Infinitesimalrechnung I und II

W. Müller

1997/98

Inhaltsverzeichnis

INFINITESIMALRECHNUNG I

Vorlesung Infinitesimalrechnung I

1. Grundbegriffe der Mengenlehre und Logik

1.1. Mathematische Logik und Beweistechnik

Inhalt dieses Abschnitts sind einige elementare Bemerkungen zur mathematischen Sprache. In der Mathematik haben wir es zunächst mit mathematischen Objekten (z.B. Zahlen) und Aussagen über diese Objekte zu tun. Aussagen nennt man in der Mathematik *Sätze* oder *Theoreme.* Weniger wichtige Aussagen werden als *Hilfssätze* oder *Lemmata* bezeichnet. Beispiele sind der Satz von Pythagoras und der Satz des Thales aus der elementaren Geometrie. Für jede Aussage in der Mathematik gibt es zwei Möglichkeiten: Sie ist entweder *wahr* oder *falsch,* d.h., wir haben es mit der sogenannten zweiwertigen Logik zu tun, in der gilt:

(1) Jede Aussage ist entweder wahr oder falsch (Satz vom ausgeschlossenen Dritten)
(2) Es gibt keine Aussage, die sowohl wahr als auch falsch ist(Satz vom ausgeschlossenen Widerspruch).

Die Klasse aller Aussagen zerfällt also in zwei disjunkte Teilklassen, die Klasse der wahren und die Klasse der falschen Aussage. Der Satz der Zweiwertigkeit besagt dabei lediglich, daß man von jeder Aussage *entscheiden* kann, ob sie *wahr* oder *falsch* ist. Aussagen, deren Wahrheitswert nicht bekannt ist, nennt man *Vermutungen.* So ist z.B. unbekannt, welchen Wahrheitswert die *Goldbachsche Vermutung* hat:

Jede gerade Zahl $n \geq 4$ läßt sich als Summe von zwei Primzahlen darstellen.

In der Mathematik werden neue Aussagen, d.h. also neue Sätze, mittels der mathematischen Schlußregeln aus bereits als *wahr* bekannten Aussagen abgeleitet. Eine solche Ableitung einer neuen Aussage aus bekannten Aussagen nennt man *Beweis.*

Eine der Hauptbeschäftigungen eines Mathematikers ist es, neue Sätze zu beweisen. Das Verfahren, neue Aussagen aus bereits bekannten wahren Aussagen abzuleiten, ist endlich. Am Anfang einer jeden Theorie stehen die *Grundannahmen* oder *Axiome,* die man im Rahmen der jeweiligen Theorie nicht weiter beweisen kann.

Alles zusammengenommen ergibt eine mathemtatische Theorie wie z.B. die Euklidische Geometrie, die Zahlentheorie, die Analysis,.... Im allgemeinen ist eine Theorie nicht abgeschlossen, d.h. nicht alle beweisbaren Sätze der Theorie sind bewiesen.

Inhalt der mathematischen Logik ist es, die grundlegenden Methoden der Mathematik zu formalisieren, mathematische Theorien zu untersuchen, u.s.w.

Bei jeder axiomatischen Theorie treten eine Reihe von wichtigen Fragen auf, die entschieden werden müssen:

a) Unabhängigkeit der Axiome
b) Widerspruchsfreiheit
c) Vollständigkeit.

Ein typisches Beispiel ist das Parallelaxiom in der Euklidischen Geometrie.

Einige elementare Dinge der mathemtischen Logik werden in der Vorlesung benutzt und sollen deshalb hier zusammengefaßt werden.

a) Aussagenlogik

In der zweiwertigen Logik ist jeder Aussage A eindeutig ein Wahrheitswert $w(A)$ zugeordnet, wobei entweder $w(A) = W$ (wahr) oder $w(A) = F$ (falsch) gilt. Aus gegebenen Aussagen kann man mit Hilfe von Bindewörtern neue Aussagen bilden. Dafür gibt es bestimmte Symbole in der Aussagenlogik, die von mir auch als Abkürzugen bei der Formulierung von mathematischen Aussagen benutzt werden. Es seien A und B Aussagen. Dann werden die folgenden Symbole häufig benutzt:

$$\neg A \qquad \text{“nicht } A\text{“, Verneinung der Aussage } A. \tag{0.1}$$

$$A \vee B \qquad \text{“}A \text{ oder } B\text{“} \tag{0.2}$$

$$A \wedge B \qquad \text{“}A \text{ und } B\text{“} \tag{0.3}$$

$$A \Rightarrow B \qquad \text{“aus } A \text{ folgt } B\text{“, “wenn } A,\text{ dann } B\text{“} \tag{0.4}$$

$$A \Leftrightarrow B \qquad \text{“}A \text{ ist äquivalent mit } B\text{“, “}A \text{ genau dann,wenn } B\text{“} \tag{0.5}$$

Dies sind **extensionale** Aussagenfunktionen. Der Wahrheitswert der Verknüpfung hängt nur vom Wahrheitswert der Argumente ab. Dies kann man sich durch eine Wahrheitstafel veranschaulichen.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

Bemerkungen:
1) Die angegebene Verwendung von $\neg$, $\wedge$, $\vee$, und $\Leftrightarrow$ entspricht etwa der im üblichen Sprachgebrauch. $\Rightarrow$ wird als **materiale Implikation** verwendet im Gegensatz zur **formalen Implikation** in der Logik. In der Mathematik wird "wenn A, so B" stets synonym für "nicht A, oder B" verwendet, d.h., es gilt die folgende Identität der Aussagenlogik

$$(A \Rightarrow B) \Leftrightarrow (\neg A \vee B).$$

2) In der mathematischen Literatur existieren für die materiale Implikation verschiedene andere Bezeichnungen und Sprechweisen: Voraussetzung und Behauptung, Prämisse und Conclusio, "A ist hinreichend für B", "B ist notwendig für A", "Aus A folgt B".

3) Die materiale Implikation hat nichts mit einem **Ursache-Wirkungs-Prinzip** zu tun. Ein solches Prinzip existiert nicht in der Mathematik.

Literatur: A. Tarski: "Einführung in die mathematische Logik".

Aus den Gesetzen der Aussagenlogik erhält man wichtige Schlußregeln, aus denen man von wahren Aussagen zu neuen wahren Aussagen gelangen kann.

Gesetze der Aussagenlogik (Auswahl)

Als Tautologie oder Gesetz der Aussagenlogik bezeichnet man Aussagenfunktionen, die für jede Wahl der Wahrheitswerte der Argumente den Wert "W" annehmen. Daraus ergeben sich z.B. wichtige Schlußregeln in der Logik. Wir geben einige Beispiele dafür an.

$((A \vee B) \wedge \neg A)) \Rightarrow B$

$A \vee \neg A$ (Satz vom ausgeschlossenen Dritten)

$\neg(A \wedge \neg A)$ (Satz vom ausgeschlossenen Widerspruch)

$\neg(\neg A) \Leftrightarrow A$ (Satz von der doppelten Verneinung)

$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ Gesetze von de

$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$ Morgan

$\neg(A \Rightarrow B) \Leftrightarrow A \wedge \neg B$

$\neg(A \Leftrightarrow B) \Leftrightarrow (A \Leftrightarrow \neg B)$

$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ (Kontrapositionsgesetz)

$A \wedge (A \Rightarrow B) \Rightarrow B$ (Abtrennungsregel)

$(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A$ (Gesetz zum modus tollens)

$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$ (Kettenschlußregel)

$A \Rightarrow A \vee B$

$B \Rightarrow A \vee B$

$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ (Distributivgesetze)

$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Der Beweis erfolgt mit Hilfe von Wahrheitstafeln. Als Beispiel betrachten wir $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$.

A	B	$A \Rightarrow B$	$\neg A \vee B$	$(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$
W	W	W	W	W
W	F	F	F	W
F	W	W	W	W
F	F	W	W	W

Es ist manchmal sinnvoll, sich die Struktur eines Beweises mit Hilfe von Schlußregeln klar zu machen.

Beweisschemata für direkte Beweise sind z.B.:

$$((A \Rightarrow B) \wedge \neg B) \Rightarrow \neg A$$

$$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$$

$$((A \Rightarrow B) \wedge A) \Rightarrow B.$$

b) Prädikatenlogik

Häufig kommen in der Mathematik auch Aussagen vor, die von freien Variablen abhängen. Beispiele sind

(1) "Es gibt eine rationale Zahl x, für die gilt: $x^2 = 2$"
(2) "Für alle reellen Zahlen x und y gilt: $x + y = y + x$."

Es handelt sich dabei um Aussagen der Prädikatenlogik. Aussage 1 ist bekanntlich falsch und 2 ist wahr. In der Prädikatenlogik benutzt man die *Quantoren* "es gibt ein" und "für alle", um aus einer Aussagenform $A(x)$ Aussagen zu bilden. Dafür verwendet man die Symbole "$\exists$" und "$\forall$." Obige Aussagen können dann wie folgt geschrieben werden.

(1) $\exists x \colon x \in \mathbb{Q} \wedge x^2 = 2, \quad \exists x \in \mathbb{Q} \colon x^2 = 2$
(2) $\forall x, y \in \mathbb{R} \colon x + y = y + x.$

Die Quantoren "$\exists$" und " $\forall$" werden von mir auch häufig zur Abkürzung bei der Formulierung von mathematischen Sätzen benutzt.

Gesetze der Prädikatenlogik

Als Gesetze der Prädikatenlogik bezeichnet man prädikatenlogische Ausdrücke, die bei jeder Belegung der Varaiblen den Wert "W" haben.

$$\exists x(A(x) \vee B(x)) \Leftrightarrow (\exists x A(x)) \vee (\exists x B(x))$$
$$\exists x(A(x) \wedge B(x)) \Rightarrow (\exists x A(x)) \wedge (\exists x B(x))$$
$$\forall x(A(x) \wedge B(x)) \Leftrightarrow (\forall x A(x)) \wedge (\forall x B(x))$$
$$\forall x\ A(x) \vee \forall x\ B(x) \Rightarrow \forall x(A(x) \vee B(x)$$

$$\neg(\forall x\ A(x)) \Leftrightarrow \exists x\ \neg A(x)$$
$$\neg(\forall x \neg\ A(x)) \Leftrightarrow \exists x\ A(x)$$
$$\neg(\exists x\ A(x)) \Leftrightarrow \forall x\ \neg A(x)$$
$$\neg(\exists x\ \neg A(x)) \Leftrightarrow \forall x\ A(x)$$

$$\forall x\ A(x) \Rightarrow A(x)$$
$$A(x) \Rightarrow \exists x\ A(x)$$

$$\exists x \forall y\ A(x, y) \Rightarrow \forall y \exists x\ A(x, y)$$
$$\exists x \exists y A(x, y) \Leftrightarrow \exists y \exists x A(x, y)$$
$$\forall x \forall y A(x, y) \Leftrightarrow \forall y \forall x\ A(x, y)$$

1.2. Mengenlehre

Im Folgenden werden wir häufig über Mengen sprechen. Deshalb ist es wichtig, uns eine gewisse Klarheit über den Begriff der Menge in der Mathematik zu verschaffen.

Die Mengenlehre dient heute weitgehend als Grundlage der Mathematik. Viele fundamentale Begriffe und Methoden lassen sich auf die Mengenlehre zurückführen.

Die Mengenlehre wurde am Ende des vorigen Jahrhunderts durch den deutschen Mathematiker *Georg Cantor* (1845-1918) begründet. Von Cantor wurde 1895 folgende Definition einer Menge formuliert: "Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen."

Da hierdurch der zu definierende Begriff "Menge" nicht auf andere definierte Begriffe zurückgeführt wird, handelt es sich nicht um eine Definition im mathematischen Sinne, sondern um eine Beschreibung, welcher Sachverhalt vorliegen muß, damit von einer Menge gesprochen werden kann.
Beispiele:

(1) Menge der Einwohner von Bonn. Dies ist eine Menge von wohlunterschiedenen Objekten unserer Anschauung.
(2) Menge der natürlichen Zahlen. Dies ist eine Menge von Objekten unseres Denkens.

Mengenbildungsprinzip: Ist eine solche Eigenschaft E gegeben, daß jedes Objekt eines bestimmten Grundbereiches diese Eigenschaft besitzt oder nicht besitzt, so sagt man, alle Dinge, die diese Eigenschaft besitzen, bilden eine Menge M. Sie werden Elemente der Menge genannt.

$$M = \{x \mid E \text{ gilt für } x\}.$$

Bezeichnung:
$x \in M := x$ ist Element der Menge M;
$x \notin M := x$ ist nicht Element der Menge M.
$\{a, b, c, ...\}$ die Menge, die aus den Elementen $a, b, c, ...$ besteht.

Axiomatische Mengenlehre

Die naive Mengenlehre führt zu Widersprüchen. Z.B. ist die Menge aller Mengen Element von sich selbst. Um diese Probleme zu vermeiden, geht man zur axiomatischen Mengenlehre über.

Die elementaren Begriffe dieser Theorie sind der Begriff der **Menge** und die Relation, **Element einer Menge** zu sein.

Axiom 1: Zwei Mengen A und B sind genau dann gleich, wenn sie die gleichen Elemente enthalten.

$$(A = B) \Leftrightarrow (\forall x\colon x \in A \Leftrightarrow x \in B).$$

Axiom 2: Für je zwei Mengen A und B existiert eine Menge $A \cup B$, die Vereinigung von A und B, die genau die Elemente A und B enthält, und sonst keine weiteren.

$$A \cup B := \{x \mid x \in A \vee x \in B\}.$$

Axiom 3: Für je zwei Mengen A und B existiert eine Menge $A \setminus B$, die genau die Elemente von A enthält, die nicht in B liegen.

$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}.$$

Axiom 4: Es existiert wenigstens eine Menge.

Die weiteren Axiome besprechen wir später. Mit Hilfe der bisherigen Axiome kann man den Durchschnitt von zwei Mengen A und B definieren als

$$A \cap B := A \setminus (A \setminus B).$$

Dann gilt

$$A \cap B = \{x \mid x \in A \wedge x \in B\}.$$

Aus Axiom 3 und 4 folgt, daß die Menge $A \setminus A$ existiert, die kein Element enthält. Dies ist die **leere Menge**, die man mit $\emptyset$ bezeichnet. Es gilt:

$$\forall x\colon x \notin \emptyset.$$

Wenn für zwei Mengen A und B gilt:

$$A \cap B = \emptyset$$

so nennt man A und B **disjunkt**.

Es gelten eine Reihe von Regeln für die Verknüpfung von $\cap$, $\cup$ und $\setminus$. So z.B.

Satz 1.1.

(1) *Kommutativgesetz*

$$A \cup B = B \cup A$$

(2) *Assoziativgesetz*

$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A \cap (B \cap C) = (A \cap B) \cap C$$

(3) *Distributivgesetz*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Beweis: Der Beweis dieser Aussagen erhält man aus den entsprechenden Gesetzen der Aussagenlogik. Als Beispiel beweisen wir das Distributivgesetz. Es gilt folgende Schlußkette.

$$\begin{aligned} x \in (A \cap (B \cup C) &\Leftrightarrow (x \in A) \wedge (x \in (B \cup C)) \\ &\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\ &\Leftrightarrow ((x \in A) \wedge (x \in B)) \vee ((x \in A) \wedge (x \in C)) \\ &\Leftrightarrow \big((x \in A \cap B) \vee (x \in A \cap C)\big) \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

Analoge Gleichungen gelten für $\setminus$. Eine weitere Relation zwischen Mengen A und B ist die **Inklusion** $A \subseteq B$, die definiert ist durch

$$A \subseteq B =_{Df} \forall x\colon (x \in A \Rightarrow x \in B).$$

Satz 1.2. *Für alle Mengen A und B gilt:*

$$(A \subseteq B) \wedge (B \subseteq A) \Rightarrow A = B.$$

Beweis:

$$\begin{aligned} (A \subseteq B) \wedge (B \subseteq A) &\Leftrightarrow \forall x(x \in A \Rightarrow x \in B) \wedge (x \in B \Rightarrow x \in A) \\ &\Leftrightarrow \forall x(x \in A \Leftrightarrow x \in B) \Leftrightarrow A = B. \end{aligned}$$

Bei der letzten Gleichung haben wir das Axiom 1 benutzt.

Satz 1.3. *(Transitivität) Aus $A \subseteq B$ und $B \subseteq C$ folgt $A \subseteq C$.*

Wir definieren jetzt noch die folgende Relation:

$$A \subsetneq B =_{Df} A \subseteq B \text{ und} B \setminus A \neq \emptyset.$$

Mengen von Mengen

Man möchte in der Mengentheorie auch Mengen bilden können, deren Elemente wieder Mengen sind. Dafür sind weitere Axiome nötig. Es gibt verschiedene Axiomensysteme.

(1) **Das Axiomensytem von Russell.** Dabei handelt es sich um eine Typentheorie. Es wird eine Hierarchie von Mengen eingeführt.

Individuen

|

Mengen (Mengen 1. Stufe)

|

Mengensysteme (Mengen 2. Stufe)

|

.

.

.

(2) **Das Axiomensystem von Zermelo-Fraenkel**. Es handelt sich dabei um eine typenfreie Mengenlehre.

(3) **Das Axiomensystem von Bernays-Gödel**.

Wir benutzen das Axiomensystem von Zermelo-Fraenkel.

Axiom 5 (Potenzmengenaxiom) Für jede Menge M gibt es eine Menge $\mathcal{P}(M)$, deren Elemente genau alle Teilmengen von M sind.

$$\forall X(X \in \mathcal{P}(M) \Leftrightarrow X \subset M).$$

Beispiel: Es sei $M = \{a, b, c\}$. Dann ist

$$\mathcal{P}(M) = \big\{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\big\}.$$

Axiom 6 (Unendlichkeitsaxiom) Es gibt eine Menge, die $\emptyset$ enthält, und mit jedem X auch $X \cup \{X\}$.

Axiom 7 (Fundierungsaxiom) Jede nichtleere Menge X enthält ein Element y, daß mit X kein Element gemeinsam hat.

$$\forall X(X \neq \emptyset \Rightarrow \exists y(y \in X \wedge y \cap X = \emptyset).$$

Daraus ergibt sich

Satz: Es gibt keine Menge X mit $X \in X$.

Beweis: Es sei X eine Menge, für die gilt $X \in X$. Dann ist $X \cap \{X\} \neq \emptyset$. Nach dem Fundierungsaxiom existiert $y \in \{X\}$ mit $y \cap \{X\} = \emptyset$. Aus dem Axiom 1 folgt $y = X$. Daher ist $X \cap \{X\} = \emptyset$. Dies ist aber ein Widerspruch zu unserer Annahme über X.

$\square$

Axiom 8 (Auswahlaxiom) Für jede Familie $\mathcal{F}$ paarweise disjunkter Mengen existiert eine Menge X, die mit jeder Menge von $\mathcal{F}$ genau ein Element gemeinsam hat.

Bemerkung: Das Auswahlaxiom wird nicht von allen Mathematikern akzeptiert. Einerseits kann mit Hilfe des Auswahlaxioms abstruse Resultate beweisen. Anderseits ist das Auswahlaxiom von großer Tragweite für die Mathematik. Ohne das Auswahlaxiom kann man im Rahmen der Zermelo-Fraenkel-Axiomatik nicht beweisen, daß jeder Vektorraum eine Basis besitzt oder, daß man jeden Körper algebraisch abschließen kann. Das Auswahlaxiom ist äquivalent zum Zornschen Lemma.

Geordnete Paare und kartesisches Produkt

Definition 1.4. *Seien A und B Mengen und sei $a \in A$, $b \in B$. Dann ist das geordnete Paar $\langle a, b\rangle$ definiert durch*

$$\langle a, b\rangle =_{Df} \{\{a, b\}, \{a\}\}\,.$$

Lemma 1.5. *Es gilt*

$$\langle a, b\rangle = \langle x, y\rangle \Leftrightarrow (a = x \wedge b = y).$$

Definition 1.6. *A und B seien Mengen. Dann ist das kartesische Produkt $A \times B$ definiert durch*

$$A \times B =_{Df} \{\langle a, b\rangle \mid a \in A \wedge b \in B\}\,.$$

Abbildungen von Mengen

Es seine X und Y Mengen. Im allgemeinen versteht man unter einer Abbildung $f : X \to Y$ von X nach Y eine Vorschrift, die jedem Element von X eindeutig ein Element von Y zuordnet. Dies ist aber keine Definition im mathematischen Sinne. Mit Hilfe der Mengenlehre kann man den Abbildungsbegriff wie folgt definieren.

Definition 1.7. *Es seien X und Y zwei Mengen. Eine Abbildung f von X nach Y ist eine Teilmenge $R \subset X \times Y$ mit folgenden Eigenschaften*

(1) $\forall x \in X\ \exists y \in Y\colon \langle x, y\rangle \in R$
(2) $\forall x \in X\ \forall y_1, y_2 \in Y\colon\ (\langle x, y_1\rangle \in R \wedge \langle x, y_2\rangle \in R) \Rightarrow y_1 = y_2.$

Anders formuliert besagt die Bedingung 2: Es sei $x \in X$. Dann existiert genau ein $y \in Y$ mit $\langle x, y\rangle \in R$.

Bezeichnung: $y = f(x)$, $f : X \to Y$.

Die Menge $X = D(f)$ heißt **Definitionsbereich** von f.

Die Menge $W(f) = \{y \in Y \mid \exists x \in X\colon \langle x, y\rangle \in R\}$ heißt **Wertevorrat** von f.

Definition 1.8. *Es sei $f\colon X \to Y$ eine Abbildung. Dann heißt f eineindeutig, wenn gilt*

$$\forall x_1, x_2 \in X\colon f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

Definition 1.9. *Zwei Mengen X und Y heißen* **gleichmächtig**, *wenn es eine 1-1-deutige Abbildung $f\colon X \to Y$ gibt mit $W(f) = Y$.*

Endliche Mengen sind gleichmächtig genau dann, wenn sie die gleiche Anzahl von Elementen enthalten. Naiv kann man die Mächtigkeit einer Menge M definieren als die Klasse aller Mengen, die mit M gleichmächtig sind. dies führt aber wieder auf logische Schwierigkeiten und man muß die Mächtigkeit anders einführen. Wir verweisen dazu auf das Buch von Ebinghaus.
Es sei $\mathbb{N} = \{1, 2, 3, ...\}$ die Menge der natürlichen Zahlen.

Definition 1.10. *Eine Menge heißt abzählbar, wenn sie gleichmächtig mit $\mathbb{N}$ ist.*

Literatur: H.-D. Ebinghaus: “Einführung in die Mengenlehre.

2. Die reellen Zahlen

Die reellen Zahlen bilden die Grundlage der Analysis. Sie sind eine wesentliche Erweiterung des Zahlbegriffes.

2.1. Aufbau des Zahlsystems

Wir verwenden folgende Bezeichnungen:

$\mathbb{N} = \{1, 2, 3, ...\}$ die Menge der natürlichen Zahlen

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

$\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$ die Menge der ganzen Zahlen

$\mathbb{Q}$ die Menge der rationalen Zahlen

$\mathbb{R}$ die Menge der reellen Zahlen

$\mathbb{C}$ die Menge der komplexen Zahlen.

Jeder dieser Zahlbereiche ist eine Erweiterung des vorhergehenden Bereiches, d.h., es gilt folgende Kette von Inklusionen:

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Die Erweiterung eines Zahlbereiches wird unter anderem dadurch motiviert, daß man im neuen Bereich Gleichungen lösen kann, die man im bisherigen Zahlbereich nicht lösen konnte.

Beispiele:

(1) In $\mathbb{Z}$ kann man die Gleichung

$$x + a = b$$

für alle $a, b \in \mathbb{N}$ lösen. In $\mathbb{N}$ ist dies nicht möglich.

(2) In $\mathbb{Q}$ kann man die Gleichung

$$ax + b = c$$

für alle $a, b, c \in \mathbb{Z}$ mit $a \neq 0$ lösen, aber nicht in $\mathbb{Z}$.

(3) die Gleichung

$$x^2 = 2$$

hat in $\mathbb{R}$ eine Lösung. In $\mathbb{Q}$ ist sie nicht lösbar.

(4) Die Gleichung

$$x^2 = -1$$

hat in $\mathbb{R}$ keine Lösung, ist aber in $\mathbb{C}$ lösbar.

Der Aufbau des Zahlsystems bis zu den rationalen Zahlen erfolgt im Rahmen der allgemeinen Mengenlehre und wird hier nicht behandelt. Wir setzen rationale Zahlen als gegeben voraus. Reelle Zahlen kann man sich intuitiv z.B. wie folgt vorstellen:

(1) **Als unendliche Dezimalbrüche** Z.B. ist

$$\pi = 3,141592653589793...$$

Dabei ist zu beachten, daß der unendliche Dezimalbruch auf der rechten Seite nicht die reelle Zahl π ist, sondern deren Darstellung im Dezimalsystem.

(2) **Als Punkte auf der Zahlengeraden**

Jeder rationalen Zahl entspricht ein Punkt auf der Zahlengeraden. Wenn wir alle rationalen Zahlen als Punkte betrachten, so sind "Lücken auf der Zahlengeraden vorhanden.

Beispiel: $\sqrt{2}$ ist nicht rational.

Beweis: Wir nehmen an, es existieren $p, q \in \mathbb{Z}$ mit $(p, q) = 1$, so daß $\sqrt{2} = p/q$ gilt.

$$\Rightarrow p^2 = 2q^2 \Rightarrow 2|p^2 \Rightarrow 2|p$$
$$\Rightarrow 4|p^2 \Rightarrow 2|q^2 \Rightarrow 2|q.$$

Wir haben damit gezeigt, daß p und q durch 2 teilbar sind. Dies steht im Widerspruch zur Annahme, daß p und q teilerfremd sind. Deshalb kann $\sqrt{2}$ nicht rational sein.

Für das Ausfüllen der Lücken der Zahlengeraden, d.h., für die Konstruktion der reellen Zahlen, gibt es verschiedene Möglichkeiten.

(1) Dedekindsche Schnitte in $\mathbb{Q}$;
(2) Intervallschachtelungen in $\mathbb{Q}$;
(3) Cauchyfolgen rationaler Zahlen.

Die genaue Ausführung jedes dieser Verfahren ist langwierig und kann hier nicht in allen Einzelheiten behandelt werden. Wir werden die Grundzüg der Konstruktionsmethode der reellen Zahlen durch Dedekindsche Schnitte in $\mathbb{Q}$ kennenlernen. Im Augenblick stellen wir uns auf den Standpunkt, daß wir über eine gewisse Vertrautheit im Umgang mit den reellen Zahlen verfügen und charakterisieren die reellen Zahlen durch ihre Eigenschaften, d.h., wir nehmen den axiomatischen Standpunkt ein.

2.2. Die Körperstruktur der reellen Zahlen

Die reellen Zahlen sind eine Menge $\mathbb{R}$, in der zwei Operationen, die Addition $+$ und die Multiplikation $\cdot$, erklärt sind, d.h., jedem Paar $\langle a, b\rangle$ von Elementen aus $\mathbb{R}$ ist genau ein Element $a+b \in \mathbb{R}$, die Summe, und genau ein Element $a \cdot b \in \mathbb{R}$ das Produkt, zugeordnet. Für diese Operationen gelten die folgenden Axiome:

(**K1**) Kommutativgesetz
Für alle $a, b \in \mathbb{R}$ gilt
$$a + b = b + a$$
$$a \cdot b = b \cdot a.$$

(**K2**) Assoziativgesetz
Für alle $a, b, c \in \mathbb{R}$ gilt:
$$(a + b) + c = a + (b + c)$$
$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$

(**K3**) Distributivgesetz
Für alle $a, b.c \in \mathbb{R}$ gilt:
$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

(**K4**) Existenz neutraler Elemente
i) Es gibt in $\mathbb{R}$ ein Element 0 mit
$$a + 0 = a$$
für alle $a \in \mathbb{R}$. ii) Es gibt in $\mathbb{R}$ ein Element $1 \neq 0$ mit:
$$a \cdot 1 = a$$
für alle $a \in \mathbb{R}$.

(**K5**) Existenz inverser Elemente
i) Für jedes $a \in \mathbb{R}$ existiert ein additiv inverses Element $(-a)$, so daß
$$a + (-a) = 0.$$
ii) Für jedes $a \in \mathbb{R} \backslash \{0\}$ existiert ein multiplikativ inverses Element $a^{-1} \in \mathbb{R}$, so daß
$$a \cdot (a^{-1}) = 1.$$

Schreibweisen: $a + (-b) = a - b,\ a \cdot (b^{-1}) = \frac{a}{b}$.

Die Axiome $(K1) - (K5)$ sind die sogenannten Körperaxiome. Eine Menge K, in der zwei Operationen $+$ und $\cdot$ definiert sind, so daß die Axiome $(K1) - (K5)$ gelten, nennt man Körper. Die reellen Zahlen bilden also insbesondere einen Körper. Ein anderes Beispiel für einen Körper ist der Körper der rationalen Zahlen $\mathbb{Q}$. Es gibt aber auch Körper, die nur endlich viele Elemente enthalten. So z.B. der Körper

$\mathbb{F}_2 = \{0, 1\}$, der aus 2 Elementen besteht. Addition und Multiplikation in $\mathbb{F}_2$ sind definiert durch

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

Aus den Axiomen $(K1) - (K5)$ kann man eine Reihe von bekannten Rechenregeln und weiteren Eigenschaften ableiten wie z.B.:

(1) Für alle $a, b \in \mathbb{R}$ existiert genau ein $x \in \mathbb{R}$ mit $a + x = b$. Und zwar ist $x = b - a$.
Beweis: Es sei $x = b - a$. Dann gilt

$$\begin{aligned} a + x &= a + (b - a) = (b - a) + a \\ &\underset{K2}{=} b + ((-a) + a) \underset{K1}{=} b + (a + (-a)) \underset{K4}{=} b. \end{aligned}$$

Sei umgekehrt x Lösung der Gleichung $a + x = b$. Dann gilt

$$\begin{aligned} b - a &= b + (-a) = (a + x) + (-a) \\ &\underset{K1}{=} (x + a) + (-a) \underset{K2}{=} x + (a + (-a)) \underset{K4}{=} x. \end{aligned}$$

Ebenso beweist man die folgenden Rechenregeln:

(2) $-(-a) = a,\ (-a) + (-b) = -(a + b)$.
Für $a, b \neq 0$ gilt

$$(a^{-1})^{-1} = a,\ a^{-1} \cdot b^{-1} = (a \cdot b)^{-1}.$$
$$a \cdot 0 = 0,\quad a \cdot (-b) = -(a \cdot b).$$
$$(-a) \cdot (-b) = a \cdot b,\ a(b - c) = ab - ac.$$

(3) $\forall a, b \in \mathbb{R}\colon\ a \cdot b = 0 \Leftrightarrow a = 0 \vee b = 0$.

(4) Die neutralen Elemente 0 und 1 sind eindeutig bestimmt: Es seien $0'$ und $1'$ andere neutrale Elemente. Dann gilt

$$0' = 0' + 0 = 0,\ \ 1' = 1' \cdot 1 = 1.$$

2.3. Die Anordnung von $\mathbb{R}$

Wir wissen bereits, daß es für rationale Zahlen die Relation $x < y$ gibt. Insbesondere können wir von positiven und negativen Zahlen sprechen. Diese Ordnungsrelation überträgt sich auf die reellen Zahlen und wird wie folgt axiomatisiert:

In der Menge der reellen Zahlen ist eine Teilmenge $\mathbb{R}_+$ ausgezeichnet, die wir die Menge der positiven Zahlen nennen. Für $a \in \mathbb{R}_+$ verwenden wir die Bezeichnung $a > 0$. Es gelten folgende Axiome:

(**A1**) Für alle $a \in \mathbb{R}$ gilt genau eine der folgenden Relationen:

$$a > 0,\ a = 0,\ (-a) > 0.$$

(**A2**) Aus $a > 0$ und $b > 0$ folgt $a + b > 0$ und $a \cdot b > 0$.

Ein Körper K, in dem eine Ordnungsrelation $<$ definiert ist, so daß die Axiome $(A1)$ und $(A2)$ gelten, nennt man einen **geordneten Körper**. Wir verwenden folgende Bezeichnungen:

$$\begin{aligned} a > b &\Leftrightarrow a - b > 0 \\ a < b &\Leftrightarrow b - a > 0 \\ a \geq b &\Leftrightarrow a > b \vee a = b \\ a \leq b &\Leftrightarrow a < b \vee a = b. \end{aligned}$$

Es gelten folgende Rechenregeln:
Für alle $a, b, c, d, \in \mathbb{R}$ gilt

(1) Transitivität
$a < b \wedge b < c \Rightarrow a < c.$
(2) $a < b \Rightarrow a + c < b + c$
(3) $a < b \Rightarrow -a > -b$
(4) $a < b \wedge c > 0 \Rightarrow a \cdot c < b \cdot c$
(5) $a < b \wedge c < 0 \Rightarrow a \cdot c > b \cdot c$
(6) $a \neq 0 \Rightarrow a^2 > 0$
(7) $a > 0 \Rightarrow a^{-1} > 0$
(8) $a < 0 \Rightarrow a^{-1} < 0$
(9) $0 < a < b \Rightarrow a/b < 1 \wedge b/a > 1$
(10) $0 < a < b \Rightarrow a^{-1} > b^{-1}$
(11) $a < b \wedge c < d \Rightarrow a \cdot c < b \cdot d$

In $\mathbb{R}$ gilt auch das sogenannte Archimedische Axiom.

(**A3**)**Archimedisches Axiom:** $\forall a \in \mathbb{R}\ \exists n \in \mathbb{N}\colon n > a.$

Einen geordneten Körper, für den zusätzlich (A3) gilt, nennt man einen **archimedisch geordneten Körper**.

Bernoullische Ungleichung:
Für alle $x \in \mathbb{R}$ mit $x > -1$ und alle $n \in \mathbb{N}$ gilt

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

Beweis: Der Beweis erfolgt durch vollständige Induktion.
Induktionsanfang: Für $n = 1$ ist die Behauptung richtig.
Induktionsschritt: Für $n \in \mathbb{N}$ gelte $(1+x)^n \geq 1 + nx$. Dann folgt

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

□

Definition. *Für $a \in \mathbb{R}$ ist der Absolutbetrag $|a|$ von a definiert durch*

$$|a| = \begin{cases} a, & \text{falls } a \geq 0; \\ -a, & \text{falls } a < 0. \end{cases}$$

Für den Absolutbetrag gelten folgende Rechenregeln. Für alle $a, b \in \mathbb{R}$ ist:

(1) Für $a \neq 0$ ist $|a| > 0$.
(2) $|a \cdot b| = |a| \cdot |b|$
(3) $|a + b| \leq |a| + |b|$ (Dreiecksungleichung)
(4) $||a| - |b|| \leq |a - b|$

2.4. Die Vollständigkeit von $\mathbb{R}$.

Definition. *Eine Teilmenge $M \subseteq \mathbb{R}$ heißt* **nach oben** *(bzw.* **unten***)* **beschränkt***, wenn ein $s \in \mathbb{R}$ existiert so, daß gilt:*

$$\forall x \in M \colon x \leq s \text{ (bzw. } x \geq s).$$

s heißt dann **obere** (bzw. **untere Schranke**) von M. Die Menge M heißt **beschränkt**, wenn sie nach oben und unten beschränkt ist.

Definition. *Es sei $M \subseteq \mathbb{R}$. Eine reelle Zahl $s \in \mathbb{R}$ heißt* **Suprenum** *(bzw.* **Infinum***) von M, wenn es die kleinste obere Schranke (bzw. die größte untere Schranke) von M ist.*

Es gilt also:

a) $s \in \mathbb{R}$ ist Suprenum von $M \Leftrightarrow s$ ist obere Schranke von M und für alle $s' \in \mathbb{R}$ mit $s' < s$ existiert ein $x \in M$ mit $s' < x$.
b) $s \in \mathbb{R}$ ist Infinum von $M \Leftrightarrow s$ ist untere Schranke von M und für alle $s' \in \mathbb{R}$ mit $s' > s$ existiert $x \in M$ mit $x < s'$.

Wenn das Suprenum bzw. Infinum einer Menge M exsistieren, so sind sie eindeutig bestimmt und werden mit

$$\sup M \quad \text{bzw.} \quad \inf M$$

bezeichnet.
Wenn $\sup M \in M$ (bzw. $\inf M \in M$), so heißt $\sup M$ **Maximum** von M (bzw. $\inf M$ **Minimum** von M).

(**V**) **Vollständigkeitsaxiom:** *Jede nichtleere, nach oben beschränkte Teilmenge M von $\mathbb{R}$ besitzt ein Suprenum.*

Einen geordneten Körper, für den das Axiom V gilt, nennt man einen vollständigen geordneten Körper.

Eine ander Möglichkeit, die Vollständigkeit von $\mathbb{R}$ zu axiomatisieren, ist das Intervallschachtelungsprinzip.

Definition. *Eine* **Intervallschachtelung** *ist eine Folge (I_k) von kompakten Intervallen $I_k = [a_k, b_k]$ mit folgenden Eigenschaften:*

(1) *Für alle $k \in \mathbb{N}$ gilt: $I_{k+1} \subset I_k$.*
(2) *Zu jedem $\epsilon > 0$ gibt es ein Intervall I_k mit einer Länge $|I_k| < \epsilon$.*

In $\mathbb{R}$ gilt das

(**I**) **Intervallschachtelungsprinzip:** Für jede Intervallschachtelung (I_k) existiert genau eine reelle Zahl $r \in \mathbb{R}$, die in allen Intervallen I_k enthalten ist.

Bewies: Es sei $([a_k, b_k])$ eine Intervallschachtelung. Dann ist die Menge $A = \{a_1, a_2, a_3, ...\}$ nach oben beschränkt. Obere Schranken sind alle b_k. Für die kleinste obere Schranke s gilt

$$a_k \leq s \leq b_k$$

für alle $k \in \mathbb{N}$. Daher liegt $s = \sup A$ in allen Intervallen $[a_k, b_k]$. Es sei s' eine weitere Zahl, die in allen Intervallen $[a_k, b_k]$ liegt. Wir nehmen an, daß $s \neq s'$. Sei $\epsilon = |s - s'|/2$. Nach Voraussetzung existiert $k \in \mathbb{N}$ mit $b_k - a_k < \epsilon$. Da $s, s' \in [a_k, b_k]$, folgt

$$|s - s'| \leq b_k - a_k < \epsilon = |s - s'|/2.$$

Dies ist ein Widerspruch und daher muß $s = s'$ gelten.

$\square$

Umgekehrt folgt auch die Supremumseigenschaft aus dem Intervallschachtelungsprinzip.

2.5. Konstruktion der reellen Zahlen

Die rationalen Zahlen haben “Lücken“ (auf dem Zahlenstrahl) und es gibt 3 Methoden, diese Lücken zu “füllen“ d.h. $\mathbb{Q}$ zu vervollständigen.

(1) Dedekindsche Schnitte
R. Dedekind (1831 - 1916)
(2) Intervallschachtelungen
(3) Cauchy - Folgen

Definition 2.12. *Ein Dedekindscher Schnitt $A \mid B$ in $\mathbb{Q}$ ist eine Zerlegung von $\mathbb{Q}$ in 2 Teilmengen A und B mit folgenden Eigenschaften:*

(1) $A \neq \emptyset,\ B \neq \emptyset$
(2) $A \cup B = \mathbb{Q}$
(3) $\forall a \in A,\ \forall b \in B\colon a < b.$

Beispiele: 1)

$$\begin{aligned} A &= \{a \in \mathbb{Q} \mid a < 1\} \\ B &= \{a \in \mathbb{Q} \mid a \geq 1\} \end{aligned}$$

Es giltinf $B = 1$, und in A existiert kein größtes Element.
2)

$$\begin{aligned} A &= \{a \in \mathbb{Q} \mid a > 0 \wedge a^2 < 2\} \cup \{a \in \mathbb{Q} \mid a \leq 0\} \\ B &= \{a \in \mathbb{Q} \mid a > 0 \wedge a^2 > 2\} \end{aligned}$$

Da $\sqrt{2}$ irrational ist, definieren A und B einen Schnitt $A \mid B$ in $\mathbb{Q}$.
Behauptung: i) In A gibt es keine größte Zahl.
ii) In B gibt es keine kleinste Zahl.
Beweis: i) Es sei $a > 0$ und $a^2 < 2,\ a \in \mathbb{Q}$. Nach dem Axiom von Archimedes für $\mathbb{Q}$ existiert $n \in \mathbb{N}$: $n > \frac{2a+1}{2-a^2}$

$$\begin{aligned} &\Rightarrow \frac{2a+1}{n} < 2 - a^2 \Rightarrow \frac{2a}{n} + \frac{1}{n^2} \leq \frac{2a+1}{n} < 2 - a^2 \\ &\Rightarrow \left(a + \frac{1}{n}\right)^2 = a^2 + \frac{2a}{n} + \frac{1}{n^2} < 2. \\ &\Rightarrow a + \frac{1}{n} \in A. \end{aligned}$$

Ebenso beweist man ii).

□

Es gibt 3 Möglichkeiten für einen Schnitt $A \mid B$:

(1) A enthält kein größtes Element, aber B enthält ein kleinstes Element.
(2) A enthält ein größtes Element, aber B enthält kein kleinstes Element.
(3) A enthält kein größtes Element und B enthält kein kleinstes Element.

Vierte Möglichkeit: A enhält ein größtes Element a und B enthält ein kleinstes Element b. Dann ist $a < b$. Daraus folgt, daß $c \in \mathbb{Q}$ existiert mit $a < c < b$. Da $a < c$ ist, ist $c \notin A$, und aus $c < b$ folgt $c \notin B$. Hieraus folgt $c \notin A \cup B = \mathbb{Q}$. Dies ist ein Widerspruch. Daher gibt es den 4. Fall nicht.
Im Falle i) oder ii) wird der Schnitt $A \mid B$ durch die rationale Zahl r definiert.
Festlegung: Wir wählen für eine rationale Zahl r den Schnitt $A \mid B$ so, daß $r \in B$.

$$A = \{q \in \mathbb{Q} \mid q < r\}, \ B = \{q \in \mathbb{Q} \mid q \leq r\}.$$

Defintion 2.13. *Ein Schnitt der Form 3) definiert eine irrationale Zahl α. Es sei*

$$\mathbb{R} = \mathbb{Q} \cup \{\alpha \mid \alpha \text{ irrational}\}.$$

Man kann jetzt zeigen, daß für die so definierte Menge $\mathbb{R}$ die Axiome (K1)-(K5), (A1), (A2) und (V) gelten. Die Ordnungsrelation kann man wie folgt einführen:

Definition 2.14. *Seien $\alpha, \beta \in \mathbb{R}$ definiert durch die Schnitte $A \mid B$ bzw. $A' \mid B'$. Dann ist $\alpha < \beta$ definiert durch*

$$\alpha < \beta \Leftrightarrow A \subsetneq A'$$

Lemma 2.15. *1) $\forall \alpha, \beta \in \mathbb{R}$ gilt genau eine der folgenden Relationen: $\alpha = \beta, \ \alpha > \beta, \ \alpha < \beta$ 2) $\forall \alpha, \beta, \gamma \in \mathbb{R}$: $\alpha < \beta \wedge \beta < \gamma \Rightarrow \alpha < \gamma$*

Beweis: Übung!

Lemma 2.16. *Für alle $\alpha, \beta \in \mathbb{R}$ mit $\alpha > \beta$ existiert $r \in \mathbb{Q}$: $\alpha > r > \beta$.*

Beweis: Seien α und β definiert durch $A \mid B$ bzw. $A' \mid B'$. Da $\alpha > \beta$, ist $A' \subsetneq A$. Daher existiert $r \in \mathbb{Q}$ so, daß

$$r \in A \wedge r \notin A'.$$

Da $A'|B'$ ein Schnitt in $\mathbb{Q}$ ist, folgt $r \in B'$. Dies bedeutet $\alpha > r > \beta$.

□

Defintion 2.17. *Ein Schnitt $A \mid B$ in $\mathbb{R}$ ist eine Zerlegung von $\mathbb{R}$ in zwei Teilmengen A und B mit folgenden Eigenschaften:*

(1) $A \neq \emptyset,\ B \neq \emptyset$
(2) $A \cup B = \mathbb{R}$
(3) $\forall a \in A,\ \forall b \in B \colon a < b$

Theorem. *(Dedekind): Für jeden Schnitt $A \mid B$ in $\mathbb{R}$ existiert $\gamma \in \mathbb{R}$ mit*

(1) $\gamma = \sup A = \inf B$
(2) $\gamma \in A\ \vee\ \gamma \in B$

Beweis: Es sei $A_1 = A \cap \mathbb{Q}$ und $B_1 = B \cap \mathbb{Q}$.
Behauptung: Die Teilmengen A_1, B_1 von $\mathbb{Q}$ definieren einen Schnitt $A_1 \mid B_1$ in $\mathbb{Q}$.
Beweis: a) Es gilt $A_1 \neq \emptyset,\ B_1 \neq \emptyset$.
Nach Voraussetzung ist $A \neq \emptyset$. Also existiert $\alpha \in A$. Die Zahl α ist definiert durch einen Schnitt $C \mid C'$ in $\mathbb{Q}$. Dabei sind C, C' nichtleere Teilmengen von $\mathbb{Q}$. Daher existiert $r \in C$. Sei $r' \in \mathbb{Q}$ mit $r' < r$. Dann folgt $r' \in C$. Hieraus folgt

$$\{r' \in \mathbb{Q} \mid r' < r\} \subset C$$

Daher gilt $r < \alpha$ und somit folgt $r < \alpha$. Da α durch den Schnitt $A \mid B$ definiert wird, ist daher $r \in A$. Daraus folgt $r \in A_1$, d.h., $A_1 \neq \emptyset$. Ebenso zeigt man $B_1 \neq \emptyset$.
b) Es gilt

$$A_1 \cup B_1 = (A \cap \mathbb{Q}) \cup (B \cap \mathbb{Q}) = (A \cup B) \cap \mathbb{Q} = \mathbb{R} \cap \mathbb{Q} = \mathbb{Q}$$

b) Es sei $a \in A_1$ und $b \in B_1$. Dann ist $a \in A$ und $b \in B$. Auf Grund der Definition eines Dedekindschen Schnittes folgt daraus $a < b$.
Damit ist $A_1 \mid B_1$ ein Schnitt in $\mathbb{Q}$ und definiert eine Zahl $\gamma \in \mathbb{R}$. Da $A \cup B = \mathbb{R}$, folgt: $\gamma \in A\ \vee\ \gamma \in B$
Annahme: Es sei $\gamma \in A$.
Zu zeigen: γ ist größtes Element von A.
Sei $\alpha \in A$ und $\alpha < \gamma$. Aus Lemma 2.16 folgt

$$\exists r \in \mathbb{Q} \colon \alpha > r > \gamma.$$

Da $\alpha \in A$ und $r < \alpha$, folgt $r \in A$. Dann ist aber $r \in A_1$. Daraus folgt

$$\{r' \in \mathbb{Q} \mid r' < r\} \subseteq A_1,$$

und daher $r \leq \gamma$. Dies ist ein Widerspruch zu $r > \gamma$. Damit ist $\gamma = \sup A$. Genauso zeigt man $\gamma = \inf B$.

□

Wir können nun den folgenden wichtigen Satz beweisen.

Satz 2.18. *Jede nichtleere nach oben beschränkte Menge reeller Zahlen besitzt ein Supremum.*

Beweis: Sei $M \subseteq \mathbb{R}$ und $M \neq \emptyset$. Sei $C \in \mathbb{R}$ eine obere Schranke von M: $\forall a \in M : a \leq C$.
Fall 1: M enthält ein größtes Element a_0.
Dann ist $a_0 = \sup M$.

Fall 2: M enthält kein größtes Element.
Wir definieren in diesem Falle einen Schnitt in $\mathbb{R}$ wie folgt. Es sei

$$B = \{b \in \mathbb{R} \mid \forall a \in M : a \leq b\} \quad \text{und} \quad A = \mathbb{R} \setminus B.$$

Behauptung: $A \mid B$ ist ein Schnitt.
Beweis: i) Nach Definition von B ist $C \in B$. Daher ist $B \neq \emptyset$. Weiter ist $M \subseteq A$ $M \neq \emptyset$. Daraus folgt $A \neq \emptyset$.
ii) Auf Grund der Definition von A ist $A \cup B = \mathbb{R}$.
iii) Seien $a \in A,\ b \in B$. Dann ist $a \notin B$. Wegen der Definition von B folgt, daß ein $m \in M$ existiert mit $a < m$. Nach der Definition von B folgt: $a < m \leq b$.
Damit haben wir gezeigt, daß A, B einen Schnitt $A \mid B$ in $\mathbb{R}$ definieren. Aus dem Theorem folgt, daß ein $\gamma \in \mathbb{R}$ existiert, das den Schnitt $A \mid B$ erzeugt, d.h., $\gamma = \sup A = \inf B$. Da M in A enthalten ist, folgt

$$\forall m \in M : m \leq \gamma.$$

$$\Rightarrow \gamma \in B \ \wedge \ (\forall b \in B : \gamma \leq b) \Rightarrow \gamma = \sup M.$$

□

Addition reeller Zahlen

Seien $\alpha, \beta \in \mathbb{R}$, definiert durch Schnitte $A \mid A'$ bzw. $B \mid B'$ in $\mathbb{Q}$.
Sei

$$C = \{c \in \mathbb{Q} \mid \exists a \in A, \exists b \in B : c \leq a + b\}$$
$$C' = \mathbb{Q} \setminus C = \{c \in \mathbb{Q} \mid \forall a \in A, \forall b \in B : c > a + b\}$$

Behauptung: $C \mid C'$ ist ein Schnitt.
Beweis:

(1) $C \neq \emptyset, C' \neq \emptyset$ klar.
(2) $C \cup C' = \mathbb{Q}$ gilt nach Definition.
(3) $c \in C, c' \in C' \Rightarrow \exists a \in A, b \in B :$

$$c \leq a + b < c'$$

Der Schnitt $C \mid C'$ definiert $\gamma \in \mathbb{R}$.

Definition. $\alpha \overset{\circ}{+} \beta = \gamma$

Seien $\alpha, \beta \in \mathbb{Q}$. Dann kann man zeigen, daß

$$\alpha + \beta = \alpha \overset{\circ}{+} \beta = \gamma$$

gilt, d.h., die Addition, die wir für reelle Zahlen definiert haben, stimmt für rationale Zahlen mit der bisherigen Addition überein. Genauso definiert man die Multiplikation. Es muß jetzt noch nachgeprüft werden, daß die Körperaxiome (K1)-(K5) erfüllt sind.

2.6. Dezimalbruchdastellung reeller Zahlen

Sei $r \in \mathbb{R}$. Dann wird r definiert durch einen Schnitt $A \mid B$ in $\mathbb{Q}$. Sei $q \in A$. Auf Grund des archimedischen Axiomes existiert $n \in \mathbb{N}\colon -q < n$. Dann ist $-n < q$. Daraus folgt $-n \in A \cup \mathbb{Z}$, d.h., $A \cup \mathbb{Z} \neq \emptyset$. Sei $b \in B$. Dann gilt

$$\forall m \in A \cup \mathbb{Z}\colon m \leq b.$$

Daraus folgt, daß $A \cap \mathbb{Z}$ ein größtes Element a_0 enthält. Daraus folgt $\Rightarrow a_0 + 1 \in B$. Deshalb ist

$$a_0 \leq r < a_0 + 1.$$

Weiter ist

$$[a_0, a_0 + 1] = \left[a_0, a_0 + \frac{1}{10}\right] \cup \left[a_0 + \frac{1}{10}, a_0 + \frac{2}{10}\right] \cup \\ \cdots \cup \left[a_0 + \frac{a}{10}, a_0 + 1\right].$$

Daraus folgt, daß ein $a_1 \in \{0, 1, \cdots, q\}$ existiert mit

$$a_0 + \frac{a_1}{10} \leq r < a_0 + \frac{a_1 + 1}{10}.$$

Wenn wir dies fortsetzen, so erhalten wir $a_1, a_2, a_3, ..., a_i \in \{0, 1, 2, ...9\}$ so, daß gilt

$$\forall n \in \mathbb{N}\colon a_0 + \sum_{j=1}^{n} \frac{a_j}{10^j} \leq r < a_0 \sum_{j=1}^{n} \frac{a_j}{10^j} + \frac{1}{10^n}.$$

Damit ist

$$r = a_0, a_1 a_2 \cdots a_n \cdots = a_0 + \sum_{n=1}^{\infty} \frac{a_j}{10^j}$$

die Dezimalbruchentwicklung von r.

Zusammenhang mit Intervallschachtelung

Es sei $r = a_0, a_1 a_2 \cdots a_k \cdots$ die Dezimalbruchdastellung von $r \in \mathbb{R}$. Sei

$$c_n = \sum_{k=0}^{n} \frac{a_k}{10^k}.$$

Dann ist $c_n \in \mathbb{Q}$. Für $n \in \mathbb{N}$ sei

$$I_n = [c_n - \frac{1}{10^n}, c_n + \frac{1}{10^n}].$$

Dann ist gilt für alle $n \in \mathbb{N}$:$I_{n+1} \subset I_n$ und $|I_n| = \frac{2}{10^n}$. Also ist $\{I_n \mid n \in \mathbb{N}\}$ eine Intervallschachtelung und es gilt

$$\bigcap_{n \in \mathbb{N}} I_n = \{r\}.$$

Bemerkung: Die Dezimalbruchentwicklung nicht eindeutig! Z.B. ist

$$3,826\overline{0} = 3,825\overline{9}.$$

2.7. Überabzählbarkeit von $\mathbb{R}$

Es sei eine A Menge. Dann heißt A **abzählbar**, wenn eine bijektive Abbildung $f\colon \mathbb{N} \to A$ existiert. Dann ist $A = \{a_1, a_2, \cdots\}$.

Definition 2.19. *Eine Menge, die weder endlich noch abzählbar ist, heißt* **überabzählbar** *oder* **nicht abzählbar**.

Satz 2.20. *Es sei M eine Menge mit wenigstens zwei Elementen und*

$$F = \{f \mid f\colon \mathbb{N} \to M\}.$$

Dann ist F nicht abzählbar.

Beweis: Wir nehmen an, daß F abzählbar ist. Sei

$$F = \{f^1, f^2, \dots\}$$

mit

$$\begin{aligned} f^1 &= (f_{11}, f_{12}, f_{13}, \dots \\ f^2 &= (f_{21}, f_{22}, f_{23}, \dots) \\ &\vdots \end{aligned}$$

Es sei $g = (g_1, g_2, \dots)$ eine Folge mit $\forall i \in \mathbb{N}\colon g_i \in M \wedge g_i \neq f_{ii}$.
Dies ist möglich, da M wenigstens 2 Elemente entält. Dann ist aber $g \neq f^k$ für alle $k \in \mathbb{N}$, denn aus $g = f^k$ folgt $g_k = f_{kk}$.

□

Korollar 2.21. $[0,1]$ *und* $\mathbb{R}$ *sind nicht abzählbar.*

2. Beweis: Sei$\mathbb{R} = \{x_1, x_2, \dots\}$. Wir konstuieren eine Intervallschachtelung rekursiv wie folgt: 1) Es sei $I_1 = [x_1 + 1, x_1 + 2]$.

2) Sei I_n konstruiert. Wir teilen $I_n = [a_n, b_n]$ in 3 gleichlange Intervalle. Sei I_{n+1} ein Teilintervall, das x_{n+1} nicht enthält. Sei

$$r \in \bigcap_{n\in\mathbb{N}} I_n.$$

Dann existiert $k \in \mathbb{N}$ mit $r = x_k$. Dann ist $x_k \in I_k$. Dies ist ein Widerspruch zur Definition von I_k.

□

Kontinuumshypothese (Cantor): *Jede unendliche Teilmenge von* $\mathbb{R}$ *hat entweder die Mächtigkeit von* $\mathbb{N}$ *oder von* $\mathbb{R}$.

Auf der Grundlage von Zermelo-Fraenkel ist die Kontinuumshypothese weder widerlegbar (Gödel 1938) noch beweisbar (Cohen 1963).

2.8. Die Potenzfunktion mit rationalen Exponenten

Sei $a \in \mathbb{R}_+$ und $r \in \mathbb{R}$. Das Ziel ist die Definition von a^r. In diesem Abschnitt definieren wir a^r nur für den Fall $r \in \mathbb{Q}$. Der allgemeine Fall wird im Abschnitt über Folgen behandelt.

Die ganzzahligen Potenzen von a definiert man rekursiv wie folgt:

(1) Es sei $n \in \mathbb{N}_0$. Dann wird a^n induktiv definiert durch:

$$a^0 = 1 \quad \text{und} \quad a^n = a \cdot a^{n-1},\ n \in \mathbb{N}.$$

(2) Sei $n \in \mathbb{Z}$, $n < 0$. Dann sei

$$a^n = (a^{-n})^{-1}.$$

Es gelten die Potenzgesetze:

(1) $\forall n, m \in \mathbb{Z}\colon a^n \cdot a^m = a^{n+m}$
(2) $\forall a, b \in \mathbb{R}_+\colon a^n \cdot b^n = (a \cdot b)^n$
(3) $\forall n, m \in \mathbb{Z}\colon (a^n)^m = a^{n\cdot m}$.

Lemma 2.22. *Sei* $a \in \mathbb{R}_+,\ p \in \mathbb{Z}$ *und* $q, n \in \mathbb{N}$. *Dann gilt*

$$\sqrt[q]{a^p} = \sqrt[qn]{a^{pn}}$$

Beweis: Sei $r = \sqrt[q]{a^p}$ und $s = \sqrt[qn]{a^{pn}}$. Dann folgt $r^p = a^p$ und $s^{qn} = a^{pn}$. Damit ist

$$r^{qn} = a^{pn} = s^{qn}$$

Aus der Eindeutigkeit der qn-ten Wurzel erhalten wir $r = s$.

$\square$

proclaim Definition 2.23. Sei $a \in \mathbb{R}_+$ und $r \in \mathbb{Q}$. Es sei $r = p/q$, $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Dann sei

$$a^r : = \sqrt[q]{a^p}.$$

Aus Lemma 2.22 folgt, daß a^r unabhängig von der Darstellung von r als Bruch p/q ist.

Satz 2.24. *Für die Potenzfunktion a^r gilt:*

(1) *Es gelten die Potenzgesetze:*
$\forall a, b \in \mathbb{R}^+,\ \forall r, s \in \mathbb{Q}$:
   i) $a^r \cdot a^s = a^{r+s}$

   ii) $a^r \cdot b^r = (a \cdot b)^r$

   iii) $(a^r)^s = a^{r \cdot s}$.

(2) *Sei $r \in \mathbb{Q}$.*

$$\begin{aligned} &i) \quad r > 0 \Rightarrow (\forall x, y \in \mathbb{R}_+ \colon x < y \Rightarrow x^r < y^r) \\ &ii) \quad r < 0 \Rightarrow (\forall x, y \in \mathbb{R}_+ \colon x < y \Rightarrow y^r < x^r) \end{aligned}$$

(d.h. für $r > 0$ ist x^r streng monoton wachsend und für $r < 0$ ist x^r streng monoton fallend)

(3) *Für alle $r, s \in \mathbb{Q}$ mit $r < s$ gilt:*

$$\begin{aligned} a^r < a^s, &\text{ falls } a > 1 \\ a^s < a^r, &\text{ falls } 0 < a < 1 \end{aligned}$$

(4) *Für jedes kompakte Intervall $J \subset \mathbb{R}$ existiert $C \in \mathbb{R}_+$ mit*

$$\forall r, s \in J \cap \mathbb{Q} \colon |a^r - a^s| \leq C|r - s|.$$

Beweis: 1) Sei $r = \frac{p}{q}$ und $s = \frac{m}{q}$ mit $p, m \in \mathbb{Z}$ und $q \in \mathbb{N}$. Dann gilt

$$a^r \cdot a^s = \sqrt[q]{a^p} \cdot \sqrt[q]{a^m} = \sqrt[q]{a^{p+m}} = a^{\frac{p+m}{q}} = a^{r+s}.$$

Dies ist der Beweis von i). Dabei haben wir benutzt, daß aus der Eindeutigkeit der Wurzel die Gleichung

$$\sqrt[q]{a} \cdot \sqrt[q]{b} = \sqrt[q]{a \cdot b}$$

für alle $a, b \in \mathbb{R}^+$ folgt. Die Gleichungen ii), iii) werden analog gezeigt. Übung!

2) Seien $x, y \in \mathbb{R}_+$ mit $x < y$, und es sei $q \in \mathbb{N}$.
Behauptung: $\sqrt[q]{x} < \sqrt[q]{y}$
Beweis: Wir nehmen an, daß das Gegenteil gilt: $\sqrt[q]{y} \leq \sqrt[q]{x}$. Daraus folgt $y = \left(\sqrt[q]{y}\right)^q \leq (\sqrt[q]{x})^q = x$. Dies ist ein Widerspruch zur Annahme $x < y$.
i) Sei $r \in \mathbb{Q}$, $r > 0$. Sei $r = \frac{p}{q}$, $p, q \in \mathbb{N}$ eine Darstellung von r als Bruch. Sei $x, y \in \mathbb{R}_+$ mit $x < y$. Da $p \in \mathbb{N}$ent **Behauptung:** $\sqrt[q]{x} < \sqrt[q]{y}$
Beweis: Wir nehmen an, daß das Gegenteil gilt: $\sqrt[q]{y} \leq \sqrt[q]{x}$. Daraus folgt $y = \left(\sqrt[q]{y}\right)^q \leq (\sqrt[q]{x})^q = x$. Dies ist ein Widerspruch zur Annahme $x < y$. i) Sei $r \in \mathbb{Q}$, $r > 0$. Sei $r = \frac{p}{q}$, $p, q \in \mathbb{N}$ eine Darstellung von r als Bruch. Sei $x, y \in \mathbb{R}_+$ mit $x < y$. Da $p \in \mathbb{N}$, folgt $x^p < y^p$. Daraus ergibt sich $x^r = \sqrt[q]{x^p} < \sqrt[q]{y^p} = y^r$.
ii) Sei $r \in \mathbb{Q}$, $r < 0$. Dann ist $-r > 0$ und aus i) folgt $x^{-r} < y^{-r}$ und somit $y^r = (y^{-r})^{-1} < (x^{-r})^{-1} = x^r$.
3) Seien $r, s \in \mathbb{Q}$, $t = r - s > 0$, $a \in \mathbb{R}$, $a > 0$.
Aus 2) folgt:

$$\begin{aligned} a^t &< 1^t = 1, 0 < a < 1, \\ a^t &> 1^t = 1, a > 1. \end{aligned}$$

Multiplikation mit a^s ergibt

$$\begin{aligned} a^r = a^{t+s} = a^t \cdot a^s &< a^s, 0 < a < 1 \\ a^s < a^t a^s = a^{t+s} &= a^r, a > 1. \end{aligned}$$

4) Sei $J = [-m, m]$, $m \in \mathbb{N}$, $a \geq 1$. Sei $r, s \in J \cap \mathbb{Q}$ und $r < s$. Dann ist $t = s - r > 0$.

Fall 1: $t \geq 1$. Dann ist

$$a^t \leq ta^t \leq ta^{t+1} < 1 + ta^{t+1} \Rightarrow a^t - 1 \leq ta^{t+1}$$

Fall 2: $t = \frac{p}{n} < 1$. Dann ist

$$a^{p/n} = \sqrt[n]{a^p} \leq \frac{p \cdot a + (n-p)}{n} = 1 + \frac{p}{n}(a-1) \leq 1 + \frac{p}{n} a^{p/n+1}.$$

Dabei haben wir folgende Ungleichung benutzt:

$$\forall x_1, \ldots, x_n \in \mathbb{R}_+ \colon \sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

Es gilt also stets:

$$a^t - 1 \leq ta^{t+1}.$$

Durch Multiplikation beider Seiten mit a^r erhalten wir

$$0 < a^s - a^r \leq (s-r)a^{s+1} \leq (s-r)a^{m+1}.$$

Für $a < 1$ ist $a^r = (a^{-1})^{-r}$ und wir können wie oben verfahen.

$\square$

3. Komplexe Zahlen

Die komplexen Zahlen wurden bereits in der Vorlesung “Lineare Algebra“ behandelt. Deshalb beschränke ich mich hier auf eine Zusammenfassung der wichtigsten Fakten, die in der Vorlesung benutzt werden.

Historische Anmerkungen

Folgende Mathematiker haben wesentlich zur Entwicklung der komplexen Zahlen beigetragen.

Cardano	(1501 -1576)
J. Bernoulli	(1667 - 1748)
L. Euler	(1707 - 1783)
C.F. Gauß	(1777 - 1855) Gaußsche Zahlenebene

Für $a < 0$ ist die Gleichung $x^2 = a$ in $\mathbb{R}$ nicht lösbar. Dies führt zur Erweiterung von $\mathbb{R}$ zu $\mathbb{C}$.

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

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{\langle a, b\rangle \mid a, b \in \mathbb{R}\}$$

(**A**)**Addition in $\mathbb{C}$:** Die Addition wird definiert durch

$$\langle a, b\rangle + \langle c, d\rangle = \langle a + c, b + d\rangle$$

Die so definierte Addition kann man als Vektoraddition in $\mathbb{R}^2$ ansehen:

(**M**)**Multiplikation in $\mathbb{C}$:** Die Multiplikation in $\mathbb{C}$ ist wie folgt definiert:

$$\langle a, b\rangle \cdot \langle c, d\rangle = \langle ac - bd, ad + bc\rangle$$

Satz 3.1. $(\mathbb{C}, +, \cdot)$ *erfüllt die Körperaxiome (K1) - (K5), wobei* $0 = \langle 0, 0\rangle$ *und* $1 = \langle 1, 0\rangle$ *ist.*

$$z = \langle a, b\rangle \neq \langle 0, 0\rangle \Rightarrow z^{-1} = \left\langle \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right\rangle$$

Sei $I = \mathbb{R} \to \mathbb{C}$ definiert durch $I(r) = \langle r, 0\rangle$. Dann wird $\mathbb{R}$ ein Teilkörper von $\mathbb{C}$. Es gilt $\mathbb{R} \subset \mathbb{C}$ als Körper.

3.2 Imaginäre Einheit

Sei $i = \langle 0, 1\rangle$. Dann gilt

$$i^2 = \langle 0, 1\rangle \cdot \langle 0, 1\rangle = \langle -1, 0\rangle = -1,$$

d.h. i ist Lösung der Gleichung $z^2 = -1$.
i heißt **imaginäre Einheit**, $i = \sqrt{-1}$. Damit gilt für $z = \langle x, y\rangle \in \mathbb{C}$:

$$z = x\langle 1, 0\rangle + y\langle 0, 1\rangle = x + iy.$$
$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}.$$

Definition: Sei $z = x + iy,\ x, y \in \mathbb{R}$. *Dann heißt x* **Realteil** *von z und y* **Imaginärteil** *von z.*

Bezeichnung: $x = \mathrm{Re}(z),\ y = \mathrm{Im}(z)$

3.3. Konjugation

Sei $z \in \mathbb{C},\ z = x+iy$. Die zu z komplex konjugierte Zahl $\overline{z}$ ist definiert als

$$\overline{z} = x - iy.$$

Rechenregeln: $\forall z, w \in \mathbb{C}$ gilt
(1) $\overline{z + w} = \overline{z} + \overline{w},\ \overline{z \cdot w} = \overline{z} \cdot \overline{w}$
(2) $z + \overline{z} = 2\mathrm{Re}(z),\ z - \overline{z} = 2i\mathrm{Im}(z)$
(3) $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$
(4) $z = x + iy,\ z\overline{z} = x^2 + y^2$.

3.4. Absolutberag

Der Betrag $|z|$ von $z = x + iy \in \mathbb{C}$ ist definiert durch

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

Für $|z|$ gilt:
(1) $|z| \geq 0\ \wedge\ (|z| = 0 \Leftrightarrow z = 0)$
(2) $|\overline{z}| = |z|$
(3) $|zw| = |z| \cdot |w|$
(4) $|z + w| \leq |z| + |w|$ Dreiecksungleichung

Beweis: Es sei $z = x + iy$
(1) $|z| = \sqrt{x^2 + y^2} \geq 0,\ = 0 \Leftrightarrow x = y = 0$
(2) $|z| = \sqrt{x^2 + y^2} = |x - iy| = |\overline{z}|$
(3) Sei $z = x + iy,\ w = u + iv$. Dann ist

$$|z \cdot w|^2 = (xu - yv)^2 + (xv + yu)^2 =$$
$$x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2$$
$$= (x^2 + y^2)(u^2 + v^2) = |z|^2|w|^2.$$

(4) $|z+w|^2 = (z+w)(\overline{z}+\overline{w}) = |z|^2 + |w|^2 + z\overline{w} + \overline{z}w =$
$|z|^2 + |w|^2 + 2\mathrm{Re}(z\overline{w}) \leq |z|^2 + |w|^2 + 2|z \cdot \overline{w}| = |z|^2 + |w|^2 + 2|z||w| =$
$(|z|^2 + |w|^2)$.

3.5. Fundamentalsatz der Algebra:

Satz. *Jedes nicht-konstante komplexe Polynom hat mindestens eine komplexe Nullstelle.*

Daraus folgt, daß in $\mathbb{C}$ jedes Polynom faktorisierbar ist. Es sei

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

wobei $a_0, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$. Dann existieren $c_1, \cdots, c_r \in \mathbb{C}$, $c_i \neq c_j$, $i \neq j$, und $m_1, \cdots, m_r \in \mathbb{N}$, $n = m_1 + \cdots + m_r$ so, daß

$$p(z) = a_n (z - c_1)^{m_1} (z - c_2)^{m_2} \cdots (z - c_r)^{m_r}.$$

4. Folgen und Grenzwerte

In diesem Abschnitt beschäftigen wir uns mit zwei fundamentalen Begriffen der Analysis: Konvergenz und Grenzwert. Darauf baut die gesamte Analysis auf.

4.1. Folgen reeller Zahlen

Wir betrachten zunächst nur Folgen reeller Zahlen.

Definition 4.1. *Eine Folge reeller Zahlen ist eine Abbildung*

$$f\colon \mathbb{N} \to \mathbb{R}$$

.

