Höhere Mathematik für Physiker 2

Vorlesung im Sommersemester 2012

Universität Heidelberg

Prof. Dr. A. Marciniak-Czochra

Dieses Skriptum basiert auf:

Skript zur Analysis I-III, Vorlesung von Prof. Dr. Willi Jäger (Wintersemester 2003/4 bis Wintersemester 2004/5), Universität Heidelberg

Analysis für Physiker, Vorlesungsskriptum von Prof. Dr. Rainer Weissauer (Sommersemester 2011 bis Wintersemetser 2011/2012), Universität Heidelberg

Mathematik für Physiker 1, Helmut Fischer, Helmut Kaul, 7. Auflage, 2011, Wiesbaden

Inhaltsverzeichnis

Kapitel 1

Mengen und Zahlen

1.1 Logische Regeln und Zeichen

1.1.1 Quantoren

- $\forall x$ "für alle x", z.B. "$\forall x \in M : A$" heißt, dass eine Aussage A für alle Elemente einer vorgegebenen Menge M gültig ist.

- $\exists x$ ës gibt (mindestens) ein x"

Entsprechend

- $\exists!$ ës gibt genau ein ..."

- $\nexists$ ës gibt kein x..."

1.1.2 Weitere Symbole

Für zwei Aussagen A, B schreiben wir:

- $A \Rightarrow B$ als Abkürzung für äus A folgt B"

Die mathematische Bedeutung einer solchen Aussage wollen wir erläutern: $A \Rightarrow B$ ist genau dann richtig, wenn <u>entweder</u> A falsch und (B falsch oder richtig) <u>oder</u> A richtig und B richtig ist.

Als sprachliche Formulierung von "$A \Rightarrow B$" wird auch verwendet:
"A ist hinreichend(e Bedingung) für B", oder
"B ist notwendig(e Bedingung) für A".

Wir schreiben

- $A \Leftrightarrow B$ für "A gilt genau dann, wenn B gilt"
- $A \wedge B$ für "A und B sind gültig"
- $\neg A$ für "A ist falsch"bzw. "A gilt nicht"
- $A \vee B$ für "A oder B sind gültig" (aber nicht ëntweder ... oder", es können also sowohl A als auch B zutreffen)

Im folgenden geben wir einen Überblick über die logische Struktur der üblichen mathematischen Beweistypen.

1.1.3 Direkter Schluss $A \Rightarrow B$

Beispiel 1.1. *m gerade Zahl $\Rightarrow m^2$ gerade Zahl*

Beweis. m gerade
$\Rightarrow \exists n$ natürliche Zahl s.d. $m = 2n$
$\Rightarrow m^2 = 4n^2 = 2(2n^2)$
$\Rightarrow m^2 = 2k$ mit $k = 2n^2$ natürliche Zahl
$\Rightarrow m^2$ gerade. □

1.1.4 Beweis des Transponierten (der Kontraposition)

Zum Beweis von $A \Rightarrow B$ zeigt man $\neg B \Rightarrow \neg A$.
Man sieht nämlich

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

Beispiel 1.2. *Sei $m \in \mathbb{N}$ dann gilt m^2 gerade $\Rightarrow m$ gerade*

Beweis. Wir zeigen stattdessen: m ist ungerade $\Rightarrow m^2$ ist ungerade
m ungerade $\Rightarrow$ es gibt eine natürliche Zahl $n : m = 2n + 1 \Rightarrow m^2 = (2n + 1)^2 = 4n^2 + 4n + 1 = 2(2n^2 + 2n) + 1 \Rightarrow m^2 = 2k + 1$, wobei k natürliche Zahl $\Rightarrow m^2$ ist ungerade □

1.1.5 Indirekter Schluss (Beweis durch Widerspruch)

Man nimmt an, dass $A \Rightarrow B$ nicht gilt, d.h. $A \wedge \neg B$ und zeigt, dass dann für eine dritte Aussage C gelten muss:
$C \wedge \neg C$, also ein Widerspruch (kurz ↯)

Beispiel 1.3. *$\nexists a$ rationale Zahl: $a^2 = 2$*
anders ausgedrückt: a rationale Zahl $\Rightarrow a^2 \neq 2$

Beweis. Wir nehmen an, dass $a^2 = 2$
dann folgt
einerseits: $\exists$ ganze Zahlen b, c teilerfremd (ohne Einschränkung, denn sonst kürze soweit wie möglich) mit $a = b/c$.
andererseits: $2 = a^2 = \left(\frac{b}{c}\right)^2 = \frac{b^2}{c^2} \Rightarrow b^2 = 2c^2$
b^2 ist gerade $\Rightarrow b$ ist gerade (wie schon gezeigt) $\Rightarrow \exists d$ natürliche Zahl : $b = 2d \Rightarrow b^2 = 4d^2$.
Außerdem ist $b^2 = 2c^2$, also $2c^2 = 4d^2 \Rightarrow c^2 = 2d^2 \Rightarrow c^2$ ist gerade $\Rightarrow c$ ist auch gerade.
Also müssen b und c beide gerade sein, insbesondere nicht teilerfremd, damit haben wir einen Widerspruch $\lightning$hergeleitet. $\square$

1.1.6 Induktionsbeweis

Um zu zeigen, dass eine Aussage $A(n)$ für alle natürlichen Zahlen gilt, beweist man:

1. $A(0)$ gilt
2. mit $A(n)$ gilt auch $A(n+1)$ für alle natürlichen Zahlen

Dieses Induktinosprinzip lautet also in formaler Schreibweise:

$$A(0) \wedge \big(\forall n : A(n) \Rightarrow A(n+1)\big) \Rightarrow \forall n : A(n)$$

Bemerkung 1.4. *Eine ausführliche Behandlung folgt später.*

Das Beweisverfahren der vollständigen Induktion
Es soll die Gültigkeit einer Aussage $A(n)$ für alle $n \in \mathbb{N}$ bewiesen werden.

Beispiel 1.5. $A(n) = 1^2 + 2^2 + ... + n^2 = \frac{1}{6}n(n+1)(2n+1)$

Beweis. Wir zeigen

- Induktionsanfang, d.h. $A(1)$ ist richtig: $A(1) : 1^2 = \frac{1}{6} \cdot 1 \cdot 2 \cdot 3 = 1$
- Induktionsschritt: wir zeigen, dass: falls $A(k)$ richtig $\Rightarrow A(k+1)$ richtig

Wir nehmen an, dass $A(k)$ richtig ist, d.h.

$$1^2 + ... + k^2 = \frac{1}{6}k(k+1)(2k+1) \tag{1.1}$$

dann folgt:

$$\begin{aligned} 1^2 + ... + k^2 + (k+1)^2 &= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2 \\ &= (k+1)\big(\frac{1}{6}k(2k+1) + k + 1\big) \\ &= (k+1)(\frac{1}{3}k^2 + \frac{1}{6}k + k + 1) \\ &= (k+1)(\frac{1}{3}k^2 + \frac{7}{6}k + 1) \\ &= \frac{1}{6}(k+1)(2k^2 + 7k + 6) \\ &= \frac{1}{6}(k+1)(2k+3)(k+2) \\ &= \frac{1}{6}(k+1)(k+2)(2(k+1)+1) \end{aligned}$$

Erläuterung: Falls $A(k)$ richtig ist, ist auch $A(k+1)$ richtig. Nun ist $A(1)$ richtig, nach dem eben bewiesenen also auch $A(2)$. Aus $A(2)$ folgt $A(3)$ und so weiter. Das Induktionsprinzip erlaubt aus dem Induktionsschluss: $A(n)$ richtig für alle $n \in \mathbb{N}$ □

Bemerkung 1.6. *Bei der vollständigen Induktion kann auch mit irgendeiner natürlichen Zahl n_0 anstelle von 0 gestartet werden*

1.1.7 Summenzeichen und Produktzeichen

Definition 1.7 (Definition (Summenzeichen):)**.** *Wir definieren für $n > 0$:*

$$\sum_{k=1}^{n} a_k := a_1 + ... + a_n.$$

Allgemeiner definieren wir

$$\sum_{k=m}^{n} a_k = a_m + ... + a_n, \text{ falls } n \geq m$$

bzw.

$$\sum_{k=m}^{n} a_k := 0, \text{ falls } n < m, \text{ sog. "leere Summe".}$$

Analog definieren das Produktzeichen.

Definition 1.8 (Definition (Produktzeichen):)**.** *Wir definieren für $n > 0$:*

$$\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdot ... \cdot a_n.$$

Allgemeiner definieren wir

$$\prod_{k=m}^{n} a_k = a_m \cdot ... \cdot a_n, \text{ falls } n \geq m$$

bzw.

$$\prod_{k=m}^{n} a_k := 1, \text{ falls } n < m, \text{ sog. "leeres Produkt".}$$

Definition 1.9 (Definition (Fakultät und Potenz):)**.** *Die Fakultät (sprich: "n Fakultät") ist definiert durch*

$$n! := \prod_{k=1}^{n} k \quad \textit{und } 0! := 1.$$

Die Potenz (sprich: "a hoch n") ist definiert durch

$$a^n := \prod_{k=1}^{n} a \ , \ a^0 := 1.$$

Bemerkung 1.10. *Die Rechenregeln*

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

ergeben sich durch Induktion.

1.2 Mengen

Den Grundbegriff der Menge können wir nach Georg Cantor (1895) "definieren".

