{b} \in \mathcal{F}_{n+l}^l(\tilde{V})$. Nun gilt für alle $(\vec{x}, \vec{y}) \in \tilde{U}$

$$(\vec{x}, \vec{y}) = \vec{G}(\vec{G}^{-1}(\vec{x}, \vec{y})) = \vec{G}(\vec{a}(\vec{x}, \vec{y}), \vec{b}(\vec{x}, \vec{y})) = (\vec{a}(\vec{x}, \vec{y}), \vec{g}(\vec{a}(\vec{x}, \vec{y}), \vec{b}(\vec{x}, \vec{y})))$$

$$\Rightarrow \vec{a}(\vec{x}, \vec{y}) = \vec{x} \text{ und } \vec{g}(\vec{a}(\vec{x}, \vec{y}), \vec{b}(\vec{x}, \vec{y})) = \vec{y}$$

$$\Rightarrow \vec{g}(\vec{x}, \vec{b}(\vec{x}, \vec{y})) = \vec{g}(\vec{a}(\vec{x}, \vec{y}), \vec{b}(\vec{x}, \vec{y})) = \vec{y}.$$

Setzen wir $\vec{f}(\vec{x}) := \vec{b}(\vec{x}, \vec{0})$ für alle $(\vec{x}, \vec{y}) \in \tilde{U}$, so ist auf $U = \{\vec{x} \in \mathbb{R}^n \mid (\vec{x}, \vec{y}) \in \tilde{U}\}$ dann $\vec{g}(\vec{x}, \vec{f}(\vec{x})) = \vec{0}$. Auf U ist $\vec{g}(\vec{x}, \vec{f}(\vec{x}))$ als Funktion $\mathbb{R}^n \rightarrow \mathbb{R}^l$ konstant der Nullvektor, also ist auch die Funktionalmatrix $\frac{d\vec{g}(\vec{x}, \vec{f}(\vec{x}))}{d\vec{x}}$ Null. Schreiben wir $g(\vec{x}, \vec{f}(\vec{x})) = (g \circ h)(\vec{x})$ mit den Funktionen $\vec{g} : \mathbb{R}^{n+l} \rightarrow \mathbb{R}^l$ und $\vec{h} : \mathbb{R}^n \rightarrow \mathbb{R}^{n+l}$ mit $\vec{h}(\vec{x}) = (x_1, \dots, x_n, f_1(\vec{x}), \dots, f_l(\vec{x}))^T$, oder symbolisch $\vec{h} = (\text{id}, \vec{f})$, so gilt nach der Kettenregel 8.3.2 aber auch

$$\frac{d\vec{g}(\vec{x}, \vec{f}(\vec{x}))}{d\vec{x}} = \frac{d\vec{g}}{d(\vec{x}, \vec{y})} \cdot \frac{d\vec{h}}{d\vec{x}} = \begin{pmatrix} \underbrace{\frac{\partial \vec{g}}{\partial \vec{x}}}_{n \text{ Spalten}} & \underbrace{\frac{\partial \vec{g}}{\partial \vec{y}}}_{l \text{ Spalten}} \end{pmatrix} \cdot \begin{pmatrix} \left\{ \frac{d \text{id}}{d\vec{x}} \right\} n \text{ Zeilen} \\ \left\{ \frac{d\vec{f}}{d\vec{x}} \right\} l \text{ Zeilen} \end{pmatrix} \stackrel{\frac{d \text{id}}{d\vec{x}} = E}{=} \frac{\partial \vec{g}}{\partial \vec{x}} + \left(\frac{\partial \vec{g}}{\partial \vec{y}} \right) \cdot \left(\frac{d\vec{f}}{d\vec{x}} \right).$$

Da der gesamte Ausdruck die Nullmatrix ist, folgt durch Auflösen nach der Ableitung von f auf U

$$\frac{d\vec{f}}{d\vec{x}}(\vec{x}) = - \left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}, \vec{f}(\vec{x})) \right)^{-1} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}, \vec{f}(\vec{x})).$$

Die Funktion \vec{f} ist auch eindeutig, denn erfüllt eine Funktion \vec{F} auf U ebenfalls die Gleichung $\vec{g}(\vec{x}, \vec{F}(\vec{x})) = \vec{0}$ und die Voraussetzung $\vec{F}(\vec{x}_0) = \vec{y}_0$, so ist nach der obigen Rechnung $\frac{d\vec{f}}{d\vec{x}} = \frac{d\vec{F}}{d\vec{x}}$ auf U , damit verschwindet die Ableitung der Differenz $\vec{F} - \vec{f}$, womit diese nach Satz 8.6.1 konstant ist. Wegen $\vec{f}(\vec{x}_0) = \vec{y}_0 = \vec{F}(\vec{x}_0)$ ist dann $\vec{f} = \vec{F}$. \square

BEISPIEL 8.8.2

Es sei $g(x, y) = x^2 + e^y$. Die Gleichung $g(x, y) = 0$ ist nicht nach y auflösbar, da es kein $(x_0, y_0)^T \in \mathbb{R}^2$ gibt mit $g(x_0, y_0) = 0$.

BEISPIEL 8.8.3

Es sei $g(x, y) = x^2 + y^2 - 1$. Es ist $\frac{\partial g}{\partial y}(x, y) = g_y(x, y) = 2y \neq 0$ für $y \neq 0$. Nach dem Hauptsatz gibt es für jeden Punkt $(x_0, y_0)^T \neq (1, 0)^T, (-1, 0)^T$ auf dem Einheitskreis eine stetig differenzierbare

implizite Funktion f und Intervalle U, V mit $x_0 \in U$ und $y_0 \in V$, so dass $f(x_0) = y_0$ ist und für $x \in U$, $y \in V$ gilt $g(x, y) = 0 \Leftrightarrow y = f(x)$. Hier ist f explizit bestimmbar mit $f(x) = \sqrt{1-x^2}$ falls $y_0 > 0$ und $f(x) = -\sqrt{1-x^2}$ falls $y_0 < 0$.

BEISPIEL 8.8.4

Die Gleichung $g(x, y) = x^3 e^y + y e^y + x^2 y = 0$ besitzt eine Auflösung $y = f(x)$ durch den Nullpunkt, weil $\frac{\partial g}{\partial y}(x, y) = x^3 e^y + (y+1)e^y + x^2 \neq 0$ ist für $(x, y)^T = (0, 0)$. Wir können $y = f(x)$ nicht explizit bestimmen, jedoch kann f' nach dem Hauptsatz durch x und y ausgedrückt werden:

$$f'(x) = -\frac{g_x(x, f(x))}{g_y(x, f(x))} = -\frac{3x^2 e^y + 2xy}{x^3 e^y + (y+1)e^y + x^2},$$

insbesondere ist $f'(0) = 0$.

BEISPIEL 8.8.5

Gegeben sei das System

$$(*) \quad \begin{cases} g_1(x_1, x_2, x_3, y_1, y_2) = x_1^2 y_1 + x_1 x_2 e^{y_2} - y_2^2 + x_3^2 = 0 \\ g_2(x_1, x_2, x_3, y_1, y_2) = e^{x_2 y_1 y_2} - y_1^2 = 0 \end{cases}.$$

Aufgabe: man zeige, dass in einer Umgebung von $(x_1, x_2, x_3, y_1, y_2)^T = (1, 0, 0, 1, 1)^T$ das System (*) durch Funktionen $y_1 = f_1(x_1, x_2, x_3)$ und $y_2 = f_2(x_1, x_2, x_3)$ aufgelöst werden kann. Man finde die partiellen Ableitungen von f_1 bzw. f_2 im Punkt $(x_1, x_2, x_3)^T = (1, 0, 0)^T$.

Lösung: Wir schreiben $\vec{x} = (x_1, x_2, x_3)^T$ und $\vec{y} = (y_1, y_2)^T$. Es sei $\vec{x}_0 = (1, 0, 0)^T$ und $\vec{y}_0 = (1, 1)^T$ sowie $\vec{g}(\vec{x}, \vec{y}) = (g_1(x_1, x_2, x_3), g_2(y_1, y_2))^T$. Dann ist $\vec{g}(\vec{x}_0, \vec{y}_0) = (0, 0)^T$. Es gilt

$$\frac{\partial g_1}{\partial y_1}(x_1, x_2, x_3) = x_1^2, \quad \frac{\partial g_1}{\partial y_2}(x_1, x_2, x_3) = x_1 x_2 e^{y_2} - 2y_2,$$

$$\frac{\partial g_2}{\partial y_1}(x_1, x_2, x_3) = x_2 y_2 e^{x_2 y_1 y_2} - 2y_1, \quad \frac{\partial g_2}{\partial y_2}(x_1, x_2, x_3) = x_2 y_1 e^{x_1 y_1 y_2}, \text{ also}$$

$$\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \frac{\partial g_1}{\partial y_2}(\vec{x}_0, \vec{y}_0) \\ \frac{\partial g_2}{\partial y_1}(\vec{x}_0, \vec{y}_0) & \frac{\partial g_2}{\partial y_2}(\vec{x}_0, \vec{y}_0) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix}$$

mit $\det \frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) = -4 \neq 0$. Daher gibt es wegen der Stetigkeit der partiellen Ableitungen von \vec{g} Umgebungen U von \vec{x}_0 und V von \vec{y}_0 , so dass $\det \frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}, \vec{y}) \neq 0$ ist für alle $(\vec{x}, \vec{y})^T \in U \times V$. Nach dem Hauptsatz 8.8.1 gibt es daher Umgebungen $U_0 \subseteq U$ von \vec{x}_0 und $V_0 \subseteq V$ von \vec{y}_0 , so dass durch die Gleichung $\vec{g}(\vec{x}, \vec{y}) = \vec{0}$ in U_0 genau eine Funktion $\vec{y} = \vec{f}(\vec{x})$, deren Werte in V_0 liegen, implizit definiert ist. Es ist

$$\frac{d\vec{f}}{d\vec{x}}(\vec{x}_0) = -\left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right)^{-1} \cdot \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0).$$

