Analysis I und II - Notizen[1]

Daniel Lenz

[1]Es handelt sich nicht um ein Skriptum zur Vorlesung. Besten Dank an alle, die zu Verbesserungen früherer Notizen zur Analysis I beigetragen haben, und besonderen Dank an Daniel Kilian, Stefan Neumann und Frank Nußbaum für systematisches Durcharbeiten und Verschönern

Inhaltsverzeichnis

Vorbemerkung: Zwei kleine Rechnungen

Eine Rechnung

Der Kurs wird mit 9 Leistungspunkten (LP) gewertet. Jeder Leistungspunkt entspricht 30h Arbeit. Damit geht es um

$$270h \text{ Arbeit}$$

Davon gehen ab:

$$-90h \quad \text{(6 h Vorlesung und Übung in 15 Wochen)}$$

$$-80h \quad \text{(2 Wochen Prüfungsvorbereitung à 40 h).}$$

Damit verbleiben noch 100 h Arbeit. Auf 15 Wochen verteilt bedeutet dies ca.

7 h Arbeit/ Woche

also

1 h Arbeit / Tag.

Eine andere Rechnung

Der Stoff der ersten drei Semester wurde beginnend mit Newton und Leibniz um 1670 bis etwa 1920 entwickelt. Es handelt sich also um

250 Jahre Entwicklung.

Bei 45 Wochen für die ersten drei Semester, wird also in einer Woche Vorlesung etwa

5 Jahre Entwicklung $\sim$ 260 Wochen

behandelt.

Folgerung

Es muß gearbeitet werden!

Grundlagen

Wir setzen Grundtatsachen der Mengenlehre voraus und erinnern in diesem Kapitel an einige Begriffe und Bezeichnungen.

1. Mengen

Ist X eine Menge so schreiben wir $x \in X$ (lies: x Element von X) falls x zu X gehört. Gehört x nicht zu X, so schreibt man $x \notin X$. Seien X und Y Mengen. Dann bedeutet $Y \subset X$ (lies: Y Teilmenge von X), daß jedes Element von Y auch zu X gehört. Es bezeichnet $X \setminus Y$ (lies: X ohne Y) die Menge der Elemente von X, die nicht zu Y gehören. Der *Durchschnitt* $X \cap Y$ der Mengen X und Y ist gegeben durch

$$X \cap Y := \{z : z \in X \text{ und } z \in Y\}.$$

Die *Vereinigung* der Mengen X und Y ist gegeben durch

$$X \cup Y := \{z : z \in X \text{ oder } z \in Y\}.$$

Die *Potenzmenge* $\mathcal{P}(X)$ einer Menge X ist die Menge aller Teilmengen von X. Die leere Menge wird mit $\emptyset$ bezeichnet.
Ist A eine Menge und zu jedem $\alpha \in A$ eine Menge X_α gegeben, so nennt man X_α, $\alpha \in A$, eine *Familie von Mengen.* Für eine Familie X_α, $\alpha \in A$, von Mengen definiert man

$$\bigcup_{\alpha \in A} X_\alpha := \{z : \text{z gehört zu (mindestens) einer der Mengen } X_\alpha\}$$

sowie

$$\bigcap_{\alpha \in A} X_\alpha := \{z : \text{z gehört zu allen Mengen } X_\alpha\}.$$

Ist X_α, $\alpha \in A$, eine Familie von Mengen und X eine Menge, so gilt (siehe Übung)

$$X \setminus \bigcup_\alpha X_\alpha = \bigcap_\alpha X \setminus X_\alpha$$

$$X \setminus \bigcap_\alpha X\alpha = \bigcup_\alpha X \setminus X_\alpha.$$

Aus zwei Objekten a, b bilden wir das *geordnete Paar* (a, b). Damit können wir aus zwei Mengen X, Y das *kartesische Produkt*

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

bilden.

2. Funktionen

Eine *Abbildung* oder *Funktion* f von einer Menge X in die Menge Y ist eine Zuordnung, die jedem Element von X genau ein Element von Y zuordnet. Wir schreiben

$$f : X \longrightarrow Y \text{ oder } X \longrightarrow Y, x \mapsto f(x).$$

Es heißt dann X der *Definitionsbereich* von f, Y der *Wertebereich* von f und $Bild(f) := \{f(x) : x \in X\} \subset Y$ das *Bild* von f. Ist $f : X \longrightarrow Y$ eine Funktion und $A \subset Y$, so heißt

$$f^{-1}(A) := \{x \in X : f(x) \in A\}$$

das *Urbild* von A unter f.

Beispiel. Sei X eine beliebige Menge. Dann ist $id_X : X \longrightarrow X$ die Abbildung, die $x \in X$ auf $x \in X$ abbildet. Etwa: $X = \{1, 2, 3\} \longrightarrow \{1, 2, 3\}, 1 \mapsto 1, 2 \mapsto 2, 3 \mapsto 3$.

Die *Komposition* $g \circ f$ der Abbildungen $f : X \longrightarrow Y$, $g : Y \longrightarrow Z$ ist gegeben durch

$$g \circ f : X \longrightarrow Z, x \mapsto g(f(x)).$$

Komposition ist assoziativ d.h. es gilt

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

(Denn $g \circ (f \circ h)(x) = g((f \circ h)(x)) = g((f(h(x)))) = (g \circ f)(h(x)) = (g \circ f) \circ h(x)$.) Damit können wir also die Klammern weglassen.
Eine Abbildung $f : X \longrightarrow Y$ heißt *surjektiv*, wenn zu jedem $y \in Y$ ein $x \in X$ existiert mit $y = f(x)$ (d.h. $Bild(f) = Y$).
Eine Abbildung $f : X \longrightarrow Y$ heißt *injektiv*, wenn aus $x \neq z$ folgt $f(x) \neq f(z)$.
Ist $f : X \longrightarrow Y$ injektiv und surjektiv, so heißt es bijektiv.

Behauptung. Sei $f : X \longrightarrow Y$ gegeben. Dann sind äquivalent:

(i) f bijektiv.
(ii) Es gibt ein $g : Y \longrightarrow X$ mit $f \circ g = id_Y$ und $g \circ f = id_X$.

In diesem Fall, ist die Funktion g aus (ii) eindeutig bestimmt. Beweis. (Übung)

Die Funktion g wird *Umkehrfunktion* von f genannt und (oft) mit f^{-1} bezeichnet.
Schließlich brauchen wir manchmal noch Einschränkungen von Funktionen: Sei $f : X \longrightarrow Y$ gegeben und A eine Teilmenge von X. Dann bezeichnen wir mit f_A oder $f|_A$ die *Einschränkung* von f auf A gegeben durch

$$f|_A : A \longrightarrow Y, x \mapsto f(x).$$

3. Relationen

Eine *Relation* auf eine Menge X ist eine Teilmenge R von $X \times X$. Statt $(x, y) \in R$ schreibt man oft auch $x \overset{R}{\sim} y$ oder $x \sim y$.
Eine Relation auf X heißt

- *reflexiv*, wenn gilt: $x \sim x$ für alle $x \in X$, (Zeichnung)
- *symmetrisch*, wenn gilt: $x \sim y \Longrightarrow y \sim x$, (Zeichnung)
- *transitiv*, wenn gilt: $x \sim y$ und $y \sim z \Longrightarrow x \sim z$. (Zeichnung in Übung)

Eine *Äquivalenzrelation* ist eine reflexive, symmetrische, transitive Relation.

Beispiel. $X =$ Bewohner von Jena und

$$R_1 = \{(x, y) : \text{x und y haben am selben Tag Geburtstag}\}$$

$$R_2 = \{(x, y) : \text{x und y kennen sich}\}$$

$$R_3 = \{(x, y) : \text{x kennt y }\}$$

Dann ist R_1 eine Aequivalenzrelation. Es ist R_2 reflexiv und symmetrisch aber nicht transitiv. Es ist $R3$ reflexiv.

Eine *Ordnungsrelation* oder *Ordnung* auf einer Menge X ist eine Relation $\leq$, die reflexiv, transitiv und antisymmetrisch ist. Dabei heißt antisymmetrisch, daß

$$x \leq y \text{ und } y \leq x \Longrightarrow x = y.$$

Ist $\leq$ eine Ordnungsrelation auf X, so heißt das Paar $(X, \leq)$ eine *geordnete Menge*. Eine geordnete Menge heißt *total geordnet*, wenn für alle $x, y \in X$ immer $x \leq y$ oder $y \leq x$ gilt.

'Beispiele'. $(\mathbb{N}, \leq)$, $(\mathbb{Z}, \leq)$, $(\mathbb{Q}, \leq)$, $(\mathbb{R}, \leq)$

Ende der 1. Vorlesung.

4. Verknüpfungen

Eine Abbildung $* : X \times X \longrightarrow X$ heißt *Verknüpfung* auf X. Man schreibt oft $x * y$ statt $*(x, y)$. Die Verknüpfung $* : X \times X \longrightarrow X$ heißt

- kommutativ, wenn gilt $x * y = y * x$ für alle $x, y \in X$
- assoziativ, wenn gilt $x * (x * z) = (x * y) * z$ für alle $x, y, z \in X$.

Ist die Verknüpfung assoziativ, so kann man die Klammern auch weglassen.

KAPITEL 1

Die natürlichen Zahlen

In diesem Kapitel lernen wir einen axiomatischen Zugang zu den natürlichen Zahlen kennen. Dieser Zugang liefert das Prinzip der vollständigen Induktion und die Möglichkeit der rekursiven Definition. Das werden wir genauer studieren. Der Zugang erlaubt es ebenfalls, die üblichen Rechenregeln und Eigenschaften zu beweisen. Das werden wir nur ansatzweise verfolgen, da wir diese im folgenden Kapitel als Nebenprodukt einfach erhalten.

Die charakteristische Struktur der natürlichen Zahlen ist folgende:

Zeichnung. $1 \overset{+1}{- - -} > 2 \overset{+1}{- - -} > 3 \overset{+1}{- - -} > 4 \overset{+1}{- - -} > \ldots$.

Das Problem sind die Punkte $'\ldots'$! An der Zeichnung lesen wir ab:

- Jede natürliche Zahl hat einen Nachfolger und verschiedene Zahlen haben verschiedene Nachfolger.
- Beginnt man bei 1 und bildet sukzessive die Nachfolger, so erhält man alle natürlichen Zahlen.
- Die Zahl 1 ist keine Nachfolger.

Eine präzise Fassung dieser Eigenschaften liefern die **Peano Axiome**.

DEFINITION. *(Peano Axiome) Ein Tripel (N, e, ν) bestehend aus einer Menge N zusammen mit einem ausgezeichneten Element e und einer Abbildung $\nu : N \longrightarrow N \setminus \{e\}$ genügt den Peanoaximonen, wenn gilt:*

(P1) *$\nu : N \longrightarrow N \setminus \{e\}$ ist injektiv.*

(P2) *(Induktionsaxiom) Enthält eine Teilmenge M von N das Element e und enthält sie mit jedem Element n immer auch $\nu(n)$, so gilt $M = N$.*

Es heißt ν die Nachfolgeabbildung und $\nu(n)$ der Nachfolger von n.

Weiteres Vorgehen: Die Peano Axiome charakterisieren die natürlichen Zahlen in folgendem Sinne: Man kann beweisen, daß es ein System gibt, daß diesen Axiomen genügt (wenn man Existenz einer unendlichen Menge voraussetzt), und daß ein solches System eindeutig bestimmt ist (bis auf Umbennenung). Dieses System werden wir später mit $\mathbb{N}$ bezeichnen. Wir werden die Eindeutigkeit eines solchen Systemes bald beweisen und die Rechenoperationen im nächsten Kapitel einführen.

Bemerkung. (a) Eine Teilmenge M von N mit $e \in M$ und $\nu(n) \in M$ für alle $n \in M$ wird auch induktiv genannt.

(b) Hinter $(P1)$ verbergen sich mehrere Forderungen, insbesondere folgende:

- ν bildet von N nach $N \setminus \{e\}$ ab.
- ν ist injektiv.
- Es ist e kein Nachfolger.

Diese Forderungen zusammen mit $(P2)$ werden liefern, daß

- $\nu : N \longrightarrow N \setminus \{e\}$ sogar bijektiv ist (s.u.).

Die vorangegangen vier Spiegelstriche werden manchmal als eigene Axiome aufgelistet.
(c) Das Induktionsaxiom liefert eine Art mit $' \ldots '$ umzugehen.
(d) In den Peano Axiomen ist die Forderung, daß e kein Nachfolger ist, wesentlich. Sie garantiert, daß die Menge unendlich viele Elemente hat. Man kann eine endliche Menge N angeben mit ausgezeichnetem Element e und einer injektiven Abbildung $\nu : N \longrightarrow N$, so daß $(P2)$ gilt (Übung).

Folgerung. *Es genüge (N, e, ν) den Peano Axiomen. Dann ist die Abbildung $\nu : N \longrightarrow N \setminus \{e\}$ bijektiv (d.h. e ist kein Nachfolger und jedes andere Element von $\mathbb{N}$ ist ein Nachfolger) und es gilt $\nu(n) \neq n$ für alle $n \in N$.*

Beweis. Bijektivität. Es reicht zu zeigen, daß ν surjektiv ist.
Sei

$$M := \{e\} \cup \{n \in N : n \text{ ist ein Nachfolger d.h. es existiert } m \in N \text{ mit } \nu(m) = n\}.$$

Dann gilt also:

$e \in M$.
$n \in M$ impliziert $\nu(n) \in M$ (da jeder Nachfolger in M ist).

Aus (P2) folgt also $M = N$.
Es gilt $\nu(n) \neq n$ für alle $n \in N$: Sei T die Menge der $n \in N$ mit $\nu(n) \neq n$. Dann gilt $e \in T$ (klar) und aufgrund der Injektivität von ν gilt auch, dass $n \in T$ impliziert $\nu(n) \in T$. Damit ist T induktiv. □

Wir kommen nun zu einer wesentliche Konsequenzen aus den Peanoaxiomen. Diese ist das Prinzip der vollständigen Induktion.

Prinzip der vollständigen Induktion. Sei (N, e, ν) induktiv. Sei für jedes $n \in N$ eine Aussage $A(n)$ gegeben, sodaß gilt:

- $A(e)$ ist wahr. (Induktionsanfang)
- Aus $A(n)$ folgt $A(\nu(n))$. (Induktionsschluss)

Dann ist $A(n)$ wahr für alle $n \in N$.
Beweis. Sei T die Menge der $n \in N$, so daß $A(n)$ wahr ist. Dann folgt aus dem Induktionsaxiom und der Voraussetzung, daß $T = N$. Das ist gerade die Aussage. □

Notation. Statt '$A(e)$ wahr' und 'Aus $A(n)$ folgt $A(\nu(n))$' schreibt man meist '$n = 1$' und '$n \Longrightarrow n + 1$'.

'Beispiel.' Die Summe der ersten n natürlichen Zahlen $S_n := \sum_{k=1}^{n} k$ ist gerade $n(n + 1)/2$.
Bew. $n = 1$: klar
$n \Longrightarrow n+1$: $S_{n+1} = (n+1)+S_n = (n+1)+(n+1)n\frac{1}{2} = (n+1)(1+n\frac{1}{2}) = (n + 1)(2 + n)\frac{1}{2}$.

Eine weitere wesentliche Konsequenz der Peano Axiome ist die Möglichkeit der rekursiven Definition. (Beispiel $n!$). Um das auszuführen, bedarf es einiger Vorbereitungen. Dabei werden wir einige Folgerungen aus den Peano Axiomen ziehen, die auch schon für sich von Interesse sind.

FOLGERUNG. *(Menge der Zahlen, die größer als n sind) Es genüge (N, e, ν) den Peanoaxiomen. Dann gibt es zu jedem $n \in N$ eine eindeutige Menge $M_n \subset N$ mit*

(1) $n \notin M_n$
(2) $\nu(n) \in M_n$
(3) *Ist* $k \in M_n$ *so auch* $\nu(k)$.

Die Mengen M_n erfüllen $M_e = N \setminus \{e\}$ und $M_{\nu(n)} = M_n \setminus \{\nu(n)\}$ für alle $n \in N$.

Beweis. Wir zeigen zunächst Existenz und Eindeutigkeit von Mengen M_n mit den angegebenen Eigenschaften (1), (2), (3). Sei T die Menge der Elemente $n \in N$ für die eine eindeutige Menge M_n mit den gewünschten Eigenschaften existiert. Wir zeigen, daß T induktiv ist.

Es gilt $e \in T$:

Existenz: Die Menge $N \setminus \{e\}$ hat die gewünschten Eigenschaften:

- $e \notin N \setminus \{e\}$: klar.
- $\nu(e) \in N \setminus \{e\}$: klar (da $\nu : N \longrightarrow N \setminus \{e\}$).
- $n \in N \setminus \{e\} \Longrightarrow \nu(n) \in M_e$: klar (da $\nu : N \longrightarrow N \setminus \{e\}$)

Eindeutigkeit: Ist M_e' eine Menge mit den gewünschten Eigenschaften, so betrachten wir $M := \{e\} \cup M_e'$. Dann gilt $e \in M$ und $n \in M$ impliziert $\nu(n) \in M$ (für $n = e$ wegen (2) und für $n \neq e$ wegen (3)). Damit folgt also nach dem Induktionsaxiom $M = N$ und damit $M_e' = N \setminus \{e\}$. Das zeigt die Eindeutgkeit.

Ende der 2. Vorlesung.

$n \in T$ impliziert $\nu(n) \in T$:

Setze $m := \nu(n)$.

Existenz: Wir betrachten die Menge $M_n \setminus \{m\}$. Diese Menge hat die gewünschten Eigenschaften:

- $m \notin M_n \setminus \{m\}$: klar .

- $\nu(m) \in M_n \setminus \{m\}$: $m = \nu(n) \in M_n$ wegen $(2)_n$, also $\nu(m) \in M_n$ wegen (3) angewendet auf M_n. Weiterhin $\nu(m) \neq m$ (s.o.).
- $l \in M_n \setminus \{m\} \Longrightarrow \nu(l) \in M_n \setminus \{m\}$: $l \in M_n$ impliziert $\nu(l) \in M_n$. Weiterhin $\nu(l) \neq m = \nu(n)$, sonst $n = l$ Widerspruch zu $n \notin M_n$.)

Eindeutigkeit: Ist M_m' eine Menge mit den gewünschten Eigenschaften, so erfüllt $\{m\} \cup M_m'$ die charakteristischen Eigenschaften der Menge M_n:

- $n \notin \{m\} \cup M_m'$: $n \neq m = \nu(n)$ (s.o.) und $n \notin M_m'$ (da sonst $m = \nu(n) \in M_m'$. Widerspruch zu $(2)_m$.
- $\nu(n) \in \{m\} \cup M_m'$: klar (da $m = \nu(n)$).
- $l \in \{m\} \cup M_m'$ impliziert $\nu(l) \in \{m\} \cup M_m'$: Für $l = m$ wegen $(2)_m$ und für $l \in M_m'$ wegen $(3)_m$.

Aufgrund der schon bewiesenen Eindeutigkeit gilt dann $\{m\} \cup M_m' = M_n$ und damit also $M_m' = M_n \setminus \{m\}$. Das zeigt die Eindeutigkeit.

Die letzte Aussage wurde mitbewiesen. □

Folgerung (Menge der Zahlen, die kleiner gleich n sind). *Es genüge (N, e, ν) den Peanoaxiomen. Dann gibt es eine eindeutige Familie von Mengen $A_n \subset N$, $n \in N$, mit*

$$A_e = \{e\} \text{ und } A_{\nu(n)} = A_n \cup \{\nu(n)\}.$$

Es gilt $n \in A_n$ sowie $\nu(n) \notin A_n$ für alle $n \in N$. Weiterhin gilt für beliebige $k, n \in N$ noch $A_k \subset A_n$ falls $k \in A_n$ und $A_n \subset A_k$ falls $k \notin A_n$.

Beweis. Wir zeigen zunächst Existenz und Eindeutigkeit solcher A_n:

Existenz. Seien M_n, $n \in N$, die Mengen aus der vorangehenden Folgerung. Nach der vorangegangenen Folgerung gilt $M_e = N \setminus \{e\}$ und $M_{\nu(n)} = M_n \setminus \{\nu(n)\}$. Dann erfüllen die Mengen $A_n := N \setminus M_n$

$$A_e = \{e\}, \text{ sowie } A_{\nu(n)} = A_n \cup \{\nu(n)\}$$

für alle $n \in N$.

Eindeutigkeit. Sei (A_n'), $n \in N$, eine Familie von Teilmengen von N mit $A_e' = \{e\}$ und $A_{\nu(n)}' = A_n' \cup \{\nu(n)\}$. Sei $L := \{n \in N : A_n = A_n'\}$. Dann sieht man sofort, daß L induktiv ist, und es folgt die Eindeutigkeit.

Die Aussage $n \in A_n$ folgt (mit der Fallunterscheidung $n = e$ und $n \neq e$) für alle $n \in N$ aus den charakteristischen Eigenschaften der A_n.

$A_k \subset A_n$ falls $k \in A_n$:
Sei L die Menge aller $n \in N$, für die gilt $A_k \subset A_n$ falls $k \in A_n$. Dann gilt $e \in L$ sowie $\nu(n) \in L$ falls $n \in L$. Damit folgt $L = N$, und es gilt $A_k \subset A_n$ falls $k \in A_n$.

$A_n \subset A_k$ falls $k \notin A_n$:

Wir müssen zeigen, daß $M_k \subset M_n$ falls $k \in M_n$. Es reicht zu zeigen, daß $M_k = M_k \cap M_n$. Dazu reicht es aufgrund der Eindeutigkeit zu zeigen, daß $M_k' := M_k \cap M_n$ die charakteristischen Eigenschaften von M_k hat. Das folgt einfach:
$k \notin M_k' : k \notin M_k$
$\nu(k) \in M_k'$: $\nu(k) \in M_k$ und $\nu(k) \in M_n$ (da $k \in M_n$ und M_n der Eigenschaft (3) geneugt.
$l \in M_k' \Longrightarrow \nu(l) \in M_k'$: klar (da es für beide Bestandteile von M_k' gilt).
Damit ist die Folgerung bewiesen. □

Bemerkung. (Uebung) Man kann zeigen, daß

$$x \leq y :\Longleftrightarrow A_x \subset A_y$$

eine Ordnungrelation auf N definiert und sogar eine Totalordnung. Bezüglich dieser Ordnungsrelation besteht dann A_n aus den Elementen 'kleiner oder gleich' n sind und M_n aus den Elementen die 'größer' als n sind. Bezüglich dieser Ordnungsrelation ist N wohlgeordnet, d.h. es gilt, daß jede nichtleere Teilmenge von N ein kleinstes Element besitzt. (Hinweise: Zur Antisymmetrie: Es muss gezeigt werden, dass y durch A_y bestimmt ist. Dazu reicht es zu zeigen, dass y das einzige Element z ist mit $z \in A_y$ und $\nu(z) \in M_y$. Dazu betrachtet man die Menge L aller $y \in N$ mit dieser Eigenschaft. Dann gehört offenbar e zu L. Weiterhin gehört mit n auch $\nu(n)$ zu L. Damit gilt die gewünschte Eigenschaft für alle $y \in N$. Zur Wohlgeordnetheit: Angenommen: M ist eine nichtleere Teilmenge von N, die kein kleinstes Element besitzt. Zeige dann durch Induktion, daß jedes A_n im Komplement von M liegt. Damit stimmt dieses Komplement mit N überein und M ist die leere Menge.)

Wir kommen nun zur rekursiven Definition (vgl. $n!$ oben).

Rekursive Definition von Funktionen. Es genuege (N, e, ν) den Peano Axiomen. Sei X eine Menge und für $n \in N$ sei X^n die Menge der Abbildungen von A_n nach X. Seien $a \in X$ sowie zu $n \in N \setminus \{e\}$ Abbildungen $V_n : X^n \longrightarrow X$ gegeben. Dann existiert eine eindeutige Funktion $f : N \longrightarrow X$ mit

- $f(e) = a$
- $f(\nu(n)) = V_n(f|_{A_n})$. Hier bezeichnet $f|_{A_n}$ die Einschränkung von f auf A_n gegeben durch $f|_{A_n} : A_n \longrightarrow X$, $x \mapsto f(x)$.

Bemerkung. Das ist eine präzise Fassung von

$$X^n = \text{Menge der Tupel } (f(1), \ldots, f(n))$$

und

$$f(1) = a, \;\; f(n+1) = V_n(f(1), \ldots, f(n))$$

für alle $n \in N$.

Beweis. (Skizze) Wir zeigen zunächst die *Eindeutigkeit* eines solchen f: Seien f und g zwei Funktionen mit den gewünschten Eigenschaften.

Sei $L := \{n \in N : f|_{A_n} = g|_{A_n}\}$. Dann ist L induktiv (einfach) und muss also mit N übereinstimmen. Das liefert die Eindeutigkeit.

Wir kommen nun zur *Existenzaussage.* Sei $\mathcal{A}(n)$ die Aussage:

$\mathcal{A}(n)$ Es existiert eine eindeutige Funktion $f_n : A_n \longrightarrow X$ mit $f_n(e) = a$ und $f_n(\nu(k)) = V_k(f_n|_{A_k})$ für $k \in A_n$ mit $\nu(k) \in A_n$.

Wir zeigen zunächst durch Induktion, daß $\mathcal{A}(n)$ für jedes $n \in N$ wahr ist:
$n = e$: klar. ($f_e(e) = a$.)
$n \Longrightarrow \nu(n)$:
Existenz: Definiere $f_{\nu(n)}$ auf A_n durch f_n und setze es auf $\nu(n)$ als $V_n(f_n)$. Dann hat $f_{\nu(n)}$ die gewünschten Eigenschaften.
Eindeutigkeit: Das ist einfach. (Auf A_n haben wir keine Wahl, da $\mathcal{A}(n)$ wahr ist, und auf $\nu(n)$ ist der Funktionswert nach der Vorschrift festgelegt.)

Nun zeigen wir, daß die f_n, $n \in N$, miteinander verträglich sind:
Seien $k, n \in N$ gegeben. Dann gilt (s.o.) $A_n \subset A_k$ oder $A_k \subset A_n$. Ohne Einschränkung $A_n \subset A_k$. Aufgrund der Eindeutigkeit müssen dann f_n und f_k auf A_n übereinstimmen. Damit kann man dann die f_n zu einem f auf N 'zusammensetzen', indem man definiert

$$f : N \longrightarrow X,\ f(n) := f_n(n).$$

Dieses f stimmt auf jedem A_n mit f_n überein und hat also die gewünschten Eigenschaften. $\square$

Bemerkung. Ähnlich wie man Funktionen rekursiv definieren kann, kann man auch Mengen rekursiv definieren (siehe Übung).

Damit können wir nun die schon angekündigte Eindeutigkeit der 'natürlichen Zahlen' beweisen.

THEOREM. *(Eindeutigkeit der natürlichen Zahlen) Es genügen* (N_1, e_1, ν_1) *und* (N_2, e_2, ν_2) *den Peano Axiomen. Dann gibt es eindeutige Abbildungen* $k : N_1 \longrightarrow N_2$ *und* $l : N_2 \longrightarrow N_1$ *mit*

$$k(e_1) = e_2 \ \textit{und}\ k(\nu_1(n)) = \nu_2(k(n)) \ \textit{für alle}\ n \in N_1$$

bzw.

$$l(e_2) = e_1 \ \textit{und}\ l(\nu_2(n)) = \nu_1(l(n)) \ \textit{für alle}\ n \in N_2.$$

Es gilt $l \circ k = id_{N_1}$ *und* $k \circ l = id_{N_2}$. *Damit sind also* l *und* k *bijektiv.*

Zeichnung. Kommutatives Diagramm.

Beweis. Wir widmen uns zunächst der Abbildung k:

Existenz von k: Das folgt durch rekursive Definition.

Eindeutigkeit von k: Seien k und k' Abbildungen mit der gewünschten Eigenschaft. Sei $L := \{n \in N_1 : k(n) = k'(n)\}$. Dann gilt nach

Definition $e_1 \in L$ und $n \in L$ impliziert $\nu_1(n) \in L$ da

$$k(\nu_1(n)) = \nu_2(k(n)) = \nu_2(k^{'}(n)) = k^{'}(\nu_1(n)).$$

Damit folgt $L = N_1$ aus dem zweiten Peanoaxiom.

Analog können wir Existenz und Eindeutigkeit von l beweisen.

Wir zeigen nun $l \circ k = id_{N_1}$: Sei $S := \{n \in N_1 : l \circ k(n) = n\}$. Dann folgt ähnlich wie eben, daß $S = N_1$.

Die Aussage $k \circ l = id_{N_2}$ lässt sich analog beweisen. □

!!! Die gute Nachricht!!! Das vorangehende Theorem zeigt, daß es (bis auf Umbenennung) nur ein Triple (N, e, ν) gibt, das den Peano Axiomen genügt. Wir bezeichnen dieses eindeutige Tripel ab jetzt als die natürlichen Zahlen und verwenden das Symbol $\mathbb{N}$. Weiterhin schreiben wir dann (meist) 1 statt e und fuer den Nachfolger von 1 schreiben wir 2.

Bemerkung. In gewissen Fällen wird es praktisch sein, noch ein weiteres Element 0 zur Verfügung zu haben und mit $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ zu arbeiten. Auf $\mathbb{N}_0$ zeichnen wir das Element 0 aus und definieren die Nachfolge Abbildung ν_0 auf N wie bisher und $\nu_0(0) = 1$. Dann erfüllt $(\mathbb{N}_0, 0, \nu_0)$ die Peanoaxiome. (Klar!)

Da die Peano Axiome das Tripel eindeutig bestimmen, legen sie auch die Mengen A_n und M_n (bis auf Umbenennung) eindeutig fest. Damit ist folgende Definition moeglich.

Definition (Endliche Menge). *Eine Menge X heißt endlich, wenn sie leer ist oder ein $n \in \mathbb{N}$ existiert und eine bijektive Abbildung $j : A_n = \{1, \ldots, n\} \longrightarrow X$. Dann heißt 0 bzw. n die Mächtigkeit oder Kardinalität oder Anzahl der Elemente von X und man nennt X auch n-elementig. Eine Menge heißt unendlich, wenn sie nicht endlich ist.*

Bemerkungen.

- Die Definition von 'n-elementig' ist sinnvoll, da fuer $n \neq m$ die Mengen A_n und A_m nicht gleichviele Elemente haben. Letzteres kann man wie folgt zeigen: Fuer $n \neq m$ sind A_n und A_m verschieden (da x durch A_x festgelegt ist nach einer obigen Bemerkung) und es gilt $A_n \subset A_m$ oder $A_m \subset A_n$ (vgl. eine obige Folgerung). Damit ist fuer $n \neq m$ dann A_n eine echte Teilmenge von A_m oder umgekehrt. Damit folgt die gewuenschte Aussage.
- Wir werden spaeter sehen, daß es verschieden 'große' unendliche Mengen geben kann.

Wir können nun $\mathbb{N}$ mit einer Addition und einer Multiplikation versehen gemäß:

Addition : $n+m$**:** Für jedes feste $n \in \mathbb{N}$ definieren wir die Abbildung $S_n : \mathbb{N} \longrightarrow \mathbb{N}$ rekursive gemäß

$$S_n(1) = \nu(n) \text{ und } S_n(\nu(m)) := \nu(S_n(m)).$$

Dann liefert $S_n(m)$ eine präzise Fassung des bisher nicht definierten $n+m$ d.h. wir definieren

$$n + m := S_n(m).$$

Damit gilt dann $\nu(n) = n + 1$.

Multiplikation : kn**:** Für jedes feste $n \in \mathbb{N}$ definieren wir die Abbildung $M_n : \mathbb{N} \longrightarrow \mathbb{N}$ rekursiv gemäß

$$M_n(1) = n \text{ und } M_n(\nu(m)) := M_n(m) + n.$$

Wir setzen $kn := M_n(k)$.

Weiterhin setzen wir $0 + n = n = n + 0$ sowie $0n = n0 = 0$ für alle $n \in \mathbb{N}_0$.

Es ist dann möglich zu zeigen, daß Addition und Multiplikation den üblichen Regeln genügen. Wir werden darauf im nächsten Abschnitt sehr genau (auf andere Art) eingehen.

Hier diskutieren wir aber schon kurz eine **Subtraktion:**
Zu jedem $k \in A_n \cup \{0\}$ existiert ein eindeutiges $m \in N_0$ mit $k+m = n$. Wir schreiben dann auch $n - k$ für m.
Bew. Sei L die Menge der $n \in N_0$ für die A_n die behauptete Eigenschaft hat. Dann gehört 0 zu L. (Denn $0 + 0 = 0$ und $0 + n = n \neq 0$ für alle $n \in N$). Weiterhin gehört mit n auch $\nu(n)$ zu L: (Übung. Sei $k \in A_{\nu(n)}$. Zu zeigen Existenz und Eindeutigkeit eines m mit $k + m = \nu(n)$).
Existenz: Falls $k \in A_n$, gibt es m' mit $k + m' = n$ und wir wählen $m := \nu(m')$. Falls $k = \nu(n)$ setzen wir $m := 0$.
Eindeutigkeit: Es gelte $k + m = \nu(n)$. Ist $m = 0$ so folgt $k = \nu(n)$. Damit ist dann $m = 0$ die einzige Lösung (denn $k + n \in M_k$ wie eine einfache Induktion zeigt und M_k enthält k nicht.)
Andernfalls ist m ein Nachfolger, also $m = \nu(l)$. Dann gilt $k + \nu(l) = \nu(n)$, also $\nu(k + l) = \nu(n)$ also $k + l = n$. Wegen $n \in L$ ist dann l eindeutig bestimmt und damit auch $m = \nu(l)$.

Hier geben wir noch einige Anwendungen:

- Die Funktion $\mathbb{N} \longrightarrow \mathbb{N}$, $n \mapsto n!$ (n-Fakultät) wird definiert durch
$$1! = 1 \text{ und } (n+1)! = (n+1)n!.$$
Damit ist sinngemäß $n! = 1 \cdot 2 \cdot ... \cdot n$. Zweckmässig: $0! = 1$.
- Für $a \in \mathbb{N}$ wird die Funktion $\mathbb{N} \longrightarrow \mathbb{N}$, $n \mapsto a^n$ definiert durch
$$a^1 := a, \quad a^{n+1} := a \cdot a^n.$$
Zweckmässig: $a^0 := 1$.

- Die Anzahl der Teilmengen einer n-elementigen Menge ist gerade 2^n. (vgl. Übung)

  Bew. $n = 1$: Es gibt zwei Teilmengen: die leere Menge und gesamte Menge.

  $n \Longrightarrow n+1$: Sei eine Menge M mit $n+1$ Elementen gegeben. Sei p eine Element von M. Dann gibt es genausoviele Teilmengen, die p enthalten, wie Teilmengen die p nicht enthalten (!). Es gibt 2^n Teilmengen von M, die p nicht enthalten (also Teilmengen der n elementigen Menge $M \setminus \{p\}$ sind). Insgesamt gibt es also $2 \times 2^n = 2^{n+1}$ Teilmengen. (Zeichnung: n weisse Kugeln und eine schwarze Kugel.)

Ende der 3. Vorlesung.

Mittels Rekursiver Definition können wir auch $\sum$ und $\prod$ definieren. Das wird später oft nützlich sein. Daher gehen wir nun darauf ein.

Definition von $\prod$:

Ziel: Präzise Version von $\prod_{k=1}^n a_k = a_1 \ldots a_n$.

Sei K eine Menge mit einer Verknüpfung $\cdot$ (Beispiel: $K = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$). Seien $a_n \in K$, $n \in N$, gegeben. Wir definieren dazu den Ausdruck $\prod_{k=1}^n a_k$ rekursiv durch

$$\prod_{k=1}^{1} a_k := a_1 \text{ und } \prod_{k=1}^{n+1} a_k := \left(\prod_{k=1}^{n} a_k\right) a_{n+1}.$$

(Die Funktion $f : N \longrightarrow K$, $f(n) := \prod_{k=1}^n a_k$, wird also durch die Bedingungen $f(1) = a_1$ und $f(n+1) = f(n)a_{n+1} =: V_n(f_{A_n})$ festgelegt.)

Spezialfall: Sind alle $a_k = q \in K$ so setzt man

$$q^n := \prod_{k=1}^{n} q.$$

Dann gilt also

$$q^1 = q, q^{n+1} = qq^n.$$

Definition von $\sum$:

Ziel: Präzise Version von $\sum_{k=1}^n a_k = a_1 + \cdots + a_n$.

Sei K eine Menge mit einer Verknüpfung $+$ (Beispiel: $K = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$). Seien $a_n \in K$, $n \in N$ gegeben.

Wir definieren dazu den Ausdruck $\sum_{k=1}^n a_k$ rekursiv durch

$$\sum_{k=1}^{1} a_k := a_1 \text{ und } \sum_{k=1}^{n+1} a_k := \left(\sum_{k=1}^{n} a_k\right) + a_{n+1}.$$

(Die Funktion $f : N \longrightarrow K$, $f(n) = \sum_{k=1}^n a_k$ wird also durch die Bedingungen $f(1) = a_1$ und $f(n+1) = f(n)+a_{n+1} =: V_n(f_{A_n})$ festgelegt.)

Spezialfall: Sind alle $a_k = q \in K$ so setzt man

$$nq := \sum_{k=1}^{n} q.$$

Dann gilt also

$$1q = q, (n+1)q = q + nq.$$

KAPITEL 2

Die reellen Zahlen

Die reellen Zahlen sind charakterisiert durch das Zusammenspiel von drei Strukturen:

- Körperaxiome ('Arithmetik')
- Anordnungsaxiome ('$\leq$')
- Vollständigkeitsaxiom ('Existenz von Suprema und Infima') (Liefert Existenz von Grenzwerten)

Die natürlichen Zahlen lassen sich als Teilmengen der reellen Zahlen auffassen und erben entsprechend Arithmetik und Anordnung. Darum geht es in diesem Kapitel. Insbesondere werden wir dabei folgende Zeichnung rechtfertigen:

Zeichnung. Linie mit
- 0 und Spiegelung (für Inversion im Körper),
- Positiv- und Negativteil (für Ordnung)
- ohne Lücken (für Vollständigkeit).
sowie
- natürlichen Zahlen.

1. Die Körperstruktur

Die reelen Zahlen mit Multiplikation und Addition sind ein Körper:

Definition. *(Körper) Eine Menge K zusammen mit den Verknüpfungen*

$$K \times K \longrightarrow K, (x, y) \mapsto x + y \ \ (Addition)$$

und

$$K \times K \longrightarrow K, (x, y) \mapsto x \cdot y \ \ (Multiplikation)$$

heißt Körper, wenn folgende Axiome gelten:
Axiome der Addition:

(A1) *Assoziativgesetz:* $x + (y + z) = (x + y) + z$ *für alle* $x, y, z \in K$.
(A2) *Kommutativgesetz:* $x + y = y + x$ *für alle* $x, y \in K$.
(A3) *Existenz der* 0*: Es gibt ein* $0 \in K$ *mit* $x + 0 = x$ *für alle* $x \in K$.
(A4) *Existenz des Negativen: Zu jedem* $x \in K$ *existiert ein* $-x \in K$ *mit* $x + (-x) = 0$.

Axiome der Multiplikation:

(M1) *Assoziativgesetz:* $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ *für alle* $x, y, z \in K$.
(M2) *Kommutativgesetz:* $x \cdot y = y \cdot x$ *für alle* $x, y \in K$.

(M3) *Existenz der* 1*: Es gibt ein* $1 \in K$ *mit* $1 \neq 0$ *und* $x \cdot 1 = x$ *fuer alle* $x \in K$.

(M4) *Existenz des Negativen: Zu jedem* $x \in K$ *mit* $x \neq 0$ *existiert ein* $x^{-1} \in K$ *mit* $xx^{-1} = 1$.

Distributivgesetz:

(D1) *Distributivgesetz:* $x(y+z) = xy + xz$ *für alle* $x, y, z \in K$.

Bemerkung. Man kann diese Definition auch so fassen: $(K, +)$ ist eine kommutative Gruppe mit neutralem Element 0. $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe mit neutralem Element 1. Es gilt das Distributivgesetz $x(y+z) = xy+xz$ für alle $x, y, z \in K$. Wir nennen 0 das *neutrale Element der Addition* und 1 das *neutrale Element der Multiplikation.*

Notation. Wir schreiben $x - y$ statt $x + (-y)$ und x/y statt $x \cdot y^{-1}$ (fuer $y \neq 0$) und xy statt $x \cdot y$.

Beispiel - (Kleinster Körper) $\mathbb{F}_2 = \{0, 1\}$ mit Addition $0 + 0 = 0, 0 + 1 = 1 + 0 = 1, 1 + 1 = 0$ und Multiplikation $0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$ und $00 = 0$ ist ein Körper. Es handelt sich gerade um das 'Rechnen modulo zwei'. Dabei steht 1 für alle ganzen Zahlen, deren Rest bei Division durch gerade 1 ist (ungerade Zahlen) und 0 für alle ganzen Zahlen, deren Rest bei Division durch 2 gerade 0 ist (gerade Zahlen). Die Rechenregeln lassen sich dann verstehen als *gerade* + *gerade* = *gerade*, *gerade* + *ungerad* = *ungerade*..... Allgemeiner liefert Rechnen modulo N einen Körper, falls N eine Primzahl ist (siehe Algebra).

'Beispiele.' $\mathbb{C}$, $\mathbb{Q}$, $\mathbb{R}$.

Proposition. *(Charakteristische Eigenschaften des Inversen) Sei* $(K, +, \cdot)$ *ein Körper. Dann gilt:*
(a) $y - x$ *ist die eindeutige Lösung* z *von* $x + z = y$. *Insbesondere sind das Inverse bzgl. der Addition zu einem* x *und das neutrale Element der Addition eindeutig bestimmt. Gilt* $x + y = 0$ *für* $x, y \in K$ *so ist* $x = -y$ *und* $y = -x$. *Insbesondere ist also* $x = -(-x)$.

(b) x/y *ist die eindeutige Lösung* z *von* $zy = x$. *Insbesondere ist das Inverse bzgl der Multiplikation zu einem* $y \neq 0$ *und das neutrale Element der Multiplikation eindeutig bestimmt. Gilt* $xy = 1$ *für* $x, y \in K$, *so gilt* $x = y^{-1}$ *und* $y = x^{-1}$. *Insbesondere ist dann also* $x = (x^{-1})^{-1}$.

Beweis. (a) *Lösung:* $x+(y-x) = (y-x)+x = y+(-x+x) = y+0 = y$. *Eindeutig:* $x = y + z \Longrightarrow x - y = (y + z) - y = -y + (y + z) = (-y + y) + z = 0 + z = z$.

Ende der 4. Vorlesung

Eindeutigkeit des Inversen der Addition: $-x$ loest $x + (-x) = 0$.

Eindeutigkeit des neutralen Elementes der Addition 0*:* Es löst 0 die Gleichung $x + 0 = x$.

Zu $x+y=0$: Nach Konstruktion gilt $x+(-x)=0$ und $(-y)+y=0$. Aus der Eindeutigkeit der Loesung folgt dann also $y=-x$ und $-y=x$. Damit folgt dann $x=-y=-(-x)$.

(b) ähnlich wie (a).

□

PROPOSITION. *(Rechnen mit* 0 *und* 1*) Sei* $(K,+,\cdot)$ *ein Körper. Dann gilt:*
(a) Es gilt $0=-0$ *und* $1=1^{-1}$.
(b) $0x=0$ *für alle* $x\in K$.
(c) $(-1)x=-x$ *für alle* $x\in K$.
(d) $(-x)(-y)=xy$ *für alle* $x,y\in K$.

Beweis.
(a) Es gilt $0=0+0$ und $1=1\cdot 1$. Damit folgt die Aussage aus der vorigen Proposition.
(b) $0x=(0+0)x=0x+0x$. Nun Addieren von $-0x$ auf beiden Seiten...
(c) Es ist zu zeigen, daß $z:=(-1)x$ das Inverse von x (bzgl. Addition) ist:
$z+x=(-1)x+x=(-1)x+1x=(-1+1)x=0x=0$ (nach (c)).
(d) Übung. □

FOLGERUNG. *(Vertauschen der Inversionen) Sei* $(K,+,\cdot)$ *ein Körper. Dann ist für jedes* $x\neq 0$ *auch* $-x\neq 0$ *und es gilt* $(-x)^{-1}=-x^{-1}$.

Beweis. Für $x\neq 0$ gilt auch $-x\neq 0$ (sonst $x=x+0=x+(-x)=0$ Widerspruch). Weiterhin gilt

$$(-x)(-x^{-1})=(-x)(-1)(x^{-1})=(-1)(-x)x^{-1}=xx^{-1}=1.$$

Damit folgt (s.o.) $(-x)^{-1}=-x^{-1}$. □

Bemerkung. (Uebung) Man kann nun folgende wohlbekannte Aussagen (Bruchrechnung) zeigen:
(a) Fuer $x,y\neq 0$ gilt $(xy)y^{-1}x^{-1}=1$ und insbesondere $xy\neq 0$.
(b) Fuer $y,w,v\neq 0$ gilt

$$\frac{x}{y}\frac{v}{w}=\frac{xv}{yw},\quad \frac{x}{y}+\frac{v}{w}=\frac{xw+yv}{yw},\quad \frac{x}{y}/\frac{u}{v}=\frac{xv}{yu}.$$

Bemerkung. Ist $(K,+,\cdot)$ ein Körper und $x\in K$ mit $x\neq 0$, so liefert rekursive Definition eine eindeutige Abbildung

$J:\mathbb{N}\longrightarrow K$ mit $J(1_{\mathbb{N}})=x$ und $J(n+1_{\mathbb{N}})=J(n)+x$ für alle $n\in\mathbb{N}$.

Wir definieren dann: $n\cdot x:=J(n)$.

Beispiel. Sei $K=\mathbb{F}_2$ und $x=1$. Dann gilt für $J:\mathbb{N}\longrightarrow\mathbb{F}_2$, $n\mapsto n1_{\mathbb{F}_2}$

$$J(n)=0 \text{ falls } n \text{ gerade d.h. } n=2k \text{ für ein } k\in\mathbb{N}$$

und

$$J(n) = 1 \text{ falls } n \text{ ungerade d.h. } n = 2k-1 \text{ für ein } k \in \mathbb{N}.$$

Die folgenden beiden wichtigen Formeln gelten in jedem Körper.

PROPOSITION. *(Geometrische Summenformel) Sei K ein Körper und $x, y \in K$ mit $x \neq y$. Dann gilt für jedes $n \in \mathbb{N}$*

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

Inbesondere gilt für $q \neq 1$ und $n \in \mathbb{N}_0$ die Formel

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

Beweis. Wir rechnen

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

Die zweite Gleichheit folgt durch Vertauschen von x und y. Das 'Inbesondere' folgt mit $x = 1$ und $y = q \neq 1$. □

Bemerkung. Man kann die geometrische Summenformel auch durch Induktion beweisen. (Übung).

PROPOSITION. *(Binomischer Satz) Sei K ein Körper, $n \in \mathbb{N}$ und $x, y \in K$. Dann gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

mit

$$\binom{n}{k} := \text{Anzahl der } k\text{-elementigen Teilmengen einer } n \text{ elementigen Menge.}$$

Diese Zahlen $\binom{n}{k}$ erfüllen die Rekusionsformel

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

Beweis. Die Rekursionsformel folgt direkt: Sei eine $n+1$ elementige Menge gegeben. Sei ein Element p aus dieser Menge fixiert. Dann gibt es $\binom{n}{k-1}$ k-elementige Teilmgenen die p enthalten und $\binom{n}{k}$ k-elementige Teilmengen, die p nicht enthalten. (Zeichnung: n weisse Kugeln und eine schwarze Kugel....)

Nun folgt die Aussage über $(x+y)^n$ durch Induktion:
$n = 1$: Klar.
$n \Longrightarrow (n+1)$: Direkte Rechnung liefert

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

(Dabei folgt die letzte Gleichung aus der Induktionsannahme für n.) Damit können wir weiter rechnen

$$\begin{aligned} \ldots &= \sum_{k=0}^{n}\binom{n}{k}x^{k+1}y^{n-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k+1} \\ (k \to k-1) \quad &= \sum_{k=1}^{n+1}\binom{n}{k-1}x^k y^{n-k+1} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k+1} \\ (\mathit{Rekursion}) \quad &= x^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}x^k y^{n-k+1} + y^{n+1} \\ &= \sum_{k=0}^{n+1}\binom{n+1}{k}x^k y^{n+1-k}. \end{aligned}$$

Damit ist der Beweis abgeschlossen. □

Ende der 5. Vorlesung

Bemerkung. (a) Alternative Deutung: 'Ausmultiplizieren' von

$$(x+y)^n = (x+y)(x+y)\cdots(x+y)$$

und bestimmen, wie oft $x^k y^{n-k}$ vorkommt.
(b) Es gilt $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. (Bew. Induktion und Rekursionsformel. Beachte dabei, daß man zunächst dem Quotienten a/b für natürliche Zahlen einen Sinn geben muss, etwa durch $a/b = c$ genau dann wenn $c \in \mathbb{N}$ die Gleichung $bc = a$ erfüllt.)

2. Die Ordnungsstruktur

Wir kommen nun zu einer weiteren Struktur auf $\mathbb{R}$, der Ordnungsstruktur.

DEFINITION. *Ein Körper K zusammen mit einer ausgezeichneten Menge K^+, den sogenannten positiven Elementen, heißt angeordnet, wenn die folgenden Eigenschaften (Ordnungsaxiome) gelten:*

(O1) $K = K^+ \dot\cup \{0\} \dot\cup \{-x : x \in K^+\}$, *wobei die Vereinigung disjunkt ist.*
(O2) $x, y \in K^+$ *impliziert* $x + y \in K^+$.
(O3) $x, y \in K^+$ *impliziert* $xy \in K^+$.

Die Elemente $x \in K$ mit $-x \in K^+$ heissen dann negativ. Die Elemente aus $K^+ \cup \{0\}$ heissen auch nicht negativ.

Bemerkungen. (a) '$\mathbb{Q}$ und $\mathbb{R}$' sind angeordnet (s.u.)
(b) $\mathbb{F}_2$ lässt sich nicht anordnen. (Denn $1 = -1$.) $\mathbb{C}$ lässt sich nicht anordnen.

Wir ziehen schon einmal eine einfache Folgerung aus (O1):

FOLGERUNG. *In einem angeordneten Koerper gilt fuer $x \neq 0$ entweder $x \in K^+$ oder $-x \in K^+$.*

Beweis. Es muß nach (01) entweder gelten $x \in K^+$ oder $x \in \{-h : h \in K^+\}$. Es ist $x = -h$ fuer $h \in K^+$ aequivalent zu $-x = h \in K^+$. □

Notation. K angeordneter Körper. Wir schreiben
$x > y$ (lies 'x groeßer als y') oder $y < x$ (lies: 'y kleiner als x') falls $x - y \in K^+$.
$x \geq y$ (lies 'x groesser gleich y') oder $y \leq x$ (lies: 'y kleiner gleich x') falls $x - y \in K^+ \cup \{0\}$.

FOLGERUNG. *($\leq$ liefert eine totale Ordnung) Sei K ein angeordneter Körper.*
(a) Sind $x, y, z \in K$ mit $x \leq y$ und $y \leq z$ so gilt $x \leq z$.
(b) Gilt für $x, y \in K$ sowohl $x \leq y$ als auch $y \leq x$ so folgt $x = y$.
(c) Für $x, y \in K$ gilt dann genau eine der drei folgenden Aussagen:

- $x < y$.
- $y < x$.
- $x = y$.

Beweis. (a) Zu zeigen $z - x \in K^+ \cup \{0\}$. Das folgt aus (O2) in folgender Weise:

$$z - x = z + (-y + y) - x = (z - y) + (y - x) \in K^+ \cup \{0\}.$$

(b) Für x, y mit $x \leq y$ nd $y \leq x$ gilt $x - y \in K^+ \cup \{0\}$ und $y - x \in K^+ \cup \{0\}$. Damit folgt

$$x - y \in (K^+ \cup \{0\}) \cap (\{-z : z \in K^+\}) \cup \{0\} = \{0\}.$$

Dabei folgt die letzte Gleichheit aus (O1).
(c) Das ist lediglich eine Umformulierung von (O1). □

Bemerkung. Die vorige Proposition besagt, daß jede Andordnung auch eine Ordnung liefert. Tatsaechlich ist eine Anordnung 'mehr' als eine Ordnung (s.u.).

Wir untersuchen nun wie die Anordnung mit Bilden des Inversen verträglich ist.

Proposition. *(Inversion und Quadrate) Sei K ein angeordneter Körper. Dann gilt:*
(a) $x < y \Longleftrightarrow -x > -y$. *Insbesondere* $x < 0 \Longleftrightarrow -x > 0$.
(b) $x \in K^+ \Longleftrightarrow x^{-1} \in K^+$.
(c) $x^2 \in K^+$ *für alle* $x \neq 0$. *Insbesondere* $1 = 1 \cdot 1 > 0$.

Beweis.
(a) $x < y$ bedeutet gerade $y - x \in K^+$; $-x > -y$ bedeutet gerade $-x + y \in K^+$. Das 'Insbesondere' folgt mit $x = x$ und $y = 0$ unter Beachten von $0 = -0$.

(b) $x \in K^+$, insbesondere $x \neq 0$. Angenommen $x^{-1} \notin K^+$. Dann $-x^{-1} \in K^+$. Damit $-1 = x(-x^{-1}) \in K^+$. Damit $1 = (-1)(-1) \in K^+$. Also $1 \in K^+$ und $(-1) \in K^+$. Widerspruch zu (O1).
(c) folgt aus $x^2 = xx = (-x)(-x)$, da für $x \neq 0$ $x \in K^+$ oder $-x \in K^+$ gilt.

□

Wir kommen nun zu einigen nützlichen Rechenregeln. Diese werden inbesondere zeigen, daß wir Ungleichungen addieren und (unter geeigneten Voraussetzungen) auch multiplizieren duerfen.

Proposition. *(Rechenregeln) Sei K ein angeordneter Körper. Dann gilt:*
(a) $x < y,\ x' \leq y' \Longrightarrow x + x' < y + y'$.
(b) $a < b$ *und* $x > 0 \Longrightarrow ax < bx$
(c) $a < b$ *und* $x < 0 \Longrightarrow ax > bx$.
(d) $0 < a < b$ *und* $0 < x < y \Longrightarrow ax < by$.
(e) $0 < x < y \Longleftrightarrow 0 < y^{-1} < x^{-1}$.

Beweis. (a) ähnlich wie (a): Es gilt nach (O2):

$$(y + y') - (x + x') = (y - x) + (y' - x') \in K^+.$$

(b) $bx - ax = (b - a)x \in K^+$.
(c) $a < b$ bedeutet $b - a \in K^+$. $x < 0$ liefert $-x \in K^+$ nach voriger Proposition. Damit gilt also

$$ax - bx = (a - b)x = (-1)(a - b)(-1)x = (b - a)(-x) \in K^+.$$

(d) $by - ax = by - bx + bx - ax = b(y - x) + x(b - a) \in K^+$.
(e) $\Longrightarrow$: Nach vorangegangener Proposition gilt $x^{-1}, y^{-1} \in K^+$. Damit können wir $0 < x < y$ mit $x^{-1}y^{-1}$ 'Durchmultiplizieren' und erhalten die Behauptung.
Die umgekehrte Richtung folgt dann durch Ersetzen von x durch x^{-1} und y durch y^{-1}. □

Folgerung. $0 \leq x < y,\ k \in \mathbb{N} \Longrightarrow 0 \leq x^k < y^k$.

In allen angeordneten Körper gilt die folgende Bernoulli Ungleichung. Um sie zu formulieren, erinnern wir noch einmal an folgendes: In einem Körper K können wir nx mit $n \in \mathbb{N}$ und $x \in K$ rekursiv definieren durch $1_{\mathbb{N}} x = x$ und $(n + 1_{\mathbb{N}})x := nx + x$.

Proposition. *(Bernoulli Ungleichung) Sei K ein angeordneter Körper. Dann gilt für $x \geq -1$ und $n \in \mathbb{N}$*

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

Beweis. Induktion nach n.
$n = 1$: $1 + x = 1 + x$.
$n \Longrightarrow n + 1$:

$$\begin{array}{rcl} (1+x)^{n+1} & = & (1+x)(1+x)^n \\ (A(n),\ 1+x \geq 0) & \geq & (1+x)(1+nx) \\ & = & 1 + (n+1)x + x(nx) \\ \text{(Kleine Induktion zwischendurch ;-)} & = & 1 + (n+1)x + nx^2 \\ \text{(Kleine Induktion zwischendurch ;-)} & \geq & 1 + (n+1)x. \end{array}$$

Zu den kleinen Induktionen zwischendurch: Fuer jedes x gilt

$$x(nx) = nx^2 \geq 0.$$

Bew. Induktion: $n = 1$: $x(1_{\mathbb{N}} x) = xx = x^2 \geq 0$. Dabei folgt die letzte Ungleichung aus dem oben gezeigten.
$n \Longrightarrow n + 1$: $x((n + 1)x) = x(nx + x) = x(nx) + xx = nx^2 + x^2 = (n + 1)x^2$. (Dabei: erste Gleichung: Rekursive Definition von $(n + 1)x$; zweite Gleichung: Distributivgesetz; dritte Gleichung: Gueltigkeit der Aussage fuer n; vierte Gleichung: Rekursive Definition von $(n + 1)x^2$.) Es gilt $x((n + 1)x = nx^2 + x^2 \geq 0$ aufgrund der Ordnungsaxiome, da es sich um eine Summe von zwei nichtnegativen Termen handelt. (Der erste Term ist nichtnegativ aufgrund der Gueltigkeit der Aussage fuer n; der zweite Term ist nichtnegativ wie oben gezeigt wurde). □

Bemerkungen. (a) Für $x \geq 0$ folgt das natürlich aus dem binomischen Satz. Denn $(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$ hat nur nichtnegative Summanden.
(b) Wie ist die Lage für $-2 \leq x \leq -1$? (Übung)

Ende der 6. Vorlesung

In einem angeordneten Körper können wir den Betrag definieren durch

$$|\cdot| : K \longrightarrow K^+ \cup \{0\}, |x| := \begin{cases} x & : \ x > 0 \\ 0 & : \ x = 0 \\ -x & : \ x < 0. \end{cases}$$

Der Betrag beschreibt so etwas wie eine Länge. Das Bilden des Betrages ist in gewisser Weise mit Addition und Multipliktion verträglich:

PROPOSITION. *(Betrag und Multiplikation) Ist K ein angeordneter Körper, so gilt für alle $x, y, z \in K$*

- $|z| = |-z|$
- $|1/x| = 1/|x|$ *(falls $x \neq 0$).*
- $|xy| = |x||y|$

Beweis. *Es gilt* $|z| = |-z|$. Das folgt leicht durch Fallunterscheidung.
Es gilt $|1/x| = 1/|x|$. Das ist klar für $x > 0$. Für $x < 0$ gilt $-x > 0$ und $x^{-1} < 0$ und damit

$$1/|x| = |x|^{-1} = |-x|^{-1} = (-x)^{-1} \overset{s.o.}{=} -x^{-1} = |x^{-1}| = |1/x|.$$

Es gilt $|xy| = |x||y|$. Gilt $x = 0$ oder $y = 0$, so folgt die Aussage sofort. Die übrigen Fälle folgen einfach durch Fallunterscheidung (4 Fälle). Etwa:
$x > 0, y > 0$: Dann gilt $xy > 0$ und damit $|xy| = xy = |x||y|$.
$x < 0, y > 0$: Dann gilt $-y > 0$. Damit folgt unter Anwendung des schon gezeigten:

$$|xy| = |(-1)xy| = |x(-y)| = |x||-y| = |x||y|.$$

etc. $\square$

PROPOSITION. *(Dreiecksungleichung) Ist K ein angeordneter Körper, so gilt für alle $x, y, z \in K$*

$$|z| = |-z| \ \textit{und} \ -|z| \leq z \leq |z|$$

und

$$|x+y| \leq |x| + |y| \ \ (\textit{Dreiecksungleichung})$$

also insbesondere

$$||x| - |y|| \leq |x - y| \ \ (2.\ \textit{Dreieckungleichung}).$$

Beweis. $-|z| \leq z \leq |z|$: Für $z = 0$ ist die Aussage klar. Für $z > 0$ folgt die Aussage leicht. Für $z < 0$ können wir $-z$ betrachten und erhalten die Aussage.
Es gilt $|x + y| \leq |x| + |y|$*:* Reicht z.z. $x + y \leq |x| + |y|$ und $-(x + y) \leq |x| + |y|$.
Nun gilt aber aufgrund des schon gezeigten: $x, -x \leq |x|$ und $y, -y \leq |y|$. Addieren unter Verwendung der Rechenregeln liefert die Aussage.
Zum 'Insbesondere': Aus $|x| = |x-y+y| \leq |x-y|+|y|$ folgt $|x|-|y| \leq |x - y|$ ähnlich folgt auch $|y| - |x| \leq |y - x|$. $\square$

Die Ordnungstruktur erlaubt es auch von Minima und Maxima einer Menge zu sprechen: Sei M eine Menge in K. Ein $x \in K$ heißt Maximum / Minimum von M, wenn gilt:

- $x \in M$.
- Fuer jedes $y \in M$ gilt $y \leq x$ / $y \geq x$.

Man kann sich leicht (wie?) klarmachen, dass ein solches Maximun / Minimum eindeutig ist (wenn es ueberhaupt existiert). Man schreibt dann $\max M$ bzw. $\min M$ fuer das Maximum bzw. Minimum von M.

Wir beenden diesen Abschnitt, in dem wir die Existenz von Maxima und Minima ueber endliche Mengen zeigen: Ist M eine Menge in K mit einem Element, m, so definieren wir $\max M = m$ und $\min M = m$. Sei nun M eine Menge in K mit $n + 1$ Elementen, also $M = M^{'} \cup \{m\}$ mit einer n elementigen Menge $M^{'}$. Dann definieren wir $\max M$ durch $\max M := m$ falls $m \geq \max M^{'}$ und $\max M := \max M^{'}$ falls $m < \max M^{'}$ und $\min M$ durch $\min M = m$ falls $m \leq \min M^{'}$ und $\min M = \min M^{'}$ falls $m > \min M^{'}$. Man kann dann leicht sehen, daß die so definierten Elemente gerade die charakteristische Eigenschaft von Maximum bzw. Minimum haben.

3. Ordnungsvollständigkeit

Wir kommen nun zur dritten Eigenschaft der reellen Zahlen, der Ordnungsvollständigkeit. Auf dieser Eigenschaft beruhen die Aussagen über Grenzwerte in der Analysis.

Definition. *(Beschränkte Mengen) Sei K ein angeordneter Körper und $M \subset K$ nichtleer.*

(a) Es heißt $S \in K$ eine obere/untere Schranke von M, wenn $m \leq S$ / $m \geq S$ für alle $m \in M$.

(b) Hat M eine obere/untere Schranke, so heißt M nach oben / unten beschränkt. Hat M obere und untere Schranke, so heißt M beschränkt.

Wir fragen nun nach kleinsten oberen / größten unteren Schränken.

Definition. *Sei K ein angeordneter Körper und M eine nach oben /unten beschränkte Menge in K. Eine obere /untere Schranke S von M heißt dann Supremum /Infimum, wenn für jede weitere obere / untere Schrankte S' gilt $S \leq S'$ / $S \geq S'$.*

Bemerkung. (a) Das Supremum/Infimum einer beschräkten Menge muss nicht existieren (so hat zum Beispiel in $\mathbb{Q}$ die Menge $\{x \in \mathbb{Q} : x^2 \leq 2\}$ keine Supremum. siehe Übung.).

(b) Wenn ein Supremum / Infimum existiert, ist es eindeutig:
(Bew. S, S' Suprema von M. Dann gilt $S \leq S'$ und $S' \leq S$ also $S = S'$.)

(c) Gehoert das Supremum / Infimum zu der Menge, so handelt es sich um das Maximum bzw. Minimum.

Notation. Wir schreiben $\sup M$ bzw. $\inf M$ für das Supremum bzw Infimum einer Menge (falls existent).

Proposition. *(Charakterisierung Supremum) Sei K ein angeordneter Körper und Menge eine Menge in K. Das Supremum der Menge M ist dadurch charakterisiert, daß gilt*

- *$m \leq S$ für alle $m \in M$. (S ist obere Schranke)*
- *Für jedes $\varepsilon > 0$ gibt es ein $m \in M$ mit $S - \varepsilon < m$. (Jede kleinere Zahl ist NICHT Schranke d.h. jede Schranke ist mindestens S)*

Für das Infimum gilt entsprechendes (!).

Beweis. Das ist eigentlich nur eine einfache Umformulierung der Definitionen: Es ist S Supremum von M, wenn es eine obere Schrankte von M ist und jede weitere obere Schranke nicht kleiner als S ist. Das bedeutet, daß S Supremum ist, wenn es eine obere Schranke ist und jede kleinere Element nicht obere Schranke ist. □

Damit können wir nun zur dritten Eigenschaft der reellen Zahlen kommen.

DEFINITION. *Ein angeordneter Körper heißt ordnungsvollständig, wenn jede nach oben beschränkte Menge ein Supremum besitzt und jede nach unten beschränkte Menge ein Infimum besitzt.*

Bemerkung. Besitzt in einem angeordneten Körper jede nach oben beschränkte Menge ein Supremum, so besitzt auch jede nach unten beschränkte Menge ein Infimum. (Übung. Nutzt $\inf M = -\sup(-M)$).

4. Die Charakterisierung

THEOREM. *(Charakterisierung von $\mathbb{R}$). Es gibt (bis auf Umbenennung) genau einen angeordneten, ordnungsvollständige Körper. Dieser Körper wird mit $\mathbb{R}$ bezeichnet und die reellen Zahlen genannt.*

Beweis. Es ist Existenz und Eineutigkeit zu zeigen. Wir geben nur eine sehr grobe Skizze:

Existenz. Natürliche Zahlen werden mit Addition und Multiplikation versehen; dann Grothendiek Konstruktion für $(\mathbb{N}, +)$. Das liefert $(\mathbb{Z}, +)$. Tatsächlich kann man auch die Multiplikation fortsetzen in der offensichtlichen Weise. Nun Grothendieck Konstruktion auf $(\mathbb{Z} \setminus \{0\}, \cdot)$. Das liefert $(\mathbb{Q}, \cdot)$. Tatsächlich kann man auch die Addition fortsetzen. Das liefert $(\mathbb{Q}, \cdot, +)$. Nun Vervollständigen.

Eindeutigkeit. Konstruiere Abbildung von $\mathbb{Q}$ in die rationalen Zahlen des Vergleichkörpers. Setze diese Abbildung fort. □

Zeichnung. Linie, 0, positive, negative Zahlen, Spiegelung. Keine Lücken! Ende der 7. Vorlesung

Wir können die natürlichen Zahlen in natürlicher Weise als eine Teilmenge von $\mathbb{R}$ auffasssen.

Proposition. *(Natürliche Zahlen als Teilmenge von $\mathbb{R}$) Sei $J : \mathbb{N} \longrightarrow \mathbb{R}$ die eindeutige Abbildung (s.o.) mit*

$$J(e) = 1_{\mathbb{R}} \text{ und } J(\nu(n)) = J(n) + 1_{\mathbb{R}} \text{ für alle } n \in \mathbb{N}.$$

Dann ist J injektiv. Weiterhin ist $J(\mathbb{N})$ abgeschlossen unter Bildung von Summen und Produkten (d.h. mit $a, b \in J(\mathbb{N})$ gehören auch $a + b$ und ab wieder zu $J(\mathbb{N})$).

Bemerkung. Um die Struktur klarer hervorzuheben verwechselungen zu vermeiden, bezeichnen wir hier das ausgezeichnete Element von $\mathbb{N}$ als e und die Nachfolgeabbildung als ν.

Beweis. Injektivität: Sei

$$L := \{n \in \mathbb{N} : J(n) \neq J(m) \text{ für alle } m \neq n\}.$$

Wir zeigen, daß L induktiv ist:

Vorüberlegung: Es gilt $J(n) > 0$ für alle $n \in \mathbb{N}$.
Bew. Induktion ($J(e) = 1_{\mathbb{R}} > 0$, $J(\nu(n)) = J(n) + 1_{\mathbb{R}} > 0$.)

Es gilt $e \in L$. Sei $m \neq e$. Dann gilt $m = \nu(k)$ für ein $k \in \mathbb{N}$. Damit folgt

$$J(m) = J(\nu(k)) = J(k) + 1_{\mathbb{R}} = J(k) + J(e) > J(e).$$

Dabei verwenden wir im letzten Schritt die Vorüberlegung. Mit $J(m) > J(1)$ folgt $J(m) \neq J(1)$.

$n \in L$ impliziert $\nu(n) \in L$. Sei $m \neq \nu(n)$.
Gilt $m = e$, so gilt nach dem schon bewiesenen $J(m) \neq J(\nu(n))$ (da $\nu(n) \neq e$).
Gilt $m \neq e$, so gilt $m = \nu(k)$. Wegen $m \neq \nu(n)$ folgt $k \neq n$. Aufgrund von $n \in L$ folgt dann also $J(n) \neq J(k)$. Unter Nutzen dieser Beziehung koennen wir dann rechnen

$$J(m) = J(\nu(k)) = J(k) + 1_{\mathbb{R}} \neq J(n) + 1_{\mathbb{R}} = J(\nu(n)).$$

Das liefert die Behauptung.

Abgeschlossenheit unter Addition. Sei

$$L := \{n \in \mathbb{N} : J(n) + J(m) \in J(\mathbb{N}) \text{ für alle } m \in N\}.$$

Dann gilt $e \in L$ da $J(e)+J(m) = 1_{\mathbb{R}}+J(m) = J(m)+1_{\mathbb{R}} = J(\nu(m)) \in J(\mathbb{N})$. Weiterhin gilt $n \in L \Longrightarrow \nu(n) \in L$, denn

$$J(\nu(n)) + J(m) = J(n) + 1_{\mathbb{R}} + J(m) = J(n) + J(\nu(m)) \in J(\mathbb{N}),$$

wobei $n \in L$ für die letzten Schritt genutzt wurde.

Abgeschlossenheit unter Multiplikation. Analog. Sei

$$L := \{n \in \mathbb{N} : J(n)J(m) \in J(\mathbb{N}) \text{ für alle } m \in N\}.$$

Dann gilt $e \in L$ da $J(e)J(m) = 1_{\mathbb{R}}J(m) = J(m) \in J(\mathbb{N})$. Weiterhin gilt $n \in L \Longrightarrow \nu(n) \in L$, denn

$$J(\nu(n))J(m) = (J(n)+1_{\mathbb{R}})J(m) = J(n)J(m)+1_{\mathbb{R}}J(m) = J(n)J(m)+J(m) \in J(\mathbb{N}),$$

wobei Abgeschlossenheit unter Addition im letzten Schritt genutzt wurde. □

Die vorangegangene Proposition bietet die Möglichkeit auf den natürlichen Zahlen eine Multiplikation und eine Addition einzuführen gemäß

$$n + m := J^{-1}(J(n) + J(m)),\ nm := J^{-1}(J(n)J(m)).$$

Diese Definition ist gerade so gemacht, dass J Addition und Multiplikation erhaelt, in dem Sinne, dass folgendes gilt:

$$J(n + m) = J(n) + J(m) \text{ und } J(nm) = J(n)J(m)$$

fuer alle $n, m \in \mathbb{N}$.
Es gilt dann (Übung. Nutzt Injektivität von J.):

$n + 1 = \nu(n)$ sowie $n + \nu(m) = \nu(n + m)$ für alle $m \in \mathbb{N}$.
$1n = n$ sowie $\nu(k)n = kn + n$ für alle $k \in \mathbb{N}$.

Damit handelt es sich genau um die in einer Bemerkung des letzten Kapitel schon einmal kurz angedeutete Addition und Multiplikation. Aus den entsprechenden Eigenschaften der Addition und Multiplikation auf $\mathbb{R}$ folgen sofort Assoziativität, Kommutativität der Addition und Multiplikation auf $\mathbb{N}$ sowie das Distributivgesetz

$$(n + m)k = nk + mk.$$

Wichtig! Wir werden im folgenden immer $\mathbb{N}$ als mit dieser Multiplikation und Addition ausgestattet voraussetzen und (oft) nicht zwischen $\mathbb{N}$ und $J(\mathbb{N})$ unterscheiden.

Neben den natürlichen Zahlen bilden noch die ganzen Zahlen

$$\mathbb{Z} := \{n : n \in \mathbb{N}\} \cup \{0\} \cup \{-n : n \in \mathbb{N}\}$$

und die rationalen Zahlen

$$\mathbb{Q} := \{\frac{n}{m} : n, m \in \mathbb{Z}, m \neq 0\}$$

wichtige Teilmengen der reellen Zahlen. Natuerliche, ganze und rationale Zahlen sind jeweils abgeschlossen unter Addition und Multiplikation. Die wesentliche zusaetzliche Eigenschaft der ganzen Zahlen (im Vergleich zu den natuerlichen Zahlen) ist die Abgeschlossenheit unter der Bildung von Differenzen. Die wesentliche zusaetzliche Eigenschaft der rationalen Zahlen (im Vergleich zu den ganzen Zahlen) ist die Abgeschlossenheit unter Bildung von Quotienten.

Bemerkung. Schaut man die obigen Beweise an, so stellt man fest, dass die Ordnungsvollstaendigkeit nicht verwendet wurde. Man kann also die natuerlichen Zahlen auf diese Weise in jeden angeordneten Koerper abbilden und entsprechend in jedem angeordneten Koerper die ganzen Zahlen und die rationalen Zahlen wiederfinden.

Gute Nachricht. Ab jetzt 'dürfen' wir in $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ rechnen, wie wir es gewohnt sind. Dann wir haben die entsprechenden Objekte und Rechenregeln eingeführt bzw. bewiesen.

Nach Hause nehmen: $\mathbb{R}$ charakterisiert durch Zusammenspiel von drei Strukuren: Körper, Anordnung, Ordnunsvollständigkeit. Die natürlichen Zahlen, die ganzen Zahlen und die rationalen Zahlen bilden Teilmengen von $\mathbb{R}$ (die unter gewissen Operationen abgeschlossen sind).

Wir zeigen nun die Existenz von Wurzeln nichtnegativer reeller Zahlen. Diese Existenz beruht wesentlich auf der Ordnungsvollständigkeit. Sie gilt nicht im Körper der rationalen Zahlen (s.o.).

THEOREM. *(Existenz k-ter Wurzeln) Sei $k \in \mathbb{N}$. Dann existiert für jedes $x \geq 0$ ein eindeutiges $y \geq 0$ mit $y^k = x$. Man definiert $\sqrt[k]{x} := y$. Es gilt $\sqrt[k]{x_1} < \sqrt[k]{x_2}$ falls $x_1 < x_2$.*

Notation. Man schreibt auch $x^{1/k}$ fuer $\sqrt[k]{x}$.
Beweis. Der Fall $x = 0$ ist klar. Wir betrachten nur noch $x > 0$.

Eindeutigkeit: Sei $y^k = \widetilde{y}^k$. Ist $y \neq \widetilde{y}$, so können wir ohne Einschränktung annehmen $y < \widetilde{y}$. Das führt auf $y^k < \widetilde{y}^k$. Widerspruch.

Existenz: Sei $M := \{z \geq 0 : z^k \leq x\}$. Dann ist M beschränkt (Falls $x \leq 1$ ist 1 eine Schranke. Falls $x > 1$ ist x eine Schranke.) Außerdem ist M nichtleer ($0 \in M$). Damit hat M ein Supremum S. Wir zeigen $S^k = x$, indem wir $S^k < x$ und $S^k > x$ zum Widerspruch führen.

Angenommen $S^k < x$: Wir zeigen, daß dann auch $(S + \varepsilon)^k < x$ für genügend kleine $\varepsilon > 0$. Widerspruch zu S obere Schranke.

Hier sind die Details: $D := x - S^k > 0$. Sei nun $\varepsilon > 0$ mit $\varepsilon < 1$ und

$$\varepsilon < \frac{D}{\sum_{l=0}^{k-1} S^l}$$

gegeben. Dann gilt also

$$0 < \sum_{l=0}^{k-1} \binom{k}{l} S^l (\varepsilon)^{k-l} \leq \left(\sum_{l=0}^{k-1} S^l \right) \varepsilon < x - S^k = D.$$

Damit folgt

$$\begin{aligned} (S + \varepsilon)^k &= \sum_{l=0}^{k} \binom{k}{l} S^l \varepsilon^{k-l} \\ &= S^k + \sum_{l=0}^{k-1} \binom{k}{l} S^l \varepsilon^{k-l} \\ &< S^k + D \\ &= x. \end{aligned}$$

Angenommen $S^k > x$: Wir zeigen ähnlich wie im ersten Fall, daß dann auch $(S-\varepsilon)^k > x$ für alle genügend kleinen $\varepsilon > 0$. Dann ist aber, wie man leicht sieht, $S-\varepsilon$ eine obere Schranke von M. Offenbar ist aber $S-\varepsilon$ kleiner als S. Widerspruch: S kleinste obere Schranke.

Monotonie: Sei $x < y$. Nach der gezeigten Eindeutigkeit ist $\sqrt[k]{x} \neq \sqrt[k]{y}$. Wäre $\sqrt[k]{x} > \sqrt[k]{y}$, so folgte $x = ()^k > ()^k = y$. Widerspruch. □

FOLGERUNG. *Fuer ein beliebiges* $a \in \mathbb{R}$ *gilt* $|a| = \sqrt{2}a^2$

Beweis. Nach Definition des Betrages gilt $|a| \geq 0$ und, wie die Fallunterscheidung $a \geq 0$ bzw. $a < 0$ zeigt, $|a|^2 = a^2$. Damit folgt die gewuenschte Aussage aus der Eindeutigkeit der Wurzel. □

Ende der 8. Vorlesung

Für rationale Zahlen $a = m/n$ mit $m, n \in \mathbb{N}$ und $x \geq 0$ kann man dann definieren

$$x^a := (x^m)^{\frac{1}{n}} = (x^{\frac{1}{n}})^m.$$

Das ist wohldefiniert i.e. es gilt die zweite Gleichung und es hängt nicht von der Darstellung der rationalen Zahl ab. Für reelles $s > 0$ und $x \geq 0$ definiert man dann

$$x^s := \sup\{x^q : q \in \mathbb{Q} : 0 < q \leq s\}.$$

Auch das ist wohldefiniert i.e. stimmt für rationale s mit der schon gegebenen Definition überein. Schliesslich definiert man für $a > 0$ und $x > 0$ noch

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

Dann kann man folgende Rechenregeln zeigen:

$$x^b x^c = x^{b+c}$$
$$(x^b)^c = x^{bc}$$
$$x^c y^c = (xy)^c.$$

für $x, y > 0$ und b, c reell. Wir werden diese Definitionen und Rechenregeln später als Nebenprodukt erhalten. Darum geben wir hier keine weiteren Details.

Bemerkung. Der Ausdruck 0^0 stellt ein Problem dar:
$a^0 = 1$ für alle $a > 0$. Das suggeriert $0^0 = 1$.
$0^s = 0$ für alle $s > 0$. Das suggeriert $0^0 = 0$.
Daher muss man diesen Fall getrennt und kontextabhängig behandeln.