Die Abbildung $f\colon \mathbb{N} \to \mathbb{R}$ wird eindeutig durch ihre Werte $f(n),\ n \in \mathbb{N}$, bestimmt. Es sei

$$a_n = f(n), \quad n \in \mathbb{N}.$$

Dann können wir die Folge, die durch die Abbildung $f\colon \mathbb{N} \to \mathbb{R}$ definiert wird, durch Aufzählung ihrer Glieder a_n angeben. Man schreibt Folgen meistens in der Form

$$a_1, a_2, a_3, \dots \quad \text{oder} \quad (a_n).$$

Wenn es erforderlich ist, den Indexbereich anzugeben, so kann man dies wie folgt kennzeichnen

$$(a_n)_{n\in\mathbb{N}} \quad \text{oder} \quad (a_n \mid n \in \mathbb{N}).$$

Bemerkung: Die Folgenglieder a_n brauchen nicht paarweise verschieden zu sein, d.h., man muß die Folge $(a_n)_{n\in\mathbb{N}}$ und die aus den Gliedern der Folge gebildete Menge $\{a_n \mid n \in \mathbb{N}\}$ sorgfältig unterscheiden. Z.B. kann die Menge $\{a_n \mid n \in \mathbb{N}\}$ nur aus einem einzigen Element bestehen.

Beispiele:

1) Sei $a \in \mathbb{R}$ und sei (a_n) die Folge, die definiert ist durch

$$\forall n \in \mathbb{N}\colon a_n = a.$$

Die Folge (a_n) nennt man *konstante Folge*. Für diese Folge gilt: $\{a_n \mid n \in \mathbb{N}\} = \{a\}$.

2) Für alle $n \in \mathbb{N}$ sei $a_n = \frac{1}{n}$. Wir erhalten die Folge

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

3) Es seien $a, b \in \mathbb{R}$.

$$\forall n \in \mathbb{N} \colon a_n = a + nb.$$

Die Folge (a_n) nennt man eine *arithmetische Folge*

4) Es seien $a, q \in \mathbb{R}$.

$$\forall n \in \mathbb{N} \colon a_n = aq^n.$$

Die Folge (a_n) nennt man eine *geometrische Folge*

5) Man kann Folgen auch rekursiv definieren. Z.B. ist die Folge der *Fibonacci*- Zahlen definiert durch:

$$a_0 = 0,\ a_1 = 1, \quad \text{und} \quad a_{n+1} = a_n + a_{n-1} \quad \text{für alle} \quad n \in \mathbb{N}.$$

4.2. Konvergente Folgen reeller Zahlen

Definition 4.2. *Eine Folge (a_n) heißt konvergent, wenn es eine Zahl $a \in \mathbb{R}$ gibt, für die folgendes gilt:*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N \colon |a_n - a| < \epsilon. \tag{4.1}$$

Die Zahl a nennt man **Grenzwert** oder **Limes** der Folge (a_n). Man sagt dann: Die Folge (a_n) hat den Grenzwert oder Limes a, bzw. die Folge (a_n) konvergiert gegen (den Grenzwert) a.

Übliche Bezeichnungen für diesen Sachverhalt sind:

$$\lim_{n\to\infty} a_n = a,\ a_n \xrightarrow[n\to\infty]{} a, \text{ oder } a_n \to a \text{ für } n \to \infty.$$

Definition 4.3. *Eine Folge (a_n), die gegen 0 konvergiert, d.h., für die gilt $\lim_{n\to\infty} a_n = 0$, nennt man Nullfolge.*

Es gilt also:

$$(a_n) \text{ ist Nullfolge } \Leftrightarrow \forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N \colon |a_n| < \epsilon.$$

Die Definition 4.2 für die Konvergenz einer Folge kann man auch geometrisch interpretieren. Dazu führen wir einige weitere Begriffe ein. Für $a \in \mathbb{R}$ und $\epsilon > 0$ sei

$$U_\epsilon(a) = (a-\epsilon, a+\epsilon).$$

Das Intervall $U_\epsilon(a)$ nennt man die ϵ-Umgebung von a. Weiter sei für alle $n \in \mathbb{N}$ eine Aussage $A(n)$ gegeben. Dann sagt man, die Aussage $A(n)$ gilt *für fast alle* $n \in \mathbb{N}$, wenn es ein $N \in \mathbb{N}$ gibt, so daß $A(n)$ für alle $n \in \mathbb{N}$ mit $n \geq N$ richtig ist. Die Definition 4.2 ist dann äquivalent zu folgender Aussage:

$$\lim_{n\to\infty} a_n = a \Leftrightarrow \forall \epsilon > 0 \colon a_n \in U_\epsilon(a) \text{ für fast alle}\, n \in \mathbb{N}.$$

Beispiele:

1) Es sei $a \in \mathbb{R}$ und (a_n) sei definiert durch $a_n = a$ für alle $n \in \mathbb{N}$. Dann gilt $\lim_{n\to\infty} a_n = a$, denn für alle $n \in \mathbb{N}$ gilt: $|a_n - a| = 0$.
2) Es sei $a_n = \frac{1}{n}, \quad n \in \mathbb{N}$.
   Behauptung: $\lim a_n = 0$, d.h., (a_n) ist eine Nullfolge.
   Beweis: Es sei $\epsilon > 0$ gegeben. Aus dem Archimedischen Axiom folgt: $\exists N \in \mathbb{N}\colon\; N > \frac{1}{\epsilon}$.
   $$\Rightarrow |a_n| = \frac{1}{n} < \epsilon \quad \text{für} \quad n \geq N.$$
3) Es sei $a > 0$ und
   $$a_n = \sqrt[n]{a}, \quad n \in \mathbb{N}.$$
   Dann gilt:
   $$\lim_{n\to\infty} \sqrt[n]{a} = 1.$$
   Beweis: Wir unterscheiden zwei Fälle.
   i) $a \geq 1$. Es sei $x_n = \sqrt[n]{a} - 1$. Dann ist $x_n \geq 0$ und die Anwendung der Bernoullischen Ungleichung ergibt
   $$a = (1 + x_n)^n \geq 1 + nx_n.$$
   Es sei $\epsilon > 0$ gegeben. Wir setzen $N = \left[\frac{a}{\epsilon}\right] + 1$. Dann gilt für alle $n \geq N$:
   $$|\sqrt[n]{a} - 1| = x_n < \frac{a}{n} < \epsilon.$$
   Damit gilt (4.1) für die Folge (a_n) und $a = 1$.

ii) $0 < a < 1$. Später werden wir zeigen, daß für Grenzwerte folgende Rechenregel gilt:

$$\lim_{n\to\infty} \sqrt[n]{a} = \left(\lim_{n\to\infty} \sqrt[n]{a^{-1}}\right)^{-1}.$$

Nach Voraussetzung ist $a^{-1} > 1$ und aus Fall i) folgt damit $\lim_{n\to\infty} \sqrt[n]{a} = 1$.

4) Es sei $a \in \mathbb{R}$ und es gelte $|a| < 1$. Dann gilt

$$\lim_{n\to\infty} a^n = 0.$$

Beweis: Es sei $|a|^{-1} = 1 + x$. Aus der Bernoullischen Ungleichung folgt

$$\frac{1}{|a|^n} = (1+x)^n \geq 1 + nx > nx. \tag{4.2}$$

Es sei $\epsilon > 0$ gegeben. Wir setzen $N = [(\epsilon x)^{-1}] + 1$. Dann gilt für alle $n \geq N$ auf Grund von (4.2):

$$|a^n| < \frac{1}{nx} < \epsilon.$$

Wir haben damit gezeigt, daß die Aussage (4.1) für die Folge (a_n) und $a = 0$ richtig ist.

Satz 4.4. *Es sei (a_n) eine konvergente Folge. Dann ist der Grenzwert dieser Folge eindeutig bestimmt.*

Beweis: Wir nehmen an, daß die Folge (a_n) zwei Grenzwerte a und a' mit $a \neq a'$ besitzt. Es sei

$$\epsilon = \frac{1}{2}|a - a'|.$$

Auf Grund unserer Annahme gilt: $\epsilon > 0$. Da a und a' Grenzwerte der Folge (a_n) sind, existiert zu diesem ϵ ein $N \in \mathbb{N}$ so, daß gilt:

$$\forall n \geq N\colon |a_n - a| < \epsilon \ \wedge \ |a_n - a'| < \epsilon.$$

Daraus folgt:

$$\begin{aligned} |a - a'| &= |a - a_n + a_n - a'| \\ &\leq |a - a_n| + |a' - a_n| < 2\epsilon = |a - a'|. \end{aligned}$$

Dies ist ein Widerspruch. Daher muß unsere Annahme falsch sein, d.h., der Grenzwert der Folge (a_n) ist eindeutig bestimmt.

$\square$

Definition 4.5. *Eine nicht konvergente Folge* (a_n) *heißt* **divergent***.*

Die Verneinung der Definition für die Konvergenz einer Folge ergibt:

$$\forall a \in \mathbb{R}\ \exists \epsilon > 0\ \forall N \in \mathbb{N}\ \exists n > N \colon |a_n - a| \geq \epsilon.$$

Anders ausgedrückt bedeutet dies: Für jedes $a \in \mathbb{R}$ existiert eine ϵ-Umgebung $U_\epsilon(a) = (a - \epsilon, a + \epsilon)$ so, daß außerhalb von $U_\epsilon(a)$, d.h., in der Menge $\mathbb{R} - (a - \epsilon, a + \epsilon)$, unendlich viele Folgenglieder a_n liegen. Dies bedeutet, daß

$$\{n \in \mathbb{N} \mid a_n \notin (a - \epsilon, a + \epsilon)\}.$$

unendlich ist.

Beispiele:

1) Es sei $a_n = (-1)^n,\ n \in \mathbb{N}$.
Behauptung: Die Folge a_n ist divergent.
Beweis: Wir nehmen an, die Folge (a_n) sei konvergent und der Grenzwert sei a. Zu $\epsilon = 1/2$ existiert dann ein $N \in \mathbb{N}$ so, daß gilt

$$|a_n - a| < \frac{1}{2} \quad \text{für alle} \quad n \geq \mathbb{N}.$$

Daraus folgt

$$\begin{aligned} 2 = |a_{n+1} - a_n| &= |a_{n+1} - a + a - a_n| \\ &\leq |a_{n+1} - a| + |a_n - a| < 1 \end{aligned}$$

für alle $n \geq N$. Dies ist ein offensichtlicher Widerspruch. Deshalb ist unsere Annahme falsch, d.h. die Folge (a_n) ist divergent.

2) Es sei $a_n = n,\ n \in N$. Diese Folge ist divergent, da für jedes $a \in \mathbb{R}$, fast alle a_n außerhalb jeder ϵ-Umgebung von a liegen.

Wie die Beispiele zeigen, gibt es zwei Formen der Divergenz. Eine Folge (a_n) heißt **bestimmt divergent** gegen ∞ bzw. gegen $-\infty$, wenn gilt:

$$\forall K \in \mathbb{R}\ \exists N \in \mathbb{N}\ \forall n \geq N \colon\ a_n > K$$

bzw.

$$\forall K \in \mathbb{R}\, \exists N \in \mathbb{N}\, \forall n \geq N\colon\ a_n < -K.$$

Für diesen Sachverhalt verwenden wir folgende Bezeichnung:

$$\lim_{n\to\infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n\to\infty} a_n = -\infty.$$

Man bezeichnet ∞ und $-\infty$ auch als uneigentliche Grenzwerte. Eine divergente Folge, die nicht bestimmt divergent ist, nennt man unbestimmt divergent. Die Folge aus Beispiel 1 ist unbestimmt divergent, und die aus Beispiel 2 ist bestimmt divergent mit dem Grenzwert ∞.

Definition 4.6. *Eine Folge (a_n) heißt beschränkt, wenn die Menge $\{a_n \mid n \in \mathbb{N}\}$ beschränkt ist, d.h., es existiert ein $C \in \mathbb{R}_+$ so, daß gilt: $|a_n| \leq C$ für alle $n \in \mathbb{N}$.*

Da für eine konvergente Folge (a_n) fast alle a_n in der ϵ-Umgebung $U_\epsilon(a)$ des Grenzwertes a liegen, ist klar, daß eine konvergente Folge beschränkt ist, d.h., es gilt der folgende Satz:

Satz 4.7. *Es sei (a_n) eine konvergente Folge. Dann ist die Folge (a_n) beschränkt.*

Beweis: Es sei $a = \lim_{n\to\infty} a_n$. Dann existiert ein $N \in \mathbb{N}$:
$|a_n - a| \leq 1$ für alle $n \geq N$.

$$\Rightarrow |a_n| = |a_n - a + a| \leq |a_n - a| + |a| \leq 1 + |a|$$

für alle $n \geq N$. Es sei

$$C = \max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, 1 + |a|\}.$$

Dann gilt:

$$\forall n \in \mathbb{N}\colon |a_n| \leq C,$$

d.h. (a_n) ist beschränkt.

$\square$

Die Umkehrung gilt nicht.

Beispiel: Für $n \in \mathbb{N}$ sei $a_n = (-1)^n$. Dann ist (a_n) beschränkt, aber nicht konvergent wie wir oben gezeigt haben.

4.3 Rechenregeln für Grenzwerte

Für das Rechnen mit Grenzwerten gelten eine Reihe von Rechenregeln, die die Berechnung von Grenzwerten vereinfachen.

Satz 4.8. *Seien* (a_n) *und* (b_n) *konvergente Folgen und es sei* $a = \lim_{n\to\infty} a_n$, $b = \lim_{n\to\infty} b_n$. *Dann gilt:*

1) $\forall \lambda \in \mathbb{R}\colon \lim_{n\to\infty} \lambda a_n = \lambda a$.
2) $\lim_{n\to\infty}(a_n + b_n) = a + b$
3) $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$
4) *Es sei* $b \neq 0$. *Dann existiert ein* $N \in \mathbb{N}$ *so, daß gilt:* $b_n \neq 0$ *für alle* $n \geq N$ *und*

$$\lim_{\substack{n\to\infty \\ n\geq N}} \frac{a_n}{b_n} = \frac{a}{b}.$$

Beweis:

1) Für $\lambda = 0$ ist die behauptete Gleichung offensichtlich richtig. Deshalb können wir annehmen, daß $\lambda \neq 0$. Es sei $\epsilon > 0$. Da die Folge (a_n) konvergent ist, existiert zu diesem ϵ ein $N \in \mathbb{N}$ so, daß gilt:

$$|a_n - a| < \frac{\epsilon}{|\lambda|} \quad \text{für alle} \quad n \geq N.$$

Für $n \geq N$ gilt daher

$$|\lambda a_n - \lambda a| = |\lambda||a_n - a| < \epsilon.$$

Dies bedeutet, daß die Folge (λa_n) konvergent ist und

$$\lim_{n\to\infty} \lambda a_n = \lambda a.$$

2) Es sei $\epsilon > 0$ gegeben. Da (a_n) und (b_n) konvergente Folgen mit Grenzwert a bzw. b sind, existieren $N, N' \in \mathbb{N}$:

$$\forall n \geq N\colon |a_n - a| < \frac{\epsilon}{2}$$
$$\forall n \geq N'\colon |b_n - b| < \frac{\epsilon}{2}.$$

Es sei $N_1 = \max\{N, N'\}$. Dann gilt

$$\forall n \geq N_1\colon |a_n - a| < \frac{\epsilon}{2} \;\wedge\; |b_n - b| < \frac{\epsilon}{2}.$$

$$\Rightarrow |a_n + b_n - (a + b)| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \epsilon$$

für alle $n \geq N_1$.

Wir haben damit gezeigt, daß für jedes $\epsilon > 0$ ein $N_1 \in \mathbb{N}$ existiert so, daß gilt

$$\forall n \geq N_1 \colon |a_n + b_n - (a+b)| < \epsilon,$$

d.h., (4.1) gilt für die Folge $(a_n + b_n)$ und den Grenzwert $a+b$.

3) Wir unterscheiden zwei Fälle.
Fall 1: $b = 0$. Da (a_n) konvergent ist, folgt aus Satz 4.7, daß ein $C \in \mathbb{R}_+$ existiert so, daß gilt:

$$\forall n \in \mathbb{N} \colon |a_n| \leq C. \tag{4.3}$$

Es sei $\epsilon > 0$ gegeben. Da nach Voraussetzung $\lim_{n\to\infty} b_n = 0$ gilt, existiert zu ϵ ein $N \in \mathbb{N}$:

$$\forall n \geq N \colon |b_n| < \frac{\epsilon}{C}.$$

Hieraus und aus (4.1) folgt, daß für alle $n \geq N$ gilt:

$$|a_n b_n| \leq |a_n||b_n| < C \cdot \frac{\epsilon}{C} = \epsilon \;\Rightarrow\; \lim_{n\to\infty} (a_n b_n) = 0.$$

Fall 2. $b \neq 0$. Es sei $C > 0$ so, daß (4.3) gilt, und es sei $\epsilon > 0$ gegeben. Da (a_n) und (b_n) gegen a bzw. b konvergieren, existiert zu ϵ ein $N \in \mathbb{N}$:

$$\forall n \geq N \colon |a_n - a| < \frac{\epsilon}{2|b|};$$
$$\forall n \geq N \colon |b_n - b| < \frac{\epsilon}{2C}.$$

Hieraus folgt, daß für alle $n \geq N$ gilt:

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| = |a_n(b_n - b) + b(a_n - a)|$$
$$\leq |a_n||b_n - b| + |b||a_n - a| < C \cdot \frac{\epsilon}{2C} + |b|\frac{\epsilon}{2|b|} = \epsilon.$$

Damit ist gezeigt, daß die Folge $(a_n \cdot b_n)$ konvergiert und den Grenzwert $a \cdot b$ hat.

4) Es sei $b \neq 0$. Da (b_n) konvergent ist, existiert ein $n_0 \in \mathbb{N}$:

$$\forall n \geq n_0 \colon |b_n - b| < \frac{|b|}{2}.$$

Daraus folgt, daß für alle $n \geq n_0$ gilt:

$$|b_n| = |b - b + b_n| \geq |b| - |b_n - b| > \frac{|b|}{2}. \tag{4.4}$$

Insbesondere: $\forall n \geq n_0 \colon b_n \neq 0$.
Es sei $\epsilon > 0$ gegeben. Da (b_n) konvergiert, existiert zu diesem ϵ ein$N \in \mathbb{N}$:

$$\forall n \geq N \colon |b_n - b| < \frac{|b|^2}{2} \cdot \epsilon. \tag{4.5}$$

Aus (4.4) und (4.5) folgt:

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{bb_n}\right| = \frac{|b - b_n|}{|b||b_n|} < \frac{2}{|b|^2} \cdot \frac{|b|^2}{2} \cdot \epsilon = \epsilon$$

für alle $n \geq n_0$. Hieraus folgt $\lim_{n\to\infty} b_n^{-1} = b^{-1}$. Aus 3) erhalten wir

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

□

Entsprechende Rechenregeln gelten für uneigentliche Grenzwerte.

Satz 4.9. *Es seien (a_n) und (b_n) bestimmt divergente Folgen mit dem uneigentlichen Grenzwert ∞. Dann gilt:*

a) *Für alle $\lambda \in \mathbb{R} - \{0\}$ gilt:*

$$\lambda > 0 \Rightarrow \lim_{n\to\infty} \lambda a_n = \infty,$$
$$\lambda < 0 \Rightarrow \lim_{n\to\infty} \lambda a_n = -\infty.$$

b) *Es existiert $n_0 \in \mathbb{N}$ so, daß $a_n \neq 0$ für alle $n \geq n_0$ und es gilt*

$$\lim_{\substack{n\to\infty \\ n\geq n_0}} a_n^{-1} = 0.$$

c) $\lim_{n\to\infty}(a_n + b_n) = \infty$.

d) $\lim_{n\to\infty} a_n \cdot b_n = \infty$.

Die entsprechenden Aussagen gelten für den uneigentlichen Grenzwert $-\infty$. Der Beweis ist dem Leser als Übung überlassen.

Beispiel:

Es seien

$$P(x) = \sum_{j=0}^{k} a_j x^j,\ Q(x) = \sum_{i=0}^{l} b_i x^i$$

reelle Polynome vom Grad k bzw. l, d.h., $a_k \neq 0$ und $b_l \neq 0$.

Behauptung:

$$\lim_{n\to\infty}\frac{P(n)}{Q(n)}=\begin{cases}\infty, & k>l,\,\frac{a_k}{b_l}>0;\\ -\infty & k>l,\,\frac{a_k}{b_l}<0;\\ a_k/b_l, & k=l;\\ 0, & k<l.\end{cases}$$

Der Beweis erfolgt durch Anwendung obiger Rechenregeln und wurde in den Übungen besprochen.
Die beiden folgenden Prinzipien sind ebenfalls sehr nützlich für die Berechnung von Grenzwerten.

Satz 4.10. *Es seien (a_n) und (b_n) konvergente Folgen und es gelte:*

1) $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} b_n = b$.
2) $\exists N \in \mathbb{N}\ \forall n \geq N\colon\ a_n \leq b_n$.

Dann gilt: $a \leq b$.

Beweis: Es sei $\epsilon > 0$ gegeben. Auf Grund der Voraussetzung 1) existiert ein $N_1 \in \mathbb{N}$:

$$\forall n \geq N_1\colon |a_n - a| < \epsilon \ \wedge\ |b_n - b| < \epsilon. \tag{4.6}$$

Es sei $N_2 = \max\{N, N_1\}$. Dann folgt aus 2) und (4.6), daß für alle $n \geq N_2$ gilt:

$$a - \epsilon < a_n \leq b_n < b + \epsilon$$

$\Rightarrow$

$$\forall \epsilon > 0\colon b - a > -2\epsilon.$$

$\Rightarrow$

$$b - a \geq \sup\{-2\epsilon \mid \epsilon > 0\} = 0.$$

□

Satz 4.11. *(Einschachtelungsprinzip) Es seien (a_n), (b_n) und (c_n) Folgen, für die folgendes gilt:*

1) *Für fast alle* $n \in \mathbb{N}$:

$$a_n \leq b_n \leq c_n$$

2) (a_n) *und* (c_n) *sind konvergent und es sei*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = a.$$

Dann ist auch (b_n) konvergent und es gilt

$$\lim_{n\to\infty} b_n = a.$$

Beweis: Es sei $\epsilon > 0$ gegeben. Aus den Voraussetzungen folgt, daß ein $N \in \mathbb{N}$ existiert so, daß gilt:

$$\forall n \geq N\colon |a_n - a| < \epsilon \;\wedge\; |c_n - a| < \epsilon \;\wedge\; a_n \leq b_n \leq c_n.$$

$\Rightarrow$

$$\forall n \geq N\colon a - \epsilon < a_n \leq b_n \leq c_n < a + \epsilon$$

$\Rightarrow$

$$\forall n \geq N\colon |b_n - a| < \epsilon.$$

Damit haben wir gezeigt, daß für die Folge (b_n) die Aussage (4.1) gilt, d.h., (b_n) ist konvergent und

$$\lim_{n\to\infty} b_n = a.$$

$\square$

Beispiel: Für alle $n \in \mathbb{N}$ sei

$$a_n = \frac{n}{\sqrt{n^2+n}},\; b_n = \sum_{j=1}^{n} \frac{1}{\sqrt{n^2+j}} \quad \text{und} \quad c_n = \frac{n}{\sqrt{n^2+1}}.$$

Dann gilt für alle $n \in \mathbb{N}$:

$$a_n < b_n < c_n.$$

Weiter gilt:

$$a_n = \frac{n}{\sqrt{n^2+n}} = \frac{1}{\sqrt{1+\frac{1}{n}}}.$$

Als erstes bestimmen wir den Grenzwert der Folge $\left(\sqrt{1+\frac{1}{n}}\right)$. Dazu verwenden wir die 3. binomische Formel. Es gilt:

$$0 < 2\left(\sqrt{1+\frac{1}{n}} - 1\right) < \left(\sqrt{1+\frac{1}{n}} + 1\right)\left(\sqrt{1+\frac{1}{n}} - 1\right) = \frac{1}{n}.$$

Da $\frac{1}{n}$ eine Nullfolge ist, folgt daraus

$$\lim_{n\to\infty} \sqrt{1+\frac{1}{n}} = 1.$$

Aus Satz 4.8, 4), ergibt sich

$$\lim_{n\to\infty} a_n = 1.$$

Ebenso zeigt man

$$c_n = \frac{1}{\sqrt{1+\frac{1}{n^2}}} \xrightarrow[n\to\infty]{} 1.$$

Anwendung des Satzes 4.11 ergibt

$$\lim_{n\to\infty} b_n = 1.$$

□

4.4 Monotone Folgen

Für eine gegebene Folge (a_n) entstehen zwei Fragen: 1) Ist die Folge konvergent? 2) Wenn (a_n) konvergiert, was ist der Limes der Folge a_n? Mittels der Definition 4.2 für die Konvergenz lassen sich diese beiden Fragen nur im Zusammenhang beantworten. Man kann lediglich prüfen, ob eine vorgegebene Zahl $a \in \mathbb{R}$ Limes der Folge (a_n) ist oder nicht. Dies ist kein effektives Verfahren, da man den Limes "erraten" muß. Wir beschäftigen uns deshalb im folgenden mit einigen Konvergenzkriterien, die uns gestatten, eine Folge auf Konvergenz zu prüfen. Ein solches Kriterium lernen wir in diesem Abschnitt kennen. Es bezieht sich auf monotone Folgen.

Definition 4.12. *Eine Folge* (a_n) *heißt*

a) **monoton wachsend**, *wenn gilt:*

$$\forall n \in \mathbb{N}\colon\ a_n \leq a_{n+1},$$

b) **monoton fallend**, *wenn gilt:*

$$\forall n \in \mathbb{N}\colon\ a_{n+1} \leq a_n.$$

Beispiele:

1) Es sei $a_n = \frac{1}{n}$, $n \in \mathbb{N}$. Die Folge (a_n) ist monoton fallend.

2) Es sei $a_n = 1 - \frac{1}{n}$, $n \in \mathbb{N}$. Die Folge (a_n) ist monoton wachsend.

Satz 4.13. *Jede beschränkte, monotone Folge* (a_n) *ist konvergent und es gilt:*

a) (a_n) *monoton wachsend* $\Rightarrow \lim_{n\to\infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\}$
b) (a_n) *monoton fallend* $\Rightarrow \lim_{n\to\infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}$.

Beweis: Es sei $A = \{a_n \mid n \in \mathbb{N}\}$. Da (a_n) eine beschränkte Menge reeller Zahlen ist, existieren

$$a = \sup A \quad \text{und} \quad b = \inf A.$$

a) (a_n) sei monoton wachsend.
Es sei $\epsilon > 0$ gegeben. Da a die kleinste obere Schranke von A ist, existiert ein $N \in \mathbb{N}$:

$$a - \epsilon < a_N.$$

Da (a_n) monoton wachsend ist, folgt

$$\forall n \geq N\colon \; a - \epsilon < a_N \leq a_n \leq a < a + \epsilon \Rightarrow \forall n \geq N\colon \; |a - a_n| < \epsilon.$$

Da $\epsilon > N$ beliebig gewählt war, haben wir gezeigt, daß für die Folge (a_n) und $a = \sup A$ die Aussage (4.1) gilt. Damit ist (a_n) konvergent und

$$\lim_{n\to\infty} a_n = a = \sup A.$$

b) (a_n) sei monoton fallend. Es sei $\epsilon > 0$. Da b die größte untere Schranke von A ist, existiert $N \in \mathbb{N}$:

$$a_N < b + \epsilon.$$

Da (a_n) monoton fallend ist, ergibt sich daraus:

$$\forall n \geq N\colon \; b - \epsilon < b \leq a_n \leq a_N < b + \epsilon \Rightarrow \forall n \geq N\colon \; |a_n - b| < \epsilon.$$

Wie im Falle a) erhalten wir daraus die Behauptung, d.h., es gilt

$$\lim_{n\to\infty} a_n = b = \inf A.$$

$\square$

Beispiele:

1) Die Zahl $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = 2,718281828459045\ldots$

Lemma. *Für alle $n \in \mathbb{N}$ und alle $x \geq -n$, $x \neq 0$, gilt:*

$$\left(1 + \frac{x}{n}\right)^n < \left(1 + \frac{x}{n+1}\right)^{n+1}$$

Beweis: Zum Beweis benutzen wir, daß das geometrische Mittel von endlich vielen positiven reellen Zahlen stets kleiner oder gleich dem arithmetischen Mittel ist, d.h.,

$$\forall x_1, \ldots, x_m \in \mathbb{R}_+ : \quad \sqrt[m]{x_1 \cdots x_m} \leq \frac{x_1 + \cdots + x_m}{m} \tag{4.7}$$

und "=" gilt genau dann, wenn $x_1 = \cdots = x_m$.
Aus dieser Ungleichung erhalten wir

$$\sqrt[n+1]{\left(1 + \frac{x}{n}\right)^n \cdot 1} < \frac{n(1 + \frac{x}{n}) + 1}{n+1} = 1 + \frac{x}{n+1}$$

Wegen der Monotonie der Potenzfunktion ergibt sich daraus die behauptete Ungleichung.

□

Für $n \in \mathbb{N}$ sei

$$a_n = \left(1 + \frac{1}{n}\right)^n \quad \text{und} \quad b_n = \left(1 - \frac{1}{n}\right)^n .$$

Aus obigem Lemma folgt, daß (a_n) und (b_n) monoton wachsende Folgen sind. Es sei

$$c_n = (b_{n+1})^{-1}.$$

Da (b_n) monoton wachsend ist, ist (c_n) monoton fallend. Weiter ist $c_n = (1 + \frac{1}{n})^{n+1}$ und

$$a_n = \left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n}\right)^{n+1} = c_n \tag{4.8}$$

für alle $n \in \mathbb{N}$. Daraus folgt:

$$\forall n \in \mathbb{N} : \; a_n < c_n \leq c_1 \quad \text{und} \quad a_1 \leq a_n < c_n .$$

Damit sind (a_n) und (c_n) beschränkte, monotone Folgen. Aus Satz 4.13 folgt, daß die Grenzwerte

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \text{ und } e' = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}$$

existieren. Aus (4.8) und Satz 4.10 folgt

$$e \leq e'.$$

Da $e' = \inf\{c_n \mid n \in \mathbb{N}\}$ und $e = \sup\{a_n \mid n \in \mathbb{N}\}$ ist, gilt:

$$\forall n \in \mathbb{N}\colon\ a_n < e < c_n.$$

Außerdem ist

$$c_n = a_n\left(1+\frac{1}{n}\right).$$

Da $(1+1/n) \to 1$ für $n \to \infty$, folgt aus den Rechenregeln für Produkte von Grenzwerten (Satz 4.8, 3)),

$$e' = \lim_{n\to\infty} c_n = \lim_{n\to\infty} a_n = e.$$

Da $c_n = (b_{n+1})^{-1}$, folgt aus Satz 4.8, 4):

$$e^{-1} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n.$$

2) Rekursive Berechnung von Quadratwurzeln
Gegeben seien $a > 0$ und $x_0 > 0$. Für $n \in \mathbb{N}_0$ sei

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \tag{4.9}$$

Satz 4.14. *Für alle $x_0 > 0$ gilt*

$$\lim_{n\to\infty} x_n = \sqrt{a}.$$

Beweis: 1) Mittels vollständiger Induktion zeigt man:

$$\forall n \in \mathbb{N}\colon\ x_n > 0.$$

Deshalb kann (4.9) angewendet werden, um die Folge (x_n) zu definieren.
2) Es gilt:

$$\begin{aligned}
x_n^2 - a &= \frac{1}{4}\left(x_{n-1} + \frac{a}{x_{n-1}}\right)^2 - a \\
&= \frac{1}{4}\left(x_{n-1}^2 + 2a + \frac{a^2}{x_{n-1}^2}\right) - a \\
&= \frac{1}{4}\left(x_{n-1}^2 - 2a + \frac{a^2}{x_{n-1}^2}\right) \\
&= \frac{1}{4}\left(x_{n-1} - \frac{a}{x_{n-1}}\right)^2 \geq 0.
\end{aligned}$$

$\Rightarrow \quad \forall n \in \mathbb{N}\colon\; x_n \geq \sqrt{a}$.
Weiter gilt für alle $n \in \mathbb{N}$:

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

Damit haben wir gezeigt, daß (x_n) monoton fallend und beschränkt ist.
Aus Satz 4.13 folgt, daß der Limes

$$x = \lim_{n\to\infty} x_n$$

existiert. Aus den Rechenregeln für Grenzwerte (Satz 4.8) ergibt sich

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \xrightarrow[n\to\infty]{} \frac{1}{2}\left(x + \frac{a}{x}\right).$$

Auf Grund der Eindeutigkeit des Grenzwertes (Satz 4.4) folgt

$$x = \frac{1}{2}\left(x + \frac{a}{x}\right)$$

$\Rightarrow \quad x^2 = a.$

$\square$

Fehlerabschätzung:
Es sei $f_n = x_n - \sqrt{a}$. Dann gilt

$$\begin{aligned}
f_{n+1} = x_{n+1} - \sqrt{a} &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - \sqrt{a} \\
&= \frac{1}{2x_n}\left(x_n^2 - 2\sqrt{a}x_n + a\right) \\
&= \frac{1}{2x_n}\left(x_n - \sqrt{a}\right)^2 = \frac{1}{2x_n} f_n.
\end{aligned}$$

Da für alle $n \in \mathbb{N}$: $x_n \geq \sqrt{a}$, folgt daraus

$$|f_{n+1}| \leq \frac{1}{2\sqrt{a}} f_n^2$$

$\Rightarrow$

$$\begin{aligned}
|f_{n+1}| &\leq \left(\frac{1}{2\sqrt{a}}\right)^{1+2+\cdots+2^{n-1}} f_1^{2^n} \\
&= \left(\frac{1}{2\sqrt{a}}\right)^{2^n-1} f_1^{2^n}.
\end{aligned}$$

Beispiel: Es sei $a = 2$ und $x_0 = 1$. Dann ist $x_1 = 1,5$ und es gilt $1,4 < \sqrt{2} < 1,5$. Hieraus folgt

$$f_1 = x_1 - \sqrt{2} < 10^{-1}.$$

Es sei $N = 2^n + 2^{n-1} - 2$. Dann erhalten wir aus obiger Formel

$$|x_{n+1} - \sqrt{2}| < \frac{1}{2^N} 10^{-2^n}.$$

Also z.B.

$$\begin{aligned}
|x_3 - \sqrt{2}| &< 2^{-4} 10^{-4} \\
|x_4 - \sqrt{2}| &< 2^{-10} 10^{-8}.
\end{aligned}$$

Wir erhalten nach wenigen Iterationsschritten bereits eine sehr gute Approximation von $\sqrt{2}$.

3) Die Exponentialfunktion

Es sei $a > 0$. Im Kapitel 2 haben wir die Potenz a^r für rationale Exponenten r definiert. Wir werden jetzt a^x mittels Grenzübergang für alle $x \in \mathbb{R}$ definieren. Dazu benötigen wir folgendes Lemma.

Lemma. *Es sei $x \in \mathbb{R}$. Dann existiert eine monoton wachsende Folge (r_n) rationaler Zahlen mit*

$$\lim_{n\to\infty} r_n = x.$$

Beweis: Es sei $x \in \mathbb{R}$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist, gilt:

$$\forall n \in \mathbb{N}\ \exists r_n \in \mathbb{Q}\colon\ x - \frac{1}{n} < r_n < x - \frac{1}{n+1}.$$

Die so definierte Folge (r_n) ist monoton wachsend und

$$\lim_{n\to\infty} r_n = x.$$

$\square$

Es sei $a > 0$ und $x \in \mathbb{R}$. Um die Potenz a^x zu definieren, wählen wir eine monoton wachsende Folge (r_n) rationaler Zahlen so, daß

$$\lim_{n\to\infty} r_n = x.$$

Eine solche Folge existiert auf Grund des obigen Lemmas. Wir unterscheiden zwei Fälle.
Fall 1: $a > 1$. Dann gilt:

$$\forall n \in \mathbb{N}\colon\ a^{r_n} < a^{r_{n+1}} \wedge r_n < [x] + 1.$$

Deshalb ist (a^{r_n}) eine monoton wachsende, beschränkte Folge. Nach Satz 4.13 existiert

$$\lim_{n\to\infty} a^{r_n}.$$

Fall 2: $0 < a < 1$. Dann gilt

$$\forall n \in \mathbb{N}\colon\ a^{r_{n+1}} > a^{r_n} \wedge a^{r_n} > 0.$$

Die Folge (a^{r_n}) ist daher monoton fallend und beschränkt. Wegen Satz 4.13 existiert der Limes

$$\lim_{n\to\infty} a^{r_n}$$

auch in diesem Falle.
Es sei jetzt (s_n) eine beliebige Folge rationaler Zahlen so, daß

$$\lim_{n\to\infty} s_n = x.$$

Dann gilt

$$|s_n - r_n| = |s_n - x + x - r_n| \le |s_n - x| + |r_n - x| \xrightarrow[n\to\infty]{} 0. \qquad (4.10)$$

Da (r_n) und (s_n) konvergente Folgen sind, sind sie beschränkt (Satz 4.7), d.h., es existiert $m \in \mathbb{N}$:

$$\forall n \in \mathbb{N}\colon r_n, s_n \in [-m, m].$$

Aus Satz 2.24, 4), folgt daß ein $C \in \mathbb{R}_+$ existiert.

$$\forall n \in \mathbb{N}\colon\ |a^{r_n} - a^{s_n}| \le C|r_n - s_n|.$$

Aus (4.10) erhalten wir

$$\lim_{n\to\infty} a^{s_n} = \lim_{n\to\infty} a^{r_n}.$$

Wir erhalten also für jede Folge (r_n) rationaler Zahlen mit $\lim_{n\to\infty} r_n = x$ den gleichen Grenzwert. Diesen Grenzwert bezeichnen wir mit a^x

$$a^x = \lim_{n\to\infty} a^{r_n}.$$

Satz 4.15. *Für die Potzenzfunktion a^x, $x \in \mathbb{R}$, gilt:*

1) $\forall a,\ b \in \mathbb{R}_+, \forall x, y \in \mathbb{R}$:

$$a^x a^y = a^{x+y},\ (a^x)^y = a^{x+y},\ a^x b^x = (a\cdot b)^x.$$

2) $a > 1 \Rightarrow x \longmapsto a^x$ *ist streng monoton wachsend*
   $a < 1 \Rightarrow x \longmapsto a^x$ *ist streng monoton fallend*
3) $\alpha > 0 \Rightarrow x \longmapsto x^\alpha$ *streng monoton wachsend*
   $\alpha < 0 \Rightarrow x \longmapsto x^\alpha$ *streng monoton fallend*

Beweis: Folgt aus Satz 2.24 und den Rechenregeln für Grenzwerte.

4.5. Der Satz von Bolzano - Weierstraß

Der Satz von Bolzano - Weierstraß ist grundlegend für die Konvergenztheorie beschränkter Folgen. Ein wichtiger Begriff in diesem Zusammenhang ist der Begriff des Häufungspunktes.

Definition 4.16. *Es sei (a_n) eine Folge. Eine Zahl $h \in \mathbb{R}$ heißt* **Häufungswert** *(oder* **Häufungspunkt***) von (a_n), wenn zu jedem $\epsilon > 0$ unendliche viele $n \in \mathbb{N}$ existieren so, daß*

$$|h - a_n| < \epsilon.$$

Mit anderen Worten: Jede ϵ-Umgebung $U_\epsilon(h)$ von h enthält unendlich viele Folgenglieder a_n. Im Gegensatz zum Grenzwert können bei einem Häufungswert auch noch unendlich viele Folgenglieder außerhalb der ϵ-Umgebung des Häufungswertes h liegen.

Beispiele:

1) Es sei (a_n) eine konvergente Folge mit Grenzwert a. Dann ist a Häufungswert von (a_n)
2) Für alle $n \in \mathbb{N}$ sei

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

Die Folge (a_n) ist nicht konvergent. Die Häufungswerte sind 1 und -1.

Definition 4.17. *Es sei (a_n) eine Folge und (n_k) eine streng monoton wachsende Folge natürlicher Zahlen, d.h., es gelte*

$$n_1 < n_2 < n_3 \cdots$$

Dann heißt die Folge $(a_{n_k})_{k \in \mathbb{N}}$ Teilfolge von (a_n).

Satz 4.18. *Es sei (a_n) eine Folge. Eine Zahl $h \in \mathbb{R}$ ist Häufungswert von (a_n) genau dann, wenn eine konvergente Teilfolge (a_{n_k}) von (a_n) existiert so, daß*

$$\lim_{k \to \infty} a_{n_k} = h.$$

Beweis: $\Leftarrow$) Es sei (a_{n_k}) eine konvergente Teilfolge von (a_n) und es sei

$$h = \lim_{k \to \infty} a_{n_k}.$$

Dann gilt

$$\forall \epsilon > 0 \, \exists N \in \mathbb{N} \, \forall k \geq N |a_{n_k} - h| < \epsilon.$$

Da $(a_{n_k})_{k \in \mathbb{N}}$ eine Teilfolge von (a_n) ist, liegen in jeder ϵ- Umgebung von h unendlich viele Glieder der Folge (a_n), d.h., h ist Häufungswert von (a_n).

$\Rightarrow$) Es sei $h \in \mathbb{R}$ Häufungswert von (a_n). Wir konstruieren schrittweise eine Teilfolge (a_{n_k}).
1. Schritt: Da h ein Häufungswert ist, folgt

$$\exists n_1 \in \mathbb{N}: \quad |a_{n_1} - h| < \frac{1}{2}$$

2. Schritt: Es existieren unendlich viele $n \in \mathbb{N}$ so, daß

$$|a_n - h| < \frac{1}{3}.$$

$\Rightarrow$

$$\exists n_2 \in \mathbb{N}: \; n_2 > n_1 \;\wedge\; |a_{n_2} - h| < \frac{1}{3}.$$

Es sei a_{n_j} für $j \leq k$ bereits definiert, wobei $n_1 < n_2 < \cdots < n_k$ gilt.
k+1. Schritt: Da h Häufungswert ist, existieren unendlich viele $n \in \mathbb{N}$ mit

$$|a_n - h| < \frac{1}{k+1}.$$

$\Rightarrow$

$$\exists n_{k+1} \in \mathbb{N}: n_{k+1} > n_k \;\wedge\; |a_{n_{k+1}} - h| < \frac{1}{k+1}$$

Dadurch erhalten wir eine rekursiv definierte Folge $(a_{n_k})_{k\in\mathbb{N}}$. Für diese Folge gilt auf Grund der Konstruktion $\forall k \in \mathbb{N}\colon |a_{n_k} - h| < \frac{1}{k}$

$$\begin{aligned}
&\Rightarrow \forall k \in \mathbb{N}\; \forall l \geq k\colon |a_{n_l} - h| < \frac{1}{k} \\
&\Rightarrow \forall \epsilon > 0\; \exists N \in \mathbb{N}\; \forall k \geq N\colon |a_{n_k} - h| < \epsilon.
\end{aligned}$$

Damit ist gezeigt, daß (a_{n_k}) konvergiert und

$$\lim_{k\to\infty} a_{n_k} = h.$$

□

Satz 4.19. *(Satz von Bolzano - Weierstraß)*
Jede beschränkte Folge (a_n) hat einen Häufungswert.

Beweis: Es sei (a_n) eine beschränkte Folge. Dann existieren $A, B \in \mathbb{R}$ mit:

a) $A < B$

b) $\forall n \in \mathbb{N}\colon\; A \leq a_n \leq B$.

Wir konstruieren eine Intervallschachtelung $\{I_k \mid k \in \mathbb{N}\}$, $I_k = [A_k, B_k]$, so daß gilt:

1) $I_1 = [A, B]$.
2) $\forall k \in \mathbb{N}\colon\; I_{k+1} \subset I_k$ und $|I_k| = 2^{-k+1}(B - A)$.
3) $\forall k \in \mathbb{N}\colon$ Es existieren unendlich viele $n \in \mathbb{N}\colon\; a_n \in I_k$.

Wir definieren I_k rekursiv.

1. Schritt: Sei $I_1 = [A, B]$.

Rekursionsschritt: Wir nehmen an, daß die Intervalle I_l, $l \leq k$, bereits konstruiert sind so, daß 1) - 3) gelten. Es sei

$$m_k = A_k + \frac{1}{2}(B_k - A_k).$$

Dann ist m_k der Mittelpunkt von I_k. Nach Voraussetzung liegen in I_k unendlich viele Glieder der Folge (a_n). Deshalb liegen in wenigstens einem der Intervalle $[A_k, m_k]$ und $[m_k, B_k]$ unendlich viele Folgenglieder a_n. Sei

$$I_{k+1} = \begin{cases} [A_k, m_k], & \text{falls } \{n \in \mathbb{N} \mid a_n \in [A_k, m_k]\} \text{ unendlich} \\ [m_k, B_k], & \text{sonst.} \end{cases}$$

Dann ist $\{I_k \mid k \in \mathbb{N}\}$ eine Intervallschachtelung und es gelten 1) - 3). Es sei

$$\{h\} = \bigcap_{k \in \mathbb{N}} I_k.$$

Dabei haben wir das Intervallschachtelungsprinzip für $\mathbb{R}$ benutzt, das die Existenz von h garantiert. Für dieses h gilt: $h \in I_k$ für alle $k \in \mathbb{N}$. Es sei $\epsilon > 0$ gegeben. Dann existiert $k \in \mathbb{N}$ so, daß

$$(B - A)2^{-k+1} < \epsilon.$$

Aus Eigenschaft 3) der Intervallschachtelung folgt:
Es existieren unendlich viele $n \in \mathbb{N}$ mit $a_n \in I_k$. Hieraus folgt, daß für unendlich viele $n \in \mathbb{N}$:

$$|h - a_n| \leq |I_k| = 2^{-k+1}(B - A) < \epsilon.$$

$\Rightarrow \quad h$ ist Häufungswert von (a_n).

$\square$

Eine äquivalente Fassung des Satzes von Bolzano - Weierstraß ist folgendes Korollar.

Korollar 4.20. *Jede beschränkte Folge besitzt eine konvergente Teilfolge.*

Beweis: Es sei (a_n) eine beschränkte Folge. Aus Satz 4.19 folgt die Existenz eines Häufungswertes h der Folge (a_n), und aus Satz 4.18 erhalten wir die Existenz einer Teilfolge (a_{n_k}) von (a_n), die gegen h konvergiert.

□

Satz 4.21. *Es sei (a_n) eine beschränkte Folge. Dann gilt:*

$$(a_n) \text{ ist konvergent} \Leftrightarrow (a_n) \text{ hat genau einen Häufungswert}$$

Beweis: $\Rightarrow$) Es sei (a_n) eine konvergente Folge und $a = \lim_{n\to\infty} a_n$. Dann ist a ein Häufungswert von (a_n). Es sei a' ein beliebiger Häufungswert von (a_n).
Annahme: $a' \neq a$.
Es sei $\epsilon = |a - a'|$. Dann ist $\epsilon > 0$. Zu diesem ϵ existiert $N \in \mathbb{N}$ so, daß

$$\forall n \geq N\colon\ |a_n - a| < \frac{\epsilon}{2}.$$

Da a' ein Häufungswert von (a_n) ist, existiert ein $m \geq N$ mit

$$|a_m - a'| < \frac{\epsilon}{2}.$$

$\Rightarrow$

$$\begin{aligned}\epsilon = |a - a'| &= |a' - a_m + a_m - a| \\ &\leq |a' - a_m| + |a_m - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

Dies ist ein Widerspruch und daher ist unsere Annahme falsch, d.h. $a = a'$.
$\Leftarrow$) Es sei (a_n) eine beschränkte Folge und a sei der einzige Häufungswert von (a_n).
Annahme: (a_n) ist nicht konvergent. Dann konvergiert insbesondere (a_n) nicht gegen a. Die Verneinung von (4.1) ergibt

$$\exists \epsilon > 0\ \forall N \in \mathbb{N}\ \exists n \geq N\colon |a_n - a| \geq \epsilon.$$

Daraus folgt, daß eine Teilfolge (a_{n_k}) von (a_n) existiert so, daß

$$|a_{n_k} - a| \geq \epsilon \ \text{ für alle } k \in \mathbb{N}.$$

Da (a_n) beschränkt ist, ist auch $(a_{n_k})_{k\in\mathbb{N}}$ eine beschränkte Folge. Aus Satz 4.19 folgt, daß (a_{n_k}) einen Häufungswert a' besitzt. Da a' Häufungswert von (a_{n_k}) ist, existiert $m \in \mathbb{N}$:

$$|a_{n_m} - a'| < \frac{\epsilon}{2}.$$

$\Rightarrow$

$$\begin{aligned} |a - a'| &= |a - a_{n_m} + a_{n_m} - a'| \\ &\geq |a - a_{n_m}| - |a' - a_{n_m}| \geq \epsilon - \frac{\epsilon}{2} = \frac{\epsilon}{2}. \end{aligned}$$

$\Rightarrow \quad a \neq a'$.
Dies ist ein Widerspruch zur Voraussetzung, daß (a_n) genau einen Häufungswert hat. Deshalb ist unsere Annahme falsch und (a_n) ist konvergent.

□

Satz 4.22. *Es sei (a_n) eine beschränkte Folge. Dann hat (a_n) einen größten Häufungswert h^* und einen kleinsten Häufungswert h_*.*

Beweis: Es sei $M = \{h \in \mathbb{R} \mid h \text{ ist Häufungswert von } (a_n)\}$
Dann gilt:

1) M ist beschränkt.
   Nach Voraussetzung existieren $A, B \in \mathbb{R}$, $A < B$, so, daß

$$\forall n \in \mathbb{N}\colon \ A \leq a_n \leq B.$$

$$\Rightarrow \ \forall h \in M\colon \ A \leq h \leq B.$$

2) $M \neq \emptyset$.
   Dies folgt aus Satz 4.19.

Da M eine nichtleere, beschränkte Teilmenge von $\mathbb{R}$ ist, existiert auf Grund der Vollständigkeit von $\mathbb{R}$ die Zahl

$$h^* = \sup M.$$

Behauptung: $h^* \in M$, d.h., h^* ist Häufungswert von M.
Beweis: Es sei $\epsilon > 0$ gegeben. Dann existiert $h \in M$:

$$h^* - \epsilon/2 < h \leq h^*.$$

Da h ein Häufungswert von (a_n) ist, folgt aus Satz 4.18, daß eine Teilfolge (a_{n_k}) von (a_n) existiert mit

$$\lim_{k\to\infty} a_{n_k} = h.$$

$\Rightarrow$

$$\exists N \in \mathbb{N}\ \forall k \geq N\colon\ |h - a_{n_k}| < \frac{\epsilon}{2}$$

$\Rightarrow$

$$\forall k \geq N\colon\ h^* - \epsilon < h - \frac{\epsilon}{2} < a_{n_k} < h + \frac{\epsilon}{2} < h^* + \epsilon.$$

Anders formuliert bedeutet dies

$$\forall k \geq N\colon\ |h^* - a_{n_k}| < \epsilon.$$

Damit haben wir gezeigt, daß in jeder ϵ- Umgebung von h^* unendlich viele Folgenglieder a_n liegen, d.h., h^* ist Häufungswert von (a_n). Auf Grund der Definition von M bedeutet dies $h^* \in M$. Da $h^* = \sup M$, ist h^* der größte Häufungswert von (a_n). Die Existenz von h_* zeigt man ebenso.

□

Bezeichnung:

$$h^* = \limsup a_n, \quad h_* = \liminf a_n.$$

Korollar 4.23. *Es sei (a_n) eine beschränkte Folge. Dann gilt*

$$(a_n)\ \textit{ist konvergent} \Leftrightarrow \limsup a_n = \liminf a_n$$

Beweis: Aus Satz 4.21 folgt:

$$\begin{aligned}(a_n)\ \text{ist konvergent} &\Leftrightarrow (a_n)\ \text{hat genau einen Häufungswert}\\ &\Leftrightarrow \limsup a_n = \liminf a_n.\end{aligned}$$

□

Beispiele:

1) Es sei $a_n = (-1)^n + \frac{1}{n}, \quad n \in \mathbb{N}$. Dann gilt

$$\limsup a_n = 1,\ \liminf a_n = -1.$$

2) $\mathbb{Q} \cap [0,1]$ ist eine abzählbare Menge. Es sei

$$\mathbb{Q} \cap [0,1] = \{a_1, a_2, \dots\}.$$

Dann ist
$[0,1] = \{r \mid r \text{ ist Häufungswert von } (a_n)\}$
$\Rightarrow \limsup a_n = 1,\ \liminf a_n = 0.$

4.6. Cauchy - Folgen

Wir lernen in diesem Abschnitt ein weiteres wichtiges Konvergenzkriterium für Folgen kennen.

Satz 4.24. *(Cauchysches Konvergenzkriterium) Sei (a_n) eine Folge. Dann gilt:*

$$(a_n) \text{ ist konvergent} \Leftrightarrow \forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N\colon\ |a_n - a_m| < \epsilon.$$

Beweis: $\Rightarrow$) Es sei (a_n) eine konvergente Folge und

$$a = \lim_{n \to \infty} a_n.$$

Gegeben sei $\epsilon > 0$. Dann existiert ein $N \in \mathbb{N}$:

$$\forall k \geq N\colon\ |a_k - a| < \frac{\epsilon}{2}.$$

Für alle $n, m \geq N$ gilt dann

$$\begin{aligned} |a_n - a_m| &= |a_n - a + a - a_m| \\ &\leq |a_n - a| + |a_m - a| < \epsilon. \end{aligned}$$

$\Leftarrow$) Es sei (a_n) eine Folge, für die das Cauchy-Kriterium gilt, d.h.,

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N\colon\ |a_n - a_m| < \epsilon. \tag{4.11}$$

1) Behauptung: (a_n) ist beschränkt.
Beweis: Nach Voraussetzung existiert ein $N \in \mathbb{N}$:

$$\forall n, m \geq N\colon\ |a_n - a_m| < 1.$$

Es sei $n \geq N$. Dann folgt daraus

$$|a_n| = |a_N + a_n - a_N| \leq |a_N| + |a_n - a_N| < |a_N| + 1.$$

Es sei

$$C = \max\{|a_1, |a_2|, \ldots, |a_{n-1}|, |a_N| + 1\}.$$

Dann gilt

$$\forall n \geq N\colon\ |a_n| \leq C.$$

Da (a_n) eine beschränkte Folge ist, folgt aus Korollar 4.20, daß (a_n) eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ besitzt. Es sei

$$a = \lim_{k\to\infty} a_{n_k}.$$

Wir zeigen jetzt, daß a Grenzwert der gesamten Folge (a_n) ist. Sei $\epsilon > 0$ gegeben. Dann existiert $k \in \mathbb{N}$ so, daß

$$|a_{n_k} - a| < \frac{\epsilon}{2}$$

und aus (4.11) folgt, daß $N_1 \in \mathbb{N}$ existiert, so daß

$$\forall n, m \geq N_1\colon\ |a_n - a_m| < \frac{\epsilon}{2}.$$

Es sei $n \geq N_1$. Dann erhalten wir aus diesen beiden Ungleichungen

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Wir haben damit gezeigt:

$$\forall n \geq N_1\colon\ |a_n - a| < \epsilon.$$

Da $\epsilon > 0$ beliebig war, folgt hieraus, daß (a_n) konvergiert und es gilt

$$a = \lim_{n\to\infty} a_n.$$

□

Definition 4.25. *Eine Folge (a_n) heißt* **Cauchyfolge**, *wenn gilt:*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N\colon\ |a_n - a_m| < \epsilon.$$

Beispiel: Für jedes $n \in \mathbb{N}$ sei b_n eine Zahl in $\{0, 1, 2, \ldots, 9\}$ und es sei $b_0 \in \mathbb{Z}$. Sei

$$a_n = \sum_{k=0}^{n} \frac{b_k}{10^k}.$$

Dann gilt:

1) $\forall n \in \mathbb{N}\colon\ a_n \in \mathbb{Q}$.
2) Es sei $n > m$. Dann ist

$$|a_n - a_m| = \sum_{k=m+1}^{n} \frac{b_k}{10^k} \leq \sum_{k=m+1}^{n} 10^{-k+1} = 10^{-m} \sum_{k=0}^{n-m-1} 10^{-k}$$
$$10^{-m} \frac{1 - 10^{-(n-m)}}{1 - 10^{-1}} \leq 2 \cdot 10^{-m}.$$

$\Rightarrow$ (a_n) ist eine Cauchyfolge. Es sei $r = \lim_{n\to\infty} a_n$. Dann ist $r \in \mathbb{R}$. Umgekehrt gibt es natürlich zu jeder reellen Zahl r eine Folge (r_n) rationaler Zahlen so, daß $r = \lim_{n\to\infty} r_n$. Auf Grund von Satz 4.24 ist (r_n) eine Cauchyfolge.

Es sei

$$\mathcal{C}(\mathbb{Q}) = \{(r_n) \mid \forall n \in \mathbb{N}\colon r_n \in \mathbb{Q} \wedge (r_n) \text{ ist Cauchyfolge } \}.$$

Das ist der Raum der Cauchyfolgen rationaler Zahlen. Es seien $(a_n), (b_n) \in \mathcal{C}(\mathbb{Q})$ und $\lambda \in \mathbb{Q}$. Wie man mittels der Definition 4.25 sofort nachprüft, sind $(a_n + b_n)$ und (λa_n) ebenfalls Cauchyfolgen rationaler Zahlen, d.h., $(a_n + b_n),\ (\lambda a_n) \in \mathcal{C}(\mathbb{Q})$. Damit können wir $\mathcal{C}(\mathbb{Q})$ mit der Struktur eines $\mathbb{Q}$- Vektorraumes versehen, indem die Addition von zwei Elementen und die Multiplikation mit einer rationalen Zahl λ definieren durch

$$(a_n) + (b_n) = (a_n + b_n)$$
$$\lambda(a_n) = (\lambda a_n).$$

Die Zuordnung $(r_n) \longmapsto \lim_{n\to\infty} r_n$ induziert eine surjektive lineare Abbildung

$$c\colon \mathcal{C}(\mathbb{Q}) \to \mathbb{R}$$

von $\mathbb{Q}$ - Vektorräumen. Für den Kern gilt

$$\operatorname{Ker}(c) = \{(r_n) \mid \lim_{n\to\infty} r_n = 0\},$$

d.h., der Kern besteht aus allen Nullfolgen rationaler Zahlen. Wir erhalten damit einen Isomorphismus von $\mathbb{Q}$ - Vektorräumen

$$\mathcal{C}(\mathbb{Q})/\mathrm{Ker}\,(c) \xrightarrow{\sim} \mathbb{R}.$$

Diese Verfahren kann man umgekehrt benutzen, um die reellen Zahlen mittels Cauchyfolgen rationaler Zahlen zu konstruieren. Dazu führen wir eine Äquivalenzrelation

$$(a_n) \sim (b_n)$$

zwischen Cauchyfolgen ein.

Definition. *Seien (a_n) und (b_n) Cauchyfolgen. Dann nennen wir (a_n) und (b_n) äquivalent, in Zeichen $(a_n) \sim (b_n)$, genau dann, wenn $(a_n - b_n)$ eine Nullfolge ist.*

Es sei

$$\tilde{\mathbb{R}} = \mathcal{C}(\mathbb{Q})/\sim .$$

In $\tilde{\mathbb{R}}$ kann man die Addition und die Multiplikation von zwei Äquivlalenzklassen $[(a_n)]$ und $[(b_n)]$ definieren durch

$$[(a_n)] + [(b_n)] = [(a_n + b_n)]$$
$$[(a_n)] \cdot [(b_n)] = [(a_n \cdot b_n)].$$

Mit Hilfe der Rechenreglen für Grenzwerte prüft man sofort nach, daß die rechten Seiten unabhängig von der Auswahl der Repräsentanten der Äquivalenzklassen sind. Entsprechend kann man in $\tilde{\mathbb{R}}$ eine Ordunungsrelation $<$ einführen. Man kann dann zeigen, daß $\tilde{\mathbb{R}}$ ein archimedisch angeordneter Körper ist.
Schießlich beweist man, daß $\tilde{\mathbb{R}}$ auch vollständig ist. Damit erfüllt $\tilde{\mathbb{R}}$ alle Axiome von $\mathbb{R}$, d.h., wir haben ein Modell für $\mathbb{R}$ konstruiert. Damit ist $\mathbb{R} \cong \tilde{\mathbb{R}}$. Die Vervollständigung von $\mathbb{Q}$ mittels Äquivalenzklassen von Cauchyfolgen ist eine sehr wichtige Konstruktion der Analysis. Sie wird in vielen anderen Fällen angewendet. So kann man z.B. den Raum der stetigen Funktionen auf einem Intervall (a, b) vervollständigen und erhält den Hilbertraum $L^2((a, b))$.
Hinsichtlich der Vollständigkeit von $\mathbb{R}$ gilt folgender wichtiger Satz.

Satz. *Die folgenden Aussagen sind äquivalent.*

a) *Zu jeder Intervallschachtelung (I_k) in $\mathbb{R}$ existiert genau ein $r \in \mathbb{R}$ so, daß $r \in I_k$ für alle $k \in \mathbb{N}$.*
b) *Jede nichtleere, beschränkte Menge reeller Zahlen hat ein Suprenum.*

c) *Jede beschränkte Folge reeller Zahlen hat eine konvergente Teilfolge.*

d) *Jede Cauchyfolge in* $\mathbb{R}$ *hat einen Grenzwert in* $\mathbb{R}$.

Beweis: Wir wissen bereits aus Kapitel 2, daß a) und b) äquivalent sind. Weiter haben wir bei den Beweisen von Satz 4.19, Korollar 4.20 und Satz 4.24 folgendes gezeigt:
a) $\Rightarrow$ *c)* $\Rightarrow$ *d)*.
Es genügt zu zeigen, daß d) $\Rightarrow$ a) gilt. Dazu sei $\{I_k \mid k \in \mathbb{N}\}$, $I_k = [a_k, b_k]$, eine Intervallschachtelung. Es sei $\epsilon > 0$ gegeben. Dann existiert $N \in \mathbb{N}$ so, daß

$$b_N - a_N < \epsilon.$$

Es sei $n > N$. Dann ist $I_n \subset I_N$. Daraus folgt:

$$\forall m, n > N\colon\ a_m, a_n \in [a_N, b_N].$$

Für alle $m, n > N$ gilt daher

$$|a_m - a_n| < b_N - a_N < \epsilon.$$

Damit haben wir folgendes gezeigt:

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall m, n > N\colon\ |a_m - a_n| < \epsilon,$$

d.h., (a_n) ist eine Cauchyfolge. Es sei

$$s = \lim_{n \to \infty} a_n.$$

Da (a_n) monoton wächst, gilt

$$\forall n \in \mathbb{N}\colon\ a_n \leq s.$$

Da für alle $n, k \in \mathbb{N}\colon\ a_n \leq b_k$, folgt

$$\forall n \in \mathbb{N}\colon\ s \leq b_n.$$

Hieraus folgt

$$s \in \bigcap_{n \in \mathbb{N}} I_n.$$

Da ein solches s eindeutig bestimmt ist, gilt

$$\{s\} = \bigcap_{n\in\mathbb{N}} I_n.$$

□

4.7. Folgen komplexer Zahlen

Definition 4.27. *Eine Folge komplexer Zahlen ist eine Abbildung*

$$f\colon \mathbb{N} \to \mathbb{C}.$$

Es sei $z_n = f(n)$. Dann schreibt man die durch f definierte Folge in der Form

$$(z_n),\ (z_n)_{n\in\mathbb{N}}, \quad \text{oder} \quad z_1, z_2, z_3, \dots$$

Definition 4.28. *Es sei z_n eine Folge in $\mathbb{C}$. Die Folge (z_n) heißt* **konvergent**, *wenn ein $z \in \mathbb{C}$ existiert so, daß gilt*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\colon\ |z_n - z| < \epsilon.$$

Dann heißt z **Grenzwert** *oder* **Limes** *der Folge (z_n).*

Bezeichnung: $z = \lim_{n\to\infty} z_n$.

Satz 4.29. *Es sei (z_n) eine Folge in $\mathbb{C}$, und für alle $n \in \mathbb{N}$ sei*

$$x_n = \mathrm{Re}(z_n), \quad y_n = \mathrm{Im}(z_n).$$

Dann gilt:

$$(z_n)\ \textit{ist konvergent}\ \Leftrightarrow (x_n)\ \textit{und}\ (y_n)\ \textit{konvergieren in}\ \mathbb{R}.$$

Beweis: $\Rightarrow$) Es sei (z_n) konvergent und $z = \lim_{n\to\infty} z_n$. Weiter sei

$$x = \mathrm{Re}(z), \quad y = \mathrm{Im}(z).$$

Sei $\epsilon > 0$ gegeben. Zu ϵ existiert ein $N \in \mathbb{N}$ mit

$$\forall n \geq N\colon \quad |z_n - z| < \epsilon.$$

Daraus folgt

$$\forall n \geq N\colon\ \sqrt{(x - x_n)^2 + (y - y_n)^2} < \epsilon$$

Da

$$|x - x_n|, |y - y_n| \leq \sqrt{(x - x_n)^2 + (y - y_n)^2},$$

erhalten wir

$$\forall n \geq N\colon\ |x - x_n| < \epsilon\ \wedge\ |y - y_n| < \epsilon.$$

Daraus folgt, daß (x_n) und (y_n) konvergente Folgen sind mit

$$x = \lim_{n\to\infty} x_n, \quad y = \lim_{n\to\infty} y_n.$$

$\Leftarrow$) Es seien (x_n) und (y_n) konvergent und es sei

$$x = \lim_{n\to\infty} x_n, \quad y = \lim_{n\to\infty} y_n.$$

Dann existiert zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$:

$$\forall n \geq N\colon\ |x - x_n| < \frac{\epsilon}{\sqrt{2}}\ \wedge\ |y - y_n| < \frac{\epsilon}{\sqrt{2}}.$$

Es sei $z = x + iy$. Dann folgt aus diesen Ungleichungen, daß für alle $n \geq N$ gilt:

$$|z - z_n| = \sqrt{(x - x_n)^2 + (y - y_n)^2} < \sqrt{\frac{\epsilon^2}{2} + \frac{\epsilon^2}{2}} = \epsilon.$$

Dies bedeutet, daß (z_n) konvergiert und es gilt

$$\lim_{n\to\infty} z_n = z.$$

□

Mit Hilfe von Satz 4.29 können alle Resultate für reelle Folgen, bei denen die Anordnung keine Rolle spielt, auf komplexe Folgen übertragen werden.

Definition 4.30. *Es sei (z_n) eine Folge in $\mathbb{C}$. Die Folge (z_n) heißt beschränkt, wenn ein $C \in \mathbb{R}_+$ existiert so, daß gilt*

$$\forall n \in \mathbb{N}\colon\ |z_n| \leq C.$$

Satz 4.31. *Es sei (z_n) eine konvergente Folge. Dann ist (z_n) beschränkt.*

Beweis: Es sei $x_n = \mathrm{Re}(z_n)$ und $y_n = \mathrm{Im}(z_n)$, $n \in \mathbb{N}$. Aus Satz 4.29 folgt, daß (x_n) und (y_n) konvergieren. Aus Satz 4.7 folgt, daß ein $C \in \mathbb{R}_+$ existiert so, daß

$$\forall n \in \mathbb{N}\colon\ |x_n|, |y_n| \leq C.$$

$\Rightarrow$

$$|z_n| = \sqrt{x_n^2 + y_n^2} \le \sqrt{2}C.$$

□

Satz 4.32. *Jede beschränkte Folge komplexer Zahlen besitzt eine konvergente Teilfolge.*

Beweis: Es sei (z_n) eine beschränkte Folge in $\mathbb{C}$. Für alle $n \in \mathbb{N}$ sei

$$x_n = \mathrm{Re}(z_n), \quad y_n = \mathrm{Im}(z_n).$$

Da für alle $n \in \mathbb{N}$ gilt

$$|x_n|, |y_n| \le |z_n|,$$

so sind (x_n) und (y_n) beschränkte Folgen reeller Zahlen. Aus Korollar 4.20 folgt, daß (x_n) eine konvergente Teilfolge (x_{n_k}) besitzt. Ohne Beschränkung der Allgemeinheit können wir deshalb annehmen, daß für die Folge (z_n) die Folge $(\mathrm{Re}(z_n))$ bereits konvergiert.
Aus Korollar 4.20 folgt, daß (y_n) eine konvergente Teilfolge (y_{n_k}) besitzt. Dann sind (x_{n_k}) und (y_{n_k}) konvergent und aus Satz 4.29 folgt, daß (z_{n_k}) konvergiert.

□

Definition 4.33. *Eine komplexe Zahl $z \in \mathbb{C}$ heißt* **Häufungswert** *(oder* **Häufungspunkt***) der Folge (z_n), wenn für alle $\epsilon > 0$ unendlich viele $n \in \mathbb{N}$ existieren so, daß*

$$|z_n - z| < \epsilon.$$

Satz 4.34. *(Satz von Bolzano - Weierstraß)*
Jede beschränkte Folge in $\mathbb{C}$ hat einen Häufungswert.

Beweis: Es sei (z_n) eine beschränkte Folge. Aus Satz 4.32 folgt, daß (z_n) eine konvergente Teilfolge (z_{n_k}) besitzt. Es sei $z = \lim_{k\to\infty} z_{n_k}$. Dann ist z ein Häufungswert von (z_{n_k}) und damit auch von (z_n).

□

Satz 4.35. *(Cauchy - Kriterium)*
Eine Folge (z_n) komplexer Zahlen ist genau dann konvergent, wenn gilt:

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \ge N\colon\ |z_n - z_m| < \epsilon.$$

Beweis: $\Leftarrow$) Es sei (z_n) konvergent. Für alle $n \in \mathbb{N}$ sei

$$x_n = \mathrm{Re}(z_n), \quad y_n = \mathrm{Im}(z_n).$$

Aus Satz 4.29 folgt, daß (x_n) und (y_n) konvergieren. Es sei $\epsilon > 0$ gegeben. Auf Grund des Cauchy - Kriteriums für reelle Folgen (Satz 4.24) existiert zu ϵ ein $N \in \mathbb{N}$ so, daß gilt

$$\forall n, m \geq N\colon\ |x_n - x_m| < \frac{\epsilon}{\sqrt{2}} \ \wedge\ |y_n - y_m| < \frac{\epsilon}{\sqrt{2}}.$$

$\Rightarrow$

$$\forall n, m \geq N\colon |z_n - z_m| = \sqrt{(x_n - x_m)^2 + (y_n - y_m)^2} < \epsilon.$$

$\Leftarrow$) Es sei (z_n) eine Cauchyfolge in $\mathbb{C}$, d.h., für (z_n) gelte das Cauchy-kriterium. Es sei $\epsilon > 0$ gegeben. Zu ϵ existiert $N \in \mathbb{N}$:

$$\forall n, m \geq N\colon\ |z_n - z_m| < \epsilon.$$

Für $n, m \geq N$ erhalten wir

$$|x_n - x_m| \leq |z_n - z_m| < \epsilon$$
$$|y_n - y_m| \leq |z_n - z_m| < \epsilon.$$

Daraus folgt, daß (x_n) und (y_n) Cauchyfolgen sind. Aus Satz 4.24. folgt, daß (z_n) konvergiert.