Hierbei ist

$$\left(\frac{\partial \vec{g}}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right)^{-1} = \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix}^{-1} = \frac{1}{-4} \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}.$$

Ferner ist

$$\frac{\partial g_1}{\partial x_1} = 2x_1 y_1 + x_2 e^{y_2}, \quad \frac{\partial g_1}{\partial x_2} = x_1 e^{y_2}, \quad \frac{\partial g_1}{\partial x_3} = 3x_3^2,$$

$$\frac{\partial g_2}{\partial x_1} = 0, \quad \frac{\partial g_2}{\partial x_2} = y_1 y_2 e^{x_2 y_1 y_2}, \quad \frac{\partial g_2}{\partial x_3} = 0, \text{ also } \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} 2 & e & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Somit gilt

$$\frac{d\vec{f}}{d\vec{x}}(\vec{x}_0) = - \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{4} \end{pmatrix} \cdot \begin{pmatrix} 2 & e & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & \frac{e}{2} + \frac{1}{4} & 0 \end{pmatrix},$$

also

$$\frac{\partial f_1}{\partial x_1}(1,0,0) = 0, \quad \frac{\partial f_1}{\partial x_2}(1,0,0) = \frac{1}{2}, \quad \frac{\partial f_1}{\partial x_3}(1,0,0) = 0,$$

$$\frac{\partial f_2}{\partial x_1}(1,0,0) = 1, \quad \frac{\partial f_2}{\partial x_2}(1,0,0) = \frac{e}{2} + \frac{1}{4}, \quad \frac{\partial f_2}{\partial x_3}(1,0,0) = 0.$$

8.9. Extrema mit Nebenbedingungen

Den Argumenten $\vec{x} \in \mathbb{R}^n$ einer Funktion f wollen wir gewisse durch ein System

$$g_1(\vec{x}) = 0, \quad g_2(\vec{x}) = 0, \quad \dots, \quad g_l(\vec{x}) = 0$$

von Gleichungen gegebene Beschränkungen auferlegen und nach den Extrema der dann vorkommenden Funktionswerte von f fragen. Man spricht von Extrema mit Nebenbedingungen. Wir geben eine notwendige Bedingung für das Vorliegen eines solchen Extremums.

SATZ 8.9.1 (Multiplikatorenregel von Lagrange)

Es seien $f \in \mathcal{F}_n^1$ sowie $\vec{g} \in \mathcal{F}_n^l$ für $n > l$ in einer Umgebung eines Punktes $\vec{x}_0 \in \mathbb{R}^n$ stetig differenzierbar. Die Funktionalmatrix $\frac{d\vec{g}}{d\vec{x}}$ besitze in \vec{x}_0 maximalen Rang, d. h.

$$\text{rg} \left(\frac{d\vec{g}}{d\vec{x}}(\vec{x}_0) \right) = l,$$

und es sei $\vec{g}(\vec{x}_0) = \vec{0}$. Wenn dann $f(\vec{x}_0)$ Extremum der Werte von $f(\vec{x})$ bei Beschränkung auf solche \vec{x} ist, für die $\vec{g}(\vec{x}) = \vec{0}$ gilt, so gibt es Zahlen $\lambda_1, \dots, \lambda_l \in \mathbb{R}$, für die in \vec{x}_0 gilt:

$$\frac{\partial f}{\partial x_\nu} + \lambda_1 \frac{\partial g_1}{\partial x_\nu} + \dots + \lambda_l \frac{\partial g_l}{\partial x_\nu} = 0, \quad \nu = 1 \dots n.$$

OHNE BEWEIS

□

BEISPIEL 8.9.1

Es sei eine Hyperebene durch

$$g(\vec{x}) = \vec{a}^T \vec{x} - c = 0, \quad \vec{a} = (a_1, \dots, a_n)^T$$

gegeben. Man bestimme den Abstand dieser Hyperebene vom Nullpunkt.

Lösung: Es geht darum, die Quadratsumme $f(\vec{x}) = x_1^2 + \dots + x_n^2$ zu minimieren. Wir haben das Gleichungssystem

$$\frac{\partial f}{\partial x_\nu}(\vec{x}_0) + \lambda \frac{\partial g}{\partial x_\nu}(\vec{x}_0) = 0, \quad \nu = 1, \dots, n$$

unter der Nebenbedingung $\vec{a}^T \vec{x}_0 - c = 0$ zu lösen. Dies ist äquivalent zu

$$\begin{array}{l} 2\vec{x} + \lambda\vec{a} = \vec{0} \\ \vec{a}^T \vec{x} - c = 0 \end{array} \Rightarrow 2\vec{a}^T \vec{x} + \lambda \|\vec{a}\|^2 = 0 \Rightarrow -\lambda \|\vec{a}\|^2 = 2c \Rightarrow \lambda = -\frac{2c}{\|\vec{a}\|^2} \Rightarrow \vec{x} = \frac{c}{\|\vec{a}\|^2} \vec{a}.$$

Tatsächlich liegt in diesem Punkt \vec{x} ein Minimum des Abstands der Hyperebene vom Nullpunkt vor.

9. Integration über mehreren Veränderlichen

9.1. Bereichsintegrale

Der Begriff des Riemannsches Integrals einer Funktion von mehreren Variablen ist eine Verallgemeinerung dieses Begriffs für den Fall einer Variablen: wieder sind Zerlegungen von Intervallen, Ober- und Untersummen grundlegend. Wir beginnen mit der Definition des n -dimensionalen Intervalls.

DEFINITION 9.1.1

Unter einem kompakten Intervall I des \mathbb{R}^n versteht man das kartesische Produkt von n kompakten Intervallen $I_\nu = [a_\nu, b_\nu] \subset \mathbb{R}$, also

$$I = I_1 \times \cdots \times I_n = \{(x_1, \dots, x_n)^T \in \mathbb{R}^n \mid \forall \nu = 1 \dots n : a_\nu \leq x_\nu \leq b_\nu\}.$$

Unter dem Inhalt von I verstehen wir die Zahl

$$|I| = \prod_{\nu=1}^n (b_\nu - a_\nu).$$

DEFINITION 9.1.2

Es sei I ein kompaktes Intervall des \mathbb{R}^n . Unter einer Zerlegung \mathcal{Z} von I versteht man eine Folge I_1, \dots, I_l von endlich vielen kompakten Intervallen, deren offene Kerne I_ν^0 paarweise disjunkt sind, und deren Vereinigung I ist. Die I_ν heißen Teilungsintervalle von \mathcal{Z} .

DEFINITION 9.1.3

Es sei I ein kompaktes Intervall des \mathbb{R}^n . $f \in \mathcal{F}_n^1(I)$ sei auf I beschränkt. \mathcal{Z} sei eine Zerlegung von I mit den Teilungsintervallen I_1, \dots, I_l . Dann sind Ober- und Untersummen der Zerlegung \mathcal{Z} definiert durch

$$\overline{S}(\mathcal{Z}) := \sum_{\nu=1}^l \overline{M}_\nu |I_\nu| \quad \text{und} \quad \underline{S}(\mathcal{Z}) := \sum_{\nu=1}^l \underline{M}_\nu |I_\nu| \quad \text{mit}$$

$$\overline{M}_\nu := \sup\{f(\vec{x}) \mid \vec{x} \in I_\nu\} \quad , \quad \underline{M}_\nu := \inf\{f(\vec{x}) \mid \vec{x} \in I_\nu\}.$$

Wir nennen

$$\overline{\int_I} f(\vec{x}) d\vec{x} := \inf_{\mathcal{Z}} \overline{S}(\mathcal{Z}) \quad , \quad \underline{\int_I} f(\vec{x}) d\vec{x} := \sup_{\mathcal{Z}} \underline{S}(\mathcal{Z})$$

das Ober- bzw. Unterintegral von f über I .

SATZ 9.1.1

Es ist stets

$$\underline{\int_I} f(\vec{x}) d\vec{x} \leq \overline{\int_I} f(\vec{x}) d\vec{x}.$$

OHNE BEWEIS

□

Der Begriff des Ober- und Unterintegrals lässt sich nun erweitern auf Funktionen, die auf beliebigen beschränkten Mengen $M \subseteq \mathbb{R}^n$ definiert sind.

SATZ 9.1.2

Es sei $M \subset \mathbb{R}^n$ beschränkt, $f \in \mathcal{F}_n^1(M)$ beschränkt. Es sei $I \subset \mathbb{R}^n$ ein kompaktes Intervall mit $M \subseteq I$ und

$$\tilde{f} : \begin{cases} I & \rightarrow & \mathbb{R} \\ \vec{x} & \mapsto & \tilde{f}(\vec{x}) \end{cases} \quad \text{definiert durch} \quad \tilde{f}(\vec{x}) := \begin{cases} f(\vec{x}) & \text{falls } \vec{x} \in M \\ 0 & \text{falls } \vec{x} \notin M \end{cases}.$$

Dann sind

$$\overline{\int_I \tilde{f}(\vec{x}) d\vec{x}} \quad \text{und} \quad \underline{\int_I \tilde{f}(\vec{x}) d\vec{x}}$$

unabhängig von der Wahl von I .

OHNE BEWEIS

□

DEFINITION 9.1.4

Es sei $M \subseteq \mathbb{R}^n$ beschränkt und $f \in \mathcal{F}_n^1(M)$ beschränkt, dann definieren wir das Ober- und Unterintegral von f über M durch

$$\overline{\int_M f(\vec{x}) d\vec{x}} := \overline{\int_I \tilde{f}(\vec{x}) d\vec{x}} \quad \text{und} \quad \underline{\int_M f(\vec{x}) d\vec{x}} := \underline{\int_I \tilde{f}(\vec{x}) d\vec{x}},$$

wobei I irgend ein kompaktes Intervall ist, das M enthält, und \tilde{f} wie in Satz 9.1.2 definiert ist (nach diesem Satz sind Ober- und Unterintegral von der Wahl von I unabhängig).