Definition 1.11. *Unter einer Menge verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen.*

Uns genügt es hier, festzuhalten, dass eine Menge M dadurch charakterisiert ist, dass von jedem vorliegendem Objekt x feststeht, ob gilt:

- $x \in M$ (x gehört zu M; x Element von M) oder
- $x \notin M$ (x gehört nicht zu M; x kein Element von M)

Eine Menge M mit genau den Elementen $x_1, ...x_n$ schreiben wir als $M = \{x_1, ..., x_n\}$
Eine Menge M für die $x \in M \Leftrightarrow A(x)$, wobei $A(x)$ eine Aussage über x ist, schreiben wir als

$$M = \{x \mid A(x)\} \text{ oder } M = \{x : A(x)\}$$

Jetzt definieren wir mit Hilfe dieser Grundbegriffe übliche Abkürzungen der Mengensprache

Definition 1.12. *1. Mengeninklusion*
$A \subset M$ (A ist eine Teilmenge von M)
$\Leftrightarrow \forall x : (x \in A \Rightarrow x \in M)$
Das schließt den Fall $A = M$ ein. Wir können auch schreiben $M \supset A$
zum Beispiel:
$\mathbb{N}$ - natürliche Zahlen
$\mathbb{Z}$ - ganze Zahlen
$\mathbb{N} \subset \mathbb{Z}$

2. $A = B \Leftrightarrow \forall x(x \in A \Leftrightarrow x \in B)$
Es gilt also $A = B \Leftrightarrow (A \subset B \wedge B \subset A)$

3. $A \subsetneq M$ (A ist echte Teilmenge von M)
$\Leftrightarrow A \subset M \wedge A \neq M$

Bildung spezieller Mengen

4. $\emptyset$ ist die Menge, die kein Element enthält, die so genannte leere Menge:

$$\nexists x : x \in \emptyset$$

Wir setzen fest, dass $\emptyset$ Teilmenge jeder Menge ist

Beispiel 1.13.

$$\{x \in \mathbb{R} : x^2 + 1 = 0\} = \emptyset \text{ mit } \mathbb{R} : \textit{reelle Zahlen}$$

5. Durchschnitt von A und B

$$A \cap B := \{x | x \in A \wedge x \in B\} \tag{1.2}$$

6. Vereinigung von A und B

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

7. *Differenz*
   auch: "A−Komplement von B" oder "Komplement von B in A" oder "A ohne B" oder "relatives/mengentheoretisches Komplement")

$$A \setminus B := \{x \mid x \in A \text{ und } x \notin B\} =: C_A B$$

8. *Potenzmenge von* $A : \mathcal{P}(A) := \{B \mid B \subset A\}$
   Die Elemente von $\mathcal{P}(A)$ *sind also selbst Mengen, nämlich gerade alle Teilmengen von* A

   Beispiel 1.14.
   $$\mathcal{P}([1,2]) = \{\{1\},\{2\},\{1,2\},\emptyset\}$$

Lemma 1.15. *Hat eine Menge* A n *Elemente, so hat die Potenzmenge* $\mathcal{P}(A)$ 2^n *Elemente*

Beweis. durch Induktion:

1. $n = 0$: Dann ist $A = \emptyset$ und $\mathcal{P}(A) = \emptyset$ hat genau $1 = 2^0$ Elemente

2. Induktionsschritt: $n \longrightarrow n+1$

   Angenommen, für jede Menge B mit n Elementen, habe $\mathcal{P}(B)$ 2^n Elemente. Sei nun A eine Menge mit $n+1$ Elementen und $A = \{B\} \cup \{x\}$, wo B n-elementig $\wedge$ $x \notin B$
   Wir unterscheiden zwei Sorten von Teilmengen von A

   (a) diejenigen, die x nicht enthalten; das sind die Teilmengen von B, also genau 2^n Stück

   (b) diejenigen, die x enthalten; das sind alle Mengen $C \cup \{x\}$ mit $C \subset B$, also nochmals genau 2^n Stück.

   Zusammen sind dies $2^n + 2^n = 2^{n+1}$ Teilmengen von A
   Daher schreiben wir auch
   $$\mathcal{P}(A) := 2^A$$

□

Definition 1.16. *Sei* I *eine Indexmenge,* $I \subset \mathbb{N}, (A_i)_{i\in I}$ *eine Familie von Mengen* A.

1. *Durchschnitt der* A_i*:*

$$\bigcap_{i\in I} A_i = \{x \mid \forall i \in I : x \in A_i\}$$

2. *Vereinigung der A_i*

$$\bigcup_{i\in I} A_i = \{x \mid \exists i \in I : x \in A_i\}$$

Bemerkung 1.17. *Diese Def. verallgemeinert die Bildung von $A_1 \cap A_2$, bzw. $A_1 \cup A_2$*

Rechenregeln
A, B, C, D seien Mengen:

1. $\emptyset \subset A$

2. Reflexivität: $A \subset A$

3. Transitivität: $A \subset B, B \subset C \Rightarrow A \subset C$

4. Kommutativität:

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

5. Assoziativität:

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

6. Distributivität:

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

7.

$$A \cap A = A$$
$$A \cup A = A$$
$$A \cup \emptyset = A$$
$$A \cap \emptyset = \emptyset$$

Eigenschaften der Komplementbildung

8. Seien $A, B \subset D$ (Notation: $C_D A := D \setminus A$), dann gilt

$$C_D(C_D A) = A,$$
$$C_D(A \cap B) = C_D A \cup C_D B,$$
$$C_D(A \cup B) = C_D A \cap C_D B$$

Beweis. Wir zeigen exemplarisch die letzte Gleichung in (8):
Sei $x \in D$
$x \in C_D(A \cup B) \Leftrightarrow x \notin A \cup B$
$\Leftrightarrow$ nicht $(x \in A \vee x \in B)$
$\Leftrightarrow (x \notin A \wedge x \notin B)$
$\Leftrightarrow x \in C_D A \wedge x \in C_D B$
$\Leftrightarrow x \in C_D A \cap C_D B$ $\square$

Definition 1.18. *Seien $x_1, x_2, ..., x_n$ (nicht notwendig verschiedene) Objekte. Ein geordnetes n-Tupel*

$$(x_1, ..., x_n) = (y_1, ..., y_n) \Leftrightarrow x_1 = y_1, \ ..., \ x_n = y_n$$

Beachte: $\{x_1, ..., x_n\} = \{y_1, ..., y_n\} \nRightarrow x_1 = y_1,$ *etc.*

Beispiel 1.19. $\{1, 2, 3\} = \{3, 2, 1\}$

Definition 1.20. ***kartesisches Produkt von Mengen***
Seien $A_1, ..., A_n$ Mengen
$A_1 \times A_2 \times ... \times A_n = \{(x_1, ..., x_n) \mid x_j \in A_j \text{ für } j = 1, ..., n\ \}$
wir identifizieren:
$(x_1, (x_2, x_3)), ((x_1, x_2), x_3)$ *und* (x_1, x_2, x_3)
Es gilt

$$A_1 \times (A_2 \times A_3) = (A_1 \times A_2) \times A_3 = A_1 \times A_2 \times A_3$$
$$A^n = A \times A ... \times A \quad n \text{ Faktoren}$$

Beispiel 1.21. $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z} \to$ *Punktgitter in der Ebene*
$\mathbb{R}^n \to$ *n-dimensionaler Raum von reellen Zahlen*

Rechenregeln
Seien A, B, C, D Mengen

1. $A \times B = \emptyset \Leftrightarrow A = \emptyset \vee B = \emptyset$
2. $(A \times C) \cup (B \times C) = (A \cup B) \times C$
3. $(A \times C) \cap (B \times D) = (A \cap B) \times (C \cap D)$

1.3 Natürliche Zahlen

Aus der Schule:

- $\mathbb{N}$ natürliche Zahlnen

- $\mathbb{Z}$ ganze Zahlen
- $\mathbb{Q} = \left\{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\right\}$ rationale Zahlen
- $\mathbb{R}$ reelle Zahlen

1.3.1 Axiome für $\mathbb{N}$ (Peano)

(A0) 0 ist natürliche Zahl

(A1) Zu jeder natürlichen Zahl gibt es genau einen "Nachfolger" $\sigma(n)$ ($= n + 1 \leftarrow$ Notation)

(A2) 0 ist nicht Nachfolger einer natürlichen Zahl

(A3) $\sigma(n) = \sigma(m) \Rightarrow n = m$

(A4) Es gilt das Induktionsprinzip

Bemerkung 1.22. *Allein mit Hilfe der "Nachfolger-Funktionünd den Axiomen (A3), (A4) lassen sich auf* $\mathbb{N}$ *Addition* $+$, *Multiplikation* $\cdot$ *und Ordnung* $\leqslant$ *einführen. Wir definieren*

$$\begin{aligned} 1 : = \sigma(0), 2 := \sigma(1) \ \textit{usw.} \\ n + 0 = n \\ n + 1 : = \sigma(n) \\ n + (m + 1) = (n + m) + 1 \end{aligned}$$

womit wir nacheinander für jedes m *die Zahl* $m + n$ *berechnen können.*
Wir definieren: $n \leqslant m \Leftrightarrow \exists$ *natürliche Zahl* $x : x + n = m$.
Produkte von natürlichen Zahlen definieren wir durch iterierte Summation.

Lemma 1.23. *Jede nichtleere Teilmenge* M *von* $\mathbb{N}$ *enthält ein kleinstes Element.*

Beweis. Sei M eine Menge natürlicher Zahlen. Wir wollen zeigen (Transposition): M hat kein kleinstes Element $\Rightarrow M$ ist leer.

Wir verdeutlichen uns

1. Die Aussage "m ist kleinstes Element von M" bedeutet: $m \in M$ und $\underbrace{\forall x, (x + 1) \leqslant m : x \notin M}_{A(m)}$.

2. Die Aussage "M ist leer" bedeutet: $\forall m : A(m)$.

3. Die Aussage "M hat kein kleinstes Element" bedeutet $\forall m \in \mathbb{N}: \; A(m) \Rightarrow m \notin M$ (d.h. entweder gilt $A(m)$ nicht und wenn $A(m)$ gilt, ist $m \notin M$).

Wir nehmen also an, M habe kein kleinstes Element und zeigen durch Induktion $\forall m : A(m)$.

Es gilt $A(0)$, denn $\nexists x : x+1 \leqslant 0$, andernfalls gäbe es nach Definition ein y mit $y + (x+1) = 0$ ↯ zu (A2).

