Analysis I für Physiker

nach den Vorlesungen von Prof. Dr. Werner Timmermann
(Wintersemester 2006/07)

Herausgegeben von

Jeffrey Kelling
Felix Lemke
Stefan Majewsky

Stand: 23. Oktober 2008

Inhaltsverzeichnis

Vorwort

Bevor Ihr beginnt, mit diesem Skript zu arbeiten, möchten wir vor allem den Studienanfängern eine wichtige Botschaft mitgeben: Dieses Skript ersetzt weder den Besuch der Vorlesung noch das selbstständige Nacharbeiten des Stoffes. Wer das nicht verstanden hat, bei dem kann die Benutzung des Skriptes im weiteren Studienverlauf für Probleme sorgen.

Wir weisen darauf hin, weil das Skript nicht als vorgekauter Wissensspeicher zu verstehen ist. Das hier ist eine Abschrift des Inhaltes, den die Vorlesung zu vermitteln versucht. Nicht enthalten sind zum Beispiel mündliche Kommentare des Professoren, auch wenn diese im individuellen Falle oft erst den Groschen fallen lassen.

Nochmals möchten wir uns an die Studienanfänger wenden: Hoffentlich begreift Ihr genauso wie wir die ersten Semester Eures Studiums als Chance, Euren persönlichen Lernfluss zu finden. Das Studium verlangt in dieser Beziehung nach eigener Erfahrung wesentlich mehr von einem Studenten ab als eine Schule. Umso wichtiger ist es, herauszufinden, wie man möglichst viel Inhalt in möglichst kurzer Zeit aufnehmen, systematisieren und abspeichern kann.

Eines ist sicher: Das bloße Durchlesen des Skriptes reicht meistens nicht. Zur Aufnahme des Stoffes und für einen ersten Überblick reicht in den meisten Fällen das klassische Mitschreiben in der Vorlesung, bei wenigen auch das aufmerksame Zuhören. Zum Systematisieren ist der Besuch der Übungen und das Bearbeiten der Übungsaufgaben unerlässlich (auch wenn einige Erstsemester das immer wieder mal anders sehen, aber das gibt sich). Eine andere gute Idee ist oft auch die Bildung von eigenen kleinen Lehrgruppen von drei bis vier Personen, auch wenn man da aufpassen muss, sich nicht abzulenken.

Bevor wir abschweifen, fassen wir noch kurz zusammen, wofür dieses Skript wirklich geeignet ist: einfach gesagt als Wissensstütze, also zum Beispiel zum schnellen Nachschlagen; außerdem zum Wiederholen früheren Stoffes, sofern ein ausreichendes Grundverständnis vorhanden ist. Nach diesen einleitenden Worten wünschen wir Euch viel Spaß bei der Arbeit mit diesem Skript und viel Erfolg beim Studium!

Die AGeS-Redaktion
`www.ages-skripte.org`

P.S. Wir suchen immer Helfer, die unsere Skripte um neue Inhalte erweitern, Fehler suchen, oder das Layout ansprechender gestalten wollen. Wenn Ihr Lust habt, meldet Euch über unsere Webseite.

1 Mengen, Abbildungen

1.1 Einführende Bemerkungen

Verwende Großbuchstaben zur Bezeichnung von Mengen $(A, B, \ldots)$ sowie folgende Schreibweisen für Mengen:

- Beschreibung durch Elementliste: $M = \{1, 2, 3\}$
- Beschreibung durch Elementeigenschaft: $M = \{x : \text{ Eigenschaft } \}$, z.B. $M = \{x : x \text{ durch 3 teilbar } \}$
- leere Menge (enthält keine Elemente): $\varnothing = \{\}$

Bezeichnungen für spezielle Zahlenmengen

- natürliche Zahlen: $\mathbb{N} = \{1, 2, 3, \ldots\}$
- natürliche Zahlen mit Null: $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$
- ganze Zahlen: $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- rationale Zahlen: $\mathbb{Q} = \left\{\frac{x}{y} : x, y \in \mathbb{Z}\right\}$
- reelle Zahlen: $\mathbb{R}$ (siehe Kapitel 2)
- komplexe Zahlen: $\mathbb{C}$ (siehe Lineare Algebra, Kapitel 1.1)

Für Binomialkoeffizienten, vollständige Induktion und Summationszeichen siehe die 1. Übung.

1.2 Mengenoperationen und -beziehungen

- ***Vereinigung***:
  Für zwei Mengen A und B ist
  $$A \cup B := \{x : x \in A \vee x \in B\}$$
  Beispiel: $A = \{1, 2, 3\}, B = \{2, 4, 7\} \Rightarrow A \cup B = \{1, 2, 3, 4, 7\}$
  Verallgemeinerung: Seien $A_1, \ldots, A_n$ Mengen, dann ist
  $$\bigcup_{i=1}^{n} A_i := \{x : \exists i : 1 \leq i \leq n \wedge x \in A_i\}$$
- ***Durchschnitt***:
  Für zwei Mengen A und B ist
  $$A \cap B := \{x : x \in A \wedge x \in B\}$$
  Beispiel: $A = \{1, 2, 3\}, B = \{2, 4, 7\} \Rightarrow A \cap B = \{2\}$
  Verallgemeinerung: Seien $A_1, \ldots, A_n$ Mengen, dann ist
  $$\bigcap_{i=1}^{n} A_i := \{x : \forall i : 1 \leq i \leq n \wedge x \in A_i\}$$

- ***Disjunktheit***:
Zwei Mengen A und B heißen **disjunkt**, wenn gilt: $A \cap B = \varnothing$.
Mengen $A_1, \ldots, A_n$ heißen **paarweise disjunkt**, wenn gilt:

$$\forall k, l : 1 \leq k, l \leq n \wedge k \neq l : A_k \cap A_l = \varnothing$$

- ***Differenz***:
Für zwei Mengen A und B ist
$$A \setminus B = \{x \in A : x \notin B\}$$

- ***Inklusion von Mengen***:
Eine Menge A ist eine Teilmenge einer Menge B, wenn gilt: $x \in A \Rightarrow x \in B$. Man schreibt $A \subset B$. Also ist
$$A \subset B \Leftrightarrow \forall x \in A : x \in B$$
Eigenschaften der Inklusionsbeziehung:
  1. Reflexivität: $A \subset A$
  2. Antisymmetrie: $A \subset B \wedge B \subset A \Leftrightarrow A = B$
  3. Transitivität: $A \subset B \wedge B \subset C \Rightarrow A \subset C$

  Um $A = B$ zu zeigen, beweist man oft $A \subset B$ und $B \subset A$.

- ***Komplementbildung***:
Seien M und $A \subset M$ Mengen. Dann ist das Komplement von A bzgl. M definiert als

$$C_M A = \{x \in M : x \notin A\} = M \setminus A$$

  Belegen Sie zur Übung die folgenden Eigenschaften:
  1. $(C_M A) \cup A = M$
  2. $(C_M A) \cap A = \varnothing$
  3. $C_M\left(C_M A\right) = A$, insbesondere $C_M \varnothing = M$ und $C_M M = \varnothing$

  Wird CA ohne Index geschrieben, so ist $C_X A$ gemeint, wobei X die Menge aller möglichen Elemente ist.

1.2.1 Einige Kombinationen dieser Operationen

- ***Distributivgesetze***:
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
Allgemeine Formulierung für endlich viele Mengen:
$$\left(\bigcup_{i=1}^{n} A_i\right) \cap B = \bigcup_{i=1}^{n} (A_i \cap B)$$
$$\left(\bigcap_{i=1}^{n} A_i\right) \cup B = \bigcap_{i=1}^{n} (A_i \cup B)$$

- ***De Morgansche Regeln***:
$$C_M\left(\bigcup_{i=1}^{n} A_i\right) = \bigcap_{i=1}^{n} (C_M A_i)$$
$$C_M\left(\bigcap_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n} (C_M A_i)$$

- ***Mengenprodukt***:
Seien A und B nichtleere Mengen ($A \neq \varnothing \wedge B \neq \varnothing$). Unter dem **kartesischen Produkt** von A und B versteht man die Menge
$$A \times B := \{(a,b) : a \in A \wedge b \in B\}$$
$A \times B$ ist die Menge der geordneten Paare von Elementen aus A und B.
Allgemeiner: $A_1, \dots, A_n$ seien nichtleere Mengen, dann hat man zwei gleichwertige Möglichkeiten, $A_1 \times \dots \times A_n$ zu bilden.
  1. rekursive Definition: $A_1 \times \dots \times A_n := (A_1 \times \dots \times A_{n-1}) \times A_n$
  2. $A_1 \times \dots \times A_n := \{(a_1, \dots, a_n) : a_1 \in A_1 \wedge \dots \wedge a_n \in A_n\}$
  Dabei gilt: $(a_1, \dots, a_n) = (b_1, \dots, b_n) \Leftrightarrow a_1 = b_1 \wedge \dots \wedge a_n = b_n$. Also ist $A_1 \times \dots \times A_n$ die Menge aller geordneten n-tupel mit Elementen aus $A_1, \dots, A_n$.

Als Beispiel zeigen wir die allgemeine Form des ersten Distributivgesetzes.

Beweis
Wie oben angedeutet zeigen wir „linker Term $\subset$ rechter Term“ und „linker Term $\supset$ rechter Term“, woraus folgt: „linker Term = rechter Term“.

$$\begin{aligned}
& x \in \left(\bigcup_{i=1}^{n} A_i\right) \cap B \\
\Rightarrow\ & x \in B \wedge x \in \left(\bigcup_{i=1}^{n} A_i\right) \\
\Rightarrow\ & \exists k, 1 \leq k \leq n : x \in B \wedge x \in A_k \\
\Rightarrow\ & \exists k, 1 \leq k \leq n : x \in (B \cap A_k) \\
\Rightarrow\ & x \in \bigcup_{i=1}^{n} (A_i \cap B)
\end{aligned}$$

$$\begin{aligned}
& x \in \bigcup_{i=1}^{n} (A_i \cap B) \\
\Rightarrow\ & \exists j, 1 \leq j \leq n : x \in (B \cap A_j) \\
\Rightarrow\ & \exists j, 1 \leq j \leq n : x \in B \wedge x \in A_j \\
\Rightarrow\ & x \in B \wedge x \in \left(\bigcup_{i=1}^{n} A_i\right) \\
\Rightarrow\ & x \in \left(\bigcup_{i=1}^{n} A_i\right) \cap B
\end{aligned}$$

■

1.2.2 Mengenproduktdefinitionen

Problem: Beide Definitionen des Mengenproduktes $A_1 \times A_2 \times A_3$ scheinen inäquivalent.
$$A_1 \times A_2 \times A_3 = \{(a_1, a_2, a_3) : a_i \in A_i\}$$
$$A_1 \times A_2 \times A_3 = (A_1 \times A_2) \times A_3 = \{((a_1, a_2), a_3) : a_i \in A_i\}$$
Man definiert deshalb, dass Paare in geordneten Paaren aufgelöst werden können, also etwa:
$$((a_1, a_2), a_3) := (a_1, a_2, a_3)$$

Beispiel 1.1

$A = \{1,2\}, B = \{1,3\} \Rightarrow A \times B = \{(1,1), (1,3), (2,1), (2,3)\}$

Achtung: Man unterscheide genau zwischen dem geordneten Tupel $(a_1, \ldots, a_n)$ und der Menge $\{a_1, \ldots, a_n\}$. Zum Beispiel ist $(1,1)$ sinnvoll, $\{1,1\}$ aber sinnlos. Beachte außerdem, dass das Mengenprodukt nicht kommutativ ist:

$$(a,b) \neq (b,a) \Rightarrow A \times B \neq B \times A \,(\text{aber}\, \{a,b\} = \{b,a\})$$

- Wenn $A_i = A \quad \forall i \in \{1, \ldots, n\}$, dann ist $A \times \ldots \times A = A^n$.
- Sei M beliebig. Unter der **Potenzmenge** $P(M)$ von M versteht man das System aller Teilmengen von M.
$$P(M) = \{A \subset M\}$$
- Hat A n Elemente, so hat $P(A)$ 2^n Elemente.

Beispiel 1.2

$A = \{1,2,3\} \Rightarrow P(A) = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\} \Rightarrow P(A)$ hat 8 Elemente.

1.3 Relationen und Abbildungen

Wir behandeln zwei wichtige Relationen und den Abbildungsbegriff.

Definition 1.1

Sei A eine nichtleere Menge. Eine **Relation** auf oder in A ist eine Teilmenge $R \subset A \times A$.

Bemerkung

$(x,y) \in R$ schreibt man als xRy und sagt „x steht in Relation zu y“ (meist suggestivere Schreibweise).

Beispiel 1.3 Ordnungsrelationen

1. Sei $A = \mathbb{N}$ (oder $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$). $xRy \Leftrightarrow x \leq y$
2. Sei M eine feste nichtleere Menge und $A = P(M)$. Für $B, C \in P(M)$ ist $BRC \Leftrightarrow B \subset C$.

   Unterschied: Zwei Zahlen stehen immer in Relation zueinander, zwei Mengen nicht unbedingt.

Definition 1.2 Ordnungsrelation

Eine Relation $\leq$ auf A heißt **Ordnungsrelation** oder einfach **Ordnung** (manchmal auch Halbordnung), wenn sie folgende Eigenschaften besitzt:

1. Reflexivität: $\forall x \in A : (x,x) \in R$, d.h. $x \leq x$
2. Antisymmetrie: $x \leq y \wedge y \leq x \Leftrightarrow x = y$
3. Transitivität: $x \leq y \wedge y \leq z \Leftrightarrow x \leq z$

Eine Ordnung heißt **linear** oder **total**, wenn gilt: $\forall x, y \in A : x \leq y \;\vee\; y \leq x$

Im obigen Beispiel ist die 1. Ordnung linear, die zweite jedoch nicht.

Beispiel 1.4 Äquivalenzrelation

In $A = \mathbb{N}_0$ sei eine mit $\sim$ bezeichnete Relation gegeben durch

$$a \sim b \;\Leftrightarrow\; a \equiv b \mod 3$$

Die rechte Seite bedeutet: a und b lassen bei Division durch 3 den gleichen Rest (a ist kongruent zu b zum Modul 3). Zum Beispiel: $0 \sim 3$ (lies: „0 ist äquivalent zu 3“), $1 \sim 10$, $1 \sim 13$ und $13 \sim 1$, aber: $1 \nsim 8$.

Definition 1.3

Eine Relation $\sim$ auf A heißt **Äquivalenzrelation**, wenn sie folgende Eigenschaften besitzt:

1. Reflexivität: $\forall x \in A : x \sim x$
2. Symmetrie: $x \sim y \Leftrightarrow y \sim x$
3. Transitivität: $x \sim y \wedge y \sim z \Leftrightarrow x \sim z$

Sei A fest, dann nennt man $[x] := \{y \in A : x \sim y\}$ die zu x gehörende **Äquivalenzklasse**. Jedes $y \in [x]$ heißt **Repräsentant** der Äquivalenzklasse.

Beispiel 1.5

Die im Beispiel 1.4 definierte Relation hat drei Äquivalenzklassen:

$$[0] = \{0, 3, 6, 9, 12, \ldots\}$$

$$[1] = \{1, 4, 7, 10, 13, \ldots\}$$

$$[2] = \{2, 5, 8, 11, 14, \ldots\}$$

1.3.1 Abbildungen und Funktionen

Wir verwenden zuerst einen äußerst pragmatischen Begriff.

Definition 1.4 Abbildung

Seien $X, Y \neq \varnothing$. Unter einer **Abbildung** f *aus* X *in* Y versteht man eine Vorschrift, die jedem $x \in D(f) \subset X$ genau ein $y \in Y$ zuordnet.

Bemerkung

Notation: $D(f)$ und $W(f) = \{y = f(x) \in Y : x \in D(f)\}$ sind Definitionsbereich und Wertebereich von f. Man schreibt für die Abbildung $f : X \to Y$ und für die Abbildungsvorschrift $f : x \mapsto y$. (Wenn nicht explizit notiert, ist $D(f) = X$.)

Unterscheide streng zwischen der Abbildung f und dem Bild $f(x)$ von x. Der Begriff „Abbildung“ wird oft als Synonym zum Begriff „Funktion“ verwendet, wir wollen beide Begriffe allerdings unterscheiden. (Der Begriff **Funktion** wird oft nur für Abbildungen auf Zahlenmengen verwendet.)

Definition 1.5 Eigenschaften von Abbildungen

Eine Abbildung f heißt:

- **injektiv**, wenn $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ („Jedes $y \in Y$ ist *höchstens* ein Bild.“)
- **surjektiv**, wenn $W(f) = Y$ („Jedes $y \in Y$ ist *mindestens* ein Bild.“)
- **bijektiv**, wenn f injektiv und surjektiv ist. („Jedes $y \in Y$ ist *genau* ein Bild.“)

Definition 1.6 Graph einer Abbildung

Unter dem **Graphen** $G(f)$ versteht man die Menge
$G(f) = \{(x, f(x)) \in X \times Y : x \in D(f)\} \subset X \times Y$.

Also zeichnet man nie eine Funktion, sondern immer den Graphen einer Funktion. Wir verwenden den Graphen für eine „exaktere“ Definition des Abbildungsbegriffes.

Definition 1.7 Abbildung

Eine Abbildung aus X in Y ist eine Teilmenge $f \subset X \times Y$ mit der Eigenschaft:
$(x, y_1), (x, y_2) \in f \Rightarrow y_1 = y_2$ (Jedem $x \in X$ ist höchstens ein $y \in Y$ zugeordnet.)

Bemerkung

$f : x \mapsto y \Leftrightarrow f(x) = y \Leftrightarrow (x, y) \in f$
$D(f) = \{x \in X : \exists y \in Y : (x, y) \in f\}$
$W(f) = \{y \in Y : \exists x \in X : (x, y) \in f\}$
$\Rightarrow G(f) = f$

Definition 1.8

Unter der **identischen Abbildung** auf X versteht man die Abbildung $id_X : X \to X$ mit $id_X(x) = x \ \ \forall x \in X$.

Definition 1.9

Seien $f : X \to Y$ eine Abbildung und $A \in X$ (bzw. $A \in D(f)$) eine nichtleere Menge.
Dann versteht man unter der **Einschränkung** von f auf A:
$f|_A : A \to Y, \quad f|_A : x \in A \mapsto f(x)$

Beispiel 1.6

Sei $A \subset D(f)$, dann heißt $f(A) := \{f(x) : x \in A\}$ **Wertebereich** von $f|_A$.
Sei $B \subset Y$, dann heißt $f^{-1}(B) := \{x \in X : f(x) \in B\}$ **Urbild** von B bzgl. f.

f^{-1} selbst wird hier nicht definiert. Man sollte nicht sofort $B \subset W(f)$ vorauszusetzen.

Beispiel 1.7

$X = \mathbb{R}^2, \quad f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x$
f^{-1} existiert nicht, da f nicht injektiv ist. Trotzdem ist $f^{-1}(B)$ definiert.

Definition 1.10 Verkettung

Sei $g : X \to Y$, $f : Y \to Z$. Dann versteht man unter der **Verkettung** $h := f \circ g$ die Abbildung

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

Definition 1.11 Umkehrabbildung

Sei $f : X \to Y$ injektiv, dann wird die **Umkehrabbildung** oder **inverse Abbildung** f^{-1} wie folgt definiert: $D\left(f^{-1}\right) = W(f)$ und $\forall y \in D\left(f^{-1}\right) : f^{-1}(y) = x$ mit $f(x) = y$
Also gilt $\forall x \in D(f) : f^{-1}(f(x)) = x$ bzw. $f^{-1} \circ f = id_{D(f)}$.

Beachte: $G(f) = \{(x, f(x)) : x \in D(f)\}$, aber $G(f^{-1}) = \{(f(x), x) : x \in D(f)\}$.
Verwechsle auf keinen Fall f^{-1} und $\frac{1}{f}$. Letzteres ist nicht definiert.

1.3.2 Rechnen mit Funktionen

Seien $f, g : X \to Y$ Abbildungen. Zu erklären sind: $f + g$, $\lambda f \ (\lambda \in \mathbb{R}, \mathbb{C})$, $f \cdot g$, $\frac{f}{g}$.
Wenn diese Operationen einen Sinn haben sollen, dann sollen sie punktweise erklärt werden.

$$(f + g)(x) := f(x) + g(x) \quad \text{u.s.w.}$$

Daraus folgt, dass die genannten Operationen in Y definiert sein müssen.

Spezialfall: Sei $Y \subset \mathbb{R}$ oder $Y \subset \mathbb{C}$. Dann sind alle Operationen definiert, wobei $\frac{f}{g}$ definiert ist $\forall x \in X: \ g(x) \neq 0$.

Beispiel 1.8 Spezialfälle

- $f : \mathbb{R}^n \to \mathbb{R}$ - Funktion von n reellen Variablen
- $f : \mathbb{R} \to \mathbb{R}^m$
- Folgen

Definition 1.12 Folge

Eine Abbildung f von $\mathbb{N}$ (oder $\mathbb{N}_0$, $\mathbb{Z}$) in Y heißt **Folge** aus Y (oder in Y oder mit Elementen aus Y).

Bemerkung

Notation: $f(1) = y_1, f(2) = y_2, \ldots; f = (y_n)$. Also bezeichnet (y_n) die Folge und y_n das n-te Glied der Folge. Die Gleichheit wird elementweise definiert:

$$(y_n) = (z_n) \Leftrightarrow y_i = z_i \ \forall i \in \mathbb{N} \, (\mathbb{N}_0, \mathbb{Z})$$

Auch mit Folgen lassen sich zu den gleichen Regeln wie oben Rechenoperationen durchführen.

Unterscheide bei Folgen streng die Folge (y_n) und den Wertebereich der Folge $\{y_n : n \in \mathbb{N}\}$. Diese Menge nennt man auch die **der Folge unterliegende Menge** und schreibt kurz $\{y_n\}$.

Beispiel 1.9

$(y_n) = (0, 1, 0, 1, 0, \ldots) \Rightarrow \{y_n\} = \{0, 1\}$

2 Die reellen Zahlen

Bis auf Modifikationen gibt es zwei Wege, um die Zahlenbereiche $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ einzuführen.

1. axiomatische Beschreibung von $\mathbb{N}$ / $\mathbb{N}_0$ ($\Rightarrow$ Peano-Axiome) und darauf basierend Konstruktion der anderen Zahlenbereiche
2. axiomatische Charakterisierung von $\mathbb{R}$, darin $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ wiederfinden

Wir nehmen einen stark vereinfachten Weg und gehen davon aus, dass $\mathbb{N}, \dots, \mathbb{R}$ bekannt sind.

$\mathbb{R}$ wird durch drei „Eigenschaften" (genauer Axiomgruppen) charakterisiert: die algebraische Struktur, die Ordnungsstruktur und die Vollständigkeit.

2.1 Die algebraische Struktur der reellen Zahlen

Kurzform: $\mathbb{R}$ **ist ein Körper.**

Das bedeutet, es existieren in $\mathbb{R}$ zwei algebraische Operationen, die Addition $(+)$ und die Multiplikation $(\cdot)$, mit den folgenden Eigenschaften (immer $\forall a, b, c \in \mathbb{R}$):

- (A1): $a + b = b + a$
- (A2): $a + (b + c) = (a + b) + c$
- (A3): $\exists n \in \mathbb{R} : a + n = n + a = a$
  n ist das **neutrale Element** der Addition. Statt n schreibt man 0.
- (A4): $\forall a \in \mathbb{R} : \exists x \in \mathbb{R} : a + x = x + a = 0$
  x ist das zu a **entgegengesetzte Element** bzgl. der Addition und wird mit $-a$ bezeichnet.

Bemerkung

Die Eigenschaften (A1) und (A2) werden **Kommutativität** und **Assoziativität** genannt.
Das neutrale Element sowie das zu einem bestimmten Element entgegengesetzten Element sind eindeutig bestimmt.
Betrachtet man $\mathbb{R}$ nur mit der Addition, so ist $\mathbb{R}$ aufgrund der Axiome (A2) bis (A4) eine Gruppe und, weil (A1) gilt, sogar eine kommutative Gruppe (Abelsche Gruppe).

- (M1): $a \cdot b = b \cdot a$
- (M2): $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- (M3): $\exists e \in \mathbb{R} : a \cdot e = e \cdot a = a$
  e ist das neutrale Element der Multiplikation. Forderung: $e \neq 0$. Statt e schreibt man 1.
- (M4): $\forall a \in \mathbb{R}, \mathbf{a \neq 0} : \exists y \in \mathbb{R} : a \cdot y = y \cdot a = 1$
  y ist das zu a entg.ges. Element bzgl. der Multiplikation und wird mit a^{-1}, $\frac{1}{a}$ bezeichnet.

Bemerkung

> Die Axiome (M1) bis (M4) sind die Analogien zu den Axiomen (A1) bis (A4) mit der Ausnahme, dass (M4) für $a = 0$ nicht gilt.
> $\mathbb{R} \setminus \{0\}$ ist bzgl. der Multiplikation eine kommutative Gruppe. Die 0 ist ausgeschlossen, da sie kein entgegengesetztes Element hat.
> Die merkwürdige Bedingung $n \neq e$ bzw. $1 \neq 0$ garantiert, dass der triviale Fall $(\{0\}, +, \cdot)$ ausgeschlossen wird. Ein Körper hat auf jeden Fall mindestens zwei Elemente.

Addition und Multiplikation werden durch das Distributivgesetz verknüpft:

- (D): $\forall a, b, c \in \mathbb{R} : a \cdot (b + c) = a \cdot b + a \cdot c$

Bemerkung

> Hat man eine Menge M mit zwei Operationen $+$ und $\cdot$, so heißt M ein **kommutativer Ring**, wenn gelten: (A1-4), (M1-*3*), (D).
> Zum Beispiel bildet die Menge $A^{n\times n}$ aller (n, n)-Matrizen mit den Operationen $+$ und $\cdot$ einen nicht-kommutativen Ring (M1 gilt nicht.)

(A1-4), (M1-4) und (D) beschreiben die algebraische Struktur von $\mathbb{R}$.

2.2 Die Ordnungsstruktur der reellen Zahlen

In $\mathbb{R}$ ist eine Ordnung $<$ mit folgenden Eigenschaften definiert:

- (O1): Für $a, b \in \mathbb{R}$ gilt genau eine der folgenden drei Beziehungen: $(a < b) \vee (a = b) \vee (b < a)$
- (O2): $a < b \wedge b < c \Rightarrow a < c$
- (O3): $\forall c \in \mathbb{R} : a < b \Rightarrow a + c < b + c$
- (O4): $\forall c \in \mathbb{R}, 0 < c : a < b \Rightarrow a \cdot c < b \cdot c$

Bemerkung

> Die Eigenschaften (O1) und (O2) heißen **Trichotomie** und **Transitivität**, (O3) und (O4) beschreiben die **Monotonie**. Man benutzt $a < b$ und $b > a$ äquivalent. Führt man zudem das Symbol $a \leq b$ mit $a \leq b \Leftrightarrow a < b \vee a = b$, dann ist $\mathbb{R}$ mit der Relation $\leq$ eine total geordnete Menge.
> **Sprechweisen:**
> - $a > 0$: „a ist positiv.“
> - $a \geq 0$: „a ist nicht-negativ.“
> - $a < 0$: „a ist negativ.“
> - $a \leq 0$: „a ist nicht-positiv.“

Die bisherigen Eigenschaften (A), (M), (D), (O) besagen: $\mathbb{R}$ ist ein **angeordneter Körper**.

2.3 Die Vollständigkeit der reellen Zahlen

Bis jetzt erfüllt auch $\mathbb{Q}$ alle bisherigen Axiome. Aber: Es existiert kein $x \in \mathbb{Q}$ mit $x^2 = 2$. Historisch hat es lange gedauert, die Vollständigkeit exakt zu formulieren. Es existieren mehrere äquivalente Möglichkeiten, das zu formulieren. Wir benötigen nun einige Begriffe.

Definition 2.1 Beschränktheit, Schranken

Eine Menge $M \subset \mathbb{R}$ heißt

- **nach oben beschränkt**, wenn $\exists b \in \mathbb{R} : \forall x \in M : x \leq b$
- **nach unten beschränkt**, wenn $\exists a \in \mathbb{R} : \forall x \in M : a \leq x$

Dann heißt a **untere Schranke** und b **obere Schranke** von M.
M heißt **beschränkt**, wenn M nach oben und nach unten beschränkt ist.

Bemerkung

Wenn a untere Schranke von M ist, dann auch jedes $a' < a$ (analog für obere Schranken).

Definition 2.2 Minima und Maxima, Grenzen

Sei $M \subset \mathbb{R}$. Man sagt, M besitzt

- ein **Minimum**, wenn $\exists c \in M : \forall x \in M : c \leq x$ (c heißt Minimum von M, $c = \min M$)
- ein **Maximum**, wenn $\exists C \in M : \forall x \in M : x \leq C$ (C heißt Maximum von M, $C = \max M$)
- ein **Infimum** (untere **Grenze**, größte untere Schranke), wenn eine untere Schranke $s \in \mathbb{R}$ von M existiert, sodass kein $s' > s$ existiert, das ebenfalls untere Schranke ist. ($s = \inf M$)
- ein **Supremum** (obere Grenze, kleinste obere Schranke), wenn eine obere Schranke $S \in \mathbb{R}$ von M existiert, sodass kein $S' < S$ existiert, das ebenfalls obere Schranke ist. ($S = \sup M$)

Bemerkung

Man muss, um die Bezeichnungen zu rechtfertigen, zeigen, dass jede Menge je höchstens ein Minimum, Maximum, Infimum und Supremum besitzt.

Satz 2.3

1. Wenn M ein Supremum besitzt, ist es eindeutig bestimmt.
2. $S = \sup M$
$\Leftrightarrow$ S ist obere Schranke von $M \wedge \forall S' < S : \exists x \in M : S' < x$
$\Leftrightarrow$ S ist obere Schranke von $M \wedge \forall \varepsilon > 0 : \exists x \in M : S - \varepsilon < x$

Beweis
Bei einer Wenn-Dann-Beziehung (Äquivalenz) müssen beide Richtungen gezeigt werden.

Teilsatz 1: Angenommen, S_1 und S_2 seien Suprema von M und es gelte $S_1 \neq S_2$ (o.E.d.A. $S_1 < S_2$). Aus $S_2 = \sup M$ folgt: S_2 ist obere Schranke von M und es existiert keine kleinere obere Schranke von M. Da aber $S_1 < S_2$ und $S_1 = \sup M$, ist S_1 eine obere Schranke von M, die kleiner als S_2 ist. Also folgt aus $S_1 \neq S_2$ ein Widerspruch. Es muss $S_1 = S_2$ sein, was heißt, dass es nur ein Supremum $S = S_1 = S_2$ gibt.

Teilsatz 2: Offenbar sind beide Formulierungen äquivalent. ($S' = S - \varepsilon$ bzw. $S - S' = \varepsilon$).

„Hin"-Richtung: Sei S Supremum und $S' < S$. Also ist S' (nach Supremumdefinition) keine obere Schranke von M. Daraus folgt (nach Def. der oberen Schranke): $\exists x \in M : S' < x$

„Rück"-Richtung: S ist obere Schranke von M und es gilt $\forall S' < S : \exists x \in M : S' < x$. Daraus folgt, dass kein $S' < S$ obere Schranke von M ist. Also ist $S = \sup M$.

Analog kann man die Eindeutigkeit von Minima, Maxima, Infima, sowie das Äquivalent zum Satz 2.3 für Infima zeigen.

■

Bemerkung

Nicht jede nach oben beschränkte Menge rationaler Zahlen hat ein Maximum oder Supremum.
Nicht jede nach unten beschränkte Menge rationaler Zahlen hat ein Minimum oder Infimum.

Beispiel 2.1

$M_1 = \left\{x \in \mathbb{Q} : x^2 < 2\right\}$ hat in $\mathbb{Q}$ weder ein Maximum noch ein Supremum.
$M_2 = \left\{-\frac{1}{n} : n \in \mathbb{N}\right\}$ hat das Supremum 0, aber kein Maximum. ($\inf M_2 = \min M_2 = -1$)

2.3.1 Das Vollständigkeitsaxiom

- (V): Jede nichtleere, nach oben beschränkte Menge $M \subset \mathbb{R}$ besitzt ein Supremum.

Bemerkung

$\exists \min M \Rightarrow \min M = \inf M$
$\exists \max M \Rightarrow \max M = \sup M$

Satz 2.4

Sei $M \subset R$ und $-M := \{-x : x \in M\}$. Dann gilt:

1. $-M$ ist genau dann nach unten (oben) beschränkt, wenn M nach oben (unten) beschränkt ist.
2. $c = \min M \Leftrightarrow -c = \max(-M)$
3. $s = \sup M \Leftrightarrow -s = \inf(-M)$

Die Axiome (A), (M), (D), (O), (V) charakterisieren $\mathbb{R}$ eindeutig. Das heißt:

- Die Axiome sind in sich schlüssig und widerspruchsfrei.
- $\mathbb{R}$ ist bis auf Isomorphie eindeutig bestimmt.
- Es existiert tatsächlich eine Menge, die alle Axiome erfüllt.

2.4 Folgerungen aus der Struktur der reellen Zahlen

Wir zeigen exemplarisch zwei Folgerungen aus der **algebraischen** Struktur:

Zu jedem $a \in \mathbb{R}$ existiert genau ein $b \in \mathbb{R}$ mit $a + b = 0$. ($b = -a$)

Beweis
Angenommen, es existieren $x, y \in \mathbb{R}$ mit $a + x = 0 \wedge a + y = 0$. Zu zeigen ist: $x = y$.
$x \overset{(A3)}{=} x + 0 \overset{(a+y=0)}{=} x + a + y \overset{(A1)}{=} a + x + y \overset{(a+x=0)}{=} 0 + y \overset{(A3)}{=} y$

■

$\forall a \in \mathbb{R} : 0 \cdot a = 0$

Beweis

$0 \cdot a \overset{(A3)}{=} (0+0) \cdot a \overset{(D)}{=} 0 \cdot a + 0 \cdot a \qquad (*)$
$0 \overset{(A3)}{=} 0 \cdot a + (-0 \cdot a) \overset{(*)}{=} 0 \cdot a + (0 \cdot a + (-0 \cdot a)) \overset{(A3)}{=} 0 \cdot a + 0 \overset{(A3)}{=} 0 \cdot a$

■

2.4.1 Folgerungen aus den Ordnungsaxiomen

1. $a > 0 \Leftrightarrow -a < 0$, $a < 0 \Leftrightarrow -a > 0$
2. $a \neq 0 \Leftrightarrow a^2 > 0$ (insbesondere: $1^2 = 1 > 0 \overset{(O3)}{\Rightarrow} 2 > 1$)
3. $a > 0 \Leftrightarrow \frac{1}{a} > 0$, $a < 0 \Leftrightarrow \frac{1}{a} < 0$
4. $a < b \wedge c < 0 \Rightarrow ac > bc$ (insbesondere: $a < b \Rightarrow -b < -a$)
5. $a < b < c < d \Leftrightarrow a + c < b + d$

Beweis

Zu 1.: $a > 0 \overset{(O3)}{\Rightarrow} a + (-a) > 0 + (-a) \overset{(A3)}{\Rightarrow} 0 > -a \Rightarrow -a < 0$

■

Beweis
Zu 2.: Fallunterscheidung:
$a > 0 \Rightarrow a \cdot a > 0 \cdot a \Rightarrow a^2 > 0$ - Beweisende
$a < 0 \overset{(s.o.)}{\Rightarrow} (-a) > 0 \overset{(s.o.)}{\Rightarrow} (-a)^2 > 0 \qquad (**)$
Man zeigt $-a = (-1) \cdot a$ und $(-1)^2 = 1$.
$(**) \Rightarrow (-1)^2 \cdot (-a)^2 > 0 \Rightarrow 1 \cdot a^2 > 0 \Rightarrow a^2 > 0$

Zu 3.: Für $a > 0$: $1 > 0 \Rightarrow a \cdot a^{-1} > a \cdot 0 \Rightarrow a \cdot a^{-1} \cdot a^{-1} > a \cdot a^{-1} \cdot 0 \Rightarrow a^{-1} > 0$

■

Lemma 2.5 Bernoulli'sche Ungleichung

Für $x \in \mathbb{R}$, $x \geq 1$, $n \in \mathbb{N}$ gilt: $(1+x)^n \geq 1 + nx$

Beweis
Beweis durch vollständige Induktion:
Induktionsanfang: $n = 1 \Rightarrow 1 + x \geq 1 + x \Rightarrow$ wahr
Induktionsschritt: Gelte 2.5 für $n = m$ (Induktionsvoraussetzung). Für $n = m+1$ gilt: $(1+x)^{m+1} = (1+x)(1+x)^m \geq (1+x)(1+nx) = 1 + nx + x + \underbrace{nx^2}_{\geq 0} \geq 1 + (n+1)x$

■

Definition 2.6 absoluter Betrag

Sei $a \in \mathbb{R}$. Setze $|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$

Lemma 2.7 Der absolute Betrag hat folgende Eigenschaften:

1. $|a| \geq 0$, $|a| = 0 \Leftrightarrow a = 0$ (Definitheit)
2. $|-a| = |a|$
3. $a \leq |a|$
4. $|ab| = |a|\,|b|$ (Multiplikativität)
5. $|a+b| \leq |a| + |b|$ (Dreiecksungleichung)
6. $||a| - |b|| \leq |a-b|$
7. $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$
8. $|a| < c \Leftrightarrow -c < a < c$
9. $|a-b| < c \Leftrightarrow b - c < a < b + c$

8. und 9. gelten analog für $\leq$.