DEFINITION 9.1.5

Es sei $M \subset \mathbb{R}^n$ beschränkt und $f \in \mathcal{F}_n^1(M)$ beschränkt. f heißt integrierbar über M , wenn Ober- und Unterintegral von f über M übereinstimmen. Den gemeinsamen Wert nennen wir das Integral

$$\int_M f(\vec{x}) d\vec{x} := \overline{\int_M f(\vec{x}) d\vec{x}} = \underline{\int_M f(\vec{x}) d\vec{x}}$$

von f über M (auch Riemannsches Integral oder Bereichsintegral).

Die meisten Sätze über Integrale von Funktionen einer Variablen lassen sich auf Funktionen mehrerer Variablen verallgemeinern.

SATZ 9.1.3 (Linearität des Bereichsintegrals)

Es sei $M \subset \mathbb{R}^n$ beschränkt, $f_j \in \mathcal{F}_n^1(M)$ beschränkt für $j = 1, 2$. Wenn die Integrale

$$\int_M f_1(\vec{x}) d\vec{x} \quad \text{und} \quad \int_M f_2(\vec{x}) d\vec{x}$$

existieren, so existiert für beliebige reelle Zahlen c_1, c_2 auch

$$\int_M (c_1 f_1(\vec{x}) + c_2 f_2(\vec{x})) d\vec{x}$$

und es gilt

$$\int_M (c_1 f_1(\vec{x}) + c_2 f_2(\vec{x})) d\vec{x} = c_1 \int_M f_1(\vec{x}) d\vec{x} + c_2 \int_M f_2(\vec{x}) d\vec{x}.$$

OHNE BEWEIS

□

SATZ 9.1.4 (Monotonie des Bereichsintegrals)

Es sei $M \subset \mathbb{R}^n$ beschränkt und $f, g \in \mathcal{F}_n^1(M)$ integrierbar über M , sowie $f(\vec{x}) \leq g(\vec{x})$ für alle $\vec{x} \in M$, dann ist

$$\int_M f(\vec{x}) d\vec{x} \leq \int_M g(\vec{x}) d\vec{x}.$$

DEFINITION 9.1.6

Eine beschränkte Menge $M \subset \mathbb{R}^n$ heißt (Jordan-) messbar, falls

$$\int_M d\vec{x} = \int_M 1 d\vec{x}$$

existiert. Im Falle der Existenz heißt die Zahl

$$|M| = \int_M d\vec{x}$$

das Maß, oder auch der Inhalt oder das n -dimensionale Volumen von M . Im Fall $n = 2$ spricht man auch von der Fläche von M , im Fall $n = 3$ vom Rauminhalt von M .

SATZ 9.1.5 (Rechenregeln für das Maß)

Es seien $A, B \subseteq \mathbb{R}^n$ messbar, dann gilt

- (i) $|A|, |B| \leq |A \cup B| \leq |A| + |B|$, $0 \leq |A \cap B| \leq |A|, |B|$ (Dreiecksungleichung),
- (ii) $|[0, 1]^n| = 1$ (Normierung),
- (iii) $|\{\vec{x} + \vec{a} \mid \vec{a} \in A\}| = |A|$ für alle $\vec{x} \in \mathbb{R}^n$ (Translationsinvarianz),
- (iv) $|\{\lambda \vec{a} \mid \vec{a} \in A\}| = \lambda^n |A|$ für alle $\lambda \geq 0$ (Skalierung).

BEWEIS

(Übungsaufgabe) □

BEMERKUNG 9.1.1

Die üblichen geometrischen Objekte (Kugeln, Quader, ...) sind messbar. Das Maß $|\cdot|$ entspricht dem „anschaulichen“ Volumenbegriff.

DEFINITION 9.1.7

Eine Menge $M \subseteq \mathbb{R}^n$ heißt Nullmenge, falls es für jedes $\varepsilon > 0$ eine (ggf. unendliche) Folge (I_n) von Intervallen gibt mit

$$M \subseteq \bigcup_{n \in \mathbb{N}} I_n \quad \text{und} \quad \left| \bigcup_{n \in \mathbb{N}} I_n \right| < \varepsilon.$$

BEISPIEL 9.1.1

Jede in einer Hyperebene des \mathbb{R}^n enthaltene Menge ist eine Nullmenge. Messbare Mengen sind genau dann Nullmengen, wenn ihr Volumen Null ist. Es gibt aber auch nicht messbare Nullmengen, beispielsweise $\mathbb{Q}^n \subseteq \mathbb{R}^n$.

Mit Hilfe von Nullmengen kann man eine ganze Reihe von Kriterien formulieren:

SATZ 9.1.6

Es sei I ein kompaktes Intervall des \mathbb{R}^n . Es gilt:

- (i) (Integrabilitätskriterium von Lebesgue) $f : I \rightarrow \mathbb{R}$ ist genau dann (Bereichs-)integrierbar über I , falls $\{\vec{x} \in I \mid f \text{ un stetig in } \vec{x}\}$ eine Nullmenge ist.
- (ii) (Messbarkeitskriterium) $M \subseteq \mathbb{R}^n$ ist genau dann (Jordan-) messbar, wenn der Rand von M eine Nullmenge ist.

OHNE BEWEIS □

Zusammenfassen der beiden Kriterien ergibt

SATZ 9.1.7

Es sei $M \subseteq \mathbb{R}^n$ messbar und $f : \overline{M} \rightarrow \mathbb{R}$ stetig auf dem Abschluss \overline{M} . Dann ist f über M integrierbar.

OHNE BEWEIS

□

Der Mittelwertsatz 6.2.1 kann nun verallgemeinert werden:

SATZ 9.1.8 (Mittelwertsatz der Integralrechnung im \mathbb{R}^n)

Es sei $M \subseteq \mathbb{R}^n$ messbar und $f : M \rightarrow \mathbb{R}$ integrierbar, mit $a \leq f(\vec{x}) \leq b$ für alle $\vec{x} \in M$, dann gilt für das Bereichsintegral

$$a \cdot |M| \leq \int_M f(\vec{x}) d\vec{x} \leq b \cdot |M|.$$

OHNE BEWEIS

□

9.2. Mehrfache Integrale

Wir kommen nun zur praktischen Berechnung von Bereichsintegralen. Diese kann auf die Berechnung von mehrfachen, und diese auf die Berechnung von einfachen Integralen zurückgeführt werden. Wir beginnen mit dem Fall $n = 2$.

DEFINITION 9.2.1

Es sei $M \subseteq \mathbb{R}^2$ beschränkt, $f \in \mathcal{F}_n^1(M)$ und $M \subseteq [a_1, b_1] \times [a_2, b_2]$. Für jedes $x_2 \in [a_2, b_2]$ sei der Schnitt von M gegeben durch

$$M(x_2) := \{x_1 \in \mathbb{R} \mid (x_1, x_2)^T \in M\}.$$

Für jedes $x_2 \in [a_2, b_2]$ existiere das Integral

$$F(x_2) := \int_{M(x_2)} f(x_1, x_2) dx_1$$

und es existiere

$$(*) \int_{a_2}^{b_2} F(x_2) dx_2.$$

Dann heißt (*) das Doppelintegral über M , Schreibweise

$$\iint_M f(x_1, x_2) dx_1 dx_2.$$

Wenn speziell

$$M := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid a_2 \leq x_2 \leq b_2, \varphi(x_2) \leq x_1 \leq \psi(x_2) \right\}$$

ist, so schreibt man auch

$$\int_{a_2}^{b_2} \int_{\varphi(x_2)}^{\psi(x_2)} f(x_1, x_2) dx_1 dx_2.$$

SATZ 9.2.1 (Satz von Fubini)

Es sei $M \subseteq \mathbb{R}^2$ beschränkt und f integrierbar über M . Falls das Doppelintegral von f über M existiert, so ist es gleich dem Bereichsintegral:

$$\int_M f(x_1, x_2) d\vec{x} = \iint_M f(x_1, x_2) dx_1 dx_2.$$

Existiert insbesondere auch das andere Doppelintegral, so gilt

$$\int_M f(x_1, x_2) d\vec{x} = \int_M \int_M f(x_1, x_2) dx_1 dx_2 = \int_M \int_M f(x_1, x_2) dx_2 dx_1 .$$

BEWEIS

Wir zeigen die Behauptung nur für den Spezialfall eines kompakten Intervalls. Es sei $I = A \times B$ mit $A = [a, b]$ bzw. $B = [c, d]$, sowie \mathcal{Z}_n die äquidistante Zerlegung von I in n Teilintervalle pro Dimension, also $\mathcal{Z}_n = (I_{j,k})$ mit $I_{j,k} = A_j \times B_k$ und

$$A_j = \left[a + \frac{j-1}{n}(b-a), a + \frac{j}{n}(b-a) \right] , \quad B_k = \left[c + \frac{k-1}{n}(d-c), c + \frac{k}{n}(d-c) \right]$$

für $j, k = 1 \dots n$. Da f bereichsintegrierbar ist, müssen auch für diese spezielle Zerlegung die Ober- und Untersummen gegen den Integralwert w von f über I konvergieren. Andererseits möge das Doppelintegral

$$w' = \iint_I f(x, y) dy dx = \int_a^b \int_c^d f(x, y) dy dx = \int_a^b F(x) dx \quad \text{mit} \quad F(x) := \int_c^d f(x, y) dy ,$$

existieren, insbesondere existiert dann das Integral $F(x)$ für alle $x \in A$, und es ist nur noch $w' = w$ zu zeigen. Es sei dazu

$$\overline{M}_{j,k} = \sup\{f(\vec{x}) \mid \vec{x} \in A_j \times B_k\} , \quad \overline{N}_{x,k} = \sup\{f(x, y) \mid y \in B_k\} , \quad \overline{L}_j = \sup\{F(x) \mid x \in A_j\} .$$