KAPITEL 3

Archimedisches Axiom und Intervallschachtelungsprinzip

Die Ordnungsvollständigkeit von $\mathbb{R}$ ist von entscheidender Bedeutung für alle weiteren Untersuchungen. Sie kann in zwei Aspekte 'zerlegt' werden, nämlich Gültigkeit des Archimedischen Axiom und Konvergenz gewisser Folgen. Eine Möglichkeit, Konvergenz von Folgen zu fassen, liefert das Intervallschachtelungsprinzip. Das behandeln wir in diesem Abschnitt. Ausführliche weitere Untersuchungen zu Konvergenz von Folgen finden sich im kommenden Kapitel.

1. Das Archimedische Axiom

In diesem Abschnitt lernen wir noch eine weitere Eigenschaft der reellen Zahlen kennen, die in unserem Zugang eine Folgerung ist. Das Archimedische Axiom liefert insbesondere Existenz von Nullfolgen und Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$.

Wie schon (mehrfach) diskutiert, kann man die natuerlichen Zahlen $\mathbb{N}$ in jedem angeordneten Koerper K via

$$\mathbb{N} \longrightarrow K, n \mapsto J(n) = n1$$

mit $J(e) = 1$ und $J(\nu(n)) = J(n) + 1$, wiederfinden. Wir schreiben dann im folgenden (meist) n statt $n1$ und auch n/m statt $n1/m1$.

Lemma (Charakterisierung Archimedisches Axiom). *Sei K ein angeordneter Körper. Dann sind äquivalent:*

(i) *Für alle $x > 0$ existiert ein $m \in \mathbb{N}$ mit $x < m$.* ***Zeichnung***
(ii) *Für alle $\epsilon > 0$ existiert ein $m \in \mathbb{N}$ mit $\frac{1}{m} < \epsilon$.* ***Zeichnung***
(iii) *Für alle $x, y > 0$ existiert ein $m \in \mathbb{N}$ mit $y < mx$ oder, äquivalent, $y/m < x$.*

Beweis. (iii)$\Longrightarrow$ (ii): Das folgt sofort mit $x = \epsilon$ und $y = 1$.

(ii)$\Longrightarrow$ (i): (Nach (ii) angewendet auf $\epsilon = 1/x$) existiert ein $m \in \mathbb{N}$ mit $1/m < 1/x$. Damit folgt dann durch Bilden des Kehrwertes $x < m$.

(i)$\Longrightarrow$ (iii): Wähle nach (i) ein $m \in \mathbb{N}$ mit $1/x < m$ also $1/m < x$. Wähle außerde, wieder nach (i), ein $l \in \mathbb{N}$ mit $y/l < 1$. Dann gilt für $k = lm \in \mathbb{N}$ also

$$\frac{y}{lm} = \frac{1}{m}\frac{y}{l} < \frac{1}{m}1 < x.$$

□

DEFINITION. *(Archimedisches Axiom) Ein angeordneter Körper K erfüllt das Archimedische Axiom, wenn eine der äquivalenten Eigenschaften des vorigen Lemma gilt.*

FOLGERUNG. *(Dichtheit von $\mathbb{Q}$) Sei K ein angeordneter Körper, der das Archimedische Axiom erfüllt. Dann gibt es zu $x, y \geq 0$ mit $x < y$ Elemente $n, m \in \mathbb{N}$ mit $x < n/m < y$. Insbesondere existiert zu jedem $x \geq 0$ und $\varepsilon > 0$ Elemente $n, m \in \mathbb{N}$ mit $|x - n1/m1| < \epsilon$. Für $x \leq 0$ gilt entsprechendes, wenn man m/n durch $-m/n$ ersetzt.*

Beweis. Sei $\delta := y - x > 0$. Dann existiert nach dem Archimedischen Axiom ein $m \in \mathbb{N}$ mit $1 < m\delta = my - mx$.

Beh. Es gibt ein $n \in \mathbb{N}$ mit $mx < n < my$.
Bew. Angenommen, die Behauptung gilt nicht. Dann gilt fuer die Menge $L := \{k \in \mathbb{N}_0 : k1 \leq mx\}$ die Gleichheit

$$L = \mathbb{N}_0.$$

(Denn $0 \in L$: klar. $k \in L \Longrightarrow k + 1 \in L$: Sonst $k \leq mx$ und $(k+1)1 > mx$. Dann muss aber, wegen $1 \leq my - mx$ gelten $mx < (k+1)1 < my$ Widerspruch). Das ist ein Widerspruch zum Archimedischen Axiom.

Sind n, m wie in der Behauptung, so folgt nach Division durch m also $x < n/m < y$. □

Bemerkung. In einem angeordneten Körper K, der das Archimedische Axiom erfüllt, gilt dann also für jedes $s \in K$ mit $s > 0$

$$s = \sup\{\frac{n}{m} : n, m \in \mathbb{N}, \frac{n}{m} < s\} = \inf\{\frac{n}{m} : n, m \in \mathbb{N}, \frac{n}{m} > s\}$$

und entsprechend für negative s.

FOLGERUNG. *Sei K ein angeordneter Körper, der das Archimedische Axiom erfüllt.*
(a) Sei $a > 1$. Dann existiert zu jedem $C \in K$ ein $n \in \mathbb{N}$ mit $a^n > C$.
(b) Sei $0 < a < 1$. Dann exisiert zu jedem $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $0 < a^n < \epsilon$.

Beweis. (a) Das folgt aus der Bernoulli Ungleichung: $a = 1 + \delta$ mit $\delta > 0$. Also

$$a^n = (1 + \delta)^n \geq 1 + n\delta > C.$$

Hier wird im letzten Schritt das Archimedische Axiom verwendet.
(b) Das folgt aus (a) durch Bilden des Kehrwertes. □

THEOREM. *(Archimedisches Axiom) In $\mathbb{R}$ gilt das archimedische Axiom.*

Beweis. Zu zeigen: Sind $x, y > 0$ so existiert ein $n \in \mathbb{N}$ mit $ny > x$. Wir nehmen an, daß die Aussage nicht gilt. Dann ist die Menge

$$M := \{ny : n \in \mathbb{N}\}$$

also nach oben beschränkt (durch x) und besitzt aufgrund der Ordnungsvollständigkeit ein Supremum S. Da S eine obere Schranke fuer M ist gilt dann also

$$(n+1)x \leq S \text{ fuer alle } n \in \mathbb{N}.$$

Damit folgt dann sofort

$$nx \leq S - x \text{ fuer alle } n \in \mathbb{N}.$$

Damit ist dann $S - x$ ist obere Schranke von M. Widerspruch. □

Bemerkungen.

- Auch in $\mathbb{Q}$ gilt das Archimedische Axiom (Warum? $n/m \leq n$).
- Nicht in jedem angeordneten Körper gilt das Archimedische Axiom ($---$ > Nichtstandard Analysis).

2. Intervallschachtelungsprinzip

In diesem Abschnitt lernen wir eine weitere Konsequenz der Ordnungsvollständigkeit kennen.

Zunächst einige Bezeichnungen. Sei K ein angeordneter Körper. Dann definiert man für $a \leq b$ die Intervalle:

$[a, b] := \{x \in K : a \leq x \leq b\}$ abgeschlossenes Intervall.
$(a, b) := \{x \in K : a < x < b)$ offenes Intervall (kann leer sein)
$(a, b] := \{x \in K : a < x \leq b\}$ nach links halboffenes Intervall.
$[a, b) := \{x \in K : a \leq x < x\}$ nach rechts halboffenes Intervall.

Es heissen dann a, b die Randpunkte des Intervalles und $|I| := b - a$ die Länge des Intervalles.

Idee zur Intervallschachtelung: Eine geschachtelte Folge von Intervallen, die sich zusammenziehen. **Zeichnung**.

DEFINITION. *Sei K ein angeordneter Körper. Eine Familie I_n, $n \in \mathbb{N}$, von Intervallen in K heißt Intervallschachtelung, wenn gilt:*

- *Jedes I_n ist abgeschlossen.*
- *$I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$. ('Geschachtelt')*
- *$|I_n| \to 0$, $n \to \infty$ (d.h. für alle $\epsilon > 0$ existiert ein $n(\epsilon) \in \mathbb{N}$ mit $|I_n| < \epsilon$ für alle $n \geq n_\epsilon$). ('Zusammenziehen')*

DEFINITION. *(Intervallschachtelungsprinzip) Sei K ein angeordneter Körper. Dann erfüllt K das Intervallschachtelungsprinzip, wenn es zu jeder Intervallschachtelung I_n, $n \in \mathbb{N}$, einen Punkt $x \in K$ gibt mit $x \in I_n$ für alle $n \in \mathbb{N}$.*

Bemerkung. Ist I_n, $n \in \mathbb{N}$, eine Intervallschachtelung, so kann es höchstens einen Punkt geben, der zu allen I_n gehört. Ein solcher Punkt ist also eindeutig.
(Bew. Seien x und y zwei solcher Punkte, so gilt $|x - y| \leq |I_n|$ für alle $n \in \mathbb{N}$ (da $x, y \in I_n$). Wegen $|I_n| \to 0$ gilt dann $|x - y| \leq \varepsilon$ für alle $\varepsilon > 0$. Damit folgt $|x - y| = 0$.)

Theorem. *In $\mathbb{R}$ gilt das Intervallschachtelungsprinzip.*

Beweis. Es bilden die Intervalle $I_n = [a_n, b_n]$ eine Intervallschachtelung. Zu zeigen: Es gibt ein $x \in \mathbb{R}$ mit $x \in I_n$ für alle $n \in \mathbb{N}$.

Es gilt (Induktion) $I_m \subset I_n$ für alle $m \geq n$. Damit folgt

$$a_m \leq b_n \quad (*)$$

für alle $n, m \in \mathbb{N}$. (Fallunterscheidung $n \leq m : a_m \leq b_m \leq b_n$ und $m < n$: $a_m \leq a_n \leq b_n$. Zeichnung).

Ende der 9. Vorlesung

Damit ist also die Menge

$$M := \{a_m : m \in \mathbb{N}\}$$

beschränkt (zum Beispiel durch b_1). Aufgrund der Ordnungsvollständigkeit existiert dann also

$$x := \sup M.$$

Da x eine obere Schranke von M ist gilt

$$a_m \leq x \quad (**)$$

für alle $m \in \mathbb{N}$. Weiterhin ist aufgrund von $(*)$ aber jedes b_n, $n \in \mathbb{N}$, eine obere Schranke von M und es gilt dann (aufgrund der Supremunseigenschaft) also

$$x \leq b_n \quad (***)$$

für jedes $n \in \mathbb{N}$. Mit $(**)$ und $(***)$ folgt $x \in I_n$ für alle $n \in \mathbb{N}$. $\square$

Bemerkung. (Uebung) Der obige Schluss nutzt nicht, daß sich die Intervalle zusammenziehen. Er funktioniert für jede Folge von ineinenander enthaltenenen abgeschlossenen Intervallen. Genauer gilt: Ist I_n, $n \in \mathbb{N}$, eine Folge von abgeschlossenen Intervallen $I_n = [a_n, b_n]$ in $\mathbb{R}$ mit $I_{n+1} \subset I_n$ für jedes $n \in \mathbb{N}$, so gilt $\sup a_n \leq \inf b_n$ und

$$S := \bigcap_n I_n = [\sup a_n, \inf b_n].$$

Insbesondere ist S also nichtleer und ein abgeschlossenes Intervall. Wenn sich die Intervalle zusammenziehen (also eine Intervallschachtelung bilden), so besteht S nur aus einem Punkt. Ein entsprechende Aussage gilt im Allgemeinen nicht, wenn die Intervalle nicht abgeschlossen sind.

3. Eine Äquivalenz

THEOREM. *Sei K ein angeordneter Körper. Dann sind äquivalent:*

(i) *Es ist K ordnungsvollständig (d.h. $K = \mathbb{R}$).*

(ii) *Für K gilt das Intervallschachtelungsprinzip und das Archimedische Axiom.*

Beweis. (i)$\Longrightarrow$ (ii): Die entsprechende Aussagen wurden in den beiden vorigen Abschnitten gezeigt.

(ii)$\Longrightarrow$ (i): Sei M eine nach oben beschränkte Menge in K. Zu zeigen: M hat ein Supremum.

Sei C eine obere Schranke von M. Sei $u \in K$ keine obere Schranke von M z.b. $u = y - 1$ für ein $y \in M$. Wir konstruieren induktiv Intervalle $I_n = [a_n, b_n]$ mit

- Alle b_n sind obere Schranken von M.
- Alle a_n sind keine oberen Schraken von M.
- $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$.
- $|I_{n+1}| = \frac{1}{2}|I_n|$.

Dann bilden die I_n, $n \in N$, eine Intervallschachtelung (die letzte Eigenschaft liefert nach dem Archimedischen Axiom, daß $|I_n| = \frac{1}{2^n} \to 0$). Für den gemeinsamen Punkt x aller I_n (der nach Voraussetzung existiert) gilt dann

- x ist obere Schranke (sonst $z \in M$ mit $x < z$ Widerspruch zu b_n beliebig nahe an x für große n und b_n obere Schranke. Zeichnung)
- Es gibt keine kleinere obere Schranke als x (sonst $m \leq z < x$, für alle $m \in M$. Widerspruch zu a_n beliebig nahe an x für große n und a_n keine obere Schranke. Zeichnung)

Nun zur Konstruktion: Wir setzen $I_1 := [u, C]$. Seien $I_1, \ldots, I_n$ wie oben schon konstruiert und $I_n = [a_n, b_n]$. Sei

$$m := (a_n + b_n)/2$$

der Mittelpunkt von I_n. Wir unterscheiden zwei Fälle (Zeichnung):

Fall 1: m ist obere Schranke von M. Wir setzen $I_{n+1} := [a_n, m]$. Zeichnung.

Fall 2: m ist keine obere Schranke von M. Wir setzen $I_{n+1} := [m, b_n]$. Dann hat I_{n+1} die gewünschten Eigenschaften.
Zeichnung 'konvergierende Intervalle'. $\square$

KAPITEL 4

Konvergenz von Folgen in $\mathbb{R}$

In diesem Kapitel lernen wir das zentrale Konzept der Analysis kennen, nämlich das Konzept der Konvergenz bzw. des Grenzwertes. Es ist grundlegend für alle weiteren Untersuchungen und (in gewisser Weise) das schwierigste Konzept der Analysis.

1. Definitionen und Rechenregeln

DEFINITION. *(Folge) Sei X eine Menge. Eine Abbildung $x : \mathbb{N} \longrightarrow X$ heißt Folge (in X).*

Notation. (x_n) oder $(x_n)_{n\in\mathbb{N}}$ oder $(x_n)_n$.

Beispiele.

- Sei $c \in \mathbb{R}$ beliebig und $x : \mathbb{N} \longrightarrow \mathbb{R}$, $n \mapsto c$, also $x_n = c$ für alle n. Dann heißt (x_n) die konstante Folge mit Wert c.
- $x : \mathbb{N} \longrightarrow \mathbb{R}$, $n \mapsto (-1)^n$, also $x_n = (-1)^n$.
- $x : \mathbb{N} \longrightarrow \mathbb{R}$, $n \mapsto \frac{1}{n}$, also $x_n = \frac{1}{n}$.

Wir kommen nun zu einem zentralen Begriff der Analysis, dem Begriff der **Konvergenz**.

Idee. Die Folge (x_n) konvergiert gegen den Wert x, wenn für alle genügend großen n die Zahl x_n der Zahl x beliebig nahe ist.

Es wird nun darum gehen, diese Idee und insbesondere 'genuegend gross' und 'beliebig nahe' praezise zu fassen. Das verlangt Arbeit, da es um das Verhalten der Folge im Unendlichen geht.

DEFINITION. *(Konvergenz) Die Folge (x_n) in $\mathbb{R}$ konvergiert gegen x, wenn zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ existiert, sodaß für alle $n \geq n_\varepsilon$ gilt $|x_n - x| < \varepsilon$. Dann heißt x Grenzwert der Folge (x_n). Eine Folge, die nicht gegen ein x konvergiert heißt divergent.*

Zeichnung ε- Falle

Zeichnung. ε- Schlauch

Wichtig. Eine Folge (x_n) kann nicht gegen zwei verschiedene Grenzwerte konvergieren (d.h. der Grenwert ist eindeutig, wenn er existiert).

Bew: Es konvergiere (x_n) gegen x und gegen y. Sei $\varepsilon > 0$. Dann gibt es also n_x mit $|x - x_n| < \varepsilon$ für $n \geq n_x$, und es gibt n_y mit $|y - x_n| < \varepsilon$ für $n \geq n_y$. Mit $n \geq n_x, n_y$ gilt dann also

$$|x - y| = |x - x_n + x_n - y| \leq |x - x_n| + |x_n - y| < 2\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt $|x - y| = 0$, also $x = y$.

Aufgrund der Eindeutigkeit kann man von **dem Grenzwert** einer Folge sprechen (falls existent).

Notation. Konvergiert (x_n) gegen x, so schreibt man auch

$$x = \lim_{n\to\infty} x_n \text{ oder } x_n \longrightarrow x, n \longrightarrow \infty$$

und nennt x den Grenzwert der Folge (x_n).

Bemerkung. Diese und ähnliche Definitionen lassen sich mit sogenannten Quantoren ausdrücken. Wir werden in dieser Vorlesung kaum Quantoren benutzen (aber die zugrundeliegenden Zusammenhaenge natürlich ständig verwenden). In Quantoren lautet die Definition von Konvergenz einer Folge:

$$\forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n \geq n_\varepsilon \; |x - x_n| < \varepsilon.$$

Hier: '$\forall \diamond$' steht für 'Für alle $\diamond$ gilt:'

'$\exists \clubsuit$' steht für 'Es existiert $\clubsuit$ mit der Eigenschaft, daß / sodaß...'. Damit ergibt sich auch, daß Quantoren immer an den Anfang der Aussage gestellt werden müssen.

Ende der 10. Vorlesung

Bemerkung. (Konvergenz entscheidet sich ganz weit draussen): $x_n \to x$. Sei $N > 0$ und (y_n) Folge mit $x_n = y_n$ für $n \geq N$. Dann gilt $y_n \to x$. Bew. Sei $\varepsilon > 0$. Es existiert ein $n_\varepsilon \in \mathbb{N}$ mit $|x_n - x| < \varepsilon$ für $n \geq n_\varepsilon$. Für $n \geq n_\varepsilon, N$ gilt also

$$|y_n - x| = |x_n - x| < \varepsilon.$$

Definition. *Eine Folge (x_n) in $\mathbb{R}$ mit $x_n \to 0$ heisst Nullfolge.*

Proposition. *Ist (x_n) eine Nullfolge in $\mathbb{R}$ und $k \in \mathbb{N}$ beliebig, so sind auch (x_n^k) und $(\sqrt[k]{|x_n|})$ Nullfolgen.*

Beweis. Wir betrachten zunaechst den Fall (x_n^k): Sei $\varepsilon > 0$ beliebig. Da (x_n) eine Nullfolge ist, existiert ein $n_\varepsilon \in \mathbb{N}$ mit $|x_n| = |x_n - 0| < \sqrt[k]{\varepsilon}$ fuer alle $n \geq n_\varepsilon$. Damit gilt dann fuer alle $n \geq \varepsilon$ auch

$$|x_n^k - 0| = |x_n^k| = |x_n|^k < \left(\sqrt[k]{\varepsilon}\right)^k = \varepsilon.$$

Wir betrachten nun den Fall $(\sqrt[k]{|x_n|}$: Sei $\varepsilon > 0$ beliebig. Da (x_n) eine Nullfolge ist, existiert ein $n_\varepsilon \in \mathbb{N}$ mit $|x_n| = |x_n - 0| < \varepsilon^k$ fuer alle $n \geq n_\varepsilon$. Damit gilt dann fuer alle $n \geq \varepsilon$ auch

$$|\sqrt[k]{|x_n|} - 0| = \sqrt[k]{|x_n|} < \sqrt[k]{\varepsilon^k} = \varepsilon.$$

Das beendet den Beweis. $\square$

Drei Beispiele.

- Sei $c \in \mathbb{R}$. Dann ist die konstante Folge $x : \mathbb{N} \longrightarrow \mathbb{R}$, $x_n = c$, konvergent gegen c.
  Bew...
- Sei $x : \mathbb{N} \longrightarrow \mathbb{R}$, $x_n = \frac{1}{n}$. Dann konvergiert x_n geben 0.
  Bew. Das folgt aus dem Archimedischen Axiom.
  In gewisser Weise ist dies die einzige explizite konvergente Folge, die wir kennen.
- Sei $x : \mathbb{N} \longrightarrow \mathbb{R}$, $x_n = (-1)^n$ d.h. $x_n = -1$ für ungerade n und $x_n = 1$ für gerade n. Dann ist (x_n) nicht konvergent.
  Bew. Wir beweisen ein allgemeines Kriterium.
  Beh. Ist (x_n) eine konvergente Folge in $\mathbb{R}$, so ist die Folge $y_n := x_{n+1} - x_n$ eine Nullfolge (d.h. konvergent gegen 0).
  Bew. ...
  Mit diesem allgemeinen Kriterium sieht man sofort, daß $x_n = (-1)^n$ nicht konvergiert, da y_n immer den Betrag 2 hat.

Wir geben jetzt noch eine leichte Umformulierung der Definition von Konvergenz. Zu $x \in \mathbb{R}$ und $\varepsilon > 0$ sei die ε-Umgebung (oder ε-Kugel um x) $U_\varepsilon(x)$ von x definiert durch

$$U_\varepsilon(x) := \{z \in \mathbb{R} : |z - x| < \varepsilon\}.$$

Es gilt also

$$U_\varepsilon(x) = (x - \varepsilon, x + \varepsilon).$$

Eine solche Umgebung ist dann nichts anderes als ein offenes Intervall um x. In diesem Sinne hätten wir das Konzept der Umgebung also nicht neu einführen müssen. Für spätere Verallgemeinerungen erweist sich aber das Denken mit Umgebungen als sehr nütztlich.

Bemerkung. Allgemein nennt man eine Menge U Umgebung von $x \in \mathbb{R}$, wenn es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(x) \subset U$.

PROPOSITION. *(Charakterisierung Konvergenz) Sei (x_n) eine Folge in $\mathbb{R}$. Dann sind äquivalent:*

(i) *(x_n) konvergiert gegen x*
(ii) *Für alle $\varepsilon > 0$ existiert ein $n_\varepsilon \in \mathbb{N}$ so daß für alle $n \geq n_\varepsilon$ gilt $x_n \in U_\varepsilon(x)$.*

(iii) *Für alle $\varepsilon > 0$ ist die Menge $\{n \in \mathbb{N} : x_n \notin U_\varepsilon(x)\}$ endlich (d.h. für jedes feste $\varepsilon > 0$ liegen bis auf endlich viele Ausnahmen alle x_n in der ε-Umgebung von x).*

Beweis. Die Äquivalenz von (i) und (ii) ist klar. (Denn: $z \in U_\varepsilon(x) \Longleftrightarrow |z - x| < \varepsilon$.)
$(ii) \Longleftrightarrow (iii)$: Eine Teilmenge von $\mathbb{N}$ ist genau dann endlich, wenn ab einem gewissen n_0 keine natürliche Zahl mehr dazu gehört. □

Bevor wir uns der Existenz konvergenter Folgen widmen, sammeln wir hier schon einmal ein paar nützliche Eigenschaften.

Folgerung. *Ist die Folge (x_n) in $\mathbb{R}$, so ist die Menge $\{x_n : n \in \mathbb{N}\}$ beschränkt.*

Beweis. Sei $x := \lim\limits_{n\to\infty} x_n$. Für $n \geq n_1$ gilt $|x - x_n| < 1$, also

$$|x_n| = |x_n - x + x| \leq |x_n - x| + |x| \leq 1 + |x|.$$

Damit folgt

$$|x_n| \leq \max\{|x_1|, \ldots, |x_{n_1-1}|, 1 + |x|\}$$

für alle $n \in \mathbb{N}$. □

Bemerkung. Beschränktheit einer Folge hängt nicht von den ersten endlich vielen Gliedern ab.

Wir werden nun beweisen, daß Konvergenz mit einer ganzen Reihe von Operationen vertraeglich ist. Die folgenden Betrachtungen zeigen insbesondere, daß Konvergenz mit den Operationen $+$, $\cdot$, $:$ und $|\cdot|$ verträglich ist.

Proposition (Rechenregeln). *Seien (x_n) und (y_n) Folgen in $\mathbb{R}$. Dann gilt.*
(a) $x_n \to x$, $y_n \to y \Longrightarrow x_n + y_n \to x + y$.
(b) $x_n \to x$, $y_n \to y \Longrightarrow x_n y_n \to xy$. Insbesondere $\alpha x_n \to \alpha x$ für $\alpha \in \mathbb{R}$.
(c) $x_n \to x$, und $y_n \to y$ mit $y \neq 0 \Longrightarrow y_n \neq 0$ für $n \geq n_0$ und $\frac{x_n}{y_n} \to \frac{x}{y}$.

Bemerkung. Später werden wir obige Proposition so ausdruecken: Die Abbildungen $+ : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \to x + y$, $\cdot : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \to xy$ und $:: \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \mapsto x/y$ sind stetig.

Beweis. (a) Sei $\varepsilon > 0$. Wegen $x_n \to x$ gibt es ein n_x mit $|x_n - x| < \frac{\varepsilon}{2}$ für $n \geq n_x$. Wegen $y_n \to y$ gibt es ein n_y mit $|y_n - y| < \frac{\varepsilon}{2}$ für $n \geq n_y$. Für $n \geq n_\varepsilon := \max\{n_x, n_y\}$ gilt dann also nach Dreiecksungleichung

$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(b) Sei $\varepsilon > 0$ beliebig.

Wähle $C > 0$ mit $|x_n| \leq C$ für alle n. Wegen $x_n \to x$ existiert $n_x \in \mathbb{N}$ mit $|x_n - x| < \frac{\varepsilon}{2|y|+1}$.Wegen $y_n \to y$ existiert $n_y \in \mathbb{N}$ mit $|y_n - y| < \frac{\varepsilon}{2C}$. Damit gilt für $n \geq \max\{n_x, n_y\}$ also

$$|x_n y_n - xy| = |x_n y_n - x_n y + x_n y - xy| \leq |x_n||y_n - y| + |y||x_n - x|...$$

(c) Wegen $y_n \to y$ existiert ein n_0 mit $|y_n - y| < \frac{|y|}{2}$ für $n \geq n_0$. Damit gilt also für $n \geq n_0$

$$|y_n| \geq |y| - |y_n - y| > |y|/2 > 0.$$

Sei nun $\varepsilon > 0$ beliebig. Wegen $x_n \to x$ existiert ein n_x mit

$$|x_n - x| < \frac{\varepsilon |y|}{4}$$

für alle $n \geq n_1$. Wegen $y_n \to y$ existiert ein n_y mit

$$|y_n - y| < \frac{\varepsilon |y|^2}{2|x| + 1}.$$

Für $n \geq \max\{n_x, n_y, n_1\}$ gilt dann

$$\begin{aligned} \left| \frac{x_n}{y_n} - \frac{x}{y} \right| &= \left| \frac{y x_n - x y_n}{y y_n} \right| \\ &\leq \frac{2}{|y|^2} |y x_n - x y_n| \\ &\leq \frac{2}{|y|^2} (|y||x_n - x| + |x||y_n - y|) \\ &< \varepsilon \end{aligned}$$

Das beendet den Beweis. □

Ende der 11. Vorlesung

Proposition (Vertraeglichkeit mit Betrag). *Sei (x_n) eine Folge in $\mathbb{R}$ mit $x_n \to x$, $n \to \infty$. Dann folgt $|x_n| \to |x|$, $n \to \infty$.*

Beweis. Das folgt sofort aus der zweiten Dreiecksungleichung

$$||x| - |x_n|| \leq |x - x_n| :$$

Ist naemlich $\varepsilon > 0$ beliebig gegeben, so existiert aufgrund von $x_n \to x$ ein $N \in \mathbb{N}$ mit $|x - x_n| < \varepsilon$ fuer alle $n \geq N$ und aufgrund der genannten Dreiecksungleichung gilt dann fuer solche n ebenfalls $||x| - |x_n|| < \varepsilon$. □

Proposition (Verträglichkeit mit Bildung von Maximum und Minimum). *Seien (x_n) und (y_n) Folgen in $\mathbb{R}$ und $x, y \in \mathbb{R}$ mit $x_n \to x$ und $y_n \to y$. Dann gilt* $\max\{x_n, y_n\} \to \max\{x, y\}$ *und* $\min\{x_n, y_n\} \to \min\{x, y\}$.

Beweis. Es gilt

$$\min\{a,b\} = \frac{a+b-|a-b|}{2}, \ \max\{a,b\} = \frac{a+b+|a-b|}{2}$$

Nun folgt die Behauptung aus den schon gezeigten Aussagen zur Vertraeglichkeit der Konvergenz mit Betragsbildung und mit Addition und Subtraktion. □

PROPOSITION (Verträglichkeit von Konvergenz mit $\leq$). *Sei (x_n) eine Folge in $\mathbb{R}$ mit $x_n \leq c$ / $x_n \geq c$ und $x_n \to x$. Dann gilt $x \leq c$ / $x \geq c$.*

Beweis. Wir betrachten $x_n \leq c$. Angenommen $x > c$. Dann ist $\varepsilon := x - c > 0$. Wegen $x_n \to x$ müsste gelten $x_n \in U_\varepsilon(x)$ für große n, also $x_n > c$. Widerspruch. □

Bemerkung. Gilt $x_n < c$ fuer alle $n \in \mathbb{N}$ sowie $x_n \to x$ so folgt im allgemeinen NICHT $x < c$. Beispiel: $x_n = 1 - \frac{1}{n}$. Dann $x_n < 1$, aber $x = \lim x_n = 1$

THEOREM (Sandwichtheorem). *Seien $L \in \mathbb{R}$ und konvergente Folgen (x_n) und (y_n) in $\mathbb{R}$ mit $\lim_{n\to\infty} x_n = L = \lim_{n\to\infty} y_n$ gegeben. Ist (z_n) eine weitere Folge in $\mathbb{R}$ mit $x_n \leq z_n \leq y_n$ für alle n ab einem gewissen $n_0 \in \mathbb{N}$, so konvergiert (z_n) ebenfalls gegen L.*

Beweis. **Zeichnung.**
Sei $\varepsilon > 0$ beliebig.
Wegen $L = \lim x_n$ gibt es ein $n_x \in \mathbb{N}$ mit $x_n \geq L - \varepsilon$ für alle $n \geq n_x$.
Wegen $L = \lim y_n$ gibt es ein $n_y \in \mathbb{N}$ mit $y_n \leq L + \varepsilon$ für alle $n \geq n_y$.
Für $n \geq n_x, n_y, n_0$ gilt dann also

$$L - \varepsilon \leq x_n \leq z_n \leq y_n \leq L + \varepsilon$$

und damit

$$|z_n - L| < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Es ist sinnvoll in gewissen Fällen auch $\pm\infty$ als Wert zuzulassen: Tatsaechlich gibt es unter den divergenten Folgen in $\mathbb{R}$ zwei Klassen von Folgen mit besonders guten Eigenschaften:

- Eine Folge (x_n) in $\mathbb{R}$ heißt bestimmt divergent gegen ∞, $x_n \to \infty$, wenn für jedes $C \in \mathbb{R}$ ein $n_C \in \mathbb{N}$ existiert mit $x_n \geq C$ für alle $n \geq n_C$.
- Eine Folge (x_n) in $\mathbb{R}$ heißt bestimmt divergent gegen $-\infty$, $x_n \to -\infty$, wenn für jedes $C \in \mathbb{R}$ ein $n_C \in \mathbb{N}$ existiert mit $x_n \leq C$ für alle $n \geq n_C$.

Bei vorsichtigem Umgang mit ∞ bleiben einige Rechenregeln für konvergente Folgen auch für bestimmt divergente Folgen noch gültig. Eine **wichtige Ausnahme** stellt der Umgang mit Termen der Form $0 \cdot \infty$ und $\infty - \infty$ dar (siehe Übung).

Bemerkung. Es gilt folgende Aequivalenz: $x_n \to \infty \iff x_n > 0$ für alle genuegend großen n und $1/x_n \to 0$. Entsprechendes gilt fuer $x_n \to -\infty$.

Beispiele. $\sqrt[n]{n!} \to \infty$, $n \to \infty$.
Bew. $n! = n(n-1)\cdots 1 \geq n/2 \cdots n/2 = (n/2)^{n/2}$. ($n/2$-Faktoren). Damit $\sqrt[n]{n!} \geq \sqrt[n]{(n/2)^{n/2}} = (n/2)^{1/2} \to \infty$.
Ähnlich kann man für Supremum und Infimum von unbeschränkten Mengen verfahren:

- Sei M eine Teilmenge von $\mathbb{R}$. Ist M nicht nach oben / unten beschränkt, so setzten wir $\sup M = \infty$ / $\inf M = -\infty$.

Bemerkung.

- Gilt $\sup M = \infty$, so existiert Folge (x_n) in M mit $x_n \to \infty$. Entsprechendes gilt für Infimum.
- Mit obiger Definition hat also jede Menge M in $\mathbb{R}$ ein Supremum bzw. ein Infimum.

Mit den bisherigen Betrachtungen können wir Konvergenz einiger Folgen untersuchen.

Beispiele. (a) Sei $a \neq 0$ und $x_n = \frac{a}{n}$. Dann konvergiert (x_n) gegen 0. (Nullfolge).
Bew. Archimedes oder Rechenregeln $a/n = a \cdot \frac{1}{n}$.

(b) Sei $0 < q < 1$ und $x_n = q^n$. Dann konvergiert (x_n) gegen 0. (Exponentielles Fallen).

Bew. Das ist eine Folgerung aus dem Archimedischen Axiom und wurde oben schon behandelt.

(c) Sei $0 < q < 1$ und $k \in \mathbb{N}$. Sei $x_n = q^n n^k$. Dann gilt $x_n \to 0$, $n \to \infty$. (Exponentielles Fallen schlägt polynomielles Wachsen)
Beachte. $q^n \to 0$, aber $n^k \to \infty$ für $k > 1$. Die Frage ist also, welcher Effekt sich durchsetzt.
Bew. Plan: Schreibe q^n also $p^{2kn} = p^{kn}p^{kn}$ mit geeignetem p mit $0 < p < 1$. Dann gilt also

$$q^n n^k = p^{kn}p^{kn}n^k = p^{kn}(p^{kn}n^k).$$

Erster Term gegen Null; zweiter Term beschränkt. Hier sind die Details: $p := \sqrt[2k]{q} < 1$.
Also $1/p = 1 + a$ mit $a > 0$. Die Bernoulli - Ungleichung impliziert $(1/p)^n \geq 1 + na$ und damit $0 < p^n < \frac{1}{na}$. Mit $0 \leq q^n = p^{2kn}$ folgt also

$$0 \leq q^n n^k \leq p^{kn}p^{kn}n^k \leq p^{kn}\left(\frac{1}{na}\right)^k n^k = p^{kn}\left(\frac{1}{a}\right)^k.$$

Damit folgt die Aussage aus (b) und dem Sandwichtheorem.

Ende der 12. Vorlesung

(d) Sei $a > 0$ und $x_n = \sqrt[n]{a}$. Dann gilt $x_n \to 1$, $n \to \infty$.

Bew. Wir unterschieden drei Fälle:

$a = 1$. Das ist einfach.

$a > 1$: Wegen $a > 1$ und $a = x_n^n$ gilt $x_n > 1$. Weiterhin gilt

$$a = x_n^n = (1 + (x_n - 1))^n \geq 1 + n(x_n - 1).$$

Damit folgt $(a-1)/n \geq x_n - 1 \geq 0$ und dann $(a-1)/n + 1 \geq x_n \geq 1$. Damit folgt nach (a) und dem Sandwichtheorem $x_n \to 1$.

$0 < a < 1$: Das folgt nach Bilden des Kehrwertes aus dem schon bewiesenen.

(e) Sei $x_n = \sqrt[n]{n}$. Dann gilt $x_n \to 1$, $n \to \infty$.

Beachte. Wettstreit zwischen $\sqrt[n]{a} \to 1$ und $n \to \infty$ für $n \to \infty$. Die Frage ist also, welcher Effekt sich durchsetzt.

Bew. Es gilt $x_n^n = n$, also insbesondere $x_n > 1$. Mit binomischem Satz folgt

$$n = (1+(x_n-1))^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k}(x_n-1)^k \geq \binom{n}{2}(x_n-1)^2 = \frac{n(n-1)}{2}(x_n-1)^2,$$

also

$$\sqrt{\frac{2}{n-1}} \geq x_n - 1 \geq 0,$$

also

$$1 \leq x_n \leq 1 + \sqrt{\frac{2}{n-1}}.$$

Mit $\frac{2}{n-1} \to 0$, $n \to \infty$, folgt auch $\sqrt{\frac{2}{n-1}} \to 0$, $n \to \infty$. Damit folgt Behauptung aus dem Sandwichtheorem.

2. Aspekte der Vollständigkeit

Die Vollständigkeit von $\mathbb{R}$ liefert die Konvergenz ganzer Klassen von konvergenten Folgen. Tatsächlich ist diese Konvergenz zusammen mit dem Archimedischen Axiom ein Charakteristikum der reellen Zahlen. Das wird in diesem Abschnitt studiert.

Definition. *Eine Folge (x_n) in $\mathbb{R}$ heißt monoton wachsend / fallend wenn $x_{n+1} \geq x_n$ / $x_{n+1} \leq x_n$ für alle $n \in \mathbb{N}$ gilt.*

Zeichungen einer nach oben beschränkten monotonen Folge: auf der Achse, oder als Graph...

Bemerkung. Ist (x_n) monoton wachsend, so gilt $x_n \leq x_m$ fuer alle $n \leq m$ (nach einer einfachen Induktion). Entsprechendes gilt fuer monoton fallende Folgen.

DEFINITION. *Eine Folge (x_n) heißt nach oben / unten beschränkt, wenn die Menge $\{x_n : n \in \mathbb{N}\}$ nach oben / unten beschränkt ist.*

THEOREM. *(Konvergenz monotoner beschränkter Folgen) Jede monoton wachsende / fallende nach oben / unten beschränkte Folge in $\mathbb{R}$ konvergiert.*

Beweis. **Zeichnung.** Sei (x_n) eine monoton wachsende, nach oben beschränkte Folge. Dann existiert also aufgrund der Ordngungsvollständigkeit

$$S := \sup\{x_n : n \in \mathbb{N}\} \in \mathbb{R}.$$

Da S eine obere Schranke von ist, gilt

$$x_n \leq S$$

fuer alle $n \in \mathbb{N}$. Sei $\varepsilon > 0$ beliebig. Dann existiert, da $S - \varepsilon$ keine obere Schranke ist, also ein $n_\varepsilon \in \mathbb{N}$ mit

$$S - \varepsilon \leq x_{n_\varepsilon}.$$

Damit folgt also für alle $n \geq n_\varepsilon$ aufgrund der Monotonie

$$S - \varepsilon \leq x_{n_\varepsilon} \leq x_n \leq S$$

also $|S - x_n| < \varepsilon$.
Der Fall monoton fallender nach unten beschränkter Folgen kann analog behandelt werden. □

Bemerkung.

- Neben der Konvergenz von $(\frac{1}{n})$ gegen 0 haben wir also in unseren Zugang zu $\mathbb{R}$ eine Methode zur Erzeugung konvergenter Folgen eingebaut.
- Ist die Folge (x_n) wachsend / fallend, so gilt $x_n \to C$, wobei $C \in \mathbb{R}$ (falls die Folge (x_n) beschraenkt ist) und $C = \infty$ / $C = -\infty$ (falls die Folge (x_n) unbeschraenkt ist).

Beispiel - die Eulersche Zahl e: $x_n := (1 + \frac{1}{n})^n$ konvergiert. Der Grenzwert wird Eulersche Zahl genannt und mit e bezeichnet.

Beachte. $1 + \frac{1}{n} \to 1$, aber $a^n \to \infty$ für $a > 1$. Die Frage ist also, welcher Effekt sich durchsetzt.

Bew: Wir zeigen, daß (x_n) wachsend und beschränkt ist. Die Bernoulli Ungleichung liefert

$$\left(1 + \frac{1}{n}\right)^n \left(1 - \frac{1}{n}\right)^n = \left(1 - \frac{1}{n^2}\right)^n \geq \left(1 - \frac{1}{n}\right)$$

also

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

für alle $n \geq 2$. Die Folge (x_n) ist also wachsend. Die Folge (x_n) ist beschränkt durch 3:

$$
\begin{aligned}
x_n &= \left(1+\frac{1}{n}\right)^n \\
(Binomi) &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{1\cdot 2\cdots k}\frac{1}{n^k} \\
(Umsortieren) &= \sum_{k=0}^{n} \frac{n(n-1)\ldots(n-k+1)}{n\ldots n}\frac{1}{k!} \\
&\leq 1+\sum_{k=1}^{n}\frac{1}{k!} \\
(Induktion:\ k! \geq 2^{k-1}) &\leq 1+\sum_{k=1}^{n}\frac{1}{2^{k-1}} \\
(Geom.Summe) &\leq 1+\frac{1-(1/2)^{n+1}}{1-1/2} \leq 3.
\end{aligned}
$$

Bemerkung. Später werden wir zeigen:

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!} =: \sum_{k=0}^{\infty}\frac{1}{k!}.$$

Bemerkung. Die Zahl e spielt bei kontinuierlichen Wachstumsvorgängen eine Rolle, z.B. stetige Verzinsung: Kapital A Zinsatz 100 prozent. Dann hat man nach einem Jahr bei

- 0 mal (äquidistant) Abheben und Einzahlen : $A(1+1)$
- 1 mal (äquidistant) Abheben und Einzahlen: $A(1+1/2)(1+1/2)$
- 2 mal (äquidistant) Abheben und Einzahlen: $A(1+1/3)(1+1/3)(1+1/3)$
- etc.

Beispiel - die Zahl $e(a)$: Sei $a > 0$. Sei $x_n := (1+a/n)^n$. Dann konvergiert (x_n). Wir nennen den Grenzwert $e(a)$. (Später: $e(a) = e^a$.).

Bew. Wir zeigen Monotonie und Beschränktheit.

Monotonie:

$$\begin{aligned}
x_n &= \sum_{k=0}^{n} \binom{n}{k} a^k \\
\text{(siehe voriges Bsp.)} \quad &= \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{n\cdot n\cdots n} \frac{a^k}{k!} \\
\text{(Fuer } 0 \le l < r \text{ gilt } l/r \le (l+1)/(r+1))) \quad &\le \sum_{k=0}^{n} \frac{(n+1)n\cdots(n+1-k+1)}{(n+1)\cdot(n+1)\cdots(n+1)} \frac{a^k}{k!} \\
&\le \sum_{k=0}^{n+1} \frac{(n+1)n\cdots(n+1-k+1)}{(n+1)\cdot(n+1)\cdots(n+1)} \frac{a^k}{k!} \\
&= x_{n+1}.
\end{aligned}$$

(Alternativer direkter Beweis der Monotonie: Sei $a > 0$:

$$\begin{aligned}
\frac{x_{n+1}}{x_n} &= \left(1 + \frac{a}{n+1}\right) \left(\frac{1+\frac{a}{n+1}}{1+\frac{a}{n}}\right)^n \\
&= \left(1 + \frac{a}{n+1}\right) \left(\frac{(n+1+a)n}{(n+a)(n+1)}\right)^n \\
&= \left(1 + \frac{a}{n+1}\right) \left(\frac{(n+1)n+na}{(n+1)(n+a)}\right)^n \\
&\ge \left(1 + \frac{a}{n+1}\right) \left(1 - \frac{a}{(n+1)(n+a)}\right)^n \\
(\mathit{Bernoulli}) \quad &\ge \left(1 + \frac{a}{n+1}\right) \left(1 - \frac{na}{(n+1)(n+a)}\right) \\
&= \left(\frac{n+1+a}{n+1}\right) \left(\frac{(n+1)(n+a)-an}{(n+1)(n+a)}\right) \\
(\mathit{Sortieren\ n.\ Potenzen}) \quad &= \frac{n^3+n^2(a+2)+n(1+2a)+a(1+a)}{n^3+n^2(a+2)+n(1+2a)+a} \\
(a > 0) \quad &\ge 1)
\end{aligned}$$

Beschränkteit:
Die Betrachtungen des vorigen Beispiels führen auf

$$x_n \le \sum_{k=0}^{n} \frac{a^k}{k!}.$$

Wähle nun $N \in \mathbb{N}$ mit $N \ge 2a$. Dann gilt

$$x_n = \sum_{k=0}^{N} \frac{a^k}{k!} + \sum_{k=N+1}^{n} \frac{a^k}{k!} = C + \frac{a^{N+1}}{N!} \sum_{k=N+1}^{n} \frac{a^{k-N-1}}{(N+1)\cdots k} \le C + \frac{a^{N+1}}{N!} 3.$$

(Grundidee $a^k/k!$ ist schließlich a/l mit l groß....)

Beispiel - die Zahl $e(-a)$**:** Sei $a > 0$ und $b = -a < 0$. Dann gilt $\lim_{n\to\infty}\left(1+\frac{b}{n}\right)^n = \frac{1}{e(a)}$.
Bew. (Übung.= Idee $\left(1+\frac{a}{n}\right)^n\left(1-\frac{a}{n}\right)^n = (1+a^2/n^2)^n$ konvergiert gegen 1...

Ende der 13. Vorlesung

Beispiel - k**-te Wurzel.** (Übung) Sei $a > 0$ beliebig. Definiere induktiv die Folge (x_n) durch $x_0 := c > 0$ beliebig und

$$x_{n+1} := \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right) = x_n + \frac{x_n}{k}\left(\frac{a}{x_n^k} - 1\right).$$

Dann konvergiert die Folge (x_n) gegen $\sqrt[k]{a}$.

Beweisskizze: Offenbar (?Induktion!) gilt $x_n > 0$ für alle $n \in \mathbb{N}$.
Bernoulli Ungleichung liefert (Wie?):

$$x_{n+1}^k \geq a$$

für alle $n \in \mathbb{N}$. Außerdem gilt nach Definition

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

Damit ist also (Warum?) $(x_n)_{n\geq 2}$ monoton fallend und durch 0 nach unten beschränkt. Also konvergiert die Folge (x_n) nach dem Satz gegen einen Grenzwert b. Dieser Grenzwert erfüllt (Wieso?)

$$b = \frac{1}{k}\left((k-1)b + \frac{a}{b^{k-1}}\right)$$

und damit gilt dann (Wieso?) auch $b^k = a$.

Bemerkung.(Fuer spaeter ;-) Diese Betrachtungen sind ein Fall des sogenannten Newton Verfahren. Damit kann man (oft) eine Nullstelle einer Funktion f (hier $x^k - a$) auf folgende Art berechnen:

$n = 0$: Wähle einen (geeigneten) Wert p.

$n \Longrightarrow n+1$: Ist x_n schon bestimmt, so berechnet man x_{n+1} wie folgt: Bilde die Tangente an $(x, f(x))$ und berechne ihren Schnitt mit der x-Achse. Dieser Schnittpunkt ist dann x_{n+1}. (Zeichnung). Rechnung liefert die Rekursion

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Im konkreten Fall geht es um die Funktion $f(x) = x^k - a$.

Der Satz zur Konvergenz monotoner Folgen mag erst einmal speziell erscheinen, da keineswegs jede Folge monoton ist. Aber er hat weitreichende Konsequenzen. Um das näher zu erläutern, brauchen wir noch einen neuen Begriff:

DEFINITION (Teilfolge). *Sei $(x_n)_n$ eine Folge und $(n_k)_k$ eine strikt wachsende Folge in $\mathbb{N}$ (d.h. $n_{k+1} > n_k$ für alle $k \in \mathbb{N}$). Dann heißt (x_{n_k}) eine Teilfolge von (x_n).*

Bemerkungen.

- Ist $n_\bullet : \mathbb{N} \longrightarrow \mathbb{N}, k \mapsto n_k$ und $x : \mathbb{N} \longrightarrow X$, so ist (x_{n_k}) gerade die Abbildung $x \circ n_.$.
- Für den Begriff der Teilfolge ist nicht wichtig, daß die Werte der Folge in $\mathbb{R}$ liegen.

Zeichnung. $x_1, x_2 \ldots$ vs $x_{n_1}, x_{n_2} \ldots$

Beispiel. $x_n = (-1)^n$. Dann ist (x_{2n}) die konstante Folge 1 und (x_{2n+1}) die konstante Folge -1. Diese Teilfolgen sind konvergent also 'schöner' als die Ursprungsfolge. Das ist ein allgemeines Phänomen (s.u.).

Das folgende Lemma zeigt, daß der Unterschied zwischen beliebigen Folgen und monotonen Folgen doch nicht so groß ist.

LEMMA. *Jede Folge in $\mathbb{R}$ enthält eine monotone Teilfolge.*

Beweis. Sei (x_n) eine Folge in $\mathbb{R}$. Ein $N \in \mathbb{N}$ heißt Gipfelpunkt von (x_n), wenn gilt $x_N \geq x_n$ für alle $n \geq N$. Wir unterscheiden zwei Fälle.
Fall 1: Es gibt unendlich viele Gipfelpunkte. Sei $n_k :=$ k-ter Gipfelpunkt. Dann ist $n_1 < n_2 < \ldots n_k < n_{k+1}$ die Folge von Gipfelpunkten. Daher gilt also

$$x_{n_k} \geq x_{n_{k+1}}$$

für alle $k \in \mathbb{N}$ und (x_{n_k}) ist eine monoton fallende Teilfolge.

Fall 2: Es gibt nur endlich viele Gipfelpunkte. Wir konstruieren induktiv streng monoton wachsende (n_k), so daß x_{n_k} monoton wachsend ist.
$k = 1$: Wähle n_1 größer als jeden Gipfelpunkt (möglich, da nur endlich viele Gipfelpunkte).
$k \Longrightarrow k+1$: Seien $n_1 < n_2 \ldots < n_k$ schon konstruiert. Da n_k kein Gipfelpunkt ist ($n_k > n_1 > jeder\, Gipfelpunkt$), gibt es ein $n_{k+1} > n_k$ mit $x_{n_k} < x_{n_{k+1}}$. □

THEOREM (Satz von Bolzano - Weierstraß). *Jede beschränkte Folge in $\mathbb{R}$ enthält eine konvergente Teilfolge.*

Beweis. Nach dem vorigen Lemma enthält die Folge eine monotone Teilfolge. Diese ist beschränkt (da die Ursprungsfolge beschränkt ist). Damit konvergiert die Teilfolge nach dem Satz über Konvergenz monotoner beschränkter Folgen. □

Wir werden nun Konvergenz von Folgen in $\mathbb{R}$ charakterisieren. Dazu benötigen wir den folgenden Begriff.

DEFINITION (Cauchy Folge). *Eine Folge (x_n) in $\mathbb{R}$ heißt Cauchy-Folge wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $n_\varepsilon \in \mathbb{N}$, sodaß für alle $n, m \geq n_\varepsilon$ gilt*

$$|x_n - x_m| < \varepsilon.$$

Idee. Folgeglieder sind beliebig nahe aneinander für genügend große n und m.

Bemerkung. In Quantorisch:

$$\forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n, m \geq n_\varepsilon \, |x_n - x_m| < \varepsilon.$$

LEMMA. *Jede konvergente Folge in $\mathbb{R}$ ist eine Cauchy-Folge.*

Beweis. (Eigentlich klar: Wenn Folgeglieder nahe am Grenzwert sind, sind sie auch nahe aneinander. **Zeichnung**.)
Details: Sei $x_n \to x$, $n \to \infty$. Sei $\varepsilon > 0$ beliebig. Dann existiert ein $n_\varepsilon \in \mathbb{N}$ mit $|x_n - x| < \frac{\varepsilon}{2}$ für $n \geq n_\varepsilon$. Also

$$|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x_m - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n, m \geq n_\varepsilon$. □

LEMMA. *Eine Cauchy-Folge in $\mathbb{R}$ mit einer konvergenten Teilfolge ist konvergent (gegen den Grenzwert der Teilfolge).*

Beweis. Sei (x_n) eine Cauchy-Folge und sei (x_{n_k}) eine gegen x konvergente Teilfolge. Wir zeigen $\lim x_n = x$.
Sei $\varepsilon > 0$ gegeben.
Es existiert $n_\varepsilon \in \mathbb{N}$ mit $|x_n - x_m| < \frac{\varepsilon}{2}$ für alle $m, n \geq n_\varepsilon$ (da Cauchy Folge).
Es existiert $k_0 \in \mathbb{N}$ mit $|x_{n_k} - x| < \frac{\varepsilon}{2}$ für $k \geq k_0$ (da Teilfolge konvergent).
Sei nun $k \geq k_0$ mit $n_k \geq n_\varepsilon$ gewählt.
Es gilt für alle $n \geq n_\varepsilon$

$$|x_n - x| = |x_n - x_{n_k} + x_{n_k} - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Das beendet den Beweis. □

Wir kommen nun zur angekündigten Charakterisierung von Konvergenz.

THEOREM (Cauchy-Kriterium). *Sei (x_n) eine Folge in $\mathbb{R}$. Dann sind äquivalent:*

(i) (x_n) *ist konvergent.*
(ii) (x_n) *ist eine Cauchy Folge.*

Beweis. Die Implikation $(i) \Longrightarrow (ii)$ haben wir in einem vorausgehenden Lemma gezeigt.

$(ii) \Longrightarrow (i)$: Sei (x_n) eine Cauchy Folge. Dann ist (x_n) beschränkt (Wähle n_1 mit $|x_n - x_{n_1}| < 1$ für $n \geq n_1$. Dann gilt für $n \geq n_1$ also $|x_n| \leq |x_n - x_{n_1}| + |x_{n_1}| < 1 + |x_{n_1}|$...) Damit hat (x_n) also nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge. Damit ist (x_n) nach dem vorigen Lemma konvergent.

□

Bemerkung. Es handelt sich um eine wesentliche Eigenschaft der reellen Zahlen (siehe folgendes Theorem). Die entscheidende Implikation ist $(ii) \Longrightarrow (i)$.

In gewisser Weise charakterisieren (fast) alle in diesem Abschnitt gegebenen Sätze die reellen Zahlen. Genauer gilt folgendes.

THEOREM (Die große Charakterisierung). *Sei K ein angeordneter Körper. Dann sind äquivalent:*

(i) *Es ist K ordnugsvollständig (d.h. $K = \mathbb{R}$).*
(ii) *Es erfüllt K das Intervallschachtelungsprinzip und das Archimedische Axiom.*
(iii) *Jede monotone beschränkte Folge konvergiert, und es gilt das Archimedische Axiom.*
(iv) *Jede beschränkte Folge hat eine konvergente Teilfolge, und es gilt das Archimedische Axiom.*
(v) *Jede Cauchy-Folge konvergiert, und es gilt das Archimedische Axiom.*

Beweis. (i)$\Longleftrightarrow$ (ii): Das wurde schon in Kapitel 3 durchgeführt.

(i)$\Longrightarrow$ (iii): Siehe oben. (Definiere $S := \sup\{x_n : n \in \mathbb{N}\}$....)

(iii)$\Longrightarrow$ (iv): Jede beschränkte Folge hat eine monotone Teilfolge (nach dem oben gegebenen Beweis). Damit folgt die gewuenschte Implikation.

(iv)$\Longrightarrow$ (v): Jede Cauchy Folge ist beschränkt und jede Cauchy Folge mit konvergenter Teilfolge konvergiert (nach den oben gegebenen Beweisen). Damit folgt die gewuenschte Implikation.

(v)$\Longrightarrow$ (ii): Ist I_n eine Intervallschachtelung, so bilden die Mittelpunkte (rechte Randpunkte, linke Randpunkte,...) eine Cauchy Folge. Der Grenzwert x hat die gewünschten Eigenschaften. □

Ende der 14. Vorlesung

3. Teilfolgen und Häufungspunkte

Wir kommen nun zu einem wichtigen Konzept, das das Konzept der Teilfolge komplementiert.

Erinnerung: Eine Menge X heißt endlich, wenn es ein $n \in \mathbb{N}$ gibt und eine bijektive Abbildung $j : \{1, \ldots, n\} \longrightarrow X$. Andernfalls heißt sie unendlich (siehe unten fuer weitere Diskussion).

LEMMA. *(Charakterisierung Häufungspunkt) Sei (x_n) eine Folge in $\mathbb{R}$ und $x \in \mathbb{R}$ gegeben. Dann sind äquivalent:*

(i) *Es gibt eine Teilfolge von (x_n), die gegen x konvergiert.*
(ii) *Für alle $\varepsilon > 0$ ist $\{n \in \mathbb{N} : |x_n - x| < \varepsilon\}$ eine unendliche. Menge*

Beweis. $(i) \Longrightarrow (ii)$: Sei $\varepsilon > 0$. Aufgrund von (i) gibt es eine Teilfolge x_{n_k} mit $x_{n_k} \to x$, $k \to \infty$. Damit folgt also für $k \geq k_\varepsilon$

$$|x_{n_k} - x| < \varepsilon.$$

Damit gilt

$$\{n \in \mathbb{N} : |x_n - x| < \varepsilon\} \supset \{n_k : k \geq k_\varepsilon\}$$

und es folgt die Behauptung.

$(ii) \Longrightarrow (i)$: Wir konstruieren induktiv eine Teilfolge mit

$$|x_{n_k} - x| < \frac{1}{k}.$$

Diese Teilfolge konvergiert dann also gegen x.

$k = 1$: Wähle n_1 mit $|x_{n_1} - x| < 1$.

$k \Longrightarrow k + 1$: Aufgrund von (ii) ist die Menge

$$A := \{n \in \mathbb{N} : |x_n - x| < \frac{1}{k+1}\} \cap \{n \in \mathbb{N} : n > n_k\}$$

nichtleer. Wir können also n_{k+1} aus A wählen z.B. als kleinstes Element von A. Dann gilt

$$|x_{n_k} - x| < \frac{1}{k}.$$

Damit folgt $(x_{n_k}) \to x$ (nach Archimedischen Axiom). □

Bemerkung: Beweis nutzt, daß $x_n \to x$ genau dann gilt, wenn fuer alle $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$ existiert mit $|x_n - x| < 1/k$ für $n \geq n_k$.

DEFINITION (Häufungspunkt). *Sei (x_n) eine Folge in $\mathbb{R}$. Ein $x \in \mathbb{R}$ heißt Häufungspunkt von (x_n), wenn eine der beiden äquivalenten Bedingungen des vorigen Lemma gilt.*

Bemerkung. Gibt es eine Teilfolge von (x_n) die gegen ∞ / $-\infty$ konvergiert, so spricht man manchmal vom uneigentlichen Häufungspunkt ∞ bzw. $-\infty$.

Erkläre im Spaziergängermodell, das Häufen.

Beispiele. (a) $(-1)^n$ hat die beiden Häufungspunkte -1 und 1.

(b) Die Folge x_n mit $x_n = \pi$ für n hat Rest 0 bei Division durch 3, $x_n = 7$ für n mit Rest 1 bei Division durch 3, $x_n = 42$ für n mit Rest 2 bei Division durch 3 hat die drei Häufungspunkte $\pi, 7$ und 42.

(c) $y_n = x_n + \frac{1}{n}$ (mit (x_n) aus (b)) hat ebenfalls die Häufungspunkte $\pi, 7, 42$ (obwohl diese Werte nicht angenommen werden; vgl. Konvergenz!)

Bemerkung. Der Satz von Bolzano/Weierstraß lässt sich nun so formulieren: Jede beschränkte Folge in $\mathbb{R}$ hat einen Häufungspunkt.

Für beschränkte Folgen in $\mathbb{R}$ gibt es zwei besondere Häufungspunkte, nämlich den größten und den kleinsten Häufungspunkt. Das werden wir jetzt genauer untersuchen:

Sei (x_n) eine Folge in $\mathbb{R}$.

Dann ist die Folge (X_N) mit

$$X_N := \sup\{x_n : n \geq N\}$$

fallend in N. Damit existiert

$$\limsup_{n\to\infty} x_n := \lim_{N\to\infty} X_N = \lim_{n\to\infty} \sup\{x_k : k \geq n\}.$$

und wird als 'Limsup' oder 'Limes superior' von (x_n) bezeichnet. Hier ist der Wert $+\infty$ möglich (wenn nämlich die Folge nach oben nicht beschränkt ist).

Analog sieht man, daß die Folge (X_M) mit

$$X_M := \inf\{x_n : n \geq M\}$$

wachsend in M ist. Damit existiert

$$\liminf_{n\to\infty} x_n := \lim_{M\to\infty} X_M = \lim_{n\to\infty} \inf\{x_k : k \geq n\}$$

und wird als 'Liminf' oder 'Limes inferior' von (x_n) bezeichnet. Hier ist der Wert $-\infty$ möglich (wenn nämlich die Folge nach unten nicht beschränkt ist).

Lemma (Charakterisierung von $\limsup/\liminf$). *Sei (x_n) eine Folge in $\mathbb{R}$. Sei $x \in \mathbb{R}$. Dann sind äquivalent:*

(i) $x = \limsup_{n\to\infty} x_n$.
(ii) *Für jedes $\varepsilon > 0$ gilt:*
   – *Es gibt ein $n_\varepsilon \in \mathbb{N}$ mit $x_k \leq x + \varepsilon$ für alle $k \geq n_\varepsilon$*
   – *Die Menge $\{n \in \mathbb{N} : x_n \geq x - \varepsilon\}$ ist unendlich.*
(iii) *Es ist (x_n) nach oben beschränkt und es ist x der größte Häufungspunkt (x_n) (d.h. x ist Häufungspunkt von (x_n) und es gibt keinen größeren Häufungspunkt).*

Analoge Aussagen gelten für $\liminf$.

Beweis. (i)$\Longrightarrow$ (ii): Sei $\varepsilon > 0$. Sei $X_N := \sup\{x_n : n \geq N\}$. Wegen $X_N \to x$ fallend, existiert ein $N_0 \in \mathbb{N}$ mit

$$x_k \leq \sup\{x_n : n \geq N_0\} < x + \varepsilon$$

für alle $k \geq N_0$ und es ist $\{n \in \mathbb{N} : x_n \geq x - \varepsilon\}$ unendlich.

(ii)$\Longrightarrow$ (iii): (x_n) *nach oben beschränkt.* Das folgt aus der ersten Eigenschaft.

x *ist Häufungspunkt.* Das folgt aus den beiden Eigenschaften und der Charakterisierung von Häufungspunkten.

Es gibt keinen größeren Häufungspunkt. Das ist klar nach der ersten Eigenschaft.

(iii)$\Longrightarrow$ (i): Sei $X_N := \sup\{x_n : n \geq N\}$.

Sei $\varepsilon > 0$. Dann gibt es ein n_ε mit $x_n \leq x+\varepsilon$ für alle $n \geq n_\varepsilon$ (andernfalls gäbe es nach Bolzano/Weiersterass eine Teilfolge, die gegen eine größere Zahl als x konvergiert. Widerspruch *größter* HP.). Damit folgt

$$X_N \leq x + \varepsilon$$

für alle $N \geq n_\varepsilon$. Damit folgt

$$\limsup x_n = \lim X_N < x + \varepsilon.$$

Umgekehrt gibt es, da x ein Häufungspunkt ist, zu jedem $\varepsilon > 0$ beliebig große n mit $x_n \geq x - \varepsilon$. Damit folgt $X_N \geq x - \varepsilon$ für alle N. Damit folgt $X \geq x - \varepsilon$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Folgerung. *Es gilt* $\limsup x_n < C$ *genau dann wenn ein* $C' < C$ *und* $N_0 \in \mathbb{N}$ *existiert mit* $x_n \leq C'$ *für alle* $n \geq N$. *Enstprechendes gilt für* $\liminf$.

Beweis. $\Longrightarrow$: Das folgt leicht aus (ii).

$\Longleftarrow$: Das folgt sofort aus der Definition. □

Folgerung. *Sei* (x_n) *eine Folge in* $\mathbb{R}$. *Dann gilt:*

(a) Sei $(x_{n_k})_k$ *eine konvergente Teilfolge. Dann gilt*

$$\liminf_n x_n \leq \lim_k x_{n_k} \leq \limsup_n x_n.$$

(b) Es konvergiert(x_n) *genau dann, wenn* $\liminf_n x_n = \limsup_n x_n \in \mathbb{R}$ *und in diesem Fall gilt* $\lim_n x_n = \limsup_n x_n = \liminf_n x_n$.

Beweis. (a) Der Grenzwert $\lim_k x_{n_k}$ ist ein Haeufungspunkt von (x_n). Damit folgt die Aussage (a) sofort aus dem vorigen Lemma.

(b) Wenn die Folge konvergiert, so konvergiert auch jede Teilfolge (mit demselben Grenzwert) und der groeßte und der kleinste Haeufungspunkt stimmen ueberein. Nach dem vorigen Lemma ($(iii) \Longrightarrow (i)$) folgt dann die gewuenschte Gleichheit von $\limsup$ und $\liminf$.

Umgekehrt folgt aus $\liminf_n x_n = \limsup_n x_n \in \mathbb{R}$ und dem vorigen Lemma ($(i) \Longrightarrow (ii)$) auch die Konvergenz. □

Wir notieren noch einige (leicht zu beweisende) Rechenregeln:

PROPOSITION (Rechenregeln). *Seien* (x_n) *und* (y_n) *Folgen in* $\mathbb{R}$. *Dann gilt:*
(a) $\limsup_n x_n = -\liminf(-x_n)$.
(b) $\limsup_n(ax_n) = a\limsup_n x_n$ *fuer* $a \geq 0$.
(c) $\limsup_n(x_n + y_n) \leq \limsup_n x_n + \limsup_n y_n$.
(d) $\limsup_n x_n + \liminf_n y_n \leq \limsup(x_n + y_n)$.
(e) $\limsup_n(x_n + y_n) = \limsup_n x_n + \lim_n y_n$ *(falls* (y_n) *konvergiert).*

Beweis. Hier folgen (a) und (b) direkt aus den Definitionen. Es folgt (c) aus der Definition unter Nutzen von $\sup\{x_n + y_n : n \geq N\} \leq \sup\{x_n : n \geq N\} + \sup\{y_n : n \geq N\}$. Es folgt (d) leicht aus den Definitionen und einem Widerspruchsbeweis oder auch direkt aus (a) und (c). Es folgt (e) einfach. □

Ende der 15. Vorlesung

Bemerkung. Ist $\{x_n : n \in \mathbb{N}\}$ nicht nach oben beschränkt, so gilt $\sup\{x_n : n \geq N\} = \infty$ für alle N. Entsprechend folgt $\limsup x_n = \infty$. In diesem Fall gibt es eine Teilfolge, die gegen ∞ konvergiert. Man kann dann also in diesem Sinne ∞ als den größten Häufungspunkt auffassen. Entsprechendes gilt für $-\infty$ und $\liminf x_n$.

Bemerkung. (Zweipunktkompaktifizierung) Wir betrachten $\pm\infty$ nicht als Punkte einer geeigneten Fortsetzung von $\mathbb{R}$. Stattdessen betrachten wir Ausdrücke wie $x_n \to \infty$ nur als Kurzschreibweise fuer gewisse Zusammenhaenge. Daß das gut möglich ist, liegt daran, daß man die Punkte $\pm$ zu $\mathbb{R}$ dazunehmen kann und damit die sogenannte Zweipunktkompaktifizierung von $\mathbb{R}$ erhält. Wir machen das an einer Zeichnung deutlich.

KAPITEL 5

Mächtigkeit

In diesem Abschnitt untersuchen wir die 'Größe' von Mengen mittels der Anzahl ihrer Elemente. Wir werden drei Abstufungen kennenlernen: endliche Mengen, abzählbar unendliche Mengen und überabzählbare Mengen.

DEFINITION. *(Mächtigkeit)*

- *Eine Menge X heißt endlich, wenn sie leer ist oder ein $n \in \mathbb{N}$ existiert und eine bijektive Abbildung $j : A_n = \{1, \ldots, n\} \longrightarrow X$. Dann heißt 0 bzw. n die Mächtigkeit oder Kardinalität von X.*
- *Eine Menge heißt unendlich, wenn sie nicht endlich ist.*
- *Eine Menge heißt abzählbar unendlich, wenn es eine bijektive Abbildung $J : \mathbb{N} \longrightarrow X$ gibt. In diesem Fall heißt J eine Abzählung und wir schreiben die Menge auch als $X = \{J(1), J(2), \ldots, \}$.*
- *Eine Menge, die weder endlich noch abzählbar unendlich ist, heißt überabzählbar.*

Bemerkung. Die ersten beiden Definitionen haben wir schon kennengelernt. Es handelt sich also lediglich um eine Erinnerung. In der Definition einer endlichen Menge nutzen wir, daß es keine Bijektion zwischen A_n und A_m gibt fuer $n \neq m$.

Beispiele.

- $\mathbb{N}$ ist abzählbar.
- $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ abzählbar.
- $\mathbb{Z}$ ist abzählbar. (Bew. $\mathbb{N}_0 \longrightarrow \mathbb{Z}, 2n \mapsto n, 2n+1 \mapsto -n.$)
- Die Menge der geraden natuerlichen Zahlen ist abzaehlbar. Ebenso ist die Menge der ungeraden natuerlichen Zahlen abzaehlbar.

Notation. In dieser Vorlesung nennen wir eine Menge abzählbar, wenn sie endlich oder abzählbar unendlich ist. Diese Verwendung des Begriffes 'abzählbar' ist nicht einheitlich.

Wir werden nun das Prinzip der Wohlordnung benutzen: Jede nichtleere Teilmenge M von $\mathbb{N}$ hat ein kleinstes Element.

Anschaulich klar: Laufe $\mathbb{N}$ beginnend bei 1 ab, bis man auf die Menge trifft. Details hier: Sei M eine solche Teilmenge. Angenommen: M hat kein kleinstes Element. Dann ist $B := \{n \in \mathbb{N} : \{1, \ldots, n\} \in \mathbb{N} \setminus M\}$ induktiv ($1 \in B$: klar. $n \in B \Longrightarrow (n+1) \in B$: klar. Damit ist M leer.)

LEMMA. *Sei X eine Menge und $H : \mathbb{N} \longrightarrow X$ surjektiv. Dann ist entweder X endlich oder abzählbar unendlich.*

Beweis. Sei X nicht endlich. Zu zeigen: Es existiert ein $J : \mathbb{N} \longrightarrow X$ bijektiv.

Wir konstruieren induktiv ein $J : \mathbb{N} \longrightarrow X$ mit

- $\{J(1), J(2), \ldots, J(n)\} \supset \{H(1), \ldots, H(n)\}$.
- Die Elemente $J(1), \ldots, J(n)$ sind paarweise verschieden.

Aufgrund des ersten Punktes ist J surjektiv. Aufgrund des zweiten Punktes ist J injektiv.

Zur Konstruktion:

$n = 1$: $J(1) = H(1)$.

$n \Longrightarrow n+1$: Betrachte $M := \{k > n : H(k) \notin \{J(1), \ldots, J(n)\}\}$. Dann ist M nichtleer, da X unendlich ist. Damit hat M ein kleinstes Element m. Dann setzt man $J(n+1) := H(m)$. Dann gelten die gewünschten Eigenschaften: Nach Konstruktion ist $J(n+1)$ von $J(1), \ldots, J(n)$ verschieden. Weiterhin zeigt die Fallunterscheidung $m = n+1$ und $m \neq n+1$, daß $H(n+1)$ in $\{J(1), \ldots, J(n+1)\}$ enthalten ist. □

LEMMA. *Es sind $\mathbb{N} \times \mathbb{N}$ und $\mathbb{N} \times \{1, \ldots, n\}$ für beliebiges $n \in \mathbb{N}$ abzählbar. Insbesondere ist $X \times Y$ abzählbar für alle abzählbaren X und Y.*

Beweis. $\mathbb{N} \times \{1, \ldots, n\}$ *abzählbar:* Zeichnung.

$\mathbb{N} \times \mathbb{N}$ *abzählbar:* Zeichnung.

Das 'Insbesondere' kann nun auf folgende Art gezeigt werden: Sei $J_X : \mathbb{N} \longrightarrow X$ bijektiv und $J_Y : \mathbb{N} \longrightarrow Y$ bijektiv. Weiterhin existiert nach dem schon gezeigten ein bijektives

$$H : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}, n \mapsto H(n) = (H_1(n), H_2(n)).$$

Dann ist die Abbildung

$$J : \mathbb{N} \longrightarrow X \times Y, J(n) = (J_X(H_1(n)), J_Y(H_2(n)))$$

surjektiv (und sogar bijektiv). □

Bemerkung. Das Lemma liefert leicht, daß auch $X_1 \times \ldots \times X_n$ abzählbar ist für abzählbare $X_1, \ldots, X_n$. (Übung: Wie ist es mit abzählbaren Produkten bestellt?)

THEOREM. *Es ist $\mathbb{Q}$ abzählbar.*

Beweis. Betrachte

$$H : \mathbb{N}_0 \times \mathbb{N} \times \{-1, 1\} \longrightarrow \mathbb{Q}, \;\; H(n, m, q) = q\frac{n}{m}.$$

Dann ist H surjektiv und nach dem vorangehenden Lemma ist $\mathbb{N}_0 \times \mathbb{N} \times \{-1, 1\}$ abzählbar. Damit folgt die Aussage aus dem ersten Lemma des Abschnitts. $\square$

Theorem. *Es ist $\mathbb{R}$ überabzählbar. Tatsächlich ist jedes Interval positiver Länge in $\mathbb{R}$ überabzählbar.*

Beweis. Angenommen: $\mathbb{R}$ ist abzählbar.
Dann gibt es eine Abbildung $J : \mathbb{N} \longrightarrow \mathbb{R}$ mit $\mathbb{R} = \{J(1), J(2), \ldots, \}$. Wir konstruieren nun rekursiv eine Intervallschachtelung (I_n), $n \in \mathbb{N}$ mit

- $J(n) \notin I_n$ für alle $n \in \mathbb{N}$.
- $|I_{n+1}| = \frac{1}{3}|I_n|$.

Nach dem Intervallschachtelungsprinzip gibt es einen Punkt x der zu allen I_n gehört. Damit stimmt x dann mit keinem der $J(n)$ überein (da $J(n) \notin I_n$). Das ist ein Widerspruch.
Es bleibt die I_n zu konstruieren:
$n = 1$: Setze $I_1 := [J(1) + 1, J(1) + 2]$.
$n \Longrightarrow (n+1)$: Teile I_n in drei gleichlange abgeschlossene Teilintervalle. Es kann $J(n + 1)$ nicht in allen drei Teilintervallen liegen. Wähle für I_{n+1} ein Teilinterval, das $J(n + 1)$ nicht enthält. $\square$

Ende der 16. Vorlesung

Folgerung. *(Dichtheit von $\mathbb{R} \setminus \mathbb{Q}$) In jedem Interval positiver Länge in $\mathbb{R}$, gibt es Punkte von $\mathbb{R} \setminus \mathbb{Q}$.*

Beweis. Nach dem vorigen Satz ist jedes Interval positiver Länge überabzählbar. Da $\mathbb{Q}$ abzählbar ist, kann auch der Schnitt von $\mathbb{Q}$ mit einem solchen Interval nur abzählbar sein. Damit folgt die Behauptuung. $\square$

Bemerkung. Auch wenn ein Intervall positiver Länge 'fast nur' aus Punkten aus $\mathbb{R} \setminus \mathbb{Q}$ besteht, kann es schwierig sein, einen solchen Punkt explizit anzugeben.

Wir diskutieren nun, dass die Potenzmenge einer Menge immer 'echt größer' als die Menge ist.

Proposition. *Sei X eine beliebige Menge und $P(X)$ die Potenzmenge von X. Dann gibt es keine surjektive Abbildung J von X nach $P(X)$.*

Beweis. Übung. $\square$

Bemerkung. Die Proposition liefert insbesondere, daß es unterschiedlich 'große' unendliche Mengen gibt. Tatsaechlich ist die Anzahl der Teilmengen einer unendlichen Menge nach der Proposition immer echt großer als die Anzahl der Element der Menge. Inbesondere hat also $\mathbb{N}$ ueberabzaehlbar viele Teilmengen (vgl. Weihnachtszettel fuer einen direkten Beweis).

KAPITEL 6

Die komplexen Zahlen

Es geht um einen Erweiterungskörper von $\mathbb{R}$, in dem $x^2 + 1 = 0$ eine Lösung hat. Nimmt man die Existenz eines solchen Koerper an und nennt diese Lösung i, so gehoert mit $a, b \in \mathbb{R}$ dann auch $a + ib$ zu diesem Körper, und es gilt nach Distributivgesetz

$$(a + ib)(a' + ib') = aa' + iba' + iab' + iibb' = aa' - bb' + i(ba' + ab').$$

Damit ist also die Menge $a + ib$, $a, b \in \mathbb{R}$, unter obiger Multiplikation abgeschlossen. Das motiviert folgenden Satz.

Theorem. *Die Menge* $\mathbb{R} \times \mathbb{R}$ *versehen mit der Addition*

$$(a, b) + (c, d) = (a + c, b + d)$$

und der Multiplikation

$$(a, b)(c, d) = (ac - bd, ad + bc)$$

ist ein Körper mit neutralem Element $(0, 0)$ *der Addition und neutralem Element* $\underline{1} = (1, 0)$ *der Multiplikation.*
Das Inverse bzgl. Addition von (a, b) *ist gegeben durch* $(-a, -b)$*.*
Das Inverse bzgl. der Multiplikation von $(a, b) \neq (0, 0)$ *ist gegeben durch*

$$\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

Beweis. Direkte Rechnungen (vgl. Algebravorlesung). □

Zeichnung / Notation. Gaußsche Zahlenebene, imaginäre Achse, reelle Achse. Obere Halbebene, untere Halbebene.

Definition. *Der Körper aus dem vorangehenden Satz wird mit* $\mathbb{C}$ *bezeichnet.*

Notation. Wir schreiben i für $(0, 1)$. Außerdem identifizieren wir ein $c \in \mathbb{R}$ mit dem Element der Form $(c, 0) \in \mathbb{C}$. Damit kann dann $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ aufgefaßt werden. Dann gilt

$$c(a, b) = (c, 0)(a, b) = (ca - 0b, cb + 0a) = (ca, cb).$$

Insbesondere gilt also mit $\underline{1} = (1, 0)$

$$ci = (c, 0)i = (c, 0)(0, 1) = (0, c) \text{ und } c\underline{1} = (c, 0)(1, 0) = (c, 0).$$

Damit lässt sich dann das Element $(a, b) \in \mathbb{C}$ schreiben als

$$(a, b) = (a, 0) + (0, b) = (a, 0)\underline{1} + (b, 0)i = a + ib.$$

Das ist gerade die zu Anfang des Kapitels gegebene Darstellung.

Bemerkung. Es ist $\mathbb{C}$ ein Vektorraum ueber $\mathbb{R}$ (der Dimension 2) mit skalarer Multiplikation $\lambda \cdot (a, b) := (\lambda a, \lambda b)(= (\lambda, 0)(a, b))$. Es ist $\{\underline{1}, i\}$ ist eine Basis.

PROPOSITION. *Es gilt* $i^2 = -1$.