Beweis

2. $|-a| = \begin{Bmatrix} -a & \text{falls } -a \geq 0 \\ a & \text{falls } -a < 0 \end{Bmatrix} = \begin{Bmatrix} -a & \text{falls } a \leq 0 \\ a & \text{falls } a > 0 \end{Bmatrix} = |a|$
5. Aus 2. und 3. folgt: $a \leq |a| \wedge b \leq |b| \Rightarrow a + b \leq |a| + |b|$
   $-a \leq |-a| = |a| \Rightarrow -a \leq |a|$ und analog $-b \leq |b| \Rightarrow -a - b \leq |a| + |b|$
   $|a+b| = \begin{Bmatrix} a+b & \text{falls } a+b \geq 0 \\ -a-b & \text{falls } a+b < 0 \end{Bmatrix} \leq |a| + |b|$
6. $a = (a-b) + b \overset{(5)}{\Rightarrow} |a| \leq |a-b| + |b| \Rightarrow \underline{|a| - |b| \leq |a-b|}$
   Analog ist $|b| - |a| \leq |b-a| = |a-b| \Rightarrow \underline{-(|a| - |b|) \leq |a-b|}$.
   Aus beiden Ergebnissen folgt die Behauptung.
8. $|a| < c \Leftrightarrow -c < \underbrace{-|a| < |a|}_{\text{einer der Werte ist a}} < c \Leftrightarrow -c < a < c$
9. folgt sofort aus 8.

■

Definition 2.8 Intervalle

Seien $a, b \in \mathbb{R}, a < b$. Dann heißt

- $[a,b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ <u>abgeschlossenes Intervall</u>
- $(a,b) := \{x \in \mathbb{R} : a < x < b\}$ <u>offenes Intervall</u>
- $[a,b) := \{x \in \mathbb{R} : a \leq x < b\}$ <u>rechtsseitig halboffenes Intervall</u>
- $(a,b] := \{x \in \mathbb{R} : a < x \leq b\}$ <u>linksseitig halboffenes Intervall</u>

mit den Endpunkten a und b.

Definition 2.9

$L = b - a$ ist die Länge jedes der obigen Intervalle.
Sei $\varepsilon > 0$, dann heißt $U_\varepsilon(a) = \{x \in \mathbb{R} : a - \varepsilon < x < a + \varepsilon\} = (a - \varepsilon, a + \varepsilon)$ die offene ε-Umgebung von a.

2.4.2 Metrische Struktur

Definition 2.10

Für beliebige $a, b \in \mathbb{R}$ nennt man $d(a, b) := |a - b|$ Abstand (bzw. Differenz) von a und b.

Lemma 2.11

Der so definierte Abstand hat folgende Eigenschaften:

- (ME1): $d(a, b) \geq 0$, $d(a, b) = 0 \Leftrightarrow a = b$ (Definitheit)
- (ME2): $d(a, b) = d(b, a)$ (Symmetrie)
- (ME3): $d(a, b) \leq d(a, c) + d(c, b)$ (Dreiecksungleichung)

Beweis
(ME1), (ME2) trivial. (ME3): $d(a, b) = |a - c + c - b| \leq |a - c| + |c - b| = d(a, c) + d(c, b)$

■

Diese Eigenschaften sind Ausgangspunkt für die Eigenschaften eines sehr wichtigen Begriffs:

Definition 2.12

Sei $M \neq \varnothing$ eine beliebige Menge und $d : M \times M \to \mathbb{R}$ eine beliebige Abbildung mit den Eigenschaften (ME1), (ME2) und (ME3). Dann nennt man d eine Metrik in M und (M, d) einen metrischen Raum.

Achtung! In M muss überhaupt keine Struktur vorhanden sein, um einen metrischen Raum zu definieren.

Die Metrik in $\mathbb{R}$ hat die zusätzliche Eigenschaft der Translationsinvarianz:

$$\forall a, b, c \in \mathbb{R} : d(a, b) = d(a + c, b + c)$$

2.4.3 Dichtheitsaussagen

Satz 2.13 Satz von Archimedes

Die Anordnung von $\mathbb{R}$ ist archimedisch, d.h. $\forall a \in \mathbb{R} : \exists n \in \mathbb{N} : a < n$
Mit anderen Worten: $\mathbb{N}$ ist als Teilmenge von $\mathbb{R}$ nicht nach oben beschränkt.

Beweis
Angenommen, $\mathbb{N}$ wäre nach oben beschränkt, d.h. $\exists S = \sup \mathbb{N}$ (Vollständigkeit!). Sei nun $S' = S - \frac{1}{2}$. Dann $\exists n \in \mathbb{N} : S' < n$. Also ist $S - \frac{1}{2} < n \Rightarrow S < n + \frac{1}{2} < n + 1 \in \mathbb{N} \Rightarrow$ Widerspruch

■

Der Satz von Archimedes gilt auch in $\mathbb{R}$. Aus diesem Satz folgt der:

Satz 2.14 Satz von Eudoxos

$\forall \varepsilon > 0 : \exists m \in \mathbb{N} : \frac{1}{m} < \varepsilon$ und erst recht $\forall n > m : \frac{1}{n} < \varepsilon$

Beweis
Angenommen, das wäre falsch. Also $\exists \varepsilon > 0$, für das es kein $m \in \mathbb{N}$ mit $\frac{1}{m} < \varepsilon$ gibt.
Also: $\exists \varepsilon > 0 : \forall m \in \mathbb{N} : \frac{1}{\varepsilon} \geq m \Rightarrow$ Widerspruch mit 2.10

■

Bemerkung

In $\mathbb{Q}$ bzw. $\mathbb{R}$ gilt: $a < b \Rightarrow a < \frac{a+b}{2} < b$. Das heißt, zwischen zwei reelen (rationalen) Zahlen liegt immer mindestens eine weitere reelle (rationale) Zahl. Der nächste Satz sagt aus, wie $\mathbb{Q}$ in $\mathbb{R}$ liegt.

Satz 2.15

$\mathbb{Q}$ liegt in $\mathbb{R}$ dicht. Das heißt: $\forall r \in \mathbb{R}, \varepsilon > 0 : \exists q \in \mathbb{Q} : d(q, r) < \varepsilon$

Beweis
Wähle zuerst $m \in \mathbb{N}$ mit $\frac{1}{m} < \varepsilon$ (siehe 2.12).
Nun reicht es, ein $q \in \mathbb{Q}$ zu finden mit $r - \frac{1}{m} < q < r + \frac{1}{m}$, denn $r - \varepsilon < r - \frac{1}{m} < q < r + \frac{1}{m} < r + \varepsilon$.
Sei $r > 0$ (der Fall $r < 0$ wird analog behandelt). Nach Satz 2.11 $\exists n \in \mathbb{N} : n > m$.
Die Menge aller n hat ein kleinstes Element, dieses sei n_0. Daraus folgt: $n_0 - 1 \leq m < n$. Setze $q := \frac{n_0}{m}$.
$q - \frac{1}{m} \leq r < q \Rightarrow q \leq r + \frac{1}{m} \land r - \frac{1}{m} < q \Rightarrow r - \frac{1}{m} < q \leq r + \frac{1}{m}$

■

Man kann noch mehr zeigen: In jedem Intervall (a, b) liegen sowohl rationale als auch irrationale Zahlen ($\{\text{irrationale Zahlen}\} = \mathbb{R} \setminus \mathbb{Q}$).

2.5 Endliche und unendliche Mengen

Definition 2.16

Zwei nichtleere Mengen A, B heißen gleichmächtig (A und B haben die gleiche Mächtigkeit),wenn es eine bijektive Abbildung $f : A \to B$ gibt.

Man schreibt $A \sim B$. Die Gleichmächtigkeit ist eine Äquivalenzrelation.

Definition 2.17

Eine Menge M heißt endlich, wenn $\exists n \in \mathbb{N} : M \sim \mathbb{N}_n := \{1, 2, \ldots, n\}$. Andernfalls heißt die Menge unendlich.
Die Menge heißt abzählbar (unendlich), wenn $M \sim \mathbb{N}$ ist, und überabzählbar, wenn M weder endlich noch abzählbar ist.

Die Äquivalenzklassen gleichmächtiger Mengen nennt man Kardinalzahlen.
Bei endlichen Mengen $M \sim \mathbb{N}_n$ ist n die Kardinalzahl von M. Man schreibt $n = \operatorname{card} M = |M|$.
Die leere Menge wird als endliche Menge betrachtet: $|\varnothing| := 0$.

Satz 2.18 1. Cantor'sches Diagonalverfahren

$\mathbb{Q}$ ist abzählbar, d.h. $\mathbb{Q} \sim \mathbb{N}$.

Beweis
Konstruktives Beweisverfahren: Man ordne die positiven rationalen Zahlen in dieses Schema:

$$\begin{matrix} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\ \frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \frac{2}{4} & \frac{2}{5} & \cdots \\ \frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \frac{3}{4} & \frac{3}{5} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{matrix} \quad = \quad \begin{matrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\ 2 & 1 & \frac{2}{3} & \frac{1}{2} & \frac{2}{5} & \cdots \\ 3 & \frac{3}{2} & 1 & \frac{3}{4} & \frac{3}{5} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{matrix}$$

Dieses Schema wird nach **irgendeinem** Schema abgezählt. Man lässt alle Zahlen, die schon einmal vorkamen, weg. (Wenn etwa $\frac{2}{1}$ gezählt wurde, wird $\frac{4}{2}$ nicht mehr gezählt.)
Eine mögliche Abzählung ergibt: $\left(1, 2, \frac{2}{2}, \frac{1}{2}, 3, \frac{3}{2}, \frac{3}{3}, \frac{2}{3}, \frac{1}{3}, 4, \frac{4}{2}, \ldots\right)$
Durch Auslassen doppelter Zahlen ergibt sich: $\left(1, 2, \frac{1}{2}, 3, \frac{3}{2}, \frac{2}{3}, \frac{1}{3}, 4, \ldots\right) =: (r_1, r_2, \ldots)$
Nun kann $\mathbb{Q}$ auf $\mathbb{N}$ wie folgt abgebildet werden: $(0, r_1, -r_1, r_2, -r_2, r_3, -r_3, \ldots) =: (f(1), f(2), f(3), \ldots)$
$\Rightarrow \mathbb{Q} \sim \mathbb{N} \Rightarrow \mathbb{Q}$ ist abzählbar.

■

Satz 2.19 2. Cantor'sches Diagonalverfahren

$\mathbb{R}$ ist überabzählbar.

Satz 2.20

Die Menge F aller Folgen aus Nullen und Einsen ist nicht abzählbar.

Bemerkung

Hat man diesen Satz gezeigt, dann hat man offenbar auch gezeigt, dass folgende Menge $M \subset \mathbb{R}$ nicht abzählbar ist.
$M = \{0, \underbrace{x_1 x_2 x_3 \ldots}_{\substack{\text{Dezimalzahl,} \\ \text{z.B. } 0{,}1101001\ldots}} \text{ mit } x_i \in \{0, 1\}\}$ Daraus folgt, dass $\mathbb{R}$ nicht abzählbar ist.

Beweis
Angenommen, F (die Menge aller Folgen mit Elementen aus $\{0, 1\}$) wäre abzählbar, d.h. F könnte durchnummeriert werden. Sei im folgenden **irgendeine** Abzählung von F gegeben:

$$\begin{matrix} \mathbf{a_1^{(1)}}, a_2^{(1)}, a_3^{(1)}, \ldots & \text{1. Folge} & \\ a_1^{(2)}, \mathbf{a_2^{(2)}}, a_3^{(2)}, \ldots & \text{2. Folge} & (*) \\ a_1^{(3)}, a_2^{(3)}, \mathbf{a_3^{(3)}}, \ldots & \text{3. Folge} & \end{matrix}$$

Hierbei ist $a_i^{(k)}$ das i-te Element der k-ten Folge. Jetzt müssen wir nur eine Folge $b = (b_1, b_2, \ldots)$ finden mit $b \in F$, aber b kommt in $(*)$ nicht vor. Daraus folgt ein Widerspruch, also ist F nicht abzählbar. Setze nun $b_n = 1$ für $a_n^{(n)} = 0$ und $b_n = 0$ für $a_n^{(n)} = 1$. Angenommen, (b_n) käme im Schema an der k-ten Stelle vor. Dann ist aber $b_k \neq a_k^{(k)} \Rightarrow$ Widerspruch

■

2.6 Potenzen mit rationalen Exponenten

Ziel: Für $a > 0, r \in \mathbb{Q}$ soll a^r erklärt werden.

Vorbemerkung: Für $n \in \mathbb{N}, m \in \mathbb{N}$ ist $a^n := \prod_{i=1}^{n} a$ und $a^{-m} = \frac{1}{a^m}$.

Sei $r = \frac{q}{p} \in \mathbb{Q}$. Man möchte setzen: $a^r = \left(a^{\frac{1}{p}}\right)^q$. Problem: Was ist $a^{\frac{1}{p}}$?

O.E.d.A. sei $p \in \mathbb{N}$ (Vorzeichen werden von q „abgefangen"). Wir zeigen die Existenz als Demonstrationsbeispiel für die Verwendung des Vollständigkeitsaxioms.

Definition und Satz 2.21

Sei $a > 0, p \in \mathbb{N}$, dann besitzt die Gleichung $x^p = a$ genau eine nichtnegative Lösung. Diese wird mit $\sqrt[p]{a}$ bezeichnet und p-te **Wurzel** aus a genannt. Im Falle $p = 2$ schreibt man statt $\sqrt[2]{a}$ kürzer $\sqrt{a}$.

Beweis
Mit $a < b \wedge c < d \Rightarrow ac < bd$ ergibt sich $x < y \Rightarrow xx < yy \Rightarrow x^3 < y^3 \Rightarrow \ldots \Rightarrow x^p < y^p \qquad (L)$

Eindeutigkeit:
Angenommen, x_1 und x_2 seien zwei unterschiedliche Lösungen von $x^p = a$, o.E.d.A. $x_1 < x_2$. Dann ist $x^p < y^p \Rightarrow a < y^p \Rightarrow$ Widerspruch: Wenn x Lösung ist, kann y nicht Lösung sein. $\Rightarrow$ *höchstens eine Lösung*

Existenz: Natürlich sind nur $p > 1$ interessant.
Betrachte $M := \{y \in \mathbb{R} : y \geq 0, y^p < a\}$. $M \neq \varnothing$, denn $0 \in M$. M ist nach oben beschränkt:
$(1+a)^p \overset{\text{(Bernoulli)}}{>} 1 + pa \overset{(p>1)}{>} a > y^p \;\Rightarrow\; \forall y \in M : (1+a)^p > y^p \overset{(L)}{\Rightarrow} 1 + a > y$
Also ist M durch $1 + a$ nach oben beschränkt. Also existiert nach (V) $\sup M =: z$. Wir zeigen $z^p = a$, indem wir $z^p < a$ und $z^p > a$ zum Widerspruch führen, woraus nach (O1) die Behauptung folgt.

Fall 1: $\mathbf{z^p < a}$
Finde m mit $\left(z + \frac{1}{m}\right)^p < a \Rightarrow z + \frac{1}{m} < y \Rightarrow$ Widerspruch zu $z = \sup M$. Benutze folgende Abschätzung:

$$\left(z + \tfrac{1}{n}\right)^p = z^p + \tfrac{1}{n}\binom{p}{1} z^{p-1} + \ldots + \tfrac{1}{n^p}\binom{p}{p} \leq (\text{ersetze alle } \tfrac{1}{n^i} \text{ durch } \tfrac{1}{n}) \leq z^p + \tfrac{1}{n}\underbrace{\left[\binom{p}{1} z^{p-1} + \ldots + \binom{p}{p}\right]}_{=:c,\text{ unabhängig von } n}$$

Also ist $= z^p + \frac{1}{n}c \qquad (*)$
$z^p < a\,[\Rightarrow a - z^p > 0] \wedge c > 0 \Rightarrow \frac{c}{a - z^p} > 0$
Nach Satz des Archimedes $\exists m \in \mathbb{N} : m > \frac{c}{a-z^p} \Rightarrow z^p + \frac{c}{m} < a$. Wende $(*)$ für $n = m$ an:
$\left(z + \frac{1}{m}\right)^p \leq z^p + \frac{1}{m}c < a \Leftrightarrow z + \frac{1}{m} \in M \Leftrightarrow$ Widerspruch

Fall 2: $\mathbf{z^p > a}$
Intuitiv $\exists n \in \mathbb{N} : \left(z - \frac{1}{n}\right)^p > a \Rightarrow z - \frac{1}{n} \notin M$
Nach Supremumsdefinition existiert ein $y \in M$: $z - \frac{1}{n} < y \Rightarrow$ Das wird zum Widerspruch geführt.
$z - \frac{1}{n} < y \left[\Rightarrow \left(z - \frac{1}{n}\right)^p < y^p\right] \wedge \left(z - \frac{1}{n}\right)^p > a \Leftrightarrow y^p > a \Leftrightarrow$ Widerspruch zu $y \in M$
Wir zeigen: $\exists z' < z : a < z'^p < z^p$. Das ist (s.o.) ein Widerspruch.
$\left(z - \frac{1}{n}\right)^p = z^p \left(1 - \frac{1}{nz}\right)^p \overset{\text{(Bernoulli)}}{>} z^p \left(1 - \frac{p}{nz}\right)$ $\left(\text{falls } z - \frac{1}{n} > 0 \left[\Rightarrow n > \frac{1}{z}\right]\right)$
Wähle außerdem $n = \frac{p \cdot z^p}{z \cdot (z^p - a)} \Rightarrow z^p - a > \frac{p}{nz} z^p \Rightarrow z^p \left(1 - \frac{p}{nz}\right) > a \Rightarrow \left(z - \frac{1}{n}\right)^p > a \Rightarrow$ Widerspruch

■

2.7 Die Symbole ∞ und $-\infty$

Es ist nützlich, zu $\mathbb{R}$ zwei Symbole hinzuzunehmen: $\infty\,(\equiv +\infty)$ und $-\infty$. Diese beiden Symbole sind *keine* reellen Zahlen. Wir vereinbaren außerdem folgende Rechenregeln:

- $\forall x$:
  - $-\infty < x < \infty$
  - $x + \infty = \infty + x = \infty$
  - $x - \infty = -\infty + x = -\infty$
  - $\frac{x}{\infty} = \frac{x}{-\infty} = 0$
- $\forall x > 0$:
  - $x \cdot \infty = \infty \cdot x = \infty$
  - $x \cdot (-\infty) = (-\infty) \cdot x = -\infty$
- $\forall x < 0$:
  - $x \cdot \infty = \infty \cdot x = -\infty$
  - $x \cdot (-\infty) = (-\infty) \cdot x = \infty$

Ausdrücke wie $\infty \cdot \infty$,$\infty - \infty$ oder $0 \cdot \infty$ werden nicht erklärt, da eine sinnvolle Definition nicht möglich ist. Sinnvollerweise vereinbart man aber:

- Wenn $M \subset \mathbb{R}$ nicht nach oben beschränkt ist, <u>setzt</u> man $\sup M := \infty$.
- Wenn $M \subset \mathbb{R}$ nicht nach unten beschränkt ist, <u>setzt</u> man $\inf M := -\infty$.

Das Vollständigkeitsaxiom wird davon nicht berührt, da ∞ und $-\infty$ keine reellen Zahlen sind.

3 Der Grenzwertbegriff

3.1 Zahlenfolgen

Im allgemeinen betrachten wir komplexe Zahlenfolgen $(a_n)\,, n \in \mathbb{N}$ oder $n = k, k+1, \ldots$.

Wiederholungen:

- $M \subset \mathbb{C}$ heißt beschränkt, wenn $\exists k > 0 : \forall z \in M : |z| \leq k$
- Jede endliche Teilmenge $M \subset \mathbb{C}$ (d.h. M hat endlich viele Elemente) ist beschränkt.
- (a_n) heißt beschränkt, wenn $\exists k > 0 : \forall n \in \mathbb{N} : |a_n| \leq k$
- $U_\varepsilon(z_0) = \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$ ist die ε-Umgebung um $z_0 \in \mathbb{C}$ (entspricht in der grafischen Darstellung einem Kreis um z_0 mit dem Radius ε)

Definition 3.1

Eine Zahl $a \in \mathbb{C}$ heißt Grenzwert (oder Limes) einer Folge (a_n), wenn gilt:
$\forall \varepsilon > 0 : \exists n_0 = n_0(\varepsilon) \in \mathbb{N} : \forall n \geq n_0 : |a_n - a| < \varepsilon$
Die Folge (a_n) heißt dann konvergent (gegen a). Man schreibt: $a = \lim\limits_{n\to\infty} a_n$ oder $a_n \to a$.
Eine Folge (a_n) mit $\lim\limits_{n\to\infty} a_n = 0$ heißt Nullfolge. Eine Folge, die nicht konvergiert, heißt divergent.

Bemerkung

Diskutieren Sie verschiedene Varianten und Formulierungen des Grenzwertbegriffes. Beispiele:

- $a_n \to a \Leftrightarrow a_n - a \to 0$ (Beweis!)
- $a_n \to a \Leftrightarrow$ In jeder ε-Umgebung von a liegen alle Folgenglieder bis auf endlich viele.

In der Definition ist es egal, ob man $n \geq n_0$ oder $n > n_0$ schreibt. Es kommt beim Nachweis der Konvergenz überhaupt nicht darauf an, das kleinste n_0 zu finden, solange man für jedes ε irgendein n_0 findet.

Beispiel 3.1

$a_n = \frac{1}{n} \to 0$, denn: Sei $\varepsilon > 0$ gegeben. Dann gilt $|a_n - a| = \left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon$, falls $n > n_0$, wobei n_0 jede beliebige natürliche Zahl $> \frac{1}{\varepsilon}$ sein kann, denn $\frac{1}{n_0} < \varepsilon \wedge n > n_0 \Rightarrow \frac{1}{n} < \frac{1}{n_0} < \varepsilon$.

Beispiel 3.2

$b_n = \left(17, 2^{22}, -2000, 0, 0, 0, \ldots\right) \to 0$ (offensichtlich)

Beispiel 3.3

$a_n = (1, 0, 1, 0, \ldots)$ ist divergent.
Angenommen, a wäre Grenzwert. Wenn $a \neq 0, 1$, dann $\exists \varepsilon > 0 : 0, 1 \notin U_\varepsilon(a)$. Also liegt in $U_\varepsilon(a)$ kein Folgenglied. Wenn $a = 0$ oder $a = 1$, wähle $\varepsilon = \frac{1}{2}$, dann liegen außerhalb von $U_\varepsilon(a)$ unendlich viele Folgenglieder.

Satz 3.2 Eindeutigkeitssatz

Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt.

Beweis
Angenommen, a und b mit $a \neq b$ wären Grenzwerte von (a_n). Dann existiert ein $\varepsilon > 0$ mit $U_\varepsilon(a) \cap U_\varepsilon(b) = \varnothing$. Weil a Grenzwert ist, liegen innerhalb von $U_\varepsilon(a)$ unendlich viele Folgenglieder, innerhalb von $U_\varepsilon(b)$ maximal endlich viele Folgenglieder. Also kann b nicht Grenzwert sein. (Widerspruch)

■

Satz 3.3

Jede konvergente Folge ist beschränkt.

Beweis
Kurzargumentation: Sei $\lim\limits_{n\to\infty} a_n = a$. Wähle irgendein $\varepsilon > 0$. Dann liegen in $U_\varepsilon(a)$ alle Folgenglieder bis auf endlich viele. Diese endlich vielen Folgenglieder bilden eine beschränkte Menge. $U_\varepsilon(a)$ ist auch beschränkt, also ist die Folge insgesamt beschränkt.

■

Die Konvergenz komplexer Folgen lässt sich leicht auf die Konvergenz reeller Folgen zurückführen.

Satz 3.4

Sei (a_n) eine komplexe Folge mit $a_n = b_n + ic_n$. Dann gilt: $a_n \to a = b + ic \Leftrightarrow b_n \to b \wedge c_n \to c$.

Satz 3.5 Rechenregeln für konvergente Folgen

Seien (a_n) und (b_n) konvergente Folgen mit $a_n \to a$, $b_n \to b$. Seien $\lambda, \mu \in \mathbb{C}$. Dann gilt:

1. $(\lambda a_n + \mu b_n)$ konvergiert gegen $\lambda a + \mu b$.
2. $(a_n \cdot b_n)$ konvergiert gegen $a \cdot b$.
3. $\forall n \in \mathbb{N} : b_n \neq 0 \wedge b > 0 \Rightarrow \left(\frac{a_n}{b_n}\right)_{n>k}$ konvergiert gegen $\frac{a}{b}$.

Beweis
1. und 3. sind Thema der Übungsaufgaben. Zu 2.:
Kurzargumentation: (a_n) ist konvergent, also beschränkt, d.h. $\exists C : |a_n| < C$. Gezeigt wird $|a_nb_n - ab| \to 0$.
$|a_nb_n - ab| = |a_nb_n - a_nb + a_nb - ab| \leq$ (Dreiecksungl.) $\leq |a_nb_n - a_nb| + |a_nb - ab|$
$= |a_n|\,|b_n - b| + |b|\,|a_n - a| \leq C\,\underbrace{|b_n - b|}_{\to 0} + |b|\,\underbrace{|a_n - a|}_{\to 0} \to 0$

Ausführlicher Beweis mit ε und n_0:
Sei $\varepsilon > 0$ gegeben. $a_n \to a \Rightarrow |a_n - a| < \frac{\varepsilon}{2|b|}$ für $n > n_1$, $b \neq 0$. Analog ist $|b_n - b| < \frac{\varepsilon}{2C}$ für $n > n_2$.
Sei $n_0 > n_1, n_2 \Rightarrow \forall n > n_0 : |a_nb_n - ab| \leq$ (siehe Kurzfassung) $\leq C\,|b_n - b| + |b|\,|a_n - a|$
$< C\frac{\varepsilon}{2C} + |b|\,\frac{\varepsilon}{2|b|} = \varepsilon \Rightarrow |a_nb_n - ab| \leq \varepsilon \Rightarrow a_nb_n \to ab$

■

Die folgenden Eigenschaften werden auch oft benutzt.

Satz 3.6

(a_n), (b_n) seien reelle Folgen mit $a_n \to a$, $b_n \to b$. Dann gilt:

1. $\forall n \in \mathbb{N} : a_n < b_n$ (oder $a_n \leq b_n$) $\Rightarrow a \leq b$ (<u>nicht</u> $a < b$)
2. („Sandwich-Theorem") Wenn $a = b$ und (c_n) eine Folge mit $a_n \leq c_n \leq b_n \; \forall n$ ist, dann $c_n \to a$.

Beweis

1. $\forall \varepsilon > 0 : \exists n_0 : \forall n > n_0 : |a_n - a| < \varepsilon, |b_n - b| < \varepsilon$
   Das heißt: $a - \varepsilon < a_n < a + \varepsilon$ und $b - \varepsilon < b_n < b + \varepsilon$.
   $a - \varepsilon < a_n \leq b_n < b + \varepsilon \Rightarrow \forall \varepsilon > 0 : a - b < 2\varepsilon \Rightarrow a - b \leq 0 \Rightarrow a \leq b$
2. $a_n \leq c_n \leq b_n \Rightarrow c_n - a_n \leq b_n - a_n \to 0 \Rightarrow c_n - a_n \to 0$
   $|c_n - a_n| = |c_n - a_n + a_n - a| \leq \underbrace{|c_n - a_n|}_{\to 0} + \underbrace{|a_n - a|}_{\to 0} \to 0 \Rightarrow c_n \to a$

■

Satz 3.7 Wichtige Beispiele für Grenzwerte

1. $\forall p \in \mathbb{N} : \frac{1}{\sqrt[p]{n}} \to 0$
2. $\forall a > 0 : \sqrt[n]{a} \to 1$
3. $\sqrt[n]{n} \to 1$
4. $\forall q \in \mathbb{C}, |q| < 1 : q^n \to 0$
5. $\forall k \in \mathbb{N}, z \in \mathbb{C}, |z| > 1 : \frac{n^k}{z^n} \to 0$

Beweis

1. $\forall \varepsilon > 0 : \exists n_0 \in \mathbb{N} : \forall n > n_0 : \frac{1}{n} < \varepsilon^p \Rightarrow \frac{1}{\sqrt[p]{n}} < \varepsilon$
2. Fallunterscheidung: ($a = 1$ ist trivial.)
   <u>$a > 1$</u>: Setze $x_n = \sqrt[n]{a} - 1$ ($\forall n : x_n > -1$).
   $a = (1 + x_p)^n \geq$ (Bernoulli) $\geq 1 + nx_n$
   $\Rightarrow$ (mit $x_n = 0$) $\Rightarrow 0 \leq x_n \leq \frac{a-1}{n} \leq \frac{a}{n} \to 0 \Rightarrow x_n \to 0 \Rightarrow \sqrt[n]{a} \to 1$ <u>$a < 1$</u>: $\frac{1}{a} > 1 \Rightarrow \sqrt[n]{a} = \left(\sqrt[n]{\frac{1}{a}}\right)^{-1} = \left(\frac{1}{\sqrt[n]{a}}\right)^{-1} \to \left(\frac{1}{1}\right)^{-1} = 1$
3. (analog zu 2.) Setze $x_n = \sqrt[n]{n} - 1$ ($x_n > 0$). Anwendung des binomischen Satzes:
   $n = (1 + x_n)^n \geq 1 + \binom{n}{2} x_n^2$ (alle Teile bis auf den ersten und dritten werden weggelassen)
   $n - 1 \geq \frac{n(n-1)}{2} x_n^2 \Rightarrow x_n \leq \frac{2}{\sqrt{n}} \to 0 \Rightarrow \sqrt[n]{n} - 1 \to 0 \Rightarrow \sqrt[n]{n} \to 1$
4. Sei $q \neq 0$ ($q = 0$ trivial). Setzt $a := \frac{1-|q|}{|q|} > 0 \Rightarrow |q| = \frac{1}{1+a}$
   $\Rightarrow$ (Bernoulli) $\Rightarrow (1 + a)^n \geq 1 + na > 0 \Rightarrow$
   $0 \leq |q^n| = |q|^n = \frac{1}{(1+a)^n} \leq \frac{1}{1+na} \leq \frac{1}{na} \to 0 \Rightarrow |q|^n \to 0 \Rightarrow q^n \to 0$

■

Im folgenden werden einige wichtige Sätze behandelt, die Aussagen über die Konvergenz von Folgen machen, ohne den Grenzwert zu kennen.

Definition 3.8

Eine reelle Folge (a_n) heißt:

- monoton wachsend (fallend), wenn $a_n \leq a_{n+1}$ $(a_n \geq a_{n+1})$ $\forall n \in \mathbb{N}$
- streng monoton wachsend (fallend), wenn $a_n < a_{n+1}$ $(a_n > a_{n+1})$ $\forall n \in \mathbb{N}$

Satz 3.9

Sei (a_n) eine monotone und beschränkte reelle Folge, dann ist (a_n) konvergent.

- Wenn (a_n) monoton wachsend ist, ist $\lim\limits_{n\to\infty} a_n = \sup\{a_n\}$
- Wenn (a_n) monoton fallend ist, ist $\lim\limits_{n\to\infty} a_n = \inf\{a_n\}$

Beweis
Wir beweisen nur 1., 2. ist eine Analogie.
Sei $a = \sup\{a_n\}$ (existiert gemäß Vollständigkeitsaxiom), $\varepsilon > 0$ beliebig.
$\Rightarrow a_n \leq a \ \forall n$. Nach Supremumsdefinition $\exists n_0 \in \mathbb{N} : a - \varepsilon < a_{n_0}$
$\Rightarrow$ Monotonie (Sei $n > n_0$.) $\Rightarrow a - \varepsilon < a_{n_0} \leq a_n \leq a < a + \varepsilon \Rightarrow a - \varepsilon < a_n < a + \varepsilon \Rightarrow a_n \to a$

∎

Der folgende Begriff ist sehr wichtig.

Definition 3.10

Sei (a_n) eine beliebige Folge und $n_1 < n_2 < \ldots$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann nennt man die Folge $(a_{n_k})_{k\in\mathbb{N}} = (a_{n_1}, a_{n_2}, \ldots)$ eine Teilfolge von (a_n).

Beispiel 3.4

Sei $(a_n) = (1, 0, 0, 2, 7, 4, 4, 3, 0, 0, 0, \ldots)$.
$(0, 7, 3, 0, 0, 0, \ldots)$ ist eine Teilfolge von (a_n), $(0, 7, 2, 3, 3, 0, 0, 0, \ldots)$ jedoch nicht.

Satz 3.11

Jede Teilfolge einer gegen a konvergenten Folge konvergiert auch gegen a.

Satz 3.12 Satz von Bolzano-Weierstrass

Jede beschränkte Folge reeller oder komplexer Zahlen besitzt eine konvergente Teilfolge.

Beweis
Wir zeigen den Satz nur für reelle Folgen (der Beweis für komplexe Folgen ist eine Übungsaufgabe).
Da (a_n) beschränkt, $\exists c, d \in \mathbb{R} : c \leq a_n \leq d$.
1. Schritt: $M = \{x \in \mathbb{R} : a_n > x$ für höchstens endlich viele $n \in \mathbb{N}\}$
Offenbar ist M nichtleer ($d \in M$) und durch c nach unten beschränkt.
2. Schritt: Setze $a := \inf M$. Zeige: Es existiert eine gegen a konvergente Teilfolge.
3. Schritt: Sei $\varepsilon > 0$ gegeben. Dann gilt: $a + \varepsilon \in M$, $a - \varepsilon \notin M$, da alle $x \in M$ erfüllen müssen: $x \geq a$
$a - \varepsilon \notin M \Rightarrow \exists$ unendlich viele a_n mit $a - \varepsilon < a_n < a + \varepsilon \qquad (*)$
4. Schritt: Die Teilfolge wird wie folgt gewählt, wobei $(*)$ auf verschiedene ε angewendet wird:

- $\varepsilon = 1 : a_{n_1}$ mit $a - 1 < a_{n_1} < a + 1$
- $\varepsilon = \frac{1}{2} : a_{n_2}$ mit $a - \frac{1}{2} < a_{n_2} < a + \frac{1}{2}$, $n_2 > n_1$
  (möglich, da im Intervall unendlich viele Elemente liegen)

- $\varepsilon = \frac{1}{k} : a_{n_k}$ mit $a - \frac{1}{k} < a_{n_k} < a + \frac{1}{k},\ n_k > n_{k-1} > \ldots > n_2 > n_1$

$\Rightarrow (a_{n_k})$ ist eine Teilfolge von (a_n) und (nach Sandwichtheorem) gilt: $a - \frac{1}{n} \to a \wedge a + \frac{1}{n} \to a \Rightarrow a_{n_k} \to a$

■

Satz 3.13 Cauchy-Kriterium

Eine Folge (a_n) komplexer Zahlen konvergiert genau dann, wenn gilt:
$\forall \varepsilon > 0 : \exists n_0 = n_0(\varepsilon) : \forall m, n > n_0 : |a_n - a_m| < \varepsilon \qquad (CK)$

Bemerkung

Dies ist das einzige notwendige und hinreichende Konvergenzkriterium.