$\square$

4.8. Uneigentliche Konvergenz

Es sei (a_n) eine Folgen in $\mathbb{R}$.

5. Reihen

Gegeben sei eine Folge (a_n) in $\mathbb{C}$. Wir ordnen der Folge (a_n) eine neue Folge (s_n) wie folgt zu:

$$s_n = \sum_{k=1}^{n} a_k, \quad n \in \mathbb{N}.$$

Definition 5.1. *Die Folge (s_n) heißt (unendliche)* **Reihe** *mit den Gliedern a_n. Die s_n heißen* **Partialsummen.**

Bezeichnung: Man schreibt für die Reihe

$$\sum_{k=1}^{\infty} a_k \quad \text{oder} \quad a_1 + a_2 + a_3 + \cdots$$

Definition 5.2. *Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt* **konvergent,** *wenn die Folge (s_n) der Partialsummen konvergiert. Andernfalls heißt die Reihe* **divergent.**

Sei (s_n) konvergent und $s = \lim_{n\to\infty} s_n$.
Bezeichnung: $s = \sum_{n=1}^{\infty} a_n$.

Bemerkung: $\sum_{n=1}^{\infty} a_n$ hat zwei Bedeutungen.

a) Bezeichnung der Reihe, die durch (a_n) definiert wird.
b) Im konvergenten Falle bezeichnet $\sum_{n=1}^{\infty} a_n$ den Grenzwert

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

Beispiele:

1) *Die geometrische Reihe*

Es sei $z \in \mathbb{C}$ und $a_n = z^n$, $n \in \mathbb{N} \cup \{0\}$. Dann ist

$$s_n = 1 + z + \cdots z^n = \frac{1 - z^{n+1}}{1 - z}.$$

Es sei $|z| < 1$. Dann gilt: $|z|^n \to 0$. Weiter ist

$$\frac{1-z^{n+1}}{1-z} = \frac{1}{1-z} - \frac{z^{n+1}}{1-z}$$

und aus den Rechenregeln für Grenzwerte folgt

$$s_n \xrightarrow[n\to\infty]{} \frac{1}{1-z}.$$

Daher ist $\sum_{k=0}^{\infty} z^k$ für $|z| < 1$ konvergent und es gilt

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

2) *Die harmonische Reihe*

$\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent, denn für $k \in \mathbb{N}$ und $n \geq 2^k$ gilt:

$$\begin{aligned} s_n = \sum_{j=1}^{n} \frac{1}{j} \geq \sum_{j=1}^{2^k} \frac{1}{j} = 1 + \sum_{l=1}^{k} \sum_{j=2^{l-1}+1}^{2^l} \frac{1}{j} \geq \\ \geq 1 + \sum_{l=1}^{k} \frac{2^{l-1}}{2^l} = 1 + \frac{k}{2} \end{aligned}$$

Daraus folgt, daß $s_n \to \infty$ für $n \to \infty$ und daher ist $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

3) *Die Reihe* $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$.
Es gilt:

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

Daraus folgt:

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

Daher ist $\lim_{n\to\infty} s_n = 1$, und somit

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

4) Es sei $a_n = (-1)^n, \quad n \in \mathbb{N}$. Dann ist

$$s_n = \begin{cases} 1, & \text{falls } n \text{ gerade} \\ 0, & \text{falls } n \text{ ungerade.} \end{cases}$$

Die Folge (s_n) ist divergent und daher ist $\sum_{n=1}^{\infty}(-1)^n$ divergent.

Lemma 5.3. *Wenn die Reihe* $\sum_{n=1}^{\infty} a_n$ *konvergiert, so ist* (a_n) *eine Nullfolge, d.h.* $a_n \xrightarrow[n\to\infty]{} 0$.

Beweis: Es sei $s_n = \sum_{k=1}^{n} a_k$ und es gelte $\lim_{n\to\infty} s_n = s$. Dann gilt:

$$a_n = s_n - s_{n-1} \xrightarrow[n\to\infty]{} s - s = 0.$$

□

Damit ist eine notwendige Bedingung für die Konvergenz einer Reihe $\sum a_n$, daß die Glieder a_n der Reihe eine Nullfolge bilden.

Bemerkung: $a_n \to 0$ ist **keine** hinreichende Bedingung für die Konvergenz von $\sum a_n$. In Beispiel 2 haben wir gezeigt, daß $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert. Die Glieder $a_n = n^{-1}$ bilden jedoch eine Nullfolge.

5.1. Konvergenzkriterien für Reihen

Da eine Reihe $\sum_{n=1}^{\infty} a_n$ durch die Folge (s_n) ihrer Partialsummen definiert wird, ergeben sich aus den Konvergenzkriterien für Folgen entsprechende Konvergenzkriterien für Reihen.

a) Monotoniekriterium

Satz 5.4. *Es sei* $\sum_{n=1}^{\infty} a_n$ *eine Reihe mit reellen Gliedern* a_n *und es gelte* $a_n \geq 0$ *für alle* $n \in \mathbb{N}$. *Dann konvergiert* $\sum_{n=1}^{\infty} a_n$ *genau dann, wenn die Folge der Partialsummen* (s_n) *beschränkt ist.*

Beweis: Es sei $s_n = \sum_{k=1}^{n} a_k, \quad n \in \mathbb{N}$. Da $a_k \geq 0$ für alle $k \in \mathbb{N}$ ist, gilt

$$s_{n+1} = s_n + a_{n+1} \geq s_n \quad \text{für alle} \quad n \in \mathbb{N}.$$

Daher ist (s_n) eine monoton wachsende Folge. Aus dem Satz 4.13 über monotone beschränkte Folgen ergibt sich damit:

$$(s_n) \text{ ist beschränkt } \Rightarrow (s_n) \text{ ist konvergent.}$$

Andererseits ist nach Satz 4.8 jede konvergente Folge beschränkt. Daraus folgt:

$$(s_n)\,\text{konvergent} \;\Rightarrow (s_n)\,\text{beschränkt.}$$

□

Beispiel: Für $s \in \mathbb{Q}$ gilt:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \quad \text{ist} \quad \begin{cases} \text{konvergent, falls} \quad s > 1 \\ \text{divergent, falls} \quad s \leq 1. \end{cases}$$

Beweis: i) Es sei $s > 1$. Für $n \in \mathbb{N}$ wählen wir $p \in \mathbb{N}$ so daß $2^p - 1 \geq n$. Da $k^{-s} > 0$ ist, folgt

$$s_{2^p-1} = s_n + \sum_{k=n+1}^{2^p-1} \frac{1}{k^s} \geq s_n.$$

Weiter ist

$$\begin{aligned} s_{2^p-1} &= \sum_{k=1}^{2^p-1} \frac{1}{k^s} = \sum_{l=0}^{p-1} \quad \sum_{k=2^l}^{2^{l+1}-1} \frac{1}{k^s} \leq \sum_{l=0}^{p-1} \frac{2^l}{2^{ls}} = \sum_{l=0}^{p-1} \left(\frac{1}{2^{s-1}}\right)^l \\ &< \sum_{l=0}^{\infty} \left(\frac{1}{2^{s-1}}\right)^l. \end{aligned}$$

Da $s > 1$, ist $2^{s-1} > 1$ und daher

$$\frac{1}{2^{s-1}} < 1.$$

Deshalb konvergiert die geometrische Reihe

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

und wir erhalten aus den Abschätzungen

$$s_n < \frac{1}{1 - 2^{1-s}} \quad \text{für alle} \quad n \in \mathbb{N}.$$

Da $n^{-s} \geq 0$ für alle $n \in \mathbb{N}$, folgt aus Satz 5.4, daß $\sum_{n=1}^{\infty} n^{-s}$ konvergiert.

ii) Es sei $s \leq 1$. Dann gilt: $n^s \leq n$ für alle $n \in \mathbb{N}$. Hieraus folgt

$$\frac{1}{n^s} \geq \frac{1}{n}$$

und wir erhalten die Abschätzung

$$s_n = \sum_{k=1}^{n} \frac{1}{k^s} \geq \sum_{k=1}^{n} \frac{1}{k} \xrightarrow[n\to\infty]{} \infty,$$

wobei die letzte Aussage aus Beispiel 2 folgt.

b) Das Leibnizkriterium

Satz 5.5. *Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine monoton fallende Folge reeller Zahlen und es gelte $a_n \xrightarrow[n\to\infty]{} 0$. Dann gilt:*

1) *Die Reihe $\sum_{n=0}^{\infty}(-1)^n a_n$ ist konvergent,*
2) *Es sei $s = \sum_{n=0}^{\infty}(-1)^n a_n$. Dann ist*

$$|s - s_n| \leq a_{n+1} \quad \textit{für alle} \quad n \in \mathbb{N}.$$

Beweis: Für die Partialsummen s_n gilt

$$s_n - s_{n-2} = (-1)^n(a_n - a_{n-1}).$$

Da $a_n \leq a_{n-1}$ für alle $n \in \mathbb{N}$ ist, folgt:

$$\begin{aligned} s_{2k} &\leq s_{2k-2} \quad \text{für} \quad k \in \mathbb{N}, \\ s_{2k+1} &\geq s_{2k-1} \quad \text{für} \quad k \in \mathbb{N}. \end{aligned}$$

Damit ist (s_{2k+1}) eine monoton wachsende Folge und (s_{2k}) eine monoton fallende Folge.

Lemma. $\forall n \in \mathbb{N}\colon a_n \geq 0$.

Beweis: Wir nehmen an, es existiert $k \in \mathbb{N}$ so, daß $a_k < 0$. Da (a_n) monoton fallend ist, gilt:

$$a_n \leq a_k < 0 \quad \text{für alle} \quad n \geq k.$$

Auf Grund des Vergleichsprinzips für Grenzwerte (Satz 4.10) folgt daraus

$$\lim_{n\to\infty} a_n \leq a_k < 0.$$

Dies steht im Widerspruch zur Annahme, daß $a_n \to 0$.

$\square$

Aus dem Lemma folgt

$$s_{2k} - s_{2k-1} = a_{2k} \geq 0.$$

Wir können annehmen, daß $a_k \neq 0$ für alle $k \in \mathbb{N}$. Dann ist

$$s_{2k} > s_{2k-1}.$$

Für $k \in \mathbb{N}$ sei

$$I_k = [s_{2k-1}, s_{2k}].$$

Dann gilt:

i) $I_{k+1} \subseteq I_k$ für alle $k \in \mathbb{N}$

ii) $|I_k| = s_{2k} - s_{2k-1} = a_{2k} \xrightarrow[k\to\infty]{} 0$.

Daher ist $\{I_k \mid k \in \mathbb{N}\}$ eine Intervallschachtelung. Es sei $s \in \mathbb{R}$ die eindeutig bestimmte Zahl mit

$$\{s\} = \bigcap_{k\in\mathbb{N}} I_k,$$

die auf Grund der Vollständigkeit von $\mathbb{R}$ existiert. Dann gilt

$$s_{2k-1} \leq s \leq s_{2k}.$$

Daraus folgt

$$\begin{aligned} |s_{2k} - s| &\leq |s_{2k} - s_{2k-1}| = a_{2k} \xrightarrow[k\to\infty]{} 0, \\ |s_{2k-1} - s| &\leq |s_{2k} - s_{2k-1}| = a_{2k} \xrightarrow[k\to\infty]{} 0. \end{aligned}$$

Damit ist

$$s = \lim_{n\to\infty} s_n,$$

und der erste Teil der Behauptung ist bewiesen. Wir müssen noch den Fehler $s - s_n$ abschätzen. Oben haben wir bereits gezeigt, daß

$$|s_{2k-1} - s| < a_{2k}$$

gilt. Weiter ist auf Grund des Lemmas

$$s_{2k} - s_{2k+1} = a_{2k+1} > 0.$$

Hieraus folgt

$$s_{2k+1} \leq s \leq s_{2k}$$

und wir erhalten

$$|s_{2k} - s| \leq |s_{2k} - s_{2k+1}| = a_{2k+1}.$$

Damit ergibt sich

$$|s - s_n| \leq a_{n+1} \quad \text{für alle} \quad n \in \mathbb{N}.$$

□

Beispiel: $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$ ist konvergent. Später zeigen wir:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln 2$$

Das Leibnizkriterium ergibt auch die Konvergenz der Reihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

Diese Reihe wurde von Leibniz betrachtet und heißt deshalb auch **Leibnizreihe**. Dies ist die berühmte Reihe für $\pi/4$, d.h. wir werden später zeigen, daß

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots$$

gilt.

c) **Cauchy-Kriterium**

Die Anwendung des Cauchy-Kriteriums für Folgen auf die Folge der Partialsummen (s_n) ergibt

Satz 5.6. *Die Reihe* $\sum_{n=1}^{\infty} a_n$ *ist konvergent genau dann, wenn gilt:*

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \colon \forall n > m \geq N \colon \left| \sum_{k=m+1}^{n} a_k \right| < \epsilon.$$

Beweis: Sei $s_n = \sum_{k=1}^{n} a_k$. Das Cauchy-Kriterium für Folgen besagt dann: Die Folge (s_n) ist konvergent genau dann, wenn gilt:

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \colon |s_n - s_m| < \epsilon \quad \text{für alle} \quad n, m \geq N.$$

O.B.d.A. können wir annehmen, daß $n > m$ ist. Dann gilt:

$$s_n - s_m = \sum_{k=1}^{n} a_k - \sum_{k=1}^{m} a_k = \sum_{k=m+1}^{n} a_k,$$

und die Anwendung des Cauchy-Kriteriums auf (s_n) ergibt die Behauptung des Satzes.

□

Beispiel: $\sum_{n=1}^{\infty} \frac{1}{n^2}$.
Sei $m > n$. Dann ist

$$\begin{aligned} s_m - s_n &= \sum_{k=n+1}^{m} \frac{1}{k^2} < \sum_{k=n+1}^{m} \frac{1}{(k-1)k} \\ &= \sum_{k=n+1}^{m} \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &= \sum_{k=n}^{m-1} \frac{1}{k} - \sum_{k=n+1}^{m} \frac{1}{k} = \frac{1}{n} - \frac{1}{m}. \end{aligned}$$

Sei $\epsilon > 0$. Dann existiert $N \in \mathbb{N}$ mit $N > \frac{1}{\epsilon}$. Für $m > n \geq N$ gilt dann $n^{-1} < \epsilon$. Daraus folgt zusammen mit der obigen Abschätzung:
Für alle $m > n \geq N$ gilt:

$$s_m - s_n < \frac{1}{n} < \epsilon,$$

und die Anwendung des Cauchy-Kriteriums ergibt die Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$, von der wir bereits wissen, daß Sie konvergent ist.

d) **Majorantenkriterium**

Dieses Kriterium ist sehr wichtig, da es die Frage nach der Konvergenz einer Reihe in bestimmten Fällen auf die Konvergenz bereits bekannter Reihen zurückführt.

Satz 5.7. *Es seien (a_n) und (c_n) Folgen in $\mathbb{C}$ und es gelte:*

1) $|a_n| \leq |c_n|$ *für $n \in \mathbb{N}$,*
2) $\sum_{n=1}^{\infty} |c_n|$ *ist konvergent.*

Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ und es gilt:

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

Beweis: i) Da $\sum_{n=1}^{\infty} |c_n|$ konvergent ist, folgt aus Satz 5.6.:

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N}\colon \sum_{k=m+1}^{n} |c_k| < \epsilon$$

für alle $n > m \geq N$.
Daraus folgt:

$$\left| \sum_{k=m+1}^{n} a_k \right| \leq \sum_{k=m+1}^{n} |a_k| \leq \sum_{k=m+1}^{n} |c_k| < \epsilon$$

für alle $n > m \geq N$. Aus dem Cauchy Kriterium (Satz 5.6) folgt daher, daß die Reihe $\sum_{n=1}^{\infty} a_n$ konvergent ist.
ii) Wir beweisen zuerst folgenden Hilfssatz.

Lemma. *Es sei (a_n) eine konvergente Folge komplexer Zahlen mit Grenzwert a. Dann gilt*

$$\lim_{n\to\infty} |a_n| = |a|.$$

Beweis: Es sei $a_n = x_n + iy_n$ die Aufspaltung von a_n in Real- und Imaginärteil $x_n \in \mathbb{R}$ bzw. $y_n \in \mathbb{R}$. Ebenso sei $x = \mathrm{Re}(a)$ und $y = \mathrm{Im}(a)$. Nach Satz 4.29 gilt:

$$a_n \to a \quad \Leftrightarrow \quad (x_n \to x) \wedge (y_n \to y).$$

Aus den Rechenregeln für Grenzwerte erhalten wir:

$$|a_n|^2 = x_n^2 + y_n^2 \to x^2 + y^2 = |a|^2.$$

Aus den Eigenschaften der Quadratwurzel folgt damit

$$|a_n| = \sqrt{|a_n|^2} \longrightarrow \sqrt{|a|^2} = |a|.$$

□

Die Anwendung des Lemmas ergibt

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

Die beiden Ungleichungen sind dabei Konsequenz des Vergleichsprinzips für Grenzwerte (Satz 4.10).

Bezeichnung: $\sum_{k=1}^{\infty} |c_k|$ heißt **Majorante** für die Reihe $\sum_{k=1}^{\infty} a_k$.

Beispiele: 1) Es sei $a_n \in \mathbb{C}$ und es gelte $|a_n| \leq 1$ für alle $n \in \mathbb{N}$. Weiter sei $z \in \mathbb{C}$ und $|z| < 1$.

Behauptung: Die Reihe $\sum_{n=1}^{\infty} a_n z^n$ ist konvergent.

Beweis: Nach Vorraussetzung gilt

$$|a_n z^n| \leq |z|^n \quad \text{für alle} \quad n \in \mathbb{N}.$$

Da $|z| < 1$ ist, konvergiert die Reihe $\sum_{n=1}^{\infty} |z|^n$ (siehe Beispiel 1 der Einleitung). Aus Satz 5.7 erhalten wir damit unmittelbar, daß die Reihe $\sum_{n=1}^{\infty} a_n z^n$ ebenfalls konvergiert.

2) $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergiert für alle $s \in \mathbb{Q}$ mit $s \geq 2$, denn für $s \geq 2$ gilt:

$$n^s \geq n^2 \quad \text{für alle} \quad n \in \mathbb{N}.$$

(siehe 2.8). Daraus folgt:

$$\frac{1}{n^s} \leq \frac{1}{n^2} \quad \text{für alle} \quad n \in \mathbb{N}.$$

Wie wir bereits wissen, ist $\sum_{n=1}^{\infty} n^{-2}$ konvergent. Aus dem Satz 5.7 folgt daraus, daß auch die Reihe $\sum_{n=1}^{\infty} n^{-s}$ konvergiert.

Wir benötigen einige Rechenregeln für Reihen. Dies ist Inhalt des folgenden Satzes.

Satz 5.8. *Es seien $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergente Reihen. Dann ist für alle $\alpha, \beta \in \mathbb{C}$ die Reihe*

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n)$$

ebenfalls konvergent und es gilt

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

Beweis: Es sei $s_n = \sum_{k=1}^{n} a_k$, $s'_n = \sum_{k=1}^{n} b_k$ und $s''_n = \sum_{k=1}^{n} (\alpha a_n + \beta b_n)$.

Dann ist

$$s''_n = \alpha s_n + \beta s'_n.$$

Nach Voraussetzung existieren $s, s' \in \mathbb{C}$ mit

$$s_n \xrightarrow[n \to \infty]{} s \quad \text{und} \quad s'_n \xrightarrow[n \to \infty]{} s'.$$

Aus den Rechenregeln für Grenzwerte von Folgen erhalten wir

$$s''_n = \alpha s_n + \beta s'_n \xrightarrow[n \to \infty]{} \alpha s + \beta s'.$$

Daraus folgt, daß die Reihe $\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n)$ konvergiert, und es gilt

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \lim_{n \to \infty} s''_n =$$

$$\alpha s + \beta s' = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

$\square$

Bemerkung: Aus der Konvergenz von $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ folgt **nicht** die Konvergenz von $\sum_{n=1}^{\infty} a_n b_n$.

Beispiel: Aus dem Leibnizkriterium folgt, daß die Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ konvergiert.

Andererseits ist

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

und wir wissen aus dem Beispiel der Einleitung, daß diese Reihe divergiert.

5.2. Absolute Konvergenz

Definition 5.9. *Es sei* $\sum_{n=1}^{\infty} a_n$ *eine Reihe mit komplexen Gliedern* a_n*. Die Reihe* $\sum_{n=1}^{\infty} a_n$ *heißt* **absolut konvergent,** *wenn die Reihe* $\sum_{n=1}^{\infty} |a_n|$ *konvergiert.*

Satz 5.10. *Es sei* $\sum_{n=1}^{\infty} a_n$ *absolut konvergent. Dann ist die Reihe* $\sum_{n=1}^{\infty} a_n$ *auch konvergent.*

Beweis: Die Reihe $\sum_{n=1}^{\infty} |a_n|$ ist eine konvergente Majorante für die Reihe $\sum_{n=1}^{\infty} a_n$. Die Behauptung folgt damit unmittelbar aus Satz 5.7.

Bemerkung: Die Umkehrung dieses Satzes gilt natürlich nicht. Z.B. ist $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ auf Grund des Leibnizkriteriums konvergent, aber $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent (siehe Beispiel 2) der Einleitung).

Satz 5.11. *(Quotientenkriterium)*
Es sei $\sum_{n=1}^{\infty} a_n$ *eine Reihe mit komplexen Gliedern* a_n *und es gelte* $a_n \neq 0$ *für fast alle* $n \in \mathbb{N}$*. Weiter existiere der Grenzwert*

$$q = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| .$$

Dann gilt:

1) *Wenn* $q < 1$*, so ist* $\sum_{n=1}^{\infty} a_n$ *absolut konvergent.*
2) *Wenn* $q > 1$*, so ist* $\sum_{n=1}^{\infty} a_n$ *divergent.*

Beweis: 1) Es sei $q = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1$.

Wir wählen $q' \in \mathbb{R}$ so daß $q < q' < 1$ gilt. Es sei $\epsilon = q' - q$. Da q Grenzwert der Folge $\left| \frac{a_{n+1}}{a_n} \right|$ ist, existiert zu diesem ϵ ein $N \in \mathbb{N}$ so daß gilt:

$$\left| q - \left| \frac{a_{n+1}}{a_n} \right| \right| < \epsilon \quad \text{für} \quad n \geq N,$$

denn da q Grenzwert der Folge $\left|\frac{a_{n+1}}{a_n}\right|$ ist, liegen in jeder Umgebung von q fast alle Glieder der Folge. Daraus folgt:

$$\left| \frac{a_{n+1}}{a_n} \right| < q + \epsilon = q' \quad \text{für alle} \quad n \geq N.$$

Sei $n \geq N$. Dann erhalten wir aus dieser Ungleichung

$$|a_n| \leq q'|a_{n-1}| \leq \cdots \leq |a_N| {q'}^{(n-N)}.$$

Daraus ergibt sich die folgende Abschätzung: Für alle $n \geq N$ gilt

$$\begin{aligned} \sum_{k=N}^{n} |a_k| \leq |a_N| \sum_{k=N}^{n} {q'}^{(k-N)} &= |a_N| \sum_{k=0}^{n-N} {q'}^k \\ < |a_N| \sum_{k=0}^{\infty} {q'}^k &= |a_N| \frac{1}{1-q'}. \end{aligned}$$

Daraus folgt wiederum:

$$\sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{N} |a_k| + |a_N| \frac{1}{1-q'}$$

für alle $n \geq N$. Die Anwendung des Monotoniekriteriums (Satz 5.4) ergibt die Konvergenz der Reihe $\sum_{n=1}^{\infty} |a_n|$.

2) Es sei $q > 1$. Wir wählen ein $q' \in \mathbb{R}$ mit $q > q' > 1$. Dann folgt wie oben, daß zu q' ein $N \in \mathbb{N}$ existiert so daß gilt:

$$\left| \frac{a_{n+1}}{a_n} \right| \geq q' \quad \text{für alle} \quad n \geq N.$$

Für $n \geq N$ erhalten wir daraus

$$|a_n| \geq q'^{(n-N)}|a_N|.$$

Da $q' > 1$, ist $q'^k \xrightarrow[k\to\infty]{} \infty$.

Daraus folgt, daß $|a_n| \xrightarrow[n\to\infty]{} \infty$, d.h. (a_n) ist insbesondere keine Nullfolge. Aus Lemma 5.3 folgt damit, daß die Reihe $\sum_{n=1}^{\infty} a_n$ divergent sein muß.

$\square$

Bemerkung: Wenn $q = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1$ ist, so kann keine Aussage über die Konvergenz oder Divergenz von $\sum_{n=1}^{\infty} a_n$ gemacht werden.

Beispiel: $\sum_{n=1}^{\infty} \frac{1}{n^s}$, $s \in \mathbb{Q}$.
Dann ist $\left|\frac{a_{n+1}}{a_n}\right| = \left(\frac{n}{n+1}\right)^s$.

Weiter gilt:

$$\frac{n}{n+1} = \frac{1}{1+\frac{1}{n}} \xrightarrow[n\to\infty]{} 1.$$

Hieraus folgt:

$$\left(\frac{n}{n+1}\right)^s \xrightarrow[n\to\infty]{} 1.$$

Andererseits gilt aber

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \begin{cases} \text{konvergent, falls} & s > 1, \\ \text{divergent, falls} & s \leq 1. \end{cases}$$

Wir betrachten jetzt ein Beispiel für die Anwendung des Quotientenkriteriums.

Beispiel: Für $s \in \mathbb{C}$ ist $\binom{s}{n}$ definiert durch

$$\binom{s}{n} = \begin{cases} \frac{s(s-1)\cdots(s-n+1)}{n!}, & n > 0; \\ 1, & n = 0; \\ 0, & n < 0. \end{cases}$$

Für $z \in \mathbb{C}$ sei

$$\begin{aligned} B_s(z) &= \sum_{n=0}^{\infty} \binom{s}{n} z^n \\ &= 1 + sz + \frac{s(s-1)}{2} z^2 + \cdots \end{aligned}$$

Behauptung: Für $s \neq \mathbb{N} \cup \{0\}$ gilt:
$B_s(z)$ konvergiert für $|z| < 1$ absolut und divergiert für $|z| > 1$.

Beweis: 1) Es sei $s \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
Dann ist

$$\binom{s}{n} = 0 \quad \text{für} \quad n > s.$$

$$\Rightarrow B_s(z) = \sum_{n=0}^{s} \binom{s}{n} z^n = (1+z)^s,$$

d.h. $B_s(z)$ ist ein Polynom.

2) Es sei $s \neq \mathbb{N}_0$. Dann ist $\binom{s}{n} \neq 0$ für alle $n \in \mathbb{N}$. Nach Definition ist $a_n = \binom{s}{n} z^n$ das n-te Glied der Reihe. Sei $z \neq 0$. Dann gilt:

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \left| \binom{s}{n+1} z^{n+1} \middle/ \binom{s}{n} z^n \right| \\ &= |z| \frac{|s(s-1)\cdots(s-n)|}{(n+1)!} \quad \frac{n!}{|s(s-1)\cdots(s-n+1)|} \\ &= |z| \frac{|s-n|}{n+1} = |z| \frac{\left|1-\frac{s}{n}\right|}{1+\frac{1}{n}} \xrightarrow[n\to\infty]{} |z|. \end{aligned}$$

Die Anwendung des Quotientenkriteriums ergibt

$$B_s(z) \begin{cases} \text{absolut konvergent für} & |z| < 1; \\ \text{divergent für} & |z| > 1. \end{cases}$$

Die Reihe $B_s(z)$ heißt die Binomialreihe.

Satz 5.12. *(Wurzelkriterium) Es sei* (a_n) *eine Folge in* $\mathbb{C}$ *und*

$$L = \limsup \sqrt[n]{|a_n|}.$$

Dann gilt:

1) $L < 1 \Rightarrow \sum_{n=1}^{\infty} a_n$ *ist absolut konvergent;*
2) $L > 1 \Rightarrow \sum_{n=1}^{\infty} a_n$ *ist divergent.*

Beweis: 1) Sei $L < 1$. Wir wählen ein $q \in \mathbb{R}$ mit $L < q < 1$. Sei $\epsilon = q - L$. Nach Definition ist L der größte Häufungspunkt der Folge $\sqrt[n]{|a_n|}$. Daraus folgt:

$$\exists N \in \mathbb{N} \colon \sqrt[n]{|a_n|} \leq L + \epsilon = q \quad \text{für alle} \quad n \geq N.$$

$$\Rightarrow \forall n \geq N \colon |a_n| \leq q^n$$
$$\Rightarrow \sum_{k=N}^{\infty} |a_k| \leq \sum_{k=N}^{\infty} q^k \leq \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

Damit haben wir gezeigt, daß die Reihe $\sum_{k=N}^{\infty} |a_k|$ die konvergente Majorante $\sum_{k=N}^{\infty} q^k$ hat. Aus Satz 5.7 folgt, daß die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert, d.h. $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent.

2) Es sei $L > 1$. Da $L = \limsup \sqrt[n]{|a_n|}$, existieren unendlich viele $n \in \mathbb{N}$ mit

$$\sqrt[n]{|a_n|} > 1.$$

$$\Rightarrow |a_n| > 1 \quad \text{für unendlich viele} \quad n \in \mathbb{N}$$
$$\Rightarrow (a_n) \quad \text{ist keine Nullfolge.}$$

Aus Lemma 5.3 folgt, daß $\sum_{n=1}^{\infty} a_n$ divergent sein muß.

□

Bemerkung: Es existiere

$$q = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$

Behauptung: Es gilt

$$\limsup \sqrt[n]{|a_n|} \leq q.$$

Beweis: Es sei $\epsilon > 0$. Dann existiert zu ϵ ein $N \in \mathbb{N}$ mit

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q + \epsilon \colon = q' \quad \text{für alle} \quad n \geq N.$$

$$\Rightarrow |a_n| \leq |a_{n-1}|q' \leq \cdots \leq |a_N|q'^{(n-N)}.$$

Da die n-te Wurzel monoton wachsend ist, folgt daraus

$$\sqrt[n]{|a_n|} \leq \sqrt[n]{|a_N|} \quad q'^{1-\frac{N}{n}} \quad \text{für} \quad n \geq N.$$

Weiter gilt

$$\sqrt[n]{|a_N|} \xrightarrow[n\to\infty]{} 1 \quad \text{und} \quad q'^{(1-\frac{N}{n})} \xrightarrow[n\to\infty]{} q'.$$

$$\Rightarrow \exists N' \in \mathbb{N} \colon \sqrt[n]{|a_n|} < q' + \epsilon = q + 2\epsilon, \quad \text{für alle} \quad n \geq N.$$

Da $\epsilon > 0$ beliebig gewählt war, folgt daraus

$$\limsup \sqrt[n]{|a_n|} \leq q.$$

□

Wenn man also mit Hilfe des Quotientenkriteriums die Konvergenz einer Reihe $\sum_{n=1}^{\infty} a_n$ beweisen kann, so kann man dies auch mit Hilfe des Wurzelkriteriums.

Bemerkung: Die Umkehrung gilt nicht.

Beispiel: Es sei

$$a_n = \begin{cases} 2^{-n}, & \text{falls } n \text{ gerade;} \\ 3^{-n}, & \text{flals } n \text{ ungerade.} \end{cases}$$

Dann ist

$$\sqrt[n]{|a_n|} = \begin{cases} \frac{1}{2}, & n \quad \text{gerade;} \\ \frac{1}{3}, & n \quad \text{ungerade.} \end{cases}$$

$$\Rightarrow \limsup \sqrt[n]{|a_n|} = \frac{1}{2}.$$

Aus Satz 5.11 folgt damit, daß die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergent ist. Andererseits ist

$$\begin{aligned} \frac{a_{2k}}{a_{2k-1}} &= \frac{3^{2k-1}}{2^{2k}} = \frac{1}{3}\left(\frac{3}{2}\right)^{2k} \xrightarrow[k\to\infty]{} \infty \\ \frac{a_{2k+1}}{a_{2k}} &= \frac{2^{2k}}{3^{2k+1}} = \frac{1}{3}\left(\frac{2}{3}\right)^{2k} \xrightarrow[k\to\infty]{} 0. \end{aligned}$$

Die Folge $\left|\frac{a_{n+1}}{a_n}\right|$ ist also nicht konvergent, und daher ist das Quotientenkriterium nicht anwendbar.

5.3. Umordnung von Reihen

Seien $a_1, \dots, a_n \in \mathbb{C}$ und

$$\sigma\colon \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$$

sei eine bijektive Abbildung, d.h. σ ist eine Permutation der Zahlen $1, \dots, n$. Dann gilt bekanntlich:

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n} a_{\sigma(k)}.$$

Diese Regel ist für Reihen im allgemeinen falsch.

Definition 5.13. *Es sei $\sum_{n=1}^{\infty} a_n$ eine Reihe komplexer Zahlen und $\sigma\colon \mathbb{N} \to \mathbb{N}$ sei eine bijektive Abbildung. Für $n \in \mathbb{N}$ sei $b_n = a_{\sigma(n)}$. Dann heißt die Reihe $\sum_{n=1}^{\infty} b_n$ eine Umordnung der Reihe $\sum_{n=1}^{\infty} a_n$.*

Im allgemeinen gilt:

$$\sum_{n=1}^{\infty} a_n \neq \sum_{n=1}^{\infty} a_{\sigma(n)}.$$

Beispiel: Wir betrachten die alternierende Reihe

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

Für die Folge der Partialsummen gilt $s_1 > s_3 > s_5 > \cdots$ und $s_2 < s_4 < s_6 < \cdots$. Weiter ist $s_{2k} < s < s_{2k+1}$ für alle $k \in \mathbb{N}$ (siehe Beweis des Satzes 5.5). Insbesondere gilt also

$$\frac{1}{2} = s_2 < s < s_3 = \frac{5}{6}.$$

Wir betrachten jetzt die Reihe

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

Dies ist eine Umordnung der alternierenden Reihe. Die Glieder dieser Reihe kann man wie folgt beschreiben

$$b_n = \begin{cases} \frac{1}{4k-3}, & n = 3k-2; \\ \frac{1}{4k-1}, & n = 3k-1; \\ -\frac{1}{2k}, & n = 3k. \end{cases}$$

a) Die Reihe $\sum_{n=1}^{\infty} b_n$ ist konvergent.

Dazu betrachten wir die Patialsummen

$$\begin{aligned} {s'}_{3n} &= \sum_{k=1}^{n} \left(\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} \right) \\ &= \sum_{k=1}^{n} \frac{8k-4}{(4k-3)(4k-1)2k} \end{aligned}$$

Da alle Glieder der Reihe positiv sind, ist

$$s'_3 < s'_6 < s'_9 < \cdots$$

Weiter ist für $k \geq 2$:

$$\frac{8k-4}{(4k-3)(4k-1)2k} < \frac{8}{32} \cdot \frac{k}{(k-1)^3} < \frac{1}{2} \frac{1}{(k-1)^2}.$$

Da $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$ ist, existiert eine Zahl $a > 0$ so daß

$$s'_{3n} < a \qquad \text{für alle} \quad n \in \mathbb{N}.$$

Damit ist (s'_{3n}) eine monotone, beschränkte Folge und hat daher einen Grenzwert

$$s' = \lim_{n \to \infty} s'_{3n}.$$

(siehe Satz über monotone, beschränkte Folgen).

Nach Definition der b_n gilt

$$s'_{3n+1} = s'_{3n} + a_{3n+1} = s'_{3n} + \frac{1}{4n-3}.$$

$$\Rightarrow \lim_{n \to \infty} s'_{3n+1} = \lim_{n \to \infty} s'_{3n} = s'.$$

Entsprechend erhalten wir

$$\begin{aligned} s'_{3n+2} &= s'_{3n} + b_{3n+1} + b_{3n+2} \\ &= s'_{3n} + \frac{1}{4n-3} + \frac{1}{4n-1}. \end{aligned}$$

Da $\frac{1}{4n-3} \xrightarrow[n \to \infty]{} 0$ und $\frac{1}{4n-1} \xrightarrow[n \to \infty]{} 0$ gilt, folgt

$$\lim_{n \to \infty} s'_{3n+2} = \lim_{n \to \infty} s'_{3n} = s'.$$

Es gilt also

$$\lim_{n \to \infty} s'_n = s'$$

d.h. die Reihe $\sum_{n=1}^{\infty} b_n$ ist konvergent. Weiter ist

$$\sum_{n=1}^{\infty} b_n > s'_3 = 1 + \frac{1}{3} - \frac{1}{2} = \frac{5}{6}.$$

Da andererseits $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} < \frac{5}{6}$, folgt

$$\sum_{n=1}^{\infty} b_n \neq \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

Satz 5.14. *(Umordnungssatz) Es sei $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe mit $a_n \in \mathbb{C}$, und $\sum_{n=1}^{\infty} b_n$ sei eine Umordnung von $\sum_{n=1}^{\infty} a_n$. Dann ist $\sum_{n=1}^{\infty} b_n$ absolut konvergent und es gilt:*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$$

Beweis: Es sei $s_n = \sum_{k=1}^{n} a_k$ und $t_n = \sum_{k=1}^{n} b_k$. Sei $\epsilon > 0$ gegeben. Da nach Voraussetzung $\sum_{k=1}^{\infty} |a_k|$ konvergent ist, folgt aus dem Cauchy-Kriterium (Satz 5.6):

$$\exists N \in \mathbb{N} \colon \sum_{k=N+1}^{N+p} |a_k| < \epsilon \quad \text{für alle} \quad p \in \mathbb{N}. \tag{5.1}$$

Es sei $\sigma \colon \mathbb{N} \to \mathbb{N}$ die Bijektion so daß

$$b_n = a_{\sigma(n)}, \quad n \in \mathbb{N}.$$

Da $\sigma \colon \mathbb{N} \to \mathbb{N}$ bijektiv ist, existiert $M \geq N$ so, daß

$$\{1, \ldots, N\} \subseteq \{\sigma(1), \ldots, \sigma(M)\}.$$

Für $n > M$ heben sich in der Differenz $s_n - t_n$ alle Glieder a_i mit einem Index $i \leq N$ heraus, weil sie sowohl in s_n als auch in t_n vorkommen. Für $k \geq N + 1$ sei

$$\epsilon_k = \begin{cases} 1, & a_k \text{ kommt in } s_n \text{ vor, aber nicht in } t_n; \\ 0, & a_k \text{ kommt in } s_n \text{ und in } t_n \text{ vor;} \\ -1, & a_k \text{ kommt in } t_n \text{ vor und nicht in } s_n. \end{cases}$$

Insbesondere ist $\epsilon_k = 0$ für fast alle $k \in \mathbb{N}$. Dann gilt

$$s_n - t_n = \sum_{k=N+1}^{\infty} \epsilon_k a_k.$$

Aus (5.1) folgt

$$|s_n - t_n| \leq \sum_{k=N+1}^{\infty} |a_k| < \epsilon.$$

Damit gilt

$$\lim_{n\to\infty} (s_n - t_n) = 0. \tag{5.2}$$

Nach Voraussetzung ist die Folge (s_n) konvergent mit

$$\lim_{n\to\infty} s_n = s = \sum_{k=1}^{\infty} a_k.$$

Aus (5.2) folgt dann

$$\lim_{n\to\infty} t_n = s.$$

Damit gilt

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n.$$

Die absolute Konvergenz von $\sum_{n=1}^{\infty} b_n$ zeigt man analog.

□

Wie das obige Gegenbeispiel zeigt, ist die absolute Konvergenz eine notwendige Voraussetzung für die Gültigkeit des Umordnungssatzes. Für nicht absolut konvergente Reihen gilt folgender bemerkenswerter Satz von B. Riemann.

Satz 5.15. *(B. Riemann) Es sei* $\sum_{n=1}^{\infty} a_n$ *eine konvergente Reihe reeller Zahlen, die* **nicht** *absolut konvergent ist. Dann existiert für jedes* $a \in \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ *eine Bijektion* $\sigma\colon \mathbb{N} \to \mathbb{N}$ *so daß*

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = a.$$

Beweisskizze: Es seien $b_1, b_2, b_3, \ldots$ die positiven Glieder und $c_1, c_2, c_3, \ldots$ die nichtpositiven Glieder der Reihe $\sum_{n=1}^{\infty} a_n$. Dann gilt:

Lemma. $\sum_{j=1}^{\infty} b_j = \infty, \quad \sum_{j=1}^{\infty}(-c_j) = \infty$

Beweis: Es sei

$$a_n^+ = \max\{a_n, 0\} \quad \text{und} \quad a_n^- = \max\{-a_n, 0\}$$

Dann gilt offensichtlich

$$\sum_{n=1}^{\infty} a_n^+ = \sum_{j=1}^{\infty} b_j \quad \text{und} \quad \sum_{n=1}^{\infty} a_n^- = -\sum_{j=1}^{\infty} c_j.$$

Weiter gilt:

$$\forall n \in \mathbb{N}\colon 0 \leq a_n^+, a_n^- \leq |a_n| \quad \text{und} \quad a_n^+ + a_n^- = |a_n|.$$

Wenn wir annehmen, daß sowohl $\sum_{n=1}^{\infty} a_n^+$ als auch $\sum_{n=1}^{\infty} a_n^-$ konvergiert, so folgt aus den Rechenregeln für Reihen:

$$\sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} a_n^+ + \sum_{n=1}^{\infty} a_n^-,$$

d.h. $\sum_{n=1}^{\infty} |a_n|$ wäre ebenfalls konvergent. Dies widerspricht der Annahme, daß $\sum_{n=1}^{\infty} a_n$ nicht absolut konvergiert. Wir nehmen jetzt an, daß eine der Reihen, z.B. $\sum_{n=1}^{\infty} a_n^+$ konvergiert. Aus den Rechenregeln folgt dann

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

Da auf Grund unserer Annahmen beide Reihen auf der rechten Seite konvergieren, fogt daraus, daß $\sum_{n=1}^{\infty} a_n^-$ ebenfalls konvergiert. Wir haben damit gezeigt:

$$\sum_{n=1}^{\infty} a_n^{\pm} \quad \text{konvergiert} \quad \Rightarrow \sum_{n=1}^{\infty} a_n^{\mp} \quad \text{konvergiert},$$

d.h., wenn eine der Reihen $\sum a_n^{\pm}$ konvergiert, so konvergiert auch die andere. Aus dem ersten Teil des Beweises folgt damit, daß $\sum_{n=1}^{\infty} a_n^+ = \infty$ und $\sum_{n=1}^{\infty} a_n^- = \infty$ gelten muß.

□

Es sei $a \in \mathbb{R}$. Wir definieren dann

$$k_1 = \min\{k \in \mathbb{N} \mid \sum_{j=1}^{k} b_j > a\}$$
$$l_1 = \min\{l \in \mathbb{N} \mid \sum_{j=1}^{k_1} b_j + \sum_{j=1}^{l} c_j < a\}.$$

Aus dem Lemma folgt, daß beide Mengen nicht leer sind, und daher existieren die Minima. Dann sei

$$k_2 = \min\{k \in \mathbb{N} \mid \sum_{j=1}^{l_1} c_j + \sum_{j=1}^{k} b_j > a\}$$
$$l_2 = \min\{l \in \mathbb{N} \mid \sum_{j=1}^{k_2} b_j + \sum_{j=1}^{l} c_j < a\}$$

Aus dem Lemma folgt die Existenz dieser Minima. Dieses Verfahren können wir beliebig fortsetzen und definieren rekursiv Folgen natürlicher Zahlen

$$k_1 < k_2 < k_3 < \cdots, \quad l_1 < l_2 < l_3 < \cdots$$

Daraus erhalten wir eine Umordnung der Reihe wie folgt:

$$b_1, \ldots, b_{k_1}, c_1, \ldots, c_{l_1}, b_{k_1+1}, \ldots, b_{k_2}, c_{l_1+1}, \ldots, c_{l_2}, \ldots,$$

d.h. eine Bijektion $\sigma\colon \mathbb{N} \to \mathbb{N}$. Nach Definition von (k_n) und (l_n) gilt für die Partialsummen z.B.:

$$\sum_{j=1}^{k_{n+1}} b_j + \sum_{j=1}^{l_n} c_j > a > \sum_{j=1}^{k_{n+1}-1} b_j + \sum_{j=1}^{l_n} c_j.$$

Daraus folgt:

$$\left| a - \left(\sum_{j=1}^{k_{n+1}} b_j + \sum_{j=1}^{l_n} c_j \right) \right| < b_{k_{n+1}} \xrightarrow[n\to\infty]{} 0,$$

wobei die Behauptung $b_{k_{n+1}} \xrightarrow[n\to\infty]{} 0$ aus Lemma 5.3. folgt, da $\sum_{n=1}^{\infty} a_n$ konvergent ist.

$\square$

5.4. Doppelreihen und Produkte von Reihen

Es sei $f\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{C}$ eine Abbildung, wobei $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ ist. Wir setzen

$$a_{ij} = f(i,j), \quad i,j \in \mathbb{N}_0.$$

Die Zahlen a_{ij} können wir in einer unendlichen Matrix (a_{ij}) anordnen:

$$(a_{ij}) = \begin{pmatrix} a_{00} & a_{01} & a_{02} & \cdot & \cdot & \cdot \\ a_{10} & a_{11} & a_{12} & \cdot & \cdot & \cdot \\ a_{20} & a_{21} & a_{22} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & \end{pmatrix}$$

Dieser Matrix kann man auf verschiedene Weise Reihen zuordnen.

1. Wir summieren über die Zeilen:

$$(a_{ij}) \longmapsto \left(\sum_{j=0}^{\infty} a_{ij} \right)_{i\in\mathbb{N}_0} \longmapsto \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right)$$

2. Wir summieren über die Spalten

$$(a_{ij}) \longmapsto \left(\sum_{i=0}^{\infty} a_{ij}\right)_{j\in\mathbb{N}_0} \longmapsto \sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} a_{ij}\right)$$

Die Reihen

$$\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right) \quad \text{und} \quad \sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} a_{ij}\right)$$

nennt man Doppelreihen.

Beispiel: Es sei $a_{ij} = \frac{1}{(1+i)^2(1+j)^2}, \quad i,j \in \mathbb{N}_0$

Dann erhält man

$$\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} \frac{1}{(1+i)^2(1+j)^2}\right) = \left(\sum_{i=1}^{\infty} \frac{1}{i^2}\right)\left(\sum_{j=1}^{\infty} \frac{1}{j^2}\right).$$

Definition 5.16. *Die Doppelreihe* $\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right)$ *heißt konvergent, wenn gilt:*

1) $\forall i \in \mathbb{N}_0\colon \sum_{j=0}^{\infty} a_{ij}$ *ist konvergent*
2) *Es sei* $A_i = \sum_{j=0}^{\infty} a_{ij}$, $i \in \mathbb{N}_0$. *Dann ist die Reihe*

$$\sum_{i=0}^{\infty} A_i$$

konvergent.

Bezeichnung:

$$A = \sum_{i=0}^{\infty} A_i = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right).$$

Genauso definiert man die Konvergenz für die Reihe $\sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} a_{ij}\right)$.

Es gibt aber noch weitere Möglichkeiten, der Matrix (a_{ij}) eine Reihe zuzuordnen. Wie wir wissen, ist $\mathbb{N}_0 \times \mathbb{N}_0$ abzählbar, d.h., es existiert eine Bijektion

$$\sigma\colon \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0.$$

Für jede solche Bijektion σ können wir die Reihe

$$\sum_{k=0}^{\infty} a_{\sigma(k)}$$

bilden. Es gibt dabei eine ausgezeichnete Bijektion, die Standardabzählung σ_0:
ZEICHNUNG

Die Doppelreihen $\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right)$, $\sum_{j=0}^{\infty}(\sum_{i=0}^{\infty} a_{ij})$ und die Reihen $\sum_{k=0}^{\infty} a_{\sigma(k)}$ werden im allgemeinen verschiedene Werte haben. Es gilt jedoch der folgende Satz:

Satz 5.17. *(Doppelreihensatz) Es sei* $\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right) < \infty$.
Dann sind die Doppelreihen und die Reihen $\sum_{k=0}^{\infty} a_{\sigma(k)}$ *für alle Bijektionen* σ *konvergent, und es gilt*

$$\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right) = \sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} a_{ij}\right) = \sum_{k=0}^{\infty} a_{\sigma(k)} = \sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_{ij}\right).$$

Beweis:
1) Es sei $A = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right)$. Nach Voraussetzung ist $A < \infty$.

Da alle Glieder der Reihe positiv sind, gilt:

$$\forall n, m \in \mathbb{N}_0\colon \sum_{i=0}^{n}\left(\sum_{j=0}^{m} |a_{ij}|\right) \le A.$$

Da wir bei endlichen Summen die Reihenfolge vertauschen können, ergibt sich

$$\forall n, m \in \mathbb{N}_0\colon \sum_{j=0}^{m}\left(\sum_{i=0}^{n} |a_{ij}|\right) \le A.$$

Für festes $m \in \mathbb{N}$ ist

$$\left(\sum_{j=0}^{m}\left(\sum_{i=0}^{n}|a_{ij}|\right)\right)_{n\in\mathbb{N}_0}$$

eine monoton wachsende, beschränkte Folge, und daher hat diese Folge einen Grenzwert $\sum_{j=0}^{m}(\sum_{i=0}^{\infty}|a_{ij}|) \leq A$. Ebenso folgt

$$\sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty}|a_{ij}|\right) \leq A.$$

2) Die Anwendung des Cauchy-Kriteriums (Satz 5.6) ergibt, daß für alle $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert mit
i) $\forall n \geq \mathbb{N}$: $\sum_{i=n}^{\infty}\left(\sum_{j=0}^{\infty}|a_{ij}|\right) < \epsilon$
ii) $\forall n \geq N$: $\sum_{j=n}^{\infty}(\sum_{i=0}^{\infty}|a_{ij}|) < \epsilon$

Aus ii) folgt:
$\forall m \geq n, \quad \forall l \in \mathbb{N}_0$: $\sum_{j=n}^{m}\left(\sum_{i=0}^{l}|a_{ij}|\right) < \epsilon$
Wie in 1) folgt daraus
iii) $\forall n \geq N$: $\sum_{i=0}^{\infty}\left(\sum_{j=n}^{\infty}|a_{ij}|\right) < \epsilon$.

3) Nach Voraussetzung ist $\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty}|a_{ij}|\right) < \infty$.
Daraus folgt, daß für jedes $i \in \mathbb{N}_0$: $\sum_{j=0}^{\infty}|a_{ij}| < \infty$. Aus Satz 5.10 ergibt sich, daß die Reihe $\sum_{j=0}^{\infty} a_{ij}$ konvergiert und es gilt

$$\left|\sum_{j=0}^{\infty} a_{ij}\right| \leq \sum_{j=0}^{\infty}|a_{ij}|.$$

Daher konvergiert die Doppelreihe $\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right)$ und es gilt:

$$\left|\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right)\right| \leq \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty}|a_{ij}|\right).$$

Es sei

$$A_1 = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right).$$

Die äußere Summe spalten wir auf in $\sum_{i=0}^{N} + \sum_{i=N+1}^{\infty}$. Dann erhalten wir

$$\begin{aligned} A_1 &= \sum_{i=0}^{N}\left(\sum_{j=0}^{\infty} a_{ij}\right) + \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right) \\ &= \sum_{i=0}^{N}\left(\sum_{j=0}^{N} a_{ij}\right) + \sum_{i=0}^{N}\left(\sum_{j=N+1}^{\infty} a_{ij}\right) + \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} a_{ij}\right). \end{aligned}$$

Aus i) und iii) erhalten wir damit

$$\begin{aligned} \left|A_1 - \sum_{i,j=0}^{N} a_{ij}\right| &\leq \sum_{i=0}^{N}\left(\sum_{j=N+1}^{\infty} |a_{ij}|\right) + \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right) \\ &\leq \sum_{i=0}^{\infty}\left(\sum_{j=N+1}^{\infty} |a_{ij}|\right) + \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right) < 2\epsilon. \end{aligned}$$

Es sei $\sigma\colon \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$ eine Bijektion. Dann existiert $M \geq N$ so daß:

$$\{(i,j) \mid i,j \in \mathbb{N}_0, \quad 0 \leq i,j \leq N\} \subseteq \{\sigma(0), \ldots, \sigma(M)\}. \tag{5.3}$$

Hieraus folgt auf Grund obiger Ungleichung:

$$\begin{aligned} \left|A_1 - \sum_{k=0}^{M} a_{\sigma(k)}\right| &\leq \left|A_1 - \sum_{i,j=0}^{N} a_{ij}\right| \\ &\quad + \left|\sum_{i,j=0}^{N} a_{ij} - \sum_{k=0}^{M} a_{\sigma(k)}\right| < 2\epsilon + \left|\sum_{i,j=0}^{N} a_{ij} - \sum_{k=0}^{M} a_{\sigma(k)}\right|. \end{aligned}$$

Wegen (5.3) ist

$$\begin{aligned} \left|\sum_{i,j=0}^{N} a_{ij} - \sum_{k=0}^{M} a_{\sigma(k)}\right| &\leq \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right) + \sum_{i=0}^{N}\left(\sum_{j=N+1}^{\infty} |a_{ij}|\right) \\ &\leq \sum_{i=N+1}^{\infty}\left(\sum_{j=0}^{\infty} |a_{ij}|\right) + \sum_{i=0}^{\infty}\left(\sum_{j=N+1}^{\infty} |a_{ij}|\right), \end{aligned}$$

und aus i) und iii) folgt, daß die rechte Seite durch 2ϵ abgeschätzt werden kann. Damit haben wir gezeigt:

$$\forall \epsilon > 0 \, \exists M \in \mathbb{N}\colon \qquad \left| A_1 - \sum_{k=0}^{M} a_{\sigma(k)} \right| < 4\epsilon,$$

d.h., die Reihe $\sum_{k=0}^{\infty} a_{\sigma(k)}$ ist konvergent und es gilt

$$\sum_{k=0}^{\infty} a_{\sigma(k)} = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right).$$

Dies gilt für jede Bijektion σ, insbesondere also die Standardbijektion σ_0. Daraus folgt

$$\sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_{ij} \right) = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right).$$

Für $i, j \in \mathbb{N}_0$ sei

$$b_{ij} = a_{ji}.$$

Dann folgt aus der soeben bewiesenen Gleichung

$$\begin{aligned} \sum_{j=0}^{\infty} \left(\sum_{i=0}^{\infty} a_{ij} \right) &= \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} b_{ij} \right) = \sum_{k=0}^{\infty} \left(\sum_{i+j=k} b_{ij} \right) \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_{ij} \right) = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right). \end{aligned}$$

$\square$

Beim Beweis dieses Satzes haben wir gezeigt, daß für jede Bijektion $\sigma\colon \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$ gilt

$$\sum_{k=0}^{\infty} a_{\sigma(k)} = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right);$$

falls die Doppelreihe absolut konvergent ist. In diesem Falle können wir für die Doppelreihe einfach schreiben

$$\sum_{i,j=0}^{\infty} a_{ij} \quad ,$$

d.h., wie können die Glieder der Reihe in beliebiger Reihenfolge summieren.

Beispiel: (Jakob Bernoulli)

$$1 + 2q + 3q^2 + \ldots = \sum_{n=0}^{\infty}(n+1)q^n = \frac{1}{(1-q)^2}$$

für $|q| < 1$.

Es gilt

$$(n+1)q^n = \sum_{\substack{i,j=0 \\ i+j=n}}^{n} q^{i+j}.$$

$$\Rightarrow \sum_{n=0}^{\infty}(n+1)q^n = \sum_{n=0}^{\infty} \sum_{i+j=n} (q^i q^i) = \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} q^i q^j \right)$$
$$\left(\sum_{i=0}^{\infty} q^i \right) \left(\sum_{j=0}^{\infty} q^i \right) = \frac{1}{(1-q)^2}.$$

Produkte von Reihen

Für endliche Summen gilt

$$(a_0 + a_1 + \cdots + a_m)(b_0 + b_1 + b \ldots + b_n) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j.$$

Dies kann wie folgt auf unendliche Reihen übertragen werden.

Satz 5.18. *Es seien* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{m=0}^{\infty} b_m$ *absolut konvergente Reihen komplexer Zahlen. Dann ist die Reihe* $\sum_{i,j=0}^{\infty} a_i b_j$ *absolut konvergent und es gilt*

$$\left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right) = \sum_{i,j=0}^{\infty} a_i b_j.$$

Beweis: Es sei

$$A = \sum_{i=0}^{\infty} |a_i| \quad \text{und} \quad B = \sum_{j=0}^{\infty} |b_j|.$$

Nach Vorraussetzung ist $A < \infty$ und $B < \infty$. Weiter gilt

$$\sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} |a_i b_j|\right) = \sum_{i=0}^{\infty} |a_i| \left(\sum_{j=0}^{\infty} |b_j|\right) = \sum_{i=0}^{\infty} |a_i| \cdot B = A \cdot B < \infty.$$

Aus Satz 5.17 folgt daher, daß gilt:

$$\left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right) = \sum_{i=0}^{\infty} a_i \sum_{j=0}^{\infty} b_j = \sum_{i=0}^{\infty}\left(\sum_{j=0}^{\infty} a_i b_j\right) = \sum_{i,j=0}^{\infty} a_i b_j.$$

□

Satz 5.19. *(Cauchy-Produkt) Es seien* $\sum_{i=0}^{\infty} a_i$ *und* $\sum_{j=0}^{\infty} b_j$ *absolut konvergent. Dann gilt*

$$\left(\sum_{i=0}^{\infty} a_i\right) \cdot \left(\sum_{j=0}^{\infty}\right) = \sum_{k=0}^{\infty}\left(\sum_{i+j=k} a_i b_j\right).$$

Beweis: Der Beweis folgt unmittelbar aus den Sätzen 5.17 und 5.18.

□

Beispiele: Es sei $s \in \mathbb{C}$ und $z \in \mathbb{C}$. Wir betrachten die **Binomialreihe**

$$B_s(z) = \sum_{n=0}^{\infty} \binom{s}{n} z^n,$$

deren Konvergenzverhalten wir bereits im Abschnitt 5.2 diskutiert haben. Es gilt

1) Für $s = k \in \mathbb{N}_0$ ist

$$B_k(z) = (1+z)^k,$$

2) Für $s \notin \mathbb{N}_0$ gilt:

$$B_s(z) \begin{cases} \text{absolut konvergent, falls} & |z| < 1; \\ \text{divergent, falls} & |z| > 1. \end{cases}$$

Satz 5.20. *(Additionstheorem) Seien $t, s \in \mathbb{C}$, $z \in \mathbb{C}$ und $|z| < 1$. Dann gilt:*

$$B_s(z) \cdot B_t(z) = B_{s+t}(z)$$

Beweis: Die Anwendung von Satz 5.19 ergibt unmittelbar

$$\left(\sum_{i=0}^{\infty} \binom{s}{i} z^i\right) \cdot \left(\sum_{j=0}^{\infty} \binom{t}{j} z^j\right) = \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \binom{s}{i} \binom{t}{k-i}\right) z^k. \tag{5.4}$$

Lemma. *Für alle $s, t \in \mathbb{C}$ und $k \in \mathbb{N}_0$ gilt:*

$$\sum_{i=0}^{k} \binom{s}{i} \binom{t}{k-i} = \binom{s+t}{k}.$$

Beweis: 1) Es sei $s = n, t = m, \quad n, m \in \mathbb{N}_0$ Dann ist einerseits

$$(1+x)^{n+m} = \sum_{k=0}^{n+m} \binom{n+m}{k} x^k.$$

Andererseits gilt

$$\begin{aligned} (1+x)^{n+m} = (1+x)^n (1+x)^m &= \left(\sum_{i=0}^{n} \binom{n}{i} x^i\right) \left(\sum_{j=0}^{m} \binom{m}{j} x^j\right) \\ &= \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i}\right) x^k \end{aligned}$$

Die behauptete Gleichung im Lemma folgt durch Koeffizientenvergleich.

2) Es sei $t \in \mathbb{N}$. Für $s \in \mathbb{C}$ definieren wir

$$P_t(s) = \sum_{i=0}^{k} \binom{s}{i} \binom{t}{k-i} \quad \text{und}$$

$$Q_t(s) = \binom{s+t}{k}, \quad k \in \mathbb{N}.$$

Dann sind $P_t(s)$ und $Q_t(s)$ Polynome und für alle $m \in \mathbb{N}_0$ gilt auf Grund von 1):

$$P_t(m) = Q_t(m),$$

Aus dem Identitätssatz für Polynome (ein Polynom $\neq 0$ vom Grade n hat genau n Nullstellen) folgt damit $P_t(s) = Q_t(s)$ für alle $s \in \mathbb{C}$.

3) Schließlich fixieren wir $s \in \mathbb{C}$ und definieren die Polynome $P_s(t) = \sum_{i=0}^{k} \binom{i}{s} \binom{t}{k-i}$ und $Q_s(t) = \binom{s+t}{k}$.

Wiederum ist wegen 2):

$$P_s(t) = Q_s(t) \quad \text{für alle} \quad t \in \mathbb{N}_0,$$

und der Identitätssatz für Polynome impliziert

$$P_s(t) = Q_s(t) \quad \text{für alle} \quad t \in \mathbb{C}.$$

□

Wenn wir das Lemma auf die Gleichung (5.4) anwenden, erhalten wir

$$B_s(z) \cdot B_t(z) = B_{s+t}(z).$$

□

Folgerung 5.21 Für alle $s \in \mathbb{Q}$ und $x \in (-1, 1)$ gilt:

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

Beweis: Für $p \in \mathbb{N}$ gilt

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

Sei $s \in \mathbb{Q}$. Wir nehmen zuerst an, daß $s > 0$ ist. Dann ist $s = p/q, \quad p, q \in \mathbb{N}$. Aus Satz 5.20 erhalten wir

$$(B_{p/q}(x))^q = B_{(p/q)\cdot q}(x) = B_p(x).$$

Aus der Reihe für $B_s(x)$ erhält man unmittelbar, daß $B_{p/q}(x) > 0$ ist. Aus der Definition der q-ten Wurzel erhält man damit

$$B_{p/q}(x) = \sqrt[q]{B_p(x)} = \sqrt[q]{(1+x)^p} = (1+x)^{p/q}.$$

Für $s < 0$ benutzt man

$$B_s(z)B_{-s}(z) = B_0(z) = 1,$$

um diesen Fall auf den Fall $s > 0$ zurückzuführen.

□

Bemerkung: Später wird gezeigt, daß

$$B_s(x) = (1+x)^s$$

für alle $s \in \mathbb{R}$ und $x \in (-1, 1)$ gilt.

Anwendung: Als Anwendung der Folgerung 5.21 erhalten wir Reihendarstellungen für die rationalen Potenzen von $(1+x)$. Z.B. gilt:

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 - \cdots$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \cdots$$

Dies kann man zur Berechnung dieser reellen Zahlen mit vorgegebener Genauigkeit benutzen.

5.5. Potenzreihen

Definition 5.22. *Es sei* $(a_n)_{n\in\mathbb{N}_0}$ *eine Folge in* $\mathbb{C}$ *und* $z \in \mathbb{C}$. *Dann heißt die Reihe*

$$P(z) = \sum_{n=0}^{\infty} a_n z^n$$

Potenzreihe in z.

Potenzreihen kann man als Verallgemeinerung von Polynomen ansehen. Wenn $a_n = 0$ ist für alle $n \geq N+1$ und $a_N \neq 0$, so ist

$$P(z) = \sum_{i=0}^{N} a_i z^i$$

ein Polynom vom Grade N.

Satz 5.23. *Es sei* $z_0 \neq 0$ *und* $P(z_0)$ *sei konvergent. Dann ist* $P(z)$ *absolut konvergent für alle* $z \in \mathbb{C}$ *mit* $|z| < |z_0|$.

Beweis: Da nach Voraussetzung die Reihe $\sum_{n=0}^{\infty} a_n z_0^n$ konvergent ist, folgt aus Lemma 5.3:

$$a_n z_0^n \xrightarrow[n\to\infty]{} 0.$$

Daraus folgt

$$\exists\, C > 0 \colon |a_n z_0^n| \leq C \quad \text{für alle} \quad n \in \mathbb{N}.$$

Es sei $q = \left|\frac{z}{z_0}\right|$. Dann gilt

$$|a_n z^n| = |a_n z_0^n| \left|\frac{z}{z_0}\right|^n \leq C q^n.$$

Für $|z| < |z_0|$ ist $q < 1$. In diesem Falle ist $C \sum_{n=0}^{\infty} q^n$ eine konvergente Majorante für $\sum_{n=0}^{\infty} |a_n z^n|$. Aus Satz 5.7 folgt, daß die Reihe $\sum_{n=0}^{\infty} a_n z^n$ absolut konvergent ist.

□

Der Satz 5.23 besagt folgendes: Wenn die Reihe $\sum_{n=0}^{\infty} a_n z^n$ in einem Punkt $z_0 \neq 0$ konvergiert, so konvergiert sie absolut in einem ganzen

Kreis vom Radius $r < |z_0|$. Wir können deshalb den größten Kreis betrachten, in dem die Reihe absolut konvergiert. Das führt zu

Definition 5.24. *Für die Potenzreihe* $P(z) = \sum_{n=0}^{\infty} a_n z^n$ *sei*

$$R(P) = R = \sup\{r \in \mathbb{R}_+ \mid P(r) \quad \textit{konvergiert}\}.$$

Die Zahl R heißt **Konvergenzradius** von $P(z)$.

Satz 5.25. *Für eine Potenzreihe* $P(z) = \sum_{n=0}^{\infty} a_n z^n$ *gilt*

1) $P(z)$ *ist absolut konvergent für alle* $z \in \mathbb{C}$ *mit* $|z| < R$;
2) $P(z)$ *divergiert für* $|z| > R$.

Beweis: 1) Es sei $|z| < R$. Dann existiert $r \in \mathbb{R}$ mit $|z| < r < R$ und $P(r)$ konvergiert (nach Definition von R). Aus Satz 5.23 folgt, daß $P(z)$ absolut konvergiert.

2) Es sei $|z| > R$.

Annahme: Wir nehmen an, daß $P(z)$ konvergiert. Wir wählen $r \in \mathbb{R}$, so daß $|z| > r > R$. Da $P(z)$ konvergiert, folgt aus Satz 5.23, daß $P(r)$ absolut konvergiert. Dies steht im Widerspruch zur Definition von R. Daher muß unsere Annahme falsch sein, d.h. $P(z)$ ist divergent.

$\square$

Den Konvergenzradius kann man aus den Koeffizienten der Potenzreihe wie folgt berechnen.

Satz 5.26. *Es gilt*

1) *Es sei* $L = \limsup \sqrt[n]{|a_n|}$. *Dann ist* $R = \frac{1}{L}$. *Dabei verwenden wir die Konvention:* $0 = \frac{1}{\infty}$ *und* $\infty = \frac{1}{0}$.
2) *Es existiere* q: $\lim_{n\to\infty} \left|\frac{a_n+1}{a_n}\right|, \quad q \in [0, \infty]$. *Dann ist* $R = \frac{1}{q}$.

Beweis:

1) Es gilt

$$\limsup \sqrt[n]{|a_n z^n|} = \limsup |z| \sqrt[n]{|a_n|} = |z| \limsup \sqrt[n]{|a_n|} = |z| \cdot L.$$

Aus dem Wurzelkriterium (Satz 5.12) erhalten wir damit

$$P(z) = \begin{cases} \text{absolut konvergent, falls} & |z| < \frac{1}{L}; \\ \text{divergent, falls} & |z| > \frac{1}{L}. \end{cases}$$

Aus Satz 5.25 folgt: $R = \frac{1}{L}$.

2) Es existiere $q = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$. Dann erhalten wir aus den Rechenregeln für Grenzwerte:

$$\lim_{n\to\infty} \left|\frac{a_{n+1}z^{n+1}}{a_n z^n}\right| = \lim_{n\to\infty} |z| \left|\frac{a_{n+1}}{a_n}\right| = |z| \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |z| \cdot q.$$

Die Anwendung des Quotientenkriteriums (Satz 5.11) ergibt

$$P(z) \begin{cases} \text{absolut konvergent, falls} & |z| < \frac{1}{q}; \\ \text{divergent, falls} & |z| > \frac{1}{q}. \end{cases}$$

Damit erhalten wir aus Satz 5.25:
$R = 1/q$.

□

Beispiel: Die Exponentialreihe.
Für $z \in \mathbb{C}$ sei

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

Behauptung: $R = \infty$

Beweis: Nach Definition sind die Koeffizienten a_n gegeben durch

$$a_n = \frac{1}{n!}, \quad n \in \mathbb{N}_0$$

Dann ist

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{n+1} \xrightarrow[n\to\infty]{} 0.$$

Aus Satz 5.26, 2) folgt unmittelbar, daß $R = \infty$.