Es verbleibt zu zeigen, dass $\forall m : A(m) \Rightarrow A(m+1)$. Wir haben angenommen, dass M kein kleinstes Element hat, folglich kann m nicht kleinstes Element von M sein, d.h. $A(m) \Rightarrow m \notin M$. Für den Induktionsschritt ist also anzunehmen: $A(m)$ und $m \notin M \Leftrightarrow \forall x, (x+1) \leqslant m : x \notin M$ und $m \notin M$.
$\Rightarrow \forall x, (x+1) \leqslant m+1 : x \notin M$ (denn $x+1 \leqslant m+1 \Leftrightarrow x+1 \leqslant m$ oder $x = m$)
$\Leftrightarrow A(m+1)$
Nach dem Induktionsprinzip folgt

$$\forall m : A(m), \text{ d.h. } M \text{ ist leer.}$$

□

1.4 Abbildungen und Relationen

Definition 1.24. *Unter einer Relation verstehen wir eine Teilmenge $R \subset X + Y$, wobei X, Y Mengen sind.*
Für $x \in X$ sei das Bild von x unter R

$$R(x) := \{y \in Y \mid (x,y) \in R\}$$
$$D(R) := \{x \in X \mid R(x) \neq \emptyset\} \text{ heißt Definitionsbereich von } R \text{ (bzgl. } X\text{)}$$

Definition 1.25. *Eine Relation $R \subset X \times Y$ heisst Graph der Abbildung (Funktion) $f : X \to Y$*
$\Leftrightarrow$

1. $D(R) = X$

2. $\forall x \in X : R(x) = \{f(x)\}$

d.h. $R(x)$ enthält genau ein Element.

- *X heißt Definitionsbereich von$f$*

- *$x \in X$ heißt Argument*

- *Y heißt Werte- oder Bildbereich von f*
- *$f(x) \in Y$ heißt Wert von f an der Stelle x.*

Beispiel 1.26.

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

dann ist

$$M = \text{ Graph } f = \left\{(x,y) \in \mathbb{R}^2 : y = x^2\right\}$$

Bemerkung 1.27. *Der an der Diagonalen gespiegelte Graph*

$$\begin{aligned} M^*(x) &= \left\{(x,y) \in \mathbb{R}^2 : x = y^2\right\} \\ &= \left\{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0, y = \sqrt{x} \text{ oder } y = -\sqrt{x}\right\} \end{aligned}$$

ist nicht Graph einer Funktion $\mathbb{R} \to \mathbb{R}$, denn

$$\begin{aligned} M^*(x) &= \emptyset \text{ für } x < 0 \\ M^*(x) &= \left\{\sqrt{x}, -\sqrt{x}\right\} \text{ für } x \geq 0 \end{aligned}$$

Wir können aber schreiben $M^* = M_1^* \cup M_2^*$ mit

$$\begin{aligned} M_1^* &= \left\{(x,y) \in \mathbb{R}_{\geq 0} \times \mathbb{R} : y = \sqrt{x}\right\} \\ M_2^* &= \left\{(x,y) \in \mathbb{R}_{\geq 0} \times \mathbb{R} : y = -\sqrt{x}\right\} \end{aligned}$$

Nun sind M_1^*, M_2^* Graphen der Funktionen

$$\begin{aligned} f_1, f_2 : \mathbb{R}_{\geq 0} &\to \mathbb{R} \\ x &\mapsto \sqrt{x} \quad \text{bzw.} \\ x &\mapsto -\sqrt{x} \end{aligned}$$

Definition 1.28. *Gegeben seien Mengen A, B und eine Abbildung $f : A \to B$*

f heisst surjektiv, wenn gilt $f(A) = B$
f heisst injektiv, wenn gilt $\forall x, y \in A : f(x) = f(y) \Rightarrow x = y$
f heisst bijektiv, wenn f surjektiv und injektiv ist

Definition 1.29. *1. Sei die Abbildung $f : A \to B$ bijektiv. Dann definieren wir die Umkehrabbildung $f^{-1} : B \to A$ durch $y \mapsto x \in A$, eindeutig bestimmt durch $y = f(x)$*

Bemerkung 1.30. *Der Graph von f^{-1} entsteht aus dem von f durch SSpiegelung an der Diagonalen"*

$$(x, y) \in \text{ Graph } f \;\Leftrightarrow (y, x) \in \text{ Graph } f^{-1}$$

2. *Seien $f : A \to B$ und $g : B \to C$ Abbildungen. Wir definieren die Komposition von g und f: $g \circ f : A \to C$ durch $x \mapsto g(f(x))$*

3. *Für jede Menge A definieren wir die identische Abbildung:*

$$\mathrm{id}(A) = I_A : A \to A \text{ durch } x \mapsto x$$

Beispiele

1. Die Menge aller Paare $(x, y), x, y \in \mathbb{R}$ mit $x^2 + y^2 = 1$.
   Die Lösungsmenge dieser Gleichung in $\mathbb{R}^2$ bezeichnet man als S^1
   In der Ebene $\mathbb{R}^2$ wird S^1 durch die Einheitskreislinie dargestellt.
   $S^{n-1} = \{(x_1, ...x_n) \in \mathbb{R}^n : \sum\limits_{i=1}^{n} x_i^2 = 1\}$ ist die $(n-1)-$ Sphäre im $\mathbb{R}^n$
   S^2 ist also die Oberfläche der Einheitskugel in $\mathbb{R}^3$.

2. Seien $X,\ Y$ Mengen, $M \subset X \times Y,\ D(M) = X$.

$$\begin{aligned} f : M &\to X \\ (x, y) &\mapsto x \end{aligned}$$

   f ist surjektiv, denn $f(M) = \{x : \exists y : (x, y) \in M\} = D(M) = X$.
   f heisst Projektion.

3. Wir fragen nach Existenz und Eindeutigkeit von Lösungen des Gleichungssystems in den Variablen x, y:

$$\begin{cases} a(x + y) = \xi \\ a(-x + y) = \eta \end{cases} \qquad a,\ \xi, \eta \in \mathbb{R}, a \neq 0$$

   Durch Äquivalenzumformungen erhalten wir:

$$y = \frac{\xi + \eta}{2a}$$
$$x = \frac{\xi - \eta}{2a},$$

   d.h. es gibt für jedes Paar (ξ, η) genau eine Lösung (x, y).

In der Sprache von Abbildungen und Umkehrabbildungen:
Wir betrachten die Abbildung

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (a(x + y), a(x + y)) \end{aligned}$$

und wollen wissen, ob für jedes $(\xi, \eta) \in \mathbb{R}^2$ genau ein $(x, y) \in \mathbb{R}^2$ existiert mit $f(x, y) = (\xi, \eta)$.
Das ist genau dann der Fall, wenn f bijektiv ist. Wie wir oben nachgerechnet haben, ist das betrachtete f bijektiv.

Wir erhalten $(x, y) = f^{-1}(\xi, \eta) = \left(\frac{\xi-\eta}{2a}, \frac{\xi+\eta}{2a}\right)$.

1.5 Abzählbarkeit

Definition 1.31. *Sei A eine Menge*

1. *A heisst endlich mit der Anzahl $|A| = n$ Elementen ist äquivalent zu*

$$\begin{cases} A = \emptyset \text{ oder } n = 0 \\ \exists f : A \to \{1, ..., n\}, \text{wobei } f \text{ bijektiv ist} \end{cases}$$

2. *A heisst abzählbar unendlich genau dann, wenn: $\exists f : A \to \mathbb{N}$ bijektiv*

3. *A heisst überabzählbar genau dann, wenn A weder endlich noch abzählbar unendlich ist.*

Beispiel 1.32. *$\mathbb{Z}$ ist abzählbar unendlich*

Beweis. Die Abbildung

$$\begin{aligned} f : \mathbb{Z} &\to \mathbb{N} \\ z &\mapsto \begin{cases} 2z & \text{für } z \geq 0 \\ -2z - 1 & \text{für } z < 0 \end{cases} \end{aligned}$$

ist bijektiv.
Zur Surjektivität:
Zu zeigen: $f(\mathbb{Z}) = \mathbb{N}$. Offenbar gilt $f(\mathbb{Z}) \subset \mathbb{N}$. Es verbleibt zu zeigen, dass auch $\mathbb{N} \subset f(\mathbb{Z})$ gilt. Sei $n \in \mathbb{N}$, finde $z \in \mathbb{Z}$ mit $f(z) = n$. Man unterscheide:

- n gerade $\Rightarrow$ wähle $z = \frac{n}{2}$

- n ungerade $\Rightarrow$ wähle $z = -\frac{n+1}{2}$

Damit gilt $\mathbb{N} \subset f(\mathbb{Z})$.

Zur Injektivität: Sei $z_1, z_2 \in Z$ und $f(z_1) = f(z_2)$
Zeige: $z_1 = z_2$ o.B.d.A $z_1 \leq z_2$
entweder sind $z_1, z_2 \geq 0$ oder $z_1, z_2 < 0$, denn sonst wäre$f(z_1)$ ungerade, $f(z_2)$ gerade
$\lightning$

$$z_1, z_2 \geq 0 : 2z_1 = f(z_1) = f(z_2) = 2z_2 \Rightarrow z_1 = z_2$$
$$z_1, z_2 < 0 : -2z_1 - 1 = f(z_1) = f(z_2) = -2z_2 - 1 \Rightarrow z_1 = z_2$$

$\square$

Beispiel 1.33.
- $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ *ist abzählbar unendlich*
- $\mathbb{Q}$ *ist abzählbar unendlich*
- $\mathbb{R}$ *ist überabzählbar*

Satz 1.34. *Sei $A \subset \mathbb{N}$: A enthalte mindestens zwei Elemente, dann ist die Menge aller Funktionen $\mathbb{N} \to A$,*

$$G := \{g \mid g : \mathbb{N} \to A\}$$

überabzählbar.