Beweis
Notwendigkeit der Aussage:
Sei $a_n \to a$ und $\varepsilon > 0$ gegeben.$\Rightarrow \exists n_0 : |a_k - a| < \frac{\varepsilon}{2}\ \forall k > n_0$
$\Rightarrow \forall m, n > n_0 : |a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a - a_m| < \varepsilon$
Das heißt: $a_n \to a \Rightarrow (CK)$

Hinlänglichkeit der Aussage:
Wir zeigen: Wenn (CK) gilt, ist (a_n) beschränkt.
Betrachte zum Beispiel $\varepsilon = 1$, d.h. $|a_m - a_n| < 1\ \forall m, n > n_0 = n_0(1)$.
Für $m = n_0 + 1$: $|a_{n_0+1} - a_n| < 1\ \forall n > n_0$
$\Rightarrow |a_n| = |a_n - a_{n_0+1} + a_{n_0+1}| \leq |a_n - a_{n_0+1}| + |a_{n_0+1}| \leq 1 + |a_{n_0+1}|$
Da $a_1, \ldots, a_n$ sicherlich beschränkt sind, ist die Folge (a_n) insgesamt beschränkt.
Nach Satz von Bolzano-Weierstrass existiert eine konvergente Teilfolge (a_{n_k}) mit $a_{n_k} \to a$.
Dann gilt auch $a_n \to a$. Im folgenden der Beweis für diesen Schritt:

Sei $\varepsilon > 0$ gegeben. $\exists n_0 : \forall m, n > n_0 : |a_m - a_n| < \frac{\varepsilon}{2}$ (nach Cauchykrit.)
$\exists n_1 : \forall n_k > n_1 : |a_{n_k} - a| < \frac{\varepsilon}{2}$ (nach Konvergenz der Teilfolge)
$\Rightarrow |a_n - a| = |a_n - a_{n_k} + a_{n_k} - a| \leq \underbrace{|a_n - a_{n_k}|}_{<\frac{\varepsilon}{2},\text{ falls } n_1 n_k > n_2} + \underbrace{|a_{n_k} - a|}_{<\frac{\varepsilon}{2},\text{ falls } n_k > n_2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon\ \forall n > n_2$
Hierbei ist $n_2 = \max\{n_0, n_1\}$.

■

Bemerkung

Die Notwendigkeit von CK benötigt nicht die Vollständigkeit von $\mathbb{R}$ (bzw. von $\mathbb{C}$), die Hinlänglichkeit hingegen schon.
Die Vollständigkeit von $\mathbb{R}$ haben wir über das Supremumsprinzip definiert. Man kann allerdings alternativ formulieren:

Eine Menge heißt vollständig, wenn jede Cauchyfolge mit Elementen aus dieser Menge einen Grenzwert besitzt.
(Hierbei ist eine Cauchyfolge eine Folge, die das Cauchykriterium erfüllt.)

Nimmt man diese Definition der Vollständigkeit als Axiom von $\mathbb{R}$, kann man das Supremumsprinzip beweisen.
Der Beweis entählt folgende Aussage, die für sich von Interesse ist: Wenn (a_n) eine Cauchyfolge mit einer gegen a konvergenten Teilfolge ist, konvergiert auch a_n gegen a.

Wir wollen das Bisherige in einer Rekursionsformel für $\sqrt{a}$ anwenden.

Satz 3.14

Sei $a > 0$ und sei (x_n) wie folgt rekursiv definiert:
$x_1 > 0$ beliebig und $x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ für $n \in \mathbb{N}$ $\quad (*)$
Dann ist (x_n) konvergent gegen $\sqrt{a}$.

Beweis Wir zeigen: (x_n) ist monoton fallend und nach unten beschränkt.
Beschränktheit: Benutze Ungleichung „geometrisches Mittel $\leq$ arithmetisches Mittel“
Mit $y := \frac{a}{x_n}$ und $x := x_n$ ist $\sqrt{xy} \leq \frac{1}{2}(x+y)$.
$\Rightarrow \sqrt{a} \leq \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = x_{n+1} \Rightarrow x_n$ nach unten beschränkt zumindest ab $n = 2$

Monotonie: $a \leq x_n^2 \Rightarrow \frac{a}{x_n} \leq x_n \Rightarrow x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \leq \frac{1}{2}(x_n + x_n) = x_n$.
Daher ist (x_n) monoton fallend.

Daraus folgt: (x_n) ist konvergent (Satz 3.9) gegen x.
$\Rightarrow$ Man kann in $(*)$ auf beiden Seiten zum Grenzwert übergehen. Dabei gilt: $x_n \to x \wedge x_{n+1} \to x$.
$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \Rightarrow \lim\limits_{n\to\infty} x_{n+1} = \lim\limits_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \Rightarrow x = \frac{1}{2}\left(x + \frac{a}{x}\right) \Rightarrow x = \frac{a}{x} \Rightarrow x = \sqrt{a}$ ■

Bemerkung

Wenn x_1 rational ist, sind alle x_n rational. Also ist (x_n) eine rationale Cauchyfolge, jedoch oft (z.B. für $a = 2$) ohne rationalen Grenzwert. Wie man sieht, ist $\mathbb{Q}$ nicht vollständig.

3.2 Exponential- und Logarithmusfunktion, allgemeine Potenzen

Sei $f : D(f) \to \mathbb{R}$ eine Funktion, so heißt f

- monoton wachsend bzw. fallend, wenn gilt: $x \leq y \Rightarrow f(x) \leq f(y)$ bzw. $f(x) \geq f(y)$
- streng monoton wachsend (fallend), wenn $x < y \Rightarrow f(x) < f(y)$ bzw. $f(x) > f(y)$

Satz 3.15

Betrachte die Folgen $a_n = \left(1 + \frac{x}{n}\right)^n$, $b_n = \left(1 - \frac{x}{n}\right)^{-n}$. Sei $x \in \mathbb{R}$ fest und n_0 so, dass $n_0 > |x|$. Dann gilt:

1. $a_n \leq a_{n+1}\ \forall n \geq n_0$ (falls $x \geq 0$, gilt dieses $\forall n$)
2. $b_n \geq b_{n+1}$
3. $a_n \leq b_n\ \forall n \geq n_0$
4. $\lim\limits_{n\to\infty} (b_n - a_n) = 0$

Bezeichnet man mit $I_n = [a_n, b_n]$ die entsprechenden Intervalle, bildet (I_n) eine sogenannte Intervallschachtelung, d.h. $I_n \supset I_{n+1}\ \forall n$ und $|I_n| = b_n - a_n \to 0$.

Beweis

1. Für $n_0 > |x|$ und $n > n_0$ gilt: $1 + \frac{x}{n} > 0$. Definiert man h durch $1 + h = \frac{1+\frac{x}{n+1}}{1+\frac{x}{n}}$, d.h. $h = -\frac{x}{n(n+1)\left(1+\frac{x}{n}\right)}$, dann folgt $1 + h > 0$. Mit der Bernoulli'schen Ungleichung folgt:

$\frac{a_{n+1}}{a_n} = \frac{\left(1+\frac{x}{n+1}\right)^{n+1}}{\left(1+\frac{x}{n}\right)^n} = \left(1+\frac{x}{n}\right)\cdot(1+h)^{n+1} \geq \left(1+\frac{x}{n}\right)\cdot(1+(n+1)\,h) = \left(1+\frac{x}{n}\right)\cdot\left(1-\frac{x}{1+\frac{x}{n}}\right) = 1.$
Da $a_n > 0$, gilt also $a_{n+1} \geq a_n$.

2. Für $n \geq n_0 > |x|$ ist $1-\frac{x}{n} > 0 \Rightarrow \frac{1}{b_n} = \left(1-\frac{x}{n}\right)^n = \left(1+\frac{-x}{n}\right)^n =$ monoton wachsend (siehe 1.) Also ist (b_n) monoton fallend.
3. $\frac{a_n}{b_n} = \left(1+\frac{x}{n}\right)^n \cdot \left(1-\frac{x}{n}\right)^n = \left(1-\frac{x^2}{n^2}\right)^n \leq 1 \overset{(b\geq 0)}{\Rightarrow} a_n \leq b_n$
4. Bernoulli anwenden: $b_n - a_n = b_n\left(1-\frac{a_n}{b_n}\right) = b_n\left(1-\left(1-\frac{x^2}{n^2}\right)^n\right) \leq b_n \cdot \frac{x^2}{n} \leq b_{n_0}\cdot\frac{x^2}{n} \to 0$

■

(a_n) und (b_n) sind monoton und beschränkt, also konvergent. Da $b_n - a_n \to 0$, haben beide Folgen denselben Grenzwert. Setze $f(x) := \lim\limits_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim\limits_{n\to\infty}\left(1-\frac{x}{n}\right)^{-n}$.

Lemma 3.16

Diese Funktion f hat die folgenden Eigenschaften:

1. $f(0) = 1$
2. $f(x+y) = f(x)\cdot f(y)$
3. $f(x) > 0, f(-x) = \frac{1}{f(x)}$

Beweis

1. $f(0) = \lim\limits_{n\to\infty}\lim\limits_{n\to\infty}\left(1+\frac{0}{n}\right)^n = \lim\limits_{n\to\infty} 1^n = 1$
2. $f(x+y) - f(x)\cdot f(y) = \lim\limits_{n\to\infty}\left[\left(1+\frac{x+y}{n}\right)^n - \left(1+\frac{x}{n}\right)^n\cdot\left(1+\frac{y}{n}\right)^n\right]$
   $[\ldots]$ hat die Gestalt: $c^n - (ab)^n = (c-ab)\cdot\sum\limits_{k=1}^{n} c^{n-k}\,(ab)^{k-1} \qquad (*)$
   Dabei ist $|ab| \leq \left|1+\frac{|x|}{n}\right|\cdot\left|1+\frac{|y|}{n}\right|$ und $|c| \leq 1+\frac{|x|}{n}+\frac{|y|}{n} \leq \left(1+\frac{|x|}{n}\right)\cdot\left(1+\frac{|y|}{n}\right)$.
   Außerdem ist $c - ab = -\frac{xy}{n^2}$.
   $\Rightarrow \left|c^{n-k}\,(ab)^{k-1}\right| \leq \left(1+\frac{|x|}{n}\right)^{n-1}\cdot\left(1+\frac{|y|}{n}\right)^{n-1} \leq \left(1+\frac{|x|}{n}\right)^n\cdot\left(1+\frac{|y|}{n}\right)^n \leq f(|x|)\cdot f(|y|)$
   $\Rightarrow |[\ldots]| \leq \frac{|x|\cdot|y|}{n^2}\cdot n\cdot f(|x|)\cdot f(|y|) = \frac{f(|x|)\cdot f(|y|)\cdot|x|\cdot|y|}{n} \to 0$
   $f(x+y) - f(x)\cdot f(y) = 0 \Rightarrow f(x+y) = f(x)\cdot f(y)$

■

Definition 3.17

Die **Euler'sche Zahl** ist $\mathrm{e} := f(1) = \lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n \approx 2{,}71828\ldots$.

Für jede rationale Zahl $x = \frac{m}{n}$, $m,n\in\mathbb{N}$ ist $f(x) = \mathrm{e}^x$.

Beweis

Sei o.E.d.A. $\frac{m}{n} > 0$ ($\frac{m}{n} < 0$ trivial).

Dann gilt: $\mathrm{e}^m = f(1)^m \overset{(3.16.2)}{=} f\underbrace{(1+\ldots+1)}_{m\text{-mal}} = f(m) = f(nx) = f\underbrace{(x+\ldots+x)}_{n-\text{mal}} = f(x)^n$

$\Rightarrow \mathrm{e}^x = \mathrm{e}^{\frac{m}{n}} = \sqrt[n]{\mathrm{e}^m} = \sqrt[n]{f(1)^m} = \sqrt[n]{f(x)^n} = f(x)$

■

3.2.1 Exponentialfunktion

Definition 3.18

Die Funktion $\mathrm{e}^x := \lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n \ \forall x \in \mathbb{R}$ heißt **Exponentialfunktion**. Notation: $\exp(x)$

Bemerkung

Man kann zeigen: Wenn $r_n \to x$, $r_n \in \mathbb{Q}$, $x \in \mathbb{R}$, dann ist $e^{r_n} = f(r_n) \to f(x)$.

Satz 3.19

1. $\mathrm{e}^0 = 1$
2. $\mathrm{e}^{x+y} = \mathrm{e}^x \cdot \mathrm{e}^y$
3. $\mathrm{e}^x \geq 1 + x$, Gleichheit nur für $x = 0$
4. $x < y \Rightarrow \mathrm{e}^x < \mathrm{e}^y$
5. $|x| < 1 \Rightarrow |\mathrm{e}^x - 1| \geq \frac{|x|}{1-|x|}$ - Daraus folgt: $x_n \to 0 \Rightarrow \mathrm{e}^{x_n} \to 1$, d.h. e^x ist stetig für $x = 0$.
6. e^x wächst stärker als jede Potenz von x. Genauer gilt für ein $n \in \mathbb{N}$: $\mathrm{e}^x > x^n \ \forall x > 4n^2$.
7. $\lim_{n\to\infty} e^{-n} = 0$

Beweis

3. Wenn $x > -1$, dann gilt nach (3.15.1): $1 + x = \left(1+\frac{x}{1}\right)^1 \overset{(Monot.)}{\leq} \left(1+\frac{x}{n}\right)^n \leq \mathrm{e}^x$.
   Für $x \leq -1$ ist ohnehin $\mathrm{e}^x > 0 \geq 1 + x$.
4. Nach 3. gilt $\mathrm{e}^x > 1$ für $x > 0$. Wenn $x < y$, dann gilt: $\mathrm{e}^y - \mathrm{e}^x = \mathrm{e}^x\left(\mathrm{e}^{y-x} - 1\right) > 0$,
   da $y - x > 0 \Rightarrow \mathrm{e}^{y-x} > 1$ und $\mathrm{e}^{y-x} > 0$. Also $\mathrm{e}^y > \mathrm{e}^x$.
5. Nimmt man in (3.15.2) $\left(b_n > b_{n+1}, b_n = \left(1-\frac{x}{n}\right)^{-n}\right)$ $n_0 > 1$, dann erhält man:
   $\left(1-\frac{x}{n}\right)^{-n}$ monoton fallend gegen e^x, also falls $|x| < 1$: $\mathrm{e}^x \leq \left(1-\frac{1}{x}\right)^{-1} = \frac{x}{1-x}$
   Für $x \geq 0$ ergibt sich die Behauptung unmittelbar: $\mathrm{e}^x - 1 \leq \frac{1}{1-x} - 1 = \frac{x}{1-x}$.
   Für $-1 < x < 0$ benutze 3.: $1 - \mathrm{e}^x \leq -x = x < \frac{|x|}{1-|x|}$
6. Sei $x > 4n^2$. Dann $\sqrt{x} > 2n \Rightarrow \sqrt{x}\cdot\sqrt{x} > 2n\cdot\sqrt{x} \Rightarrow \frac{x}{2n} > \sqrt{x}$.
   $\mathrm{e}^x > a_n \ \forall n \Rightarrow \mathrm{e}^x > a_{2n} = \left(1+\frac{x}{2n}\right)^{2n} > \left(1+\sqrt{x}\right)^{2n} \geq \sqrt{x}^{2n} = x^n$
7. Nach 3. gilt: $\mathrm{e}^n > 1 + n$ (für $n \neq 0$!) $\Rightarrow \mathrm{e}^{-n} < \frac{1}{n+1} \overset{(\mathrm{e}^{-n}>0)}{\Rightarrow} \mathrm{e}^{-n} \to 0$

■

3.2.2 Logarithmusfunktion

Satz 3.20

Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$ ist bijektiv. Die Umkehrfunktion wird mit $\log : x \mapsto \log(x)$ für $x > 0$ bezeichnet und heißt **(natürliche(r)) Logarithmus(funktion)**. (Statt log schreibt man auch ln.) log ist streng monoton wachsend.

Beweis

Da exp streng monoton wachsend ist, muss exp injektiv sein. Zeigen: exp ist surjektiv auf $\mathbb{R}_+ = (0, \infty)$. Sei $y > 0$. Zeigen: $\exists a \in \mathbb{R} : \mathrm{e}^a = y$. Betrachte dazu $A := \{x \in \mathbb{R} : \mathrm{e}^x < y\}$. $A \neq \varnothing$ und A ist nach oben beschränkt, denn $x < \mathrm{e}^x - 1$ (3.19.3), also $x < \mathrm{e}^x - 1 = y - 1$. Also existiert ein Supremum $a = \sup A$.
Zeigen: $\mathrm{e}^a = y$. Nach der Supremumsdefinition: $\forall x \in \mathbb{N} : \exists x_n \in A : a - \frac{1}{n} < x_n \leq a$.
Aus Monotonie von exp folgt: $\mathrm{e}^{a-\frac{1}{n}} < \mathrm{e}^{x_n} < y < \mathrm{e}^{a+\frac{1}{n}}$. $(a + \frac{1}{n} \notin A)$ $(*)$.
Aus (3.19.5) folgt: $\mathrm{e}^{a-\frac{1}{n}} = \mathrm{e}^a \mathrm{e}^{-\frac{1}{n}} \to \mathrm{e}^a$, $\mathrm{e}^{a+\frac{1}{n}} \to \mathrm{e}^a$. Aus $(*)$ folgt (Sandwichtheorem): $\mathrm{e}^y = a$.
Monotonie von log: Angenommen, aus $0 < x < y$ folgt $\log y \leq \log x$.
$\Rightarrow \exp(\log y) \leq \exp(\log x) \Rightarrow y \leq x \Rightarrow$ Widerspruch $\Rightarrow (0 < x < y \Rightarrow \log x < \log y)$

■

Die folgenden Eigenschaften folgen direkt aus (3.19) und werden hier nicht explizit bewiesen.

Satz 3.21 Eigenschaften des natürlichen Logarithmus

1. $\mathrm{e}^{\log y} = y \; \forall y > 0$, $\log \mathrm{e}^x = x \forall x \in \mathbb{R}$ - Speziell: $\log 1 = 0$, $\log \mathrm{e} = 1$.
2. $\log(ab) = \log a + \log b \; \forall a, b > 0$, $\log\left(\frac{a}{b}\right) = \log a - \log b \; \forall a, b > 0$
3. $\log(x^n) = n \log(x) \;\; \forall n \in \mathbb{N}$
4. $\log(1 + x) < x \; \forall x > -1, x \neq 0$
5. Der Logarithmus wächst langsamer als jede Linearfunktion $x \mapsto cx$.
   Genauer für $n \in \mathbb{N}$ mit $\frac{1}{n} < c$ ist $\log x < \frac{x}{n} < c \; \forall x > 4n^2$

3.2.3 Allgemeine Potenzen

Definition 3.22

Für $x > 0$, $y \in \mathbb{R}$ setzen wir $x^y := \mathrm{e}^{y \cdot \log x}$.

Bemerkung

Es gelten die üblichen Potenzgesetze: $(xy)^r = x^r y^r$, $(x^r)^s = x^{rs}$, $x^{r+s} = x^r x^s$.

Beweis

Wir zeigen, dass diese Definition für $x > 0$, $y \in \mathbb{Q}$ mit der bisherigen Definition kongruent ist.
Für $y = \frac{m}{n} > 0$ gilt wegen $\log x^m = m \log x$:
$n \log x^{\frac{m}{n}} = n \log(\sqrt[n]{x})^m = mn \log \sqrt[n]{x} = m \log(\sqrt[n]{x})^n = m \log x \Rightarrow \log x^{\frac{m}{n}} = \frac{m}{n} \log x \; \forall m, n \in \mathbb{N}$.
Für $y < 0$ benutze $\log\left(\frac{1}{a}\right) = -\log a \Rightarrow$ Rest analog. Also: $\log x^y = y \log x \Rightarrow x^y = \mathrm{e}^{y \cdot \log x} \; \forall x \in \mathbb{R}_+, y \in \mathbb{Q}$.

■

3.3 Häufungswerte von Zahlenfolgen

Wir wollen die Struktur divergenter Folgen etwas genauer untersuchen.

Definition 3.23

Sei a_n eine beliebige komplexe Folge. Die Zahl $a \in \mathbb{C}$ heißt **Häufungswert** von (a_n), wenn es eine Teilfolge (a_{n_k}) von (a_n) gibt mit $a_{n_k} \to a$.
Die Menge aller Häufungswerte von (a_n) heißt dann $H = H(a_n)$. (Diese Menge aller Häufungswerte heißt manchmal auch **Limesmenge** von (a_n).)

Unmittelbar aus der Definition folgt: (Beweisen Sie dieses Lemma als Übung!)

Lemma 3.24

$a \in H(a_n) \Leftrightarrow$ in jeder ε-Umgebung $U_\varepsilon(a)$ liegen unendlich viele Glieder von (a_n).

Beispiel 3.5

1. $(a_n) = (0, 1, 0, 1, ...) \Rightarrow H = \{0, 1\}$
2. Sei (a_n) irgend eine Durchnummerierung der rationalen Zahlen in $(0, 1) \Rightarrow H(a_n) = [0, 1]$
3. $(a_n) = (0, 1, 0, 2, 0, 3, ...) \Rightarrow H = \{0\}$
4. $(a_n) = (1, 2, 3, 4, ...) \Rightarrow H = \varnothing$

Eigenschaften von Häufungswerten

Ist (a_n) durch C beschränkt, so ist jede Teilfolge mit C beschränkt: $\forall a \in H(a_n) : |a| \leq C$.
Ist (a_n) eine reelle Folge mit $a_n \in (-C, C)\ \forall n$, so gilt: $H(a_n) \subset [-C, C] \Rightarrow \exists \inf H(a_n), \sup H(a_n)$.
Setze $\inf H(a_n) =: \alpha, \sup H(a_n) =: \beta$. Dann gilt $\alpha, \beta \in H(a_n)$ für reelle Folgen.

Beweis
Sei $\varepsilon > 0$ beliebig gegeben. Nach der Infimumsdefinition gilt: $\exists b \in H(a_n) : \alpha \leq b < \alpha + \varepsilon$
Wenn $b = \alpha$, folgt daraus unmittelbar die Behauptung. Wenn $b \neq \alpha$, dann existiert ein $\delta > 0$ mit:
$\alpha \leq b - \delta < b < b + \delta < \alpha + \varepsilon$
Nach Lemma liegen in der δ-Umgebung $U_\delta(b)$ unendlich viele a_n.
$\Rightarrow$ In $U_\varepsilon(\alpha)$ liegen auch unendlich viele Folgenglieder $\Rightarrow \alpha \in H(a_n)$.

■

Diese Überlegungen fasst der folgende Satz zusammen.

Definition und Satz 3.25

Jede reelle beschränkte Folge besitzt einen größten und einen kleinsten Häufungswert. Der größte Häufungswert heißt **limes superior** (oberer Limes). Man schreibt $\limsup a_n$. Der kleinste Häufungswert heißt **limes inferior** (unterer Limes). Man schreibt $\liminf a_n$. Kurzschreibweise: $\overline{\lim}\, a_n$, $\underline{\lim}\, a_n$

Satz 3.26

Eine reelle Folge (a_n) konvergiert genau dann (gegen a), wenn gilt: $\liminf a_n = \limsup a_n (= a)$

Der Beweis dieses Satzes ist eine Übung. Nützlich ist folgender Sprachgebrauch:

Definition 3.27

Eine reelle Folge (a_n) heißt

- divergent gegen ∞, wenn gilt: $\forall C > 0 : \exists n_0 : a_n > C \forall n > n_0$
- divergent gegen $-\infty$, wenn gilt: $\forall C > 0 : \exists n_0 : a_n < -C \forall n > n_0$

In beiden Fällen nennt man (a_n) bestimmt divergent.
Wenn (a_n) nicht nach oben (bzw. unten) beschränkt ist, dann setzt man $\limsup a_n = \infty$ bzw. $\liminf a_n = -\infty$.

3.4 Unendliche Reihen (Zahlenreihen)

Wir wenden den Grenzwertbegriff zur Untersuchung unendlicher Reihen an. Mit der ersten Definition soll dem Ausdruck $\sum\limits_{k=0}^{\infty} a_k$ ein Sinn gegeben werden.

Definition 3.28

1. Sei (a_k) eine unendliche Folge komplexer Zahlen. Dann heißt die Folge (s_n) der **Partialsummen** $s_n := \sum\limits_{k=0}^{n} a_k$ **unendliche Reihe** mit den Reihengliedern a_k. Man schreibt $\sum\limits_{k=0}^{\infty} a_k$ oder $a_1 + a_2 + \ldots$.
2. Die unendliche Reihe heißt **Konvergent**, wenn die Folge der Partialsummen konvergiert. Der Grenzwert $s := \lim s_n$ heißt **Summe** der unendlichen Reihe. Man schreibt $s = \sum\limits_{k=0}^{\infty} a_k$ oder einfach $s = \sum a_k$.
3. Wenn (s_n) divergiert, heißt die unendliche Reihe **divergent**.

Bemerkung

Manchmal beginnt die Summation bei einer unendlichen Reihe nicht bei $k = 0$, sondern bei $k = l > 0$.

Um mit dem Begriff arbeiten zu können, formulieren wir jetzt das Konvergenzkriterium, das Cauchy-kriterium und das Monotoniekriterium.

Satz 3.29

1. $\sum\limits_{k=0}^{\infty} a_k = s \Leftrightarrow \forall \varepsilon > 0 : \exists n_0(\varepsilon) : \forall n > n_0 : |s - s_n| < \varepsilon$
$$\Leftrightarrow \left| \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n} a_k \right| = \left| \sum_{k=n+1}^{\infty} a_k \right| < \varepsilon \Leftrightarrow \sum_{k=n+1}^{\infty} a_k \to 0 \text{ für } n \to \infty$$
2. $\sum\limits_{k=0}^{\infty} a_k$ konvergent $\Leftrightarrow \forall \varepsilon > 0 : \exists n_0(\varepsilon) : \forall m, n > n_0 : |s_n - s_m| < \varepsilon$
$\Leftrightarrow$ o.E.d.A. $m < n \Leftrightarrow \left| \sum\limits_{k=0}^{n} a_k - \sum\limits_{k=0}^{m} a_k \right| < \varepsilon \Leftrightarrow \left| \sum\limits_{k=m+1}^{n} a_k \right| < \varepsilon$, d.h. $\forall p \in \mathbb{N}, m > n_0 :$ $\left| \sum\limits_{i=1}^{p} a_{m+i} \right| < \varepsilon$
3. Seien alle $a_n \geq 0$. Dann ist (s_n) **monoton wachsend** und es gilt:
$\sum\limits_{k=0}^{\infty} a_k$ ist konvergent $\Leftrightarrow (s_n)$ ist nach oben beschränkt

Bemerkung

Der Ausdruck $\sum\limits_{k=n+1}^{\infty} a_k$ heißt **Reihenrest**. Aus dem Cauchykriterium folgt der folgende Satz.

Satz 3.30 notwendiges Konvergenzkriterium

$\sum\limits_{k=0}^{\infty} a_k$ konvergent $\Rightarrow a_n \to 0$ (bzw. $a_n \nrightarrow 0 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ nicht konvergent)

Beweis
Wende Cauchykriterium für $p = 1$ an:
$\sum a_k$ konvergent $\Rightarrow \forall \varepsilon > 0 : \exists n_0(\varepsilon) : \forall m > n_0 : \left|\sum\limits_{i=1}^{1} a_{m+i}\right| = |a_{m+1}| < \varepsilon \Leftrightarrow a_n \to 0.$ ■

Beispiel 3.6 Geometrische Reihe

Sei $a_n = q^n$ mit $q \in \mathbb{C}$. Dann ist $s_n = \sum\limits_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}$ (hier ohne Beweis).
Da $q^n \to 0$ nur für $|q| < 1$, ist $\sum\limits_{k=0}^{\infty} q^k$ für $|q| \geq 1$ nicht konvergent. Für $|q| < 1$ ist: $\sum\limits_{k=0}^{\infty} q^k = \lim\limits_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q}$

Beispiel 3.7 Harmonische Reihe - Beispiel, dass $a_n \to 0$ kein hinreichendes Kriterium ist

Sei $a_n = \frac{1}{n}$, dann $a_n \to 0$, aber $\sum\limits_{k=1}^{\infty} a_k = 1 + \frac{1}{2} + \frac{1}{3} + \ldots \to \infty$, da Cauchy-Kriterium verletzt:
$$|s_{2n} - s_n| = \underbrace{\frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}}_{n \text{ Summanden, alle } \geq \frac{1}{2n}} \geq n \cdot \frac{1}{2n} = \frac{1}{2}$$
Die Summe von n bis $2n$ wird mit steigenden n nicht beliebig klein, sondern ist immer größer als $\frac{1}{2}$.

3.4.1 Rechnen mit konvergenten Reihen

- Wenn $\sum a_k = a$, $\sum b_k = b$, $\lambda \in \mathbb{C}$, dann ist $\sum(\lambda a_k + b_k) = \lambda a + b$.
  Die Menge aller konvergenten Reihen bildet also einen $\mathbb{C}$-Vektorraum.
- Wenn $\sum a_k$ konvergiert, dann darf man in der Summe beliebig Klammern setzen, ohne die Konvergenz zu ändern. Allerdings darf man keine Klammern wegnehmen.
  Zum Beispiel ist $(1-1)+(1-1)+(1-1)+\ldots$ konvergent, $1-1+1-1+1-1+\ldots$ jedoch divergiert.
- Multiplikation und Umordnung von Reihen erfordern einen stärkeren Konvergenzbegriff:

Definition 3.31

Man sagt, $\sum a_k$ **konvergiert absolut**, wenn $\sum |a_k|$ konvergiert. Wenn $\sum a_k$ konvergiert, aber $\sum |a_k|$ divergiert, heißt $\sum a_k$ auch **nicht absolut konvergent** oder **bedingt konvergent**.

Satz 3.32 Allgemeine Dreiecksungleichung

Eine absolut konvergente Reihe ist konvergent. Es gilt: $|\sum a_k| \leq \sum |a_k|$

Beweis
Cauchy-Kriterium: Wenn $|a_k|$ konvergiert, gilt: $\forall \varepsilon > 0 : \exists n_0 : \forall p \in \mathbb{N}, n > n_0 : |a_{n+1}| + \ldots + |a_{n+p}| < \varepsilon$.

Mit der Dreiecksungleichung ist: $|a_{n+1} + \ldots + a_{n+p}| \leq |a_{n+1}| + \ldots + |a_{n+p}| < \varepsilon$.
Also ist das Cauchy-Kriterium auch für die Ausgangsreihe erfüllt.

■

Beispiel 3.8 Aus Konvergenz folgt nicht unbedingt absolute Konvergenz.

Sei $a_k = \frac{(-1)^k}{k}$. Dann ist $\sum\limits_{k=1}^{\infty} a_k = -1 + \frac{1}{2} - \frac{1}{3} + \ldots$ konvergent (Beweis folgt).
$\sum\limits_{k=1}^{\infty} |a_k| = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$ ist jedoch divergent (harmonische Reihe).

Definition 3.33

Sei $\varphi : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann nennt man $\sum\limits_{k=1}^{\infty} a_{\varphi(k)}$ eine **Umordnung** von $\sum\limits_{k=1}^{\infty} a_k$.

Satz 3.34 Der erste Teilsatz ist der Riemann'sche Umordnungssatz.

1. Seien $a_n \in \mathbb{R}$. Wenn $\sum a_k$ konvergiert, aber nicht absolut, dann existiert $\forall S \in \mathbb{R} \cup \{\infty, -\infty\}$ eine Umordnung $\sum a_{\varphi(k)}$ mit $\sum a_{\varphi(k)} = S$.
2. Wenn $\sum a_k$ absolut konvergiert und $\sum a_k = s$, dann ist $\sum a_{\varphi(k)} = s\ \forall \varphi$.
   Umgekehrt: Wenn jede Umordnung konvergiert, konvergiert $\sum a_k$ absolut.

Multiplikation unendlicher Reihen - Worin besteht das Problem?
Man möchte $(a_1 + \ldots + a_n) \cdot (b_1 + \ldots + b_n)$ ausmultiplizieren und dann aufsummieren. Dabei entsteht ein Schema der Art

$$\begin{array}{cccc} a_0b_0 & a_0b_1 & a_0b_2 & \cdots \\ a_1b_0 & a_1b_1 & a_1b_2 & \cdots \\ a_2b_0 & a_2b_1 & a_2b_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array} \qquad (*)$$

Aufsummieren von $(*)$ heißt: $(*)$ abzählen. Jede Art dieses Abzählens liefert eine Produktreihe.
Sinnvoller Wunsch: Wenn $\sum a_n = a \wedge \sum b_n = b$, dann soll jede Produktreihe gegen ab konvergieren.

Satz 3.35

Seien $\sum a_n$ und $\sum b_n$ absolut konvergent mit $\sum a_n = a$ und $\sum b_n = b$, dann konvergiert jede ihrer Produktreihen absolut und zwar gegen $a \cdot b$.

Die üblichste und wichtigste Form der Produktreihe ist die sogenannte **Cauchy'sche Produktreihe**, die entsteht, wenn man $(*)$ „diagonal" aufsummiert, d.h. man fasst so zusammen:

$$c_n = a_0b_n + a_1b_{n-1} + \ldots + a_{n-1}b_1 + a_nb_0 = \sum_{j=0}^{n} a_jb_{n-j}$$

Dann ist $\sum a_n \cdot \sum b_n = a \cdot b = \sum c_n$. (Die c_n entstehen zum Beispiel bei Multiplikation von Polynomen.) Für die Konvergenz von $\sum c_n = ab$ reicht es auch, wenn nur eine der beiden Folgen absolut konvergiert.

Satz 3.36 Leibniz-Kriterium

Strebt (a_n) *monoton* gegen Null, dann konvergiert die alternierende Reihe $\sum (-1)^n a_n = a_0 - a_1 + a_2 - \ldots$

Beweis
Aus der Monotonie folgt: Alle a_n haben dasselbe Vorzeichen. O.E.d.A. $a_n > 0\ \forall n$. Anwendung des Cauchy-Kriteriums für die Folge (s_n) der Partialsummen:
$|s_{n+p} - s_n| = (-1)^{n+1} \left[a_{n+1} - a_{n+2} + \ldots + (-1)^{p-1}\, a_{n+p}\right]$

Fall 1: p ist gerade: $[\ldots] = (a_{n+1} - a_{n+2}) + \ldots + (a_{n+p-1} + a_{n+p})$
Da $a_n \to 0$ monoton, ist jeder dieser Summanden > 0. Also ist $[\ldots] > 0$. Setze die Klammern nun anders:
$0 < [\ldots] = a_{n+1} - (a_{n+2} + a_{n+3}) - \ldots - (a_{n+p-2} - a_{n+p-1}) - a_{n+p} < a_{n+1}$
$\Rightarrow |s_{n+p} - s_n| < a_{n+1}\ \forall p \in \mathbb{N} \Rightarrow s_{n+p} - s_n \to 0 \Rightarrow (s_n)$ ist Cauchyfolge.

Fall 2: p ist gerade: Analoge Abschätzung liefert, dass $|s_{n+p} - s_n| < a_{n+1}$.

■

Satz 3.37

Sei $\sum\limits_{n}^{\infty} (-1)^n\, a_n$ eine alternierende Reihe mit der monotonen Folge $a_n \to 0$ und $a_n > 0$.
Dann gilt die folgende Abschätzung für den Reihenrest: $\left|\sum\limits_{k=n+1}^{\infty} -1^k a_k\right| < a_{n+1}$.

Beweis
Im obigen Beweis wurde gezeigt: $|s_{n+p} - s_n| < a_{n+1}$. Für $p \to \infty$ ist:
$s_{n+p} \to s = \sum\limits_{k=1}^{\infty} -1^k a_k \Rightarrow |s - s_n| = \left|\sum\limits_{k=1}^{\infty} -1^k a_k - \sum\limits_{k=1}^{n} -1^k a_k\right| = \left|\sum\limits_{k=n+1}^{\infty} -1^k a_k\right| \leq a_{n+1}$

■

Satz 3.38 Majoranten- und Minorantenkriterium

$\sum a_n$, $\sum b_k$ seien unendliche Reihen mit nichtnegativen Gliedern. Sei $a_n \leq b_n$ zumindest $\forall n > N \in \mathbb{N}$.
Dann heißt $\sum b_k$ **Majorante(nreihe)** für $\sum a_n$ und $\sum a_n$ **Minorante(nreihe)** für $\sum b_k$. Es gilt:

1. Wenn $\sum b_k$ konvergiert, konvergiert auch $\sum a_n$.
2. Wenn $\sum a_n$ divergiert, divergiert auch $\sum b_k$.

Beweis
Offenbar folgt 2. aus 1. Deshalb beweisen wir hier nur 1.
$\sum b_n$ konvergent $\Leftrightarrow$ Folge der zugehörigen Partialsummen (s_n) ist monoton wachsend und beschränkt:
$\lim s_n = s = \sum b_n \quad (s_n \leq s\ \forall n \in \mathbb{N})$
Für die Folge (t_n) der Partialsummen von (a_n) gilt: (t_n) ist monoton wachsend und $t_n \leq s_n$.
Wenn man o.E.d.A. $a_k \leq b_k\ \forall k \in \mathbb{N}$ annimmt, ist $t_n \leq s_n \leq s\ \forall n \in \mathbb{N}$.
Also ist (t_n) auch beschränkt und damit konvergent.

■

Bemerkung

Dieses Kriterium wird häufig so angewendet: $\sum a_k$ ist eine beliebige Reihe und (b_k) eine Folge mit $b_k \geq 0$ und $|a_k| \leq b_k\ \forall k$. Wenn $\sum b_k$ konvergiert, konvergiert auch $\sum |a_k|$, also konvergiert $\sum a_k$ absolut.
Um das Kriterium effektiv anwenden zu können, benötigt man viele konvergente Vergleichsreihen, etwa:

– geometrische Reihen $\sum q^n$
– Dirichletreihen $\sum\limits_{n=1}^{\infty} \frac{1}{n^\alpha}$ (konvergieren für $\alpha > 1$, divergieren für $\alpha \leq 1$)

Um die Konvergenz der Dirichletreihen zu beweisen, verwendet man den folgenden Satz.