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

$$\exp(z) \cdot \exp(w) = \exp(z + w).$$

Beweis: Aus dem Produktsatz 5.19 erhält man

$$\begin{aligned}
\exp(z) \cdot \exp(w) &= \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \frac{z^i}{i!} \frac{w^{k-i}}{(k-i)!} \right) \\
&= \sum_{k=0}^{\infty} \left(\frac{1}{k!} \sum_{i=0} \frac{k!}{i!(k-i)!} z^i w^{k-i} \right) \\
&= \sum_{k=0}^{\infty} \frac{(z+w)^k}{k!} = \exp(z + w).
\end{aligned}$$

□

Es sei $e_1 = \exp(1)$. Durch wiederholte Anwendung des Lemmas erhält man

$$e_1^n = \exp(n) \quad \text{für alle} \quad n \in \mathbb{N}.$$

Durch Wiederholung des Beweises von Satz 5.20 ergibt sich daraus

$$e_1^q = \exp(q) \quad \text{für alle}\, q \in \mathbb{Q}.$$

Die Zahl

$$e_1 = \sum_{k=0}^{\infty} \frac{1}{k!}$$

kann man mit der Zahl

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n} \right)^n$$

identifizieren, d.h. es gilt

Lemma. $e = e_1$

Beweis: Es sei $\epsilon > 0$. Wir wählen $K > 0$ so daß

$$\sum_{k=K+1}^{\infty} \frac{1}{k!} < \frac{\epsilon}{3}$$

Dann folgt

$$\begin{aligned}\left| \left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^{\infty} \frac{1}{k!} \right| &\leq \sum_{k=0}^{K} \left| \binom{n}{k} \frac{1}{n^k} - \frac{1}{k!} \right| \\ &+ \sum_{k=K+1}^{n} \binom{n}{k} \frac{1}{n^k} + \sum_{k=K+1}^{\infty} \frac{1}{k!}\end{aligned}$$

Die letzte Summe ist auf Grund der Wahl von k kleiner als $\epsilon/3$. Die mittlere Summe schätzen wir wie folgt ab:

$$\begin{aligned}\binom{n}{k} \frac{1}{n^k} &= \frac{1}{k!} \frac{n(n-1) \cdots n-k+1}{n \cdot n \cdots n} \\ &= \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \leq \frac{1}{k!}.\end{aligned}$$

Damit ergibt dies die Abschätzung

$$\sum_{k=K+1}^{n} \binom{n}{k} \frac{1}{n^k} < \sum_{k=K+1}^{n} \frac{1}{k!} < \frac{\epsilon}{3}.$$

Schließlich gilt

$$\binom{n}{k} \frac{1}{n^k} = \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \xrightarrow[n \to \infty]{} \frac{1}{k!}.$$

Deshalb konvergiert die erste Summe für $n \to \infty$ gegen 0, d.h. es existiert $N > K$ so daß

$$\sum_{k=0}^{K} \left| \binom{n}{k} \frac{1}{n^k} - \frac{1}{k!} \right| < \frac{\epsilon}{3} \quad \text{für alle} \quad n \geq N.$$

Daraus folgt

$$\left| \left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^{\infty} \frac{1}{k!} \right| < \epsilon \quad \text{für} \quad n \geq N.$$

Damit gilt

$$e^q = \exp(q) \quad \text{für alle}\, q \in \mathbb{Q}.$$

□

Da $\exp(x)$ für alle $x \in \mathbb{R}$ definiert ist, können wir damit e^x für alle $x \in \mathbb{R}$ definieren.

Man kann aber auch e^x als Grenzwert in der folgenden Weise definieren. Dazu betrachten wir allgemeiner die Potenzen a^x, $a > 0$. Wir nehmen an, daß $a > 1$ gilt. Für $x \in \mathbb{R}$ existiert eine Folge (x_i) rationaler Zahlen, die monoton wächst und für die gilt:

$$\lim_{i\to\infty} x_i = x.$$

Nach Satz 2.24 ist (a^{x_i}) eine monoton wachsende Folge, die nach oben beschränkt ist. Deshalb existiert

$$a^x = \lim_{i\to\infty} a^{x_i}.$$

Es sei (y_i) eine beliebige Folge in $\mathbb{Q}$ mit $\lim_{i\to\infty} y_i = x$. Aus Satz 2.24, 4) folgt:

$$\exists\, C > 0\colon |a^{x_i} - a^{y_i}| \leq C|x_i - y_i|, \quad i \in \mathbb{N}.$$

Da $|x_i - y_i| \xrightarrow[i\to\infty]{} 0$, folgt daraus, daß

$$\lim_{i\to\infty} a^{x_i} = \lim_{i\to\infty} a^{y_i}.$$

Damit ist a^x unabhängig von der gewählten Folge (x_i), die gegen x konvergiert.

Satz 5.27. *(Cauchy-Produktsatz)*
Es seien $P(z) = \sum_{n=0}^{\infty} a_n z^n$ *und* $Q(z) = \sum_{m=0}^{\infty} b_m z^m$ *absolut konvergent im Punkte* $z \in \mathbb{C}$. *Dann gilt*

$$P(z) \cdot Q(z) = \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) z^k$$

und diese Reihe ist absolut konvergent.

Beweis: Der Beweis folgt unmittelbar aus Satz 5.19.

□

6. Stetige Funktionen und Grenzwerte

6.0. Der Funktionsbegriff

In vielen Bereichen der Wissenschaft und der Praxis hat man mit Problemen zu tun, bei denen Größen in bestimmter Weise von anderen Größen abhängen. Man spricht dann von Funktionen.

Beispiele: 1) Sei $F(r)$ der Flächeninhalt eines Kreises vom Radius r. Dann gilt bekanntlich

$$F(r) = \pi r^2,$$

d.h. der Flächeninhalt ist eine Funktion des Radius.

2) *Freier Fall*

Sei $g \approx 9,81 m/s^2$ die Erdbeschleunigung. Dann gilt für den Weg eines frei fallenden Körpers

$$s(t) = \frac{g}{2} t^2.$$

3) *Boylsches Gesetz für ideales Gas*

Es sei $p =$ Druck, $V =$ Volumen und $T =$ Temperatur. Dann gilt

$$p = \frac{c}{V}, \quad \text{für} \quad T = \text{Konstant}$$

4) *Barometrische Höhenformel*

Es sei $p = p(h)$ der Luftdruck in der Höhe h über dem Meeresspiegel. Dann gilt

$$p(h) = p_0 e^{-kh},$$

wobei $p_0 =$ Luftdruck auf dem Meeresspiegel.

In der Mathematik löst man sich von der speziellen Bedeutung der physikalischen Größen und untersucht Funktionen im abstrakten Sinne. Der Funktionsbegriff hat sich im Laufe der Zeit entwickelt. Wesentliche Beiträge dazu wurden unter anderm von Leibniz, Newton, Euler, Cauchy, Weierstraß geleistet.

Von Euler stammt die erste zusammenfassende Darstellung einer Funktionenlehre "Introductio in analysin infinitorum". Dort gibt Euler die folgende Definition einer Funktion:

"Eine Funktion einer veränderlichen Größe ist ein analytischer Ausdruck, der in irgendeiner Weise aus der Veränderlichen und aus Zahlen oder konstanten Größen zusammengesetzt ist."

Der heutige Funktionenbegriff ist viel allgemeiner gefaßt und gründet sich auf die Mengenlehre. Wir haben im Abschnitt über Mengenlehre ganz allgemein den Begriff der Abbildung $f\colon X \to Y$ von Mengen X, Y definiert. Darauf stützt sich der moderne Funktionenbegriff.

Definition 6.1.

1) *Es sei $D \subset \mathbb{R}$ eine Teilmenge. Eine Abbildung $f\colon D \to \mathbb{R}$ heißt (reellwertige)* **Funktion einer reellen Veränderlichen.** *$D = D(f)$ heißt der* **Definitionsbereich** *von f.*
2) *Es sei $D \subset \mathbb{C}$. Eine Abbildung $h\colon D \to \mathbb{C}$ heißt* **komplexwertige Funktion** *einer komplexen Veränderlichen. $D = D(h)$ heißt* **Definitionsbereich** *von h.*

Bezeichnung: $f(x)$, $x \in D(f)$, $h(z)$, $z \in D(h)$.

Bemerkung: Da wir bisher keine Voraussetzung über die Teilmenge $D \subset \mathbb{R}$ machen, kann man jede Funktion $f\colon D \to \mathbb{R}$ auch als Funktion $f\colon D \to \mathbb{C}$ auffassen.

Beispiele: 1) Sei $c \in \mathbb{R}$ und sei $f\colon \mathbb{R}$ definiert durch

$$f(x) = c \quad \text{für alle} \quad x \in \mathbb{R}.$$

Dann ist f eine konstante Funktion.
2) Seien $a_0, \dots, a_n \in \mathbb{C}$ und

$$h(z) = \sum_{j=0}^{n} a_j z^j, \quad z \in \mathbb{C},$$

Dann ist $h(z)$ ein Polynom, das als Funktion von z aufgefaßt werden kann. Dies ist eine Funktion im Sinne der Definition von Euler.

3) (Dirichlet) Sei $q\colon \mathbb{R} \to \mathbb{R}$ definiert durch

$$q(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ 1, & x \notin \mathbb{Q} \end{cases}$$

Dies ist keine Funktion im Eulerschen Sinne.

Definition 6.2.

1) *Es sei* $f: D \to \mathbb{R}$ *eine Funktion. Die Menge*
$$\Gamma_f = \{(x, f(x)) \mid x \in D\} \subset \mathbb{R} \times \mathbb{R}$$
heißt **Graph der Funktion** f.
2) *Seien* $D_1, D_2 \subset \mathbb{R}$ *und* $f_i: D_i \to \mathbb{R}$ *Funktionen. Dann sind* $f_1 + f_2$ *und* $f_1 \cdot f_2$ *die Funktionen auf* $D_1 \cap D_2$*, die definiert sind durch*
$$(f_1 + f_2)(x) = f_1(x) + f_2(x); \quad x \in D_1 \cap D_2$$
$$(f_1 \cdot f_2)(x) = f_1(x) \cdot f_2(x), \quad x \in D_1 \cap D_2.$$
Falls $f_2(x) \neq 0$ *für alle* $x \in D_1 \cap D_2$*, so ist* f_1/f_2 *definiert durch*
$$\frac{f_1}{f_2}(x) = \frac{f_1(x)}{f_2(x)} \quad \text{für alle} \quad x \in D_1 \cap D_2.$$

Definition 6.3. *(Komposition)*
Seien $f_1: D_1 \to \mathbb{R}$ *und* $f_2: D_2 \to \mathbb{R}$ *Funktionen und es sei*
$$D = \{x \in D_1 \mid f(x) \in D_2\}.$$
Dann ist $f_2 \circ f_1: D \to \mathbb{R}$ *definiert durch*
$$(f_2 \circ f_2)(x) = f_2(f_1(x)), \quad \forall x \in D.$$

Analoge Definitionen gelten für komplexe Funktionen.

Beispiel: Sei $s(x) = \sqrt{x}$, $x \in \mathbb{R}_+$, und $q(x) = x^2$, $x \in \mathbb{R}$. Dann ist
$$(s \circ q)(x) = |x|, \quad x \in \mathbb{R}.$$

6.1. Stetigkeit reeller Funktionen

Ein fundamentaler Begriff in der Theorie der Funktionen ist der Begriff der Stetigkeit. Die modernen Definition der Stetigkeit stammt von Bolzano (1817) und Weierstraß (ϵ-δ-Sprache).

Definition 6.4. *1) Eine Funktion* $f\colon D \to \mathbb{R}$ *heißt* **stetig in** $x_0 \in D$, *wenn gilt*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in D\colon |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

oder äquivalente Formulierung:

$$\forall \epsilon > 0\ \exists \delta > 0\colon |f(x) - f(x_0)| < \epsilon \text{ für alle } x \in D \text{ mit } |x - x_0| < \epsilon.$$

2) Eine Funktion $f\colon D \to \mathbb{R}$ *heißt stetig, wenn* f **in jedem Punkt** $x \in D$ *stetig ist.*

Für gegebenes $\epsilon > 0$ hängt $\delta > 0$ im allgemeinen vom Punkt $x \in D$ ab.

Eine Funktion $f\colon D \to \mathbb{R}$, die nicht stetig in $x_0 \in D$ ist, heißt **unstetig in** $x_0 \in D$. Durch Verneinung obiger Aussage erhalten wir:

$$\begin{aligned} &f \text{ ist unstetig in } x_0 \Leftrightarrow \\ &\quad \exists \epsilon > 0\ \forall \delta > 0\ \exists x \in D\colon |x - x_0| < \delta\ \wedge\ |f(x) - f(x_0)| \geq \epsilon. \end{aligned}$$

Beispiele: 1) Sei $c \in \mathbb{R}$ und $f(x) = c$ für alle $x \in \mathbb{R}$. Dann ist

$$f(x) - f(y) = 0, \quad \forall x, y \in \mathbb{R}.$$

Damit ist f stetig.

2) Sei $a > 1$ und $f(x) = a^x, \quad \forall x \in \mathbb{R}$.

Behauptung: a^x ist stetig

Beweis: Seien $x, y \in \mathbb{R}$ und $y > x$. In Kapitel 4 haben wir gezeigt:

$$0 < \sqrt[n]{a} - 1 < \frac{a-1}{n} \quad \text{für alle} \quad n \in \mathbb{N}.$$

Sei $\epsilon > 0$ gegeben. Wir setzen

$$n = \left[\frac{(a-1)a^x}{\epsilon}\right] + 1, \quad \epsilon = \frac{1}{n}.$$

Dann gilt für alle $x, y \in \mathbb{R}$ mit $y > x$: Wenn $|x - y| < \delta$, so ist

$$\begin{aligned}|a^y - a^x| &= a^x(a^{y-x} - 1) < a^x(a^{\frac{1}{n}} - 1) \\ &< \frac{a^x(a-1)}{n} < \frac{a^x(a-1)}{a^x(a-1)}\epsilon = \epsilon.\end{aligned}$$

□

3) Sei

$$f(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ 1, & x \in \mathbb{R}\backslash\mathbb{Q}, \end{cases},$$

Behauptung: f ist in jedem Punkt $x \in \mathbb{R}$ unstetig.

Beweis: Sei $x_0 \in \mathbb{R}$ und sei $\epsilon < 1$.
a) $x_0 \in \mathbb{Q}$. Es gilt: Für alle $\delta > 0$ existiert $x \in \mathbb{R}\backslash\mathbb{Q}$: $|x - x_0| < \delta$.

$$\Rightarrow |f(x) - f(x_0)| = 1 > \epsilon.$$

b) $x_0 \in \mathbb{R}\backslash\mathbb{Q}$. Da $\mathbb{Q}$ dicht ist in $\mathbb{R}$, existiert zu jedem $\delta > 0$ ein $x \in \mathbb{Q}$: $|x - x_0| < \delta$.

$$\Rightarrow |f(x) - f(x_0)| = 1 > \epsilon.$$

□

Definition 6.5. *Eine Funktion $f\colon D \to \mathbb{R}$ heißt* **Lipschitz-stetig**, *wenn gilt:*
$\exists L > 0\ \forall x, y \in D$:

$$|f(x) - f(y)| \leq L|x - y|.$$

Lemma 6.6. *Sei $f\colon D \to \mathbb{R}$ Lipschitz - stetig. Dann ist f stetig.*

Beweis: Seien $x_0 \in D$ und $\epsilon > 0$ gegeben. Sei $\delta = \frac{\epsilon}{L}$. Dann gilt für $x \in D$:

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| \leq L|x - x_0| < L \cdot \delta = \epsilon.$$

□

Beispiel: $\exp(x)$ ist auf jedem abgeschlossenen Intervall $[a, b]$, $b < \infty$, Lipschitz-stetig.

Beweis: 1) Sei $c > 0$ und $|x| \le c$. Dann gilt:

$$|\exp(x) - 1| \le \sum_{n=1}^{\infty} \frac{|x|^n}{n!} \le |x| \sum_{n=1}^{\infty} \frac{c^{n-1}}{n!} \le \exp(c)|x|.$$

2) Seien $x, x_0 \in [a, b]$, $-\infty < a < b < \infty$, $x > x_0$. Dann folgt aus 1):

$$\begin{aligned} |\exp(x) - \exp(x_0)| &= \exp(x_0)|\exp(x - x_0) - 1| \\ &\le \exp(b - a)\exp(b)|x - x_0| \\ &= \exp(2b - a) \cdot |x - x_0|. \end{aligned}$$

Damit haben wir gezeigt, daß $\exp(x)$ Lipschitz-stetig auf $[a, b]$ ist. Insbesondere ist $\exp\colon \mathbb{R} \to \mathbb{R}_+$ stetig.

Rechenregeln

Satz 6.7. *Seien $f, g\colon D \to \mathbb{R}$ stetig in $x_0 \in D$, und sei $\lambda \in \mathbb{R}$. Dann sind auch $f + g$, $f \cdot g$ und λf stetig in x_0. Wenn $g(x_0) \ne 0$ ist, so ist f/g stetig in x_0.*

Beweis: Sei $\epsilon > 0$ gegeben. Aus der Stetigkeit von f in x_0 folgt, daß zu ϵ ein $\delta_1 > 0$ existiert so, daß

$$\forall x \in D\colon |x - x_0| < \delta_1 \Rightarrow |f(x) - f(x_0)| < \frac{\epsilon}{2}.$$

Aus der Stetigkeit von g in x_0 folgt: Zu ϵ existiert $\delta_2 > 0$:

$$\forall x \in D\colon |x - x_0| < \delta_2 \Rightarrow |g(x) - g(x_0)| < \frac{\epsilon}{2}.$$

Dann gilt für alle $x \in D$ mit $|x - x_0| < \delta$:

$$\begin{aligned} |(f + g)(x) - (f + g)(x_0)| &\le |f(x) - f(x_0)| + |g(x) - g(x_0)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Die anderen Fälle können analog behandelt werden. Übung!

□

Folgerung 6.8. *Jedes Polynom ist stetig.*

Satz 6.9. *(Kompositionsregel)*
Seien $f_1\colon D_1 \to \mathbb{R}$ und $f_2\colon D_2 \to \mathbb{R}$ Funktionen. Es sei $x_0 \in D_1$ und $f(x_0) \in D_2$. Weiter sei f stetig in x_0 und g stetig in $f(x_0)$. Dann ist $g \circ f$ stetig in x_0.

Beweis: Sei $\epsilon > 0$ gegeben. Da g stetig in $f(x_0)$ ist, existiert zu ϵ ein $\delta_1 > 0$:

$$\forall y \in D_2\colon |y - f(x_0)| < \delta_1 \Rightarrow |g(y) - g(f(x_0))| < \epsilon.$$

Da f stetig in $x_0 \in D_1$ ist, existiert zu δ_1 ein $\delta_2 > 0$:

$$\forall x \in D_1\colon |x - x_0| < \delta_2 \Rightarrow |f(x) - f(x_0)| < \delta_1.$$

Damit gilt für alle $x \in D_1$:

$$\begin{aligned} |x - x_0| < \delta_2 &\Rightarrow |f(x) - f(x_0)| < \delta_1 \\ &\Rightarrow |g(f(x)) - g(f(x_0))| < \epsilon. \end{aligned}$$

□

Wir betrachten jetzt zwei andere Möglichkeiten, die Stetigkeit einer Funktion zu definieren.

a) **Stetigkeit und Umgebungsbegriff**
Um die Betrachtung zu vereinfachen, beschränken wir uns auf Funktionen, die auf ganz $\mathbb{R}$ definiert sind.
Es sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Funktion und $x_0 \in \mathbb{R}$. Nach Definition gilt

$$\begin{aligned} &f \text{ ist stetig in } x_0 \Leftrightarrow \\ &\quad \forall \epsilon > 0\ \exists\, \delta > 0\ \forall x \in \mathbb{R}\colon |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon. \end{aligned}$$

Weiterhin gilt:

$$\begin{aligned} \{x \in \mathbb{R} \mid |x - x_0| < \delta\} &= (x_0 - \delta, x_0 + \delta) \\ \{y \in \mathbb{R} \mid |y - f(x_0)| < \epsilon\} &= (f(x_0) - \epsilon, f(x_0) + \epsilon). \end{aligned}$$

Deshalb können wir die obige Definition der Stetigkeit von f in x_0 wie folgt umformulieren:

$$\begin{aligned} &f \text{ ist stetig in } x_0 \Leftrightarrow \\ &\quad \forall \epsilon > 0\ \exists \delta > 0\colon f((x_0 - \delta, x_0 + \delta)) \subset (f(x_0) - \epsilon, f(x_0) + \epsilon). \end{aligned}$$

Jetzt wollen wir uns von der ϵ-δ-Definition lösen. Dazu führen wir den Umgebungsbegriff ein.

Definition. *Es sei $x_0 \in \mathbb{R}$. Eine Teilmenge $U \subset \mathbb{R}$ heißt Umgebung von x_0 (in $\mathbb{R}$), wenn ein $\delta > 0$ existiert so, daß $(x_0 - \delta, x_0 + \delta) \subset U$.*

Beispiel: Es sei $x_0 \in \mathbb{R}$. Weiter sei $a, b \in \mathbb{R}$, $a < x_0 < b$. Dann ist $U = (a, b)$ eine Umgebung von x, denn für $\delta < \min(b - x_0, x_0 - a)$. ist $(x_0 - \delta, x_0 + \delta) \subset (a, b) = U$.

Bezeichnung: Die Menge $U_\delta(x_0) = (x_0 - \delta, x_0 + \delta)$ heißt δ-Umgebung von x_0.

Satz 6.11. *Es sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Funktion und $x_0 \in \mathbb{R}$. Dann gilt:*
f ist stetig in x_0 $\Leftrightarrow$ Für jede Umgebung U von $f(x_0)$ existiert eine Umgebung V von x_0 mit $f(V) \subset U$.

Beweis: $\Rightarrow$) Es sei U eine Umgebung von $f(x_0)$. Dann folgt

$$\exists \epsilon > 0\colon U_\epsilon(f(x_0)) \subset U$$

Wegen der Stetigkeit von f in x_0 existiert $\delta > 0$:

$$f(U_\delta(x_0)) \subset U_\epsilon(f(x_0)) \subset U.$$

Die Menge $V = U_\delta(x_0)$ ist nach Definition eine Umgebung von x_0 und es gilt $f(V) \subset U$.

$\Leftarrow$) Es sei $\epsilon > 0$. Wir betrachten die Umgebung $U = U_\epsilon(f(x_0))$. Nach Voraussetzung existiert zu U eine Umgebung V von x_0 so daß $f(V) \subset U$. Da V Umgebung von x_0 ist, existiert $\delta > 0$ mit $U_\delta(x_0) \subset V$. Daraus folgt $f(U_\delta(x_0)) \subset U_\delta(f(x_0))$. Wie wir oben gesehen haben, ist dies äquivalent zur Stetigkeit von f in x_0.

$\square$

Anwendung: Es seien $f\colon \mathbb{R} \to \mathbb{R}$ und $g\colon \mathbb{R} \to \mathbb{R}$ Funktionen, wobei f stetig in x_0, und g stetig in $f(x_0)$ ist. In Satz 6.9. haben wir gezeigt, daß dann $g \circ f$ stetig in x_0 ist. Mit Hilfe des Umgebungsbegriffes kann man den Beweis wie folgt führen: Es sei U eine Umgebung von $g(f(x_0))$. Da g stetig in $f(x_0)$ ist, existiert eine Umgebung V von $f(x_0)$ so, daß $g(V) \subset U$. Da f stetig in x_0 ist, existiert eine Umgebung W von x_0 so, daß $f(W) \subset V$. Hieraus folgt $(g \circ f)(W) \subset U$ Aus Satz 6.11 folgt daher, daß $g \circ f$ stetig in x_0 ist.

b) **Folgenkriterium für die Stetigkeit**

Satz 6.12. *Sei $f\colon D \to \mathbb{R}$ eine Funktion. Dann gilt: f ist stetig in $x_0 \in D$ genau dann, wenn für jede Folge (x_n) in D gilt:*

$$x_n \xrightarrow[n\to\infty]{} x_0 \quad \Rightarrow f(x_n) \xrightarrow[n\to\infty]{} f(x_0).$$

Beweis: $\Rightarrow$) Wir setzen voraus, daß f in $x_0 \in D$ stetig ist. Es sei (x_n) eine Folge in D und es gelte $\lim_{n\to\infty} x_n = x_0$. Wir möchten jetzt zeigen, daß $\lim_{n\to\infty} f(x_n) = f(x_0)$ gilt. Dazu wählen wir $\epsilon > 0$. Zu ϵ existiert $\delta > 0$:

$$|f(x) - f(x_0)| < \epsilon \quad \text{für alle} \quad x \in D \quad \text{mit} \quad |x - x_0| < \delta.$$

Da $x_n \xrightarrow[n\to\infty]{} x_0$ gilt, existiert $N \in \mathbb{N}$:

$$|x_n - x_0| < \delta \quad \text{für alle} \quad n \geq N.$$

Hieraus folgt

$$|f(x_n) - f(x_0)| < \epsilon \quad \text{für} \quad n \geq N.$$

Damit haben wir folgendes gezeigt:

$$\forall \epsilon > 0 \; \exists n \in \mathbb{N} \; \forall n \geq N\colon |f(x_n) - f(x_0)| < \epsilon.$$

Dies bedeutet aber, daß $(f(x_n))$ eine konvergente Folge ist und

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

$\Leftarrow$) Wir setzen jetzt voraus, daß für alle Folgen (x_n) in D gilt:

$$x_n \xrightarrow[n\to\infty]{} x_0 \quad \Rightarrow f(x_n) \xrightarrow[n\to\infty]{} f(x_0).$$

Annahme: f sei unstetig in x_0 Dann gilt:

$$\exists \epsilon > 0 \; \forall \delta > 0 \; \exists x \in D\colon |x - x_0| < \delta \;\wedge\; |f(x) - f(x_0)| \geq \epsilon.$$

Hieraus folgt

$$\forall n \in \mathbb{N} \; \exists x_n \in D\colon |x_n - x_0| < \frac{1}{n} \;\wedge\; |f(x_n) - f(x_0)| \geq \epsilon.$$

Damit haben wir eine Folge (x_n) in D konstruiert mit $x_n \to x_0$, für die die Folge $(f(x_n))$ nicht gegen $f(x_0)$ konvergiert. Dies ist ein Widerspruch zu Voraussetzung.

$\square$

Folgerung 6.13. Für alle $x \in \mathbb{R}$ gilt:

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

Beweis: Wir haben bereits gezeigt:

$$\forall q \in \mathbb{Q}\colon e^q = \exp(q).$$

Es sei $x \in \mathbb{R}$. Dann existiert eine Folge (x_i) rationaler Zahlen mit $x_i \xrightarrow[i\to\infty]{} x$. Da wir bereits wissen, daß $\exp(x)$ und e^x stetige Funktionen sind, folgt aus Satz 6.12:

$$e^x = \lim_{i\to\infty} e^{x_i} = \lim_{i\to\infty} \exp(x_i) = \exp(x).$$

$\square$

Definition 6.14. *Eine Funktion $f\colon [a,b] \to \mathbb{R}$ heißt* **streng monoton**

$$\begin{bmatrix}\textbf{wachsend}\\ \textbf{fallend}\end{bmatrix} \Leftrightarrow \begin{bmatrix}\forall x,y \in [a,b]\colon x < y \Rightarrow f(x) < f(y)\\ \forall x,y \in [a,b]\colon x < y \Rightarrow f(y) < f(x)\end{bmatrix}$$

f heißt **streng monoton**, *wenn f entweder streng monoton wachsend oder streng monoton fallend ist.*

Sei $f\colon [a,b] \to \mathbb{R}$ streng monoton, und sei

$$D_1 = f([a,b]) = \{f(x) \mid x \in [a,b]\}$$

der Wertebereich von f. Da f streng monoton ist, gilt:

$$\forall y \in D\ \exists! x \in [a,b]\colon f(x) = y,$$

d.h. f ist eine eineindeutige (oder bijektive) Abbildung $f\colon [a,b] \to D_1$. Die **Umkehrfunktion** $f^{-1}\colon D_1 \to \mathbb{R}$ ist dann definiert durch

$$f^{-1}(y) = x, \quad \text{falls} \quad y = f(x) \quad \text{gilt.}$$

Satz 6.15. *Es sei $f\colon [a,b] \to \mathbb{R}$ streng monoton und stetig. Dann ist die Umkehrfunktion $f^{-1}\colon D_1 \to \mathbb{R}$ ebenfalls streng monoton und stetig.*

Beweis: Sei f streng monoton wachsend. Den Fall, daß f streng monoton fallend ist, kann man durch den Übergang $f \longmapsto -f$ auf obigen Fall zurückführen.
1) f^{-1} ist streng monoton wachsend.

Seien $y_1, y_2 \in D_1$ und es gelte $y_1 < y_2$. Wir nehmen an, daß $f^{-1}(y_1) \geq f^{-1}(y_2)$ gilt. Da f streng monoton wachsend ist, folgt daraus

$$y_1 = f(f^{-1}(y_1)) \geq f(f^{-1}(y_2)) = y_2.$$

Dies ist ein Widerspruch zur Voraussetzung.

2) f^{-1} ist stetig.

Um dies zu zeigen wenden wir das Folgenkriterium an. Es sei $x_0 \in [a, b]$. Weiter sei (y_n) eine Folge in D_1 und es gelte

$$y_n \xrightarrow[n\to\infty]{} f^{-1}(x_0).$$

Zu zeigen: $f^{-1}(y_n) \xrightarrow[n\to\infty]{} f(f^{-1}(x_0)) = x_0$. Auf Grund der Definition von D_1 gilt:

$$\forall n \in \mathbb{N}: \quad f(a) \leq y_n \leq f(b).$$

Daraus folgt

$$a \leq f^{-1}(y_n) \leq b \quad \text{für alle} \quad n \in \mathbb{N},$$

d.h., $(f^{-1}(y_n))$ ist eine beschränkte Folge. Aus dem Satz von Bolzano-Weierstraß folgt, daß $(f^{-1}(y_n))$ mindestens einen Häufigkeitspunkt besitzt. Es sei x_1 ein Häufungspunkt von $(f^{-1}(y_n))$. Dann existiert eine Teilfolge $(f^{-1}(y_{n_j}))$ mit

$$f^{-1}(y_{n_j}) \xrightarrow[j\to\infty]{} x_1.$$

Da f stetig ist, erhalten wir aus Satz 6.12, daß

$$y_{n_j} = f(f^{-1}(y_{n_j})) \xrightarrow[j\to\infty]{} f(x_1)$$

gilt. Nach Voraussetzung gilt aber $y_n \xrightarrow[j\to\infty]{} f(x_0)$. Daraus folgt

$$f(x_1) = f(x_0)$$

Da f eineindeutig ist, folgt daraus $x_1 = x_0$. Wir haben damit gezeigt, daß die Folge $(f^{-1}(y_n))$ genau einen Häufungspunkt besitzt, und zwar x_0. Aus Satz 4.21 folgt dann

$$f^{-1}(y_n) \xrightarrow[n\to\infty]{} x_0.$$

□

Satz 6.16. *(Zwischenwertsatz) Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann existiert zu jedem y zwischen $f(a)$ und $f(b)$ ein $c \in [a,b]$ mit $f(c) = y$.*

Beweis: Ohne Beschränkung der Allgemeinheit können wir annehmen, daß gilt:

$$f(a) \le y < f(b).$$

Andernfalls wäre $y = f(b)$. Wir konstruieren eine Intervallschachtelung $I_n = [a_n, b_n]$ mit folgenden Eigenschaften:

1) $I_1 = [a,b]$ und $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$.
2) $\forall k \in \mathbb{N}\colon\ b_k - a_k = \frac{1}{2^{k-1}}(b-a)$.
3) $\forall k \in \mathbb{N}\colon\ f(a_k) \le y \le f(b_k)$.

Wir konstruieren die Intervallschachtelung rekursiv.
1. Anfangsschritt: Es sei $I_1 = [a,b]$.
2. Schritt von k auf $k+1$: Es sei für $k \in \mathbb{N}$ das Intervall I_k bereits konstruiert. Sei

$$m_k = a_k + \frac{1}{2}(b_k - a_k)$$

der Mittelpunkt von I_k. Wir unterscheiden 2 Fälle:
1. Fall $f(m_k) > y$.
Dann sei $a_{k+1} = a_k$ und $b_{k+1} = m_k$.
2. Fall: $f(m_k) \le y$.
Dann sei $a_{k+1} = m_k$ und $b_{k+1} = b_k$.
Damit erhalten wir das Intervall I_{k+1}. Nach Konstruktion von I_{k+1} gilt:

(1) $I_{k+1} = [a_{k+1}, b_{k+1}] \subset I_k$;
(2) $b_{k+1} - a_{k+1} = \frac{1}{2}(b_k - a_k) = \frac{1}{2^k}(b-a)$
(3) $f(a_{k+1}) \le y \le f(b_{k+1})$.

Durch diese rekursive Definition erhalten wir eine Intervallschachtelung (I_n) mit den gewünschten Eigenschaften 1) - 3). Die Anwendung des Intervallschachtelungsprinzips ergibt:

$$\exists! c \in \mathbb{R}\colon\ c \in \bigcap_{n\in\mathbb{N}} I_n.$$

Für dieses c gilt:

(1) $c \in [a,b]$

(2) $a_k \xrightarrow[k\to\infty]{} c$
(3) $b_k \xrightarrow[k\to\infty]{} c.$

Da f stetig in c ist, folgt aus Satz 6.12:

$$\lim_{k\to\infty} f(a_k) = f(c) = \lim_{k\to\infty} f(b_k). \tag{6.1}$$

Auf Grund von Eigenschaft 3) der Folge (I_n) gilt:

$$\forall k \in \mathbb{N}\colon f(a_k) \le y \le f(b_k)$$

Hieraus folgt

$$\lim_{k\to\infty} f(a_k) \le y \le \lim_{k\to\infty} f(b_k),$$

und aus (6.1) erhalten wir

$$y = f(c).$$

□

Mit Hilfe des Zwischenwertsatzes kann man z.B. die Existenz von Lösungen für Gleichungen beweisen. Als Beispiel erwähnen wir den folgenden bekannten Satz:

Satz 6.17. *Es sei $n \in \mathbb{N}$ ungerade und es seien $a_0, \dots, a_n \in \mathbb{R}$, wobei $a_n \neq 0$. Dann hat das Polynom $P(x) = \sum_{j=0}^{n} a_j x^j$ eine reelle Nullstelle.*

Beweis: Da P und $a_n^{-1}P$ die gleichen Nullstellen haben, können wir annehmen, daß $a_n = 1$. Es sei

$$r = 1 + |a_{n-1}| + \cdots + |a_0|.$$

Dann ist $r > 1$ und es gilt

$$\begin{aligned} P(r) &= r^n + a_{n-1}r^{n-1} + \cdots + a_0 \ge \\ &\ge r^n - (|a_{n-1}|r^{n-1} + \cdots + |a_0|) \\ &\ge r^n - r^{n-1}(|a_{n-1}| + \cdots + |a_0|) \\ &= r^n - r^{n-1}(r-1) = r^{n-1} > 0. \end{aligned}$$

Da n ungerade ist, erhalten wir

$$\begin{aligned} P(-r) &= -r^n + (a_{n-1}r^{n-1} + \cdots + a_0) \\ &\le -r^n + r^{n-1}(|a_{n-1}| + \cdots + |a_0|) \\ &< -r^n + r^{n-1}(r-1) = -r^{n-1} < 0. \end{aligned}$$

Aus dem Zwischenwertsatz folgt, daß ein $x_0 \in [-r, r]$ existiert mit $P(x_0) = 0$.

$\square$

Satz 6.18. *Die Exponentialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}^+$ *ist streng monoton wachsend, stetig und bijektiv.*

Beweis: Wir haben bereits gezeigt, daß exp streng monoton wachsend und stetig ist. Wegen der strengen Monotonie ist exp injektiv. Es bleibt noch zu zeigen, daß $\exp\colon \mathbb{R} \to \mathbb{R}^+$ surjektiv ist. Es sei $y > 0$ gegeben. Wir wählen $n \in \mathbb{N}$ so, daß $1/n < y < n$. Dann gilt

$$e^n = 1 + n + \frac{n^2}{2} + \cdots > n > y$$

und

$$e^{-n} = \frac{1}{e^n} < \frac{1}{n} < y.$$

Nach dem Zwischenwertsatz existiert ein $x \in [-n, n]$ so, daß $y = e^x$.

$\square$

Definition 6.19. *Es sei* $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ *die Umkehrfunktion zu* $\exp\colon \mathbb{R} \to \mathbb{R}^+$. ln *heißt* **natürlicher Logarithmus**.

Bemerkung: Häufig wird die Funktion ln auch mit log bezeichnet.

Satz 6.20. $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ *ist streng monoton wachsend, stetig und bijektiv. Weiter gilt*

$$1)\ \ln(1) = 0 \qquad 2)\ \forall x, y \in \mathbb{R}_+ :\ \ln(xy) = \ln x + \ln y.$$

Beweis: Aus Satz 6.15 folgt, daß ln streng monoton und stetig ist. Außerdem ist ln bijektiv. Weiter gilt $e^0 = 1$. Daraus folgt

$$\ln(1) = \exp^{-1}(1) = 0.$$

Es seien $x, y \in \mathbb{R}^+$. Aus Satz 6.18 folgt, daß eindeutig bestimmte Zahlen $u, v \in \mathbb{R}$ existieren so, daß $e^u = x$ und $e^v = y$. Hieraus folgt

$$\ln(xy) = \ln(e^u e^v) = \ln(e^{u+v}) = u + v = \ln(x) + \ln(y).$$

$\square$

Satz 6.21. *Es sei $I \subset \mathbb{R}$ ein Intervall (endlich oder unendlich) und $f\colon I \to \mathbb{R}$ streng monoton und stetig. Dann gilt*

1) *$J = f(I)$ ist ein Intervall.*
2) *$f\colon I \to J$ ist bijektiv*
3) *$f^{-1}\colon J \to I$ ist stetig und streng monoton.*

Beweis: Wir betrachten den Fall, daß f monoton wachsend ist. 2) und 3) haben wir bereits gezeigt. Es bleibt noch 1) zu zeigen. Wir betrachten den Fall

$$I = (a,b),\ a,b \in \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\},\ a < b.$$

Es sei

$$c = \inf\{f(x) \mid x \in I\}, \quad d = \sup\{f(x) \mid x \in I\}.$$

Sei $x \in (a,b)$. Dann existieren $x_1, x_2 \in (a,b)$ so, daß $x_1 < x < x_2$. Aus der strengen Monotonie von f folgt

$$c \leq f(x_1) < f(x) < f(x_2) \leq d.$$

$\Rightarrow$

$$c < d \quad \text{und} \quad f\left((a,b)\right) \subset (c,d).$$

Es sei $y \in (c,d)$. Dann gilt auf Grund der Definition von c und d:

$$\exists x_1, x_2 \in (a,b)\colon\ c \leq f(x_1) < y < f(x_2) \leq d.$$

Aus dem Zwischenwertsatz folgt

$$\exists x \in [x_1, x_2]\colon \quad f(x) = y.$$

Hieraus erhalten wir

$$f\left((a,b)\right) = (c,d).$$

Die Fälle $I = [a,b),\ (a,b]$ und $[a,b]$ behandelt man ebenso.

□

6.2. Stetigkeit komplexer Funktionen

Definition 6.22. *Sei $D \subset \mathbb{C}$. Eine Funktion $f\colon D \to \mathbb{C}$ heißt* **stetig in** *$z_0 \in D$, wenn gilt*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall z \in D\colon |z - z_0| < \delta \Rightarrow |f(z) - f(z_0)| < \epsilon.$$

Die Funktion $f\colon D \to \mathbb{C}$ heißt **stetig***, wenn sie in jedem Punkt $z_0 \in D$ stetig ist.*

Alle Sätze, bei denen die Anordnung von $\mathbb{R}$ nicht benutzt wurde, übertragen sich.

Satz 6.23. *(Folgenkriterium) $f\colon D \to \mathbb{C}$ ist stetig in $z_0 \in D$ genau dann, wenn für alle Folgen (z_n) in D gilt:*

$$z_n \xrightarrow[n\to\infty]{} z_0 \;\Rightarrow\; f(z_n) \xrightarrow[n\to\infty]{} f(z_0).$$

Beweis: Analog zum Beweis von Satz 6.12.

Der Umgebungsbegriff wird genauso wie im reellen Fall definiert.
Beispiel: $\exp(z)$ ist stetig. <u>Übung!</u>

Definition 6.24. *Für $z \in \mathbb{C}$ sei*

$$e^z =_{Df} \exp(z).$$

Aus der Folgerung 6.13 ergibt sich, daß für $x \in \mathbb{R}$, e^x mit der früheren Definition übereinstimmt. Wir haben die Potenzreihe für e^x benutzt, um e^x zu einer komplexen Funktion fortzusetzen.

Satz 6.25. *Für die Funktion e^z gilt*

1) $\forall z \in \mathbb{C}\colon\ \overline{e^z} = e^{\overline{z}}$.
2) $\forall z, w \in \mathbb{C}\colon\ e^z e^w = e^{z+w}$.
3) $\forall x \in \mathbb{R}\colon\ |e^{ix}| = 1$.

Beweis:

1)

$$\overline{\exp(z)} = \overline{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \sum_{n=0}^{\infty} \frac{\overline{z}^n}{n!} = \exp(\overline{z}).$$

2) Wurde schon gezeigt.
3) Aus 1) folgt $e^{\overline{ix}} = e^{-ix}$. Daraus und aus 3) erhalten wir

$$|e^{ix}|^2 = e^{ix} e^{-ix} = 1.$$

$\square$

Es gelten die gleichen Rechenregeln wie im Falle reeller Funktionen.

6.3. Kompakte Mengen und Extremwerte stetiger Funktionen

Definition 6.26. *Eine Teilmenge $A \subset \mathbb{C}$ heißt* **abgeschlossen**, *wenn für alle konvergenten Folgen (a_n) in A gilt:*

$$\lim_{n\to\infty} a_n \in A.$$

Beispiele: $[a, b], \mathbb{R},\ \{z \in \mathbb{C} \mid |z| \le r\}$, $\mathbb{C}$ sind abgeschlossen.
$[a, b), (a, b),\ \{z \in \mathbb{C} \mid |z| < r\}$ sind nicht abgeschlossen.

Satz 6.27.

1) *Es seien $A_1, \dots , A_k \subset \mathbb{C}$ abgeschlossen. Dann ist $\bigcup_{j=1}^{k} A_j$ abgeschlossen.*
2) *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*

Beweis: Übung!

Die Vereinigung unendlich vieler abgeschlossener Mengen ist im allgemeinen nicht abgeschlossen.

Beispiel: $\bigcup_{n=1}^{\infty}[\frac{1}{n}, 1] = (0, 1]$.

Definition 6.28. *Eine Menge $K \subset \mathbb{C}$ heißt* **kompakt**, *wenn gilt:*

1) *A ist abgeschlossen.*
2) *A ist beschränkt, d.h., $\exists R \in \mathbb{R}_+ :\ A \subset \{z \in \mathbb{C} \mid |z| \le R\}$.*

Beispiel: $[a, b]$ ist kompakt für $a, b \in \mathbb{R}$, $\mathbb{R}$ ist nicht kompakt.

Satz.

1) *Es seien $K_1, \dots , K_m \subset \mathbb{C}$ kompakte Teilmengen. Dann ist $\bigcup_{j=1}^{m} K_j$ kompakt.*
2) *Der Durchschnitt beliebig vieler kompakter Mengen ist kompakt.*

Beweis: Folgt aus Satz 6.27.

Bolzano-Weierstraß-Charakterisierung kompakter Mengen

Satz 6.29. *Eine Menge $K \subset \mathbb{C}$ ist kompakt genau dann, wenn jede Folge (z_n) von Elementen aus K eine konvergente Teilfolge (z_{n_j}) besitzt so, daß $\lim_{j\to\infty} z_{n_j} \in K$.*

Beweis: $\Rightarrow$) Sei $K \subset \mathbb{C}$ kompakt, und sei (z_n) eine Folge in K. Da K beschränkt ist, ist (z_n) eine beschränkte Folge. Aus Korollar 4.20 folgt, daß (z_n) eine konvergente Teilfolge (z_{n_j}) besitzt. Sei

$$z_0 = \lim_{j\to\infty} z_{n_j}.$$

Da K abgeschlossen ist, ist $z_0 \in K$.

$\Leftarrow$) i) K ist beschränkt.

Annahme: K sei nicht beschränkt.

Dann existiert eine Folge (z_n) in K mit $|z_n| \geq n$ für alle $n \in \mathbb{N}$. Diese Folge besitzt offensichtlich keine konvergente Teilfolge und dies ist ein Widerspruch zur Voraussetzung.

ii) K ist abgeschlossen.

Sei (z_n) eine konvergente Folge in K und sei $z_0 = \lim_{n\to\infty} z_n$. Dann ist zu zeigen, daß $z_0 \in K$. Nach Voraussetzung besitzt (z_n) eine konvergente Teilfolge (z_{n_j}) mit $\lim_{j\to\infty} z_{n_j} \in K$. Dann ist aber

$$z_0 = \lim_{j\to\infty} z_{n_j} \in K.$$

$\square$

Satz 6.30. *Sei $K \subset \mathbb{C}$ kompakt und $f\colon K \to \mathbb{C}$ stetig. Dann ist $f(K)$ kompakt.*

Beweis: Es sei (w_n) eine Folge in $f(K)$. Dann existiert eine Folge (z_n) in K mit $w_n = f(z_n)$ für alle $n \in \mathbb{N}$. Da K kompakt ist, hat (z_n) eine konvergente Teilfolge (z_{n_j}) mit

$$z_0 = \lim_{j\to\infty} z_{n_j} \in K.$$

Da f stetig ist, folgt aus Satz 6.12

$$\lim_{j\to\infty} f(z_{n_j}) = f(z_0) \in K.$$

Aus Satz 6.29 erhalten wir, daß $f(K)$ kompakt ist.

$\square$

Satz 6.31. *(Satz vom Minimum u. Maximum)*
Sei $D \subset \mathbb{C}$ kompakt und $f\colon D \to \mathbb{R}$ stetig. Dann existieren $z_1, z_2 \in D$ so, daß gilt:

$$\forall z \in D\colon\ f(z_1) \leq f(z) \leq f(z_2),$$

d.h. f nimmt sein Maximum und Minimum an.

Beweis: Da D kompakt ist, so ist nach Satz 6.30 $f(D) \subset \mathbb{R}$ ebenfalls kompakt. Insbesondere ist $f(D)$ beschränkt. Daraus folgt, daß $m =_{Df} \inf f(D)$ und $M =_{Df} \sup f(D)$ existieren. Nach Definition von $\inf f(D)$ und $\sup f(D)$ existieren Folgen (z_n) und (w_n) in D mit $f(z_n) \to m$ und $f(w_n) \to M$ für $n \to \infty$. Da $f(D)$ abgeschlossen ist, ist $m, M \in f(D)$. Dies bedeutet aber

$$\exists z_1, z_2 \in D\colon\ m = f(z_1),\ M = f(z_2).$$

□

Folgerung 6.32. Sei $D \subset \mathbb{C}$ kompakt und $f\colon D \to \mathbb{C}$ stetig. Dann ist f beschränkt, d.h., es exixstiert $C > 0$ so, daß

$$\forall z \in D\colon |f(z)| \leq C.$$

Beweis: $|f|\colon D \to \mathbb{R}$ ist stetig. Die Behauptung folgt aus Satz 6.31.

□

Bemerkung: Die Kompaktheit von D und die Stetigkeit von f sind notwendige Voraussetzungen.

Beispiele: 1) $f\colon [-1,1] \to \mathbb{R}$ sei definiert durch

$$f(x) = \begin{cases} \frac{1}{x},\ x \neq 0 \\ 0,\ x = 0. \end{cases}$$

Die Funktion f ist nicht beschränkt. Sie ist aber unstetig in $x = 0$.

2) $f\colon (0,1] \to \mathbb{R}$ sei definiert durch $f(x) = \frac{1}{x}$. Die Funktion f ist ebenfalls nicht beschränkt. In diesem Falle ist f stetig, aber der Definitionsbereich ist nicht kompakt.

Definition 6.33. *Eine Funktion* $f\colon D \to \mathbb{C}$ *heißt* **gleichmäßig stetig**, *wenn gilt:*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall z, w \in D\colon |z - w| < \delta \Rightarrow |f(z) - f(w)| < \epsilon.$$

Beispiele: 1) Es sei $f\colon D \to \mathbb{C}$ Lipschitz-stetig, d.h., es gilt

$$\exists L > 0\ \forall z, w \in D\colon\ |f(z) - f(w)| \le L|z - w|.$$

Dann ist f gleichmäßig stetig.
2) $f(x) = \frac{1}{x},\ x \in (0,1]$. Die Funktion $f\colon (0,1] \to \mathbb{R}$ ist nicht gleichmäßig stetig.

Satz 6.34. *(Heine) Sei* $D \subset \mathbb{C}$ *kompakt und* $f\colon D \to \mathbb{C}$ *stetig. Dann ist* f *gleichmäßig stetig.*

Beweis: Wir nehmen an, daß f nicht gleichmäßig stetig ist. Aus der Definition 6.33 folgt durch Verneinung des logischen Ausdruckes

$$\exists \epsilon > 0\ \forall \delta > 0\ \exists z, w \in D\colon |z - w| < \delta\ \wedge\ |f(z) - f(w)| \ge \epsilon.$$

Wir wählen $\delta_n = 1/n,\ n \in \mathbb{N}$. Dann folgt

$$\exists \epsilon > 0\ \forall n \in \mathbb{N}\ \exists z_n, w_n \in D(|z_n - w_n| < \frac{1}{n}\ \wedge\ |f(z_n) - f(w_n)| \ge \epsilon)$$

Da D kompakt ist, existiert eine konvergente Teilfolge (z_{n_j}) von (z_n) mit

$$z_0 =_{Df} \lim_{j\to\infty} z_{n_j} \in D.$$

Weiter gilt

$$|z_{n_k} - w_{n_k}| < \frac{1}{n_k} \le \frac{1}{k}.$$

Daraus folgt

$$z_0 = \lim_{k\to\infty} w_{n_k}.$$

Da f stetig ist, folgt aus Satz 6.12 (Folgenkriterium)

$$\lim_{j\to\infty} f(z_{n_j}) = f(z_0) = \lim_{j\to\infty} f(w_{n_j}).$$

Dies steht aber im Widerspruch zu

$$|f(z_n) - f(w_n)| \ge \epsilon \quad \text{für alle} \quad n \in \mathbb{N}.$$

$\square$

6.4. Potenzreihen und Stetigkeit

Viele wichtige Funktionen der Analysis werden durch Grenzprozesse wie z.B. Reihen definiert. Deshalb ist es wichtig zu wissen, wann solche Funktionen stetig sind.

Definition 6.35. *Eine Folge (f_n) von Funktionen $f_n\colon D \to \mathbb{C}$ heißt* **punktweise konvergent**, *wenn für alle $z \in D$ die Folge $(f_n(z))$ konvergiert.*

Es sei (f_n) punktweise konvergent. Für $z \in D$ sei

$$f(z) =_{Df} \lim_{n\to\infty} f_n(z).$$

Dann ist $f\colon D \to \mathbb{C}$ eine Funktion.

Bemerkung: Aus der Stetigkeit der Funktionen $f_n\colon D \to \mathbb{C}$ folgt nicht die Stetigkeit von f.

Beispiel: Für $n \in \mathbb{N}$ sei

$$f_n(x) = e^{-nx^2},\ x \in \mathbb{R}.$$

Für $n \neq 0$ ist $\lim_{n\to\infty} f_n(x) = 0$. Andererseits ist $f_n(0) = 1$ für alle $n \in \mathbb{N}$. Daher ist $\lim_{n\to\infty} f_n(0) = 1$. Die Folge (f_n) ist punktweise konvergent. Es sei $f = \lim_{n\to\infty} f_n$. Dann ist

$$f(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0. \end{cases}$$

Definition 6.36. *Für eine Funktion $f\colon D \to \mathbb{C}$ sei*

$$\parallel f \parallel_D =_{Df} \sup\{f(x) \mid x \in D\}.$$

f heißt **beschränkt**, *wenn $\parallel f \parallel_D < \infty$. $\parallel f \parallel_D$ heißt* **Norm** *von f.*

Lemma 6.37. *Für die Norm gilt*

1: $\parallel f \parallel_D = 0 \Leftrightarrow f(z) = 0$ für alle $z \in D$.
2: $\forall c \in \mathbb{C}\colon\ \parallel cf \parallel_D = |c| \parallel f \parallel_D$.
3: $\forall f, g\colon\ \parallel f + g \parallel_D \leq \parallel f \parallel_D + \parallel g \parallel_D$.

Beweis: 1) und 2) folgen unmittelbar aus der Definition der Norm.

Zum Beweis von 3) bemerken wir, daß für alle $z \in D$ gilt:

$$\begin{aligned} |f(z) + g(z)| &\leq |f(z)| + |g(z)| \\ &\leq \| f \|_D + \| g \|_D \,. \end{aligned}$$

Hieraus folgt

$$\| f + g \|_D \leq \| f \|_D + \| g \|_D \,.$$

□

Es sei (f_n) eine Folge von Funktionen. Dann verstehen wir unter der Reihe $\sum_{n=1}^{\infty} f_n$ die Folge der Partialsummen

$$s_n = \sum_{k=1}^{n} f_k.$$

Definition 6.38. *Seien* $f_n \colon D \to \mathbb{C}$, $n \in \mathbb{N}$ *Funktionen. Die Reihe* $\sum_{n=1}^{\infty} f_n$ *heißt* **normal konvergent** *auf* D, *wenn gilt:*

1) $\forall n \in \mathbb{N}\colon$ f_n *ist auf* D *beschränkt.*
2) $\sum_{n=1}^{\infty} \| f_n \|_D < \infty$.

Beispiel: Eine Potenzreihe $P(z) = \sum_{n=0}^{\infty} a_n z^n$ mit Konvergenzradius $R > 0$ konvergiert normal auf jedem Kreis $K_r(0)$ mit $r < R$, denn $\sum_{n=0}^{\infty} |a_n| r^n$ ist konvergent und $|a_n z^n| \leq |a_n| r^n$ für $|z| \leq r$.

Es sei $\sum_{n=0}^{\infty} f_n$ normal konvergent. Dann gilt für alle $z \in D$:

$$\forall n \in \mathbb{N}\colon |f_n(z)| \leq \| f_n \|_D \,.$$

Hieraus folgt

$$\forall z \in D\colon \sum_{n=0}^{\infty} |f_n(z)| < \infty.$$

Sei

$$f(z) = \sum_{n=0}^{\infty} f_n(z), \quad z \in D.$$

Bezeichnung: $f = \sum_{n=0}^{\infty} f_n$.

Wie im Falle von Folgen hat $\sum_{n=0}^{\infty} f_n$ wieder eine doppelte Bedeutung.

Satz 6.39. *Es sei (f_n) eine Folge von Funktionen $f_n\colon D \to \mathbb{C}$ und $\sum_{n=0}^{\infty} f_n$ sei normal konvergent. Es sei $z_0 \in D$ und für alle $n \in \mathbb{N}$ sei f_n stetig in z_0. Dann ist $f = \sum_{n=0}^{\infty} f_n$ stetig in z_0.*

Beweis: Sei $\epsilon > 0$ gegeben. Dann müssen wir zeigen:

$$\exists \delta > 0\ \forall z \in D\colon |z - z_0| < \delta \Rightarrow |f(z) - f(z_0)| < \epsilon.$$

1) Da $\sum_{k=0}^{\infty} \| f_n \|_D < \infty$, existiert zu ϵ ein $N \in \mathbb{N}\colon$ so, daß

$$\sum_{k=N+1}^{\infty} \| f_k \|_D < \frac{\epsilon}{3}.$$

2) Da $f_0, \dots, f_N$ stetig in z_0 sind, folgt aus den Rechenregeln für stetige Funktionen, daß $\sum_{j=0}^{N} f_j$ stetig in z_0 ist. Daraus folgt

$$\exists \delta > 0\colon \left| \sum_{k=0}^{N} f_k(z) - \sum_{k=0}^{N} f_k(z_0) \right| < \frac{\epsilon}{3}$$

für alle $z \in D$ mit $|z - z_0| < \delta$.

Es sei $z \in D$ und $|z - z_0| < \delta$. Dann folgt aus 1) und 2):

$$\begin{aligned} \left| \sum_{k=0}^{\infty} f_k(z) - \sum_{k=0}^{\infty} f_k(z_0) \right| &\le \left| \sum_{k=0}^{N} f_k(z) - \sum_{k=0}^{N} f_k(z_0) \right| + \sum_{k=N+1}^{\infty} |f_k(z)| \\ &\quad + \sum_{k=N+1}^{\infty} |f_k(z_0)| < \frac{\epsilon}{3} + 2 \sum_{k=N+1}^{\infty} \| f_k \|_D = \epsilon. \end{aligned}$$

□

Für $r \in \mathbb{R}_+$ sei $K_r(0) = \{z \in \mathbb{C} \mid |z| < r\}$.

Satz 6.40. *Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist $P\colon K_R(0) \to \mathbb{C}$ stetig.*

Beweis: Es sei $f_n(z) = a_n z^n,\ n \in \mathbb{N},\ z \in \mathbb{C}$. Sei $r < R$. Für $|z| < r$ gilt

$$|f_n(z)| \le |a_n| r^n.$$

Daraus folgt

$$\| f_n \|_{K_r(0)} \le |a_n| r^n.$$

Da $P(r)$ absolut konvergiert, gilt

$$\sum_{n=0}^{\infty} \| f_n \|_{K_r(0)} < \infty.$$

Aus Satz 6.40 folgt, daß P stetig auf $K_r(0)$ ist. Daraus folgt, daß f stetig auf $K_R(0)$ ist.

□

An Hand von Beispielen haben wir gesehen, daß bei punktweiser Konvergenz einer Folge (f_n) stetiger Funktionen $f_n\colon D \to \mathbb{C}$ gegen eine Funktion f, die Grenzfunktion nicht stetig zu sein braucht. Um die Stetigkeit der Grenzfunktion zu sichern, führt man einen stärkeren Konvergenzbegriff für Funktionen ein.

Definition 6.41. *Die Folge (f_n) von Funktionen $f_n\colon D \to \mathbb{C}$* **konvergiert gleichmäßig** *gegen die Funktion $f\colon D \to \mathbb{C}$, wenn gilt*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\colon |f_n(z) - f(z)| < \epsilon \quad \textit{für alle} \quad z \in D.$$

Lemma 6.42. *Die Folge (f_n) von Funktionen $f_n\colon D \to \mathbb{C}$ konvergiert gleichmäßig gegen die Funktion $f\colon D \to \mathbb{C}$, wenn gilt*

$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\colon\ \| f_n - f \|_D < \epsilon.$$

Beweis: Übung.

Satz 6.43. *Sei (f_n) eine Folge stetiger Funktionen $f_n\colon D \to \mathbb{C}$ und (f_n) konvergiere gleichmäßig gegen die Funktion $f\colon D \to \mathbb{C}$. Dann ist f stetig.*

Beweis: Seien $z_0 \in D$ und $\epsilon > 0$ gegeben.

1) Zu $\epsilon > 0$ existiert $n \in \mathbb{N}$:

$$\| f_n - f \| < \frac{\epsilon}{3}.$$

2) Da f_n stetig ist, existiert zu n und z_0 ein $\delta > 0$:

$$\forall z \in D\colon |z - z_0| < \delta \Rightarrow |f_n(z) - f_n(z_0)| < \frac{\epsilon}{3}.$$

Sei $z \in D$ und es gelte $|z - z_0| < \delta$. Daraus folgt

$$\begin{aligned} |f(z) - f(z_0)| &\leq |f(z) - f_n(z)| + |f_n(z) - f_n(z_0)| \\ &\quad + |f_n(z_0) - f(z_0)| \\ &< \| f_n - f \|_D + \frac{\epsilon}{3} + \| f_n - f \|_D < \epsilon. \end{aligned}$$

□

6.5. Grenzwerte und stetige Fortsetzung von Funktionen

In diesem Abschnitt untersuchen wir das Verhalten einer Funktion $f\colon D \to \mathbb{C}$ bei Annäherung von $z \in D$ an einen sogenannten **Häufungspunkt** z_0 von D. Dies führt auf den Begriff des **Grenzwertes einer Funktion** und hängt eng zusammen mit der Frage der **stetigen Fortsetzbarkeit der Funktion** f **in den Punkt** z_0.

Beispiele: 1) Es sei $f(x) = x^2,\ x \in (-1,1)$. Dann hat die Funktion $f\colon (-1,1) \to \mathbb{C}$ eine stetige Fortsetzung in die Punkte -1 und 1.

2) $f(x) = \sin\frac{1}{x},\ x \neq 0$.
a) Sei $x_n = \frac{1}{\pi n}$, $n \in \mathbb{N}$. Dann ist $f(x_n) = \sin(\pi n) = 0$ für alle $n \in \mathbb{N}$. Insbesondere ist $(f(x_n))$ konvergent.
b) Sei $x_n = \frac{2}{(2n-1)\pi}$, $n \in \mathbb{N}$. Dann ist

$$f(x_n) = \sin(n\pi - \frac{\pi}{2}) = (-1)^{n+1},\ n \in \mathbb{N}.$$

Dann ist $(f(x_n))$ unbestimmt divergent. In diesem Beispiel hängt das Verhalten von $f(x)$ bei Annäherung von x an $x = 0$ von der gewählten Folge (x_n) mit $x_n \to 0$ ab. Insbesondere kann f nicht stetig in $x = 0$ fortgesetzt werden.

Definition 6.44. *Sei $D \subset \mathbb{C}$. Ein Punkt $z_0 \in \mathbb{C}$ heißt* **Häufungspunkt von** *D, wenn jede Umgebung von z_0 unendlich viele Punkte aus D enthält.*

Bemerkung: Ein Häufungspunkt einer Menge D kann, aber muß nicht Element der Menge D sein.

Beispiele:

(1) Die Menge der Häufungspunkte von (a,b) ist gleich $[a,b]$.
(2) Die Menge der Häufungspunkte von $\mathbb{Q}$ ist $\mathbb{R}$
(3) Es sei $D = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$. D hat den einzigen Häufungspunkt 0.

Lemma 6.45. *Sei $D \subset \mathbb{C}$ und $z_0 \in \mathbb{C}$. Dann sind folgende Aussagen äquivalent:*

(1) *z_0 ist Häufungspunkt von D.*
(2) *Jede Umgebung von z_0 in $\mathbb{C}$ enthält einen von z_0 verschiedenen Punkt in D.*
(3) *Es gibt eine Folge (z_n) in $D - \{z_0\}$ mit $z_n \to z_0$.*

Beweis: Übung.

Definition 6.46. *Ein Punkt $z_0 \in D$ heißt isolierter Punkt von D, wenn ein $\epsilon > 0$ existiert, so daß $U_\epsilon(z_0) \cap D = \emptyset$.*

Aus Lemma 6.45 folgt, daß ein Punkt $z_0 \in D$ entweder ein Häufungspunkt oder ein isolierter Punkt von D ist. Bei der Untersuchung von Grenzwerten von Funktionen interessieren uns nur Häufungspunkte.

Definition 6.47. *Die Funktion $f\colon D \to \mathbb{C}$ hat im Häufungspunkt z_0 von D den* **Grenzwert** *a, wenn gilt*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall z \in D - \{z_0\}\colon |z - z_0| < \delta \Rightarrow |f(z) - a| < \epsilon.$$

Schreibweise: $a = \lim_{z \to z_0} f(z)$, "$f(z)$ konvergiert gegen a für $z \to z_0$."

Beispiele:

(1) $f\colon \mathbb{C} - \{0\} \to \mathbb{C}$ sei definiert durch

$$f(z) = \frac{e^z - 1}{z}.$$

Dann gilt

$$\lim_{z \to 0} f(z) = 1.$$

2)

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1, \quad \lim_{x \to 0} \frac{\cos(x) - 1}{x} = 0.$$

Übung!

Satz 6.48. *(Folgenkriterium für Grenzwerte)*
Sei z_0 Häufungspunkt von D und sei $a \in \mathbb{C}$. Sei $f\colon D \to \mathbb{C}$ eine Funktion. Dann gilt $\lim_{z \to z_0} f(z) = a$ genau dann, wenn für alle Folgen (z_n) in $D - \{z_0\}$ mit $z_n \xrightarrow[n \to \infty]{} z_0$ gilt: $f(z_n) \xrightarrow[n \to \infty]{} a$.

Beweis: Es sei

$$F\colon D \cup \{z_0\} \to \mathbb{C} \quad \text{definiert durch}$$

$$F(z) = \begin{cases} f(z), & z \in D - \{z_0\}, \\ a, & z = z_0. \end{cases}$$

Dann gilt

$$\lim_{z \to z_0} f(z) = a \Leftrightarrow F \text{ ist stetig in} z_0.$$

Aus Satz 6.12 (Folgenkriterium für stetige Fkt.) folgt die Behauptung.

□

Satz 6.49. *(Cauchy - Krterium für Grenzwerte)*
Sei z_0 ein Häufungspunkt von D und $f\colon D \to \mathbb{C}$ eine Funktion. Dann existiert $\lim_{z\to z_0} f(z)$ genau dann, wenn gilt

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall z, z' \in D - \{z_0\}\colon$$
$$|z - z_0| < \delta \ \wedge\ |z' - z_0| < \delta \Rightarrow |f(z) - f(z')| < \epsilon.$$

Beweis: $\Rightarrow$) Wir nehmen an, daß

$$a = \lim_{z\to z_0} f(z)$$

existiert. Sei $\epsilon > 0$ gegeben. Zu ϵ existiert $\delta > 0$ so, daß

$$\forall z \in D - \{z_0\}\colon |z - z_0| < \delta \Rightarrow |f(z) - a| < \epsilon.$$

Seien $z, z' \in D - \{z_0\}$ und es gelte $|z - z_0| < \delta$ und $|z' - z_0| < \delta$. Dann ist

$$|f(z) - f(z')| \leq |f(z) - a| + |f(z') - a| < 2\epsilon.$$

Also gilt das Cauchy-Kriterium.

$\Leftarrow$) Es gelte das Cauchy-Kriterium in z_0.
Sei (z_n) Folge in $D - \{z_0\}$ mit $z_n \to z_0$. Sei $\epsilon > 0$ gegeben. Zu ϵ existiert $\delta > 0$ so, daß

$$\forall z, z' \in D - \{z_0\}\colon |z - z_0| < \delta \wedge |z' - z_0| < \delta \Rightarrow |f(z) - f(z')| < \epsilon \quad (6.2)$$

Da $z_0 = \lim_{n\to\infty} z_n$, existiert zu ϵ ein $N \in \mathbb{N}$ so, daß gilt:

$$\forall n \geq N\colon\ |z_n - z_0| < \delta.$$

Seien $n, m \geq N$. Dann gilt

$$|z_n - z_0| < \delta,\ |z_m - z_0| < \delta,\ z_n, z_m \neq z_0.$$

Aus (6.2) folgt

$$|f(z_n) - f(z_m)| < \epsilon.$$

Daher ist $(f(z_n))$ eine Cauchy-Folge. Sei

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

Sei (z'_n) eine weitere Folge in $D - \{z_0\}$ mit $z'_n \to z_0$. Dann existiert

$$a' = \lim_{n\to\infty} f(z'_n).$$

Zu zeigen: $\quad a = a'$.

Sei (z_n'') die Folge

$$z_1, z_1', z_2, z_2', z_3, z_3', \ldots$$

Dann gilt $z_n'' \to z_0$ und (z_n'') ist eine Folge in $D - \{z_0\}$. Deshalb existiert

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

Da (z_n) und (z_n') Teilfolgen von (z_n'') sind, folgt

$$a = a'' = a'.$$

□

Rechenregeln für Grenzwerte

Satz 6.50. *Es seien $f\colon D \to \mathbb{C}$ und $g\colon D \to \mathbb{C}$ Funktionen und z_0 sei ein Häufungspunkt von D. Es existiere $a = \lim_{z\to z_0} f(z)$ und $b = \lim_{z\to z_0} g(z)$. Dann gilt*

(1) $\lim_{z\to z_0}(f(z) + g(z)) = a + b$
(2) $\lim_{z\to z_0} f(z) \cdot g(z) = a \cdot b$
(3) *Es sei $b \neq 0$. Dann ist*

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = \frac{a}{b}.$$

Beweis: Folgt aus Satz 6.48 (Folgenkriterium) und den Rechenregeln für Grenzwerten von Folgen.

□

Satz 6.51. *Es seien $D, E \subset \mathbb{C}$ und $f\colon D \to E$, $g\colon E \to \mathbb{C}$ seien Funktionen. Es sei z_0 ein Häufungspunkt von D, es existiere $a = \lim_{z\to z_0} f(z)$ und es sei $a \in E$. Weiter sei g stetig in a. Dann existiert $\lim_{z\to z_0} g(f(z))$ und es gilt:*

$$g(a) = \lim_{z\to z_0} g(f(z)).$$

Beweis: Es sei (z_n) eine Folge in $D - \{z_0\}$ mit $z_n \to z_0$. Dann ist

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

Da g stetig in a ist, folgt aus Satz 6.12:

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

□

6.6. Einseitige und uneigentliche Grenzwerte

6.6.1. Einseitige Grenzwerte

Beispiel: Sei $f\colon \mathbb{R} \to R$ definiert durch

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Für $x \to 0$ hat f keinen Grenzwert. Es existieren aber die links- und rechtsseitigen Grenzwerte.

$$\lim_{\substack{x \to 0 \\ x<0}} f(x) = 0, \quad \lim_{\substack{x \to 0 \\ x>0}} f(x) = 1.$$

Allgemeiner definieren wir einseitige Grenzwerte wie folgt. Es sei $D \subset \mathbb{R}$ und $f\colon D \to \mathbb{C}$ eine Funktion.

Definition 6.52. *Sei x_0 Häufungspunkt von $D \cap (-\infty, x_0)$ (bzw. $D \cap (x_0, \infty)$). f hat* **linksseitig** *(bzw.* **rechtsseitig***) den Grenzwert a, wenn gilt*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in D \cap (x_0 - \delta, x_0)\colon |f(x) - a| < \epsilon$$
$$(\text{bzw. } \forall x \in D \cap (x_0, x_0 + \delta)\colon |f(x) - a| < \epsilon).$$

Bezeichnung:

$$a = \lim_{x \uparrow x_0} f(x) = f(x_0-) \text{ (linksseitig)}$$
$$a = \lim_{x \downarrow x_0} f(x) = f(x_0+) \text{ (rechsseitig)}$$

Sei $x_0 \in D$ und es gelte $f(x_0-) = f(x_0)$ bzw. $f(x_0+) = f(x_0)$. Dann heißt f **linksseitig** bzw. **rechtsseitig** stetig in x_0.

Bemerkung: Für einseitige Grenzwerte gelten die gleichen Rechenregeln wie für gewöhnliche Grenzwerte.