Beweis. Wir nehmen an, dass es eine bijektive Abbildung h gibt, mit $h : \mathbb{N} \to G$
Sei $h(n) =: g_n$
Wähle $a_0, a_1 \in A$, $a_0 \neq a_1$. Definiere $g^* \in G$ durch

$$g^*(n) = \begin{cases} a_0 & falls \quad g_n(n) \neq a_0 \\ a_1 & falls \quad g_n(n) = a_0 \end{cases}$$

Nach der Voraussetzung (Bijektion) existiert $m \in \mathbb{N}$ mit

$$g^* = h(m) = g_m$$

Nach Definition gilt aber $g^*(m) \neq g_m(m)$ $\lightning$ $\square$

1.5.1 Ordnung

Definition 1.35. *Sei A eine Menge. Eine Relation $R \subset A \times A$ heisst Teilordnung auf A, wenn für alle $x, y, z \in A$ gilt:*

(O1) $x \leq x$ *(Reflexivität)*

(O2) $x \leq y$ *und* $y \leq x \Rightarrow x = y$ *(Symmetrie)*

(O3) $x \leq y$ *und* $y \leq z \Rightarrow x \leq z$ *(Transitivität)*

Bemerkung 1.36. *Notation:* $x \leq y :\Leftrightarrow (x, y) \in \mathbb{R}$

Wenn ausserdem noch für alle $x, y \in A$ *gilt:*

(O4) $x \leq y$ *oder* $y \leq x$ *(Vergleichbarkeit je zweier Elemente),*

so heisst R (totale) Ordnung auf A. $(A, \leq)$ *heisst teilweise bzw. (total) geordnete Menge.*

Beispiele

1. $(\mathbb{R}, \leq)$ mit der üblichen Ordnung ist eine total geordnete Menge
2. Wir definieren auf der Potenzmenge $\mathcal{P}(A)$ einer Menge A eine Teilordnung "$\leq$".
$$B \leq C :\Leftrightarrow B \subset C, \quad \forall B, C \subset \mathcal{P}(A)$$
Beweis. (O1)-(O3) sind trivial
(O4) gilt nicht, d.h. es liegt keine Totalordnung vor: Wähle B, $C \in \mathcal{P}(A)$, $B, C \neq \emptyset$, $B \cap C = \emptyset$. Dann gilt weder $B \subset C$ noch $C \subset B$. □
3. Sei $F := \{f \mid f : A \to \mathbb{R}\}$ die Menge aller Funktionen $A \to \mathbb{R}$ für eine Menge $A \subset \mathbb{R}$. Wir definieren
$$f \leq g \Leftrightarrow \forall x \in A : f(x) \leq g(x)$$
(O1)-(O3) trivial
Auch diese Ordnung ist im Allgemeinen nicht vollständig: Falls A mehr als ein Element hat, gibt es Funktionen, die nicht miteinander verglichen werden können.

Definition 1.37. *Sei* $(A, \leq)$ *eine teilweise geordnete Menge,* $a \in A$
$a = \max A$ *("a ist Maximum von A")* $\Leftrightarrow \forall x \in A : x \leq a$
$a = \min A$ *("a Minimum von A")* $\Leftrightarrow \forall x \in A : a \leq x$

Bemerkung 1.38. *Durch diese Aussagen ist a eindeutig bestimmt, denn seien* $a_1, a_2 \in A$*:*

$$\forall x \in A \begin{cases} x \leq a_1 \\ x \leq a_2 \end{cases} \quad \textit{bzw.} \quad \begin{cases} a_1 \leq x \\ a_2 \leq x \end{cases} \quad \Rightarrow \quad \begin{cases} a_2 \leq a_1 \\ a_1 \leq a_2 \end{cases} \quad \Rightarrow \quad a_1 = a_2$$

Definition 1.39. *Sei* $(A, \leq)$ *eine (total) geordnete Menge,* $B \subset A$

1. *$S \in A$ heisst <u>obere Schranke</u> zu $B \Leftrightarrow \forall x \in B : x \leq S$*
   $S \in A$ heisst <u>untere Schranke</u> zu $B \Leftrightarrow \forall x \in B : S \leq x$

2. *Die Menge der oberen bzw. unteren Schranken von B bezeichnen wir mit*

$$\begin{aligned} &\bar{S}(B) = \{S \in A : S \textit{ ist untere Schranke zu } B\} \\ \textit{bzw.} \quad &\underline{S}(B) = \{S \in A : S \textit{ ist untere Schranke zu } B\} \end{aligned}$$

3. *Existiert $g := \min \bar{S}(B)$ bzw. $g := \max \underline{S}(B)$,*
   so sagen wir:

$$\begin{aligned} &g = \sup B \quad \textit{(kleinste obere Schranke; \underline{Supremum}; obere Grenze von B in A)} \\ \textit{bzw.} \quad &g = \inf B \quad \textit{(größte untere Schranke; \underline{Infimum}; untere Grenze von B in A)} \end{aligned}$$

Bemerkung 1.40. *1. Existiert $\max B = \bar{b}$, so folgt $\sup B$ existiert und $\sup B = \bar{b}$, denn*

   (a) $\bar{b} \in \bar{S}(B)$ nach Definition von $\max B$ und $\bar{S}(B)$

   (b) $s \in \bar{S}(B) \Rightarrow \bar{b} \leq s$, da $\bar{b} \in B$

   Ebenso gilt:

2. Existiert $\min B = \underline{b}$, so folgt, dass $\inf B$ existiert und $\inf B = \underline{b}$

<u>Beispiele</u>

1. $B = \left\{\frac{1}{n} : n \in \mathbb{N}^+\right\}, \quad A = \mathbb{R}$

   - Es gilt: $1 \in B$. $\forall n \in \mathbb{N}^+$ gilt: $\frac{1}{n} \leq 1$, daher folgt $1 = \max B = \sup B$.
   - Sei $s \leq 0$, dann gilt $\forall n \in \mathbb{N} : s \leq \frac{1}{n}$, also $s \in \underline{S}(B)$.
   - Sei $s > 0$, dann gilt: $s > \frac{1}{n} \Leftrightarrow n > \frac{1}{s}$, also $s \notin \underline{S}(B)$.

   Es folgt: $\underline{S}(B) = \{s \in \mathbb{R} : s \leq 0\}$, insbesondere $0 \in \underline{S}(B)$. Ferner gilt $\forall s \in \underline{S}(B) : s \leq 0 \Rightarrow 0 = \max \underline{S}(B) = \inf B$

2. $A = \mathbb{Q}, \quad B = \{x \in \mathbb{Q} : 0 \leq x, \quad x^2 \leq 2\}$
   Es gilt $0 = \min B = \inf B$, aber $\sup B$ existiert nicht (in $\mathbb{Q}$).

1.6 Reelle Zahlen

1.6.1 Axiome für $\mathbb{R}$

Wir haben schon angedeutet, wie Addition und Multiplikation auf $\mathbb{N}$ mit Hilfe der Peano-Axiome definiert werden kann. Wir wiederholen an dieser Stelle den aus der linearen Algebra bekannten Begriff des Körpers.

Eine Gruppe ist eine Menge zusammen mit einer zweistelligen inneren Verknüpfung, wenn diese Verknüpfung assoziativ ist und ein neutrales Element besitzt, sowie zu jedem Element ein Inverses existiert.

Beispiel 1.41. *1.* $(\mathbb{Z}, +)$

2. Die Lösung von Gleichungen $x + a = b$ *bzw.* $x \cdot c = d$ *ist in* $\mathbb{N}$ *nicht uneingeschränkt möglich* $\Rightarrow (\mathbb{N}, +)$ *bzw.* $(\mathbb{N}, \cdot)$ *sind keine Gruppen*

3. $(\mathbb{N}, +)$ *kann aber zur Gruppe* $(\mathbb{Z}, +)$ *erweitert werden, durch "Hinzunahme negativer Zahlen".* $(\mathbb{Z}, \cdot)$ *ist immer noch keine Gruppe.* $\mathbb{Z}$ *kann aber durch "Hinzunahme der Quotienten"*$\frac{p}{q}$, $p, q \in \mathbb{Z}$, $q \neq 0$ *erweitert werden zu* $\mathbb{Q}$, *so dass auf* $\mathbb{Q}$ *zwei Gruppenstrukturen* $(\mathbb{Q}, +)$, $(\mathbb{Q}, \cdot)$ *existieren.*

Definition 1.42. $\mathbb{K}$ *sei eine Menge auf der eine Addition und eine Multiplikation erklärt sind.* $\mathbb{K}$ *heisst Körper wenn die folgenden Axiome erfüllt sind.*

- *Addition*
  $(\mathbb{K}, +)$ *ist eine kommutative Gruppe, d.h.* $\forall a, b, c \in \mathbb{K}$

  A1 $(a + b) + c = a + (b + c)$ *(Assoziativgesetz der Addition)*

  A2 $a + b = b + a$ *(Kommutativgesetz der Addition)*

  A3 $\exists 0 \in \mathbb{K} : a + 0 = a$ *(Existenz des Nullelements)*

  A4 $\exists x \in \mathbb{K} : a + x = 0$ *(Existenz des Negativen)*

- *Multiplikation*
  $(\mathbb{K} \setminus \{0\}, \cdot)$ *ist eine kommutative Gruppe, d.h.* $\forall a, b, c \in \mathbb{K}$

  M1 $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ *(Assoziativgesetz der Multiplikation)*

  M2 $a \cdot b = b \cdot a$ *(Kommutativgesetz der Multiplikation)*

  M3 $\exists 1 \in \mathbb{K} : a \cdot 1 = a$ *(Existenz des Einselements)*

  M4 *Für* $a \neq 0$, $\exists y \in \mathbb{K}$ *mit* $a \cdot y = 1$ *(Existenz des Inversen)*

- *Verträglichkeit* $\forall a, b, c \in \mathbb{K}$

AM $a \cdot (b{+}c) = (a \cdot b) + (a \cdot c)$

Satz 1.43. *$\mathbb{Q}$ ist ein Körper. Definieren wir auf $\mathbb{Q}$ eine Ordnung "$\leq$"durch*

$$x \leq y \Leftrightarrow \exists m \in \mathbb{N},\ \ n \in \mathbb{N}^+ \ \ : \ \ y - x = \frac{m}{n}$$

dann ist diese Ordnung mit der Addition und der Multiplikation in $\mathbb{Q}$ im folgenden Sinne verträglich:

(AO) $a \leq b \Rightarrow a + c \leq b + c$

(MO)

$$\begin{cases} 0 \leq a \\ 0 \leq b \end{cases} \Rightarrow 0 \leq a \cdot b$$

(ohne Beweis)

1.6.2 Einführung der Reellen Zahlen

Motivation: *Wir haben schon gesehen, dass die Gleichung $x^2 - 2$ in $\mathbb{Q}$ keine Lösung hat. Allerdings können wir $\sqrt{2}$ "beliebig gut"durch $y \in \mathbb{Q}$ approximieren, d.h. $\forall \epsilon > 0$ beliebig klein $\exists y \in \mathbb{Q}:\ \ 2 - \epsilon \leq y^2 \leq 2 + \epsilon$.*

Das motiviert folgende Vorstellungen:

1. *$\mathbb{Q}$ ist "unvollständig"*
2. *$\mathbb{Q}$ ist "dicht in $\mathbb{R}$"*

Vollständigkeitsaxiom (Archimedes)

(V) Jede nach oben (unten) beschränkte Teilmenge von $\mathbb{R}$ hat ein Supremum (Infimum).