Für jedes feste $x \in A$ konvergiert nun

$$\overline{S}_n(f(x, \cdot)) = \sum_{k=1}^n \overline{N}_{x,k} \cdot |B_k| = \sum_{k=1}^n \overline{N}_{x,k} \cdot \frac{d-c}{n} \xrightarrow[n \rightarrow \infty]{\geq} F(x) ,$$

wobei mit „ \geq “ gemeint ist, dass die Obersumme in n fallend ist. Andererseits konvergiert wegen der Existenz des Bereichsintegrals

$$\overline{S}_n(f(\cdot)) = \sum_{j,k=1}^n \overline{M}_{j,k} \cdot |I_{j,k}| = \sum_{j,k=1}^n \overline{M}_{j,k} \cdot \frac{(d-c)(b-a)}{n^2} \xrightarrow[n \rightarrow \infty]{\geq} w .$$

Nun ist $\overline{M}_{j,k}$ nicht kleiner als $\overline{N}_{x_j,k}$ wenn $x_j = a + \frac{j}{n}(b-a)$ der j -te Teilungspunkt der äquidistanten Zerlegung von A ist, da das Supremum über eine größere Menge gebildet wird. Man erhält

$$\begin{aligned} \overline{S}_n(f(\cdot)) &= \sum_{j,k=1}^n \overline{M}_{j,k} \cdot \frac{(d-c)(b-a)}{n^2} = \sum_{j=1}^n \left(\sum_{k=1}^n \overline{M}_{j,k} \cdot \frac{d-c}{n} \right) \frac{b-a}{n} \\ &\geq \sum_{j=1}^n \left(\sum_{k=1}^n \overline{N}_{x_j,k} \cdot \frac{d-c}{n} \right) \frac{b-a}{n} = \sum_{j=1}^n \overline{S}_n(f(x_j, \cdot)) \frac{b-a}{n} \geq \sum_{j=1}^n F(x_j) \frac{b-a}{n} \xrightarrow[n \rightarrow \infty]{} w' . \end{aligned}$$

Dabei gilt die letzte Konvergenz wegen Satz 6.1.9. Da die Relation \geq mit Grenzwertbildung verträglich ist folgt durch den Übergang $n \rightarrow \infty$ mit $\overline{S}_n(f(\cdot)) \rightarrow w$, dass $w \geq w'$ ist. Analog ist das zweidimensionale Infimum $\underline{M}_{j,k}$ nicht größer als $\underline{N}_{x_j,k}$, die obige Ungleichungskette kann daher mit umgekehrten Ungleichheitszeichen für die Untersummen geführt werden, woraus durch Grenzwertübergang $w \leq w'$ und damit die Behauptung folgt. \square

Einen wichtigen Spezialfall, für den die Voraussetzungen von Satz 9.2.1 erfüllt sind, bilden die Parameterintegrale:

DEFINITION 9.2.2

Ein Integral

$$\int_a^b f(x, u) dx,$$

bei dem Integrand neben der Integrationsvariablen x noch von einem Parameter u (aus einer beliebigen Menge) abhängt, heißt Parameterintegral.

SATZ 9.2.2

Es sei $R = [a, b] \times [c, d]$ und

$$f : \begin{cases} R & \rightarrow \mathbb{R} \\ (x, u) & \mapsto f(x, u) \end{cases}$$

stetig, dann ist das Parameterintegral

$$F(u) = \int_a^b f(x, u) dx$$

stetig (in u).

OHNE BEWEIS

□

BEMERKUNG 9.2.1

Da ebenso auch

$$G(x) := \int_c^d f(x, u) du$$

stetig (in x) ist folgt nach Satz 9.2.1

$$\int_R f(x, u) d\vec{x} = \int_a^b \int_c^d f(x, u) du dx = \int_c^d \int_a^b f(x, u) dx du.$$

BEISPIEL 9.2.1

Es sei $M = \{(x, y)^T \in \mathbb{R}^2 \mid x \in [0, 1], y \in [0, 1], x \leq y\}$, berechne

$$\int_M xy^2 d\vec{x}.$$

1. Lösungsansatz: Es ist $M = \{(x, y)^T \in \mathbb{R}^2 \mid 0 \leq x \leq 1, x \leq y \leq 1\}$. Also

$$\iint_M xy^2 dx dy = \int_0^1 \int_x^1 xy^2 dy dx = \int_0^1 x \left[\frac{y^3}{3} \right]_{y=x}^{y=1} dx = \frac{1}{3} \int_0^1 x(1 - x^3) dx = \frac{1}{3} \left(\frac{1}{2} - \frac{1}{5} \right) = \frac{1}{10}.$$

2. Lösungsansatz: Es ist $M = \{(x, y)^T \in \mathbb{R}^2 \mid 0 \leq y \leq 1, 0 \leq x \leq y\}$. Also

$$\iint_M xy^2 dx dy = \int_0^1 \int_0^y xy^2 dx dy = \int_0^1 y^2 \left[\frac{x^2}{2} \right]_{x=0}^{x=y} dy = \frac{1}{2} \int_0^1 y^4 dy = \frac{1}{10}.$$

Für beliebige $n \in \mathbb{N}$ werden n -fache Integrale nun rekursiv definiert:

DEFINITION 9.2.3

Es sei $n \geq 2$ und $M \subset \mathbb{R}^n$ beschränkt, $f \in \mathcal{F}_n^1(M)$ sowie $M \subseteq I = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Für alle $x_n \in [a_n, b_n]$ sei

$$M(x_n) := \{(x_1, \dots, x_{n-1})^T \in \mathbb{R}^{n-1} \mid (x_1, \dots, x_{n-1}, x_n) \in M\}.$$

Es existiere das $(n-1)$ -fache Integral

$$F(x_n) := \overbrace{\int \cdots \int}^{(n-1)\text{-mal}}_{M(x_n)} f(x_1, \dots, x_{n-1}, x_n) dx_1 \cdots dx_{n-1}$$

\uparrow
fest

sowie das Integral

$$(*) \int_{a_n}^{b_n} F(x_n) dx_n,$$

dann heißt $(*)$ das n -fache Integral, geschrieben

$$\overbrace{\int \cdots \int}^{n\text{-mal}}_M f(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Es gilt die Verallgemeinerung von Satz 9.2.1:

SATZ 9.2.3

Es sei $M \subset \mathbb{R}^n$ beschränkt und f integrierbar über M . j_1, \dots, j_n seien die Zahlen $1, \dots, n$ in irgend einer Reihenfolge. Dann ist

$$\int \cdots \int_M f(\vec{x}) dx_{j_1} \cdots dx_{j_n} = \int_M f(\vec{x}) d\vec{x}$$

falls das betreffende n -fache Integral existiert.

OHNE BEWEIS

□

BEISPIEL 9.2.2

Man berechne die z -Koordinate des Schwerpunkts des Tetraeders

$$\mathcal{T} := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x, y, z \geq 0, x + y + z \leq 1 \right\}$$

gegeben durch

$$\iiint_{\mathcal{T}} z \, dx dy dz.$$

Lösung: Man kann die Punktmenge \mathcal{T} auch schreiben als

$$\mathcal{T} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid 0 \leq z \leq 1, 0 \leq y \leq 1 - z, 0 \leq x \leq 1 - y - z \right\}.$$

Damit ist

$$\begin{aligned} \iiint_{\mathcal{T}} z \, dx \, dy \, dz &= \int_0^1 \int_0^{1-z} \int_0^{1-y-z} z \, dx \, dy \, dz = \int_0^1 z \left(\int_0^{1-z} (1-y-z) \, dy \right) dz = \int_0^1 z \left[y - \frac{y^2}{2} - zy \right]_{y=0}^{y=1-z} dz \\ &= \int_0^1 z \left(1-z - \frac{(1-z)^2}{2} - z(1-z) \right) dz = \frac{1}{2} \int_0^1 z(1-z)^2 dz = \frac{1}{2} \int_0^1 (z - 2z^2 + z^3) dz = \frac{1}{24}. \end{aligned}$$

Es stellt sich die Frage, ob in diesem Fall Tripelintegral und Bereichsintegral übereinstimmen:

$$\iiint_{\mathcal{T}} z \, dx \, dy \, dz = \int_{\mathcal{T}} z \, d\vec{x}$$

gilt nach Satz 9.2.1, falls $f(x, y, z) = z$ über \mathcal{T} integrierbar ist. Wegen der Stetigkeit von f folgt dies nach Satz 9.1.7 aus der Messbarkeit von \mathcal{T} . Die Messbarkeit von \mathcal{T} ist nun ein Spezialfall der Messbarkeit von Mengen in Normalform, die wir im folgenden Satz formulieren wollen:

Satz 9.2.4

Für $\nu = 1, \dots, n$ seien $a_\nu < b_\nu$ in \mathbb{R} . Für $\nu = 2, \dots, n$ sei $I_\nu = [a_\nu, b_\nu] \times \dots \times [a_{\nu-1}, b_{\nu-1}]$ und $\varphi_\nu, \psi_\nu : I_\nu \rightarrow \mathbb{R}$ stetig auf I_ν mit $\varphi_\nu(x_1, \dots, x_{\nu-1}) \leq \psi_\nu(x_1, \dots, x_{\nu-1})$ für alle $(x_1, \dots, x_{\nu-1})^T \in I_\nu$. Dann ist die Menge

$$\begin{aligned} M := \{ \vec{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid & a_1 \leq x_1 \leq b_1, \\ & \varphi_2(x_1) \leq x_2 \leq \psi_2(x_1), \\ & \vdots \\ & \varphi_\nu(x_1, \dots, x_{\nu-1}) \leq x_\nu \leq \psi_\nu(x_1, \dots, x_{\nu-1}), \\ & \vdots \\ & \varphi_n(x_1, \dots, x_{n-1}) \leq x_n \leq \psi_n(x_1, \dots, x_{n-1}) \} \end{aligned}$$

messbar mit Inhalt

$$|M| = \int_{a_1}^{b_1} \int_{\varphi_2(x_1)}^{\psi_2(x_1)} \dots \int_{\varphi_\nu(x_1, \dots, x_{\nu-1})}^{\psi_\nu(x_1, \dots, x_{\nu-1})} \dots \int_{\varphi_n(x_1, \dots, x_{n-1})}^{\psi_n(x_1, \dots, x_{n-1})} 1 \, dx_n \dots dx_\nu \dots dx_1.$$