Beweis. Das folgt direkt durch Nachrechnen:

$$i^2 = (0, 1)(0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cot 0) = (-1, 0) = -1.$$

Hier nutzen wir im letzten Schritt die schon besprochene Identifikation von $\mathbb{R}$ mit einer Teilmenge von $\mathbb{C}$. $\square$

DEFINITION. *Für* $z = a+ib \in \mathbb{C}$ *definieren wir den Imaginärteil* $\Im z := b$ *und den Realteil* $\Re z := a$, *sowie die zu* z *konjugiert komplexe Zahl* $\overline{z} := a - ib$ *und* $|z| := \sqrt{z \cdot \overline{z}} = \sqrt{a^2 + b^2}$.

Beachte. (a) Komplex Konjugieren bedeutet gerade Spiegeln an der reellen Achse.
(b) Es ist $|z|$ die *Laenge* von z. Es gilt $|z|^2 = z\overline{z}$.

Folgende Regeln sind einfach zu beweisen.

PROPOSITION. *Für alle* $z, w \in \mathbb{C}$ *gilt:*
(a) $\overline{z + w} = \overline{z} + \overline{w}$.
(b) $\overline{zw} = \overline{z}\,\overline{w}$ *und* $\overline{\frac{1}{z}} = \frac{1}{\overline{z}}$ *für* $z \neq 0$.
(c) $\Re z = \frac{1}{2}(z + \overline{z})$, $\Im z = \frac{1}{2i}(z - \overline{z})$.
(d) $z \neq 0 \Longrightarrow z^{-1} = \frac{\overline{z}}{|z|^2}$.
(e) $|zw| = |z||w|$ *und für* $z \neq 0$ $|1/z| = 1/|z|$.

Beweis. Hausaufgabe. $\square$

Bemerkung. Aus (d) ergibt sich ein allgemeines Verfahren zum Umgang mit Bruechen der Form A/B mit komplexen Zahlen A, B: Durch Erweitern mit $\overline{B}$ ergibt sich der Ausdruck $A\overline{B}/|B|^2$, der keine komplexen Zahlen mehr im Nenner enthaelt.

PROPOSITION. *(Dreiecksungleichung) Für* $z, w \in \mathbb{C}$ *gilt*

$$|z + w| \leq |z| + |w|.$$

Beweis. Sei $z = a + ib$ und $w = c + id$.
Zwischenbehauptung. Es gilt $|ac + bd| \leq |z||w|$.
Bew. Ohne Einschraenkung $|z|, |w| \neq 0$. Reicht nun zu zeigen

$$|\frac{a}{|z|}\frac{c}{|w|} + \frac{b}{|z|}\frac{d}{|w|}| \leq 1.$$

Das folgt leicht unter Anwenden von $\alpha\beta \leq \frac{1}{2}(\alpha^2 + \beta^2)$ auf das erste und zweite Produkt im Betrag.

Mit der Zwischenbehauptung und einer kleinen Rechnung ergibt sich nun folgendes:

$$\begin{aligned}
|z+w|^2 &= (a+c)^2 + (b+d)^2 \\
&= a^2 + b^2 + 2(ac+bd) + c^2 + d^2 \\
&\leq a^2 + b^2 + 2|ac+bd| + c^2 + d^2) \\
&\leq a^2 + b^2 + 2|z||w| + c^2 + d^2 \\
&= |z|^2 + 2|z||w| + |w|^2 \\
&= (|z| + |w|)^2.
\end{aligned}$$

Nach Ziehen der Wurzel folgt die gewuenschte Beziehung. □

Zeichnung / Bemerkung. In $\mathbb{C}$ kann man Dreiecksungleichung gut deuten.

Der Betrag erlaubt es uns ähnlich wie in $\mathbb{R}$ das Konzept der Konvergenz und der Cauchy Folge zu definieren. Diesem Thema widmen wir uns als nächstes.

Definition. *Eine Folge (z_n) in $\mathbb{C}$ heißt konvergent gegen $z \in \mathbb{C}$, wenn für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ existiert, sodaß für alle $n \geq n_\varepsilon$ gilt $|z_n - z| < \varepsilon$. Es heißt dann z Grenzwert der Folge (z_n).*

Wie schon in $\mathbb{R}$ kann auch in $\mathbb{C}$ eine Folge höchstens gegen einen Wert konvergieren und dieser Wert heißt dann DER Grenzwert und wir schreiben $z = \lim z_n$ oder $z_n \to z$, $n \to \infty$.

Man definiert für $\varepsilon > 0$ die *ε-Umgebung von w* durch

$$U_\varepsilon(w) := \{z \in \mathbb{C} : |z - w| < \varepsilon\}.$$

Es heißt dann $U_\varepsilon(w)$ auch die *offene ε-Kugel* um w.

Allgemein nennt man eine Menge U eine *Umgebung von w*, wenn ein $\varepsilon > 0$ existiert mit $U_\varepsilon(w) \subset U$.

Lemma. *Für eine Folge (z_n) in $\mathbb{C}$ und $z \in \mathbb{C}$ sind äquivalent:*

(i) *Es konvergiert (z_n) gegen z.*
(ii) *Für alle $\varepsilon > 0$ existiert ein $n_\varepsilon > 0$ mit $z_n \in U_\varepsilon(z)$ für alle $n \geq n_\varepsilon$.*
(iii) *$|z - z_n| \to 0$, $n \to \infty$.*

Beweis. Es bedeutet $z_n \in U_\varepsilon(z)$ gerade $|z_n - z| < \varepsilon$. Damit folgt sofort die Aequivalenz von (i) und (ii).

Es bedeutet $|z_n - z| \to 0$ gerade, daß fuer alle $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ existiert mit

$$||z_n - z| - 0| < \varepsilon$$

fuer alle $n \geq n_\varepsilon$. Mit

$$||z_n - z| - 0| = |z_n - z|$$

folgt dann die Aequivalenz von (iii) und (i). $\square$

Konvergenz in $\mathbb{C}$ und Konvergenz in $\mathbb{R}$ haben viel miteinander zu tun. Tatsächlich lassen sich wesentliche Betrachtungen zu Konvergenz in $\mathbb{C}$ auf die entsprechenden Betrachtungen in $\mathbb{R}$ zurückführen. Dazu dient folgende Proposition.

Proposition (Abschaetzung Real- Imaginaerteil via Betrag). *Für $z = a + ib \in \mathbb{C}$ gilt*

$$|z| \leq |a| + |b| \ \textit{ sowie } |a|, |b| \leq |z|.$$

Beweis. Nach Definition gilt $|z| = \sqrt{a^2 + b^2}$.
Erste Ungleichung: Es gilt

$$a^2 + b^2 = |a|^2 + |b|^2 \leq |a|^2 + 2|a||b| + |b|^2.$$

Damit folgt also

$$|z|^2 \leq (|a| + |b|)^2.$$

Da die Wurzelfunktion monoton ist ($x < y \Longrightarrow \sqrt{x} \leq \sqrt{y}$) folgt damit die Behauptung durch Wurzelziehen.
Zweite Ungleichung: Mit $a^2, b^2 \leq a^2 + b^2$ folgt wieder aufgrund der Monotonie der Wurzelfunktion

$$|a|, |b| \leq |z|.$$

Das beendet den Beweis. $\square$

Als erste Folgerung zeigen wir:

Proposition. *Sei (z_n) eine Folge in $\mathbb{C}$ und $z_n = a_n + ib_n$ mit $a_n, b_n \in \mathbb{R}$ d.h. $a_n = \Re z_n$ und $b_n = \Im z_n$. Dann sind äquivalent:*

(i) *Die Folge (z_n) konvergiert in $\mathbb{C}$.*
(ii) *Die Folgen (a_n) und (b_n) konvergieren in $\mathbb{R}$.*

In diesem Fall gilt $\lim z_n = \lim a_n + i(\lim b_n)$ d.h.

$$\Re(\lim_n z_n) = \lim_n \Re(z_n) \ \textit{ und } \ \Im(\lim_n z_n) = \lim_n(\Im z_n).$$

Beweis. (i)$\Longrightarrow$ (ii): Es gelte $z_n \to z = a + ib$. Sei $\varepsilon > 0$. Dann existiert also ein $n_\varepsilon \in \mathbb{N}$ mit $|z_n - z| < \varepsilon$ für $n \geq n_\varepsilon$. Dann gilt für $n \geq n_\varepsilon$ also nach voriger Proposition

$$|a_n - a|, |b_n - b| \leq |z_n - z| < \varepsilon.$$

(ii) $\Longrightarrow$ (i). $a_n \to a$, $b_n \to b$. Sei $z := a + ib$. Sei $\varepsilon > 0$. Dann existiert also ein $n_a \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}$ für $n \geq n_a$ und ein $n_b \in \mathbb{N}$ mit $|b_n - b| < \frac{\varepsilon}{2}$ für $n \geq n_b$. Für $n \geq \max\{n_a, n_b\}$ gilt also nach voriger Proposition

$$|z - z_n| = |a + ib - (a_n + ib_n)| \leq |a - a_n| + |b - b_n| < \varepsilon.$$

Zur letzten Aussage: Das wurde beim Beweis von (i)$\Longrightarrow$ (ii) schon mitgezeigt. $\square$

Ende der 17. Vorlesung

Damit kann man aus den Rechenregeln für Konvergenz in $\mathbb{R}$ leicht die folgenden Rechenregeln für Konvergenz in $\mathbb{C}$ ableiten.

PROPOSITION. *(Rechenregeln)*
(a) $z_n \to z$, $w_n \to w \Longrightarrow z_n + w_n \to z + w$.
(b) $z_n \to z$, $w_n \to w \Longrightarrow z_n w_n \to zw$.
(c) $z_n \to z$, $z \neq 0$, $z_n \neq 0$ *alle* $n \Longrightarrow 1/z_n \to 1/z$.
(d) $z_n \to z \Longrightarrow |z_n| \to |z|$.

Ähnlich wie in $\mathbb{R}$ definiert man folgende Konzepte.

DEFINITION (Cauchy Folge). *Eine Folge* (z_n) *in* $\mathbb{C}$ *heißt Cauchy Folge, wenn zu jedem* $\varepsilon > 0$ *ein* $n_\varepsilon \in \mathbb{N}$ *existiert, so daß für alle* $n, m \geq n_\varepsilon$ *gilt* $|x_n - x_m| < \varepsilon$.

THEOREM. *Für eine Folge* (x_n) *in* $\mathbb{C}$ *sind äquivalent:*
(i) (z_n) *ist eine Cauchy Folge.*
(ii) (z_n) *ist konvergent.*

Beweis. Sei $z_n = a_n + ib_n$ mit $a_n, b_n \in \mathbb{R}$. Dann gilt: (z_n) Cauchy Folge $\Longleftrightarrow (a_n), (b_n)$ Cauchy Folgen in $\mathbb{R} \Longleftrightarrow (a_n), (b_n)$ konvergent in $\mathbb{R} \Longleftrightarrow z_n$ konvergent. (Dabei verwenden wir im ersten Schritt die obige Proposition zur Abschaetzung von Real- und Imaginaerteil via Betrag, im zweiten Schritt die Vollstaendigkeit von $\mathbb{R}$ und im dritten Schritt die vorangegangene Proposition.) $\square$

FOLGERUNG. *Es ist* $\mathbb{C}$ *vollständig d.h. jede Cauchy-Folge in* $\mathbb{C}$ *konvergiert.*

Bemerkung. Die Vollständigkeit ist eine fundamentale analytische Eigenschaft der komplexen Zahlen.
Da es in $\mathbb{C}$ keine Anordnung gibt (Warum? $x^2 + 1 = 0$ hat Lösung!), gibt es auch keine monotonen Folgen und also auch keine Konvergenz monotoner Folgen. Aber es gibt eine komplexe Version des Satzes von Bolzano-Weierstraß. Dazu führen wir noch folgenden Begriff ein: Eine Folge $(z_n)_n$ in $\mathbb{C}$ heißt beschränkt, wenn ein $C > 0$ existiert mit $|z_n| \leq C$ für alle $n \in \mathbb{N}$.

THEOREM. *(Bolzano - Weierstraß - komplex). Sei* (z_n) *eine beschränkte Folge in* $\mathbb{C}$*. Dann hat* (z_n) *eine konvergente Teilfolge.*

Beweis. Sei $z_n = a_n + ib_n$ mit $a_n, b_n \in \mathbb{R}$. Da (z_n) beschränkt ist, sind auch die Folgen (a_n) und (b_n) beschränkt (nach der Proposition zur Abschaetzung von Real- und Imaginaerteil via Betrag). Da (a_n) beschränkt ist, gibt es nach der reellen Version des Satz von Bolzano

- Weierstraß eine konvergente Teilfolge $(a_{n_k^{(1)}})_k$. Dann ist aber auch $(b_{n_k^{(1)}})$ beschränkt und hat also eine konvergente Teilfolge (b_{n_k}). Dann konvergiert sowohl (a_{n_k}) also auch (b_{n_k}). Damit konvergiert dann auch (z_{n_k}). $\square$

KAPITEL 7

Summen und Reihen

Reihen liefern (eigentlich nur) eine spezielle Art, Folgen darzustellen. Diese Art tritt in vielerlei Zusammenhängen auf und ist entsprechend von besonderem Interesse.

Ziel: Gegeben eine Folge (a_n) in $\mathbb{C}$. Definiere

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots .$$

Problem: Das Problem sind wieder die '...' oder anders gesagt, die Tatsache, dass es unendlich viele Summanden gibt.

Lösung. Summiere über endlich viele Summanden und bilde (falls moeglich) den Grenzwert:

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

Zeichnung.
S_1: a_1
S_2: $a_1 + a_2$
S_3: $a_1 + a_2 + a_3$
S_4: $a_1 + a_2 + a_3 + a_4$
....

Definition (Reihe gleich Folge der Partialsummen). *Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{C}$. Zu $n \in \mathbb{N}$ ist dann die n-te Partialsumme der (a_n) definiert durch*

$$S_n := \sum_{k=1}^{n} a_k.$$

Die Folge (S_n) dieser Partialsummen wird dann als Reihe mit den Gliedern a_n bezeichnet. Diese Reihe heißt konvergent (mit Grenzwert S), wenn die Folge (S_n) konvergiert (mit Grenzwert S).

Bemerkung. Voellig entsprechend definiert man Reihen ueber Folgen, deren Indexmenge nicht $\mathbb{N}$ sonderen $\mathbb{N} + N$ fuer ein $N \in \mathbb{Z}$ ist.

Notation. Wir schreiben

$$\sum_{k\geq 1} a_k \quad \text{für die Reihe, d.h. die Folge der Partialsummen}$$

und

$$\sum_{k=1}^{\infty} a_k \quad \text{für den Grenzwert der Reihe, d.h. } \lim_{n\to\infty} S_n$$

(falls dieser existiert).

Bemerkung (Reihen entsprechen Folgen). Sei (x_n) eine beliebige Folge in $\mathbb{C}$. Dann ist $x_n = \sum_{k=1}^{n} a_k$ für $a_1 = x_1$, $a_n := x_n - x_{n-1}$ $n \geq 2$. In diesem Sinne lässt sich jede Folge (in $\mathbb{C}$) als Reihe darstellen.

Beispiel. (Geometrische Reihe) Ist $\alpha \in \mathbb{C}$ beliebig und $q \in \mathbb{C}$ mit $|q| < 1$ so ist $\sum_{k\geq 0} \alpha q^k$ konvergent gegen $\frac{\alpha}{1-q}$.
(Bew. Es gilt $|q^n| = |q|^n \to 0$. Damit folgt die Aussage aus der Formel für die geometrische Summe.)
Anwendung. Sei $(a_k)_{k\geq N}$ eine beliebige Folge mit $a_{k+1}/a_k = q$ mit $|q| < 1$. Dann gilt

$$\sum_{k=N}^{\infty} a_k = a_N \frac{1}{1-q}.$$

Insbesondere gilt also $\sum_{k=N}^{\infty} \alpha\beta^{k+l} = \alpha\beta^{N+l} \sum_{k=0}^{\infty} \beta^k = \alpha\beta^{N+1} \frac{1}{1-\beta}$ für $|\beta| < 1$.

Beispiel. $\sum_{k\geq 2} \frac{1}{k(k-1)}$ konvergiert gegen 1.
Bew. $S_n = \sum_{k=2}^{n} \frac{1}{k(k-1)} = \sum_{k=2}^{n} (\frac{1}{k-1} - \frac{1}{k}) = 1 - \frac{1}{n} \to 1$. Die in der zweiten Gleichung genutzte Technik ist unter dem Namen Partialbruchzerlegung bekannt (s. später).

Da es sich um Folgen handelt, gelten für konvergente Reihen natürlich weiterhin die Rechenregeln für konvergente Folgen. Insbesondere gilt folgendes.

Proposition. *Sind $\sum_{k\geq 1} a_k$ und $\sum_{k\geq 1} b_k$ konvergente Reihen, so ist fuer jedes $\lambda \in \mathbb{C}$ auch die Reihe $\sum_{k\geq 1}(a_k + \lambda b_k)$ konvergent und es gilt*

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

Die Grundfrage im Umgang mit Reihen betrifft natuerlich die Konvergenz der Reihe. Das werden wir nun untersuchen.

Theorem. *(Charakterisierung Konvergenz - Cauchy Kriterium) Sei (a_k) in $\mathbb{C}$ gegeben. Dann sind äquivalent:*

(i) *Die Reihe $\sum_{k\geq 1} a_k$ konvergiert.*
(ii) *Zu jedem $\varepsilon > 0$ existiert $n_\varepsilon \in \mathbb{N}$ mit $|\sum_{k=m+1}^{n} a_k| < \varepsilon$ für alle $n_\varepsilon \leq m < n$. (**Zeichnung.** Endstück der Reihe.)*

Bemerkung. Es ist $|\sum_{k=m+1}^{n} a_k| = |S_n - S_m|$ gerade die Summe über ein 'Endstück' der Reihe.

Beweis. Nach Definition bedeutet Konvergenz einer Reihe (d.h. (i)) gerade die Konvergenz der Folge der Partialsummen. Aufgrund der Vollstaendigkeit von $\mathbb{C}$ ist diese Konvergenz aequvalent dazu, daß die Partialsummen eine Cauchy-Folge bilden. Das ist aber gerade Bedingung (ii). □

Das Theorem liefert eine **notwendige** Bedingung für Konvergenz.

FOLGERUNG (Erster Test auf Konvergenz). *Ist $\sum_{k\geq 0} a_k$ konvergent, so ist $|a_k|$ eine Nullfolge (d.h. $|a_k| \to 0$, $k \to \infty$).*

Beweis. $|a_k| = |\sum_{l=k}^{k} a_l| = |S_k - S_{k-1}| \to 0, k \to \infty$. □

Ende der 18. Vorlesung

Diese Bedingung ist nicht hinreichend:

Gegenbeispiel.(Harmonische Reihe ist divergent) $\sum_{k\geq 1} 1/k$ ist divergent (**obwohl** $\frac{1}{n} \to 0$).
Bew. $1+1/2+(1/3+1/4)+(1/5+\ldots+1/8)+\ldots+(\frac{1}{2^n+1}+\ldots+\frac{1}{2^{n+1}})\ldots.$

Besonders wichtig sind Reihen mit nichtnegativen Gliedern. Denn darauf lassen sich viele Konvergenzbetrachtungen zurückführen. Für diese Reihen gilt folgendes Lemma.

LEMMA (Konvergenz von Reihen mit nichtnegativen Gliedern). *Sei Folge (a_k) mit $a_k \geq 0$ gegeben. Sei $S_n := \sum_{k=1}^{n} a_k$. Dann sind äquivalent:*

(i) *Die Reihe $\sum_{k\geq 1} a_k$ ist konvergent (d.h. (S_n) konvergent).*
(ii) *Die Reihe $\sum_{k\geq 1} a_k$ ist beschränkt (d.h. $\sup S_n < \infty$).*
(iii) *Für alle $\varepsilon > 0$ existiert ein $n_\varepsilon \in \mathbb{N}$ mit*

$$a_n + a_{n+1} + \ldots + a_m < \varepsilon$$

für alle $n, m \in \mathbb{N}$ mit $n_\varepsilon \leq n < m$.

Beweis. Wegen $a_k \geq 0$ ist (S_n) monoton wachsend. Damit sind Konvergenz (i) und Beschränktheit (ii) äquivalent. Weiterhin ist Konvergenz in $\mathbb{R}$ äquivalent dazu, daß (S_n) eine Cauchy Folge ist. Das bedeutet aber gerade (iii) (nach dem Cauchy-Kriterium). □

Notation. Ist $a_k \geq 0$, so schreiben wir $\sum_{k=1}^{\infty} a_k < \infty$, wenn eine der Bedingungen des vorigen Lemma gilt. Andernfalls schreiben wir $\sum_{k=1}^{\infty} a_k = \infty$.

Für Reihen erweist sich eine Verschärfung des Begriff der Konvergenz als sinnvoll. Dies ist der Punkt, an dem sich die Theorie der Reihen von der Theorie der Folgen unterscheidet.

Definition. *Sei (a_k) eine Folge in $\mathbb{C}$. Dann heißt die Reihe $\sum_{k\geq 1} a_k$ absolut konvergent, wenn $\sum_{k\geq 1} |a_k|$ konvergiert, d.h. wenn gilt*

$$\sum_{k=1}^{\infty} |a_k| < \infty.$$

Bemerkung.

- Für Reihen mit nichtnegativen Gliedern ist Konvergenz gleichbedeutend mit absoluter Konvergenz.
- Absolute Konvergenz / Konvergenz ändert sich nicht, wenn man endlich viele Glieder der Reihe abändert. (Es geht immer nur um a_k mit großen k.)

Beispiel. (Geometrische Reihe) Ist $q \in \mathbb{C}$ mit $|q| < 1$ und $\alpha \in \mathbb{C}$, so ist $\sum_{k\geq 1} \alpha q^k$ absolut konvergent (gegen $\frac{\alpha}{1-q}$). Ist $|q| \geq 1$, so ist die Reihe nicht konvergent.
Bew. Es gilt $|\alpha q^n| = |\alpha||q|^n$. Damit folgt absolute Konvergenz der Reihe für $|q| < 1$ (s.o.). Ist $|q| \geq 1$, so ist q^n keine Nullfolge und es folgt Divergenz.

Theorem (Absolute Konvergenz impliziert Konvergenz). *Sei (a_k) eine Folge in $\mathbb{C}$. Ist $\sum_{k\geq 0} a_k$ absolut konvergent, so ist $\sum_{k\geq 0} a_k$ konvergent.*

Beweis. Das folgt eigentlich sofort aus dem Cauchy Kriterium. Hier sind die Details: Sei $S_n := \sum_{k=1}^{n} a_k$. Zu zeigen: (S_n) ist Cauchy Folge. Sei $\varepsilon > 0$ beliebig. Aufgrund der absoluten Konvergenz, konvergiert $\sum_{k\geq 1} |a_k|$. Damit existiert also nach dem Lemma ein $n_\varepsilon \in \mathbb{N}$ mit

$$|a_n| + |a_{n+1}| + \ldots + |a_m| < \varepsilon$$

für alle $n, m \geq n_\varepsilon$. Damit gilt also für $n, m \geq n_\varepsilon$ mit $n < m$

$$|a_n + a_{n+1} + \ldots a_m| \leq |a_n| + |a_{n+1}| + \ldots + |a_m| < \varepsilon.$$

Damit konvergiert die Reihe nach dem Cauchy-Kriterium. □

Bemerkung. Es ist absolute Konvergenz tatsaechlich strenger als Konvergenz. Denn es gibt Reihen, die konverigeren aber nicht absolut konvergieren. Ein Beispiel ist gegeben durch die Reihe ueber $a_k := (-1)^{k+1}\frac{1}{k}$. Dann gilt:

- $\sum |a_k|$ nicht konvergent (harmonische Reihe).
- Es ist aber $\sum_{k\geq 1} a_k$ konvergent (nach Leibniz Kriterium, s.u.)

Bemerkung. (Absolute Konvergenz stabil bei Umordnungen) Absolute Konvergenz ist stabil unter Umordnungen (s.u.). Konvergente aber nicht absolut konvergente Reihen sind extrem instabil unter Umordnung (s.u.). Daher ist absolute Konvergenz oft wesentlich nützlicher als Konvergenz, wenn es um Reihen geht.

Proposition (Dreieckungleichung für Reihen). *. Sei (a_k) eine Folge in $\mathbb{C}$. Existiert $\sum a_k$, so gilt $|\sum_{k=1}^{\infty} a_k| \leq \sum_{k=1}^{\infty} |a_k|$, wobei die rechte Seite den Wert unendlich hat, wenn die Summe nicht absolut konvergiert.*

Beweis. Es reicht den Fall zu betrachten, daß $\sum a_k$ absolut konvergiert. Da Betrag mit Konvergenz von Folgen verträglich ist, gilt

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

Hier verwenden wir die Dreiecksungleichung fuer endliche Summen im vorletzten Schritt. □

Theorem (Majorantenkriterium). *Sei (a_k) eine Folge in $\mathbb{C}$. Gibt es $k_0 \in \mathbb{N}$ und $b_k \geq 0$ mit*

- *$|a_k| \leq b_k$ für $k \geq k_0$ und*
- *$\sum_{k\geq k_0} b_k < \infty$,*

so ist $\sum a_k$ absolut konvergent.

Beweis. Nach dem Lemma zur Konvergenz von Reihen mit nichtnegativen Gliedern reicht es

$$\sup \sum_{k=1}^{n} |a_k| < \infty$$

zu zeigen. Nach der zweiten Voraussetzung gibt es $C \geq 0$ mit $\sum_{k=k_0}^{n} b_k \leq C$ für all $n \geq k_0$. Damit folgt fuer $n \geq k_0$ dann

$$\sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{k_0-1} |a_k| + \sum_{k=k_0}^{n} |a_k| \leq \sum_{k=1}^{k_0-1} |a_k| + \sum_{k=k_0}^{n} b_k \leq \sum_{k=1}^{k_0-1} |a_k| + C.$$

Fuer $n \leq k_0$ gilt natuerlich sowieso

$$\sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{k_0-1} |a_k| + C.$$

Damit ergibt sich die gewuenschte Beschraenktheit. □

Beispiel. (Dezimalzahldarstellung) Ist (a_n) eine Folge mit Werten in $\{0, 1, 2, \ldots, 9\}$, so konvergiert

$$\sum_{k\geq 1} a_k 10^{-k}$$

absolut.
Bew. Sei $b_k := 9/10^k$. Dann gilt $|a_k 10^{-k}| \leq b_k$ und $\sum b_k$ existiert (da geometrische Reihe).

Weitere Beispiele.

- $\sum_{k\geq 1} \frac{1}{k^2}$ konvergiert absolut.
  Bew. $0 \leq 1/k^2 \leq 1/k(k-1)$ und $\sum 1/k(k-1)$ konvergent. Nun folgt Beh. aus Majorantenkriterium.

- $\sum_{k\geq 1}\frac{n!}{n^n}$ konvergiert absolut.

  Bew. $\frac{n!}{n^n}=\frac{12...n}{nn...n}\leq 2/n^2$ und $\sum\frac{1}{n}^2$ konvergiert. Nun folgt Beh. aus Majorantenkriterium.

Man kann das Majorantenkriterium auch umdrehen.

Folgerung. *(Minorantenkriterium) Sei (a_k) eine Folge in $\mathbb{C}$ und $b_k\geq 0$, $k\in\mathbb{N}$. Gilt $b_k\geq |a_k|$ für alle $k\in\mathbb{N}$ ab einem k_0 und ist $\sum a_k$ divergent, so ist auch $\sum b_k$ divergent.*

Beweis. Angenommen $\sum b_k$ konvergent. Dann folgt aus dem Majorantenkriterium die absolute Konvergenz von $\sum |a_k|$ absolut konvergent. Damit folgt nach dem Theorem die Konvergenz von $\sum a_k$. Das ist ein Widerspruch. □

Beispiele.

- $\sum\frac{1}{\sqrt{n(n+1)}}$ ist divergent.

  Bew. $\frac{1}{\sqrt{n(n+1)}}\geq\frac{1}{\sqrt{(n+1)(n+1)}}=\frac{1}{n+1}$ und $\sum\frac{1}{n+1}$ divergent (harmonische Reihe).
- Sei
$$a_k := (\text{k-te ungerade natürliche Zahl})^{-1}=\frac{1}{2k-1}$$
  und
$$b_k := (\text{k-te gerade natürliche Zahl})^{-1}=\frac{1}{2k}.$$
  Dann ist $\sum a_k$ und $\sum b_k$ divergent.

  Bew. Wäre eine der beiden Summen konvergent, so müsste auch die andere konvergieren nach dem Majorantenkriterium. Dann wäre aber auch die harmonische Reihe konvergent.

Theorem (Quotientenkriterium). *Sei (a_k) eine Folge in $\mathbb{C}$ mit $a_k\neq 0$ für alle k ab einem k_0. Gilt $q:=\limsup_{k\to\infty}\frac{|a_{k+1}|}{|a_k|}<1$, so konvergiert $\sum a_k$ absolut. Gilt $p:=\liminf_{k\to\infty}\frac{|a_{k+1}|}{|a_k|}>1$, so divergiert $\sum a_k$.*

Beweis. $q<1$: Sei $\widetilde{q}$ eine Zahl mit $q<\widetilde{q}<1$. (Zeichnung.) Wegen $q=\limsup_{k\to\infty}\frac{|a_{k+1}|}{|a_k|}$ gibt es $N\in\mathbb{N}$ mit
$$\frac{|a_{k+1}|}{|a_k|}\leq\widetilde{q}$$
für alle $k\geq N$. Das impliziert
$$|a_n|\leq\widetilde{q}|a_{n-1}|\leq\cdots\leq\widetilde{q}^{n-N}|a_N|$$
für alle $n\geq N$. Weiterhin ist $\sum_{n\geq N}|a_N|\widetilde{q}^{n-N}$ konvergent (als geometrische Reihe). Damit folgt nach dem Majorantenkriterium die absolute Konvergenz der ursprünglichen Reihe.

$p > 1$: Ähnlich wie im vorigen Fall schließt man, daß ein $N \in \mathbb{N}$ existiert mit

$$|a_{k+1}|/|a_k| \geq 1$$

für alle $k \geq N$. Damit folgt

$$|a_k| \geq |a_{k-1}| \geq \ldots \geq |a_N| > 0$$

für alle $k \geq N$. Also ist $|a_k|$ keine Nullfolge. Damit konvergiert die Reihe nicht. □

Ende der 19. Vorlesung

Bemerkung. Für $q = 1$ bzw. $p = 1$ ist jedes Konvergenzverhalten möglich.
- $a_k = \frac{1}{k}$. Dann $q = p = \lim \frac{1/(k+1)}{1/k} = \lim \frac{k}{k+1} = 1$ und $\sum a_k$ divergent.
- $a_k = \frac{1}{k^2}$. Dann $q = p = \lim \frac{1/(k+1)^2}{1/k^2} = \lim \frac{k^2}{(k+1)^2} = 1$ und $\sum a_k$ konvergent.

THEOREM (Wurzelkriterium). *Sei (a_k) eine Folge in $\mathbb{C}$ und $q := \limsup_{k\to\infty} \sqrt[k]{|a_k|}$. Gilt $q < 1$, so ist $\sum a_k$ absolut konvergent. Gilt $q > 1$ so ist $\sum a_k$ divergent.*

Beweis. $q < 1$: Sei $\widetilde{q}$ mit $q < \widetilde{q} < 1$ und $N \in \mathbb{N}$ mit

$$\sqrt[k]{|a_k|} \leq \widetilde{q}$$

für alle $k \geq N$. Damit gilt also

$$|a_k| \leq \widetilde{q}^k$$

für alle $k \geq N$. Die Reihe $\sum_{k \geq N} \widetilde{q}^k$ konvergiert (geometrische Reihe). Damit folgt aus dem Majorantenkriterium die absolute Konvergenz von $\sum a_k$.
$q > 1$: Es gibt unendlich viele k mit $\sqrt[k]{|a_k|} \geq 1$, also $|a_k| \geq 1$. Damit ist $|a_k|$ keine Nullfolge.

□

Bemerkung. Für $q = 1$ ist jedes Konvergenzverhalten möglich.
- $a_k = 1/k$. Dann $q = \lim \sqrt[k]{1/k} = 1$ und $\sum a_k$ divergent.
- $a_k = 1/k^2$. Dann $q = \lim \sqrt[k]{1/k^2} = 1$ und $\sum a_k$ konvergent.

Beispiel - Exponentialreihe. Sei $z \in \mathbb{C}$ beliebig und $a_k := z^k/k!$. Dann ist die Exponentialreihe

$$\sum_{k \geq 0} a_k = \sum_{k \geq 0} \frac{z^k}{k!}$$

absolut konvergent und der Grenzwert wird als e^z bezeichnet.
Bew. Das kann man mit Quotientenkriterium oder Wurzelkriterium zeigen. Wir geben (zur Sicherheit ;-) beide Beweise.

Quotientenkriterium: Aufgrund von

$$\frac{|a_{k+1}|}{|a_k|} = \frac{|z^{k+1}/(k+1)!|}{|z^k/k!|} = \frac{|z|}{|k+1|} \to 0 < 1$$

ist die Voraussetzung des Quotientenkriterium erfuellt.

Wurzelkriterium: Aufgrund von

$$\sqrt[k]{\left|\frac{z^k}{k!}\right|} = \frac{|z|}{\sqrt[k]{k!}} \to 0 < 1$$

ist die Voraussetzung des Wurzelkriterium erfuellt. (Hier nutzen wir im letzten Schritt die uns schon bekannte bestimmte Divergenz $\sqrt[k]{k!} \to \infty, k \to \infty$).

Bemerkung. Es ist kein Zufall, daß im vorigen Beispiel sowohl Quotienten- als auch Wurzelkriterium zum Ziel fuehren. Tatsaechlich ist das Wurzelkriterium echt stärker als das Quotientenkriterium d.h.

- wenn das Quotientenkriterium Konvergenz liefert, dann tut dies auch das Wurzelkriterium,
- es gibt Faelle, in denen das Wurzelkriterium Konvergenz liefert aber nicht das Quotientenkriterium.

Allerdings ist das Quotientenkriterium (oft) leichter anzuwenden als das Wurzelkriterium.

Hier skizzieren wir noch kurz Begruendungen zu den beiden Aussagen ueber die Wurzelkriterium:

- Ist $b_n > 0$, so gilt $\limsup \sqrt[n]{b_n} \leq \limsup \frac{b_{n+1}}{b_n} = q$. (siehe Weihnachtszettel). Idee :

$$\sqrt[n]{b_n} = \sqrt[n]{b_1 \frac{b_2}{b_1} \cdots \frac{b_n}{b_{n-1}}} \approx \sqrt[n]{\frac{b_{N+1}}{b_N} \cdots \frac{b_n}{b_{n-1}}} \preceq \sqrt[n]{q^{n-N}} \approx q.$$

- Es gibt Reihen, deren Konvergenz aus dem Wurzelkriterium folgt aber nicht aus dem Quotientenkriterium. Ein Beispiel ist folgendes:

$$a_k := \begin{cases} 2^{-k} & : \text{ k gerade} \\ 3^{-k} & : \text{ k ungerade} \end{cases}$$

Dann ist $\sqrt[k]{a_k}$ gleich 1/2 oder 1/3, also $\limsup \sqrt[k]{a}_k = 1/2 < 1$. Aber es gilt

$$\frac{a_{k+1}}{a_k} = \begin{cases} \frac{2^{k-1}}{3^{k+1}} & : \text{ k gerade} \\ \frac{3^k}{2^{k+1}} & : \text{ k ungerade} \end{cases}$$

Damit wird dann also $\frac{a_{k+1}}{a_k}$ fuer grosse ungerade k beliebig gross.

Beispiele. Sei $a_n = (1+\frac{1}{n})^{n^2}$. Dann ist die Reihe nicht konvergent, da (a_n) keine Nullfolge ist. Sei $b_n = \frac{1}{(1+\frac{1}{n})^{n^2}}$. Dann folgt aus dem Wurzelkriterium die absolute Konvergenz der entsprechenden Reihe, wegen $\lim_{n\to\infty} \sqrt[n]{b_n} = 1/e < 1$.

Für reelle Reihen gibt es zwei weitere Konvergenzkriterien, die von der Ordnungsstruktur von $\mathbb{R}$ Gebrauch machen.

Theorem (Leibnizkriterium). *Sei (a_k) eine monoton fallende Nullfolge in $\mathbb{R}$ (d.h. $a_k \geq a_{k+1} \geq \cdots \geq 0$ und $a_k \to 0$). Dann ist die alternierende Reihe $\sum(-1)^k a_k$ konvergent.*

Notation. Eine Reihe heißt alternierend, wenn die zugrundeliegenden Folgeglieder abwechselnd positives und negatives Vorzeichen haben.
Beweis. **Zeichnung.** Sei $n > l$. Dann gilt

$$\begin{aligned} \left|\sum_{k=l}^{n}(-1)^k a_k\right| &= |a_l - a_{l+1} + a_{l+2} - \ldots a_n| \\ &= a_l - a_{l+1} + ...a_n \\ &\leq a_l \\ &\to 0. \end{aligned}$$

Hier: Erste Gleichung: Kein Vorzeichen bei a_l, da Betrag; Zweite Gleichung: betrachte Paare beginnend mit a_l und eventuell einen einzelnen positiven Summanden am Ende

$$0 \leq (a_l - a_{l+1}) + (a_{l+2} - a_{l+3}) + ...$$

Denn es ist jeder einzelne Term nichtnegativ aufgrund der Monotonie. Dritte Gleichung: betrachte Paare beginnend mit a_{l+1} und eventuel einem einzelnen negativen Summanden am Ende

$$a_l \geq a_l + (-a_{l+1} + a_{l+2}) + ...$$

Vierte Gleichung: Nullfolge.
Damit ist $\sum(-1)^k a_k$ eine Cauchy-Folge, also konvergent. □

Bemerkung - Alternativer Beweis. Aufgrund der Monotonie der Folge (a_n) ist die Folge der geraden Partialsummen (S_{2n}) fallend und die Summe der ungeraden Partialsummen (S_{2n+1}) wachsend und es gilt

$$S_{2n+2} = S_{2n+1} + a_{2n} \geq S_{2n+1}.$$

Damit konvergieren sowohl (S_{2n}) als auch (S_{2n+1}) als beschraenkte monotone Folgen. Ihr Grenzwert stimmt ueberein, da die Differenz gerade eine Nullfolge ist.

Bemerkung. Die beiden Voraussetzungen des Theorem sind nötig: Tatsaechlich ist die Voraussetzung deiner Nullfolge immer noetig fuer die Konvergenz der entsprechenden Reihe. Gilt nun die Monotonie nicht, kann die Reihe divergieren, wie folgendes Beispiel zeigt: Sei $a_n =$

$2/n$, n-gerade; $a_n = 0$, sonst. Dann ist $\sum(-1)^n a_n = 1 + 1/2 + 1/3 + \dots$ divergent.

Beispiel - alternierende harmoische Reihe. Die alternierende harmonische Reihe
$$\sum_{n\geq 1} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + ...$$
ist konvergent. Diese Reihe ist nicht absolut konverent, da die harmonische Reihe divergiert.

THEOREM. *(Verdichtungskriterium) Sei (a_n) eine nichtnegative fallende Folge in $\mathbb{R}$ (d.h. $a_k \geq a_{k+1} \geq \dots 0$.) Dann konvergiert $\sum_{k\geq 1} a_k$ genau dann wenn $\sum_{p\geq 1} 2^p a_{2^p}$ konvergiert.*

Beweis. **Zeichnung.** Einteilung von $\mathbb{N}$ in die Abschnitte von $[2^p, 2^{p+1})$.

Für $p \in \mathbb{N}$ setze $C_p := \sum_{k=2^p}^{2^{p+1}-1} a_k$. ($2^p$-Terme). Aufgrund der Monotonie gilt
$$(*) \quad 2^p a_{2^{p+1}} \leq C_p \leq 2^p a_{2^p}.$$
Aufgrund der Nichtnegativität der C_p und a_k reicht es jeweils Beschränktheit zu zeigen. Es gilt also

$\sum_{p\geq 1} 2^p a_{2^p}$ konvergent $\Longleftrightarrow \sum_{p\geq 1} 2^p a_{2^p}$ beschränkt $\overset{(*)}{\Longleftrightarrow} \sum C_p$ beschränkt $\Longleftrightarrow \sum a_k$ beschränkt $\Longleftrightarrow \sum a_k$ konvergent.

$\square$

Beispiel. Sei $\alpha > 0$. Dann gilt $\sum \frac{1}{n^\alpha}$ konvergent $\Longleftrightarrow \alpha > 1$.
Bew. Nach dem Verdichtungskriterium ist $\sum \frac{1}{n^\alpha}$ konvergent genau dann wenn $\sum 2^p/(2^p)^\alpha = \sum 1/(2^{\alpha-1})^p$ konvergiert. Letzteres ist eine geometrische Reihe mit $q = 1/2^{\alpha-1}$.
Ist $\alpha > 1$, so ist $2^{\alpha-1} > 1$ also $0 < q < 1$ und die geometrische Reihe konvergiert.
Ist $\alpha \leq 1$, so ist $2^{\alpha-1} \leq 1$ also $1 \leq q$ und die geometrische Reihe divergiert.

Ende der 20. Vorlesung.

Wir betrachten jetzt noch Doppelsummen. Es geht also um
$$a : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{C}, (i, j) \mapsto a_{ij} = a(i, j)$$
und wir fragen,

- ob $\sum a(n, m)$ in irgendeinem Sinne existiert
- und, wenn ja, ob es egal ist, wie man summiert d.h. ob gilt:

$$\lim_{N\to\infty} \sum_{i,j=1}^{N} a(i, j) = \sum_{N=1}^{\infty} \sum_{k=1}^{N} a(k, N - k + 1) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}.$$

Zeichnung zu den verschiedenen Summationsarten!

Das ist von eigenem Interesse und nützlich zum Studium von Produkten von Reihen und Umordnungen.

Achtung. Es geht um das Vertauschen von Grenzwerten!

Gegenbeispiel. Sei a_{ij} gegeben, so daß alle Einträge Null sind ausser auf der Diagonalen und der ersten unteren Nebendiagonalen. Seien die Einträge auf der Diagonalen $1, 2, 4, 8, \ldots$ und die Einträge auf der ersten Nebendiagonalen $-1, -2, -4, -8, \ldots$. Dann sind alle Spaltensummen absolut konvergent und haben den Wert 0. Die k-te Zeilensumme ist absolut konvergent und hat den Wert 2^{k-1}.

Idee der folgenden Betrachtungen: Gibt es eine gleichmäßige Schranke an die (endlichen Teile) der Doppelsummen, so sitzt die wesentliche 'Masse' in einem (großen) Quadrat. Damit liefern dann alles Summationsarten dasselbe Ergebnis.

Proposition. *Sei $a : \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{C}$. Es gebe $C \geq 0$ mit $\sum_{i,j=1}^m |a_{ij}| \leq C$ für alle $m \in \mathbb{N}$. Dann gilt:*

(a) $\sum a_{\tau(k)}$ für jedes injektive $\tau : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ absolut konvergent.

(b) Für jedes $\varepsilon > 0$ existiert ein $L = L(\varepsilon) \in \mathbb{N}$ mit $\sum_{k=0}^\infty |a_{\sigma(k)}| \leq \varepsilon$ für jedes injektive $\sigma : \mathbb{N} \longrightarrow \mathbb{N}\times\mathbb{N}$ mit $\sigma(k) \notin \{1, \ldots, L\}^2$ für jedes $k \in \mathbb{N}$.

Beweis. (a) Zu $n \in \mathbb{N}$ existiert $m \in \mathbb{N}$ mit $\{\tau(1), \ldots, \tau(n)\} \subset \{1, \ldots, m\}^2$. Damit folgt

$$\sum_{k=0}^n |a_{\tau(k)}| \leq \sum_{i,j=1}^m |a_{ij}| \leq C.$$

Da $n \in \mathbb{N}$ beliebig war, folgt (a).

(b) Sei $\varepsilon > 0$. Sei fuer $N \in \mathbb{N}$

$$S_N := \sum_{k,j=1}^N |a_{kj}|.$$

Dann ist $(S_N)_N$ monoton wachsend und (nach Voraussetzung) beschraenkt. Damit existiert

$$S := \lim_{N\to\infty} \sum_{k,l=1}^N |a_{kl}|.$$

(Das ist die Summe über alle $|a_{ij}|$.) Insbesondere ist also (S_N) eine Cauchy-Folge. Daher existiert also ein $L \in \mathbb{N}$ mit

$$\sum_{k,l\in\{1,\ldots,N\}^2\setminus\{1,\ldots,L\}^2} |a_{kl}| \leq \varepsilon. \ (*)$$

fuer alle $N \geq L$. Ist nun $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ injektiv mit $\sigma(k) \in \mathbb{N} \times \mathbb{N} \setminus \{1, \ldots, L\}^2$, so folgt aus $(*)$

$$\sum_{k=0}^{m} |a_{\sigma(k)}| \leq \lim_{N \to \infty} \sum_{k,l=L}^{N} |a_{kl}| \leq \varepsilon.$$

Das ist die gewuenschte Behauptung. □

Theorem (Konvergenz von Doppelsummen). *Sei $a : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{C}$ gegeben. Es gebe ein $C \geq 0$ mit $\sum_{ij=1}^{m} |a_{ij}| \leq C$ für alle $m \in \mathbb{N}$. Dann existiert $\sum_{j=1}^{\infty} a_{ij}$ für jedes $i \in \mathbb{N}$ und es existiert $\sum_{i=1}^{\infty} a_{ij}$ für jedes $j \in \mathbb{N}$ und die Reihen*

$$\sum_{i \geq 1} \left(\sum_{j=1}^{\infty} a_{ij} \right), \sum_{j \geq 1} \left(\sum_{i=1}^{\infty} a_{ij} \right), \sum_{k \geq 1} \left(\sum_{l=1}^{k} a_{k-l+1,l} \right)$$

sind absolut konvergent und haben den gleichen Grenzwert, nämlich $\lim_{N \to \infty} \sum_{i,j=1}^{N} a_{i,j}$.

Zeichnung der Summen, vgl. oben.

Beweis. Das folgt aus (a) und (b) der vorigen Proposition:

Existenz der 'inneren' Summen: Das folgt aus (a) der vorigen Proposition (z.B. $\tau : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, $\tau(k) = (i, k)$...)

Absolute Konvergenz der Doppelsummmen: Das ist klar nach Voraussetzung, etwa

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{\infty} a_{ij} \right| = \sum_{i=1}^{m} \left| \lim_{n \to \infty} \sum_{j=1}^{n} a_{ij} \right| = \lim_{n \to \infty} \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} \right| \leq C.$$

(Grenzwert vertauscht mit Beträgen und mit endlichen Summen.)

Gleichheit der Grenzwerte: Das folgt aus (b) der vorigen Proposition: In allen Summen wird über beliebig große Quadrate in $\mathbb{N} \times \mathbb{N}$ summiert. Nach (b) der vorigen Proposition sind Summen über Terme ausserhalb solcher Quadrate beliebig klein. Damit folgt die Aussage. □

Wir ziehen nun einige Folgerungen aus dem Satz.

Folgerung (Cauchy-Produkt). *Seien (b_j), (c_i) Folgen in $\mathbb{C}$ sodaß $\sum b_j$ und $\sum c_i$ absolut konvergent sind. Dann gilt*

$$\left(\sum_{i=1}^{\infty} c_i \right) \left(\sum_{j=1}^{\infty} b_j \right) = \sum_{k=1}^{\infty} \sum_{l=1}^{k} c_{k-l+1} b_l$$

Bemerkung. $\sum_{l=1}^{k} c_{k-l+1} b_l = \sum_{l=1}^{k} c_l b_{k-l+1}$.

Beweis. $C := \sum_{i=1}^{\infty} |c_i| < \infty$, $B := \sum_{j=1}^{\infty} |b_j| < \infty$. Setze $a_{ij} := c_i b_j$. Dann gilt

$$\sum_{i,j=1}^{m} |a_{ij}| = \sum_{i,j=1}^{m} |c_i b_j| = \sum_{i,j=1}^{m} |c_i||b_j| = \sum_{i=1}^{m} |c_i| \sum_{j=1}^{m} |b_j| \leq CB < \infty$$

für alle $m \in \mathbb{N}$. Damit erfüllt a die Voraussetzung des vorigen Satz. Dessen Aussage liefert dann die Behauptung. □

Anwendung (Funktionalgleichung der Exponentialfunktion):
Für die Exponentialfunktion $\exp : \mathbb{C} \longrightarrow \mathbb{C}, \exp(z) = e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$ gilt

- $\exp(0) = 1$ und
- $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$ für alle $z_1, z_2 \in \mathbb{C}$ (Funktionalgleichung)

Insbesondere verschwindet exp nirgends und es gilt

$$\frac{1}{\exp(z)} = \exp(-z).$$

Beweis. $\exp(0) = 1$. Das ist klar.

Es gilt die Funktionalgleichung. Das folgt mittels Cauchy-Produkt:

$$\begin{aligned}
\exp(z_1)\exp(z_2) &= \left(\sum_{k=0}^{\infty} z_1^k/k!\right)\left(\sum_{l=0}^{\infty} z_2^l/l!\right) \\
\text{(Cauchy-Produkt)} &= \sum_{n=0}^{\infty}\sum_{m=0}^{n} \frac{z_1^m z_2^{n-m}}{m!(n-m)!} \\
\text{(Erweitern mit } 1 = n!/n!) &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} z_1^m z_2^{n-m} \\
\text{(Binomi)} &= \sum_{n=0}^{\infty} \frac{1}{n!}(z_1 + z_2)^n \\
&= \exp(z_1 + z_2).
\end{aligned}$$

Zum 'Insbesondere'. Nach dem schon Bewiesenen gilt

$$1 = \exp(0) = \exp(z - z) = \exp(z)\exp(-z).$$

Das liefert die Aussage. □

Bemerkung. (Übung) Eine stetige (s.u.) Funktion $E : \mathbb{R} \longrightarrow \mathbb{R}$ mit $E(x+y) = E(x)E(y)$ ist durch ihren Wert an der Stelle $x = 1$ eindeutig bestimmt.

Wir kommen nun zur schon angesprochenen Stabilität von absolut konvergenter Reihen unter Umordnung.

Definition (Umordnung). *Ist $\sum_{k\geq 1} a_k$ eine Reihe und $\tau : \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv, so heißt $\sum_{k\geq 1} a_{\tau(k)}$ eine Umordnung der Reihe $\sum_{k\geq 1} a_k$.*

Bemerkung. Umordnung summiert über 'dieselben' Terme in einer anderen Reihenfolge.

FOLGERUNG (Absolute Konvergenz impliziert Stabilität unter Umordnung). *Sei (a_k) eine Folge in $\mathbb{C}$. Ist $\sum_{k\geq 1} a_k$ absolut konvergent, so ist jede Umordnung absolut konvergent und hat denselben Grenzwert.*

Beweis. Setze $c_{ij} := a_{\tau(j)} = a_i$ falls $i = \tau(j)$ und 0 sonst. **Zeichnung** Dann gilt

- j-te Spalte enthält $a_{\tau(j)}$ an der $\tau(j)$-ten Stelle (und sonst Nullen).
- i-te Zeile enthält a_i an der Stelle j mit $\tau(j) = i$ (und sonst Nullen).

Damit gilt

$$\sum_{j=1}^{\infty} c_{ij} = a_i$$

für jedes $i \in \mathbb{N}$ und

$$\sum_{i=1}^{\infty} c_{ij} = a_{\tau(j)}$$

für jedes $j \in \mathbb{N}$ sowie

$$\sum_{i,j=1}^{m} |c_{ij}| \leq \sum_{k=1}^{\infty} |a_k| = C < \infty.$$

Damit liefert der vorige Satz die Behauptung. □

Nach Hause nehmen. Wenn die Reihen absolut konvergieren, darf man summieren wie man will und erhält immer den gleichen Grenzwert!

Gegenbeispiel. (Voraussetzung der absoluten Konverenz ist nötig) Betrachte die Reihe

$$1 - 1 + 1/2 - 1/2 + 1/3 - 1/3 \ldots.$$

Dann gilt für die Partialsummen S_n also $S_n = 0$ falls n gerade und $S_n = 1/k$ falls $n = 2k - 1$. Damit folgt $S_n \to 0$, $n \to \infty$. Andererseits kann man die Summe so umordnen, daß sie divergiert:

$$(1 + 1/2) - 1 + (1/3 + \ldots + 1/k) - 1/2 + (1/(k+1) + \ldots + \frac{1}{n}) - 1/3..$$

so daß man immer wieder Summanden $\geq 1/2$ erhält...

Dieses Verfahren lässt sich verallgemeinern und führt auf den folgenden Satz:

$\longleftrightarrow$
Ende der 21. Vorlesung.

THEOREM (Riemannscher Umordnungsatz). *Ist* (a_k) *eine Folge in* $\mathbb{R}$ *und* $\sum a_k$ *konvergent aber nicht absolut konvergent, so existiert zu jedem* $s \in \mathbb{R}$ *eine Umordnung* τ *mit* $s = \sum_{k=1}^{\infty} a_{\tau(k)}$. *Ebenso existieren bestimmt divergente Umordnungen zu* $+\infty$ *und* $-\infty$.

Beweis. Die Grundidee ist, daß Konvergenz der Reihe ohne absolute Konvergenz nur auftreten kann, wenn die Reihe in zwei Teile zerlegt werden kann, die sich zu $+\infty$ bzw. $-\infty$ addieren. Durch geeignetes 'Verschieben' der Balance zwischen diesen beiden Teilen lässt sich dann die gewünschte Konvergenz der Umordnung erzwingen. Hier sind die Details: Teile die Folge (a_k) in positive und nichtpositive Terme wie folgt: Sei

$$a_k^+ := \text{ k-tes positives Element von } (a_k)$$

und

$$a_k^- := -k\text{-tes nichtpositives Element von } (a_k).$$

Setze

$$A^+ := \lim_{n\to\infty} \sum_{k=1}^{n} a_k^+, \quad A^- := \lim_{n\to\infty} \sum_{k=1}^{n} a_k^-,$$

wobei der Wert ∞ möglich ist.

Behauptung. Es gilt:

- $A^+ := \lim_{n\to\infty} \sum_{k=1}^n a_k^+ = \infty$ und $A^- := \lim_{n\to\infty} \sum_{k=1}^n a_k^- = \infty$.
- (a_n^+) und (a_n^-) sind Nullfolgen.

Bew: $A^+, A^- = \infty$*:* Angenommen A^+ und A^- endlich: Widerspruch zu $\sum a_n$ nicht absolut konvergent. Angenommen: A^+ endlich und A^- unendlich (oder umgekehrt): Widerspruch zu $\sum a_n$ konvergent.
Nullfolge: klar (da Summe konvergiert).
Damit ist die Behauptung beweisen.

Nun gehe so vor: Summiere a_k^+ bis gerade s überschritten wird; summiere nun a_k^- bis s gerade unterschritten wird etc. Wegen $A^+, A^- = \infty$ kann dieses Verfahren beliebig fortgesezt werden. Da (a_k^+) und (a_k^-) Nullfolgen sind, folgt die Behauptung. □

Bemerkung. Mit einer leichten Modifikation lassen sich dann Umordnungen erzeugen, deren zugehoerige Reihen zwei Haeufungspunkte haben.

FOLGERUNG (Absolute Konvergenz äquivalent zu unbedingter Konvergenz). *Sei* (a_n) *eine Folge in* $\mathbb{C}$. *Dann sind äquivalent:*

(i) *Es ist* $\sum a_n$ *absolut konvergent.*
(ii) *Es konvergiert jede Umordnung von* $\sum a_n$ *in* $\mathbb{C}$. *('Unbedingte Konvergenz der Reihe')*

In diesem Fall haben alle Umordnungen denselben Grenzwert.

Beweis. (i)$\Longrightarrow$ (ii): s.o.

(ii)$\Longrightarrow$ (i): Anwenden des Riemanschen Umordnungssatzes auf Realteil und Imaginärteil der Summe liefert absolute Konvergenz von $\sum \Re a_n$ und $\sum \Im a_n$. Damit folgt die Aussage. $\square$

KAPITEL 8

Stetigkeit und Grenzwerte von Funktionen

Stetigkeit von Funktionen ist ein fundamentales Konzept der Analysis, das man als Anwendung des Grenzwertkonzeptes sehen kann. Die **Grundidee** ist folgende: Eine Funktion f heißt stetig im Punkt p, wenn sie Punkte q, die nahe an p liegen, auf Werte abbildet, die nahe an $f(p)$ liegen d.h. kurz gefaßt: *q nahe p impliziert $f(q)$ nahe $f(p)$.* Natuerlich muss praezise gefaßt werden, was unter Naehe verstanden wird.

DEFINITION (Stetigkeit). *Sei $D \subset \mathbb{C}$ und $f : D \longrightarrow \mathbb{C}$. Dann heißt f stetig in $p \in D$, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert mit*

$$|f(q) - f(p)| < \varepsilon$$

für alle $q \in D$ mit $|q - p| < \delta$. Ist f in jedem $p \in D$ stetig, so heißt f stetig auf D.

Bemerkungen.

- In dieser Definition wird also *q nahe an p* gefaßt durch $|q-p| < \delta$ und *$f(q)$ nahe an $f(p)$* durch $|f(q) - f(p)| < \varepsilon$.
- Unter Nutzen des Konzeptes der offenen r- Kugel (r-Umgebung)

  $$U_r(z) = \{w : |w - z| < r\}$$

  von z mit Radius $r > 0$ lässt sich obige Definition der Stetigkeit offenbar auch so ausdrücken: Es heißt f stetig in p, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

  $$f(U_\delta(p) \cap D) \subset U_\varepsilon(f(p)),$$

  d.h. zu jeder offenen Kugel um $f(p)$ existiert eine offene Kugel um p, die in die Kugel um $f(p)$ abgebildet wird. **Zeichnung.**
- Offenbar schließt diese Definition die Fälle ein, daß $D \subset \mathbb{R}$ gilt und/oder f nur reelle Werte annimmt. Diesen Fällen werden wir uns später noch einmal besonders widmen.
- Ist $p \in D$ ein isolierter Punkt von D (d.h. es existiert ein $r > 0$ mit $\{p\} = D \cap U_r(p)$ Zeichnung), so ist jedes $f : D \longrightarrow \mathbb{C}$ in p stetig. (Bew. Wähle $\delta > 0$ mit $\delta < r$ beliebig.)

Wir betrachten nun einige **Beispiele**. Tatsaechlich wird sich zeigen, daß alle 'gaengigen' Funktionen stetig sind.

Konstante Funktion. Die konstante Funktion $D \longrightarrow \mathbb{C}\ x \mapsto c$, ist stetig.

Beweis. Es kann $\delta > 0$ beliebig gewählt werden.

Identitaet. Die Identitaet $id : D \subset \mathbb{R}^N \longrightarrow \mathbb{C}$, $x \mapsto x$, ist stetig.

Beweis. Es kann $\delta = \varepsilon$ gewählt werden.

Betrag. Der Betrag $|\cdot| : D \subset \mathbb{C} \longrightarrow \mathbb{C}$, $x \mapsto |x|$, ist stetig.

Beweis. Dreiecksungleichung $||x|-|y|| \leq |x-y|$ zeigt, daß $\delta = \epsilon$ gewählt werden kann.

k-te Wurzel. Für jedes $k \in \mathbb{N}$ ist $\sqrt[k]{\ } : [0,\infty) \longrightarrow [0,\infty)$ ist stetig

Beweis. Wir zeigen zunächst folgende

Zwischenbehauptung. $|\sqrt[k]{y} - \sqrt[k]{x}| \leq \sqrt[k]{|y-x|}$ für alle $x, y \geq 0$.

Beweis der Zwischenbehauptung: Ohne Einschränkung sei $x < y$. Dann gilt $\sqrt[k]{y} > \sqrt[k]{x}$. Es ist also

$$\sqrt[k]{y} \leq \sqrt[k]{y-x} + \sqrt[k]{x}$$

zu zeigen. Es reicht also zu zeigen, daß

$$(\sqrt[k]{y})^k \leq \left(\sqrt[k]{y-x} + \sqrt[k]{x}\right)^k.$$

Das ist wahr nach dem binomischen Satz. Damit ist die Zwischenbehauptung bewiesen.

Nach der Zwischenbehauptung kann $\delta = \varepsilon^k$ gewählt werden.
Bemerkung. In diesen Beispielen kann (bei gegebenem $\varepsilon > 0$) das δ unabhaengig von x gewaehlt werden. Das fuehrt auf eine Verschaerfung des Konzeptes der Stetigkeit, naemlich die gleichmaeßige Stetigkeit (s.u.). Im allgemeinen kann (zu gegebenem $\varepsilon > 0$) das δ nicht unabhaengig von x gewaehlt werden, wie man auch am folgenden Beispiel sieht.

Exponentialfunktion. Die Exponentialfunktion $\exp : \mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto \sum_{k=0}^{\infty} \frac{z^k}{k!}$ ist stetig.
Bew. Es gilt nach Funktionalgleichung

$$\exp(z)-\exp(z_0) = \exp(z_0)(\exp(z-z_0)-\exp(0)) = \exp(z_0)(\exp(z-z_0)-1).$$

Es reicht also die Stetigkeit von exp bei 0 zu zeigen. Dazu rechnen wir

$$\begin{aligned}
|\exp(z) - 1| &= |\sum_{k=1}^{\infty} \frac{1}{k!} z^k| \\
&= |z \sum_{k=1}^{\infty} \frac{1}{k!} z^{k-1}| \\
&\leq |z| \sum_{k=1}^{\infty} \frac{1}{k!} |z|^{k-1} \\
(|z| < 1) \quad &= |z| \sum_{k=0}^{\infty} \frac{1}{k!} \\
&= |z| \exp(1).
\end{aligned}$$

Daraus folgt

$$|\exp(z) - 1| \leq \varepsilon$$

fuer $z \in \mathbb{C}$ mit

$$|z| < 2 \ \text{ und } \ |z| < \frac{\varepsilon}{\exp(1)}.$$

Damit ergibt sich die Stetigkeit in 0 mit

$$\delta := \min\{2, \frac{\varepsilon}{\exp(1)}\}.$$

Ende der 22. Vorlesung.

Die Argumentation des letzten Beispiel laeßt sich auf eine große und sehr wichtige Klasse von Funktionen verallgemeinern. Diese diskutieren wir als naechstes.

Potenzreihen. Sei die Folge (a_n) in $\mathbb{C}$ mit

$$\limsup_n \sqrt[n]{|a_n|} =: \rho < \infty$$

gegeben und sei zu dieser Folge R definiert durch

$$R := 1/\rho > 0$$

(mit $R = \infty$ falls $\rho = 0$). Dann ist die Reihe $\sum_{k=0}^{\infty} a_k z^k$ fuer alle $z \in U_R = U_R(0)$ absolut konvergent und die Funktion

$$f : U_R \subset \mathbb{C} \longrightarrow \mathbb{C}, \ \ f(z) := \sum_{k=0}^{\infty} a_k z^k$$

ist stetig.

Ein **Spezialfall** von Potenzreihen ergibt sich, wenn $a_n = 0$ fuer alle $n \geq N$ gilt. In diesem Fall ist die Potenzreihe also ein Polynom. Dann ist $\rho = 0$ und damit $R = \infty$ und man erhaelt Stetigkeit auf ganz $\mathbb{C}$.

Beweis. Wir betrachten nur den Fall $R < \infty$. (Der andere Fall ist leichter.)

Reihe ist absolut konvergent: Das folgt aus dem Wurzelkriterium mit

$$\limsup_k \sqrt[k]{|a_k z^k|} = \limsup_k |z| \sqrt[k]{|a_k|} = |z|\rho < R\rho = 1.$$

Stetigkeit: Aehnlich wie im Beispiel der Exponentialfunktion ergibt sich durch direkte Rechnung

$$|f(z) - f(z_0)| \leq \sum_{k=1}^{\infty} a_k (z - z_0) \sum_{l=0}^{k-1} z^{k-1-l} z_0^l.$$

Für z mit $|z| < \frac{|z_0|+R}{2} =: c < R$ (Zeichnung) ergibt sich dann

$$|f(z) - f(z_0)| \leq |z - z_0| \sum_{k=1}^{\infty} |a_k| k c^{k-1} = C|z - z_0|$$

mit $C := \sum_{k=1}^{\infty} |a_k| k c^{k-1} < \infty$. (Hier folgt Endlichkeit von C nach dem Wurzelkriterum, Check!) Damit folgt Stetigkeit (mit $\delta = \min\{\varepsilon/C, c - |z_0|\}$).

Es heißt R der *Konvergenzradius* der Potenzreihe $\sum_{k=0}^{\infty} a_k z^k$. Es ist R charakterisiert durch folgende Eigenschaft:

- Für $|z| < R$ ist $\sum a_k z^k$ absolut konvergent. (Bew. s.o.)
- Für $|z| > R$ ist is $\sum a_k z^k$ divergent. (Bew. Klar, da dann $|a_k z^k|$ keine Nullfolge ist.)

(Fuer z mit $|z| = R$ ist keine allgemeine Aussage moeglich.) Ist insbesondere $\sum a_k z^k$ konvergent für ein z, so folgt $|z| \leq R$. Damit erhaelt man also folgende alternative Formel fuer den Konvergenzradius

$$R = \sup\{r : \sum_{k=0}^{\infty} a_k z^k \text{ konvergent für ein } z \text{ mit } |z| = r \}.$$

Wir diskutieren nun noch kurz eine Verschärfung des Konzeptes der Stetigkeit, die durch die Beispiele nahegelegt wird:

Definition (Gleichmäßige Stetigkeit). *Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt gleichmäßig stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit*

$$|f(p) - f(q)| < \varepsilon$$

für alle $p, q \in D$ mit $|p - q| < \delta$.

Kurzfassung: Stetigkeit: $\delta = \delta(\varepsilon, x)$. Gleichmässige Stetigkeit $\delta = \delta(\varepsilon)$.

Beispiele. Wie schon weiter oben diskutiert sind zum Beispiel die Funktionen $id : \mathbb{C} \longrightarrow \mathbb{C}$, $|\cdot| : \mathbb{C} \longrightarrow [0, \infty)$ und $\sqrt[k]{\ } : [0, \infty) \longrightarrow [0, \infty)$ gleichmaeßig stetig.

Eine spezielle Klasse gleichmaeßig stetiger Funktionen sind die folgenden Funktionen:

Lipschitzstetige Funktionen. Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt lipschitzstetig, wenn ein $C \geq 0$ existiert mit

$$|f(x) - f(y)| \leq C|x - y|$$

für alle $x, y \in D$. Offenbar ist jede lipschitzstetige Funktion gleichmäßig stetig (mit $\delta = \frac{\varepsilon}{C}$).

Typische Beispiele von lipschitzstetigen Funktionen:

- Lineare Funktionen $f : \mathbb{C} \longrightarrow \mathbb{C}$, $f(x) = ax + b$.
- $|\cdot|$, $\Re$, $\Im$, id auf $\mathbb{C}$.
- (Übung) Ist $A \subset \mathbb{C}$, so ist die zugehörige Distanzfunktion

  $$d = d_A : \mathbb{C} \longrightarrow [0, \infty),\ d(z) := \inf\{|z - w| : w \in A\}$$

  Lipschitzstetig mit $C = 1$.

Gegenbeispiel. Fuer $k \geq 2$ ist die Wurzel $\sqrt[k]{\ } : [0, \infty) \longrightarrow)$ ist nicht lipschitzstetig, aber gleichmäßig stetig. (Gleichmässig stetig: s.o.; Angenommen lipschitzstetig d.h. $|\sqrt[k]{x} - \sqrt[k]{y}| \leq C|x - y|$. Betrachte $x = 0$ und $y = 1/n$ fuer $n \in \mathbb{N}$...)

Bemerkung. Eine noch allgemeinere Klasse stetiger Funktionen bilden die *hölderstetige Funktionen.* Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt hölderstetig mit Exponent $\alpha > 0$, wenn ein $C > 0$ existiert mit $|f(x) - f(y)| \leq C|x - y|^\alpha$ für alle $x, y \in D$ mit $|x - y| < 1$. Jede Hölderstetige Funktion ist stetig.
(Bew. Ist $\alpha = 1/k$ so folgt dies aus der Stetigkeit der k-ten Wurzel. Zur Behandlung eines beliebigen $\alpha > 0$ wählt man ein $k \in \mathbb{N}$ mit $\frac{1}{k} < \alpha$ und nutzt $|x - y|^\alpha < |x - y|^{1/k}$ für $|x - y| < 1$.)
Übung: Ist $f : \mathbb{R} \longrightarrow \mathbb{C}$ Hölderstetig mit Exponent $\alpha > 1$, so ist f konstant. (**Hinweis.** Manchmal wird auch in der Definition der Hölderstetigkeit die Einschränkung $|x - y| < 1$ weggelassen. Dann würde man die oben definierte Klasse als lokal hölderstetige Funktionen bezeichnen. Auf beschränkten Mengen D stimmen die beiden Definitionen überein. Für unbeschränkte Mengen ist das im allgemeinen nicht der Fall (so sind etwa lineare nichtkonstante Funktionen g auf $\mathbb{R}$ hölderstetig im obigen Sinne, erfüllen aber nicht $|g(x) - g(y)| \leq C|x - y|^\alpha$ für $0 \leq \alpha < 1$ für alle $x, y \in \mathbb{R}$). Dann hat obige Definition den Vorteil, daß jede lipschitzstetige Funktion automatisch hölderstetig zu jedem Exponente $0 \leq \alpha \leq 1$ ist.)

Zur Abgrenzung geben wir noch einige Beispiele von unstetigen Funktionen:

Beispiel (Heaviside Funktion) Die Funktion $H : \mathbb{R} \longrightarrow \{0, 1\}$ mit $H(x) = 0$ für $x \leq 0$ und $H(x) = 1$ für $x > 0$ ist unstetig in $x = 0$ und stetig in allen anderen Punkten.

Beispiel. Eine Funktion, die in keinem Punkt stetig ist, wird gegeben durch

$$1_{\mathbb{Q}} : \mathbb{R} \longrightarrow \{0,1\}, 1_{\mathbb{Q}}(x) := \begin{cases} 1 & : \quad x \in \mathbb{Q} \\ 0 & : \quad x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Bew. Das folgt, da $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ dicht in $\mathbb{R}$ sind, also jede ε-Kugel um ein $p \in \mathbb{R}$ sowohl Punkte von $\mathbb{Q}$ als auch von $\mathbb{R} \setminus \mathbb{Q}$ enthält.

Beispiel. (Übung) Die Funktion

$$N : \mathbb{R}\backslash\{0\} \longrightarrow \mathbb{R},\ N(x) := \begin{cases} 0 & : \quad \text{x irrational} \\ \frac{1}{p} & : \quad x = q/p \text{ mit } p \in \mathbb{N} \text{ und } q \in \mathbb{Z} \text{ teilerfremd}, \end{cases}$$

ist stetig in allen irrationalen Punkten und unstetig in den rationalen Punkten.

Es lässt sich Stetigkeit einer Funktion mit Konvergenz von Folgen charakterisieren und das ist sehr nützlich (da wir gut mit konvergenten Folgen umgehen können).

Lemma (Folgencharakterisierung der Stetigkeit). *Sei $D \subset \mathbb{C}$ und $f : D \longrightarrow \mathbb{C}$ gegeben. Dann sind für $p \in D$ äquivalent:*

(i) *Es ist f stetig in p.*
(ii) *Für jede Folge (p_n) in D mit $p_n \to p$ gilt $f(p_n) \to f(p)$.*

Beweis. Sei $c = f(p)$.

(i)$\Longrightarrow$ (ii): $p_n \to p$ (Zu zeigen: $f(p_n) \to c$.) Sei $\varepsilon > 0$. Dann existiert nach (i) ein $\delta > 0$ mit $|f(q) - c| < \varepsilon$ für alle $q \in D$ mit $|q - p| < \delta$. Wegen $p_n \to p$ existiert ein $N = N_\delta \in \mathbb{N}$ mit $|p_n - p| < \delta$ für $n \geq N_\delta$. Damit gilt für $n \geq N$ also $|f(p_n) - c| \leq \varepsilon$.

(ii)$\Longrightarrow$ (i): Angenommen nein: Dann existiert ein $\varepsilon > 0$ 'ohne δ', d.h. mit der Eigenschaft, daß für jedes $n \in \mathbb{N}$ ein $p_n \in D$ existiert mit $|p_n - p| < \frac{1}{n}$ aber $|f(p_n) - c| \geq \varepsilon$. Dann gilt $p_n \to p$ aber nicht $f(p_n) \to c$. Widerspruch. □

Ende der 23. Vorlesung

Aus dieser Charakterisierung von Stetigkeit erhalten wir sofort einige Rechenregeln für stetige Funktionen (die wir natürlich auch direkt zeigen könnten).

Proposition (Rechenregeln). *$D \subset \mathbb{C}$ und $f, g : D \longrightarrow \mathbb{C}$ gegeben. Seien f, g stetig in p. Dann gilt:*
(a) Fuer jedes $\alpha \in \mathbb{C}$ ist $f + \alpha g : D \longrightarrow \mathbb{C}$ stetig in p.
(b) Es ist $fg : D \longrightarrow \mathbb{C}$ stetig in p.
(c) Gilt $g(p) \neq 0$, so existiert ein $r > 0$ mit $g(q) \neq 0$ für alle $q \in U_r(p)$ und es ist $f/g : U_r(p) \cap D \longrightarrow \mathbb{C}$ stetig in p.

Beweis. Es folgen (a) und (b) direkt aus den Recheneregeln für konvergente Folgen.

(c) Wir zeigen die Existenz eines $r > 0$ mit $g(q) \neq 0$ fuer alle $q \in U_r(p) \cap D$. (Dann folgt gewuenschte Stetigkeitsaussage direkt aus den Recheneregeln für konvergente Folgen.) Sei $\varepsilon := |g(p)| > 0$. Dann existiert also aufgrund der Stetigkeit von g in p ein $\delta > 0$ mit $|g(q) - g(p)| < \varepsilon$ fuer alle $q \in U_\delta(p)$. Damit gilt dann also fuer $q \in U_\delta(p)$

$$|g(q)| = |q(p) - (g(p) - g(q))| \geq |g(q)| - |g(p) - g(q)| = \varepsilon - |g(p) - g(q)| > 0.$$

Damit kann man also $r = \delta > 0$ waehlen. $\square$

Beispiel. Ist P ein Polynom d.h. $P(z) = \sum_{j=0}^{n} a_j z^j$, so ist $f(z) = P/\exp$ stetig.

Auch Komposition stetiger Funktionen ist stetig.

PROPOSITION. *Seien $u : D \subset \mathbb{C} \longrightarrow \mathbb{C}$ und $v : E \longrightarrow \mathbb{C}$ und $p \in D$ gegeben mit $u(D) \subset E$. Ist u stetig in p und v stetig in $u(p)$, so ist $v \circ u : D \longrightarrow \mathbb{C}$ stetig in p.*

Beweis. Auch das folgt leicht mit der Folgencharakterisierung der Stetigkeit: $p_n \to p \Longrightarrow u(p_n) \to u(p) \Longrightarrow v(u(p_n)) \to v(u(p))$. Das beendet den Beweis. $\square$

FOLGERUNG. *Sind $f, g : D \longrightarrow \mathbb{R}$ stetig, so sind auch $\max\{f, g\}$ und $\min\{f, g\}$ stetig.*

Beweis. Es gilt

$$\max\{a, b\} = \frac{a+b}{2} + \frac{|a-b|}{2}$$

und

$$\min\{a, b\} = \frac{a-b}{2} - \frac{|a-b|}{2}.$$

Damit folgt die Aussage leicht aus den vorangehenden beiden Propositionen. $\square$

Wir systematisieren jetzt die obigen Betrachtungen durch das Konzept des **Grenzwertes einer Funktion in einem Punkt**. Dabei werden wir auch Punkte zulassen, in denen die Funktionen nicht definiert ist (in deren Nähe sie aber definiert ist). Für solche Punkte ist die Frage nach Existenz von Grenzwerten gerade die Frage nach stetiger Fortsetzbarkeit der Funktion. Für Punkte, in denen die Funktion definiert ist, ist Frage nach der Existenz von Grenzewerten gerade die Frage nach Stetigkeit der Funktion.

Wir führen zunächst die Punkte ein, um die es uns geht.

DEFINITION (Berührpunkt). *Sei $D \subset \mathbb{C}$. Ein $x \in \mathbb{C}$ heißt Berührpunkt von D, wenn es eine Folge in D gibt die gegen x konvergiert.*

Bemerkung.

- Es ist x Berührpunkt von D genau dann, wenn $D \cap U_r(x) \neq \emptyset$ für alle $r > 0$. (Einfach)
- Es gibt zwei Typen von Berührpunkten von D: (Zeichnung)
  - Punkte aus D (Wähle $x_n \equiv x$.)
  - Punkte aus $\mathbb{C} \setminus D$, die von D den 'Abstand 0 haben' d.h. für die in jeder Umgebung ein Punkt von D liegt. Beispiel: $D = (a, b)$ hat a, b als Beruehrpunkte.
- Ein Punkt p heißt Haeufungspunkt von D, wenn $D \cap U_r(x)$ unendlich viele Element hat. Dann ist jeder Beruehrpunkt von D, der nicht zu D gehoert, ein Haeufungspunkt von D (einfach). Allgemeiner gilt:

$$\text{p Beruehrpunkt von } D \setminus \{p\} \iff \text{p Haeufungspunkt von } D.$$

Lemma (Grenzwert einer Funktion). *Sei $D \subset \mathbb{C}$ und $f : D \longrightarrow \mathbb{C}$ gegeben. Sei p ein Berührpunkt von D. Für $c \in \mathbb{C}$ sind äquivalent:*

(i) *Für jede Folge (p_n) in D mit $p_n \to p$ gilt $f(p_n) \to c$.*
(ii) *Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|f(q) - c| \leq \varepsilon$ für alle $q \in D$ mit $|q - p| < \delta$.*

In diesem Fall gilt $c = f(p)$ falls p zu D gehoert.

Beweis. Die Aeuqivaelenzaussage verallgemeinert die Folgencharakterisierung der Stetigkeit. Der Beweis kann wortwörtlich von dort übernommen werden. Da er so schoen (und wichtig ist) fuegen wir das hier ein:

(i)$\Longrightarrow$ (ii): $p_n \to p$ (Zu zeigen: $f(p_n) \to c$.) Sei $\varepsilon > 0$. Dann existiert nach (i) ein $\delta > 0$ mit $|f(q) - c| < \varepsilon$ für alle $q \in D$ mit $|q - p| < \delta$. Wegen $p_n \to p$ existiert ein $N = N_\delta \in \mathbb{N}$ mit $|p_n - p| < \delta$ für $n \geq N_\delta$. Damit gilt für $n \geq N$ also $|f(p_n) - c| \leq \varepsilon$.

(ii)$\Longrightarrow$ (i): Angenommen nein: Dann existiert ein $\varepsilon > 0$ 'ohne δ', d.h. mit der Eigenschaft, daß für jedes $n \in \mathbb{N}$ ein $p_n \in D$ existiert mit $|p_n - p| < \frac{1}{n}$ aber $|f(p_n) - c| \geq \varepsilon$. Dann gilt $p_n \to p$ aber nicht $f(p_n) \to c$. Widerspruch.