Satz. *Sei $f\colon (a, b) \to \mathbb{R}$ beschränkt und monoton wachsend. Dann exitiert $f(x_0-)$ für jedes $x_0 \in (a, b]$.*

Beweis: Sei $s = \sup\{f(x) \mid x \in (a, x_0)\}$. Sei $\epsilon < 0$ gegeben. Dann exitiert $x \in (a, x_0)$ mit $s - \epsilon < f(x)$. Da f monoton wachsend ist, gilt

$$\forall y \in (x, x_0)\colon s - \epsilon < f(y) \leq s.$$

Hieraus folgt

$$\lim_{x \uparrow x_0} f(x) = s.$$

$\square$

6.6.2. Grenzwerte bei $\pm\infty$

Sei $D \subset \mathbb{R}$ nach oben unbeschränkt. Sei $f\colon D \to \mathbb{C}$ eine Funktion und $a \in \mathbb{C}$. Dann heißt a Grenzwert von f bei ∞, wenn gilt:

$$\forall \epsilon > 0 \; \exists K \in \mathbb{R} \colon |f(x) - a| < \epsilon \text{ für alle } x \in D \text{ mit } x > K.$$

Schreibweise:

$$a = \lim_{x \to \infty} f(x), \quad f(x) \to a \text{ für } x \to \infty.$$

Entsprechend definiert man $\lim_{x \to -\infty} f(x)$.

Beispiel: 1) Sei $r > 0$. Dann ist

$$\lim_{x \to \infty} \frac{1}{x^r} = 0.$$

2)

$$\lim_{x \to \infty} \frac{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0}{a_n x^n} = 1, \quad a_n \neq 0.$$

6.6.3. Uneigentliche Grenzwerte

Sei $D \subset \mathbb{R}$ und $f\colon D \to \mathbb{R}$ eine Fkt.

Definition 6.53. *f hat in $x_0 \in D$ den uneigentlichen Grenzwert ∞ bzw. $-\infty$, wenn zu jedem $K \in \mathbb{R}$ eine Umgebung U von x_0 in $\mathbb{R}$ existiert mit*

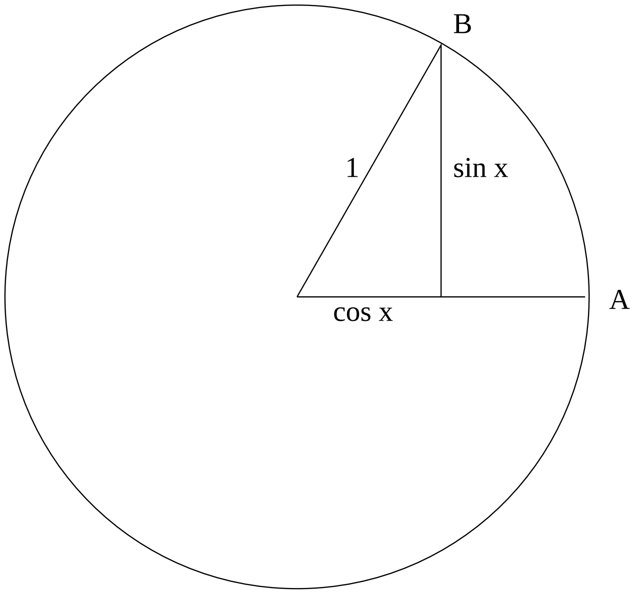
$$f(x) > K \quad \text{bzw.} \quad f(x) < K$$

für alle $x \in D \cap U, \quad x \neq x_0$.
Bez.: $\lim_{x \to x_0} f(x) = \infty$, $\lim_{x \to x_0} f(x) = -\infty$.

6.7. Elementare Funktionen

(1) $e^z = \exp(z) = \sum_{n=1}^{\infty} \frac{z^n}{n!}$. $z \in \mathbb{C}$
(2) $\ln(x) = \exp^{-1}(x)$, $x \in \mathbb{R}_+$.
e^z und $\ln(x)$ sind stetig. (schon gezeigt).
3 $\sin(x)$, $\cos(x)$

Üblicherweise definiert man diese Funktionen wie folgt:



Dabei ist A der Punkt $\langle 1, 0\rangle$ und B der eindeutig bestimmte Punkt auf dem Einheitskreis so, daß x gleich der Länge des positiv orientierten Kreisbogens AB auf dem Einheitskreis ist. Die Schwierigkeit besteht darin, daß wir die Länge eines Kreisbogens bisher nicht definiert haben. Dies wird später ausführlich besprochen. Im Augenblick setzen wir voraus, daß wir wissen, was die Länge eines Kreisbogens ist. Dann können wir $\sin(x)$ und $\cos(x)$ wie üblich definieren. Es sei dann

$$\varphi(x) = \cos(x) + i \sin(x), \quad x \in \mathbb{R}.$$

Diese Funktion hat folgende Eigenschaften:

(1) $\varphi(0) = 1$
(2) $\forall x, y \in \mathbb{R}\colon \varphi(x + y) = \varphi(x)\varphi(y)$
Dies folgt aus den Additionstheoremen für $\cos(x)$ und $\sin(x)$
(3) $\lim_{x\to 0} \frac{\varphi(x)-1}{x} = i$.

Beweis von 3): Es gilt

$$\frac{\varphi(x) - 1}{x} = \frac{\cos(x) - 1}{x} + i\frac{\sin(x)}{x}.$$

Die Behauptung ist dann äquivalent zu

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1, \quad \lim_{x\to 0} \frac{\cos(x) - 1}{x} = 0.$$

Dies kann man wie folgt beweisen: Auf Grund der Definition der Bogenlänge x folgt

$$\sin(x) < x < \tan x, \quad 0 \leq x < \pi/2.$$

Daraus folgt

$$1 < \frac{x}{\sin(x)} < \frac{1}{\cos(x)}.$$

Dies impliziert

$$0 < 1 - \frac{\sin(x)}{x} < 1 - \cos(x) = 2\sin^2\frac{x}{2} < \frac{x^2}{2}.$$

Daraus ergeben sich die Behauptungen über obige Grenzwerte.
Aus 2) und 3) folgt

$$\lim_{x\to 0}\frac{\varphi(x+y)-\varphi(y)}{x} = \varphi(y)\lim_{x\to 0}\frac{\varphi(x)-1}{x} = i\varphi(y).$$

d.h.

$$\varphi'(y) = i\varphi(y), \quad y \in \mathbb{R}.$$

Daraus erhalten wir

$$\begin{aligned}\frac{d}{dx}\left(\varphi(x)e^{-ix}\right) &= \varphi'(x)e^{-ix} - ie^{-ix}\varphi(x)\\ &= i\left(\varphi(x)e^{-ix} - e^{-ix}\varphi(x)\right) = 0.\end{aligned}$$

Daraus folgt, daß eine Konstante $c \in \mathbb{C}$ existiert so, daß

$$\varphi(x)e^{-ix} = c.$$

Aus 1) folgt $1 = \varphi(0) = c$. Dies ergibt

$$\varphi(x) = e^{ix}.$$

Dami haben wir gezeigt, daß die Differentialgleichung

$$\varphi'(x) = i\varphi(x), \quad \varphi(0) = 1,$$

als einzige Lösung e^{ix} hat.

Satz. *(Eulersche Formel)*

$$e^{ix} = \cos(x) + i\sin(x), \quad x \in \mathbb{R}.$$

Damit können wir $\cos(x)$ und $\sin(x)$ auch wie folgt charakterisieren:

$$\begin{aligned}\cos(x) &= \mathrm{Re}(e^{ix}) = \frac{e^{ix}+e^{-ix}}{2}\\ \sin(x) &= \mathrm{Im}(e^{ix}) = \frac{e^{ix}+e^{-ix}}{2i}.\end{aligned}$$

Aus der Potenzreihenentwicklung der Exponentialfunktion ergibt sich

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$
$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Für $z \in \mathbb{C}$ definieren wir dann $\cos(z)$ und $\sin(z)$ durch die entsprechenden Potenzreihen:

$$\sin(z) = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = \frac{e^{iz} - e^{-iz}}{2i}$$
$$\cos(z) = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} = \frac{e^{iz} + e^{-iz}}{2}.$$

Der Konvergenzradius dieser Reihen ist ∞. Weiter sei

$$\tan(z) = \frac{\sin(z)}{\cos(z)} = \frac{1}{i} \frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}}$$

für alle $z \in \mathbb{C}$ mit $\cos(z) \neq 0$. Entsprechend können wir die Umkehrfunktion einführen.

$$\arctan = \tan^{-1} \colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$
$$\arccos = \cos^{-1} \colon [-1, 1] \to [0, \pi]$$
$$\arcsin = \sin^{-1} \colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}].$$

Schließlich führen wir noch die hyperbolischen Funktionen für komplexe z ein:

$$\cosh(z) = \frac{e^z + e^{-z}}{2} = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}$$
$$\sinh(z) = \frac{e^z - e^{-z}}{2} = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!}$$
$$\tanh(z) = \frac{\sinh(z)}{\cosh(z)}$$
$$\coth(z) = \frac{\cosh(z)}{\tanh(z)}.$$

7. Differentialrechnung

Die von Leibniz und Newton begründete Differential- und Integralrechnung bildet den Kern der gesamten Analysis.

7.1. Ableitung einer Funktion

Der Begriff der Ableitung einer Funktion ist von fundamentaler Bedeutung für die Differentialrechnung. Er spielt in den Naturwissenschaften, z.B. in der Physik, eine wichtige Rolle bei der Formulierung von Gesetzmäßigkeiten.

Beispiel: 2. *Newtonsches Grundgesetz*

$$m\frac{d^2x}{dt^2} = F.$$

Definition 7.1. *Sei $I = (a,b)$. Eine Funktion $f\colon I \to \mathbb{C}$ heißt* **differenzierbar im Punkte** $x_0 \in I$, *wenn der Grenzwert*

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Der Limes heißt dann **Ableitung** *oder* **Differentialquotient von** f **in** x_0.

Bezeichnung: $f'(x_0)$, $\frac{df}{dx}(x_0)$, $Df(x_0)$. Es gilt

$$f'(x_0) = \lim_{n\to 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

Geometrische Interpretation

$$L_h(x) = f(x_0) + \frac{f(x_0+h) - f(x_0)}{h}(x - x_0),\ x \in \mathbb{R},$$

ist die Sekante durch $P_0 = (x_0, f(x_0))$ und $P_1 = (x_0 + h,\ f(x_0 + h))$.
Für $h \to 0$ konvergiert $L_h(x)$ gegen die Tangente

$$t(x) = f(x_0) + f'(x_0)(x - x_0).$$

Beispiele: 1) $s(t)$ Weg eines Massepunktes als Funktion der Zeit.
Die mittlere Geschwindigkeit im Zeitintervall $[t_0, t_0 + h]$ ist $\frac{s(t_0+h)-s(t_0)}{h}$.

$\frac{ds}{dt}(t_0) = \dot{s}(t_0)$ ist die Geschwindigkeit zum Zeitpunkt t_0.

2) $f(x) = c,\ x \in I.\ f' = 0.$

3) Sei $f\colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = x^n,\ n \in \mathbb{N}$.
Behauptung: $f'(x) = nx^{n-1}$

Beweis:

$$\frac{(x_0+h)^n - x_0^n}{n} = nx_0^{n-1} + \binom{n}{2} hx_0^{n-2} + \cdots + h^{n-1} \xrightarrow[n\to 0]{} nx_0^{n-1}.$$

4) Es sei $f(x) = e^{cx}$, $c \in \mathbb{C}$. Dann ist $f' = cf$, denn aus

$$\frac{e^{c(x_0+h)} - e^{cx_0}}{h} = e^{cx_0}\frac{e^{ch}-1}{h} = e^{cx_0}\left(c + \frac{hc^2}{2} + \cdots\right)$$

folgt, daß der Differenzenquotient für $h \to 0$ gegen ce^{cx_0} konvergiert. Daher ist

$$\frac{d}{dx}e^{cx} = ce^{cx}.$$

Satz 7.2. *Sei $f\colon I \to \mathbb{C}$ in $x_0 \in I$ differenzierbar. Dann ist f stetig in x_0.*

Beweis:

$$\begin{aligned}\lim_{x\to x_0} f(x) &= \lim_{x\to x_0}\left(f(x_0) + \frac{f(x)-f(x_0)}{x-x_0}(x-x_0)\right)\\ &= f(x_0) + f'(x_0)\lim_{x\to x_0}(x-x_0) = f(x_0).\end{aligned}$$

Daraus folgt, daß f stetig in x_0 ist.

□

Äquivalente Definitionen der Differenzierbarkeit

Satz 7.3. *Es sei $f\colon I \to \mathbb{C}$ eine Funktion und $x_0 \in I$. Dann sind folgende Aussagen äquivalent:*

(1) *f ist in x_0 differenzierbar.*

(2) *Es existiert $\lambda \in \mathbb{C}$ und eine Funktion $\rho\colon T \to \mathbb{C}$ mit $\rho(x_0) = 0$ und $\lim_{x\to x_0}\frac{\rho(x)}{x-x_0} = 0$ so, daß*

$$f(x) = f(x_0) + \lambda(x-x_0) + \rho(x).$$

Es ist dann $f'(x_0) = \lambda$.

(3) *Es gibt eine in x_0 stetige Funktion $r\colon I \to \mathbb{C}$ mit*

$$f(x) = f(x_0) + r(x)(x-x_0).$$

Es ist $f'(x_0) = r(x_0)$.

Beweis: 1)⇒ 2) Es sei $\lambda = f'(x_0)$ und $\rho(x) = f(x) - f(x_0) - \lambda(x - x_0)$. Dann gilt $\rho(x_0) = 0$ und

$$\lim_{x\to x_0}\frac{\rho(x)}{x-x_0} = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} - \lambda = f'(x_0) - \lambda = 0.$$

Nach Definition von ρ ist

$$f(x) = f(x_0) + \lambda(x - x_0) + \rho(x).$$

2) $\Rightarrow$ 3) Wir setzen

$$r(x) = \begin{cases} \lambda + \frac{\rho(x)}{x - x_0}, & x \neq x_0; \\ \lambda, & x = x_0. \end{cases}$$

Dann gilt

$$\lim_{x \to x_0} r(x) = \lim_{x \to x_0} \frac{\rho(x)}{x - x_0} + \lambda = \lambda = r(x_0).$$

$\Rightarrow r$ ist stetig in x_0. Weiter ist auf Grund der Definition von r:

$$f(x) = f(x_0) + r(x)(x - x_0).$$

2) $\Rightarrow$ 1) Es gilt

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} r(x) = r(x_0),$$

d.h. der Grenzwert des Differentialquotienten existiert.

□

Bemerkung: 2) kann wie folgt geschrieben werden

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \rho(x).$$

Dies bedeutet, daß f in x_0 eine Taylorentwicklung 1. Ordnung hat. Sei $L(h) = f'(x_0)h$, $h \in \mathbb{R}$. Dann ist $L\colon \mathbb{R} \to \mathbb{C}$ eine lineare Abbildung und es gilt

$$f(x_0 + h) = f(x_0) + L(h) + \rho(h), \quad \lim_{h \to 0} \frac{\rho(h)}{h} = 0.$$

Defintion. *Eine Funktion $f\colon I \to \mathbb{C}$ heißt* **differenzierbar**, *wenn f in jedem Punkt $x \in I$ differenzierbar ist.*

Bisher haben wir nur offenen Intervalle als Definitionsbereich von f betrachtet. Es sei $f\colon [a, b) \to \mathbb{R}$ eine Funktion. Dann heißt f **rechtsseitig differenzierbar in** a, wenn der Grenzwert

$$f'_+(a) = \lim_{x \downarrow a} \frac{f(x) - f(a)}{x - a}$$

existiert. $f'_+(a)$ heißt dann **rechtsseitige Ableitung von** f **in** a. Entsprechend heißt $f\colon (a,b] \to \mathbb{C}$ **linksseitig differenzierbar in** b, wenn

$$f'_-(b) = \lim_{x \uparrow b} \frac{f(x) - f(b)}{x - b}$$

existiert. $f'_-(b)$ heißt dann **linksseitige Ableitung von** f **in** b. Wir können jetzt beliebige Intervalle (a,b), $(a,b]$, $[a,b)$, $[a,b]$ als Definitionsbereich von f zulassen.

Extremalwerte und Ableitungen

Sei $f\colon I \to \mathbb{R}$ eine Funktion.

Definition.

1) *f hat in $x_0 \in I$ ein* **globales Maximum** *(bzw.* **Minimum***), wenn $\forall x \in I\colon f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$).*
2) *f hat in $x_0 \in I$ ein* **lokales Maximum** *(bzw.* **Minimum***), wenn*

$$\exists \epsilon > 0\ \forall x \in I \cap (x_0 - \epsilon, x_0 + \epsilon)\colon f(x) \leq f(x_0)\, (\text{bzw. } f(x) \geq f(x_0)).$$

Satz 7.4. *Es sei $f\colon (a,b) \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und f habe in x_0 ein lokales Extremum. Dann ist $f'(x_0) = 0$.*

Beweis: Wir betrachten den Fall, daß f in x_0 ein lokales Maximum hat. Sei $\epsilon > 0$ so, daß $(x_0 - \epsilon, x_0 + \epsilon) \subset I$ und $f \restriction (x_0 - \epsilon, x_0 + \epsilon)$ ein Maximum in x_0 hat.
Für $x_0 + \epsilon > x > x_0$ gilt dann

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \Rightarrow f'(x_0) \leq 0.$$

Für $x_0 - \epsilon < x < x_0$ ist andererseits

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0 \Rightarrow f'(x_0) \geq 0 \Rightarrow f'(x_0) = 0.$$

□

Bemerkungen:

(1) Der Satz 7.4. ist falsch für Randpunkte.
Beispiel: $f(x) = x,\ x \in [0,1]$. Dann sind 0 und 1 Extrema von f, aber $f'(x) = 1 \neq 0$.
(2) Aus $f'(x_0)$ folgt nicht, daß x_0 ein Extremum von f ist.
Beispiel: $f(x) = x^3,\ f'(0) = 0$, aber 0 ist kein Extremwert.

(3) Sei $-\infty < a < b < \infty$ und sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann nimmt f sein Minimum und Maximum auf $[a,b]$ an. Für Extrema von f gibt es dann folgende Möglichkeiten:
i) Die Randpunkte a und b;
ii) Die Punkte $x \in (a,b)$, in denen f diffenzierbar ist und für die $f'(x) = 0$ gilt.
iii) Punkte $x \in (a,b)$, in denen f nicht differenzierbar ist.

7.2. Ableitungsregeln

Satz 7.5. *Es seien $f, g\colon I \to \mathbb{C}$ in $x_0 \in I$ differenzierbar. Dann sind $f+g$, $f \cdot g$, und falls $g(x_0) \neq 0$, auch f/g in x_0 differenzierbar, und es gilt:*

(1) $(f+g)'(x_0) = f'(x_0) + g'(x_0)$;
(2) $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$;
(3) $\left(\dfrac{f}{g}\right)(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$.

Beweis: Diese Rechenregeln erhält man durch Anwendung der Rechenregeln für Grenzwerte.

1)

$$\begin{aligned}\lim_{h\to 0} \frac{(f+g)(x_0+h) - (f+g)(x_0)}{h} &= \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \\ &\quad + \lim_{h\to 0} \frac{g(x_0+h) - g(x_0)}{h} \\ &= f'(x_0) + g'(x_0).\end{aligned}$$

2)

$$\begin{aligned}(f \cdot g)'(x_0) &= \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h) - f(x_0)g(x_0)}{h} \\ &= \lim_{h\to 0} \frac{1}{h}\{f(x_0+h)g(x_0+h) - f(x_0)g(x_0+h) \\ &\qquad + f(x_0)g(x_0+h) - f(x_0)g(x_0)\} \\ &= \lim_{h\to 0} \left(\frac{f(x_0+h) - f(x_0)}{h} g(x_0+h)\right) \\ &\qquad + \lim_{h\to 0} \left(f(x_0)\frac{g(x_0+h) - g(x_0)}{h}\right) \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0).\end{aligned}$$

3)

$$\begin{aligned}
\left(\frac{f}{g}\right)'(x_0) &= \lim_{h\to 0}\frac{1}{h}\left\{\frac{f(x_0+h)}{g(x_0+h)}-\frac{f(x_0)}{g(x_0)}\right\} \\
&= \lim_{h\to 0}\frac{1}{g(x_0+h)g(x_0)}\left\{\frac{f(x_0+h)g(x_0)-f(x_0)g(x_0+h)}{h}\right\} \\
&= \lim_{h\to 0}\frac{1}{g(x_0+h)g(x_0)}\left\{\frac{f(x_0+h)-f(x_0)}{h}g(x_0)\right. \\
&\qquad\qquad\left. -f(x_0)\frac{g(x_0+h)-g(x_0)}{h}\right\} \\
&= \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g^2(x_0)}.
\end{aligned}$$

$\square$

Beispiele: 1) $\cos'(x) = -\sin(x)$, $\sin'(x) = \cos(x)$.
Es gilt:

$$e^{ix} = \cos(x) + i\sin(x) \quad \text{und} \quad (e^{ix})' = ie^{ix}.$$

Daraus folgt

$$\begin{aligned}
(\cos(x)+i\sin(x))' = \cos'(x)+i\sin'(x) &= i(\cos x + i\sin x) \\
&= -\sin x + i\cos x.
\end{aligned}$$

Dies ergibt die behaupteten Gleichungen für die Ableitungen.
2)

$$(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

3)

$$\frac{d}{dx}x^n = \frac{d}{dx}(x\cdots x) = nx^{n-1}.$$

Satz 7.6. *(Kettenregel) Seien $I, J \subset \mathbb{R}$ Intervalle, $f\colon I \to J$ und $g\colon J \to \mathbb{C}$ Funktionen. Weiter sei f in x_0 und g in $y_0 = f(x_0)$ differenzierbar. Dann ist $g\circ f\colon I \to \mathbb{C}$ in x_0 differenzierbar und es gilt:*

$$(g\circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis: Nach Satz 7.3, 3), existieren Funktionen $p\colon I \to J$, $q\colon J \to \mathbb{C}$, die in x_0 bzw. $y_0 = f(x_0)$ stetig sind so, daß

a) $f(x) - f(x_0) = p(x)(x - x_0)$, $x \in I$.
b) $g(y) - g(y_0) = q(y)(y - y_0)$, $y \in J$.

und $p(x_0) = f'(x_0)$, $q(y_0) = g'(y_0)$. Daraus folgt

$$g(f(x)) - g(f(x_0)) = q(f(x))(f(x) - f(x_0)) = q(f(x))p(x)(x - x_0).$$

Die Funktion $q(f(x))p(x)$ ist in x_0 stetig und es gilt

$$q(f(x_0))p(x_0) = g'(f(x_0))f'(x_0).$$

Aus Satz 7.3, 3) folgt die Behauptung.

□

Satz 7.7. *(Ableitung der Umkehrfunktion) Sei $f\colon I \to \mathbb{R}$ streng monoton, differenzierbar in $x_0 \in I$ und es sei $f'(x_0) \neq 0$. Dann ist die Umkehrfunktion $g\colon f(I) \to I$ in $y_0 = f(x_0)$ differenzierbar und es gilt*

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$

Beweis: Aus Satz 7.3, 3) folgt: Es existiert eine in x_0 stetige Funktion $p\colon I \to \mathbb{R}$ mit $p(x_0) = f'(x_0)$ und

$$f(x) - f(x_0) = p(x)(x - x_0). \tag{7.1}$$

Da f streng monoton ist, folgt $p(x) \neq 0$ für alle $x \neq x_0$. Weiter ist $p(x_0) = f'(x_0) \neq 0$. Damit gilt $p(x) \neq 0$ für alle $x \in I$. Sei $y = f(x)$, $y_0 = f(x_0)$. Dann ist $x = g(y)$, $x_0 = g(y_0)$. Aus (7.1) folgt

$$y - y_0 = p(g(y))(g(y) - g(y_0)).$$

Daraus erhalten wir

$$g(y) - g(y_0) = \frac{1}{p(g(y))}(y - y_0).$$

Da f in x_0 differenzierbar ist, folgt aus Satz 7.2, daß f in x_0 stegig ist. Daraus folgt, daß g in $y_0 = f(x_0)$ stetig. Da p stetig in $x_0 = g(y_0)$ ist und $p(x) \neq 0$ für alle $x \in I$, folgt, daß $p(g(y))^{-1}$ stetig in y_0 ist und

$$\frac{1}{p(g(y_0))} = \frac{1}{f'(x_0)}.$$

Aus Satz 7.3 folgt die Behauptung.

□

Beispiele: 1) $\ln(x) = \exp^{-1}(x)$.

$$(\ln(x))' = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$

2) Sei $x > 0, a \in \mathbb{R}$. Dann ist $x^a = e^{a \ln x}$.
Anwendung der Kettenregel ergibt

$$(x^a)' = (e^{a \ln x})' = e^{a \ln x} \cdot a(\ln x)' = ax^a \frac{1}{x} = ax^{a-1}.$$

7.3. Mittelwertsätze

Ein wichtiges Hilfsmittel bei der Untersuchung von lokalen Eigenschaften von differenzierbaren Funktionen ist der Mittelwertsatz. Er besagt, daß für eine stetige Funktion $f\colon [a,b] \to \mathbb{R}$, die auf dem offenen Intervall differenzierbar ist, immer ein Punkt $\xi \in (a,b)$ existiert so, daß die Tangente an den Graphen der Funktion f im Punkte $(\xi, f(\xi))$ den gleichen Anstieg hat wie die Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$, d.h., es gilt

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

Ein Spezialfall ist der **Satz von Rolle**.

Satz 7.8. *(Satz von Rolle). Sei* $a < b$, $a, b \in \mathbb{R}$. *Sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und differenzierbar auf* (a,b). *Weiter sei* $f(a) = f(b)$. *Dann existiert* $x \in (a,b)$ *mit* $f'(x) = 0$.

Beweis: Da $f\colon [a,b] \to \mathbb{R}$ stetig ist, folgt aus Satz 6.31, daß $x_1, x_2 \in [a,b]$ existieren so, daß gilt:

$$\forall x \in [a,b]\colon\ m = f(x_1) \le f(x) \le f(x_2) = M,$$

d.h., $m = \min\{f(x) \mid x \in [a,b]\}$ und $M = \max\{f(x) \mid x \in [a,b]\}$. Wir unterscheiden zwei Fälle:

Fall 1: $m = M$.

Dann ist $f(x) = m$ für alle $x \in [a,b]$. Daraus folgt $f'(x) = 0$ für alle $x \in (a,b)$.

Fall 2: $m < M$.

O.B.d.A. können wir annehmen, daß $f(a) < M$ gilt. Dann ist auch $f(b) < M$. Daraus folgt, daß ein $x \in (a,b)$ existiert mit $f(x) = M$. Aus Satz 7.4 erhalten wir $f'(x) = 0$.

□

Satz 7.9. *(Mittelwertsatz)*
Sei $f\colon [a,b] \to \mathbb{R}$ *stetig und differenzierbar auf* (a,b). *Dann existiert* $x \in (a,b)$ *mit*

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

Beweis: Wir reduzieren diesen Fall auf den Satz von Rolle. Sei $F\colon [a,b] \to \mathbb{R}$ definiert durch

$$F(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a).$$

Dann ist F stetig auf $[a,b]$, differenzierbar auf (a,b) und es gilt

$$F(a) = f(a) = F(b).$$

Aus dem Satz von Rolle folgt, daß $x \in (a,b)$ existiert mit $F'(x) = 0$. Daraus erhalten wir

$$0 = F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

□

Anwendungen des Mittelwertsatzes

Satz 7.10. *(Monotoniekriterium) Es sei $f\colon (a,b) \to \mathbb{R}$ differenzierbar. Dann gilt:*

(1) *$f'(x) > 0$ für alle $x \in (a,b) \Rightarrow f$ streng monoton wachsend.*
(2) *$f'(x) \geq 0$ für alle $x \in (a,b) \Leftrightarrow f$ monoton wachsend.*
(3) *$f'(x) < 0$ für alle $x \in (a,b) \Rightarrow f$ streng monoton fallend.*
(4) *$f'(x) \leq 0$ für alle $x \in (a,b) \Rightarrow f$ monoton fallend.*

Ist f stetig auf $[a,b)$ oder $(a,b]$, so gelten die rechts stehenden Aussagen auf $[a,b)$ bzw. $(a,b]$.

Beweis: 1) Seien $x_1, x_2 \in (a,b)$ mit $x_2 > x_1$. Aus Satz 7.9 folgt, daß ein $x \in (x_1, x_2)$ existiert mit

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x).$$

Nach Voraussetzung ist $f'(x) > 0$. Daraus folgt $f(x_2) > f(x_1)$.

Der Beweis von 2) - 4) ist analog.

□

Korollar 7.11. *Es sei $f\colon (a,b) \to \mathbb{R}$ differenzierbar und in $x_0 \in (a,b)$ gelte $f'(x_0) = 0$. Dann hat f in x_0 ein*

(1) *Minimum, wenn $f' \leq 0$ in (a, x_0) und $f' \geq 0$ in (x_0, b);*
(2) *Maximum, wenn $f' \geq 0$ in (a, x_0) und $f' \leq 0$ in (x_0, b).*

Beweis: Aus Satz 7.10 folgt, daß f in $(a, x_0]$ monoton fallend und in $[x_0, b)$monoton wachsend ist. Daraus folgt

$$f(x) \geq f(x_0) \text{ für alle } a < x \leq x_0,$$
$$f(x_0) \leq f(x) \text{ für alle } x_0 \leq x < b.$$

Dies bedeutet, daß f in x_0 ein Minimum hat.

Der Beweis von 2) ist analog.

□

Satz 7.12. *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b). Dann gilt:*

$$f'(x) = 0 \text{ für alle } x \in (a,b) \Leftrightarrow f \text{ konstant.}$$

Beweis: $\Leftarrow$) klar

$\Rightarrow$) Seien $x_1, x_2 \in [a,b]$ und sei $x_1 < x_2$. Aus dem Mittelwertsatz folgt, daß ein $x_0 \in (x_1, x_2)$ existiert mit

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_0).$$

Da $f'(x_0) = 0$, ergibt sich daraus $f(x_1) = f(x_2)$.

□

Satz 7.13. *(Schrankensatz)*
Es sei $f\colon I \to \mathbb{C}$ differenzierbar und es existiere $L \in \mathbb{R}_+$ so, daß $|f'(x)| \leq L$ für alle $x \in I$. Dann gilt

$$\forall x_1, x_2 \in I\colon |f(x_1) - f(x_2)| \leq L|x_1 - x_2|,$$

d.h. f ist Lippschitzstetig.

Beweis: Seien $x_1, x_2 \in I$ mit $x_1 < x_2$. Aus dem Mittelwertsatz folgt, daß ein $x_0 \in (x_1, x_2)$ existiert mit

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(x_0)$$

Hieraus folgt

$$|f(x_1) - f(x_2)| \leq |x_1 - x_2||f'(x_0)| \leq L|x_1 - x_2|.$$

□

Defintion 7.14. *Eine Funktion $f\colon I \to \mathbb{R}$ heißt stetig differenzierbar, wenn f differenzierbar und f' stetig ist.*

Korollar 7.15. *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig differenzierbar. Dann ist f Lipschitz-stetig.*

Beweis: Da $f'\colon [a,b] \to \mathbb{R}$ stetig ist, folgt aus Korollar 6.32, daß ein $L > 0$ existiert so, daß $|f'(x)| \leq L$ für alle $x \in [a,b]$. Die Behauptung folgt dann aus Satz 7.13.

□

Beispiele: 1. Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit

$$f' = cf \quad \text{und} \quad f(0) = 1.$$

Dann ist $f(x) = e^{cx}$.

Beweis: Die Funktion $h : \mathbb{R} \to \mathbb{R}$ sei definiert durch $h = f(x)e^{-cx}$. Die Anwendung der Produktregel ergibt

$$h'(x) = f'(x)e^{-cx} - cf(x)e^{-cx} = cf(x)e^{-cx} - cf(x)e^{-cx} = 0.$$

Aus Satz 7.12 folgt, daß ein $C \in \mathbb{R}$ existiert mit $h(x) = C$ für alle $x \in \mathbb{R}$. Daraus folgt

$$f(x) = Ce^{cx}.$$

Wenn wir in dieser Gleichung $x = 0$ setzen, ergibt sich auf Grund der Voraussetzung $C = f(0) = 1$.

2. Methode der kleinsten Quadrate

Seien $a_1, \dots, a_n \in \mathbb{R}$ gegeben. Gesucht ist $a \in \mathbb{R}$, so daß $(a - a_1)^2 + \cdots + (a - a_n)^2$ minimal wird.

Um dieses Problem zu lösen, betrachten wir die Funktion

$$f(x) = (x - a_1)^2 + \cdots + (x - a_n)^2.$$

Dann ist

$$f'(x) = 2(x - a_1) + \cdots + 2(x - a_n).$$

Es sei $f'(x_0) = 0$. Dann folgt aus dieser Gleichung

$$2nx_0 = 2(a_1 + \cdots + a_n),$$

d.h.,

$$x_0 = \frac{a_1 + \cdots + a_n}{n}.$$

Weiter gilt

$$x < \frac{a_1 + \cdots + a_n}{n} \Rightarrow f'(x) < 0$$
$$x > \frac{a_1 + \cdots + a_n}{n} \Rightarrow f'(x) > 0$$

Daraus folgt, daß $\frac{a_1 + \cdots + a_n}{n}$ das eindeutig bestimmte Minimum ist.

Satz 7.16. *(Verallgemeinerter Mittelwertsatz)*
Seien $f, g \colon [a, b] \to \mathbb{R}$ stetig und differenzierbar in (a, b). Weiter sei $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist $g(b) \neq g(a)$, und es existiert $x_0 \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}$$

Beweis: 1. Wir nehmen an, daß $g(b) = g(a)$ gilt. Aus dem Satz von Rolle folgt, daß ein $x \in (a, b)$ existiert mit $g'(x) = 0$. Dies ist ein

Widerspruch zur Voraussetzung, daß $g'(x) \neq 0$ für alle $x \in (a, b)$. Daher ist $g(b) \neq g(a)$.
2. Sei

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

Dann gilt $F(a) = F(b)$ und $F\colon [a, b] \to \mathbb{R}$ ist stetig und differenzierbar in (a, b). Aus dem Satz von Rolle folgt, daß ein $x_0 \in (a, b)$ existiert mit $F'(x_0) = 0$. Daraus folgt

$$f'(x_0) = \frac{f(b) - f(a)}{g(b) - g(a)} g'(x_0).$$

□

Als Anwendung ergibt sich

Satz 7.17. *(Regel von L'Hospital)*
Für $-\infty \leq a < b \leq \infty$ seien $f, g\colon (a, b) \to \mathbb{R}$ differenzierbar und es gelte $g'(x) \neq 0$ für alle $x \in (a, b)$. In jedem der beiden Fälle

(1) $f(x) \to 0$ *und* $g(x) \to 0$ *für* $x \downarrow a$
(2) $f(x) \to \infty$ *und* $g(x) \to \infty$ *für* $x \downarrow a$ *gilt:*

Wenn $\lim\limits_{x \downarrow a} \dfrac{f'(x)}{g'(x)}$ *existiert, so existiert* $\lim\limits_{x \downarrow a} \dfrac{f(x)}{g(x)}$ *und*

$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = \lim_{x \downarrow a} \frac{f'(x)}{g'(x)}.$$

Entsprechende Aussagen gelten für $x \uparrow b$, $x \to \infty$ und $x \to -\infty$.

Beweis: 1. Wir können f und g zu stetigen Funktionen $f, g\colon [a, b) \to \mathbb{R}$ fortsetzen mit $f(a) = g(a) = 0$. Aus dem verallgemeinerten Mittelwertsatz folgt: Für alle $x \in (a, b)$ existiert $\xi \in (a, x)$ so, daß

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

Aus $x \to a$ folgt $\xi \to a$ und damit die Behauptung.
2. Sei $\lim_{x \downarrow a} \frac{f'(x)}{g'(x)} = A$. Zu $\epsilon > 0$ wählen wir $\delta > 0$ so, daß

$$\left| \frac{f'(l)}{g'(l)} - A \right| < \epsilon$$

für alle $t \in (a, a + \delta)$. Nach dem verallgemeinerten Mittelwertsatz gilt für alle $x, y \in (a, a + \delta)$ mit $x \neq y$:

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - A \right| < \epsilon.$$

Nun ist

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{1-g(y)/g(x)}{1-f(y)/f(x)}.$$

Nach Voraussetzung gilt: $f(x) \to \infty$ und $g(x) \to \infty$ für $x \downarrow a$. Daraus folgt, daß für alle $y \in (a, a+\delta)$ gilt:

$$\lim_{x \downarrow a} \frac{1-g(y)/g(x)}{1-f(y)/f(x)} = 1.$$

Hieraus folgt, daß ein $\delta_1 > 0$ existiert so, daß für alle $x \in (a, a+\delta_1)$ gilt:

$$\left| \frac{f(x)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)} \right| < 2\epsilon$$

Für alle $x \in (a, a+\min(\delta, \delta_1)$ gilt dann

$$\left| \frac{f(x)}{g(x)} - A \right| < 2\epsilon.$$

Dies beweist die Behauptung im Falle b). Der Grenzprozess $x \to \infty$ kann auf den bewiesenen Grenzprozess $y \downarrow 0$ durch die Substituion $x = 1/y$ zurückgeführt werden.

$\square$

Beispiele: 1) Berechnung von $\lim_{x \downarrow 0} x \ln x$. Die Anwendung von Satz 7.17 ergibt

$$\lim_{x \downarrow 0} x \ln x = \lim_{x \downarrow 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x \downarrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0.$$

2)

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \cos x = 1.$$

$$\lim_{x \to 0} \frac{1-\cos x}{x} 0 \lim_{x \to 0} \sin x = 0.$$

4)

$$\lim_{x \downarrow 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right) = \lim_{x \downarrow 0} \frac{x - \sin x}{x \cdot \sin x} = \lim_{x \downarrow 0} \frac{1-\cos x}{x \cos x + \sin x}$$

$$= \lim_{x \downarrow 0} \frac{\sin x}{2\cos x - x \sin x} = 0.$$

Beim letzten Beispiel haben wir die l'Hospitalsche Regel zweimal angewendet.

7.4. Höhere Ableitungen

Definition 7.18. *Sei $f\colon I \to \mathbb{C}$ differenzierbar. Ist die Funktion $f'\colon I \to \mathbb{C}$ in $x_0 \in I$ differenzierbar, so definiert man die* **zweite Ableitung** *von f in x_0 als*

$$f''(x_0) = (f')'(x_0).$$

Rekursiv definiert man die n-te Ableitung $f^{(n)}(x_0)$ von f in x_0, falls die $(n-1)$-te Ableitung $f^{(n-1)}$ auf I erklärt ist und in x_0 differenzierbar ist durch

$$f^{(n)}(x_0) = \left(f^{(n-1)}\right)'(x_0).$$

*Wir sagen f ist n-***mal stetig differenzierbar***, wenn $f^{(n)}\colon I \to \mathbb{R}$ existiert und stetig ist. Existiert für jedes $n \in \mathbb{N}$ die Ableitung n-ter Ordnung, so heißt f* **beliebig oft differenzierbar***.*

Bezeichnungen: $f^{(n)}(x_0)$, $\frac{d^n f}{dx^n}(x_0)$, $D^n f(x_0)$.

$$\begin{aligned} C^0(I) &= \{f\colon I \to \mathbb{C} \mid f \text{ stetig}\} \\ C^n(I) &= \{f\colon I \to \mathbb{C} \mid f\ n-\text{mal stetig differenzierbar}\} \\ C^\infty(I) &= \{f\colon I \to \mathbb{C} \mid f \text{ ist } \infty-\text{oft differenzierbar}\}. \end{aligned}$$

Die Räume $C^0(I)$, $C^n(I)$ und $C^\infty(I)$ sind komplexe Vektorräume unendlicher Dimension.

Beispiele: 1) Sei

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Dann ist $f \in C^\infty(\mathbb{R})$.

Beweis: Für $x \neq 0$ ist $f(x)$ differenzierbar und es gilt

$$f'(x) = -\frac{2}{x^3} e^{-1/x^2}.$$

Weiter ist für $x \neq 0$:

$$\frac{e^{-1/x^2}}{|x|} = \frac{1}{|x|} \frac{1}{1 + \frac{1}{x^2} + \cdots} < \frac{x^2}{|x|} = |x|.$$

Daraus folgt, daß $f'(0)$ existiert und

$$f'(0) = \lim_{x \to 0} x^{-1} e^{-1/x^2} = 0.$$

Ebenso zeigt man

$$\lim_{x\to 0} f'(x) = 0.$$

Damit ist $f'\colon \mathbb{R} \to \mathbb{R}$ stetig. Durch Induktion zeigt man, daß $f^{(n)}$ für alle n existiert und stetig ist. Insbesondere folgt

$$f^{(n)}(0) = 0 \text{ für alle } n \in \mathbb{N}.$$

□

Eine besonders wichtige Rolle spielen die zweiten Ableitungen. Z.B. sei für einen Bewegungsvorgang der Weg $s(t)$ als Funktion der Zeit gegeben. Dann ist die zweite Ableitung $s''(t)$ die Beschleunigung zum Zeitpunkt t.

Eine andere wichtige Rolle spielen die zweiten Ableitungen bei der Charakterisierung von Extremalwerten.

Satz 7.19. *Es sei $f\colon (a,b) \to \mathbb{R}$ zweimal stetig differenzierbar, und sei $x_0 \in (a,b)$. Es sei $f'(x_0) = 0$. Dann gilt:*

(1) *$f''(x_0) > 0 \Rightarrow f$ hat ein lokales Minimum in x_0.*
(2) *$f''(x_0) < 0 \Rightarrow f$ hat ein lokales Maximum in x_0.*

Beweis: 1) Da f'' stetig ist, existiert $\delta > 0$, so daß $(x_0 - \delta, x_0 + \delta) \subset (a,b)$ und $f''(x) > 0$ für alle $x \in (x_0 - \delta, x_0 + \delta)$.

Da $f'' = (f')'(x) > 0$, folgt aus Satz 7.10, daß f' streng monoton wachsend auf $(x_0 - \delta, x_0 + \delta)$ ist. Da $f'(x_0) = 0$ gilt, folgt

$$f'(x) < 0 \text{ für alle } x \in (x_0 - \delta, x_0)$$
$$f'(x) > 0 \text{ für alle } x \in (x_0, x_0 + \delta).$$

Aus Korollar 7.11 folgt, daß f ein lokales Minimum in x_0 hat. Analog beweist man 2).

□

Der Satz 7.19 liefert uns ein hinreichendes Kriterium für die Existenz eines lokalen Extremwertes einer Funktion.

7.5. Konvexität

Der Begriff der Konvexität einer Funktion spielt eine wichtige Rolle in der Analysis.

Definition 7.20. *Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f\colon I \to \mathbb{R}$ heißt* **konvex**, *wenn für jedes Tripel $x_0, x, x_1 \in I$ mit $x_0 < x < x_1$ gilt*

$$(K) \qquad f(x) \le \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1).$$

Wir nennen f **strikt konvex**, *falls für alle solche Tripel die Ungleichung (K) streng ist. Die Funktion f heißt* **konkav**, *wenn $-f$ konvex ist.*

Die Ungleichung (K) kann wie folgt umformuliert werden. Für $x \in (x_0, x_1)$ sei

$$\lambda = \frac{x - x_0}{x_1 - x_0}.$$

Dann ist $\lambda \in (0, 1)$ und

$$x = \lambda x_1 + (1 - \lambda)x_0.$$

Umgekehrt liegt jeder solcher Punkt in (x_0, x_1). Die Gleichung (K) ist dann äquivalent zu der folgenden Aussage

Für alle $x_0, x_1 \in I$ mit $x_0 \neq x_1$ und alle $\lambda \in (0, 1)$ gilt:

$$f(\lambda x_1 + (1 - \lambda)x_0) \leq \lambda f(x_1) + (1 - \lambda)f(x_0).$$

Lemma 7.21. *Es sei $f\colon I \to \mathbb{R}$ eine Funktion. Dann sind folgende Aussagen äquivalent:*

(1) *f ist konvex*

(2) *Für alle Tripel $x_0, x, x_1 \in I$ mit $x_0 < x < x_1$ gilt:*

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x)}{x_1 - x}.$$

(3) *Für alle Tripel $x_0, x, x_1 \in I$ gilt*

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_1) - f(x)}{x_1 - x}.$$

Beweis: 1) $\Rightarrow$ 3). Es sei $x_0, x, x_1 \in I$ ein Tripel mit $x_0 < x < x_1$. Da f konvex ist, folgt aus der Ungleichung (K)

$$(x_1 - x_0)f(x) \leq (x_1 - x)f(x_0) + (x - x_0)f(x_1).$$

Daraus folgt

$$(x_1 - x_0)f(x) \leq ((x_1 - x_0) + (x_0 - x))f(x_0) + (x - x_0)f(x_1).$$

Die Gleichung ist äuquivalent zu

$$(x_1 - x_0)(f(x) - f(x_0)) \leq (x - x_0)(f(x_1) - f(x_0)).$$

und diese wiederum zu

$$\frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Dies ergibt die linke Ungleichung in 3). Entsprechend folgt aus (K)

$$(x_1 - x_0)f(x) \le (x_1 - x)f(x_0) + ((x - x_1) + (x_1 - x_0))f(x_1)$$

Umordnung ergibt

$$(x_1 - x)(f(x_1) - f(x_0)) \le (x_1 - x_0)(f(x_1) - f(x))$$

und daraus folgt die rechte Ungleichung von 3).

3) $\Rightarrow$ 2): klar

2) $\Rightarrow$ 1) Sei $x_0, x, x_1 \in I$ ein Tripel mit $x_0 < x < x_1$. Aus der Ungleichung von 2) folgt

$$(x_1 - x)(f(x) - f(x_0)) \le (f(x_1) - f(x_0))(x - x_0).$$

Daraus folgt

$$(x_1 - x_0)f(x) \le (x - x_0)f(x_1) + (x_1 - x)f(x_0).$$

und dies ergibt (K).

□

Für differenzierbare Funktionen hängt die Konvexität mit dem Wachstum der Ableitung zusammen. Und zwar gilt

Korollar 7.22. *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b). Dann ist f genau dann konvex, wenn f' monoton wachsend ist.*

Beweis: Sei f konvex. Für $x_0, x_1 \in (a,b)$ mit $x_0 \le x_1$ folgt aus 3):

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \le \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

und

$$f'(x_1) = \lim_{x \to x_0} \frac{f(x_1) - f(x)}{x_1 - x} \ge \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Daraus folgt

$$f'(x_0) \le f'(x_1).$$

Sei umgekehrt f' monoton wachsend auf (a,b) und sei $x_0, x, x_1 \in [a,b]$ ein Tripel mit $x_0 < x < x_1$. Aus dem Mittelwertsatz, angewendet auf $f\colon [x_0, x] \to \mathbb{R}$ bzw. $f\colon [x, x_1] \to \mathbb{R}$, folgt die Existenz von $y_0 \in (x_0, x)$ und $y_1 \in (x, x_1)$ so, daß

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(y_0) \quad \text{und} \quad \frac{f(x_1) - f(x)}{x_1 - x} = f'(y_1).$$

Da $y < y_1$ und f' monoton wachsend ist, folgt

$$\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(x_1)-f(x)}{x_1-x}$$

und aus Lemma 7.21 die Ungleichung (K).

□

Bemerkung: Wenn f' streng monoton wachsend ist, so ist f streng konvex.

Satz 7.23. *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) zweimal differenzierbar. Dann gilt*

(1) *f ist genau dann konvex, wenn $f'' \geq 0$.*
(2) *Wenn $f'' > 0$, so ist f streng konvex.*

Beweis:

(1) Aus Korollar 7.22 folgt:

$$f \;\text{ konvex} \Leftrightarrow f' \;\text{ monoton wachsend.}$$

Aus Lemma 7.10 folgt:
f' ist monoton wachsend $\Leftrightarrow f'' \geq 0$ auf (a,b).

(2) Es sei $f'' > 0$ auf (a,b). Aus Lemma 7.10 folgt, daß f' auf (a,b) streng monoton wächst. Aus Korollar 7.22 folgt dann, daß f auf (a,b) streng konvex ist.

□

Defintion 7.24. *Sei $f\colon (a,b) \to \mathbb{R}$ stetig. Der Punkt $(x_0, f(x_0))$ heißt* **Wendepunkt** *von f, wenn es Intervalle (x_1, x_0) und (x_0, x_2) gibt so, daß eine der beiden folgenden Bedingungen erfüllt ist:*

(1) *f ist auf (x_1, x_0) konvex und auf (x_0, x_2) konkav.*
(2) *f ist auf (x_1, x_0) konkav und auf (x_0, x_2) konvex.*

Beispiel: Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ sei definiert durch $f(x) = x^3,\ x \in \mathbb{R}$. Dann ist $f''(x) = 6x$. Aus Satz 7.23 folgt, daß f auf $(-\infty, 0)$ konkav und auf $(0, \infty)$ konvex ist. Daher ist 0 ein Wendepunkt von f.

Satz 7.25. *Es sei $f\colon (a,b) \to \mathbb{R}$ dreimal stetig differenzierbar und sei $x_0 \in (a,b)$. Wenn $f''(x_0) = 0$ und $f^{(3)}(x_0) \neq 0$, so hat f einen Wendepunkt in x_0.*

Beweis: Es sei $f^{(3)}(x_0) > 0$. Da $f^{(3)}$ auf (a,b) stetig ist, existiert $\delta > 0$ so, daß $f^{(3)}(x) > 0$ für alle $x \in (x_0 - \delta, x_0 + \delta)$. Aus Lemma 7.10 folgt, daß f'' auf $(x_0 - \delta, x_0 + \delta)$ streng monoton wachsend ist. Da $f''(x_0) = 0$, ist, $f''(x) < 0$ für alle $x \in (x_0 - \delta, x_0)$ und $f''(x) > 0$ für alle $x \in (x_0, x_0 + \delta)$. Aus Satz 7.23 folgt, daß f'' auf $(x_0 - \delta, x_0)$

konkav und auf $(x_0, x_0 + \delta)$ konvex ist. Analog behandelt man den Fall $f^{(3)}(x_0) < 0$.

□

7.6. Differentation von Reihen

Im Abschnitt 6.5 haben wir gezeigt, daß eine normal konvergente Reihe stetiger Funktionen eine stetige Funktion definiert. Eine normal konvergente Reihe differenzierbarer Funktionen ist jedoch im allgemeinen nicht differenzierbar.

Beispiel: Für $n \in \mathbb{N}$ sei die Funktion $f_n\colon \mathbb{R} \to \mathbb{R}$ definiert durch

$$f_n(x) = \sqrt{x^2 + \frac{1}{n}}.$$

Dann ist f_n für alle $n \in \mathbb{N}$ eine differenzierbare Funktion. Es sei

$$f(x) = f_1(x) + \sum_{n=2}^{\infty} (f_n(x) - f_{n-1}(x)). \tag{7.2}$$

Für die n-te Partialsumme s_n gilt $s_n = f_n$. Daher ist

$$f(x) = \lim_{n\to\infty} \sqrt{x^2 + \frac{1}{n}} = \sqrt{x^2} = |x|,\ x \in \mathbb{R}.$$

Weiter ist

$$\begin{aligned}
0 &\le 2\sqrt{\frac{1}{n+1}} \left(\sqrt{x^2 + \frac{1}{n}} - \sqrt{x^2 + \frac{1}{n+1}} \right) \\
&\le \left(\sqrt{x^2 + \frac{1}{n}} - \sqrt{x^2 + \frac{1}{n+1}} \right) \left(\sqrt{x^2 + \frac{1}{n}} + \sqrt{x^2 + \frac{1}{n+1}} \right) \\
&= \frac{1}{(n+1)n}.
\end{aligned}$$

Daraus folgt

$$\left| \sqrt{x^2 + \frac{1}{n+1}} - \sqrt{x^2 + \frac{1}{n}} \right| \le \frac{1}{n\sqrt{n+1}}$$

für alle $n \in \mathbb{N}$. Dies wiederum ergibt

$$\sum_{n=2}^{\infty} \| f_n - f_{n-1} \| \le \sum_{n=2}^{\infty} \frac{1}{n\sqrt{n+1}} \le \sum_{n=1}^{\infty} n^{-3/2} < \infty.$$

Die Reihe (7.2) ist also normal konvergent, die Grenzfunktion $f(x) = |x|$ ist aber in $x = 0$ nicht differenzierbar.

Der folgende Satz ist ein hinreichendes Kriterium für die Differenzierbarkeit der Grenzfunktion.

Satz 7.26. *Sei $f_n \colon I \to \mathbb{C}$ eine Folge differenzierbarer Funktionen und es gelte:*

(1) $\sum_{n=1}^{\infty} f_n$ *konvergiert punktweise auf I.*
(2) $\sum_{n=1}^{\infty} f_n'$ *konvergiert normal auf I.*

Dann ist die Funktion $f = \sum_{n=1}^{\infty} f_n$ auf I differenzierbar und es gilt

$$f' = \sum_{n=1}^{\infty} f_n'.$$

Beweis: Es sei $\epsilon > 0$. Auf Grund der Voraussetzung 2. existiert zu ϵ ein $N \in \mathbb{N}$ so, daß

$$\sum_{n=N+1}^{\infty} \| f_n \| < \frac{\epsilon}{3}. \tag{7.3}$$

Seien $x_0, x \in I,\ x > x_0$. Aus dem Mittelwertsatz folgt, daß für jedes $n \in \mathbb{N}$ ein $\xi_n \in [x_0, x]$ existiert mit

$$\frac{f_n(x) - f_n(x_0)}{x - x_0} = f_n'(\xi_n)$$

Daraus folgt

$$\left| \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| \leq \| f_n' \|$$

für alle $n \in \mathbb{N}$. Aus (7.3) erhalten wir daher

$$\sum_{n=N+1}^{\infty} \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| < \frac{\epsilon}{3}. \tag{7.4}$$

Aus der Voraussetzung 1) zusammen mit (7.3) und (7.4) folgt

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - \sum_{n=1}^{\infty} f_n'(x_0) \right| \leq \sum_{n=1}^{N} \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - f_n'(x_0) \right|$$
$$+ \sum_{n=N+1}^{\infty} \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| + \sum_{n=N+1}^{\infty} \| f_n' \| \leq \sum_{n=1}^{N} \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - f_n'(x_0) \right| + \frac{2}{3}\epsilon.$$

Da jede Funktion $f_n \colon I \to \mathbb{C}$ differenzierbar ist, existiert $\delta > 0$ so, daß für alle x mit $|x - x_0| < \delta$ und $x \neq x_0$ gilt

$$\sum_{n=1}^{N} \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - f_n'(x_0) \right| < \frac{\epsilon}{3}.$$

Daraus folgt, daß f in x_0 differenzierbar ist und

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

□

Korollar 7.27. *Es sei* $P(x) = \sum_{n=1}^{\infty} a_n x^n$ *eine Potenzreihe mit Konvergenzradius* $R > 0$. *Dann ist die Funktion* $P\colon (-R, R) \to \mathbb{C}$ *differenzierba und es gilt*

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

Beweis: Da $\sqrt[n]{n} \to 1$ für $n \to \infty$, gilt

$$\limsup \sqrt[n]{n|a_n|} = \limsup \sqrt[n]{|a_n|} = R.$$

Damit hat die abgeleitete Reihe

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

ebenfalls den Konvergenzradius R. Es sei $r < R$. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} n a_n x^n$ normal im Intervall $(-r, r)$. Aus Satz 7.26 folgt, daß

$$P'(x) = \sum_{n=1}^{\infty} n a_n x^n$$

für alle $x \in (-r, r)$. Da $r < R$ beliebig gewählt war, gilt die Gleichung für alle $x \in (-R, R)$.

□

8. Integralrechnung

8.1. Das Integral von Treppenfunktionen

Definition 8.1. *Eine Funktion* $\varphi\colon [a, b] \to \mathbb{C}$ *heißt* **Treppenfunktion**, *wenn es eine Zerlegung (oder Unterteilung)*

$$a = x_0 < x_1 < \cdots < x_n = b$$

von $[a, b]$ *gibt so, daß* φ *auf jedem Intervall* (x_{k-1}, x_k), $k = 1, \ldots, n$, *konstant ist.*

Die Werte der Funktion φ in den Teilungspunkten $x_0, \ldots, x_n$ unterliegen keiner Einschränkung. Den Vektorraum aller Treppenfunktionen auf $[a, b]$ bezeichnen wir mit $T[a, b]$.

Definition 8.2. *Es sei $\varphi \in T[a,b]$ und es sei $Z = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a,b]$ so, daß $\varphi(x) = c_k \in \mathbb{C}$ für $x \in (x_{k-1}, x_k)$. Dann definieren wir*

$$\int_a^b \varphi(x)dx =_{Df} \sum_{k=1}^{n} c_k(x_k - x_{k-1}).$$

Lemma 8.3. *$\int_a^b \varphi(x)dx$ ist unabhängig von der Zerlegung Z.*

Beweis: Für eine Zerlegung $Z = \{x_0, \dots, x_n\}$ mit $\varphi(x) = c_k$ für $x \in (x_{k-1}, x_k)$ sei

$$I_Z(\varphi) = \sum_{k=1}^{n} c_k(x_k - x_{k-1}).$$

Es sei Z' eine weitere Zerlegung von $[a,b]$ so, daß φ konstant ist auf den offenen Teilintervallen. Wir bilden eine neue Zerlegung $Z'' = Z \cup Z'$. Dies ist die gemeinsame Verfeinerung von Z und Z'. Wir zeigen dann, daß $I_{Z''}(\varphi) = I_Z(\varphi)$ gilt. Die Zerlegung Z'' entsteht aus Z durch Hinzunahme endlich vieler neuer Teilungspunkte. Durch Induktion läßt sich die Behauptung zurückführen auf den Fall, daß Z'' aus Z durch Hinzunahme eines neuen Teilungspunktes $y \in (x_{k-1}, x_k)$ entsteht. Da

$$c_k(x_k - x_{k-1}) = c_k(x_k - y) + c_k(y - x_{k-1}),$$

folgt unmittelbar aus der Defintion, daß

$$I_{Z''}(\varphi) = I_Z(\varphi)$$

gilt. Analog zeigt man $I_{Z''}(\varphi) = I_{Z'}(\varphi)$.

□

Durch Lemma 8.3 wird die Bezeichnung $\int_a^b \varphi(x)dx$ für $I_Z(\varphi)$ gerechtfertigt. Zur Abkürzung schreiben wir auch

$$\int_a^b \varphi dx.$$

für $\int_a^b \varphi(x)dx$.
Für eine Funktion $f\colon [a,b] \to \mathbb{C}$ sei

$$\| f \| = \sup\{|f(x)| \mid x \in [a,b]\}$$

die Suprenumsnorm.

Lemma 8.4. *Es seien $\varphi, \psi \in [a,b]$ und $\alpha, \beta \in \mathbb{C}$. Dann gilt*

(1) $\int_a^b (\alpha\varphi + \beta\psi)dx = \alpha \int_a^b \varphi dx + \beta \int_a^b \psi dx$ *(Linearität)*

(2) $|\int_a^b \varphi dx| \leq \int_a^b |\varphi| dx \leq (b-a) \parallel \varphi \parallel$ *(Beschränktheit)*

(3) *Falls φ, ψ reell sind und $\varphi \leq \psi$, so gilt*

$$\int_a^b \varphi dx \leq \int_a^b \psi \, dx \qquad \textit{(Monotonie)}$$

Beweis: Wir wählen eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ so, daß φ und ψ auf jedem Teilintervall (x_{k-1}, x_k), $k = 1, \ldots, n$, konstant sind. Dann folgen 1) und 3) unmittelbar aus den entsprechenden Eigenschaften endlicher Summen komplexer Zahlen. Ebenso folgt die 1. Ungleichung in 2) aus der Dreiecksungleichung. Die zweite Ungleichung folgt aus

$$\max\{|c_k| \mid 1 \leq k \leq n\} \leq \parallel \varphi \parallel .$$

□

8.2. Das Integral von Regelfunktionen

Definition 8.5. *Sei $I \subset \mathbb{R}$ ein Intervall mit Anfangspunkt a und dem Endpunkt b. Eine Funktion $f\colon I \to \mathbb{C}$ heißt* **Regelfunktion**, *wenn gilt:*

(1) *In jedem Punkt $x_0 \in (a,b)$ hat f einen linksseitigen und einen rechtsseitigen Grenzwert.*

(2) *Falls $a \in I$, so hat f in a einen rechtsseitigen Grenzwert, und falls $b \in I$, so hat f in b einen linksseitigen Grenzwert. Sei*

$$R(I) = \{f\colon I \to \mathbb{C} \mid f \text{ Regelfunktion}\}$$

Beispiele:

(1) Treppenfunktionen

(2) Stetige Funktionen

(3) monotone Funktionen.

Satz 8.6. *Es sei $f\colon [a,b] \to \mathbb{C}$ eine Funktion. Dann gilt*

$$f \in R[a,b] \Leftrightarrow \forall \epsilon > 0 \; \exists \varphi \in [a,b]\colon \; \parallel f - \varphi \parallel < \epsilon.$$

Beweis: $\Rightarrow$) Sei $f \in R[a,b]$.

Annahme: Es existiert $\epsilon > 0$ so, daß f keine ϵ- approximierende Treppenfunktion hat.

Wir konstruieren eine Intervallschachtelung $I_n = [a_n, b_n]$, $n \in \mathbb{N}$, mit folgenden Eigenschaften

(1) $|I_n| = (b-a)2^{-n+1}$ für alle $n \in \mathbb{N}$.

(2) Für alle $\varphi \in T(I_n)$: $\parallel \varphi - f \restriction I_n \parallel \geq \epsilon$.
Die Konstruktion erfolgt rekursiv.
1. Schritt: Sei $I_1 = [a, b]$
n+1. Schritt: Sei I_n bereits konstruiert und sei m der Mittelpunkt von I_n. Sei I_{n+1} eines der Teilintervalle $[a_n, m]$ und $[m, b_n]$ so, daß auf diesem Teilintervall keine ϵ-approximierende Treppenfunktion zu f existiert. Sei

$$\{x_0\} = \bigcap_{n \in \mathbb{N}} I_n.$$

Wir betrachten den Fall $x_0 \in (a, b)$. Die Fälle $x_0 = a$ bzw. $x_0 = b$ behandelt man analog. Sei

$$c_l = \lim_{x \uparrow x_0} f(x), \quad c_r = \lim_{x \downarrow x_0} f(x).$$

Zu $\epsilon > 0$ existiert $\delta > 0$ so, daß

$$|f(x) - c_l| < \epsilon$$

für alle $x \in [x_0 - \delta, x_0)$ und

$$|f(x) - c_r| < \epsilon \text{ für alle } x \in (x_0, x_0 + \delta].$$

Zu $\delta > 0$ existiert $n \in \mathbb{N}$ mit $I_n \subset [x_0 - \delta, x_0 + \delta]$. Es sei $\varphi \colon I_n \to \mathbb{C}$ definiert durch

$$\varphi(x) = \begin{cases} c_l, & x < x_0 \\ f(x_0), & x = x_0 \\ c_r, & x > x_0. \end{cases}$$

Dann gilt auf I_n: $\parallel \varphi - f \restriction I_n \parallel < \epsilon$. Dies ist ein Widerspruch zur Konstruktion von I_n. Also ist die Annahme falsch.
$\Leftarrow$) Zu jedem $\epsilon > 0$ existiere $\varphi \in T[a, b]$ mit $\parallel \varphi - f \parallel < \epsilon$.
Es sei $x_0 \in (a, b]$. Wir zeigen, daß f in x_0 einen linksseitigen Grenzwert hat. Sei $\epsilon > 0$ gegeben und sei $\varphi \in T[a, b]$ mit $\parallel \varphi - f \parallel < \frac{\epsilon}{2}$. Wir wählen $x_1 < x_0$ so, daß φ auf (x_1, x_0) konstant ist. Seien $x, x' \in (x_1, x_0)$. Dann gilt

$$\begin{aligned} |f(x) - f(x')| &\leq |f(x) - \varphi(x)| + |\varphi(x) - \varphi(x')| + |\varphi(x') - f(x')| \\ &\leq 2 \parallel f - \varphi \parallel < \epsilon. \end{aligned}$$

Aus dem Cauchy-Kriterium für Grenzwerte folgt, daß f in x_0 einen linksseitigen Grenzwert besitzt.

□

Lemma 8.7. *Es sei $f \in R[a,b]$ und es sei (φ_n) eine Folge in $T[a,b]$ mit $\| \varphi_n - f \| \to 0$ für $n \to \infty$. Dann gilt*

(1) *Die Folge $(\int_a^b \varphi_n(x)dx)$ konvergiert*
(2) *$\lim_{n\to\infty} \int_a^b \varphi_n(x)dx$ ist unabhängig von der Folge (φ_n) in $T[a,b]$ mit $\| \varphi_n - f \| \to 0$.*

Beweis:

(1) Für $n, m \in \mathbb{N}$ gilt

$$\left| \int_a^b \varphi_n dx - \int_a^b \varphi_m dx \right| = \left| \int_a^b (\varphi_n - \varphi_m) dx \right|$$
$$\leq (b-a) \| \varphi_n - \varphi_m \| \leq (b-a)(\| \varphi_n - f \| + \| \varphi_m - f \|).$$

Daraus folgt, daß $(\int_a^b \varphi_n dx)$ eine Cauchy-Folge ist.

(2) Es sei (ψ_n) eine weitere Folge in $T[a,b]$ mit $\| \psi_n - f \| \to 0$. Dann gilt

$$\left| \int_a^b \varphi_n dx - \int_a^b \psi_n \, dx \right| \leq (b-a) \| \varphi_n - \psi_n \|$$
$$\leq (b-a)(\| \varphi_n - f \| + \| f - \psi_n \|) \xrightarrow[n\to\infty]{} 0.$$

□

Definition 8.8. *Es sei $f \in R[a,b]$ und (φ_n) sei eine Folge in $T[a,b]$ mit $\| \varphi_n - f \| \to 0$ für $n \to \infty$. Dann sei*

$$\int_a^b f(x)dx =_{Df} \lim_{n\to\infty} \int_a^b \varphi_n(x)dx.$$

Aus Lemma 8.7. folgt, daß der Grenzwert existiert und unabhängig von der gewählten Folge (φ_n) ist.

Eigenschaften des Integrals

Satz 8.9. *Seien $f, g \in R[a,b]$ und $\alpha, \beta, \in \mathbb{C}$. Dann gilt:*

(1) $\int_a^b (\alpha f + \beta g)dx = \alpha \int_a^b f\, dx + \beta \int_a^b g dx$ *(Linearität)*
(2) $| \int_a^b f\, dx| \leq \int_a^b |f|\, dx \leq (b-a) \| f \|$ *(Beschränktheit)*
(3) *Falls f, g reell sind und $f \leq g$, so ist*

$$\int_a^b f\, dx \leq \int_a^b g\, dx. \qquad \textit{(Monotonie)}$$

Beweis: Folgt durch Grenzübergang aus Lemma 8.4.

Satz 8.10. *Seien $a < b < c$ und $f \in R[a,b]$. Dann gilt:*

$$\int_a^c f\,dx = \int_a^b f\,dx + \int_b^c f\,dx.$$

Beweis: Der Satz ist klar für Treppenfunktionen. Der Grenzübergang ergibt den Beweis für Regelfunktionen.

□

Bemerkung: Wir setzen $\int_a^a f(x)\,dx = 0$. Für $a > b$: $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$.

Mittelwertsatz der Integralrechnung

Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann heißt

$$\frac{1}{b-a}\int_a^b f(x)\,dx$$

der Mittelwert von f.

Satz 8.11. *Sei $f\colon [a,b] \to R$ stetig und $p \in R[a,b]$ mit $p \geq 0$. Dann existiert ein $\xi \in [a,b]$ so, daß*

$$f(\xi)\int_a^b p(x)\,dx = \int_a^b p(x)f(x)\,dx.$$

Beweis: Es sei $m = \min f$ und $M = \max f$. Aus $p \geq 0$ folgt

$$mp \leq fp \leq Mp.$$

Aus Satz 8.9, 3) folgt

$$m\int_a^b p\,dx \leq \int_a^b fp\,dx \leq M\int_a^b p\,dx.$$

Hieraus folgt, daß ein $\lambda \in [m,M]$ existiert mit

$$\lambda\int_a^b p\,dx = \int_a^b fp\,dx.$$

Da f stegig ist, folgt aus dem Zwischenwertsatz die Existenz von $\xi \in [a,b]$ mit $f(\xi) = \lambda$. Damit ist

$$f(\xi)\int_a^b pdx = \int_a^b fp\,dx.$$

□

Korollar 8.12. *Sei $f\colon [a,b] \to R$ stetig. Dann existiert $\xi \in [a,b]$ mit*

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

8.3. Hauptsatz der Differential- und Integralrechnung

Auf der Grundlage der Definition des Integrals ist es im allgemeinen sehr schwierig, das Integral einer Funktion zu berechnnen. Der Hauptsatz gibt uns eine effektive Methode zur Berechnung von Integralen in die Hand. Er beruht auf der Tatsache, daß die Integration als Umkehrung der Differentation angesehen werden kann.

Definition 8.13. *Es sei $f\colon [a,b] \to \mathbb{C}$ eine gegebene Funktion. Eine differenzierbare Funktion $F\colon [a,b] \to \mathbb{C}$ heißt* **Stammfunktion** *von f, wenn gilt:*

$$F' = f.$$

Bemerkung: Manchmal fordert man nur $F'(x) = f(x)$ für fast alle $x \in [a,b]$.

Lemma 8.14. *Es seien F, G Stammfunktionen von f. Dann ist $F-G$ konstant.*

Beweis: $(F-G)' = F' - G' = f - f = 0$.

□

Hauptsatz 8.15. *Es sei $f\colon I \to \mathbb{C}$ stetig.*

(1) *Sei $x_0 \in I$. Dann ist die Funktion*

$$F(x) = \int_{x_0}^x f(y)\;dy$$

eine Stammfunktion von f.

(2) *Es sei umgekehrt G eine beliebige Stammfunktion zu f. Für alle $a, b \in I$ gilt dann*

$$\int_a^b f(x)\,dx = G(b) - G(a).$$

Beweis: Es sei $F(x) = \int_{x_0}^x f(y)\,dy$.