OHNE BEWEIS

□

9.3. Die Substitutionsformel

Die Substitutionsregel, die in Satz 6.3.1 für unbestimmte Integrale gegeben wurde, lässt sich für bestimmte Integrale (in einer Variable) folgendermaßen formulieren: $f : [a, b] \rightarrow \mathbb{R}$ sei auf $[a, b]$ stetig. Wird das Intervall $[\alpha, \beta]$ durch die stetig differenzierbare und bijektive Funktion $u = g(x)$ auf das Intervall $[a, b]$ abgebildet, so gilt

$$\int_{\alpha}^{\beta} f(g(x))g'(x)dx = \int_a^b f(u)du.$$

Die Bedingung der Bijektivität folgt sicherlich aus der stärkeren Bedingung $g'(x) \neq 0$. Wie lautet die Verallgemeinerung dieser Regel auf \mathbb{R}^n , wenn das Intervall $[\alpha, \beta]$ durch eine kompakte Menge $\mathcal{T} \subset \mathbb{R}^n$ ersetzt wird, die durch \vec{g} bijektiv auf $\mathcal{T}' = \vec{g}(\mathcal{T})$ abgebildet wird? Wie lässt sich das Integral

$$\int_{\mathcal{T}'} f(\vec{u}) d\vec{u}$$

als Integral über \mathcal{T} ausdrücken?

BEMERKUNG 9.3.1

Es sei D der Term, der g' aus der eindimensionalen Substitutionsformel verallgemeinert. Der folgende einfache Spezialfall führt auf den richtigen Ausdruck für D : Es sei $\vec{g}(\vec{x}) = A\vec{x}$ linear für eine Matrix $A \in \mathbb{R}^{n \times n}$ sowie $T = [0, 1]^n$ der normierte Würfel mit Maß Eins. Wir nehmen an, dass D konstant in \vec{x} ist, da die totale Ableitung von \vec{g} konstant ist. Anschaulich ist $\vec{g}(T)$ das von den Spaltenvektoren von A aufgespannte Parallelepiped, dessen Volumeninhalt $|\det(A)|$ ist (siehe Lineare Algebra). Dann ist (mit $f = 1$ konstant)

$$\int_{\vec{g}(T)} f(\vec{u}) d\vec{u} = |\vec{g}(T)| = |\det(A)|$$

und andererseits

$$\int_T (f(\vec{g}) \cdot D) d\vec{x} = \int_T D d\vec{x} = D \cdot |T| = D.$$

Beide Integrale sollen übereinstimmen, also $D = |\det(A)|$.

Die allgemeine Antwort gibt der folgende Satz. Er zeigt, dass die richtige Verallgemeinerung von g' tatsächlich die Funktionaldeterminante $\det\left(\frac{d\vec{g}}{d\vec{x}}(\vec{x})\right)$ ist.

SATZ 9.3.1

Die Funktion $\vec{g} \in \mathcal{F}_n^n$ sei auf der offenen Menge $M \subseteq \mathbb{R}^n$ injektiv und stetig differenzierbar, sowie $\det\left(\frac{d\vec{g}}{d\vec{x}}(\vec{x})\right)$ auf M ständig positiv oder ständig negativ. Ferner sei \mathcal{T} eine kompakte, messbare Teilmenge von M und f eine auf $\vec{g}(T)$ stetige Funktion. Dann ist $\vec{g}(T)$ messbar, f auf $\vec{g}(T)$ integrierbar, und

$$\int_{\vec{g}(T)} f(\vec{u}) d\vec{u} = \int_T f(\vec{g}(\vec{x})) \left| \det \frac{d\vec{g}}{d\vec{x}}(\vec{x}) \right| d\vec{x}.$$

Diese Aussage gilt auch dann noch, wenn die obigen Voraussetzungen der Injektivität von \vec{g} und des Nichtverschwindens der Funktionaldeterminante auf einer Nullmenge $N \subseteq M$ nicht erfüllt sind.

OHNE BEWEIS

□

Wir führen die wichtigsten Spezialfälle dieser Integrationstechnik ein:

Integration in Polarkoordinaten:

Der Punkt $(x, y)^T \in \mathbb{R}^2 \setminus \{(0, 0)^T\}$ hat die Polarkoordinaten (r, ϕ) , wobei r und ϕ eindeutig bestimmt sind durch

$$x = r \cdot \cos(\phi) \quad , \quad 0 \leq \phi < 2\pi$$

$$y = r \cdot \sin(\phi) \quad , \quad r > 0.$$

Die Abbildung

$$\vec{g}(r, \phi) = \begin{pmatrix} r \cos(\phi) \\ r \sin(\phi) \end{pmatrix}$$

sowie die Mengen M und \mathcal{T} mögen die Voraussetzungen von Satz 9.2.4 erfüllen. Dann gilt mit $\vec{p} = (r, \phi)^T$ die Substitutionsregel

$$\int_{\vec{g}(\mathcal{T})} f(x, y) d\vec{x} = \int_{\mathcal{T}} f(r \cos \phi, r \sin \phi) r d\vec{p}$$

für jede auf $\vec{g}(\mathcal{T})$ stetige und reellwertige Funktion. Das folgt aus Satz 9.3.1, da

$$\det \frac{d\vec{g}}{d\vec{x}}(\vec{x}) = \det \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} = r(\cos^2 \phi + \sin^2 \phi) = r > 0$$

ist. Spezialfall: es sei \mathcal{T} das Rechteck $\mathcal{T} = \{\vec{p} = (r, \phi)^T, 0 \leq r \leq R, 0 \leq \phi \leq 2\pi\}$, dann ist $\vec{g}(\mathcal{T})$ die Kreisscheibe $\{(x, y)^T \in \mathbb{R}^2 \mid x^2 + y^2 \leq R^2\}$ mit Mittelpunkt $(0, 0)^T$ und Radius R . \vec{g} ist injektiv auf $\mathcal{T} \setminus \{(r, 2\pi)^T \mid 0 \leq r \leq R\} \cup \{(0, \phi) \mid 0 \leq \phi \leq 2\pi\}$, und es ist $\det \frac{d\vec{g}}{d\vec{p}} \neq 0$ für $\vec{p} \neq (0, \phi)^T$. Für jede auf $\vec{g}(\mathcal{T})$ stetige Funktion f gilt

$$\iint_{x^2+y^2 \leq R^2} f(x, y) dx dy = \int_0^{2\pi} \int_0^R f(r \cos \phi, r \sin \phi) r dr d\phi.$$

Beispielsweise ist

$$\iint_{x^2+y^2 \leq R^2} (x^2 + y^2) dx dy = \int_0^{2\pi} \int_0^R r^3 dr d\phi = 2\pi \frac{R^4}{4}.$$

BEISPIEL 9.3.1

Es soll das uneigentliche Integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

berechnet werden. Dazu sei

$$I(R) = \int_{-R}^R e^{-x^2} dx, \quad \mathcal{R}(R) = \{(x, y)^T \in \mathbb{R}^2 \mid |x| \leq R, |y| \leq R\}.$$

Dann ist

$$I(R)^2 = \left(\int_{-R}^R e^{-x^2} dx \right)^2 = \int_{-R}^R e^{-x^2} dx \int_{-R}^R e^{-y^2} dy = \int_{-R}^R \int_{-R}^R e^{-(x^2+y^2)} dy dx.$$

Man zeigt leicht, dass

$$I^2 = \lim_{R \rightarrow \infty} I(R)^2 = \lim_{R \rightarrow \infty} \int_{x^2+y^2 \leq R^2} e^{-(x^2+y^2)} d\vec{x}.$$

Nun ist

$$\int_{x^2+y^2 \leq R^2} e^{-(x^2+y^2)} d\vec{x} = \int_0^{2\pi} \int_0^R \exp(-r^2 \cos^2 \phi - r^2 \sin^2 \phi) dr = 2\pi \int_0^R r e^{-r^2} dr = \pi(1 - e^{-R^2}).$$

Damit ist $I^2 = \pi$ und

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Integration in Zylinderkoordinaten:

Die Zylinderkoordinaten eines Punktes $\vec{x} = (x, y, z)^T \in \mathbb{R}^3$ sind gegeben durch die Polarkoordinaten der „Projektion“ $(x, y)^T$ und die z -Koordinate. Also

$$x = r \cos \phi \quad , \quad y = r \sin \phi \quad , \quad z = z .$$

Die Abbildung

$$\vec{g} \begin{pmatrix} r \\ \phi \\ z \end{pmatrix} := \begin{pmatrix} r \cos \phi \\ r \sin \phi \\ z \end{pmatrix}$$

und die Mengen M und \mathcal{T} mögen die Voraussetzungen von Satz 9.3.1 erfüllen. Dann gilt

$$\int_{\vec{g}(\mathcal{T})} f(\vec{x}) d\vec{x} = \int_{\mathcal{T}} f(r \cos \phi, r \sin \phi, z) r d\vec{\zeta}$$

mit $\vec{\zeta} = (r, \phi, z)^T$ für jede auf $\vec{g}(\mathcal{T})$ stetige Funktion f .