Axiomatischer Standpunkt
Es gibt eine Menge $\mathbb{R}$ (genannt Menge der reellen Zahlen) mit Addition, Multiplikation, Ordnung, die (A1)-(A4), (M1)-(M4), (AM), (A0), (M0) und (V) erfüllt.

Bemerkung 1.44. *1. Bis auf Isomorphie gibt es höchstens ein solches System $\mathbb{R}$, d.h. wenn $\tilde{\mathbb{R}}$ ein weiteres System der reellen Zahlen ist, dann existiert eine bijektive Abbildung $f : \mathbb{R} \to \tilde{\mathbb{R}}$ die bezüglich der Addition, Multiplikation und Ordnung ein Homomorphismus ist im folgenden Sinne:*
$\forall x, y \in \mathbb{R}$

(a) $f(x + y) = f(x) + f(y)$

(b) $f(xy) = f(x) \cdot f(y)$

(c) $x \leq y \Rightarrow f(x) \leq f(y)$

2. $\mathbb{N}$ *und damit auch* $\mathbb{Z}, \mathbb{Q}$ *lassen sich durch injektive Homomorphismen*

$$g : \mathbb{N}, \mathbb{Z}, \mathbb{Q} \to \mathbb{R} \text{ in } \mathbb{R} \text{ einbetten}$$

$$\begin{aligned} g(\tilde{0}) &= 0 \\ g(\tilde{1}) &= 1 \\ g(\underbrace{\tilde{n}+1}_{\mathbb{N}}) &= \underbrace{g(\tilde{n})}_{\mathbb{R}} + 1 \end{aligned}$$

Wir schreiben $g(\tilde{n}) = n$ *und können* g *auf* $\mathbb{Z}$ *definieren durch* $g(-\tilde{n}) = -n$ *und auf* $\mathbb{Q}$ *definieren durch* $g(\pm\frac{\tilde{n}}{\tilde{m}}) = \pm\frac{n}{m}$ $(\tilde{m} \neq 0)$.

Konstruktiver Standpunkt
Wir können $\mathbb{R}$ *ausgehend von* $\mathbb{Q}$ *konstruieren.*

1. *Methode der Abschnitte Jede reelle Zahl wird charakterisiert durch ein "rechts offenes, unbeschränktes Intervall", dessen "rechte Grenze"die Zahl darstellt*

$$\mathbb{R} := \left\{ A \subset \mathbb{Q} : \begin{cases} A \neq \emptyset \\ x \in A, \;\; y \leq x \Rightarrow y \in A \\ \forall x \in A \exists y \in A; \;\; x < y \end{cases} \right\}$$

2. *Methode der Cauchy-Folgen Jede reelle Zahl wird charakterisiert als "Grenzwertëiner Klasse "äquivalenter Cauchy-Folgenäus* $\mathbb{Q}$ *(⇒ spätere Vorlesungen)*

1.6.3 Grundlegende Strukturen und Funktionen in $\mathbb{R}$

Definition 1.45.

$$x \in \mathbb{R} \text{ heißt } \begin{cases} \text{positiv} & \Leftrightarrow \;\; x > 0 \\ \text{nichtnegativ} & \Leftrightarrow \;\; x \geq 0 \\ \text{negativ} & \Leftrightarrow \;\; x < 0 \\ \text{nichtpositiv} & \Leftrightarrow \;\; x \leq 0 \end{cases}$$

$\mathbb{R}^+ := \{x \in \mathbb{R} : x \text{ positiv}\}$

Definition 1.46. *Die Betragsfunktion* $|\cdot| : \mathbb{R} \to \mathbb{R}$ *wird definiert durch*

$$|x| = \max\{x, -x\} = \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0 \end{cases}$$

Die Vorzeichen- oder Signumfunktion sgn : $\mathbb{R} \to \mathbb{R}$ *wird definiert durch*

$$\operatorname{sgn} x = \begin{cases} \frac{x}{|x|} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases} = \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -1 & \text{für } x < 0 \end{cases}$$

Satz 1.47. *Sei* $x, y \in \mathbb{R}$

1. $|x \cdot y| = |x| \cdot |y|$
2. $|x + y| \leq |x| + |y|$
3. $|x + y| = |x| + |y| \Leftrightarrow x \cdot y \geq 0$ *(*$\Leftrightarrow x, y$ *haben gleiches Vorzeichen)*

Beweis.

$$\begin{aligned} 2. \quad |x+y|^2 = (x+y)^2 = x^2 + 2xy + y^2 = |x^2| + 2xy + |y|^2 \\ \leq |x|^2 + 2|xy| + |y|^2 = |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2 \\ \Rightarrow |x+y| \leq \big||x| + |y|\big| = |x| + |y| \geq 0 \end{aligned}$$

□

Satz 1.48. $\forall x, y \in \mathbb{R}$

1. $\big||x| - |y|\big| \leq |x - y|$
2. $|x - y| \leq \epsilon \Leftrightarrow \begin{cases} x - \epsilon \leq y \leq x + \epsilon \\ y - \epsilon \leq x \leq y + \epsilon \end{cases}$

Beweis. (1)

$$\begin{aligned} |x| = |x - y + y| \leq |x - y| + |y| \Rightarrow |x| - |y| \leq |x - y| \\ |y| = |y - x + x| \leq |y - x| + |x| \Rightarrow |y| - |x| \leq |y - x| \end{aligned}$$

$\Rightarrow$

$$\big||x| - |y|\big| = \max\{|x| - |y|, |y| - |x|\} \leq |x - y|$$

(2)

$$|x - y| = \max\{x - y, y - x\} \leq \epsilon$$

$$\Leftrightarrow \begin{cases} x-y \leq \epsilon \\ y-x \leq \epsilon \end{cases} \Leftrightarrow \begin{cases} x \leq y+\epsilon \\ y \leq x+\epsilon \end{cases} \Leftrightarrow \quad y-\epsilon \leq x \leq y+\epsilon$$

Vertausche x, y:

$$|y-x| = |x-y| \leq \epsilon \Leftrightarrow x-\epsilon \leq y \leq x+\epsilon$$

□

Definition 1.49. *Sei $a, b \in \mathbb{R}, \quad a \leq b$*

- $[a, b] = \{x \in \mathbb{R} : \quad a \leq x \leq b\}$ *abgeschlossenes Intervall*
- $(a, b) =]a, b[= \{x \in \mathbb{R} : \quad a < x < b\}$ *offenes Intervall*
- $[a, b) = [a, b[= \{x \in \mathbb{R} : \quad a \leq x < b\}$ *rechts-halboffenes Intervall*
- $(a, b] =]a, b] = \{x \in \mathbb{R} : \quad a < x \leq b\}$ *links-halboffenes Intervall*
- $\epsilon > 0, \quad I_\epsilon(a) := (a-\epsilon, a+\epsilon)$
- $I_\epsilon(x) := \{y \in \mathbb{R} : \quad |x-y| < \epsilon\}$
  Dann gilt: $y \in I_\epsilon(x) \Rightarrow \exists \delta > 0 \quad : \quad I_\delta(y) \subset I_\epsilon(x)$

  Beweis. Sei $y \in I_\epsilon(x)$, dann gilt $|x-y| < \epsilon \Leftrightarrow \epsilon - |x-y| > 0$. Wähle $0 < \delta < \epsilon - |x-y|$. Es ist nun zu zeigen: $I_\delta(y) \subset I_\epsilon(x)$, d.h. $z \in I_\delta(y) \Rightarrow z \in I_\epsilon(x)$
  Es gilt:

$$\begin{aligned} z \in I_\delta(y) &\Rightarrow |z-y| < \delta \\ \Rightarrow |z-x| = |z-y+y-x| \leq |z-y| + |y-x| &< \delta + |x-y| < \epsilon \\ &\Rightarrow z \in I_\epsilon(x) \end{aligned} \tag{1.3}$$

□

Definition 1.50. *A, B seien geordnete Mengen*
$f : A \to B$ heißt

$$\textit{monoton} \begin{cases} \textit{wachsend} & \Leftrightarrow x \leq y \Rightarrow f(x) \leq f(y) \\ \textit{fallend} & \Leftrightarrow x \leq y \Rightarrow f(x) \geq f(y) \end{cases}$$

$$\textit{streng monoton} \begin{cases} \textit{wachsend} & \Leftrightarrow x < y \Rightarrow f(x) < f(y) \\ \textit{fallend} & \Leftrightarrow x < y \Rightarrow f(x) > f(y) \end{cases}$$

Satz 1.51. *Die Abbildung*

$$\begin{aligned} \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto x^n \end{aligned}$$

ist streng monoton wachsend $\forall n \in \mathbb{N}^+$

Beweis. Induktion
$n = 1$ klar
$n \to n+1: \quad x < y \Rightarrow x^n < y^n$ nach Induktionsvorraussetzung.

$$\begin{cases} x \geq 0 \\ x^n < y^n \end{cases} \quad \Rightarrow \quad x^{n+1} < xy^n \text{ (M0)}$$

$$\begin{cases} y^n > 0 \\ x < y \end{cases} \quad \Rightarrow \quad xy^n < y^{n+1} \text{ (M0)}$$

$$\Rightarrow \quad x^{n+1} < y^{n+1}$$

□

Lemma 1.52. *Sei $M, N \subset \mathbb{R}$*
$f : M \to N$ streng monoton und bijektiv
$\Rightarrow f^{-1}$ streng monoton

Beweis. Wir betrachten den Fall:
f streng monoton wachsend. Seien

$$\begin{aligned} &y_1, y_2 \in \mathbb{N}, \ \ y_1 < y_2 \\ &x_1 = f^{-1}(y_1) \\ &x_2 = f^{-1}(y_2) \end{aligned}$$

Behauptung: $x_1 < x_2$. Sonst wäre $x_1 \geq x_2$:

Falls $x_1 > x_2 \ \Rightarrow y_1 = f(x_1) > f(x_2) = y_2$ wegen der Monotonie von f, ↯(zu $y_1 < y_2$)
Falls $x_1 = x_2 \ \Rightarrow y_1 = y_2$ ↯(Widerspruch zur Annahme $y_1 < y_2$)

□

Ein wichtiges Hilfsmittel für Beweise von Ungleichungen ist das Kriterium für die Positivität einer quadratischen Form zweier Variablen.