Satz 3.39 Cauchy'scher Verdichtungssatz

Sei (a_n) monoton fallend mit $a_n \geq 0 \; \forall n \in \mathbb{N}$. Dann gilt:
$\sum\limits_{k=1}^{\infty} a_k$ ist konvergent $\Leftrightarrow \sum\limits_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + \ldots$ ist konvergent

Beweis
Es genügt, zu zeigen, dass beide Folgen der Partialsummen gleichzeitig beschränkt (d.h. konvergent) oder unbeschränkt (d.h. divergent) sind. Sei $s_n = a_1, \ldots, a_n$ und $t_k = a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k}$.

Sei $n < 2^k$. Dann gilt wegen Monotonie: $s_n \leq a_1 + (a_2 + a_3) + (a_4 + \ldots + a_7) + \ldots + \left(a_{2^k} + \ldots + a_{2^{k+1}-1}\right)$
$\leq a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k} = t_k \Rightarrow s_n \leq t_k$ für $n < 2^k$.

Für $n > 2^k$ gilt andererseits: $s_n \geq a_1 + a_2 + (a_3 + a_4) + \ldots + \left(a_{2^{k-1}+1} + \ldots + a_{2^k}\right)$
$\geq \frac{1}{2} a_1 + a_2 + 2a_4 + \ldots + 2^{k-1} a_{2^k} = \frac{1}{2} t_k \Rightarrow s_n \geq \frac{1}{2} t_k$ für $n > 2^k$.

■

Beispiel 3.9 Anwendung auf die Dirichletreihen

Sei $a_n = \frac{1}{n^\alpha}$. Dann ist $2^k a_{2^k} = 2^k \cdot \left(\frac{1}{2^k}\right)^\alpha = \left(\frac{1}{2^k}\right)^{\alpha-1} = \left(\frac{1}{2^{\alpha-1}}\right)^k$.
$\sum\limits_{k=1}^{\infty} \left(\frac{1}{2^{\alpha-1}}\right)^k$ konvergiert als geometrische Reihe genau dann, wenn $0 \leq \frac{1}{2^{\alpha-1}} < 1 \Rightarrow \alpha > 1$
Also konvergiert a_n genau dann, wenn $\alpha > 1$.

Satz 3.40 Wurzelkriterium

Sei $\sum a_k$ mit $a_k \in \mathbb{C}$ eine unendliche Reihe.

1. Wenn ein $\vartheta \in (0,1)$ existiert, sodass gilt: $\sqrt[n]{|a_n|} \leq \vartheta \; \forall n \in \mathbb{N}$, dann ist $\sum a_n$ absolut konvergent. (Es reicht, die Beschränktheit des Wurzeltermes ab einem festen $N \in \mathbb{N}$ zu zeigen.)
2. Gilt $\sqrt[n]{|a_n|} \geq 1 \; \forall n \in \mathbb{N}$, dann divergiert $\sum a_n$.
   Darstellung des Kriteriums in Limesform: Sei $\limsup \sqrt[n]{|a_n|} =: A$. Dann gilt:
3. $0 \leq A < 1 \Rightarrow \sum a_n$ konvergiert absolut.
4. $A > 1 \Rightarrow \sum a_n$ divergiert.
5. Für $A = 1$ ist keine eindeutige Aussage möglich.
   (Man kann $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} =: A$ setzen, wenn dieser Grenzwert existiert.)

Beweis
Zu 1.: $\sqrt[n]{|a_n|} \leq \vartheta \Rightarrow |a_n| \leq \vartheta^n \Rightarrow$ Die geometrische Reihe $\sum \vartheta^n$ ist eine konvergente Majorante für $\sum |a_k|$, d.h. $\sum |a_k|$ konvergiert absolut.
Zu 2.: $\sqrt[n]{|a_n|} \geq 1 \Rightarrow |a_n| \geq 1 \Rightarrow (a_k)$ ist keine Nullfolge, also divergiert $\sum |a_k|$.
3. und 4. werden nicht hier, sondern in den Übungsaufgaben behandelt.
Zu 5.: Für $\sum \frac{1}{n}$ und $\sum \frac{(-1)^n}{n}$ gilt: $\sqrt[n]{|n|} \to 1$, die erste Reihe ist aber divergent, die zweite konvergent.

■

Beispiel 3.10

1. $\sum\limits_{n=2}^{\infty} \frac{1}{(\log n)^n}$ ist konvergent denn $\sqrt[n]{|n|} = \frac{1}{\log n} \to 0$.
2. $\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \sum\limits_{i=1}^{\infty} \left(\frac{1}{2^i} + \frac{1}{3^i}\right)$ konvergiert,
   denn $\lim \sqrt[n]{|a_n|}$ existiert nicht (Idealfall), aber $\overline{\lim} \sqrt[n]{|a_n|} = \frac{1}{\sqrt{2}} < 1$

Satz 3.41 Quotientenkriterium

1. $\sum a_n$ mit $a_n \in \mathbb{C}$. Wenn $\exists \vartheta \in (0,1)$ mit $\left|\frac{a_{n+1}}{a_n}\right| \leq \vartheta \; \forall n$ (bzw. $\forall n > N \in \mathbb{N}$), dann konvergiert $\sum a_n$ absolut.
2. Ist $\left|\frac{a_{n+1}}{a_n}\right| \geq 1 \; \forall n$ (bzw. $\forall n > N \in \mathbb{N}$), so divergiert $\sum a_n$.
   Darstellung des Kriteriums in Limesform: Sei $\lim \left|\frac{a_{n+1}}{a_n}\right| = A$. Dann gilt:
3. $A < 1 \Rightarrow \sum a_n$ konvergiert absolut.
4. $A > 1 \Rightarrow \sum a_n$ divergiert.
5. Für $A = 1$ ist keine eindeutige Aussage möglich.

Beweis
Analog zum Wurzelkriterium:
$\left|\frac{a_{n+1}}{a_n}\right| \leq \vartheta \Rightarrow a_{n+1} \leq \vartheta \, |a_n| \leq \vartheta^2 \, |a_n| \leq \ldots \leq \vartheta^n \, |a_n|$ - geometrische Reihe als konvergente Majorante ■

Bemerkung

Man kann das Quotientenkriterium auch mit $\overline{\lim}$ und $\underline{\lim}$ darstellen. Dann gilt:

- $\overline{\lim} \left|\frac{a_{n+1}}{a_n}\right| < 1 \Rightarrow$ absolute Konvergenz
- $\underline{\lim} \left|\frac{a_{n+1}}{a_n}\right| > 1 \Rightarrow$ Divergenz

Das Wurzelkriterium ist stärker als das Quotientenkriterium.

3.5 Potenzreihen

Definition 3.42

Seien $a_n \in \mathbb{C}$, $a \in \mathbb{C}$, $n = 0, 1, \ldots$. Unter einer **Potenzreihe** (PR) in $z - a$ (oder nach Potenzen von $(z - a)$) mit den Koeffizienten a_n versteht man die unendliche Reihe $\sum\limits_{k=0}^{\infty} a_n (z-a)^n$ (1)

(Wie üblich wird vereinbart: $(z-a)^0 := 1$, auch für $z = a$.)

Ziel: Bestimme alle $z \in \mathbb{C}$, für die (1) konvergiert.

Beispiel 3.11

Jede PR (1) konvergiert für $z = a$. Gibt es Reihen, die nur für $z = a$ konvergieren?
a_n muss so schnell wachsen, das keine Nullfolge von $(z-a)^n$ das Wachstum kompensieren kann.

1. $\sum\limits_{n=0}^{\infty} n^n (z-a)^n$ kann aufgrund des Wurzelkriteriums nur für $z = a$ konvergieren:
   $\sqrt[n]{n^n (z-a)^n} = n(z-a) \to \infty \; \forall z \neq a$
2. $\sum\limits_{n=0}^{\infty} \frac{(z-a)^n}{n!}$ konvergiert für alle $z \in \mathbb{C}$ aufgrund des Quotientenkriteriums:
   $\left| \frac{(z-a)^{n+1}}{(n+1)!} \frac{n!}{(z-a)^n} \right| = \frac{|z-a|}{n+1} \to 0 \; \forall z \in \mathbb{C}$

Wie sieht außer in diesen Extremfällen die Menge aller z, für die (1) konvergiert, aus? Die Menge ist (in der grafischen Darstellung) immer kreisförmig, wie aus dem folgenden Satz unmittelbar folgt.

Lemma 3.43

1. Die PR (1) möge für ein $z_0 \neq a$ konvergieren. Dann konvergiert (1) absolut für alle $z \in \mathbb{C}$ mit $|z-a| < |z_0 - a|$.
2. Die PR (1) möge für ein z_1 divergieren. Dann divergiert (1) auch für alle z mit $|z-a| > |z_1 - a|$.

Beweis
2. folgt aus 1. Zu 1.: Wenn $\sum a_n (z_0 - a)^n$ konvergiert, dann gilt $a_n (z_0-a)^n \to 0$. Also ist $a_n (z_0 - a)^n$ beschränkt,
zum Beispiel $|a_n (z_0 - a)^n| \leq M \; \forall n$. Sei nun $|z-a| < |z_0 - a|$. Also ist $q := \frac{|z-a|}{|z_0-a|} < 1$.

$$|a_n (z-a)^n| = \left| a_n \frac{|z-a|^n}{|z_0-a|^n} |z_0-a|^n \right| \leq |a_n |z_0-a|^n| \cdot \left| \frac{|z-a|^n}{|z_0-a|^n} \right| \leq Mq^n$$

Aufgrund der konvergenten Majorante Mq^n konvergiert $\sum a_n (z-a)^n$ absolut.

■

Definition und Satz 3.44

Die PR $\sum a_n (z-a)^n$ (1) möge nicht nur für $z = a$ und auch nicht für alle z konvergieren. Dann existiert eine positive Zahl ϱ mit der Eigenschaft:

- $\forall z, |z-a| < \varrho \Rightarrow$ (1) ist für z konvergent.
- $\forall z, |z-a| > \varrho \Rightarrow$ (1) ist für z divergent.
- $\forall z, |z-a| = \varrho \Rightarrow$ Die Konvergenz muss näher untersucht werden.

ϱ heißt **Konvergenzradius** der Potenzreihe. Der Kreis um a mit dem Radius ϱ heißt **Konvergenzkreis** der PR (1). Konvergiert (1) nur für $z = a$, so setze $\varrho := 0$. Konvergiert (1) für alle z, so setze $\varrho := \infty$.

Bemerkung

Werden Potenzreihen nur im Reellen betrachtet, wird aus dem Konvergenzkreis das Konvergenzintervall (z.B. $(a-\varrho, a+\varrho)$, der Typ des Intervalls muss näher untersucht werden).

Beweis
Sei $M := \{r > 0 : \exists z \in \mathbb{C}, |z-a| = r : (1) \text{ konvergiert für } z\}$. $M \neq \varnothing$, denn PR konvergiert nicht nur für $z = a$. M ist nach oben beschränkt, denn PR konvergiert nicht für alle z, d.h. $\exists z_1$: PR divergiert

für z_1. Aufgrund des Lemmas gehören alle S mit $|z_1 - a| < S$ nicht zu M.
Setze $\varrho := \sup M$. Dann hat ϱ alle Eigenschaften des Satzes.

1. (Wende Supremumsdefinition an.) Sei z so, dass $|z-a| < \varrho$, d.h. $\exists r \in M : |z-a| < r < \varrho$. Dann $\exists z_0 : |z_0 - a| = r$ und PR konvergent für z_0. Also konvergiert die PR für z.
2. Sei z so, dass $|z-a| > \varrho$, d.h. $\exists r\,(\notin M) : |z-a| > r > \varrho \Rightarrow \forall z_2$ mit $|z_2 - a| = r$ folgt aus dem Lemma, dass die PR für z_2 divergent ist.

■

Beispiel 3.12

Gesucht ist ϱ der Potenzreihe $\sum\limits_{n=1}^{\infty} \frac{z^n}{n}$ (d.h. $a = 0$, $a_n = \frac{1}{n}$, $n > 1$)
Anwendung des Wurzelkriteriums: $\sqrt[n]{\left|\frac{z^n}{n}\right|} = \frac{|z|}{\sqrt[n]{n}} \to |z| \Rightarrow \varrho = 1$, da PR für $|z| < 1$ konvergent
Für $|z| = 1$ muss das Verhalten näher untersucht werden: $z = 1 \Rightarrow \sum\limits_{n=1}^{\infty} \frac{1}{n}$ divergent; $z = -1 \Rightarrow \sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n}$ konvergent
Also ist das Konvergenzintervall $[-1, 1)$.

Beispiel 3.13 Kurzbeispiele

$\sum\limits_{n=1}^{\infty} nz^n \Rightarrow \varrho = 1$ (z.B. Wurzelkriterium), $\sum\limits_{n=1}^{\infty} \frac{z^n}{n^2} \Rightarrow \varrho = 1$

Aus dem Wurzelkriterium folgt: $\limsup \sqrt[n]{|a_n (z-a)^n|} = |z-a| \cdot \limsup \sqrt[n]{|a_n|} \begin{cases} > 1 & \Rightarrow \text{ Divergenz} \\ < 1 & \Rightarrow \text{ Konvergenz} \end{cases}$

Satz 3.45

Für die Potenzreihe $\sum\limits_{n} a_n (z-a)^n$ setze $l := \limsup \sqrt[n]{|a_n|}$. Dann gilt:

- $l = 0 \Rightarrow \varrho = \infty$, $l = \infty \Rightarrow \varrho = 0$
- $0 < l < \infty \Rightarrow \varrho = \frac{1}{l}$

Satz 3.46 Rechnen mit Potenzreihen

Seien $\sum a_n (z-a)^n$, $\sum b_n (z-a)^n$ zwei PR mit den Konvergenzradien ϱ_1 und ϱ_2. Ferner seien $\alpha, \beta \in \mathbb{C}$.

- Die Potenzreihe $\sum (\alpha a_n + \beta b_n)(z-a)^n$ hat den Konvergenzradius $\varrho \geq \min\{\varrho_1, \varrho_2\}$ und:
  $\forall z, |z-a| < \min\{\varrho_1, \varrho_2\} : \alpha \left(\sum a_n (z-a)^n\right) + \beta \left(\sum b_n (z-a)^n\right) = \sum (\alpha a_n + \beta b_n)(z-a)^n$
- Sei $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$. Dann hat die PR $\sum c_n (z-a)^n$ den Konvergenzradius $\varrho \geq \min\{\varrho_1, \varrho_2\}$ und:
  $\forall z, |z-a| < \min\{\varrho_1, \varrho_2\} : \left(\sum a_n (z-a)^n\right)\left(\sum b_n (z-a)^n\right) = \sum c_n (z-a)^n$

Beweis
Folgt unmittelbar aus den Rechenregeln für Zahlenfolgen.

■

Es folgt ein fundamentaler Begriff.

Definition 3.47

Sei f wie folgt definiert: $f : \{z : |z-a| < \varrho\} \to \mathbb{C}, z \mapsto \sum_n a_n (z-a)^n$, wobei ϱ der Konvergenzradius der Potenzreihe $f(z)$ sei.
Dann heißt f **analytische Funktion**. (Wenn $a_n, a, z \in \mathbb{R}$, heißt f reell analytische Funktion.)

Satz 3.48 Identitätssatz für Potenzreihen und analytische Funktionen

Seien $\sum c_n (z-a)^n$ und $\sum d_n (z-a)^n$ zwei Potenzreihen mit den Konvergenzradien $\varrho_1, \varrho_2 > 0$. Sei ferner (z_i) eine Folge mit den Eigenschaften:

1. $0 < |z_i - a| < \min\{\varrho_1, \varrho_2\}$, $\lim z_i = a$
   (alternative Formulierung: $|\lim z_i - a| < \min\{\varrho_1, \varrho_2\}$)
2. $\sum c_n (z_i - a)^n = \sum d_n (z_i - a)^n \ \forall i$

Dann sind beide Potenzreihen gleich, d.h. $c_n = d_n \ \forall n$.
Mit anderen Worten: Wenn zwei analytische Funktionen auf einer Punktfolge aus dem gemeinsamen Definitionsbereich übereinstimmen (Punkt 2) und diese Punktfolge innerhalb dieses Definitionsbereiches konvergiert (Punkt 1), sind beide Funktionen gleich.

Zum Beweis verwenden wir das folgende Lemma.

Lemma 3.49

Sei $\sum_{n=0}^{\infty} a_n (z-a)^n$ eine Potenzreihe mit dem Konvergenzradius $\varrho > 0$ und der Summe $f(z)$. Sei (z_i) eine Folge mit $\lim z_i = a$, $|z_i - a| < \varrho$. Dann gilt: $\lim_{i\to\infty} f(z_i) = a_0$.
Kurz: $\lim_{i\to\infty} \sum_n a_n (z_i - a)^n = a_0$ für $\lim_{i\to\infty} z_i = a$

$$\lim_{i\to\infty} \lim_{N\to\infty} \sum_{n=0}^{N} a_n (z_i - a)^n = \lim_{N\to\infty} \lim_{i\to\infty} \sum_{n=0}^{N} a_n (z_i - a)^n = \lim_{N\to\infty} \sum_{n=0}^{N} a_n \underbrace{\lim_{i\to\infty} (z_i - a)^n}_{=0 \text{ für } n\neq 0} = \lim_{N\to\infty} a_0 = a_0$$

Die Gleichheit der ersten zwei Terme ist zu zeigen. Der Satz sagt also etwas über die Vertauschung von Grenzprozessen aus.

Beweis zum Lemma

$\lim_{i\to\infty} z_i = a \Rightarrow \exists n_1 : |z_i - a| < \frac{\varrho}{2} \ \forall i > n_1$
Ferner sei M so gewählt, dass gilt: $M > \sum_{n=1}^{\infty} |a_n| \left(\frac{\varrho}{2}\right)^{n-1} = \frac{2}{\varrho} \sum_{n=1}^{\infty} |a_n| \left(\frac{\varrho}{2}\right)^{n}$.
Sei $\varepsilon > 0$ gegeben. Dann $\exists n_2 : |z_i - a| < \frac{\varepsilon}{M} \ \forall i > n_2$.
Setze $n_0 := \max\{n_1, n_2\}$. Dann gilt: $\forall i > n_0 : |f(z_i) - a_0| = \left|(z_i - a) \sum_{n=1}^{\infty} a_n (z_i - a)^{n-1}\right|$
$\leq |z_i - a| \cdot \sum_{n=1}^{\infty} |a_k| \cdot |z_i - a|^{n-1} \leq \frac{\varepsilon}{M} \cdot \sum_{n=1}^{\infty} |a_k| \left(\frac{\varrho}{2}\right)^{n-1} \leq \frac{\varepsilon}{M} \cdot M = \varepsilon$

■

Beweis zum Satz

Betrachte den Fall $z_i \to a$, also $z_i - a \to 0$.
Sei $a_n := c_n - d_n$ und die PR $\sum_n a_n (z-a)^n$ mit dem Konvergenzradius $\varrho = \min\{\varrho_1, \varrho_2\}$.
$\Rightarrow \sum_n a_n (z_i - a)^n = 0 \ \forall n$. Induktiv wird gezeigt: $a_n = 0 \ \forall n$.

Für $n=0$ ist $\lim\limits_{i\to\infty} \underbrace{\sum a_0 \left(z_i-a\right)^0}_{=0} = a_0 \Rightarrow a_0 = 0.$

Sei gezeigt: $a_0 = a_1 = \ldots = a_{k-1} = 0$. Dann gilt für $n=k$:

$\sum\limits_{n=k}^{\infty} a_n \left(z_i-a\right)^n = 0 = a_k \left(z_i-a\right)^k + a_{k+1}\left(z_i-a\right)^{k+1} + \ldots$

$\Rightarrow 0 = \left(z_i-a\right)^k \cdot \sum\limits_{n=0}^{\infty} a_{k+n}\left(z_i-a\right)^n \Rightarrow \Big($beachte: $0<|z_i-a| \Rightarrow$ Div. durch $\left(z_i-a\right)^k \neq 0\Big)$

$\Rightarrow 0 = \sum\limits_{n=0}^{\infty} a_{k+n}\left(z_i-a\right)^n \overset{(Lemma)}{\Rightarrow} 0 = \lim\limits_{i\to\infty} \sum\limits_{n=0}^{\infty} a_{k+n}\left(z_i-a\right)^n = (\text{0-tes Glied}) = a_k = 0$

■

4 Der metrische Raum, die Topologie des $\mathbb{R}^n$

4.1 Der Begriff des metrischen Raumes

Definition 4.1 Wiederholung aus 2.4

Sei $M \neq \varnothing$. Eine Abbildung $d : M \times M \to \mathbb{R}$ heißt **Metrik** von M, wenn $\forall x, y, z \in M$ gilt:

- (ME1): $d(a,b) \geq 0$, $d(a,b) = 0 \Leftrightarrow a = b$ (Definitheit)
- (ME2): $d(a,b) = d(b,a)$ (Symmetrie)
- (ME3): $d(a,b) \leq d(a,c) + d(c,b)$ (Dreiecksungleichung)

Das Paar (M,d) heißt dann **metrischer Raum**.
Sei (M,d) ein metrischer Raum und $A \subset M$, $A \neq \varnothing$. Dann ist $(A, d|_A)$ ein metrischer Raum und $d|_A$ heißt die von d in A **induzierte Metrik**.

Beispiel 4.1

Sei $M \neq \varnothing$ beliebig. Die **diskrete Metrik** d in M ist gegeben durch $d(x,y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$.
Beachte: In einer Menge können sehr viele verschiedene Metriken existieren.

Definition 4.2

Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\| : V \to \mathbb{R}$ heißt **Norm**, wenn $\forall x, y \in V, \lambda \in \mathbb{K}$ gilt:

- (N1): $\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0$ (Definitheit)
- (N2): $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$ (Homogenität)
- (N3): $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

Das Paar $(V, \|\|)$ heißt dann **normierter Raum**.

Lemma 4.3

Sei $(V, \|\|)$ ein normierter Raum. Dann wird durch $d(x,y) = \|x - y\|$ eine Metrik in V definiert. d heißt **durch die Norm induzierte Metrik**.

Beispiel 4.2 Zahlen, d.h. $M = \mathbb{C}$, $M = \mathbb{R}$, $M = \mathbb{Q}$, ...

Durch $d : (x,y) \mapsto |x - y|$ ist eine Metrik definiert. In $\mathbb{R}$ und $\mathbb{C}$ ist $x \mapsto |x|$ auch eine Norm.

Beispiel 4.3 $V = \mathbb{R}^n$ und $V = \mathbb{C}^n$ mit $n \in \mathbb{N}$

$V = \mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\} \Rightarrow \|x\| := \sqrt{x_1^2 + \ldots + x_n^2}$
$V = \mathbb{C}^n = \{(z_1, \ldots, z_n) : z_i \in \mathbb{C}\} \Rightarrow \|z\| := \sqrt{|z_1|^2 + \ldots + |z_n|^2}$
Beweis von (N1) und (N2) trivial, (N3) ist schwieriger und wird vertagt.
Diese Normen heißen auch **euklidische Normen** (sie kommen von einem Skalarprodukt, siehe eines der nächsten Kapitel).

Beispiel 4.4

Sei $B\,[a,b]$ die Menge aller reell- oder komplexwertigen beschränkten Funktionen auf $[a,b]$. Das heißt, f ist auf $[a,b]$ beschränkt $\Rightarrow \exists c > 0 : |f(x)| \leq c\ \forall x \in [a,b]$.
Dann ist durch $\|f\| := \sup\limits_{x \in [a,b]} |f(x)|$ eine Norm definiert. (Das Infimum kann nicht benutzt werden, da sonst (N1) nicht gelte. Das Nullelement von $B\,[a,b]$ ist hierbei $f_0 : [a,b] \to \mathbb{R},\ x \mapsto 0$.
Aufgrund der Rechenregeln für den Betrag kann man leicht zeigen, dass $\|f\|$ eine Norm ist.

Beispiel 4.5

Sei $l^\infty := \{(x_n) \in \mathbb{C} : (x_n) \text{ beschränkt}\}$. Dann ist $\|x\| := \sup\limits_n |x_n|$ eine Norm (Beweis ebenfalls einfach).

Beispiel 4.6 Matrizen

Sei $M_n(\mathbb{K})$ oder kurz M_n die Menge aller (n,n)-Matrizen mit Elementen aus $\mathbb{K}$, also $M_n(\mathbb{K}) = \mathbb{K}^{n \times n}$. Dann sind folgende Abbildungen Normen:

1. $\|A\|_1 := \max\limits_{i,j} |a_{ij}|$ - größtes Element
2. $\|A\|_2 := \sqrt{\sum\limits_{i,j=1}^{n} |a_{ij}|^2}$ - euklidische Norm
3. $\|A\|_3 := \max\limits_{1 \leq i \leq n} \sum\limits_{j=1}^{n} |a_{ij}|$ - größte Quersumme einer Zeile („Zeilensumme“)
4. $\|A\|_4 := \max\limits_{1 \leq j \leq n} \sum\limits_{i=1}^{n} |a_{ij}|$ - größte Quersumme einer Spalte („Spaltensumme“)

4.1.1 Konvergenz in metrischen Räumen

Definition 4.4

Sei (M,d) ein metrischer Raum.

1. Eine Folge $(x_n) \subset M$ **konvergiert** gegen $x \in M$, wenn gilt:
   $\forall \varepsilon > 0 : \exists n_0 = n_0(\varepsilon) \in \mathbb{N} : \forall n \geq n_0 : d(x, x_n) < \varepsilon$ (Verwende die üblichen Schreibweisen.)
2. Eine Folge $(x_n) \subset M$ heißt **Cauchy-Folge**, wenn gilt:
   $\forall \varepsilon > 0 : \exists n_0 = n_0(\varepsilon) \in \mathbb{N} : \forall m,n \geq n_0 : d(x_m, x_n) < \varepsilon$
3. (M,d) heißt **vollständig**, wenn jede Cauchy-Folge (x_n) aus M gegen ein $x \in M$ konvergiert.

Wie in $\mathbb{R}$ und $\mathbb{C}$ zeigt man:

Satz 4.5

Sei (M, d) ein metrischer Raum.

1. Der Grenzwert einer konvergenten Folge aus M ist eindeutig bestimmt.

2. Jede konvergente Folge aus M ist eine Cauchy-Folge.

Bemerkung

Nach Definition gilt die Umkehrung von Punkt 2 nur in vollständigen Räumen. Bisher kennen wir nur primitive Beispiel unvollständiger Räume, zum Beispiel $\mathbb{Q}$.
In einem normierten Raum $(V, \|\|)$ gilt $x_n \to x \Leftrightarrow \|x_n - x\| \to 0$ aufgrund der durch die Norm induzierten Metrik.

Definition 4.6

Eine Menge $B \subset (M, d)$ heißt **beschränkt**, wenn: $\forall a \in M : \exists K = K(a, B) > 0 : d(a, b) \leq K \ \forall b \in M$

Mit anderen Worten: B ist beschränkt, wenn für jeden Punkt a aus M die Menge B vollständig in (je nach Mengentyp) einem Intervall, Kreis, einer Kugel oder (im Allgemeinen) einer ε-Umgebung um a liegt.

Satz 4.7

1. $B \subset M$ ist genau dann beschränkt, wenn die Bedingung aus der Def. für ein festes $a_0 \in M$ erfüllt ist.

2. Sei $(V, \|\|)$ ein normierter Raum. Dann ist $B \subset V$ genau dann beschränkt, wenn gilt: $\exists C > 0 : \|x\| \leq C \ \forall x \in B \qquad (*)$

Beweis
Zu 1.: Von links nach rechts nach Definition erfüllt.
Zeigen, dass aus Beschränktheit für $a_0 \in M$ Beschränktheit $\forall a \in M$ folgt.
Sei $d(a_0, b) \leq K \ \forall b \in B$. Dann gilt für ein beliebiges $a \in M$ und $\forall b \in B$:
$d(a, b) \leq d(a, a_0) + \underbrace{d(a_0, b)}_{\leq K} \leq \underbrace{d(a, a_0) + K}_{=K(a,B)}$

■

Beweis
Zu 2.: Beachte $d(x, y) = \|x - y\|$, d.h. in $(*)$ steht: $d(x, 0) \equiv \|x\| \leq C \ \forall x \in B$
Nach 1. folgt mit $a_0 = 0 \in V$ die Behauptung.

■

Satz 4.8

Jede konvergente Folge (a_n) aus M ist beschränkt.

Beweis
Analog zu Zahlenfolgen: Sei $x_n \to x$. Setze (gemäß Satz 4.7, Punkt 1) $a_0 = x$. Für $\varepsilon = 1$ gilt: $\exists n_0 : \forall n \leq n_0 : d(x, x_n) < 1 \Rightarrow \forall n : d(x, x_n) \leq \max\{1, d(x, x_1), \ldots, d(x, x_{n_0 - 1})\} = $ konst.. Also ist (x_n) beschränkt.

■

Definition 4.9 Umgebungen

Sei (M, d) ein metrischer Raum und $a \in M$.

1. Eine Menge $U_\varepsilon(a) = \{x \in M : d(x, a) < \varepsilon\}$ heißt **ε-Umgebung** von a.
2. Eine Menge $U \subset M$ heißt **Umgebung** von a, wenn gilt: $\exists U_\varepsilon(a) \subset U$.

Bemerkung

Jede Obermenge einer Umgebung von a ist wieder Umgebung von a.
Eine Umgebung braucht weder zusammenhängend noch um a zentriert zu sein.

Mit Umgebungen werden sehr wichtige Typen von Mengen definiert.

Definition 4.10

Sei (M, d) ein metrischer Raum.

1. $O \subset M$ heißt **offen**, wenn $\forall x \in O$ eine Umgebung $U(x) \subset O$ existiert.
   $(\Leftrightarrow \forall x \in O : \exists \varepsilon = \varepsilon(x) > 0 : U_\varepsilon(x) \subset O)$
2. $A \subset M$ heißt **abgeschlossen**, wenn $M \setminus A$ offen ist.

Man setzt sinnvollerweise, dass die leere Menge offen ist (da „für alle Elemente von $\varnothing$“ die Bedingung für Elemente offener Mengen erfüllt ist). Achtung: Eine Menge kann sowohl offen als auch abgeschlossen sein.

Beispiel 4.7

Sei $M := [-1, 1] \cup (2, 3]$ und $X := [-1, 1]$.
X ist offen! (Achtung: Die ε-Umgebungen von -1 und 1 *in* M können nur in M liegen.)
Gleichzeitig ist X abgeschlossen, da $M \setminus X = (2, 3]$ offen ist.

Definition 4.11

Sei (M, d) ein metrischer Raum und $B \subset M$.

1. $x \in M$ heißt **Häufungspunkt** von B, wenn in jeder Umg. von x Punkte $y \in B$ liegen mit $y \neq x$.
2. $y \in B$ heißt **innerer Punkt** von B, wenn es eine Umgebung $U(y) \subset B$ gibt.
3. $z \in M$ heißt **Randpunkt** von B, wenn jede Umg. $U(z)$ Punkte aus B und aus $M \setminus B$ enthält.
4. $a \in B$ heißt **isolierter Punkt** von B, wenn a kein Häufungspunkt von B ist.
   Das heißt, es existiert eine Umgebung $U(a)$, sodass gilt: $U(a) \cap B = \{a\}$

Randpunkte und isolierte Punkte von B bilden zusammen die **Berührungspunkte** von B.

Beispiel 4.8

Sei $M = \mathbb{R}$ und $B = \{0\} \cup [1, 2]$.

- Häufungspunkte: $[1, 2]$
- innere Punkte: $(1, 2)$
- Randpunkte: $\{0, 1, 2\}$
- isolierte Punkte: $\{0\}$

Satz 4.12

Eine Teilmenge $A \subset M$ ist genau dann abgeschlossen, wenn jeder Häufungspunkt von A zu A gehört.

Beweis

Hinrichtung: Sei A abgeschlossen und $x \notin A$. Da A abgeschlossen ist, ist $M \setminus A$ offen, und da $x \in M \setminus A$, existiert ein $U(x) \subset M \setminus A$, also ist $U(x) \cap A$ leer und x kein Häufungspunkt von B.

Rückrichtung: Mögen alle Häufungspunkte von A zu A gehören. Ein Punkt $x \in M \setminus A$ ist kein Häufungspunkt, also existiert ein $U(x)$ mit $U(x) \cap A = \varnothing \Rightarrow U(x) \subset M \setminus A$. Da für jedes $x \in M \setminus A$ ein solches $U(x)$ existiert, ist $M \setminus A$ ist offen.

■

Bemerkung

Problem: Sind $\varnothing$ und M offen oder abgeschlossen? Wir wissen bereits: M ist offen (folgt aus Definition). Außerdem gehören alle Häufungspunkte von M zu M (da nur M betrachtet wird). Also ist M abgeschlossen. $\varnothing$ ist offen (folgt aus Definition, siehe oben). Da M offen ist, ist $\varnothing$ abgeschlossen. Aufgrund des obigen Satzes **setzt** man $\varnothing$ als offen und M als abgeschlossen.

Satz 4.13

Sei $A \subset M$, $x \in M$. Dann sind äquivalent:

1. x ist Häufungspunkt von A.
2. Jede Umgebung von x enthält unendlich viele Punkte von A.
3. $\exists\, (a_n) \subset A : a_n \to x, a_n \neq x\ \forall n$

Beweis

Aus 3. folgt 1.: In jeder Umgebung von x liegen unendlich viele Folgenglieder, also unendlich viele von x verschiedene Punkte von A. Also ist x ein Häufungspunkt von A.

Aus 1. folgen 2. und 3.: Sei x ein Häufungspunkt von A. Definieren induktiv eine Folge $0 < \varepsilon_n \to 0$.
Nun definieren wir induktiv eine Folge $a_n \to a$, $a_m \neq a_n\ \forall m \neq n$, $a_n \neq x\ \forall n$, $a_n \in U_{\varepsilon_n} \cap A$.
Sei nun $U(x)$ beliebig. Sei $\varepsilon_1 > 0$ so, dass $U_{\varepsilon_1}(x) \subset U(x)$.
Da x ein Häufungspunkt von A ist, $\exists a_1 \subset U_{\varepsilon_1}(x) \cap A$, $a_1 \neq x$.
Seien nun $\varepsilon_1 > \varepsilon_2 > \ldots > \varepsilon_n > 0$ und $a_1, \ldots, a_n$ gewählt wie oben gefordert.
Wir wählen ε_{n+1} und a_{n+1} wie folgt: Setze $\varepsilon_{n+1} := \frac{d(x,a_n)}{2}$, $a_{n+1} \in U_{\varepsilon_{n+1}} \cap A$, $a_{n+1} \neq x$.
Also gilt $\varepsilon_n \to 0$ und a_n erfüllt alle der geforderten Eigenschaften, insbesondere $a_n \to x$.
Dann sind 2. und zugleich 3. erfüllt, da für die unendliche Menge $\{a_n\} = \{a_i : i \in \mathbb{N}\}$ gilt:
$\{a_n\} \subset A$, $\{a_n\} \subset U_{\varepsilon_1}(x)$, $\{a_n\} \subset U(x)$

Aus 2. folgt 3.: Sei ε_n eine beliebige Folge mit $0 < \varepsilon_n \to 0$, $\varepsilon_1 > \varepsilon_2 > \ldots > 0$. In $U_{\varepsilon_1}(x)$ liegen unendlich viele Elemente aus A. Wähle $a_1 \in U_{\varepsilon_1}(x) \cap A$. Seien nun $a_1, \ldots, a_n$ gewählt, mit $a_i \neq x$, $a_i \neq a_j\ \forall i, j$.
$U_{\varepsilon_{n+1}}(x)$ enthält unendlich viele Elemente aus $A \Rightarrow \exists a_{n+1} \in U_{\varepsilon_{n+1}}(x) \cap A$, $a_{n+1} \neq x$, $a_{n+1} \neq a_i$ mit $1 \leq i \leq n$. Dann ist $a_n \to x$.

■

4.2 Die Topologie des $\mathbb{R}^n$

Das Wort „Topologie“ stammt von den topologischen Räumen, einer Obermenge der metrischen Räume.