Integration in Kugelkoordinaten:

Der Punkt $P : \vec{x} = (x, y, z)^T \in \mathbb{R}^3$ wird senkrecht auf die xy -Ebene projiziert. Dabei entsteht der Punkt $P^* : (x, y, 0)^T$. Als Kugelkoordinate von P nehmen wir den Radius $r = \|\vec{x}\|$, die Winkelkoordinate φ des Punktes P^* , und als dritte Koordinate den Winkel ψ zwischen den Strecken \overline{OP} und $\overline{OP^*}$. Also

$$x = r \cos \varphi \cos \psi \quad , \quad y = r \sin \varphi \cos \psi \quad , \quad z = r \sin \psi .$$

Mit $\vec{k} = (r, \varphi, \psi)^T$ ist die Funktionaldeterminante

$$\det \frac{d\vec{x}}{d\vec{k}} = \det \begin{pmatrix} \cos \varphi \cos \psi & -r \sin \varphi \cos \psi & -r \cos \varphi \sin \psi \\ \sin \varphi \cos \psi & r \cos \varphi \cos \psi & -r \sin \varphi \sin \psi \\ \sin \psi & 0 & r \cos \psi \end{pmatrix} = r^2 \cos \psi .$$

Die Abbildung

$$\vec{g} : \begin{pmatrix} r \\ \varphi \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} r \cos \varphi \cos \psi \\ r \sin \varphi \cos \psi \\ r \sin \psi \end{pmatrix} ,$$

und die Mengen M und \mathcal{T} mögen den Voraussetzungen von Satz 9.3.1 erfüllen. Dann gilt

$$\int_{\vec{g}(\mathcal{T})} f(x, y, z) d\vec{x} = \int_{\mathcal{T}} f(r \cos \varphi \cos \psi, r \sin \varphi \cos \psi, r \sin \psi) r^2 \cos(\psi) d\vec{k}$$

für jede auf $\vec{g}(\mathcal{T})$ stetige Funktion f . Spezialfall: es sei \mathcal{T} der Quader

$$\mathcal{T} := \left\{ \begin{pmatrix} r \\ \varphi \\ \psi \end{pmatrix} \in \mathbb{R}^3 \mid 0 \leq r \leq R, 0 \leq \varphi \leq 2\pi, -\frac{\pi}{2} \leq \psi \leq \frac{\pi}{2} \right\} .$$

Dann ist $\vec{g}(\mathcal{T})$ die Kugel mit Mittelpunkt $\vec{0}$ und Radius R :

$$\vec{g}(\mathcal{T}) = \{ \vec{x} = (x, y, z) \mid x^2 + y^2 + z^2 \leq R^2 \} .$$

Für jede auf $\vec{g}(\mathcal{T})$ stetige Funktion f gilt

$$\int_{x^2+y^2+z^2 \leq R^2} f(x, y, z) dx dy dz = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^R f(r \cos \varphi \cos \psi, r \sin \varphi \cos \psi, r \sin \psi) r^2 \cos(\psi) dr d\varphi d\psi .$$

9.4. Kurven- und Oberflächenintegrale

DEFINITION 9.4.1

Auf $M \subseteq \mathbb{R}^n$ sei ein stetiges Vektorfeld

$$\vec{f} \in \mathcal{F}_n^n(M) \quad \vec{x} \mapsto \vec{f}(\vec{x}) = (f_1(\vec{x}), \dots, f_n(\vec{x}))^T$$

gegeben. Ferner sei

$$\gamma : \begin{cases} [a, b] & \rightarrow M \\ t & \mapsto \vec{x}(t) \end{cases}$$

eine (stückweise) glatte Kurve, deren Träger komplett in M liegt. Dann erklären wir das Kurvenintegral durch

$$\int_{\gamma} \vec{f}(\vec{x}) d\vec{x} := \int_a^b \vec{f}(\vec{x}(t)) \frac{d\vec{x}(t)}{dt} dt.$$

BEISPIEL 9.4.1

Es sei

$$\gamma : \begin{cases} [0, 2\pi] & \rightarrow \mathbb{R}^3 \\ t & \mapsto \vec{x}(t) \end{cases}$$

gegeben durch $\vec{x}(t) = (\cos t, \sin t, t)^T$, sowie $\vec{f}(x, y, z) = (x, z, z^2)^T$. Dann ist

$$\begin{aligned} \int_{\gamma} \vec{f}(\vec{x}) d\vec{x} &= \int_0^{2\pi} \begin{pmatrix} \cos t \\ t \\ t^2 \end{pmatrix}^T \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix} dt = \int_0^{2\pi} (-\sin t \cos t + t \cos t + t^2) dt \\ &= \left[-\frac{\sin^2(t)}{2} + t \sin t + \cos t + \frac{t^3}{3} \right]_{t=0}^{t=2\pi} = \frac{8}{3} \pi^3. \end{aligned}$$

DEFINITION 9.4.2

Es sei $K \neq \emptyset$ ein kompaktes Intervall des \mathbb{R}^2 . Unter einer Fläche Φ mit dem Parameterbereich K verstehen wir eine stetig partiell differenzierbare Abbildung $\Phi : K \rightarrow \mathbb{R}^3$, deren Funktionalmatrix den Rang 2 besitzt. Die Bildmenge $S = \Phi(K)$ wird ein Flächenstück mit der Parameterdarstellung $\vec{v} = \Phi(u, v)$ und dem Parameterbereich K genannt.

BEMERKUNG 9.4.1

Man kann Flächen auch mit allgemeineren Parameterbereichen definieren, der Einfachheit halber beschränken wir uns hier auf kompakte Intervalle $K = I \times J$.

BEISPIEL 9.4.2

Die Oberfläche der Kugel mit Radius R um $\vec{0}$ kann parametrisiert werden durch

$$\vec{r} = \Phi(u, v) = (R \cos u \cos v, R \sin u \cos v, R \sin u)^T$$

mit $0 \leq u \leq 2\pi$ und $-\frac{\pi}{2} \leq v \leq \frac{\pi}{2}$.

DEFINITION 9.4.3

Unter dem Vektorprodukt (oder Kreuzprodukt) zweier Vektoren $\vec{a} = (a_1, a_2, a_3)^T$ und $\vec{b} = (b_1, b_2, b_3)^T$ versteht man den Vektor

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

BEMERKUNG 9.4.2

Die komplizierte Definition lässt sich leicht merken, wenn man sie symbolisch schreibt als

$$\vec{a} \times \vec{b} = \det \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}.$$

Man kann zeigen: $\vec{a} \times \vec{b}$ steht senkrecht auf seinen Faktoren \vec{a} und \vec{b} . Seine Länge $\|\vec{a} \times \vec{b}\|$ ist gleich dem Flächeninhalt des von \vec{a} und \vec{b} aufgespannten Parallelogramms.

DEFINITION 9.4.4

Es sei eine Fläche Φ mit dem Parameterbereich K gegeben. Dann nennt man den Vektor

$$\vec{N}(u, v) = \frac{\partial \Phi}{\partial u} \times \frac{\partial \Phi}{\partial v}$$

einen Normalenvektor von Φ im Flächenpunkt $\Phi(u, v)$.

DEFINITION 9.4.5

Es sei eine Φ Fläche mit dem Parameterbereich K und $\vec{f}(\vec{x})$ ein stetiges Vektorfeld auf $\Phi(K)$. Man nennt

$$\int_{\Phi} \vec{f}(\vec{x}) \cdot \vec{n} db := \int_K \vec{f}(\Phi(u, v)) \cdot \vec{N}(u, v) dudv$$

das Oberflächenintegral 2. Art von f über Φ . Ist f eine stetige reellwertige Funktion auf $\Phi(K)$, so nennt man

$$\int_{\Phi} f d\sigma := \int_K f(\Phi(u, v)) \|\vec{N}(u, v)\| dudv$$

das Oberflächenintegral 1. Art von f über Φ . Der Flächeninhalt von Φ ist definiert als

$$\int_{\Phi} 1 d\sigma.$$

BEISPIEL 9.4.3

Wir wollen den Flächeninhalt der Kugeloberfläche $S = \{(x, y, z)^T \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = R^2\}$ mit Radius R berechnen. Es ist $S = \Phi(K)$ mit

$$\Phi(u, v) = R \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ \sin v \end{pmatrix}, \quad K = [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Damit ist

$$\frac{\partial \Phi}{\partial u} = R \begin{pmatrix} -\sin u \cos v \\ \cos u \cos v \\ 0 \end{pmatrix}, \quad \frac{\partial \Phi}{\partial v} = R \begin{pmatrix} -\cos u \sin v \\ -\sin u \sin v \\ \cos v \end{pmatrix},$$

$$\vec{N}(u, v) = R^2 \begin{pmatrix} \cos(u) \cos(v)^2 \\ \sin(u) \cos(v)^2 \\ \cos(v) \sin(v) \end{pmatrix}, \quad \|\vec{N}(u, v)\| = R^2 \cos(v).$$

Damit ergibt sich das Oberflächenintegral 1. Art (mit $f = 1$ konstant in Definition 9.4.5) zu

$$\int_{\Phi} d\sigma = \int_K \|\vec{N}(u, v)\| d(u, v) \stackrel{9.2.1}{=} \int_0^{2\pi} \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} R^2 \cos(v) dv du = R^2 \cdot \int_0^{2\pi} [\sin(v)]_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} du = R^2 \cdot \int_0^{2\pi} 2 du = 4\pi R^2.$$

DEFINITION 9.4.6

Die Rotation $\text{rot } \vec{f}$ eines Vektorfeldes $\vec{f} = (f_1, f_2, f_3)^T$ ist

$$\text{rot } \vec{f} = \begin{pmatrix} \frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \\ \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \\ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \end{pmatrix}$$

oder als symbolische Merkgel $\text{rot } \vec{f} = \text{grad} \times \vec{f}$. Ein Vektorfeld $\vec{f} = (f_1, f_2, f_3)^T$ heißt rotationsfrei auf $M \subseteq \mathbb{R}^3$, falls $\text{rot } \vec{f} = 0$ auf M ist.