Satz 1.53. *Seien $a, b, c \in \mathbb{R}$*

$$F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
$$\textit{definiert durch } F(x,y) = ax^2 + 2bxy + cy^2$$

Dann gilt:

(i)

$$F(x,y) \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R}\setminus\{0\}) \Leftrightarrow \begin{cases} a & \geq 0 \;\; (> 0) \\ c & \geq 0 \\ ac - b^2 & \geq 0 \;\; (> 0) \end{cases}$$

(ii)

$$\begin{aligned} &F(x,y) \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R}\setminus\{0\}) \\ &\qquad \Rightarrow a^2x^2 + 2abxy + acy^2 \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R}\setminus\{0\}) \\ &\qquad \Leftrightarrow (ax+by)^2 + (ac-b^2)y^2 \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R}\setminus\{0\}) \end{aligned}$$

Beweis. (i)

1. "≥ 0"

<u>Fall 1</u>: $a = 0$

"$\Rightarrow$"

$F(x,y) = 2bxy + cy^2 \geq 0 \;\; \forall x,y \in \mathbb{R}$

$$\begin{aligned} &\Rightarrow F(0,1) \geq 0 \Leftrightarrow c \geq 0 \\ &\quad F(-c,b) \geq 0 \Leftrightarrow -b^2c \geq 0 \quad \Leftrightarrow \quad b = 0, \quad \text{falls } c > 0 \end{aligned}$$

Falls $c = 0$ gilt:

$$\begin{aligned} F(1,1) \geq 0 &\Leftrightarrow 2b \geq 0 \text{ und} \\ F(-1,1) \geq 0 &\Leftrightarrow -2b \geq 0 \\ &\Rightarrow b = 0 \end{aligned}$$

"$\Leftarrow$"

Umgekehrt

$$\begin{cases} c \geq 0 \\ b = 0 \end{cases} \quad \Rightarrow F(x,y) = cy^2 \geq 0 \;\; \forall (x,y) \in \mathbb{R}^2$$

<u>Fall 2</u>: $a \neq 0$

"$\Rightarrow$"

$$\begin{aligned} F(x,y) &= a\big(x^2 + 2\frac{b}{a}xy + \left(\frac{b}{a}y\right)^2\big) + \big(c - \frac{b^2}{a}\big)y^2 \\ &= a\big(x + \frac{b}{a}y\big)^2 + \frac{1}{a}(ac-b^2)y^2 \end{aligned}$$

$F(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2 \Rightarrow$

$$\begin{cases} F(1,0) & \geq 0 \\ F(-\frac{b}{a},1) & \geq 0 \end{cases} \Leftrightarrow \begin{cases} a & \geq 0 \\ \frac{1}{a}(ac-b^2) & \geq 0 \end{cases} \Leftrightarrow \begin{cases} a & \geq 0 \\ ac-b^2 & \geq 0 \end{cases} \begin{cases} a & \geq 0 \\ c & \geq \frac{b^2}{a} \geq 0 \\ ac-b^2 & \geq 0 \end{cases}$$

" $\Leftarrow$ "
Umgekehrt, sei $(x,y) \in \mathbb{R}^2$ beliebig:

$$a \geq 0 \quad \Rightarrow \quad a\big(x + \frac{b}{a}y\big)^2 \geq 0$$
$$a > 0, \;\; ac - b^2 \geq 0 = \frac{1}{a}(ac-b^2)y^2 \geq 0$$

2. ">"ähnlich

(ii) Aus $F(x,y) \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R} \setminus \{0\})$ folgt $F(1,0) = a \geq 0 \;\; (> 0)$. Somit folgt auch $a \cdot F(x,y) \geq 0 \;\; \forall x,y \in \mathbb{R} \;\; (> 0 \;\; \forall x,y \in \mathbb{R} \setminus \{0\})$.

Da

$$\begin{aligned} a \cdot F(x,y) &= a^2x^2 + 2abxy + acy^2 \\ &= (ax+by)^2 + (ac-b^2)y^2 \end{aligned}$$

folgt die Behauptung. □

Bemerkung 1.54. *Wir können die quadratische Form schreiben als*

$$F(x,y) = (x,y) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

Wir wissen aus der linearen Algebra: Eine Matrix A heißt positiv (semi-)definit, wenn $x'Ax > 0 \; (\geq 0) \quad \forall x \neq 0$. Eine Matrix ist genau dann positiv definit, wenn alle führenden Hauptminoren positiv sind. Das entspricht genau dem oben bewiesenen (s. (i) für den Fall "$F(x,y) > 0$"). Für positiv semidefinite Matritzen gibt es ein vergleichbares, jedoch etwas komplizierteres Kriterium. Diese Verbindung zur lineraen Algebra erlaubt ein zwangloses verallgemeinern des obigen Resultates auf quadratische Formen in mehr als zwei Variablen. Positiv definite Matritzen spielen bei der Definition von Skalarprodukten (s. später) eine wichtige Rolle.

1.6.4 Vollständigkeit

$\mathbb{Q}$ ist nicht vollständig (es gibt "Löcher" innerhalb der rationalen Zahlen), z.B. ist die Gleichung $x^2 = 2$ in $\mathbb{Q}$ nicht lösbar. In $\mathbb{R}$ gibt es aber das Vollständigkeitsaxiom, aus dem wir folgern:

Satz 1.55. $\forall m \in \mathbb{N}^+, \ \ a \in [0, \infty)$ *gilt:*

$$\exists x \in [0, \infty) : x^m = a.$$

Wir schreiben:

$$x := \sqrt[m]{a} = a^{\frac{1}{m}}$$

Beweis. Wir geben ein Iterationsverfahren zur Berechnung von x, o.B.d.A. $a > 0, \ m \geq 2$, x muss die Gleichung $x^m - a = 0$, lösen, d.h. Nullstelle der Funktion

$$f : (0, \infty) \to \mathbb{R}$$
$$x \mapsto x^m - a$$

sein.
Diese approximieren wir nach dem Newton-Verfahren

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^m - a}{m x_n^{m-1}} = x_n \left(1 - \frac{1}{m}\left(1 - \frac{a}{x_n^m}\right)\right)$$

$$f'(x_n) = \frac{f(x_n)}{x_n - x_{n-1}} = \tan(\alpha)$$

Also $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
Wir hoffen, dass $x_n \to x^*$
Sei $x_0^m \geq a$. Wir zeigen: $\forall n \in \mathbb{N}$

1. $x_n > 0$
2. $x_n^m \geq a$
3. $x_{n+1} \leq x_n$

Induktion:
(1),(2): $n = 0$, Induktionsannahme: $x_0^m \geq a \Rightarrow x_0 > 0$, da $a > 0, x_0 \geq 0$
$n \to n+1 :$ $x_n > 0, x_n^m \geq a \Rightarrow 1 \geq \frac{1}{m}\left(1 - \frac{a}{x_n^m}\right) \geq 0 \Rightarrow x_{n+1} \geq 0$
$x_{n+1}^m = \underbrace{x_n^m}_{\geq 0} \left(1 + \left[-\frac{1}{m}\left(1 - \frac{a}{x_n^m}\right)\right]\right)^m \underbrace{\geq}_{\text{Bernoulli-Ungleichung}} x_n^m \left(1 - m\left(\frac{1}{m}\left(1 - \frac{a}{x_n^m}\right)\right)\right) =$
$x_n^m \left(1 - 1 + \frac{a}{x_n^m}\right) = a$
$\Rightarrow x_{n+1} > 0$, da $a > 0; x_{n+1} \geq 0$
(3): $x_{n+1} = x_n \left(1 - \frac{1}{m}\left(1 - \frac{a}{x_n^m}\right)\right)$
nach (2): $x_n^m \geq a \Rightarrow 0 \leq 1 - \frac{1}{m}\left(1 - \frac{a}{x_n^m}\right) \leq 1$
nach (1): $x_n > 0 \Rightarrow x_n \left(1 - \frac{1}{m}\left(1 - \frac{a}{x_n^m}\right)\right) \leq x_n$

Wegen (1) ist $M = \{x_n : \ n \in \mathbb{N}\}$ nach unten beschränkt, so dass $x := \inf M$ existiert. Wir wollen zeigen, dass $x^m = a$:

$$x \leq x_{n+1} = \left(1 - \frac{1}{m}\right) x_n + \frac{1}{m} \frac{a}{x_n^{m-1}} \leq \left(1 - \frac{1}{m}\right) x_n + \frac{a}{m} \underbrace{\sup \left\{ \frac{1}{x_n^{m-1}} \middle| n \in \mathbb{N} \right\}}_{:=y}$$

$$\Rightarrow \quad x \leq \left(1 - \frac{1}{m}\right) x + \frac{a}{m} y$$

$$\Rightarrow x \leq ay$$

Dann nutzen wir, dass für eine Folge $(y_k)_{k \in \mathbb{N}}$, $y_k > 0 \quad \forall n \in \mathbb{N}$

$$\inf \{y_n^k : \ n \in \mathbb{N}\} = (\inf \{y_n : \ n \in \mathbb{N}\})^k$$

(Beweis s. Übungen).

Es gilt nach (2):

$$a \leq \inf \{x_n^m : \ n \in \mathbb{N}\} = (\inf \{x_n : \ n \in \mathbb{N}\})^m = x^m \quad (4)$$

und damit $x > 0$.
Ferner gilt:

$$y = \sup \left\{ \frac{1}{x_n^{m-1}} : \ n \in \mathbb{N} \right\} s = \left(\inf \{x_n^{m-1} : \ n \in \mathbb{N}\}\right)^{-1} = (*),$$

weil für $y_n > 0$, $n \in \mathbb{N}$ und $\inf \{y_n : \ n \in \mathbb{N}\} > 0$ gilt: $\sup \{\frac{1}{y_n} : \ n \in \mathbb{N}\} = (\inf \{y_n : \ n \in \mathbb{N}\})^{-1}$.