Zur letzten Aussage: Waehle $x_n = p$. $\square$

Definition (Grenzwert einer Funktion). *In der Situation des Lemma, heißt c der Grenzwert von f bei p. Man schreibt $c = \lim_{x \to p} f(x)$ oder $f(x) \to c$, $x \to p$.*

Bemerkung. Wieder lässt sich (ii) mittels offener Kugeln ausdrücken: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit

$$f(U_\delta(p) \cap D) \subset U_\varepsilon(c).$$

Kurzfassung: 'Zu jeder offenen Kugel um c existiert eine offene Kugel um p, die durch f in die Kugel um c abgebildet wird.' **Zeichnung.**

Existenz des Grenzwertes in einem Punkt ist äquivalent zu Stetigkeit (falls der Punkt zum Definitionsbereich gehört) und zu Fortsetzbarkeit zu einer im Punkt stetigen Funktion (falls der Punkt nicht zum Definitionsbereich gehört):

Folgerung (Existenz des Grenzwertes und Stetigkeit). *Sei $f : D \longrightarrow \mathbb{C}$ gegeben. Sei p ein Berührpunkt von D. Dann gilt:*
(a) Gehört p zu D, so sind äquivalent:
(i) *Es existiert* $\lim_{x\to p} f(x)$.
(ii) *f ist stetig in $p \in D$.*
(b) Gehört p zu $\mathbb{C} \setminus D$, so sind äquivalent:
(i) *Es existiert* $\lim_{x\to p} f(x)$.
(ii) *f besitzt eine stetige Fortsetzung $\widetilde{f}$ auf $D \cup \{p\}$.*
In diesem Fall ist diese Fortsetzung eindeutig bestimmt und gegeben durch

$$\widetilde{f}(x) = f(x) \text{ für } x \in D \text{ und } \widetilde{f}(p) = c.$$

Beweis. Das folgt sofort aus den Definitionen und dem vorangegangenen Lemma. □

Wir betrachten nun noch eine besondere Situation, naemlich den Fall $D \subset \mathbb{R}$. In diesem Fall gibt es noch eine weitere Charakterisierung der Existenz des Grenzwertes. Das ist eine recht spezielle Eigenschaft von $\mathbb{R}$. (Fuer $\mathbb{C}$ ist die analoge Aussage falsch; vgl. Analysis II.)

Definition (Grenzwerte von links und von rechts). *Sei $D \subset \mathbb{R}$, $f : D \longrightarrow \mathbb{C}$ und $p \in D$ gegeben.*
(a) Ist p ein Berührpunkt von $D \cap (p, \infty)$, so nennt man, falls existent, $\lim_{q\to p} f|_{D\cap(p,\infty)}$ den rechtsseitigen Grenzwert von f in p und bezeichnet ihn auch als

$$\lim_{q\to p+} f(x).$$

(Bsp: $D = (p, \infty)$).
(b) Ist p ein Berührpunkt von $D\cap(-\infty, p)$, so nennt man, falls existent, $\lim_{q\to p} f|_{D\cap(-\infty,p)}$ den linksseitigen Grenzwert von f in p und bezeichnet ihn auch als

$$\lim_{q\to p-} f(x).$$

(Bsp. $D = (-\infty, p)$).

Bemerkung. Nach dem vorangehenden Lemma lautet das ausgeschrieben so:

$$\lim_{q\to p+} f(q) = c.$$

$$\Longleftrightarrow$$

Für alle $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|f(q) - c| < \varepsilon$ für alle $q \in D$ mit $|q - p| < \delta$ **und** $q > p$

$$\Longleftrightarrow$$

Es gilt $f(p_n) \to f(p)$ für jede Folge (p_n) in D mit $p_n \to p$ **und** $p_n > p$.

$\lim_{q\to p-} f(q) = c$: Analog (zur Übung überlassen).

Beispiele. (Existenz links- und rechtsseitiger Grenzwerte ohne Stetigkeit)

- $f = 1_{(-\infty,0]} : \mathbb{R} \longrightarrow \{0,1\}$. Links- und rechtsseitige Grenzwerte existieren in 0 und sind ungleich; es ist f aber nicht stetig in 0. **Zeichnung.**
- $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = 0$ für $x < 0$, $f(x) = 1/2$ für $x = 0$ und $f(x) = 1$ für $x > 0$. Links und rechtsseitige Grenzwerte existieren in 1 und sind ungleich; es ist f aber nicht stetig in 0. **Zeichnung.**
- $f : (0,2) \longrightarrow \mathbb{R}$, $f(x) = x$ für $x \neq 1$, $f(1) = 2$. Links und rechtsseitige Grenzwerte existieren in 1 und sind gleich; es ist f aber nicht stetig in 1.

Eine wesentliche Anwendung des Konzeptes von links- und rechtseitigem Grenzwert ist das folgende Theorem. (Die andere wesentliche Anwendung werden wir bei der Diskussion monotoner Funktionen kennenlernen.)

Theorem (Charakterisierung Stetigkeit mit links- und rechtsseitigen Grenzwerten). *Sei $D \subset \mathbb{R}$, $f : D \longrightarrow \mathbb{C}$ und $p \in D$ gegeben. Dann sind äquivalent:*

(i) *Es ist f stetig in p.*
(ii) *Es existieren $\lim_{q\to p+} f(q)$ und $\lim_{q\to p-} f(q)$ und stimmen mit $f(p)$ überein.*

Beweis. $(i) \Longrightarrow (ii)$: f stetig $\Longrightarrow \lim f(p_n) = f(p)$ für JEDE Folge $(p_n) \to p$. Damit folgt (i).

$(ii) \Longrightarrow (i)$: Sei $\varepsilon > 0$ beliebig. Dann gilt:

- $\lim_{q\to p+} f(q) = f(p)$.
- $\lim_{q\to p-} f(q) = f(p)$.

Damit existiert ein δ_+ mit $|f(q) - f(p)| < \varepsilon$ für alle $q > p$ mit $|q - p| < \delta_+$ und es existiert ein δ_- mit $|f(q) - f(p)| < \varepsilon$ für alle $q < p$ mit $|q - p| < \delta_-$. Mit $\delta := \min\{\delta_+, \delta_-\}$ gilt dann

$$|f(p) - f(q)| < \varepsilon$$

für alle q mit $|q - p| < \delta$. □

Bemerkung. (Übung) Analog lässt sich für $f : D \subset \mathbb{R} \longrightarrow \mathbb{C}$ zeigen: $\lim f(p) = c \Longleftrightarrow$ Es existieren links- und rechtsseitiger Grenzwert $\lim_{q\to p+} f(q)$ und $\lim_{q\to p+} f(q)$ und sind gerade c.

Ende der 24. Vorlesung

Wir betrachten nun noch **Grenzwerte, die $\pm\infty$ involvieren**. Die **Grundidee** ist dabei, daß weiterhin

$$\lim_{q\to p} f(q) = c$$

genau dann gilt, wenn fuer jede Folge (q_n) mit $q_n \to p$ gilt $f(q_n) \to c$. Im Unterschied zu den bisherigen Betrachtungen ist hier allerdings der Fall $p = \pm\infty$ bzw. $c = \pm\infty$ erlaubt.

Hier sind die Details:

In der Situation $D \subset \mathbb{R}$ gibt es noch **Grenzwerte bei $\pm\infty$**: Sei $f : D \longrightarrow \mathbb{C}$.

Grenzwert bei ∞: Ist $D \subset \mathbb{R}$ nach oben unbeschränkt (und in diesem Sinne also ∞ ein Beruehrpunkt von D), und $c \in \mathbb{C}$, so hat f bei ∞ den Grenzwert c, geschrieben als $\lim_{x\to\infty} f(x) = c$, oder $f(x) \to c$, $x \to \infty$, falls fuer jedes $\varepsilon > 0$ ein $C \in \mathbb{R}$ existiert mit $|f(x) - c| < \varepsilon$ für alle $x \in D$ mit $x \geq C$. Analog zum Beweis des entsprechenden obigen Lemma zeigt man dann:
$\lim_{x\to\infty} f(x) = c \Longleftrightarrow$ Für jede Folge (x_n) in D mit $x_n \to \pm\infty$ gilt $f(x_n) \to c$.
Grenzwert bei $-\infty$: Ist $D \subset \mathbb{R}$ nach unten unbeschränkt (und in diesem Sinne also $-\infty$ ein Beruehrpunkt von D), und $c \in \mathbb{C}$, so hat f bei $-\infty$ den Grenzwert c, geschrieben als $\lim_{x\to-\infty} f(x) = c$, oder $f(x) \to c$, $x \to -\infty$, falls fuer jedes $\varepsilon > 0$ ein $C \in \mathbb{R}$ existiert mit $|f(x) - c| < \varepsilon$ für alle $x \in D$ mit $x \leq C$. Analog zum Beweis des entsprechenden obigen Lemma zeigt man dann:
$\lim_{x\to-\infty} f(x) = c \Longleftrightarrow$ Für jede Folge (x_n) in D mit $x_n \to -\infty$ gilt $f(x_n) \to c$.

Beispiel. Betrachte $\exp : \mathbb{R} \longrightarrow \mathbb{R}$. Dann gilt $\lim_{x\to-\infty} \exp(x) = 0$. (Bew. Offenbar gilt $\exp(x) \geq x$ fuer $x \geq 0$....)

Grenzwert ist $\pm\infty$: Fuer Funktionen mit Werten in $\mathbb{R}$ gibt es noch die Moeglichkeit daß der Grenzwert von f bei $p \in \mathbb{C}$ gerade $\pm\infty$ ist:
Sei $f : D \longrightarrow \mathbb{R}$, p Berührpunkt von D. Dann definiert man: $\lim_{x\to p} f(x) = \pm\infty :\Longleftrightarrow$ Für jedes $C \in \mathbb{R}$ existiert $\delta > 0$ mit $f(x) > C$ bzw. $f(x) < C$ für alle $x \in D$ mit $|x - p| \leq \delta :\Longleftrightarrow$ Für jede Folge (x_n) in D mit $x_n \to \pm p$ gilt $f(x_n) \to \pm\infty$.

Beispiel. Betrachte $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ mit $f(x) = 1/x$. **Zeichnung.** Dann gilt

$$\lim_{x\to 0+} f(x) = \infty, \quad \lim_{x\to 0-} f(x) = -\infty.$$

Grenzwert bei ∞ bzw $-\infty$ ist $\pm\infty$: Für $f : D \longrightarrow \mathbb{R}$ mit $D \subset \mathbb{R}$ gibt es noch den **weiteren Grenzwert** $\lim_{q\to\infty} f(q) = \pm\infty$: $\lim_{x\to\infty} f(x) = \pm\infty \Longleftrightarrow$ Für alle $C \in \mathbb{R}$ existiert ein $S \in \mathbb{R}$ mit $f(x) \geq C$ bzw.

$f(x) \leq C$ für alle $x \geq S \iff$ Für jede Folge (x_n) in $\mathbb{R}$ mit $x_n \to \infty$ gilt $f(x_n) \to \pm\infty$.
Entsprechend definiert man $\lim_{x\to-\infty} f(x) = \pm\infty$.

Beispiel. $\exp : \mathbb{R} \longrightarrow \mathbb{R}$, $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Dann gilt $\lim_{x\to\infty} f(x) = \infty$. (Bew. $\exp(x) \geq x$...).

Zum Abschluss des Kapitels kommen wir nun noch zu einer Stabilitätseigenschaft der Menge aller stetigen Funktionen auf einer Menge.

Definition (Gleichmäßige Konvergenz). *Seien $D \subset \mathbb{C}$ und $f_n : D \longrightarrow \mathbb{C}$, $n \in \mathbb{N}$ und $f : D \longrightarrow \mathbb{C}$ gegeben. Dann heißt (f_n) gleichmäßig gegen f konvergent, wenn zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit*

$$|f_n(x) - f(x)| \leq \varepsilon$$

für alle $x \in D$ und $n \geq N_\varepsilon$.

Bemerkung. Eine Folge $f_n : D \subset \mathbb{C} \longrightarrow \mathbb{C}$ heißt punktweise konvergent gegen $f : D \longrightarrow \mathbb{C}$, wenn $f_n(x) \to f(x)$, $n \to \infty$, fuer jedes $x \in D$ gilt. Offenbar impliziert gleichmaßige Konvergenz die punktweise Konvergenz. Es gilt aber nicht die Umkehrung (s.u.).

Theorem (Gleichmäßiger Grenzwert stetiger Funktionen ist stetig). *Seien $D \subset \mathbb{C}$ und $f_n : D \longrightarrow \mathbb{C}$, $n \in \mathbb{N}$, und $f : D \longrightarrow \mathbb{C}$ gegeben. Sind alle f_n stetig und konvergiert (f_n) gleichmäßig gegen f, so ist auch f stetig.*

Beweis. Der Beweis wird mit einem sogenannten $\frac{\varepsilon}{3}$-Argument geführt. Sei $p \in D$ gegeben. Sei $\varepsilon > 0$ beliebig.

- Aufgrund der gleichmäßigen Konvergenz existiert ein $N \in \mathbb{N}$ mit $|f(q) - f_N(q)| < \varepsilon/3$ für alle $q \in D$.
- Da f_N stetig ist, existiert ein $\delta > 0$ mit $|f_N(q) - f_N(p)| < \varepsilon/3$ für alle $q \in D$ mit $|q - p| < \delta$.

Damit gilt für q mit $|q - p| < \delta$ also

$$\begin{aligned} |f(q) - f(p)| &= |f(q) - f_N(q) + f_N(q) - f_N(p) + f_N(p) - f(p)| \\ &\leq |f(q) - f_N(q)| + |f_N(q) - f_N(p)| + |f_N(p) - f(p)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon \end{aligned}$$

Das beendet den Beweis. □

Bemerkung. Der Beweis zeigt eigentlich noch mehr, nämlich eine punktweise Aussage: Konvergiert (f_n) gleichmäßig gegen f und sind alle f_n in p stetig, so ist auch f in p stetig.

Beispiele.

- Sei $f_n : \mathbb{R} \longrightarrow \mathbb{R}$, $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$ und $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = |x|$. Dann konvergiert f_n gleichmäßig gegen f.

Bew. $|f_n(x) - f(x)| \leq \sqrt{\frac{1}{n}} \to 0$, $n \to \infty$ (unabhängig von x!).

- Sei $f_n : [0,1] \longrightarrow [0,1]$, $x \mapsto x^n$ und $f : [0,1] \longrightarrow \mathbb{R}$, $f(x) = 0$, $x < 1$ und $f(1) = 1$. Dann gilt $f_n(x) \to f(x)$ für jedes $x \in [0,1]$, f_n stetig für jedes n, aber f_n konvergiert nicht gleichmäßig. (In diesem Beispiel liegt also punktweise Konvergenz vor, aber keine gleichmäßige Konvergenz.)

Bew. Das kann man direkt sehen durch Untersuchen von x, die nahe an 1 liegen (**Zeichnung.**), oder aus dem vorigen Satz folgern, da f nicht stetig ist.

KAPITEL 9

Funktionen auf Intervallen

In diesem Kapitel geht es um Funktionen

$$f : I \longrightarrow \mathbb{R},$$

wobei $I \subset \mathbb{R}$ ein Intervall ist. Dabei untersuchen wir zunächst stetige Funktionen. Für diese Funktionen gibt es (eigentlich nur) zwei Sätze und noch einen weiteren ;-) Diese Sätze lernen wir in diesem Abschnitt kennen.
Anschließend diskutieren wir noch eine weitere Klasse von Funktionen, nämlich die monotonen Funktionen. Für diese lässt sich Stetigkeit der Funktion und der Umkehrfunktion leicht charakterisieren. Außerdem erweist es sich, daß diese Funktionen automatisch gewisse Stetigkeitseigenschaften haben.

Zeichnung. Stetige Funktion auf einem Intervall.

THEOREM (Zwischenwertsatz). *Sei $-\infty < a < b < \infty$ und $f : [a, b] \longrightarrow \mathbb{R}$ stetig. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an.* ***Zeichnung.***

Bemerkung. Man mache sich klar, daß die Aussage im allgemeinen falsch ist, wenn f nicht stetig ist oder der Definitionsbereich von f kein Intervall ist.
Beweis. Ohne Einschränkung $f(a) < f(b)$. Sei $f(a) < \gamma < f(b)$ beliebig. Zu zeigen: Es gibt ein $c \in [a, b]$ mit $f(c) = \gamma$.
Beginnend mit $[a_0, b_0] = [a, b]$ konstruieren wir induktiv durch Halbierung der Intervalle eine Intervallschachtelung $[a_n, b_n]$ mit

- $f(a_n) \leq \gamma \leq f(b_n)$,
- $|b_n - a_n| = \frac{1}{2^n}|b - a|$.

auf folgende Weise: $n = 0$: $[a_0, b_0] = [a, b]$. $n \Longrightarrow (n + 1)$: Setze $m_n := \frac{1}{2}(b_n - a_n)$ und betrachte $f(m_n)$. Gilt $f(m_n) \geq \gamma$, so setzt man $[a_{n+1}, b_{n+1}] = [a_n, m_n]$. Gilt $f(m_n) < \gamma$, so setzt man $[a_{n+1}, b_{n+1}] = [m_n, b_n]$.
Offenbar haben dann die so konstruieren $[a_n, b_n]$ die gewuenschten Eigenschaften. Die Folgen (a_n) und (b_n) konvergieren dann gegen den (eindeutigen) Punkt

$$c \in \bigcap_n [a_n, b_n].$$

Aufgrund der (Folgen)-Stetigkeit gilt:

$$\gamma \leq \lim_n f(b_n) = f(c) = \lim_n f(a_n) \leq \gamma.$$

Damit folgt also $\gamma = f(c)$. Das beendet den Beweis. $\square$

Wir geben jetzt noch eine Umformulierung des Satzes.

THEOREM (Zwischenwertsatz - Variante). *Sei I ein Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ stetig. Dann ist $J := f(I)$ ebenfalls ein Intervall. Genauer gilt*

$$(\inf f(I), \sup f(I)) \subset J \subset [\inf f(I), \sup f(I)]$$

falls $\inf f(I)$ *und* $\sup f(I)$ *endlich sind.*

Beweis. Ohne Einschränkung seien $\inf f(I)$ und $\sup f(I)$ endlich.
Zweite Inklusion: klar.
Erste Inklusion: Reicht für jedes $\varepsilon > 0$ zu zeigen, daß gilt

$$(\inf f(I) + \varepsilon, \sup f(I) - \varepsilon) \subset f(I).$$

Wähle dazu $a \in I$ mit $f(a) < \inf f(I) + \varepsilon$ und $b \in I$ mit $f(b) > \sup f(I) - \varepsilon$. Nach dem Zwischenwertsatz gilt dann $[f(a), f(b)] \subset f(I)$. $\square$

Ende der 25. Vorlesung.

Bemerkung.

- Aus der Variante folgt der Zwischenwertsatz (und umgekehrt).
- Ist I ein offenes Intervall, so kann man über Offenheit / Abgeschlossenheit von $f(I)$ im allgemeinen keine Aussage treffen: (Identität auf $(0, 1)$; Hut auf $(0, 1)$). Im Falle von abgeschlossenen Intervallen ist die Lage anders. Das werden wir gleich untersuchen.

Offenbar impliziert der Zwischenwertsatz die folgende Aussage.

FOLGERUNG. *Sei $-\infty < a < b < \infty$ und $f : [a, b] \longrightarrow \mathbb{R}$ stetig. Gilt $f(a) \leq 0 \leq f(b)$, so hat f in $[a, b]$ eine Nullstelle.* ***Zeichnung***

Anwendung - Existenz der Wurzel. Ist $\alpha > 0$ so hat

$$P : [0, \infty) \longrightarrow \mathbb{R}, \; P(x) = x^n - \alpha$$

ein Nullstelle.
Beweis. $P(0) = -\alpha < 0$ und $P(1+\alpha) \geq 1$ (nach Bernoulli-Ungleichung). Anwendung des Folgerung (oder der Zwischenwertsatzes) auf die Einschränkung von P auf $[0, 1 + \alpha]$ liefert dann die Behauptung.

Anwendung- Fixpunkt von Selbstabbildungen Jede stetige Funktion $f : [a, b] \longrightarrow [a, b]$ hat einen Fixpunkt (d.h. es gibt ein $c \in [a, b]$ mit $f(c) = c$). **Zeichnung**
Bew. Betrachte

$$g : [a, b] \longrightarrow \mathbb{R}, \;\; g(x) = f(x) - x.$$

Dann gilt (wegen $f([a,b]) \subset [a,b]$) aber

$$g(a) = f(a) - a \geq 0, \quad g(b) = f(b) - b \leq 0.$$

Aus der Folgerung (oder dem Zwischenwertsatz) folgt Existenz eines c mit $g(c) = 0$ d.h. $f(c) = c$.

Wir kommen nun zum anderen Satz über stetige Funktionen auf Intervallen.

Theorem (Existenz Minimum und Maximum stetiger Funktionen auf abg. beschränkten Intervall). *Sei $-\infty < a < b < \infty$ und $f : [a,b] \longrightarrow \mathbb{R}$ stetig. Dann nimmt f auf $[a,b]$ Minimum und Maximum an, d.h. es gibt p_m und p_M in $[a,b]$ mit*

$$f(p_m) \leq f(p) \leq f(p_M)$$

für alle $p \in [a,b]$.

Bemerkung. Es ist nötig, daß I ein abgeschlossenes und beschränktes Intervall ist (vgl. id auf $(0,1)$ oder exp auf $\mathbb{R}$).

Beweis. Wir zeigen nur die Aussage zum Maximum. Die Aussage zum Minimum kann ähnlich bewiesen werden. Alternativ folgt die Aussage zum Minimum auch durch Anwenden der Aussage zum Maximum auf die Funktion $-f$.

Sei $M := \sup f(I)$. Sei (p_n) eine Folge in $[a,b]$ mit $f(p_n) \to M$.

- Da $[a,b]$ beschränkt ist, existiert nach dem Satz von Bolzano/Weierstraß eine konvergente Teilfolge (p_{n_k}) von (p_n).
- Da $[a,b]$ abgeschlossen ist, gehört der Grenzwert p der konvergenten Teilfolge (p_{n_k}) wieder zu $[a,b]$.

Es gilt also

$$p_{n_k} \to p \in [a,b].$$

Da f in p stetig ist, folgt dann

$$f(p) = \lim_{k\to\infty} f(p_{n_k}) = \lim_n f(p_n) = M.$$

Das beendet den Beweis. □

Nach diesen beiden Sätzen kommen wir nun zum dritten Satz über stetige Funktionen auf Intervallen.

Theorem (Gleichmässige Stetigkeit von stetigen Funktionen auf abgeschlossenen beschränkten Intervallen). *Sei $-\infty < a < b < \infty$ und $f : [a,b] \longrightarrow \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig.*

Beweis. Angenommen Nein! Dann existiert ein $\varepsilon > 0$ 'ohne δ' d.h. mit der Eigenschaft, daß für jedes $n \in \mathbb{N}$ Punkte $x_n, y_n \in [a,b]$ existieren mit

$$|x_n - y_n| < \frac{1}{n} \text{ aber } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Durch Übergang zu einer Teilfolge können wir ohne Einschränkung annehmen, daß

$$x_n \to x \in [a,b].$$

Wegen $|y_n - x_n| \leq \frac{1}{n}$ gilt dann auch $y_n \to x \in [a,b]$. Damit folgt

$$\begin{aligned} \varepsilon &\leq |f(x_n) - f(y_n)| \\ &= |f(x_n) - f(x) + f(x) - f(y_n)| \\ &\leq |f(x_n) - f(x)| + |f(x) - f(y_n)| \\ &\to 0, n \to \infty. \end{aligned}$$

Widerspruch. □

Bemerkung. Für die beiden vorangegangenen Schlüsse ist die entscheidende Eigenschaft der abgeschlossenen beschränkten Intervalle I die folgende:

(K) Jede Folge mit Werten in I hat eine konvergente Teilfolge, deren Grenzwert wieder in I liegt.

Die Schlüsse (und damit auch die Aussagen) gelten für jede andere Menge I mit dieser Eigenschaft ebenfalls. Solche Mengen heißen kompakt (s.u.).

Wir kommen nun zu monotonen Funktionen auf Intervallen.

Definition (Monotonie). *Sei $D \subset \mathbb{R}$ und $f : D \longrightarrow \mathbb{R}$. Dann heißt f monoton wachsend/fallend, wenn für alle $x, y \in D$ mit $x < y$ gilt*

$$f(x) \leq f(y) \;\; / f(x) \geq f(y).$$

Gilt eine strikte Ungleichung, so heißt f streng oder strikt monoton wachsend/fallend.

Beispiele.

- Die **k-te Potenz** auf $[0,\infty)$ d.h. die Funktion $(\cdot)^k : [0,\infty) \longrightarrow [0,\infty), x \mapsto x^k$ ist streng monoton.
  Bew. $y > x$ impliziert $y = x + h$ mit $h > 0$. Damit folgt Aussage aus binomischem Satz:
- Die **k-te Wurzel** $\sqrt{k} : [0,\infty) \longrightarrow [0,\infty)$ ist streng monoton wachsend.
  Bew. Das wissen wir schon.
- Auf $\mathbb{R}$ ist die **Exponentialfunktion** $\exp : \mathbb{R} \longrightarrow [0,\infty), x \mapsto \sum_{n\geq 0} \frac{1}{n!}x^n$ streng monoton wachsend.
  Bew. Das folgt mit der Funktionalgleichung: $y > x$ impliziert $y = x + h$ mit $h > 0$. Damit folgt

  $$\exp(y) = \exp(x+h) = \exp(x)\exp(h) > \exp(x)$$

  da für $h > 0$ gilt $\exp(h) = 1 + ... > 1$.

Lemma (Hauptlemma monotone Funktionen). *Sei $D \subset \mathbb{R}$ gegeben und $f : D \longrightarrow \mathbb{R}$ monoton wachsend. Ist $c \in D$ ein Beruehrpunkt von $D \cap (-\infty, c)$, so existiert der halbseitige Grenzwerte $f(c_-) := \lim_{x\to c-} f(x)$ und es gilt $f(c_-) \leq f(c)$. Ist $c \in D$ ein Beruehrpunkt von $D \cap (c, \infty)$, so existiert der halbseitige Grenzwert $f(c_+) := \lim_{x\to c+} f(x)$ und es gilt $f(c) \leq f(c_+)$. Entsprechendes gilt fuer monoton fallende Funktionen.*

Beweis. Wir zeigen nur die Aussage zum Grenzwert von links. Die andere Aussage folgt analog.

Es ist $\{f(x) : x < c\}$ durch $f(c)$ nach oben beschränkt und besitzt daher ein Supremum M (und dieses Supremum erfuellt $M \leq f(c)$). Wir zeigen $\lim_{x\to c-} f(x) = M$.

Ist nun $\varepsilon > 0$ beliebig, so existiert ein $x_\varepsilon < c$ mit $M - \varepsilon \leq f(x_\varepsilon) \leq M$. Setzt man $\delta := c - x_\varepsilon$, so gilt dann also für alle $x < c$ mit $|x - c| < \delta$ auch $x_\varepsilon < x < c$. **Zeichnung.** Damit gilt für solche x dann aufgrund der Monotonie

$$M - \varepsilon < f(x_\varepsilon) \leq f(x) \leq M$$

also $|f(x) - M| < \varepsilon$. □

Ende der 26. Vorlesung

Theorem (Stetigkeitseigenschaft monotoner Funktionen). *Sei $D \subset \mathbb{R}$ gegeben und $f : D \longrightarrow \mathbb{R}$ monoton. Dann ist f in einem Punkt $c \in D$ entweder stetig, oder hat dort eine Sprungstelle (*

Beweis. Das folgt sofort aus dem vorigen Lemma. □

Bemerkung. Fuer monoton wachsende f und ein $c \in D$, das Beruehrpunkt von $D \cap (-\infty, c)$ und $D \cap (c, \infty)$ ist gibt es vier Moeglichkeiten: **Zeichnung.**

- Es gilt $f(c_-) = f(c_+)$. Dann ist f stetig.
- Es gilt $f(c_-) = f(c)$ und $f(c) < f(c_+)$ und f ist unstetig.
- Es gilt $f(c_-) < f(c) = f(c_+)$ und f ist unstetig.
- Es gilt $f(c_-) < f(c) < f(c_+)$ und f ist unstetig.

Folgerung. *Sei I ein Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ monoton. Dann ist f genau dann stetig, wenn $f(I)$ ein Intervall ist.*

Beweis. Es sind zwei Implikationen zu zeigen.

Sei f stetig. Dann ist $f(I)$ ein Intervall nach dem Zwischenwertsatz. (Das hat nichts mit Monotonie zu tun).

Sei $f(I)$ ein Intervall. Dann kann f keine Sprungstelle haben. Damit folgt aus dem vorigen Theorem die Stetigkeit von f. □

Wir untersuchen die Umkehrfunktionen monotoner Funktionen und insbesondere deren Stetigkeit: Ist $D \subset \mathbb{R}$ und

$$f : D \longrightarrow f(D) \subset \mathbb{R}$$

streng monoton, so ist f offenbar injektiv (und natuerlich surjektiv). Also existiert die Umkehrfunktion

$$g = f^{-1} : f(D) \longrightarrow \mathbb{R}, \text{ mit } g \circ f = \mathrm{id}_D \text{ und } f \circ g = id_{f(D)},$$

(d.h. $g(y) = x$ falls $f(x) = y$). Ist f streng monoton wachsend/fallend, so ist auch g streng monoton wachsend/fallend.
(Bew. Ohne Einschränkung f streng monoton wachsend. Sei $y < y'$ und x, x' mit $f(x) = y$, $f(x') = y'$. Dann gilt also $x \neq x'$. Wäre $x > x'$, so folgte $y = f(x) > f(x') = y'$. Widerspruch.)

Der Graph der Funktion f ist gegeben durch

$$\text{Graph von f} = \{(x, f(x)) : x \in D\}.$$

Der Graph der Umkehrfunktion g ist dann gegeben durch

$$\text{Graph von g} = \{(y, g(y)) : y \in f(D)\} = \{(f(x), x) : x \in D\}.$$

Es entsteht also der Graph der Umkehrfunktion aus dem Graphen der Funktion durch Vertauschen der Koordinaten. Geometrisch entspricht das Vetauschen der Koordination gerade dem Spiegeln an der Winkelhalbierenden **Zeichnung.**. Es entsteht die Umkehrfunktion also durch Spiegeln an der Winkelhalbierenden.

Beispiel: k-te Potenz und k-te Wurzel. $f : [0, \infty) \longrightarrow \mathbb{R}$, $f(x) = x^k$. Dann ist f streng monoton wachsend mit $f([0, \infty) = [0, \infty)$ und die Umkehrfunktion $g : [0, \infty) \longrightarrow \mathbb{R}$ ist gegeben durch $g(y) = \sqrt[k]{y}$.

Theorem (Umkehrfunktion streng monotoner Funktionen auf Intervallen ist stetig). *Sei I ein Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ streng monoton. Dann ist die Umkehrfunktion $g = f^{-1} : f(I) \longrightarrow I \subset \mathbb{R}$ stetig.*

Beweis. Es ist g streng monoton (wie oben gezeigt). Wegen $g(f(I)) = I$ hat g keine Sprungstellen. Also ist g stetig nach dem Theorem zu den Stetigkeitseigenschaften monotoner Funktionen. □

Bemerkung. (Übung) Ist $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ stetig, so gilt:

$$f \text{ invertierbar} \Longleftrightarrow f \text{ streng monoton.}$$

(Ist I kein Intervall, so gilt diese Äquivalenz nicht.)

Beispiel: Exponentialfunktion und Logarithmus. Sei $\exp : \mathbb{R} \longrightarrow (0, \infty)$, $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Dann ist exp streng monoton wachsend mit $\exp(\mathbb{R}) = (0, \infty)$ und die Umkehrfunktion $\ln : (0, \infty) \longrightarrow \mathbb{R}$ ist stetig. **Zeichnung.**
Bew. Es reicht zu zeigen, daß $\exp(\mathbb{R}) = (0, \infty)$ gilt: Wir wissen schon

$$\exp(x)^{-1} = \exp(-x) \text{ und } \exp(x) > 0$$

für alle $x \in \mathbb{R}$. Weiterhin gilt offenbar für alle $k \in \mathbb{N}$

$$\exp(k) = \sum_{n=0}^{\infty} \frac{k^n}{n!} \geq k$$

und damit also auch

$$\exp(-k) < \frac{1}{k}.$$

Also enthält (nach Zwischenwertsatz) $\exp(\mathbb{R})$ das Intervall $[1/k, k]$ für jedes $k \in \mathbb{N}$. Da $k \in \mathbb{N}$ beliebig ist, folgt die gewünschte Behauptung.

Die Umkehrfunktion der Exponentialfunktion heißt *Logarithmus* bzw. mit vollem Namen *natuerlicher Logarithmus* oder auf Latein, *logarithmus naturalis*). Sie wird mit ln bezeichnet. Der Logarithmus wird uns immer wieder begegenen (wie auch die Exponentialfunktion). Die Funktionalgleichung der Exponentialfunktion führt zur Gültigkeit der folgenden Gleichung

$$\ln x + \ln y = \ln(xy) \text{ und } \ln x = -\ln \frac{1}{x}$$

für $x, y > 0$.
Bew. $x = e^a$, $y = e^b$. Dann gilt

$$\ln x + \ln y = \ln e^a + \ln e^b = a + b = \ln(e^{a+b}) = \ln(xy).$$

Auch Monotonie ist stabil unter Konvergenz von Funktionen und zwar sogar unter punktweiser Konvergenz.

Proposition. *Sei I ein Intervall und $f_n : I \longrightarrow \mathbb{R}$, $n \in \mathbb{N}$, seien monotone wachsende Funktionen. Gibt es ein $f : I \longrightarrow \mathbb{R}$ mit $f_n(x) \to f(x)$ für alle $x \in I$, so ist f ebenfalls monoton wachsend. Entsprechendes gilt für monoton fallende Funktionen.*

Beweis. Sei $x < y$. Dann gilt

$$f(x) = \lim_n f_n(x) \leq \lim_n f_n(y) = f(y).$$

Das beendet den Beweis. □

Man könnte denken, daß Funktionen auf einem Intervall, die stetig und monoton sind, besonders einfach sind. Das ist nicht der Fall, wie man an folgendem Beispiel sieht.

Beispiel - Teufelstreppe. Sei $I = [0, 1]$.
Setze $C_0 := I$ und konstruiere rekursiv durch Herausnehmen der offenen mittleren Drittelintervalle abgeschlossene Mengen $C_n \subset I$ mit $C_n \subset C_{n-1}$. **Zeichnung.** Damit besteht also C_n aus 2^n Intervallen der Länge $1/3^n$. Die 'Gesamtlänge' von C_n ist damit

$$L_n = \frac{1}{3^n} 2^n$$

und die 'Gesamtlänge' des Komplementes $I \setminus C_n$ ist gegeben durch

$$L_n = 1 - \frac{1}{3^n} 2^n.$$

Sei

$$C := \bigcap_{n=1} C_n.$$

Dann ist die 'Gesamtlänge' des Komplementes $I \setminus C$ also gerade 1 und C hat die 'Länge' 0. Nun definieren wir für $n \in \mathbb{N}$ die Funktion

$$f_n : I \longrightarrow [0,1]$$

auf folgende Weise: **Zeichnung**

- $f_n(0) = 0$, $f_n(1) = 1$.
- Auf den herausgenommenen $2^n - 1$ ($= 1+2+\cdot+2^{n-1}$) Intervallen ist die Funktion konstant mit den Werten $1/2^n, 2/2^n, \ldots, (2^n - 1)/2^n$.
- Auf den noch verbliebenen Intervallen wird die Funktion linear interpoliert.

Dann gilt offenbar $|f_n(x) - f_m(x)| \leq 1/2^n$ für $m \geq n$. Damit ist für jedes $x \in I$ also $(f_n(x))$ eine Cauchy-Folge und damit konvergent und für die Grenzwert $f(x)$ gilt

$$|f(x) - f_n(x)| = \lim_{m \to \infty} |f_m(x) - f_n(x)| \leq \frac{1}{2^n}.$$

Damit konvergiert (f_n) also gleichmäßig gegen f und f ist stetig und monoton mit $f(0) = 0$ und $f(1) = 1$ und f konstant auf den einzelnen Stücken von $I \setminus C$. Diese Funktion kann also nicht gezeichnet werden.

Bemerkung.

- In gewisser Weise sind Funktionen wie obige die typischen stetigen monotonen Funktionen.
- Unsere Anschauung vom Zeichnen von Funktionen ist insofern irreführend, als wir (eigentlich) nur Funktionen zeichnen können, die - von wenigen Ausnahmepunkten abgesehen - lokal lipschitzstetig sind.
- Im nächsten Abschnitt werden wir differenzierbare Funktionen kennenlernen. Für diese Klasse 'gelten' unsere Zeichnungen.

KAPITEL 10

Differenzierbare Funktionen

Eine Funktion ist in einem Punkt stetig, wenn sie in einer Umgebung des Punktes gut durch eine konstante Funktion approximiert wird.
Eine Funktion ist in einem Punkt differenzierbar, wenn sie in der Nähe des Punktes sehr gut durch eine linear affine Funktion, ihre Tangente, approximiert wird. Die zugehörige lineare Funktion (gegeben durch eine 1×1 Matrix d.h. eine Zahl) heißt die Ableitung.
Etwas praeziser laeßt sich das so formulieren: Sei $f : (a,b) \longrightarrow \mathbb{R}$ und $p \in (a,b)$ gegeben. Dann gilt:

- f stetig in $p : f(q) = f(p) + \psi(q)$, wobei der Fehler $\psi(q)$ klein wird für $q \to p$, d.h. f ist gut durch eine konstante Funktion approximierbar in p.
- f differenzierbar in p: $f(q) = f(p) + b(q-p) + \varphi(q)$, wobei der φehler $\varphi(q)$ sehr klein wird für $q \to p$, nämlich $\varphi(q) = \psi(q)(q-p)$ mit $\psi(q) \to 0$, $q \to p$. Damit ist f in p sehr gut durch eine lineare Funktion, naemlich
$$T(x) = f(p) + b(x-p)$$
approximierbar. Diese Funktion wird Tangente genannt.

Um über 'Nähe' eines Punktes sprechen zu können, brauchen wir Platz. Daher wird es sich meist um Funktionen auf offenen Intervalle handeln. 'Sehr gute Approximation' wird bedeuten, daß der 'Fehler' klein ist. Differenzierbare Funktionen haben viele gute Eigenschaften. Im wesentlichen liefern unsere Zeichnungen meist differenzierbare Funktionen.

1. Definition und grundlegende Eigenschaften von Differenzierbarkeit in einem Punkt.

Es stellt sich heraus, daß die erwaehnte sehr gute Approximierbarkeit aequivalent ist zur Existenz eines Grenzwertes. Das ist der Inhalt des folgenden Lemma.

Lemma (Charakterisierung der Differenzierbarkeit in einer Dimension). *Sei $I \subset \mathbb{R}$ offenes Intervall und $f : I \longrightarrow \mathbb{R}$ und $p \in I$ gegeben. Dann sind äquivalent:*

(i) *Es existiert ein $b \in \mathbb{R}$ und ein $\varphi : I \longrightarrow \mathbb{R}$ mit*
$$f(x) = f(p) + b(x-p) + \varphi(x),$$

fuer alle $x \in I$, wobei gilt

$$\lim_{x \to p, x \neq p} \frac{\varphi(x)}{|x-p|} = 0.$$

(ii) *Der Grenzwert* $\lim_{x \to p, x \neq p} \frac{f(x)-f(p)}{x-p}$ *existiert.*

In diesem Fall gilt $b = \lim_{x \to p, x \neq p} \frac{f(x)-f(p)}{x-p}$.

Ende der 27. Vorlesung.

Bemerkungen.

- Die Funktionen $x \mapsto \frac{f(x)-f(p)}{x-p}$ und $x \mapsto \frac{\varphi(x)}{x-p}$ sind auf micht auf I sondern auf $I \setminus \{p\}$ definiert. Um daran zu erinnern, schreiben wir die entsprechenden Grenzwerte als $\lim_{x \to p, x \neq p}$. Spaeter werden wir dazu uebergehen, den Zusatz $x \neq p$ wegzulassen.
- Die Forderung
$$\lim_{x \to p, x \neq p} \frac{\varphi(x)}{|x-p|} = 0$$
an φ in (i) spiegelt gerade wieder, dass es sich um eine sehr gute Approximation handeln soll. Ohne eine entsprechende Forderung an φ ist die Aussage in (i) komplett trivial: Fuer jedes $b \in \mathbb{R}$ laßt sich ein φ mit $f(x) = f(p)+b(x-p)+\varphi(x)$ finden. (Denn man braucht ja lediglich φ durch $\varphi(x) = f(x)-f(p)-b(x-p)$ zu definieren.)
- Betrachtet man den sogenannten *Differenzenquotienten*
$$\frac{f(q)-f(p)}{q-p}$$
für $q \to p$, so fällt auf, daß der Nenner gegen Null konvergiert. Damit der Quotient endlich bleibt und gegen einen Grenzwert strebt, muss also der Zähler auch gegen Null streben und zwar im wesentlichen im 'gleichen Masse'. Damit folgt aus Differenzierbarkeit also sinngemaaß die Stetigkeit. Einen praezisen Beweis geben wir unten.
- Wir erwaehnen noch, daß die Ableitung auch manchmal als Differentialquotient bezeichnet wird.

Beweis. $(i) \Longrightarrow (ii)$: Es gilt

$$\frac{f(x)-f(p)}{x-p} = b + \frac{\varphi(x)}{x-p} \to b, \quad x \to p.$$

$(ii) \Longrightarrow (i)$: Setze $b := \lim_{x \to p, x \neq p} \frac{f(x)-f(p)}{x-p}$. Dann erfüllt φ mit

$$f(x) = f(p) + b(x-p) + \varphi(x)$$

also
$$\frac{\varphi(x)}{(x-p)} = \frac{f(x)-f(p)}{x-p} - b \to 0, x \to p.$$
Weitere Aussage: Wurde schon gezeigt und zwar sowohl in $(ii) \Longrightarrow (i)$ als auch in $(i) \Longrightarrow (ii)$. $\square$

DEFINITION (Differenzierbarkeit). *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : (a,b) \longrightarrow$ gegeben. Dann heißt f in $p \in I$ differenzierbar, wenn der Grenzwert*
$$\lim_{x\to p, x\neq p} \frac{f(x)-f(p)}{x-p}$$
existiert. Dieser Grenzwert heißt dann die Ableitung von f in p und wird mit $f^{'}(p)$ oder $Df(p)$ oder $\frac{df}{dx}(p)$ bezeichnet. Ist f in allen Punkten von I differenzierbar, so heißt f differenzierbar.

Hinweis: Ausrechnen des Grenzwertes. Es gilt
$$\lim_{x\to p, x\neq p} \frac{f(x)-f(p)}{x-p} = \lim_{h\to 0, h\neq 0} \frac{f(p+h)-f(p)}{h}.$$

Geometrische Deutung.

- Der Differenzenquotient $\frac{f(x)-f(p)}{x-p}$ ist gerade die Steigung der Sekanten durch $(x, f(x))$ und $(p, f(p))$. **Zeichnung.**
- Die Gerade durch $(p, f(p))$ mit Steigung $f^{'}(p)$ (falls die Ableitung existiert) heißt Tangente an f im Punkt p. Sie ist gegeben durch
$$T(x) = f(p) + f^{'}(p)(x-p).$$
Zeichnung.
- Die Existenz der Ableitung bedeutet dann gerade, daß die Sekanten (punktweise) gegen die Tangente konvergieren.
- Die Gleichung
$$f(x) = f(p) + f^{'}(p)(x-p) + \varphi(x)$$
mit $\varphi(x)/(x-p) \to 0$, $x \to p$, bedeutet gerade, daß f sehr gut durch die Tangente approximiert wird.
- Es gilt
$$\lim_{x\to p, x\neq p} \frac{\varphi(x)}{|x-p|} = 0$$
genau dann, wenn gilt
$$\varphi(x) = \psi(x)(x-p)$$
mit $\psi(x) \to 0$ für $x \to p$, nämlich
$$\psi(x) := \begin{cases} \frac{\varphi(x)}{x-p} & : \quad x \neq p \\ 0 & : \quad x = p. \end{cases}$$

Proposition (Differenzierbarkeit impliziert Stetigkeit). *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \longrightarrow \mathbb{R}$ differenzierbar in $p \in I$. Dann ist f stetig in p.*

Beweis. Es gilt
$$f(q) = f(p) + b(q - p) + \varphi(q)$$
mit $\varphi(q) = 0$ für $q = p$ und $\varphi(q) = \frac{\varphi(q)}{q-p}(q-p)$ für $q \neq p$. Damit gilt also $\varphi(q) \to 0$ für $q \to p$ (und die Konvergenz ist sogar ziemlich schnell). Es folgt durch Betrachten der drei Terme in der Gleichung dann sofort $f(q) \to f(p)$ für $q \to p$ d.h. Stetigkeit von f in p. □

Wir kommen nun zu einigen Rechenregeln im Umgang mit Ableitungen.

Proposition (Rechenregeln). *Sei I ein offenes Intervall in $\mathbb{R}$ und $f, g : (a, b) \to \mathbb{R}$ differenzierbar in $p \in I$. Dann gilt:*

(a) *$f + \alpha g$ ist differenzierbar in p mit $(f + \alpha g)'(p) = f'(p) + \alpha g'(p)$.*

(b) *(Produktregel) Das Produkt fg ist differenzierbar in p mit*
$$(fg)'(p) = f'(p)g(p) + f(p)g'(p).$$

(c) *(Quotientenregel) Gilt $g(p) \neq 0$, so verschwindet g in einer Umgebung von p nicht und es ist ist auch f/g auf dieser Umgebung definiert und differenzierbar in p mit*
$$\left(\frac{f}{g}\right)'(p) = \frac{f'(p)g(p) - f(p)g'(p)}{g^2(p)}.$$
Insbesondere gilt
$$(1/g)'(p) = -g'(p)/g^2(p).$$

Beweis. (a) Das folgt direkt aus den Rechenregeln fuer Grenzwerte.

(b) Es gilt nach Addieren einer geeigneten Null
$$\frac{f(x)g(x) - f(p)g(p)}{x - p} = \frac{f(x) - f(p)}{x - p} g(x) + f(p) \frac{g(x) - g(p)}{x - p}.$$
Bilden des Grenzwertes $x \to p$ liefert dann sofort die Behauptung.

(c) Es reicht das 'Insbesondere' zu zeigen. (Dann folgt erste Aussage durch Anwenden der Produktregel.) Es gilt
$$\frac{1/g(x) - 1/g(p)}{x - p} = -\frac{g(p) - g(x)}{x - p} \frac{1}{g(x)g(p)}.$$
Bilden des Grenzwertes $x \to p$ liefert nun die Behauptung. □

Proposition (Kettenregel). *Seien I, J offene Intervalle in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ und $g : J \longrightarrow I$ gegeben. Ist g differenzierbar in p und f differenzierbar in $g(p)$, so ist $f \circ g : I \longrightarrow \mathbb{R}$ differenzierbar in p, und es gilt*

$$(f \circ g)'(p) = f'(g(p))g'(p).$$

Bemerkung. In diesem Zusammenhang bezeichnet man $f^{'}(g(p))$ als die *aeußere Ableitung* und $g^{'}(p)$ als die *innere Ableitung.*

Beweis. Die Idee ist klar:

$$\frac{f \circ g(x) - f \circ g(p)}{x - p} = \frac{f \circ g(x) - f \circ g(p)}{g(x) - g(p)} \frac{g(x) - g(p)}{x - p} \to f'(g(p))g'(p).$$

Da $g(x) - g(p) = 0$ möglich ist, muss man etwas genauer sein. Man ersetzt dazu für $g(x) = g(p)$ den Quotienten $(f(g(x)) - f(g(p)))/(g(x) - g(p))$ durch $f'(g(p))$. Genauer definiert man die die Funktion f^* auf I durch

$$f^*(\eta) = \begin{cases} \frac{f(\eta) - f(g(p))}{\eta - g(p)}, & : \quad \eta \neq g(p) \\ f'(g(p)) & : \quad \eta = g(p), \end{cases}$$

(wobei $\eta = g(x)$ sein wird).
Dann gilt

- $f(\eta) - f(g(p)) = f^*(\eta)(\eta - g(p))$ für **alle** η (nach Konstruktion).
- $\lim_{\eta \to q} f^*(\eta) = f'(q) = f^*(q)$ (da f in $g(p)$ differenzierbar ist).

Damit können wir nun mit $\eta = g(x)$ fuer $x \neq p$ schließen

$$\frac{f \circ g(x) - f \circ g(p)}{x - p} = f^*(g(x)) \frac{g(x) - g(p)}{x - p}.$$

Nun folgt die gewuenschte Behauptung sofort durch den Grenzuebergang $x \to p$. □

Ende der 28. Vorlesung.

Beispiele differenzierbarer Funktionen.

Konstante Funktion. Die konstante Funktion hat Ableitung 0.
Bew. klar.

k-te Potenz. Die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^k$ hat die Ableitung $f'(x) = kx^{k-1}$.
Bew. $\frac{f(x) - f(p)}{x - p} = \frac{x^k - p^k}{x - p} = \sum_{j=0}^{k-1} x^j p^{k-1-j} \to kp^{k-1}$.

Polynom. Das Polynom $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = \sum_{j=0}^{n} a_j x^j$ hat Ableitung $\sum_{j=1}^{n} a_j j x^{j-1}$.
Bew. Linearkombination von Potenzen.

Inverse Potenz. Die Funktion $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$, $f(x) = 1/x^k$ hat Ableitung $f'(x) = -k \frac{1}{x^{k+1}}$.
Bew. Quotientenregel liefert $f'(x) = -\frac{kx^{k-1}}{x^{2k}}$.

Exponentialfunktion. Es gilt für jedes $z_0 \in \mathbb{C}$

$$\lim_{z \to z_0, z \neq z_0} \frac{\exp(z) - \exp(z_0)}{z - z_0} = \exp(z_0).$$

Insbesondere hat die Funktion $\exp : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto \exp x$ die Ableitung $\exp(x)$.
Bew. Aufgrund der Funktionalgleichung gilt

$$\frac{\exp(z) - \exp(z_0)}{z - z_0} = \exp(z_0) \frac{\exp(z - z_0) - 1}{z - z_0}.$$

Daher reicht es $\lim_{h \to 0} \frac{\exp(h)-1}{h-0} = 1$ zu zeigen. Dazu:

$$\begin{aligned} \frac{\exp(h) - 1}{h} &= \frac{1}{h} \left(\sum_{k=0}^{\infty} \frac{h^k}{k!} - 1 \right) \\ &= \frac{1}{h} \sum_{k=1}^{\infty} \frac{h^k}{k!} \\ &= 1 + \sum_{k=2}^{\infty} \frac{h^{k-1}}{k!} \\ &= 1 + h \sum_{k=2}^{\infty} \frac{h^{k-2}}{k!}. \end{aligned}$$

Für $|h| < 1$ gilt $|\sum_{k=2}^{\infty} \frac{h^{k-2}}{k!}| \leq \sum_{k=2}^{\infty} \frac{1}{k!} \leq \exp(1)$. Damit folgt

$$\frac{\exp(h) - 1}{h} \to 1, \quad h \to 0.$$

Die Funktionen $\sin$ **und** $\cos$**.** Die Funktionen $\sin : \mathbb{R} \to \mathbb{R}$, $\sin(x) = \Im \exp(ix)$ und $\cos : \mathbb{R} \longrightarrow \mathbb{R}$, $\cos(x) = \Re(\exp(ix)$ sind differenzierbar mit $\sin' = \cos$ und $\cos' = -\sin$. Insbesondere gilt $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

Bemerkung (Beschreibung von $\sin$ **und** $\cos$ **am Einheitskreis)**: Eine kleine Rechnung zeigt, daß für $u = \exp(ix)$ gilt $|u| = 1$. Damit handelt es sich also bei sin und cos in der Tat um die Verhältnisse der Seiten eines rechtwinklichen Dreiecks am Einheitskreis. **Zeichnung.**

Bew. Wir betrachten nur sin. (Der andere Fall kann analog behandelt werden.)

$$\begin{aligned} \frac{\sin(x) - \sin(p)}{x - p} &= \Im \frac{\exp(ix) - \exp(ip)}{x - p} \\ &= \Im i \frac{\exp(ix) - \exp(ip)}{ix - ip} \\ &= \Re \frac{\exp(ix) - \exp(ip)}{ix - ip} \\ &\to \Re \exp(ip) = \cos(p). \end{aligned}$$

Zum 'Insbesondere':

$$\frac{\sin x}{x} = \frac{\sin x - 0}{x - 0} = \frac{\sin x - \sin 0}{x - 0} \to \sin'(0) = \cos(0) = 1.$$

Potenzreihe. Sei (a_n) Folge in $\mathbb{C}$ mit $\rho := \limsup \sqrt[n]{|a_n|} < \infty$ und $R := \frac{1}{\rho}$ (falls $\rho > 0$) bzw. $R = \infty$ (falls $\rho = 0$) und

$$f : U_R(0) \longrightarrow \mathbb{C}, f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Dann gilt

$$\lim_{w \to z, w \neq z} \frac{f(w) - f(z)}{w - z} = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

Sind die a_n, $n \in \mathbb{N}$, alle reell, so ist die Einschränkung von f auf $(-R, R)$ differenzierbar mit $f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$.

Bemerkung. Es handelt sich bei der Ableitung um die durch formale Differentiation der Summanden gewonnene Funktion. Das Problem besteht darin, die Vertauschung von Summenbildung und Ableitung zu begründen. Wie üblich verlangt diese Vertauschung von Grenzwerten Arbeit!

Bew. Der Konvergenzradius von $\sum_{n=1}^{\infty} n a_n z^{n-1}$ ist offenbar derselbe wie der Konvergenzradius von $\sum_{n=1}^{\infty} n a_n z^n$. Letzterer ist aber nach den schon bekannten Formeln gerade gegeben als Kehrwert von

$$\limsup_{n \to \infty} \sqrt[n]{n|a_n|} = \limsup_{n \to \infty} \sqrt[n]{|a_n|} = \rho$$

(mit der Vereinbarung, dass dieser Kehrwert als ∞ zu gelten hat, wenn $\rho = 0$ gilt.) Insbesondere ist also

$$g : U_R(0) \longrightarrow \mathbb{C}, z \mapsto \sum_{n=1}^{\infty} n a_n z^{n-1}$$

wohldefiniert durch eine absolut konvergente Reihe. Wir betrachten nun

$$\frac{f(z) - f(w)}{z - w} = \frac{1}{z - w} \sum_{n=1}^{\infty} a_n (z^n - w^n) = \sum_{n=1}^{\infty} a_n \left(\sum_{k=0}^{n-1} z^k w^{n-k-1} \right).$$

Damit folgt für jedes $N \in \mathbb{N}$

$$\begin{aligned} &\frac{f(z) - f(w)}{z - w} - g(w) \\ &= \left[\sum_{n=1}^{\infty} a_n \sum_{k=0}^{n-1} z^k w^{n-k-1} \right] - \sum_{n=1}^{\infty} n a_n w^{n-1} \\ &= \left[\sum_{n=1}^{N} a_n \sum_{k=0}^{n-1} \left(z^k w^{n-k-1} - w^{n-1} \right) \right] + \left[\sum_{n=N+1}^{\infty} a_n \sum_{k=0}^{n-1} z^k w^{n-k-1} \right] - \sum_{n=N+1}^{\infty} n a_n w^{n-1}. \end{aligned}$$

Sei $S > 0$ mit $|w| < S < R$ beliebig. Ist z so nahe an w, daß auch $|z| < S$ gilt, so kann man jeden einzelnen der beiden letzten Summen im Betrag abschätzen durch

$$\sum_{n=N+1}^{\infty} n|a_n|S^{n-1}.$$

Dieser Term geht für $N \to \infty$ gegen 0. Außerdem geht bei festem N aber der erste Term für $z \to w$ gegen 0. Damit folgt die gewünschte Konvergenz

$$\left|\frac{f(z)-f(w)}{z-w} - g(w)\right| \to 0$$

direkt. (Hier sind die Details: Wähle zu beliebigem $\varepsilon > 0$ zunächst $N \in \mathbb{N}$ so groß, daß die obigen beiden letzten Terme kleiner als $\varepsilon/3$ sind für $|z| < S$. Wähle anschließend $\delta > 0$ so klein, daß für $|z-w| < \delta$ sowohl $|z| < S$ also auch

$$\left|\sum_{n=1}^{N} a_n (\sum_{k=0}^{n-1} z^k w^{n-k-1}) - nw^{n-1})\right| < \varepsilon/3$$

gilt. (Das ist möglich, da es sich um eine endliche Summe stetiger Funktionen handelt).)

Wir diskutieren nun einige Beispiele nicht differenzierbarer Funktionen.

Beispiel - nichtdifferenzierbare Funktion

- Die Heavisidefunktion $H : \mathbb{R} \longrightarrow \mathbb{R}$, $H(x) = 1$ fuer $x \geq 0$ und $H(x) = 0$ fuer $x < 0$ ist in $x \neq 0$ differenzierbar mit $H^{'}(x) = 0$. In $x = 0$ ist sie nicht differenzierbar.
  Zeichnung.
- Die Funktion abs : $\mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto |x|$ ist in $x \neq 0$ differenzierbar und in $x = 0$ nicht differenzierbar. (Dort existieren die links und rechtsseitigen Ableitungen und sind verschieden).
  Zeichnung.

  Bew. Für $p > 0$ gilt $\frac{\text{abs}(x) - \text{abs}(p)}{x-p} = \frac{x-p}{x-p} = 1 \to 1$, $x \to p$.

  Für $p < 0$ gilt $\frac{\text{abs}(x) - \text{abs}(p)}{x-p} = \frac{-x-(-p)}{x-p} = -1 \to -1$, $x \to p$.

  Für $p = 0$ gilt $\frac{\text{abs}(x) - \text{abs}(p)}{x-p} = |x|/x = \pm 1$ für $x > 0$ bzw. $x < 0$. Damit existiert die Ableitung nicht.
- Die Funktion $u : \mathbb{R} \longrightarrow \mathbb{R}$ mit $u(x) = 0$ falls x rational ist und $u(x) = x^2$ falls x irrational ist, ist in allen $x \neq 0$ unstetig also nicht differenzierbar und in $x = 0$ differenzierbar mit Ableitung Null. (Übung).

Proposition. *(Ableitung der Umkehrfunktion) Sei $f : (a, b) \longrightarrow \mathbb{R}$ streng monoton, stetig und differenzierbar in p mit $f'(p) \neq 0$. Dann ist die Umkehrfunktion $g := f^{-1} : f(a, b) \longrightarrow \mathbb{R}$ differenzierbar in $q = f(p)$, und es gilt*

$$g'(q) = \frac{1}{f'(p)} = \frac{1}{f'(g(q))}.$$

Zeichnung. Spiegeln

Bemerkung. Die Voraussetzung $f'(p) \neq 0$ ist nötig, andernfalls hat man eine waagerechte Tangente. (**Beispiel.** Quadrat und Wurzel bei 0).

Beweis. Sei $y = f(x)$, $q = f(p)$, also $x = g(y)$, $p = g(q)$. Dann folgt $x \to p$ aus $y \to q$. Damit erhält man

$$\frac{g(y) - g(q)}{y - q} = \frac{x - p}{f(x) - f(p)} \to \frac{1}{f'(p)}$$

für $x \to p$. □

Ende der 29. Vorlesung.

Beispiele (Umkehrfunktion)

- $k - te$ **Wurzel als Umkehrfunktion:** Die Funktion $g : (0, \infty) \longrightarrow \mathbb{R}$, $g(y) = \sqrt[k]{y} = y^{1/k}$ ist differenzierbar mit Ableitung $g'(y) = \frac{1}{k} y^{(1/k)-1}$.

  Bew. Es ist g Umkehrfunktion von $f(x) = x^k$. Damit gilt mit $y = f(x)$, $x = g(y) = y^{1/k}$ also

  $$g'(y) = \frac{1}{f'(x)} = \frac{1}{kx^{k-1}} = \frac{1}{k} \frac{1}{y^{(k-1)/k}}.$$

- **Der Logarithmus als Umkehrfunktion:** Die Funktion $\ln : (0, \infty) \longrightarrow \mathbb{R}$, $g(y) = \ln y$ ist differenzierbar mit $\ln'(y) = 1/y$.

  Bew. ln ist Umkehrfunktion von exp. Damit gilt mit $y = \exp(x)$, $x = \ln y$ also

  $$\ln'(y) = \frac{1}{\exp'(x)} = \frac{1}{\exp(x)} = \frac{1}{y}.$$

Nachdem wir nun den Logarithmus kennen, können wir noch die allgemeine Potenz einführen und auf Differenzierbarkeit untersuchen.

Beispiel. (Allgemeine Potenz) Für $\alpha \in \mathbb{R}$ und $x > 0$ definiert man

$$x^\alpha := e^{\alpha \ln x}$$

und nennt dies die α-te Potenz von x. Dann folgt sofort

$$\ln x^\alpha = \alpha \ln x.$$

Wir koennen nun dies nun bei festem α als Funktion von x sehen oder bei festem x als Funktion von α. Das werden wir nun untersuchen.

Die Funktion $x \mapsto x^\alpha$. Es ist

$$P_\alpha : (0,\infty) \longrightarrow (0,\infty), P_\alpha(x) = x^\alpha$$

differenzierbar mit

$$P_\alpha'(x) = \alpha P_{\alpha-1}(x) = \alpha x^{\alpha-1}.$$

Zeichnung. (für $\alpha < 0$, $0 < \alpha < 1$, $\alpha = 1$, $1 < \alpha$).
Bew. Nach Kettenregel gilt

$$P_\alpha'(x) = e^{\alpha \ln x}\alpha\frac{1}{x} = e^{\alpha \ln x}\alpha\frac{1}{e^{\ln x}} = e^{(\alpha-1)\ln x}\alpha.$$

Die Funktion $\alpha \mapsto x^\alpha$.
Für $a > 0$ ist

$$\exp_a : \mathbb{R} \longrightarrow (0,\infty),\ x \mapsto a^x$$

differenzierbar mit

$$\exp_a'(x) = \ln a \exp_a(x).$$

Bew. Es gilt $\exp_a(x) = e^{x \ln a}$. Damit folgt die Behauptung sofort aus der Kettenregel.

Rechenregeln für die allgemeine Potenz.

$$x^\alpha x^\beta = x^{\alpha+\beta} \text{ sowie } x^\alpha y^\alpha = (xy)^\alpha$$

Bew. Erste Aussage folgt direkt aus der Definition:

$$x^\alpha x^\beta = e^{\alpha \ln x}e^{\beta \ln x} = e^{(\alpha+\beta)\ln x} = x^{\alpha+\beta}.$$

Zweite Aussage folgt mit

$$x^\alpha x^\beta = e^{\alpha \ln x}e^{\alpha \ln y} = e^{\alpha(\ln x + \ln y)} = e^{\alpha \ln(xy)}.$$

(Nutzt Rechenregel $\ln(xy) = \ln x + \ln y$): Bew. $x = e^a$, $y = e^b$. Dann gilt

$$\ln x + \ln y = \ln e^a + \ln e^b = a + b = \ln(e^{a+b}) = \ln(xy).)$$

Bemerkungen.

- Für $k \in \mathbb{Z}$ haben wir nun x^k und $x^{1/k}$ auf zwei Arten definiert. Man kann sich aber leicht überlegen, daß beide Arten übereinstimmen:
  Sei $k \in \mathbb{N}$. Dann gilt $P_k(x) = e^{k \ln x} = e^{\ln x} \dots e^{\ln x} = x^k$. Ebenso gilt für $y := e^{\frac{1}{k}\ln x}$ offenbar $y > 0$ sowie $y^k = \dots = e^{\ln x} = x$. Damit gilt also $y = x^{1/k}$.
- Der Ausdruck 0^0 ist erstmal nicht definiert. Je nach Kontext sind fuer ihn die Werte $0 = \lim_{\alpha\to 0+} 0^\alpha$ bzw. $1 = \lim_{x\to 0+} x^0$ sinnvoll.

Notation. Wir definieren die Eulersche Zahl e als $e = \sum_{k=0}^{\infty} \frac{1}{k!} = \exp(1)$. Dann gilt also $1 = \ln(e)$. Damit folgt dann (mit dem gerade definierten Konzept von allgemeiner Potenz)

$$e^x = \exp(x \ln(e)) = \exp(x).$$

DEFINITION. *Ist $f : (a,b) \longrightarrow \mathbb{R}$ differenzierbar und $f' : (a,b) \longrightarrow \mathbb{R}$ stetig, so heißt f stetig differenzierbar.*

Unter Umständen ist es praktisch auch die Ableitungen in Randpunkten bilden zu können. Dazu definiert man (falls existiert) für $f : [a,b) \longrightarrow \mathbb{R}$ bzw. $f : (a,b] \longrightarrow \mathbb{R}$

$$f'(a) = \lim_{x \to a+} \frac{f(x) - f(a)}{x - a}$$

und

$$f'(b) = \lim_{x \to b-} \frac{f(x) - f(b)}{x - b}.$$

2. Differenzierbare Funktionen auf Intervallen

In diesem Abschnitt wollen wir differenzierbare Funktionen auf Intervallen detaillierter behandeln. Wir werden sehen, daß die Ableitung wichtige Informationen über Monotonie und Extremwerte von Funktionen kodiert. Eine wesentliche Rolle in den Betrachtungen spielt der Mittelwertsatz. Konkret betrachten wir Funktionen

$$f : [a,b] \to \mathbb{R}$$

und nehmen meist an:

- f stetig auf $[a,b]$
- f differenzierbar in (a,b)

DEFINITION (Lokale Extrema). *Sei $f : I \to \mathbb{R}$ gegeben.*

- *Es hat f in $\xi \in I$ ein lokales $\substack{\text{Maximum} \\ \text{Minimum}}$ wenn ein $\delta > 0$ existiert mit $\substack{f(x) \le f(\xi) \\ f(x) \ge f(\xi)}$, für alle $x \in I$ mit $|x - \xi| < \delta$.*
- *Gilt $\substack{f(x) \le f(\xi) \\ f(x) \ge f(\xi)}$ für alle $x \in I$, so hat f in ξ ein globales $\substack{\text{Maximum} \\ \text{Minimum}}$.*

Sind die entsprechenden Ungleichungen strikt für $x \neq \xi$, so spricht man von einem strikten $\substack{\text{Maximum} \\ \text{Minimum}}$. Maxima und Minima werden auch als Extrema bezeichnet.

THEOREM (Notwendige Bedingung für lokale Extrema). *Sei I ein offenes Intervall in $\mathbb{R}$. Sei $f : I \to \mathbb{R}$ differenzierbar. Hat f in $\xi \in I$ ein lokales Extremum, so gilt $f'(\xi) = 0$.* ***Zeichnung.***

Beweis. Wir betrachten nur den Fall, daß f in ξ ein lokales Maximum besitzt (für Minimum ist der Beweis analog).

Für $x > \xi$ und x nahe an ξ gilt: $\frac{f(x)-f(\xi)}{x-\xi} \leq 0$. Damit folgt

$$f'(\xi) = \lim_{x\to\xi} \frac{f(x) - f(\xi)}{x - \xi} \leq 0$$

Für $x < \xi$ und x nahe an ξ gilt: $\frac{f(x)-f(\xi)}{x-\xi} \geq 0$. Damit folgt

$$f'(\xi) = \lim_{x\to\xi} \frac{f(x) - f(\xi)}{x - \xi} \geq 0$$

Damit ergibt sich insgesamt $f'(\xi) = 0$. □

Bemerkung zu Chancen und Schwächen des vorigen Satzes. Sei $-\infty < a < b < \infty$.

- Es handelt sich um eine notwendige Bedingung, aber keine hinreichende Bedingung, wie das Beispiel $f: \ (-1,1) \to \mathbb{R}, \ f(x) = x^3$ in $\xi = 0$ zeigt.
- Eine differenzierbare Funktion muss auf (a,b) kein Extremum haben (siehe etwa $f: \ (0,1) \to \mathbb{R}, \ f(x) = x$).
- Eine stetige Funktion $f: \ [a,b] \to \mathbb{R}$ muss Minimum und Maximum annehmen (s.o.). Allerdings könnten diese Extrema am Rand liegen.
- Ist $f: [a,b] \longrightarrow \mathbb{R}$ stetig und differenzierbar auf (a,b) und liegen die Extrema nicht am Rand, so werden sie (sowie möglicherweise weitere Punkte) durch Lösen der Gleichung $f^{'}(\xi) = 0$ gefunden.

THEOREM (Rolle). *Sei $-\infty < a < b < \infty$. Sei $f: \ [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b) mit $f(a) = f(b)$. Dann existiert mindestens ein $\xi \in (a,b)$ mit $f'(\xi) = 0$.*

Zeichnung.

Beweis. Idee: Es sind Extrema im Inneren unvermeidbar und in einem solchen Extremum verschwindet die Ableitung nach dem vorigen Satz.

f ist stetig auf dem kompakten Intervall $[a,b]$. Daher nimmt f ein Minimum und ein Maximum an.

Fall 1: Es ist f konstant $= 0$. Dann kann man jedes $\xi \in (a,b)$ wählen.

Fall 2: Es ist f nicht konstant. Sei $c := f(a) = f(b)$. Dann gilt $\min f \neq c$ oder $\max f \neq c$. Sei o.E. $\max f \neq c$ und sei $\xi \in (a,b)$ ein Punkt in dem f das Maximum annimmt. Dann gilt aber $f'(\xi) = 0$. □

THEOREM (Mittelwertsatz). *Sei $-\infty < a < b < \infty$ und sei $f: \ [a,b] \to \mathbb{R}$ stetig und differenzierbar in (a,b). Dann existiert ein $\xi \in (a,b)$ mit*

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

Zeichnung.

Beweis. Idee: Ziehe von f die Gerade durch $(a, f(a))$ und $(b, f(b))$ ab
Zeichnung: Betrachte

$$F: [a,b] \to \mathbb{R}, \ F(x) = f(x) - \underbrace{\left(f(a) + \frac{f(b) - f(a)}{b-a}(x-a)\right)}_{\text{Gerade durch } (a,f(a)),(b,f(b))}.$$

Dann ist F stetig und differenzierbar auf (a,b) und es gilt $F(a) = 0 = F(b)$. Nach dem Satz von Rolle existiert daher ein $\xi \in (a,b)$ mit $0 = F'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}$. □

Bemerkung. Natuerlich kann man den Satz auch auf Teilintervalle $[c,d]$ von $[a,b]$ anwenden.

Ende der 1. Vorlesung

THEOREM (Verallgemeinerter Mittelwertsatz). *Sei $-\infty < a < b < \infty$ und seien $f, g: [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann existiert ein $\xi \in (a,b)$ mit*

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

Bemerkung. Das kann man (wenn entsprechende Terme nicht verschwinden) auch so schreiben:

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

In diesem Sinne handelt es sich um eine 'simultane' Mittelwertbildung.

Beweis. Betrachte

$$F(x) = (f(b) - f(a))(g(x) - g(a)) - (g(b) - g(a))(f(x) - f(a))$$

Damit ist F stetig, differenzierbar auf (a,b) und $F(a) = 0, \ F(b) = 0$. Der Satz von Rolle liefert $\xi \in (a,b)$ mit

$$0 = F'(\xi) = (f(b) - f(a))g'(\xi) - (g(b) - g(a))f'(\xi)$$

und damit die Behauptung. □

Bemerkungen. Es sind Mittelwertsatz, Satz von Rolle und verallgemeinerter Mittelwertsatz ganz end miteinander verwandt in folgendem Sinne:

- Aus dem Satz von Rolle folgt direkt die Verallgemeinerung des Mittelwertsatzes (siehe Beweis).
- Aus der Verallgemeinerung des Mittelwertsatzes folgt der Mittelwertsatz mit $f = f$ und $g = \mathrm{id}$.
- Aus dem Mittelwertsatz folgt der Satz von Rolle (Klar.)

THEOREM (Charakterisierung von Monotonie mittels Ableitungen). *Sei $f: [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b). Dann gilt:*

(a) f *konstant* $\Leftrightarrow$ $f' \equiv 0$ *auf* (a,b)

(b) f *monoton* ${}^{\textit{wachsend}}_{\textit{fallend}}$ $\Leftrightarrow$ ${}^{f'\geq 0}_{f'\leq 0}$ *auf* (a,b)

(c) ${}^{f'>0}_{f'<0}$ $\Rightarrow$ f *ist strikt* ${}^{\textit{wachsend}}_{\textit{fallend}}$

Beweis. (a) Das folgt sofort aus (b). (Übung: Direkter Beweis.)

(b) Wir zeigen zunächst $\Longrightarrow$: Wir behandeln nur monoton wachsendes f. (Der andere Fall kann analog behandelt werden). Für $\xi \in (a,b)$ und $x \neq \xi$ gilt aufgrund der Monotonie dann

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

Damit folgt

$$f'(\xi) = \lim_{x\to\xi} \frac{f(x)-f(\xi)}{x-\xi} \geq 0$$

Wir zeigen nun $\Longleftarrow$: Seien $x, y \in [a,b]$ mit $x < y$. Aus dem Mittelwertsatz folgt

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

und damit

$$f(y)-f(x) \geq 0.$$

(c) Wir betrachten nur $f' > 0$. (Der andere Fall kann analog behandelt werden). Sei $x < y$. Nach dem Mittelwertsatz gilt

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

also

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

Das beendet den Beweis. □

Bemerkung. Die Umkehrung in c) gilt nicht. Betrachte dazu $f : (-1,1) \to \mathbb{R},\ f(x) = x^3$ in 0.

Folgerung (Ableitung bestimmt Funktion bis auf eine Konstante). *Seien $f, g : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar mit $f' = g'$. Dann ist $f - g$ konstant. Gibt es in diesem Fall ein c mit $f(c) = g(c)$, so folgt $f = g$.*

Beweis. Setze $h := f - g$. Dann gilt $h' = f' - g' = 0$ auf (a,b) und nach obigem Satz ist $h =$const. □

Nachdem wir nun die erste Ableitung untersucht haben, wollen wir sehen, was man aus der zweiten Ableitung lernen kann.

Definition. *Ist $f : (a,b) \to \mathbb{R}$ differenzierbar und $f' : (a,b) \to \mathbb{R}$ ebenfalls differenzierbar, so heißt f zweimal differenzierbar und man nennt $f'' := (f')'$ die zweite Ableitung von f.*

Bemerkung. Offenbar ist $f^{'}$ stetig, wenn f zweimal differenzierbar ist.

THEOREM (Hinreichende Bedingung für strikte lokale Extrema). *Sei $f:\ (a,b) \to \mathbb{R}$ zweimal differenzierbar. Gilt für ein $\xi \in (a,b)$*

$$f'(\xi) = 0 \text{ und } {}^{f''(\xi)<0}_{f''(\xi)>0},$$

so hat f in ξ ein striktes lokales ${}^{\text{Maximum}}_{\text{Minimum}}$. ***Zeichnung.*** *($f$, $f^{'}$, $f^{''}$ in beiden Fällen)*

Beweis. Wir betrachten $f'(\xi) = 0,\ f''(\xi) < 0$ (anderer Fall analog). Wegen

$$0 > f''(\xi) = \lim_{x\to\xi} \frac{f'(x) - f'(\xi)}{x-\xi} = \lim_{x\to\xi} \frac{f'(x)}{x-\xi}$$

gilt für alle x nahe an ξ also

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

Damit gibt es ein $\delta > 0$ mit

- $f'(x) < 0$ für $x \in (\xi, \xi+\delta)$
- $f'(x) > 0$ für $x \in (\xi-\delta, \xi)$

Also ist f strikt ${}^{\text{fallend}}_{\text{wachsend}}$ in ${}^{(\xi,\xi+\delta)}_{(\xi-\delta.\xi)}$
Daher hat f striktes Maximum in ξ. □

Bemerkung. Diese Bedingung ist nicht notwendig. (Beispiel: $f:\ (-1,1) \to \mathbb{R},\ f(x) = x^4$. Dann hat f ein striktes Minimum in $\xi = 0$. Aber $f''(\xi) = 12\xi^2 = 0$.)

Die Betrachtungen zum Verhalten von Extremwerten mit $f^{''}(\xi) < 0$ bzw. $f^{''}(\xi) > 0$ lassen sich verallgemeinern. Dazu brauchen wir noch einen Begriff.

DEFINITION (konkav und konvex). *$f:\ [a,b] \to \mathbb{R}$ heißt ${}^{\text{konkav}}_{\text{konvex}}$, wenn für alle $c,d \in [a,b]$ mit $c < d$ und $\lambda \in (0,1)$ gilt*

$$f(c + \lambda(d-c) \gtreqless f(c) + \lambda(f(d) - f(c)).$$

Geometrische Deutung. Sei $\lambda \in [0,1]$ und $\mu = 1-\lambda$. Dann gilt

$$c + \lambda(d-c) = \lambda d + (1-\lambda)c = (1-\mu)d + \mu c = d + \mu(c-d)$$

und mit der Geraden

$$g: \mathbb{R} \longrightarrow \mathbb{R},\ g(x) = f(c) + \frac{f(d)-f(c)}{d-c}(x-c)$$

durch $(c, f(c))$ und $(d, f(d))$ gilt

$$g(c + \lambda(d-c)) = f(c) + \lambda(f(d) - f(c)) = \lambda f(d) + (1-\lambda)f(c).$$

Zeichnung. Damit bedeutet dann Konvexitaet gerade, daß der Sekantenabschnitt oberhalb der Funktion verläuft und Konkavitaet bedeutet, daß der Sekantenabschnitt unterhalb der Funktion verläuft.

Bemerkung. f konvex $\Leftrightarrow -f$ konkav

PROPOSITION. *Sei I ein offenes Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ gegeben. Ist f konvex oder konkav, so ist f stetig.*

Bemerkung. Die Aussage gilt nicht fuer abgeschlossene Intervalle (Uebung).

Beweis. Wir betrachten nur konvexes f. (Der andere Fall kann analog behandelt werden.) Sei $\xi \in I$ beliebig. Wähle $x < \xi < y$ aus I beliebig. Sei g_1 die Sekante durch $(x, f(x))$ und $(\xi, f(\xi))$ und g_2 die Sekante durch $(\xi, f(\xi))$ und $(y, f(y))$. Dann verläuft nach Konvexität also für $s > \xi$ **Zeichnung.**

- g_2 oberhalb des Graphen von f auf $[\xi, y]$
- g_1 unterhalb des Graphen von f auf $[\xi, y]$.

Damit folgt leicht $\lim_{s\to\xi+} f(s) = f(\xi)$. Der Grenwert von links kann analog untersucht werden. $\square$

Wir kommen nun zur schon angekündigten Verallgemeinerung der Betrachtungen um Minima/Maxima zweimal differenzierbarer Funktionen.

THEOREM (Charakterisierung konkaver und konvexer Funktionen). *Sei $f : (a, b) \to \mathbb{R}$ stetig. Dann gilt:*

(a) Ist f differenzierbar, so ist f $\substack{konkav \\ konvex}$ genau dann, wenn f' monoton $\substack{fallend \\ wachsend}$ ist.