In $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ wird durch $\langle x,y\rangle := \sum\limits_{k=1}^{n} x_k y_k$ bzw. $\langle x,y\rangle := \sum\limits_{k=1}^{n} \overline{x_k} y_k$ ein **Skalarprodukt** definiert, das sogenannte **Standardskalarprodukt**. Jedes Skalarprodukt (d.h. eine Abbildung $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{R}\,(\mathbb{C})$; V ist ein beliebiger Vektorraum) hat folgende definierenden Eigenschaften:

- (SP1): $\langle x,x\rangle \geq 0, = 0 \Leftrightarrow x = 0$ (positive Definitheit)
- (SP2): $\langle x,y\rangle = \overline{\langle y,x\rangle}$ (Symmetrie)
- (SP3): $\langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle = \alpha_1\langle x,y_1\rangle + \alpha_2\langle x,y_2\rangle$ (Linearität in der 2. Komponente)
  Aus (SP2) und (SP3) folgt die konjugierte Linearität in der 1. Komponente:
  $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \overline{\langle y, \alpha_1 x_1 + \alpha_2 x_2\rangle} = \overline{\alpha_1\langle y_1,x\rangle + \alpha_2\langle y_2,x\rangle} = \overline{\alpha_1}\langle x,y_1\rangle + \overline{\alpha_2}\langle x,y_2\rangle$

Bemerkung

Es ist eine Vereinbarung, ob man das Skalarprodukt in der 1. oder 2. Komponente als linear definiert. In der Physik wird die Linearität in der 2. Komponente gesetzt (so wie oben), in der Mathematik in der 1. Komponente. Für $\mathbb{R}^n$ ist das Skalarprodukt in beiden Komponenten linear, da $\overline{\alpha} = \alpha$.

Hat man ein Skalarprodukt $\langle\cdot,\cdot\rangle$, so wird durch $\|x\| := \langle x,x\rangle^{\frac{1}{2}}$ eine Norm definiert, die durch $\langle\cdot,\cdot\rangle$ induzierte Norm. In $\mathbb{R}^n$ ist also $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$, für $\mathbb{C}^n$ gilt $\|z\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$. Der Beweis der Eigenschaften (N1) und (N2) sind einfach, für die Dreiecksungleichung benötigt man einen fundamentalen Satz.

Satz 4.14 Cauchy-Schwarz'sche Ungleichung

Für jedes Skalarprodunkt (und die induzierte Norm) gilt: $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$.
Speziell gilt für $\mathbb{R}^n$ bzw. $\mathbb{C}^n$: $\left|\sum\limits_{i=1}^{n} \overline{x_i} y_i\right| \leq \sqrt{|x_1|^2 + \ldots + |x_n|^2}\cdot\sqrt{|y_1|^2 + \ldots + |y_n|^2}$

Für den Beweis siehe die Vorlesung über Lineare Algebra. Dann zeigt man leicht, dass $\|\|$ die Dreiecksungleichung erfüllt.

Beweis

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle = \langle x,x\rangle + \left(\langle x,y\rangle + \overline{\langle x,y\rangle}\right) + \langle y,y\rangle$$

$$\leq \|x\|^2 + 2\,|\langle x,y\rangle| + \|y\|^2 \overset{\text{(C.S.U.)}}{\leq} \|x\|^2 + 2\,\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

■

Für den nächsten Satz benötigen wir folgende Abschätzung in $\mathbb{R}^n$ bzw $\mathbb{R}^n$. Dabei sei $1 \leq i \leq n$ beliebig (bei (1) wurde jeder Summand durch das Maximum ersetzt).

$$|x_i - y_i| \leq \|x-y\| = \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2} \overset{(1)}{\leq} \sqrt{n\cdot \max_{1\leq k\leq n} |x_k - y_k|^2} = \sqrt{n}\cdot \max_{1\leq i\leq n} |x_k - y_k| \qquad (*)$$

Satz 4.15

Die Räume $\mathbb{R}^n$ und $\mathbb{C}^n$ sind bzgl. der induz. Metrik $d(x,y) = \|x-y\| = \sqrt{\sum\limits_{i=1}^{n} |x_i - y_i|^2}$ vollständig.

Beweis Hier aus Platzgründen nur für $\mathbb{R}^n$ (der Beweis für $\mathbb{C}^n$ läuft völlig analog).
Sei $\left(x^{(k)}\right)$ eine Cauchyfolge. $\left(x^{(k)}\right) = \left(x_1^{(k)}, \dots, x_n^{(k)}\right)$ und $\forall \varepsilon > 0 : \exists n_0 : \left\|x^{(k)} - x^{(l)}\right\| < \varepsilon \ \forall k, l > n_0$
Anwendung von $(*)$ (linke Seite): $\left|x_i^{(k)} - x_i^{(l)}\right| \le \left\|x^{(k)} - x^{(l)}\right\| < \varepsilon \ \forall k, l > n_0, 1 \le i \le n$
Also ist die Koordinatenfolge $\left(x_i^{(k)}\right)$ eine Cauchyfolge in $\mathbb{R}$ $\forall i = 1, \dots, n$.
Da $\mathbb{R}$ vollständig ist, $\exists x_i = \lim\limits_{k\to\infty} x_i^{(k)}$. Also ist jede Koordinatenfolge konvergent.
Sei $x = (x_1, \dots, x_n)$. Wir zeigen $x^{(k)} \to x$ in $\mathbb{R}^n$. Benutze $(*)$ (rechte Seite):
$0 \le \left\|x^{(k)} - x\right\| \le \sqrt{n} \underbrace{\max\limits_{1\le i\le n} \left|x_l^{(k)} - x_l\right|}_{\text{Maximum von } n \text{ Nullfolgen}} \to 0 \Rightarrow \left\|x^{(k)} - x\right\| \to 0 \Rightarrow x^{(k)} \to x$

■

Folgerung 4.16

Eine Folge $\left(x^{(k)}\right)$ aus dem $\mathbb{R}^n/\mathbb{C}^n$ konvergiert genau dann gegen $x = (x_1, \dots, x_n)$, wenn die Koordinatenfolgen $\left(x_i^{(k)}\right)$ gegen x_i konvergieren $(i = 1, \dots, n)$.

Unser nächstes Ziel ist der Satz von Bolzano-Weierstraß für $\mathbb{R}^n$.

Satz 4.17 Intervallschachtelung

Sei (I_n) eine Intervallfolge $I_n = [a_n, b_n]$ mit $I_{n+1} \subset I_n$ und $\underbrace{|b_n - a_n|}_{\text{Länge von } I_n} \to 0$. Es ex. genau ein $a \in \bigcap\limits_n I_n$.

Beweis
$I_1 \subset I_2 \subset I_3 \subset \dots \Rightarrow (a_n)$ ist monoton wachsend und (z.B. durch b_1) nach oben beschränkt, (b_n) ist monoton fallend und (z.B. durch a_1) nach unten beschränkt. Also sind beide Folgen konvergent, d.h. $\exists \lim a_n, \lim b_n$ und (da $b_n - a_n \to 0$), ist $\lim a_n = \lim b_n =: a$. Die Existenz folgt aus $a_n \le a \le b_n \ \forall n \Rightarrow a \in I_n \ \forall n$. Eindeutigkeit (Kurzform): Gebe es zwei Zahlen $a^{(1)}$ und $a^{(2)}$, die die Bedingung des Satzes erfüllen. Dann müsste die Länge aller Intervalle größer als $\left|a^{(1)} - a^{(2)}\right|$ sein, was aber ein Widerspruch zur Konvergenz ist.

■

Unter einem abgeschlossenen Quader $Q \in \mathbb{R}^n$ versteht man $Q := \{x = (x_1, \dots, x_n) : a_i \le x_i \le b_i,\ i = 1, \dots, n\}$.
Dabei sind die a_i und b_i fest vorgegeben. Das heißt mit $I_i := [a_i, b_i]$: $Q = I_1 \times \ldots \times I_n$.
Der Durchmesser des Quaders sei die Länge der Raumdiagonale: $d(Q) := \sqrt{\sum\limits_{i=1}^{n} (b_i - a_i)^2}$

Satz 4.18 Quaderschachtelung

Eine **Quaderschachtelung** ist eine Folge $Q^{(k)}$ von Quadern mit $Q^{(k+1)} \subset Q^{(k)} \ \forall k$ und $d\left(Q^{(k)}\right) \to 0$. Dann existiert genau ein $x^0 \in \mathbb{R}^n$ mit $x^0 \in \bigcap\limits_k Q^{(k)}$.

Beweis
Eindeutigkeit einfach. Zeigen Existenz. Aus den Voraussetzungen folgt:
Die $I_i^{(k)}$ mit $I_i^{(k)} = \left\{x \in \mathbb{R} : a_i^{(k)} \le x_i \le b_i^{(k)}\right\},\ i = 1, \dots, n$ bilden eine Intervallschachtelung.
Das heißt, es existiert genau ein x_i^0 mit $x_i^0 \in \bigcap\limits_k I_i^{(k)} \Rightarrow x^0 = \left(x_1^0, \dots, x_n^0\right) \in \bigcap\limits_k Q^{(k)}$.

■

Satz 4.19 Satz von Bolzano-Weierstraß in $\mathbb{R}^n$

Jede unendliche beschränkte Menge $B \subset \mathbb{R}^n$ besitzt einen Häufungspunkt.

Beweis

Benutze die sogenannte „Löwenfangmethode“ (der Einfachheit halber hier nur für $\mathbb{R}^2$).
B ist beschränkt, also existiert ein Quadrat $Q^{(1)}$ mit $B \subset Q^{(1)}$. Wir vierteilen nun $Q^{(1)}$ und erhalten vier gleichgroße Teilquadrate. Da B unendlich ist, existiert (mindestens) ein Teilquadrat, dass unendlich viele Elemente aus B enthält. Dieses sei $Q^{(2)}$. $Q^{(2)}$ wird wiederum geviertelt und so weiter.
Auf diese Weise erhalten wir eine Quadratschachtelung $Q^{(1)} \supset Q^{(2)} \supset Q^{(3)} \supset \ldots$ mit
$d\left(Q^{(k)}\right) = \frac{1}{2} d\left(Q^{(k-1)}\right) = \ldots = \frac{1}{2^{k-1}} d\left(Q^{(1)}\right) \to 0$.
Nach Satz 4.18 existiert genau ein $x^0 \in \bigcap\limits_k Q^{(k)}$. Zu zeigen: x^0 ist ein Häufungspunkt.
Sei $U_\varepsilon\left(x^0\right)$ die ε-Umgebung von x^0. Da $d\left(Q^{(k)}\right) \to 0$, existiert ein k_0 mit $Q^{(k_0)} \subset U_\varepsilon\left(x^0\right)$. Die linke Menge, und dadurch die rechte Menge, enthält unendlich viele Elemente aus B. Also ist x^0 Häufungspunkt von B.

■

Mit der gleichen Methode zeigt man:

Satz 4.20

Sei (b_n) eine beschränkte Folge aus $\mathbb{R}^n$. Dann besitzt (b_n) eine konvergente Teilfolge.

Bemerkung

Die Sätze 4.19 und 4.20 gelten auch in $\mathbb{C}^n$.

4.3 Kompaktheit

Definition 4.21

Sei (M, d) ein metrischer Raum.

- Unter einer **offenen Überdeckung** einer Menge $A \subset M$ versteht man eine Familie (unendlich große Menge von Mengen) $\underline{\mathcal{O}}$ von offenen Mengen $\mathcal{O}$ mit $A \subset \bigcup\limits_{\mathcal{O} \in \underline{\mathcal{O}}} \mathcal{O}$.
- Man sagt: $\underline{\mathcal{O}}$ enthält eine **endliche Teilüberdeckung** $\mathcal{O}_1, \ldots, \mathcal{O}_n \in \underline{\mathcal{O}}$, wenn gilt: $A \subset \bigcup\limits_{i=1}^{n} \mathcal{O}_i$

Definition 4.22

Sei (M, d) ein metrischer Raum. Eine Menge $K \subset M$ heißt

- **kompakt**, wenn *jede* offene Überdeckung von K eine endliche Teilüberdeckung enthält.
- **folgenkompakt**, wenn *jede* Folge $(x_n) \subset K$ eine Teilfolge (x_{n_k}) enthält mit $\lim\limits_{k\to\infty} x_{n_k} \in K$.

Bemerkung

In den Anwendungen bestehen die Teilmengen der offenen Überdeckungen häufig als Umgebungen (spezieller Art) der Punkte aus K.

Beispiel 4.9

$A = (0, 1] \subset \mathbb{R}$ ist nicht kompakt. Zum Beweis muss eine offene Überdeckung ohne endliche Teilüberdeckung $\supset A$ gefunden werden. Hier ist die Kreativität grenzenlos.
Sei $\underline{\mathcal{O}} = \left\{ \left(\frac{1}{2}, 2\right), \left(\frac{1}{4}, \frac{1}{2}\right), \left(\frac{1}{n}, 1\right) \text{ mit } n = 2, 3, 4, \ldots \right\}$. Diese offene Überdeckung von M enthält keine endliche Teilüberdeckung $\supset A$.

Beispiel 4.10

$B = [0, 1] \subset \mathbb{R}$ ist kompakt. Den Beweis führt man per Widerspruch: Angenommen, B sei nicht kompakt. Dann existiert eine offene Überdeckung $\underline{\mathcal{O}}$, die keine endliche Teilüberdeckung enthält. Halbiere B zu $[0, \frac{1}{2}]$ und $[\frac{1}{2}, 1]$, eines dieser Teilintervalle lässt sich nicht durch endlich viele Mengen aus $\underline{\mathcal{O}}$ überdecken, dieses sei I_1. I_1 wird wiederum halbiert und I_2 gewählt und so weiter. Es entsteht eine Intervallschachtelung (I_n) aus abgeschlossenen Intervallen mit $l(I_n) = 2^{-n}$ und kein I_n lässt sich durch endlich viele Mengen aus $\underline{\mathcal{O}}$ überdecken. Wegen Satz 4.17 existiert genau ein $x \in I_n\ \forall n$. Natürlich existiert ein $\mathcal{O} \in \underline{\mathcal{O}}$ mit $x \in \mathcal{O}$. Da $\mathcal{O}$ offen ist, existiert ein $\varepsilon > 0$ mit $(x - \varepsilon, x + \varepsilon) \subset \mathcal{O}$. Wähle n so groß, dass $2^{-n} < \varepsilon$, dann ist $I_n \subset (x - \varepsilon, x + \varepsilon) \subset \mathcal{O}$. Also wird I_n sogar von nur einer Menge aus $\mathcal{O}$ überdeckt: Widerspruch!

Beispiel 4.11

$A = (0, 1] \subset \mathbb{R}$ ist nicht folgenkompakt. Zum Beispiel hat $(a_n) = \left(\frac{1}{n}\right)$ keine in I konvergente Teilfolge.

Beispiel 4.12

$B = [0, 1] \subset \mathbb{R}$ ist folgenkompakt. Sei $(x_n) \subset [0, 1]$ beliebig. Dann (Satz B-WS) existiert eine konvergente Teilfolge (x_{n_k}), also $x_{n_k} \to x \in \mathbb{R} \Rightarrow x \in H(x_n) = [0, 1]$ (da $[0, 1]$ abgeschlossen).

Beispiel 4.13

$\mathbb{R}$ ist nicht kompakt (obgleich abgeschlossen). Betrachte die Mengenfamilie $\left(n - \frac{3}{4}, n + \frac{3}{4}\right)\ \forall n \in \mathbb{Z}$. Diese überdeckt $\mathbb{R}$, enthält jedoch keine endliche Teilüberdeckung von $\mathbb{R}$.

Satz 4.23

Jede kompakte Teilmenge K eines metrischen Raumes ist beschränkt und abgeschlossen.

Bemerkung

Achtung: Die Umkehrung gilt nicht (unbedingt).

Beweis
Beschränktheit:
Betrachte die offene Überdeckung $\mathcal{U} = \{U_1(x) : x \in K\} = \{\{y \in M : d(x, y) < 1\} : x \in K\}$ von K. K ist kompakt, also existieren endlich viele $x_1, \ldots, x_n$ mit $U_1(x_1) \cup \ldots \cup U_1(x_n) \supset K$.
Sei nun $y \in M$ fest und $c := \max\{d(y, x_1), \ldots, d(y, x_n)\}$. Eine Kugel um y mit dem Radius c umhüllt alle x_i. Eine Kugel um y mit einem Radius $r \geq c + 1$ umhüllt die 1-Umgebungen aller x_i und damit auch die Menge K, d.h. $U_r(y) \subset K \Rightarrow K$ ist beschränkt.

Abgeschlossenheit: Angenommen, es existiert ein Häufungspunkt $x \notin K$ von K. Zu jedem $y \in K$ existieren offene Umgebungen $U(y)$ (Umgebung um y), $V_y(x)$ (Umgebung um x) mit $U(y) \cap V_y(x) = \varnothing$. Die Menge aller $U(y)$ ist eine offene Überdeckung von K.
Da K kompakt ist, existiert eine endliche Teilüberdeckung, also existieren $y_1, \ldots, y_n$ mit $K \in \bigcup_{i=1}^{n} U(y_i)$.
Betrachte die zugehörigen Umg. $V_{y_i}(x)\ \ (1, \ldots, n) \Rightarrow V := \bigcap_{i=1}^{n} V_{y_i}(x)$ ist eine offene Umg. von x.

Es gilt: $V \cap U(y_i) = \varnothing \ \forall i = 1, \ldots, n \Rightarrow V \cap \underbrace{\left(\bigcup_{i=1}^{n} U(y_i) \right)}_{\supset K} = \varnothing \Rightarrow V \cap K = \varnothing \Rightarrow$ Widerspruch $\Rightarrow$ Jeder Häufungspunkt von K ist Element von K. $\Rightarrow K$ ist abgeschlossen.

■

Satz 4.24

Jede unendliche Teilmenge B eines kompakten metrischen Raumes (M, d) besitzt einen Häufungspunkt.

Beweis

Angenommen, B hat keinen Häufungspunkt, d.h. $x \in M$ existiert eine Umgebung $U(x)$ mit $U(x) \cap B = \varnothing$ oder $U(x) \cap B = \{x\}$. Dann ist $\{U(x) : x \in M\}$ eine offene Überdeckung von M. Da M kompakt ist, existiert eine endliche Teilüberdeckung $U(x_1), \ldots, U(x_m)$ von M.

$B = B \cap M = B \cap \left(\bigcup_{i=1}^{m} U(x_i) \right) = \bigcup_{i=1}^{m} \underbrace{(B \cap U(x_i))}_{\text{entweder } \varnothing \text{ oder } \{x_i\}} \subset \{x_1, \ldots, x_n\} \Rightarrow B$ ist endlich $\Rightarrow$ Widerspruch

■

Satz 4.25

Sei (M, d) ein metrischer Raum und $K \subset M$. Dann ist K genau dann kompakt, wenn K folgenkompakt ist. (Also fallen in metrischen Räumen beide Kompaktheitsbegriffe zusammen.)

Beweis

„Folgenkompaktheit $\Rightarrow$ Kompaktheit" ist schwierig (siehe Literatur). **Kompaktheit $\Rightarrow$ Folgenkompaktheit:** Sei $(x_n) \subset K$ (soll heißen: $x_n \in K \ \forall n \in \mathbb{N}$). Wenn $\{x_n\}$ endlich ist, ist es einfach, aus (x_n) eine in K konvergente Teilfolge auszuwählen, denn ein x_{n_0} muss in (x_n) unendlich oft vorkommen. Wähle als Teilfolge dann einfach die konstante Teilfolge (x_{n_k}) mit $x_{n_k} = x_{n_0} \ \forall k$.

Sei nun $\{x_n\}$ unendlich. Nach Satz 4.24, angewendet auf den komplexen Raum (K, d), existiert ein Häufungs-punkt x von (x_n). Damit existiert eine Folge (y_n) aus Elementen von $\{x_n\}$ mit $y_n \to x$. (Beachte: Die y sind irgendwelche der x.) Jetzt muss man sich noch überlegen (durch Betrachten der Indizes, wie man (y_n) als Teilfolge von (x_n) wählen kann. Da K abgeschlossen (als kompakte Teilmenge), muss $x \in K$ sein.

■

Die für uns wichtigste Charakterisierung kompakter Mengen aus $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ enthält der...

Satz 4.26

Eine Teilmenge $A \subset \mathbb{R}^n$ (bzw. $\mathbb{C}^n$) ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Beweis

Die Hin-Richtung folgt aus Satz 4.23, die Rück-Richtung der Äquivalenz ist der Satz von Bolzano-Weierstraß (unendlich beschränkt $\Rightarrow \exists$ Häufungspunkt $\overset{\text{(abgeschlossen)}}{\Rightarrow \ldots \Rightarrow}$ kompakt).

■

---

5 Stetigkeit

5.1 Begriff der Stetigkeit

Aus der Schule ist Stetigkeit wie folgt bekannt.

1. Sei $f : D(f) \to \mathbb{R}$ eine Funktion. f heißt in $x_0 \in D(f)$ stetig, wenn gilt:
   $\forall \varepsilon > 0 : \exists \delta = \delta(\varepsilon, x_0) : \forall x \in D(f), |x - x_0| < \delta : |f(x) - f(x_0)| < \varepsilon$
2. f heißt in $x_0 \in D(f)$ stetig, wenn gilt: $\forall (x_n) \subset D(f), x_n \to x_0 : f(x_n) \to f(x_0)$
3. „Grenzwert = Funktionswert“, d.h. „$\lim\limits_{x \to x_0} f(x) = f(x_0)$“ (basiert auf 1. und 2.)

Wir definieren Stetigkeit allgemein wie folgt:

Definition 5.1 Stetigkeit

Sei f eine Abbildung des metrischen Raumes (X, ϱ) in den metrischen Raum (Y, σ). f heißt **stetig** in $x_0 \in X$, wenn zu jeder Umgebung V von $f(x_0)$ eine Umgebung U von x_0 gibt mit $f(U) \subset V$. Kurz:

$$\forall V(f(x_0)) : \exists U(x_0) : f(U(x_0)) \subset V(f(x_0))$$

f heißt auf X **stetig**, wenn f in jedem Punkt von X stetig ist.

Bemerkung

1. Die Stetigkeit ist eine lokale Eigenschaft (Punkteigenschaft) einer Abbildung.
2. In Definition 5.1 ist automatisch $D(f) = X$. Wir werden später $D(f) \subsetneq X$ betrachten. (Dabei sind nur unwesentliche Modifikationen nötig.)
3. Denkt man an die Definition der Umgebung eines Punktes (d.h. W ist Umgebung von $z \Leftrightarrow W$ enthält eine gewisse Kugel um z), erhält man die folgende äquivalente Definition.

Satz 5.2

Eine Abbildung $f : X \to Y$ (X,Y wie in 5.1) ist genau dann in x_0 stetig, wenn gilt:
$\forall \varepsilon > 0 : \exists \delta = \delta(\varepsilon, x_0) > 0 : \forall x \in D(f) : [\varrho(x, x_0) < \delta \Rightarrow \sigma(f(x), f(x_0)) < \varepsilon]$

Satz 5.3

Eine Abbildung $f : X \to Y$ (X,Y wie in 5.1) ist genau dann in x_0 stetig, wenn $x_n \to x_0 \Rightarrow f(x_n) \to f(x_0)$
Das heißt: f in x_0 stetig $\Leftrightarrow [\varrho(x_0, x_n) \to 0 \Rightarrow \sigma(f(x_0), f(x_n)) \to 0]$

Beweis
Hin-Richtung: (Sei f im Sinne von 5.1 definiert.) Sei V eine *beliebige* Umgebung von $f(x_0)$ und U eine solche Umgebung von x_0, dass gilt: $f(U) \subset V$. Sei $x_n \to x_0 \Rightarrow \exists n_0 : x_n \in U \; \forall n > n_0 \Rightarrow f(x_n) \in V \; \forall n > n_0 \Rightarrow f(x_n) \to f(x_0)$
Rück-Richtung: Möge gelten: $x_n \to x_0 \Rightarrow f(x_n) \to f(x_0)$. Angenommen, f sei nicht stetig in x_0. Dann existiert eine Umgebung V von $f(x_0)$, für die für alle Umgebungen U von x_0 gilt: $f(U) \not\subset V$. Betrachte $\forall n$ die spezielle Umgebung $U_{\frac{1}{n}}(x_0) \Rightarrow f\left(U_{\frac{1}{n}}(x_0)\right) \not\subset V$, d.h. $\forall n : \exists x_n \in U_{\frac{1}{n}}(x_0) : f(x_n) \notin V$.

Also hat x_n die Eigenschaft $x_n \to x_0$, aber $f(x_n) \notin V\ \forall n \Rightarrow f(x_n) \not\to f(x_0) \Rightarrow$ Widerspruch

■

Satz 5.4

$f : X \to Y$ ist genau dann stetig, wenn das Urbild offener (abgeschl.) Mengen offen (abgeschl.) ist.

Bemerkung

Unter stetigen Abbildungen muss das Bild einer offenen (abgeschl.) Menge nicht offen (abgeschl.) sein.

Beispiel 5.1

Sei $y_0 \in Y$ fest. Betrachte $f(x) := y_0\ \forall x \in X$. Die Menge $\{y_0\}$ ist stets abgeschlossen, jedoch bilden auch offene Mengen $O \subset X$ in $\{x_0\}$ ab.

Wir werden nun einige Funktionen auf Stetigkeit untersuchen.

Beispiel 5.2

$f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist stetig für alle $x \in \mathbb{R}$ (analog in $\mathbb{C}$).
Beweis mit der Folgencharakterisierung: $x_n \to x_0 \Rightarrow x_n \cdot x_n \to x_0 \cdot x_0 \Leftrightarrow x_n^2 \to x_0^2$

Beispiel 5.3

$f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$ ist stetig für alle $x \in \mathbb{R} \setminus \{0\}$. Für $x = 0$: $(x_n) = \left(1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \dots\right) \to 0$, aber $(f(x_n)) = (1, -1, 1, -1, \dots)$

Beispiel 5.4 Dirichlet-Funktion

$f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$ ist in keinem Punkt stetig.
Sei $x_0 \in \mathbb{R}$ beliebig, wähle (z.B.) $\varepsilon = \frac{1}{2}$. Dann existiert kein δ mit $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \frac{1}{2}$, denn: Sei x_0 rational, dann liegt in jeder Umgebung $U_\delta(x_0)$ ein irrationales x mit $|f(x) - f(x_0)| = 1 \nleq \frac{1}{2}$. Analog für irrationale x_0.

Beispiel 5.5

Sei $f : \mathbb{R}^m \to \mathbb{R}^k$ mit $m, k \in \mathbb{N}$ beliebig. Diese Abbildung hat die Gestalt $f = (f_1, \dots, f_k)$ mit den Koordinatenfunktionen $f_i : \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, k$. Sei $x = (x_1, \dots, x_m)$ und $y = (y_1, \dots, y_k)$. Dann ist $f(x) = (f_1(x_1, \dots, x_m), \dots, f_k(x_1, \dots, x_m)) = (y_1, \dots, y_k)$.
Sei zum Beispiel $m = 2$, $k = 3$, $f(x_1, x_2) := \begin{pmatrix} x_1^2 + x_2 \\ \sin x_1 \cdot \tan x_2 \\ x_1^3 \cdot x_2^{97} \end{pmatrix} = (y_1, y_2, y_3)$. Stetigkeit in einem Punkt $x^{(0)} = (x_1^{(0)}, x_2^{(0)}) \in D(f)$ heißt: Für jede Folge $\left(x^{(n)}\right) \subset D(f)$ mit $x^{(n)} \to x^{(0)}$ gilt $f(x^{(n)}) \to f(x^{(0)})$. Ausführlich:
$x^{(n)} = \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ x_3^{(n)} \end{pmatrix} \to x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{pmatrix} \Leftrightarrow x_i^{(n)} \to x_i^{(0)},\ i = 1, \dots, n$ (koordinatenweise Konvergenz; Standardmetrik!).
Daraus muss folgen:
$f(x^{(n)}) = \begin{pmatrix} f_1(x_1^{(n)}, \dots, x_m^{(n)}) \\ \vdots \\ f_k(x_1^{(n)}, \dots, x_m^{(n)}) \end{pmatrix} \to f(x^{(0)}) = \begin{pmatrix} f_1(x_1^{(0)}, \dots, x_m^{(0)}) \\ \vdots \\ f_k(x_1^{(0)}, \dots, x_m^{(0)}) \end{pmatrix} \Leftrightarrow f(x_1^{(n)}, \dots, x_m^{(n)}) \to f(x_1^{(0)}, \dots, x_m^{(0)})$,
$i = 1, \dots, n$ (wiederum koordinatenweise Konvergenz).

Satz 5.5

Eine Abbildung $f : \mathbb{R}^m \to \mathbb{R}^k$ ist genau dann in $x^{(0)}$ stetig, wenn alle Koordinationenfunktionen $f_1, \dots, f_k$ in $x^{(0)}$ stetig sind.

Spezialfall: Lineare Abbildungen $f : \mathbb{R}^m \to \mathbb{R}^k$ sind immer $\forall x \in \mathbb{R}^m$ stetig.

Beispiel 5.6 Stetigkeit von Potenzreihen

Sei $f(z) := \sum\limits_{n=0}^{\infty} a_n \left|z-a\right|^n$, $z \in U_\varrho(a)$ (offener Konvergenzkreis)
Dann gilt: f ist stetig $\forall z \in U_\varrho(a) = D(f)$. Mit Lemma 3.49 ist ein Spezialfall schon gezeigt: Wenn $z_n \to a$, dann $f(z_n) \to f(a) = a_0$.
Der allgemeine Fall wird nicht hier bewiesen, sondern fällt später bei einem Fundamentalsatz ab (siehe gleichmäßige Konvergenz).

Bemerkung

1. Sei jetzt $f : X \to Y$, aber $D(f) \subsetneq X$. Was bedeutet Stetigkeit in $x_0 \in D(f)$?
Es ändert insofern nichts, als D(f) wieder ein metrischer Raum ist und $f : D(f) \to Y$ auf Stetigkeit untersucht wird. Wichtige Übung: Formulieren Sie den Stetigkeitsbegriff für diesen Fall. Zum Beispiel für Folgen: Sei $x_0 \in D(f)$ und $(x_n) \subset D(f)$. $x_n \to x_0 \Rightarrow f(x_n) \to f(x_0)$.
2. Beachte: Jedes f ist in jedem isolierten Punkt x_0 des Definitionsbereiches stetig, da die einzigen Folgen $(x_n) \subset D(f)$ mit $x_n \to x_0$ diejenigen sind, für die ab einem $N \in \mathbb{N}$ gilt: $x_n = x_0 \ \forall n \geq N$.
Bsp.: $f(x) = \begin{cases} 1 & x \in [1,2] \\ 2 & x \in \{3\} \end{cases}$

5.2 Grenzwerte von Funktionen und Abbildungen

Sei $f : D(f) \subset X \to Y$ stetig und $x_0 \notin D(f)$. Kann man $f(x_0)$ so zusätzlich definieren, dass die entstehende Abbildung in x_0 stetig wird. Genauer: Definiere eine Abbildung $\widetilde{f}$ so, dass sie stetig ist und gilt: $\widetilde{f}(x) = \begin{cases} f(x) & x \in D(f) \\ ? & x = x_0 \end{cases}$.

Im einfachsten Falle ist x_0 ein isolierter Punkt von $D(f) \cup \{x_0\}$, d.h. es existiert eine Umgebung $U(x_0)$ mit $U(x_0) \cap D(f) = \varnothing$. Setze $\widetilde{f}(x_0) = y_0 \in Y$ beliebig, da $\widetilde{f}$ wie bereits gesehen auf jeden Fall in x_0 stetig ist.

Schwieriger ist es, wenn x_0 Häufungspunkt von $D(f)$ ist. Dann muss man garantieren:
$\forall (x_n) \subset D(f), x_n \to x_0 : f(x_n) \to y_0$ für ein festes y_0.

Beispiel 5.7

1. $f(x) = \frac{\sin x}{x}$, $x_0 = 0$
2. $f(x) = \sin \frac{1}{x}$, $x_0 = 0$
3. $f(x,y) = \frac{xy}{x^2+y^2}$, $(x_0, y_0) = (0,0)$

Unabhängig von der obigen Fortsetzungsproblematik definieren wir:

Definition 5.6

Seien (X, ϱ) und (Y, σ) metrische Räume, $f : D(f) \subset X \to Y$ und x_0 ein Häufungspunkt von $D(f)$.
Man sagt, f hat in x_0 den **Grenzwert** y, wenn gilt: $\forall (x_n) \subset D(f), x_n \to x_0 : f(x_n) \to y$
Notation: $\lim\limits_{x \to x_0} f(x) = y$ oder ausführlich $\lim\limits_{\substack{x \to x_0 \\ x \in D(f)}} f(x) = y$

Beispiel 5.8

Sei $f(x) = \sin\frac{1}{x}$. Existiert $\lim\limits_{x\to 0} \sin\frac{1}{x}$? Durch geeignete Wahl von Folgen (x_n) mit $x_n \to x_0$ findet man zum Beispiel:

- $\left(f\left(x_n^{(1)}\right)\right) = (1,1,1,1,\dots) \to 1$
- $\left(f\left(x_n^{(2)}\right)\right) = (-1,-1,-1,-1,\dots) \to -1$
- $\left(f\left(x_n^{(3)}\right)\right) = (1,-1,1,-1,\dots)$ ist divergent

Wie man sofort sieht, existiert $\lim\limits_{x\to 0} \sin\frac{1}{x}$ nicht. Also ist f an $x=0$ nicht stetig ergänzbar.

Satz 5.7

Es liege die in Definition 5.6 beschriebene Situation vor. Dann gilt:

- Wenn der Grenzwert $\lim\limits_{x\to 0} f(x)$ existiert, ist er eindeutig bestimmt.
- $\lim\limits_{x\to x_0} f(x) = y \Leftrightarrow$ Für alle Umgebungen $V(y)$ existiert eine Umgebung $U(x_0)$ mit $f(x) \in V(y)\ \forall x \in U(x_0) \cap D(f)$
  Mit anderen Worten: $\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in D(f) : \varrho(x, x_0) < \delta \Rightarrow \sigma(f(x), f(y)) < \varepsilon$
- Sei $x_0 \in D(f)$ und $x_0 \in H\left(D(f)\right)$. Dann ist f in x_0 genau dann stetig, wenn $\lim\limits_{x\to x_0} f(x) = f(x_0)$.

Übung: Beweisen Sie diesen Satz (mit Blick auf die verschiedenen Formulierungen von Stetigkeit)!

Die eingangs formulierte Fortsetzungsproblematik hat folgende Lösung:

Sei $x_0 \notin D(f)$, $x_0 \in H\left(D(f)\right)$ und $\lim\limits_{x\to x_0} f(x) =: y$, dann ist $\widetilde{f}$ def. durch $\widetilde{f}(x) = \begin{cases} f(x) & x \in D(f) \\ y & x = x_0 \end{cases}$.

Beispiel 5.9

Sei $f(x) = \frac{\sin x}{x}$ und x_0. Mit $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$ ist $f(x) = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots$
Also ist $\lim\limits_{x\to 0} \frac{\sin x}{x} =$ (man beachte die Stetigkeit der Potenzreihe in $x_0 = 0$) $= 1 - 0 + 0 - \dots = 1$.
$\Rightarrow \widetilde{f}(x) = \begin{cases} f(x) & x \in D(f) \\ y & x = x_0 \end{cases}$

Beispiel 5.10

Sei $f(x,y) = \frac{xy}{x^2+y^2}$ und $(x_0, y_0) = (0,0)$. Betrachte die folgenden Nullfolgen:

$$a_n = (\frac{1}{n}, 0), b_n = (\frac{1}{n}, \frac{1}{n}) \Rightarrow f(a_n) = \frac{\frac{1}{n}\cdot 0}{\frac{1}{n}^2 + 0} = 0 \to 0 \wedge f(b_n) = \frac{\frac{1}{n}\cdot\frac{1}{n}}{\frac{1}{n}^2 + \frac{1}{n}^2} = \frac{1}{2} \to \frac{1}{2}$$

Da der Grenzwert $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ nicht existiert, kann f nicht stetig ergänzt werden.

5.3 Eigenschaften stetiger Funktionen und Abbildungen

Satz 5.8

Sei X ein metrischer Raum. Mit $C(X)$ (**algebraischer Teil**) bezeichnen wir die Menge aller stetigen Funktionen aus X mit Werten aus $\mathbb{K}$ (d.h. $\mathbb{R}$ oder $\mathbb{C}$).

1. $C(X)$ ist eine Algebra, d.h. $C(X)$ ist ein Vektorraum über $\mathbb{K}$ und ein Ring. Mit anderen Worten, für $f, g \in C(X)$ und $\alpha, \beta \in \mathbb{K}$ gilt: $\alpha f + \beta g \in C(X) \wedge f \cdot g \in C(X)$, es gelten Assoziativ- und Kommutativgesetze (auch für die Multiplikation) sowie das Distributivgesetz.
2. Wenn $\mathbb{K} = \mathbb{C}$, gilt $f \in C(X) \Leftrightarrow \overline{f} \in C(X)$. Dabei ist $\overline{f}(x) = \overline{f(x)}\ \forall x \in X$.
3. In allen Punkten $x \in X$ mit $g(x) \neq 0$ ist $\frac{f}{g}$ stetig.