Wir bekommen:

$$(*) = \left(\frac{1}{\inf\{x_n : \ n \in \mathbb{N}\}} \right)^{m-1} = \frac{1}{x^{m-1}} \Rightarrow y \leq \frac{1}{x^{m-1}} \quad \Rightarrow ay \leq \frac{a}{x^{m-1}}$$

Von oben wissen wir $x \leq ay$, daher erhalten wir: $x \leq ay \leq \frac{a}{x^{m-1}}$. $\Rightarrow$
(5) $\quad x^m \leq a$. Aus (4) und (5) folgt $x^m = a$. $\square$

Bemerkung: Gerade hatten wir es mit einem “Approximationsproblem” und “Konvergenz einer Folge” zu tun.

Definition 1.56. *Eine Folge von reellen Zahlen wird gegeben durch eine Abbildung*

$$\mathbb{N} \to \mathbb{R}$$
$$n \mapsto x_n$$

Wir bezeichnen die Folge auch mit $\left((x_n)_{n \in \mathbb{N}}\right)$.

Definition 1.57. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Wir sagen x_n konvergiert gegen $x \in \mathbb{R}$*

$$x = \lim_{n\to\infty} x_n \quad \textit{oder} \quad x_n \rightarrow_{n\to\infty} x$$
$$\Leftrightarrow \forall \epsilon > 0 \ \exists n_0 \in \mathbb{N} : \ \forall n \geq n_0 \quad x_n \in I_\epsilon(x)$$

Folgerung
Sei $(x_n)_{n\in\mathbb{N}}$ eine monoton wachsende bzw. fallende Folge reeller Zahlen, $M = \{x_n : n \in \mathbb{N}\}$ sei nach oben bzw. nach unten beschränkt.
Dann gilt:

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

Für den Beweis benutzen wir folgendes Lemma:

Lemma 1.58. *Sei $M \in \mathbb{R}$ nach oben bzw. nach unten beschränkt. Dann gilt für $s \in \bar{S}(M)$ bzw. $l \in \underline{S}(M)$:*

1. $s = \sup M \Leftrightarrow \forall \epsilon > 0 \ \exists x \in M : \ s - \epsilon < x \ (\leq s)$
2. $l = \inf M \Leftrightarrow \forall \epsilon > 0 \ \exists x \in M : \ (l \leq) x < l + \epsilon$

Beweis von 1.
$s \neq \sup M \Leftrightarrow s$ ist nicht kleinste obere Schranke von $M \Leftrightarrow$ Es gibt eine kleinere obere Schranke $s' = s - \epsilon$ von $M \Leftrightarrow$ nicht $(\forall \epsilon > 0 \ \exists x \in M : \ x > s - \epsilon)$
Zum Beweis der Folgerung:

Beweis. Sei $s = \sup M$
$\forall \epsilon > 0 \ \exists x \in M : \ s - \epsilon < x \leq s; \quad x = x_{n_0}, \ n_0 \in \mathbb{N}$
Aus der Monotonie folgt: $\forall n \geq n_0 : s - \epsilon < x_n \leq s$
d.h. $x_n \in I_\epsilon(s)$. □

Definition 1.59. *Seien $\emptyset \neq A \subset B \subset \mathbb{R}$*
A heißt dicht in $B \Leftrightarrow \forall x \in B \quad \forall \epsilon > 0$
$I_\epsilon(x) \cap A \neq \emptyset$

Definition 1.60. *Die durch $[x] := \max\{z \in \mathbb{Z} : z \leq x\}$ definierte Funktion $[\cdot] : \mathbb{R} \to \mathbb{Z}$ heißt Gauß-Klammer*

Satz 1.61. *$\mathbb{Q}$ ist dicht in $\mathbb{R}$*

Beweis. Wir müssen zeigen:
$\forall x \in \mathbb{R} \ \forall \epsilon > 0 \ \exists y \in \mathbb{Q} : |x - y| < \epsilon$
Wir wählen bei gegebenem $x \in \mathbb{R}$, $\epsilon > 0$ ein $n \in \mathbb{N}$ mit $n > \frac{1}{\epsilon}$ (n existiert,

weil $\mathbb{N}$ unbeschränkt in $\mathbb{R}$)
Dann gilt:

$$\begin{aligned} &[nx] \leq nx \leq [nx] + 1 \\ \Rightarrow\ &\frac{[nx]}{n} \leq x \leq \frac{[nx]}{n} + \frac{1}{n} < \frac{[nx]}{n} + \epsilon \\ \Rightarrow\ &|x - \frac{[nx]}{n}| < \epsilon \end{aligned}$$

und $\frac{[nx]}{n} \in \mathbb{Q}$ $\square$

Definition 1.62. *Sei* $\hat{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$.

Bemerkung 1.63. *Auf* $\hat{\mathbb{R}}$ *können wir die Ordnung von* $\mathbb{R}$ *fortsetzen durch*

$$-\infty < x < +\infty \quad \forall x \in \mathbb{R}$$

Bemerkung 1.64. *Sei* $M \subset \hat{\mathbb{R}}$ *eine beschränkte Menge, dann gilt*

$$-\infty \in \underline{S}(M), \quad \infty \in \bar{S}(M).$$

Falls $s \in \bar{S}(M)$ *für* $s \in \mathbb{R} \Rightarrow$ M *ist in* $\mathbb{R}$ *nach oben beschränkt,* $\sup M \in \mathbb{R}$ *existiert.*
Falls

$$\bar{S}(M) = \{\infty\}$$

gilt

$$\sup M = \infty,$$

M *ist nach oben unbeschränkt. Analoges gilt für* inf.

Beispiel 1.65.

$$\begin{aligned} \sup \mathbb{N} = \sup \mathbb{Z} = \sup \mathbb{R} = \infty \\ \inf \mathbb{Z} = \inf \mathbb{Q} = \inf \mathbb{R} = -\infty \end{aligned}$$

1.7 Vektorraum

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}}_{n}$$

In der linearen Algebra werden die Eigenschaften von $\mathbb{R}^n$ als n-dimensoionaler Vektorraum untersucht (mit Addition sowie skalarer Multiplikation mit $\lambda \in \mathbb{R}$). Im Allgemeinen ist es unmöglich auf $\mathbb{R}^n$ eine geeignete Multiplikation einzuführen, so dass die Körpereigenschaften erfüllt sind.

- für $n = 2 \quad \Rightarrow \quad$ komplexe Zahlen
- für $n = 4 \quad \Rightarrow \quad$ nicht-kommutative Multiplikation (SSchiefkörper der Quaternionen")

Definition 1.66. *Sei* $\mathbb{K}$ *ein Körper.*
$|\cdot| : \mathbb{K} \to \mathbb{R}$ *heißt Betragsfunktion auf* $\mathbb{K}$ $:\Leftrightarrow$

(B1) $\forall x \in \mathbb{K}: \ |x| \geq 0$ *und* $(|x| = 0 \ \Leftrightarrow \ x = 0)$

(B2) $\forall x, y \in \mathbb{K}: \ |xy| = |x| \cdot |y|$

(B3) $\forall x, y \in \mathbb{K}: \ |x + y| \leq |x| + |y|$

Beispiel 1.67.

$$x \to \sqrt{x\bar{x}} =: |x| \ \textit{ auf } \mathbb{C}$$

Definition 1.68. *Sei* $\mathbb{K}$ *ein Körper mit einer Betragsfunktion* $|\cdot|$. *Sei* V *ein Vektorraum über* $\mathbb{K}$
$||\cdot|| : V \to \mathbb{R}$ *heißt eine Norm auf* V $:\Leftrightarrow$

(N1) $\forall x \in V; \ ||x|| \geq 0$ *und* $(||x|| = 0 \Leftrightarrow \ x \equiv 0)$

(N2) $\forall x \in V, \ \alpha \in \mathbb{K}: \ ||\alpha x|| = |\alpha| \cdot ||x||$

(N3) $\forall x, y \in V: \ ||x + y|| \leq ||x|| + ||y||$ *(Dreieckungleichung)*

$(V, ||\cdot||)$ *heißt normierter Vektorraum*

(N1),(N3) analog zu (B1),(B3)
Bei (N2) handelt es sich um eine Abschwächung von (B2)

Beispiel 1.69.
- $||x|| := |x|$ *auf* $V = \mathbb{R}^2 = \mathbb{C}-$ *Vektorraum über* $\mathbb{K} = \mathbb{C}$
  In diesem Fall sind die Norm-Eigenschaften gerade die Eigenschaften der Betragsfunktion.
- $V = \mathbb{R}^n, \ \mathbb{K} = \mathbb{R}, \ x = (x_1, ..., x_n) \in \mathbb{R}^n$ *(∞-Norm)*

$$||x||_\infty := \max |x_1|, |x_2|, ..., |x_n|$$

(N1),(N2) sind trivial
(N3):

$$\begin{aligned} ||x + y||_\infty &= \max\{|x_1 + y_1|, ..., |x_n, y_n|\} \\ &\leq \max\{|x_1| + |y_1|, ..., |x_n| + |y_n|\} \\ &\leq \max\{|x_1|, ..., |x_n|\} + \max\{|y_1|, ..., |y_n|\} \\ &= ||x||_\infty + ||y||_\infty \end{aligned}$$

- $V = \mathbb{R}^n, \ \ \mathbb{K} = \mathbb{R}, \ \ p \in \mathbb{Q}, \ \ p \geq 1$ *(p-Norm)*
  $x \in \mathbb{R}^n$
  $$||x||_p := \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}$$
  Der allgemeine Beweis von (N3) ist lang.
  Wir werden später den Fall $p = 2$ betrachen und zeigen auch, dass
  $$\forall x \in \mathbb{R}^n \ \ ||x||_p \to ||x||_\infty, \quad \textit{für} \quad p \to \infty$$

Mit Normen können wir Längen von Vektoren und Abstände zwischen Punkten messen. Wir können auch Umgebungen in Vektorräumen definieren.