BEISPIEL 9.4.4

SATZ 9.4.1 (Stokesscher Integralsatz)

Es sei Φ eine Fläche mit Parameterbereich K . Φ sei zweimal stetig differenzierbar auf einer K enthaltenden offenen Menge. Die Kurve γ sei eine stückweise glatte Kurve, welche eine positive Orientierung des Randes $\text{Rd}K$ darstellt. Das Vektorfeld $\vec{f} = (f_1, f_2, f_3)^T$ sei stetig differenzierbar auf einer offenen Menge, die $\Phi(K)$ enthält. Dann gilt

$$\int_{\Phi} \text{rot } \vec{f} \cdot \vec{n} db = \int_{\Phi \circ \gamma} \vec{f}(\vec{x}) d\vec{x}.$$

OHNE BEWEIS

□

BEMERKUNG 9.4.3

Der Integralsatz von Stokes kann als Verallgemeinerung des Hauptsatzes der Integral- und Differentialrechnung

$$\int_a^b f'(x) dx = f(b) - f(a)$$

angesehen werden, wenn man $\Phi = [a, b]$ mit $\text{Rd}([a, b]) = \{a, b\}$ und $\text{rot}(f) = f'$ im eindimensionalen Fall setzt. Sein Beweis ist jedoch ungleich komplizierter.

SATZ 9.4.2 (Gaußscher Integralsatz im Raum)

Es sei V eine (Jordan-) messbare Teilmenge des \mathbb{R}^3 und der Rand $\text{Rd}V$ sei ein Flächenstück mit einer Parameterdarstellung $\vec{v} = \Phi(u, v)$ und Parameterbereich K derart, dass Φ auf K (bis auf eine Nullmenge) bijektiv ist, und dass $h\vec{v} \in V$ für hinreichend kleines $h > 0$. \vec{f} sei ein stetig differenzierbares Vektorfeld auf einer offenen, V umfassenden Menge. Dann gilt

$$\int_V \text{div } \vec{f} d\vec{x} = \int_{\text{Rd}V} \vec{f} \cdot \vec{n} db.$$

OHNE BEWEIS

□

Anstelle eines Beweises geben wir

BEISPIEL 9.4.5 (Physikalische Deutung der Divergenz)

Wir betrachten eine *stationäre Strömung* im Raum, d. h. eine Flüssigkeitsströmung, deren Geschwindigkeit in jedem Punkt $(x, y, z)^T$ nicht von der Zeit abhängt. Dabei wird jedem Raumpunkt die *Flussdichte* $\vec{f}(x, y, z) \in \mathbb{R}^3$ mit der Einheit $\frac{\text{kg}}{\text{m}^2 \cdot \text{sec}}$ zugeordnet, die das Produkt aus der Geschwindigkeit der Flüssigkeit im Punkt $(x, y, z)^T$ und der Dichte in diesem Punkt ist. Bei nicht komprimierbaren

Flüssigkeiten kann die Dichte als nicht von $(x, y, z)^T$ abhängige Konstante vernachlässigt werden. Im Gegensatz dazu entsteht der Begriff der *Quelldichte* $q(\vec{x}_0)$ in einem Punkt \vec{x}_0 dadurch, dass man das Strömungsverhalten durch einen gleichseitigen achsenparallelen Würfel um \vec{x}_0 für den Übergang der Kantenlänge l gegen Null betrachtet. In jeder der drei Achsenrichtungen weist der Würfel zwei quadratische Flächen auf. Die Summe der in x -Richtung ein- bzw. austretenden Masse pro Zeiteinheit muss nicht Null sein, es könnte beispielsweise mehr in x -Richtung herausfließen als in x -Richtung aus dem Würfel hineinfließt: man spricht von der *Quellenstärke* des Würfels in x -Richtung. Beim Grenzwertübergang $l \rightarrow 0$ geht die Quellenstärke in x -Richtung im Punkt \vec{x}_0 über in $\frac{\partial \vec{f}}{\partial x}(\vec{x}_0)$. Die Summe der Quellenstärken in den drei Achsenrichtungen ist dann im Grenzwert die *Quelldichte*

$$q(\vec{x}_0) = \frac{\partial \vec{f}}{\partial x}(\vec{x}_0) + \frac{\partial \vec{f}}{\partial y}(\vec{x}_0) + \frac{\partial \vec{f}}{\partial z}(\vec{x}_0) = \operatorname{div} \vec{f}(\vec{x}_0) \left[\frac{\text{kg}}{\text{m}^3 \cdot \text{sec}} \right].$$

Dabei wird $q(\vec{x}_0) > 0$ so interpretiert, dass im Punkt \vec{x}_0 Zufluss von neuer Flüssigkeit stattfindet. Für $q(\vec{x}_0) < 0$ findet in \vec{x}_0 ein Abfluss von Flüssigkeit statt.

BEISPIEL 9.4.6 (Physikalische Deutung des Gaußschen Integralsatzes)

Wieder sei $\vec{f}(\vec{x})$ ein Vektorfeld, das die Flussdichte eines stationären Flusses einer nicht komprimierbaren Flüssigkeit beschreibe. Es sei $V \subseteq \mathbb{R}^3$ irgend ein räumlicher Körper, den man sich innerhalb des Stromes denkt, der aber den Fluss nicht stört. Die linke Seite des Satzes 9.4.2

$$\int_V \operatorname{div} \vec{f}(\vec{x}) d\vec{x}$$

beschreibt in diesem Kontext die gesamte Quellenstärke von V , d. h. die Flüssigkeitsmenge, die dem Gesamtsystem innerhalb V hinzufließt. Die rechte Seite des Satzes

$$\int_{\operatorname{Rd}(V)} \vec{f} \cdot \vec{n} db$$

beschreibt die Flüssigkeitsmenge, die entlang der Normalenvektoren auf V durch Passierung des Randes von V in das System abgegeben/aufgenommen wird. Da die Flüssigkeit nicht komprimierbar ist, muss anschaulich die in V hinzugefügte Flüssigkeitsmenge mit der aus V in das System fließenden Menge übereinstimmen, was die Aussage des Gaußschen Integralsatzes ist.

BEISPIEL 9.4.7

Wir wollen das Volumen der Kugel mit Radius R berechnen. Es ist

$$\vec{v} = \Phi(u, v) = R \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ \sin v \end{pmatrix}, \quad K = [0, 2\pi] \times [0, 2\pi].$$

Wir setzen $\vec{f} = (x, y, z)^T = \vec{v}$ mit $\operatorname{div} \vec{f} = 3$, dann gilt

$$\iiint_V 3 dx dy dz = R^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \left\| \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ \sin v \end{pmatrix} \right\|^2 \cos v du dv = 4\pi R^3 \Rightarrow |V| = \iiint_V dx dy dz = \frac{4\pi}{3} R^3$$

für das Volumen der Kugel.

Wir betrachten nun den Stokesschen Integralsatz für den Spezialfall eines Vektorfeldes in der Ebene und eines ebenen Flächenstücks. Wir nehmen in Satz 9.4.1 an, dass $\vec{f} = (f_1, f_2, 0)^T$ und $\Phi: K \rightarrow M$

durch $\Phi(u, v) = (u, v, 0)^T$ gegeben ist. Dann ist

$$\operatorname{rot} \vec{f} = \det \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{\partial f_2}{\partial x} \\ \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \end{pmatrix}.$$

Es gilt

$$\frac{\partial \Phi}{\partial u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial \Phi}{\partial v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{n} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Nach Satz 9.4.1 ist dann

$$\int_{\Phi} \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) d\sigma = \int_{\gamma} \vec{f}(\vec{x}) d\vec{x},$$

und es gilt

SATZ 9.4.3 (Gaußscher Integralsatz in der Ebene)

Es sei V eine (Jordan-) messbare Teilmenge des \mathbb{R}^2 und der Rand $\operatorname{Rd}V$ sei eine Kurve mit einer Parameterdarstellung $\gamma : t \mapsto \vec{x}(t)$ für $t \in [a, b]$ derart, dass γ bis auf eine Nullmenge bijektiv ist und $h\vec{x} \in V$ gilt für genügend kleines h . \vec{f} sei ein stetig differenzierbares Vektorfeld auf einer offenen, V umfassenden Menge. Dann gilt

$$\int_V \left(\frac{\partial f_2}{\partial x}(\vec{x}) - \frac{\partial f_1}{\partial y}(\vec{x}) \right) d\vec{x} = \int_{\gamma} \vec{f} d\vec{x}.$$

BEISPIEL 9.4.8

Mit diesem Satz kann man den Inhalt von V durch ein Kurvenintegral über den Rand ausdrücken:

$$|V| = \frac{1}{2} \int_{\gamma} \begin{pmatrix} -y \\ x \end{pmatrix} d\vec{x}.$$

Index

- ε -Umgebung 26, 100
- n -faches Integral 129
- n -te Ableitung 53
- n -te Potenz 20
- n -te Wurzel 45
- $n \rightarrow n + 1$, 22

- Abbildung** 23
- abgeschlossen 42, 43, 101
- abgeschlossene Hülle 101
- Ableitung 53
- abnehmend 45
- Abschätzung 14
- Abschluss 101
- absolut konvergent 35, 93
- Abstand 99
- Addition 8, 98
- additives Inverses 11
- alternierende Reihe 37
- Anfangspunkt 112
- angeordneter Körper 10
- äquivalent 112
- Äquivalenzrelation 8
- Assoziativgesetz 8, 9
- ausgezeichnet 80
- Axiome der Addition 8
- Axiome der Multiplikation 9
- Approximation 55

- Basis** 20
- bedingt Konvergent 38
- beschränkt 10, 27, 100, 102
- beschränkt (nach oben/unten) 10
- Bereichsintegral 124
- Besetzung 83
- Betrag 26, 72, 98
- Beweise 5
- Beweise durch Induktion 20
- bijektiv 25
- Bild 23, 24
- Binomialkoeffizient 21
- Bogen 112

- Cauchyfolge** 30
- Cauchy Kriterium 73
- Cauchyprodukt 39

- Dedekindscher Schnitt** 18
- Definitheit (Betrag) 26
- Definitheit (Matrix) 117
- Definitionsbereich 23
- Differentiationsklasse 110
- Differentiationsoperator 53
- Differentiationsregeln 56
- Differenz 6
- Differenzenquotient 53
- differenzierbar 53, 106, 112
- disjunkt 8
- Distributivgesetz, 9
- Divergenz (von Folgen) 26, 33, 91, 92
- Divergenz (der Operator) 107
- Divergenz (uneigentliche) 32
- Doppelintegral 126
- Dreiecksungleichung 26, 98, 99, 125
- Durchschnitt 6