(1) Nach dem Mittelwertsatz (Korollar 8.12) existiert zu $h \in \mathbb{R}$ mit $|h| \ll 1$ ein $\vartheta_h,\ 0 \le \vartheta_h \le 1$, mit

$$F(x+h) - F(x) = \int_x^{x+h} f(y)\,dy = f(x + \vartheta_h \cdot h)h.$$

Da $\lim_{h\to 0} f(x+\vartheta_h h) = f(x)$, existiert

$$F'(x) = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h},$$

und es gilt

$$F'(x) = f(x),\ x \in I.$$

(2) Sei G eine Stammfunktion zu f. Nach Lemma 8.14 existiert $C \in \mathbb{C}$ so, daß

$$G(x) = \int_{x_0}^{x} f(y)\,dy + C.$$

Daraus folgt

$$\begin{aligned}\int_a^b f(y)\,dy &= \int_{x_0}^b f(y)\,dy + \int_a^{x_0} f(y)\,dy \\ &= \int_{x_0}^b f(y)\,dy - \int_{x_0}^a f(y)\,dy = G(b) - G(a).\end{aligned}$$

□

Bemerkungen:

(1) Den Hauptsatz kann man allegemeiner für Regelfunktionen formulieren.
(2) Manchmal werden die Aussagen 1) und 2) auch getrennt als 1. und 2. Hauptsatz bezeichnet.
(3) Die 2. Aussage des Hauptsatzes besagt, daß es zur Berechnung eines Integrals ausreicht, eine Stammfunktion zu kennen.

Definition. *Es sei $f\colon I \to \mathbb{C}$ stetig und F sei eine Stammfunktion zu f. Dann nennt man F auch unbestimmtes Integral von f.*

Bezeichnung: $\int f(x)\,dx$.

Grundintegrale

Zur Integration von Funktionen ist es wichtig, die unbestimmten Integrale elementarer Funktionen zu kennen. Dabei ließt man die Formeln für die Ableitungen der elementaren Funktionen in umgekehrter Richtung.

Beispiele:

(1) $x > 0,\ \alpha \neq 1.$
$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1}$
(2) $\int \frac{dx}{x} = \ln|x|$
(3) $\int \frac{dx}{1+x^2} = \arctan x$

(4) $\int \sin x \, dx = -\cos x$
(5) $\int \cos x \, dx = \sin x$

Integrationstechniken

Bei der Differentation hat man einfache Regeln, nach denen die Ableitung elementarer Funktionen berechnet werden kann.
Beim Integrieren ist die Situation prinzipiell anders: Es existieren elemetare Funktionen, die nicht ″elementar inegriert″ werden können, d.h. die keiner elemenatare Stammfunmktion besitzen. Zur Integration benutzt man eine Kombination verschiedener Methoden wie z.B.

(1) Liste von Grundintegralen,
(2) partielle Integration
(3) Variablensubstitution
(4) Integraltafeln
(5) Computerprogramme wie z.B. Mathematica, Maple,...

Partielle Integration

Satz 8.16. *Sind $f, g\colon I \to \mathbb{C}$ stetig differenzierbar, so gilt*

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx$$

bzw.

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\,\Big|_a^b - \int_a^b f'(x)g(x)\,dx$$

für alle $a, b \in I$, $a < b$.

Beweis: Folgt aus Produktregel und Hauptsatz. Übung!
Beispiel:

$$\int \sin x \cos x \; dx = \sin^2 x - \int \cos x \sin x \, dx \Rightarrow \int \sin x \cos x = \frac{1}{2}\sin^2 x.$$

Substitutionsregel

Satz 8.17. *Sei $f\colon I \to \mathbb{C}$ stetig und $\varphi\colon [a,b] \to I$ stetig differenzierbar. Dann gilt*

$$\int_a^b f(\varphi(u))\varphi'(u)\; du = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

Beweis: Folgt aus Kettenregel und Hauptsatz. Übung!

Beispiel: $\int \sqrt{1-x^2}\,dx$, Substitution: $x = \cos t$

$$= - \textit{f} \sin^2 t\, dt \underset{\text{partiell}}{=} \frac{1}{2}(\sin t \cos t - t)$$
$$= \frac{1}{2}(x\sqrt{1-x^2} - \arccos x) = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x).$$

Damit kann man den Flächeninhalt F des Einheitskreises berechnen.

$$F = 2\int_{-1}^{1} \sqrt{1-x^2}\,dx = (x\sqrt{1-x^2} + \arcsin x)\;|_{-1}^{1} = \arcsin(1) - \arcsin(-1) = \pi.$$

8.4. Integration rationaler Funktionen

Es seien

$$p(z) = \sum_{i=0}^{n} a_i z^i \quad \text{und} \quad q(z) = \sum_{j=0}^{m} b_j z^j$$

Polynome in z.

Definition. *Die Funktion* $R(z) = \dfrac{p(z)}{q(z)}$ *nennt man* **rationale Funktion**.

Euklidischer Algorithmus für Polynome

Seien $p(z) = \sum_{i=0}^{n} a_i z^i$ und $q(z) = \sum_{j=0}^{m} b_j z^j$ Polynome mit $a_n, b_m \neq 0$. Dann gibt es eindeutig bestimmte Polynome h und r mit

$$1)\ p = qh + r \quad 2)\ \text{Grad}\,(r) < \text{Grad}\,(q).$$

Daraus folgt, daß jede rationale Funktion p/q geschrieben werden kann als

$$\frac{p}{q} = h + \frac{r}{q}$$

wobei h ein Polyom ist und Grad $(r) <$ Grad (q). Die Funktion r/q heißt **echte rationale Funktion**.

Partialbruchzerlegung

Satz 8.18. *Es sei $R = p/q$ eine rationale Funktion mit Grad $(p) <$ Grad (q) so, daß p und q keine gemeinsame Nullstelle besitzen. Weiter sei*

$$q(z) = a(z-\alpha_1)^{n_1} \cdots (z-\alpha_s)^{n_s}$$

die Faktorisierung von q mit $a, \alpha_1, \ldots, \alpha_s \in \mathbb{C}$, $\alpha_i \neq \alpha_j$, für $i \neq j$. Dann hat p/q die folgende eindeutige Patialbruchzerlegung

$$\begin{aligned}\frac{p(z)}{q(z)} &= \frac{a_{11}}{z-\alpha_1} + \frac{a_{12}}{(z-\alpha_1)^2} + \cdots + \frac{a_{1,n_1}}{(z-\alpha_1)^{n_1}} \\ &+ \frac{a_{21}}{(z-\alpha_2)} + \frac{a_{22}}{(z-\alpha_2)^2} + \cdots + \frac{a_{2,n_2}}{(z-\alpha_2)^{n_2}} \\ &\vdots \\ &+ \frac{a_{s1}}{(z-\alpha_s)} + \frac{a_{s2}}{(z-\alpha_s)^2} + \cdots + \frac{a_{s,n_s}}{(z-\alpha_s)^{n_s}}\end{aligned}$$

mit $a_{11}, \ldots, a_{s,n_s} \in \mathbb{C}$.

Beweis: Sei α eine m- falche Nullstelle von $q(z)$ mit $m \geq 1$. Dann ist

$$q(z) = (z-\alpha)^m s(z)$$

mit $s(\alpha) \neq 0$. Sei $a = p(\alpha)/q(\alpha)$. Dann erhalten wir

$$\frac{p(z)}{q(z)} - \frac{a}{(z-\alpha)^m} = \frac{p(z) - as(z)}{q(z)}$$

und α ist eine Nullstelle des Polynoms $p(z) - as(z)$.

Fall 1: $p(z) - a(z) \equiv 0$. Dann ist die Partialbruchzerlegung bereits erreicht.

Fall 2: $p(z) - as(z) \equiv 0$. Dann existiert $\tilde{p}(z)$ mit

$$p(z) - as(z) = (z-\alpha)\tilde{p}(z)$$

und Grad $(\tilde{p})$ = Grad $(p) - 1$. Sei $\tilde{q}(z) = q(z)/(z-\alpha)$. Dann gilt

$$\frac{p(z)}{q(z)} - \frac{a}{(z-\alpha)m} = \frac{\tilde{p}(z)}{\tilde{q}(z)}$$

wobei Grad $(\tilde{q})$ < Grad $(\tilde{p})$. Wir können jetzt rekursiv den Beweis weiterführen.

□

Die Koeffizienten a_{ij} der Partialbruchzerlegung kann man durch Koeffizientenvergleich ermitteln.

Beispiel:

$$\frac{5z+3}{(z-2)^2(z-1)} = \frac{a}{(z-2)^2} + \frac{b}{z-2} + \frac{c}{z-1}.$$

(1) Multipliziere beide Seiten mit $(z-2)^2$ und setze $z = 2$. Daraus erhalten wir $a = 13$.

(2)

$$\frac{5z+3}{(z-2)^2(z-1)} - \frac{13}{(z-2)^2} = \frac{5z+3-13z+13}{(z-2)^2(z-1)} = \frac{-8z+16}{(z-2)^2(z-1)}$$
$$= -8\frac{z-2}{(z-2)^2(z-1)} = -\frac{8}{(z-2)(z-1)} = \frac{b}{z-z} + \frac{c}{z-1}.$$

Multiplikation mit $(z-2)$ und $z=2$ setzen ergibt $b=-8$. Ebenso folgt $c=8$. Damit ist

$$\frac{5z+3}{(z-2)^2(z-1)} = \frac{13}{(z-2)^2} + \frac{-8}{z-2} + \frac{8}{z-1}.$$

Die elementaren rationalen Funktionen werden wie folgt integriert:

(1) $k>1$.

$$\int \frac{dx}{(x-a)^k} = \frac{1}{1-k}\,\frac{1}{(x-a)^{k-1}}.$$

(2) $k=1$. Wir nehmen an, daß die Koeffizienten von p und q reell sind. Dann treten in der Patialbruchzerlegung Paare von Brüchen der Form

$$\frac{a}{x-\alpha} + \frac{\overline{a}}{x-\overline{\alpha}} = \frac{Ax+B}{x^2+2cx+d}$$

auf.

Diese Funktion integriert man wie folgt: Es sei $Q(x)=x^2+2cx+d$. Dann ist

$$Ax+B = \frac{A}{2}(2x+2c) + (B-Ac) = \frac{A}{2}Q'(x) + (B-Ac).$$

Hieraus folgt

$$\frac{Ax+B}{x^2+2cx+d} = \frac{A}{2}\frac{Q'(x)}{Q(x)} + (B-Ac)\frac{1}{Q(x)}.$$

Damit erhalten wir

$$\int \frac{Ax+B}{x^2+2cx+d}dx = \frac{A}{2}\int \frac{Q'(x)}{Q(x)}dx + (B-Ac)\int \frac{dx}{x^2+2cx+d}.$$

Weiter ist

$$\int \frac{Q'(x)}{Q(x)}dx = \log|x^2+2cx+d|.$$

Es bleibt noch das Integral $\int \frac{dx}{x^2+2cx+d}$ zu berechnen. Es sei

$$x^2+2cx+d = (x+c)^2 + (d-c^2).$$

Fall 1: $D = d - c^2 > 0$.

$$\int \frac{dx}{x^2 + 2cx + d} = \frac{1}{\sqrt{D}} \arctan\left(\frac{x+c}{\sqrt{D}}\right).$$

Übung!

Fall 2: $d = c^2$.

$$\int \frac{dx}{x^2 + 2cx + d} = -\frac{1}{x+c}$$

Fall 3: $D = d - c^2 < 0$.

$$\int \frac{dx}{x^2 + 2cx + d} = \frac{1}{2\sqrt{-D}} \log\left|\frac{x+c-\sqrt{-D}}{x+c+\sqrt{-D}}\right|$$

Übung!

8.5. Integration von Folgen und Reihen von Funktionen

Der Einfachheit halber beschränken wir uns auf Folgen stetiger Funktionen. Die Sätze gelten aber auch für Regelfunktionen.

Satz 8.19. *Es sei (f_n) eine Folge stetiger Funktionen $f_n\colon [a,b] \to \mathbb{C}$ und (f_n) konvergiere gleichmäßig gegen die Funktion $f\colon [a,b] \to \mathbb{C}$. Dann ist f stetig und es gilt*

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

Beweis: Es sei $\epsilon > 0$ gegeben. Zu ϵ existiert $N \in \mathbb{N}$ so, daß

$$\forall n \geq \mathbb{N}\ \forall x \in [a,b]\colon |f_n(x) - f(x)| < \frac{\epsilon}{b-a}.$$

Daraus folgt

$$\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| \leq (b-a) \parallel f - f_n \parallel < (b-a)\frac{\epsilon}{b-a} = \epsilon$$

für $n \geq N$.

□

Satz 8.20. *Sei (f_n) eine Folge stetiger Funktionen $f_n\colon [a,b] \to \mathbb{C}$ und $f(x) = \sum_{n=1}^{\infty} f_n(x)$ konvergiere gleichmäßig auf $[a,b]$. Dann ist f stetig und es gilt*

$$\int_a^b f(x)dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)dx.$$

Beweis. Folgt aus Satz 8.19 durch Anwendung auf die Folge der Partialsummen

$$s_n(x) = \sum_{k=1}^{n} f_k(x).$$

□

Gliedweise Differentation

Satz 8.21. *Es sei* (f_n) *eine Folge stetig differenzierbarer Funktionen* $f_n\colon [a,b] \to \mathbb{C}$ *und es existiere* $f(x) = \lim_{n\to\infty} f_n(x)$ *punktweise. Die Folge* (f_n') *sei gleichmäßig konvergent mit*

$$g(x) = \lim_{n\to\infty} f_n'(x).$$

Dann ist f *stetig differenzierbar auf* $[a,b]$ *und es gilt*

$$f'(x) = (\lim_{n\to\infty} f_n(x))' = \lim_{n\to\infty} f_n'(x)$$

in $[a,b]$.

Beweis: Da (f_n') gleichmäßig konvergiert, folgt aus Satz 6.43, daß

$$g(x) = \lim_{n\to\infty} f_n'(x)$$

stetig ist.

Es sei $x_0 \in [a,b]$. Aus Satz 8.19 folgt für alle $x \in [a,b]$:

$$\begin{aligned}\int_{x_0}^{x} g(x)dx &= \lim_{n\to\infty} \int_{x_0}^{x} f_n'(x)dx \\ &\underset{Hauptsatz}{=} \lim_{n\to\infty} [f_n(x) - f_n(x_0)] = f(x) - f(x_0).\end{aligned}$$

Aus dem Hauptsatz 8.16 folgt

$$g(x) = \frac{d}{dx}\left(\int_{x_0}^{x} g(y)d(y)\right).$$

Daher ist $f(x)$ differenzierbar und es gilt für alle $x \in [a,b]$:

$$f'(x) = g(x) = \lim_{n\to\infty} f_n'(x).$$

□

Korollar 8.22. *Für alle* $n \in \mathbb{N}_0$ *sei* $f_n\colon [a,b] \to \mathbb{C}$ *stetig differenzierbar. Weiter gelte:*

(1) $f(x) = \sum_{n=0}^{\infty} f_n(x)$ *konvergiere für alle* $x \in [a,b]$.
(2) $\sum_{n=0}^{\infty} f_n'(x)$ *konvergiert gleichmäßig in* $[a,b]$.

Dann ist f stetig differenzierbar und es gilt

$$f'(x) = \left(\sum_{n=0}^{\infty} f_n(x) \right)' = \sum_{n=0}^{\infty} f_n'(x)$$

für $x \in [a, b]$.

Beweis: Folgt durch Anwendung von Satz 8.21 auf die Folge der Partialsummen. □

Korollar 8.23. *Eine Potenzreihe $P(x) = \sum_{n=0}^{\infty} a_n x^n$ mit Konvergenzradius $r > 0$ kann für $|x| < r$ beliebig of gliedweise integriert und differenziert werden.*

Beispiele:

(1) arctan(x).

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

für $|x| < 1$. Aus Korollar 8.23 folgt

$$\arctan(x) = \int_0^x \frac{1}{1+x^2} dx = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

für $|x| < 1$.

(2) $\ln(1+x)$. Es gilt

$$\ln'(1+x) = \frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n,$$

für $|x| < 1$. Aus Korollar 8.23 folgt

$$\ln(1+x) = \int_0^x \frac{1}{1+x} dx = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

für $|x| < 1$.

Der Grenzwertsatz von Abel

Satz 8.24. *(Abel) Die Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ habe den Konvergenzradius $r > 0$ und sei für $x = r$ konvergent. Dann ist die Reihe im abgeschlossenen Intervall $[0, r]$ gleichmäßig konvergent und die Funktion $f(x) = \sum_{n=0}^{\infty} a_n x^n$ ist in $x = r$ linksseitig stetig:*

$$\lim_{x \to r-0} f(x) = f(r) = \sum_{n=0}^{\infty} a_n r^n.$$

Entsprechendes gilt, wenn die Reihe für $x = -r$ konvergiert.

Beweis: Wir können annehmen, daß $r = 1$ ist. Andernfalls betrachten wir die Reihe

$$g(x) = f(xr) = \sum_{n=0}^{\infty} a_n r^n x^n.$$

Es sei also $r = 1$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$. Für $n > m$ sei

$$s_{n,m} = \sum_{k=m+1}^{n} a_k.$$

Aus dem Cauchy-Kriterium folgt:

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n > m \geq N\colon \ |s_{n,m}| < \epsilon \tag{8.1}$$

Insbesondere konvergiert die Reihe $\sum_{n=m+1}^{\infty} s_{n,m} x^n$ für $|x| < 1$. Weiter gilt

$$(1-x) \sum_{n=m+1}^{\infty} s_{n,m} x^n = \sum_{n=m+1}^{\infty} (s_{n,m} - s_{n-1,m}) x^n = \sum_{n=m+1}^{\infty} a_n x^n. \tag{8.2}$$

Es sei $r_m(x) = \sum_{n=m+1}^{\infty} a_n x^n$. Dann folgt aus (8.1) und (8.2)

$$|r_m(x)| \leq (1-x)\epsilon \sum_{n=m+1}^{\infty} x^n \leq \epsilon$$

für $0 \leq x < 1$. Dies gilt auch für $x = 1$. Damit konvergiert $r_m(x)$ gleichmäßig in $[0,1]$ gegen 0. Dies bedeutet, daß die Reihe $\sum_{n=0}^{\infty} a_n x^n$ gleichmäßig in $[0,1]$ konvergiert. Daraus folgt, daß $f(x) = \sum_{n=0}^{\infty} a_n x^n$ stetig in $[0,1]$ ist.

□

Beispiele: 1)

$$\arctan(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}, \quad |x| < 1.$$

Auf Grund des Leibnitz-Kriteriums konvergiert die Reihe für $x = 1$. Aus Satz 8.24 folgt dann

$$\frac{\pi}{4} = \arctan(1) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

Dies ist die berühmte Leibniz-Reihe für $\pi/4$.

2)

$$\ln(1+x) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1}, \quad |x| < 1.$$

Aus dem Leibnitz-Kriterium folgt, daß die Reihe für $x = 1$ konvergiert. Aus Satz 8.24 folgt

$$\ln(2) = \sum_{n=0}^{\infty}(-1)^n \frac{1}{n+1} = 1 - \frac{1}{2} + \frac{1}{3} \cdots$$

9. Taylorpolynome und Taylorreihen

Aus dem Abschnitt 7.1 wissen wir, das die Differenzierbarkeit einer Funktion $f\colon I \to \mathbb{C}$ im Punkte $x_0 \in I$ äquivalent dazu ist, daß f in einer Umgebung von x_0 durch eine lineare Funktion approximiert werden kann. Genauer gilt

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \rho(x)$$

mit

$$\lim_{x \to x_0} \frac{\rho(x)}{x - x_0} = 0.$$

Wir untersuchen jetzt, inwieweit f durch Polynome höherer Ordunung approximiert werden kann.

Defintion 9.1. *Es sei $f\colon I \to \mathbb{C}$ n-mal differenzierbar, $n \geq 1$, und sei $x_0 = I$. Dann heißt die Funktion*

$$T_n f(x; x_0) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

das n-te **Taylorpolynom** *von f in x_0.*

Bemerkung: Wenn klar ist, um welchen Punkt x_0 es sich handelt, schreiben wir $T_n f(x)$ anstelle von $T_n f(x; x_0)$. Um festzustellen wie gut das Taylorpolynom $T_n f(x; x_0)$ die Funktion f im Punkte x_0 approximiert, müssen wir den Fehler oder das sogenannte Restglied

$$R_{n+1}(x) = f(x) - T_n f(x; x_0)$$

abschätzen. Das Restglied $R_{n+1}(x)$ kann auf verschiedene Weise charakterisiert werden.

Satz 9.2. *(Integralform für R_{n+1}): Die Funktion $f: I \to \mathbb{C}$ sei $(n+1)$-mal stetig differenzierbar auf dem Intervall I und es sei $x_0 \in I$. Dann gilt für alle $x \in I$:*

$$R_{n+1}(x) = \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt.$$

Beweis: Der Beweis erfolgt durch vollständige Induktion.
Induktionsanfang: $n = 0$. Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt.$$

Schluß von n-1 auf n: Nach Voraussetzung gilt

$$f(x) = T_{n-1}f(x;x_0) + \frac{1}{(n-1)!} \int_{x_0}^{x} (x-t)^{n-1} f^{(n)}(t)dt.$$

Daraus erhält man durch partielle Integration:

$$\begin{aligned} f(x) &= T_{n-1}f(x;x_0) - \frac{1}{n!}(x-t)^n f^{(n)}(t) \,|_{x_0}^{x} + \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt \\ &= T_n f(x;x_0) + \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt. \end{aligned}$$

Hieraus folgt die behauptete Form für das Restglied.

□

Korollar 9.3. *(Lagrange-Form für R_{n+1}):*
Es sei $f: I \to \mathbb{R}$ $(n+1)$- mal stetig differenzierbar. Seien $x, x_0 \in I$, $x \neq x_0$. Dann gibt es ein ξ zwischen x_0 und x so, daß

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

Beweis: Es sei $p(t) = (x-t)^n$. Wenn $x > x_0$, so ist $p(t) \geq 0$ für alle $t \in [x_0, x]$ und falls $x < x_0$, ist $p(t) \leq 0$ für alle $t \in [x, x_0]$. Aus dem Mittelwertsatz der Integralrechnung folgt, daß ein ξ zwischen x_0 und x existiert mit

$$\begin{aligned} R_{n+1}(x) &= \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt \\ &= \frac{f^{(n+1)}(\xi)}{n!} \int_{x_0}^{x} (x-t)^n dt = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \end{aligned}$$

□

Definition 9.4. *Sei $f : I \to \mathbb{C}$ unendlich oft differenzierbar und sei $x_0 \in I$. Dann heißt*

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

die **Taylorreihe** *zu f in x_0.*

Man beachte: 1) Die Taylorreihe ist für $x \neq x_0$ im allgemeinen **nicht** konvergent.
2) Wenn die Taylorreihe in einer Umgebung von $x_0 \in I$ konvergiert, so folgt daraus **nicht**, daß die Taylorreihe mit f übereinstimmt.
Beispiel: Es sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Für $x > 0$ ist $f'(x) = x^{-2}e^{-1/x}$ und daher gilt für die rechtsseitige Ableitung

$$f'_{+}(x) = 0.$$

Da $f(x) = 0$ für $x < 0$, ist

$$f'_{-}(x) = 0.$$

Daher stimmen der links- und rechtsseitige Ableitung von $f(x)$ in $x = 0$ überein, d.h., f ist in $x = 0$ differenzierbar und damit auf ganz $\mathbb{R}$. Durch Induktion zeigt man, daß f unendlich oft differenzierbar ist und es existieren Polynome $P_n(x)$, $n \in \mathbb{N}$, mit

$$f^{(n)}(x) = \begin{cases} P_n(\frac{1}{x})e^{-1/x}, & x > 0; \\ 0, & \leq 0; \end{cases}$$

für alle $n \in \mathbb{N}$. Insbesondere folgt daraus

$$\forall n \in \mathbb{N}_0 : \ f^{(n)}(0) = 0.$$

Deshalb ist die Taylorreihe von f in 0 identisch Null, aber $f \neq 0$.

Die Frage, wann $f(x) = Tf(x; x_0)$ gilt, wird durch den folgenden Satz beantwortet.

Satz 9.5. *Es seien $f \in C^{\infty}(I)$ und $x, x_0 \in I$. Dann gilt*

$$f(x) = Tf(x; x_0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

genau dann, wenn für das Restglied gilt:

$$\lim_{n\to\infty} R_n(x, x_0) = 0.$$

Beweis: Für jedes $n \in \mathbb{N}$ ist

$$f(x) = T_n f(x; x_0) + R_{n+1}(x; x_0).$$

Deshalb ist die Bedingung

$$\lim_{n\to\infty} R_n(x; x_0) = 0$$

äquivalent zu

$$\lim_{n\to\infty} T_n f(x; x_0) = f(x).$$

□

Satz 9.6. *Es sei $f \in C^\infty(I)$. Wir nehmen an, daß Konstanten $\alpha, C > 0$ existieren so, daß für alle $n \in \mathbb{N}_0$:*

$$|f^{(n)}(x)| \leq \alpha C^n \quad \text{für alle } x \in I.$$

Dann gilt

$$\lim_{n\to\infty} R_n(x; x_0) = 0$$

für alle $x \in I$.

Beweis: Wir verwenden das Restglied in Lagrange-Form. Zu jedem $n \in \mathbb{N}$ existiert $\xi_n \in I$ mit

$$R_n(x; x_0) = \frac{f^{(n)}(\xi_n)}{n!}(x - x_0)^n.$$

Daraus folgt

$$|R_n(x; x_0)| = \frac{|x - x_0|^n}{n!}|f^{(n)}(\xi_n)| \leq \alpha \frac{C^n|x - x_0|^n}{n}.$$

Da $\exp(C|x - x_0|) = \sum_{n=0}^\infty \frac{C^n|x-x_0|^n}{n!}$ konvergiert, folgt

$$\frac{C^n|x - x_0|^n}{n!} \to 0$$

für $n \to \infty$.

□

Vorlesung Infinitesimalrechnung II

10. Differentialrechnung im $\mathbb{R}^n$

10.1. Topologie des $\mathbb{R}^n$

Es ist $\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}\}$ und für $x \in \mathbb{R}^n$ ist die Norm $\| x \|$ definiert durch

$$\| x \| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Für die Norm $\| \cdot \|$ gilt:

(1) $\forall x \in \mathbb{R}^n\colon \ \| x \| \geq 0$ und $\| x \| = 0 \Leftrightarrow x = 0$.
(2) $\forall x \in \mathbb{R}^n$, $\forall \lambda \in \mathbb{R}\colon \ \| \lambda x \| = |\lambda| \, \| x \|$.
(3) $\forall x, y \in \mathbb{R}^n\colon \ \| x + y \| \leq \| x \| + \| y \|$.

$(\mathbb{R}^n, \| \cdot \|)$ ist ein **normierter Vektorraum**

Allgemeiner: Sei $d(x,y) = \| x - y \|$ für $x, y \in \mathbb{R}^n$. Dann gilt für alle $x, y, z \in \mathbb{R}^n$:

(1) $d(x,y) \geq 0$ und $d(x,y) = 0 \Leftrightarrow x = y$
(2) $d(x,y) = d(y,x)$
(3) $d(x,z) \leq d(x,y) + d(y,z)$.

$(\mathbb{R}^n, d)$ ist dann ein **metrischer Raum**.

Definition 10.1. *Für $\epsilon > 0$ und $x \in \mathbb{R}^n$ sei*

$$U_\epsilon(x) = \{y \in \mathbb{R}^n \mid \| x - y \| < \epsilon\}.$$

Für $n = 1$ ist $U_\epsilon(x) = (x - \epsilon, x + \epsilon)$ ein Intervall.
Für $n + 2$ ist $U_\epsilon(x)$ die Kreisscheibe mit Mittelpunkt x und Radius ϵ.
Im allgemeinen ist $U_\epsilon(x)$ die n-dimensionale Kugel vom Radius ϵ mit Zentrum x.
Sei $U \subset \mathbb{R}^n$.

$$U \text{ heißt } \begin{cases} \text{offen} & \Leftrightarrow \forall x \in U \ \exists \epsilon > 0\colon U_\epsilon(x) \subset U; \\ \text{abgeschlossen} & \Leftrightarrow \mathbb{R}^n \setminus U \text{ ist offen;} \\ \text{beschränkt} & \Leftrightarrow \exists C \in \mathbb{R}_+\colon \forall x \in U\colon \ \| x \| \leq C; \\ \text{kompakt} & \Leftrightarrow U \text{ ist beschränkt und abgeschlossen.} \end{cases}$$

Sei $f\colon \mathbb{N} \to \mathbb{R}^n$ eine Abbildung. Wir setzen $x_k = f(k)$ für $k \in \mathbb{N}$. (x_k) heißt **Folge in** $\mathbb{R}^n$.

Defintion 10.2. *Eine Folge (x_k) in $\mathbb{R}^n$ heißt* **konvergent** $\Leftrightarrow$

$$\exists a \in \mathbb{R}^n \ \forall \epsilon > 0 \ \exists N \in \mathbb{N}\colon \ \| x_k - a \| < \epsilon$$

für alle $k \geq N$.

Satz 10.3. *$A \subset \mathbb{R}^n$ ist abgeschlossen $\Leftrightarrow$ Für jede Folge (x_k) in A, die in $\mathbb{R}^n$ gegen ein x konvergiert, gilt $x \in A$.*

Beweis: Übung!

Satz 10.4. *(Heine-Borel) $K \subset \mathbb{R}^n$ ist kompakt $\Leftrightarrow$ Jede Folge (x_k) in K hat eine konvergente Teilfolge (x_{k_j}) mit $\lim_{j\to\infty} x_{k_j} \in K$.*

Beweis: Übung!

10.2. Stetige Abbildungen

Defintion 10.5. *Sei $U \subset \mathbb{R}^n$. Eine Abbildung $f\colon U \to \mathbb{R}^m$ heißt stetig in $x_0 \in U$, wenn gilt:*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U\colon\ \| x - x_0 \| < \delta \Rightarrow \| f(x) - f(x_0) \| < \epsilon.$$

$f\colon U \to \mathbb{R}^m$ ist stetig $\Leftrightarrow$ f ist in jedem Punkt $x \in U$ stetig.

Alle Sätze über stetige Funktionen in $\mathbb{R}$ übertragen sich sinngemäß.

Satz 10.6. *(Folgenkriterium) $f\colon U \to \mathbb{R}^m$ ist stetig in $a \in U$ $\Leftrightarrow$ Für jede Folge (x_k) in U gilt*

$$x_k \xrightarrow[k\to\infty]{} a \ \Rightarrow\ f(x_k) \xrightarrow[k\to\infty]{} f(a).$$

Beweis: Analog zum Fall $n = m = 1$. Übung!

Grenzwerte

Sei $U \subset \mathbb{R}^n$. Ein Punkt $x \in \mathbb{R}^n$ heißt **Häufungspunkt** von U, wenn jede Umgebung von x einen von x verschiedenen Punkt aus U enthält.

Satz. *$A \subset \mathbb{R}^n$ ist abgeschlossen $\Leftrightarrow$ A enthält alle Häufungspunkte von A.*
Beweis: *Übung!*

Defintion. *Sei $f\colon U \to \mathbb{R}^m$ eine Abbildung und sei $x_0 \in \mathbb{R}^n$ ein Häufungspunkt von U. Die Abbildung f hat den* **Grenzwert** *$a \in \mathbb{R}^m$ in x_0 $\Leftrightarrow$*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U - \{x_0\}\colon\ \| x - x_0 \| < \delta \Rightarrow \| f(x) - a \| < \epsilon.$$

Bezeichnung. $\lim_{x\to x_0} f(x) = a$.

Satz. *$f\colon U \to \mathbb{R}^m$ ist stetig in $x_0 \in U$ genau dann, wenn*

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

Beispiel: Sei $L\colon \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Dann ist L stetig.
Beweis: Übung!

10.3. Differenzierbare Abbildungen

Es sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ eine Abbildung. Wir möchten jetzt definieren, was es heißt, daß f in $x_0 \in U$ differenzierbar ist. Dafür gibt es verschiedene Möglichkeiten. Wir erinnern uns an die folgende äquivalente Defintion der Differenzierbarkeit im Falle $n = m = 1$.

$f\colon I \to \mathbb{R}$ ist in $x_0 \in I$ differenzierbar genau dann, wenn gilt: $\exists \lambda \in \mathbb{R}$ und $\exists \rho\colon I \to \mathbb{R}$ mit $\rho(x_0) = 0$ und $\lim_{x\to x_0} \rho(x)/(x - x_0) = 0$ so, daß

$$f(x) = f(x_0) + \lambda(x - x_0) + \rho(x).$$

Dann ist $\lambda = f'(x_0)$. Sei $L\colon \mathbb{R} \to \mathbb{R}$ definiert durch

$$L(x) = \lambda x.$$

Dann ist $L \in \operatorname{Hom}(\mathbb{R}, \mathbb{R})$ und f wird durch L in x_0 approximiert:

$$f(x) = f(x_0) + L(x - x_0) + \rho(x),\ \rho(x) = o(x - x_0).$$

Dies verallgemeinern wir für beliebige n, m wie folgt:

Definition 10.7. *Sei $U \in \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ eine Abbildung. Sei $x_0 \in U$. Dann heißt f in x_0 (total) differenzierbar, wenn ein $L \in Hom(\mathbb{R}^n, \mathbb{R}^m)$ und eine in einer hinreichend kleinen Umgebung $U_\epsilon(0)$ definierte Abbildung $\varphi\colon U_\epsilon(0) \to \mathbb{R}^m$ existieren so, daß*

(1) *$f(x_0 + h) = f(x_0) + L(h) + \varphi(h)$ für alle $h \in U_\epsilon(0)$.*
(2) $\lim_{h\to 0} \dfrac{\varphi(h)}{\| h \|} = 0.$

Wenn L existiert, so ist L eindeutig bestimmt und heißt **Differential** *oder* **Ableitung** *von f in x_0.*

Bezeichnung: $L = df(x_0) = f'(x_0)$.

$f\colon U \to \mathbb{R}^m$ heißt differenzierbar $\Leftrightarrow$ f ist in jedem Punkt $x \in U$ differenzierbar.

Beispiel: Sei $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ und $f\colon D \to \mathbb{R}$ sei definiert durch

$$f(x, y) = \sqrt{1 - x^2 - y^2}.$$

Dann ist f differenzierbar. Der Graph der linearen Funktion $t(h) = f(x_0) + L(h)$, $h \in \mathbb{R}^n$ ist die Tangentialhyperebene an den Graph Γ_f im Punkte $(x_0, f(x_0))$.

Bemerkung: $\lim_{h\to 0} \frac{\varphi(h)}{\|h\|}$ kann auch kurz geschrieben werden als $\varphi(h) = o(\| h \|)$.

Lemma. *$f\colon U \to \mathbb{R}^m$ ist differenzierbar in $x_0 \in U$ genau dann, wenn eine Abbildung*

$$\Lambda\colon U \to \ Hom(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{n\cdot m}$$

existiert so, daß

(1) *Λ ist in x_0 stetig.*
(2) *$f(x_0 + h) = f(x_0) + \Lambda(x_0 + h)$ für alle $h \in \mathbb{R}^n$ mit $x_0 + h \in U$.*

Beweis: Analog zum Fall $n = m = 1$. Übung.

Lemma 10.8. *Wenn $f\colon U \to \mathbb{R}^m$ differenzierbar ist, so ist f stetig.*

Beweis: Sei $x_0 \in U$. Da f in x_0 differenzierbar ist, existieren $L \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ und $\varphi\colon U_\epsilon(0) \to \mathbb{R}^m$ so, daß

$$f(x_0 + h) = f(x_0) + L(h) + \varphi(h)$$

und $\varphi(h) = o(\| h \|)$. Da L stetig ist und $\lim_{h\to 0} \varphi(h) = 0$, folgt

$$\begin{aligned}\lim_{h\to\infty} f(x_0 + h) &= f(x_0) + \lim_{h\to 0} L(h) + \lim_{h\to 0} \varphi(h)\\ &= f(x_0) + L(0) = f(x_0).\end{aligned}$$

$\square$

Elementare Eigenschaften der Ableitung

Satz 10.9. *(Reduktionslemma): Sei $U \subset \mathbb{R}^n$ offen und seien $f\colon U \to \mathbb{R}^m$, $g\colon U \to \mathbb{R}^k$ Abbildungen. Es sei $(f, g)\colon U \to \mathbb{R}^m \times \mathbb{R}^k$ definiert durch*

$$(f, g)(x) = (f(x), g(x)), \quad x \in U.$$

Dann ist (f, g) in $x_0 \in U$ differenzierbar genau dann, wenn f und g in x_0 differenzierbar sind. Falls dies der Fall ist, gilt

$$d(f, g)(x_0) = (df(x_0), dg(x_0)).$$

Beweis: $\Leftarrow$) Sei $L = df(x_0)$ und $M = dg(x_0)$. Nach Voraussetzung existieren Abbildungen $\varphi\colon U_\epsilon(0) \to \mathbb{R}^m$ und $\psi\colon U_\epsilon(0) \to \mathbb{R}^k$ mit $\varphi(h) = o(\| h \|)$ und $\psi(h) = o(\| h \|)$ für $h \to 0$ so, daß gilt

$$\begin{aligned}f(x_0 + h) &= f(x_0) + L(h) + \varphi(h),\\ g(x_0 + h) &= g(x_0) + M(h) + \psi(h),\end{aligned}$$

Sei $(L, M)\colon \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^k$ die Abbildung, die definiert ist durch $(L, M)(h) = (L(h), M(h))$. Entsprechend sei $(\varphi, \psi) : U_\epsilon(0) \to \mathbb{R}^{m+k}$ definiert durch $(\varphi, \psi)(h) = (\varphi(h), \psi(h))$. (L, M) ist linear und es gilt:

$$\begin{aligned}(f, g)(x_0 + h) &= (f(x_0 + h), g(x_0 + h)) \\ &= (f(x_0), g(x_0)) + (L(h), M(h)) + (\varphi(h), \psi(h)).\end{aligned}$$

Weiter ist

$$\| (\varphi(h), \psi(h)) \|^2 = \| \varphi(h) \|^2 + \| \psi(h) \|^2 = o(\| h \|^2).$$

Daraus folgt

$$\lim_{h\to 0} \frac{(\varphi(h), \psi(h))}{\| h \|} = 0.$$

Gemäß Definition 10.7 ist (f, g) daher differenzierbar in x_0 und $d(f, g)(x_0) = (L, M) = (df(x_0), dg(x_0))$.

$\Rightarrow$) Der Beweis dieser Richtung ist analog.

□

Es sei $pr_j\colon \mathbb{R}^m \to \mathbb{R}$ definiert durch $pr_j(x_1, \dots, x_m) = x_j$, d.h., pr_j ist die j-te Koordinatenprojektion. Sei $f\colon U \to \mathbb{R}^m$ eine Abbildung und sei $f_j = pr_j \circ f$. Dann sind $f_j\colon U \to \mathbb{R}$, $j = 1, ..m$, Funktionen und es gilt

$$f(x) = (f_1(x), \dots, f_m(x)).$$

Korollar 10.10. *Sei $f\colon U \to \mathbb{R}^m$ eine Abbildung und sei $x_0 \in U$. Dann ist f differenzierbar in x_0 genau dann, wenn für alle j, $1 \le j \le m$, f_j in x_0 differenzierbar ist.*

Mit Hilfe dieses Korollars können wir Fragen der Differenzierbarkeit von Abbildungen $f\colon U \to \mathbb{R}^k$ auf den Fall von Funktionen reduzieren.

Satz 10.11. *Seien $f, g\colon U \to \mathbb{R}^m$ differenzierbar in $x_0 \in U$ und seien $\alpha, \beta \in \mathbb{R}$. Dann ist $\alpha f + \beta g$ differenzierbar in x_0.*

Beweis: Übung!

Satz 10.12. *(Kettenregel) Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen, $f\colon U \to \mathbb{R}^m$ und $g\colon V \to \mathbb{R}^l$ seien Abbildungen mit $f(U) \subset V$. Sei f differenzierbar in $x_0 \in U$ und g differenzierbar in $y_0 = f(x_0) \in V$. Dann ist $g \circ f$ differenzierbar in x_0 und es gilt*

$$d(g \circ f)(x_0) = dg(y_0) \circ df(x_0).$$

Beweis: Der Beweis ist analog zum Falle $n = m = 1$. Auf Grund des oben bewiesenen Lemmas existieren Abbildungen

$$\Lambda\colon U \to \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m) \ \text{ und } M\colon V \to \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^l)$$

so, daß gilt

(1) Λ ist stetig in x_0 und $\Lambda(x_0) = df(x_0)$, M ist stetig in y_0 und $M(x_0) = dg(y_0)$.

(2)
$$f(x_0 + h) = f(x_0) + \Lambda(x_0 + h)(h);$$
$$g(y_0 + k) = g(y_0) + M(y_0 + k)(k).$$

Daraus folgt
$$(g \circ f)(x_0 + h) = g(f(x_0) + \Lambda(x_0 + h)(h)) =$$
$$g(f(x_0)) + M(f(x_0) + \Lambda(x_0 + h)(h))(\Lambda(x_0 + h)(h)).$$

Weiter folgt aus 1), daß
$$\lim_{h\to 0} M(f(x_0) + \Lambda(x_0 + h)(h))(\Lambda(x_0 + h)(h)) = M(f(x_0))\Lambda(x_0)$$
$$= dg(f(x_0))df(x_0).$$

Nochmalige Anwendung des obigen Lemmas folgt, daß $g \circ f$ in x_0 differenzierbar ist und das Differential die behauptete Form hat.

□

Abbildung von Tangentialvektoren

Sei $\gamma\colon I \to U \subset \mathbb{R}^n$ eine differenzierbare Abbildung. Dann ist γ ist eine differenzierbare Kurve in U. Es sei
$$\gamma(t) = (x_1(t), \dots, x_n(t)).$$

Dann ist
$$\frac{d\gamma}{dt}(t_0) = \left(\frac{dx_1}{dt}(t_0), \dots, \frac{dx_n}{dt}(t_0)\right) \in \mathbb{R}^n$$
der Tangentialvektor an die Kurve γ in $\gamma(t_0)$.

Sei $f\colon U \to \mathbb{R}^m$ eine differenzierbare Abbildung. Dann ist $f\circ\gamma\colon I \to \mathbb{R}^m$ eine differenzierbare Kurve im $\mathbb{R}^m$. Nach der Kettenregel gilt
$$\frac{d}{dt}(f \circ \gamma)(t_0) = df(\gamma(t_0)) \left(\frac{d\gamma}{dt}(t_0)\right).$$

Dies ist der Tangentialvekotor an die Bildkurve $f \circ \gamma$. Wir haben also gezeigt, daß $df(\gamma(t_0))$ Tangentialvektoren an γ in $\gamma(t_0)$ in Tangentialvektoren an $f \circ \gamma$ in $f(\gamma(t_0))$ abbildet.

10.4. Richtungsableitung und partielle Ableitungen

Sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ in $x \in U$ differenzierbar. Sei $v \in \mathbb{R}^n$ gegeben. Da U offen ist, existiert $\epsilon > 0$ so, daß $x + tv \in U$ für alle $t \in \mathbb{R}$ mit $|t| < \epsilon$. Sei

$$g(t) = f(x + tv), \ |t| < \epsilon,$$

und sei $\varphi\colon U_\epsilon(0) \to \mathbb{R}^m$ die Funktion aus der Definition 10.7. Dann ist $g\colon (-\epsilon, \epsilon) \to \mathbb{R}^m$ differenzierbar und es gilt

$$\begin{aligned}\frac{dg}{dt}(t) &= \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} = \lim_{t\to 0}\left(\frac{df(x)(tv)}{t} + \frac{\varphi(tv)}{t}\right) \\ &= df(x)(v) \pm \lim_{t\to 0} \frac{\varphi(tv)}{\| tv \|} \| v \| = df(x)(v).\end{aligned}$$

Also gilt:

$$df(x)(v) = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}.$$

Defintion 10.13. *Sei $f\colon U \to \mathbb{R}^m$ eine Abbildung und sei $v \in \mathbb{R}^n$. Wenn der Grenzwert*

$$\partial_v f(x) = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}$$

existiert, so heißt $\partial_v f(x)$ die **Richtungsableitung** *in x von f in Richtung v. Insbesondere sei $v = e_i = (0, \dots, 0, 1, 0, \dots, 0)$ der i-te Einheitsvektor. Dann heißt*

$$\partial_i f(x) = \partial_{e_i} f(x).$$

die i-te partielle Ableitung von f.

Wir betrachten also $f(x_1, \dots, x_n)$ als Funktion von x_i, wobei alle x_j, $j \neq i$ fest bleiben. Dann ist

$$\partial_i f(x) = \frac{\partial f}{\partial x_i}(x_1, \dots, x_i, \dots, x_n).$$

Definition. *Die Matrix*

$$\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{i,j=1}^{m,n}$$

heißt **Jacobi - Matrix** *von f in x_0.*

Satz 10.14. *Sei $U \subset \mathbb{R}^n$ offen und sei $f\colon U \to \mathbb{R}^m$ in x_0 differenzierbar. Dann ist die Darstellung von $df(x_0)$ in den Standardbasen von $\mathbb{R}^n$ und $\mathbb{R}^m$ gleich der Jacobi - Matrix.*

Beweis: Es gilt

$$df(x_0)(e_j) = \frac{\partial f}{\partial x_j}(x_0) = \sum_{i=0}^{n} \frac{\partial f_i}{\partial x_j}(x_0)e_i.$$

□

Ein Spezialfall ist $m = 1$.

Definition 10.15. *Sei $f\colon U \to \mathbb{R}$ in $x_0 \in U$ differenzierbar. Dann heißt*

$$\operatorname{grad} f(x_0) := (\partial f(x_0), \dots, \partial_n f(x_0))$$

der **Gradient von f in x_0.**

Bezeichnung: Häufig verwendet man auch die Bezeichnung $\nabla f(x_0)$ anstelle von $gradf(x_0)$.

Es sei $f\colon U \to \mathbb{R}$ differenzierbar. In jedem Punkt $x \in U$ ist $\operatorname{grad} f(x)$ ein Vektor in $\mathbb{R}^n$, d.h.,

$$x \in U \mapsto \operatorname{grad} f(x) \in \mathbb{R}^n$$

ist ein **Vektorfeld**. Man nennt es das **Gradientenvektorfeld** zu f.

Bedeutung des Gradientenvektorfeldes

Sei $v \in \mathbb{R}^n,\ v = \sum_{i=1}^n v_i e_i$. Dann ist

$$\begin{aligned}\partial_v f(x_0) = df(x_0)(v) &= \sum_{i=1}^{n} v_i df(x_0)(e_i) = \sum_{i=1}^{n} v_i \partial_i(x) \\ &= \langle v, \operatorname{grad} f(x_0)\rangle = \| v \| \| \operatorname{grad} f(x_0) \| \cos\varphi,\end{aligned}$$

wobei φ der Winkel zwischen v und $\operatorname{grad} f(x_0)$ ist. Hieraus folgt

(1) $\| \operatorname{grad} f(x_0) \| = \max\{\partial_v f(x_0) \mid \| v \| = 1\} =: M$

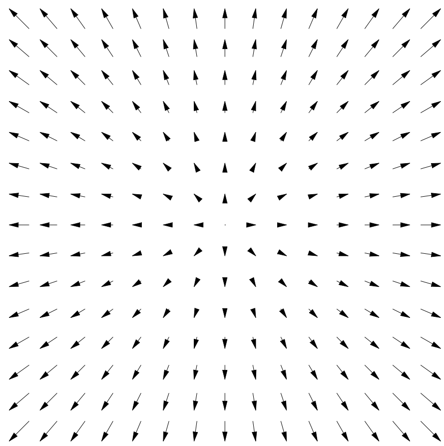
(2) Sei $M \neq 0$. Dann existiert genau ein $v \in \mathbb{R}^n$, $\| v \| = 1$, mit

$$\operatorname{grad} f(x_0) = Mv.$$

Der Gradient zeigt also in Richtung des maximalen Wachstums von f.

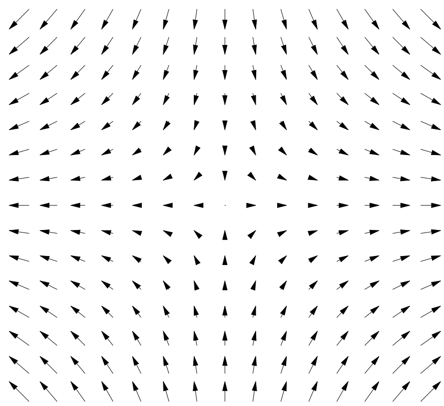
Beispiele:

(1) Es sei $f(x, y) = x^2 + y^2$. Dann ist $\operatorname{grad} f(x, y) = 2(x, y)$. Das Gradientenvektorfeld im Punkte (x_0, y_0) ist orthogonal zur Tangente an die Niveaulinie $f(x, y) = f(x_0, y_0)$ im Punkte (x_0, y_0).



Das Gradientenvektorfeld von $f(x,y) = x^2 + y^2$.

(2) Sei $f(x,y) = x^2 - y^2$. Dann ist $\mathrm{grad} f(x,y) = 2(x,-y)$.



Das Gradientenvektorfeld von $f(x,y) = x^2 - y^2$.

Bemerkung: Partielle Differenzierbarkeit **impliziert nicht** totale Differenzierbarkeit.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = \begin{cases} 0, & (x,y) = (0,0) \\ \frac{xy}{x^2+y^2}, & \text{sonst.} \end{cases}$$

Dann existieren die partiellen Ableitungen von f in (0,0):

$$\partial_x f(0,0) = 0 = \partial_y f(0,0).$$

Aber f ist nicht stetig in (0,0):

$$f(x,x) = \frac{1}{2},\ x \neq 0,\ f(0,0) = 0.$$

Die Frage, unter welchen zusätzlichen Voraussetzungen die totale Differenzierbarkeit aus der partiellen Differenzierbarkeit folgt, wird durch folgenden Satz beantwortet.

Satz 10.16. *Sei $U \subset \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^m$ und $x \in U$. Wenn in U alle partiellen Ableitungen existieren und in x stetig sind, so ist f in x total differenzierbar.*

Beweis: Aus Satz 10.9. folgt, daß wir $m = 1$ annehmen können. Da U offen ist, existiert ein $\delta > 0$ so, daß $U_\delta(x) \subset U$. Sei $h \in U_\delta(0)$. Wir setzen

$$y_0 = x,\ y_i = y_{i-1} + h_i e_i,\ i = 1, \dots, n.$$

Dann ist $y_n = x + h$. Weiter gilt

$$f(x+h) - f(x) = \sum_{i=1}^{n} (f(y_i) - f(y_{i-1})).$$

Aus dem Mittelwertsatz der Differentialrechnung folgt, daß ein $\theta_i \in [0, h_i]$ existiert mit

$$f(y_i) - f(y_{i-1}) = f(y_{i-1} + h_i e_i) - f(y_{i-1}) = h_i \partial_i f(y_{i-1} + \theta_i e_i).$$

Sei $L = \sum_{i=1}^n \partial_i f(x) h_i$ und $\xi_i = y_{i-1} + \theta_i e_i$. Dann ist

$$f(x+h) - f(x) - L(h) = \sum_{i=1}^{n} (\partial_i f(\xi_i) - \partial_i f(x)) h_i$$

Hieraus folgt

$$|f(x+h) - f(x) - L(h)| \leq \| h \| \left(\sum_{i=1}^{n} |\partial_i f(\xi_i) - \partial_i f(x)|^2 \right)^{1/2}.$$

Für $h \to 0$ gilt $\xi_i = y_{i-1} + h_i e_i \to x$. Da $\partial_i f$ stetig in x ist, folgt

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - L(h)|}{\| h \|} = 0.$$

□

Die Beziehungen zwischen Stetigkeit und den verschiedenen Formen der Differenzierbarkeit kann man sich in folgendem Diagramm vergegenwärtigen:

$$\begin{tikzcd}
\textit{stetig partiell differenzierbar} \arrow[d] & \\
\textit{total differenzierbar} \arrow[r] \arrow[d] & \textit{stetig} \\
\textit{partiell differenzierbar} &
\end{tikzcd}$$

Die Folgepfeile gelten nur in dieser Richtung.

Bemerkung: Sei $U \subset \mathbb{R}^2 = \mathbb{C}$ offen, $f\colon U \to \mathbb{R}^2 = \mathbb{C}$ stetig partiell differenzierbar. Dies **impliziert nicht**, daß $f\colon U \to \mathbb{C}$ komplex differenzierbar ist.

Beispiel: Sei $f\colon \mathbb{C} \to \mathbb{C}$ definiert durch $f(z) = \overline{z}$, d.h., wenn $z = x+iy$ ist, so ist

$$f(x, y) = (x, -y).$$

Dann ist offensichtlich f stetig partiell differenzierbar. Sei $z = re^{i\varphi}$ die Darstellung von z in Polarkoordinaten. Dann ist

$$\frac{f(z)}{z} = \frac{\overline{z}}{z} = e^{-2i\varphi}.$$

Daraus folgt, daß der Limes von $\frac{f(z)-f(0)}{z} = \frac{f(z)}{z}$ für $z \to 0$ nicht existiert. $f(z)$ ist also nicht komplex differenzierbar.

10.5. Mittelwert - und Schrankensatz

Für zwei beliebige Punkte $x, y \in \mathbb{R}^n$ sei

$$[x, y] = \{z \in \mathbb{R}^n \mid \exists t \in [0,1]\colon\ z = (1-t)x + ty\}.$$

Dies ist die Verbindungsstrecke von x und y.

Satz 10.17. *Sei $U \in \mathbb{R}^n$ offen und sei $f\colon U \to \mathbb{R}$ differenzierbar. Seien $x, y \in U$ Punkte so, daß $[x, y] \subset U$. Dann existiert $\xi \in [x, y]$ mit*

$$f(y) - f(x) = df(\xi)(y - x).$$

Beweis: Wir definieren $g\colon [0,1] \to \mathbb{R}$ durch

$$g(t) = f((1-t)x + ty).$$

Nach der Kettenregel ist g differenzierbar und es gilt

$$\begin{aligned} g'(t) &= \frac{df}{dt}((1-t)x+ty) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}((1-t)x+ty)(y_j - x_j) \\ &= df((1-t)x+ty)(y-x). \end{aligned}$$

Aus dem Mittelwertsatz für Funktionen einer reellen Veränderlichen folgt, daß ein $t_0 \in (0,1)$ existiert so, daß

$$\begin{aligned} g(1) - g(0) = g'(t_0) &= \\ &= df((1-t_0)x + t_0 y)(y-x). \end{aligned}$$

Sei $\xi = (1-t_0)x + t_0 y$. Dann ist

$$f(y) - f(x) = g(1) - g(0) = df(\xi)(y-x).$$

□

Korollar 10.18. *Sei $U \subset \mathbb{R}^n$ offen und zusammenhängend. Sei $f\colon U \to \mathbb{C}$ differenzierbar und sei $df(x) \equiv 0$. Dann ist f konstant.*

Beweis: O.B.d.A. könnnen wir annehmen, daß f reellwertig ist. Seien $x, y \in U$. Wir wählen Punkte $x_0 = x, x_1 \dots, x_k = y$ so, daß $[x_{i-1}, x_i] \subset U$ für alle $i, i = 1, \dots, k$. Aus Satz 10.17 folgt

$$f(x_i) - f(x_{i-1}) = df(\xi_i)(x_i - x_{i-1}) = 0$$

für alle i. Daher ist $f(x) = f(y)$.

□

Satz 10.19. *Sei $f\colon U \to \mathbb{R}^m$ stetig partiell differenzierbar und sei $\gamma\colon [a,b] \to \mathbb{R}^n$ stetig differenzierbar mit $\gamma(a) = x$ und $\gamma(b) = y$. Dann gilt:*

$$f(y) - f(x) = \int_a^b df(\gamma(t))(\frac{d\gamma}{dt}(t))dt.$$

Beweis: Aus dem Hauptsatz der Differential und Integralrechnung und der Kettenregel folgt

$$\begin{aligned} f(y) - f(x) = f(\gamma(b)) - f(\gamma(a)) &= \int_a^b \frac{d}{dt}(f \circ \gamma)(t)dt \\ &= \int_a^b df(\gamma(t))(\frac{d\gamma}{dt}(t))dt. \end{aligned}$$

□

Eine Teimenge $K \subset \mathbb{R}^n$ heißt konvex, wenn für alle $x, y \in K$ gilt: $[x,y] \subset K$.

Satz 10.20. *(Schrankensatz) Sei $f\colon U \to \mathbb{R}^m$ stetig differenzierbar und sei $K \subset U$ kompakt und konvex. Dann ist f auf K Lipschitz-stetig. Sei*

$$L = \max_{\xi \in K} \| \, df(\xi) \, \| \, .$$

Dann gilt für alle $x, y \in K$:

$$\| \, f(y) - f(x) \, \| \leq L \, \| \, y - x \, \| \, .$$

Beweis: Seien $x, y \in K$ und sei

$$\gamma(t) = (1-t)x + ty, \quad t \in [0,1].$$

Da K konvex ist, ist γ eine Kurve in K. Wir wenden Satz 10.19 auf die Kurve γ an. Dann ist

$$\begin{aligned} \| \, f(y) - f(x) \, \| &= \| \int_0^1 df((1-t)x + ty)(y-x)dt \, \| \\ &\leq \int_0^1 \| \, df((1-t)x + ty) \, \| \, dt \, \| \, y - x \, \| \, . \end{aligned}$$

Da K kompakt und $\| \, df(\xi) \, \|$ stetig auf K ist, folgt aus Satz 6.31, daß

$$L = \max_{\xi \in K} \| \, df(\xi) \, \|$$

existiert und endlich ist. Daraus erhalten wir

$$\| \, f(x) - f(y) \, \| \leq L \, \| \, x - y \, \| \, .$$

□

10.6. Höhere Ableitungen

Definition 10.21. *Sei $U \in \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ partiell differenzierbar. Wenn die partiellen Ableitungen*

$$\partial_i f\colon U \to \mathbb{R}, \ \ i = 1, \dots, n,$$

wieder differenzierbar sind, so nennt man

$$\partial_{ij} f := \partial_i(\partial_j f) = \partial_i \partial_j f$$

partielle Ableitungen 2. Ordnung.

Bezeichnungen:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \ f_{x_j x_i}.$$

Analog bildet man partielle Ableitungen k-ter Ordnung nach $x_1, \dots, x_k$:

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}.$$

Vertauschung der Reihenfolge der Differentation

Im allgemeinen hängen die höheren partiellen Ableitungen von der Reihenfolge ab, in der differenziert wird, d.h., im allgemeinen ist

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \neq \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \text{für} \quad i \neq j.$$

Beispiel: Die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2}, & (x,y) \neq (0,0); \\ 0, & (x,y) = (0,0). \end{cases}$$

Dann ist

$$\operatorname{grad} f(x,y) = \left(\frac{y^3(y^2-x^2)}{(x^2+y^2)^2}, \frac{xy^2(3x^2+y^2)}{(x^2+y^2)^2} \right).$$

Insbesondere ist

$$\partial_x f(0,y) = y, \; \partial_y f(x,0) = 0.$$

Daraus folgt

$$f_{xy}(0,y) = 1 \text{ für alle } y;$$
$$f_{yx}(x,0) = 0 \text{ für alle } x.$$

Da dies auch für $x = y = 0$ gilt, folgt

$$f_{xy}(0,0) = 1, \quad f_{yx}(0,0) = 0.$$

□

Satz 10.22. *(H.A. Schwarz): Sei $U \subset \mathbb{R}^n$ offen und sei $f\colon U \to \mathbb{R}$ eine Funktion, für die die partiellen Ableitungen $\partial_i f$, $\partial_j f$ und $\partial_{ij} f$ in U existieren. Weiter sei $\partial_{ij} f$ in $x_0 \in U$ stetig. Dann existiert $\partial_{ji} f$ in x_0 und es gilt*

$$\partial_{ij} f(x_0) = \partial_{ji} f(x_0).$$

Beweis: Sei $e_1, \dots, e_n \in \mathbb{R}^n$ die Standardbasis. Da U offen ist, existiert $c > 0$ so, daß $x_0 + se_i + te_j \in U$ für alle $s, t \in \mathbb{R}$ mit $|s|, |t| < c$. Sei

$$\varphi(s,t) = f(x_0 + se_i + te_j), \; s,t \in (-c,c).$$

Dann ist $\varphi\colon (-c,c) \times (-c,c) \to \mathbb{R}$ eine Funktion, für die gilt:

$\partial_1\varphi$, $\partial_2\varphi$ und $\partial_{12}\varphi$ existieren in $(-c,c)\times(-c,c)$ und $\partial_{12}\varphi(s,t)$ ist stetig in $(0,0)$.

Zu zeigen: $\partial_{21}\varphi(0,0)$ existiert und

$$\partial_{21}\varphi(0,0) = \partial_{12}\varphi(0,0).$$

Für $s,t \in (-c,c),\ s,t \neq 0$, sei

$$D\varphi(s,t) = \frac{(\varphi(s,t)-\varphi(0,t)) - (\varphi(s,0)-\varphi(0,0))}{st}$$

Sei $s \neq 0$ fest und sei $g\colon (-c,c) \to \mathbb{R}$ definiert durch

$$g(t) = \varphi(s,t) - \varphi(0,t).$$

Dann ist

$$D\varphi(s,t) = \frac{1}{s}\left(\frac{g(t)-g(0)}{t}\right).$$

Aus dem Mittelwertsatz folgt, daß ein $\theta \in (0,1)$ existiert, so daß

$$D\varphi(s,t) = \frac{1}{s}g'(\theta t) = \frac{\partial_2\varphi(s,\theta t) - \partial_2\varphi(0,\theta t)}{s}.$$

Anwendung des Mittelwertsatzes bezüglich s ergibt:

$$\exists \eta \in (0,1)\colon D\varphi(s,t) = \partial_{12}\varphi(\eta s, \theta t).$$

Da $\partial_{12}\varphi(s,t)$ in $(0,0)$ stetig ist, folgt:

$$\forall \epsilon > 0\ \exists \delta > 0\colon |D\varphi(s,t) - \partial_{12}\varphi(0,0)| < \epsilon$$

für alle s,t mit $|s|,|t| < \delta$. Weiter gilt:

$$\begin{aligned}\lim_{s\to 0} D\varphi(s,t) &= \lim_{s\to 0}\frac{1}{t}\left(\frac{\varphi(s,t)-\varphi(0,t)}{s} - \frac{\varphi(s,0)-\varphi(0,0)}{s}\right)\\ &= \frac{\partial_1\varphi(0,t) - \partial_1\varphi(0,0)}{t}\end{aligned}$$

Daraus folgt

$$\left|\frac{\partial_1\varphi(0,t)-\partial_1\varphi(0,0)}{t} - \partial_{12}\varphi(0,0)\right| < \epsilon$$

für $|t| < \delta$. Dies bedeutet, daß

$$\partial_{21}\varphi(0,0) = \lim_{t\to 0}\frac{\partial_1\varphi(0,1) - \partial_1\varphi(0,0)}{t}$$

existiert und es gilt:

$$\partial_{21}\varphi(0,0) = \partial_{12}\varphi(0,0).$$

$\square$

Beispiel-Anwendung: Sei $V\colon \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch $V(x,y) = (-y,x)$.
Frage: Existiert $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit $\mathrm{grad} f = V$?
Angenommen f existiert. Dann ist

$$\partial_x f(x,y) = -y,\ \partial_y f(x,y) = x.$$

Daraus folgt

$$\partial_{yx} f = -1 \quad \text{und} \quad \partial_{xy} f = 1.$$

Diese Funktionen sind stetig. Daraus erhalten wir einen Widerspruch zu Satz 10.22 und daher existiert f nicht.

Definition 10.23. *Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $f\colon U \to \mathbb{C}$ heißt* **k-mal stetig differenzierbar**, *wenn alle partiellen Ableitungen*

$$\partial_{i_1} \dots \partial_{i_k} f,\ i_1, \dots, i_k \in \{1, \dots, n\},$$

auf U existieren und stetig sind. Weiter sei

$$C^k(U) := \{f\colon U \to \mathbb{C} \mid f \text{ k-mal stetig differenzierbar}\}$$
$$C^\infty(U) := \bigcap_{k=1}^{\infty} C^k(U).$$

$C^\infty(U)$ ist der Raum der beliebig oft differenzierbaren Funktionen auf U. Sei $f \in C^k(U)$. Man sagt auch “f ist von der Klasse C^k”.

Korollar 10.24. *Sei $f \in C^k(U)$ und sei $\sigma \in S_k$ eine Permutation. Dann ist*

$$\partial_{i_k} \dots \partial_{i_k} f = \partial_{i_{\sigma(1)}} \dots \partial_{i_{\sigma(k)}} f.$$

Lineare Differentialoperatoren

Sei $P(x) \in \mathbb{C}[x_1, \dots, x_n]$ ein Poynom vom Grade m:

$$P(x) = \sum_{k_1 + \dots + k_n \le m} a_{k_1 \cdots k_n} x_1^{k_1} \cdots x_n^{k_n}, \quad a_{k_1 \cdots k_n} \in \mathbb{C}.$$

Die Ersetzung $x_i \longmapsto \partial_i$ ergibt

$$P(D) = \sum_{k_1 + \dots + k_n \le m} a_{k_1 \cdots k_n} \partial_1^{k_1} \cdots \partial_n^{k_n}.$$

Dies definiert eine lineare Abbildung $P(D)\colon C^m(U) \to C^0(U)$ durch

$$(P(D)f)(x) = \sum_{k_1 + \dots + k_n \le m} a_{k_1 \cdots k_n} \frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}}(x).$$

$P(D)$ ist ein linearer Differentialoperator mit konstanten Koeffizienten.

Bemerkung: Für $n = 1$ ist $P(D)$ ein gewöhnlicher Differentialoperator.

Beispiele:

(1) Laplace-Operator

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}.$$

(2) Wellenoperator

$$\frac{\partial^2}{\partial t^2} - \Delta.$$

(3) Wärmeleitungsoperator

$$\frac{\partial}{\partial t} - \Delta.$$

Eine Gleichung der Art

$$\boxed{\text{P(D) f=0}}$$

nennt man **lineare partielle Differentialgleichung**.

Eines der Hauptprobleme der Analysis ist die Untersuchung des Lösungsverhaltens von partiellen Differentialgleichungen.

Viele Gleichungen der Physik sind partielle Differentialgleichungen.

Beispiele:

1.Wellengleichung:

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi = \Delta\psi, \quad \psi(0) = \psi_0,\ \psi'(0) = \psi_1.$$

2.Laplace-Gleichung:

$$\Delta\psi = \lambda\psi.$$

3.Wärmeleitungsgleichung:

$$\frac{\partial}{\partial t}\psi = \Delta\psi, \quad \lim_{t\to 0}\psi(t,x) = \psi_0(x)$$

4.Schrödinger-Gleichung:

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi, \quad H = \text{Hamiltonoperator}$$

Z.B.: $H = -\Delta + V,\ V \in C^\infty(\mathbb{R}^n)$.

10.7. Taylorapproximation

Zunächst erinnern wir uns an den 1-dimensionalen Fall. Es sei $f \in C^{p+1}(I)$, und seien $0, t \in I$. Dann gilt

$$f(t) = T_p f(t) + R_{p+1}(t)$$

wobei

$$T_p f(t) = \sum_{k=0}^{p} \frac{f^{(k)}(0)}{k!} t^k$$

und

$$R_{p+1}(t) = \frac{f^{(p+1)}(\theta t)}{(p+1)!} t^{p+1}$$

für ein $\theta \in (0,1)$. Dies löst das Problem, die Funktion f in der Umgebung von $t = 0$ durch ein Polynom zu approximieren. Wir möchten dies für den Fall von mehreren Variablen verallgemeinern. Es sei $U \subset \mathbb{R}^n$ offen und $f \in C^{p+1}(U)$. Seien $x_0, x \in U$ so, daß $[x_0, x] \subset U$. Sei $h = x - x_0$. Wir definieren $g\colon [0,1] \to \mathbb{R}$ durch

$$g(t) = f(x_0 + th).$$

Dann ist $g \in C^{p+1}((0,1))$ und

$$\begin{aligned} f(x) &= f(x_0 + h) = g(1); \\ f(x_0) &= g(0). \end{aligned}$$

Anwendung der Taylorformel für Funktionen einer reellen Variablen ergibt: Es existiert $\theta \in (0,1)$ mit

$$g(1) = T_p g(1) + R_{p+1}(1) = \sum_{k=0}^{p} \frac{g^{(k)}(0)}{k!} + \frac{g^{(p+1)}(\theta)}{(p+1)!}.$$

Berechnung von $g^{(k)}(0)$.

1. $g(0) = f(x_0)$
2. Nach der Kettenregel ist

$$\frac{dg}{dt}(t) = \frac{df(x_0 + th)}{dt} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0 + th) h_j = \langle \mathrm{grad} f(x_0 + th), h \rangle.$$

Sei $\nabla_h\colon C^k(U) \to C^{k-1}(U)$ definiert durch

$$(\nabla_h f)(x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) h_j.$$

Dann ist

$$g'(t) = (\nabla_h f)(x_0 + th).$$

3. Durch Anwendung der Kettenregel erhalten wir

$$\frac{d^2g}{dt^2}(t) = \frac{d}{dt}(\nabla_h)f(x_0+th) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(\nabla_h f)(x_0+th)h_j.$$

Hieraus folgt $g''(t) = ((\nabla_h)^2 f)\,(x_0 + th)$. Durch Induktion erhalten wir

$$\frac{d^kg}{dt^k}(t) = \big((\nabla_h)^k f\big)\,(x_0 + th).$$

Wir haben damit den folgenden Satz bewiesen.

Satz 10.25. *Sei $U \subset \mathbb{R}^n$ offen und sei $f \in C^{p+1}(U)$. Seien $x, x_0 \in U$ so, daß $[x_0, x] \subset U$. Dann existiert $\theta \in (0,1)$:*

$$\begin{aligned} f(x) = f(x_0) &+ \frac{(\nabla_h f)(x_0)}{1!} + \frac{((\nabla_h)^2 f)(x_0)}{2!} \\ &+ \cdots + \frac{((\nabla_h)^p f)\,(x_0)}{p!} + \frac{((\nabla_h)^{p+1} f)\,(x_0+\theta h)}{(p+1)!}. \end{aligned}$$

Dies ist noch keine brauchbare Formel. Wir müssen $(\nabla_h)^k f$ durch die partiellen Ableitungen von f ausdrücken. Um die Formeln, die wir daraus erhalten, übersichtlich aufschreiben zu können, führen wir einige Bezeichnungen ein.