(b) Ist f zweimal differenzierbar, so ist so ist f $\substack{konkav \\ konvex}$ genau dann, wenn gilt $f'' \substack{\leq 0 \\ \geq 0}$.

Beweis. Offenbar folgt (b) aus (a) und der schon gegebenen Charakterisierung von Monotonie. Es bleibt also (a) zu zeigen. Wir betrachten nur konvexe f bzw. monoton wachsende f'.

Sei f' monoton wachsend: Sei $x < \xi < y$ in (a, b). Sei g_1 die Gerade durch $(x, f(x))$ und $(\xi, f(\xi))$ und g_2 die Gerade durch $(\xi, f(\xi))$ und $(y, f(y))$. **Zeichnung.** Nach dem Mittelwertsatz und der Monotonie von f' ist die Steigung von g_1 kleiner gleich der Steigung von g_2. Damit folgt, daß $(\xi, f(\xi))$ unterhalb der Geraden durch $(x, f(x))$ und $(y, f(y))$ liegt.

Ende der Vorlesung.

Sei f konvex. **Zeichnung.** Sei $x < y$. Sei c die Steigung der Geraden durch $(x, f(x))$ und $(y, f(y))$. Sei $\xi \in (x, y)$. Dann ist die Steigung der Geraden durch $(\xi, f(\xi))$ und $(x, f(x))$ kleiner oder gleich c aufgrund der

Konvexität. Da $\xi \in (x, y)$ beliebig war, folgt, daß $f'(x) \leq c$. Ähnlich sieht man aber $c \leq f'(y)$. □

Beispiele

- Die Funktion $\ln: (0, \infty) \to \mathbb{R},\ x \mapsto \ln(x)$ ist konkav.
  Bew. $\ln'(x) = \frac{1}{x} \Rightarrow \ln''(x) = -\frac{1}{x^2} < 0$.
- Die Exponentialfunktion $\exp: \mathbb{R} \to \mathbb{R},\ x \mapsto \exp x$ ist konvex.
  Bew. $\exp''(x) = \exp x > 0$.

Bemerkung. Ist f konvex und streng monoton wachsend mit der Umkehrfunktion g, so ist g konkav und umgekehrt. Ist f nicht monoton wachsend so gilt die Aussage im allgemeinen nicht (vgl. etwa die Funktion $f(x) = 1/x$.)

DEFINITION (Stetig Differenzierbar). *Sei I ein Intervall in $\mathbb{R}$ und $f: I \longrightarrow \mathbb{R}$ eine Funktion. Dann heißt f stetig differenzierbar, wenn f in jedem Punkt von I differenzierbar ist und die Ableitung $f': I \longrightarrow \mathbb{R}$ eine stetige Funktion ist.*

Bemerkung. Gehoert der rechte (linke) Randpunkt zum Intervall, so bedeutet Differenzierbarkeit von f in diesem Punkt gerade Differenziebarkeit von links (bzw. rechts).

LEMMA. *Sei $-\infty < a < b < \infty$ und $f: [a, b] \longrightarrow \mathbb{R}$ gegeben. Dann sind aequivalent:*

(i) *Es ist f stetig differenzierbar auf $[a, b]$.*
(ii) *Es ist f auf (a, b) differenzierbar und es gibt eine stetige Funktion $g: [a, b] \longrightarrow \mathbb{R}$ mit $f' = g$ auf (a, b).*

Beweis. (i)$\Longrightarrow$ (ii): Das folgt sofort mit $g := f'$.
(ii)$\Longrightarrow$ (i): Es bleibt zu zeigen, daß f in a rechtsseitig differenzierbar ist mit Ableitung $g(a)$ und in b linksseitig differenzierbar ist mit Ableitung $g(b)$. Das folgt leicht aus dem Mittelwertsatz und der Stetigkeit von g. □

3. Taylorscher Satz und L'Hospitalsche Regel

In diesem Abschnitt diskutieren wir Taylorschen Satz und L'Hospitalsche Regel. Dabei geht es um

- Approximation von Funktionen durch Polynome (Taylorscher Satz)
- Verhalten von Termen der Form $0/0$ bzw ∞/∞ bzw 0∞ (L'Hospitalsche Regel).

Wir werden beides als Konsequenzen aus dem verallgemeinerten Mittelwertsatz erhalten.

Wir beginnen mit dem Taylorschen Satz.

Definition (k-te Ableitung). *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \longrightarrow \mathbb{R}$ eine Funktion. Man definiert induktiv die k-te Ableitung $f^{(k)}$ durch*

$$f^{(0)} := f, \quad f^{(k+1)} := (f^{(k)})'$$

(falls $f^{(k)}$ auf I differenzierbar ist). Existiert $f^{(k)}$ für jedes $k \in \mathbb{N}$, so heißt f beliebig oft differenzierbar oder auch unendlich oft differenzierbar.

Idee - Taylorsches Theorem. Approximation von f durch Polynome.

- f stetig in p:
$$f(x) = f(p) + r(x)$$
mit $r(x) \to 0$, $x \to p$.
- f differenzierbar in p:
$$f(x) = f(p) + f'(p) \cdot (x - p) + r(x)(x - p)$$
mit $r(x) \to 0$ für $x \to p$ (nämlich $r(x) = \frac{\varphi(x)}{|x-p|} \to 0, \;\; x \to p$)
- f k-mal differenzierbar in p:
$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(p)}{j!}(x - p)^j + r(x)(x - p)^k$$
mit $r(x) \to 0$, $x \to p$.

Bemerkung. Es gibt bei der Approximation einer Funktion f durch Polynome zwei apriori verschiedene Varianten. Zum einen kann man Polynome suchen, die mit f in einem Punkt bis zu einer hohen Ableitung uebereinstimmen. Zum anderen kann man Polynome suchen, die mit f bis auf einen sehr kleinen Fehler in einer Umgebung des Punktes uebereinstimmen. Es zeigt sich, daß beide Varianten auf dieselben Polynome fuehren.

Definition (Das n-te Taylorpolynom). *Sei I ein offenes Intervall und $f : I \longrightarrow \mathbb{R}$ n-mal differenzierbar. Dann definiert man zu $p \in I$ das n-te Taylorpolynom von f im Punkt p durch*

$$P_{n,p}(x) := \sum_{k=0}^{n} \frac{f^{(k)}(p)}{k!}(x - p)^k.$$

Proposition. *Sei $n \in \mathbb{N}_0$ und sei P ein Polynom vom Grad hoechstens n mit*

$$P(c) = P'(c) = \ldots = P^{(n)}(c) = 0$$

für ein c. Dann gilt $P \equiv 0$.

Beweis. Das folgt durch Induktion. Der Fall $n = 0$ ist klar. Wir diskutieren nun den Schluss von n auf $n + 1$. Sei P ein Polynom vom Grad $n+1$ mit $P(c) = \ldots P^{(n+)}(c) = 0$. Dann ist offenbar $Q := P'$ ein Polynom vom Grad n mit $Q(c) = ... = Q^{(n)}(c) = 0$. Damit folgt also $Q = 0$ (aufgrund der Gueltigkeit der gewuenschten Aussage fuer n). Aus dem Mittelwertsatz folgt aber aus $P' = Q = 0$ sofort $P = const$. Mit $P(c) = 0$ folgt die gewuenschte Aussage fuer $n + 1$.

□

LEMMA (Charakterisierung Taylorpolynom). *Sei I ein offenes Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ n-mal differenzierbar und $p \in I$. Dann ist $P_{n,p}$ das eindeutige Polynom vom Grad höchstens n mit*

$$P_{n,p}^{(k)}(p) = f^{(k)}(p) \quad , k = 0, \ldots, n.$$

('P und f stimmen bis zur n-ten Ableitung in p überein.')

Beweis. *Eindeutigkeit.* Das folgt aus vorangehender Proposition.

Gewünschte Eigenschaften: Das folgt durch eine direkte Rechnung.

Diese Rechnung nutzt, daß die l-te Ableitung von $(x - p)^k$ an der Stelle p gegeben ist durch

$$\left((x-p)^k\right)^{(l)} = \begin{cases} k! & l = k \\ 0 & l > k \\ 0 & l < k \quad (x = p) \end{cases}.$$

Das beendet den Beweis. □

Die Frage ist nun, wie sich f von $P_{n,p}$ unterscheidet. Eine erste Antwort gibt der folgende Satz.

THEOREM (Taylorsche Formel mit Lagrange Restglied). *Sei I ein offenes Intervall in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ $(n + 1)$-mal differenzierbar und $p \in I$. Dann existiert zu jedem $x \in (a, b)$ ein $t = t(x)$ zwischen x und p mit*

$$f(x) = P_{n,p}(x) + R_{n+1,p}(t)$$

mit dem Lagrange Restglied

$$R_{n+1,p}(t) = \frac{f^{(n+1)}(t)}{(n+1)!}(x-p)^{n+1}$$

(und natürlich $P_{n,p}(x) = \sum_{k=0}^n \frac{f^{(k)}(p)}{k!}(x - p)^k$.)

Beweis. Beachte x, p gegeben! Wir entwickeln um ξ. Sei dazu

$$h(\xi) := (x - \xi)^{n+1}$$

$$g(\xi) := \sum_{k=0}^{n} \frac{f^{(k)}(\xi)}{k!}(x-\xi)^k.$$

Dann gilt:

·) $g(x) = f(x) \quad (k = 0)$
·) $g(p) = P_{n,p}(x)$
·) $g'(\xi) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n$ (!) *Teleskopsumme*

Nach verallgemeinertem Mittelwertsatz existiert ein t zwischen x und p mit:

$$(g(x) - g(p))h'(t) = (h(x) - h(p))g'(t).$$

Also gilt

$$(f(x) - P_{n,p}(x))(n+1)(-1)(x-t)^n = (-1)(x-p)^{n+1}\frac{f^{(n+1)}(t)}{n!}(x-t)^n$$

$$\Rightarrow f(x) - P_{n,p}(x) = \underbrace{\frac{f^{(n+1)}(t)}{(n+1)!}(x-p)^{n+1}}_{=R_{n+1,p}(t)}$$

$$\Rightarrow f(x) = P_{n,p}(x) + R_{n+1,p}(t)$$

(Es bleibt noch (!) zu begründen:

$$\begin{aligned} g(\xi) &= \sum_{k=0}^{n} \frac{f^{(k)}(\xi)}{k!}(x-\xi)^k \\ g'(\xi) &= \sum_{k=0}^{n} \frac{f^{(k+1)}(\xi)}{k!}(x-\xi)^k - \sum_{k=1}^{n} \frac{f^{(k)}(\xi)}{k!}k(x-\xi)^{k-1} \\ &= \sum_{k=1}^{n+1} \frac{f^{(k)}(\xi)}{(k-1)!}(x-\xi)^{k-1} - \sum_{k=1}^{n} \frac{f^{(k)}(\xi)}{(k-1)!}(x-\xi)^{k-1} \\ &= \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n \end{aligned}$$

Damit ist der Beweis von (!) abgeschlossen). □

Korollar 1. *Ist $g : I \longrightarrow \mathbb{R}$ $(n+1)$-mal differenzierbar mit $g^{(n+1)} = 0$, so ist g ein Polynom vom Grad höchstens n.*

Theorem (Taylorsche Formel mit Abschätzung). *Sei $f : (a,b) \longrightarrow \mathbb{R}$ n-mal differenzierbar mit $n \geq 1$ und $p \in (a,b)$. Dann gibt es eine Funktion $r : (a,b) \longrightarrow \mathbb{R}$ mit $r(x) \to 0$, $x \to p$ und*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(p)}{k!}(x-p)^k + r(x)\ (x-p)^n.$$

Bemerkung.

- Für $n = 1$ ist uns das wohlbekannt:

$$f(x) = f(p) + f'(p)\ (x-p) + \underbrace{\frac{\varphi(x)}{(x-p)}}_{=r(x)}(x-p).$$

- Im Unterschied zum vorangehenden Satz ist in diesem Satz die Voraussetzung schwaecher (da eine Differenzierbarkeitsstufe weniger gefordert wird). Dafuer ist allerdings auch die Aussage etwas schwaecher (da keine explizite Abschaetzung gegeben wird).

Beweis. Wir setzen $g(x) = f(x) - P_{n,p}(x)$. Es reicht zu zeigen: Es existiert $r : (a, b) \longrightarrow \mathbb{R}$ mit $r(x) \to 0$, $x \to p$ und $|g(x)| \leq r(x)\ |x - p|^n$.

Nach Definition von g (und der charakteristischen Eigenschaft des Taylorpolynomes) gilt

$$g(p) = g'(p) = \ldots = g^{(n)}(p) = 0.$$

Es gilt nun

$$|g(x)| = \frac{\left|g^{(n-1)}(t_x)\right|}{(n-1)!}|x-p|^{n-1} \qquad (\Diamond)$$

mit $t_x = x$ falls $n = 1$ und t_x zwischen x und p nach Anwendung der Taylorformel mit $(n - 2)$ fuer $n \geq 2$.
Nun ist $g^{(n-1)}$ differenzierbar, d.h.

$$g^{(n-1)}(t_x) = \underbrace{g^{(n-1)}(p)}_{=0} + \underbrace{g^{(n)}(p)(t_x - p)}_{=0} + \varphi(t_x) \qquad (\Diamond\Diamond)$$

mit

$$\frac{\varphi(y)}{y-p} \to 0, \quad y \to p.$$

$$\begin{aligned}
\left\{ {(\Diamond) \atop (\Diamond\Diamond)} \right\} \Rightarrow |g(x)| &\leq \frac{\varphi(t_x)}{(n-1)!}|x-p|^{n-1} \\
&= \underbrace{\frac{1}{(n-1)!}\underbrace{\frac{|\varphi(t_x)|}{|t_x-p|}}_{\to 0,\ x\to p}\underbrace{\frac{|t_x-p|}{|x-p|}}_{\leq 1}}_{\to 0,\ x \to p}|x-p|^n \\
&=: r(x)\ |x-p|^n
\end{aligned}$$

Das beendet den Beweis. □

Anwedung - Extrema von Funktionen mit vielen verschwindenden Ableitungen. Mithilfe des Satzes von Taylor können wir Funktionen mit 'vielen' verschwindenden Ableitungen auf Extrema untersuchen: Sei also $f : (a, b) \longrightarrow \mathbb{R}$ $n \geq 2$-mal differenzierbar mit

$$f'(p) = \ldots = f^{(n-1)}(p) = 0, \quad f^{(n)}(p) \neq 0.$$

Dann gilt:

- Ist n ungerade, so hat f in p einen Wendepunkt (d.h. $f(x) - f(p)$ wechselt in p das Vorzeichen).

- Ist n gerade, so hat f in p ein lokales $\frac{Maximum}{Minimum}$, falls $\frac{f^{(n)}(p)<0}{f^{(n)}(p)>0}$.

Beweis. Voriger Satz liefert

$$f(x) = f(p) + \left(\frac{f^{(n)}(p)}{n!} + r(x)\right)(x-p)^n$$

F'ur x nahe p hat $(\ldots)$ dasselbe Vorzeichen wie $f^{(n)}(p)$ (da $r(x) \to 0,\ x \to p$) Damit folgt, daß f im wesentlichen aussieht wie

$$f(p) + (Vorzeichen)\ (x-p)^n.$$

und das liefert die Aussage. □

Ende der Vorlesung.

Notation - Landausymbole. Eine passende Notation für obige Resultate liefern die Landausymbole: o 'klein ohhh von' und $\mathcal{O}$ 'groß Ohhh von' definiert wie folgt: Sei $f, g : D \longrightarrow \mathbb{R}$, p Berührpunkt von D. Dann:

$$f(x) = o(g(x)),\quad x \to p :\Leftrightarrow \lim_{x\to p}\left|\frac{f(x)}{g(x)}\right| = 0$$

$$f(x) = \mathcal{O}(g(x)),\quad x \to p :\Leftrightarrow\ \exists\, C > 0\ mit\ |f(x)| \le C\cdot|g(x)| \text{für}\quad \text{x nahe p.}$$

Damit können wir die Abschätzungen aus dem Taylorschen Satz wie folgt formulieren.

Folgerung. *Sei I ein offenes Intervall und $f : I \longrightarrow \mathbb{R}$ gegeben. Dann gilt:*
(a) Ist f n-mal differenzierbar (mit $n \ge 1$), so gilt $f(x) = P_{n,p}(x) + o((x-p)^n)$.
(b) Ist f (n+1)-mal differenzierbar (mit $n \ge 1$) so gilt $f(x) = P_{n,p}(x) + \mathcal{O}((x-p)^{n+1})$

Beweis. (a) folgt sofort aus der vorangehenden Variante des Satz von Taylor.

(b) Das folgt leicht wegen

$$f(x) = P_{n,p}(x) + \underbrace{\frac{f^{(n+1)}(p)}{(n+1)!}(x-p)^{n+1} + o((x-p)^{n+1})}_{=\mathcal{O}((x-p)^{n+1})}.$$

Das beendet den Beweis. □

Wir kommen nun noch zu einem weiteren Thema (das oft mit dem Taylorschen Satz verwechselt wird).

Definition (Taylor-Reihe). *Ist $f : (a,b) \longrightarrow \mathbb{R}$ beliebig oft differenzierbar, so können wir die sogenannte* ***Taylor-Reihe*** *von f im Entwicklungspunkt p*

$$T_{f,p}(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(p)}{k!}(x-p)^k.$$

bilden.

Bemerkungen.

- ! Der Satz von Taylor behandelt NICHT die Taylorreihe: Beim Satz von Taylor geht es (bei festem n) um die Frage, wie nahe $P_{n,p}(x)$ an $f(x)$ ist für x nahe p. Bei Taylorreihen geht es (bei festem x) um die Frage, wie nahe $P_{n,p}(x)$ an $f(x)$ ist für große n.
- ! Es ist <u>nicht</u> klar, daß Reihe für $x \neq p$ konvergiert.
- ! Selbst wenn $T_f(x)$ konvergiert, ist nicht automatisch auch $f(x) = T_f(x)$. Dazu ein Beispiel (Uebung): Sei

$$f : \mathbb{R} \longrightarrow \mathbb{R},\ f(x) = \begin{cases} e^{\frac{-1}{x}} & : x > 0 \\ 0 & : x \leq 0 \end{cases}$$

Dann ist f für $x < 0$ beliebig oft differenzierbar mit Ableitung 0. Ebenso ist f für $x > 0$ beliebig oft differenzierbar und die Ableitung hat die Form

$$f^{(n)}(x) = P_n\left(\frac{1}{x}\right) \cdot e^{\frac{-1}{x}}$$

mit geeigneten Polynomen P_n (Induktion). Damit folgt (Induktion), daß $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}_0$. (Denn: $n = 0$: klar.

$n \Longrightarrow (n+1)$:

$$\begin{aligned} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} &= \frac{f^{(n)}(x)}{x} \\ \frac{f^{(n)}(x)}{x} &= \frac{0}{x} = 0, \qquad x < 0 \\ \frac{f^{(n)}(x)}{x} &= \frac{P_n\left(\frac{1}{x}\right)\ e^{\frac{-1}{x}}}{x} = \frac{1}{x}P_n\left(\frac{1}{x}\right)\ e^{\frac{-1}{x}} \\ &\overset{!}{\to} 0 \quad,\ x \to 0^+ \end{aligned}$$

$$!\ \lim_{x\to 0^+} \frac{1}{x}P_n\left(\frac{1}{x}\right)\ e^{\frac{-1}{x}} \overset{(y=\frac{1}{x})}{=} \lim_{y\to\infty} yP_n(y)\ e^{-y} = 0.$$

Insgesamt gilt also

$$f \neq 0 \text{ aber } T_{f,0}(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}(x)^k = \sum_{k=0}^{\infty} \frac{0}{k!}(x)^k \equiv 0.$$

- Ist $R_n(x) = f(x) - P_{n,p}(x)$, so gilt

$$f(x) = T_{f,p}(x) \Leftrightarrow R_n(x) \to 0, \ n \to \infty.$$

Die vorhergehende Bemerkung zeigt, daß Taylorreihen nur mit großer Vorsicht zu behandeln sind. Im Falle von Potenzreihen ist die Lage (wie üblich) sehr schön:

Proposition. *(Potenzreihe ist Taylorreihe) Sei (a_n) eine Folge in $\mathbb{R}$ mit $\rho = \limsup \sqrt[n]{|a_n|} < \infty$ und $R := \frac{1}{\rho}$ falls $\rho > 0$ und $R = \infty$ falls $\rho = 0$ und*

$$f : (-R+p, R+p) \longrightarrow \mathbb{R}, \ f(x) = \sum_{n=0}^{\infty} a_n(x-p)^n.$$

Dann ist f beliebig oft differenzierbar und es gilt $f^{(n)}(p) = a_n n!$ für alle $n \in \mathbb{N}$. Insbesondere ist die Taylorreihe von f gerade die f definierende Potenzreihe.

Beweis. Wir wissen schon, daß f differenzierbar ist mit $f'(x) = g(x) = \sum_{n=1}^{\infty} n a_n (x-p)^n$, wobei die Reihe für g auch auf $(p-R, p+R)$ absolut konvergiert. Damit gilt also $f'(p) = a_1$ und eine Induktion liefert die erste Aussage. Das 'Insbesondere' folgt dann sofort. □

Wir kommen nun zu den Regeln von **L'Hospital**. Dabei geht es um Grenzwerte der Form 0/0 bzw ∞/∞ (bzw. $0 \times \infty$). Wir betrachten also Funktionen f und g mit $f(x) \to 0, g(x) \to 0$ bzw. $f(x) \to \infty$ und $g(x) \to \infty$ und untersuchen $f(x)/g(x)$.
Zur **Einstimmung** betrachten wir eine einfache Situation: Seien f, g differenzierbar in $a \in \mathbb{R}$ mit

$$f(a) = g(a) = 0, \ g'(a) \neq 0.$$

Dann gilt

$$\begin{aligned}
\lim_{x \to a} \frac{f(x)}{g(x)} &= \lim_{x \to a} \frac{f(x) - f(a)}{g(x) - g(a)} \\
&= \lim_{x \to a} \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} \\
&= \frac{\lim_{x \to a} \frac{f(x)-f(a)}{x-a}}{\lim_{x \to a} \frac{g(x)-g(a)}{x-a}} \\
&= \frac{f'(a)}{g'(a)}.
\end{aligned}$$

Allgemein können wir Ausdrücke der Form '$\frac{0}{0}$' und '$\frac{\infty}{\infty}$' mit folgenden Sätzen behandeln.

THEOREM (L'Hospital für '$\frac{0}{0}$' bei $a \in \mathbb{R}$). *Seien $a < b$ und $f, g : [a, b) \longrightarrow \mathbb{R}$ stetig und differenzierbar in (a, b) mit $g'(x) \neq 0$ für alle $x \in (a, b)$) gegeben, und es gelte*

$$f(a) = 0 = g(a).$$

Existiert der Grezwert $\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ (ggf. mit Wert $\pm\infty$), so existiert auch $\lim_{x\to a^+} \frac{f(x)}{g(x)}$ und es gilt

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

Eine entsprechende Aussage gilt für $f, g : (a, b] \longrightarrow \mathbb{R}$.

Bemerkung. Die Voraussetzungen implizieren $g(x) \neq 0$ für $x \neq a$ wegen

$$g(x) = g(x) - 0 = g^{'}(\xi)(x - a) \neq 0.$$

Beweis. Es gilt nach dem verallgemeinerten Mittelwertsatz

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} \overset{vMWS}{=} \frac{f'(\xi)}{g'(\xi)}$$

mit einem ξ zwischen x und a. Nach Voraussetzung existiert

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$$

und damit auch

$$\lim_{x\to a^+} \frac{f'(\xi)}{g'(\xi)}.$$

Damit folgt die Behauptung. $\square$

THEOREM (L'Hospital für '$\frac{0}{0}$' bei $\underline{+}\infty$). *Seien $c > 0$ und differenzierbare f, g in (c, ∞) mit $g^{'}(x) \neq 0$ auf (c, ∞) gegeben und es gelte*

$$\lim_{x\to\infty} f(x) = 0 = \lim_{x\to\infty} g(x).$$

Existiert $\lim_{x\to+\infty} \frac{f'(x)}{g'(x)}$ (ggf. mit Wert $\pm\infty$), so existiert auch

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)}$$

und es gilt

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}.$$

Analoge Aussage gilt bei $-\infty$.

Beweis. Falls jeweils der Grenzwert auf der rechten Seite existiert, gilt

$$\begin{aligned}
\lim_{x\to\infty}\frac{f(x)}{g(x)} &\overset{(x=\frac{1}{t})}{=} \lim_{t\to 0^+}\frac{f\left(\frac{1}{t}\right)}{g\left(\frac{1}{t}\right)}\\
(\textit{voriges Theorem}) \quad &= \lim_{t\to 0^+}\frac{\frac{-1}{t^2}f'\left(\frac{1}{t}\right)}{\frac{-1}{t^2}g'\left(\frac{1}{t}\right)}\\
&= \lim_{t\to 0^+}\frac{f'\left(\frac{1}{t}\right)}{g'\left(\frac{1}{t}\right)}\\
&\overset{(x=\frac{1}{t})}{=} \lim_{x\to\infty}\frac{f'(x)}{g'(x)}.
\end{aligned}$$

Da der letzte Grenzwert nach Voraussetzung existiert, folgt die Behauptung. □

Nachdem wir den Fall $'\frac{0}{0}'$ behandelt haben, kommen wir nun zum Fall $'\frac{\infty}{\infty}'$.

Theorem (L'Hospital für '$\frac{\infty}{\infty}$' von links bzw. rechts). *Sei $a \in \mathbb{R}$ bzw. $a = \infty$. Seien f, g stetig und differenzierbar mit $g^{'}(x) \neq 0$ auf einem offenen Intervall links von a (d.h. auf $(a-r, a)$ für ein $r > 0$ falls $a \in \mathbb{R}$ bzw. (c, ∞) für ein $c > 0$ falls $a = \infty$) und es gelte*

$$\infty = \lim_{x\to a^-} f(x) = \lim_{x\to a^-} g(x).$$

Existiert $\lim_{x\to a^-}\frac{f'(x)}{g'(x)}$ (ggf. mit Wert $\pm\infty$), so existiert auch $\lim_{x\to a^-}\frac{f(x)}{g(x)}$ und es gilt

$$\lim_{x\to a^-}\frac{f(x)}{g(x)} = \lim_{x\to a^-}\frac{f'(x)}{g'(x)}.$$

Entsprechende Aussagen gelten für Grenzwerte von rechts.

Bemerkung. Die Voraussetzungen implizieren, daß g streng monoton ist. (Denn $g(x)$ konvergiert gegen $\pm\infty$ und kann keine lokalen Extrema haben wegen $g^{'} \neq 0$.)

Beweis. Wir betrachten nur den Fall $a = \infty$, $\lim_{x\to a^-}\frac{f'(x)}{g'(x)} = \infty$.

Nach Voraussetzung existiert zu jedem $S > 0$ ein $d > 0$ mit $\frac{f'(\xi)}{g'(\xi)} \geq S$ für $\xi \geq d$. Nach dem verallgemeinerten Mittelwertsatz gilt dann

$$\frac{f(x)-f(d)}{g(x)-g(d)} \overset{vMWS}{=} \frac{f'(\xi)}{g'(\xi)} \geq S.$$

Es ist g streng monoton wachsend; vgl. Bemerkung vor dem Beweis. Damit gilt also $g(x) - g(d) > 0$, und es folgt

$$f(x) - f(d) \geq S(g(x) - g(d))$$

für $x > d$.

$$\frac{f(x)}{g(x)} \geq S - S \underbrace{\frac{g(d)}{g(x)}}_{\to 0, x\to\infty} + \underbrace{\frac{f(d)}{g(x)}}_{\to 0, x\to\infty}$$

und nach Bilden des Grenzwertes also

$$\liminf_{x\to\infty} \frac{f(x)}{g(x)} \geq S.$$

Da $S > 0$ beliebig war, folgt

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty.$$

Das beendet den Beweis. $\square$

Bemerkungen.

- **!** Die Sätze von L'Hospital haben zwei Vorausetzungen: (1) Konvergenz von f und g gegen 0 bzw. $\pm\infty$ und (2) Konvergenz von $\frac{f'}{g'}$. Beide Vorausetzungen sind nötig, wie man an folgendem Beispiel sieht: $f(x) = 1$, $g(x) = x + 7$. Dann gilt $\lim_{x\to 0} f(x)/g(x) = 1/7$ und $\lim_{x\to 0} f'(x)/g'(x) = 0$.
- Ist f stetig differenzierbar in (a, b) und $p \in (a, b)$, so liefert die L'Hospitalsche Regel

$$\lim_{x\to p, x\neq p} \frac{f(x) - f(p)}{x - p} = \lim_{x\to p} \frac{f'(x)}{1} = f'(p).$$

  Tatsächlich kann man das auch ohne L'Hospitalsche Regel und ohne Stetigkeit der Ableitung direkt aus der Definition der Differenzierbarkeit folgen ;-)

Beispiele - L'Hospital:

- $\lim_{\substack{x\to 0\\ x\neq 0}} \frac{1-\cos(x)}{x^2} = \lim_{\substack{x\to 0\\ x\neq 0}} \frac{\sin(x)}{2x} = \frac{1}{2} \lim_{\substack{x\to 0\\ x\neq 0}} \frac{\sin(x)}{x} = \frac{1}{2}.$
- $\lim_{x\to\infty} \frac{\log(x)}{2\sqrt{x}} = \lim_{x\to\infty} \frac{\frac{1}{x}}{\frac{1}{\sqrt{x}}} = \lim_{x\to\infty} \frac{1}{\sqrt{x}} = 0.$

Zur Anwendung noch zwei Tipps.

- Gegebenenfalls kann man L'Hospital auch iterieren (d.h. mehrmals hintereinander anwenden):

$$\lim_{x\to\infty} \frac{x^2}{e^x} \overset{falls\ ex.}{=} \lim_{x\to\infty} \frac{2x}{e^x} \overset{falls\ ex.}{=} \lim_{x\to\infty} \frac{2}{e^x} = 0.$$

- Man kann L'Hospital auch verwenden, um Ausdrücke der Form $0 \times \infty$ zu untersuchen. Dazu schreibt man diese in der Form ∞/∞ bzw. $0/0$: Sei $\alpha > 0$:

$$\begin{aligned}
\lim_{x\to 0^+} x^\alpha \log(x) &= \lim_{x\to 0^+} \frac{\log(x)}{x^{-\alpha}} = \lim_{x\to 0^+} \frac{\frac{1}{x}}{-\alpha x^{-\alpha-1}} \\
&= \lim_{x\to 0^+} -\frac{1}{\alpha}\frac{x^{\alpha+1}}{x} \\
&= -\frac{1}{\alpha} \lim_{x\to 0^+} x^\alpha = 0
\end{aligned}$$

$\longleftrightarrow$
Ende der Vorlesung.

KAPITEL 11

Das Riemann Integral in einer Dimension

Ziel. Ordne für geeignetes $f : [a, b] \to \mathbb{R}$ der Fläche zwischen dem Graphen von f und der x-Achse einen Wert zu (wobei unterhalb der x-Achse negativ gezählt wird). Der Wert der Fläche wird mit $\int_a^b f(x)dx$ bezeichnet. Dabei soll gelten:

- $f(x) \equiv c \ \Rightarrow\ \int_a^b f(x)dx = c(b-a)$
- $c \in [a, b] \ \Rightarrow\ \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$
- $f \leq g \ \Rightarrow\ \int_a^b f(x)dx \leq \int_a^b g(x)dx$

Außerdem soll folgende Stetigkeitseigenschaft gelten:

- Sind f_n, g_n, $n \in \mathbb{N}$, zulässig und gilt $f_n \leq f \leq g_n$ sowie $\int_a^b (g_n - f_n)dx \to 0$, so ist auch f zulässig und es gilt
$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

Es zeigt sich, daß durch diese Forderungen $\int_a^b f dx$ eindeutig bestimmt ist. Es wird als das Riemann-Integral von f bezeichnet. Das Riemann-Integral ist dann *linear* d.h.
$$\int_a^b f(x) + \lambda g(x)dx = \int_a^b f(x)dx + \lambda \int_a^b g(x)dx$$
und *positiv* d.h.
$$f \geq 0 \ \Rightarrow\ \int_a^b f(x)dx \geq 0$$
und mit gleichmäßiger Konvergenz verträglich. Stetige Funktionen und monotone Funktionen sind Riemann-integrierbar. Ebenso stückweise stetige und stückweise monotone Funktionen.

Bemerkung. Die entscheidende Einschränkung des Riemannintegrals ist die vergleichsweise unflexible Stetigkeitseigenschaft. Es gibt eine Reihe von zum Teil deutlich flexibleren Integralbegriffen (Lebesgü-Integral, Bochner-Integral, Daniell-Integral, Darboux-Integral, Petis-Integral). Das wird zumindest zum Teil in späteren Vorlesungen behandelt.

Um über Riemann-Integrale sprechen zu können, müssen wir voraussetzen:

- Das Intervall $[a, b]$ ist beschränkt und abgeschlossen d.h. $-\infty < a < b < \infty$,
- Die Funktion $f : [a, b] \to \mathbb{R}$ ist beschränkt.

Wir gehen nun in zwei Schritten vor: Im ersten Schritt zerlegen wir das Intervall in Teilintervalle und approximieren f auf den Teilstücken durch konstante Funktionen. Für diese Approximation ist das Integral durch obige Forderungen eindeutig festgelegt. Im zweiten Schritt führen wir einen Grenzübergang durch. Dieser Grenzübergang muss gerechtfertigt werden. Diese Rechtfertigung wird durch eine Definition geliefert: Die Funktionen, wo er gelingt, heissen Riemann-integrierbar.

1. Grundlegendes zu Riemann Integration

Definition (Zerlegung). *Sei $-\infty < a < b < \infty$. Ein Tupel*

$$Z = (x_0, \ldots, x_n)$$

mit $a = p < x_1 < \ldots < x_n = b$ heißt Zerlegung von $[a, b]$.
Die Feinheit $|Z|$ der Zerlegung Z ist definiert als

$$|Z| = \max\{x_i - x_{i-1},\ i = 1, \ldots, n\}.$$

Eine Zerlegung $Z' = (y_0, \ldots, y_k)$ heißt Verfeinerung von $Z = (x_0, \ldots, x_n)$, wenn gilt

$$\{y_i :\ i = 0, \ldots, k\} \supset \{x_i :\ i = 0, \ldots, n\}.$$

Man schreibt dann $Z' \supset Z$. Zur Zerlegung $Z = (x_0, \ldots, x_n)$ und $Z' = (q, \ldots, y_k)$ definieren wir die Vereinigung $Z \cup Z'$ als die Zerlegung mit den Punkten $\{x_i :\ i = 0, \ldots, n\} \cup \{y_i :\ i = 0, \ldots, k\}$.

Approximiert man f entlang einer Zerlegung Z, so gibt es mehrere Möglichkeiten **Zeichnung.**

- ''Durch Rechtecke von unten'' (jeweils kleinster Funktionswert auf $[x_{i-1}, x_i]$)
- ''Durch Rechtecke von oben'' (jeweils größter Funktionswert auf $[x_{i-1}, x_i]$)
- ''Durch Rechtecke dazwischen'' (irgendein Funktionswert auf $[x_{i-1}, x_i]$)

Uns wird es um Funktionen gehen, bei denen alle drei Möglichkeiten im Grenzwert den selben Wert ergeben.

Definition (Obersumme und Untersumme). *Zu $f : [a, b] \to \mathbb{R}$ beschränkt und Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a, b]$ definieren wir die Obersumme von f bezüglich Z durch*

$$O_Z(f) := \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x_{i-1} \le x \le x_i} f(x)$$

die Untersumme von f bezüglich Z durch

$$U_Z(f) := \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x_{i-1} \le x \le x_i} f(x)$$

Proposition (Eigenschaften von U_Z, O_Z). *Es gilt:*
(a) $U_Z(f) \le O_Z(f)$ *für alle* Z
(b) $Z_1 \subseteq Z_2 \;\Rightarrow\; U_{Z_1}(f) \le U_{Z_2}(f),\ O_{Z_1}(f) \ge O_{Z_2}(f)$
(c) $Z_1,\ Z_2$ *beliebig* $\Rightarrow\; U_{Z_1}(f) \le O_{Z_2}(f)$

Beweis. (a) klar.
(b) Wir betrachten nur U_Z (O_Z kann analog behandelt werden).
o.E.: Z_2 hat genau einen Punkt mehr als Z_1.
$Z_1:\ a = p < \ldots < \alpha < \beta < \ldots = b$
$Z_2:\ a = p < \ldots < \alpha < \gamma < \beta < \ldots = b$
In Z_1: $(\beta-\alpha) \inf_{\alpha \le x \le \beta} f(x) = (\beta-\gamma) \inf_{\alpha \le x \le \beta} f(x) + (\gamma-\alpha) \inf_{\alpha \le x \le \beta} f(x)$
In Z_2: $(\beta-\alpha) \inf_{\alpha \le x \le \beta} f(x) \le (\gamma-\alpha) \inf_{\alpha \le x \le \gamma} f(x) + (\beta-\gamma) \inf_{\gamma \le x \le \beta} f(x)$

Damit folgt die Behauptung (b).

(c) $U_{Z_1}(f) \overset{\text{b)}}{\le} U_{Z_1 \cup Z_2}(f) \overset{\text{a)}}{\le} O_{Z_1 \cup Z_2}(f) \le O_{Z_2}(f)$ Das beendet den Beweis. □

Definition (Riemann-Summe). *Sei* $-\infty < a < b < \infty$. *Ist* $Z = (x_0, \ldots, x_n)$ *eine Zerlegung von* $[a,b]$ *und* $\xi = (\xi_1, \ldots, \xi_n)$ *mit* $\xi_i \in [x_{i-1}, x_i]$ *gegeben.*
So definiert man die Riemann-Summe von f bezüglich der Zerlegung Z und den Stützstellen ξ als

$$S_{Z,\xi}(f) := \sum_{i=1}^{n} (x_i - x_{i-1}) f(\xi_i)$$

Lemma (Charakterisierung Riemann-integrierbarkeit). *Sei* $-\infty < a < b < \infty$ *und* $f:\ [a,b] \to \mathbb{R}$ *beschränkt. Dann sind äquivalent:*

(i) $\sup_Z U_Z(f) = \inf_Z O_Z(f)$
(ii) *Für alle* $\varepsilon > 0$ *existiert eine Zerlegung* Z *von* $[a,b]$ *mit* $O_Z(f) - U_Z(f) \le \varepsilon$
(iii) *Es existiert ein* $I_f \in \mathbb{R}$, *sodaß zu jedem* $\varepsilon > 0$ *ein* $\delta > 0$ *existiert mit*

$$|S_{Z,\xi}(f) - I_f| \le \varepsilon$$

für alle Zerlegungen Z *mit Feinheit* $|Z| \le \delta$.

In diesem Fall gilt:

$$I_f = \sup_Z U_Z(f) = \inf_Z O_Z(f)$$

Beweis. *(ii)* $\Rightarrow$ *(i)*: Da $U_{Z_1}(f) \le O_{Z_2}(f)$ für alle Z_1, Z_2 (s.o.) folgt

$$\sup_{Z_1} U_{Z_1}(f) \le \inf_{Z_2} O_{Z_2}(f).$$

Sei die Zerlegung Z zu $\varepsilon > 0$ gemäß (ii) gewaehlt. Wähle nun $Z_1 = Z_2 = Z$. Damit ergibt sich:

$$0 \leq O_Z(f) - U_Z(f) \leq \varepsilon$$

Da $\varepsilon > 0$ beliebig, gilt (i).
$(i) \Rightarrow (ii)$: Zu $\varepsilon > 0$ existieren Z_1 und Z_2 mit

$$0 \leq O_{Z_2}(f) - U_{Z_1}(f) \leq \varepsilon$$

Da $O_{Z_2}(f) \geq O_{Z_1 \cup Z_2}(f)$ und $U_{Z_1}(f) \leq U_{Z_1 \cup Z_2}(f)$ folgt

$$O_{Z_1 \cup Z_2}(f) - U_{Z_1 \cup Z_2}(f) \leq \varepsilon$$

und damit (ii) mit $Z = Z_1 \cup Z_2$.
$(i)/(ii) \Rightarrow (iii)$: Sei $C > 0$ mit
$|x| \leq C$ für alle $x \in [a, b]$. Sei

$$I_f := \sup_Z U_Z(f) = \inf_Z O_Z(f)$$

und $\varepsilon > 0$ beliebig. Nach $(i)/(ii)$ existiert eine Zerlegung $\tilde{Z} = (x_0, \ldots, x_N)$ mit

$$I_f - \frac{\varepsilon}{2} \leq U_{\tilde{Z}}(f) \leq O_{\tilde{Z}}(f) \leq I_f + \frac{\varepsilon}{2}.$$

Für eine sehr feine Zerlegung Z liegen nun die Terme von $S_{Z,\xi}$ zwischen den Termen von $O_{\tilde{Z}}(f)$ und $U_{\tilde{Z}}(f)$ außer an den Punkten $p, \ldots, x_N$ **Zeichnung**. Für $Z = (y_0, \ldots, y_k)$ mit Feinheit δ gilt:

$$U_{\tilde{Z}}(f) - \delta NC \leq S_{Z,\xi}(f) \leq O_{\tilde{Z}}(f) + \delta NC$$

wobei N die Anzahl der Intervalle von $\widetilde{Z}$ ist. Nach der Definition von I_f ergibt sich

$$I_f - \frac{\varepsilon}{2} - \delta NC \leq U_{\tilde{Z}}(f) - \delta NC \leq S_{Z,\xi}(f) \leq O_{\tilde{Z}}(f) + \delta NC \leq I_f + \frac{\varepsilon}{2} + \delta NC$$

Für alle $\delta > 0$ mit $\delta NC \leq \frac{\varepsilon}{2}$ ergibt sich also

$$|S_{Z,\xi}(f) - I_f| \leq \varepsilon$$

$(iii) \Rightarrow (ii)$: Sei $\varepsilon > 0$ beliebig und $\delta > 0$ gemäß (iii) gewählt. Dann gilt fuer jede Zerlegung Z mit Feinheit $|Z| < \delta$

$$I_f + \varepsilon \geq S_{Z,\xi}(f) \geq I_f - \varepsilon$$

für jede Wahl der Stützstellen ξ. Damit folgt auch

$$I_f + \varepsilon \geq O_Z(f) \geq U_Z(f) \geq I_f - \varepsilon$$

Das liefert gerade

$$O_Z(f) - U_Z(f) \leq 2\varepsilon$$

und da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Bemerkungen.

- Hinter (iii) verbirgt sich Konvergenz eines *Netzes*. Man schreibt dann auch $\lim_{|Z| \to 0} S_{Z,f}(\xi)$ fuer den Grenzwert I_f.

- Mit der entsprechendn Definition des Grenzwertes und einer leichten Modifikation des obigen Argumentes laßt sich auch folgendes zeigen (Uebung):

$$\inf O_Z(f) = \lim_{|Z|\to 0} O_Z(f), \text{ and } \sup U_Z(f) = \lim_{|Z|\to 0} U_Z(f).$$

Ende der Vorlesung.

Definition (Riemann-Integrierbarkeit). *Sei $-\infty < a < b < \infty$. Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ heißt Riemann-integrierbar, wenn eine der Bedingungen des vorigen Lemmas erfüllt ist. Dann definiert man das Riemann-Integral von f in $[a,b]$ als*

$$\int_a^b f(x)dx := \int_a^b f dx := \sup_Z U_Z(f) = \inf_Z O_Z(f) = I_f =: \lim_{|Z|\to 0} S_{Z,\xi}(f)$$

Weiterhin setzt man $\int_b^a f(x)dx := -\int_a^b f(x)dx$ sowie $\int_a^a f(x)dx = 0$.

Es gibt im wesentlichen drei Klassen von Riemann-integrierbaren Funktionen:

- Treppenfunktionen
- stetige Funktionen
- monotone Funktionen

Definition (Treppenfunktion). *Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt Treppenfunktion, wenn es eine Zerlegung $Z = (x_0, \dots, x_n)$ von $[a,b]$ und $c_i \in \mathbb{R}$, $i = 1, \dots, n$ gibt mit*

$$f \equiv c_i \text{ auf } (x_{i-1}, x_i)$$

Bemerkung. Die Werte einer Treppenfunktion in den Stellen x_j sind unerheblich.

Proposition (Treppenfunktionen sind Riemann-integrierbar). *Jede Treppenfunktion ist Riemann-integrierbar und es gilt*

$$\int_a^b f(x)dx = \sum_{i=1}^n (x_i - x_{i-1})c_i$$

wenn $f \equiv c_i$ auf (x_{i-1}, x_i) für die Zerlegung $Z = (x_0, \dots, x_n)$ von $[a,b]$.

Beweis. Übung □

Proposition (Stetige Funktionen sind Riemann-integrierbar). *Jede stetige Funktion ist Riemann-integrierbar.*

Beweis. Da $[a,b]$ beschränkt und abgeschlossen ist, ist das stetige f auf $[a,b]$ beschränkt und gleichmäßig stetig, d. h. für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|f(x) - f(\tilde{x})| \le \varepsilon$ für $|x - \tilde{x}| < \delta$.
Sei $\varepsilon > 0$ gegeben. Ziel: Finde Z mit $O_Z(f) - U_Z(f) < \varepsilon$.

Da f gleichmäßig stetig ist, existiert ein $\delta > 0$ mit

$$(\star) \qquad |f(x) - f(\tilde{x})| < \frac{\varepsilon}{b-a} \text{ für } |x - \tilde{x}| < \delta$$

Wähle nun eine beliebige Zerlegung Z mit $|Z| < \delta$. Dann gilt

$$\begin{aligned} O_Z(f) - U_Z(f) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \underbrace{\Big(\sup_{x_{i-1} \le x \le x_i} f(x) - \inf_{x_{i-1} \le x \le x_i} f(x) \Big)}_{\le \frac{\varepsilon}{b-a}, \text{ nach } (\star)} \\ &\le \sum_{i=1}^{n} (x_i - x_{i-1}) \frac{\varepsilon}{b-a} \\ &= (b-a) \frac{\varepsilon}{b-a} \\ &= \varepsilon. \end{aligned}$$

Das beendet den Beweis. □

Proposition (Monotone Funktionen sind Riemann-integrierbar). *Jede monotone Funktion $f : [a,b] \to \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. Sei f o.E. monoton wachsend (anderer Fall analog). Sei

$$Z = \left(a, a + \frac{b-a}{n}, \ldots, a + (n-1)\frac{b-1}{n}, b \right)$$

eine äquidistante Zerlegung mit Feinheit $\frac{b-a}{n}$. Dann gilt wegen der Monotonie:

$$\sup_{x_{i-1} \le x \le x_i} f(x) = f(x_i) \qquad \inf_{x_{i-1} \le x \le x_i} f(x) = f(x_{i-1})$$

Damit folgt

$$\begin{aligned} O_Z(f) - U_Z(f) &= \sum_{i=1}^{n} (x_i - x_{i-1})(f(x_i) - f(x_{i-1})) \\ &= \frac{b-a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ (\textit{Teleskopsumme}) \quad &= \frac{b-a}{n} (f(b) - f(a)) \\ &\to 0, \ n \to \infty \end{aligned}$$

Damit folgt die Riemann-Integrierbarkeit. □

Bemerkung. Die vorangehenden beiden Aussagen gelten auch, wenn man das Intervall $[a,b]$ in endlich viele abgeschlossene Intervalle teilen kann, auf denen f die entsprechende Eigenschaft hat.

Proposition (Eigenschaften: Linear und positiv). *Seien* $f, g : [a, b] \to \mathbb{R}$ *Riemann-integrierbar und* $\alpha \in \mathbb{R}$. *Dann gilt:*
(a) $f + \alpha g$ *ist Riemann-integrierbar mit*

$$\int_a^b f + \alpha g dx = \int_a^b f dx + \alpha \int_a^b g dx$$

(b) $f \geq 0 \;\Rightarrow \int_a^b f dx \geq 0$
Insbesondere gilt also $g \leq f \;\Rightarrow\; \int_a^b g dx \leq \int_a^b f dx$

Beweis. (a) folgt aus Betrachtungen zu Riemann-Summen durch Grenzübergang.
(b) folgt aus Betrachtungen zu Riemann-Summen durch Grenzübergang.
Das 'Insbesondere' folgt dann mit (a) und (b). ($g \leq f \;\Rightarrow\; 0 \leq f - g \;\Rightarrow \int_a^b f - g dx \geq 0 \;\Rightarrow\; \int_a^b f dx - \int_a^b g dx$.) □

Bemerkung. Die Riemann-integrierbaren Funktionen bilden einen unendlichdimensionalen Vektorraum (vgl. Lineare Algebra).

Proposition (Eigenschaften: Zusammensetzen von Intervallen). *Seien* $-\infty < a < b < \infty$ *und* $c \in (a, b)$ *gegeben. Dann gilt:*

- *Ist* $f : [a, b]$ *Riemann-integrierbar, so sind auch die Einschränkungen* $f|_{[a,c]}, f|_{[c,b]}$ *Riemann-integrierbar und es gilt*
$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx$$
- *Ist* $f : [a, c] \to \mathbb{R}$ *Riemann-integrierbar und* $g : [c, b] \to \mathbb{R}$ *Riemann-integrierbar, so ist auch*
$$h : [a, b] \to \mathbb{R},\; h(x) = \begin{cases} f(x) : & x \in [a, c) \\ 0 : & x = c \\ g(x) : & x \in (c, b] \end{cases}$$
Riemann-integrierbar und es gilt:
$$\int_a^b h dx = \int_a^c f dx + \int_c^b g dx$$

Beweis.

- Einfach (Einschränkung von ''guten'' Zerlegungen auf $[a, c]$ bzw. $[c, b]$)
- Einfach (Zusammensetzen von ''guten'' Zerlegungen)

Das beendet den Beweis. □

Proposition (Rechenregeln: Einfache Operationen). *Seien* $f, g : [a, b] \to \mathbb{R}$ *Riemann-integrierbar. Dann gilt:*

- $|f|$ *ist Riemann-integrierbar und*

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f|dx$$

- *Die Funktionen* $\min\{f,g\}, \max\{f,g\}: [a,b] \to \mathbb{R}$ *sind Riemann-integrierbar und es gilt*

$$\int_a^b \min\{f,g\}dx \leq \min\{\int_a^b f(x)dx, \int_a^b g(x)dx\}$$

sowie

$$\int_a^b \max\{f,g\}dx \geq \max\{\int_a^b f dx, \int_a^b g dx\}.$$

Beweis. Zum ersten Punkt: Eine einfache Fallunterscheidung zeigt

$$\sup_J |f(x)| - \inf_J |f(x)| \leq \sup_J f(x) - \inf_J f(x)$$

für jedes Teilintervall J von $[a,b]$. Damit folgt

$$O_Z(|f|) - U_Z(|f|) \leq O_Z(f) - U_Z(f).$$

Da f Riemann-integrierbar ist, folgt nun, daß $|f|$ Riemann integrierbar ist. Wegen $-|f| \leq f \leq |f|$ folgt dann aus der Positivität des Integrals

$$-\int_a^b |f|dx \leq \int_a^b f dx \leq \int_a^b |f|dx.$$

Das liefert die Behauptung. (Alternativ kann die Abschätzung auch mit der Dreiecksungleichung für Riemann-Summen beweisen werden: $|S_{Z,\xi}(f)| \leq S_{Z,\xi}(|f|)$. Nach Grenzübergang folgt dann die Behauptung.)

Zum zweiten Punkt: Es gilt

$$\min\{f,g\} = \frac{f+g-|f-g|}{2}, \ \max\{f,g\} = \frac{f+g+|f-g|}{2}$$

Nun folgt die Riemann-Integrierbarkeit aus der Linearitaet und dem ersten Punkt. Die Abschaetzung ist klar aufgrund der Positivitaet des Integrals. □

Ende der Vorlesung.

Beispiel - nicht Riemann-integrierbare Funktion. Sei

$$f: [0,1] \to \mathbb{R}, \ f(x) = \begin{cases} 1: & x \text{ rational} \\ 0: & x \text{ irrational} \end{cases}$$

gegeben. Dann gilt für jede Zerlegung z von $[0,1]$

$$U_Z(f) = 0 \text{ und } O_Z(f) = 1$$

Insbesondere ist f nicht Riemann-integrierbar.

Beweis. Da $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ dicht in $\mathbb{R}$ sind, gilt

$$U_Z(f) = \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x_{i-1} \le x \le x_i} f(x) = 0$$

$$O_Z(f) = \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x_{i-1} \le x \le x_i} f(x) = 1$$

□

Beispiel- Ausrechnen einer Riemannsumme. Sei $f : [a,b] \to \mathbb{R},\ f = c \equiv$ const. gegeben. Dann gilt:

$$\int_a^b f(x)dx = c(b-a)$$

Beweis. Es handelt sich um eine Treppenfunktion. □

Beispiel - Ausrechnen einer Riemannsumme. Sei $f : [a,b] \to \mathbb{R},\ f = x$ gegeben. Dann gilt:

$$\int_a^b f(x)dx = \frac{b^2 - a^2}{2}$$

Zeichnung - Differenz der Flächen zweier Dreiecke.

Beweis. Es ist f stetig und damit Riemann-integrierbar. Damit ergibt sich das Integral aus

$$\int_a^b f(x)dx = \lim_{|z| \to 0} S_{Z,\xi}(f)$$

Seien $Z_n = \left(a, a + \frac{b-a}{n}, a + \frac{2(b-a)}{n}, \dots, b\right)$ und $\xi_n = (x_1, \dots, x_n)$. Dann gilt:

$$\begin{aligned}
S_{Z_n,\xi_n}(f) &= \sum_{i=1}^{n} \frac{b-a}{n} f\left(a + i\frac{b-a}{n}\right) = \sum_{i=1}^{n} \frac{b-a}{n}\left(a + i\frac{b-a}{n}\right) \\
&= \sum_{i=1}^{n} \left(a + i\frac{b-a}{n}\right)\frac{b-a}{n} = \frac{b-a}{n}\left(na + \frac{b-a}{n}\sum_{i=1}^{n} i\right) \\
&= (b-a)a + \frac{b-a}{n}\frac{b-a}{n}\frac{n(n+1)}{2} = ba - a^2 + \frac{b^2 - 2ab + a^2}{2}\frac{n(n+1)}{n^2} \\
&\overset{n\to\infty}{\longrightarrow} ba - a^2 + \frac{b^2}{2} - ab + \frac{a^2}{2} \\
&= \frac{b^2}{2} - \frac{a^2}{2}
\end{aligned}$$

□

Das war mühsam!!!

2. Der Hauptsatz der Differential- und Integralrechnung

Für stetige Funktionen gibt es eine andere Methode als das Bilden von Riemann Summen, um das Riemann-Integral zu berechnen. Diese zeigt, daß die Riemann-Integration im wesentlichen so etwas wie die Umkehrung der Differentiation ist. Das lernen wir jetzt kennen.

DEFINITION (Stammfunktion). *Sei* $f: [a,b] \to \mathbb{R}$ *gegeben. Eine Funktion* $F: [a,b] \longrightarrow \mathbb{R}$ *heißt Stammfunktion zu* f*, wenn* F *differenzierbar ist mit* $F^{'} = f$*.*

Bemerkungen.

- Zwei Stammfunktionen unterscheiden sich durch eine Konstante (denn $F^{'} = G^{'}$ impliziert $F - G = constant$ (nach Mittelwertsatz s.o.)).
- In obiger Definition bedeutet natuerlich Differenzierbarkeit in a bzw. b gerade Existenz der linksseitigen bzw. rechtsseitigen Ableitung in a bzw. b.
- Seien $F, f: [a,b] \to \mathbb{R}$ stetig und F auf (a,b) differenzierbar mit $F' = f$ auf (a,b). Dann gilt: In a existiert die rechtsseitige Ableitung von F und es gilt $F'(a) = f(a)$. In b existiert die linksseitige Ableitung von F und es gilt $F'(b) = f(b)$.
  (Bew. Mittelwertsatz: $\frac{F(x)-F(a)}{x-a} = f(\xi)$ und f ist stetig....)
- $f = F^{'}$ impliziert, daß $f(I)$ ein Intervall ist (Zwischenwertsatz für Ableitungen, vgl. Übung). Damit hat man einen ersten Test, ob ein f eine Stammfunktion haben kann.

Nun zur angekündigten Methode zur Integration stetiger Funktionen:

THEOREM (Hauptsatz der Differential- und Integralrechnung - HDI). *Sei* $-\infty < a < b < \infty$ *und* $f: [a,b] \to \mathbb{R}$ *stetig. Dann gilt:*

- *Die Funktion* $G: [a,b] \to \mathbb{R}$, $G(x) := \int_a^x f(t)dt$ *ist eine Stammfunktion von* f*.*
- *Ist* F *eine Stammfunktion von* f*, so gilt*

$$\int_a^b f(t)dt = F(b) - F(a)$$

Bemerkung. Der HDI besagt insbesondere, daß jede stetige Funktion eine Stammfunktion besitzt und liefert dann eine Methode zur Integration, nämlich '"Finde Stammfunktion". In diesem Sinne ist die Integration die Umkehrung der Differentiation. Den Differentiationsregeln 'Kettenregel' und 'Produktregel' entsprechen dabei die Substitutionsregel und die partielle Integration (s.u.).

Beweis. Zum ersten Punkt: Wir bilden den Differenzenquotienten von G in x. Dazu untersuchen wir die Werte von $G(x)$ und $G(x+h)$ ohne Einschraenkung nehmen wir $h > 0$ an.

Aus der Definition von G folgt
$$G(x+h)-G(x)=\int_x^{x+h} f(t)dt.$$
Idee. $G(x+h)-G(x) \sim h(f(x)+\text{kleiner Fehler})$.
Durch Multiplizieren von $\frac{1}{h}$ und **Trick!** Subtrahieren von
$$f(x)=\frac{1}{h}\int_x^{x+h} f(x)dt$$
erhält man
$$\frac{G(x+h)-G(x)}{h}-f(x)=\frac{1}{h}\int_x^{x+h}\left(f(t)-f(x)\right)dt$$
Daraus folgt dann durch Bilden des Betrages
$$\begin{aligned}\left|\frac{G(x+h)-G(x)}{h}-f(x)\right| &\le \frac{1}{h}\int_x^{x+h}\left|f(t)-f(x)\right|dt\\ &\le \frac{1}{h}\int_x^{x+h}\sup_{x\le s\le x+h}\left|f(s)-f(x)\right|dt\\ &= \frac{1}{h}\sup_{x\le s\le x+h}\left|f(s)-f(x)\right|\int_x^{x+h}dt\\ &= \sup_{x\le s\le x+h}\left|f(s)-f(x)\right|\\ &\to 0,\ h\to 0,\ \text{da } f \text{ stetig}\end{aligned}$$
Nun zum zweiten Punkt: Ist F eine Stammfunktion zu f, so gilt $F'=f=G'$ also $(F-G)'=0$. Nach dem Mittelwertsatz existiert dann also eine Konstante c mit
$$F-G=c$$
also
$$F=G+c.$$
Damit folgt
$$F(b)-F(a)=(G(b)+c)-(G(a)+c)=G(b)-G(a)=G(b)=\int_a^b f(t)dt.$$
Das beendet den Beweis. □

Bemerkung. Zum Beweis des ersten Punktes des Haupsatzes der Differential und Integralrechnung kann man auch nutzen:
$$h\min\{f(t): x\le t\le x+h\}\le\int_x^{x+h} f(t)dt\le h\max\{f(t): x\le t\le x+h\}.$$
Bemerkung. Integrale unstetiger Funktionen sind im allgemeinen nicht differenzierbar: Betrachtet man zum Beispiel
$$f(x)=\begin{cases}0: & x<0\\ 1: & x\ge 0\end{cases}$$

So ist $G(x) = \int_{-1}^{x} f(t)dt$ gegeben durch

$$G(x) = \begin{cases} 0: & x \leq 0 \\ x: & x > 0 \end{cases}.$$

Damit ist G in $x = 0$ nicht differenzierbar.

Notation. Man schreibt $F = \int f(x)dx$ oder $F + C = \int f(x)dx$, falls F eine Stammfunktion zu f ist also $\int_a^b = F(b) - F(a)$ gilt. Ein solches F ist nur bis auf eine Konstante eindeutig bestimmt. Weiterhin schreibt man dann $F|_a^b$ für $F(b) - F(a)$.

Beispiele - Stammfunktionen

- $\int x^\alpha dx = \frac{1}{1+\alpha} x^{\alpha+1} + C$ für $\alpha \neq -1$
- $\int \frac{1}{x} dx = \ln|x| + C$. Beachte: $\int_a^b \frac{1}{x} dx = \ln|b| - \ln|a| = \ln \frac{b}{a}$ für $0 < a < b$ bzw. $a < b < 0$.
- $\int e^x dx = e^x + C$
- $\int a^x dx = \frac{1}{\ln(a)} a^x + C$ für $a > 0$
- $\int \sin x dx = -\cos x + C$
- $\int \cos x dx = \sin x + C$
- $\int \frac{1}{\cos^2 x} dx = \tan x + C$ mit $\tan(x) = \sin x / \cos x$.
- $\int \frac{1}{\sin^2 x} dx = -\cot x + C$ mit $\cot x = \cos x / \sin x$.
- $\int \frac{1}{1+x^2} dx = \arctan x + C$
- $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin + C$
- $\int \sinh x dx = \cosh x + C$
- $\int \cosh x dx = \sinh x + C$
- $\int \frac{1}{\sqrt{x^2+1}} dx = \mathrm{arsinh} x + C$
- $\int \frac{1}{\sqrt{x^2-1}} dx = \mathrm{arcosh} x + C$

Zum Finden von Stammfunktionen gibt es drei Methoden:

- Nachschauen in einer Liste (z. B. Bronstein),
- Substitutionsregel,
- partielle Integration.

Oft sind auch Kombinationen dieser Methoden nützlich.

Theorem (Substitutionsregel). *Sei $f : [a,b] \to \mathbb{R}$ stetig und $\varphi : [c,d] \to [a,b]$ stetig differenzierbar. Dann gilt*

$$\int_c^d f \circ \varphi(t) \cdot \varphi'(t) dt = \int_{\varphi(c)}^{\varphi(d)} f(x) dx$$

Beweis. Sei F eine Stammfunktion zu f. Dann ist nach der Kettenregel $F \circ \varphi$ differenzierbar mit

$$(F \circ \varphi)'(t) = F'(\varphi(t)) \cdot \varphi'(t) = f(\varphi(t)) \cdot \varphi'(t)$$

d. h. $F \circ \varphi$ ist eine Stammfunktion zu $(f \circ \varphi) \cdot \varphi'$. Damit folgt

$$\int_c^d f(\varphi(t)) \cdot \varphi'(t)dt = F \circ \varphi(d) - F \circ \varphi(c) = \int_{\varphi(c)}^{\varphi(d)} f(x)dx.$$

Das beendet den Beweis. $\square$

Bemerkung.

- Die Substitutionsregel ist also die Umkehrung der Kettenregel.
- Es wird nicht voraussgesetzt, daß φ injektiv oder surjektiv ist. Es ist möglich, daß $\varphi(c) > \varphi(d)$.

Die Substitutionsregel kann in zwei verschiedenen Varianten angewandt werden:

"Von links nach rechts": Im Integral stehen bereits die Funktion φ und die Ableitung φ', sodaß man direkt substituieren kann. Die wichtigste Anwendung dieser Variante ist die *logarithmische Integration*:

Beispiele

- Logarithmische Integration:
$$\int_a^b \frac{f'(t)}{f(t)}dt \overset{\varphi(t)=f(t)}{=} \int_{f(a)}^{f(b)} \frac{1}{x}dx = \ln \frac{f(b)}{(f(a)}.$$
(falls f auf $[a, b]$ stetig ist und festes Vorzeichen hat (also insbesondere nicht verschwindet).
- $\int_a^b \tan x dx = -\int_a^b \frac{-\sin x}{\cos x}dx = \ln \frac{\cos(a)}{\cos(b)}$.
- $\int_a^b (2t)e^{t^2}dt \overset{\varphi(t)=t^2}{=} \int_a^b \varphi'(t)e^{\varphi(t)}dt = \int_{\varphi(a)}^{\varphi(b)} e^x dx = e^{b^2} - e^{a^2}$.

"Von rechts nach links": Durch geschicktes Substituieren versucht man das Integral zu vereinfachen. Dazu wird zunächst x durch $\varphi(t)$ ersetzt und die Ableitung gebildet

$$\frac{dx}{dt} = \varphi'(t) \qquad \text{“}dx = \varphi'(t)dt\text{”}.$$

Anschliessend ersetzt man dann die Grenzen $a = \varphi(c)$ und $b = \varphi(d)$ durch c mit $a = \varphi(c)$ und d mit $b = \varphi(d)$.

Beispiel $\int_a^b 2ye^{y^2} \overset{y=\sqrt{t}}{=} \int_{a^2}^{b^2} 2\sqrt{t}e^t \frac{1}{2\sqrt{t}}dt = \int_{a^2}^{b^2} e^t dt = e^{a^2} - e^{b^2}$

Ende der Vorlesung.

Wir kommen nun zur 'Umkehrung' der Produktregel.

THEOREM (Partielle Integration). *Seien $f, g : [a, b] \to \mathbb{R}$ stetig und g stetig differenzierbar und F Stammfunktion zu f. Dann gilt:*

$$\int_a^b f(x)g(x)dx = F(x)g(x)|_a^b - \int_a^b F(x)g'(x)dx$$

Beweis. Nach Produktregel gilt $(Fg)'(x) = F'g + Fg' = fg + Fg'$. Also folgt $fg = (Fg)' - Fg'$. Integrieren liefert dann

$$\int_a^b f(x)g(x)dx = \int_a^b (Fg)'(x)dx - \int_a^b F(x)g'(x)dx = Fg|_a^b - \int_a^b F(x)g'(x)dx.$$

Das beendet den Beweis. □

Beispiele.

- Es gilt

$$\begin{aligned}\int_a^b x\sin x dx &= x(-\cos x)|_a^b - \int_a^b 1(-\cos x)dx \\ &= -x\cos x|_a^b + \int_a^b \cos x dx \\ &= -b\cos c + a\cos a + \sin b - \sin a\end{aligned}$$

- Es gilt

$$\begin{aligned}\int_a^b x^2e^x dx &= x^2e^x\Big|_a^b - \int_a^b 2xe^x dx \\ &= x^2e^x\Big|_a^b - \left(2xe^x|_a^b - \int_a^b 2e^x dx\right) \\ &= x^2e^x\Big|_a^b - 2xe^x|_a^b + 2e^x|_a^b \\ &= (x^2 - 2x + 2)e^x\Big|_a^b\end{aligned}$$

Manchmal liefert partielle Integration eine Iterationsformel für ein Integral. **Beispiel**

$$\begin{aligned}\int_a^b \sin^n x dx &= \int_a^b \underbrace{\sin x}_{f}\underbrace{\sin^{n-1} x}_{g} dx \\ &= -\cos x \sin^{n-1} x\Big|_a^b - \int_a^b (-\cos x)(n-1)\sin^{n-2} x \cos x dx \\ &= -\cos x \sin^{n-1} x\Big|_a^b + \int_a^b (n-1)\sin^{n-2} x \cos^2 x dx \\ &= -\cos x \sin^{n-1} x\Big|_a^b + (n-1)\int_a^b (\sin^{n-2} x - \sin^n x)dx\end{aligned}$$

$$\Rightarrow n\int_a^b \sin^n x dx = -\cos x \sin^{n-1} x\Big|_a^b + (n-1)\int_a^b \sin^{n-2} x dx$$

$$\Rightarrow \int_a^b \sin^n x dx = \frac{-\cos x \sin^{n-1} x}{n}\Bigg|_a^b + \frac{n-1}{n}\int_a^b \sin^{n-2} x dx$$

Es kann auch nützlich sein, die konstante Funktion 1 ins Integral zu schreiben. **Beispiel**

$$\begin{aligned}\int_a^b \ln x dx &= \int_a^b \underbrace{1}_{f} \underbrace{\ln x}_{g} dx \\ &= x \ln x|_a^b - \int_a^b x \frac{1}{x} dx \\ &= x \ln x|_a^b - x|_a^b\end{aligned}$$

Bemerkung. Wir haben partielle Integration und Substitution bisher nur für bestimmte Integrale eingeführt. Für die unbestimmten Integrale (Stammfunktionen) gelten analoge Regeln.

Erinnerung: Sind (h_n) und h Funktionen auf $[a,b]$, so heisst (h_n) gleichmaessig konvergent gegen h, falls zu jedem $\varepsilon > 0$ ein $N = N_\varepsilon \in \mathbb{N}$ existiert mit

$$|h_n(x) - h(x)| \leq \varepsilon$$

fuer alle $n \geq N$ und $x \in [a,b]$.

Als Anwendung des HDI lernen wir jetzt einen Satz ueber Vertauschung von Ableiten und Grenzwert kennen.

Theorem. *Sei $f : [a,b] \to \mathbb{R}$ gegeben. Gibt es stetig differenzierbare Funktionen $f_n : [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$, und eine Funktion $g : [a,b] \to \mathbb{R}$ mit*

- *$f_n(x) \to f(x)$ für jedes $x \in [a,b]$,*
- *$f_n' \to g$ gleichmäßig,*

so ist f stetig differenzierbar und $f' = g$.

Beweis. Da (f_n') stetig und gleichmäßig gegen g konvergent, folgt Stetigkeit von g. Weiterhin gilt nach dem HDI (angewendet auf f_n)

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)dt$$

Bildet man nun Grenzwert und nutzt (!) gleichmäßig Konvergenz der f_n' gegen g, so folgt

$$f(x) = f(a) + \int_a^x g(t)dt$$

Nach dem HDI (beachte g stetig) ist dann f differenzierbar mit $f' = g$.

! Wir schaetzen ab

$$\left|\int_a^x f_n'(t)dt - \int_a^x g(t)dt\right| \leq \int_a^b |f_n'(t) - g(t)|dt \leq \sup_{x\in[a,b]} |f_n'(t) - g(t)|(b-a) \to 0, n \to \infty.$$

□

Bemerkung.

- Der Beweis zeigt, dass gleichmaessige Konvergenz der (f_n') und Konvergenz der f_n in einem Punkt a, schon punktweise Konvergenz der f_n auf ganz $[a,b]$ implizieren.
- Punktweise Konvergenz von f_n und f_n' reicht nicht, um Differenzierbarkeit von f sicherzustellen: Betrachte dazu
$$f_n : \ [-1,1] \to \mathbb{R}, \ f_n(x) = \sqrt{\frac{1}{n} + x^2}.$$
Dann gilt $f_n \to |\cdot|$ punktweise und $f_n'(x) = \frac{x}{\sqrt{\frac{1}{n}+x^2}} \to \operatorname{sgn}(x)$ punktweise. Aber $|\cdot|$ ist nicht differenzierbar (in 0).

Nach dieser Anwendung des HDI, kommen wir noch zu einer Abgeschlossenheitseigenschaft des Vektorraumes der Riemann-integrierbaren Funktionen.

Theorem. *Sei $-\infty < a < b < \infty$. Sei $h : [a,b] \longrightarrow \mathbb{R}$ gegeben. Wenn es Riemann-integrierbare Funktionen $f_n, g_n : [a,b] \longrightarrow \mathbb{R}$ gibt mit*

- *$g_n \leq h \leq f_n$ fuer alle $n \in \mathbb{N}$ und*
- *$\int_a^b (f_n - g_n) dx \to 0, n \to \infty$,*

so ist auch h Riemann-integrierbar und es gilt
$$\int_a^b h dx = \lim_{n\to\infty} \int_a^b f_n dx = \lim_{n\to\infty} \int_a^b g_n dx.$$

Beweis. Sei $\varepsilon > 0$ beliebig. Das Ziel ist es nun zunaechst
$$O_Z(h) - U_Z(h) \leq \varepsilon$$
fuer eine geeignete Zerlegung Z zu zeigen.
Waehle nun ein $N \in \mathbb{N}$ mit
$$(*) \quad \int_a^b f_N dx - \int_a^b g_N dx = \int_a^b (f_N - g_N) dx \leq \frac{\varepsilon}{3}.$$
Da f_N und g_N Riemann-integrierbar sind, gibt es eine Zerlegung Z von $[a,b]$ mit
$$\int_a^b g_N dx - \frac{\varepsilon}{3} \leq U_Z(g_N) \left(\leq \int_a^b g_N dt \right)$$
und
$$\int_a^b f_N dx + \frac{\varepsilon}{3} \geq O_Z(f_N) \left(\geq \int_a^b f_N dt \right).$$
Wegen $g_N \leq h \leq f_N$ gilt also
$$(**) \quad \int_a^b g_N dx - \frac{\varepsilon}{3} \leq U_Z(g_N) \leq U_Z(h) \leq O_Z(h) \leq O_Z(f_N) \leq \int_a^b f_N dx + \frac{\varepsilon}{3}.$$
Wegen $(*)$ folgt dann
$$O_Z(h) - U_Z(h) \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, ergibt sich die Riemannintegrierbarkeit von h. Weiterhin folgt aus $(**)$ auch

$$(***)\qquad \left|\int_a^b hdx - \int_a^b f_N dx|, |\int_a^b hdx - \int_a^b g_N dx\right| \leq 2\varepsilon$$

fuer alle N, sodass $(*)$ gilt. Da $(*)$ fuer alle genuegend grossen N gilt, folgt $(***)$ fuer alle genuegend grossen N. Da $\varepsilon > 0$ beliebig ist, folgt die Konvergenzaussage. □

Bemerkung. Nach Definition ist ein $f : [a,b] \longrightarrow \mathbb{R}$ genau dann Riemann-integrierbar, wenn es Treppenfunktionen f_n, g_n gibt, die die beiden im Theorem genannten Punkte erfuellen. Da alle Treppenfunktionen Riemann-integrierbar sind, sagt das Theorem damit, dass die Menge der Riemann-integrierbaren Funktionen, die kleinste Menge ist, die alle Treppenfunktionen enthaelt und under der in den beiden Punkten beschriebenen Konvergenz abgeschlossen ist.

Folgerung. *Seien $h_n : [a,b] \to \mathbb{R}$, Riemann-integrierbar und h_n konvergiere gleichmäßig gegen $h : [a,b] \to \mathbb{R}$. Dann ist h Riemann-integrierbar und es gilt*

$$\int_a^b h(x)dx = \lim_{n\to\infty} h_n(x)dx$$

Beweis. Es ist zweierlei zu zeigen: Die Riemann-Integrierbarkeit von f und die Aussage ueber den Grenzwert.

Sei

$$\delta_n := \sup\{|h(x) - h_n(x)| : x \in [a,b]\}.$$

Die gleichmaessige Konvergenz der h_n gegen h bedeutet gerade, dass $\delta_n \to 0, n \to \infty$. Weitherhin sind die Funktionen $g_n := h_n - \delta_n$ und $f_n := h_n + \delta_n$ Riemann-integrierbar, und es gilt

- $g_n \leq h \leq f_n$ fuer alle $n \in \mathbb{N}$
- $\int_a^b (f_n - g_n)dx \leq |b-a| 2\delta_n \to 0, n \to \infty$.

Damit folgt aus dem vorigen Theorem, dass h Riemannintegrierbar ist mit

$$\int_a^b hdx = \lim_{n\to\infty} \int_a^b f_n dx = \lim_{n\to\infty} \int_a^b g_n dx.$$

Offenbar gilt

$$\left|\int_a^b f_n dx - \int_a^b h_n dx\right| = \int_a^b (f_n - h_n)dx = \delta_n(b-a) \to 0, n \to \infty.$$

Damit folgt die gewuenschte Konvergenzaussage. □

Erinnerung - Cauchy-Schwarz Ungleichung. Sei V ein reeller Vektorraum und

$$\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{R} \ \ (v,w) \mapsto \langle v,w\rangle$$

sei

- bilinear (d.h. $\langle v+\alpha w, u\rangle = ..., \langle u, v+\alpha w\rangle = ...$)
- symmetrisch (d.h $\langle u, v\rangle = \langle v, u\rangle$)
- positiv (d.h. $\langle u, u\rangle \geq 0$ fuer alle $u \in V$).

Dann gilt die Cauchy-Schwarz-Ungleichung

$$|\langle v, w\rangle| \leq \langle v, v\rangle^{1/2} \langle w, w\rangle^{1/2}$$

fuer alle $v, w \in V$.
(Bew. Es gilt

$$(*) \quad 0 \leq \langle v+\lambda w, v+\lambda w\rangle = \langle v, v\rangle + 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle$$

fuer alle $\lambda \in \mathbb{R}$. Wir unterscheiden zwei Faelle:

Fall 1: $\langle w, w\rangle = 0$. Aus $(*)$ fuer $\lambda \to \infty$ und $\lambda \to -\infty$ folgt dann $\langle v, w\rangle = 0$. Damit folgt die gewuenschte Ungleichung.

Fall 2: $\langle w, w\rangle > 0$. Dann koennen wir in $(*)$ dividieren und erhalten

$$0 \leq \frac{\langle v, v\rangle}{\langle w, w\rangle} + 2\lambda \frac{\langle v, w\rangle}{\langle w, w\rangle} + \lambda^2.$$

Quadratisches Ergaenzen ergibt dann

$$0 \leq \left(\lambda + \frac{\langle v, w\rangle}{\langle w, w\rangle}\right)^2 + \frac{\langle v, v\rangle}{\langle w, w\rangle} - \frac{|\langle v, w\rangle|^2}{|\langle w, w\rangle|^2}.$$

Da dies fuer alle λ gelten soll, folgt

$$0 \leq \frac{\langle v, v\rangle}{\langle w, w\rangle} - \frac{|\langle v, w\rangle|^2}{|\langle w, w\rangle|^2}.$$

Damit folgt die gewuenschte Ungleichung.)

Theorem (Cauchy-Schwarz Ungleichung). *Sind $f, g : [a, b] \longrightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch fg Riemann integrierbar und es gilt*

$$\left|\int_a^b fg dx\right| \leq \left(\int_a^b f^2 dx\right)^{1/2} \left(\int_a^b g^2 dx\right)^{1/2}.$$

Beweis. Fuer $f, g \geq 0$ Riemann-integrierbar, folgt die Riemannintegrierbarkeit vonfg leicht. Der allgemeine Fall laesst sich darauf zurueckfuehren durch Zerlegen von f, g in Positiv- und Negativteil. Diese entstehen durch geeignete Maxima und Minimabildung und sind daher Riemann-integrierbar.