- Eindeutigkeit** 12
- Eins 11
- Einschränkung 25
- einseitiger Grenzwert 47
- einseitige Stetigkeit 47
- endliche Überdeckung 44, 102
- Endpunkt 112
- Endstück 27
- entgegengesetzt orientiert 112
- Euklidische Metrik 99
- Euklidischer Raum 98
- Existenz des Inversen 8
- Exponenten 20
- Exponentialfunktion 39
- Extrema (eindimensional) 75
- Extrema (mehrdimensional) 117
- Extrema mit Nebenbedingungen 122

- Fakultät** 21
- fast immer 28
- Feinheit 80
- Fläche 134
- Flächeninhalt 125, 135
- Flächenproblem 79
- Flächenstück 134
- Fortsetzung 25
- Fubini (Satz von) 126
- Funktion 23
- Funktionaldeterminante 107
- Funktionalmatrix 107
- Funktionentheorie 73

Gebiet 113
 glatt 112
 gleiche Funktionen 24
 gleichmäßig konvergent 50, 104
 gleichmäßig stetig 44, 102
 globales Minimum 54
 globales Maximum 54
 größte untere Schranke 11
 Gradient 107
 Graph 24
 Grenzfunktion 50
 Grenzwert 26, 46, 48, 72, 99
 Gruppe 9

Häufungspunkt 46, 101
 Häufungswert 30, 100
 Hauptsatz (der Integral-/Differentialrechnung) 87
 Hauptsatz (über implizite Funktionen) 119
 halboffen 10
 harmonische Reihe 34
 Hessesche Matrix 117
 l'Hôpital (Regel von) 74

Imaginärachse 72
 imaginäre Einheit 72
 Imaginärteil 72
 implizit 119
 indefinit (Matrix) 117
 Indexverschiebung 21
 Induktionsanfang 20
 Induktionsschluss 20
 induktive Menge 16
 Infimum 11
 Inhalt 123, 125
 injektiv 24
 Inneres 101
 innerer Kern 101
 innere Funktion 25
 innerer Punkt 54, 101
 Integrabilitätskriterien 82, 125
 Integral 82, 124
 Integralsatz (Gauß) 136
 Integralsatz (Stokes) 136
 Integration 88
 Integration in Kugelkoordinaten 133
 Integration in Polarkoordinaten 131
 Integration in Zylinderkoordinaten 133
 Integrationsbereich 82
 Integrationsvariable 82
 integrierbar 82, 124
 Intervallschachtelung 30
 inverse Funktion 25
 Inverses 8
 isolierter Punkt 46, 102

Jacobische Determinante 107
 Jensensche Ungleichung 75

Körper 8, 9
 Kartesisches Produkt 23
 Kettenregel 88, 109
 kleinstes Element 10

kleinste obere Schranke 11
 Koeffizientenvergleich 90
 kommutativer Ring 9
 Kommutativgesetz 8, 9
 kompakt 42, 43, 101
 kompaktes Intervall 123
 Komplement 7
 Komplementierungsregeln 7
 komplexe Koeffizienten 90
 komplexe Zahlen 71
 komplexe Zahlenfolgen 72
 Komponenten 23
 Komposition 25
 Kompositum 25
 konjugierte komplexe Zahl 72
 konkav 75
 konvergent 26, 33, 91, 92
 konvergente Majorante 94
 Konvergenzbereich 61
 Konvergenzintervall 73
 Konvergenzkreis 73
 Konvergenzkriterien 29, 34
 Konvergenzradius 63, 73
 konvex 75, 113
 Konvexität 78
 Kosinus 68
 Kotangens 70
 Kreuzprodukt 134
 Kurve 112
 Kurvendiskussion 77
 Kurvenintegral 134

Länge 113
 Laplace-Operator 111
 leere Menge 6
 Leibnizsche Regel 37
 Lehrsätze 5
 Limes 26, 99
 Limes inferior 31
 Limes superior 31
 lineare Approximation 55
 Linearität des Integrals 84
 linksseitig stetig 48
 linksseitiger Grenzwert 47
 Logarithmusfunktion 52
 lokale Extrema und Monotonie 78
 lokales Maximum 54
 lokales Minimum 54

Maß 125
 Majorante 36
 Maximalstelle 44
 Maximum 10
 mehrere Veränderliche 24, 98
 Menge 6
 Mengensystem 6
 messbar 125
 Messbarkeitskriterium 125
 Metrik 99
 metrischer Raum 99
 Minimalstelle 44
 Minimum 10

Minorante 36
 Mittelwertsatz (der Differentialrechnung) 58, 113
 Mittelwertsatz (der Integralrechnung) 86, 126
 monoton 29, 45
 monoton fallend 29
 monoton wachsend 29
 Monotoniegesetz 10
 Monotoniekriterium 34
 Multiplikation 8
 multiplikatives Inverses 11
 Multiplikativität 26

Näherung 55
 natürliche Exponentialfunktion 39
 natürlicher Logarithmus 52
 negativ 10
 neutrales Element 8, 9
 Neutralität der Eins 5
 nichtnegativ 10
 nichtpositiv 10
 Normalenvektor 134
 Normalform 130
 Normierung 125
 Null 11
 Nullfolge 28, 34
 Nullmenge 125
 Nullpolynom 29
 Nullstelle 29

Oberflächenintegral 1. Art 135
 Oberflächenintegral 2. Art 135
 Oberintegral 81, 123
 Oberklasse 18
 Obermenge 6
 Obersumme 79, 123
 offen 42, 43, 101
 offene Überdeckung 44, 102
 Ordnungssaxiome 73
 orientierte Kurve 112

Parameterbereich 134
 Parameterdarstellung 112, 134
 Parameterintegral 128
 Partialbruchzerlegung 89
 Partialsommen 33
 partiell differenzierbar 105
 partielle Ableitung 105
 partielle Integration 88
 Partition 8
 Polardarstellung 74
 Polynom 29
 positiv 10
 positiv definit (Betrag) 98
 positiv definit (Matrix) 117
 Potenz 20
 Potenzreihe 61, 73
 Produkt 8, 20
 Produktregel, 88
 Produktreihe 38
 Punkte 98
 punktweise Konvergenz 50

Quadratwurzel 45
 Quotientenkriterium 36

Rand 101
 Randpunkt 101
 rationale Funktion 29
 rationale Zahlen 6
 Rauminhalt 125
 Realteil 72
 Rechenoperationen 6
 rechtsseitig stetig 48
 rechtsseitiger Grenzwert 47
 reelle Achse 72
 reelle Zahlen 6
 Reflexivität 8
 Reihe 33
 Reihenwert 33
 rein-imaginär 72
 rekursive Definitionen 20
 relativ offen 42, 101
 relatives Maximum 117
 relatives Minimum 117
 Reste 34
 Reziprokenregel 56
 Richtungsableitung 105
 Riemannsche Summe 83, 123
 Riemannsches Integral 82, 124
 Ring 9
 Rolle (Satz von) 58
 Rotation 134
 rotationsfrei 136

Schnitt (bzgl. Koordinate) 126
 Schranke 10, 82
 semidefinit (Matrix) 117
 Sinus 68
 Skalarmultiplikation 98
 Skalierung 125
 Spiegelbild 72
 stückweise 112
 Stammfunktion 86
 sternförmig 113
 stetig 41, 102
 stetig differenzierbar 110
 stetig ergänzbar 46
 stetiges Vektorfeld 134
 streben 26, 99, 102
 streng abnehmend 45
 streng monoton 45
 streng wachsend 45
 strenges Extremum 117
 Substitution 21
 Substitutionsregel 88, 131
 Summationsgrenzen 21
 Summationsindex 20
 Summe 8, 20, 33
 Superposition 81
 Supremum 11
 Supremumsprinzip 11
 surjektiv 24
 Symmetrie (Relation) 8
 Symmetrie (Metrik) 99

Tangens 70
 Tangente 55
 Tangentenproblem 79
 Tangentenvektor 112
 Taylorpolynom 66
 Taylorreihe 67, 114
 Taylor (Satz von) 65, 115
 Teilfolge 27
 Teilmenge 6
 Teilungsintervalle 123
 Teilungspunkte 79
 topologische Begriffe 98
 total differenzierbar 106
 Träger 112
 Transitivität 8, 10
 Translationsinvarianz 125
 Trennungszahl 18
 Trichotomie 9
 trigonometrische Funktionen 68
 Tupel 23

$U_\varepsilon(x)$ 26, 100
 Überdeckungseigenschaft 44
 Umkehrabbildung 25
 umkehrbar eindeutig 24
 Umkehrfunktion 25
 Umkehrregel 109
 Umkehrsatz 45
 Umordnung 38
 Umparametrisierung 112
 unbedingt konvergent 38
 unbestimmtes Integral 86
 uneigentlicher Grenzwert 32
 uneigentliches Integral 91, 92
 unendlich viele 30
 unendliche Intervalle 10
 unendliche Reihe 33
 Unterintegral 81, 123
 Unterklasse 18
 Untersumme 79, 123
 Untersummen 123
 Urbild 23, 24

Variable 23
 Vektorprodukt 134
 Vektorraum 98
 Veränderliche 23
 Verbindungsstrecke 112
 Vereinigung 6
 Verfeinerung 79
 Vertauschung der Grenzwertbildung 60, 96
 vollständige Induktion 20
 Volumen 125

wachsend 45
 Wendepunkt 76
 Wert 23, 34, 48
 Widerspruch 15
 Wurzelfunktion 25

Zahlenebene 72
 Zahlenfolge 23

Zerlegung 79, 123
 Zielmenge 23
 zusammenhängend 113
 zweimal Differenzierbar 53
 zweite Ableitung 53
 Zwischenwertsatz 44