Bemerkung

Dieser Satz gilt auch für die Menge aller Funktionen $f : Y \to \mathbb{K}$, wobei Y eine beliebige Menge ist. Stetigkeit wird dabei nicht verlangt (da i.A. keine Metrik vorhanden ist).

Beweis

Trivial: $x_n \to x\ (x_n, x \in X) \Rightarrow f(x_n) \to f(x), g(x_n) \to g(x) \xrightarrow[\text{Grenzwerte}]{\text{Rechenregeln}}$ Behauptung

■

Satz 5.9 Stetigkeit zusammengesetzter Funktionen

Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen und X, Y, Z metrische Räume. Ist f in x_0 und g in $f(x_0)$ stetig, dann ist $h := g \circ f$ in x_0 stetig.

Beweis

$$x_n \to x_0 \overset{(f \text{ stetig})}{\Rightarrow} f(x_n) \to f(x_0) \overset{(g \text{ stetig})}{\Rightarrow} g(f(x_n) \to g(f(x_0)) \Leftrightarrow h(x_n) \to h(x_0)$$

■

Satz 5.10

Sei X ein metrischer Raum, $f : X \to \mathbb{R}$ stetig und $f(x_0) > 0$. Dann existiert eine Umgebung $U(x_0)$ mit $f(x) > 0\ \forall x \in U(x_0)$.

Beweis

Da $f(x_0) > 0$, existiert eine Umgebung $V(f(x_0))$ mit $y > 0\ \forall y \in V(f(x_0))$. Aufgrund der Stetigkeit von f existiert eine Umgebung $U(x_0)$ mit $f(U(x_0)) \subset V(f(x_0))$. Diese erfüllt die geforderte Bedingung.

■

Eventuell aus der Schule bekannt: Sei f stetig auf $[a, b]$.

1. f beschränkt
2. f hat ein Minimum und ein Maximum (d.h. Inf. und Supr. werden angenommen)
3. f nimmt jeden Wert zwischen Minimum und Maximum an
   (oder: $f(a) < 0, f(b) > 0 \Rightarrow \exists c \in (a, b) : f(c) = 0$)

Wir werden die Frage beantworten, welche Eigenschaften von $[a,b]$ diese Beziehungen (für stetige f) implizieren.

Satz 5.11 Stetiges Bild kompakter Mengen

Seien X,Y metrische Räume, $f : X \to Y$ stetig und $K \subset X$ kompakt. Dann ist $f(K) \subset Y$ kompakt. Kurz: Jedes stetige Bild kompakter Mengen ist kompakt.

Beweis
Sei $(y_n) \subset f(K)$. Zeigen: $\exists (y_{n_k}) : y_{n_k} \to y \in f(K)$. Es ist $y_n =: f(x_n)$ mit $x_n \in K$ Da K kompakt ist, $\exists (x_{n_k})$ mit $x_{n_k} \to x \in K$. Da f stetig ist, gilt: $f(x_{n_k}) \to f(x) \in f(K)$. $f(x_{n_k}) = y_{n_k}$

■

Daraus folgen unmittelbar die nächsten beiden Sätze.

Satz 5.12

Sei $f : X \to Y$ stetig. Wenn $K \subset X$ kompakt ist, ist $f(K)$ beschränkt. Spezialfälle:

1. Sei $f : K \to \mathbb{C}$ stetig. K kompakt $\Rightarrow f$ beschränkt
2. $f : [a,b] \to \mathbb{R}$ oder $\mathbb{C}$ stetig $\Rightarrow f$ beschränkt

Bemerkung

$[a,b]$ ist kompakt. Die Beschränktheit eines Intervalls reicht nicht.
Beispiel: $f : (0,1) \to \mathbb{R}, x \mapsto \frac{1}{x}$ ist stetig, aber unbeschränkt.

Satz 5.13

Sei $K \subset X$ kompakt und $f : X \to Y$ stetig. Dann hat f auf K ein Minimum und ein Maximum, das heißt:
$\exists x_0, x_1 \in K : f(x_0) = \sup\limits_{x\in K} f(x) = \max\limits_{x\in K} f(x) \wedge f(x_1) = \inf\limits_{x\in K} f(x) = \min\limits_{x\in K} f(x)$

Beweis
$f(K)$ ist abgeschlossen und beschränkt (weil kompakt $\subset \mathbb{R}$), also $\exists \sup\limits_{x\in K} f(x) = y_0$.
Da $f(K)$ abgeschlossen ist, ist $y_0 \in f(K) \Rightarrow \exists x_0 \in K : f(x_0) = y_0$.
Analog für Infimum und Minimum.

■

Bemerkung

Maximum und Maximum müssen natürlich nicht eindeutig sein, d.h. (z.B. für das Maximum) es können mehrere
$x_0^{(1)}, \dots, x_0^{(n)} \in K$ existieren mit $f(x_0^{(1)}) = \ldots = f(x_0^{(n)}) = y_0$.
Wenn f beschränkt, aber K nicht kompakt ist, muss kein Maximum oder Minimum existieren.

Beispiel 5.11

$f : (0,1) \subset \mathbb{R}, x \mapsto 2x \Rightarrow f$ hat weder Maximum noch Minimum: $f((0,1)) = (0,2)$
Aber: $\inf\limits_{x\in(0,1)} 2x = 0$ und $\sup\limits_{x\in(0,1)} 2x = 2$.

Definition 5.14

Seien (X,ϱ) und (Y,σ) metr. Räume. Eine Abbildung $f : X \to Y$ heißt auf X **gleichmäßig stetig**, wenn gilt: $\forall \varepsilon > 0 : \exists \delta > 0 : \forall x_1, x_2 \in X : \varrho(x_1,x_2) < \delta \Rightarrow \sigma(f(x_1), f(x_2)) < \varepsilon$

Bemerkung

Aus gleichmäßiger Stetigkeit folgt Stetigkeit. Also ist die gleichmäßige Stetigkeit ein stärkerer Begriff.

Beispiel 5.12

Sei $f(x) = \frac{1}{x}$ auf $[a,b]$ mit $a > 0$ definiert. Dann ist f gleichmäßig stetig, denn: Sei $\varepsilon > 0$ gegeben.
$|f(x_1) - f(x_2)| = \left|\frac{1}{x_1} - \frac{1}{x_2}\right| = \frac{|x_1-x_2|}{x_1x_2} = \frac{|x_1-x_2|}{a^2} \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$, falls $|x_1 - x_2| < \delta$ und $\delta = a^2\varepsilon$.

Beispiel 5.13

Betrachte das gleiche f auf $(0,1)$ oder $(0,\infty)$. Dann ist f nicht gleichmäßig stetig.
Angenommen, F wäre gleichmäßig stetig. Wähle $\varepsilon = 1$. Angenommen, $\exists \delta > 0 : |x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < 1$.
Wähle $x_1 = \frac{\delta}{1+\delta}$ und $x_2 = \frac{\delta}{2(1+\delta)}$. Dann gilt $|x_1 - x_2| < \delta$, aber $|f(x_1) - f(x_2)| = \left|\frac{1+\delta-2(1+\delta)}{\delta}\right| = \frac{1+\delta}{\delta} > 1$.

Satz 5.15

Jede auf einer kompakten Menge (oder einem kompakten metrischen Raum) K definierte stetige Abbildung $f : K \to Y$ ist auf K gleichmäßig stetig.

Beweis

Sei $\varepsilon > 0$ gegeben: $\forall x \in K : \exists \delta_x > 0 : f(U_{\delta_x}(x)) \subset U_{\frac{\varepsilon}{2}}(f(x))$. Dann ist $\left\{U_{\frac{\delta_x}{2}}(x) : x \in K\right\}$ eine offene Überdeckung von K. Da K kompakt ist, existieren $x_1, \ldots, x_n \in K$, sodass die Mengen $U_{\frac{\delta_{x_1}}{2}}(x_1), \ldots, U_{\frac{\delta_{x_n}}{2}}(x_n)$ K bereits überdecken. Setze $\delta := \min\left\{\frac{1}{2}\delta_{x_1}, \ldots, \frac{1}{2}\delta_{x_n}\right\}$. Betrachte nun beliebige $x \in K$ sowie $y \in U_\delta(x)$.
Dann ist für ein k: $x \in U_{\frac{1}{2}\delta_{x_k}}(x_k)$. Mithin gilt: $\varrho(y, x_k) \leq \varrho(y,x) + \varrho(x, x_k) < \delta + \frac{1}{2}\delta_{x_k} \leq \delta_{x_k}$, also $y \in U_{\delta_{x_k}}(x_k)$. Somit gilt $U_\delta(x) \subset U_{\delta_{x_k}}(x_k)$, insbesondere $x \in U_{\delta_{x_k}}(x_k)$. Das heißt: $\sigma(f(x), f(x_k)) < \frac{\varepsilon}{2}$, also $f(U_\delta(x)) \subset f(U_{\delta_{x_k}}(x_k)) \subset U_{\frac{\varepsilon}{2}}(f(x_k)) \subset U_\varepsilon(f(x))$.

■

Wir haben bereits zwei der drei bekannten Eigenschaften auf die Kompaktheit zurückführen können. Die dritte Eigenschaft hängt nicht mit der Kompaktheit von $[a,b]$ zusammen!

Beispiel 5.14

Sei $K = [1,2] \cup [3,4]$ und $f(x) = \begin{cases} -1 & x \in [1,2] \\ 1 & x \in [3,4] \end{cases}$
K ist kompakt und f stetig auf K, jedoch nimmt f keine Werte zwischen -1 und 1 an.

Die entscheidende Eigenschaft ist der *Zusammenhang* von $D(f)$ (siehe 12. Übungsaufgabe).

Satz 5.16

Sei f auf $[a,b]$ stetig, $f(a) < 0$ und $f(b) > 0$ (oder umgekehrt). Dann existiert (mindestens) ein $c \in (a,b)$ mit $f(c) = 0$.

Beweis

Sei $M := \{x_1 \in [a,b] : f(x) < 0 \forall x \in [a, x_1]\}$. M ist durch b nach oben beschränkt und $M \neq \varnothing$. Also $\exists \sup M =: c$. Zeigen: $f(c) = 0$.
Angenommen, $f(c) < 0$. Nach Satz 5.10 $\exists \delta > 0$, sodass f im gesamten Intervall $(c-\delta, c+\delta)$ noch < 0 ist. Also ist f in $[a, c+\delta)$ negativ, also auch in $\left[a, c+\frac{\delta}{2}\right]$, d.h. $c + \frac{\delta}{2} \in M$
$\Rightarrow$ Widerspruch zu $c = \sup M$! Analog für $f(c) > 0$.

■

Satz 5.17 Zwischenwertsatz

Sei f auf $I = [a,b]$ stetig (und reellwertig), sei $g := \min\limits_{x\in I} f(x)$ und $G := \max\limits_{x\in I} f(x)$.
Dann gilt: $\forall h \in [g,G] : \exists c = c(h) : f(c) = h$. (Mit anderen Worten: $[g,G] \subset W(f)$.)

Beweis
Nach 5.13 $\exists a_1, b_1 \in I : f(a_1) = g, f(b_1) = G$. Betrachten $\overline{f} = f|_{[a_1,b_1]}$ (o.E.d.A. $a_1 < b_1$).
Betrachte nun g mit $g(x) = \overline{f}(x) - h$. Dann ist g auf $[a_1, b_1]$ stetig.
$\Rightarrow g(a_1) = g - h < 0, g(b_1) = G - h > 0$ für $h \in (g,G)$.
Nach 5.16 existiert ein $c \in (a_1, b_1)$ mit $g(c) = 0 \Rightarrow f(c) = h$.

■

Satz 5.18 Stetigkeit der Umkehrfunktion

Sei f auf $[a,b]$ streng monoton wachsend und stetig. Dann ist f^{-1} in $[f(a), f(b)]$ streng monoton wachsend und stetig. (Analog für streng monoton fallende f, dann mit $[f(b), f(a)]$.)

Beweis
Da f streng monoton wachsend ist, ist f injektiv, $f(a) < f(b)$, $f(a), f(b) \in W(f)$, $f(a) = \min f$ und $f(b) = \max f$. Also ist nach 5.17 $W(f) = [f(a), f(b)] \Rightarrow f$ ist bijektiv von $[a,b]$ auf $[f(a), f(b)]$.
Betrachte nur $y_0 \in (f(a), f(b))$. (Randpunkte dann analog.) Sei $x_0 \in (a,b)$ mit $f(x_0) = y_0$. Sei $(y_n) \subset [f(a), f(b)]$ und $y_n \to y_0$. Zeigen: $\underbrace{f^{-1}(y_n)}_{=:x_n} \to f^{-1}(y_0) = x_0$. *Angenommen*, $x_n \not\to x_0$.
Das heißt, $\exists \varepsilon > 0$, sodass unendlich viele $x_{n_k} \notin (x_0 - \varepsilon, x_0 + \varepsilon)$. O.E.d.A. $x_{n_k} < x_0 - \varepsilon$. Wegen der Monotonie gilt: $y_{n_k} = f(x_{n_k}) < f(x_0 - \varepsilon) < f(x_0) = y_0$. Widerspruch: $y_{n_k} \to y_0$, aber $y_{n_k} \le \underbrace{f(x_0) - \varepsilon}_{\text{Schranke}} < y_0$.

■

5.4 Erweiterung des Stetigkeits- und Grenzwertbegriffes

Definition 5.19

1. f heißt im Punkt $x_0 \in D(f)$ $(f : D(f) \subset \mathbb{R} \to \mathbb{R})$
   - **rechtsseitig stetig**, wenn gilt: $\forall (x_n) \subset D(f), x_n > x_0$ und $x_n \to x_0 : f(x_n) \to f(x_0)$
   - **linksseitig stetig**, wenn gilt: $\forall (x_n) \subset D(f), x_n < x_0$ und $x_n \to x_0 : f(x_n) \to f(x_0)$
2.
   - Sei x_0 ein Häufungspunkt der Menge $M_r(x_0) := \{x \in D(f) : x > x_0\}$. f hat in x_0 den **rechtsseitigen Grenzwert** A, wenn gilt: $\forall (x_n) \subset M_r(x_0), x_n \to x_0 : f(x_n) \to A$.
   - Sei x_0 ein Häufungspunkt der Menge $M_l(x_0) := \{x \in D(f) : x < x_0\}$. f hat in x_0 den **linksseitigen Grenzwert** a, wenn gilt: $\forall (x_n) \subset M_l(x_0), x_n \to x_0 : f(x_n) \to a$.

   Man schreibt: $\lim\limits_{x\to x_0 - 0} f(x) = a$, $\lim\limits_{x \to x_0 + 0} f(x) = A$.

Bemerkung

Man schreibt $\lim\limits_{x\to x_0-0} f(x) = f(x_0)$ bei linksseitiger und $\lim\limits_{x\to x_0+0} f(x) = f(x_0)$ bei rechtsseitiger Stetigkeit.
Normalerweise ist $D(f)$ ein Intervall und $M_l(x)$ sowie $M_r(x)$ damit überflüssig.

Wir definieren zwei Prototypen von unendlichen Grenzwerten (alle anderen Varianten analog):

- $\lim\limits_{x \to x_0} f(x) = \infty \Leftrightarrow \forall (x_n) \subset D(f), x_n \neq x_0, x_n \to x_0 : f(x_n) \to \infty$
- $D(f)$ sei nach oben unbeschränkt: $\lim\limits_{x \to \infty} f(x) = A \Leftrightarrow \forall (x_n) \subset D(f), x_n \to \infty : f(x_n) \to A$

Definition 5.20 Klassifikation der Unstetigkeitsstellen

f sei in einem Häufungspunkt von $x_0 \in D(f)$ unstetig.

1. Wenn $\lim\limits_{x \to x_0} f(x) = A$ existiert, nennt man x_0 eine **hebbare Unstetigkeit**. Dann kann man f in x_0 so (re)definieren $\Rightarrow f(x_0) := A$, sodass eine an x_0 stetige Funktion entsteht.
   Wenn $\exists \lim\limits_{x \to x_0 - 0} f(x) = A_l$, $\lim\limits_{x \to x_0 + 0} f(x) = A_r$ mit $A_l \neq A_r$, dann sagt man, f hat an x_0 einen **Sprung** (der Höhe $|A_l - A_r|$).
   Hebbare Unstetigkeiten und Sprungstellen heißen **Unstetigkeiten erster Art**.
2. Wenn $\lim\limits_{x \to x_0} |f(x)| = \infty$ an der Unstetigkeitsstelle x_0 ist oder ein einseitiger Grenzwert von x_0 nicht existiert, dann heißt x_0 **Unstetigkeit zweiter Art**.

Man möchte häufig das Wachstum zweier Funktionen in einem Punkt oder im Unendlichen vergleichen.

Beispiel 5.15

1. $\sqrt{x^3+1}$ geht schneller $\to \infty$ als $x+4$
2. x^3 geht für $x \to 0$ schneller $\to 0$ als $3x$
3. $\log x$ geht für $x \to \infty$ langsamer $\to \infty$ als x^c $\forall c > 0$.
4. $\sqrt{x^2+5} \approx x$ für $x \to \infty$
5. $\sin x \approx x$ für $x \to 0$

Diese Situationen werden durch die **Landau-Symbole** o („kleines Oh“) und O („großes Oh“) exakt erfasst.

Definition 5.21 Landau-Symbole

- Gilt für Funktionen f, g: $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} = 0$ bzw. $\lim\limits_{x \to \infty} \frac{f(x)}{g(x)} = 0$ bzw. $\lim\limits_{x \to -\infty} \frac{f(x)}{g(x)} = 0$, so schreibt man $f(x) = o(g(x))$ für $x \to x_0$ bzw. $x \to \infty$ bzw. $x \to -\infty$.
- Gilt für Funktionen f, g: $\left|\frac{f(x)}{g(x)}\right| \leq K$ in einer (kleinen) Umgebung von x_0 bzw. für $x > C$ bzw. für $x < -C$ für ein gewisses $C > 0$, so schreibt man $f(x) = O(g(x))$ für $x \to x_0$ bzw. $x \to \infty$ bzw. $x \to -\infty$.

Dabei sollen f, g immer dort definiert sein, wo die entsprechenden Ausdrücke betrachtet werden.

Beispiel 5.16

1. $x+4 = o(\sqrt{x^3+1})$ für $x \to \infty$ (d.h. $\frac{x+4}{\sqrt{x^3+1}} \to 0$ für $x \to \infty$)
2. $x^3 = o(3x)$ für $x \to 0$ (d.h. $\frac{x^3}{3x} = \frac{1}{3}x^2 \to 0$ für $x \to 0$)
3. $\log x = o(x^c)$ für $x \to \infty, c > 0$
4. $\sqrt{x^2+5} = O(x)$ für $x \to \infty$ (d.h. $\frac{\sqrt{x^2+5}}{x} = \sqrt{1+\frac{5}{x^2}} \to 1 \Leftrightarrow \left|\sqrt{1+\frac{5}{x^2}}\right| \leq K$, $x > C$)
5. $\sin x = O(x)$ für $x \to 0$ (d.h. $\frac{\sin x}{x}$ in der Nähe von $x = 0$ beschränkt)
6. $\sin x = O(1)$ für $x \to \infty$ (d.h. $\sin x$ auf $\mathbb{R}$ beschränkt)

6 Differentialrechnung

6.1 Der Begriff der Ableitung

Definition 6.1

Wir betrachten $f : D(f) \subset \mathbb{R} \to \mathbb{R}$.

1. f heißt in $x_0 \in D(f)$ **differenzierbar**, wenn gilt:
$$\exists \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{f(x)-f(x_0)}{x-x_0} =: f'(x_0) \qquad (6.1)$$
Man nennt $f'(x_0)$ **Ableitung** von f in x_0 und schreibt auch $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$. Wenn f $\forall x_0 \in A \subset D(f)$ differenzierbar ist, so heißt f in A differenzierbar.
2. Wenn $\lim\limits_{x\to x_0-0} \frac{f(x)-f(x_0)}{x-x_0} =: f'_l(x_0)$ bzw. $\lim\limits_{x\to x_0+0} \frac{f(x)-f(x_0)}{x-x_0} =: f'_r(x_0)$ existiert, heißt f an x_0 **linksseitig** bzw. **rechtsseitig differenzierbar**.

Bemerkung

Die Differenzierbarkeit ist wie die Stetigkeit eine lokale Eigenschaft.
f ist differenzierbar $\Leftrightarrow$ f ist links- und rechtsseitig differenzierbar und $f'_l(x_0) = f'_r(x_0)\,(=: f'(x_0))$

Beispiel 6.1

Sei $f(x) = x^n \Rightarrow f'(x_0) = nx_0^{n-1}$. Beweis: $\frac{f(x_0+h)-f(x_0)}{h} = \frac{(x_0+h)^n - x_0^n}{h} = \frac{\binom{n}{1}x_0^{n-1}\cdot h + \binom{n}{2}x_0^{n-2}\cdot h^2 + \ldots + \binom{n}{n}h^n}{h}$
$= nx_0^{n-1} + \underbrace{\binom{n}{2} x_0^{n-2}\cdot h + \ldots + \binom{n}{n} h^{n-1}}_{\to 0} \to nx_0^{n-1}$

Beispiel 6.2

Sei $f(x) = |x|$. Für alle $x \neq 0$ ist f natürlich differenzierbar, für $x = 0$ jedoch nicht:
$\frac{f(0+h)-f(0)}{h} = \frac{|h|}{h} \to \begin{cases} 1 & h \to 0+0 \\ -1 & h \to 0-0 \end{cases}$
$\Rightarrow f'_r(0) = 1, f'_l(0) = -1 \Rightarrow f'(0)$ existiert nicht.

Im Hinblick auf Ableitungen von $f : \mathbb{R}^n \to \mathbb{R}^m$ ist folgende Umformulierung der Differenzierbarkeit nützlich:

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0) \Rightarrow \frac{f(x)-f(x_0)-f'(x_0)\cdot(x-x_0)}{x-x_0} \to 0 \text{ für } x \to x_0$$

Satz 6.2 Weierstraß'sche Zerlegungsformel

Eine Funktion f ist in x_0 genau dann differenzierbar, wenn eine Zahl $a = a(x_0)$ existiert mit:

$$f(x) = f(x_0) + a \cdot (x - x_0) + r(x, x_0) \qquad (6.2)$$

mit $r(x, x_0) = o(x - x_0)$, d.h. $\lim\limits_{x \to x_0} \frac{r(x,x_0)}{x-x_0} = 0 \qquad (6.3)$.

a ist durch diese Bedingung eindeutig bestimmt: $a = f'(x_0)$.

Beweis

Hin-Richtung (f stetig $\Rightarrow$ (6.2, 6.3)): Siehe oben mit $r(x, x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0)$.
Rück-Richtung: Wenn (6.2, 6.3) gelten, dann: $\frac{f(x)-f(x_0)}{x-x_0} = a + \frac{r(x,x_0)}{x-x_0} \to a$ für $x \to x_0$

■

- Die durch die Gleichung $y = f(x_0)+a(x-x_0)$ bestimmte Gerade heißt **Tangente** an den Graphen von f im Punkt $(x_0, f(x_0))$.
- Ein wesentlicher Gesichtspunkt bei (6.2) besteht darin, a nicht als Zahl anzusehen, sondern als lineare Abbildung von $\mathbb{R}$ nach $\mathbb{R}$ zu betrachten.

Satz 6.3 Differenziationsregeln

Seien f, g in x_0 differenzierbar, $\lambda, \mu \in \mathbb{R}$ (bzw. $\mathbb{C}$).

1. $\lambda f + \mu g$ ist in x_0 differenzierbar, $(\lambda f + \mu g)'(x_0) = \lambda f'(x_0) + \mu g'(x_0)$
2. Produktregel: $f \cdot g$ ist in x_0 differenzierbar, $(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$
3. Quotientenregel: $\frac{f}{g}$ ist in x_0 differenzierbar $\Leftrightarrow g(x_0) \neq 0$, $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)\cdot g(x_0) - f(x_0)\cdot g'(x_0)}{g(x_0)^2}$
4. Kettenregel: Ist f in x_0 und F in $y_0 = f(x_0)$ differenzierbar, dann ist $F \circ f$ in x_0 differenzierbar, $(F \circ f)'(x_0) = F'(y_0) \cdot f'(x_0) = F'(f(x_0)) \cdot f'(x_0)$.
5. Ableitung der Umkehrfunktion: Sei f stetig und streng monoton in $I = (a, b)$, d.h. f^{-1}existiert, ist stetig und monoton. Wenn für ein $x_0 \in I$ die Ableitung $f'(x_0)$ existiert und $f'(x_0) \neq 0$ ist, dann ist f^{-1} in $y_0 = f(x_0)$ differenzierbar und es gilt: $(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}$

Bemerkung

Diesen Satz wollen wir hier nicht beweisen. Die Beweise sind zum Teil Übungsaufgaben und auf jeden Fall aus der Schule bekannt.
Aus den Punkten 1 und 2 folgt: Die auf (a, b) differenzierbaren Funktionen bilden einen Vektorraum und eine Algebra.

Beispiel 6.3 Kettenregel

Sei $F(y) = \sin y$ und $f(x) = x^3$, also $(F \circ f)(x) = \sin x^3$. Dann ist $(F \circ f)'(x) = (\sin x^3)' = \cos x^3 \cdot 3x^2$.

Beispiel 6.4 Umkehrfunktion

Sei $f(x) = \mathrm{e}^x$, also $f^{-1}(y) = \log y$. Da $f'(x) = \mathrm{e}^x$, ist $(\log y)' = (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{\mathrm{e}^{\log y}} = \frac{1}{y}$.

Vorsicht: Die Formel für die Ableitung der Umkehrfunktion könnte man wie folgt herleiten.

$$y = f(f^{-1}(y)) \xrightarrow{\text{Kettenregel}} 1 = f'(f^{-1}(y)) \cdot (f^{-1})'(y) \Rightarrow (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Dabei geht man aber fälschlicherweise von der Differenzierbarkeit von f^{-1} aus. Trotzdem eignet sich diese Beziehung als Merkhilfe.

Wir definieren zusätzlich zur ersten Ableitung höhere Ableitungen: Wenn f in $U(x_0)$ differenzierbar, dann ist $f'(x)$ in $U(x_0)$ erklärt. Wenn f' in x_0 differenzierbar, dann heißt $(f')'(x_0) =: f''(x_0) =: f^{(2)}(x_0) =: \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_0)$ **2. Ableitung** von f an x_0. Analog wird induktiv die n-te Ableitung $f^{(n)} = \frac{\mathrm{d}^n f}{\mathrm{d}x^n}(x_0)$ erklärt.

Nun soll $f'(x_0)$ für $f : \mathbb{R}^m \to \mathbb{R}^n$ erklärt werden. Dazu benötigen wir die folgende Definition.

Definition 6.4 Partielle Ableitung

Sei $G \subset \mathbb{R}^m$ offen, $f : G \to \mathbb{R}$ eine gegebene Funktion und $x^0 = (x_1^0, \dots, x_m^0) \in G$. Wenn

$$\exists \lim_{h \to 0} \frac{f(x_1^0, \dots, x_{l+1}^0, x_l^0 + h, x_{l+1}^0, \dots, x_m^0) - f(x_1^0, \dots, x_l^0, \dots, x_m^0)}{h} = A,$$

heißt f an x^0 partiell nach x_l differenzierbar und A heißt partielle Ableitung von f nach x_l an x^0: $A = \frac{\partial f}{\partial x_l}(x^0) = \mathrm{D}_l f(x^0) = \partial_l f(x^0) = f_{x_l}(x^0)$
(Das heißt, der obige Grenzwert wird längs der l-ten Koordinatenachse betrachtet.)

Bemerkung

Um $\frac{\partial f}{\partial x_l}$ in der Praxis zu berechnen, wird nach x_l so abgeleitet, als wäre f nur eine Funktion von x_l. Die anderen Variablen werden wie Parameter behandelt. Höhere partielle Ableitungen werden wiederum induktiv als partielle Ableitungen partieller Ableitungen erklärt.

Beispiel 6.5

$f(x_1, x_2) = x_1 \cdot \cos(x_1 \cdot x_2) + x_1, \quad x = (x_1, x_2)$
$\frac{\partial f}{\partial x_1}(x_1, x_2) = \cos(x_1 x_2) + x_1 \cdot (-\sin(x_1 x_2)) \cdot x_2 + 1 = \cos(x_1 x_2) - x_1 x_2 \cdot \sin(x_1 x_2) + 1$
$\frac{\partial f}{\partial x_2}(x_1, x_2) = -x_1^2 \sin(x_1 x_2)$

Höhere partielle Ableitungen werden nun in offensichtlicher Weise erklärt. Zum Beispiel:

$$\frac{\partial^3 f}{\partial x_2 \partial x_1^2}(x_1, x_2) = \frac{\partial}{\partial x_2}\left(\frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_1} f\right)\right)(x_1, x_2)$$

Beispiel 6.6

Sei $f(x_1, x_2) = x_1 e^{x_2}$.
Erste partielle Ableitungen von f: $\frac{\partial f}{\partial x_1}(x_1, x_2) = e^{x_2}$ $\quad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1 e^{x_2}$
Zweite partielle Ableitungen von f: $\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) = 0$ $\quad \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) = x_1 e^{x_2}$
Gemischte zweite partielle Ableitungen von f:
$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)(x_1, x_2) = e^{x_2}$ $\quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right)(x_1, x_2) = e^{x_2}$

Satz 6.5 Satz von Schwarz

Seien $f : G \to \mathbb{R}$ sowie k und l so, dass die zweiten partiellen Ableitungen $\frac{\partial^2 f}{\partial x_l \partial x_k}, \frac{\partial^2 f}{\partial x_k \partial x_l}$ für alle $x \in G$ existieren und stetig sind. Dann sind beide Ableitungen in ganz G gleich. (Analog für höhere Ableitungen.)
(Kurz: Es kommt auf Reihenfolge der partiellen Ableitungen nicht an!)

Sprechweisen und Bezeichnungen:

- Sei $G \subset \mathbb{R}^m$ offen. Dann heißt $f : G \to \mathbb{R}$ in G **stetig differenzierbar**, wenn alle ersten partiellen Ableitungen von f existieren und in G stetig sind. f heißt in G **k-mal stetig differenzierbar**, wenn alle partiellen Ableitungen bis einschließlich k-ter Ordnung in G existieren und stetig sind.
- Der Vektorraum der k-mal stetig differenzierbaren Funktionen in G heißt $C^k(G)$.

Definition 6.6 Ableitung

Sei $G \subset \mathbb{R}^m$ offen und $f : G \to \mathbb{R}^n$. f heißt in $x^0 \in G$ **(total) differenzierbar**, wenn eine **lineare Abbildung** $A = A(x^0) : \mathbb{R}^m \to \mathbb{R}^n$ existiert, so dass $\lim\limits_{h \to o} \frac{\left\|f(x^0+h)-f(x^0)-Ah\right\|}{\|h\|} = 0$ (6.4)
$A(x^0) =: \mathrm{D}f(x^0) =: \mathrm{d}f(x^0) =: f'(x^0)$ heißt **(Fréchet-)Ableitung** oder **totales Differential** von f an x^0. f heißt in G differenzierbar, wenn f an allen $x^0 \in G$ differenzierbar ist.

Im linken Term sind $h, o \in \mathbb{R}^m$, rechts ist $0 \in \mathbb{R}$. Die Norm im Zähler ist eine Funktion $\mathbb{R}^n \to \mathbb{R}$, die Norm im Nenner bildet $\mathbb{R}^m \to \mathbb{R}$ ab. Beachte: h muss so klein sein, dass $x^0 + h \in G$. Das ist möglich, da G offen ist.

Es folgen einige äquivalente Formulierungen zur Differenzierbarkeit (vgl. Gleichung 6.2):

f ist in x^0 genau dann differenzierbar, wenn es eine lineare Abbildung $A : \mathbb{R}^m \to \mathbb{R}^n$ gibt mit

$$\lim_{x \to x^0} \frac{\left\|f(x) - f(x^0) - A(x - x^0)\right\|}{\|x - x^0\|} = 0 \qquad (6.4a)$$

$$\text{bzw. } f(x) = f(x^0) + A(x - x^0) + r(x, x^0) \qquad (6.4b)$$

wobei $r(x, x^0) \in \mathbb{R}^n, r(x, x^0) = o(\left\|x - x^0\right\|)$ für $x \to x^0$, d.h. $\lim\limits_{x \to x^0} \frac{r(x,x^0)}{\|x-x^0\|} = o \in \mathbb{R}^n$.

$$\text{Mit } h\text{: } f(x^0 + h) = f(x^0) + Ah + r(h), r(h) = o(\|h\|)\ (h \to 0)$$

$$\frac{f(x) - f(x^0) - A(x - x^0)}{\|x - x^0\|} = \frac{r(x, x^0)}{\|x - x^0\|} \to o \in \mathbb{R}^n \qquad \widehat{=}\ (6.4a)$$

Manchmal ist es günstig, zu schreiben: $r(x, x^0) =: \left\|x - x^0\right\| \cdot R(x, x^0)$ mit $R(x, x^0) \in \mathbb{R}^n$ und $R(x, x^0) \to o$ für $x \to x^0$ (6.5).

Satz 6.7

1. Wenn f an x^0 differenzierbar ist, ist f an x^0 stetig.
2. Wenn f an x^0 differenzierbar ist, ist die Ableitung $f'(x^0)$ eindeutig bestimmt.

Beweis

1. Eine Übungsaufgabe war: Jede lineare Abbildung $B : \mathbb{R}^m \to \mathbb{R}^n$ ist $\forall x \in \mathbb{R}^m$ stetig $(*)$.

$$f(x) - f(x^0) = (6.4b, 6.5) = \underbrace{A(x - x^0)}_{=Ax-Ax^0 \xrightarrow[(*)]{} 0} + \left\|x - x^0\right\| \cdot R(x, x^0) \to 0 \text{ für } x \to x^0$$

2. wird aus den folgenden Betrachtungen zur Berechnung von $f'(x^0)$ folgen.

■

Wie bestimmt man $f'(x^0)$? Wir suchen bzgl. der Standardbasis $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ in $\mathbb{R}^m$ und $\mathbb{R}^n$ die zugehörige Matrixdarstellung $f'(x^0) = A = (a_{ij}),\ i, j = 1, \ldots, n$.

Dazu benutze (6.4b) in der Formulierung mit h:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \Rightarrow f(x^0+h)-f(x^0) = \begin{pmatrix} f^1(x^0+h) - f^1(x^0) \\ \vdots \\ f^n(x^0+h) - f^n(x^0) \end{pmatrix} = Ah+r(h) = \begin{pmatrix} \sum\limits_{j=1}^{m} a_{1j}h_j \\ \vdots \\ \sum\limits_{j=1}^{m} a_{nj}h_j \end{pmatrix} + r(h) \qquad (6.6)$$

Dividiere (6.6) durch $\|h\|$ und beachte: $\frac{r(h)}{\|h\|} \to 0$ für $h \to 0$. Dann ergibt sich für die k-te Komponente:

$$\frac{1}{\|h\|} \left[f_k(x^0+h) - f_k(x^0) - \sum_{j=1}^{m} a_{kj}h_j \right] \to 0$$

Wähle nun spezielles $h \to 0$: $h := t \cdot e_l$, also $\|h\| = |t|$ ($h \to 0$ längs l-ter Koordinatenachse).

$$\Rightarrow \frac{1}{|t|} \left[f_k(x^0 + te_l) - f_k(x^0) - a_{kl}t \right] \to 0 \text{ für } t \to 0$$

$$\Leftrightarrow \frac{1}{t} \left[f_k(x^0 + te_l) - f_k(x^0) - a_{kl}t \right] \to 0 \text{ für } t \to 0$$

Mit allen Koordinaten: $a_{kl} = \lim\limits_{t \to 0} \frac{f_k(x_1^0,\ldots,x_l^0+t,\ldots,x_m^0) - f_k(x_1^0,\ldots,x_m^0)}{t} = \frac{\partial f_k}{\partial x_l}(x^0)$

Satz 6.8

Sei $G \subset \mathbb{R}^m$ offen und $f : G \to \mathbb{R}^n$ in $x^0 \in G$ differenzierbar. Dann existieren alle partiellen Ableitungen $\frac{\partial f_k}{\partial x_l}(x^0)$ und bzgl. der Standardbasen in $\mathbb{R}^m$ bzw. $\mathbb{R}^n$ gilt

$$\mathrm{D}f(x^0) = f'(x^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & \cdots & \frac{\partial f_1}{\partial x_m}(x^0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^0) & \cdots & \frac{\partial f_n}{\partial x_m}(x^0) \end{pmatrix} = \left(\frac{\partial f_k}{\partial x_l}(x^0) \right) =: \frac{\partial\,(f_1, \ldots, f_n)}{\partial\,(x_1, \ldots, x_m)}(x^0) \qquad (6.7)$$

Die obige Matrix heißt **Jacobi-Matrix** von f an x^0. Eine weitere Bezeichnung: $\det \mathrm{D}(f(x^0)) = \frac{\mathrm{D}(f_1,\ldots,f_n)}{\mathrm{D}(x_1,\ldots,x_m)}$

Die Herleitung dieses Satzes zeigt auch die Eindeutigkeit der Ableitung A von f an x^0.