Bezeichnungen: Sei $\alpha \in \mathbb{N}_0^n$, $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_i \in \mathbb{N}_0$, und $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Dann definieren wir

$$\alpha! = \alpha_1! \cdots \alpha_n!,\ |\alpha| = \alpha_1 + \cdots + \alpha_n,\ x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

Lemma. *Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt*

$$(x_1 + \cdots + x_n)^k = k! \sum_{|\alpha|=k} \frac{x^\alpha}{\alpha!}$$

Beweis: Der Beweis erfolgt durch Induktion.

1. $n = 1$: trivial

2. $n = 2$:

$$(x_1 + x_2)^k = k! \sum_{p=0}^{k} \frac{x_1^p x_2^{k-p}}{p!(k-p)!} = \sum_{p=0}^{n} \binom{n}{p} x_1^p x_2^{k-p}.$$

3. **Schluß von** $(n-1)$ **auf** n: Sei $x = (y, x_n)$, $y = (x_1, \dots, x_{n-1})$, $\alpha = (\beta, \alpha_n)$, $\beta = (\alpha_1, \dots, \alpha_{n-1})$. Dann ist

$$\begin{aligned}
(x_1 + \cdots + x_n)^k &= \sum_{\alpha_n=0}^{k} \binom{k}{\alpha_n} (x_1 + \cdots + x_{n-1})^{k-\alpha_n} x_n^{\alpha_n} \\
&= \sum_{\alpha_n=0}^{k} \sum_{|\beta|=k-\alpha_n} \frac{k!}{\alpha_n!(k-\alpha_n)!} \frac{(k-\alpha_n)!}{\beta!} y^\beta x_n^{\alpha_n} \\
&= \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha.
\end{aligned}$$

□

Aus dem Lemma folgt:

$$\frac{(\nabla_h)^k}{k!} = \frac{(h_1\partial_1 + \cdots + h_n\partial_n)^k}{k!} = \sum_{|\alpha|=k} \frac{h^\alpha}{\alpha!} \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

Für $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ sei

$$\boxed{D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.}$$

Definition 10.26. *Für $f \in C^p(U)$ sei*

$$T_p f(h; x_0) = \sum_{|\alpha| \le p} \frac{D^\alpha f(x_0)}{\alpha!} h^\alpha$$

das Taylorpolynom der Ordnung p.

Satz 10.27. *(Satz von Taylor) Sei $f \in C^{p+1}(U)$ und seien $x_0, x \in U$ mit $[x_0, x] \subset U$. Sei $h = x - x_0$. Dann existiert $\theta \in (0,1)$:*

$$f(x_0 + h) = T_p f(h; x_0) + R_{p+1}(x)$$

und

$$R_{p+1}(x) = \sum_{|\alpha|=p+1} \frac{D^\alpha f(x_0 + \theta h)}{\alpha!} h^\alpha.$$

Korollar 10.28. *Seien $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ p-mal stetig differenzierbar. Weiter sei $x \in U$ und $\delta > 0$ so, daß $U_\delta(x) \subset U$. Dann* *gilt für* $\| h \| < \delta$:

$$f(x+h) = \sum_{|\alpha| \le p} \frac{D^\alpha f(x)}{\alpha!} h^\alpha + o(\| h \|^p).$$

d.h. es gilt

$$\lim_{h \to 0} \frac{f(x+h) - T_p f(h;x)}{\| h \|^p} = 0.$$

Beweis: Sei $h \in U_\delta(0)$. Dann existiert $\theta \in (0,1)$:

$$f(x+h) = T_{p-1} f(h;x) + \sum_{|\alpha|=p} \frac{D^\alpha f(x+\theta h)}{|\alpha|!} h^\alpha.$$

Da

$$D^\alpha f(x) = \frac{\partial^{\alpha_1 + \cdots + \alpha_n} f}{\partial x_n^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(x)$$

in U stetig ist, existiert zu $\epsilon > 0$ ein $\eta > 0$:

$$\sum_{|\alpha|=p} \frac{1}{\alpha!} |D^\alpha f(x+h) - D^\alpha f(x)| < \epsilon \quad \text{für} \quad \| h \| < \eta.$$

Aus Satz 10.27 folgt:

$$\begin{aligned} f(x+h) = T_{p-1}(h;x) + \sum_{|\alpha|=p} \frac{D^\alpha f(x)}{\alpha!} h^\alpha \\ + \sum_{|\alpha|=p} \frac{1}{\alpha!} \left(D^\alpha f(x+\theta h) - D^\alpha f(x) \right) h^\alpha \\ = T_p f(h;x) + \sum_{|\alpha|=p} \frac{1}{\alpha!} \left(D^\alpha f(x+\theta h) - D^\alpha f(x) \right) h^\alpha. \end{aligned}$$

Weiter ist

$$|h^\alpha| \le \max\{|h_1|, \ldots, |h_n|\}^p \le \| h \|^p .$$

Daraus folgt für $0 < \| h \| < \eta$:

$$\frac{|f(x+h) - T_p f(h;x)|}{\| h \|^p} \le \sum_{|\alpha|=p} \frac{1}{\alpha!} |D^\alpha f(x+\theta h) - D^\alpha f(x)| < \epsilon.$$

$\square$

10.8. Extrema von Funktionen

Definition 10.29. *Sei $X \subset \mathbb{R}^n$ und $f\colon X \to \mathbb{R}$ eine Funktion.*

(1) *$x_0 \in X$ heißt* **Maximum** *(bzw.* **Minimum***) von f, wenn gilt*

$$\forall x \in X: \; f(x_0) \geq f(x) \text{ (bzw. } \forall x \in X\colon f(x) \geq f(x_0)).$$

Der Punkt x_0 heißt **Extremum** *von f.*

(2) *$x_0 \in X$ heißt* **lokales Maximum** *(bzw.* **lokales Minimum***) von f , wenn $\epsilon > 0$ existiert, so daß x_0 Maximum (bzw. Minimum) von $f\colon U_\epsilon(x_0) \cap X \to \mathbb{R}$ ist . x_0 heißt in diesem Falle (lokales) Extremum von f.*

Der folgende Satz liefert eine notwendige Bedingung für die Existenz eines Extremums.

Satz 10.30. *Sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$. Es sei $x_0 \in U$ ein Extremum von f und f sei in x_0 partiell differenzierbar. Dann gilt*

$$\operatorname{grad} f(x_0) = 0,$$

d.h., $\partial_1 f(x_0) = \cdots = \partial_n f(x_0) = 0$. Wenn f in x_0 total differenzierbar ist, so ist $df(x_0) = 0$.

Beweis: Es sei $e_1 \ldots , e_n \in \mathbb{R}^n$ die Standardbasis und

$$g_i(t) = f(x_0 + te_i), \quad i = 1, \ldots , n.$$

Dann hat $g_i(t)$ in $t = 0$ ein lokales Extremum und g_i ist in $t = 0$ differenzierbar. Daraus folgt

$$0 = \frac{dg_i}{dt}(t) = \partial_i f(x_0).$$

□

Definition 10.31. *Sei $U \subset \mathbb{R}^n$ offen und $f \in C^1(U)$. Der Punkt $x_0 \in U$ heißt* **kritischer Punkt** *von f, wenn $df(x_0) = 0$. Für einen kritischen Punkt $x_0 \in U$ heißt $f(x_0)$* **kritischer Wert***.*

Man beachte: Lokale Extrema sind kritische Punkte. Die Umkehrung gilt nicht.

Beispiel: Sei $f\colon \mathbb{R}^3 \to \mathbb{R}$ definiert durch $f(x, y) = x^2 - y^2$. Dann ist $df(0, 0) = 0$. Aber $(0, 0)$ ist kein Extemum von f.

Hinreichendes Kriterium für Extrema

Wir erinnern uns an die Situation im Falle $n = 1$. Es sei $f\colon (a, b) \to \mathbb{R}$ 2-mal stetig differenzierbar. Sei $x_0 \in (a, b)$ und $f'(x_0) = 0$. Dann gilt:

(1) $f''(x_0) > 0 \Rightarrow f$ hat ein lokales Minimum in x_0.

(2) $f''(x_0) < 0 \Rightarrow f$ hat ein loklaes Maximum in x_0.

Frage: Wie kann man dies auf den Fall mehrerer Veränderlicher verallgemeinern?

Die Antwort liefert den Satz von Taylor. Im Falle $n = 1$ ist

$$f(x+h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + o(h^2).$$

Mit Hilfe dieser Gleichung kann man 1) und 2) beweisen. Sei jetzt n beliebig. Sei $f \in C^2(U)$ und $x \in U$. Aus Korollar 10.28 folgt: Es existiert $\delta > 0$ so, daß für alle $h \in \mathbb{R}^n$ mit $\| h \| < \delta$ gilt:

$$f(x+h) = f(x) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)h_j + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)h_i h_j + R(h)$$

mit

$$\lim_{h \to 0} \frac{R(h)}{\| h \|^2} = 0.$$

Dies ist die Taylorapproximation bis zur 2. Ordnung. Sei

$$d^2 f(x)(u,v) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)u_i v_j,\ u, v \in \mathbb{R}^h.$$

Die so definierte Funktion ist eine ymmetrische Bilinearform in $\mathbb{R}^n$. Es sei

$$d^2 f(x)(h,h) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} h_i h_j$$

die assoziierte quadratische Form.

Definition 10.32. *Die quadratische Form $d^2 f(x)(h,h)$ heißt* **Hesse-Form** *von f in $x \in U$. Die Matrix*

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1}^{n}$$

heißt **Hesse-Matrix** *von f in x.*

Bezeichnung: Für die Hesse-Matrix ist auch die Bezeichnung Hess$f(x)$ üblich.

Bemerkung: Sei $f \in C^2(U)$. Aus Satz 10.22. folgt: $H_f(x)$ ist symmetrisch. Es gilt

$$d^2 f(x)(h,h) = \langle h, H_f(x)h \rangle,$$

d.h., die symmetrische Matrix $H_f(x)$ repräsentiert die quadratische Form $d^2 f(x)(h,h)$.

Die Taylorapproximation 2. Ordnung kann man auch wie folgt schreiben:

$$\boxed{f(x+h) = f(x) + \langle \mathrm{grad} f(x), h \rangle + \frac{1}{2}\langle h, H_f(x)(h) \rangle + R(h).}$$

Sei $x \in U$ kritischer Punkt. Aus Satz 10.30 folgt $\mathrm{grad}\, f(x) = 0$. Daher hat die Taylorapproximation im kritischen Punkte x die Form

$$f(x+h) = f(x) + \frac{1}{2}\langle h, H_f(x)(h) \rangle + R(h).$$

Da $R(h) = o(\| h \|^2)$, wird das Verhalten von f in der Umgebung von x durch die quadratische Form $\langle h, H_f(x)(h) \rangle$ bestimmt. Wir erinnern uns an den folgenden Satz aus der linearen Algebra.

Satz. *Sei $A \in Mat_n(\mathbb{R})$ symmetrisch. Dann existiert eine orthonormierte Basis $v_1, \dots, v_n$ von $\mathbb{R}^n$ so, daß*

$$Av_i = \lambda_i v_i, \quad \lambda_i \in \mathbb{R}, \quad i = 1, \dots, n,$$

d.h., jede symmetrische Matrix kann diagonlalisiert werden.

Insbesondere folgt daraus für $H_f(x)$, daß $U \in O(n)$ existiert, so daß

$$U^t H_f(x) U = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Betrachten wir die quadratische Form:

$$Q(h) = \lambda_1 h_1^2 + \dots + \lambda_n h_n^2.$$

Es genügt, das Verhalten von $Q(h)$ in Umgebung von $h = 0$ zu untersuchen.

Fallunterscheidung:

(1) $\forall i$: $\lambda_i > 0$. Dann ist 0 isoliertes Minimum von Q.
(2) $\forall i$: $\lambda_i < 0$. Dann ist 0 isoliertes Maximum von Q.

(3) $\exists i, j, i \neq j\colon\ \lambda_i > 0,\ \lambda_j < 0$. Dann ist 0 kein Extremum von Q.
Beispiel: $Q(h) = h_1^2 - h_2^2$
In den Fällen 1) - 3) ist Q nicht ausgeartet.
(4) $\exists i,\ 1 \leq i \leq n\colon \lambda_i = 0$. Dann ist $Q(h)$ ausgeartet.
Beispiel: Sei $Q(h) = \lambda_1 x_1^2 + \cdots \lambda_k x_k^2$ mit $k < n$ und $\lambda_1, \ldots, \lambda_k > 0$. Dann ist zwar 0 ein Minimum von Q, aber es ist kein isoliertes Minimum.

Es sei $q\colon \mathbb{R}^n \to \mathbb{R}$ eine quadratische Form. Dann existiert eine symmetrische Matrix A so, dass $q(x) = \langle x, Ax \rangle,\ x \in \mathbb{R}^n$, d.h., A repräsentiert die quadratische Form q.

Definition. *Q heißt:*

positiv definit	$\Leftrightarrow \forall x \in \mathbb{R}^n :$	$Q(x) > 0$
positiv semidefinit	$\Leftrightarrow x \in \mathbb{R}^n :$	$Q(x) \geq 0$
negativ definit	$\Leftrightarrow x \in \mathbb{R}^n :$	$Q(x) < 0$
negativ semidefinit	$\Leftrightarrow \forall x \in \mathbb{R}^n :$	$Q(x) \leq 0$
indefinit	$\Leftrightarrow \exists x, y :$	$Q(x) > 0$ *und* $Q(y) < 0$,
nicht ausgeartet	$\Leftrightarrow$	$\det A \neq 0$.

Satz 10.34. *Sei $f \in C^2(U)$ und $x \in U$ mit $df(x) = 0$. Dann gilt:*
(1) *$d^2 f(x) > 0 \Rightarrow x$ ist lokales isoliertes Minimum.*
(2) *$d^2 f(x) < 0 \Rightarrow x$ ist lokales isoliertes Maximum.*
(3) *$d^2 f(x)$ indefinit $\Rightarrow x$ ist kein lokales Extremum.*

Beweis:
1. Sei $d^2 f(x) > 0$. Nach Voraussetzung ist $df(x) = 0$. Wie wir bereits oben gesehen haben, hat dann die Taylorformel 2. Ordnung die Gestalt

$$f(x+h) = f(x) + \frac{1}{2}\langle h, H_f(x)(h)\rangle + R(h),$$

wobei $\lim_{h \to 0} R(h)/ \parallel h \parallel^2 = 0$ gilt. Da $H_f(x)$ symmetrisch ist, existiert $U \in O(n)$ so, daß

$$UH_f(x)U^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

und $\lambda_i \in \mathbb{R}$. Nach Voraussetzung ist $H_f(x) > 0$. Daraus folgt: $\forall i\colon \lambda_i > 0$. Deshalb existiert $m > 0$ so, daß

$$\langle h, H_f(x)(h)\rangle \geq m \parallel h \parallel^2$$

für alle $h \in \mathbb{R}^n$. Da $\lim_{h \to 0} R(h)/ \parallel h \parallel^2 = 0$, existiert $\epsilon > 0$:

$$\forall h \in \mathbb{R}^n\colon\ \parallel h \parallel < \epsilon \Rightarrow |R(h)| < \frac{m}{4} \parallel h \parallel^2 .$$

Sei $\| h \| < \epsilon$. Dann folgt

$$\begin{aligned} f(x+h) &= f(x) + \frac{1}{2}\langle h, H_f(x)h\rangle + R(h) \\ &\geq f(x) + \frac{m}{2} \| h \|^2 - |R(h)| \\ &\geq f(x) + \frac{m}{4} \| h \|^2 . \end{aligned}$$

Aus dieser Ungleichung folgt, daß x ein lokales isoliertes Minimum von f ist.

2. Dieser Fall wird analog behandelt.

3. Wenn $d^2 f(x)$ indefinit ist, so existieren $u, v \in \mathbb{R}^n$ mit $\langle u, H_f(x)(u)\rangle > 0$ und $\langle v, H_f(x)(v)\rangle < 0$.

Sei $F_u(t) = f(x+tu)$, $t \in \mathbb{R}$, $|t| < \epsilon$. Dann ist $F_u\colon (-\epsilon, \epsilon) \to \mathbb{R}$ 2-mal stetig differenzierbar und es gilt:

$$\begin{aligned} F_u'(0) &= \partial_u f(x) = df(x)(u) = 0, \\ F_u''(0) &= \langle u, H_f(x)(u)\rangle > 0. \end{aligned}$$

Daraus folgt, daß F_u in $t = 0$ ein isoliertes lokales Minimum hat. Entsprechend folgt: $F_v(t) = f(x+tv)$ hat in $t = 0$ ein isoliertes lokales Maximum. Daher ist x kein Extremum von f.

□

Im Falle $n = 2$ ist es leicht, zu entscheiden, ob eine gegebene Matrix positiv oder negativ definit ist. Dies wird durch folgendes Lemma beantwortet.

Lemma 10.35. *Sei $n = 2$, $Q(x) = \langle x, Ax\rangle$ mit $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Dann gilt*

$$\begin{aligned} &Q > 0 \Leftrightarrow a > 0 \textit{ und } \det A > 0 \\ &Q < 0 \Leftrightarrow a < 0 \textit{ und } \det A > 0 \\ &Q \textit{ indefinit} \Leftrightarrow \det A < 0. \end{aligned}$$

Beweis:

$$\begin{aligned} Q(x) &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\ &= a\left(x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{b^2}{a^2}x_2^2\right) + \left(c - \frac{b^2}{a}\right)x_2^2 \\ &= a\left(x_1 + \frac{b}{a}x_2\right)^2 + \frac{1}{a}\left(ac - b^2\right)x_2^2. \end{aligned}$$

(1) $a > 0$, $\det A > 0 \Rightarrow Q > 0$. $Q > 0 \Rightarrow a > 0$, $\det A > 0$.
(2) $a < 0$, $\det A > 0 \Rightarrow Q < 0$. $Q < 0 \Rightarrow a < 0$, $\det A > 0$.

Für $n > 2$ gibt es den folgenden Satz von Sylvester, der für eine symmetrische reelle Matrix A ein notwendiges und hinreichendes Kriterium dafür angibt, daß $A > 0$ gilt.

Satz. *Es sei $A \in Mat_n(\mathbb{R})$ symmetrisch. Dann ist A positiv definit genau dann, wenn gilt:*

$$\forall k \in \{1, ...n\} : \det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0.$$

10.9. Der Umkehrsatz

In diesem Abschnitt untersuchen wir, wann eine stetig differenzierbare Abbildung $\phi \colon U \to V$ von offenen Mengen $U, V \subset \mathbb{R}^n$ eine stetige differenzierbare Umkehrung $\psi = \phi^{-1}$ besitzt. Abbildungen dieser Art spielen in der Analysis die gleiche Rolle wie die Isomorphismen in der Algebra.

Definition 10.36. *Es seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen und $\phi \colon U \to V$ sei eine C^1-Abbildung. Dann heißt ϕ* **Diffeomorphismus**, *wenn eine C^1-Abbildung $\psi \colon V \to U$ existiert, so daß gilt:*

$$\psi \circ \phi = Id_U, \quad \phi \circ \psi = Id_V. \tag{10.1}$$

Dann ist $\psi = \phi^{-1}$ die Umkehrabbildung von ϕ.

Beispiele:

(1) Wie wir aus Kapitel 6 wissen, ist $\exp \colon \mathbb{R} \to \mathbb{R}^+$ eine eineindeutige C^∞-Abbildung. Die Umkehrabbildung $\ln = \exp^{-1}$ ist eine C^∞-Abbildung von $\mathbb{R}^+$ auf $\mathbb{R}$. Damit ist $\exp \colon \mathbb{R} \to \mathbb{R}^+$ ein Diffeomorphismus.

(2) Es sei $f \colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = x^3,\ x \in \mathbb{R}$. Dann ist f eine eineindeutige C^∞-Abbildung von $\mathbb{R}$ auf $\mathbb{R}$. Die Umkehrabbildung $f^{-1}(x) = \sqrt[3]{x}$ ist jedoch in $x = 0$ nicht differenzierbar. Deshalb ist f **kein** Diffeomorphismus.

Wir stellen zunächst einige elementare Eigenschaften eines Diffeomorphismus $\phi \colon U \to V$ zusammen. Es sei $\psi = \phi^{-1}$ die Umkehrabbildung. Es sei $x \in U$ und $y = \phi(x) \in V$. Die Anwendung der Kettenregel auf (10.1) ergibt

$$d\psi(y) \circ d\phi(x) = \mathrm{Id}_{\mathbb{R}^n},\ d\phi(x) \circ d\psi(y) = \mathrm{Id}_{\mathbb{R}^m}.$$

Aus der linearen Algebra erhalten wir

Lemma 10.37. *Es sei $\phi\colon U \to V$ ein Diffeomorphismus offener Mengen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$. Dann gilt:*

(1) *$n = m$.*
(2) *Für alle $x \in U$ ist $d\phi(x)\colon \mathbb{R}^n \to \mathbb{R}^n$ ein Isomorphismus und es gilt*

$$(d\phi^{-1})(y) = (d\phi(x))^{-1},$$

wobei $y = \phi(x)$.

Bemerkung:

(1) Aus dem Lemma folgt insbesondere, daß für $n \neq m$, $\mathbb{R}^n$ nicht diffeomorph auf $\mathbb{R}^m$ abgebildet werden kann.
(2) Wie Lemma 10.37 zeigt, ist die Invertierbarkeit der Differentiale $d\phi(x)\colon \mathbb{R}^n \to \mathbb{R}^n$ in jedem Punkt $x \in U$ eine notwendige Bedingung dafür, daß $\phi\colon U \to V$ ein Diffeomorphismus ist. Für $n = 1$ ist dies auch hinreichend. Denn es sei $\phi\colon I \to J$ eine surjektive C^1-Abbildung eines offenen Intervalls $I \subset \mathbb{R}^1$ auf ein offenes Intervall $J \subset \mathbb{R}^1$. Weiter sei $\phi'(x) \neq 0$ für alle $x \in I$. Dann gilt entweder $\phi'(x) > 0$ für alle $x \in I$ oder $\phi'(x) < 0$ für alle $x \in I$, d.h. ϕ ist streng monoton. Nach Satz 7.7 ist dann die Umkehrfunktion $\psi\colon J \to I$ stetig differenzierbar.

Für $n > 1$ ist dies nicht mehr richtig, d.h., die Invertierbarkeit der Differentiale $d\phi(x)$ für jedes $x \in U$ ist **nicht** hinreichend für die Existenz einer differenzierbaren Umkehrabbildung. Dies zeigt folgendes Beispiel. Es sei

$$P\colon \mathbb{R}^2 \to \mathbb{R}^2$$

definiert durch

$$P(r, \varphi) = (r\cos\varphi, r\sin\varphi). \tag{10.2}$$

Dann bildet P den offenen Streifen $\mathbb{R}_+ \times (-\pi, \pi)$ **homöomorph** auf $\mathbb{R}^2 - \{(x, 0) \mid x \leq 0\}$ ab. Homöomorph bedeutet dabei, daß die Umkehrabbildung existiert und stetig ist. Die Abbildung

$$P_1\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R} - \{(x, 0) \mid x \leq 0\},$$

die durch die Einschränkung von P auf $\mathbb{R}^+ \times \mathbb{R}$ definiert wird, ist eine surjektive C^1-Abbildung. Für das Differential gilt

$$dP_1(r, \varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Die Determinante dieser Matrix ist gleich r. Daher ist $dP_1(r, \varphi)$ in jedem Punkt $(r, \varphi) \in \mathbb{R}^+ \times \mathbb{R}$ invertierbar. Andererseits gilt

$$P_1(r, \varphi + 2\pi) = P_1(r, \varphi)$$

für alle $(r, \varphi) \in \mathbb{R}^+ \times \mathbb{R}$. Daher besitzt P_1 nicht einmal eine Umkehrabbildung.

Wie wir in obigem Beispiel gesehen haben reicht die Invertierbarkeit der Differentiale $d\phi(x)$, $x \in U$, im allgemeinen nicht aus, damit die C^1-Abbildung $\phi\colon U \to V$ ein Diffeomorphismus ist. Wenn wir jedoch zusätzlich annehmen, daß ϕ eine stetige Umkehrung besitzt, so ist die Invertierbarkeit der Differentiale $d\phi(x)$, $x \in U$, hinreichend für die Existenz einer stetig differenzierbaren Umkehrung. Und zwar gilt

Satz 10.38. *Es sei $\phi\colon U \to V$ eine C^1-Abbildung einer offenen Menge $U \subset \mathbb{R}^n$ auf eine offene Menge $V \subset \mathbb{R}^n$ mit folgenden Eigenschaften:*

(1) *ϕ besitzt eine stetige Umkehrung.*

(2) *Für alle $x \in U$ ist $d\phi(x)$ ein Isomorphismus.*

Dann ist φ ein Diffeomorphismus.

Beweis:

(1) Wir zeigen, daß $\psi = \phi^{-1}$ in jedem Punkt $y \in V$ differenzierbar ist. Es sei $k \in \mathbb{R}^n$ so, daß $y + k \in V$. Wir setzen

$$x := \psi(y), \quad h := \psi(y+k) - \psi(y). \tag{10.3}$$

Da ϕ in $x \in U$ differenzierbar ist, gilt:

$$\phi(x+h) = \phi(x) + d\phi(x)(h) + R(h)$$

mit

$$\lim_{h \to 0} \frac{R(h)}{\| h \|} = 0. \tag{10.4}$$

In dieser Gleichung ersetzen wir x und h durch die rechten Seiten von (10.3) und benutzen, daß $\phi \circ \psi = \mathrm{Id}$ gilt. Dann erhalten wir

$$k = d\phi(x)(\psi(y+k) - \psi(y)) + R(\psi(y+k) - \psi(y)).$$

Da $d\phi(x)$ invertierbar ist, folgt daraus

$$\psi(y+k) = \psi(y) + (d\phi(x))^{-1}(k) + R_1(k) \tag{10.5}$$

wobei

$$R_1(k) := -(d\phi(x))^{-1} R(\psi(y+k) - \psi(y)). \tag{10.6}$$

Wir schätzen jetzt $R_1(k)$ für $k \to 0$ ab. Es sei

$$c = \| (d\phi(x))^{-1} \| .$$

Da für $R(h)$ die Bedingung (10.4) gilt, existiert $r > 0$ so, daß

$$\| R(h) \| \leq \frac{1}{2c} \| h \|$$

für alle $h \in \mathbb{R}^n$ mit $\| h \| \leq r$. Da ψ stetig ist, existiert ein $\delta > 0$ so, daß

$$\| h \| = \| \psi(y+k) - \psi(y) \| \leq r$$

für alle $k \in \mathbb{R}^n$ mit $\| k \| \leq \delta$. Aus (10.6) folgt für diese k:

$$\begin{aligned} \| R_1(k) \| &\leq \| (d\phi(x))^{-1} \| \, \| R(\psi(y+k) - \psi(y)) \| \\ &\leq c\frac{1}{2c} \| \psi(y+k) - \psi(y) \| = \frac{1}{2} \| \psi(y+k) - \psi(y) \| \,. \end{aligned}$$

Weiter folgt aus (10.5)

$$\| \psi(y+k) - \psi(y) \| \leq c \| k \| + \| R_1(k) \| \,.$$

Aus diesen beiden Ungleichungen erhalten wir

$$\| \psi(y+k) - \psi(y) \| \leq 2c \| k \| \,.$$

Aus (10.6) erhalten wir zusammen mit dieser Abschätzung und (10.3)

$$\begin{aligned} \frac{\| R_1(k) \|}{\| k \|} &\leq c \, \frac{\| R(\psi(y+k) - \psi(y) \|}{\| k \|} \\ &= c \, \frac{\| \psi(y+k) - \psi(y) \|}{\| k \|} \frac{\| R(\psi(y+k) - \psi(y)) \|}{\| \psi(y+k) - \psi(y) \|} \\ &\leq 2c^2 \, \frac{\| R(h) \|}{\| h \|} \end{aligned}$$

für alle $k \neq 0$ mit $\| k \| \leq \delta$. Wegen der Stetigkeit von ψ folgt

$$\lim_{k \to 0} \psi(y+k) - \psi(y) = 0.$$

Nach Definition von h folgt daraus, daß $h \to 0$ für $k \to 0$ gilt. Aus obiger Abschätzung folgt damit

$$\lim_{k \to 0} \frac{\| R_1(k) \|}{\| k \|} = 0.$$

Aus (10.5) folgt daher, daß ψ differenzierbar ist.
Es bleibt noch zu zeigen, daß ψ stetig differenzierbar ist, d.h., daß die Abbildung

$$y \in V \longmapsto d\psi(y) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$$

stetig ist. Auf Grund von (10.5) gilt

$$d\psi(y) = (d\phi(\psi(y)))^{-1} \,.$$

Dies bedeutet, daß $d\psi$ die Komposition der Abbildungen $\psi\colon V \to U$, $d\phi\colon U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ und $A \in GL(\mathbb{R}^n) \longmapsto A^{-1} \in GL(\mathbb{R}^n)$ ist. Diese Abbildungen sind alle stetig. Daher ist auch $d\psi\colon V \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ stetig.

□

Ohne die Voraussetzung, daß ϕ eine stetige Umkehrung besitzt, ist der obige Satz nicht mehr richtig. Der Umkehrsatz gilt jedoch in der lokalen Form, d.h., daß ϕ in der Umgebung jedes Punktes $x \in U$ eine Umkehrung besitzt. Dazu benötigen wir den Banachschen Fixpunktsatz. Wir formulieren diesen Satz allgemein für metrische Räume. Deshalb müssen wir zuerst einige Begriffe aus der Theorie der metrischen Räume einführen.

Definition 10.39. *Eine Menge M zusammen mit einer Abbildung*

$$d \colon M \times M \to \mathbb{R}$$

heißt **metrischer Raum**, *wenn folgende Eigenschaften gelten:*

(1) $\forall x, y \in M \colon d(x,y) \geq 0$ *und* $d(x,y) = 0 \Leftrightarrow x = y$.
(2) $\forall x, y \in M \colon d(x,y) = d(y,x)$.
(3) $\forall x, y, z \in M \colon d(x,z) \leq d(x,y) + d(y,z)$.

Es sei (M, d) ein metrischer Raum. Eine Folge (x_n) von Elementen aus M heißt konvergent, wenn gilt:

$$\exists x \in M \, \forall \epsilon > 0 \, \exists N \in \mathbb{N} \colon d(x_n, x) < \epsilon \quad \text{für alle} \quad n \geq N.$$

Das Element x heißt **Grenzwert** der Folge (x_n):

$$x = \lim_{n \to \infty} x_n.$$

Eine Folge (x_n) in M heißt Cauchy-Folge, wenn gilt:

$$\forall \epsilon > 0 \, \exists N \in \mathbb{N} \, \forall n, m \geq N \colon$$

$$d(x_n, x_m) < \epsilon.$$

Ein metrischer Raum (M, d) heißt **vollständig**, wenn jede Cauchy-Folge (x_n) von Elementen aus M einen Grenzwert $x \in M$ besitzt.

Definition 10.40. *Es sei (M, d) ein metrischer Raum und $\varphi \colon M \to M$ eine Abbildung. Die Abbildung φ heißt* **Kontraktion**, *wenn ein $\lambda < 1$ existiert so, daß gilt:*

$$\forall x, y \in M \colon d(\varphi(x), \varphi(y)) \leq \lambda d(x,y).$$

Satz 10.41. (Banachscher Fixpunktsatz): *Es sei (M, d) ein vollständiger metrischer Raum und $\varphi \colon M \to M$ eine Kontraktion. Dann hat φ genau einen Fixpunkt $y \in M$, d.h., es gibt genau einen Punkt $y \in M$ mit $\varphi(y) = y$. Für jeden Startwert $x_0 \in M$ konvergiert die Folge (x_n), die rekursiv durch*

$$x_{n+1} := \varphi(x_n), \quad n \in \mathbb{N}_0,$$

definiert ist, gegen den Fixpunkt y.

Beweis: Es sei $x_0 \in M$. Wir definieren die Folge (x_n) rekursiv durch

$$x_{n+1} = \varphi(x_n), \quad n \in \mathbb{N}_0.$$

Durch Induktion zeigt man

$$\forall n \in \mathbb{N}\colon d(x_n, x_{n+1}) \leq \lambda^n d(x_0, x_1).$$

Es seien $n, k \in \mathbb{N}$. Dann folgt daraus

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{n+p-1}, x_{n+p}) \\ &\leq (\lambda^n + \cdots + \lambda^{n+p-1}) d(x_0, x_1) \\ &\leq \lambda^n \left(\sum_{p=0}^{\infty} \lambda^p \right) d(x_0, x_1) = \frac{\lambda^n}{1-\lambda} d(x_0, x_1). \end{aligned}$$

Für $\lambda < 1$ gilt $\lim_{n\to\infty} \lambda^n = 0$. Daher folgt aus dieser Ungleichung, daß (x_n) eine Cauchy-Folge ist. Wegen der Vollständigkeit von (M, d) existiert der Grenzwert

$$y = \lim_{n\to\infty} x_n.$$

Da φ eine Kontraktion ist, gilt

$$\forall n \in \mathbb{N}\colon d(\varphi(x_n), \varphi(y)) \leq \lambda d(x_n, y).$$

Daraus folgt

$$\varphi(y) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = y,$$

d.h., y ist Fixpunkt von φ. Es sei jetzt $z \in M$ irgendein Fixpunkt von φ. Dann gilt

$$d(y, z) = d(\varphi(y), \varphi(z)) \leq \lambda \varphi(y, z).$$

Da $\lambda < 1$ ist, folgt daraus

$$d(y, z) = 0, \quad \text{d.h.}, y = z.$$

□

Wir können jetzt den Satz über die lokale Umkehrbarkeit beweisen.

Satz 10.42. *Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $\phi\colon U \to \mathbb{R}^n$ eine C^1-Abbildung. Es sei $a \in U$ und $d\phi(a)\colon \mathbb{R}^n \to \mathbb{R}^n$ sei ein Isomorphismus. Dann existiert eine offene Umgebung $U_0 \subset U$ von a so, daß gilt*

(1) *$V = \phi(U_0)$ ist eine offenen Umgebung von $b = \phi(a)$.*
(2) *$\phi\colon U_0 \to V$ ist ein Diffeomorphismus.*
(3) *Die Umkehrabbildung $\psi\colon V \to U_0$ hat in b das Differential*

$$d\psi(b) = (d\phi(a))^{-1}.$$

Beweis: Es sei $b = \phi(a)$. Wir konstruieren zunächst in einer geeigneten Umgebung von b eine stetige Umkehrabbildung. O.B.d.A. können wir annehmen, daß gilt:

$$a = 0 \quad \text{und} \quad \phi(a) = 0.$$

Andernfalls betrachten wir die Abbildung

$$\widetilde{\phi}(x) = \phi(x + a) - \phi(a).$$

Diese Abbildung ist differenzierbar in einer Umgebung von $0 \in \mathbb{R}^n$ und es gilt

$$d\widetilde{\phi}(0) = d\phi(a).$$

Daher ist $d\widetilde{\phi}(0)$ ein Isomorphismus. Es sei $L = (d\phi(0))^{-1}$. Da L ein Isomorphismus des $\mathbb{R}^n$ ist, genügt es den Satz für die Abbildung

$$L \circ \phi \colon U \to \mathbb{R}^n$$

zu beweisen, denn wenn $L \circ \phi$ ein lokaler Diffeomorphismus ist, so ist auch $\phi = L^{-1}(L \circ \phi)$ ein lokaler Diffeomorphismus. Nach der Kettenregel gilt:

$$d(L \circ \phi) = dL \circ d\phi = L \circ d\phi = \mathrm{Id}_{\mathbb{R}^n}.$$

Wir können damit O.B.d.A. folgende Annahmen machen:

$$a = \phi(a) = 0 \quad \text{und} \quad d\phi(0) = \mathrm{Id}.$$

1) Die Konstruktion der Umkehrabbildung zu $\phi \colon U \to \mathbb{R}^n$ in einer Umgebung von 0 bedeutet, für eine geeignete Kugel $U_\delta(0)$ und für jedes $y \in U_\delta(0)$ die Gleichung

$$y = \phi(x)$$

zu lösen. Dies kann wie folgt auf die Lösung einer Fixpunktgleichung zurückgeführt werden. Für $y \in \mathbb{R}^n$ sei $\varphi_y \colon U \to \mathbb{R}^n$ definiert durch

$$\varphi_y(x) = y + x - \phi(x).$$

Dann ist x ein Fixpunkt von φ_y genau dann, wenn x Lösung der Gleichung $y = \phi(x)$ ist.
Wir wählen $r > 0$ so, daß $\overline{U_{2r}(0)} \subset U$ und daß gilt:

$$\forall x \in \overline{U_{2r}}(0) \colon \quad \| \mathrm{Id} - d\phi(x) \| \leq \frac{1}{2}.$$

Da $d\phi(0) = \mathrm{Id}$ und $d\phi$ stetig ist, existiert ein solches r. Nach Defintion von φ_y gilt

$$d\varphi_y = \mathrm{Id} - d\phi.$$

Aus dem Schrankensatz (Satz 10.20) folgt:

$$\forall x_1, x_2 \in U_{2r}(0)\colon \ \| \varphi_y(x_1) - \varphi_y(x_2) \| \leq \frac{1}{2} \| x_1 - x_2 \| . \tag{10.7}$$

Hieraus folgt für $\| y \| < r$ und $\| x \| \leq 2r$

$$\begin{aligned} \| \varphi_y(x) \| &\leq \| \varphi_y(x) - \varphi_y(0) \| + \| \varphi_y(0) \| \\ &= \| \varphi_y(x) - \varphi_y(0) \| + \| y \| \\ &< 2r. \end{aligned}$$

Dabei haben wir benutzt, daß $\varphi_y(0) = y$ auf Grund unserer Annahme $\phi(0) = 0$ gilt. Daraus folgt, daß für alle $y \in \mathbb{R}^n$ mit $\| y \| < r$ die Abbildung φ_y die Kugel $\overline{U_{2r}}(0)$ in sich abbildet und wegen (10.7) sind diese Abbildungen Kontraktionen mit $\lambda = 1/2$. Da die abgeschlossene Kugel $\overline{U_{2r}}(0)$ kompakt ist, ist $\overline{U_{2r}}(0)$ ein vollständiger metrischer Raum. Aus Satz 10.41 folgt daher, daß für $\| y \| < r$ jede Abbildung

$$\varphi_y\colon \overline{U_{2r}}(0) \to \overline{U_{2r}}(0)$$

genau einen Fixpunkt $x \in \overline{U_{2r}}(0)$ hat. Wie wir oben gezeigt haben, liegt x in der offenen Kugel $U_{2r}(0)$. Damit folgt, daß zu jedem $y \in U_r(0)$ genau ein $x \in U_{2r}(0)$ existiert mit $\varphi_y(x) = x$. Diese Gleichung ist äquivalent zu $\phi(x) = y$. Wir setzen

$$\psi(y) = x, \quad V = U_r(0) \quad \text{und} \quad U_0 = \phi^{-1}(V) \cap U_{2r}(0).$$

Dann ist $\psi\colon V \to U_0$ die Umkehrabbildung zu $\phi \mid U_0$.
2)Wir zeigen jetzt, daß ψ stetig ist. Es seien $y_1, y_2 \in V$. Für die Bildpunkte $x_1 = \psi(y_1)$ und $x_2 = \psi(y_2)$ gilt dann

$$x_2 - x_1 = \varphi_0(x_2) - \varphi_0(x_1) + \phi(x_2) - \phi(x_1).$$

Aus (10.7) folgt damit

$$\| x_2 - x_1 \| \leq \frac{1}{2} \| x_2 - x_1 \| + \| \phi(x_2) - \phi(x_1) \| .$$

Da $\phi \circ \psi = Id$ gilt, ist $\phi(x_i) = y_i, i = 1, 2$. Damit erhalten wir aus obiger Ungleichung

$$\| \psi(y_2) - \psi(y_1) \| \leq 2 \| y_2 - y_1 \|,$$

d.h., ψ ist stetig.
3) Wir zeigen, daß das Differential $d\phi(x)$ in jedem Punkt $x \in U_0$ ein Isomorphismus ist. Auf Grund der Wahl von r und wegen $\| x \| < 2r$ gilt

$$\| (\mathrm{Id} - d\phi(x))v \| \leq \frac{1}{2} \| v \| \tag{10.8}$$

für alle $v \in \mathbb{R}^n$. Es sei $v \in \ker(d\phi(x))$. Dann folgt aus (10.8)

$$\parallel v \parallel \leq \frac{1}{2} \parallel v \parallel,$$

d.h., $v = 0$. Daraus folgt, daß $d\phi(x)$ ein Isomorphismus ist.

Wir haben damit gezeigt, daß für $\phi \mid U_0$ die Voraussetzungen von Satz 10.38 gelten. Daraus folgt, daß $\phi \colon U_0 \to V$ ein Diffeomorphismus ist.

Die Formel für das Differential folgt aus $\phi \circ \psi = \mathrm{Id}_{U_0}, \quad \psi \circ \phi = \mathrm{Id}_V$ und der Kettenregel.

$\square$

Beispiel: Für die Polarkoordinatenabbildung $P_2 : \mathbb{R}^2 \to \mathbb{R}^2$, die durch

$$P_2(r, \varphi) = (r \cos\varphi, r \sin\varphi)$$

definiert ist, gilt $\det(dP_2(r,\varphi)) = r$. Für alle (r, φ) mit $r \neq 0$ ist daher $dP_2(r,\varphi)$ ein Isomorphismus. Zum Beispiel ist

$$P_2 : \mathbb{R}^+ \times (-\pi, \pi) \to \mathbb{R}^2 - \{(x, 0) \mid x \leq 0\}$$

ein Diffeomorphismus. Die Umkehrabbildung ist

$$P_2^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, \operatorname{sign}(y) \arccos\left(\frac{x}{\sqrt{x^2 + y^2}} \right) \right).$$

Definition 10.43. *Es sei $U \subset \mathbb{R}^n$ offen und $\phi : U \to \mathbb{R}^n$ sei eine C^1-Abbildung. Dann heißt ϕ im Punkte $x_0 \in U$ ein* **lokaler Diffeomorphismus**, *wenn es offene Umgebungen U_0 von x_0 und V von $\phi(x_0)$ gibt, so daß die Einschränkung $\phi \mid U_0$ ein Diffeomorphismus von U_0 auf V ist.*

Damit kann der Umkehrsatz auch wie folgt formuliert werden:

Eine C^1-Abbildung $\phi : U \to \mathbb{R}^n$ ist in einem Punkte $x_0 \in U$ genau dann ein lokaler Diffeomophismus, wenn das Differential $d\phi(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ ein Isomorphismus ist.

Eine wichtige Anwendung des Umkehrsatzes ist der folgende Offenheitssatz.

Satz 10.44. *Es sei $U \subset \mathbb{R}^n$ offen und $\phi : U \to \mathbb{R}^n$ eine C^1-Abbildung derart, daß für alle $x \in U$ das Differential $d\phi(x) : \mathbb{R}^n \to \mathbb{R}^n$ ein Diffeomorphismus ist. Dann gilt:*

(1) *$\phi(U)$ ist offen.*
(2) *Wenn ϕ injektiv ist, so ist $\phi : U \to \phi(U)$ ein Diffeomorphismus.*

Beweis: 1) Es sei $x \in U$. Nach Satz 10.42 existiert eine offene Umgebung U_x von x so, daß die Einschränkung von $\phi \mid U_x$ ein Diffeomorphismus von U_x auf die offene Menge $\phi(U_x)$ ist. Dann ist

$$\phi(U) = \bigcup_{x \in U} \phi(U_x)$$

als Vereinigung offener Mengen ebenfalls offen.
2)Es sei $\phi : U \to \mathbb{R}^n$ injektiv. Dann existiert die Umkehrabbildung $\psi : \phi(U) \to U$. Es sei $U' \subset U$ eine offene Teilmenge. Nach 1) ist $\psi^{-1}(U') = \phi(U')$ offen. Daher ist ψ stetig. Aus Satz 10.38 folgt, daß $\phi : U \to \phi(U)$ ein Diffeomorphismus ist.

□

10.10. Der Satz über implizite Funktionen

Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine Funktion. In diesem Abschnitt untersuchen wir die Frage, wann man die Gleichung

$$f(x, y) = 0$$

in die Umgebung eines Punktes (x_0, y_0), der Lösung dieser Gleichung ist, nach y auflösen kann, d.h., wir suchen eine Funktion

$$g : (x_0 - \epsilon, x_0 + \epsilon) \to \mathbb{R}$$

so, daß gilt:

$$g(x_0) = y_0 \quad \text{und} \quad f(x, g(x)) = 0 \quad \text{für} \quad |x - x_0| < \epsilon.$$

Es sei

$$X = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$$

die Nullstellenmenge von f und es sei $(x_0, y_0) \in X$. Anders formuliert lautet unsere Frage:

Existieren $\epsilon > 0$ und eine Funktion $g : (x_0 - \epsilon, x_0 + \epsilon) \to \mathbb{R}$ so, daß $X \cap U_\epsilon(x_0, y_0)$ der Graph von g ist?

Beispiele:

(1) Es sei $f(x, y) = x - y^2$. Es sei $(x_0, y_0) \neq (0, 0)$. Dann existiert g. Für $|x| < \epsilon$ ist

$$g(x) = \begin{cases} \sqrt{x}, & y_0 > 0; \\ -\sqrt{x}, & y_0 < 0. \end{cases}$$

Andererseits hat die Gleichung $f(x,y)=0$ in keiner Umgebung von $(0,0)$ eine Auflösung nach y, da zu jedem $x>0$ zwei y-Werte existieren, $y=\sqrt{x}$ und $y=-\sqrt{x}$, so daß

$$f(x,y)=0.$$

(2) Es sei $f(x,y)=x^2-y^2=(x-y)(x+y)$.
In der Nähe von $(0,0)$ hat die Kurve $f(x,y)=0$ zwei Zweige und ist deshalb nicht auflösbar.

(3) Es sei $f(x,y)=x^2(1-x^2)-y^2=0$.
Die Gleichung $f(x,y)=0$ ist in der Umgebung von $(0,1),(-1,0)$ und $(0,0)$ nicht nach y auflösbar.

Die Nullstellenmenge von $f(x,y)=x^2(1-x^2)-y^2$.

Die Frage nach der Existenz einer Auflösung der Gleichung $f(x,y)=0$ wird wie folgt auf den Umkehrsatz zurückgeführt. Wir betrachten für $u\in(-\epsilon,\epsilon)$ die Gleichung

$$f(x,y)=u$$

und fragen nach der Existenz einer Funktion

$$g:(x_0-\epsilon,x_0+\epsilon)\times(-\epsilon,\epsilon)\longrightarrow\mathbb{R}$$

so, daß

$$f(x,g(x,u))=u$$

gilt. Es sei

$$\psi(x,u)=(x,g(x,u)),\quad (x,u)\in(x_0-\epsilon,x_0+\epsilon)\times(-\epsilon,\epsilon)$$

und

$$\phi(x,y)=(x,f(x,y)),\quad (x,y)\in\mathbb{R}^2.$$

Dann gilt

$$\phi\circ\psi(x,u)=\phi(x,g(x,u))=(x,f(x,g(x,u))=(x,u),$$

d.h., es gilt

$$\phi\circ\psi=\mathrm{Id}.$$

Um die Existenz von $g(x,u)$ zu beweisen, genügt es also zu zeigen, daß in einer Umgebung von $(x_0,0)$ die Umkehrabbildung $\psi=\phi^{-1}$ existiert. Aus dem Umkehrsatz wissen wir, daß

$$\det(d\phi(x_0,y_0))\neq 0 \qquad (10.9)$$

eine hinreichende Bedingung für die Existenz von ψ ist.

Wir betrachten jetzt eine etwas allgemeinere Situation und führen einige Bezeichnungen ein, um (10.9) geeignet formulieren zu können. Es sei $U \subset \mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k$ offen und $f : U \to \mathbb{R}^m$ sei eine C^1-Abbildung. Wir definieren die partiellen Differentiale

$$d_x f(x,y) : \mathbb{R}^n \to \mathbb{R}^m$$
$$d_y f(x,y) : \mathbb{R}^k \to \mathbb{R}^m$$

durch

$$d_x f(x,y)(h) = df(x,y)(h,0)$$
$$d_y f(x,y)(k) = df(x,y)(0,k).$$

Dann gilt für das Differential

$$df(x,y) : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$$

$$df(x,y)(h,k) = d_x f(x,y)(h) + d_y f(x,y)(k). \tag{10.10}$$

Bemerkungen: Es sei $k = m$.

(1) Wenn $d_y f(x,y) : \mathbb{R}^k \to \mathbb{R}^k$ invertierbar ist, so ist $df(x,y)$ surjektiv.

(2) Es seien $f_1, \dots, f_m$ die Komponenten von f, wobei $f_i : \mathbb{R}^{n+m} \to \mathbb{R}$ Funktionen sind. Dann entspricht die Gleichung $f(x,y) = 0$ dem Gleichungssystem

$$\begin{gathered} f_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \\ \vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) = 0. \end{gathered}$$

Die Jacobi-Matrix von f ist gleich

$$\begin{pmatrix} \partial_{x_1} f_1 \cdots \partial_{x_n} f_1 & \partial_{y_1} f_1 \cdots \partial_{y_m} f_1 \\ \vdots \qquad\quad \vdots & \vdots \qquad\quad \vdots \\ \partial_{x_1} f_m \cdots \partial_{x_n} f_m & \partial_{y_1} f_m \cdots \partial_{y_m} f_m \end{pmatrix}$$

Entsprechend ist

$$d_x f(x,y) = \begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_n} f_1 \\ \vdots & & \vdots \\ \partial_{x_1} f_m & \cdots & \partial_{x_n} f_m \end{pmatrix}$$

$$d_y f(x,y) = \begin{pmatrix} \partial_{y_1} f_1 & \cdots & \partial_{y_1} f_1 \\ \vdots & & \vdots \\ \partial_{y_1} f_m & \cdots & \partial_{y_m} f_m \end{pmatrix}.$$

Satz 10.45. *(Satz über implizite Funktionen) Es sei $U \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^m$ eine C^1-Abbildung. Es sei $(x_0, y_0) \in U$ eine Nullstelle von f, d.h., $f(x_0, y_0) = 0$, und es sei $d_y f(x_0, y_0) : \mathbb{R}^m \to \mathbb{R}^m$ invertierbar. Dann gibt es Umgebungen $U_1 \subset \mathbb{R}^n$ von x_0 und $U_2 \subset \mathbb{R}^m$ von y_0 sowie eine C^1-Abbildung $g : U_1 \to U_2$ mit $g(x_0) = y_0$ so, daß gilt:*

$$\{(x, y) \in U_1 \times U_2 \mid f(x, y) = 0\} = \{(x, g(x)) \mid x \in U_1\}.$$

Für das Differential $dg(x_0)$ gilt

$$dg(x_0) = -\left(d_y f(x_0, y_0)\right)^{-1} \circ d_x f(x_0, y_0).$$

Beweis: Wir betrachten die Abbildung

$$\phi : U \to \mathbb{R}^{n+m},$$

die definiert ist durch

$$\phi(x, y) = (x, f(x, y)).$$

Dann ist ϕ eine C^1-Abbildung und für das Differential

$$d\phi(x, y) : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$$

gilt

$$d\phi(x, y)(h, k) = (h, d_x f(x, y)(h) + d_y f(x, y)(k)),$$

d.h.,

$$d\phi(x_0, y_0) = \begin{pmatrix} \mathrm{Id}_{\mathbb{R}^n} & 0 \\ d_x f(x_0, y_0) & d_y f(x_0, y_0) \end{pmatrix}.$$

Hieraus folgt

$$\det(d\phi(x_0, y_0)) = \det(d_y f(x_0, y_0)) \neq 0.$$

Aus dem Umkehrsatz (Satz 10.42) folgt, daß es Umgebungen U_0 von $(x_0, y_0) \in \mathbb{R}^{n+m}$ und V von $\phi(x_0, y_0) = (x_0, 0)$ gibt so, daß $\phi \mid U_0$ ein Diffeomorphismus von U_0 auf V ist. Sei

$$\phi^{-1} : V \to U_0$$

die Umkehrabbildung. Dann ist

$$\phi^{-1}(x, y) = (h_1(x, y), h_2(x, y))$$

und es gilt

$$(x, y) = \phi \circ \phi^{-1}(x, y) = \left(h_1(x, y), f(h_1(x, y), h_2(x, y))\right).$$

Hieraus folgt $h_1(x, y) = x$, d.h., ϕ^{-1} hat die Form

$$\phi^{-1}(x, y) = (x, h(x, y))$$

wobei $h : V \to \mathbb{R}^m$ eine C^1-Abbildung ist. Für $(x, y) \in U_0$ gilt damit

$$f(x, y) = 0 \Leftrightarrow \phi(x, y) = (x, 0) \Leftrightarrow \phi^{-1}(x, 0) = (x, y) \Leftrightarrow y = h(x, 0).$$

Zusammenfassug der Äquivalenzen ergibt

$$f(x, y) = 0 \Leftrightarrow y = h(x, 0).$$

Insbesondere ist $y_0 = h(x_0, 0)$. Da h stetig ist, existieren Umgebungen $U_1 \subset \mathbb{R}^n$ von x_0 und $U_2 \subset \mathbb{R}^m$ von y_0 so, daß gilt:

(1) $U_1 \times U_2 \subset U_0$
(2) $\forall x \in U_1 : h(x, 0) \in U_2$.

Es sei die Abbildung $g : U_1 \to U_2$ definiert durch

$$g(x) := h(x, 0).$$

Dann ist g eine C^1-Abbildung und es gilt

$$f(x, g(x)) = f(x, h(x, 0)) = 0, \quad x \in U_1.$$

Dieses g ist damit unsere gesuchte Abbildung. Aus der Kettenregel folgt

$$\forall x \in U_1 : df(x, g(x)) \circ (\mathrm{Id}_{\mathbb{R}^n}, dg(x)) = 0.$$

Diese Gleichung ist äquivalent zu

$$d_x f(x, g(x)) + d_y f(x, g(x)) \circ dg(x) = 0.$$

Da $y_0 = g(x_0)$, folgt daraus

$$dg(x_0) = -(d_y f(x_0, y_0))^{-1} \circ d_x f(x_0, y_0).$$

□

Wir haben den Beweis des Satzes über implizite Funktionen auf den Umkehrsatz zurückgeführt. Es gibt aber auch andere Beweise für diesen Satz. Wir betrachten der Einfachheit halber den Fall $n = m = 1$. Es sei $U \subset \mathbb{R}^2$ eine offene Teilmenge und $f : U \to \mathbb{R}$ stetig differenzierbar. Weiter sei $(x_0, y_0) \in U$ so, daß $f(x_0, y_0) = 0$ und es gelte $\partial_y f(x_0, y_0) \neq 0$. Da nach Voraussetzung $\partial_y f$ stetig ist, existiert eine offene Teilmenge $U_0 \subset U$ so, daß gilt:

$$\forall (x, y) \in U_0 : \quad \partial_y f(x, y) \neq 0.$$

Gesucht sind offene Intervalle $I_1, I_2 \subset \mathbb{R}$ und eine C^1-Funktion $g : I_1 \to I_2$ so, daß gilt:

(1) $x_0 \in I_1$, $y_0 \in I_2$, $I_1 \times I_2 \subset U_0$;
(2) $g(x_0) = y_0$;
(3) $\forall x \in I_1 \colon f(x, g(x)) = 0$.

Für die Ableitung der gesuchten Funktion g muß gelten

$$g'(x) = -\frac{1}{\partial_y f(x, g(x))} \partial_x f(x, g(x)) \tag{10.11}$$

für alle $x \in I_1$. Es sei

$$F(x, y) = -\frac{1}{\partial_y f(x, y)} \partial_x f(x, y)$$

für $(x, y) \in U_0$. Dann erhalten wir aud (10.11) und 2) die folgende gewöhnliche Differentialgleichung für die gesuchte Funktion g :

$$g'(x) = F(x, g(x)), \quad g(x_0) = y_0. \tag{10.12}$$

Aus der Theorie der gewöhnlichen Differentialgleichungen folgt, daß die Gleichung (10.12) eine eindeutig bestimmte Lösung $g(x)$ besitzt, die in einem Intervall $(x_0 - \epsilon, x_0 + \epsilon), \quad \epsilon > 0$, definiert ist. Für diese Lösung g gilt dann

$$\frac{d}{dx} f(x, g(x)) = 0, \quad g(x_0) = y_0.$$

Daraus folgt, daß $c \in \mathbb{R}$ existiert mit

$$f(x, g(x)) = c, \quad |x - x_0| < c.$$

Da nach Voraussetzung $f(x_0, g(x_0)) = 0$ ist, muß $c = 0$ gelten. Damit haben wir die Existenz der Auflösung von $f(x, y) = 0$ nach y auf den Existenz- und Eindeutigkeitssatz für gewöhnliche Differentialgleichungen zurückgeführt.

Beispiele:

(1) Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x, y) = e^x + e^y$. Dann ist

$$\partial_x f(x, y) = e^x \neq 0, \quad \partial_y f(x, y) = e^y \neq 0$$

für alle $(x, y) \in \mathbb{R}^2$. Die Gleichung $f(x, y) = 0$ hat jedoch keine Lösung! Wir können Satz 10.45 nicht anwenden, da wir keine Anfangslösung (x_0, y_0) zur Verfügung haben.

(2) Es sei $F : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch

$$F(x, y, z) = \begin{pmatrix} x^2 + y^2 - z^2 \\ x^2 + y^2 + z^2 - 2 \end{pmatrix}.$$

Dann gilt $F(0, 1, 1) = 0$ und

$$\partial_{(y,z)} F(0, 1, 1) = \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix},$$

d.h., $\det(\partial_{(y,z)}F(0,1,1)) \neq 0$. Die Voraussetzungen von Satz 10.45 sind damit erfüllt und es existieren daher Funktionen $y = f(x)$, $z = g(x)$, die für $|x| < \epsilon$ definiert sind so, daß $f(0) = 1 = g(0)$ und

$$F(x, f(x), g(x)) = 0, \quad |x| < \epsilon.$$

10.11. Differenzierbare Untermannigfaltigkeiten

Unter einer differenzierbaren Mannigfaltigkeit X der Dimension n versteht man einen topologischen Raum, der sich lokal so wie der Euklidische Raum $\mathbb{R}^n$ verhält. Genauer kann man dies wie folgt formulieren.

Definition 10.46. *Es sei X ein topologischer Raum. X heißt n-dimensionale differenzierbare Mannigfaltigkeit, wenn folgendes gilt: Es existiert ein System $\{U_\alpha\}_{\alpha\in I}$ von offenen Teilmengen von X und für jedes $\alpha \in I$ existiert eine offene Teilmenge $V_\alpha \subset \mathbb{R}^n$ und ein Homöomorphismus $\varphi_\alpha : U_\alpha \to V_\alpha$ so, daß gilt:*

(1) $\cup_{\alpha\in I} U_\alpha = X$.

(2) *Für alle $\alpha, \beta \in I$: Wenn $U_\alpha \cap U_\beta \neq \emptyset$, so ist*

$$\varphi_\beta \circ \varphi_\alpha^{-1} : V_\alpha \cap V_\beta \to V_\alpha \cap V_\beta$$

eine differenzierbare Abbildung von offenen Mengen des $\mathbb{R}^n$.

Das Paar $(U_\alpha, \varphi_\alpha)$, $\alpha \in I$, nennt man eine **Karte** oder **Koordinatensystem von** X und $\mathcal{A} = \{(U_\alpha, \varphi_\alpha) \mid \alpha \in I\}$ heißt **Atlas** von X. Mit anderen Worten bedeutet dies, daß man in der Umgebung jedes Punktes $x \in X$ Koordinaten einführen kann und die Koordinatentransformationen zwischen verschiedenen Koordinatensystemen sind differenzierbar. Das Standardbeispiel einer differenzierbaren Mannigfaltigkeit ist natürlich der $\mathbb{R}^n$ selbst.

Wir betrachten in diesem Abschnitt Mannigfaltigkeiten X, die Teilmengen des $\mathbb{R}^n$ sind. Man sagt dann, daß X eine Untermannigfaltigkeit des $\mathbb{R}^n$ist. Das Standardbeispiel einer d-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^n$ ist der Unterraum

$$\mathbb{R}^d_0 := \{x \in \mathbb{R}^n \mid x_{d+1} = \cdots = x_n = 0\}.$$

Der Unterraum $\mathbb{R}^d_0 \subset \mathbb{R}^n$ ist das lokale Modell einer Untermannigfaltigkeit des $\mathbb{R}^n$.

Definition 10.47. *Sei $M \subset \mathbb{R}^n$ eine nichtleere Teilmenge. M heißt* **d-dimensionale C^k-Untermannigfaltigkeit von** $\mathbb{R}^n$, *wenn für alle $x \in M$ eine Umgebung $U \subset \mathbb{R}^n$ von x in $\mathbb{R}^n$ und ein C^k-Diffeomorphismus $\varphi : U \to V$, $V \subset \mathbb{R}^n$ offen, existieren so, daß*

$$\varphi(M \cap U) = \mathbb{R}^d_0 \cap V.$$

Das Paar (U, φ) *heißt* **Karte** *der Untermannigfaltigkeit* M. *Sei* $\mathcal{A} = \{(U_i, \varphi_i) \mid i \in I\}$ *eine Menge von Karten von* M *mit*

$$\bigcup_{i \in I} U_i = M.$$

Dann heißt $\mathcal{A}$ **Atlas** *von* M.

Beispiel: Es sei

$$S^{n-1} = \{x \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}$$

die Einheitssphäre im $\mathbb{R}^n$. Dann ist $S^{n-1} \subset \mathbb{R}^n$ eine Untermannigfaltigkeit. Dies kann man wie folgt zeigen. Für $p \in \mathbb{R}^n$ sei

$$i_p(x) = p + \frac{2}{\| x - p \|^2}(x - p), \quad x \in \mathbb{R}^n - \{p\}.$$

Dann ist

$$i_p : \mathbb{R}^n - \{p\} \to \mathbb{R}^n - \{p\}$$

ein Diffeomorphismus mit $i_p^{-1} = i_p$. Deshalb nennt man i_p die **Inversion** mit dem Pol p. Insbesondere sei $N = (0, \dots, 0, 1)$ der Nordpol und $S = (0, \dots, 0, -1)$ der Südpol von S^{n-1}. Die Inversion $i_N : \mathbb{R}^n - \{N\} \to \mathbb{R}^n - \{N\}$ ist ein Diffeomorphismus, der $S^{n-1} - \{N\}$ auf die Hyperebene $\mathbb{R}_0^{n-1} \subset \mathbb{R}^n$ abbildet, d.h., $(\mathbb{R}^n - \{N\}, i_N)$ ist eine Karte für S^{n-1}. Eine zweite Karte erhält man aus der Inversion i_S bezüglich des Südpols $S \in S^{n-1}$. Diese beiden Kartengebiete ergeben eine Überdeckung von S^{n-1}. Damit haben wir gezeigt, daß $S^{n-1} \subset \mathbb{R}^n$ eine Untermannigfaltigkeit ist.

Nach Definition ist S^{n-1} die Lösungsmenge der Gleichung

$$\| x \|^2 = 1, \quad x \in \mathbb{R}^n.$$

Viele wichtige Untermannigfaltigkeiten ergeben sich ebenfalls als Lösungen von Gleichungen. Wie das Beispiel

$$X = \{(x, y) \in \mathbb{R}^2 \mid x^2(1 - x^2) - y^2 = 0\}$$

zeigt, ist eine solche Lösungsmenge im allgemeinen keine Untermannigfaltigkeit. Die Gleichung

$$x^2(1 - x^2) - y^2 = 0$$

hat in der Umgebung von $(0, 0)$ zwei Zweige als Lösungen und ist deshalb keine Untermannigfaltigkeit.

Wir beweisen jetzt einen allgemeinen Satz, der angibt, wann die Lösungsmenge einer Gleichung eine Umkehrmannigfaltigkeit ist.

Definition 10.48. *Es sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ sei differenzierbar. Ein Punkt $x \in U$ heißt* **regulärer Punkt** *von f, wenn*

$$df(x) : \mathbb{R}^n \to \mathbb{R}^m$$

surjektiv ist. Ferner heißt $y \in \mathbb{R}^m$ **regulärer Wert von** *f, wenn alle $x \in f^{-1}(y)$ reguläre Punkte sind. (y heißt auch regulärer Wert, wenn $f^{-1}(y)$ leer ist.) $x \in U$ heißt* **singulär**, *wenn $df(x) : \mathbb{R}^n \to \mathbb{R}^m$ nicht surjektiv ist. In diesem Falle heißt $f(x)$* **singulärer Wert von** *f.*

Anders formuliert bedeutet dies: Der Punkt $y \in \mathbb{R}^m$ ist regulärer Wert von $f : U \to \mathbb{R}^m$ genau dann, wenn für alle $x \in f^{-1}(y)$ die Funktionalmatrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

den Rang m hat. Insbesondere heißt dies für $m = 1$:

$$df(x) \neq 0 \quad \text{für alle } x \in f^{-1}(y).$$

Satz 10.49. *(Satz vom regulären Wert) Es sei $U \subset \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}$ eine C^1-Abbildung. Weiter sei $c \in \mathbb{R}^m$ ein regulärer Wert von f und $M := f^{-1}(c)$. Wenn $M \neq \emptyset$, so ist $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit der Dimension*

$$\dim M = n - m.$$

Beweis: Sei $M = f^{-1}(c)$ und sei $a \in M$. Sei weiter $Z = \ker(df(a))$ und $W = Z^{\perp}$. Da $df(a) : \mathbb{R}^n \to \mathbb{R}^m$ surjektiv ist, gilt

(1) $df(a) \mid W$ ist eine Isomorphismus von W auf $\mathbb{R}^m$.
(2) Die Abbildung $i : Z \times W \to \mathbb{R}^n$, die definiert ist durch $i(x, y) = x + y$, ist ein Isomorphismus.

Es sei $k = n - m$. Dann ist $Z \cong \mathbb{R}^k, W \cong \mathbb{R}^m$ und $i : \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\cong} \mathbb{R}^n$. Es genügt zu zeigen, daß $i^{-1}(M)$ eine Untermannigfaltigkeit ist.
Es sei $\widetilde{f} = f \circ i$. Für $(x, y) \in \mathbb{R}^k \times \mathbb{R}^m$ und $(h, k) \in \mathbb{R}^k \times \mathbb{R}^m$ gilt

$$d\widetilde{f}(x, y)(h, k) = df(i(x, y))(i(h, k)) = df(x + y)(h + k).$$

Es sei $(x_0, y_0) = i^{-1}(a)$. Dann gilt für das partielle Differential

$$d_y\widetilde{f}(x_0, y_0)(k) = d\widetilde{f}(x_0, y_0)(0, k) = df(a)(k), \ k \in \mathbb{R}^m.$$

Aus 1) folgt, daß

$$d_y f(x_0, y_0) : \mathbb{R}^m \to \mathbb{R}^m$$

ein Isomorphismus ist. Aus dem Satz über implizite Funktionen folgt daher, daß offene Umgebungen $U_1 \subset \mathbb{R}^k$ von x_0 und $U_2 \subset \mathbb{R}^m$ von y_0 sowie eine C^1-Abbildung $g : U_1 \to U_2$ existieren so, daß

$$\widetilde{f}^{-1}(c) \cap (U_1 \times U_2) = i^{-1}(M) \cap (U_1 \times U_2) = \{(x, g(x)) \mid x \in U_1\}. \quad (10.13)$$

Es sei $\varphi : U_1 \times \mathbb{R}^m \to U_1 \times \mathbb{R}^m$ definiert durch

$$\varphi(x, y) = (x, y - g(x)).$$

Entsprechend sei $\psi : U_1 \times \mathbb{R}^m$ definiert durch

$$\psi(x, y) = (x, y + g(x)).$$

Dann sind ϕ und ψ C^1-Abbildungen mit $\psi \circ \phi = \mathrm{Id}$ und $\phi \circ \psi = \mathrm{Id}$, d.h. ϕ und ψ sind Diffeomorphismen und $\phi = \psi^{-1}$. Es sei $V = \varphi(U_1 \times U_2)$. Dann ist $V \subset \mathbb{R}^k \times \mathbb{R}^m$ offen und

$$\varphi : U_1 \times U_2 \to V$$

ist ein Diffeomorphismus. Ferner folgt aus (10.13)

$$\varphi(U_1 \times U_2) \cap i^{-1}(M)) = U_1 \times \{0\}.$$

Damit ist $(U_1 \times U_2, \varphi)$ eine Karte von $i^{-1}(M)$ in der Umgebung von (x_0, y_0).

□

Beispiele:

(1) Wir betrachten nochmals die Einheitssphäre $S^{n-1} \subset \mathbb{R}^n$. Es sei $f : \mathbb{R}^n \to \mathbb{R}$ definiert duch $f(x) = \| x \|^2$. Dann ist $S^{n-1} = f^{-1}(1)$. Für $x \in \mathbb{R}^n$ ist

$$df(x) = 2(x_1, \dots, x_n)$$

und

$$df(x) = 2(x, v), \quad v \in \mathbb{R}^n.$$

Daraus folgt, daß für alle $x \in S^{n-1}$

$$df(x) : \mathbb{R}^n \to \mathbb{R}$$

surjektiv ist. Aus Satz 10.49 erhalten wir damit, daß S^{n-1} eine Untermannigfaltigkeit ist.

(2) Es sei $E \in \mathrm{Mat}(n, \mathbb{R})$ die Einheitsmatrix und

$$O(n) = \{A \in \mathrm{Mat}(n, \mathbb{R}) \mid AA^t = E\}$$

die orthogonale Gruppe. Da $\mathrm{Mat}(n, \mathbb{R}) \cong \mathbb{R}^{n^2}$, können wir $O(n)$ als Teilmenge von $\mathbb{R}^{n^2}$ auffassen.

Beh.: $O(n) \subset \mathbb{R}^2$ ist eine Untermannigfaltigkeit der Dimension $n(n-1)/2$.