Es bleibt die Ungleichung zu beweisen. Dazu betrachtet man auf dem Vekorraum R der Riemann-integrierbaren Funktionen auf $[a, b]$ die Abbildung

$$\langle \cdot, \cdot \rangle : R \times R \longrightarrow \mathbb{R}, \langle f, g\rangle = \int_a^b fg dt.$$

Diese Abbildung ist offenbar bilinear, symmetrisch und positiv. Nun folgt die Ungleichung aus der Cauchy-Schwarz Ungleichung. □

Folgerung. *Sei $-\infty < a < b < \infty$. Ist $w : [a,b] \longrightarrow [0,\infty)$ Riemann-integrierbar, so ist auch $\sqrt{w}$ Riemann-integrierbar und es gilt*

$$\int_a^b \sqrt{w}dx \leq (b-a)^{1/2} \left(\int_a^b wdt\right)^{1/2}.$$

Beweis. Es reicht zu zeigen, dass $\sqrt{w}$ Riemann-integrierbar ist. Die letzte Aussage folgt dann aus der Cauchy-Schwarz Ungleichung. Da w Riemann-integrierbar ist, existieren zu jedem $\varepsilon > 0$ Treppenfunktionen φ, ψ mit

$$0 \leq \varphi \leq w \leq \psi$$

und $\int_a^b \psi - \varphi dx \leq \varepsilon$. Wegen der Monotonie der Wurzel sind dann $\sqrt{\varphi}, \sqrt{\psi}$ Treppenfunktionen mit

$$0 \leq \sqrt{\varphi} \leq \sqrt{w} \leq \sqrt{\psi}.$$

Weiterhin gilt

$$\int_a^b (\sqrt{\psi} - \sqrt{\varphi})dx \leq \int_a^b \sqrt{\psi - \varphi}dx \leq \left(\int_a^b (\psi - \varphi)dx\right)^{1/2} (b-a)^{1/2}.$$

Daraus folgt die gewuenschte Integrierbarkeit nach dem Theorem (zur Abgeschlossenheit R.i. Funktionenen). $\square$

Bemerkung. Tatsaechlich ist fuer jede stetige Funktion w auf $\mathbb{R}$ und jedes Rieman-integrierbare f auf $[a,b]$ auch $w \circ f$ Riemann-integrierbar (Uebung: Nutze, dass jedes stetige w auf Kompakte gleichmaesig durch Polynome approximiert werden kann).

Zum Abschluss dieses Abschnittes diskutieren wir noch Integration komplexwertiger Funktionen

$$f : [a,b] \longrightarrow \mathbb{C}, \;\; f(x) = u(x) + iv(x)$$

mit $u, v : [a,b] \longrightarrow \mathbb{R}$. Es heisst ein solches f Riemann-integrierbar, wenn $u = \Re f$ und $v = \Im f$ Riemann-integrierbar sind. Dann definiert man

$$\int_a^b f(x)dx := \int_a^b udx + i\int_a^b vdx.$$

Das Integral ist dann (Nachrechnen) linear d.h. es gilt

$$\int_a^b (f + zg)dx = \int_a^b fdx + z\int_a^b gdx.$$

Ist $F = U + iV$ (mit reellewertigen U, V) eine Stammfunktion von $f = u + iv$ (d.h. es gilt $U' = u$ und $V' = v$), so folgt aus Anwendung des HDI auf Real- und Imaginaerteil sofort

$$F(b) - F(a) = \int_a^b f(x)dx.$$

PROPOSITION. *Ist* $f : [a,b] \longrightarrow \mathbb{C}$ *Riemann-integrierbar, so ist auch* $|f| : [a,b] \longrightarrow \mathbb{R}$ *Riemann-integrierbar, und es gilt*

$$\left| \int_a^b f dx \right| \leq \int_a^b |f| dx.$$

Beweis. Sei $f = u + iv$ mit reellwertigem u und v. Dann ist

$$|f| = \sqrt{u^2 + v^2}$$

Riemann-integrierbar (nach den oben diskutierten Saetzen). Setze nun $z := \int_a^b f dx$. Ist $z = 0$, so ist nichts mehr zu zeigen. Andernfalls setzen wir

$$\alpha := \frac{\overline{z}}{|z|^2}.$$

Dann gilt nach Konstruktion von α also

$$\left| \int_a^b f dx \right| = \alpha \int_a^b f dx = \int_a^b \alpha f dx.$$

Da die linke Seite reell ist, folgt also

$$|\int_a^b f dx| = \int_a^b \Re(\alpha f) dx \leq \int_a^b |\Re(\alpha f)| dx \leq \int_a^b |\alpha f| dx = \int_a^b |f| dx.$$

Das beendet den Beweis. □

3. Partialbruchzerlegung und Integration rationaler Funktionen

In diesem Abschnitt diskutieren wir die Integration rationaler Funktionen.

Eine rationale Funktion ist eine Funktion der Form $\frac{P_1}{P_2}$ mit Polynomen P_1 und P_2. (Diese Funktion ist dann ueberall definiert ausser in den Nullstellen von P_2.) Ist der Grad von P_1 größer (oder gleich) dem Grad von P_2, so kann man nach Polynomdivision $\frac{P_1}{P_2}$ schreiben als $\frac{P_1}{P_2} = Q + \frac{R}{P_2}$ mit geeigneten Polynomen Q und R, wobei der Grad von R kleiner ist als der Grad von P_2.

Beispiel - Polynomdivision

$$\frac{x^3 + 5x^2 - 7x + 4}{2x - 8} \qquad P_1 = x^3 + 5x^2 - 7x + 4,\ P_2 = 2x - 8$$

$$(x^3 + 5x^2 - 7x + 4) : (2x - 8) = \frac{1}{2}x^2 + 4,5x + 14,5 + \frac{120}{2x - 8}$$

Für die Integration rationaler Funktionen reicht es also Funktionen der Form $\frac{R}{P}$ mit $\mathrm{Grad} R < \mathrm{Grad} P$ zu betrachten.

Hat P den Grad n, so hat P nach dem Fundamentalsatz der Algebra genau n Nullstellen (über $\mathbb{C}$), d. h.

$$P(x) = a\prod_{i=1}^{n}(x - z_i) \quad \text{mit } z_i \in \mathbb{C}, a \neq 0$$

Ist P ein reelles Polynom, d. h. Koeffizienten von P sind reell, so ist mit z auch $\overline{z}$ eine Nullstelle von P ($\overline{P(z)} = P(\overline{z})$). Weiterhin gilt mit $z = \xi + i\mu$ mit $\mu \neq 0$ dann

$$(x - z)(x - \overline{z}) = (x - \xi - i\mu)(x - \xi + i\mu) = (x - \xi)^2 + \mu^2,$$

wobei die rechte Seite ein reelles Polynom ist. Damit ist dann mit P auch das durch $(x - z)(x - \overline{z})$ dividierte Polynom reell. Induktiv folgt, dass dann sogar die Vielfachheit von der Nullstelle z und der Nullstelle $\overline{z}$ überein stimmen muessen. Ist diese Vielfachheit l so gilt natuerlich

$$(x - z)^l(x - \overline{z})^l = \left((x - \xi)^2 + \mu^2\right)^l .$$

Ein reelles Polynom hat daher die Produktdarstellung

$$P(x) = a\prod_{i=1}^{r}(x - x_i)^{k_i}\prod_{j=1}^{s}\left[(x - \xi_j)^2 + \mu_j^2\right]^{l_j}$$

mit den reellen Nullstellen $x_1, \ldots, x_r$ der Vielfachheit k_i und den komplexen Nullstellen $\xi_j \pm i\mu_j$ $j = 1, \ldots, s$ der Vielfachheit l_j. Der Grad $GradP$ von P ist dann gerade

$$\sum_{i=1}^{r} k_i + 2\sum_{j=1}^{s} l_j$$

Theorem. *(Partialbruchzerlegung) Zu reellen Polynomen R und P mit*

$$P(x) = \prod_{i=1}^{r}(x - x_i)^{k_i}\prod_{j=1}^{s}\left[(x - \xi_j)^2 + \mu_j^2\right]^{l_j}$$

und $GradR < GradP$ gibt es Koeffizienten a_{ik}, b_{jl}, c_{jl} mit

$$\frac{R(x)}{P(x)} = \sum_{i=1}^{r}\sum_{k=1}^{k_i}\frac{a_{ik}}{(x - x_i)^k} + \sum_{j=1}^{s}\sum_{l=1}^{l_j}\frac{b_{jl}x + c_{jl}}{\left((x - \xi_j)^2 + \mu_j^2\right)^l}$$

Beweis. Übung □

Die Zahlen a_{ik}, b_{jl}, c_{jl} werden durch Koeffizientenvergleich oder Einsetzen spezieller Werte für x bestimmt.

Beispiel

$$\frac{-x^2 - 6x + 3}{(x - 1)^2(x^2 + 1)^2} = \frac{a_{11}}{x - 1} + \frac{a_{12}}{(x - 1)^2} + \frac{b_{11}x + c_{11}}{x^2 + 1} + \frac{b_{12}x + c_{12}}{(x^2 + 1)^2}$$

Nun Multiplikation mit dem Hauptnenner

$$= \frac{a_{11}(x-1)(x^2+1)^2 + a_{12}(x^2+1)^2 + (b_{11}x+c_{11})(x-1)^2(x^2+1) + (b_{12}x+c_{12})(x-1)^2}{(x-1)^2(x^2+1)^2}$$

Koeffizientenvergleich ergibt:

$$\begin{aligned}
x^5: &\quad 0 = a_{11} + b_{11} \\
x^4: &\quad 0 = -a_{11} + a_{12} - 2b_{11} + c_{11} \\
x^3: &\quad 0 = 2a_{11} + 2b_{11} - 2c_{11} + b_{12} \\
x^2: &\quad -1 = -2a_{11} + 2a_{12} - 2b_{11} + 2c_{11} - 2b_{12} + c_{12} \\
x^1: &\quad -6 = a_{11} + b_{11} - 2c_{11} + b_{12} - 2c_{12} \\
x^0: &\quad 3 = a_{11} + a_{12} + c_{11} + c_{12}
\end{aligned}$$

Auflösen des Gleichungssystems liefert:

$$a_{11} = 0,\ a_{12} = -1,\ b_{11} = 0,\ b_{12} = 2,\ c_{11} = 1,\ c_{12} = 3$$

Damit folgt:

$$\frac{-x^2-6x+3}{(x-1)^2(x^2+1)^2} = -\frac{1}{(x-1)^2} + \frac{1}{x^2+1} + \frac{2x+3}{(x^2+1)^2}$$

Beachte. Man kann die Koeffizienten auch durch Einsetzen geeigneter Werte von x bestimmen. Insbesondere gilt: Ist x_{i_0} eine k-fache Nullstelle von P, so gilt für a_{i_0k}:

$$a_{i_0k} = \frac{R(x)}{P(x)}(x-p)^k \bigg|_{x=p}$$

Im Beispiel:

$$a_{12} = \frac{-x^2-6x+3}{(x^2+1)^2}\bigg|_{x=1} = -1$$

Mit dem Theorem wird die Integration von rationalen Funktionen auf eine Reihe von bekannten Integralen zurueckgefuehrt. Hier geben wir die Ergebnisse an:

- $\int \frac{a}{x-p}dx = a\ln|x-p| + C.$
- $\int \frac{a}{(x-p)^n}dx = \frac{1}{1-n}\frac{a}{(x-p)^{n-1}} + C$ (
- $\int \frac{bx+c}{(x-\xi)^2+\mu^2}dx = \frac{b}{2}\int \frac{2x-2\xi}{(x-\xi)^2+\mu^2}dx + \frac{b\xi+c}{\mu}\int \frac{\frac{1}{\mu}}{\left(\frac{x-\xi}{\mu}\right)^2+1}$
  $= \frac{b}{2}\log\left((x-\xi)^2+\mu^2\right) + \frac{b\xi+c}{\mu}\arctan\left(\frac{x-\xi}{\mu}\right) + C$. Dabei wurde zunaechst das Intgral so umgeschrieben, dass der Zaehler die Ableitung des Nenners ist und anschliessend die Substitutionsregel angewendet.

- $\int \frac{bx+c}{((x-\xi)^2+\mu^2)^n}dx = \frac{b}{2(1-n)}\frac{1}{((x-\xi)^2+\mu^2)^{n-1}} + \frac{b\xi+c}{\mu^{2n-1}}\int\frac{dt}{(t^2+1)^n}$ mit $t = \frac{x-\xi}{\mu}$. Hier wurde aehnlich wie im vorigen Fall das Integral so umgeschrieben, dass der Zaehler die Ableitung des Terms unter der Klammer im Nenner ist, und anschliessend die Substitutionsregel angewendet. Fuer die auftretenden Terme
$$I_n = \int \frac{dt}{(t^2+1)^n}$$
gibt es eine Rekursion wie folgt:
$$I_{n+1} = \frac{1}{2n}\left(\frac{t}{(t^2+1)^n} + (2n-1)I_n(t)\right)$$
Insbesondere
$$I_2 = \frac{1}{2}\left(\frac{t}{t^2+1} + \arctan t\right) + C$$

4. Uneigentliche Integrale

Wir kommen nun zur uneigentlichen Integration. Bisher hatten wir immer vorausgesetzt:

- $[a,b]$ kompakt, d. h. a,b endlich
- $f:\ [a,b]\to\mathbb{R}$ beschränkt

Wir schwächen das nun ab.

Definition. *(Uneigentliche Riemann-Integrierbarkeit)*

- *Sei $-\infty \le a < b < \infty$. Die Funktion $f:\ (a,b]\to\mathbb{R}$ heißt uneigentlich Riemann-integrierbar bei a, wenn für $x > a$ die Einschränkung von f auf $[x,b]$ Riemann-integrierbar ist und der Grenzwert*
$$\lim_{x\to a+}\int_x^b f(t)dt$$
existiert. ***Zeichnung*** *In diesem Fall bezeichnen wir diesen Grenzwert mit*
$$\int_a^b f(t)dt$$
- *Sei $b \le \infty$ und $-\infty < a < b$. Die Funktion $f:\ [a,b)\to\mathbb{R}$ heißt uneigentlich Riemann-integrierbar bei b, wenn für $x < b$ die Einschränkung von f auf $[a,x]$ Riemann-integrierbar ist und der Grenzwert*
$$\lim_{x\to b-}\int_a^x f(t)dt$$
existiert. ***Zeichnung.*** *In diesem Fall bezeichnen wir diesen Grenzwert mit $\int_a^b f(t)dt$.*

- *Ist $f: (a,b) \to \mathbb{R}$ gegeben und sind für ein (jedes) $c \in (a,b)$ die Einschränkungen von f auf $(a,c]$ bzw. $[c,b)$ uneigentlich Riemann-integrierbar, so nennen wir f bei a und b uneigentlich Riemann-integrierbar* ***Zeichnung.*** *und setzen*

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$$

Bemerkung

- Die Definition enthält zwei relevante Situationen:
  - a bzw. b ist endlich und f hat Singularität, d. h. ist unbeschränkt bei a bzw. b, aber so, dass das Integral noch existiert.
  - Es gilt $a = -\infty$ bzw. $b = \infty$ und f faellt schnell genug bzw. oszilliert stark genug, um Existenz des Grenzwertes zu ermoeglichen.
- Auf die im Grenzwerte auftretenden Integrale können die üblichen Rechenreglen (Subtitutionsregel, partielle Integration...) angewendet werden. In diesem Sinne bleiben diese Regeln auch für uneigentliche Integrale gültig.

Beispiele.

- $f: [0,\infty) \to \mathbb{R}$, $f(t) = \exp(-t)$. Dann gilt

$$\int_0^\infty \exp(-t)dt = \lim_{x\to\infty} \int_0^x \exp(-t)dt = \lim_{x\to\infty} -\exp(-t)|_0^s = 1.$$

- Sei $\alpha > 0$, $f: (0,1] \to \mathbb{R}$, $f(t) = \frac{1}{t^\alpha}$. Es gilt

$$\int_0^1 \frac{1}{t^\alpha}dt = \lim_{x\to 0+} \int_x^1 \frac{1}{t^\alpha}dt = \lim_{x\to 0+} \frac{1}{1-\alpha} t^{1-\alpha}\Big|_x^1 \quad \text{bzw.} \quad \lim_{x\to 0+} \ln x|_x^1$$
$$= \begin{cases} \frac{1}{1-\alpha}: & 0 < \alpha < 1 \\ \infty: & \alpha \geq 1 \end{cases}$$

- Sei $\alpha > 0$, $f: [1,\infty) \to \mathbb{R}$, $f(t) = \frac{1}{t^\alpha}$. Es gilt

$$\int_0^\infty \frac{1}{t^\alpha}dt = \lim_{x\to\infty} \int_1^x \frac{1}{t^\alpha}dt = \lim_{x\to\infty} \frac{1}{1-\alpha} t^{1-\alpha}\Big|_1^x \text{; bzw.} \quad \lim_{x\to 0+} \ln x|_1^x$$
$$= \begin{cases} \frac{1}{1-\alpha}: & 1 < \alpha \\ \infty: & 0 < \alpha \leq 1 \end{cases}$$

- (Uebung): Es existiert $\int_1^\infty \frac{1}{x}\sin x dx$, aber nicht $\int_1 \infty \left|\frac{1}{x}\sin x\right| dx$.

Bemerkung. Wir halten fest: $\int_0^\infty \frac{1}{t^\alpha}dt$ existiert fuer kein $\alpha \in \mathbb{R}$. (Fuer $\alpha > 0$ wurde das gerade oben diskutiert. Fuer $\alpha \leq 0$ ist die Aussage sowieso klar.)

Eine wichtige Funktion (für 'später') wird mittels uneigentlichem Integral definiert.

PROPOSITION. *Für $x > 0$ existiert das uneigentliche Integral*

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt.$$

Es gilt

$$\Gamma(1) = 1 \;\; \textit{sowie} \;\; \Gamma(x+1) = x\Gamma(x).$$

Insbesondere folgt $\Gamma(n+1) = n!$ *für alle* $n \in \mathbb{N}$

Beweis. Wir zeigen zunächst Existenz des Integrals bei 0 und ∞. Da der Integrand positiv ist, geht es letzten Endes darum geeignete Beschraenktheiten zu zeigen.
Existenz bei 0 *(für* $0 < x < 1$*):* Für $x < 1$ ist der Integrand unbeschränkt. Es existiert $\int_0^1 t^{x-1} dt$ für alle $x - 1 > -1$ (nach obigem). Da e^{-t} auf $[0, 1]$ beschränkt und stetig ist, existiert dann das uneigentliche Integral bei 0.
Existenz bei ∞*:* Für genügend große t gilt

$$\left|t^{x-1}e^{-t}\right| = \left|\underbrace{t^{x-1}e^{-\frac{t}{2}}}_{\leq 1}\right| \left|e^{-\frac{t}{2}}\right| \leq \left|e^{-\frac{t}{2}}\right|$$

Es ergibt sich

$$\int_1^R t^{x-1}e^{-t}dt = \int_1^{R_0} t^{x-1}e^{-t}dt + \int_{R_0}^R t^{x-1}e^{-t}dt \leq C + \int_{R_0}^R e^{-t/2}dt \leq \tilde{C} < \infty.$$

Damit existiert

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt = \lim_{r\to 0} \int_r^1 t^{x-1}e^{-t}dt + \lim_{R\to\infty} \int_1^R t^{x-1}e^{-t}dt$$

Wir kommen nun zu den behaupteten Eigenschaften: Es gilt (s.o.) $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$. Weiterhin gilt mit partieller Integration:

$$\Gamma(x+1) = \lim_{r\to 0} \int_r^1 t^x e^{-t}dt + \lim_{R\to\infty} \int_1^R t^x e^{-t}dt$$

$$= \lim_{r\to 0} \left(-t^x e^{-t}|_r^1 + \int_r^R x t^{x-1} e^{-t}dt \right) + \lim_{R\to\infty} \left(-t^x e^{-t}|_1^R + \int_1^R x t^{x-1} e^{-t}dt \right)$$

$$= 0 + x\Gamma(x).$$

Das 'Insbesondere' folgt sofort aus den Eigenschaften und einer Induktion. □

DEFINITION. *Die Abbildung*

$$\Gamma(0, \infty) \to \mathbb{R}, \; x \mapsto \Gamma(x),$$

heißt Gammafunktion.

KAPITEL 12

Die Exponentialfunktion und ihre Verwandten

In diesem Kapitel untersuchen wir noch einmal systematisch die Exponentialfunktion und von ihr abgeleitete Funktionen.

Wir erinnern zunaechst an einige Eigenschaften. Die Exponentialfunktion ist definiert durch die auf ganz $\mathbb{C}$ absolut konvergente Reihe

$$\exp : \mathbb{C} \to \mathbb{C},\ \exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Sie hat folgende Eigenschaften.

- $\exp(0) = 1$ und $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$. Insbesondere verschwindet exp nirgends und es gilt also $\exp(-z) = \frac{1}{\exp(z)}$.
- Die Exponentialfunktion ist stetig und es existiert

$$\lim_{z \to z_0} \frac{\exp(z) - \exp(z_0)}{z - z_0} = \exp(z_0).$$

Proposition (Restgliedabschätzung der Exponentialreihe). *Sei das Restglied der Exponentialreihe fuer $N \in \mathbb{N}$ definiert durch $R_N(z) = \sum_{k=N+1}^{\infty} \frac{z^k}{k!}$. Dann gilt für $|z| \leq 1 + \frac{N}{2} = \frac{2+N}{2}$*

$$|R_N(z)| \leq 2\frac{|z|^{N+1}}{(N+1)!}$$

Beweis. *Sei $|z| \leq 1 + \frac{N}{2}$ und $M > N$ beliebig. Dann gilt:*

$$\begin{aligned}
\sum_{k=N+1}^{M} \frac{|z|^k}{k!} &= \frac{|z|^{N+1}}{(N+1)!}\left(1 + \underbrace{\frac{|z|}{N+2}}_{\leq \frac{1}{2}} + \ldots + \underbrace{\frac{|z|^{M-N-1}}{(N+1)\ldots M}}_{\leq \left(\frac{1}{2}\right)^{M-N-1}}\right)\\
&\leq \frac{|z|^{N+1}}{(N+1)!}\sum_{l=0}^{\infty}\left(\frac{1}{2}\right)^{l}\\
&= 2\frac{|z|^{N+1}}{(N+1)!}
\end{aligned}$$

Das liefert die Behauptung. $\square$

Bemerkung. Bei festem z wird z/n beliebig klein (fuer n gross) und daher konvergiert $R_N(z)$ schliesslich extrem schnell gegen 0 fuer $N \to \infty$ (auch wenn es fuer die kleinen N nicht unbedingt so aussieht, siehe etwa $z = 1000...$).

FOLGERUNG. *Die Zahl* $e = \exp(1)$ *ist irrational.*

Beweis. Annahme: $e = \frac{p}{q}$ mit $p, q \in \mathbb{N}$. Sei o. E. $q \geq 2$ Dann gilt also

$$\underbrace{q!e}_{\in\mathbb{Z}} = \underbrace{q! \sum_{k=0}^{q} \frac{1}{k!}}_{\in\mathbb{Z}} + \underbrace{\lim_{M\to\infty} \left(q! \sum_{k=q+1}^{M} \frac{1}{k!} \right)}_{=q!R_q(1)}$$

und damit

$$q!R_q(1) \in \mathbb{Z}.$$

Weiterhin gilt aber nach der gerade gezeigten Abschaetzung fuer das Restglied:

$$0 < q!R_q(1) \leq 2\frac{|1|^{q+1}}{(q+1)!} \leq \frac{2}{3} < 1.$$

Das ist ein Widerspruch. $\square$

Wir kommen nun zu einer weiteren Darstellung für $\exp(z)$:

THEOREM (Exponentialfunktion als Grenzwert). *Für* $z \in \mathbb{C}$ *gilt*

$$\exp(z) = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$$

Bemerkung. Die Konvergenz ist sehr langsam. Daher ist die Formel für praktische Anwendungen nicht so nützlich wie die Exponentialreihe. Für konzeptuelle Fragen ist es aber eine sehr nützliche Formel (Wachstumsprozesse).

Beweis. Es gilt

$$\left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} \frac{z^k}{n^k} = \sum_{k=0}^{n} \frac{n!}{(n-k)n^k} \frac{z^k}{k!}.$$

Vergleicht man mit der Exponentialreihe $\sum_{k=0}^{\infty} \frac{z^k}{k!}$, so sieht man, dass es darum gehen wird den Abstand von $\frac{n!}{(n-k)n^k}$ zu 1 untersuchen. Offenbar gilt

- $\frac{n!}{(n-k)n^k} \leq 1$ fuer alle $k, n \in \mathbb{N}$ mit $k \leq n$.
- $\frac{n!}{(n-k)n^k} \to 1$, $n \to \infty$ (bei festem $k \in \mathbb{N}$).

Ende der 3. Vorlesung.

Das werden wir nun nutzen: Für $n, N \in \mathbb{N}$ mit $N < n$ gilt:

$$\left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{N} \frac{n!}{(n-k)n^k} \frac{z^k}{k!} + \sum_{k=N+1}^{n} \frac{n!}{(n-k)n^k} \frac{z^k}{k!}$$

sowie

$$\exp(z) = \sum_{k=0}^{N} \frac{z^k}{k!} + \sum_{k=N+1}^{\infty} \frac{z^k}{k!}.$$

Damit folgt aus der Dreicksungleichung

$$\left|\left(1+\frac{z}{n}\right)^n - \exp(z)\right| \leq \left|\sum_{k=0}^{N}\left(\frac{n!}{(n-k)!n^k}-1\right)\frac{z^k}{k!}\right| + \left|\sum_{k=N+1}^{n}\frac{n!}{(n-k)n^k}\frac{z^k}{k!}\right| + \left|\sum_{k=N+1}^{\infty}\frac{z^k}{k!}\right|.$$

Nochmaliges Anwenden der Dreiecksungleichung und Nutzen von $\frac{n!}{(n-k)n^k} \leq 1$ liefert

$$\left|\left(1+\frac{z}{n}\right)^n - \sum_{k=0}^{\infty}\frac{z^k}{n^k}\right| \leq \sum_{k=0}^{N}\left|\frac{n!}{(n-k)!n^k}-1\right|\frac{|z|^k}{k!} + 2R_N(|z|).$$

Ist nun $\varepsilon > 0$ gegeben, so koennen wir nun zunächst N so groß waehlen, dass $2R_N(|z|) < \frac{\varepsilon}{2}$ gilt. Anschliessend wählen wir dann n_ε so groß, sodass der erste Term kleiner als $\frac{\varepsilon}{2}$ wird fuer alle $n \geq n_\varepsilon$. Dann gilt fuer $n \geq n_\varepsilon$ also

$$\left|\left(1+\frac{z}{n}\right)^n - \exp(z)\right| \leq \varepsilon$$

und die Behauptung folgt. □

Bemerkung. (Übung) Zu $a > 0$ existiert genau eine stetige Funktion $F: \mathbb{R} \to \mathbb{R}$ mit

- $F(x+y) = F(x)F(y)$ alle $x, y \in \mathbb{R}$
- $a = F(1)$

Diese ist gegeben durch $F(x) = \exp_a(x)$ *Beweis. Existenz:* $F(x) = \exp(x \ln a)$ erfuellt offenbar die gewuenschten Eigenschaften.

Eindeutigkeit - Skizze: Zeige F erfüllt notwendigerweise $F(p/q) = a^{p/q}$. Nutze dann die Stetigkeit von F und die Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$... □

Die Exponentialfunktion auf $\mathbb{R}$ liefert auch die sogenannten **hyperbolische Funktionen**. Das diskutieren wir nun.

cosinus hyperbolicus	$\cosh: \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{1}{2}(e^x + e^{-x})$
sinus hyperbolicus	$\sinh: \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{1}{2}(e^x - e^{-x})$
tangens hyperbolicus	$\tanh: \mathbb{R} \to (-1,1),\ x \mapsto \frac{\sinh x}{\cosh x}$
cotangens hyperbolicus	$\coth: \mathbb{R} \setminus \{0\} \to (-\infty,1) \cup (1,\infty),\ x \mapsto \frac{\cosh x}{\sinh x}$

Es gilt:

$$\begin{aligned} 1 &= \cosh^2(x) - \sinh^2(x) \\ \cosh' &= \sinh, \ \sinh' = \cosh \\ \tanh' &= \frac{1}{\cosh^2} = 1 - \tanh^2 \\ \coth' &= -\frac{1}{\sinh^2} = 1 - \coth^2 \end{aligned}$$

Auf entsprechenden Intervallen existieren auch die Umkehrfunktionen (auch Area-Funktionen genannt):

- $\operatorname{arcosh}: \ [1, \infty) \to [0, \infty), \ x \mapsto \ln(x\sqrt{x^2 - 1})$
  $\operatorname{arcosh}'(y) = \frac{1}{\sinh x} = \frac{1}{\sqrt{\cosh^2 x - 1}} = \frac{1}{\sqrt{y^2-1}} \quad (y = \cosh x)$
- $\operatorname{arsinh}: \ \mathbb{R} \to \mathbb{R}, \ x \mapsto \ln(x + \sqrt{x^2 + 1})$.
  $\operatorname{arsinh}'(y) = \frac{1}{\sqrt{y^2+1}}$
- $\operatorname{artanh}: (-1, 1) \to \mathbb{R}, x \mapsto \frac{1}{2}\ln(1 + x)/(1 - x)$.
  $\operatorname{artanh}'(y) = \frac{1}{1-y^2} \quad (y \in (-1, 1))$
- $\operatorname{arcoth}: \ (-\infty, -1) \cup (1, \infty) \to \mathbb{R}, \ x \mapsto \frac{1}{2}\ln(x + 1)/(x - 1)$.
  $\operatorname{arcoth}'(y) = \frac{1}{1-y^2} \quad (y \in \mathbb{R} \setminus [-1, 1])$

Wir untersuchen nun die Exponentialfunktion auf dem Einheitskreis

$$S = \{x \in \mathbb{C} : \ |z| = 1\} \tag{1}$$

Wir betrachten also:

$$u: \ \mathbb{R} \to \mathbb{C}, \ u(x) = e^{ix} \tag{2}$$

und definieren:

$$\Re u = \cos x \tag{3}$$

$$\Im u = \sin x \tag{4}$$

Wir werden zeigen, dass man e^{ix} erhält, indem man auf dem Einheitskreis gegen den Uhrzeigersinn die Strecke x entlang geht. **Zeichnung.** Aus den Differenzierbarkeitseigenschaften der Exponentialfunktion folgt leicht, dass cos und sin differenzierbar sind mit

$$\cos' = -\sin, \ \ \sin' = \cos .$$

Weiterhin folgt aus der Funktionnalgleichung der Exponentialfunktion die Additionstheoreme für sin und cos

$$\sin(u \pm v) = \sin u \cos v \pm \cos u \sin v$$

$$\cos(u \pm v) = \cos u \cos v \mp \sin u \sin v$$

(Bew. Bilde Real- bzw. Imaginarteil auf beiden Seiten von $e^{i(u+v)} = e^{iu}e^{iv} \ldots$.)

Wir sammeln nun einige weitere Eigenschaften von sin und cos.

Lemma (Eigenschaften von sin und cos). (a) *Für alle $t \in \mathbb{R}$ gilt:*

$$\begin{aligned} \sin t &= \textstyle\sum_{k=0}^{\infty}(-1)^k \frac{t^{2k+1}}{(2k+1)!} &&= t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots \\ \cos t &= \textstyle\sum_{k=0}^{\infty}(-1)^k \frac{t^{2k}}{(2k)!} &&= 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots \end{aligned}$$

(b) *Es ist* sin *ungerade, d. h.* $\sin(-t) = -\sin(t)$ *und* cos *gerade, d. h.* $\cos(t) = -\cos(t)$

(c) *Es gilt* $\sin^2 t + \cos^2 t = 1 \quad \forall t \in \mathbb{R}$

Beweis. (a) Das folgt aus der Reihendarstellung für $\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ und den Definitionen $\sin t = \Im e^{it}$, $\cos t = \Re e^{it}$

(b) Mit $\overline{e^{ix}} = e^{-ix}$ folgt dies direkt aus den Definitionen von sin und cos. Alternativ kann es auch direkt aus (a) geschlossen werden.

(c) Es gilt

$$\sin^2 t + \cos^2 t = (i \sin t + \cos t)^2 = \left(e^{it}\right)^2 = e^{it}\overline{e^{it}} = e^{it}e^{\overline{it}} = e^{it}e^{-it} = e^0 = 1.$$

□

Proposition. *(Abschätzung für* sin *und* cos*)*
Für $t \in \mathbb{R}$ gilt $\cos t \leq 1 - \frac{t^2}{2!} + \frac{t^4}{4!}$.
Für $t > 0$ gilt $\sin t \geq t - \frac{t^3}{3!}$.

Beweis. Hier wird nur die Abschätzung für cos bewiesen (sin kann analog behandelt werden). Wir unterscheiden zwei Fälle:

$\underline{|t| < 7:}$ $\cos t = 1 - \frac{t^2}{2!} + \frac{t^4}{4!} \underbrace{-\frac{t^6}{6!} + \frac{t^8}{8!}}_{<0} \underbrace{- \dots}_{<0} \leq 1 - \frac{t^2}{2!} + \frac{t^4}{4!}$

$\underline{|t| \geq 7:}$ $1 - \frac{t^2}{2!} + \frac{t^4}{4!} = 1 + t^2\left(\frac{t^2}{24} - \frac{1}{2}\right) \geq 1 \geq \cos t$ □

Proposition (Analytische Beschreibung von π). *Die Funktion* sin *erfüllt* $\sin 0 = 0$ *und ist strikt positiv auf* $(0, 2]$. *Die Funktion* cos *erfüllt* $\cos 0 = 1$ *und* $\cos(2) < 0$ *und ist auf* $[0, 2]$ *strikt fallend. Insbesondere hat* cos *in* $(0, 2)$ *genau eine Nullstelle und in dieser Nullstelle nimmt* sin *den Wert* 1 *an.*

Bemerkung. Zeichnung zeigt, dass die Aussage der Proposition sofort folgt, wenn wir schon wuessten, dass e^{ix} am Enheitskreis durch Bewegung gegen den Uhrzeigersinn der Laenge x entsteht.
Beweis. Aussagen zu sin. Es gilt $\sin 0 = \Im e^{i0} = \Im 1 = 0$ sowie nach dem vorigen Proposition (fuer $0 \leq t \leq 2$

$$\sin t \geq t - \frac{t^3}{3!} = t\left(1 - \underbrace{\frac{t^2}{3!}}_{<1}\right) > 0.$$

Aussagen zu $\cos$*:* Es gilt $\cos 0 = 1$ da $1 = e^{i0} = \cos 0 + i \sin 0$. Weiterhin ist $\cos$ wegen $\cos' = -\sin$ und dem schon gezeigeten $\sin > 0$ also strikt fallend auf $(0,2]$ (Mittelwertsatz). Außerdem gilt nach der vorigen Proposition

$$\cos 2 \leq 1 - \frac{2^2}{2!} + \frac{2^4}{24} = 1 - 2 + \frac{16}{24} = -1/3 < 0.$$

Nach dem Zwischenwertsatz hat also $\cos$ auf $(0,2)$ eine Nullstelle. Da $\cos$ strikt fallend ist, ist diese Nullstelle eindeutig. □

Definition (Analytische 'Definition' von π). *Sei $\pi \in \mathbb{R}$ so, dass $\frac{\pi}{2}$ die eindeutige Nullstelle von* $\cos$ *in* $(0,2)$ *ist.*

Theorem (Exponentialfunktion auf dem Einheitskreis). *Die Abbildung*

$$e^{i\cdot} : \mathbb{R} \longrightarrow S, x \mapsto e^{ix},$$

erfüllt $e^{i(x+y)} = e^{ix}e^{iy}$ sowie

$$e^{i0} = 1,\ e^{i\frac{\pi}{2}} = i, e^{i\pi} = -1, e^{i\frac{3}{2}\pi} = -i\ \ e^{i2\pi} = 1.$$

Insbesondere ist die Abbildung also 2π periodisch d.h. es gilt $e^{i(x+2\pi)} = e^{ix}$ für alle $x \in \mathbb{R}$. Schliesslich ist die Einschränkung der Abbildung auf $[0, 2\pi)$ bijektiv (auf S).

Zeichnung. Bilder von $(e^{i\frac{\pi}{2}})^k$. Zerlege $[0, 2\pi)$ entsprechend in vier Intervalle, zeichne deren Bild am Einheitskreis.

Beweis. Die Funktionalgleichung der Exponentialfunktion liefert

$$(*)\ \ e^{i(x+y)} = e^{ix}e^{iy}.$$

Nach Konstruktion gilt $e^{i\frac{\pi}{2}} = \cos(\pi/2) + i\sin(\pi/2) = i$. Damit folgt fuer $k \in \mathbb{N}$ also aus $(*)$

$$e^{ik\frac{\pi}{2}} = (e^{i\frac{\pi}{2}})^k = (i)^k.$$

Das liefert die entsprechenden Aussagen. Die Periodizitaet folgt dann aus $(*)$ wegen

$$e^{i(x+2\pi)} = e^{ix}e^{i2\pi} = e^{ix}1 = e^{ix}.$$

Es bleibt die Bijektivitaet zu zeigen. Nach dem schon gezeigten $(*)$ reicht es, die Bijektivitaet der Abbildung

$$e^{i\cdot} : \left[0, \frac{\pi}{2}\right] \to S \cap \{u + iv : u, v \geq 0\}$$

zu zeigen. Dazu zeigen wir zunaechst, dass $\cos : [0, \frac{\pi}{2}] \longrightarrow [0,1]$ bijektiv ist:

Bew: Nach voriger Proposition und Definition ist $\cos : \left[0, \frac{\pi}{2}\right]$ strikt fallend mit $\cos(0) = 1$ und $\cos\left(\frac{\pi}{2}\right) = 0$. Damit ist $\cos : \left[0, \frac{\pi}{2}\right] \to \mathbb{R}$ injektiv (da strikt fallend) und sein Bild ist nach dem Zwischenwertsatz $[0,1]$.

Injektivität von $e^{i\cdot}$ auf $[0, \frac{\pi}{2}]$: Das folgt sofort, da $\cos t = \Re e^{it}$ injektiv ist.

Surjektivität von $e^{i\cdot}$ auf $[0, \frac{\pi}{2}]$: Sei $a, b \geq 0$ mit $a^2 + b^2 = 1$ gegeben. Dann existiert nach dem schon Bewiesenen ein $t \to \left[0, \frac{\pi}{2}\right]$ mit $a = \cos t$. Dann gilt: $b^2 = 1 - a^2 = 1 - \cos^2 t = \sin^2 t \overset{b \geq 0, \sin t \geq 0}{\Rightarrow} b = \sin t \Rightarrow a + ib = \cos t + i \sin t = e^{it}$ Wegen $e^{i(t+s)} = e^{it}e^{is}$ folgen nun die Aussagen für die übrigen drei Quadranten einfach. □

Bemerkung. Die Länge des Bogens auf S zwischen 1 und e^{it} ist gerade t. Zum Beweis betrachten wir den Polygonzug von 1 nach e^{it} mit Stützstellen $e^{i\frac{t}{n}k}$, $k = 0, 1, \dots, n$ **Zeichnung**. Seine Länge ist gegeben durch

$$\sum_{k=0}^{n-1} \left| e^{it\frac{k+1}{n}} - e^{it\frac{k}{n}} \right| = \sum_{k=0}^{n-1} \underbrace{\left| e^{itk} \right|}_{=1} \left| e^{i\frac{t}{n}} - 1 \right| = \sum_{k=0}^{n-1} \left| e^{i\frac{t}{n}} - 1 \right|$$

$$= n \left| e^{i\frac{t}{n}} - 1 \right| = \frac{\left| e^{i\frac{t}{n}} - 1 \right|}{\frac{1}{n}} = t \left| \frac{e^{i\frac{t}{n}} - 1}{i\frac{t}{n}} \right|.$$

Verfeinert man diesen Polygonzug d.h. bildet man den Grenzwert $n \to \infty$, so erhaelt man die sogenannte Bogenlaenge. Diese ist dann also gegeben durch

$$L = \lim_{n \to \infty} t \underbrace{\left| \frac{e^{i\frac{t}{n}} - 1}{i\frac{t}{n}} \right|}_{\to 1,\ n \to \infty} = t.$$

(Hier nutzen wir $\frac{e^z - 1}{z} \to 1,\ z \to 0$.)

Ende der Vorlesung

Folgerung (Periodizität von sin und cos). *Die Funktionen* sin *und* cos *sind 2π-periodisch und es gilt*

$$\cos x = 0 \ \Leftrightarrow\ x = \frac{\pi}{2} + k\pi \ \textit{mit}\ k \in \mathbb{Z}$$
$$\sin x = 0 \ \Leftrightarrow\ x = k\pi \ \textit{mit}\ k \in \mathbb{Z}$$

Beweis. Argumentation am Einheitskreis □

Mittels des vorigen Theorems können wir die sogenannte Polarzerlegung einführen. Polarzerlegung:

Theorem (Polarzerlegung). *Zu jedem $z \in \mathbb{C}\setminus\{0\}$ existieren eindeutige $\rho > 0$ und $\phi \in [0, 2\pi)$ mit $z = \rho e^{i\phi}$.*

Beweis. Existenz: $z = |z| \underbrace{\frac{z}{|z|}}_{\in S}$ Dann existiert $\phi \in [0, 2\pi)$ mit $e^{i\phi} = \frac{z}{|z|}$.

Mit diesem ϕ und $\rho = |z|$ folgt dann $z = \rho e^{i\phi}$.

Eindeutigkeit: Es gelte $z = \rho e^{i\phi}$. Dann folgt $|z| = \rho \left| e^{i\phi} \right| = \rho$. Damit ist also ρ eindeutig bestimmt. Sei nun $e^{i\phi} = e^{i\psi}$. Dann gilt also

$$1 = e^{i(\phi - \psi)}$$

und damit

$$\phi - \psi = k2\pi$$

mit $k \in \mathbb{Z}$. Wegen $\phi, \psi \in [0, 2\pi)$ folgt $\phi = \psi$. $\square$

DEFINITION. *Ist $z = \rho e^{i\phi}$ mit $\rho > 0$ und $\phi \in \mathbb{R}$, so nennt man ρ den Betrag von z, und ϕ das Argument von z.*

Bemerkung Es gilt $z_1 z_2 = \rho_1 e^{i\phi_1} \rho_2 e^{i\phi_2} = \rho_1 \rho_2 e^{i(\phi_1+\phi_2)}$ Der Betrag des Produkts ist damit gleich dem Produkt der Beträge und das Argument des Produktes ist die Summe der Argumente.
Damit können wir nun eine ganze Reihe weiterer Funktionen einführen. Das diskutieren wir nun.

Tangens und Arcustangens. Für x mit $\cos x \neq 0$, d. h. $x \in \mathbb{R} \setminus \{k\pi + \frac{\pi}{2},\ k \in \mathbb{Z}\}$ definiert man $\tan x = \frac{\sin x}{\cos x}$. Dann ist $\tan x$ π-periodisch und stetig differenzierbar (Quotientenregel anwenden) mit:

$$tan'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x > 0$$

Wegen $\tan' > 0$ existiert die Umkehrfunktion

$$\arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Es ist arctan stetig differenzierbar mit

$$\arctan'(y) = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2} \quad (y = \tan x)$$

Contangens und Arcuscotangens. Für $x \in \mathbb{R} \setminus \{k\pi : \quad k \in \mathbb{Z}\}$ definiert man den Cotangens durch

$$\cot x = \frac{\cos x}{\sin x}$$

Dann ist cot π-periodisch und stetig differenzierbar mit

$$\cot' x = -1 - \cot^2 x = -\frac{1}{\sin^2 x} < 0$$

Insbesondere ist cot strikt fallend und es existiert die Umkehrfunktion

$$\operatorname{arccot} : \mathbb{R} \to (0, \pi)$$

Es ist arccot stetig differenzierbar mit

$$\operatorname{arccot}'(y) = \frac{1}{-1 - \cot^2 x} = \frac{1}{-1 - y^2} = -\frac{1}{1 + y^2} \quad (y = \cot x)$$

KAPITEL 13

Metrische Räume und topologische Grundbegriffe

Bisher haben wir Analysis (Grenzwert, Konvergenz, Stetigkeit, Differenzierbarkeit) auf $\mathbb{R}$ betrieben. Tatsächlich interessiert man sich oft für allgemeinere Situationen. Dabei geht es dann zum Beispiel um (Teilmengen) von

- $\mathbb{R}^N = \{(\xi_1, \ldots, \xi_N) : \xi_j \in \mathbb{R}\} = \{x : \{1, \ldots, N\} \longrightarrow \mathbb{R}\}$ bzw. $\mathbb{C}^N = \{(\xi_1, \ldots, \xi_N) : \xi_j \in \mathbb{C}\} = \{x : \{1, \ldots, N\} \longrightarrow \mathbb{C}\}$
- $M^{N \times N}$ = N× N-Matrizen über $\mathbb{R}$ bzw $\mathbb{C}$.
- $\ell^2 = \{x : \mathbb{N} \longrightarrow \mathbb{C} : \sum |x_j|^2 < \infty\}$.

Will man auf diesen Mengen Grenzwerte zu untersuchen / Analysis betreiben, so erweist sich das Konzept der Metrik als grundlegend.

1. Metrische Räume

In diesem Abschnitt geht es um Metrik. Das Konzept der Metrik gibt einen quantitativen Abstandsbegriff. Mit diesem Abstandsbegriff kann dann Konvergenz und damit Stetigkeit gefaßt werden.

Der Abstand zwischen zwei Punkten sollte folgende Eigenschaften haben: **Zeichnung.**
- Abstand zwischen Punkten haengt nicht von der Reihenfolge der Punkte ab,
- Abstand erfuellt Dreiecksungleichung
- Verschiedene Punkte haben positiven Abstand.

Eine praezise Fassung gibt folgende Definition.

Definition. *Sei X eine Menge. Eine Metrik auf X ist eine Abbildung $d : X \times X \longrightarrow [0, \infty)$ fuer die gilt:*

(M1) *$d(x, y) = d(y, x)$ fuer alle $x, y \in X$ (Symmetrie)*
(M2) *$d(x, y) \leq d(x, z) + d(z, y)$ fuer alle $x, y, z \in X$. (Dreiecksungleichung)*
(M3) *$d(x, y) = 0$ genau dann wenn $x = y$ (Nichtausgeartet)*

Ein Paar (X, d) bestehend aus einer Menge X und einer Metrik d auf X heisst metrischer Raum.

Beispiel. Ist (X, d) ein metrischer Raum und Y eine Teilmenge von X, so ist auch (Y, d) ein metrischer Raum.

Beispiele.

- Auf $\mathbb{R}$ definiert $d(x,y) = |x-y|$ eine Metrik. Ähnlich definiert auf $\mathbb{C}$ dann $d(v,w) = |v-w|$ eine Metrik.
- Auf $\mathbb{Z}^N$ ist die Manhatten Metrik / Blockmetrik gegeben durch
$$d(x,y) = \sum_{j=1}^{N} |x_j - y_j|.$$
(Zeichnung mit Bloecken)
- Auf $\mathbb{R}^N$ (bzw. $\mathbb{C}^N$) ist die ℓ^1 Metrik gegeben durch
$$d_1(x,y) = \sum_{j=1}^{n} |x_j - y_j|.$$
(Das verallgemeinert die Manhatten Metrik.)
- Auf $\mathbb{R}^N$ (bzw. $\mathbb{C}^N$) ist die ℓ^∞ Metrik gegeben durch
$$d_\infty = \max\{|x_j - y_j| : j = 1, \ldots, N\}.$$
- Erwaehne l^p-Metrik. Zeige Uebung $d_p(x,y) \to d_\infty(x,y)$.
- Auf der endlichen Menge $\mathcal{A}$ (Alphabet) wird durch
$$d_D(a,b) := 1 - \delta_{a,b}$$
eine Metrik gegeben.
 Auf beliebigem X ist die diskrete Metrik gegeben durch
$$d_D(x,y) := 1 \quad \text{falls } x \neq y, \; d_D(x,y) = 0 \;\text{ falls } x = y.$$
- Sei $\mathcal{A}$ eine endliche Menge (Alphabet). Sei
$$\mathcal{W} : \mathcal{W}_{\mathcal{A}} = \{x : \mathbb{N} \longrightarrow \mathcal{A}\} = \{\text{Folgen mit Werten in } \mathcal{A}\}.$$

 - Ist $\mathcal{A}$ das uebliche Aphabet mit Satzzeichen, so handelt es sich bei $\mathcal{W}$ um die Menge aller 'Bücher'.
 - Ist $\mathcal{A} = \{U, C, A, G\}$, so handelt es sich bei $\mathcal{W}$ um die Menge aller (unendlichen) Gensequenzen.
 - Ist $\mathcal{A} = \{0, 1\}$, so handelt es sich bei $\mathcal{W}$ um die Menge aller (unendlichen) Computerprogramme.

 Es definiert
$$d(v,w) := \sum_{j=1}^{\infty} \frac{d_D(v_j, w_j)}{2^j}$$
eine Metrik auf $\mathcal{W}$. Damit kann man dann den Abstand zwischen zwei Genen oder zwischen zwei Büchern ;-) berechnen.

PROPOSITION. *(Umgekehrte Dreiecksungleichung) Sei (X,d) ein metrischer Raum. Dann gilt*
$$|d(x,y) - d(x,z)| \leq d(y,z)$$
fuer alle $x, y, z \in X$.

Beweis. Es gilt

$$d(x,y) \leq d(x,z) + d(z,y)$$

sowie nach Vertauschen von y und z

$$d(x,z) \leq d(x,y) + d(y,z).$$

Damit folgt die Aussage. □

Auf Vektorräumen werden Metriken meist von Normen induziert. Eine Norm liefert ein Konzept von Länge eines Vektors. Normen stellen also eine Verallgemeinerung der Betragsfunktion auf mehrdimensionale Situationen dar. Der Abstand zwischen zwei Vektoren ist dann die Länge ihrer Differenz. **Zeichnung.**

DEFINITION. *Sei V ein Vektorraum über dem Körper $K = \mathbb{R}$ oder $K = \mathbb{C}$. Eine Abbildung $\|\cdot\| : V \longrightarrow [0,\infty)$ heisst Norm, wenn gilt*

(N1) $\|\alpha x\| = |\alpha|\|x\|$ *fuer alle $\alpha \in K$ und $x \in V$.*
(N2) $\|x+y\| \leq \|x\| + \|y\|$ *fuer alle $x, y \in V$.*
(N3) $\|x\| = 0$ *genau dann, wenn $x = 0$.*

PROPOSITION. *(Umgekehrte Dreiecksungleichung) Ist $\|\cdot\|$ eine Norm auf V, so gilt die umgekehrte Dreiecksungleichung*

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Beweis. Nach Dreiecksungleichung gilt

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$$

und nach Vertauschen von x und y ebenso

$$\|y\| = \|y - x + x\| \leq \|y - x\| + \|x\|.$$

Damit folgt die Aussage leicht. □

DEFINITION. *Sei V eine Vektorraum mit Norm $\|\cdot\|$. Dann heisst*

$$d_{\|\cdot\|} : V \times V \longrightarrow [0,\infty), (v,w) \mapsto d_{\|\cdot\|}(v,w) := \|v - w\|$$

die durch $\|\cdot\|$ induzierte Metrik.

Beispiele.

- Auf $V = \mathbb{R}$ bzw. $V = \mathbb{C}$ wird durch $|\cdot|$ eine Norm gegeben.
- Auf $V = \mathbb{R}^N$ (bzw. $V = \mathbb{C}^N$) wird durch $\|x\|_1 := \sum_{j=1}^N |x_j|$ eine Norm, die sogenannte ℓ^1-Norm definiert. Die induzierte Metrik ist die ℓ^1 Metrik.
- Auf $V = \mathbb{R}^N$ (bzw. $V = \mathbb{C}^N$) wird durch $\|x\|_\infty := \max\{|x_j| : j = 1,\ldots,N\}$ eine Norm definiert, die sogenannte ℓ^∞-Norm. Die induzierte Metrik ist die ℓ^∞ Metrik.

Jedes Skalarprodukt liefert eine Norm (und damit eine Metrik).

PROPOSITION. *Sei V eine Vektorraum über $K = \mathbb{R}$ oder $K = \mathbb{C}$. Sei $\langle \cdot, \cdot \rangle : V \times V \longrightarrow K$ eine Skalarprodukt. Dann ist*

$$\| \cdot \| : V \longrightarrow [0, \infty), \|x\| := \langle x, x \rangle^{1/2}$$

eine Norm.

Beweis. Wir zeigen die Dreiecksungleichung. Die uebrigen Aussagen folgen einfach. Es gilt nach Cauchy Schwarz Ungleichung:

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Das beendet den Beweis. □

Beispiel - bekannteste Metrik und Norm. Die *euklidische Metrik* auf $\mathbb{R}^N$ bzw. $\mathbb{C}^N$ ist gegeben durch

$$d_2(x, y) = \left(\sum_{j=1}^N |x_j - y_j|^2 \right)^{1/2}.$$

Sie wird durch das Standardskalarprodukt

$$\langle x, y \rangle := \sum_{j=1}^N x_j \overline{y_j}$$

bzw. die dazugehoerige Norm

$$\|x\|_2 := \left(\sum_{j=1}^N |x_j|^2 \right)^{1/2}$$

induziert.

Ende der Vorlesung

Beispiel - Der Vektorraum ℓ^2. Wir betrachten die Menge

$$\ell^2 = \{x : \mathbb{N} \longrightarrow \mathbb{C} : \sum |x(j)|^2 < \infty\}.$$

Dann ist ℓ^2 ein Untervektorraum des Raumes aller komplexwertigen Funktionen auf $\mathbb{N}$. Fuer beliebige $x, y \in \ell^2$ ist

$$\sum_{j=1}^\infty x(j)\overline{y(j)}$$

absolut konverent und durch

$$\langle \cdot, \cdot \rangle : \ell^2 \longrightarrow \mathbb{C}, \langle f, g \rangle := \sum_{j=1}^\infty f(j)\overline{g(j)}$$

wird ein Skalarprodukt auf ℓ^2 definiert.

Wir zeigen zunächst, dass es sich um einen Unterraum handelt: Sei $x, y \in \ell^2$ und $\alpha \in \mathbb{C}$. Dann gilt

- $\alpha x \in \ell^2$: klar.
- $x + y \in \ell^2$: Es gilt
$$|x(j) + y(j)|^2 \leq 2|x(j)|^2 + 2|y(j)|^2.$$
Damit folgt die Aussage sofort.

Wir zeigen nun, dass die Summe absolut konvergiert: Es gilt
$$|x(j)\overline{y}(j)| \leq \frac{1}{2}(|x(j)|^2 + |y(j|)^2).$$
Damit folgt die Aussage.

Nun folgt leicht, dass $\langle \cdot, \cdot \rangle$ ein Skalarprodukt ist.

Ein nuetzliches Konzept zur weiteren Untersuchung (und zur Visualisierung von Konverenz) sind Kugeln und Umgebungen bzw. offene Mengen.

Definition. *(Kugel) Sei (X, d) metrischer Raum und $x \in X$. Dann heisst fuer $r \geq 0$*

$B_r(x) := \{y \in X : d(x, y) \leq r\}$ *abgeschlossene Kugel um x mit Radius r*

$U_r(x) := \{y \in X : d(x, y) < r\}$ *offene Kugel um x mit Radius r.*

Beispiele.

- Sei $\mathbb{R}$ mit der Metrik $d(x, y) = |x-y|$ versehen. Dann ist $x \in \mathbb{R}$ und $r > 0$ die Kugel $U_r(x) = (x - r, x + r)$. Entsprechend ist $B_r(x) = [x - r, x + r]$. **Zeichnung.**
- Sei $\mathbb{C}$ mit der Metrik $d(x, y) = |x - y|$ versehen. Dann ist fuer $z \in \mathbb{C}$ und $r > 0$ die Kugel $U_r(z)$ gegeben durch den (offenen) Kreis um z mit Radius r. Entsprechend ist $B_r(z)$ der volle Kreis um z mit Radius r. **Zeichnung.**
- Sei $X = \mathbb{R}^2$ mit der Euklidischen Metrik d_2 versehen. Dann ist fuer $x = 0$ und $r > 0$ die Kugel $U_r(x)$ ($B_r(x)$ gegeben durch den offenen (vollen) Kreis um 0 mit Radius r. (Vom Standpunkt des metrischen Raumes aus, sind also $\mathbb{R}^2$ und $\mathbb{C}$ nicht zu untercheiden.)
- Sei $X = \mathbb{R}^2$ mit $d = d_\infty$. Dann ist fuer $x = 0$ und $r > 0$ die offenen Kugel $U_r(x)$ gegeben durch...
- Sei $X = \mathbb{R}^2$ mit $d = d_1$. Dann ist fuer $x = 0$ und $r > 0$ die offenen Kugel $U_r(x)$ gegeben durch...
- Sei X eine beliebige Menge und $d = d_D$. Dann ist fuer $r = 1/2$ die offene Kugel...

Metriken erlauben es, den Begriff der Konvergenz von Folgen zu definieren.

DEFINITION. *(Konvergenz im metrischen Raum) Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in X heisst konvergent gegen $x \in X$, wenn gilt*

$$d(x_n, x) \to 0, n \to \infty.$$

Wir schreiben dann $x_n \to x$ oder $\lim_{n\to\infty} x_n = x$ und nennen x Grenzwert der Folge (x_n).

Bemerkung. Grenzwert einer Folge ist eindeutig. ($x_n \to x$, $x_n \to y$ impliziert $d(x, y) \leq d(x, x_n) + d(x_n, y) \to 0$...).

LEMMA. *Sei (X, d) ein metrischer Raum und (x_n) eine Folge in X und $x \in X$. Dann gilt $x_n \to x$ genau dann, wenn es zu jeder offenen nichtleeren Kugel K um x ein $n_K \in \mathbb{N}$ gibt mit $x_n \in K$ fuer alle $n \geq n_K$.*

Beweis. Eine offene nichtleere Kugel um x hat die Form $K = U_r(x)$ mit $r > 0$. Damit folgt die Aussage sofort. □

FOLGERUNG. *(Stetigkeit von Metrik und Norm)*

- *Sei (X, d) ein metrischer Raum. Gilt $x_n \to x$ und $y_n \to y$, so folgt $d(x_n, y_n) \to d(x, y)$.*
- *Sei V ein Vektorraum mit Norm $\|\cdot\|$ und zugehöriger Metrik d. Gilt $x_n \to x$ so folgt $\|x_n\| \to \|x\|$.*

Beweis. Das folgt jeweils aus der umgekehrten Dreiecksungleichung: Zum ersten Punkt:

$$\begin{aligned} |d(x_n, y_n) - d(x, y)| &= |d(x_n, y_n) - d(x_n, y) + d(x_n, y) - d(x, y)| \\ &\leq |d(x_n, y_n) - d(x_n, y)| + |d(x_n, y) - d(x, y)| \\ \text{(Umgek. } \triangle\text{-Ugl)} \quad &\leq d(y_n, y) + d(x_n, x). \end{aligned}$$

Zum zweiten Punkt: Nach der umgekehrten Dreiecksungleichung gilt

$$|\|x_n\| - \|x\|| \leq \|x_n - x\|.$$

Damit folgt die Aussage sofort. □

Wir untersuchen nun Konvergenz in den oben gegebenen Beispielen.
Beispiele.

Konvergenz in $\mathcal{A}$. Sei $\mathcal{A}$ eine beliebige Menge mit der diskreten Metrik d_D. Dann gilt $x_n \to x$ genau dann, wenn ein N existiert mit $x_n = x$ fuer alle $n \geq N$.

Konvergenz in $\mathcal{W}_\mathcal{A}$: Sei $\mathcal{W}_\mathcal{A}$ wie oben mit der Metrik

$$d(x, y) = \sum_{k=1}^{\infty} \frac{d_D(x(k), y(k))}{2^k}.$$

Dann gilt $x_n \to x$ genau dann, wenn fuer jedes $k \in \mathbb{N}$ gilt $x_n(k) \to x(k$ bzgl. d_D auf $\mathcal{A}$ (d.h. wenn fuer jedes $k \in \mathbb{N}$ schliesslich $x_n(k) = x(k)$ gilt).
Bew. $\Longrightarrow$: Sei $k \in \mathbb{N}$ beliebig. Offenbar gilt

$$d_D(x_n(k), x(k)) \leq 2^k d(x_n(k), x(k)).$$

Damit folgt die Aussage sofort.
$\Longleftarrow$: Betrachte $L \in \mathbb{N}$ beliebig. Dann gibt es ein n_L mit $x_n(k) = x(k)$ fuer alle $k = 1, \ldots, L$, falls $n \geq N_L$. Damit gilt dann fuer solche n

$$d(x_n, x) = \sum_{k=1}^{\infty} \frac{d_D(x_n(k), x(k)}{2^k} = \sum_{k=L+1}^{\infty} \frac{d_D(x_n(k), x(k)}{2^k} \leq \sum_{k=L+1} \frac{1}{2^k} = \frac{1}{2^L}.$$

Da L beliebig war, folgt die gewuenschte Konvergenz.

Konvergenz in $\mathbb{R}^N$ bzw. $\mathbb{C}^N$. Sei $x^{(n)} = (\xi_1^{(n)}, \ldots, \xi_N^{(n)})$ eine Folge in $\mathbb{R}^N$ und $x = (\xi_1, \ldots, \xi_N) \in \mathbb{R}^N$. Dann gilt:
$x^{(n)} \stackrel{d_1}{\to} x \Longleftrightarrow \xi_j^{(n)} \to \xi_j$ für jedes $j = 1, \ldots, N$. (Bew: ....)
$x^{(n)} \stackrel{d_\infty}{\to} x \Longleftrightarrow \xi_j^{(n)} \to \xi_j$ für jedes $j = 1, \ldots, N$. (Bew: ....)
$x^{(n)} \stackrel{d_2}{\to} x \Longleftrightarrow \xi_j^{(n)} \to \xi_j$ für jedes $j = 1, \ldots, N$. (Bew: ....)
Das ist kein Zufall. Es gilt vielmehr folgender bemerkenswerter Satz.

THEOREM. *(Äquivalenz aller Normen im $\mathbb{R}^N$ bzw. $\mathbb{C}^N$) Ist $\|\cdot\| : \mathbb{R}^N \longrightarrow [0, \infty)$ eine Norm, so gibt es $c, C \geq 0$ mit*

$$\|x\| \leq C\|x\|_1 \;\; und \;\; \|x\|_1 \leq c\|x\|$$

fuer alle $x \in \mathbb{R}^N$. Entsprechendes gilt in $\mathbb{C}^N$.

Ende der Vorlesung

Bemerkung. Es sind damit alle Normen im $\mathbb{R}^N$ aequivalent in dem Sinne, daß zu beliebigen Normen $\|\cdot\|_\clubsuit$ und $\|\cdot\|_\diamond$ Konstanten $\lambda, \mu > 0$ existieren mit

$$\|\dot{\|}_\clubsuit \leq \lambda \|\cdot\|_\diamond$$

und

$$\|\cdot\|_\diamond \leq \mu \|\cdot\|_\clubsuit.$$

Beweis. Wir zeigen zunächst die erste Ungleichung: Sei (e_j) die Standardbasis im $\mathbb{R}^N$. Dann gilt also

$$x = (\xi_1, \ldots, \xi_n) = \sum_{j=1}^{N} \xi_j e_j.$$

Dann folgt

$$\|x\| = \|\sum_{j=1}^{N} \xi_j e_j\| \leq \sum_{j=1}^{N} |\xi_j| \|e_j\| \leq C \sum_{j=1}^{N} |\xi_j| = C\|x\|_1$$

mit

$$C := \max\{\|e_j\| : j = 1, \ldots, N\}.$$

Wir zeigen nun die zweite Ungleichung

$$\|x\|_1 \leq c\|x\|.$$

Wir nehmen an, dass sie nicht gilt. Dann existiert also zu jedem $n \in \mathbb{N}$ ein x_n mit

$$\|x_n\|_1 > n\|x_n\|.$$

Ohne Einschraenkung gelte $\|x_n\|_1 = 1$ fuer alle $n \in \mathbb{N}$. Dann folgt also

$$\|x_n\| \leq \frac{1}{n}.$$

Damit folgt sofort

$$(*)\ \ x_n \to 0 \text{ bzgl. } \|\cdot\|.$$

Weiterhin sind wegen

$$1 = \|x_n\|_1 = \sum_{j=1}^{N} |\xi_j^{(n)}|$$

die Folgen $(\xi_j^{(n)})_n$ fuer jedes $j \in \{1, \ldots, N\}$ beschraenkt. Daher können wir nach Bolzano Weierstrass ohne Einschraenkung annehmen, dass x_n bzgl. $\|\cdot\|_1$ konvergiert gegen ein $x = (\xi_1, \ldots, \xi_N)$. (Genauer: $(\xi_1^{(n)})_n$ beschränkt, also gibt es nach BW eine konvergente Teilfolge. Ohne Einschraenkung sei die Folge $(\xi_1^{(n)})_n$ selber konvergent. Betrachte nun $(\xi_2^{(n)})_n$. Diese Folge ist beschraenkt. Also gibt es nach nach BW eine konvergente Teilfolge. Ohne Einschraenkung sei die Folge $(\xi_2^{(n)})_n$ selber konvergent....) Dann gilt

$$\|x\|_1 = \sum_{j=1}^{N} |\xi_j| = \lim \|x_n\|_1 = 1.$$

Insbesondere gilt $x \neq 0$. Weiterhin folgt aus $x_n \to x$ bzgl. $\|\cdot\|_1$ und der schon bewiesenen Ungleichung $\|\cdot\| \leq C\|\cdot\|_1$ aber auch $x_n \to x$ bzgl. $\|\cdot\|$. Wegen $x \neq 0$ ist dies ein Widerspruch zu $(*)$. □

Wir kommen nun noch zu einem fundamentalen Konzept, das spaeter von grosser Bedeutung sein wird.

DEFINITION (Cauchy-Folge). *Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in X heisst Cauchy-Folge, wenn zu jedem $\varepsilon > 0$ ein $n_\varepsilon > 0$ existiert mit*

$$d(x_n, x_m) \leq \varepsilon$$

fuer alle $n, m \geq n_\varepsilon$.

PROPOSITION. *Sei (X, d) ein metrischer Raum. Dann ist jede konvergente Folge eine Cauchy-Folge.*

Beweis. Sei die Folge (x_n) in X konverent gegen x. Es gilt

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m).$$

Damit folgt die Aussage leicht. $\square$

DEFINITION (Vollstaendigkeit). *Ein metrischer Raum heisst vollstaendig, wenn jede Cauchy Folge in ihm einen Grenzwert besitzt.*

Bemerkung. Vollstaendigkeit hängt von der Metrik ab und nicht nur von der Toplogie (siehe Uebung).

Beispiel - vollständiger Räume.

- Die Raum $\mathbb{R}$ mit der (euklidischen Metrik) $d(x, y) = |x - y|$ ist vollständig. Bew. Analysis I.
- Die Raum $\mathbb{C}$ mit der (euklidischen Metrik) $d(x, y) = |x - y|$ ist vollständig. Bew. Analysis I.
- Der Raum $\mathbb{R}^N$ mit der Euklidischen Metrik ist vollstaendig. Bew: Sei (x_n) eine Cauchy Folge bzgl. d_2 und $x_n = (\xi_1^{(n)}, \ldots, \xi_N^{(n)})$. Wegen

$$|\xi_j^{(n)} - \xi_j^{(m)}| \leq \left(\sum_{k=1}^{N} |\xi_k^{(n)} - \xi_k^{(m)}|^2 \right)^{1/2} = d_2(x_n, x_m)$$

ist dann $(\xi_j^{(n)})$ eine Cauchy Folge in $\mathbb{R}$ fuer jedes $j = 1, \ldots, N$. Damit existiert also nach dem im ersten Beispiel diskutieren fuer jedes $j = 1, \ldots, N$ ein $\xi_j \in \mathbb{R}$ mit $\xi_j^{(n)} \to \xi_j \in \mathbb{R}$ d.h.

$$|\xi_j^{(n)} - \xi_j| \to 0, n \to \infty.$$

Fuer $x = (\xi_1, \ldots, \xi_N)$ gilt dann also

$$d_2(x_n, x) = \left(\sum_{k=1}^{N} |\xi_k^{(n)} - \xi_k^{(m)}|^2 \right)^{1/2} \to 0, n \to \infty.$$

Damit konvergiert also (x_n) gegen x.

Wichtig. Tatsächlich haben auf dem $\mathbb{R}^N$ aufgrund des Satzes von der Aequivalenz aller Normen alle von Normen induzierten Metriken dieselben Cauchy-Folgen und dieselben konvergenten Folgen. Insbesondrere ist also der Raum $\mathbb{R}^N$ mit jeder von einer Norm induzierten Metrik vollstaendig.
- Der Raum $\mathbb{C}^N$ mit der Euklidischen Metrik ist vollständig. Bew. Wie fuer $\mathbb{R}^N$.
- Der Raum $M^{N \times N}$ der $N \times N$ Matrizen ueber $\mathbb{R}$ (ueber $\mathbb{C}$) ist vollstaendig bzgl. der durch die Norm (!)

$$\|A\| := \sup\{\|Ax\|_2 : \|x\|_2 \leq 1\}$$

induzierten Metrik.

Bew. Es handelt sich um den Vektorraum $\mathbb{R}^{N^2}$ auf dem alle Normen aequivalent sind. Es reicht also zu zeigen, dass $\|\cdot\|$ eine Norm ist. Das folgt aus den Betrachtungen der Übung.

- Sei X eine beliebige Menge und d_D die diskrete Metrik auf X. Dann ist (X, d_D) vollstaendig.

  Bew. Sei (x_n) eine Cauchy Folge. Dann gilt fuer alle genuegend grossen n und m $d_D(x_n, x_m) \leq 1/2$. Damit gilt fuer alle genuegend grossen n und m also $x_n = x_m$. Es gibt also ein $x \in X$ mit $x_n = x$ fuer alle genügend grossen n. Damit folgt die Konvergenz von (x_n) gegen x.
- Sei $\mathcal{A}$ endlich und $\mathcal{W}$ der Raum der Woerer ueber $\mathcal{A}$ mit der Metrik $d(x,y) = \sum_{j=1}^{\infty} \frac{d_D(x(j),y(j))}{2^j}$. Dann ist $\mathcal{W}$ vollstaedig.

  Bew. Sei (x_n) eine Cauchy-Folge. Wegen
  $$d_D(x_n(j), x_m(j)) \leq 2^j d(x_n, x_m)$$
  ist dann $(x_n(j))_n$ fuer jedes feste $j \in \mathbb{N}$ eine Cauchy-Folge bzgl. d_D. Also existiert nach dem schon bewiesenen zu jedem $j \in \mathbb{N}$ ein x_j mit
  $$x_n(j) \to x_j.$$
  Dann gilt also $x_n \to x$ fuer $x : \mathbb{N} \longrightarrow \mathcal{A}$, $j \mapsto x_j$.
- ℓ^2 ist vollstaendig. Bew. Nicht hier.

Beispiele - unvollständiger Räume

- Das Interval $(0,1)$ mit der Euklidischen Metrik ist nicht vollstaendig. (Uebung: Finden Sie eine Metrik d auf $(0,1)$, so dass $(0,1)$ vollstaendig bzgl d ist und eine bzgl. der Euklidischen Metrik in $(0,1)$ konvergente Folge auch bzgl d konvergiert und zwar gegen denselben Grenzwert.)
- Die Menge $\mathbb{Q}$ mit der Euklidischen Metrik ist nicht vollstaendig.

2. Etwas Topologie metrischer Räume

Neben dem Konzept der Kugel um einen Punkt gibt es auch noch das Konzept der Umgebung eines Punktes, um die Naehe zweier Punkte zu untersuchen. Der Vorteil der Umgebungen über Kugeln liegt darin, dass sie mit dem Anwenden von Funktionen besser verträglich sind. Es gibt nun Mengen, die Umgebung jedes ihrer Punkte sind. Solche Mengen spielen eine besondere Rolle. Das untersuchen wir nun.

Definition (Umgebung). *Sei (X,d) metrischer Raum und $x \in X$. Dann heisst eine Menge V Umgebung von x, wenn ein $r > 0$ existiert mit $U_r(x) \subset V$ (oder aequivalent, wenn ein $r' > 0$ existiert mit $B_{r'}(x) \subset V$.)*

Bemerkung. Es ist also V Umgebung von x, es eine Obermenge einer Kugel um x ist, d.h. wenn um x herum noch etwas 'Platz in V ist'. Die Elemente von V sind dann in gewisser Weise 'nahe' an x.

Beispiel. In einem metrischen Raum sind $U_r(x)$ und $B_r(x)$ Umgebungen von x fuer jedes $r > 0$.

Definition (Offene und abgeschlossene Mengen). *Sei (X, d) ein metrischer Raum. Dann heißt eine Teilmmenge U von X offen, wenn sie Umgebung jedes ihrer Punkte ist d.h. wenn zu jedem $x \in U$ ein $r_x > 0$ existiert mit $U_{r_x}(x) \subset U$. Eine Teilmenge A von X heisst abgeschlossen, wenn $X \setminus A$ offen ist.*

Bemerkungen.

- Es ist die Teilmenge Y des metrischen Raumes genau dann offen, wenn $X \setminus Y$ offen ist.
- Es ist NICHT so, daß jede Teilmenge eines metrischen Raumes genau eine der beiden Eigenschafen Offenheit bzw. Abgeschlossenheit haben muß.
  - Es kann eine Teilmenge offen und abgeschlossen sein.
  - Es kann eine Teilmenge weder offen noch abgeschlossen sein.

Ende der Vorlesung

Beispiele. Jede offene Kugel ist offen. Jede abgeschlossene Kugel ist abgeschlossen.
Bew. Zur Offenheit von $U_r(x)$: Diese folgt aus der Dreiecksungleichung: Ist $y \in U_r(x)$, so gilt $d(y, x) < r$. Damit folgt nach Dreiecksungleichung / Zeichnung mit $s := r - d(x, y) > 0$ also

$$U_s(y) \subset U_r(x).$$

(Denn fuer $v \in U_s(y)$ gilt $d(v, x) \leq d(v, y) + d(y, x) < s + d(y, x) = r$.)

Zur Abgeschlossenheit von $B_r(x)$: Es ist die Offenheit des Komplementes zu zeigen. Diese folgt wieder aus der Dreiecksungleichung: Sei $y \notin B_r(x)$. Dann gilt also $d(x, y) > r$. Damit gilt nach Dreiecksungleichung / Zeichnung fuer $s := d(x, y) - r$ also

$$U_s(y) \subset X \setminus B_r(x).$$

(Denn fuer $v \in U_s(y)$ gilt $d(v, x) \geq d(y, x) - d(v, y) > d(y, x) - s = r$.)
Da $y \in X \setminus B_r(x)$ beliebig war folgt die Behauptung.

Lemma. *(Charakterisierung Konvergenz) (X, d) metrischer Raum, (x_n) Folge in X, $x \in X$. Dann sind aequivalent:*

(i) *Es konvergiert (x_n) gegen x (d.h. $\lim_{n\to\infty} d(x_n, x) = 0$).*
(ii) *Fuer alle $\varepsilon > 0$ existiert ein $n_\varepsilon \in \mathbb{N}$ mit $x_n \in U_\varepsilon(x)$ fuer alle $n \geq n_\varepsilon$*

(iii) *Fuer jede Umgebung* U *von* x *existiert ein* $n_U \in \mathbb{N}$ *mit* $x_n \in U$ *fuer alle* $n \geq n_U$.

(iv) *Fuer jede offene Menge* V *mit* $x \in V$ *existiert ein* $n_V \in \mathbb{N}$ *mit* $x_n \in V$ *fuer alle* $n \geq n_U$.

Bemerkung.

- Alle vier Aussagen geben eine praezise Version davon, dass die (x_n) dem Punkt x beliebig nahe kommen, wenn n beliebig gross wird.
- Natuerlich koennte man in (ii) auch U_ε durch B_ε ersetzen, da $U_\varepsilon \subset B_\varepsilon$ und $B_{\varepsilon/2} \subset U_\varepsilon$ gilt.

Beweis. (iv)$\Longrightarrow$ (iii): Jede Umgebung von x enthaelt eine offene Menge V mit $x \in V$.

(iii)$\Longrightarrow$ (ii): Es ist $U_\varepsilon(x)$ eine Umgebung von x.

(ii)$\Longrightarrow$ (i): Sei $\varepsilon > 0$. Zu zeigen: $d(x_n, x) < \varepsilon$ fuer alle genuegend grossen n. Das ist klar nach Definition von $d \to 0$ und $U_\varepsilon(x)$.

(i)$\Longleftrightarrow$ (iv). Sei V eine offene Menge mit $x \in V$. Dann existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset V$. Wegen (i) gibt es ein $n_\varepsilon > 0$ mit $d(x_n, x) < \varepsilon$ fuer $n \geq n_\varepsilon$. Dann gilt fuer solche n also

$$x_n \in U_\varepsilon(x) \subset V.$$

Das beendet den Beweis. □

Definition. *Sei* (X, d) *ein metrischer Raum. Dann heisst die Familie aller offenen Mengen auf* X *die durch* d *induzierte Topologie und wird mit* $\mathcal{T} = \mathcal{T}_d$ *bezeichnet.*

Das System aller offenen Mengen in einem metrischen Raum hat einige bemerkenswerte Eigenschaften.

Proposition. *Sei* (X, d) *ein metrischer Raum mit von* d *erzeugter Topologie* $\mathcal{T}$. *Dann gilt:*

(T1) *Es gehoeren* $\emptyset$ *und* X *zu* $\mathcal{T}$.

(T2) *Gehören* $V_1, \ldots, V_n$ *zu* $\mathcal{T}$, *so auch* $V := \bigcap_{j=1}^n V_j$. *'Endlische Schnitte offener Mengen sind offen.'*

(T3) *Ist* I *eine Menge und gehoeren* V_α, $\alpha \in I$, *zu* $\mathcal{T}$, *so auch* $\bigcup_{\alpha \in I} V_\alpha$. *'Beliebige Vereinigungen offener Menge sind offen'.*

Beweis. Es sind $(T1)$ und $(T3)$ einfach nachzuweisen. Nun zu $(T2)$: Seien $V_1, \ldots, V_n \in \mathcal{T}$ und $V := \cap_{j=1}^n V_j$. Fuer $x \in V$ gilt $x \in V_j$ fuer alle j. Wegen $V_j \in \mathcal{T}$ existieren also $r_j > 0$ mit $U_{r_j}(x) \subset V_j$. Sei

$$r := \min\{r_j : j = 1, \ldots, n\} > 0.$$

Dann gilt

$$U_r(x) \subset U_{r_j}(x) \subset V_j$$

fuer alle $j = 1, \dots, n$ und daher

$$U_r(x) \subset V.$$

Das beweist $(T2)$. □

Bemerkung.

- Ist X eine beliebige Menge und $\mathcal{T}$ ein Familie von Teilmengen von X, die die Eigenschaften (T1), (T2)und (T3) hat, so heisst $\mathcal{T}$ eine Topologie und $(X, \mathcal{T})$ ein topologischer Raum.
- Auf topologischen Räumen kann man Konvergenz, Umgebung, Stetigkeit etc definieren. Im allgemeinen wird man aber nicht mehr alles auf Folgen zurueckführen können, sondern wird sogenannte Netze brauchen. Alle in dieser Vorlesung gegebenen Formulierungen mit Folgen nutzen also die spezielle Struktur eines metrischen Raumes. Die Formulierungen mit offenen Mengen und Umgebungen sind auch im allgemeinen gueltig. Daher rührt auch ihre Relevanz.
- Im allgemeinen ist der Schnitt von abzaehlbar vielen offenen Mengen nicht mehr offen. Bsp. Sei $U_n = U_{1/n}(0)$ in $\mathbb{R}$ mit der Euklidischen Metrik. Dann gilt $\cap U_n = \{0\}$.