Beispiel 6.7 Spezialfälle

1. $m = n = 1 \Rightarrow f : \mathbb{R} \to \mathbb{R} \Rightarrow$ übliche Ableitung
2. $m = 1, n$ beliebig. Da partielle Ableitungen bei Funktionen einer Variable mit üblichen Ableitungen zusammenfallen, ergibt sich:

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} \Rightarrow f'(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix}$$

Anwendung: Später werden wir Abbildungen $(a,b) \ni t \mapsto x(t) \in \mathbb{R}^n$ als Kurve in $\mathbb{R}^n$ denken. Die Ableitung $x'(t_0)$ wird Tangentenvektor an der Kurve $t \mapsto x(t)$ im Punkt $x(t_0)$ (zum Zeitpunkt t_0) genannt. ($x'(t_0) \mathrel{\widehat{=}}$ Geschwindigkeitsvektor)

3. m beliebig, $n = 1 : f : \mathbb{R}^m \to \mathbb{R} \Rightarrow f(x) = f(x_1, \ldots, x_m) \Rightarrow f'(x^0) = \mathrm{D}f(x^0) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_m}\right)\Big|_{x=x^0}$
$\Rightarrow$ lineare Abbildung $\mathbb{R}^m \to \mathbb{R}$ (Lineare Abbildungen aus einem Vektorraum in $\mathbb{R}/\mathbb{C}$ heißen **linear funktional**.)

$$\nabla f = \operatorname{grad} f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{pmatrix}$$

4. **Gegenbeispiel:** Die Existenz der partiellen Ableitungen von f an x^0 sichert noch nicht die Existenz der Ableitung von f an x^0:

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

f ist an $(0,0)$ nicht stetig, aber $\frac{\partial f}{\partial x_1}(0,0) = \frac{\partial f}{\partial x_2}(0,0) = 0$. f ist an $(0,0)$ nicht differenzierbar. (Beweis selber führen!) Das positive Resultat dieser Beispiele fasst der folgende Satz zusammen.

Satz 6.9

Sei $G \subset \mathbb{R}^m$ offen, $x^0 \in G$ und $f : G \to \mathbb{R}^n$. Wenn alle ersten partiellen Ableitungen $\frac{\partial f_k}{\partial x_l}$ (in x^0) existieren und (in x^0) stetig sind, dann existiert auch $f'(x^0)$.

Es folgt eine Bemerkung zu häufig benutzten Schreibweisen.

Sei $G \in \mathbb{R}^m$ offen und $f : G \to \mathbb{R}$ eine differenzierbare Funktion (in G), $f(x) = f(x_1, \ldots, x_m)$. Dann heißt der Ausdruck

$$\mathrm{d}f := \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \mathrm{d}x_i \qquad (6.8)$$

totales Differential von f bzw.

$$\mathrm{d}f(x^0) := \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0) \mathrm{d}x_i \qquad (6.8a)$$

totales Differential von f an x^0. Was sollen (6.8) und (6.8a) bedeuten und wo ist der Zusammenhang mit $f'(x^0)$ $\big(= \mathrm{d}f(x^0)\big)$? Insbesondere: Was sind die $\mathrm{d}x_i$?

Seien $p_i : \mathbb{R}^m \to \mathbb{R}$ die i-ten Koordiatenfunktionen $p_i(x_1, \ldots, x_m) = x_i$ (= Projektion auf die i-te Koordinatenachse) mit:

$$p_i(x) = x_i \Rightarrow p_i(\lambda x + \mu y) = \lambda x_i + \mu y_i = \lambda p_i(x) + \mu p_i(y)$$

p_i ist eine lineare Abbildung $\Rightarrow \mathrm{d}p_i(x^0) = p_i \ \forall x^0$. Ausgerechnet ergibt sich:

$$\mathrm{d}p_i(x^0) = \left(\frac{\partial p_i}{\partial x_1}(x^0), \ldots, \frac{\partial p_i}{\partial x_m}(x^0)\right) = (0, \ldots, 0, \overset{i}{1}, 0, \ldots, 0)$$

Wie wirkt diese lineare Abbildung? $\mathrm{d}p_i(x^0)x = x_i$ für ein beliebiges $x_i \in \mathbb{R}$. Also lässt man häufig x^0 weg, weil $\mathrm{d}p_i$ gar nicht von x^0 abhängt und man schreibt vereinfacht statt p_i einfach x_i. Dann kann man $\forall h \in \mathbb{R}^m$ schreiben:

$$\mathrm{d}f(x^0)h = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0)h_i = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0)\underbrace{\mathrm{d}x_i(h)}_{=h_i} = \left(\sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0)\mathrm{d}x_i\right)(h)$$

$$\mathrm{d}f(x^0) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0)\mathrm{d}x_i$$

Später am Ende des zweiten Semesters wird (6.8) als Differentialform wieder auftauchen.

Bemerkung

Höhere partielle Ableitungen sind schon erklärt. Höhere Ableitungen kann man auch „strukturell" erklären: Statt linearer entstehen multilineare Abbildungen. (Es existiert ein Zusammenhang mit dem Begriff des Tensors.)

6.2 Eigenschaften differenzierbarer Funktionen und Anwendungen

6.2.1 Differentiationsregeln

Satz 6.10 Allgemeine Differentiationsregeln

Seien $f, g : G \subset \mathbb{R}^m \to \mathbb{R}^n$ in $x^0 \in G$ differenzierbar und $\lambda, \mu \in \mathbb{R}$.

1. $\lambda f + \mu g$ ist in x^0 differenzierbar und $\mathrm{D}(\lambda f + \mu g)(x^0) = \lambda \mathrm{D}f(x^0) + \mu \mathrm{D}g(x^0)$
2. **Produktregel** ($n = 1!$): $f \cdot g$ ist in x^0 differenzierbar und $\mathrm{D}(f \cdot g)(x^0) = g(x^0)\mathrm{D}f(x^0) + f(x^0)\mathrm{D}g(x^0)$
   Quotientienregel ($n = 1!$): $\frac{f}{g}$ ist in x^0 differenzierbar und $\mathrm{D}\left(\frac{f}{g}\right)(x^0) = \frac{g(x^0)\mathrm{D}f(x^0) - f(x^0)\mathrm{D}g(x^0)}{g(x^0)^2}$

Beweis

1. wie bei einer Variablen (Übung: Formulieren Sie diesen Beweis!)
2. Produktregel: Verwende (6.4b) in der Formulierung mit h:

$$F(x^0 + h) - F(x^0) = F'(x^0) \cdot h + r(h)$$

Angewendet auf f, g und $f \cdot g$:

$$f(x^0+h)g(x^0+h) - f(x^0)g(x^0) = \underbrace{\left[f(x^0) + f'(x^0) + r_1(h)\right]}_{=f(x^0+h)} \cdot \underbrace{\left[g(x^0) + g'(x^0) + r_2(h)\right]}_{=g(x^0+h)} - f(x^0)g(x^0)$$

(geeignet zusammenfassen)

$$= \left[f(x^0)g'(x^0) + g(x^0)f'(x^0)\right] h + r(h)$$

mit $r(h) = \left[f(x^0) + f'(x^0)h\right] r_2(h) + \left[g(x^0) + g'(x^0)h\right] r_1(h) + r_1(h)r_2(h) + \left(f'(x^0)\right)\left(g'(x^0)\right)$. Beachte: $r(h) \in \mathbb{R}$. Zu zeigen: $r(h) = o(\|h\|)$, d.h. $\frac{r(h)}{\|h\|} \to 0$ für $h \to 0$. Für die ersten Summanden von $r(h)$ ist das einfach. Beachte dabei $\frac{r_1(h)}{\|h\|} \to 0$, $\frac{r_2(h)}{\|h\|} \to 0$ und für jede lineare Abbildung $B : \mathbb{R}^m \to \mathbb{R}^n$ gilt: $\|By\| \leq C\,\|y\|\,, C = C(B)$.

$$\frac{\left|f(x^0) + f'(x^0)h\right| \cdot |r_2(h)|}{\|h\|} \leq \underbrace{\left|f(x^0)\right|}_{=const.} \underbrace{\frac{|r_2(h)|}{\|h\|}}_{\to 0} + \underbrace{\frac{\left|f'(x^0)h\right| \cdot |r_2(h)|}{\|h\|}}_{\leq \frac{C\|h\| \cdot |r_2(h)|}{\|h\|} = C|r_2(h)| \to 0} \to 0$$

Analog für den zweiten und dritten Summand. Für den vierten Summanden gilt:

$$\left|f'(x^0)h\right| \left|g'(x^0)h\right| \leq C_1\,\|h\|\,C_2\,\|h\| \Rightarrow \frac{\left|f'(x^0)h\right| \left|g'(x^0)h\right|}{\|h\|} \leq C_1 C_2\,\|h\| \to 0$$

Also $\frac{r(h)}{\|h\|} \to 0$. Daraus folgt die Produktregel.

■

Bemerkung

Aus 1. folgt:

- Die in $G \in \mathbb{R}^m$ differenzierbaren Funktionen $f : G \to \mathbb{R}^n$ bilden einen Vektorraum.
- Die Ableitung D ist eine lineare Abbildung, da $\mathrm{D}(\lambda f + \mu g) = \lambda \mathrm{D} f + \mu \mathrm{D} g$.

Satz 6.11 Verallgemeinerte Kettenregel

Seien (mit geeigneten Definitionsbereichen) $g : G \subset \mathbb{R}^m \to \mathbb{R}^n$ in $x^0 \in G$ und $f : H \subset \mathbb{R}^n \to \mathbb{R}^k$ in $y^0 = g(x^0) \in H$ differenzierbar, wobei G, H offene Mengen sind. Dann ist $F := f \circ g$ in x^0 differenzierbar und es gilt:

$$(\mathrm{D}F)(x^0) = \mathrm{D}(f \circ g)(x^0) = (\mathrm{D}f)(y^0) \cdot (\mathrm{D}g)(x^0)$$

Die zugehörigen Jacobimatrix ist

$$\frac{\partial(F_1, \ldots, F_k)}{\partial(x_1, \ldots, x_m)}(x^0) = \frac{\partial(f_1, \ldots, f_k)}{\partial(y_1, \ldots, y_n)}(y^0) \cdot \frac{\partial(g_1, \ldots, g_n)}{\partial(x_1, \ldots, x_m)}(x^0)$$

Bemerkung

Mit $f \circ g = F : G \subset \mathbb{R}^m \to \mathbb{R}^k$ ergibt sich für $(\mathrm{D}F)(x^0)$:

$$\left.\begin{array}{r} (\mathrm{D}g)(x^0) : \mathbb{R}^m \to \mathbb{R}^n \\ (\mathrm{D}f)(y^0) : \mathbb{R}^n \to \mathbb{R}^k \end{array}\right\} \Rightarrow (\mathrm{D}F)(x^0) : \mathbb{R}^m \xrightarrow{\text{linear}} \mathbb{R}^k$$

Beweis

Nach Voraussetzung gilt:

$$\begin{aligned} g(x^0 + h) &= g(x^0) + g'(x^0)h + \|h\| \cdot R_1(h) \text{ mit } R_1(h) \to 0 \text{ für } h \to 0 \\ f(y^0 + k) &= f(y^0) + f'(y^0)k + \|k\| \cdot R_2(k) \text{ mit } R_2(k) \to 0 \text{ für } k \to 0 \end{aligned}$$

Wir machen Gebrauch davon, dass wir die Freiheit haben, wie $k \to 0$ realisiert wird. Dazu setzen wir $k := g'(x^0) \cdot h + \|h\| \cdot R_1(h)$ $(\in \mathbb{R}^n)$. Wenn $h \to 0$, dann auch $k \to 0$! Damit folgt:

$$(f \circ g)(x^0 + h) = f(g(x^0 + h)) = f(g(x^0)) + f'(y^0)\left[g'(x^0)h + \|h\| R_1(h)\right] + \|k\| R_2(k)$$
$$= f(g(x^0)) + f'(y) \cdot g'(x^0)h + \|h\| \underbrace{\left[f'(y^0) \cdot R_1(h) + \frac{\|k\|}{\|h\|} R_2(k)\right]}_{=:R(h)}$$

Zu zeigen: $R(h) \to 0$ für $h \to 0$. Zunächst gilt:

$$\|k\| \leq \left\|g'(x^0)h\right\| + \|h\| \cdot \|R_1(h)\| \leq \text{ benutze Stetigkeit der linearen Abbildung } g'(x^0)$$
$$\leq C \|h\| + \|h\| \cdot \|R_1(h)\| \Rightarrow \frac{\|k\|}{\|h\|} \leq C + \|R_1(h)\| \qquad (*)$$

Wenn $h \to 0$, dann gilt aber:

1. $R_1(h) \to 0 \Rightarrow f'(y^0) \cdot R_1(h) \to 0$ (mit Stetigkeit von $f'(y^0)$)
2. $k \to 0 \Rightarrow R_2(k) \to 0 \Rightarrow \frac{\|k\|}{\|h\|} R_2(k) \to 0$

Insgesamt ergibt sich $R(h) \to 0$ für $h \to 0$. Damit ist der 1. Teil des Satzes gezeigt. Der 2. Teil, das Produkt der Jacobi-Matrizen, ergibt sich unmittelbar, da das Hintereinanderausführen linearer Abbildungen der Multiplikation der entsprechenden Matrizen entspricht.

■

Beispiel 6.8 $m = n = k = 2$

$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}$ $\quad g_1(x_1, x_2) = x_1 x_2$, $\quad g_2(x_1, x_2) = x_1 + x_2$ $\qquad f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ $\quad f_1(y_1, y_2) = \sin y_1 + \cos y_2$, $\quad y_1 \cdot \cos y_2$

Die Funktion $F = f \circ g$ ist also von der Gestalt:

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \qquad \begin{aligned} F_1(x_1, x_2) &= f_1(g_1(x_1, x_2), g_2(x_1, x_2)) = \sin(x_1 x_2) + \cos(x_1 + x_2) \\ F_2(x_1, x_2) &= f_2(g_1(x_1, x_2), g_2(x_1, x_2)) = x_1 x_2 \cdot \cos(x_1 + x_2) \end{aligned}$$

Gesucht: $F'(x) = f(g(x))'$. Man kann entweder einfach $\mathrm{D}F(x)$ bestimmen oder die Kettenregel anwenden.
Dazu muss man zuerst $(\mathrm{D}f)(y)$ und $(\mathrm{D}g)(x)$ bestimmen.

$$f'(y) = \begin{pmatrix} \cos y_1 & -\sin y_2 \\ \cos y_2 & -y_1 \sin y_2 \end{pmatrix} \text{ und } g'(x) = \begin{pmatrix} x_2 & x_1 \\ 1 & 1 \end{pmatrix}$$

$$F'(x) = (f \circ g)'(x) = (\mathrm{D}f \cdot \mathrm{D}g)(x) = \begin{pmatrix} x_2 \cos(x_1 x_2) - \sin(x_1 + x_2) & x_1 \cos(x_1 x_2) - \sin(x_1 + x_2) \\ x_2 \cos(x_1 + x_2 - x_1 x_2 \sin(x_1 + x_2) & x_1 \cos(x_1 + x_2) - x_1 x_2 \sin(x_1 + x_2) \end{pmatrix}$$

Beispiel 6.9 Sehr wichtiger Fall der Kettenregel: m=k, n=1

$$f : \mathbb{R}^m \to \mathbb{R}, \quad f(y) = f(y_1, \ldots, y_m), \quad g : \mathbb{R} \to \mathbb{R}^m, \quad g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}$$

$$\Rightarrow F : \mathbb{R} \to \mathbb{R}, \quad F(x) = f(g_1(x), \ldots, g_m(x))$$

$$F'(x) = f'(g(x)) \cdot g'(x) = \left(\frac{\partial f}{\partial y_1}(g(x)), \ldots, \frac{\partial f}{\partial y_m}(g(x))\right) \cdot \begin{pmatrix} g_1'(x) \\ \vdots \\ g_m'(x) \end{pmatrix} = \frac{\partial f}{\partial y_1}(g(x)) \cdot g_1'(x) + \ldots + \frac{\partial f}{\partial y_m}(g(x)) \cdot g_m'(x)$$

Beispiel 6.10 „Implizite Differentiation" - **Formale** Betrachtung

f sei eine Funktion von 2 Variablen ($f(y_1, y_2)$ und $\varphi : (a,b) \to \mathbb{R}$. Es sei $f(x, \varphi(x)) = 0 \; \forall x \in (a,b)$.
Gesucht ist $\varphi'(x)$. Anwendung des obigen Beispiels: Betrachte $g : (a,b) \to \mathbb{R}^2$ mit $g(x) = \begin{pmatrix} x \\ \varphi(x) \end{pmatrix}$.

$$f : \mathbb{R}^2 \to \mathbb{R}, (f \circ g)(x) = f(x, \varphi(x)) = 0 \; \forall x \in (a,b)$$

$$0 = \frac{\partial f}{\partial y_1} g_1'(x) + \frac{\partial f}{\partial y_2} g_2'(x) = \frac{\partial f}{\partial y_1}(x, \varphi(x)) \cdot 1 + \frac{\partial f}{\partial y_2}(x, \varphi(x)) \cdot \varphi'(x)$$

$$\varphi'(x) = \frac{\frac{\partial f}{\partial y_1}(x, \varphi(x))}{\frac{\partial f}{\partial y_2}(x, \varphi(x))} = \frac{f_{y_1}(x, \varphi(x))}{f_{y_2}(x, \varphi(x))} \qquad (6.13)$$

Bemerkung

1. Diese Problematiken werden uns im Kapitel über „implizite Funktionen" genauer beschäftigen.
2. In (6.13) muss der Nenner von 0 verschieden sein.
3. Häufig wird diese Betrachtung angewendet, wenn man eine Gleichung der Form $f(x,y) = 0$ nach y auflösen möchte. Dann wird $y = y(x) =: \varphi(x)$, dann will man auf φ' die Gleichung (6.13) anwenden. **Achtung:** $\varphi : (a,b) \to \mathbb{R}$ muss auch wirklich existieren!

Beispiel 6.11 In diesen beiden Fällen existiert $\varphi(x)$ nicht.

Sei $f(x,y) = 1 + x^2 + y^2 = 0$. Wende formal (6.13) an. Dann ist $y'(x) = \varphi'(x) = -\frac{x}{y}$. Beachte: Dies ist **Unsinn**, da $f(x,y) = 0$ in $\mathbb{R}$ nicht lösbar ist. $\Rightarrow y'(x) = \varphi'(x) = -\frac{x}{y}$ *Unsinn!* $f(x,y) = 0$ nie zu erfüllen.
Anderes Beispiel: Sei $f(x,y) = x^2 + y^2 - 1 = 0.$, also $y = \sqrt{1 - x^2}$ und $y' = -\frac{x}{y} = \frac{x}{\sqrt{1-x^2}}$
Formal erfüllt der Punkt $(1,0)$ die Gleichung $f(1,0) = 0$, aber $y'(1)$ existiert nicht.

6.2.2 Mittelwertsätze und Anwendungen

Satz 6.12 Satz von Rolle

Sei f in $[a,b]$ stetig, in (a,b) differenzierbar und $f(a) = f(b) = 0$. Dann $\exists c \in (a,b) : f'(c) = 0$.

Beweis

Gelte o.E.d.A. $f \neq 0, \exists x_0 \in (a,b) : f(x_0) > 0$. Dann sei $G = \sup\limits_{x \in [a,b]} f(x) > 0$. Nach Sätzen über stetige Funktionen gilt: $G < \infty$ und $\exists c : f(c) = G$. Wir behaupten $f'(c) = 0$ und betrachten dazu den Differenzenquotienten $q(c,h) := \frac{f(c+h) - f(c)}{h}$. Für kleine h gilt: $f(c) \geq f(c+h)$. Außerdem existiert $\lim\limits_{h \to 0} q(c,h) = f'(c)$, da f in $(a,b) \ni c$ differenzierbar ist.

$$\left.\begin{array}{l} h > 0 \Rightarrow q(c,h) \leq 0 \Rightarrow f'(c) \leq 0 \\ h < 0 \Rightarrow q(c,h) \geq 0 \Rightarrow f'(c) \geq 0 \end{array}\right\} \Rightarrow f'(c) = 0$$

■

Satz 6.13 Mittelwertsatz

Sei f in $[a,b]$ stetig, in (a,b) differenzierbar. Dann existiert ein $c \in (a,b)$ mit $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Beweis durch Zurückführung auf Satz von Rolle

Betrachte die Hilfsfunktion g mit $g(x) := f(x) - f(a) - \frac{x-a}{b-a}(f(b) - f(a))$. Dann ist $g(a) = 0$ und $g(b) = 0$, d.h. g erfüllt die Vorbedingungen des Satzes von Rolle. Also $\exists c \in (a, b) : g'(c) = 0$. Man erhält:

$$0 = g'(c) = f'(c) - \frac{1}{b-a}(f(b) - f(a)) \Rightarrow f'(c) = \frac{f(b) - f(a)}{b-a}$$

■

Bemerkung

1. c ist i. A. nicht eindeutig bestimmt.
2. Der Satz von Rolle gilt auch für $f(a) = f(b) \neq 0$. Zum Beweis muss man einfach f um $f(a) = f(b)$ verschieben.
3. Eine Umformung des Mittelwertsatzes ergibt sich, indem man c anders beschreibt: $\exists \vartheta \in (0, 1)$ mit $c = a + \vartheta(b - a)$ (etwa $\vartheta = \frac{c-a}{b-a}$)
4. Dann lautet der Mittelwertsatz wie folgt: $\exists \vartheta \in (0, 1) : f(b) - f(a) = f'(a + \vartheta(b - a)) \cdot (b - a)$.

Satz 6.14 Quotientenmittelwertsatz

Seien f, g in $[a, b]$ stetig, in (a, b) differenzierbar und $g'(x) \neq 0 \; \forall x \in (a, b)$. Dann gilt:

$$\exists c \in (a, b) : \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

Achtung: Das c ist im Zähler und Nenner dasselbe.

Beweis

Aus dem MWS und $g'(x) \neq 0 \; \forall x \in (a, b)$ folgt $g(b) - g(a) \neq 0$. Wende den Satz von Rolle auf

$$h(x) = f(x) - f(a) - \frac{g(x) - g(a)}{g(b) - g(a)}(f(b) - f(a))$$

an. Also ist $h(a) = h(b) = 0 \Rightarrow \exists c \in (a, b)\colon h'(c) = 0$. Daraus folgt durch Auflösen nach $\frac{f'(x)}{g'(x)}$ der Satz.

■

Eine Anwendung ist die folgende sehr wichtige Regel.

Satz 6.15 Regel von de l'Hospital

Seien f, g in $[a, b]$ stetig, in (a, b) differenzierbar, $f(a) = g(a) = 0$ und $g'(x) \neq 0 \; \forall x \in (a, b)$. Dann gilt:

$$\lim_{x \to a+0} \frac{f'(x)}{g'(x)} = A \Rightarrow \lim_{x \to a+0} \frac{f(x)}{g(x)} = A$$

Beweis

Aus dem Satz von Rolle und $g(a) = 0, g'(x) \neq 0 \; \forall x \in (a, b)$ folgt: $g(x) \neq 0 \; \forall x \in (a, b)$. Wähle nun eine Folge (x_n) mit $a < x_n < b$ und $x_n \to a + 0$. Auf die Intervalle $[a, x_n]$ wende jeweils den Quotientenmittelwertsatz an.

$$\exists c_n \in (a, x_n) : \frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f'(c_n)}{g'(c_n)} \to A$$

Die Konv. ergibt sich aus $x_n \to a + 0 \Rightarrow c_n \to a + 0$. Oben sieht man, dass $\lim\limits_{x \to a+0} \frac{f(x)}{g(x)} = A$.

■

Bemerkung

Diese Regel hat sehr viele Modifikationen.

1. A kann $\pm\infty$ sein. a kann ebenfalls $-\infty$ sein. Der Satz gilt in diesem Falle analog mit $\lim\limits_{x\to b-0}\ldots$. Auch $b=\infty$ ist möglich.
2. Statt $f(a)=g(a)=0$ reicht $\lim\limits_{x\to a+0} f(x) = \lim\limits_{x\to a+0} g(x) = 0$ (dasselbe für ∞).
3. Gegebenenfalls muss man die Regel mehrfach anwenden, wenn man $\lim\limits_{x\to\ldots} \frac{f'(x)}{g'(x)}$ noch nicht bestimmen kann.

Beispiel 6.12

$\lim\limits_{x\to\infty} \frac{e^{\alpha x}}{x} \left(= \frac{\infty}{\infty} = \text{ n. def.}\right)$ Betrachte die Ableitungen: $\lim\limits_{x\to\infty} \frac{\alpha e^{\alpha x}}{1} = \infty \Rightarrow \lim\limits_{x\to\infty} \frac{e^{\alpha x}}{x} = \infty$

Der Einfachheit halber schreibt man zwischen beiden Grenzwerten ein Gleichheitszeichen, auch wenn noch nicht gesichert ist, ob die Grenzwerte existieren.

$\lim\limits_{x\to 0} \frac{1-\cos x}{x} = \lim\limits_{x\to 0} \frac{\sin x}{1} = 0$

Gegebenenfalls muss man erst umformen, bevor man die Regel anwenden kann.

$\lim\limits_{x\to 0+0} x^x = \lim\limits_{x\to 0+0} e^{x\log x}$

Betrachten den Exponenten: $\lim\limits_{x\to 0+0} x\log x = \lim\limits_{x\to 0+0} \frac{\log x}{\frac{1}{x}} = \lim\limits_{x\to 0+0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim\limits_{x\to 0+0} -x = 0$

$\Rightarrow \lim\limits_{x\to 0+0} e^{x\log x} = e^0 = 1$

Die folgenden Begriffe sind bereits aus der Schule bekannt.

Definition 6.16

Sei $f:(a,b)\to\mathbb{R}$ und $x_0\in(a,b)$

1. f hat in x_0 ein **lokales (relatives) Extremum**, wenn es eine Umgebung $U(x_0)$ gibt, für die gilt:
   - für relative Minima: $f(x) > f(x_0)\ \forall x\in U(x_0)\setminus\{x_0\}$
   - für relative Maxima: $f(x) < f(x_0)\ \forall x\in U(x_0)\setminus\{x_0\}$
2. f hat in x_0 einen **Wendepunkt**, wenn in x_0 ein lokales Extremum der Ableitung f' vorliegt.

Im folgenden Satz sei f in $[a,b]$ so oft differenzierbar, wie es im Satz gebraucht wird.

Satz 6.17

1. f ist in (a,b) genau dann monoton wachsend (fallend), wenn $f'(x)\geq 0 \quad \left(f'(x)\leq 0\right) \quad \forall x\in(a,b)$.
2. f ist in (a,b) genau dann streng monoton wachsend (fallend), wenn f in (a,b) monoton wachsend (fallend) ist und *kein* Teilintervall $I\subset(a,b)$ existiert, für dass gilt: $f'(x)=0\ \forall x\in I$.

Beweis Alle Teilbeweise sind ähnlich. Wir zeigen 1. für wachsende Monotonie.

Hin-Richtung: Betrachte $\frac{f(x_0+h)-f(x_0)}{h}\ \forall x_0\in(a,b)$. Für $h>0$ ist $f(x_0+h)-f(x_0)\geq 0$, für $h<0$ ist $f(x_0+h)-f(x_0)\leq 0$. Also ist $\frac{f(x_0+h)-f(x_0)}{h}\geq 0 \Rightarrow f'(x_0)\geq 0$.

Rück-Richtung: Sei $a<x_1<x_2<b$. Wende den Mittelwertsatz auf $[x_1,x_2]$ an:

$$f(x_2)-f(x_1) = \underbrace{f'(\underbrace{x_1+\vartheta(x_2-x_1)}_{\in[x_1,x_2]})}_{\geq 0}\cdot\underbrace{(x_2-x_1)}_{\geq 0}\geq 0 \Rightarrow f(x_2)\geq f(x_1)$$

■

Bemerkung

Zu 2.: In einzelnen Punkten darf die Ableitung verschwinden. $f(x) = x^3$ zum Beispiel ist streng monoton wachsend, aber $f'(0) = 0$.

Satz 6.18

Sei f in (a, b) zweimal stetig differenzierbar und $x_0 \in (a, b)$.

1. notwendige Bedingung für relative Extrema: $f'(x_0) = 0$
   hinreichende Bedingung für relative Minima: $f'(x_0) = 0 \wedge f''(x_0) > 0$
   hinreichende Bedingung für relative Maxima: $f'(x_0) = 0 \wedge f''(x_0) < 0$
2. notwendige Bedingung für Wendepunkte: $f''(x_0) = 0$
   hinreichende Bedingung für Wendepunkte: $f''(x_0) = 0 \wedge f'''(x_0) \neq 0$

Man kann weitere Bedingungen aufschreiben (unter Einbeziehung höherer Ableitungen), etwa für relative Extrema: $f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \wedge f^{(n)}(x_0) \neq 0$

Im Folgenden wird der Mittelwertsatz für Funktionen mehrerer Variablen diskutiert.

Satz 6.19 Mittelwertsatz

Sei $G \subset \mathbb{R}^m$ offen, $f : G \to \mathbb{R}$ differenzierbar und $x^0, x^0 + h \in G$ so, dass auch noch die Verbindungsstrecke zwischen x^0 und $x^0 + h$ in G liegt. Dann gilt:

$$\exists \vartheta \in (0,1) : f(x^0 + h) = f(x^0) + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \left(x^0 + \vartheta h\right) \cdot h_i$$

Bemerkung

Vergleiche diesen Mittelwertsatz mit dem Mittelwertsatz für eine Variable:

$f(b) - f(a) = f'(a + \vartheta(b-a)) \cdot (b-a) \Rightarrow$ mit $a = x^0$ und $b = x^0 + h \Rightarrow f(x^0+h) = f(x^0) + f'(x^0 + \vartheta h) \cdot h$

Beweis Zurückführung auf den Mittelwertsatz für Funktionen einer Variablen

Setze $g : [0,1] \to G, g(t) = x^0 + th$ (also ist $W(g)$ die Verbindungsstrecke zwischen x^0 und $x^0 + h$) und $\varphi = f \circ g : [0,1] \to \mathbb{R}, \varphi(t) = f(x^0 + th)$. φ ist in $[0,1]$ differenzierbar. Mit Kettenregel ist:

$$\varphi'(t) = f'(g(t)) \cdot g'(t)$$

Wende den Mittelwertsatz auf φ und $[0,1]$ an.

$$\exists \vartheta \in (0,1) : \frac{\varphi(1) - \varphi(0)}{1 - 0} = \varphi'(0 + \vartheta(1-0)) = \varphi'(\vartheta)$$

$$\varphi'(t) = f'(g(t)) \cdot \underbrace{g'(t)}_{=h} = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_m}\right)(g(t)) \cdot h$$

Also insgesamt:

$$\underbrace{f(x^0 + h)}_{=\varphi(1)} = \underbrace{f(x^0)}_{=\varphi(0)} + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \left(x^0 + \vartheta h\right) \cdot h_i$$

■

Bemerkung

1. Mit dem Mittelwertsatz kann man den Satz von Schwarz zeigen.
2. **Achtung:** Für Abbildungen $f : \mathbb{R}^m \to \mathbb{R}^n$ gilt ein ähnlicher Satz nicht, aber es existieren sogenannte Schrankensätze. Man kann die Differenz der Funktionswerte durch die Ableitung abschätzen. (Dafür brauchen wir Hilfsmittel aus der Integralrechnung.)

6.3 Taylorformel und Taylorscher Satz

Motivation: Sei $f(x) = a_0 + a_1(x - x_0) + \ldots + a_n(x - x_0)^n$, wobei $x_0 \in \mathbb{R}$ fest und $x \in \mathbb{R}$ beliebig gegeben. Wodurch sind die a_i bestimmt?
Aus dem Fundamentalsatz der Algebra folgt: Die a_i sind bestimmt durch die Funktionswerte an $(n+1)$ Stellen.

Betrachte eine beliebig oft differenzierbare analytische Funktion $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ für ein x aus dem Konvergenzintervall der Potenzreihe. Unmittelbar folgt: $a_n = \frac{f^{(n)}(x_0)}{n!}$

Umgekehrt: Sei $f \in C^\infty(a, b)$, also in (a, b) beliebig oft differenzierbar. Bilde formal:

$$a_n := \frac{f^{(n)}(x_0)}{n!} \text{ für ein } x_0 \in (a, b) \text{ und } \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

Konvergiert diese Potenzreihe? Wenn ja, hat dann ihre Summe etwas mit $f(x)$ zu tun?
Ja, manchmal.

Theorem 6.20 Taylorformel für Funktionen einer Variablen

Sei $I = (a, b)$, $f \in C^{n+1}$ (I). Dann gilt für alle $x, x_0 \in I$:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(x_0 + \vartheta(x - x_0))}{(n+1)!}(x - x_0)^{n+1} \quad (6.14)$$

mit einem $\vartheta \in (0, 1)$, wobei $\vartheta = \vartheta(x, x_0)$. Dabei heißt

- $R_n(x, x_0) := \frac{f^{(n+1)}(x_0+\vartheta(x-x_0))}{(n+1)!}(x-x_0)^{n+1}$ **Restglied der Taylorformel** in der Lagrange-Form
- $T_n(x, x_0) := f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$ **n-tes Taylorpolynom** von f an x_0

Bemerkung

1. Die Darstellung als $f(x) = T_n(x, x_0) + R$ mit einem beliebigen Rest wäre witzlos, denn das gilt immer.
   (Setze $R := f(x) - T_n(x, x_0)$.) Das Wertvolle an (6.14) besteht in der expliziten Beschreibung des Restgliedes.
2. Es existieren noch andere Darstellungen des Restes (siehe Literatur / spätere Übungen).
3. Für $n = 0$ ist (6.14) der Mittelwertsatz.
4. (6.14) heißt: f wird durch T_n mit Fehler R_n approximiert.

Beispiel 6.13 Warnung vor Fehlinterpretationen

1. $f(x) = (x - x_0)^n \Rightarrow T_n(x, x_0) = 0, R_n(x, x_0) = (x - x_0)^{n+1}$
2. $f(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$
   Dieser Fall ist noch schlimmer. f ist beliebig oft differenzierbar $\forall x \in \mathbb{R}$. Für $x_0 = 0$ erhält man: $f^{(n)}(0) = 0 \ \forall n$. Das heißt, dass $T_n(x, x_0) = 0 \ \forall n$. Alle Informationen über $f(x = 0)$ stecken im Restglied, also wird f durch T_n ganz schlecht „approximiert“. Mit anderen Worten: Eine beliebig oft differenzierbare Funktion muss sich durchaus <u>nicht</u> in eine Potenzreihe entwickeln lassen.

Beweis des Theorems
Seien n, x, x_0 fest. Trivialerweise existiert ein $\varrho \in \mathbb{R}$ mit:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{(x-x_0)^{n+1}}{(n+1)!}\varrho \quad (6.15)$$

Natürlich ist $\varrho = \varrho(x, x_0)$. Wir bestimmen ϱ geschickt und betrachten dazu

$$F(t) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k - \frac{(x-t)^{n+1}}{(n+1)!}\varrho \quad (6.16)$$

Hierbei liegt t in dem Intervall I_1 mit den Grenzen x und x_0. F ist in I_1 stetig differenzierbar und aufgrund von (6.15) und (6.16) ist $F(x) = F(x_0) = 0$. Nach Satz von Rolle existiert ein $\vartheta \in (0,1)$ mit $F'(x_0 + \vartheta(x-x_0)$.

$$F'(t) = 0 - \underbrace{f'(t)}_{\text{für } k=0} - \sum_{k=1}^{n}\left[\frac{f^{(k+1)}(t)}{k!}(x-t)^k + \frac{f^{(k)}(t)}{k!}k\cdot(x-t)^{k-1}\right] + \frac{(x-t)^n}{n!}\varrho$$

Beachte, dass sich in der Summe fast alles weghebt.

$$= -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(x-t)^n}{n!}\varrho \Rightarrow F'(x_0+\vartheta(x-x_0)) = 0 \Rightarrow \underline{\varrho = f^{(n+1)}(x_0+\vartheta(x-x_0))}$$

■

Im Idealfall ist $\lim\limits_{n\to\infty} R_n(x, x_0) = 0$.

Satz 6.21 Satz über die Taylorentwicklung von Funktionen einer Variable

Sei I ein offenes Intervall um x_0 und $f \in C^\infty(I)$. Dann hat f für $x \in I$ genau dann die Taylorentwicklung

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \qquad (6.17),$$

wenn das Restglied $R_n(x)$ in der Taylorformel gegen 0 geht für $n \to \infty$. (6.17) heißt Taylorreihe von f an x_0.