Beweis: Es sei

$$S(n) = \{X \in \mathrm{Mat}(n, \mathbb{R}) \mid X^t = X\}$$

der Raum der symmetrischen $n \times n$ Matrizen. Dann ist $S(n) \cong \mathbb{R}^{\frac{n(n+1)}{2}}$. Es sei

$$f : \mathrm{Mat}(n, \mathbb{R}) \to S(n)$$

definiert durch

$$f(X) = X^t X.$$

Nach Definition von $O(n)$ ist $O(n) = f^{-1}(E)$. Wir müssen also zeigen, daß die Voraussetzungen von Satz 10.49 für f erfüllt sind. Nach der Produktregel ist f stetig differenzierbar. Es sei $A \in \mathrm{Mat}(n, \mathbb{R})$. Dann gilt für das Differential von f in A:

$$df(A)(H) = A^t H + H^t A, \quad H \in \mathbb{R}^{n^2}.$$

Sei $Y \in S(n)$. Wir setzen

$$H = \frac{1}{2} AY.$$

Dann ist

$$df(A)(H) = A^t(\frac{1}{2}AY) + (\frac{1}{2}AY)^t A = \frac{1}{2}A^t AY + \frac{1}{2}Y^t A^t A = Y,$$

da nach Voraussetzung $Y = Y^t$ und $A^t A = E$ gilt. Damit haben wir gezeigt, daß für jedes $A \in O(n)$ das Differential

$$df(A) : \mathbb{R}^{n^2} \to \mathbb{R}^{\frac{n(n+1)}{2}}$$

surjektiv ist. Aus Satz 10.49 folgt, daß $O(n)$ eine Untermannigfaltigkeit von $\mathbb{R}^{n^2}$ ist. Für die Dimension gilt

$$\dim O(n) = n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1).$$

□

(3) Es sei

$$SL(n, \mathbb{R}) = \{A \in \mathrm{Mat}(n, \mathbb{R}) \mid \det A = 1\}$$

die **spezielle lineare Gruppe.**

Beh.: $SL(n, \mathbb{R})$ ist eine Untermannigfaltigkeit von $\mathrm{Mat}(n, \mathbb{R} \cong \mathbb{R}^{n^2}$ der Dimension $n^2 - 1$.

Beweis: Es sei $f : \mathrm{Mat}(n, \mathbb{R}) \to \mathbb{R}$ die Funktion $f(A) = \det(A)$. Dann ist $SL(n, \mathbb{R}) = f^{-1}(1)$. Wir berechnen das Differential

$df(A),\ \ A \in \mathrm{Mat}(n, \mathbb{R})$ wie folgt. Eine Matrix $X \in \mathrm{Mat}(n, \mathbb{R})$ fassen wir als n-Tupel der Spaltenvektoren auf:

$$X = (x_1, \ldots, x_n), \quad x_j = \begin{pmatrix} x_{1_j} \\ \vdots \\ x_{n_j} \end{pmatrix}.$$

Dann ist

$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$$

eine multilineare alternierende Funktion. Für $X, H \in \mathrm{Mat}(n, \mathbb{R})$ gilt daher

$$\begin{aligned} \det(X + H) &= \det(x_1 + h_1, \ldots, x_n + h_n) \\ &= \det(x_1, \ldots, x_n) + \det(h_1, x_2, \ldots, x_n) \\ &\quad + \cdots + \det(x_1, \ldots, x_{n-1}, h_n) + R(X, H). \end{aligned}$$

Dabei ist $R(X, H)$ eine Summe von Determinanten, bei denen mindestens zwei Spaltenvektoren von H auftreten. Daraus folgt

$$|R(X, H)| \leq C \parallel H \parallel^2 .$$

Zusammen mit (10.14) ergibt dies

$$\begin{aligned} &d(\det(X))(H) \\ &\quad = \det(h_1, x_2, \ldots, x_n) + \cdots + \det(x_1, \ldots, , x_{n-1}, h_n). \end{aligned}$$

Für $A, Y \in \mathrm{Mat}(n, \mathbb{R})$ folgt daraus

$$\begin{aligned} d(\det(A))(AY) &= \det(y_{11}a_1 + \cdots + y_{n1}a_n, a_2, \ldots, a_n) \\ &\quad + \det(a_1, y_{12}a_1 + \cdots + y_{n2}a_n, a_3, \ldots, a_n) + \cdots \\ &= y_{11} \det(a_1, \ldots, a_n) + \cdots + y_{nn} \det(a_1, \ldots, a_n) \\ &= \mathrm{Tr}(Y) \det(A). \end{aligned}$$

Für $A \in SL(n, \mathbb{R})$ und $Y \in \mathrm{Mat}(n, \mathbb{R})$ ergibt sich daraus

$$d(\det(A))(AY) = \mathrm{Tr}(Y).$$

Es folgt, daß für $A \in SL(n, \mathbb{R})$

$$d(\det(A)) : \mathbb{R}^{n^2} \to \mathbb{R}$$

surjektiv ist und die Behauptung folgt aus Satz 10.49.

□

Differenzierbare Untermannigfaltigkeiten kann man wie folgt charakterisieren.

Satz 10.50. *Eine nicht leere Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine d-dimensionale Untermannigfaltigkeit, wenn es zu jedem $a \in M$ eine offene Umgebung U von a in $\mathbb{R}^n$ und C^1-Funktionen*

$$\varphi_{d+1}, \dots, \varphi_n : U \to \mathbb{R}$$

gibt derart, daß gilt

(1) $M \cap U = \{x \in U \mid \varphi_{d+1}(x) = \cdots = \varphi_n(x) = 0\}$,
(2) *Die Vektoren* $\operatorname{grad}\varphi_{d+1}(x), \dots, \operatorname{grad}\varphi_n(x)$ *sind in jedem Punkt* $x \in U$ *linear unabhängig.*

Beweis: 1) Es sei $M \subset \mathbb{R}^n$ eine d-dimensionale Untermannigfaltigkeit. Es sei $a \in M$. Wir wählen eine Karte (U, φ) in a, d.h.. $a \in U$ und

$$\varphi : U \to V$$

ist ein Diffeomorphismus auf eine offene Menge $V \subset \mathbb{R}^n$ so, daß

$$\varphi(M \cap U) = \mathbb{R}^d_0 \cap V. \tag{10.15}$$

Es sei $\varphi = (\varphi_1, \dots, \varphi_n)$, wobei $\varphi_i : U \to \mathbb{R}$ die i-te Komponente von φ ist. Dann sind $\varphi_{d+1}, \dots, \varphi_n$ die gesuchten Funktionen, denn aus (10.15) folgt, daß für $\varphi_{d+1}, \dots, \varphi_n$ die Bedingung 1) gilt und 2) folgt, da φ ein Diffeomorphismus ist und daher für $x \in U$

$$d\varphi(x) = (d\varphi_1(x), \dots, d\varphi_n(x)) : \mathbb{R}^n \to \mathbb{R}^n$$

ein Isomorphismus ist.
2) Es sei $a \in M$ und

$$\varphi_{d+1}, \dots, \varphi_n : U \to \mathbb{R}$$

seien Funktionen in einer Umgebung U von a, für die 1) und 2) gelten. Wir betrachten die Abbildung $f : U \to \mathbb{R}^{n-d}$, die definiert ist durch

$$f = (\varphi_{d+1}, \dots, \varphi_n).$$

Wegen 1) ist $M \cap U = f^{-1}(0)$. Aus 2) folgt, daß $df(x)$ für jedes $x \in U$ surjektiv ist. Deshalb ist 0 ein regulärern Wert von f. Aus Satz 10.49 folgt, daß $M \cap U$ eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ ist. Da jeder Punkt $q \in M$ eine solche Umgebung besitzt, ist M eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$.

$\square$

10.12. Der Tangentialraum

Defintion 10.51. *Es sei $M \subset \mathbb{R}^n$ eine nicht leere Teilmenge und $x \in M$. Der Vektor $v \in \mathbb{R}^n$ heißt* **Tangentialvektor** *an M in x, wenn es eine C^1-Kurve $\gamma : (-\epsilon, \epsilon) \to M$ gibt mit $\gamma(0) = x$ und $\dot{\gamma}(0) = v$. Die Menge der Tangentialvektoren an M in x heißt* **Tangentialkegel** *T_xM von M in x. Falls T_xM ein Vektorraum ist, so heißt T_xM* **Tangentialraum** *an M in x.*

Satz 10.52. *Es sei $M \subset \mathbb{R}^n$ eine d-dimensionale Untermannigfaltigkeit. Dann gilt für alle $x \in M$:*

(1) *T_xM ist ein d-dimensionaler $\mathbb{R}$-Vektorraum.*

(2) *Falls $M = f^{-1}(c)$ die Niveaumenge einer C^1-Abbildung $f : U \to \mathbb{R}^m$ zu einem regulären Wert $c \in \mathbb{R}^m$ ist, so gilt*

$$T_xM = \ker df(x).$$

Beweis: 1. Es sei $x \in M$ und $\varphi : U \to V$ sei eine Karte mit $x \in U$. Dann ist

$$\varphi(U \cap M) = \mathbb{R}^d_0 \cap V.$$

Weiter sei $\gamma : (-\epsilon, \epsilon) \to M$ eine C^1-Kurve mit $\gamma(0) = x$. Durch Übergang zu einem kleineren $\epsilon > 0$ können wir annehmen, daß $\gamma(t) \in U$ für alle $t \in (-\epsilon, \epsilon)$. Dann gilt $\varphi \circ \gamma(t) \in \mathbb{R}^d_0$ für alle $t \in (-\epsilon, \epsilon)$, d.h.,

$$\varphi \circ \gamma : (-\epsilon, \epsilon) \to \mathbb{R}^d_0$$

ist eine C^1-Kurve in $\mathbb{R}^d_0$. Daher ist $\frac{d}{dt}(\varphi \circ \gamma)(0) \in \mathbb{R}^d_0$. Nach der Kettenregel ist

$$\frac{d}{dt}(\varphi \circ \gamma)(0) = d\varphi(x)(\dot{\gamma})(0).$$

Es sei

$$L = (d\varphi(x))^{-1}(\mathbb{R}^d_0).$$

Dann haben wir gezeigt:

$$T_xM \subset L. \tag{10.18}$$

Sei andererseits $v \in \mathbb{R}^d_0$ und $\epsilon > 0$ so, daß $\varphi(x) + tv \in V$ für alle t mit $|t| < \epsilon$. Wir setzen

$$\widetilde{\gamma}(t) = \varphi(x) + tv, \quad t \in (-\epsilon, \epsilon).$$

Es sei $\gamma = \varphi^{-1} \circ \widetilde{\gamma}$. Dann ist

$$\gamma : (-\epsilon, \epsilon) \to M$$

eine C^1-Kurve in M mit $\gamma(0) = x$. Daher gilt $\dot{\gamma}(0) \in T_xM$. Aus der Kettenregel folgt

$$\dot{\gamma}(0) = d(\varphi^{-1})(\varphi(x))(\dot{\widetilde{\gamma}}(0)) = (d\varphi(x))^{-1}(v).$$

Daraus folgt $L \subset T_xM$. Zusammen mit (10.18) erhalten wir
$$T_xM = L.$$
Somit ist T_xM ein Vektorraum.
2. Sei $\gamma : (-\epsilon, \epsilon) \to M$ eine C^1-Kurve mit $\gamma(0) = x$. Da $M = f^{-1}(c)$, ist $f \circ \gamma(t) = c$ für alle $t \in (-\epsilon, \epsilon)$. Daraus folgt
$$0 = \frac{d}{dt}(f \circ \gamma)(0) = df(x)(\dot\gamma(0)),$$
d.h., $T_xM \subset \ker df(x)$. Nach 1) ist $\dim T_xM = d$ und daher
$$\dim T_xM = d = \dim \ker df(x).$$
Daraus folgt
$$T_xM = \ker df(x).$$
□

Beispiele:
1. S^n.
Sei $f(x) = \| x \|^2$ und sei $a \in S^n$. Dann ist $df(a)(v) = 2\langle a, v\rangle$. Aus Satz 10.52 erhalten wir damit
$$T_aS^n = \{v \in \mathbb{R}^{n+1} \mid \langle a, v\rangle = 0\}.$$

2. $O(n)$.
Sei $f : \mathrm{Mat}(n, \mathbb{R}) \to S(n)$ die Abbildung $f(X) = X^tX$. Dann ist $O(n) = f^{-1}(1)$. Im letzten Abschnitt haben wir gezeigt
$$df(A)H = A^tH + H^tA.$$
Es sei $H \in \ker df(A)$. Wir setzen $Y = A^tH$. Dann ist Y schiefsymmetrisch und $H = AY$. Sei umgekehrt Y schiefsymmetrisch und $H = AY$. Dann ist $H \in \ker df(A)$. Daraus folgt
$$T_AO(n) = \{X \in \mathrm{Mat}(n, \mathbb{R}) \mid \exists Y \in \mathrm{Mat}(n, \mathbb{R}) : Y^t = -Y, X = AY\}.$$
Insbesondere erhalten wir für den Tangentialraum an die Einheitsmatrix $E \in O(n)$
$$T_EO(n) = \{X \in \mathrm{Mat}(n, \mathbb{R}) \mid X^t = -X\}.$$
Das ist der Raum der schiefsymmetrischen Matrizen.
3. $SL(n, \mathbb{R})$.
Für $A \in SL(n, \mathbb{R})$ ist
$$\begin{aligned} T_ASL(n, \mathbb{R}) = \{X \in \mathrm{Mat}(n, \mathbb{R}) \mid \exists Y \in \mathrm{Mat}(n, \mathbb{R}) :& \\ tr(Y) = 0 \text{ und } X = AY\}.& \end{aligned}$$

Insbesondere ist

$$T_E SL(n,\mathbb{R}) = \{X \in \mathrm{Mat}(n,\mathbb{R}) \mid tr(X) = 0\}.$$

der Raum der Matrizen mit Spur Null. Es sei $H \in T_E O(n)$. Wie wir aus Beispiel 2) wissen, ist H schiefsymmetrisch: $H^t = -H$. Es sei

$$e^H = \sum_{k=0}^{\infty} \frac{H^k}{k!}.$$

Dann ist $(e^H)^t = e^{H^t}$ und daher gilt

$$e^H \cdot (e^H)^t = e^H \cdot e^{H^t} = e^{H+H^t} = e^0 = E.$$

Dabei haben wir benutzt, daß H und H^t vertauschbare Matrizen sind und daher $e^H \cdot e^{H^t} = e^{H+H^t}$ gilt. Die Exponentialfunktion für schiefsymmetrische Matrizen liefert daher eine Abbildung.

$$\exp : T_E = O(n) \to O(n).$$

Für $H \in T_E O(n)$ sei

$$\gamma : \mathbb{R} \to O(n)$$

die Kurve $\gamma(t) = \exp(tH)$. Dies ist eine C^1-Kurve, für die gilt:

$$\gamma(0) = E, \quad \dot{\gamma}(0) = H.$$

Das Gleiche gilt für die Gruppe $SL(n,\mathbb{R})$. Aus der Formel

$$\det(e^X) = e^{tr(X)}$$

folgt, daß für $X \in T_E SL(n,\mathbb{R})$ die Matrix e^X in $SL(n,\mathbb{R})$ liegt. Damit definiert $X \longmapsto e^X$ eine Abbildung

$$\exp : T_E SL(n,\mathbb{R}) \to SL(n,\mathbb{R}).$$

Für jedes $X \in T_E SL(n,\mathbb{R})$ ist $\gamma(t) = \exp(tx)$, $t \in \mathbb{R}$, eine C^1-Kurve in $SL(n,\mathbb{R})$ mit

$$\gamma(0) = E, \quad \dot{\gamma}(0) = X.$$

Bemerkung: $O(n)$ und $SL(n,\mathbb{R})$ sind Beispiele für Liesche Gruppen.

Unter einer Lieschen Gruppe versteht man eine Gruppe, die zugleich eine differenzierbare Mannigfaltigkeit ist so, daß Multiplikation und Inversenbildung differenzierbare Abbildungen sind. Ein weiteres Beispiel einer Lieschen Gruppe ist die allgemeine lineare Gruppe

$$GL(n,\mathbb{R}) = \{X \in \mathrm{Mat}(n,\mathbb{R}) \mid \det X \neq 0\}.$$

Zu jeder Lieschen Gruppe G kann man den Tangentialraum $T_e G$ im Einselement $e \in G$ bilden. Es gibt dann ebenfalls eine Exponentialabbildung

$$\exp : T_e G \to G.$$

Wir können jetzt noch eine andere Charakterisierung von Untermannigfaltigkeiten angeben.

Defintion 10.53. *Eine $(n-1)$-dimensionale differenzierbare Untermannigfaltigkeit $H \subset \mathbb{R}^n$ heißt* **glatte Hyperfläche**.

Nach Satz 10.50 ist $H \subset \mathbb{R}^n$ genau dann eine glatte Hyperfläche, wenn es zu jedem $x \in H$ eine Umgebung $U \subset \mathbb{R}^n$ von x und eine C^1-Funktion $f : U \to \mathbb{R}$ gibt so, daß gilt:

(1) $\forall x \in U : df(x) \neq 0$.
(2) $H \cap U = \{x \in U \mid f(x) = 0\}$.

Definition 10.54. *Es seien $H_1, \dots, H_k \subset \mathbb{R}^n$ glatte Hyperflächen und $x \in H_1 \cap \dots \cap H_k$. Wir sagen, daß die Hyperflächen $H_1, \dots, H_k$ sich transvesal in x schneiden, wenn*

$$\dim(T_x H_1 \cap \dots \cap T_x H_k) = n - k.$$

Satz 10.55. *Eine nicht leere Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine d-dimensionale Untermannigfaltigkeit, wenn es zu jedem $x \in M$ eine Umgebung U und $n-d$ glatte Hyperflächen $H_{d+1}, \dots, H_n$ gibt so, daß*

(1) $M \cap U = H_{d+1} \cap \dots \cap H_n$
(2) *In jedem Punkt $x \in M \cap U$ schneiden sich diese Hyperflächen transversal.*

Beweis:$\Rightarrow$) Sei $M \subset \mathbb{R}^n$ eine d-dimensionale Untermannigfaltigkeit und sei $x \in M$. Wir wählen eine Umgebung U von x in $\mathbb{R}^n$ und C^1-Funktionen $\varphi_{d+1}, \dots, \varphi_n : U \to \mathbb{R}$ wie in Satz 10.50. Dann sind die Nullstellenmengen

$$H_j = \{x \in U \mid \varphi_j(x) = 0\}$$

glatte Hyperflächen in U mit den gewünschten Eigenschaften.

$\Leftarrow$) Es seien umgekehrt $H_{d+1}, \dots, H_n$ glatte Hyperflächen in U mit den im Satz angegebenen Eigenschaften. Nach Voraussetzung existieren C^1-Funktionen $\varphi_{d+1}, \dots, \varphi_n : U \to \mathbb{R}$ mit $d\varphi_j(x) \neq 0$ für alle $j = d+1, \dots, n$ und für alle $x \in U$ so, daß

$$H_j = \{x \in U \mid \varphi_j(x) = 0\}.$$

Aus Satz 10.52, 2), folgt

$$T_x H_j = \ker d\varphi_j(x).$$

Wegen der Voraussetzung

$$\dim(T_x H_{d+1} \cap \dots \cap T_x H_n) = d, \; x \in U,$$

folgt, daß die Differentiale $d\varphi_{d+1}(x), \dots , d\varphi_n(x)$ in jedem Punkt $x \in U$ linear unabhängig sind. Aus Satz 10.50 folgt, daß M eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ ist.

□

10.13. Der Normalenraum

Defintion 10.56. *Es sei $M \subset \mathbb{R}^n$ eine d-dimensionale Untermannigfaltigkeit und $x \in M$. Dann heißt*

$$N_xM := (T_xM)^{\perp}$$

der **Normalenraum** *zu M in x. Die Elemente von N_xM heißen* **Normalenvektoren** *in x.*

Satz 10.57. *Sei $f = (f_1, \dots , f_n) : U \to \mathbb{R}^m$ eine C^1-Abbildung mit $d = n-m \geq 0$ und $c \in \mathbb{R}^m$ sei ein regulärer Wert von f. Sei $M = f^{-1}(c)$. Dann sind für jedes $x \in M$ die Vektoren $\{\mathrm{grad} f_1(x), \dots , \mathrm{grad} f_m(x)\}$ eine Basis von N_xM.*

Beweis: Nach Satz 10.52 ist

$$T_xM = \ker df(x) = \ker df_1(x) \cap \dots \cap \ker df_m(x).$$

Also gilt:

$$v \in T_xM \Leftrightarrow \langle\, \mathrm{grad} f_i(x), v\rangle = 0 \text{ für alle } i,\ 1 \leq i \leq m.$$

Nach Definition von N_xM folgt:

$$\mathrm{grad} f_1(x), \dots , \mathrm{grad} f_m(x) \in N_xM.$$

Die Gradientenvektoren, als Zeilenvektoren aufgefaßt, sind die Zeilen der Jacobi-Matrix $(\partial f_i/\partial x_j)$. Da nach Voraussetzung $\mathrm{Rang}\, df(x)) = m$ ist, sind die Gradienten $\mathrm{grad} f_1(x), \dots , \mathrm{grad} f_m(x)$ linear unabhängig. Da andererseits

$$\dim N_xM = n - \dim T_xM = m$$

ist, bilden die Gradientenvektoren eine Basis von N_xM.

□

10.14. Extrema unter Nebenbedingungen

Bei vielen Extremalaufgaben handelt es sich um die Bestimmung von Extrema einer Funktion unter Nebenbedingungen. Z.B. in der Mechanik die Bestimmung des Minimums der Wirkungsfunktion, wobei die Phasenbahn in einer festen Energiefläche verläuft.

Dieses Problem kann wie folgt formuliert werden:
Gegeben seien $U \subset \mathbb{R}^n$, eine Funktion $f : U \to \mathbb{R}$ und eine Abbildung $\varphi : U \to \mathbb{R}^m$. Es sei

$$M = \{x \in U \mid \varphi(x) = 0\}$$

die Nullstellenmenge von φ.

Definition 10.58. *Der Punkt $x_0 \in M$ heißt*

(1) **Maximum (Minimum) von f unter der Nebenbedingung** $\varphi = 0$, *wenn gilt:*

$$\forall x \in M : f(x) \leq f(x_0) (\forall x \in M : f(x) \geq f(x_0)).$$

x_0 *heißt dann* **Extremum von f unter der Nebenbedingung** $\varphi = 0$.

Satz 10.59. *(Multiplikatorregel von Lagrange)*
Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar, und $\varphi = (\varphi_1, \ldots, \varphi_k) : U \to \mathbb{R}^k$ eine Abbildung so, daß $\varphi_i \in C^1(U)$, $i = 1, \ldots, k$. Wenn $a \in U$ ein Extremum von f unter der Nebenbedingung $\varphi = 0$ ist und die Gradienten $\mathrm{grad}\varphi_1(a), \ldots, \mathrm{grad}\varphi_k(a)$ linear unaghängig sind, so existieren $\lambda_1, \ldots, , \lambda_k \in \mathbb{R}$ mit

$$\mathrm{grad} f(a) = \sum_{i=1}^{k} \lambda_i \, \mathrm{grad}\varphi_i(a).$$

Die Zahlen $\lambda_1, \ldots, \lambda_k$ heißen **Lagrange-Multiplikatoren**.

Beweis: Es sei $M = \{x \in U \mid \varphi(x) = 0\}$. Nach Voraussetzung hat die lineare Abbildung

$$d\varphi(a) : \mathbb{R}^n \to \mathbb{R}^k$$

den Rang k. Nach eventueller Umnumerierung der Koordinaten können wir annehmen, daß bezüglich der Zerlegung $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ das partielle Differential $d_x\varphi(x_0, y_0) : \mathbb{R}^k \to \mathbb{R}^k$ invertierbar ist, wobei $a = (x_0, y_0)$. Nach dem Satz über implizite Funktionen existieren offenen Mengen $U_1 \subset \mathbb{R}^k$ von x_0 und $U_2 \subset \mathbb{R}^{n-k}$ von y_0 sowie eine C^1-Abbildung $g : U_2 \to U_1$ so, daß $g(y_0) = x_0$ und

$$M \cap (U_1 \times U_2) = \{(g(y), y) \mid y \in U_2\}.$$

Es sei $h : U_2 \to \mathbb{R}$ definiert durch

$$h(y) = f(g(y), y).$$

Nach Annahme hat h in y_0 ein Extremum. Daher ist

$$0 = dh(y_0) = d_x f(x_0, y_0) \circ (dg(y_0)) + d_y f(x_0, y_0).$$

Aus dem Satz über implizite Funktionen folgt

$$dg(y_0) = -\left(d_x\varphi(x_0, y_0)\right)^{-1} \circ d_y\varphi(x_0, y_0).$$

Zusammen mit obiger Gleichung erhalten wir daraus

$$d_x f(x_0, y_0) \circ (d_x\varphi(x_0, y_0))^{-1} \circ d_y(x_0, y_0) = d_y f(x_0, y_0). \tag{10.19}$$

Wir setzen nun

$$L = d_x f(x_0, y_0) \circ (d_x\varphi(x_0, y_0))^{-1}.$$

Dann ist L eine lineare Abbildung $L : \mathbb{R}^k \to \mathbb{R}$, d.h., es existieren $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit

$$L \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \lambda_1 x_1 + \cdots + \lambda_k x_k.$$

Nach Definition von L ist

$$L \circ d_x\varphi(x_0, y_0) = d_x f(x_0, y_0),$$

und aus (10.19) folgt

$$L \circ d_y\varphi(x_0, y_0) = d_y f(x_0, y_0).$$

Zusammenfassung dieser Gleichungen ergibt

$$L \circ d\varphi(x_0, y_0) = df(x_0, y_0).$$

Wenn wir die linearen Abbildungen

$$L : \mathbb{R}^k \to \mathbb{R},\ d\varphi(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}^k \text{ und } df(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}$$

in den Standardbasen durch Matrizen darstellen, folgt

$$\mathrm{grad} f(a) = \sum_{i=1}^{k} \lambda_i \ \mathrm{grad}\varphi_i(a).$$

$\square$

Beispiel: Es sei $A : \mathbb{R}^n \to \mathbb{R}^n$ linear und symmetrisch, d.h., es gilt

$$\langle Ax, y \rangle = \langle x, Ay \rangle, \quad \forall x, y \in \mathbb{R}^n.$$

Sei $Q(x) := \langle Ax, x \rangle$ die zugehörige quadratische Form. Wir wollen Q auf S^{n-1} maximieren, d.h., die Nebenbedingung lautet

$$\varphi(x) := \| x \|^2 - 1 = 0.$$

Für die Gradienten von Q und φ erhält man

$$\mathrm{grad} Q(x) = 2A(x), \quad \mathrm{grad}\varphi(x) = 2x.$$

Da $Q(x)$ eine stetige Funktion auf der kompakten Menge S^{n-1} ist, existiert ein Maximum $v_1 \in S^{n-1}$ von $Q(x)$. Aus der Multiplikatorregel von Lagrange folgt:

$$\exists \lambda \in \mathbb{R} : \quad A(v_1) = \lambda v_1,$$

d.h., v_1 ist ein Eigenvektor von A. Da A symmetrisch ist, ist das othogonale Komplement von $\mathbb{R}v_1$ in $\mathbb{R}^n$ wieder A invariant. Durch Induktion erhalten wir damit den bekannten Satz aus der linearen Algebra:

Satz. *Es sei $A \in Mat(n, \mathbb{R})$ eine symmetrische Matrix. Dann existiert eine orthonormierte Basis $v_1, \ldots, v_n$ von $\mathbb{R}^n$ mit*

(1) *$v_1, \ldots, v_n$ sind Eigenvektoren von A*
(2) *Es sei $H_k = \{v_1, \ldots, v_k\}^\perp, \quad k \geq 1$. Dann ist der Eigenwert λ_k das Maximum von Q auf $S^{n-1} \cap H_{k-1}$ für $k = 1, \ldots, n$.*

Beweis: Wir konstruieren $v_1, \ldots, v_n$ induktiv. v_1 haben wir bereits konstruiert. Es ist ein Maximum von Q auf S^{n-1}. Es seien $v_1, \ldots, v_k$ paarweise orthogonale Eigenvektoren von A mit den Eigenwerten $\lambda_1, \ldots, \lambda_k$. Wir suchen dann ein Maximum von Q auf $S^{n-1} \cap H_k$, d.h., ein Maximum von Q unter den Nebenbedingungen

$$\begin{aligned} \varphi_0(x) &:= \varphi(x) = \| x \|^2 - 1 = 0, \\ \varphi_1(x) &:= \langle v_1, x \rangle = 0, \\ &\vdots \\ \varphi_k(x) &:= \langle v_k, x \rangle = 0. \end{aligned}$$

Wegen der Kompaktheit von $S^{n-1} \cap H_k$ hat $Q : S^{n-1} \cap H_k \to \mathbb{R}$ ein Maximum v_{k+1}. Für die Differentiale $d\varphi_i$ gilt

$$d\varphi_0(x) = 2x, \; d\varphi_i(x) = v_i, \quad i = 1, \ldots, k.$$

Für $x \in S^{n-1} \cap H_k$ sind daher $d\varphi_0(x), d\varphi_1(x), \ldots, , d\varphi_k(x)$ linear unabhängig. Nach der Multiplikatorregel existieren $\mu_0, \ldots, \mu_k \in \mathbb{R}$ mit

$$df(v_{k+1}) = \sum_{i=0}^{k} \mu_i d\varphi_i(v_{k+1}),$$

d.h., es gilt

$$2A(v_{k+1}) = 2\mu_0 v_{k+1} + \sum_{i=1}^{k} \mu_i v_i. \tag{10.20}$$

Nach Voraussetzung ist $v_{k+1} \in H_k$, d.h., $\langle v_{k+1}, v_i \rangle = 0$ für alle $i, 1 \leq i \leq k$. Daher ist

$$\langle Av_{k+1}, v_i \rangle = \langle v_{k+1}, Av_i \rangle = \lambda_i \langle v_{k+1}, v_i \rangle = 0 \quad \text{für alle} \quad i = 1, \ldots, k.$$

Aus (10.21) folgt

$$0 = 2\langle Av_{k+1}, v_i\rangle = \mu_i \quad \text{für alle} \quad i = 1, \ldots, k,$$

und damit

$$Av_{k+1} = \mu_0 v_{k+1}.$$

v_{k+1} ist also ein Eigenvektor von A mit dem Eigenwert μ_0. Wegen

$$\mu_0 = \mu_0\langle v_{k+1}, v_{k+1}\rangle = \langle Av_{k+1}, v_{k+1}\rangle = Q(v_{k+1})$$

ist μ_0 das Maximum von Q auf $S^{n-1} \cap H_k$.

□

Die Multiplikatorregel von Lagrange kann man mit Hilfe des Normalenraumes wie folgt umformulieren:

Satz 10.60. *Es seien $f : U \to \mathbb{R}$ und $\varphi : U \to \mathbb{R}^k$ C^1-Abbildungen. Weiter sei $c \in \mathbb{R}^k$ ein regulärer Wert von φ und*

$$M = \{x \in U \mid \varphi(x) = c\}.$$

Wenn f in $a \in M$ unter der Nebenbedingung $\varphi = c$ ein Extremum hat, so ist $\mathrm{grad} f(a) \in N_aM$.

Beweis: Da c ein regulärer Wert von φ ist, ist wegen Satz 10.49 $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit. Nach Satz 10.57 sind für jedes $a \in M$ die Vektoren $\mathrm{grad}\varphi_1(a), \ldots, \mathrm{grad}\varphi_k(a)$ eine Basis von N_aM. Wenn a ein Extremum von $f : M \to \mathbb{R}$ ist, so folgt aus Satz 10.59, daß $\mathrm{grad}(a) \in N_aM$.

□

Beispiel: Es sei $M \subset \mathbb{R}$ eine C^1-Untermannigfaligkeit und $x_0 \in \mathbb{R}^n - M$. Sei $a \in M$ ein Punkt, der unter allen Punkten von M minimalen (oder maximalen) Abstand zu x_0 hat, d.h., a ist ein Extremum der Funktion

$$f(x) = \| x - x_0 \|$$

auf M. Die Funktion f ist auf $\mathbb{R}^n - \{x_0\}$ differenzierbar und es gilt

$$\mathrm{grad} f(x) = \frac{x - x_0}{\| x - x_0 \|}.$$

Aus Satz 10.16 folgt daher $a - x_0 \in N_aM$.

10.15. Differentation parameterabhängiger Integrale

Es sei $U \subset \mathbb{R}^n$ offen und $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall. Weiter sei $f : U \times [a, b] \to \mathbb{C}$ eine stetige Funktion. Für jedes $x \in U$ können wir dann die Funktion

$$f_x : [a, b] \to \mathbb{C}$$

betrachten, die durch $f_x(t) = f(x,t)$ definiert ist. Dann ist f_x für jedes $x \in U$ eine stetige Funktion auf $[a,b]$ und wir können deshalb $f_x(t)$ über $[a,b]$ integrieren. Es sei $F : U \to \mathbb{C}$ die Funktion, die wir dadurch erhalten, d.h.,

$$F(x) := \int_a^b f(x,t)dt, \quad x \in U.$$

Dies ist ein Integral, das von dem Parameter $x \in U$ abhängt. Wir nehmen jetzt an, daß für jedes $t \in [a,b]$ die Funktion $x \longmapsto f(x,t)$ nach x_j partiell differenzierbar ist. Es entsteht dann die Frage, unter welchen Voraussetzungen $F(x)$ partiell nach x_j differenzierbar ist und ob Differentation und Integration vertauschbar sind. Diese Frage wird durch den folgenden Satz beantwortet.

Satz 10.61. *Es sei $U \subset \mathbb{R}^n$ offen, $f : U \times [a,b] \to \mathbb{C}$ stetig und $F : U \to \mathbb{C}$ definiert durch*

$$F(x) = \int_a^b f(x,t)dt.$$

Für jedes $t \in [a,b]$ sei die Funktion $x \longmapsto f(x,t)$ nach x_j partiell differenzierbar und $\partial_i f : U \times [a,b] \to \mathbb{C}$ sei stetig. Dann ist $F(x)$ nach x_j stetig partiell differenzierbar und es gilt

$$\frac{\partial F}{\partial x_j}(x) = \int_a^b \frac{\partial f}{\partial x_j}(x,t)dt.$$

Beweis: Wir können annehmen, daß f reellwertig ist. Sei $x_0 \in U$ und $\epsilon > 0$. Wir setzen

$$\psi(x,t) := \frac{\partial f}{\partial x_j}(x,t) - \frac{\partial f}{\partial x_j}(x_0,t).$$

Dann ist $\psi : U \times [a,b] \to \mathbb{R}$ stetig und ψ verschwindet auf $\{x_0\} \times [a,b]$. Es sei

$$W := \{(x,t) \in U \times [a,b] \mid |\psi(x,t)| < \epsilon\},$$

d.h., $W = \psi^{-1}(-\epsilon,\epsilon)$. Da ψ stetig ist, ist $W \subset U \times [a,b]$ offen und $\{x_0\} \times [a,b] \subset W$.

Lemma. *Es existiert eine offenen Umgebung $V \subset U$ von x_0 so, daß $V \times [a,b] \subset W$.*

Beweis: Da $W \subset U \times [a,b]$ offen ist, existiert zu jedem $t \in [a,b]$ eine offene Umgebung V_t von x_0 in U und eine offenen Umgebung I_t von t in $[a,b]$ mit

$$V_t \times I_t \subset W.$$

Dann ist $\{I_t \mid t \in [a,b]\}$ eine offene Überdeckung von $[a,b]$. Da $[a,b]$ kompakt ist, existieren $t_1, \ldots, t_r \in [a,b]$ so, daß

$$[a,b] = I_{t_1} \cup \cdots \cup I_{t_r}.$$

Es sei

$$V := V_{t_1} \cap \cdots \cap V_{t_r}.$$

Dann ist $V \subset U$ eine offene Umgebung von x_0 und es gilt

$$V \times [a,b] = V \times (\bigcup_{i=1}^{r} I_{t_i}) = \bigcup_{i=1}^{r} (V \times I_{t_i}) \subset \bigcup_{i=1}^{r} (V_{i_i} \times I_{t_i}) \subset W.$$

□

Wir wählen die offene Umgebung $V \subset U$ von x_0 gemäß des Lemmas. Da V offen ist, existiert $\delta > 0$ mit $x_0 + se_j \in V$ für alle $s \in \mathbb{R}$ mit $|s| < \delta$. Wir betrachten nun die Funktionen $s \longmapsto F(x_0 + se_j)$ bzw. $s \longmapsto f(x_0 + se_j, t)$, $|s| < \delta$. Nach Voraussetzung ist $s \longmapsto f(x_0 + se_j, t)$ für jedes $t \in [a,b]$ stetig differenzierbar. Aus dem Mittelwertsatz folgt: Zu s, t existiert $\theta \in (0,1)$ so, daß

$$f(x_0 + se_j, t) - f(x_0, t) = \frac{\partial f}{\partial x_j}(x_0 + \theta se_j, t)s \qquad (10.21).$$

Sei $|s| < \delta$. Dann ist $x_0 + \theta se_j \in V$ und daher $(x_0 + \theta se_j, t) \in W$. Auf Grund der Definition von W ist

$$\left| \frac{\partial f}{\partial x_j}(x_0 + \theta se_j, t) - \frac{\partial f}{\partial x_j}(x_0, t) \right| < \epsilon \qquad (10.22)$$

für alle $t \in [a,b]$. Aus (10.21) und (10.22) erhalten wir

$$\begin{aligned}
&\left| \frac{F(x_0 + se_j) - F(x_0)}{s} - \int_a^b \frac{\partial f}{\partial x_j}(x_0, t)dt \right| \\
&\quad = \left| \int_a^b \left\{ \frac{f(x_0 + se_j, t) - f(x_0, t)}{s} - \frac{\partial f}{\partial x_j}(x_0, t) \right\} dt \right| \\
&\quad \leq \int_a^b \left| \frac{f(x_0 + se_j, t) - f(x_0, t)}{s} - \frac{\partial f}{\partial x_j}(x_0, t) \right| dt \\
&\quad = \int_a^b \left| \frac{\partial f}{\partial x_j}(x_0 + \theta se_j, t) - \frac{\partial f}{\partial x_j}(x_0, t) \right| dt < \epsilon(b-a).
\end{aligned}$$

Daher ist

$$\frac{\partial F}{\partial x_j}(x_0) = \int_a^b \frac{\partial f}{\partial x_j}(x_0, t)dt.$$

Die Stetigkeit folgt leicht auf Grund der Stetigkeit von $\partial_i f(x,t)$. Übung!

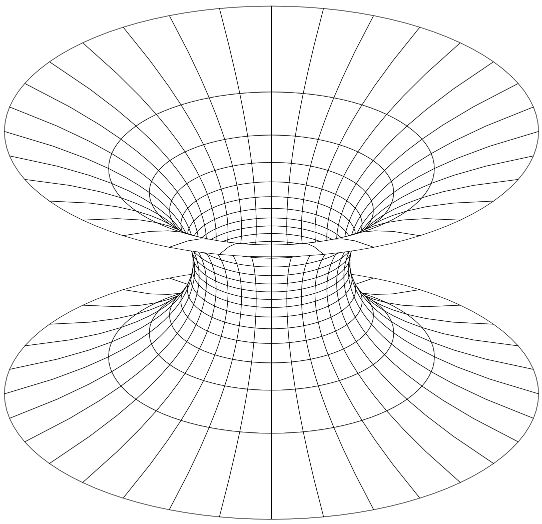
□

10.16. Die Eulersche Differentialgleichung der Variationsrechnung

Der Satz 10.61 über die Differentation von parameterabhängigen Integralen findet unter anderem in der Variationsrechnung wichtige Anwendungen. In Naturwissenschaft und Technik werden viele Gleichgewichtszustände und Bewegungsvorgänge durch Minimalprinzipien charakterisiert. Dazu gehört z.B. das **Prinzip der kleinsten Wirkung** in der klassischen Mechanik. Es handelt sich dabei um die Untersuchung von kritischen Punkten von Funktionen in "unendlichen vielen Variablen". Solche Probleme treten auch in der Geometrie auf wie z.B. bei der Untersuchung von Minimalflächen.

Beispiel: *Die Rotationsminimalfläche.*

Gegeben seien zwei koaxiale Kreise im $\mathbb{R}^3$. Gesucht ist eine Rotationsfläche, deren Rand diese Kreise bilden und deren Flächeninhalt minimal ist. Eine solche Fläche nennt man Katenoid.



Eine Rotationsminimalfläche (Katenoid).

Mathematische Formulierung: Zu gegebenen Punkten $A = (a, \alpha)$ und $B = (b, \beta)$ mit $a < b$ wird eine C^1-Funktion $y : [a, b] \to \mathbb{R}_+$ mit den Randwerten $y(a) = \alpha$ und $y(b) = \beta$ gesucht so, daß die von dieser Funktion erzeugte Rotationsfläche einen minimalen Flächeninhalt hat. Der Flächeninhalt der durch $y : [a, b] \to \mathbb{R}_+$ definierten Rotationsfläche ist gegeben durch

$$A(y) = 2\pi \int_a^b y(x)\sqrt{1 + y'(x)^2}\, dx.$$

Der Beweis dieser Formel erfolgt in einem späteren Kapitel.

Allgemein läßt sich dieses Problem wie folgt formulieren. Gegeben sei eine C^2-Funktion

$$L : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

sowie Randwerte $\alpha, \beta \in \mathbb{R}$. Die Funktion L heißt häufig **Lagrange-Funktion** oder, in der Physik, **Wirkungsfunktion**. Es sei

$$\mathcal{K} = \{y \in C^2([a, b]) \mid y \,\mathrm{reell},\ y(a) = \alpha,\ y(b) = \beta\}.$$

Weiter sei

$$\mathcal{K}_0 = \{y \in C^2([a, b]) \mid y \,\mathrm{reell},\ y(a) = 0 = y(b)\}.$$

Für alle $\varphi, \psi \in \mathcal{K}_0$, und $\lambda, \rho \in \mathbb{R}$ gilt $\lambda\varphi + \rho\psi \in \mathcal{K}_0$, d.h., $\mathcal{K}_0$ ist ein Vektorraum über $\mathbb{R}$, der allerdings unendlichdimensional ist. Sei $\varphi \in \mathcal{K}$. Dann gilt

$$\mathcal{K} = \varphi + \mathcal{K}_0,$$

d.h., $\mathcal{K}$ ist ein affiner Raum. Die Lagrange Funktion $L(x, y, z)$ induziert eine Funktion

$$J : \mathcal{K} \to \mathbb{R},$$

die definiert ist durch

$$J(y) = \int_a^b L(x, y(x), y'(x))dx, \quad y \in \mathcal{K}.$$

Gesucht wird dann ein Element $\varphi \in \mathcal{K}$ so, daß J in φ sein Minimum (oder Maximum) annimmt, d.h.:

$$\forall y \in \mathcal{K} : J(\varphi) \leq J(y)\, (J(\varphi) \geq J(y))\, .$$

Wir haben es jetzt mit einem Extremalproblem zu tun, dessen Definitionsbereich ein unendlichdimensionaler Raum ist. Wir leiten jetzt eine notwendige Bedingung für einen extremalen Punkt dieser Funktion $J : \mathcal{K} \to \mathbb{R}$ her.

Es sei $\varphi \in \mathcal{K}$. Für jedes $h \in \mathcal{K}_0$ ist $\varphi + th \in \mathcal{K}_0$ für alle $t \in \mathbb{R}$. Es sei

$$F_h(t) = J(\varphi + th) = \int_a^b (L(x, \varphi(x) + th(x), \varphi'(x) + th'(x))dx.$$

Nach Satz 10.61 ist $F_h : \mathbb{R} \to \mathbb{R}$ stetig differenzierbar und es gilt

$$
\begin{aligned}
F_h'(t) &= \int_a^b \frac{d}{dt}\,(L(x,\varphi(x)+th(x),\varphi'(x)+th'(x))dx \\
&= \int_a^b \left(\frac{\partial L}{\partial y}(x,\varphi(x)+th(x),\varphi'(x)+th'(x))h(x)\right. \\
&+ \left.\frac{\partial L}{\partial z}\,(x,\varphi(x)+th(x),\varphi'(x)+th'(x))h'(x)\right)dx.
\end{aligned}
$$

Auf Grund der Voraussetzung über L, φ und h ist der Integrand stetig in (x,t).

Definition 10.62. *Die Ableitung*

$$\frac{\delta J}{\delta h}(\varphi) := F_h'(0)$$

heißt die **erste Variation von** J **in Richtung** h. *Weiter heißt* φ *stationärer Punkt von* J*, wenn gilt*

$$\frac{\delta J}{\delta h}(\varphi) = 0 \text{ für alle } h \in \mathcal{K}_0.$$

Ist $\varphi \in \mathcal{K}$ ein Extremum von J, so hat jede Funktion $F_h(t),\ h \in \mathcal{K}_0$, in $0 \in \mathbb{R}$ ein Extremum. Dann gilt $F_h'(0) = 0$ für alle $h \in \mathcal{K}_0$, d.h., φ ist stationärer Punkt von J. Damit haben wir gezeigt:

Wenn die Funktion J *in* φ *ein Extremum annimmt, so ist* φ *stationär.*

Die Situation ist also analog zum endlichdimensionalen Fall, wo jeder extremale Punkt $x \in U$ einer Funktion $f : U \to \mathbb{R}$ notwendig ein kritischer Punkt ist, d.h., für den $\operatorname{grad} f(x) = 0$ ist.

Stationäre Punkte von J werden durch die **Eulersche Differentialgleichung der Variationsrechnung** charakterisiert. Das ist der Inhalt des folgenden Satzes.

Satz 10.63. $\varphi \in \mathcal{K}$ *ist ein stationärer Punkt von* J *genau dann, wenn für* φ *auf* $[a,b]$ *die Gleichung*

$$\frac{d}{dx}\frac{\partial L}{\partial z}(x,\varphi(x),\varphi'(x)) = \frac{\partial L}{\partial y}(x,\varphi(x),\varphi'(x)) \tag{10.23}$$

gilt.

Beweis: Aus obiger Berechnung von $F_h'(t)$ folgt, daß J in φ genau dann stationär ist, wenn für jedes $h \in \mathcal{K}_0$ gilt

$$\int_a^b \left(\frac{\partial L}{\partial y}(x,\varphi(x),\varphi'(x)h(x) + \frac{\partial L}{\partial z}(x,\varphi(x),\varphi'(x))h'(x)\right)dx = 0.$$

Durch partielle Integration folgt

$$\int_a^b \frac{\partial L}{\partial z}(x, \varphi(x), \varphi'(x))h'(x)dx = -\int_a^b \left(\frac{d}{dx}\frac{\partial L}{\partial z}(x, \varphi(x), \varphi'(x))\right) h(x)dx.$$

Dabei haben wir benutzt, daß $h(a) = h(b) = 0$ gilt. Daraus folgt, daß J in φ stationär ist genau dann, wenn für alle $h \in \mathcal{K}_0$ gilt:

$$\int_a^b \left(\frac{\partial L}{\partial y}(x, \varphi(x), \varphi'(x)) - \frac{d}{dx}\frac{\partial L}{\partial z}(x, \varphi(x), \varphi'(x))\right) h(x)dx = 0. \tag{10.24}$$

Lemma 10.64. *Es sei $f : [a, b] \to \mathbb{R}$ stetig und für jede C^2-Funktion $h : [a, b] \to \mathbb{R}$ mit $h(a) = h(b) = 0$ gelte*

$$\int_a^b f(x)h(x)dx = 0.$$

Dann ist $f = 0$.

Beweis: Angenommen es existiere $x_0 \in (a, b)$ mit $f(x_0) \neq 0$. Sei z.B. $f(x_0) > 0$. Wir wählen $[a', b'] \subset (a, b)$ so, daß $x_0 \in [a', b']$ und

$$f(x) \geq \frac{1}{2}f(x_0) \quad \text{für alle} \quad x \in [a', b'].$$

Da $f : [a, b] \to \mathbb{R}$ stetig ist, existiert ein solches Intervall. Wir wählen eine C^2-Funktion $h : [a, b] \to \mathbb{R}$ mit

(1) $h(x) > 0$ für alle $x \in (a', b')$.
(2) $h(x) = 0$ für alle $x \in [a, b] - (a', b')$.

Zum Beispiel ist

$$h(x) := \begin{cases} (x - a')^4(x - b')^4, & x \in (a', b'); \\ 0, & x \notin (a', b'). \end{cases}$$

eine solche Funktion. Dann gilt

$$\int_a^b f(x)h(x)dx \geq \frac{1}{2}f(x_0)\int_a^b h(x)dx > 0.$$

Dies ergibt einen Widerspruch zur Voraussetzung. Also ist $f = 0$.

□

Aus (10.24) zusammen mit dem Lemma folgt

$$\frac{\partial L}{\partial y}(x.\varphi(x), \varphi'(x)) = \frac{d}{dx}\frac{\partial L}{\partial z}(x, \varphi(x), \varphi'(x)).$$

□

Die Eulersche Gleichung für eine stationäre Lösung ist eine Differentialgleichung 2. Ordnung. Ausgeschrieben lautet sie

$$\frac{\partial^2 L}{\partial z^2}(x,\varphi(x),\varphi'(x))\varphi''(x) + \frac{\partial^2 L}{\partial y \partial z}(x,\varphi(x),\varphi'(x))\varphi'(x)$$
$$+ \frac{\partial^2 L}{\partial x \partial z}(x,\varphi(x),\varphi'(x)) - \frac{\partial L}{\partial y}(x,\varphi(x),\varphi'(x)) = 0.$$

Daraus erhalten wir folgendes nützliche Lemma.

Lemma 10.65. *Wenn L nicht von x abhängt, so gilt für jede Lösung φ von (10.23)*

$$E_\varphi(x) := \frac{\partial L}{\partial z}(\varphi(x),\varphi'(x))\varphi'(x) - L(\varphi(x),\varphi'(x))$$

ist konstant.

Beweis: Wegen $\dfrac{\partial L}{\partial x} = 0$ folgt aus obiger Gleichung

$$\begin{aligned}
\frac{d}{dx}E_\varphi(x) &= \frac{\partial^2 L}{\partial y \partial z}(\varphi(x),\varphi'(x))\varphi'(x)^2 + \frac{\partial^2 L}{\partial z^2}(\varphi(x),\varphi'(x))\varphi'(x)\varphi''(x) \\
&\quad + \frac{\partial L}{\partial z}(\varphi(x),\varphi'(x))\varphi''(x)) - \frac{\partial L}{\partial y}(\varphi(x),\varphi'(x))\varphi'(x) \\
&\qquad\qquad - \frac{\partial L}{\partial z}(\varphi(x),\varphi'(x))\varphi''(x) \\
&= \varphi'(x)\left(\frac{\partial^2 L}{\partial y \partial z}(\varphi(x),\varphi'(x))\varphi'(x) + \frac{\partial^2 L}{\partial z^2}(\varphi(x),\varphi'(x))\varphi''(x)\right. \\
&\qquad\qquad \left. - \frac{\partial L}{\partial y}(\varphi(x),\varphi'(x))\right) = 0.
\end{aligned}$$

□

Beispiel: *Rotationsminimalfläche.*

Das zu minimierende Integral ist bis auf eine Konstante

$$J(y) = \int_a^b y(x)\sqrt{1+y'(x)^2}dx.$$

Die Lagrange-Funktion ist demnach gegeben durch

$$L(x,y,z) = y\sqrt{1+z^2}.$$

Wenn eine Minimallösung φ existiert, so gilt nach Satz 10.63 für φ die Differentialgleichung

$$\frac{d}{dx}\left(\varphi(x)\frac{\varphi'}{\sqrt{1+\varphi'(x)^2}}\right) - \sqrt{1+\varphi'(x)^2} = 0 \qquad (10.25).$$

Da die Funktion L nicht von x abhängt, folgt aus dem letzten Lemma

$$\frac{\partial L}{\partial z}(\varphi, \varphi')\varphi' - L(\varphi, \varphi') = c$$

für eine Konstante $c \in \mathbb{R}$. Daraus folgt

$$\begin{aligned} c &= \frac{\varphi\varphi'}{\sqrt{1+\varphi'^2}}\varphi' - \varphi\sqrt{1+\varphi'^2} \\ &= \varphi\left(\frac{\varphi'^2}{\sqrt{1+\varphi'^2}} - \frac{1}{\sqrt{1+\varphi'^2}}\right) = \frac{\varphi}{\sqrt{1+\varphi'^2}}. \end{aligned}$$

Wir setzen voraus, daß $\alpha, \beta > 0$ gilt. Dann ist $c > 0$. Durch Einsetzen von $\varphi/\sqrt{1+\varphi'^2} = c$ in (10.25) erhalten wir die Differentialgleichung

$$\varphi'' = \frac{1}{c^2}\varphi.$$

Als allgemeine Lösung dieser Differentialgleichung ergibt sich

$$\varphi(x) = c\cosh\left(\frac{1}{c}(x - x_0)\right).$$

Diese Funktionen stellen die sogenannte **Kettenlinie** dar. Die von ihnen erzeugten Rotationsflächen heißen **Katenoide**.

Man muß nun noch die Konstanten c und x_0 so bestimmen, daß die Randbedingungen $\varphi(a) = \alpha$ und $\varphi(b) = \beta$ erfüllt sind. Wir beschränken uns auf den Spetzialfall $\alpha = \beta > 0$. O.B.d.A. nehmen wir dann $a = -b$ und $b > 0$ an.

Aus Symmetriegründen folgt $x_0 = 0$. Zur Bestimmung von c müssen wir dann die Gleichung

$$\alpha = c\cosh\left(-\frac{b}{c}\right) = c\cosh\left(\frac{b}{c}\right)$$

lösen. Diese Gleichung ist äquivalent zu

$$\frac{\cosh(\frac{b}{c})}{(\frac{b}{c})} = \frac{\alpha}{b}. \tag{10.26}$$

Wir betrachten die Funktion

$$f(t) = \frac{1}{t}\cosh(t).$$

Dann zeigt man leicht, daß f folgende Eigenschaften hat:

1. Es existiert $t_0 > 0$ so, daß f in $(0, t_0]$ streng monoton fällt und in $[t_0, \infty)$ streng monoton wächst.
2. $\lim_{t\downarrow 0} f(t) = \infty$ und $\lim_{t\to\infty} f(t) = \infty$. Damit ist $f(t_0)$ das absolute Minimum von $f(t)$ in $[0, \infty)$. Aus 1) und 2) folgt:

1. Für $\alpha/b < f(t_0)$ hat die Gleichung (10.26) keine Lösung.
2. Für $\alpha/b = f(t_0)$ gibt es genau eine Lösung $c = b/t_0$ von (10.26).
3. Für $\alpha/b > f(t_0)$ existieren zwei Lösungen.

Da die Euler-Lagrange Gleichung (10.23) nur ein notwendiges Kriterium für die Existenz eines Minimums ist, haben wir auch im Falle $\alpha/b > f(t_0)$ noch nicht gezeigt, daß ein Minimum wirklich existiert. Wir haben lediglich gezeigt, daß es im Falle $\alpha/b > f(t_0)$ **höchstens** zwei Lösungen geben kann. Der Existenznachweis für das Vorliegen eines Minimum ist im allgemeinen wesentlich schwieriger und kann hier nicht diskutiert werden.

2. Hamiltonsches Extremalprinzip

In der klassischen Mechanik ist die allgemeine Formulierung des Bewegungsgesetzes eines mechanischen Systems durch das sogenannte **Prinzip der kleinsten Wirkung** auch **Hamiltonsches Extremalprinzip** genannt, gegeben.

Extremalprinzip: Jedem mechanischen System mit m Freiheitsgraden $q = (q_1, \dots , q_m)$ ist eine bestimmte Funktion

$$L(t, q_1, \dots , q_m, \dot{q}_1, \dots , \dot{q}_m),$$

die **Lagrangefunktion** des Systems, zugeordnet. Weiter sei eine physikalische Bahnkurve, d.h., eine Lösung der Bewegungsgleichungen

$$\varphi(t) = \{\varphi_1(t), \dots , \varphi_m(t)\}, \quad t_1 \leq t \leq t_2$$

gegeben, die die vorgegebenen Randwerte $\varphi(t_1) = a$ und $\varphi(t_2) = b$ annimmt. Diese Bahnkurve minimiert das **Wirkungsintegral**

$$S(q) = \int_{t_1}^{t_2} L(t, q(t), \dot{q}(t))\, dt,$$

wobei über alle C^2-Funktionen $q : [t_1, t_2] \to \mathbb{R}^m$ mit $q(t_1) = a$, $q(t_2) = b$ variiert wird.

Wir haben es jetzt mit einer Lagrangefunktion $L(t, q, \dot{q})$ zu tun, für die $q, \dot{q} \in \mathbb{R}^m$ ist, d.h., gegeben ist eine C^2-Funktion

$$L : [a, b] \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}.$$

Satz 10.63 kann leicht auf diesen Fall verallgemeinert werden. Da alle $q_i(t)$ unabhängig voneinander variiert werden, erhält man als notwendige Bedingung für die Extremalität von $S(q)$ in $q = \varphi$, daß $\varphi(t)$ Lösung der Euler-Lagrange Gleichungen

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots , m, \tag{10.27}$$

ist. Dies sind die gesuchten Bewegungsgleichungen des mechanischen Systems.

Als Beispiel betrachten wir ein System von n Teilchen mit Potentialkräften. Das Potential sei $U(r_1, \dots, r_n, t)$. Dann hat die Lagrangefunktion dieses Systems die natürliche Form

$$L = T - U = \frac{1}{2} \sum_{i=1}^{n} m_i \dot{r}_i^2 - U(r_1, \dots, r_n, t). \tag{10.28}$$

Hierbei ist $m = 3n$ und

$$(q_1, \dots, q_{3n}) = (r_1, \dots, r_n).$$

Dann ist

$$\frac{\partial L}{\partial q_k} = -\frac{\partial U}{\partial q_k}, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = m_{i(k)} \frac{d^2 q_k}{dt^2}$$

Die Notation $m_{i(k)}$ deutet dabei an, daß man beim Abzählen der q_k die jeweils richtige Masse des zugehörigen Teilchens einsetzen soll. Anders geschrieben ergeben diese Gleichungen

$$m_i \frac{d^2 r_i}{dt^2} = -\nabla_i U.$$

Die Euler-Lagrangegleichungen sind also genau die Newtonschen Gleichungen.

Wenn die Lagrangefuntkion nicht von der Zeit abhängt, so erhält man als Verallgemeinerung von Lemma 10.65, daß die Funktion

$$H(q, \dot{q}) = \sum_{k=1}^{m} \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L(q, \dot{q})$$

für jede Lösung $\varphi(t)$ der Euler-Lagrange Gleichungen (10.27) konstant ist. In der Physik nennt man $H(q, \dot{q})$ eine Konstante der Bewegung. Für ein n-Teilchensystem mit Potentialkräften und einem zeitunabhängigen Potential U ergibt sich

$$H(r, \dot{r}) = 2T - (T - U) = T + U = \frac{1}{2} \sum_{i=1}^{m} m_i \dot{r}_i^2 + U(r_1, \dots, r_n).$$

Dies ist die Energie des Systems.

$$\frac{\partial H}{\partial t} = 0$$

ist der **Energieerhaltungssatz**. Man nennt die Größe

$$p_k := \frac{\partial L}{\partial \dot{q}_k}$$

den zu q_k **kanonisch konjugierten Impuls**. Aus dem Satz über implizite Funktionen folgt, daß die Funktionen

$$p_k = \frac{\partial L}{\partial \dot{q}_k}(t, q, \dot{q}), \quad k = 1, \ldots, m, \tag{10.29}$$

genau dann lokal nach $\dot{q}_k$ aufgelöst werden können, wenn die Bedingung

$$\det\left(\left(\frac{\partial^2 L}{\partial \dot{q}_k \partial \dot{q}_l}\right)_{k,l=1}^{m}\right) \neq 0$$

erfüllt ist. Es ist dann $\dot{q}_k = \dot{q}_k(t, q, p)$ und die **Hamiltonfunktion** ist durch

$$H(t, q, p) = \sum_{k=1}^{m} p_k \dot{q}_k - L(t, q, \dot{q})$$

gegeben. Daraus folgt

$$\frac{\partial H}{\partial p_k} = \dot{q}_k, \quad k = 1, \ldots, m,$$

und aus den Euler-Lagrangegleichungen (10.27) und (10.29) folgt

$$\frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k} = -\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = -\dot{p}_k.$$

Diese Gleichungen sind die sogenannten kanonischen Bewegungsgleichugen des Systems:

$$\boxed{\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad k = 1, \ldots, m.}$$

Man kann nun umgekehrt die Hamiltonfunktion als Ausgangspunkt wählen. Die kanonischen Bewegungsgleichungen ergeben sich als Euler-Lagrangegleichungen für die Funktion

$$F(t, q, p, \dot{q}, \dot{p}) = \sum_{k=1}^{m} p_k \dot{q}_k - H(t, p, q).$$

10.17. Kurven im Euklidischen Raum

Es sei $I \subset \mathbb{R}$ ein Intervall und

$$\gamma : I \to \mathbb{R}^n$$

sei eine stetige Abbildung. γ heißt parametrisierte (stetige) Kurve im $\mathbb{R}^n$. Sei $I = [a, b]$ und sei

$$Z = \{t_0, t_1, \dots, t_m\} \subset [a, b] \text{ mit } t_0 = a < t_1 < \cdots < t_m = b.$$

Z heißt *Zerlegung* von I. Sei

$$\delta(Z) = \max\{t_i - t_{i-1} \mid i = 1, \dots, m\}.$$

$\delta(Z)$ heißt Feinheit der Zerlegung Z. Für eine Zerlegung Z von I sei

$$L_Z(\gamma) = \sum_{i=1}^{m} \parallel \gamma(t_i) - \gamma(t_{i-1}) \parallel .$$

$L_Z(\gamma)$ ist die Länge des Sehnenpolynoms, das zu Z gehört. Sei $Z' \subset I$ eine weitere Zerlegung von I. Z' heißt *Verfeinerung von* Z, wenn $Z \subset Z'$. Für eine Verfeinerung Z' von Z folgt aus der Dreiecksungleichung

$$L_{Z'}(\gamma) \geq L_Z(\gamma).$$

Definition 10.66. *Die Kurve γ heißt rektifizierbar, wenn*

$$L(\gamma) := \sup_Z\{L_Z(\gamma) \mid Z \subset I \text{ Zerlegung}\} < \infty.$$

$L(\gamma)$ heißt Länge von γ.

Satz 10.67. *Sei $\gamma : I \to \mathbb{R}^n$ stetig differenzierbar, dann gilt*

1) γ ist rektifizierbar.
2) $L(\gamma = \int_a^b \parallel \gamma'(t) \parallel dt$.

Beweis: 1) Es sei $Z = \{t_0, \dots, t_m\}$ eine Zerlegung von $[a, b]$. Dann gilt

$$\begin{aligned} L_Z(\gamma) &= \sum_{i=1}^{m} \parallel \gamma(t_i) - \gamma(t_{i-1}) \parallel = \sum_{i=1}^{m} \parallel \int_{t_{i-1}}^{t_i} \gamma'(t) dt \parallel \\ &\leq \sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} \parallel \gamma'(t) \parallel dt = \int_a^b \parallel \gamma'(t) \parallel dt. \end{aligned}$$

2) Sei $\epsilon > 0$ und sei $Z = \{t_0, \dots, t_m)$ $\delta(Z) = \max\{t_i, t_{i-1} \mid i = 1, \dots, n\}$ eine Zerlegung von I. Weiter sei $\gamma = (\gamma_1, \dots, \gamma_n)$. Dann ist

$$\begin{aligned} L_Z(\gamma) &= \sum_{i=1}^{m} \parallel \gamma(t_i) - \gamma(t_{i-1}) \parallel = \sum_{i=1}^{m} \sqrt{\sum_{j=1}^{n} (\gamma_j(t_i) - \gamma_j(t_{i-1}))^2} \\ &= \sum_{i=1}^{m} \sqrt{\sum_{j=1}^{n} \left(\frac{\gamma_j(t_i) - \gamma_j(t_{i-1})}{t_i - t_{i-1}} \right)^2 (t_i - t_{i-1})}. \end{aligned} \tag{0.6}$$

Da für alle j, $j = 1, \ldots, n$, $\gamma_j : [a,b] \to \mathbb{R}$ stetig differenziebar ist, existieren nach dem Mittelwertsatz $\xi_{ij} \in [t_{i-1}, t_i]$ mit

$$\frac{\gamma_j(t_i) - \gamma_j(t_{i-1})}{t_i - t_{i-1}} = \gamma_j'(\xi_{ij}).$$

Da $\gamma_j' : [a,b] \to \mathbb{R}$ stetig ist und $[a,b] \subset \mathbb{R}$ kompakt ist, ist γ_j' gleichmäßig stetig auf $[a,b]$. Daraus folgt, daß zu $\epsilon > 0$ ein $\delta > 0$ existiert so, daß

$$|\gamma_j'(s) - \gamma_j'(t)| < \frac{\epsilon}{\sqrt{n}(b-a)}$$

für alle $s, t \in [a,b]$ mit $|s-t| < \delta$. Sei $\delta(Z) < \delta$. Dann folgt aus (10.30)

$$\begin{aligned}
\left|\int_a^b \| \gamma'(t) \| \, dt - L_Z(\gamma)\right| &= \left|\int_a^b \| \gamma'(t) \| \, dt - \sum_{i=1}^m \sqrt{\sum_{j=1}^n \gamma_j'(\xi_{ij})^2 (t_i - t_{i-1})}\right| \\
&= \left|\sum_{i=1}^m \int_{t_{i-1}}^{t_i} \left(\| \gamma'(t) \| - \sqrt{\sum_{j=1}^n \gamma_j'(\xi_{ij})^2} \right) dt\right| \\
&\le \sum_{i=1}^m \int_{t_{i-1}}^{t_i} |\| \gamma'(t) \| - \| (\gamma_1'(\xi_{i1}), \ldots, \gamma_n'(\xi_{in})) \| | \, dt \\
&\le \sum_{i=1}^m \int_{t_{i-1}}^{t_i} \| \gamma'(t)_(\gamma_1'(\xi_{i1}), \ldots, \gamma_n'(\xi_{in})) \| \, dt \\
&= \sum_{i=1}^m \int_{t_{i-1}}^{t_i} \sqrt{\sum_{j=1}^n (\gamma_j'(t) - \gamma_j'(\xi_{ij}))^2 dt} \\
&\le \sum_{i=1}^m \int_{t_{i=1}}^{t_i} \frac{\epsilon}{b-a} dt = \epsilon.
\end{aligned}$$

□

Sei $\gamma : I \to \mathbb{R}^n$ eine C^1-Kurve und sei $\varphi : J \to I$ ein Diffeomorphismus. Weiter sei $\widetilde{\gamma} : J \to \mathbb{R}$ definiert durch

$$\widetilde{\gamma} = \gamma \circ \varphi.$$

$\widetilde{\gamma}$ heißt *Umparametrisierung* von γ.

Satz 10.68. *Es gilt*

$$L(\gamma) = L(\widetilde{\gamma}).$$

Beweis: 1) Es sei $\varphi'(t) > 0$ für alle $t \in J$. Sei $J = [\alpha, \beta]$. Da φ streng monoton wachsend ist, ist

$$\varphi(a) = \alpha, \; \varphi(b) = \beta.$$

Daraus folgt

$$\begin{aligned} L(\gamma) &= \int_a^b \parallel \gamma'(t) \parallel dt = \int_\alpha^\beta \parallel \gamma'(\varphi(t)) \parallel \varphi'(t)dt \\ &= \int_\alpha^\beta \parallel \widetilde{\gamma}'(\tau) \parallel d\tau = L(\widetilde{\gamma}). \end{aligned}$$

2) Sei $\varphi'(t) < 0$ für alle $t \in J$. Dann ist φ streng monoton fallend und daher

$$\varphi(a) = \beta \text{ und } \varphi(b) = \alpha.$$

Daraus folgt

$$\begin{aligned} L(\gamma) &= \int_a^b \parallel \gamma'(t) \parallel dt = \int_\beta^\alpha \parallel \gamma'(\varphi(\tau)) \parallel \varphi'(\tau))d\tau \\ &= -\int_\beta^\alpha \parallel \gamma'(\varphi(\tau)) \parallel |\varphi'(\tau)|d\tau \\ &= \int_\alpha^\beta \parallel \widetilde{\gamma}'(\tau) \parallel d\tau = L(\widetilde{\gamma}). \end{aligned}$$

□

Sei $\gamma : [a, b] \to \mathbb{R}^n$ eine C^1-Kurve. Dann heißt

$$s(t) = \int_a^t \parallel \gamma'(u) \parallel du,\ t \in [a, b]$$

die Bogenlänge von γ.
γ heißt nach der Bogenlänge parametrisiert, wenn

$$\parallel \gamma'(t) \parallel = 1 \text{ für alla } t \in [a, b] \Leftrightarrow s(t) = t - a.$$

Definition 10.69. *Sei $\gamma : [a, b] \to \mathbb{R}^n$ eine C^1-Kurve. γ heißt regulär in $t_0 \in [a, b]$ wenn $\gamma'(t_0) \neq 0$.*

γ heißt regulär, wenn γ in jedem Punkt $t \in [a, b]$ regulär ist.

Satz 10.70. *Sei $\gamma : [a, b] \to \mathbb{R}$ eine reguläre C^1-Kurve. Dann existiert eine Umparametrisierung $\widetilde{\gamma}$ von γ so, daß nach der Bogenlänge parametrisiert ist.*

Beweis: Es gilt

$$s'(t) = \parallel \gamma'(t) \parallel > 0,\ t \in [a, b] \Rightarrow s : [a, b] \to [0, L(\gamma)]$$

ist Diffeomorphismus.

Sei $\varphi = s^{-1} : [0, L(\gamma)] \to [a, b]$. φ ist eine C^1-Funktion und es gilt

$$\varphi'(\tau) = \frac{1}{\| \gamma'(\varphi(\tau)) \|}.$$

Sei $\widetilde{\gamma} = \gamma \circ \varphi$. Dann folgt

$$\| \widetilde{\gamma}'(\tau) \| = \| \gamma'(\varphi(\tau))\varphi'(\tau) \| = |\varphi'(\tau)| \, \| \gamma'(\varphi(\tau)) \| = 1.$$

$\square$