Folgerung. *Ist (X, d) ein metrischer Raum, so sind beliebige Schnitte von abgeschlossenen Mengen abgeschlossen.*

Beweis. Das folgt durch Komplementbildung □

Folgerung. *Sei (X, d) ein metrischer Raum und $A \subset X$ abgeschlossen und $U \subset X$ offen. Dann ist $U \setminus A$ offen und $A \setminus U$ abgeschlossen.*

Beweis. Es gilt $U \setminus A = U \cap (X \setminus A)$ und $A \setminus U = A \cap (X \setminus U)$. Damit folgen die Aussagen dann aus dem schon bewiesenen. □

Eine Teilmenge eines metrischen Raumes ist im allgemeinen weder offen noch abgeschlossen. Man kann aus einer Menge aber jeweils abgeschlossene und offene Menge 'machen'.

Definition. *(X, d) ein metrischer Raum. Dann heisst:*

$$M^\circ := \bigcup_{U \subset M, U \text{ offen}} U$$

das Innere oder der offene Kern von M und

$$\overline{M} := \bigcap_{M \subset A,\ A \text{ abgeschlossen}} A$$

der Abschluss oder die abgeschlossene Huelle von M und

$$\partial M := \overline{M} \setminus M^\circ$$

der Rand von M.

Bemerkung.

- Es ist M° die groesste offene Teilmenge, die in M enthalten ist (da Vereinigung von offenen Mengen offen sind.) Insbesondere stimmt jede offene Menge mit ihrerem Inneren ueberein.
- Es ist $\overline{M}$ die kleinste abgeschlossene Menge, die M enthaelt. Insbesondere stimmt jede abgeschlossene Menge mit ihrer Huelle ueberein.
- Es gilt (siehe Uebung):

$$\begin{aligned} M^\circ &= \{x \in X : M \text{ ist Umgebung von } x\} \\ &= \{x \in X : \text{es existiert } r_x > 0 \text{ mit } U_{r_x}(x) \subset M\} \end{aligned}$$

$$\begin{aligned} \overline{M} &= \{x \in X : \text{jede Umgebung von } x \text{ enthaelt Punkt von } M\} \\ &= \{x \in X : \text{ es gibt } (x_n) \text{ in } M \text{ mit } x_n \to x\}. \end{aligned}$$

$$\partial M = \{x \in X : \text{fuer jede Umgebung } U \text{ von } x \text{ gilt } U \cap M \neq \emptyset \text{ und } U \cap X \setminus M \neq \emptyset\}.$$

Beispiele.

- Betrachte $\mathbb{K}^m$ mit der Euklidischen Metrik und der induzieren Topologie. Dann ist das Innere von $B_r(x)$ gerade $U_r(x)$ und der Abschluss von $U_r(x)$ ist $B_r(x)$. (Uebung.)
- Fuer jeden metrischen Raum gilt, dass $U_r(x)$ enthalten ist im Inneren von $B_r(x)$ (da $U_r(x)$ offen) und, dass der Abschluss von $U_r(x)$ enthalten ist in $B_r(x)$ (da $B_r(x)$ abgeschlossen). Diese Inklusionen koennen strikt sein! Beispiele:
  - X = abgeschlossene Einheitskugel in $\mathbb{R}^2$. Dann ist $X = B_1(0)$ offen, also sein Inneres.
  - X = abgeschlossene Einheitskugel in $\mathbb{R}^2$ ohne eine punktierte Umgebung von $(1,0)$. Dann enthaelt der Abschluss von $U_1(0)$ nicht den Punkt $(1,0)$. Er gehoert aber zu $B_1(0)$.
- Sei $\mathbb{R}$ mit der durch die Euklidische Metrik erzeugten Topologie versehen. Dann ist der Rand von $\mathbb{Q}$ gerade $\partial\mathbb{Q} = \overline{\mathbb{Q}} \setminus \mathbb{Q}^\circ = \mathbb{R} \setminus \emptyset = \mathbb{R}$ (Vgl. Uebung).

Zum Abschluss noch eine praktische Charakterisierung von Abgeschlossenheit.

PROPOSITION. *Sei (X, d) ein metrischer Raum. Eine Teilmenge A von X ist genau dann abgeschlossen, wenn der Grenzwert jeder (in X) konvergenten Folge aus A wieder in A liegt.*

Ende der Vorlesung

Beweis. Sei A abgeschlossen und (x_n) eine Folge in A mit $x_n \to x$. Zu zeigen: $x \in A$. Waere x nicht in A, so gehoerte es zu dem offenen

$X \setminus A$. Damit waere also x_n fuer alle groessen n ebenfalls in $X \setminus A$. Widerspruch.

Es habe A die angegebene Eigenschaft. Wir muessen zeigen, daß $X \setminus A$ offen ist. Sei $x \in X \setminus A$ beliebig. Zu zeigen: ein $\delta > 0$ existiert mit $U_\delta(x) \cap A = \emptyset$. Angenommen: Fuer jedes $\delta > 0$ gilt $U_\delta(x) \cap A \neq \emptyset$. Dann existiert also ein Folge (x_n) mit $x_n \in U_{1/n}(x) \cap A$. Damit ist (x_n) eine Folge in A mit $x_n \to x$. Aus der angegebenen Eigenschaft folgt $x \in A$. Das ist ein Widerspruch. □

3. Konvergenz und Stetigkeit

Mithilfe des Begriffes der Konvergenz koennen wir nun die Stetigkeit von Abbildungen fassen. Darum geht es in diesem Abschnitt.

Definition. *Seien (X_1, d_1), (X_2, d_2) metrische Raeume und $f : X_1 \longrightarrow X_2$ eine Abbildung. Es heisst f stetig in $x \in X_1$, wenn fuer jede Folge (x_n) in X_1 mit $x_n \to x$ bzgl. d_1 gilt $f(x_n) \to f(x)$ bzgl. d_2.*

Kurzfassung: Es ist f stetig in x genau dann, wenn gilt: $x_n \to x \Longrightarrow f(x_n) \to f(x)$.

Lemma. *Seien (X_1, d_1), (X_2, d_2) metrische Raeume und $f : X_1 \longrightarrow X_2$ eine Abbildung und $x \in X_1$. Dann sind aquivalent:* ***Zeichung.***

(i) *Fuer jede Folge (x_n) in X_1 mit $x_n \to x$ gilt $f(x_n) \to f(x)$.*
(ii) *Fuer alle $\varepsilon > 0$ existiert ein $\delta > 0$ mit $f(U_\delta(x)) \subset U_\varepsilon(f(x))$. (Beachte: Es handelt sich um Kugeln in verschiedenen Räumen!)*
(iii) *Fuer jede Umgebung V von $f(x)$ existiert eine Umgebung U von x mit $f(U) \subset V$.*
(iv) *Fuer jede Umgebung V von $f(x)$ ist auch $f^{-1}(V)$ eine Umgebung von x. (Erinnerung: Das Urbild $f^{-1}(V)$ von V unter f ist gegeben als $f^{-1}(V) = \{x \in X_1 : f(x) \in V\}$)*

Bemerkung.

- Alle vier Aussagen geben eine praezise Version davon, dass Punke nahe an x auf Punkte nahe an $f(x)$ abgebildet werden.
- In (ii) kann man natuerlich statt mit offenen Kugeln auch mit abgeschlossenen Kugeln argumentieren.
- Die Formulierung in (iii) und (iv) nehmen keinen direkten Bezug auf eine Metrik.
- Die Formulierung (iv) zeigt die Nuetzlichkeit des Umgebungsbegriffes im Vergleich mit dem Begriff der Kugel: Im allgemeinen werden die (Ur)bilder von Kugeln auch fuer 'schoene' Funktionen keine Kugeln sein. So sind schon bei linearen Funktionen die Urbilider von Kugeln im allgemeinen Ellipsen, die keine Kugeln sind.

Beweis. (iv)$\Longrightarrow$ (iii): Man kann $U = f^{-1}(V)$ setzen.
(iii)$\Longrightarrow$ (ii): Es ist $U_\varepsilon(f(x))$ eine Umgebung von $f(x)$. Daher existiert nach (iii) eine Umgebung U von x mit $f(U) \subset B_\varepsilon(f(x))$. Da U Umgebung von x ist, existiert also ein $\delta > 0$ mit $U_\delta(x) \subset U$. Damit gilt also
$$f(U_\delta(x)) \subset f(U) \subset U_\varepsilon(f(x)).$$
(ii)$\Longrightarrow$ (i): Sei $\varepsilon > 0$. (Z.z. $d_2(f(x_n), f(x)) \leq \varepsilon$ fuer groesse n.) Sei zu diesem ε ein $\delta > 0$ gemaess (ii) gewaehlt. Wegen $x_n \to x$ existiert ein $n_\delta \in \mathbb{N}$ mit $d_1(x_n, x) \leq \delta$ fuer alle $n \geq n_\delta$. Damit gilt dann nach (ii) $d_2(f(x_n), f(x)) \leq \varepsilon$ fuer $n \geq n_\delta$.
(i) $\Longrightarrow$ (iv): Sei V eine Umgebung von $f(x)$. Angenommen: Es ist $f^{-1}(V)$ keine Umgebung von x. Dann gilt also fuer kein $\delta > 0$ die Inklusion $U_\delta(x) \subset f^{-1}(V)$. Daher gibt es also (mit $\delta = 1/n$) zu jedem $n \in \mathbb{N}$ eine $x_n \in U_{1/n}(x)$ mit $f(x_n) \notin V$. Dann gilt aber $x_n \to x$ und $f(x_n)$ konvergiert nicht gegen $f(x)$. Das ist ein Widerspruch. □

Stetigkeit ist verträglich mit den 'üblichen' Operationen.

Proposition. *Seien (X_j, d_j), $j = 1, 2, 3$, metrische Räume und $f : X_1 \longrightarrow X_2$, $g : X_2 \longrightarrow X_3$ gegeben. Ist f stetig in $x \in X_1$ und g stetig in $f(x)$, so ist $g \circ f : X_1 \longrightarrow X_3$ stetig in x.*

Beweis. Das folgt aus der Folgencharakterisierung der Stetigkeit wie folgt: $x_n \to x$ impliziert, $f(x_n) \to f(x)$ (da f stetig in x) und das impliziert $g(f(x_n)) \to g(f(x))$ (da g stetig in $f(x)$).

□

Proposition. *Sei (X, d) ein metrischer Raum und $f, g : X \longrightarrow \mathbb{C}$ seien stetig in $x \in X$. Dann sind auch $f + g : X \longrightarrow \mathbb{C}$ und $fg : X \longrightarrow \mathbb{C}$ stetig in x*

Beweis. Das folgt sofort aus der Folgencharakterisierung der Stetigkeit □

Proposition. *Ist (X, d) ein metrischer Raum und sind $f, g : X \longrightarrow \mathbb{R}$ stetig (in $p \in X$), so sind auch* $\max\{f, g\}$, $\min\{f, g\}$ *und* $|f|$ *stetig (in $p \in X$).*

Beweis. (Uebung) □

Definition. *Seien (X_1, d_1), (X_2, d_2) metrische Raeume und $f : X_1 \longrightarrow X_2$ eine Abbildung. Dann heisst f stetig, wenn es in jedem Punkt von X_1 stetig ist.*

Theorem. *(Charakterisierung Stetigkeit) Seien (X_1, d_1), (X_2, d_2) metrische Räume und $f : X_1 \longrightarrow X_2$ eine Abbildung. Dann sind äquivalent:*

(i) *f ist stetig.*
(ii) *Die Urbilder offener Mengen unter f sind offen, d.h. fuer jedes $V \in \mathcal{T}_2$ gehoert $f^{-1}(V)$ zu $\mathcal{T}_1$.*

Bemerkung. Dieser Satz allein rechtfertigt schon die Einführung des Konzeptes der offenen Menge.

Beweis. (i)$\Longrightarrow$ (ii): Sei $V \in \mathcal{T}_2$ beliebig und $U := f^{-1}(V)$. Zu zeigen: U ist Umgebung von jedem $x \in U$. Sei also $x \in U$ beliebig. Dann folgt $f(x) \in V$. Da V offen ist, ist es eine Umgebung von $f(x)$ und damit folgt nach (i) die Existenz einer Umgebung U_x von x mit

$$f(U_x) \subset V.$$

Damit folgt also $x \in U_x \subset U$ (nach Definition von U). Damit ist U Umgebung von x.
(ii) $\Longrightarrow$ (i): Zu zeigen: f ist stetig in jedem $x \in X_1$. Sei also $x \in X_1$ beliebig. Sei V eine Umgebung von $f(x)$. Dann existiert also eine offene Menge W mit

$$f(x) \in W \subset V.$$

Nach (ii) ist dann $U := f^{-1}(W)$ offen. Weiterhin gilt nach Definition $x \in U$. Damit ist also U eine Umgebung von x mit $f(U) \subset W \subset V$. $\square$

Beispiel. Alle lineaeren Funktionen $A : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ sind stetig.
Bew. Uebung.

Zum Abschluss des Abschnittes disktuieren wir noch eine Anwendung, naemlich **Stetigkeit und Produkträume.** Seien (X_j, d_j), $j = 1, \ldots, N$ metrische Räume und sei

$$X := X_1 \times \ldots \times X_N = \{(\xi_1, \ldots, \xi_N) : \xi_j \in X_j, j = 1, \ldots, N\}$$

ihr Produkt. Ähnlich wie man $\mathbb{R}^N$ mittels der Euklidischen Metrik von $\mathbb{R}$ metrisiert, kann man dann auch den Raum X metrisieren. So ist zum Beispiel

$$d(\xi_1, \ldots, \xi_N), (\eta_1, \ldots, \eta_N)) := \sum_{j=1}^{N} d_j(\xi_j, \eta_j)$$

eine Metrik auf X. Es gilt dann offenbar

$$x_n = (\xi_1^{(n)}, \ldots, \xi_N^{(n)}) \to x = (\xi_1, \ldots, \xi_N)$$

bezueglich d genau dann, wenn $\xi_j^{(n)} \to \xi_j$ bzgl d_j fuer jedes $j = 1, \ldots, N$.
Offenbar gilt fuer d die Gleichheit

$$d((\xi_1, \ldots, \xi_N), (\eta_1, \ldots, \eta_N)) = \|((d_1(\xi_1, \eta_1), \ldots, d_N(\xi_N, \eta_N))\|_1.$$

Allgemeiner kann man zu jeder Norm $\| \cdot \|$ auf $\mathbb{R}^N$ eine Metrik $e_{\|\cdot\|}$ auf X einführen durch

$$e_{\|\cdot\|}(\xi_1, \ldots, \xi_N), (\eta_1, \ldots, \eta_N)) = \|((d_1(\xi_1, \eta_1), \ldots, d_N(\xi_N, \eta_N))\|.$$

Aufgrund der Aquivalenz aller Normen im $\mathbb{R}^N$, gibt es dann fuer jede Norm $\|\cdot\|$ Konstanten $c, C > 0$ mit

$$cd \le e_{\|\cdot\|} \le Cd.$$

Darum spielt es fuer die Untersuchungen von konvergenter Folgen bzw Cauchy-Folgen keine Rolle, welche dieser Metrik wir tatsaechlich verwenden.

Es ist (Uebung) (X, d) genau dann vollständig, wenn (X_j, d_j) vollstaendig ist fuer jedes $j = 1, \dots, N$.

Fuer Stetigkeit von Funktionen erhalten wir im Zusammenhang mit Produkträumen folgendes.

Sei (Z, e) ein metrischer Raum und (X, d) der obige Produktraum. Dann ist $f : X \longrightarrow Z$ stetig genau dann, wenn aus $\xi_j^{(n)} \to \xi_j$, $j = 1, \dots, N$ folgt $f((\xi_1^{(n)}, \dots, \xi_N^{(n)})) \to f(\xi_1, \dots \xi_N))$.

Es ist $f : Z \longrightarrow X, f(z) = (f_1(z), \dots, f_N(z))$ genau dann stetig, wenn jede einzelne Komponente $f_j : Z \longrightarrow X_j$ stetig ist.

Beispiele.

- Ist (Z, e) ein beliebiger metrischer Raum, so ist $e : Z \times Z \longrightarrow \mathbb{R}$ stetig. (Bew. Die entsprechende Aussage fuer Folgenkonvergenz wurde oben schon gezeigt.)
- Es sind $+ : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ und $\cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ stetig. Entsprechendes gilt auf $\mathbb{R}$. (Bew. IN der Folgencharakterisierung ist das schon bekannt).

4. Stetigkeit und Grenzwerte - Bonusmaterial

Seien (X, d) und (Y, e) metrische Raeume und $f : X \longrightarrow Y$ eine Funktion. Sei $p \in X$ und $q \in Y$. Dann hat f bei p den Grenzwert q geschrieben als

$$\lim_{x \to p} f(x) = q$$

genau dann, wenn gilt

- $f(x_n) \to q$ fuer jede Folge (x_n) in X mit $x_n \to p$.

bzw. aequivalent, wenn gilt

- zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $e(f(x), q) \le \varepsilon$ fuer alle $x \in U$ mit $d(x, p) \le \delta$.

Offenbar ist dann also f stetig in p, genau dann wenn der Grenzwert von f in p existiert.

Unter Umstaenden hat man es mit Funktionen zu tun, die in p nicht definiert sind. Dann hat f auf in p den Grenzwert q, geschrieben $\lim_{x \to p, x \neq p} f(x) = q$, falls

- $f(x_n) \to q$ fuer jede Folge (x_n) in $X \setminus \{p\}$ mit $x_n \to p$.

bzw. aequivalent, wenn

- zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $e(f(x), q) \leq \varepsilon$ fuer alle $x \in X \setminus \{p\}$ mit $d(x, p) \leq \delta$.

5. Kompaktheit

In diesem Abschnitt lernen wir ein fundamentales Konzept der Topologie kennen naemlich das Konzept der Kompaktheit. Kompaktheit ist eine ausgesprochen nützliche Eigenschaft (verlangt aber eine etwas abstrakte Definition).

DEFINITION. *Sei (X, d) ein metrischer Raum. Eine Teilmenge $K \subset M$ heisst kompakt, wenn jede Folge (x_n) in K ein Teilfolge (x_{n_k}) bezitzt, die gegen einen Grenzwert aus K konvergiert.*

Bemerkung.

- Das so definierte Konzept ist auch als Folgenkompaktheit bekannt. In metrischen Räumen stimmt es mit anderen Konzepten von Kompaktheit ueberein (s.u.). Da uns nur metrische Räume interessieren, werden wir nicht zwischen Folgenkompaktheit und Kompaktheit unterscheiden.
- Der Fall $K = X$ in obiger Definition ist natuerlich möglich. In 'Anwendungen' hat man es aber oft mit der Situation zu tun, dass X nicht kompakt ist, man sich aber fuer die kompakten Mengen in X interessiert.

Beispiele.

- Sei $I = [a, b]$ ein abgeschlossenes beschränktes Intevall in $\mathbb{R}$. Dann ist I kompakt. (Bew. Nach Bolzano/Weierstrass hat jede Folge in I eine (in $\mathbb{R}$) konvergente Teilfolge. Da das Intervall abgeschlossen ist, liegt der Grenzwert der Teilfolge in I).
- Sei A eine abgeschlossene bzgl. der Euklidischen Metrik beschraenkte Menge in $\mathbb{R}^N$ (oder $\mathbb{C}^N$). Dann ist A kompakt. Das werden wir unten genauer diskutieren. Hier weisen wir schon darauf hin, daß wir das eigentlich kuerzlich mitbewiesen haben, als es um die Aequivalenz aller Normen im $\mathbb{R}^N$ ging.

Ende der Vorlesung

Kompaktheit ist eine nuetzliche Eigenschaft, wie die folgenden Resultate zeigen. Diese Resultate haben wir fuer stetige Funktionen auf abgeschlossenen beschraenkten Intervallen in $\mathbb{R}$ schon untersucht. Die dort gegebenen Beweise uebertraegen sich wortwoertlich. Weil sie so schoen sind, geben wir sie hier noch einmal.

THEOREM. *Sei (K, d) kompakter metrischer Raum und $f : K \longrightarrow \mathbb{R}$ stetig. Dann nimmt f Minimum und Maximum an.*

Beweis. Wir behandeln nur das Maximum. (Das Minimum kann analog behandelt werden.) Waehle eine Folge (x_n) in K mit

$$\lim_n f(x_n) = \sup_{x \in X} f(x).$$

(Hier ist apriori auch die bestimmte Divergenz gegen ∞ moeglich.) Da K kompakt ist, hat (x_n) eine in K konvergente Teilfolge, d. h. es gibt $(x_{n_k})_k$ und $x \in K$ mit

$$x_{n_k} \to x.$$

Da f stetig ist folgt dann

$$f(x) = \lim_k f(x_{n_k}) = \sup_{x \in K} f(x).$$

Das beendet den Beweis. □

THEOREM. *Sei (K, d) ein kompakter metrischer Raum und (Y, e) ein metrischer Raum und $f : K \longrightarrow Y$ stetig. Dann ist f gleichmaessig stetig.*

Beweis. Sei $\varepsilon > 0$. Zu zeigen: Es gibt ein $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon$ fuer alle $x, y \in K$ mit $d(x, y) \leq \delta$. Angenommen Nein! Dann existieren also zu jedem $n \in \mathbb{N}$ Elemente $x_n, y_n \in K$ mit $d(x_n, y_n) \leq \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \varepsilon$. Es hat dann (x_n) eine in K konvergente Teilfolge. Ohne Einschraenkung sei x_n konvergent gegen $x \in K$. Dann konvergiert auch (y_n) gegen x. Damit folgt dann

$$0 = |f(x) - f(x)| = \lim_{n \to \infty} |f(x_n) - f(y_n)| \geq \varepsilon.$$

Das ist ein Widerspruch. □

THEOREM. *Sei (K, d) ein kompakter metrischer Raum und (Y, e) ein metrischer Raum und $f : K \longrightarrow Y$ stetig. Dann ist $f(K)$ komapkt.*

Bemerkung.

- Das ist einer der beiden einzigen globalen Saetze ueber stetige Funktionen.
- Anders als bei den Betrachtungen zur Stetigkeit, geht es hier um die Bilder (und nicht die Urbilder) von f.

PROPOSITION. *Sei (X, d) metrischer Raum. Ist $K \subset X$ kompakt und $A \subset K$ abgeschlossen, so ist A kompakt.*

Beweis. Sei (x_n) eine Folge in A. Dann ist (x_n) eine Folge in K. Daher hat also (x_n) eine in K konvergente Teilfolge. Da A abgeschlossen ist, liegt der Grenzwert der Teilfolge sogar in A. □

PROPOSITION. *Sei (X, d) metrischer Raum und $K \subset X$ kompakt. Dann ist K abgeschlossen.*

Beweis. Sei (x_n) eine Folge in K, die gegen $x \in X$ konvergiert. Aufgrund der Kompaktheit enthaelt (x_n) eine Teilfolge, die gegen ein $y \in K$ konvergiert. Dann muss aber gelten $x = y \in K$. $\square$

Bemerkung. Als Konsequenz der beiden vorangehenden Propositionen ist der Schnitt einer kompakten Menge mit einer abgeschlossenen Menge wieder kompakt.

THEOREM. *(Automatische Stetitgkeit der Inversen) Sei (K, d) ein kompakter metrischer Raum und (L, e) ein metrischer Raum und $f : K \longrightarrow L$ sei bijektiv und stetig. Dann ist die inverse Abbildung $g = f^{-1} : L \longrightarrow K$ stetig.*

Beweis. Zu zeigen $g^{-1}(U)$ offen fuer jedes offene $U \subset K$. Es reicht zu zeigen: $g^{-1}(A)$ abgeschlossen fuer jedes abgeschlossene $A \subset K$. Ist $A \subset K$ abgeschlossen, so ist A kompakt. Daher ist aufgrund der Stetigkeit von f auch

$$g^{-1}(A) = f(A)$$

kompakt, also nach voriger Proposition abgeschlossen. $\square$

Wir kommen nun zur konzeptuellen Untersuchung und Charakterisierung von Kompaktheit.

DEFINITION. *Sei (X, d) ein metrischer Raum. Eine Teilmenge $B \subset X$ (z.B. $B = X$) heisst total beschraenkt, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ und $x_1, \ldots, x_N \in X$ (aequivalent $x_1, \ldots, x_N \in B$) existieren mit*

$$B \subset \bigcup_{k=1}^{N} B_\varepsilon(x_k).$$

Bemerkung.

- Statt mit abgeschlossenen Kugeln koennte man auch offenen Kugeln arbeiten.
- Wir wissen schon, dass jede Metrik d zu einer Metrik e aeqivalent ist, in der der ganze Raum eine beschraenkte Menge ist (Z.B. $e = \frac{d}{1+d}$). Daher ist Beschraenktheit einer Menge keine besonders kanonische Eigenschaft. Das Konzept der total Beschraenktheit ersetzt das Konzept der Beschraenktheit. Ist ein Raum bgzl. zweier aequivalenter Metriken vollstaendig, so haben beide Metriken dieselben total beschraenkten Mengen (s.u.).

Beispiel - Total beschraenkte Mengen im $\mathbb{R}^N$ bzw. $\mathbb{C}^N$. Sei $\mathbb{R}^N$ mit der Euklidischen Metrik versehen. Dann ist ein $B \subset \mathbb{R}^N$ genau dann total beschraenkt, wenn es beschraenkt ist.
Bew. $\Longrightarrow$: Sei B beschraenkt. Betrachte zu $\varrho > 0$ das Gitter $\Gamma_\varrho := (\varrho\mathbb{Z})^N$. Dann schneidet B aufgrund der Beschraenktheit (und des archimedischen Axioms) nur endllich viele Maschen des Gitters. Jede

dieser Maschen kann mit einer Kugel vom Radius $N\rho$ ueberdeckt werden. Damit folgt die gewuenschte Implikation (da $\varrho > 0$ beliebig klein gemacht werden kann).

$\Longleftarrow$: Das ist klar.

Bemerkung. Eine total beschraenkte Menge muss nicht vollstaendig sein. Bsp: $(0,1)$ mit der ueblichen Euklidischen Metrik.

Fuer die folgenden Betrachtungen werden noch zwei Stuecke Notation von Nutzen sein.

Notation. Sei (X,d) ein metrischer Raum.

- Ist (x_n) eine Folge in X und B eine Teilmenge von X, so sagen wir, dass in B unendlich viele Folgeglieder liegen, wenn die Menge $\{n \in \mathbb{N} : x_n \in M\}$ unendlich viele Elemente hat.
- Ist B eine Teilmenge von X und sind U_α, $\alpha \in I$ Teilmengen von X so sagen wir, dass die U_α die Menge B ueberdecken, wenn $B \subset \cup_{\alpha\in I} U_\alpha$ gilt.

Wir erinnern an den Satz von Bolzano / Weierstrass. Dieser besagt, dass eine Menge I in $\mathbb{R}$ genau dann beschraenkt ist, wenn jede Folge eine konvergente Teilfolge hat. (Hier ist die schwere Richtung die eigentliche Aussage des Satzes von Bolzano / Weierstrass und die andere Richtung ist klar). Die total beschraenkten Menge lassen sich mit folgender Variante des Satzes von Bolzano/Weierstrass charakterisieren.

Lemma. *(Charakterisierung total beschraenkt). Sei (M,d) ein metrischer Raum und $B \subset M$. Dann sind aequivalent:*

(i) *Es ist B total beschraenkt.*
(ii) *Jede Folge in B hat eine Teilfolge, die eine Cauchy Folge ist.*

Beweis. (i)$\Longrightarrow$ (ii): Sei (x_n) eine Folge in B. Wir ueberdecken nun B durch endlich viele Kugeln vom Radius 1. In einer von diesen Kugeln muessen dann uendlich viele Folgeglieder von (x_n) liegen.
Bezeichne den Schnitt dieser Kugel mit B als B_1. Wir ueberdecken nun B_1 (das als Teilmenge des total beschränkten B ebenfalls total beschraenkt ist) durch endlich viele Kugeln vom Radius $1/2$. Der Schnitt einer dieser Kugeln mit B_1 muss dann unendlich viele Folgeglieder von (x_n) enthalten. Nenne diesen Schnitt B_2. Induktiv koennen wir so eine Folge von Mengen B_n, $n \in \mathbb{N}$, konstruieren mit

- $B \supset B_1 \supset \ldots \supset B_{n-1} \supset B_n$
- $\sup_{x,y\in B_n} d(x,y) =: \mathrm{diam} B_n \leq \frac{1}{2^n}$,
- Jedes B_n enthaelt unendlich viele Folgeglieder.

(Vgl. Intervallhalbierungsverfahren zum Beweis des Satzes von Bolzano / Weierstrass im ersten Semester). Aufgrund der dritten Eigenschaft koennen wir dann eine Teilfolge (x_{n_k}) von (x_n) konstruieren mit $x_{n_k} \in$

B_k fuer alle $k \in \mathbb{N}$. Aufgrund der ersten beiden Eigenschaften handelt es sich um eine Cauchy Folge.

Ende der Vorlesung

(ii)$\Longrightarrow$ (i): Angenommen B ist nicht total beschraenkt. Dann existiert also ein $\varepsilon > 0$ so dass B nicht mit endlich vielen Kugeln mit dem Radius ε ueberdeckt werden kann. Dann koennen wir induktiv eine Folge (x_n), $n \in \mathbb{N}$, in B von Kugelmittelpunkten konstuieren mit

$$x_{n+1} \notin \bigcup_{k=1}^{n} B_\varepsilon(x_n).$$

$n = 1$: Waehle x_1 beliebig.
$n \Longrightarrow n+1$: Es wird B nicht von $\cup_{k=1}^{n} B_\varepsilon(x_n)$ ueberdeckt nach Annahme.

Die Elemente dieser Folge erfuellen

$$d(x_n, x_k) \geq \varepsilon$$

fuer alle $n, k \in \mathbb{N}$ mit $n \neq k$. Daher kann diese Folge keine Teilfolge haben, die eine Cauchy Folge ist. Widerspruch zu (ii). $\square$

Damit koennen wir nun Kompaktheit charakterisieren.

Theorem. *(Charakterisierung Kompaktheit). Sei (X, d) ein metrischer Raum und $K \subset X$. Dann sind aequivalent:*

(i) *Es ist K total beschraenkt und vollstaendig.*
(ii) *Jede Folge in K hat eine Teilfolge, die in K konvergiert.*

Beweis. Das folgt leicht aus dem vorigen Lemma.
(i)$\Longrightarrow$ (ii): Da K total beschraenkt ist, hat nach dem vorigen Lemma jede Folge in K eine Teilfolge, die Cauchy Folge ist. Aufgrund der Vollstaendigkeit von K konvergiert diese dann in K.

(ii)$\Longrightarrow$ (i): Nach (ii) hat insbesondere jede Folge in K eine Teilfolge, die Cauchy Folge ist. Daher ist nach dem Lemma also die Menge K total beschraenkt. Noch zu zeigen: K vollstaendig. (Hier kommt das einzig neue im Beweis, das nicht schon aus dem vorigen Lemma folgt.) Sei (x_n) eine Cauchy Folge in K. Dann enthaelt (x_n) nach (ii) eine gegen ein $x \in K$ konvergente Teilfolge (x_{n_k}). Damit ist (nach Standardschluessen) die Folge (x_n) selber konvergent gegen x. Denn es gilt

$$d(x, x_n) \leq d(x, x_{n_k}) + d(x_{n_k}, x_n).$$

Es konvergiert nun $(d(x, x_{n_k}))$ gegen 0, da $x_{n_k} \to x$ und es wird $d(x_{n_k}, x_n)$ beliebig klein fuer k, n gross, da es sich um eine Cauchy Folge handelt. $\square$

Wir erinnern im Zusammenhang mit Vollstaendigkeit von Teilmengen noch an folgenden Zusammenhang.

LEMMA. *Sei (X, d) ein vollständiger metrischer Raum. Sei A eine Teilmenge von X. Dann ist A abgeschlossen (in X) genau dann, wenn der metrische Raum (A, d) vollstaendig ist.*

Beweis. Sei A abgeschlossen. Sei (x_n) eine Cauchy-Folge in (A, d). Dann ist (x_n) eine Cauchy Folge in X. Damit konvergiert (x_n) in X gegen ein x. Da A abgeschlossen ist, gehoert x zu A.

Sei (A, d) vollständig. Sei (x_n) eine gegen $x \in X$ konvergente Folge in A. Zu zeigen $x \in A$: Es ist (x_n) eine Cauchy Folge in X und damit auch in A. Daher hat (x_n) einen Grenzwert in A. Dieser muss dann mit x uebereinstimmen.

□

Beispiel - Kompaktheit in $\mathbb{R}^N$ bzw. $\mathbb{C}^N$. In $\mathbb{R}^N$ bzw. $\mathbb{C}^N$ ist eine Menge total beschraenkt genau dann, wenn sie beschraenkt ist. Sie ist vollstaendig, genau dann, wenn sie abgeschlossen ist. Damit erhalten wir folgendes Resultat: Ein $K \subset \mathbb{R}^N$ ist kompakt genau dann, wenn es abgeschlossen und beschränkt ist.

Eine Variante der Betrachtungen von $\mathbb{R}^N$ führt auf folgende Charakterisierung von Totaler Beschränktheit mittels Kompaktheit.

FOLGERUNG. *Sei (M, d) ein vollstaendiger metrischer Raum. Sei $A \subset M$. Dann sind äquivalent:*

(i) *A ist total beschraenkt.*
(ii) *Es ist $\overline{A}$ kompakt.*

Beweis. Man sieht leicht, dass eine Menge A von X total beschränkt ist genau dann wenn $\overline{A}$ total beschränkt ist. Damit können wir wie folgt weiterschliessen:

Ist $\overline{A}$ kompakt, so ist es also total beschraenkt. Dann ist auch A total beschraenkt.

Sei umgekehrt A total beschraenkt. Dann ist auch $\overline{A}$ total beschraenkt (einfach) und vollstaendig (als abgeschlossene Menge eines vollstaedigen Raumes). Daher ist $\overline{A}$ kompakt. □

DEFINITION. *Sei (X, d) ein metrischer Raum. Eine Menge A in X heisst relativ kompakt, wenn ihr Abschluss kompakt ist.*

Wir kommen nun zu einer weiteren Charakterisierung von Kompaktheit. Diese lässt sich auch fuer allgemeine topologische Räume geben.

THEOREM. *Sei (M, d) ein metrischer Raum und $K \subset M$. Dann sind aequivalent:*

(i) *Jede Folge in K hat eine in K konvergente Teilfolge.*

(ii) *Jede offenen Ueberdeckung von K hat eine endliche Teilueberdeckung. (Das heisst: Zu allen U_α, $\alpha \in A$, offen mit $K \subset \cup U_\alpha$ existiert $N \in \mathbb{N}$ und $\alpha_1, \ldots, \alpha_N$ mit $K \subset \cup_{j=1}^N U_{\alpha_j}$.) 'K ist ueberdeckungskompakt'.*

Bemerkung.

- Die Eigenschaft (ii) kann auch als Definition von Kompaktheit genommen werden. Man spricht dann von Ueberdeckungskompaktheit (oder der Heine-Borel Eigenschaft). Der Satz besagt dann, dass in metrischen Räumen Ueberdeckungskompaktheit aequivalent zu Folgenkompaktheit ist.
- In der Eigenschaft (ii) wird kein Bezug auf die Metrik genommen sondern nur auf die erzeugte Topologie. Insbesondere haben also zwei Metriken, die dieselbe Topologie erzeugen, dieselben kompakten Mengen.

Beweis. (ii) $\Longrightarrow$ (i): Sei (x_n) eine Folge in K. Angenommen (x_n) enthaelt keine konvergente Teilfolge. Dann hat (x_n) also keinen Haeufungspunkt. Dann existiert zu jedem $x \in K$ also ein $\delta_x > 0$ so dass in $U_{\delta_x}(x)$ nur endlich viele Folgeglieder liegen. Dann ist $U_{\delta_x}(x)$ eine offene Ueberdeckung von K und hat also eine endliche Teilueberdeckung

$$U_{\delta_{p_j}}(p_j), \;\; j = 1, \ldots, N.$$

Da jedes $U_{\delta_{p_j}}$ nur endlich viele Folgeglieder enthaelt, gibt es also insgesamt nur endlich viele Folgeglieder. Das ist ein Widerspruch.

(i)$\Longrightarrow$ (ii): Sei U_α, $\alpha \in A$, eine offene Ueberdeckung von K. Wir gehen in drei Schritten vor.

Schritt 1: Es gibt ein $\delta > 0$, so dass fuer jedes $x \in K$ ein $\alpha_x \in A$ existiert mit $x \subset U_\delta(x) \subset U_{\alpha_x}$. 'Jede δ-Kugel liegt in einem U_α'.

Bew. Angenommen nein! Dann gibt es also zu jedem $n \in \mathbb{N}$ (fuer $\delta = 1/n$) ein $x_n \in K$ sodass $U_{1/n}(x_n)$ nicht in U_α enthalten ist fuer alle $\alpha \in A$. Aufgrund von (i) hat (x_n) eine konvergente Teilfolge. Ohne Einschraenkung sei $x_n \to x \in K$. Dann gibt es ein $\alpha \in A$ mit $x \in U_\alpha$. Da U_α offen ist, ist dann auch $U_r(x) \subset U_\alpha$ fuer ein geeignetes $r > 0$. Fuer hinreichend grosse n ist dann aber $d(x_n, x) < r/2$ und $1/n < r/2$ also

$$U_{1/n}(x_n) \subset U_{r/2}(x_n) \subset U_r(x) \subset U_\alpha.$$

Das ist ein Widerspruch zur Konstruktion der (x_n).

Schritt 2: Zu $\delta > 0$ aus Schritt 1, gibt es ein $N \in \mathbb{N}$ und $x_1, \ldots, x_N \in K$ mit $K \subset \cup_{j=1}^N U_\delta(x_j)$.

Bew. Nach (i) und der Charakterisierung von Kompaktheit, ist K total beschraenkt. Damit folgt Schritt 2 (fuer jedes $\delta > 0$).

Schritt 3: Es gilt (ii).

Bew. Seien $\delta > 0$ und $x_1, \ldots, x_N$ aus Schritt 2 gegeben. Dann gilt also

$$K \subset \cup_{j=1}^N U_\delta(x_j).$$

Nach Schritt 1 ist jedes $U_\delta(x_j)$ in einem U_{α_j} enthalten und (ii) folgt. $\square$

PROPOSITION. *Seien (X,d) und (Y,e) metrische Raeume und $K \subset X$ und $L \subset Y$ kompakt. Dann ist $K \times L \subset X \times Y$ kompakt.*

Beweis. Sei (x_n, y_n) eine Folge in $K \times L$. Dann ist (x_n) eine Folge in K. Damit hat (x_n) eine in K gegen ein $x \in K$ konvergente Teilfolge (x_{n_k}). Dann ist (y_{n_k}) eine Teilfolge in L. Daher hat es eine gegen ein $y \in L$ konvergente Teilfolge (y_{m_k}). Dann gilt

$$x_{m_k} \to x \text{ und } y_{m_k} \to y.$$

Damit konvergiert (x_{m_k}, x_{n_k}) gegen $(x,y) \in K \times L$. $\square$

Bemerkung. Alternativ hier noch ein Ueberdeckungsbeweis: Sei W_α, $\alpha \in A$, eine offene Ueberdeckung von $K \times L$.
Der Beweis wird nun in drei Schritten gefuehrt. (Zeichnung)
Schritt 1: Es gibt zu jedem $(x,y) \in K \times L$ ein $\alpha(x,y) \in A$ und Umgebungen $U_{(x,y)}$ von x und $V_{(x,y)}$ von y mit

$$(x,y) \in U_{(x,y)} \times V_{(x,y)} \subset W_{\alpha(x,y)}.$$

Bew. Das folgt aus der Ueberdeckungseigenschaft under Definition der Produkttopologie.
Schritt 2: Fuer festes $x \in K$ ist $(V_{(x,y)})_{y \in L}$ eine offene Ueberdeckung von Y. Also gibt es $y_1(x), \ldots, y_{N(x)}(x)$ mit

$$L \subset v_{(x,y_1(x)} \cup \ldots \cup V_{(x,y_{N(x)}(x)}.$$

Setze nun

$$U_x := U_{(x,y_1(x)} \cap \ldots \cap U_{x,y_{N(x)}(x)}.$$

Schritt 3: Aufgrund der Kompaktheit von K gibt $x_1, \ldots, x_M$ in K mit

$$K \subset U_{x_1} \cup \ldots \cup U_{x_M}.$$

Damit ueberdecken dann die endlich vielen

$$W_{\alpha(x_j, y_k(x_j))},\ j = 1, \ldots, M, k = 1, \ldots, N(x_j)$$

die Menge $K \times L$.

6. Zusammenhang

In diesem Abschnitt diskutieren wir ein weiteres topologisches Konzept, naemlich Zusammenhang.

DEFINITION. *Ein metrischer Raum (X,d) heisst zusammenhaengend, wenn er nicht in zwei offene disjunkte nichtleere Teilmengen zerlegt werden kann (d.h. wenn aus $X = U \cup V$ mit U, V offen, $U \cap V = \emptyset$ folgt $U = X$ oder $V = X$). Eine Teilmenge A eines metrischen Raumes heisst zusammenhaengend, wenn (A,d) als metrischer Raum zusammenhaengend ist.*

Zeichnung.

Bemerkung. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Ist dann U eine Teilmenge des metrischen Raumes (A, d) so ist U offen (im metrischen Raum (A, d)) genau dann, wenn es eine offene Menge U' in (X, d) gibt mit $U = A \cap U'$. **Zeichnung.**
(Bew. Es gilt offenbar

$$U_r^A(x) := \{a \in A : d(a, x) < r\} = U_r^X(x) \cap A.$$

Damit folgt die Aussage leicht.)

Hier ist das Hauptergebnis ueber stetige Funktionen auf Zusammenhaengenden Mengen.

THEOREM. *Seien (X, d) und (Y, e) metrische Raeume und (X, d) zusammenhaengend und $f : X \longrightarrow Y$ stetig, dann ist $f(X)$ zusammenhaengend.*

Beweis. Seien U, V disjunkte offene Mengen in $(f(X), d)$ mit $U \cup V = f(X)$. Dann sind aufgrund der Stetigkeit von f also

$$U_X := f^{-1}(U), \;\; V_X := f^{-1}(V)$$

offene disjunkte Mengen in X mit $X = U_X \cup V_X$. Da X zusammenhängend ist, folgt $U_X = \emptyset$ oder $V_X = \emptyset$. Das liefert die Behauptung. □

Bemerkung. Wie wir schon erwaehnt hatten, gibt es eigentlich nur zwei globale Saetze zu stetigen Funktionen. Diese besagen:

- Stetige Bilder kompakter Mengen sind kompakt.
- Stetige Bilder zusammenhaengender Mengen sind zusammenhaengend.

Beispiel - zusammenhaengenden Mengen in $\mathbb{R}$. Wir untersuchen nun noch die Struktur der zusammenhaengende Mengen in $\mathbb{R}$. Das wird ein strukturelles Verstaendnis des Zwischenwertsatzes liefern.

THEOREM. *Die zusammenhaengenden Mengen in $\mathbb{R}$ sind gerade die Intervalle.*

Beweis. Sei $I \subset \mathbb{R}$ zusammenhaengend. Zu zeigen: Mit $x, y \in I$ gehoert auch jeder Punkt zwischen x und y zu I.
Sei ohne Einschraenkung $x < y$. Betrachte p mit $x < p < y$. Angenommen $p \notin I$. Dann bilden $U := (-\infty, p) \cap I$ und $V : (p, \infty) \cap I$ eine offene disjunkte Zerlegung von I in nichtleere Mengen. Widerspruch.

Sei nun $I \subset \mathbb{R}$ ein Intervall. Sei $I = U \cup V$ eine Zerlegung von I in offene disjunkte Mengen. Betrachte die charakteristische Funktion

$$1_U : I \longrightarrow \mathbb{R},$$

mit $1_U(x) = 1$ falls $x \in U$ und $1_U(x) = 0$ sonst. Dann sind die Urbilder von beliebigen Mengen in $\mathbb{R}$ unter 1_U offenbar entweder $\emptyset$ oder U oder

V oder I. Damit ist 1_U stetig. Da 1_U nur Werte in $\{0,1\}$ annimmt, folgt nach dem Zwischenwertsatz, dass 1_U konstant sein muss. □

Definition. *(Wegzusammenhaengend) Sei (X,d) ein metrischer Raum. Dann heisst (X,d) wegzusammenhaengend, wenn zu beliebigen $x,y \in X$ eine stetige Abbildung $\gamma : [0,1] \longrightarrow X$ existiert mit $\gamma(0) = x$, $\gamma(1) = y$. Ein solches γ heisst dann Weg von x nach y.*

Beispiel. Eine $C \subset \mathbb{R}^N$ heisst konvex, wenn mit $x,y \in C$ auch die Verbindungsstrecke $V = \{tx + (1+t)y : 0 \leq t \leq 1\}$ zwischen x und y zu C gehoert. Jede konvexe Menge in $\mathbb{R}^N$ ist offenbar wegzusammenhaengend (Klar!?). Insbesondere ist jede Kugel in $\mathbb{R}^N$ wegzusammenhaengend.

Folgerung. *Ist (X,d) wegzusammenhaengend, so ist (X,d) zusammehaengend.*

Beweis. Sei $X = U \cup V$ mit U,V offen und disjunkt. Angenommen $U \neq \emptyset$ und $V \neq \emptyset$. Dann gibt es also $x \in U$ und $y \in V$. Da (X,d) wegzusammenhaengend ist, gibt es einen Weg γ von x nach y. Dann ist $I := [0,1] = \gamma^{-1}(U) \cup \gamma^{-1}(V)$ eine Zerlegung in offene (in I) disjunkte nichtleere Teilmengen. Das ist ein Widerspruch. □

Ende der Vorlesung

Im allgemeinen gilt die Umkehrung nicht. In speziellen Faellen aber gilt eine Umkehrung. Das wird im folgenden Theorem behandelt.

Theorem. *Sei $U \subset \mathbb{R}^N$ (bzw. $U \subset \mathbb{C}^N$) offen und zusammenhaengend. Dann ist U auch wegzusammenhängend.*

Beweis. Sei $p \in U$ beliebig. Sei C_p die Menge der $x \in U$ fuer die eine stetige Kurve $\gamma : [0,1] \longrightarrow U$ existiert mit $\gamma(0) = p$ und $\gamma(1) = x$. Dann gilt: Sei $R_p := U \setminus C_p$.

- C_p ist offen, da U offen ist und jede Kugel wegzusammenhaengend. **Zeichnung**
- R_p ist offen, da U offen ist und jede Kugel wegzusammenhaengend. **Zeichnung**

Da U zusammenhaengend ist und $C_p \neq \emptyset$ gilt, folgt nun $R_p = \emptyset$. □

Proposition. *Ist (X,d) ein wegzusammenhaengender metrischer Raum, (Y,e) ein metrischer Raum und $f : X \longrightarrow Y$ stetig, so ist $(f(X),e)$ wegzusammenhaengend.*

Beweis. Das ist klar. **Zeichnung.** (Ein Weg zwischen $f(x)$ und $f(y)$ wird gerade durch $f \circ \gamma$ gegeben, wobei γ ein Weg von x nach y ist). □

7. Anwendungen - Der Banachsche Fixpunktsatz (Bonusmaterial)

In diesem Abschnitt lernen wir den Banachschen Fixpunktsatz kennen, der an vielen Stellen nuetzlich sein kann. Wir skizzieren kurz eine Anwendung (Erzeugung von Fraktalen), bei der kompakte Mengen und Metriken eine grosse Rolle spielen.

THEOREM. *(Banachsche Fixpunktsatz) Sei* (X, d) *ein vollständiger metrischer Raum und* $f : X \longrightarrow X$ *gegeben. Gibt es eine ein* $0 < c < 1$ *mit*

$$d(f(x), f(y)) \leq cd(x, y)$$

fuer alle $x, y \in X$, *so existiert ein eindeutiges* $p \in X$ *mit* $f(p) = p$. *Ist* $x \in X$ *beliebig und* (x_n) *induktiv definiert durch*

$$x_0 := x, \text{ und } x_{n+1} := f(x_n)$$

(d.h. $x_n = f \circ \ldots \circ f(x)$*), so konvergergiert* (x_n) *gegen* p, *und es gilt die apriori Abschaetzung*

$$d(p, x_n) \leq c^n \frac{d(x_0, x_1)}{1 - c}.$$

Bemerkung Eine Abbildung f wie im Theorem heisst Kontraktion mit Kontraktionsfaktor c. Ein p wie im Theorem heisst Fixpunkt von f.

Beweis. Wir zeigen Eindeutigkeit, Existenz und die Apriori-Abschaetzung.

Eindeutigkeit. Seien $x, y \in X$ mit $f(x) = x$ und $f(y) = y$ geben. Dann gilt

$$d(x, y) = d(f(x), f(y)) \leq cd(x, y).$$

Wegen $c < 1$ und $d \geq 0$ folgt $d(x, y) = 0$.

Existenz. Sei $x \in X$ beliebig und x_n wie in der Aussage des Theorem. Dann gilt (Induktion)

$$d(x_j, x_{j+1}) \leq c^j d(x_0, x_1)$$

fuer alle $j \in \mathbb{N}$ und damit fuer $n < m$

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{j=n}^{m-1} d(x_j, x_{j+1}) \\ &\leq \sum_{j=n}^{m-1} c^j d(x_0, x_1) \\ &\leq c^n d(x_0, x_1) \sum_{j=0}^{m-n-1} c^j \\ &\leq c^n d(x_0, x_1) \frac{1}{1-c}. \end{aligned}$$

Damit ist also (x_n) eine Cauchy Folge. Aufgrund der Vollstaendigkeit konvergiert dann (x_n) und fuer den Grenzwert p gilt

$$d(p, f(p)) = \lim_{n\to\infty} d(x_n, f(x_n)) = \lim_{n\to\infty} d(x_n, x_{n+1}) = 0.$$

Damit ist p ein Fixpunkt.

Abschaetzung. Schliesslich gilt aufgrund des schon gezeigten

$$d(p, x_n) = \lim_{m\to\infty} d(x_m, x_n) \leq c^n d(x_0, x_1) \frac{1}{1-c}.$$

Das beendet den Beweis. □

Anwendung - Hausdorffmetrik. Sei (X, d) ein metrischer Raum. Wir werden eine Abstandsfunktion auf den (kompakten) Teilmengen von (X, d) einführen und damit die Menge der kompakten Teilmengen zu einem metrischen Raum machen.

Fuer eine Teilmenge A von X und $x \in X$ definiert man

$$d(x, A) := \inf\{d(x, a) : a \in A\}.$$

Aufgrund der Dreiecksungleichung gilt dann (einfach!)

$$d(x, A) - d(y, A) \leq d(x, y)$$

fuer alle $x, y \in X$. Damit ist also fuer gegebenes A die Abstandsfunktion

$$d_A := d(\cdot, A) : X \longrightarrow \mathbb{R}$$

stetig auf X. Ist A abgeschlossen, so gilt weiterhin

$$d(x, A) = 0 \Longleftrightarrow x \in A.$$

($\Longrightarrow$: $x \notin A$ impliziert $U_r(x) \in X \setminus A$ da A abgeschlosse ist. Damit folgt $d(x, A) \geq r$. $\Longleftarrow$: klar).

Ist A kompakt, so gibt es ein $a \in A$ mit $d(x, A) = d(x, a)$ und man kann das Infimum durch ein Minimum ersetzen (da die stetig Funktion $d(x, \cdot) : K \longrightarrow \mathbb{R}$ ihr Minimum annimmt). Sind K, L kompakte

Teilmengen von X, so koennen wir nun d_K und d_L auf L bzw K untersuchen. Diese Funktionen sind stetig (s.o.) und nehmen also auf L bzw. K ihr Maxmimum.

- Es gilt $\max_K d_L := \max\{d(x,L) : x \in K\} \leq \varepsilon$ genau dann, wenn zu jedem $x \in K$ ein $y \in L$ existiert mit $d(x,y) \leq \varepsilon$. **Zeichnung.**
- Es gilt $\max_L d_K := \max\{d(K,y) : y \in L\} \leq \varepsilon$ genau dann, wenn zu jedem $y \in L$ ein $x \in K$ existiert mit $d(x,y) \leq \varepsilon$. **Zeichnung.**

Fuer kompakte Teilmengen K, L von X definiert man dann den Haussdorffabstand

$$d_H(K,L) := \max\{\max_K d_L, \max_L d_K\}.$$

Damit gilt $d_H(K,L) \leq \varepsilon$ genau dann, wenn zu jedem $x \in K$ ein $y_x \in L$ existiert mit $d(x,y_x) \leq \varepsilon$ und umgekehrt zu jedem $y \in L$ ein $x_y \in K$ existiert mit $d(x_y,y) \leq \varepsilon$.

THEOREM. *Ist (X,d) ein vollstaendiger metrischer Raum, so ist die Menge $\mathcal{K}(X)$ der kompakten Teilmengen von X mit der Hausdorffmetrik d_H ein vollstaendiger metrischer Raum.*

Beweis. Beweis nicht hier. (Evtl. in Uebung). $\square$

FOLGERUNG. *(Erzeugung von Fraktalen) Ist (X,d) ein vollstaendiger metrischer Raum und $F : \mathcal{K}(X) \longrightarrow \mathcal{K}$ eine Kontraktion. Dann gibt genau eine kompakte Menge K in X mit $F(K) = K$.*

Beweis. Das folgt sofort aus dem Banachschen Fixpunktsatz und dem vorangegangenen Theorem. $\square$

Eine kleine Rechnung zeigt folgende interessante Eigenschaft der Haussdorffmetrik.

PROPOSITION.

$$d_H(A \cup B, C \cup D) \leq \max\{d_H(A,C), d_H(B,D)\}.$$

Anwendung. Sei $X = \mathbb{R}$ mit der Euklidischen Metrik und

$$F : \mathcal{K} \longrightarrow \mathcal{K},\ \ F(K) = \frac{1}{3}K + (2/3 + \frac{1}{3}K).$$

Dann ist gilt (kleine Rechnung)

$$d_H(F(L), F(K)) \leq \frac{1}{3}d(K,L).$$

Damit hat also F nach dem Banachschen Fixpunktsatz einen Fixpunkt C und fuer C gilt

$$C = \lim_{n\to\infty} F \circ \ldots \circ F(K)$$

fuer jedes kompakte K in $\mathbb{R}$.
Wir 'bestimmen' den Fixpunkt auf zwei Arten, naemlich durch Wahl von $K_0 = [0,1]$ und $K_0 = \{0\}$. **Zeichnung.**

Der Fixpunkt ist die uns aus dem ersten Semester schon vertraute Cantormenge.

KAPITEL 14

Differenzierbarkeit im Höherdimensionalen

In diesem Kapitel geht es um Funktionen von Teilmengen des $\mathbb{R}^N$ nach $\mathbb{R}^M$. Dabei werden $\mathbb{R}^N$ und $\mathbb{R}^M$ als metrische Räume mit der Euklidischen Metrik aufgefasst. Ein wesentlicher Inhalt dieses Kapitels ist die Untersuchung von Differenzierbarkeit solcher Funktionen. Anschaulich bedeutet die Differenzierbarkeit von f, dass sich f gut durch eine lineare Abbbildung, seine Ableitung, approximieren lässt.

1. Zum Aufwärmen: Funktionen von $\mathbb{R}^N$ nach $\mathbb{R}^M$

Es geht um Funktionen von Teilmengen U des $\mathbb{R}^N$ nach $\mathbb{R}^M$ (mit geeigneten N und M). Den Fall $N = M = 1$ haben wir im vergangenen Semester schon untersucht.
Eine Funktion $f : U \longrightarrow \mathbb{R}^M$ laesst sich offenbar schreiben als $f(x) = (f_1(x), \ldots, f_M(x))$ mit $f_j : U \longrightarrow \mathbb{R}$. Es ist f stetig, genau dann, wenn jedes f_j stetig ist (siehe Uebung). Einige spezielle Situationen haben eigene Namen:

- Ist $U \subset \mathbb{R}$ ein Intervall, so heisst $\gamma : U \longrightarrow \mathbb{R}^M$ eine Kurve. **Zeichnung.**
  Beispiel: Es ist $\mathbb{R} \longrightarrow \mathbb{R}^M, t \mapsto p + tv$ eine stetige Kurve (fuer gegebenes $p \in \mathbb{R}^M$ und $v \in \mathbb{R}^M$).
- Ist $U \subset \mathbb{R}^N$, so heisst $f : U \longrightarrow \mathbb{R}$ eine skalare Funktion. **Zeichnung.**
  Beispiel: Es sind die Koordinatenfunktionen $\pi_j : \mathbb{R}^N \longrightarrow \mathbb{R}, (\xi_1, \ldots, \xi_N) \mapsto \xi_j$ stetige skalare Funktionen.
- Ist $U \subset \mathbb{R}^N$, so heisst $f : U \longrightarrow \mathbb{R}^N$ ein Vektorfeld auf U. **Zeichnung.**
  Beispiel: Es liefert die Identitaet $id : U \longrightarrow \mathbb{R}^N, x \mapsto x$, ein stetiges Vektorfeld.

Natuerlich sind Funktionen von einer Variable besonders leicht zu handhaben. Die folgende Prozedur erlaubt es, aus Funktionen mehrerer Variablen, mehrere Funktionen einer Variablen zu machen. Damit lassen sich viele (aber nicht alle) Betrachtungen im Mehrdimensionalen auf Betrachtungen im eindimensionalen zurueckfuehren: Haelt man fuer $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ alle Variablen bis auf eine - z.B. die j-te fest, so erhaelt man die *partiellen Funktionen*

$$t \mapsto f(\xi_1, \ldots, \xi_{j-1}, \xi_j + t, \xi_{j+1}, \ldots, \xi_N).$$

Genauer definiert man mit der Standard Normalbasis (e_j) im $\mathbb{R}^N$ zu $p \in U$ und jedem $j \in \{1, \ldots, N\}$ die sogenannte partielle Funktion

$$f_{p,j} : \{t \in \mathbb{R} : p + te_j \in U\} \longrightarrow \mathbb{R}^M, \ f_{p,j}(t) := f(p + te_j).$$

Dann heisst f *partiell stetig in* p, wenn $f_{p,j}$ stetig in 0 ist fuer alle $j = 1, \ldots, N$. Gilt das fuer alle $p \in U$, so heisst f *partiell stetig.*

Allgemeiner kann man zu jedem Einheitsvektor $e \in \mathbb{R}^N$ und jedem $p \in U$ die Funktion

$$f_{p,e} : \{t \in \mathbb{R} : p + te \in U\} \longrightarrow \mathbb{R}^M, f_{p,e}(t) := f(p + te),$$

definieren. Offenbar gilt dann also $f_{p,e_j} = f_j$, $j = 1, \ldots, N$. Es heisst f *richtungsstetig in* p, wenn fuer alle Einheitsvektoren $e \in \mathbb{R}^N$, die Funktion $f_{p,e}$ stetig in 0 ist. Gilt dies fuer alle $p \in U$, so heisst f *richtungsstetig.*

Offenbar gilt folgende Proposition. (Uebung)

PROPOSITION. *f stetig (in p) $\Longrightarrow$ f richtungsstetig (in p)$\Longrightarrow$ f partiell stetig (in p).*

Beweis. Die Funktionen $f_{p,e}$ koennen als Verknuepfung von f mit geeigneten Kurven γ dargestellt werden. □

Bemerkung. Die Umkehrung ist aber nicht wahr. (Uebung)

Schliesslich erinnern wir zum Abschluss noch an das Konzept des Grenzwertes (vgl. Analysis I, bzw. Betrachtungen zu metrischen Raeumen): Sei $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}^M$ gegeben. Sei $p \in \overline{U}$ (d.h. p ist Grenzwert einer Folge in U). Dann hat f bei p den Grenzwert $c \in \mathbb{R}^M$ geschrieben als

$$\lim_{x \to p} f(x) = c$$

genau dann, wenn gilt

- $f(x_n) \to c$ fuer jede Folge (x_n) in U mit $x_n \to p$.

bzw. aequivalent, wenn gilt

- zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $|f(x) - c| \leq \varepsilon$ fuer alle $x \in U$ mit $|x - p| \leq \delta$.

Entsprechend schreibt man $\lim_{x \to p, x \neq p} f(x) = c$, falls fuer die Einschraenkung $f|_{U \setminus \{p\}}$ von f auf $U \setminus \{p\}$ gilt

$$\lim_{x \to p} f|_{U \setminus \{p\}}(x) = c.$$

2. Definition der Ableitung und einfache Eigenschaften

Es geht um lineare Approximationen. Die Ableitung ist die lineare Approximation.

Aus dem ersten Semester wissen wir, dass fuer $f : (a, b) \longrightarrow \mathbb{R}$ und $p \in (a, b)$ aequivalent sind:

- $f^{'}(p) := \lim_{x\to p}(f(x) - f(p))/(x - p)$ existiert.
- Es gilt $f(x) = f(p) + c(x - p) + \varphi(x)$ mit $\varphi(x)/|x - p| \to 0$, $x \to p$.

In diesem Fall heisst f differenzierbar in p. Wir werden das nun auf hoeherdimensionale Situationen verallgemeinern. Dabei gehen wir von der im zweiten Punkt praesentierten Variante aus (da die Division durch $x - p$ im hoeherdimensionalen keinen Sinn hat).

Definition. *Sei $U \subset \mathbb{R}^N$ offen und $f : U \to \mathbb{R}^M$ gegeben. Sei $p \in U$. Dann heißt f in p differenzierbar, wenn eine lineare Abbildung $L : \mathbb{R}^N \to \mathbb{R}^M$ existiert mit*

$$f(x) = f(p) + L(x - p) + \varphi_p(x),$$

sodass für den Fehler

$$\varphi_p(x) = f(x) - f(p) - L(x - p)$$

gilt

$$\frac{\varphi_p(x)}{|x - p|} \to 0, x \to p.$$

In diesem Fall heißt L Ableitung von f in p. Ist f in jedem $p \in U$ differenzierbar, so heißt f auf ganz U differenzierbar.

Bemerkung.

- Ist f differenzierbar in p, so wird es also 'sehr gut' durch die linear (affine) Funktion
$$f(p) + L(\cdot - p)$$
approximiert.
- Die Offenheit von U kommt ins Spiel, da wir f von allen 'Seiten' approximieren koennen wollen. Das werden wir spaeter im Konzept der Richtungsableitung noch genauer machen (s.u.).
- Im Falle einer Funktion von $U \subset \mathbb{R} \longrightarrow \mathbb{R}$ ist die lineare Abbildung L gerade durch eine Zahl gegeben (lineare Abbildung von $\mathbb{R}$ nach $\mathbb{R}$). Diese Zahl ist die uns schon aus dem ersten Semester bekannte Ableitung.

Proposition. *In der Situation der Definition ist L eindeutig.*

Beweis. Es gebe L_1 und L_2 mit

$$f(x) = f(p) + L_1 \cdot (x - p) + \varphi_1(x)$$

$$f(x) = f(p) + L_2 \cdot (x - p) + \varphi_2(x)$$

und $\varphi_j(x)/|x - p| \to 0$, $x \to p$, $j = 1, 2$. Dann gilt

$$(L_1 - L_2)(x - p) = \varphi_1(x) - \varphi_2(x) = \frac{\varphi_1(x) - \varphi_2(x)}{|x - p|}|x - p|$$

Betrachtet man $x = p + he$ mit $h \in \mathbb{R}$, $h \neq 0$ und einem beliebigen Einheitsvektor $e \in \mathbb{R}^N$, so gilt also $x - p = he$ und

$$(L_1 - L_2)(he) = \frac{\varphi_1(x) - \varphi_2(x)}{|he|}|x - p|$$

und damit konvergiert die rechte Seite von

$$(L_1 - L_2)e = \frac{\varphi_1(x) - \varphi_2(x)}{|he|}\frac{|he|}{h}$$

fuer $h \to 0$ gegen 0. Es folgt also

$$(L_1 - L_2)e = 0.$$

Da e ein beliebiger Einheitsvektor ist, folgt $L_1 = L_2$. $\square$

Notation. Die vorigen Propostion erlaubt es uns, von DER Abeltung von f in p zu sprechen. Wir bezeichnen diese dann mit $Df(p)$.

Beispiel - Lineare Funktion. Sei $A : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ linear, $a \in R^M$ und $q \in \mathbb{R}^N$ gegeben. Dann ist die Funktion

$$f : \mathbb{R}^N \longrightarrow \mathbb{R}^M, f(x) = a + A(x - q)$$

differenzierbar und es gilt $Df(x) = A$ fuer alle $x \in \mathbb{R}^N$.
Bew. Es gilt $f(x) - f(p) = A(x - p)$, also

$$f(x) = f(p) + A(x - p) + 0.$$

Damit folgt die Differenzierbarkeit (mit $L = A$ und $\varphi = 0$).

Theorem. *(Differenzierbarkeit impliziert Stetigkeit) Sei $U \subset \mathbb{R}^N$ offen und $f : U \to \mathbb{R}^M$ gegeben. Ist f in p differenzierbar, so ist es dort auch stetig.*

Beweis. Es gilt (fuer $x \neq p$)

$$f(x) = f(p) + L(x - p) + \varphi(x) = f(p) + L(x - p) + \frac{\varphi(x)}{|x - p|}|x - p|.$$

Damit folgt sofort, dass die rechte Seite gegen $f(p)$ konvergiert für $x \to p$. $\square$

Wir haben nun die Ableitung L *charakterisiert*. Als Nächstes soll es darum gehen, wie man diese lineare Funktion *berechnet*.

Theorem. *(Berechnen der Ableitung) Sei $U \subset \mathbb{R}^N$ offen und $f : U \to \mathbb{R}^M$ differenzierbar in $p \in U$. Dann existieren die partiellen Ableitungen*

$$D_j f_i(p) := \partial_j f_i(p) := \frac{\partial f_i}{\partial x_j}(p) := \lim_{h \to 0, h \neq 0} \frac{1}{h}(f_i(p + he_j) - f_i(p))$$

für $i = 1, \ldots, M$, $j = 1, \ldots, N$ und die Ableitung L wird bzgl. der Standardbasen in $\mathbb{R}^N$ bzw. $\mathbb{R}^M$ durch die Matrix

$$\left(\frac{\partial f_i}{\partial x_j}(p)\right)_{i,j} = (....).$$

gegeben.

Bemerkungen.

- Die partielle Ableitung von f_i nach x_j in p ist also gerade die 'gewoehnliche' Ableitung der oben betrachteten partiellen Funktion
$$t \mapsto f_i(p + te_j)$$
an der Stelle 0.
- Die partiellen Ableitungen werden gebildet, indem man alle Variablen bis auf eine festhaelt und dann nach dieser einen so ableitet, wie wir es im ersten Semester gelernt haben.
- Der Satz beschreibt, wie man die Ableitung von f ausrechnet, wenn man schon weiss, dass f differenzierbar ist. Der Satz besagt NICHT, dass f differenzierbar ist, wenn die partiellen Ableitungen existieren (und das ist im allgemeinen auch falsch (siehe Uebung)).

Beweis. Nach Voraussetzung gilt
$$f(x) = f(p) + L(x - p) + \varphi(x)$$
mit $\frac{\varphi(x)}{|x-p|} \to 0$ und $L = (L_{i,j})$. Wir setzen $x = p + he_j$ und betrachten die i-te Komponente. Fuer diese gilt
$$f_i(p + he_j) = f_i(p) + (L(he_j))_i + \varphi_i(x).$$
Das impliziert
$$f_i(p + he_j) - f_i(p) = h(Le_j)_i + \varphi_i(x).$$
Damit folgt nach Division durch h also
$$\frac{f_i(p + he_j) - f_i(p)}{h} = L_{i,j} + \frac{\varphi_i(p + he_j)}{h}$$
Wegen $\frac{|\varphi_i(p+he_j)|}{|h|} = \frac{|\varphi_i(p+he_j)|}{|x-p|} \to 0$, $h \to 0$, folgt
$$\frac{f_i(p + he_j) - f_i(p)}{h} \to L_{i,j},\ h \to 0.$$
Das beendet den Beweis. □

DEFINITION. *Sei $U \subset \mathbb{R}^N$ offen. Eine Funktion $f : U \longrightarrow \mathbb{R}^M$ heißt in $p \in U$ partiell differenzierbar, wenn die im vorigen Theorem definierten partiellen Ableitungen in p existieren. In diesem Fall heißt $\left(\frac{\partial f_i}{\partial x_j}(p)\right)_{i,j}$ die Jacobi-Matrix oder auch Differentialmatrix von f in p.*

Ende der Vorlesung

Notation (Gradient). Existieren fuer $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}$ die partiellen Ableitungen in in $p \in U$, so nennt man den Vektor $(\partial_1 f(p), \ldots, \partial_N f(p))^t$ den Gradienten von f in p und schreibt ihn als $\nabla f(p)$. Ist f differenzierbar, so ist also $Df(p) = \nabla f(p)^t$.

Beispiele.

- $f : U \subset \mathbb{R}^N \to \mathbb{R}$ differenzierbar in p. Dann ist:
$$\begin{aligned} f(x) &= f(p) + Df(p)(x-p) + \varphi(x) \\ &= f(p) + \langle \nabla f(p), x-p \rangle + \varphi(x) \end{aligned}$$
mit $\varphi(x)/|x-p| \to 0$, $x \to p$.
- Sei $\gamma : (a,b) \to \mathbb{R}^N$ gegeben. Dann ist γ differenzierbar in $t_0 \in (a,b)$ genau dann, wenn
$$\gamma^{'}(t_0) := \lim_{t \to t_0} \frac{1}{t-t_0}(\gamma(t) - \gamma(t_0))$$
existiert (Bew....). Dann gilt
$$\gamma(t) = \gamma(t_0) + \gamma^{'}(t_0)(t-t_0) + \varphi(t)$$
mit $\varphi(t)/(t-t_0) \to 0$, $t \to t_0$. Insbesondere haben wir für $\gamma(t) = t_0 + vt$,
$$\gamma'(t) = v.$$
Tatsaechlich kann man in diesem Fall auch $\gamma : [a,b] \longrightarrow U$ betrachten und die Ableitung von γ in a bzw. b durch den obigen Grenzwert definieren (wenn der Grenzwert existiert) (vgl. Uebung).
- Ist $f : \mathbb{R}^N \backslash \{0\} \to \mathbb{R}$, $f(x) = |x|$, so gilt (Uebung)
$$\partial_j f(x) = \tfrac{x_j}{|x|}, \text{ also } \nabla f(x) = \tfrac{x}{|x|}$$
Tatsaechlich (s.u.) ist f in diesem Fall differenzierbar.

Wir lernen als nächstes einige Rechenregeln für die Ableitung kennen.

THEOREM. *(Kettenregel) Sei $U \subset \mathbb{R}^N$, $V \subset \mathbb{R}^M$ offen und $f : U \to V$, $g : V \to \mathbb{R}^L$ gegeben. Sind f in p und g in $f(p)$ differenzierbar, so ist $g \circ f : U \to \mathbb{R}^L$ differenzierbar in p, und es gilt*
$$D(g \circ f)(p) = Dg(f(p))Df(p).$$
Insbesondere gilt
$$D(g \circ f)_{ij}(p) = \partial_j (g \circ f)_i = \sum_{k=1}^{M} \partial_k g_i(f(p)) \partial_j f_k(p).$$

Bemerkung. Die Ableitung der Komposition ist also die Komposition der Ableitungen. Diese Komposition wird auf der Ebene von Matrizen durch das Produkt der Matrizen gegeben. Die Reihenfolge der Matrizen ist dabei natürlich wichtig.

Beweis. Es reicht die erste Aussage zu zeigen. Das 'Insbesondere' folgt dann durch Einsetzen. Mit $q = f(p)$ gilt nach Voraussetzung:
$$f(x) = f(p) + Df(p)(x-p) + \varphi_f(x)$$
mit $\frac{\varphi_f(x)}{|x-p|} \to 0$, $x \to p$ und
$$g(y) = g(q) + Dg(q)(y-q) + \varphi_g(x)$$

mit $\frac{\varphi_g(y)}{|y-q|} \to 0,\ y \to q$

also gilt für $h(x) = g\,(f(x))$ mit $y = f(x)$ und weiterhin $f(p) = q$ dann

$$\begin{aligned} h(x) &= \underbrace{g(f(p))} + Dg(f(p))(f(x) - f(p)) + \varphi_g(f(x)) \\ &= h(p) + Dg(f(p))(Df(p)(x-p) + \varphi_f(x)) + \varphi_g(f(x)) \\ &= h(p) + Dg(f(p))Df(p)(x-p) + Dg(f(p))\varphi_f(x) + \varphi_g(f(x)) \\ &= h(p) + Dg(f(p))Df(p)(x-p) + \varphi(x) \end{aligned}$$

mit

$$\varphi(x) = Dg(f(p))\varphi_f(x) + \varphi_g(f(x)) =: T(x) + S(x).$$

Es bleibt zu zeigen, dass gilt

$$\frac{\varphi(x)}{|x-p|} \to 0, x \to p.$$

Das machen wir nun: Es gilt

$$\frac{T(x)}{|x-p|} = \frac{1}{|x-p|} Dg(f(p))\varphi_f(x) = \underbrace{Dg(f(p))}_{beschr\ddot{a}nkt} \underbrace{\frac{\varphi_f(x)}{|x-p|}}_{\to 0} \to 0.$$

Fuer den zweiten Term gilt

$$\frac{S(x)}{|x-p|} = \frac{1}{|x-p|}\varphi_g(f(x)) = \frac{\varphi_g(f(x))}{|f(x)-f(p)|} \frac{|f(x)-f(p)|}{|x-p|}$$

falls $f(x) \neq f(p)$. (Falls $f(x) = f(p)$ gilt $\varphi_g(f(x)) = \varphi_g(f(p)) = 0$ und es ist nichts zu zeigen.) Nun gilt aber

$$\frac{\varphi_g(f(x))}{|f(x)-f(p)|} \to 0, x \to p \ \ (\text{da } f(x) \to f(p) \text{ für } x \to p)$$

und es bleibt

$$\frac{|f(x)-f(p)|}{|x-p|} \leq \frac{|Df(p)(x-p) + \varphi_f(x)|}{|x-p|} \leq \frac{|Df(p)(x-p)|}{|x-p|} + \frac{|\varphi_f(x)|}{|x-p|}$$

beschränkt für $x \to p$. Damit folgt

$$\frac{|T(x) + S(x)|}{|x-p|} \to 0, x \to p.$$

Das beendet den Beweis. □

Beispiel - Verknuepfung von Funktion mit Kurve. Sei $U \subset \mathbb{R}^N$ offen und $\gamma : (a,b) \longrightarrow U$ differenzierbar in t_0 und $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}$ differenzierbar in $\gamma(t_0)$. Dann ist $f \circ \gamma$ differenzierbar und es gilt

$$(f \circ \gamma)'(t) = \sum_{k=1}^{N} \partial_k f(\gamma(t_0))\gamma_k'(t_0) = \langle \nabla f(\gamma(t_0)), \gamma'(t_0)\rangle.$$

Fuer $\gamma : [a, b] \longrightarrow U$ gilt eine entsprechende Formel auch in a bzw. b, wenn γ in a bzw. b differenzierbar ist (s.o.).

Beispiel. $\gamma : (a, b) \longrightarrow \mathbb{R}^{N+1}$, $\gamma(t) = (\gamma_1(t), \ldots, \gamma_N(t), t)$.
$F : \mathbb{R}^{N+1} \longrightarrow \mathbb{R}$, $(x, t) \mapsto F(x, t)$, differenzierbar. Dann gilt

$$(F \circ \gamma)'(t) = \sum_{k=1}^{N} \partial_k F(\gamma(t_0))\gamma_k'(t_0) + \partial_{N+1} F(\gamma(t_0)).$$

(Das wird manchmal unter dem Begriff der totalen Ableitung von F gefasst.)

Beispiel - $g(|x|)$. $g : \mathbb{R} \longrightarrow \mathbb{R}$ differenzierbar, $f : \mathbb{R}^N \longrightarrow \mathbb{R}$, $f(x) = g(|x|)$. Dann ist f in $x \neq 0$ differenzierbar, und es gilt $Df(p) = g'(|p|)\frac{1}{|p|}p$. (Hier haben wir die weiter unten bewiesene Differenzierbarkeit der Betragsfunktion in $p \neq 0$ benutzt.)

Theorem. *(Produktregel) Sei $U \subset \mathbb{R}^N$ offen und $f, g : U \longrightarrow \mathbb{R}$ differenzierbar in $p \in U$. Dann ist auch $fg : U \longrightarrow \mathbb{R}$ differenzierbar in p, und es gilt*

$$D(fg)(p) : \mathbb{R}^N \longrightarrow \mathbb{R},\ y \mapsto g(p)\langle \nabla f(p), y\rangle + f(p)\langle \nabla g(p), y\rangle.$$

Insbesondere ist also

$$\partial_j(gf)(p) = f(p)(\partial_j g)(p) + g(p)\partial_j f(p)$$

fuer alle $j = 1, \ldots, N$.

Beweis. Es reicht die erste Aussage zu beweisen. Die 'Insbesondere'-Aussage folgt dann durch Einsetzen von e_j.

Nach Voraussetzung gilt

$$f(x) = f(p) + Df(p)(x - p) + \varphi_f(x)$$

mit $\varphi_f(x)/|x - p| \to 0$, $x \to p$, und

$$g(x) = g(p) + Dg(p)(x - p) + \varphi_g(x)$$

mit $\varphi_g(x)/|x - p| \to 0$, $x \to p$. Damit folgt

$$g(x)f(x) = g(p)f(p) + (Dg(p)(x - p))f(p) + g(p)Df(p)(x - p) + \varphi,$$

wobei φ aus Termen besteht, die mindestens den Faktor $\varphi_f(x)$ bzw. $\varphi_g(x)$ oder zweimal den Faktor $(x - p)$ enthalten. Also gilt

$$\frac{\varphi(x)}{|x - p|} \to 0, x \to p$$

und die Aussage folgt. $\square$

Wir gehen nun noch auf eine geometrische Deutung des Gradienten ein. Dazu benoetigen wir noch ein Konzept.

DEFINITION (Richtungsableitung). *Sei $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}$ gegeben. Sei $p \in U$ und e ein beliebiger Einheitsvektor in $\mathbb{R}^N$. Sei $\delta > 0$ mit $p + te \in U$ fuer $t \in (-\delta, \delta)$. Sei $g : (-\delta, \delta) \longrightarrow \mathbb{R}$, $g(t) = f(p + te)$. Ist g differenzierbar in 0, so heisst der Grenzwert*

$$\partial_e f(p) := g'(0) = \lim_{t \to 0} \frac{1}{t}(g(t) - g(0)) = \lim_{t \to 0} \frac{1}{t}(f(p + te) - f(p))$$

die Richtungsableitung von f in p in Richtung e. ***Zeichnung.***

LEMMA (Geometrische Deutung Gradient). *Sei $f : U \longrightarrow \mathbb{R}$ differenzierbar in p. Dann gilt:*

- *Es ist $\nabla f(p)$ senkrecht auf der Konstanzflaeche $M := \{y : f(y) = f(p)\}$ (d.h. ist $\gamma : I \longrightarrow M$ stetig differenzierbar, mit $\gamma(0) = p$, so gilt $\langle \gamma'(0), \nabla f(p) \rangle = 0$.* ***Zeichnung.*** *(siehe unten)*
- *Es gilt $\partial_e f(p) = \langle \nabla f(p), e \rangle$ für alle $e \in \mathbb{R}^N$.*
- *Ist $\nabla f(p) \neq 0$, so zeigt $\nabla f(p)$ in die Richtung in der f am staerksten waechst (d.h. die Richtungsableitung von f in der Richtung von $\nabla f(p)$ ist maximal unter den Richtungsableitungen).*

$\longleftrightarrow$ Ende der Vorlesung

Bemerkung. Da f differenzierbar ist, gilt

$$f(q) = f(p) + t\langle \nabla F(p), q - p \rangle + \text{kleiner Fehler}.$$

Fuer die Funktion $q \mapsto f(p) + \langle \nabla f(p), (1 - p) \langle$ sind die entsprechenden Aussagen des Lemma klar.