Beweis
Wenn $R_n \to 0$, wird aus (6.14) die Taylorreihe. Wenn (6.17) gilt, dann hat (6.14) die Gestalt:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \underbrace{\sum_{k=n+1}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}_{\text{Rest} \to 0, \text{ weil Reihe konvergiert}}$$

■

Bemerkung

1. Wenn (6.17) gilt, sagt man: „f ist an x_0 in eine Taylorreihe entwickelbar“. Funktionen, die durch Potenzreihen bzw.
   Taylorreihen darstellbar sind, heißen **analytisch** (siehe Kapitel 3.4). Wenn f gegeben ist durch

   $$f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k \text{ mit } x, x_0 \in I,$$

   dann ist die Potenzreihe auch genau die Taylorreihe von f an x_0, also: $a_k = \frac{f^{(k)}}{k!}$.

2. Achtung: „Analytisch“ ist stärker als „beliebig oft differenzierbar“ (siehe zum Beispiel $e^{-\frac{1}{x^2}}$).
3. Um zu zeigen, dass $R_n \to 0$, muss man das Verhalten der Ableitungen von f kontrollieren.
   Günstiger Fall: $\left|f^{(n)}(x)\right| \leq \alpha C^n \; \forall x \in I$ mit $\alpha, C > 0$ fest
   $\Rightarrow |R_n(x)| = \frac{\left|f^{(n+1)}(x_0+\vartheta(x-x_0))\right|}{(n+1)!} \cdot \left|(x-x_0)^{n+1}\right| \leq \frac{\alpha \cdot C}{(n+1)!} \cdot \left|x-x_0^{n+1}\right| \to 0$ für $n \to \infty$

Beispiel 6.14

Sei $f(x) = \frac{1}{1-x}$ in einem Intervall I um 0 mit $1 \notin I$. Gesucht ist die Taylorreihe an $x_0 = 0$.
$\frac{1}{1-x}$ ist gerade die Summe der geometrischen Reihe $\sum\limits_{n=0}^{\infty} x^n$. Dies ist auch die gesuchte Taylorreihe.
Zum eigenen Üben: Zeigen Sie, dass das Restglied der Taylorformel für $n \to \infty$ gegen 0 geht.

Beispiel 6.15 Exponentialfunktion (analog für Sinus und Cosinus)

Sei $f(x) = \mathrm{e}^x$ und $x_0 = 0$. Wir verwenden hier nur $(\mathrm{e}^x)^{(n)} = \mathrm{e}^x \Rightarrow f^{(n)}(0) = 1 \; \forall n$.
Daraus ergibt sich die Taylor-Entwicklung $\mathrm{e}^x = \sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$, wenn $R_n(x,0) \to 0$ für $n \to \infty$.
$R_n(x,0) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)} \underbrace{(x_0 + \vartheta(x-x_0))}_{x_0=0} = \frac{x^{n+1}}{(n+1)!} e^{\vartheta x} \to 0$

Jetzt wollen wir die Taylorformel für Funktionen mehrerer Variablen finden. Das einzige wirkliche Problem ist geschicktes Aufschreiben. Dies soll an einer Funktion zweier Variablen demonstriert werden.

Sei $x^0 = (x_1^0, x_2^0)$ und $x^0 + h = (x_1^0 + h_1, x_2^0 + h_2)$. Außerdem setzen wir $\frac{\partial}{\partial x_i} = \mathrm{D}_i$, also zum Beispiel $\frac{\partial f}{\partial x_i} = \mathrm{D}_i f$. f sei in der offenen Menge $G \subset \mathbb{R}^2$ definiert und hinreichend oft stetig differenzierbar. Für hinreichend kleine $\|h\|$ liegen $x^0, x^0 + h$ und die Verbindungsstrecke zwischen diesen Vektoren ganz in G.

Analog zum Beweis der MWS definieren wir $g(t) := f(x^0 + th) \quad t \in [0,1]\,, g : [0,1] \to \mathbb{R}$, also $g(0) = f(x^0)$ und $g(1) = f(x^0 + h)$ Wir werden $g', \dots, g^{(k)}$ geschickt aufschreiben und auf g (eine Funktion *einer* Variablen) die Taylorformel anwenden.

Kettenregel: $g'(t) = \frac{\partial f}{\partial x_1}(x^0 + th) \cdot \frac{\partial(x_1^0 + th_1)}{\partial t} + \frac{\partial f}{\partial x_2}(x^0 + th) \cdot \frac{\partial(x_2^0 + th_2)}{\partial t}$
$g'(t) = \frac{\partial f}{\partial x_1}(x^0 + th) \cdot h_1 + \frac{\partial f}{\partial x_2}(x^0 + th) \cdot h_2 = \mathrm{D}_1 f \cdot h_1 + \mathrm{D}_2 f \cdot h_2 = (h_1 \cdot \mathrm{D}_1 + h_2 \cdot \mathrm{D}_2)\, f$

Hierbei sind $\mathrm{D}_1 f$ und $\mathrm{D}_2 f$ wiederum Funktionen von $x^0 + th$, weshalb sich analog ergibt:
$g''(t) = h_1(\mathrm{D}_1 f)' + h_2(\mathrm{D}_2 f)' = h_1\left[h_1 \mathrm{D}_1(\mathrm{D}_1 f) + h_2 \mathrm{D}_2(\mathrm{D}_1 f)\right] + h_2\left[h_1 \mathrm{D}_1(\mathrm{D}_2 f) + h_2 \mathrm{D}_2(\mathrm{D}_2 f)\right]$
$g''(t) = \left[h_1^2 \mathrm{D}_1^2 + h_1 h_2 \mathrm{D}_1 \mathrm{D}_2 + h_1 h_2 \mathrm{D}_2 \mathrm{D}_1 + h_2^2 \mathrm{D}_2^2\right] f = (h_1 \mathrm{D}_1 + h_2 \mathrm{D}_2)^2\, f$

Durch Induktion zeigt man $g^{(k)}(t) = (h_1 \mathrm{D}_1 + h_2 \mathrm{D}_2)^k\, f(x^0 + th)$. Hierbei vereinbaren wir: $(h_1 \mathrm{D}_1 + h_2 \mathrm{D}_2)^0 f = f$. Sei nun $f \in C^{n+1}(G) \Rightarrow g \in C^{n+1}([0,1])$. Wende die Taylorformel für g an $t = 0$ an. Dann existiert ein $\vartheta \in (0,1)$:

$$\underbrace{g(1)}_{\widehat{=} f(x^0+h)} = g(0) + g'(0) + \frac{g''(0)}{2!} + \ldots + \frac{g^{(n)}}{n!} + \frac{g^{(n+1)}\vartheta}{(n+1)!}$$

$$\Rightarrow f(x^0 + th) = \sum_{k=0}^{n} \frac{1}{n!}(h_1 \mathrm{D}_1 + h_2 \mathrm{D}_2)^k f(x^0) + \frac{1}{(n+1)!}(h_1 \mathrm{D}_1 + h_2 \mathrm{D}_2)^{n+1} f(x^0 + \vartheta h)$$

Theorem 6.22 Taylorformel für Funktionen mehrerer Variablen

Sei $G \subset \mathbb{R}^m$ offen und $f \in C^{n+1}(G)$. Die Punkte x^0, $x^0 + h$ und die Verbindungsstrecke zwischen diesen Punkten liege in G. Dann existiert ein $\vartheta \in (0,1)$:

$$f(x^0+h) = \sum_{k=0}^{n} \frac{1}{k!}(h_1\mathrm{D}_1+\ldots+h_m\mathrm{D}_m)^k f(x^0) + \underbrace{\frac{1}{(n+1)!}(h_1\mathrm{D}_1+\ldots+h_m\mathrm{D}_m)^{n+1} f(x^0+\vartheta h)}_{\text{Restglied}} \tag{6.18}$$

Für kleine n ergeben sich wichtige Spezialfälle. Für $n = 0$ ist (6.18) der Mittelwertsatz:

$$f(x^0+h) = f(x^0) + (h_1\mathrm{D}_1+\ldots+h_m\mathrm{D}_m) f(x^0+\vartheta h)$$

$$= f(x^0) + \sum_{i=1}^{m} \frac{\partial f(x^0+\vartheta h)}{\partial x_i} h_i = f(x^0) + \left\langle (f'(x^0+\vartheta h))^T, h \right\rangle \tag{6.19}$$

Für $n = 1$ folgt:

$$f(x^0+h) = f(x^0) + \left\langle (f'(x^0))^T, h \right\rangle + \frac{1}{2}\sum_{j,k=1}^{m} \frac{\partial^2 f(x^0+\vartheta h)}{\partial x_j \partial x_k} h_j h_k \tag{6.20}$$

Bemerkung

1. Wenn das Restglied gegen 0 geht, erhält man eine Taylorreihe, also eine Potenzreihe mit mehreren Variablen.
2. Die Matrix $H(x) := \left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)_{jk}$ der zweiten Ableitungen von f heißt Hesse-Matrix.
   (6.20) kann so geschrieben werden: $f(x^0+h) = f(x^0) + \left\langle f'(x^0)^T, h \right\rangle + \frac{1}{2}\langle Hh, h \rangle$

$$H_{jk} = \frac{\partial^2 f}{\partial x_j \partial x_k}, (Hh)_k = \sum_j H_{kj} h_j \Rightarrow \langle Hh, h \rangle = \sum_k (Hh)_k h_k = \sum_{j,k} \frac{\partial^2 f}{\partial x_j \partial x_k} h_j h_k$$

Beispiel 6.16

Entwickle $f(x,y) = e^{x+y}\sin xy$ an $(0,1)$ bis zur dritten Ordnung. Man kann alle dritten Ableitungen ausrechnen und in (6.18) einsetzen *oder* auf bekannte Reihen zurückgreifen und dann ausmultiplizieren.

$$e^{x+y} = 1 + (x+y) + \frac{(x+y)^2}{2!} + \ldots \qquad \sin(xy) = xy - \frac{(xy)^3}{3!} + \ldots$$

6.4 Lokale Extrema von Funktionen mehrerer Variablen

Definition 6.23

Sei $G \subset \mathbb{R}^m$ offen und $f: G \to \mathbb{R}$. Man sagt, f hat an $x^0 \in G$ ein **lokales/relatives Extremum**, wenn eine Umgebung $U = U(x^0) \subset G$ existiert, sodass $\forall x \in U$ gilt:

- für lokale Minima: $f(x) \geq f(x^0)$
- für lokale Maxima: $f(x) \leq f(x^0)$

Punkte aus G, in denen f' verschwindet, heißen **kritische** oder **stationäre Punkte**, d.h. x^0 ist kritisch, wenn $f'(x^0) = 0$. Dabei ist 0 die lineare Abbildung $x \mapsto 0$.

Satz 6.24 Notwendige Bedingung

f sei wie in 6.23 definiert und möge partielle Ableitungen nach allen Variablen besitzen. Wenn x^0 ein lokales Extremum von f ist, dann gilt $f_{x_k}(x^0) = 0$ mit $k = 1, \dots, m$.
Falls $f'(x^0)$ existiert (z.B. weil $\frac{\partial f}{\partial x_k}$ für alle k stetig ist), dann heißt das, dass $f'(x^0) = 0$. x^0 ist also ein kritischer Punkt von f.

Beweis
Trick: Betrachte die Hilfsfunktionen g_k mit $k = 1, \dots, m$, die definiert sind als

$$g_k(x) = f(x_1^0, \dots, x_{k-1}^0, x, x_{k+1}^0, \dots, x_m^0)$$

Die g_k sind Funktionen einer Variablen, die jeweils an x_k^0 ein relatives Extremum haben.
Daraus folgt für die notwendige Bedingung: $0 = \frac{\partial g_k}{\partial x}(x_k^0) \mathrel{\widehat{=}} \frac{\partial f}{\partial x_k}(x^0)$

■

Für die Formulierung hinreichender Bedingungen benötigen wir Kenntnisse aus der linearen Algebra.

Definition 6.25

Sei $A = (a_{ij})$ eine symmetrische Matrix, d.h. alle Eigenwerte von A sind reell. Wir benutzen in $\mathbb{R}^m$ das Standardskalarprodukt. Beachte: $\langle Ax, x\rangle = \sum\limits_{j,k} a_{jk}x_j x_k$. A heißt

- **positiv definit**, wenn $\langle Ax, x\rangle > 0 \; \forall x \in \mathbb{R}^m \setminus \{0\}$
- **negativ definit**, wenn $\langle Ax, x\rangle < 0 \; \forall x \in \mathbb{R}^m \setminus \{0\}$
- **indefinit**, wenn $\exists x, y \in \mathbb{R}^m : \langle Ax, x\rangle > 0 \wedge \langle Ay, y\rangle < 0$

Satz 6.26 aus der linearen Algebra

1. Für A sind äquivalent:
   (a) A ist positiv definit.
   (b) Alle Eigenwerte von A sind größer als 0.
   (c) $\exists\, C > 0 : \langle Ax, x\rangle \geq C \|x\|^2 \;\; \forall x \in \mathbb{R}^m \quad \left(\Leftrightarrow \left\langle A\frac{x}{\|x\|}, \frac{x}{\|x\|}\right\rangle \geq C \; \forall x \in \mathbb{R}^m \setminus \{0\}\right)$
2. Für A sind äquivalent:
   (a) A ist negativ definit.
   (b) Alle Eigenwerte von A sind kleiner als 0.
   (c) $\exists\, C > 0 : \langle Ax, x\rangle \leq C \|x\|^2 \;\; \forall x \in \mathbb{R}^m$
3. A ist genau dann indefinit, wenn es sowohl positive als auch negative Eigenwerte gibt.

Beweis Für den ganzen Beweis siehe die Lineare-Algebra-Vorlesung.
Wir zeigen 1. $(b) \Rightarrow (c)$. Seien $\lambda_1, \dots, \lambda_m$ Eigenwerte von A und $f_1, \dots, f_m$ eine Orthonormalbasis aus den zugehörigen Eigenvektoren.

$$x = x_1 \cdot f_1 + \ldots + x_m \cdot f_m \Rightarrow Ax = x_1 \cdot Af_1 + \ldots + x_m \cdot Af_m = x_1\lambda_1 f_1 + \ldots + x_m\lambda_m f_m$$

$$\langle Ax, x\rangle = \langle x_1\lambda_1 f_1 + \ldots + x_m\lambda_m f_m, x_1 \cdot f_1 + \ldots + x_m \cdot f_m\rangle = \lambda_1 x_1^2 + \ldots + \lambda_m x_m^2 \geq \min_{i=1,\dots,m} \lambda_i \|x\|^2$$

■

Satz 6.27 Hinreichende Bedingungen

Sei $f \in C^2(G)$ und $x^0 \in G$ eine kritische Stelle, d.h. $f'(x^0) = 0$. Mit $H(x^0) := \left(\frac{\partial^2 f}{\partial x_j x_k}(x^0)\right)$ sei die Hesse-Matrix von f an x^0 bezeichnet. Dann gilt:

- Ist $H(x^0)$ positiv definit, so hat f an x^0 ein (echtes) lokales Minimum.
- Ist $H(x^0)$ negativ definit, so hat f an x^0 ein (echtes) lokales Maximum.
- Ist $H(x^0)$ indefinit, so hat f an x^0 kein lokales Extremum.

Beweis
Wir beschränken uns auf den Fall, dass $H(x^0)$ positiv definit ist. Die anderen Fälle folgen völlig analog. Benutze nun die Taylorformel (6.19) und beachte hierbei, dass $f'(x^0) = 0$. $f(x^0+h) = f(x^0) + \frac{1}{2}\sum\limits_{i,j=1}^{m} \frac{\partial^2 f(x^0+\vartheta h)}{\partial x_i \partial x_j} h_i h_j$

$$f(x^0+h) = f(x^0) + \tfrac{1}{2}\sum_{i,j=1}^{m} \frac{\partial^2 f(x^0)}{\partial x_i \partial x_j} h_i h_j + \underbrace{\frac{1}{2}\sum_{i,j=1}^{m}\left(\frac{\partial^2 f(x^0+\vartheta h)}{\partial x_i \partial x_j} - \frac{\partial^2 f(x^0)}{\partial x_i \partial x_j}\right) h_i h_j}_{=:R(h)}$$

$f(x^0+h) = f(x^0) + \frac{1}{2}\left\langle H(x^0)h, h\right\rangle + R(h)$

Umschreiben ergibt: $\frac{f(x^0+h)-f(x^0)}{\|h\|^2} = \frac{1}{2\|h\|^2}\left\langle H(x^0)h, h\right\rangle + \underbrace{\frac{R(h)}{\|h\|^2}}_{=:r(h)} \qquad (*)$

wir zeigen, dass ein $\delta > 0$ existiert, sodass für alle h mit $\|h\| < \delta$ die linke Seite von (*) positiv ist. Es gilt aufgrund der Dreiecksungleichung :

$$|r(h)| \le \frac{1}{2}\sum_{i,j}\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(x^0+\vartheta h) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0)\right| \frac{|h_i|\cdot|h_j|}{\|h\|^2} \le \frac{1}{2}\sum_{i,j}\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(x^0+\vartheta h) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0)\right|$$

Dieser Ausdruck konvergiert für $h \to 0$ gegen 0, da $\frac{\partial^2 f}{\partial x_i \partial x_j}$ an x^0 stetig ist.
Dass $H(x^0)$ positiv definit ist, bedeutet: $\exists C > 0 : \left\langle H(x^0)\frac{h}{\|h\|}, \frac{h}{\|h\|}\right\rangle \ge C > 0 \quad (**)$. Also existiert ein $\delta > 0$, für das für alle h mit $\|h\| < \delta$ gilt: $x^0 + h \in G$ und $|r(h)| < \frac{C}{2}$, d.h. $-\frac{C}{2} < r(h) < \frac{C}{2}$. Aus $(*)$, $(**)$ folgt also, dass für $\|h\| < \delta$ gilt:

$$\frac{f(x^0+h)-f(x^0)}{\|h\|^2} \ge \frac{1}{2}C + r(h) > \frac{1}{2}C - \frac{1}{2}C = 0$$

Das heißt, in der Umgebung $U(x^0) = \left\{x : \left\|x - x^0\right\| < \delta\right\}$ gilt:

$$\underbrace{f(x)}_{\widehat{=} f(x^0+\vartheta h)} - f(x^0) > 0$$

■

Beispiel 6.17

1. $f(x,y) := x^3 + y^3 - 3xy$, $f_x = 3x^2 - 3y$, $f_y = 3y^2 - 3x$. Dann sind $(0,0)$ und $(1,1)$ kritische Punkte.
   Bilden der Ableitungen: $f_{xx} = 6x$, $f_{xy} = f_{yx} = -3$, $f_{yy} = 6y$
   Die Hessematrizen lauten: $H(x,y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$ $H(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$ $H(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$
   Da $H(0,0)$ indefinit ist, liegt an diesem kritischen Punkt kein lokales Extremum von f vor.
   $H(1,1)$ hingegen ist mit den Eigenwerten 3 und 9 positiv definit, also gibt es an $(1,1)$ ein lokales Minimum von f.
2. $f(x,y) = x^2 - y^2$ hat nur den kritischen Punkt $(0,0)$ einzige kritische Stelle.
   Da $H(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ indefinit sind, liegt dort kein Minimum an f vor.
3. $f(x,y) = x^3 - y^3$ hat den kritischen Punkt $(0,0)$. Das Kriterium versagt:
   $H(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, aber an $(0,0)$ liegt (wie man leicht sieht) kein lokales Extremum von f vor.

6.5 Anfangsgründe der Vektoranalysis

6.5.1 Richtungsableitung mit Gradient

Sei $G \subset \mathbb{R}^m$ offen und $f : G \to \mathbb{R}, x^0 \in G$. Um $\frac{\partial f}{\partial x_i}(x^0)$ zu bilden, betrachtet man den Differenzenquotienten $x \to x^0$ längs der i-ten Koordinatenachse. Sei nun $v = (v_1, \ldots, v_m)^T$ ein Vektor, der eine Richtung repräsentiert.

Definition 6.28

Sei $v \in \mathbb{R}^m$ fest. Wenn der Grenzwert

$$\lim_{t \to 0} \frac{f(x^0 + tv) - f(x^0)}{t} \qquad (6.21)$$

existiert, heißt er **Richtungsableitung** von f in Richtung v im Punkt x^0. Für $\|v\| = 1$ nennt man den Grenzwert Richtungsableitung von f nach v im Punkt x^0. In diesem Falle schreibt man $(\mathrm{D}_v f)(x^0)$, $\mathrm{D}_v f(x^0)$ oder $\frac{\partial f}{\partial v}(x^0)$.

Satz 6.29

Sei f in x^0 differenzierbar, dann existiert in x^0 die Ableitung in jeder Richtung v und es gilt:

$$\frac{\partial f}{\partial v}(x^0) = \frac{\partial f}{\partial x_1}(x^0) \cdot v_1, \ldots + \frac{\partial f}{\partial x_m}(x^0) v_m = f'(x^0) v \qquad (6.22)$$

Das heißt, die lineare Abbildung $f'(x^0)$ wird auf v angewendet. (Das Ergebnis ist eine Zahl.)

Beweis Zurückführung auf Funktionen einer Variablen

Sei $\varphi(t) := f(x^0 + tv)$. (6.21) bedeutet: $\varphi'(0)$ wird betrachtet. Da f differenzierbar ist, kann die Kettenregel angewendet werden.

$$\varphi'(0) = \frac{\partial f}{\partial v}(x^0) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0) \cdot v_i = \left(\frac{\partial f}{\partial x_1}(x^0), \ldots, \frac{\partial f}{\partial x_m}(x^0) \right) v$$

■

Folgerung 6.30 Einfachste Variante der Gradientendefinition

In $\mathbb{R}^m$ sei $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt, (e_i) die Standardorthonormalbasis, f und x^0 wie oben. Dann existiert ein eindeutig bestimmter Vektor $w \in \mathbb{R}^m$ mit

$$f'(x^0) v = \frac{\partial f}{\partial v}(x^0) = \langle w, v \rangle \ \forall v \in \mathbb{R}^m \quad (\|v\| = 1)$$

w heißt **Gradient** von f an x^0. Man schreibt $w = \operatorname{grad} f(x^0)$.

Im Falle des Standardskalarproduktes und bezüglich (e_i) ist also:

$$\operatorname{grad} f(x^0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x^0) \\ \vdots \\ \frac{\partial f}{\partial x_m}(x^0) \end{pmatrix} \in \mathbb{R}^m$$

$$f'(x^0) \in L(\mathbb{R}^m, \mathbb{R}) \equiv (\mathbb{R}^m)' \equiv (\mathbb{R}^m)^*$$

$(\mathbb{R}^m)'$ bzw. $(\mathbb{R}^m)^*$ heißt **Dualraum** zu $\mathbb{R}^m$ bzw. **Raum der linearen Funktionale**.

$$\frac{\partial f}{\partial v}(x^0) = \left\langle \operatorname{grad} f(x^0), v \right\rangle \qquad (6.23)$$

Folgerung 6.31 Anspruchsvollere Variante der Gradientendefinition

Sei $\langle \cdot, \cdot \rangle'$ ein beliebiges Skalarprodukt in $\mathbb{R}^m$ und f und x^0 wie oben. Dann existiert ein eindeutig bestimmter Vektor $w \in \mathbb{R}^m$ mit $f'(x^0)v = \langle w, v \rangle'$ $\forall v \in \mathbb{R}^m$.
w heißt **Gradient** von f an x^0 bzgl. $\langle \cdot, \cdot \rangle'$. Man schreibt manchmal $w = \operatorname{grad}_{\langle \cdot, \cdot \rangle'} f(x^0)$.

$f'(x^0)v$ hängt überhaupt nicht von $\langle \cdot, \cdot \rangle'$ ab, aber w.

Satz 6.32

Sei $f : G \subset \mathbb{R}^m \to \mathbb{R}$ in $x^0 \in G$ stetig differenzierbar.

1. Ist $\operatorname{grad} f(x^0) = 0 (\in \mathbb{R}^m)$, so verschwinden alle Richtungsableitungen.
2. Ist $\operatorname{grad} f(x^0) \neq 0$, so gibt es unter allen Richtungsableitungen $\frac{\partial f}{\partial v}(x^0)$ (also $\|v\| = 1$) eine größte, nämlich in Richtung des Gradienten, d.h. für $v = \frac{\operatorname{grad} f(x^0)}{\|\operatorname{grad} f(x^0)\|}$ ist $\frac{\partial f}{\partial v}(x^0) = \operatorname{grad} f(x^0)$.
Sprachgebrauch: grad zeigt in Richtung des größten Anstiegs.

Beweis

1. Folgt unmittelbar aus der Gradientendefinition.
2. Anwendung der Cauchy-Schwarzschen Ungleichung $|\langle a, b \rangle| \leq \|a\| \cdot \|b\|$

$$\left| \frac{\partial f}{\partial v}(x^0) \right| = \left| \left\langle \operatorname{grad} f(x^0), v \right\rangle \right| \leq \left\| \operatorname{grad} f(x^0) \right\| \cdot \underbrace{\|v\|}_{=1} = \left\| \operatorname{grad} f(x^0) \right\|$$

$$\Rightarrow -\left\| \operatorname{grad} f(x^0) \right\| \leq \frac{\partial f}{\partial v}(x^0) \leq \left\| \operatorname{grad} f(x^0) \right\|$$

$$\text{Für } v = \frac{\operatorname{grad} f(x^0)}{\|\operatorname{grad} f(x^0)\|} \text{ ist } \frac{\partial f}{\partial v} = \left\langle \operatorname{grad} f(x^0), \frac{\operatorname{grad} f(x^0)}{\|\operatorname{grad} f(x^0)\|} \right\rangle = \left\| \operatorname{grad} f(x^0) \right\|$$

■

Der Nabla-Operator

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_m} \end{pmatrix}, \quad \nabla f(x^0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x^0) \\ \vdots \\ \frac{\partial f}{\partial x_m}(x^0) \end{pmatrix} = \operatorname{grad} f(x^0)$$

6.5.2 Felder, Divergenz, Rotation

Sei $G \subset \mathbb{R}^m$ offen.

- Eine Abbildung $\varphi : G \to \mathbb{R}$ heißt **skalares Feld** in G.
- Eine Abbildung $v : G \to \mathbb{R}^m$ heißt **Vektorfeld** in G.

Die Stetigkeit und Differenzierbarkeit von Feldern ist bereits erklärt. Im folgenden sei $\mathbb{R}^m$ mit dem Standardskalarprodukt und der Standard-Orthonormalbasis (e_i) versehen. Wenn φ ein skalares Feld in G ist, dann ist $\operatorname{grad}\varphi$ ein Vektorfeld in G.
Allgemein heißt φ **Potential** zum Vektorfeld v, wenn gilt:

$$v = \operatorname{grad}\varphi \text{ oder (v.a. in der Physik) } v = -\operatorname{grad}\varphi$$

Später werden wir sehen, wie man zu gegebenem v ein Potential bestimmen kann.

Sei v ein stetig differenzierbares Vektorfeld in G, dann heißt das skalare Feld

$$\operatorname{div} v := \frac{\partial v_1}{\partial x_1} + \ldots + \frac{\partial v_m}{\partial x_m}$$

Divergenz oder **Quellenfeld** von v. Wenn $\operatorname{div} v = 0$, heißt v **quellenfrei**.

Beispiel 6.18

Sei B ein magnetisches Feld. Ist $\operatorname{div} B = 0$, dann heißt das, dass keine magnetischen Ladungen existieren.

Sei jetzt $m = 3$ und v ein in G stetig differenzierbares Vektorfeld. Dann heißt das Vektorfeld

$$\operatorname{rot} v = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\right) e_1 + \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\right) e_2 + \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) e_3$$

Rotation des Vektorfeldes v. Ein Vektorfeld v mit $\operatorname{rot} v = 0$ heißt **wirbelfrei**.

$$\text{Eselsbrücke: } \operatorname{rot} v = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Mit dem Nablaoperator vereinfachen sich die Formeln für Gradient, Divergenz und Rotation:

$$\operatorname{grad}\varphi = \nabla\varphi$$

$$\operatorname{div} v = \nabla \cdot v = \langle \nabla, v \rangle = \sum_i \frac{\partial v_i}{\partial x_i}$$

$$\operatorname{rot} v = \nabla \times v$$

6.6 Lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten

Unter einer **linearen Differentialgleichung** (DGL) n-ter Ordnung mit konstanten Koeffizienten versteht man eine Gleichung der Form

$$x^{(n)} + a_{n-1}x^{(n-1)} + \ldots + a_1\dot{x} + a_0x = f \qquad (1)$$

Für $f = 0$ erhält man die zu (1) gehörende **homogene** DGL

$$x^{(n)} + \ldots + a_1\dot{x} + a_0 x = 0 \qquad (1_{hom})$$

Hierbei sind $a_i, f \in \mathbb{R}/\mathbb{C}$ gegeben. x ist eine mindestens stetige Funktion.

$$x = x(t) \text{ und } x^{(n)} = \frac{\mathrm{d}^n x}{\mathrm{d}t^n} \quad \text{(häufig ist } t \text{ die Zeit)}$$

Unter einer **Lösung** von (1) bzw. (1_{hom}) versteht man eine n-fach stetig differenzierbare reellwertige (bzw. komplexwertige) Funktion, definiert auf einem t-Intervall $I \subset \mathbb{R}/\mathbb{C}$, sodass (1) für alle $t \in I$ erfüllt ist.
Unter einem **Anfangswertproblem** (AWP) an der Stelle t_0 versteht man folgende Aufgabe: Gesucht ist eine Lösung von (1), die auf einem Intervall I definitert ist, sodass für ein fest vorgegebenes $t_0 \in I$ die folgenden Anfangsbedingungen erfüllt sind:

$$x(t_0) = x_0, \dot{x}(t_0) = x_1, \ldots, x^{(n-1)}(t_0) = x_{n-1}$$

Dabei sind $x_1, \ldots, x_{n-1}$ n vorgegebene Zahlen. Ein Anfangswertproblem ist also eine Differentialgleichung n-ter Ordnung mit n Anfangsbedingungen.

Beispiel 6.19 Schwingungsgleichung für eine eindimensionale Schwingung

$$m\ddot{x}(t) + d\dot{x}(t) + cx(t) = F(t)$$

Dabei ist m die Masse, d der Dämpfungsfaktor, c die Federkonstante und F die äußere Anregung. Umgeschrieben:

$$\ddot{x}(t) + 2k\dot{x}(t) + \omega_0^2 x(t) = f(t) \text{ mit } 2k = \tfrac{d}{m} \geq 0, \omega_0^2 = \tfrac{c}{m} > 0, f = \tfrac{F}{m}$$

Typische Aufgabenstellungen:

- Existenz von Lösungen von (1) bzw. (1_{hom})
- Struktur der Menge aller dieser Lösungen
- Existenz und Eindeutigkeit der Lösung des AWP

Es ist günstig, (1) bzw. (1_{hom}) in kompakterer Form aufzuschreiben. Abkürzung: $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$

$$(1), (1_{\text{hom}}): \quad \underbrace{\mathrm{D}^n x + a_{n-1}\mathrm{D}^{n-1}x + \ldots + a_1\mathrm{D}x + a_0 x}_{=P(\mathrm{D})x \text{ - Polynom}} = f \text{ bzw. } 0$$

$$P(\mathrm{D}) = \mathrm{D}^n + a_{n-1}\mathrm{D}^{n-1} + \ldots + a_1\mathrm{D} + a_0$$
$$\Rightarrow P(\mathrm{D})x = 0 \qquad (2) \quad (\widehat{=}\ (1_{\text{hom}}))$$

Hierbei sind $D : x \mapsto \frac{\mathrm{d}x}{\mathrm{d}t}$ und $D^k : x \mapsto \frac{\mathrm{d}^k x}{\mathrm{d}t^k}$ lineare Abbildungen, die Funktionen auf Funktionen abbilden.

Was bedeutet in dieser Symbolik, dass (1) bzw. (1_{hom}) **lineare** DGL sind?: $P(\mathrm{D})$ ist eine lineare Abbildung.

$P(\mathrm{D})(\lambda x + y) = \lambda P(\mathrm{D})x + P(\mathrm{D})y$ für zwei n-mal stetig differenzierbare Funktionen x, y und $\lambda \in \mathbb{R}/\mathbb{C}$

Lemma 6.33

Die Menge aller Lösungen der homogenen DGL (2) bzw. (1_{hom}) bildet einen Vektorraum.

Beweis
Seien x, y Lösungen (d.h. $P(\mathrm{D})x = P(\mathrm{D})y = 0$) und $\lambda \in \mathbb{C}$. Daraus folgt:

$$P(\mathrm{D})(\lambda x + y) = \lambda P(\mathrm{D})x + P(\mathrm{D})y = 0 \Rightarrow \lambda x + y \text{ ist Lösung}$$

■

$x \equiv 0$ ($\Leftrightarrow x(t) = 0 \ \forall t \in I$) heißt **triviale Lösung** von (2). Es gilt außerdem das **Superpositionsprinzip**:

1. Die Lösungen von (2) bilden einen Vektorraum.
2. Wenn x, y Lösungen von (1) sind, dann ist $x - y$ Lösung von (1_{hom}):

$$P(\mathrm{D})x = P(\mathrm{D})y = f \Rightarrow P(\mathrm{D})(x - y) = P(\mathrm{D})x - P(\mathrm{D})y = 0$$

3. Sei x_{inh} eine beliebige, aber feste Lösung von (1). Dann gilt: Jede beliebige Lösung von (1) hat die Form $x = x_{\text{inh}} +$ beliebige Lösung der homogenen DGL (1_{hom}).

Welche Dimension hat der Lösungsraum von (1_{hom})? Wenn er endlichdimensional ist, wie findet man eine Basis? Diese beiden Fragen können völlig befriedigend beantwortet werden. Dazu wird das zu $P(\mathrm{D})$ gehörende charakteristische Polynom betrachtet. Dieses hat die Form $p(\lambda)$, d.h.

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 \qquad (3)$$

Dieses Polynom muss auf Nullstellen und deren Vielfachheit untersucht werden.

Einige Bemerkungen zu Polynomen

1. **Fundamentalsatz der Algebra:** Jedes Polynom p vom Grade $n \geq 1$ hat in $\mathbb{C}$ mindestens eine Nullstelle.
2. Aus (1) folgt: Jedes Polynom n-ten Grades hat in $\mathbb{C}$ genau n Nullstellen, wenn man jede Nullstelle mit ihrer Vielfachheit zählt. λ_0 hat als Nullstelle die Vielfachheit $\alpha_0 \geq 1$, wenn

$$p(\lambda) = (\lambda - \lambda_0)^{\alpha_0} \cdot q(\lambda) \text{ mit } q(\lambda_0) \neq 0$$

Hilfsbetrachtung: Für jede Nullstelle λ_0 von p existiert ein Polynom r mit $p(\lambda) = (\lambda - \lambda_0)r(\lambda)$.

Beweis
Wende folgende Beziehung an:

$$\lambda^k - \lambda_0^k = (\lambda - \lambda^0)(\lambda^{k-1} + \lambda^{k-2}\lambda_0 + \ldots + \lambda\lambda_0^{k-2} + \lambda_0^{k-1}) \qquad (*)$$

Wenn $p(\lambda_0) = 0$, dann gilt:

$$p(\lambda) = p(\lambda) - p(\lambda_0) = (\lambda^n - \lambda_0^n) + a_{n-1}(\lambda^{n-1} - \lambda_0^{n-1}) + \ldots + a_1(\lambda - \lambda_0)$$

Setze $(*)$ für alle $(\lambda^k - \lambda_0^k)$ ein. Daraus folgt $p(\lambda) = (\lambda - \lambda_0)r(\lambda)$.

■

Betrachtet man die Nullstellen einschließlich ihrer Vielfachheit: $\mu_1, \ldots, \mu_n$, so hat jedes p die Darstellung

$$p(\lambda) = (\lambda - \mu_1)(\lambda - \mu_2)\ldots(\lambda - \mu_n)$$

Andere Zerlegung: Seien $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Nullstellen von p mit den Vielfachheiten $\alpha_1, \ldots, \alpha_k$, dann gilt offenbar

$$p(\lambda) = (\lambda - \lambda_1)^{\alpha_1} (\lambda - \lambda_2)^{\alpha_2} \cdot \ldots \cdot (\lambda - \lambda_k)^{\alpha_k}$$

Spezialfall: p habe nur reelle Koeffizienten. Trotzdem kann p komplexe Nullstellen haben, aber: Wenn $\lambda_0 \in \mathbb{C}$ eine Nullstelle von p mit der Vielfachheit α_0 ist, dann ist auch $\overline{\lambda_0} \in \mathbb{C}$ eine Nullstelle von p mit der Vielfachheit α_0. Das ist für die Anwendungen auf DGL wichtig! („Man kann dadurch komplexe Lösungen vermeiden.“)