Beweis. Nach der Kettenregel gilt für jedes differenzierbare $\gamma : (-\delta, \delta) \longrightarrow U$ mit $\gamma(0) = p$

$$\frac{d}{dt}(f \circ \gamma)(0) = \langle \nabla f(p), \gamma'(0) \rangle.$$

Mit $\gamma(t) = p + te$ folgt also

$$\partial_e f(p) = \frac{d}{dt} f \circ \gamma(t_0) = \langle \nabla f(p), e \rangle.$$

Das beweist den zweiten Punkt.

Ist andererseits γ eine Kurve mit Werten in M so gilt also $f \circ \gamma = constant$ und es folgt (mit obigem)

$$0 = \frac{d}{dt}(f \circ \gamma)(t) = \langle \nabla f(p), \gamma'(0) \rangle.$$

Das beweist den ersten Punkt.

Für jeden Einheitsvektor e gilt

$$|\partial_e f(p)| = |\langle \nabla f(p), e \rangle| \leq |\nabla f(p)|$$

und für $e = \frac{1}{|\nabla f(p)|}\nabla f(p)$ gilt dann

$$|\partial_e f(p)| = |\langle \nabla f(p), \frac{1}{|\nabla f(p)|}\nabla f(p)\rangle| = \frac{1}{|\nabla f(p)|}|\nabla f(p)|^2 = |\nabla f(p)|.$$

Damit folgt die letzte Aussage. □

Beispiel. Sei $f : \mathbb{R}^N \setminus \{0\}$, $f(x) = |x|$. Dann gilt $Df(x) = \nabla f(x) = \frac{1}{|x|}x$. Das ist in der Tat der Einheitsvektor in Richtung des staerksten Anstieges von f. **Zeichnung.** (Konstanzflaechen sind Sphaeren).

Bisher haben wir unter der Annahme der Differenzierbarkeit von f gearbeitet. Hier lernen wir nun ein (d a s) Kriterium kennen, das die Differenzierbarkeit von f liefert.

THEOREM. *(Stetigkeit der partiellen Ableitung impliziert Differenzierbarkeit) Sei $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}^M$ gegeben. Existieren die partiellen Ableitungen*

$$\partial_j f_i(q) = \frac{\partial f_i}{\partial x_j}(q) = \lim_{h\to 0, h\neq 0} \frac{1}{h}(f_i(q + he_j) - f_i(q))$$

fuer alle $q \in U$, $i = 1, \ldots, M$, $j = 1, \ldots, N$ und sind stetig in $p \in U$, so ist f differenzierbar in p.

Bemerkung. In der Situation des Theorem ist die Ableitung natuerlich durch die Jacobimatrix / Differentialmatrix $\left(\frac{\partial f_i}{\partial x_j}(p)\right)_{i,j}$ gegeben.

Beweis. Sei L die durch die Matrix $\left(\frac{\partial f_i}{\partial x_j}(p)\right)_{i,j}$ induzierte Abbildung. Zu zeigen:

$$\frac{1}{|x-p|}|f(x) - f(p) - L(x-p)| \to 0, x \to p.$$

Wir zeigen: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit

$$\frac{1}{|x-p|}|f(x) - f(p) - L(x-p)| \leq \varepsilon$$

für alle $|x - p| \leq \delta$. Sei also $\varepsilon > 0$ gegeben.

Ziel. Wir setzen

$$t = (\tau_1, \ldots, \tau_N) := x - p.$$

Fuer $i = 1, \ldots, M$ ist dann die i-te Komponente von $f(x) - f(p) - L(x-p)$ gegeben durch

$$f_i(x) - f_i(p) - \sum_{j=1}^{N} \partial_j f_i(p)\tau_j.$$

Es gilt also diesen Term abzuschaetzen.

Aufgrund der Stetigkeit der $\partial_j f_i$ existiert ein $\delta > 0$ mit

$$|\partial_j f_i(x) - \partial_j f_i(p)| \leq \frac{\varepsilon}{\sqrt{M} \cdot N} \quad (*)$$

für alle $x \in U$ mit $|x - p| < \delta$. Sei nun $x \in U_\delta(p)$ und

$$t = (\tau_1, \ldots, \tau_N) := x - p$$

und

$$t_j := (\tau_1, \ldots, \tau_j, 0, \ldots, 0) = \sum_{l=1}^{j} \tau_l e_l, \; t_0 = 0$$

für $j = 0, \ldots, N$. Insbesondere ist also $t_N = t = x - p$. Dann gilt (Teleskopsumme)

$$f_i(x) - f_i(p) = \sum_{j=1}^{N} \left(f_i(p + t_j) - f_i(p + t_{j-1})\right). \; (T)$$

Zeichnung. Nach dem 1-dimensionalen Mittelwertsatz der Differentialrechnung angewendet auf die Funktion

$$s \mapsto f_i(x + t_{j-1} + se_j)$$

exisistieren $\sigma_{j,i}$ zwischen 0 und τ_j mit

$$f_i(p+t_j)-f_i(p+t_{j-1}) = f_i(p+t_{j-1}+\tau_j e_j)-f_i(p+t_{j-1}) = \tau_j \partial_j f_i(p+t_{j-1}+\sigma_{j,i} e_j)$$

für $j = 1, \ldots, N$, $i = 1, \ldots, M$. Mit $(*)$ folgt dann

$$|f_i(p+t_j)-f_i(p+t_{j-1})-\tau_j \partial_j f_i(p)| \leq |\tau_j||\partial_j f_i(p+t_{j-1}+\sigma_{j,i} e_j)-\partial_j f_i(p)| \overset{(*)}{\leq} |x-p|\frac{\varepsilon}{\sqrt{M}N}.$$

Mit (T) und dieser Abschaetzung folgt

$$|f_i(x)-f_i(p)-\sum_{j=1}^{N} \partial_j f_i(p)\tau_j| \leq \sum_{j=1}^{N} |f_i(p+t_j)-f_i(p+t_{j-1})-\tau_j \partial_j f_i(p)| \leq |x-p|\frac{\varepsilon}{\sqrt{M}}.$$

Damit ergibt sich für x mit $|x - p| < \delta$

$$\begin{aligned} |f(x) - f(p) - L(x-p)|^2 &= \sum_{i=1}^{M} |f_i(x) - f_i(p) - \sum_{j=1}^{N} \partial_j f_i(p)\tau_j|^2 \\ &\leq \sum_{i=1}^{M} \left| |x-p| \frac{\varepsilon}{\sqrt{M}} \right|^2 \\ &\leq |x-p|^2 \varepsilon^2. \end{aligned}$$

Damit folgt

$$\frac{1}{|x-p|}|f(x) - f(p) - L(x-p)| \leq \varepsilon.$$

Das beendet den Beweis. $\square$

Beispiel. Die Funktion $f : \mathbb{R}^N \setminus \{0\} \longrightarrow \mathbb{R}$, $x \mapsto |x|$ ist differenzierbar.

Definition (Stetig differenzierbar). *Sei $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}^M$ gegeben. Existieren für jedes $p \in U$ die partiellen Ableitungen*

$$\partial_j f_i(p) = \frac{\partial f_i}{\partial x_j}(p) := \lim_{h \to 0, h \neq 0} \frac{1}{h}(f_i(p + he_j) - f_i(p))$$

fuer $i = 1, \ldots, M$, $j = 1, \ldots, N$ und sind stetig in $p \in U$, so heisst f stetig differenzierbar (auf U).

Bemerkungen. Nach dem vorigen Theorem ist jede stetig differenzierbare Funktion auch differenzierbar (wie es ja auch - rein sprachlich - sein sollte ;-)

Wir schließen diesen Abschnitt mit einer Diskussion von Mittelwertsätzen ab.

Theorem. *(Mittelwertsatz) Sei $U \subset \mathbb{R}^N$ offen, $f : U \to \mathbb{R}$ differenzierbar, und seien $x, y \in U$ mit $x + t(y - x) \in U$ für $t \in [0, 1]$ gegeben. Dann existiert ein $\vartheta \in (0, 1)$, sodass*

$$f(y) - f(x) = Df(x + \vartheta(y - x))(y - x).$$

Bemerkung.

- Ist U konvex, so gilt $x + t(y - x) \in U$ fuer alle $x, y \in U$ (und umgekehrt). Da alle Kugeln konvex sind, gilt der obige Satz also insbesondere lokal (d.h. jedes $x \in U$ hat eine Umgebung $U_r(x)$, in der der Satz gilt).
- Es ist wesentlich (siehe Uebung), dass der Wertebereich von f im eindimensionalen $\mathbb{R}$ enthalten ist.

Beweis. Sei $\gamma : [0, 1] \to U$, $\gamma(t) = x + t(y - x)$ und

$$h : [0, 1] \to \mathbb{R},\ h = f \circ \gamma.$$

Dann ist h stetig auf $[0, 1]$ und nach Kettenregel differenzierbar auf $(0, 1)$ mit

$$h'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle .$$

Der eindimensionale Mittelwertsatz liefert dann

$$\begin{array}{rcl} f(y) - f(x) & = & h(1) - h(0) \\ (1 - dim\ MWS) & = & h'(\vartheta)(1 - 0) \\ (\gamma'(t) = y - x) & = & \langle \nabla f(x + \vartheta(y - x)), y - x \rangle \\ & = & Df(x + \vartheta(y - x))(y - x). \end{array}$$

□

Der Satz hat Konsequenzen für Funktionen $f : U \longrightarrow \mathbb{R}^M$.

Folgerung. *Sei $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}^M$ differenzierbar. Dann gilt:*

(a) *Gilt* $|\partial_k f_i(z)| \leq C$ *für alle* $z \in U$ *und* $1 \leq k \leq N$, $1 \leq i \leq M$, *so gilt*

$$|f(y) - f(x)| \leq C\sqrt{MN}|y - x|$$

für alle x, y, *deren Verbindungstrecke ganz in* U *liegt.*

(b) *Für alle* $x, y \in U$, *deren Verbindungstrecke ganz in* U *liegt existiert ein* ξ *der Form* $x + \vartheta(y - x)$ *fuer* $\vartheta \in (0, 1)$ *mit*

$$|f(y) - f(x)| \leq \|Df(\xi)\||y - x|.$$

(c) *Ist* U *zusammenhängend und* $Df(x) = 0$ *fuer alle* $x \in U$, *so folgt* $f \equiv$ *constant auf* U.

Erinnerung: $\|L\| := \sup\{|Lx| : |x| \leq 1\}$. Es gilt $|Lx| \leq \|L\||x|$.)

Beweis. (a) Nach dem vorangehenden Mittelwertsatz gibt es $\vartheta_i \in (0, 1)$, $i = 1, \ldots, d$ mit

$$|f_i(y)-f_i(x)| = |\langle \nabla f_i(x+\vartheta_i(y-x), (y-x)\rangle| \leq \sum_{j=1}^{N} C|y_j-x_j| \leq C\sqrt{N}|y-x|.$$

(Hier wurde Cauchy Schwartz in der letzten Abschätzung benutzt). Damit folgt dann sofort

$$|f(y) - f(x)| = \left(\sum_{i=1}^{M} 1|f_i(y) - f_i(x)|^2\right)^{1/2} \leq C\sqrt{NM}|y - x|.$$

Ende der Vorlesung.

(b) *Spezialfall.* Es gilt $f_i(x) = f_i(y)$ für alle $i \neq 1$. Dann folgt aus obigem Mittelwertsatz

$$|f(y)-f(x)| = |f_1(y)-f_1(x)| = |\langle \nabla f_1(x+\vartheta(y-x)), (y-x)\rangle| \leq \|Df(\xi)\||y-x|.$$

Allgemeiner Fall. Sei U eine Drehung, sodass $Uf(y)$ und $Uf(x)$ sich hoechstens in der ersten Komponente unterscheiden (drehe zunaechst so, dass die durch $0, f(x), f(y)$ gebildete Ebene gerade die durch die ersten beiden Koordinatenachsen gegebene Ebene ist. Drehe anschliessend weiter.) Dann gilt mit $f^* := Uf$ nach dem Spezialfall

$$|f(y)-f(x)| = |f^*(y)-f^*(x)| \leq \|Df^*(\xi)\||y-x| = \|UDf(\xi)\||y-x| = \|Df(\xi)\||y-x|.$$

(c) Da U offen und zusammenhaengend ist, ist es wegzusammenhaengend. Sei nun $\gamma : [0, 1] \longrightarrow U$ mit $\gamma(0) = x$, $\gamma(1) = y$ gegeben. Da U offen ist und γ stetig, können wir ohne Einschränkung annehmen, dass γ stueckweise linear ist, d.h. es gibt $p = x, x_1, \ldots, x_n = y \in U$, so dass γ durch die Verbindungsstrecken zwischen x_i und x_{i+1} beschrieben ist . Dann gilt nach (a) oder (b) $f(x_i) = f(x_{i+1})$. Damit folgt die Behauptung. $\square$

3. Höhere Ableitungen und der Satz von Taylor

In diesem Abschnitt untersuchen wir, wie Funktionen gut durch Polynome approximiert werden koennen, wenn sie genuegend oft differenzierbar sind.

Wie schon bekannt, heisst f stetig differenzierbar, wenn die partiellen Ableitungen aller f_i, $i = 1, \ldots, m$ existieren und stetig sind. Wir koennen das Iterieren: Ist f einmal stetig differenzierbar und existieren alle ersten Ableitungen der partiellen Ableitungen $D_j f_i$ und sind stetig, so heißt f zweimal stetig differenzierbar und man nennt die $D_i D_j f_k$ die zweiten partiellen Ableitungen von f. Induktiv definiert man dann die $(q+1)$-ten partiellen Ableitungen von f als die ersten Ableitungen der p-ten Ableitungen $D_{i_1} \ldots D_{i_q} f$ (falls existent).

Definition. *Sei $U \subset \mathbb{R}^N$ offen. Die Funktion $f : U \longrightarrow \mathbb{R}^M$ heisst q mal stetig differenzierbar, wenn f $q-1$ mal stetig differenzierbar ist und die $(q-1)$-ten partiellen Ableitungen $D_{i_{q-1}} \ldots D_{i_1} f$, $i_1, \ldots, i_{q-1} \in \{1, \ldots, N\}$, existieren und stetig sind. In diesem Fall nennt man die entstehenden partiellen Ableitungen $D_{i_q}(D_{i_{q-1}} \ldots D_{i_1} f)$ die q-ten partiellen Ableitungen von f. Der Vektorraum der q-mal stetig differenzierbaren Funktionen wird mit $C^q(U, \mathbb{R}^M)$ bezeichnet. Die Elemente aus*

$$C^\infty(U, \mathbb{R}^M) := \bigcap_{q=1}^{\infty} C^q(U, \mathbb{R}^M)$$

heißen beliebig oft differenzierbare Funktionen.

Es zeigt sich, dass die Reihenfolge der partiellen Ableitungen keine Rolle spielt (wenn f genuegend oft s t e t i g differenzierbar ist).

Theorem. *(Satz von Schwarz) Sei $U \subset \mathbb{R}^N$ und $f : U \to \mathbb{R}^M$. Ist f zweimal stetig differenzierbar, so gilt fuer jedes $p \in U$, $k \in \{1, \ldots, M\}$ und $i, j \in \{1, \ldots, N\}$*

$$D_i D_j f_k(p) = D_j D_i f_k(p).$$

Beweis. Wir können die Situation ohne Einschraenkung etwas vereinfachen:

- Es genügt eine Komponente $k \in \{1, \ldots, n\}$ zu betrachten, d.h. ohne Einschraenkung setzen wir $M = 1$ voraus.
- Für $i = j$ gibt es nichts zu beweisen
- Wir setzen ohne Einschraenkung voraus $N = 2$ und setzen $i = 1$, $j = 2$ und $N = 2$. (Ansonsten umnummerieren.)
- Betrachte o. B. d. A. $p = 0$

Es ist somit zu zeigen: $D_1 D_2 f(0) = D_2 D_1 f(0)$.

Wir zeigen dies im wesentlichen, indem wir den entsprechenden Differenzenquotienten zweiter Ordnung auf zwei verschiedene Weisen ausrechnen und anschliessend einen Grenzuebergang durchfuehren.

Seien $t, s > 0$ beliebig. Wir betrachten nun

$\Delta(s,t) := (f(s,t) - f(s,0)) - (f(0,t) - f(0,0)) = (f(s,t) - f(0,t)) - (f(s,0) - f(0,0)).$

Dann gilt fuer diesen Ausdruck also

$$\begin{aligned}
\Delta(s,t) &= (f(s,t) - f(s,0)) - (f(0,t) - f(0,0)) \\
(g(r) := f(r,t) - f(r,0)) \quad &= g(s) - g(0) \\
(MWS) \quad &= g'(s_1)(s-0) \qquad \text{für } s_1 \in (0,s) \\
(g^{'} = ...) \quad &= s(D_1 f(s_1,t) - D_1 f(s_1,0)) \\
(h(r) := D_1 f(s_1,r)) \quad &= s(h(t) - h(0)) \\
&= st h'(t_1) \qquad \text{mit } t_1 \in (0,t) \\
&= st D_2 D_1 f(s_1,t_1).
\end{aligned}$$

Die andere Darstellung von Δ liefert nach analoger Rechnung dann

$$\begin{aligned}
\Delta(s,t) &= (f(s,t) - f(0,t)) - (f(s,0) - f(0,0)) \\
&= \dots \\
&= st D_1 D_2 f(s_2,t_2) \qquad \text{mit } s_2 \in (0,s),\ t_2 \in (0,t).
\end{aligned}$$

Fuer $s, t \neq 0$ kann man dividieren und erhaelt

$$D_2 D_2 f(s_1,t_1) = D_1 D_2 f(s_2,t_2).$$

Betrachtet man nun den Grenzwert $(s,t) \to (0,0)$ so folgt aus der Stetigkeit der zweiten Ableitungen also

$$D_1 D_2 f(0,0) = D_2 D_1 f(0,0).$$

Das beendet den Beweis. $\square$

FOLGERUNG. *(Vertauschbarkeit der partiellen Ableitungen) Sei $U \subset \mathbb{R}^N$ offen und $f : U \to \mathbb{R}^M$ p-mal stetig differenzierbar, so können in*

$$D_{j_1}, \dots, D_{j_p} f$$

die partiellen Ableitungen beliebig vertauscht werden.

Beweis. Die Behauptung folgt aus dem Satz von Schwarz und Induktion. $\square$

Fuer die weiteren Betrachtungen brauchen wir nun etwas Notation. Dabei geht es im wesentlichen um gute 'Buchhaltung':
Für $U \subset \mathbb{R}^N$ offen und $f : U \longrightarrow \mathbb{R}^M$ q-mal stetig differenzierbar. Dann definiert man fuer $r \leq q$ und $j = 1, \dots, N$

$$D_j^r f = \underbrace{D_j \dots D_j}_{r\text{-mal}} f \qquad (D_j^0 f := f)$$

Für $(j_1, \ldots, j_q) \in \{1, \ldots, N\}^q$ bezeichne α_i die Anzahl des Auftretens der Zahl i unter den j_k. Dann gilt nach dem Satz von Schwarz bzw. seinem Korollar

$$D_{j_1} \ldots D_{j_q} f = D_1^{\alpha_1} \ldots D_N^{\alpha_N} f.$$

Für *Multiindizes* $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}_0^N$ sei im folgenden

$$|\alpha| := \sum_{j=1}^{N} \alpha_j \qquad \text{der Betrag/die Länge von } \alpha,$$

$$\alpha! := \prod_{j=1}^{N} \alpha_j! \qquad \text{die Fakultät von } \alpha,$$

$$x^\alpha := \xi_1^{\alpha_1} \cdot \ldots \cdot \xi_N^{\alpha_N} = \prod_{j=1}^{N} \xi_j^{\alpha_j} \qquad \text{mit } x = (\xi_1, \ldots, \xi_N) \in \mathbb{R}^N$$

Sei weiterhin fuer ein q mal stetige differenziebares f und einen Multiindex α mit $|\alpha| \leq q$

$$D^\alpha f := D_1^{\alpha_1} \ldots D_N^{\alpha_N} f = (\prod_{j=1}^{N} D_j^{\alpha_j}) f.$$

Anwendung. (Uebung) Die Multiindexschreibweise kann anhand des *Polynomischen Satzes* verdeutlicht werden (auf einen Beweis wird an dieser Stelle verzichtet):

$$\left(\sum_{j=1}^{N} \xi_j\right)^k = \sum_{\alpha:|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

Wendet man diese Formel zum Beispiel auf $(a+b)^3$ an, so ergibt sich:

$$\begin{aligned}(a+b)^3 &= \sum_{\alpha:|\alpha|=3} \frac{3!}{\alpha!} x^\alpha = \frac{3!}{(3,0)!} x^{(3,0)} + \frac{3!}{(2,1)!} x^{(2,1)} + \frac{3!}{(1,2)!} x^{(1,2)} + \frac{3!}{(0,3)!} x^{(0,3)} \\ &= \frac{3!}{3!} a^3 b^0 + \frac{3!}{2!} a^2 b^1 + \frac{3!}{2!} a^1 b^2 + \frac{3!}{3!} a^0 b^3 = a^3 + 3a^2 b + 3ab^2 + b^3\end{aligned}$$

Mit den vorigen Resultaten und den obigen Beziehungen sind wir nun in der Lage den Taylor'schen Satz auf Funktionen von N Variablen zu übertragen.

THEOREM (Taylor). *Sei $U \subset \mathbb{R}^N$ offen, $q \geq 1$, $f \in C^{q+1}(U, \mathbb{R})$ und $p \in U$. Dann gilt für alle $x \in U$, für die die Verbindungsstrecke von p und x in U liegt:*

$$f(x) = \sum_{|\alpha| \leq q} \frac{1}{\alpha!} D^\alpha f(p)(x-p)^\alpha + R_q(x)$$

mit

$$R_q(x) = \sum_{|\alpha|=q+1} \frac{1}{\alpha!} D^\alpha f(p+\vartheta(x-p))(x-p)^\alpha \qquad \textit{mit } \vartheta \in (0,1)$$

Bemerkung / Verdeutlichung. Zur Verdeutlichung geben wir dieses Resultat für $q = 0, 1, 2$ explizit an. Sei dazu $x = (\xi_1, \dots, \xi_N)$ und $p = (p_1, \dots, p_N)$. Dann gilt:

$$\begin{aligned}
\underline{q=0}: \qquad f(x) &= f(p) + \sum_{j=1}^N D_j f(p+\vartheta(x-p))(\xi_j - p_j) \\
&= f(p) + \langle \nabla f(p+\vartheta(x-p)), x-p \rangle \\
\underline{q=1}: \qquad f(x) &= f(p) + \sum_{j=1}^N D_j f(p)(\xi_j - p_j) + \frac{1}{2} \sum_{j,k=1}^N D_j D_k f(p+\vartheta(x-p))(\xi_j - p_j)(\xi_k - p_k) \\
\underline{q=2}: \qquad f(x) &= f(p) + ... \\
&= f(p) + \langle \nabla f(p), (x-p) \rangle + \frac{1}{2} \langle H(p)(x-p), (x-p) \rangle + R_2(x)
\end{aligned}$$

mit der *Hesse-Matrix*

$$H(f)(\cdot) = H(\cdot) = (D_j D_k f(\cdot))_{j,k=1}^N$$

Beweis. Bonusmaterial: Der Satz wird bewiesen durch Rückführung auf den eindimensionalen Fall. Dazu sei

$$\begin{gathered}
y = (\eta_1, \dots, \eta_m) := (x-p) \\
g: \ [-\varepsilon, 1+\varepsilon] \to \mathbb{R} \\
g(t) = f(p+ty) = (f \circ h)(t), \text{ mit } h(t) = p + ty
\end{gathered}$$

wobei $\varepsilon > 0$ so klein gewählt sei, dass das Gesamtstück

$$\{p + ty \mid t \in [-\varepsilon, 1+\varepsilon]\}$$

ganz in U liegt (möglich, da U offen).
Offensichtlich ist g stetig differenzierbar und mit der Kettenregel folgt

$$\begin{aligned}
g'(t) &= Df(h(t))h'(t) \\
&= \sum_{j=1}^N D_j f(h(t)) \underbrace{h_j'(t)}_{\eta_j} \\
&= \sum_{j=1}^N \eta_j D_j f(p+ty)
\end{aligned}$$

Ist $q \geq 1$, so ist auch g' stetig differenzierbar und es gilt:

$$
\begin{aligned}
g''(t) &= \sum_{j=1}^{N} \eta_j \frac{d}{dt} D_j f(p+ty) \\
&= \sum_{j=1}^{N} \eta_j \sum_{i=1}^{N} D_i D_j f(p+ty)\eta_i \\
&= \left(\sum_{j=1}^{N} \eta_j D_j \right)^2 f(p+ty)
\end{aligned}
$$

wobei der Term $(\sum \eta_j D_j)^2$ formal auszumultiplizieren ist. Allgemeiner erhlt man durch induktive Anwendung des Verfahrens für $k \leq q+1$

$$
g^{(k)}(t) = \left(\sum_{j=1}^{N} \eta_j D_j \right)^2 f(p+ty)
$$

wobei auch hier die k-te Potenz formal auszurechnen ist.
Nach dem Polynomischen Satz (siehe oben) gilt für beliebige $z = (\zeta_1, \ldots, \zeta_N) \in \mathbb{R}^N$ und $k \in \mathbb{N}$

$$
\left(\sum_{j=1}^{N} \zeta_j \right)^k = \sum_{\alpha : |\alpha| = k} \frac{k!}{\alpha!} z^\alpha
$$

In dieser Formel ersetzen wir $(\zeta_1, \ldots, \zeta_N)$ durch $(\eta_1 D_1, \ldots, \eta_N D_N)$. Damit erhalten wir

$$
\begin{aligned}
\frac{1}{k!} g^{(k)}(t) &= \frac{1}{k!} \left(\sum_{j=1}^{N} \eta_j D_j \right)^k f(p+ty) \\
&= \left(\sum_{|\alpha|=k} \frac{1}{\alpha!} (\eta_1 D_1, \ldots, \eta_N D_N)^\alpha \right) f(p+ty) \\
&= \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} D^\alpha f(p+ty)
\end{aligned}
$$

Die Taylorsche Formel liefert für $t \in [0,1]$

$$
g(t) = \sum_{k=1}^{q} \frac{1}{k!} g^{(k)}(0) t^k + \widetilde{R_q}(t)
$$

mit $\widetilde{R_q}(t) = \frac{1}{(q+1)!} t^q g^{(q)}(\vartheta t)$ mit $0 < \vartheta < 1$. Damit erhalten wir:

$$\begin{aligned} f(x) = g(1) &= \sum_{k=0}^{q} \frac{1}{k!} g^{(k)}(0) + \widetilde{R_q}(1) \\ &= \sum_{k=0}^{q} \sum_{|\alpha|=k} \frac{(x-p)^\alpha}{\alpha!} D^\alpha f(p) + \widetilde{R_q}(1) \\ &= \sum_{|\alpha| \le q} \frac{1}{\alpha!} D^\alpha f(p)(x-p)^\alpha + R_q(x) \end{aligned}$$

mit

$$\begin{aligned} R_q(x) = \widetilde{R_q}(1) &= \frac{1}{(q+1)!} g^{(q+1)}(\vartheta) \\ &= \sum_{|\alpha|=q+1} \frac{y^\alpha}{\alpha!} D^\alpha f(p + \vartheta y) \\ &= \sum_{|\alpha|=q+1} \frac{1}{\alpha!} D^\alpha f(p + \vartheta(x-p))(x-p)^\alpha \end{aligned}$$

Das liefert die gewuenschte Aussage. □

4. Extrema von Funktionen

Wie im eindimensionalen Fall werden wir Kriterien für (lokale) Extrema von Funktionen mehrerer Veränderlicher (im Inneren eines Gebietes U) angeben.

Definition (Extrema im mehrdimensionalen Fall). *Sei $U \subset \mathbb{R}^N$ gegeben, $f : U \to \mathbb{R}$ und $p \in U$. Dann hat f in p ein lokales $\substack{\text{Maximum} \\ \text{Minimum}}$, falls ein $\delta > 0$ existiert, sodass gilt*

$$f(x) \substack{\leq \\ \geq} f(p)$$

fuer alle $x \in U$ mit $|x - p| < \delta$. Es hat f h in p in strenges lokales $\substack{\text{Maximum} \\ \text{Minimum}}$, falls ein $\delta > 0$ existiert, sodass gilt

$$f(x) \substack{< \\ >} f(p)$$

fuer alle $x \in U$ mit $0 < |x - p| < \delta$. Ein $p \in U$ ist ein Extremum von f, wenn es ein Minimum oder ein Maximum ist.

Bemerkung. Gelten die entsprechenden Ungleichunen fuer alle $x \in U$, so handelt es sich um globale Extrema.

Theorem. *(Notwendige Kriterien für Extrema) Sei $U \subset \mathbb{R}^N$ offen, $p \in U$ und die Funktion $f : U \to \mathbb{R}$ sei differenzierbar in p. Hat f in p ein lokales Maximum bzw. Minimum, so gilt*

$$\nabla f(p) \equiv 0$$

Beweis. Da f bei p differenzierbar ist, sind die Funktionen

$$g_j : \ (-\varepsilon, \varepsilon) \to \mathbb{R}, \ g_j(t) = f(p + te_j)$$

in $t = 0$ differenzierbar fuer jedes $j = 1, \dots, N$. Außerdem haben sie bei 0 ein Extremum. Damit folgt

$$g_j'(0) = 0 = \frac{\partial}{\partial x_j} f(p).$$

Das beendet den Beweis. □

Bemerkung. Ist $M \subset \mathbb{R}^N$ beliebig und $f : M \longrightarrow \mathbb{R}$ differenzierbar im Inneren von M, so liefet dieses Theorem eine notwendige Bedingung fuer Extremwerte im Inneren von M. Das Verhalten von f auf dem Rand muss dann natuerlich noch getrennt untersucht werden.

Ende der Vorlesung

Erinnerung. Eine symmetrische $N \times N$ Matrix A heißt positiv definit, falls für alle $x \in \mathbb{R}^N \setminus \{0\}$ gilt

$$\langle Ax, x\rangle > 0.$$

Sie heisst positiv semidefinit, wenn fuer alle $x \in \mathbb{R}^N \setminus \{0\}$ gilt

$$\langle Ax, x\rangle \geq 0.$$

Durch Diagonalisieren von A sieht man leicht, dass gilt

$$\langle Ax, x\rangle = \lambda_m \|x\|^2$$

mit $\lambda_m :=$ kleinste Eigenwert von A, wobei Gleichheit gilt fuer x aus dem Eigenraum zu λ_m. Damit folgt, dass A positiv (semi)definit istk, genau dann, wenn die Eigenwerte von A strikt positiv bzw. nichtnegativ sind. Entsprechend definiert man negative (semi)definite Matrizen.

Theorem. *(Kriterien für Extrema) Sei $U \subset \mathbb{R}^N$ offen, $f \in \mathfrak{C}^2(U, \mathbb{R})$ und $p \in U$. Dann gilt:*

- *(Notwendiges Kriterium) Hat f bei p ein (lokales) $\substack{\text{Maximum} \\ \text{Minimum}}$, so gilt $\nabla f(p) = 0$ und die Hesse-Matrix $(D_k D_j f(p))_{j,k=1}^N$ ist $\substack{\text{negativ} \\ \text{positiv}}$ semidefinit.*
- *(Hinreichendes Kriterium) Gilt $\nabla f(p) = 0$ und die Hesse-Matrix ist $\substack{\text{negativ} \\ \text{positiv}}$ definit in p, so hat f ein striktes lokales $\substack{\text{Maximum} \\ \text{Minimum}}$ in p.*

Bemerkung. Obiges Theorem liefert im eindimensionalen Fall gerade die uns schon bekannte Charakterisierung.

Beweis. **Idee.** Es gilt nach Taylor und dem vorigen Satz

$$f(x) = f(p) + \langle \nabla f(p), (x-p)\rangle + \frac{1}{2}\langle H(f)(\xi)(x-p), (x-p)\rangle = f(p) + \frac{1}{2}\langle H(f)(\xi)(x-p), (x-p)\rangle$$

mit einer Zwischenstelle ξ zwischen x und p. Damit wird also 'alles' durch das Verhalten von $H(f)(\xi)$ entschieden. Dieses Verhalten ist aber (im wesentlichen) das Verhalten von $H(f)(p)$.

Wir zeigen zunaechst die notwendige Bedingung: Nach obigen Satz gilt $\nabla f(p) = 0$. Angenommen $H(p)$ ist nicht negativ semidefinit, dann existiert ein $y \in \mathbb{R}^N$, $|y| = 1$ und $\langle H(p)y, y\rangle > 0$.
Da $f \in C^2(U, \mathbb{R})$ gilt, existiert ein $\delta > 0$ mit

$$\langle H(x)y, y\rangle > 0$$

fuer alle $x \in U$ mit $|x - p| < \delta$.
Dann folgt mit dem Satz von Taylor für alle $t \in (0, \delta]$

$$\begin{aligned} f(p + ty) &= f(p) + 0 + \sum_{|\alpha|=2} \frac{1}{\alpha!} D^\alpha f(p + \vartheta ty) t^{|\alpha|} y^\alpha \\ &= f(p) + \frac{1}{2} \sum_{i,j=1}^{N} t^2 D_i D_j f(p + \vartheta ty) \eta_i \eta_j \\ &= f(p) + \frac{1}{2} t^2 \langle H(p + \vartheta ty)y, y\rangle \\ &> f(p) \end{aligned}$$

Wir zeigen nun die hinreichende Bedinung:
Sei $Df(p) = 0$ und $(D_i D_j f(p))_{i,j}$ negativ definit, dann gilt also mit $y = (\eta_1, \ldots, \eta_m)$

$$0 > \sum_{i,j=1}^{N} D_i D_j f(p) \eta_i \eta_j = \langle H(f)(p)y, y\rangle$$

für alle y mit $y \neq 0$.
Da die Abbildung

$$S^N = \{y \in \mathbb{R}^N \ : \ |y| = 1\} \to \mathbb{R}, y \mapsto \langle H(f)(p)y, y\rangle$$

stetig ist, nimmt sie auf dem kompakten S^N ihr Maximum an. Es gibt also ein $\lambda > 0$ mit $\langle H(f)(p)y, y\rangle \leq -\lambda$ für alle y mit $|y| = 1$. Da alle $D_i D_j f$ stetig sind, existiert ein $\delta > 0$ mit

$$|D_i D_j f(x) - D_i D_j f(p)| \leq \frac{\lambda}{2N^2}$$

für alle $x \in U$ mit $|x - p| < \delta$. Damit folgt

$$\begin{aligned}
\langle H(f)(x)y,y\rangle &= \sum_{i,j=1}^{N} D_iD_jf(x)\eta_i\eta_j \\
&= \sum_{i,j=1}^{N} (D_iD_jf(x) - D_iD_jf(p))\eta_i\eta_j + \langle H(f)(p)y,y\rangle \\
&\leq \frac{\lambda}{2} - \lambda \qquad\qquad = -\frac{\lambda}{2}
\end{aligned}$$

für alle y mit $|y| = 1$. Damit folgt dann sofort

$$\langle H(f)(x)y,y\rangle \leq -\frac{\lambda}{2}|y|^2 < 0$$

für alle $|y| \neq 0$, und x mit $|x-p| < \delta$. Nachg dem Satz von Taylor gilt aber

$$f(x) = f(p) + Df(p)(x-p) + \frac{1}{2}\langle H(f)(p+\vartheta(x-p))(x-p),(x-p)\rangle$$

für ein $\vartheta \in (0,1)$. Mit der vorigen Abschaetzung und $Df(p) = 0$ ergibt sich dann sofort

$$f(x) - f(p) < 0$$

für $|x-p| < \delta$, $x \neq p$. Damit hat f also in p ein striktes Maximum. Die Aussage über Minima kann analog bewiesen werden. Alternativ kann man auch die schon bewiesene Aussage auf $-f$ anwenden. □

Bemerkung.

- Der Beweis nutzt: A nahe B $\Rightarrow$ $\langle Ay,y\rangle$ nahe $\langle By,y\rangle$
- Der Beweis zeigt: A negativ definit, größter Eigenwert ist $-\lambda$, B nahe A, B symmetrisch ;$\Rightarrow$ größter Eigenwert von B ist nahe $-\lambda$,

KAPITEL 15

Implizite Funktionen und lokale Invertierbarkeit

In diesem Abschnitt untersuchen wir (lokale) Loesbarkeit von Gleichungen der Form

$$F(x, y) = 0$$

nach x bzw y bei gegebenen y bzw. x.

Wir beginnen mit einer Diskussion von lokaler Invertierbarkeit. Zur Einstimmung betrachten wir zwei Spezialfalle.

Beispiel. Sei I ein Interval in $\mathbb{R}$ und $f : I \longrightarrow \mathbb{R}$ stetig. Dann hat f eine Umkehrfunktion genau dann wenn f streng monoton ist. Das ist eine gute Charakterisierung aber unter Umstaenden nicht einfach zu pruefen. Ist f stetig differenzierbar, so wird die Lage einfacher: Ist $f'(p) \neq 0$ so ist f lokal invertierbar. (Denn dann hat f' in der Naehe von p festes Vorzeichen und ist damit nach dem Mittelwertsatz streng monoton.) Die (lokale) Inverse g ist stetig diffbar mit

$$g'(y) = \frac{1}{f(x)}$$

mit $f(x) = y$.

Beispiel. Sei L eine lineare Abbidlung von $\mathbb{R}^N$ nach $\mathbb{R}^N$ und $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N, f(x) = a + Lx$. Dann ist f invertierbar genau dann, wenn L invertierbar ist. In diesem Fall gilt fuer die Umkehrabbildung g also

$$g : \mathbb{R}^N \longrightarrow \mathbb{R}^N, g(y) = -L^{-1}a + L^{-1}y.$$

Insbesondere ist g stetig differenzierbar mit

$$Dg(y) = Df(x)^{-1}.$$

In beiden Faellen gilt also: Ist f stetig differenzierbar und $Df(x)$ invertierbar, so ist f lokal invertierbar und seine Umkehrabbildung ist stetig differenzierbar mit $Dg(y) = Df(x)^{-1}$. Tatsaechlich gilt folgender allgemeiner Satz.

Theorem (Satz ueber die Umkehrfunktion). *Sei $U \subset \mathbb{R}^N$ offen, $f : U \longrightarrow \mathbb{R}^N$ stetig differenzierbar, $p \in U$, $q := f(p)$ und $Df(p)$ invertierbar. Dann gibt es eine offene Umgebung V von p eine offene Umgebung W von q so dass folgendes gilt:*

- *$f|_V : V \longrightarrow W$ ist bijektiv,*

- *die Umkehrabbildung* $g := (f|_V)^{-1} : W \longrightarrow V$ *ist stetig diffbar,*
- *fuer alle* $x \in V$ *ist* $Df(x)$ *invertierbar und es gilt* $Dg(y) = Df(x)^{-1}$ *fuer* $y = f(x)$.

Zeichnung.

Bemerkung.

- Ist die Umkehrfunktion g von f diffbar, so folgt mit $g \circ f = id$ aus der Kettenregel, dass
$$1 = Dg(y)Df(x)$$
fuer $y = f(x)$. Die Invertierbarkeit von $Df(x)$ und die Formel fuer $Dg(y)$ sind also Konsequenzen der Diffbarkeit von g.
- In einer Dimension gilt: Ist $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig differenzierbar mit $Df(x) = f'(x) \neq 0$, so hat also f' festes Vorzeichen, f ist streng monoton und damit auf ganz $\mathbb{R}$ invertierbar. Wir zeigen nun an zwei Beispielen, dass in beliebigen Dimensionen globalen Invertierbarkeit i. a. nicht gilt, auch wenn Df ueberall invertierbar ist.

  Beispiel. Sei $f : \mathbb{R}^2 \setminus \{0\} \longrightarrow \mathbb{R}^2 \setminus \{0\}$, $f(\xi_1, \xi_2) = (\xi_1^2 - \xi_2^2, 2\xi_1\xi_2)$. Dann gilt
$$Df(\xi_1, \xi_2) = (2\xi_1 \ -2\xi_2, 2\xi_2 \ 2\xi_1).$$
  Also $\det Df(\xi_1, \xi_2) = 4(\xi_1^2 + \xi_2^2) \neq 0$. Damit ist Df ueberall invertierbar. Aber es ist f nicht injektiv, denn es gilt $f(-\xi_1, -\xi_2) = f(\xi_1, \xi_2)$. (Setzt man $z = \xi_1 + i\xi_2$, so ist $f(\xi_1, \xi_2) = (\Re z^2, \Im z^2)$. Es handelt sich also bei f um die komplexe Quadratfunktion und bei Df umd die Multiplikation mit $2z$.)

  Beispiel. Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $f(\xi_1, \xi_2) = e^{\xi_1}(\cos\xi_2, \sin\xi_2)$. Dann ist f nicht injektiv. Aber die Ableitung ist ueberall invertierbar: Es ist $Df(\xi_1, \xi_2) = e^{\xi_1}(\cos\xi_2 \ -\sin\xi_2, \sin\xi_2 \ \cos\xi_2)$. Also ist $\det Df = e^{\xi_1} \neq 0$. Aber es ist $f(\xi_1, \xi_2) = f(\xi_1, \xi_2 + 2\pi)$. (Setzt man $z = \xi_1 + i\xi_2$ so ist $f(z) = (\Re e^z, \Im e^z)$. Es handelt sich also um die komplexe Exponentialfunktion und ihre Ableitung.)

Zur Vorbereitung auf den Beweis des Satzes beweisen wir zwei Resultate.

LEMMA. *Sei* $A : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ *linear. Dann sind aequivalent:*

(i) *Es ist* A *invertierbar.*
(ii) *Es gibt ein* $C > 0$ *mit* $|Ax| \geq C|x|$ *fuer alle* $x \in \mathbb{R}^N$
(iii) *Es ist* 0 *nicht Eigenwert von* A.

In diesem Fall kann $C = \frac{1}{\|A^{-1}\|}$ *gewaehlt werden.*

Beweis. (i)$\Longrightarrow$ (ii): Sei $y = Ax$. Dann gilt nach (i) also $x = A^{-1}y$. Damit folgt
$$|x| \leq \|A^{-1}\||y|.$$
Das liefert dann
$$|Ax| = |y| \geq \frac{1}{\|A^{-1}\|}|x|.$$
(ii)$\Longrightarrow$ (iii): Das ist klar.
(iii):$\Longrightarrow$ (i): Nach (iii) ist A injektiv. Damit ist es bijektiv. □

$\longleftrightarrow$
Ende der Vorlesung

PROPOSITION. *Sei $h : B_R(0) \subset \mathbb{R}^N \longrightarrow \mathbb{R}^N$ stetig und differenzierbar in $U_R(0)$ und*
$$\psi : U_R(0) \longrightarrow \mathbb{R}, \psi(x) = |h(x)|^2.$$
Hat ψ ein Extremum in p in U_R, so gilt $Dh(p)^t h(p) = 0$. Insbesondere verschwindet also h in p, wenn Dh auf ganz $U_R(0)$ invertierbar ist.

Beweis. Das folgt direkt durch Bilden des Gradienten:
$$0 = \partial_j \psi(p) = \partial_j \sum_{k=1}^{N} h_k(p)^2 = 2 \sum_{k=1}^{N} \partial_j h_k(p) \cdot h_k(p).$$
□

Bemerkung.
- Die Aussage ist bemerkenswert, weil auf den Funktionswert im Extremum geschlossen werden kann (und nicht nur auf den Gradienten).
- (Uebung) Ist $h : \mathbb{R}^N \longrightarrow \mathbb{R}$ stetig differenzierbar mit
  - $Dh(x)$ invertierbar fuer alle $x \in \mathbb{R}^N$,
  - $|h| \to \infty$ fuer $x \to \infty$,

  so gibt es ein $x \in \mathbb{R}^N$ mit $h(x) = 0$.

Wir kommen nun zum Beweis des Satzes ueber die Umkehrfunktion.

Beweis. Wir setzen
$$\lambda := \frac{1}{2}\|Df(p)^{-1}\|^{-1} > 0.$$
Es ist λ die einzige relevante Groesse in unserem Beweis!
Da f stetig differenzierbar ist, existiert ein $\widetilde{\delta} > 0$ mit
$$\|Df(x) - Df(p)\| \leq \lambda$$
fuer alle $x \in U$ mit $|x - p| \leq \widetilde{\delta}$. Da U offen ist, gibt es ein $\delta > 0$ mit $\delta < \widetilde{\delta}$ und $U_\delta(p) \subset U$. Setze
$$V := U_\delta(p) \subset U.$$
Dann gilt also
$$\|Df(x) - Df(p)\| \leq \lambda \ \ (*)$$
fuer alle $x \in V$.

Zwischenbehauptung: Fuer alle $x, x^{'} \in V$ gilt $|f(x) - f(x^{'})| \geq \lambda |x - x^{'}|$.
Bew. Sei $\varphi(x) = f(x) - Df(p)(x-p) - f(p)$, also

$$f(x) = f(p) + Df(p)(x-p) + \varphi(x).$$

Dann gilt:

$$\begin{array}{rcl} |f(x) - f(x^{'}| & = & |Df(p)(x - x^{'}) + \varphi(x) - \varphi(x^{'})| \\ (\Delta - Ugl) & \geq & |Df(p)(x-x^{'})| - |\varphi(x) - \varphi(x^{'})| \\ (Lemma) & \geq & 2\lambda |x - x^{'}| - |\varphi(x) - \varphi(x^{'})| \\ (!) & \geq & 2\lambda |x-x^{'}| - \lambda |x - x^{'}| \\ & = & \lambda |x - x^{'}|. \end{array}$$

(!: Das folgt aus dem Mittelwertsatz:

$$\begin{array}{rcl} |\varphi(x) - \varphi(x^{'})| & \leq & \|D\varphi(\xi)\| |x - x^{'}| \\ & = & \|Df(\xi) - Df(p)\| |x - x^{'}| \\ (\xi \in V,\ Def\ V) & \leq & \lambda |x - x^{'}|.) \end{array}$$

Nach diesen Vorbereitungen koennen wir nun die Aussagen des Satzes (in anderer Reihenfolge) beweisen.

Behauptung. Fuer alle $x \in V$ ist $Df(x)$ invertierbar.
Bew. Es gilt

$$\begin{array}{rcl} |Df(x)y| \geq |Df(p)y| - |(Df(x) - Df(p))y| & & \\ (Lemma, \text{Def. } V, \lambda) & \geq & 2\lambda|y| - \lambda|y| \\ & = & \lambda|y| \end{array}$$

Damit folgt die Aussage nach dem Lemma zur Invertierbarkeit.

Behauptung. $W := f(V)$ ist offen.
Bew. Sei $y^{'} := f(x^{'}) \in W$ beliebig. Sei $\delta^{'} := \delta - |x^{'} - p|, \varepsilon := \frac{\lambda}{2}\delta^{'}$. Wir zeigen $U_\varepsilon(y^{'}) \subset W$. **Zeichnung.**
Sei also $z \in U_\varepsilon(y^{'})$ beliebig. Damit folgt also

$$z \in U_{\frac{\lambda r}{2}}(y^{'}) \ (*)$$

fuer ein $0 < r < \delta^{'}$. Wir betrachten

$$\psi : B_r(x^{'}) \longrightarrow [0, \infty), \psi(x) = |f(x) - z|^2.$$

Da ψ (offenbar) stetig ist, nimmt es auf dem kompakten $B_r(x^{'})$ sein Minimum an. (Wir muessen zeigen, dass dieses Minimum 0 ist.) Wir zeigen nun, dass das Minimum im Inneren angenommen wird: Fuer x

auf dem Rand von $B_r(x^{'})$ gilt $|x - x^{'}| = r$ und damit

$$\begin{aligned} \psi(x) = |f(x) - z|^2 &\geq (|f(x) - f(x^{'})| - |y^{'} - z|)^2 \\ (ZB, (*)) &\geq (\lambda r - \frac{1}{2}\lambda r)^2 \\ &= (\frac{1}{2}\lambda r)^2. \end{aligned}$$

Fuer den Mittelpunkt $x^{'}$ von $B_r(x^{'})$ gilt aber

$$\psi(x^{'}) = |f(x^{'}) - z|^2 = |y^{'} - z|^2 \stackrel{(*)}{<} (\frac{1}{2}\lambda r)^2.$$

Damit wird also das Minimum im Inneren von $B_r(x^{'})$ angenommen. Dann muss aber nach dem vorigen Lemma (mit $h = f - z$) fuer den entsprechenden Punkt x_m gelten

$$0 = Df(x_m)^t(f(x_m) - z).$$

Da $Df(x_m)$ invertierbar ist (nach vorigen Behauptung) folgt also $f(x_m) - z = 0$. Das liefert die gewuenschte Offenheit.

Behauptung. $f : V \longrightarrow W$ ist bijektiv.
Bew. Nach Definition von W ist f surjektiv. Nach ZB ist f injektiv.

Behauptung. Es ist $g = (f|_V)^{-1}$ differenzierbar mit $Dg(y) = Df(x)^{-1}$ fuer $f(x) = y$.
Bew. Wir betrachten nur den Punkt $y = q = f(p)$. (In den anderen Punkten von V koennen analoge Betrachtungen gemacht werden, da Differenzierbarkeit eine lokale Eigenschaft ist.) Sei $\varepsilon = \frac{\lambda}{2}\delta$ wie oben zu $x^{'} = p$ gewaehlt. Dann gilt also $U_\varepsilon(q) \subset W$, $U_\delta(p) \subset V$. **Zeichnung.** Fuer $y \in \mathbb{R}^N$ mit $|y| < \varepsilon$ gilt dann

$$g(q + y) = p + x$$

mit $|x| < \delta$. Insgesamt hat mal also

$$g(q + y) = p + x,\ g(q) = p,\quad f(p + x) = q + y,\ f(p) = q.$$

Wir koennen nun abschatzen

$$|y| = |(y + q) - q| = |f(p + x) - f(p)| \stackrel{ZB}{\geq} \lambda|x|\ \ (**).$$

Weiterhin gilt

$$f(p + x) = f(p) + Df(p)x + \varphi_f(p + x)$$

also

$$q + y = q + Df(p)x + \varphi_f(p + x)\ \ (***)$$

mit

$$\frac{\varphi(p + x)}{|x|} \to 0, x \to 0.$$

Damit erhaelt man insgesamt:

$$\begin{aligned}
\frac{\varphi_g(q+y)}{|y|} &= \frac{1}{|y|}|g(q+y)-g(q)-Df(p)^{-1}y| \\
&= \frac{1}{|y|}|x-Df(p)^{-1}y| \\
&= \frac{1}{|y|}|Df(p)^{-1}(Df(p)x-y)| \\
(***) \quad &= \frac{1}{|y|}|Df(p)^{-1}\varphi_f(p+x)| \\
&\leq \frac{1}{|y|}\|Df(p)^{-1}\||\varphi_f(p+x)| \\
(**) \quad &\leq \frac{\|Df(p)^{-1}\|}{\lambda|x|}|\varphi_f(p+x)| \\
&= \frac{\|Df(p)^{-1}\|}{\lambda}\frac{|\varphi_f(p+x)|}{|x|} \\
&\to 0, |x| \to 0.
\end{aligned}$$

Fuer $|y| \to 0$ folgt aber aus $(**)$ sofort $|x| \to 0$. Damit ergibt sich aus der vorigen Abschaetzung

$$\frac{\varphi_g(q+y)}{|y|} \to 0, \ |y| \to 0.$$

Das beendet den Beweis. $\square$

In der bisherigen Diskussion haben wir versucht, die Gleichung

$$g(x,y) = f(x) - y = 0$$

bei gegebenen y nach x aufzuloesen, d.h. x als Graph einer Funktion von y darzustellen. Wir fragen nun, wieweit das fuer allgemeine Funktionen g moeglich ist. Dazu vertauschen wir die Rollen von x und y und betrachten $g : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^M$ und seine Nullstellenmenge

$$\mathcal{N} = \mathcal{N}(g) := \{(x,y) : g(x,y) = 0\}.$$

Die Frage lautet, ob $\mathcal{N}$ lokal als Graph darstellbar ist d.h. ob man y als Funktion von x schreiben kann $y = g(x)$. Es geht also um das Aufloesen der Gleichung $g(x,y) = 0$ nach y als Funktion von x.

Beispiel. (Hoehenlinien einer Landkarte) Sei g : *Landkarte* $\longrightarrow \mathbb{R}$, $g(p) =$ Hoehe von p ueber dem Meeresspiegel. Sei fuer $c \in \mathbb{R}$

$$\mathcal{N}_c := \{p \in \textit{Landkarte} : g(p) - c = 0\}.$$

Dann ist $\mathcal{N}_c$ gerade die Hoehenlinie zu c. **Zeichnung**

Beispiel. (Einheitskreislinie) Sei $g : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$, $g(x,y) = x^2 + y^2 - 1$.

$$\mathcal{N} := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Dann gilt:**Zeichnung.**
$\mathcal{N} \cap \{y > 0\}$: Oberer Halbkreis $y = \sqrt{1 - x^2}$.
$\mathcal{N} \cap \{x < 0\}$: Linker Halbkreis $x = -\sqrt{1 - y^2}$.
$\mathcal{N} \cap \{y < 0\}$. Unterer Halbkreis $y = -\sqrt{1 - x^2}$.
$\mathcal{N} \cap \{x > 0\}$: Rechter Halbkreis. $x = \sqrt{1 - y^2}$.
Die Gleichung kann also lokal nach x bezw y aufgeloest werden. Das Aufloesen nach y ist dabei in Bereichen mit $\partial_y g(x, y) = 2y \neq 0$ moeglich und das Aufloesn nach x in Bereichen mit $\partial_x g(x, y) = 2x \neq 0$.

Ende der Vorlesung

Notation. Fuer $z \in \mathbb{R}^N \times \mathbb{R}^M$ schreiben wir $z = (x, y)$ mit $x \in \mathbb{R}^N$ und $y \in \mathbb{R}^M$. Ist $W \subset \mathbb{R}^N \times \mathbb{R}^M$ offen und $g : W \longrightarrow \mathbb{R}^M$ diffbar, so setzen wir

$$D_x g(z) := (\partial_j g_i(z))_{i=1,\dots,M, j=1,\dots,N} = Matrixmalen$$

$$D_y g(z) := (\partial_{j+N} g_i(z))_{i=1,\dots,M, j=1,\dots,M} = Matrixmalen.$$

Damit gilt dann also

$$Dg(z) = (D_x g(z)\ D_y g(z)).$$

Wichtig. $D_y g(z))$ ist eine quadratische Matrix ($M \times M$ Matrix).

THEOREM. *(Satz ueber implizite Funktionen) Sei* $W \subset \mathbb{R}^N \times \mathbb{R}^M$ *offen und* $g : W \longrightarrow \mathbb{R}^M$ *stetig diffbar. Sei*

$$\mathcal{N} := \{(x, y) \in W : g(x, y) = 0\}.$$

Ist $z_0 := (p, q) \in M$ *und* $D_y g(z_0)$ *invertierbar, so existieren offene Umgebungen* U *von* p *und* V *von* q *und ein stetig diffbares* $f : U \longrightarrow V$ *mit*

$$\mathcal{N} \cap (U \times V) = \textit{Graph von } f = \{(x, f(x)) : x \in U\}$$

sowie $Df(x) = -D_y g(x, f(x))^{-1} D_x g(x, f(x))$.

Kurz: Ist $g(p, q) = 0$ und $D_y g(p, q)$ invertierbar, so kann man die Loesung von $g(x, y) = 0$ in einer Umgebung von p als Graph darstellen.

Zeichnung.

Beweis. *Wir zeigen zunaechst die Existenz von* f.
Idee (vgl. Zeichnung). Nach dem Satz zur Umkehrfunktion ist um (p, q) herum ist jeder Punkt eindeutig durch seine x Koordinate und den Wert von g beschrieben. Daher koenen wir (x, g) als Koordinatensystem verwenden. Das liefert eine Zerlegung der Umgebung von (p, g) in Konstanzflaechen von g. Diese sind nach Konstruktion (Bilder von) Graphen.

Hier sind die Details: Sei $G : W \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$, $(x, y) \mapsto (x, g(x, y))$. Dann gilt

$$G(p, q) = (p, 0)$$

sowie

$$DG(p,q) \equiv \begin{pmatrix} 1_N & 0_{NM} \\ D_x g(p,q) & D_y g(p,q) \end{pmatrix}.$$

Damit folgt

$$\det DG(p,q) = \det D_y g(p,q) \neq 0$$

nach unserer Vorrausetzung an die Invertierbarkeit von $D_y g(p,q)$. Nach dem Satz ueber die Umkehrfunktion gibt es also offenen Umgebungen W von (p,q) und $\widetilde{W}$ von $G(p,q) = (p,0)$, so dass $G : W \longrightarrow \widetilde{W}$ bijektiv mit stetig diffbarer Inverser ist. Ohne Einschraenkung (sonst Verkleinern) koennen wir

$$\widetilde{W} = \widetilde{U_p} \times \widetilde{U_0}$$

mit Umgebungen $\widetilde{U}_p$ von p und $\widetilde{U}_0$ von 0 annehmen.

Weiteres Vorgehen in Worten: Wir koennen $\widetilde{W}$ offenbar in Konstanzflaechen (der zweiten Koordinate) zerlegen. Zurueckziehen mit G liefert dann eine Zerlegung von $\mathcal{N} \cap W$ in Konstanzflaechen von g.**Zeichnung.** Nach Konstruktion sind diese Konstanzflaechen aber gerade Graphen. Hier sind die Details:

Sei

$$\pi : \mathbb{R}^N \times \mathbb{R}^M \longrightarrow \mathbb{R}^M, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto y = (0, 1_M) \begin{pmatrix} x \\ y \end{pmatrix},$$

$$\iota : \mathbb{R}^N \longrightarrow \mathbb{R}^N \times \mathbb{R}^M, x \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 1_N \\ 0 \end{pmatrix} x.$$

Dann hat

$$f = \pi \circ G^{-1} \circ \iota : U \longrightarrow V$$

die gewuenschten Eigenschaften.

- ι bildet U auf $U \times \{0\}$ ab, $x \mapsto (x, 0)$.
- G^{-1} bildet $U \times \{0\}$ ab nach $U \times V$, $(x,0) \mapsto (x,y)$ mit $g(x,y) = 0$.
- π bildet $U \times V$ nach V ab $(x,y) \mapsto y$, wobei fuer die in Frage kommenden (x,y) gilt $g(x,y) = 0$.

Es bleibt die Ableitung zu berechnen. Nach Kettenregel ist f differenzierbar und es gilt

$$Df(x) = D\pi(G^{-1}(\iota(x))DG^{-1}(\iota(x))D\iota(x).$$

Es ist (beachte π, ι linear):

$$D\iota = \iota, D\pi = \pi.$$

Weiterhin gilt

$$DG(x,y) \equiv \begin{pmatrix} 1_N & 0_{NM} \\ D_x g(x,y) & D_y g(x,y) \end{pmatrix} =: \begin{pmatrix} 1_N & 0_{NM} \\ L_1 & L_2 \end{pmatrix}$$

mit $L_1 = D_x g(x,y)$ und $L_2 = D_y g(x,y)$. Damit ergibt sich das Inverse (Probe!) zu

$$DG^{-1} = \begin{pmatrix} 1_N & 0_{NM} \\ -L_2^{-1}L_1 & L_2^{-1} \end{pmatrix}.$$

Insgesamt folgt

$$Df(x) = (0, 1_M) \begin{pmatrix} 1_N & 0_{NM} \\ -L_2^{-1}L_1 & L_2^{-1} \end{pmatrix} \begin{pmatrix} 1_N \\ 0 \end{pmatrix} = -L_2^{-1}L_1$$

und das ist die Behauptung. □

Eine Reihe von Kommentaren sind angemessen.

Bemerkung.

- Sind f, g diffbar mit $g(x, f(x)) \equiv 0$ und $D_y g(x, f(x))$ invertierbar, so folgt aus der Kettenregel

$$0 \equiv Dg(x, f(x)) \begin{pmatrix} 1_m \\ Df(x) \end{pmatrix} = D_x g(x, f(x)) + D_y g(x, f(x)) Df(x).$$

  Damit ergibt sich die Formel fuer die Ableitung von f

$$Df(x) = -D_y g(x, f(x))^{-1} D_x g(x, f(x)).$$

  Auch hier folgt also die Formel fuer die Ableitung von f automatisch aus der Differenzierbarkeit von f und der Kettenregel.
- Im allgemeinen werden wir die Funktion f aus dem Satz nicht explizit bestimmen koennen. Es ist bemerkenswert, dass wir trotzdem die Ableitung von f berechnen koennen.
- Ist $D_y g(p,q)$ nicht invertierbar, so ist i.a. eindeutiges Aufloesen von y nach x in Umgebung von p nicht moeglich:

  Beispiel (kein eindeutiges Aufloesen moeglich):

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R}, g(x,y) = x^2 + y^2 - 1, (p,q) = (1,0).$$

  Beispiel (gar kein Aufloesen moeglich)

$$h : \mathbb{R}^2 \longrightarrow \mathbb{R}, g(x,y) = x^2 + y^2, (p,q) = (0,0).$$

- Statt $g \equiv 0$ kann man auch $g \equiv c$ betrachten. (Dazu ersetzt man einfach g durch $\widetilde{g} = g - c$...)
- Ist $W \subset \mathbb{R}^N \times \mathbb{R}^M$ offen und $g : W \longrightarrow \mathbb{R}^M$ stetig diffbar mit $g(z_0) = 0$ und $Rang Dg(z_0) = M$, so kann man den Satz nach Umsortieren der Koordinaten anwenden. (Vgl. Beispiel mit Einheitskreis.)
- Wir haben den Satz ueber implizite Funktionen mittels des Satzes ueber die Umkehrfunktion gezeigt. Man kann auch umgekehrt vorgehen. Dazu betrachtet man $g(x,y) = f(x) - y$ in einer Umgebung von (p,q) mit $f(p) = q$ d.h. $g(p,q) = 0$. Ist $D_x f(p)$ invertierbar, so ist $D_x g(p,q)$ invertierbar und damit

nach dem Satz ueber implizite Funktionen also x nahe an p als Funktion von y nahe q darstellbar. (Details: Uebung).

KAPITEL 16

Untermannigfaltigkeiten, bedingte Extrema und all das

Untermannigfaltigkeiten sind die glatten Gebilde im Raum. Sie sind lokal Nullstellenmengen von glatten Funktionen oder aequivalent lokale Graphen. **Zeichnungen.** Kreis, Sphaere, Torus, Parabel.

Notation. Die Nullstellenmenge einer Funktion g bezeichnen wir weiterhin mit $\mathcal{N}(g)$, d.h. zu $g : W \subset \mathbb{R}^N \longrightarrow \mathbb{R}^L$ definieren wir

$$\mathcal{N}(g) := \{z \in W : g(z) = 0\} = \cap_{j=1}^{L} \mathcal{N}(g_j).$$

Den Graphen einer Funktion f bezeichen wir mit $\mathcal{G}(f)$ d.h. zu $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ definieren wir

$$\mathcal{G}(f) := \{(x, f(x)) : x \in U\} \subset \mathbb{R}^N \times \mathbb{R}^M.$$

Definition. *Eine stetig differenzierbare Funktion $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ heisst regular in x, wenn $Df(x)$ maximalen Rang hat. Ist die Funktion in jedem $x \in U$ regular, so heisst sie regulaer.*

Bemerkung. Ist $f : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^M$ regulaer, so gilt also

- falls $M \leq N$: $Rang Df(x) = M$ d.h. Zeilen von $Df(x)$ sind linear unabhaengig.
- falls $N \leq M$: $Rang Df(x) = N$ d.h. Spalten von $Df(x)$ sind linear unabhaengig.
- falls $N = M$: $Df(x)$ invertierbar.

Wie man am letzten Punkt sieht, ist Regularitaet eine Verallgemeinerung von Invertierbarkeit.

Beispiel.

- Ist $g : W \longrightarrow \mathbb{R}$ stetig diffbar, so ist g genau dann regular, wenn $\nabla g(x) \neq 0$ in jedem $x \in W$.
- $U \subset \mathbb{R}^N$ offen, $f : U \longrightarrow \mathbb{R}^M$ stetig differenzierbar. Dann ist $F : U \longrightarrow \mathbb{R}^N \times \mathbb{R}^M$, $F(x) = (x, f(x))$ regular. (Denn $DF(x) = (1, Df(x))^t$ und damit $Rang DF = N$.)

Lemma. *Sei $X \subset \mathbb{R}^N$, $p \in X$ und $L \leq N$. Aequivalent:*

(i) *Es existiert eine offene Umgebung W von p in $\mathbb{R}^N$ und regulaeres $g : W \longrightarrow \mathbb{R}^{N-L}$ mit $X \cap W = \mathcal{N}(g)$. ('X ist lokal eine Nullstellenmenge.')*

(ii) *Es existiert eine offenen Umgebung W von $p \in \mathbb{R}^N$ und eine offenen Menge $U \subset \mathbb{R}^L$ und $f : U \longrightarrow \mathbb{R}^{N-L}$, und eine Permutation $P : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ mit*

$$\mathcal{N} \cap W = P\mathcal{G}(f).$$

('X ist lokal Graph.')

Beweis. Die Richtung (i)$\Longrightarrow$ (ii) folgt aus dem Satz ueber implizite Funktionen. Die Implikation (ii)$\Longrightarrow$ (i) folgt einfach durch Betrachten von $g(x, y) = y - f(x)$ (wenn man $P = Id$ annimmt). $\square$

DEFINITION. *(Untermannigfaltigkeit) Eine Teilmenge X von $\mathbb{R}^N$ heisst Untermannigfaltigkeit der Dimension L, wenn sie in jedem Punkt $p \in X$ eine der Bedingungen des vorangehenden Lemma erfuellt.*

Bemerkung.

- Untermannigfaltigkeiten der Dimension $N - 1$ heissen Hyperflaechen. Sie sind also (lokal) Nullstellengebilde von reellwertigen Funktionen.
- Alle Untermannigfaltigkeiten sind lokal Schnitte von Hyperflaechen (da $\mathcal{N}(g) = \cap_k^{N-L} \mathcal{N}(g_k)$).

Beispiele. Offenbar sind alle Graphen und alle Nullstellengebilde regulaerer Funktionen also Untermannigfaltigkeiten. Insbesondere gehoeren dazu:

- Kreis $(g(x, y) = x^2 + y^2 - 1 = 0)$
- Die Sphaere $S^{N-1} := \{x \in \mathbb{R}^N : |x|^2 - 1 = 0\}$ $(g(\xi_1, \ldots, \xi_N) = \sum \xi_j^2 - 1 = 0.)$
- Torus (Uebung)
- Parabel (Graph).

Wir fuehren nun die Tangentialflaechen und Normalen an Untermannigfaltigkeiten ein. **Zeichnung.**

DEFINITION. *Sei $X \subset \mathbb{R}^N$ eine Untermannigfaltigkeit und $p \in X$. Dann heisst*

$$T_pX := \{\gamma'(0) : \gamma : (-1, 1) \longrightarrow \mathbb{R}^N, \textit{stetig diffbar}, \ \gamma(0) = p, \gamma(-1, 1) \subset X\}$$

der Tangentialraum von X in p. Der Normalraum N_pX von X in p ist definiert als $N_pX := T_pX^\perp$.

Bemerkung. Man koennte auch Kurven γ auf $(-\varepsilon_\gamma, \varepsilon_\gamma)$ betrachten.

THEOREM. *(Beschreibung Tangentialraum und Normalraum) Sei $X \subset \mathbb{R}^N$ eine Untermannigfaltigkeit der Dimension L. Sei $p \in X$ und $W \subset \mathbb{R}^N$ offen mit*

$$p \in X \cap W = \mathcal{N}(g) = \mathcal{G}(f) = Bild(F),$$

wobei $U \subset \mathbb{R}^L$, offen, $f : U \longrightarrow \mathbb{R}^{N-L}$ stetig diffbar, $F : U \longrightarrow \mathbb{R}^N$, $F(x) = (x, f(x))$, $g : W \longrightarrow \mathbb{R}^{N-L}$, stetig diffbar und regulaer. Dann gilt

$$T_pX = KerDg(p) = BildDF(p_0)$$

fuer $p = F(p_0) = (p_0, f(p_0))$. Inbesondere gilt

$$T_pX = \bigcap_{j=1}^{N-L} \{\nabla g_j(p)\}^{\perp}, \;\; N_pX = Lin\{\nabla g_j(p) : j = 1, \ldots, N-L\}$$

sowie $dimT_pX = L$, $dimN_pX = N - L$.

Bemerkung.

- Es zeigt ∇g_i in Richtung des staerksten Wachstum von g_i. In Richtung von T_pX gibt es umgekehrt kein Wachstum der g_i (da Nullstellenmenge). Daher ist der Normalraum gerade durch die Gradienten der g_i aufgespannt. **Zeichnung**.
- Der Satz zeigt, dass T_pX ein Unterraum ist.

Beweis. $T_pX \subset KerDg$: Sei $v \in T_pX$ beliebig und γ zugehoerige Kurve d.h $\gamma : (-1, 1) \longrightarrow X$, $\gamma(0) = p$, $\gamma'(0) = v$. Dann gilt $g \circ \gamma \equiv 0$ also

$$0 = (g \circ \gamma)'(0) = Dg(\gamma(0))\gamma'(0) = Dg(p)v.$$

$BildDF \subset T_pX$: Sei $w \in \mathbb{R}^L$ beliebig, $\gamma : (-1, 1) \longrightarrow X$, $\gamma(t) = F(p_0 + tw)$. Dann gilt

$$DF(p_0)w = \gamma'(0) \in T_pX.$$

Damit folgt die Aussage.

Insgesamt zeigt dies

$$BildDF(p_0) \subset T_pX \subset KerDg(p).$$

Weiterhin gilt:

$$dimKerD(g) = L$$

(nach Dimensionsformel $N = \dim KerDg + dimBildDg$ und g regulaer also $dimBildDg = N - L$)

$$RangBildDF(p_0) = L$$

(klar, betrachte Ableitung). Damit folgt dann

$$BildDF(p_0) = T_pX = KerDg(p).$$

$T_pX = \bigcap_{j=1}^{N-L}\{\nabla g_j(p)\}^\perp$*:* Das folgt sofort aus $T_pX = KerDg(p)$.

$N_pX = Lin\{\nabla g_j(p) : j = 1, \ldots, N-L\}$*:* Es gilt

$$Lin\{\nabla g_j(p) : j = 1, \ldots, N-L\}^\perp = \bigcap_{j=1}^{N-L}\{\nabla g_j(p)\}^\perp = T_pX.$$

Damit folgt dann durch Bilden des orthogonalen Komplementes die Aussage. □

Bemerkung/Achtung. Die Gleichheit $T_pX = KerDg = BildDF$ ist gerade eine linearisierte Version von $X = Ker(g) = BildF$. Aehnlich ist $Df(x)$ eine linearisierte Version von f. Tatsaechlich kann man fuer $f : U \subset \mathbb{R}^N \longrightarrow V \subset \mathbb{R}^M$ die Ableitung $Df(x)$ deuten als eine Abbildung des Tangentialraumes $\mathbb{R}^N = T_xU$ in den Tangentialraum $\mathbb{R}^M = T_{f(x)}V$.

Betrachtung von Untermannigfaltigkeiten ist in vielerlei Hinsicht nuetzlich. Wir lernen jetzt eine Anwendung auf sogenannte bedingte Exterema kennen. Diese ist unter dem Namen 'Methode der Lagrange Multiplikatoren' bekannt.

DEFINITION. *Sei $U \subset \mathbb{R}^N$ offen, $X \subset U$, $f : U \longrightarrow \mathbb{R}$, $p \in X$. Es nimmt f in p ein bedingtes lokales Maximum / Minimum auf X an, wenn ein $\delta > 0$ existiert mit*

$$f(x) \leq f(p)/f(x) \geq f(p)$$

fuer alle $x \in X$ mit $|x - p| \leq \delta$.

Da X i.a. nicht offen ist, koennen die bisher entwickelten Methoden zur Untersuchung von bedingten Extrema nicht angewendet werden. Wir untersuchen nun den Fall, dass X eine Nullstellenmenge ist.

THEOREM. *Seien $K \leq N$ aus $\mathbb{N}$ gegeben. Sei $U \subset \mathbb{R}^N$ offen, $g : U \longrightarrow \mathbb{R}^K$ stetig diffbar, $L := N - K$, (d.h. $K = N - L$)*

$$X := \mathcal{N}(g) = \bigcap_{i=1}^{K}\mathcal{N}(g_i).$$

Sei p ein regulaerer Punkt von g also $rangDg(p) = K$ (d.h. $\nabla g_i(p)$, $i = 1, \ldots, K$, linear unabhaengig). Sei $f : U \longrightarrow \mathbb{R}$ stetig diffbar. Hat f in p ein bedingtes lokales Extremum auf X, so gibt es reelle Zahlen $\lambda_1, \ldots, \lambda_K$ mit

$$\nabla f(p) = \sum_{j=1}^{K}\lambda_j \nabla g_j(p).$$

Ist insbesondere X ein Untermannigfaltigkeit (d.h. jeder Punkt von X ist regulaer), so gilt $\nabla f(p) \in N_pX = T_pX^\perp$.

Interpretation. Zeichnung. Haette $\nabla f(p)$ eine Komponente in Richtung von T_pX so koennte man auf X in dieser Richtung e wanderen und dabei den Wert von f wegen

$$\partial_e f(p) = \langle \nabla f(p), e \rangle$$

vergroessern und durch Wandern in Gegenrichtung verkleinern. In einem Extremum kann also $\nabla f(p)$ keine Komponente in Richtung von T_pX haben, muss also in N_pX liegen. Es ist aber N_pX die lineare Huelle der Gradienten der g_i (nach dem vorigen Satz).

Beweis. Ohne Einschraenkung ist X UMK (sonst Verkleinern). Ist $\gamma : (-\varepsilon, +\varepsilon) \longrightarrow X$ eine stetig diffbare Kurve in X mit $\gamma(0) = p$, so hat

$$h : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R},\ h(t) = f(\gamma(t))$$

in 0 ein lokales Extremum. Daher muss gelten (**Zeichnung**.)

$$0 = h'(0) = \langle \nabla f(p), \gamma'(0) \rangle.$$

Da γ beliebige Kurve in X durch p war, folgt $\nabla f(p) \perp T_P X$, also $\nabla f(p) \in N_pX$. Nach dem vorangehenden Satz folgt die Aussage. □

Verfahren/Methode der Lagrange Multiplikatoren. Um die Punkte zu finden, in denen *moeglicherweise* Extrema vorliegen, hat man das Gleichungssystem (∗)

$$g_j(p) = 0, j = 1, \ldots, K,\ \ (K - Gleichungen)$$

$$\nabla f(p) = \sum_{j=1}^{K} \lambda_j \nabla g_j(x)\ \ (N - Gleichungen)$$

nach $p = (p_1, \ldots, p_N)$ und $(\lambda_1, \ldots, \lambda_K)$ aufzuloesen. Das sind $K + N$ Gleichungen und $K + N$ Unbekannte.

Bemerkung. Man erhaelt das Gleichungssystem (∗) auch durch Betrachten von

$$F : U \times \mathbb{R}^K \longrightarrow \mathbb{R},\ \ F(x, \lambda_1, \ldots, \lambda_K) = f(x) - \sum_{j=1}^{K} \lambda_j g_j(x)$$

und Bilden des Gradienten

$$\nabla F(p, \lambda) = (\nabla f(p) - \sum_{j=1}^{L} \lambda_j \nabla g_j(p), g_1(p), \ldots, g_K(p))$$

und Nullsetzen des Gradienten. Das ist unter dem Stichwort 'Ankoppeln der Nebenbedingungen' bekannt. Die sich ergebenenden λ_j heissen Lagrangemultiplikatoren. In Beispielen koennen sie eine konkrete Interpretation haben (Zwangskraefte, Schattenpreise....).

Bemerkung. Die gesuchte Menge der bedingten Extrema und die berechnete Menge, in der $(*)$ gilt, koennen sehr verschieden sein: Seien

$$M_1 := \{\text{Extrempunkt von } f \text{ auf } X\}.$$

$$M_2 := \{\text{Extrempunkt von } f \text{ auf } X, \text{ in denen } Dg \text{ vollen Rang hat}\}.$$

$$M_3 := \{x \in X : (*)\}.$$

Dann gilt

$$M_1 \supset M_2 \subset M_3.$$

Es gilt:

- M_1 nicht leer, falls X kompakt.
- Erste Inklusion ist Gleichheit, falls Dg ueberall auf X vollen Rang hat, also X eine Untermannigfaltigkeit ist.
- Es koennen in M_3 Punkte dazukommen.

Folgerung. *Sei $U \subset \mathbb{R}^N$ offen, $f : U \longrightarrow \mathbb{R}$ stetig diffbar und $X \subset U$ eine kompakte Untermannigfaltigkeit. Dann nimmt f sein Maximum/Minimum auf X an und zwar in denjenigen Punkten von*

$$\{p \in X : \nabla f(p) \in N_pX\}$$

in denen es maximalen / minimalen Funktionswert hat.

Beispiel. $U = \mathbb{R}^2$, $g(\xi_1, \xi_2) = \xi_1^2 + \xi_2^2 - 1$, $X = \mathcal{N}(g) = S$. Gesucht Extrema von $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x) = \xi_1 + 2\xi_2$ auf X.

Da f stetig und X kompakt ist, nimmt f auf X auf jeden Fall Minimum und Maximum an. Wir berechnen nun Gradienten von f und g. Es gilt

$$\nabla f(x) = (1,2)^t, \ \ \nabla g(x) = 2(\xi_1, \xi_2)^t.$$

Insbesondere gilt $\nabla g \neq 0$ auf ganz X und X ist eine UMK. Aufloesen von $(*)$ liefert dann die Extrempunkte von f auf M (und moeglicherweise noch mehr Punkte). Es wird $(*)$ zu

$$\xi_1^2 + \xi_2^2 - 1 = 0 \quad (1,2) - \lambda 2(\xi_1, \xi_2) = 0.$$

Aufloesen liefert (nach endlicher Rechnung):

$$\lambda = \pm\frac{\sqrt{5}}{2}, \ \ (\xi_1, \xi_2) = \pm(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}).$$

Es gilt $f(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}) = \sqrt{5}$ (Maximum) und $f(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}) = -\sqrt{5}$ (Minimum).