Bemerkung 1.70. *Im $\mathbb{R}^n$ betrachten wir Vektoren aus n Komponenten. Im nächsten Kapitel werden wir Folgen kennen lernen. Folgen können in gewisser Weise als "Vektoren mit abzählbar unendlich vielen Komponenten" betrachtet werden. Folgen können dann ebenfalls mit den obigen Normen ausgestattet werden, mit dem Unterschied, dass nicht mehr über endlich viele sondern über unendlich viele Komponenten (Folgenglieder) summiert wird. Der Folgenraum l^p ist dann der Raum genau der Folgen mit endlicher $p-$Norm. Als Spezialfälle erhält man l^1 als den Raum der absolut summierbaren Folgen und l^∞ als den Raum der beschränkten Folgen.*

Bemerkung 1.71. *In späteren Kapiteln werden wir Funktionen betrachten, deren Absolutbetrag zur $p-$ten Potenz integirerbar ist, d.h. dass das Integral einen endlichen Wert annimmt. Diese Funktionen bilden den Raum L^p. Wir werden in Zukunft Zusammenhänge zwischen Integralen und Grenzwerten von Summen kennenlernen. Die sog. L^p- Norm kann dann als Verallgemeinerung der oben eingeführten $p-$Norm erkannt werden. Umgekehrt kann man obige $p-$Norm als Integral in einem "diskreten" Raum erhalten und so ihren Zusammenhang mit der L^p bzw. l^p Norm erkennen.*

Definition 1.72. *Sei V ein normierter Vektorraum, $x \in V, \ \ r > 0$*
$$K_r(x) := \{y \in V : ||y - x|| < r\}$$
heißt offene r-Kugel um x.

Beispiel 1.73.

1. $\mathbb{R}^n, \ \ ||\cdot||_2$ *(euklidische Norm)*
   $k_1(0) = \{x : \sum\limits_{i=1}^{n} {y_i}^2 < 1\}$
2. $\mathbb{R}^n, \ \ ||\cdot||_\infty$
   $k_1(0)=\{y : |y_1| < 1, ..., |y_n| < 1\}$
3. $\mathbb{R}^n, \ \ ||\cdot||_1$
   $k_1(0) = \{y : \sum\limits_{i=1}^{n} |y_i| < 1\}$

Die euklidische Norm ist mit einem Skalarprodukt verubunden.

Definition 1.74. *Sei das Skalarprodukt von $x, y \in \mathbb{R}$ definiert als*

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$$

Lemma 1.75. *Es gilt:*
$\forall x, y, z \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}$

(SP1) $\langle x, y\rangle = \langle y, x\rangle$

(SP2) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$

(SP3) $\langle x, \alpha y\rangle = \alpha\langle x, y\rangle$

(SP4)
1. $\langle x, x\rangle \geq 0$
2. $\langle x, x\rangle = 0 \quad \Leftrightarrow \quad x = 0$

Die euklidische Norm erhalten wir als

$$||x||_2 = \sqrt{\langle x, x\rangle}$$

Verallgemeinerung

Definition 1.76. *Sei V ein Vektorraum über $\mathbb{R}$. Sei $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ eine Funktion, die (SP1-4) erfüllt. Dann heißt $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf V.*

Bemerkung 1.77. *Das Skalarprodukt kann geschrieben werden als $\langle x, x\rangle = x'x$, wobei $'$ die Transposition bedeutet. Sei A eine positiv definite symmetrische $(n \times n)$ reelle Matrix, dann ist $x'Ax$ ebenfalls ein Skalarprodukt auf $\mathbb{R}^n$ (vgl. Bemerkung 1.54).*

Satz 1.78. *Cauchy-Schwarz-Ungleichung (CSU)*
Sei V ein reeller Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.
Dann gilt $\forall x, y \in V$:

1. $|\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle$
2. $|\langle x, y\rangle|^2 = \langle x, x\rangle\langle y, y\rangle \quad \Leftrightarrow \quad x, y$ *linear abhängig*

Beweis. Betrachte die quadratische Form

$$\begin{aligned} F : \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (\alpha, \beta) &\mapsto \underbrace{\langle x, x\rangle}_{a} \alpha^2 + 2\underbrace{\langle x, y\rangle}_{b} \alpha\beta + \underbrace{\langle y, y\rangle}_{c} \beta^2 \end{aligned}$$

$$\begin{aligned}\forall(\alpha,\beta)\in\mathbb{R}^2: F(\alpha,\beta)=\langle\alpha x+\beta y,\alpha x+\beta y\rangle\geq 0 \quad &\text{nach} \quad \text{(SP4,2)}\\ \Rightarrow ac-b^2\geq 0 \quad\Leftrightarrow\quad &\langle x,x\rangle\langle y,y\rangle-\langle x,y\rangle^2\geq 0\\ &\Rightarrow |\langle x,y\rangle|^2\leq\langle x,x\rangle\langle y,y\rangle\end{aligned}$$

(s. Satz 1.53)

2.

$$\begin{aligned}\exists(\alpha,\beta)\in\mathbb{R}^2\setminus\{(0,0)\}:\quad F(\alpha,\beta)=0\Leftrightarrow\\ \text{nicht}:(\forall(\alpha,\beta)\in\mathbb{R}^2\setminus\{(0,0)\}:\quad F(\alpha,\beta)>0)\Leftrightarrow\\ \langle x,x\rangle=0 \text{ oder } \langle y,y\rangle=0 \text{ oder } \langle x,x\rangle\langle y,y\rangle-\langle x,y\rangle^2=0, \text{ d.h. } |\langle x,y\rangle|^2=\langle x,x\rangle\langle y,y\rangle\end{aligned}$$

(vgl. Satz 1.53)

$\Leftrightarrow$

$$\begin{cases}x=0\\ \text{oder}\\ y=0\\ \text{oder}\\ \exists(\alpha,\beta)\in\mathbb{R}^2\setminus\{(0,0)\}:\quad\langle\alpha x+\beta y,\alpha x+\beta y\rangle=0\end{cases}$$

$\Leftrightarrow$

$$\begin{cases}x=0\\ \text{oder}\\ y=0\\ \text{oder}\\ \exists(\alpha,\beta)\in\mathbb{R}^2\setminus\{(0,0)\}:\quad\alpha x+\beta y=0\quad\text{nach (SP4 (ii))}\end{cases}$$

$\Leftrightarrow$ x,y linear abhängig □

Definition 1.79. *Sei V reeller Vektorraum mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann heißt*

$$||\cdot||: x\to\sqrt{\langle x,x\rangle}$$

die zum Skalarprodukt $\langle\cdot,\cdot\rangle$ gehörige Norm auf V.

Satz 1.80. *Die oben genannte Definition definiert eine Norm auf V*

Beweis. (N1) $\Leftrightarrow$ (SP4)
(N2)

$$\begin{aligned} ||\alpha x|| = \sqrt{\langle \alpha x, \alpha x\rangle} \underbrace{=}_{\text{(SP3)}} \sqrt{\alpha\langle \alpha x, x\rangle} \underbrace{=}_{\text{(SP1)}} \sqrt{\alpha\langle x, \alpha x\rangle} \\ \underbrace{=}_{\text{(SP3)}} \sqrt{\alpha^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = |\alpha| \cdot ||x|| \end{aligned}$$

(N3)

$$\begin{aligned} ||x+y||^2 = \langle x+y, x+y\rangle = ||x||^2 + 2\langle x, y\rangle + ||y||^2 \\ \underbrace{\leq}_{\text{CSU}} ||x||^2 + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + ||y||^2 \\ = ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2 \\ = (||x|| + ||y||)^2 \end{aligned}$$

□

Beispiel 1.81.

$$V = \mathbb{R}^n \quad \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

1. *Seien*

$$\begin{aligned} e_1 &= (1, 0, ..., 0) \\ e_2 &= (0, 1, ..., 0) \\ &\dots \\ e_n &= (0, 0, ..., 1) \end{aligned}$$

Dann gilt

$$\langle e_j, e_k\rangle = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases} =: \delta_{jk}$$

$\forall x, y \in \mathbb{R}^n$ *gilt*

$$\begin{aligned} \langle x, y\rangle &= \langle ||x|| \frac{x}{||x||}, ||y|| \frac{y}{||y||}\rangle \\ &= ||x|| ||y|| \langle \frac{x}{||x||}, \frac{y}{||y||}\rangle \end{aligned}$$

2. *Geometrische Interpretation:*
$x, y \in \mathbb{R}^n \quad ||x|| = 1$
Betrachte die Funktion

$$c : \mathbb{R} \to \mathbb{R} \quad c(s) := ||sx - y||^2$$
$$c(s) = s^2 - 2s\langle x, y\rangle + ||y||^2 = (s - \langle x, y\rangle)^2 + \underbrace{||y||^2 - \langle x, y\rangle^2}_{c(\langle x,y\rangle)}$$
$$\Rightarrow \begin{cases} c(\langle x, y\rangle) = ||y||^2 - \langle x, y\rangle^2 \\ c(s) = (s - \langle x, y\rangle)^2 + c(\langle x, y\rangle)\underline{\geq c(\langle x, y\rangle)} \end{cases}$$

Die Funktion $c(s) = ||y - \xi||$ *mit* $\xi = sx$ *gibt den Abstand an zwischen* y *und dem Punkt* ξ *auf der von* x *aufgespannten Geraden. Sie nimmt an der Stelle* $s_0 = \langle x, y\rangle$ *das Minimum an,* $\langle x, y\rangle x = \xi_0$ *ist also die "Projektion"von* y *auf diese Gerade,*
$$c(\langle x, y\rangle) = ||y - \langle x, y\rangle x||^2$$
der Abstand von y *zur Geraden.*

Das Dreieck $(0, \xi_0, y)$ *ist "rechtwinklig", denn es gilt der Satz von Pythagoras*

$$\begin{aligned} ||\xi_0||^2 + ||y - \xi_0||^2 &= \langle x, y\rangle^2 + c(\langle x, y\rangle) \\ &= \langle x, y\rangle^2 + ||y||^2 - \langle x, y\rangle^2 \\ &= ||y||^2 \end{aligned}$$

Kapitel 2

Konvergente Folgen und Reihen

Bisher haben wir uns mit *algebraischen Operationen* (Addition, Multiplikation) in geeigneten Strukturen (z.B. Körpern) beschäftigt. Im Folgenden werden wir eine *topologische Struktur* auf Mengen einführen, die es uns erlaubt, *Umgebungen* zu definieren und *Abstände* zu messen. Dieses soll die Begriffe des Betrages und des ϵ-Intervalles, wie wir sie für Zahlen bzw. die Begriffe der Norm und der ϵ-Kugel, wie wir sie für Vektorräume eingeführt haben, verallgemeinern. Wir werden dann den Begriff der *Konvergenz* einführen, der eng mit den Begriffen der *Umgebung* und des *Abstandes* zusammenhängt.

Für den $\mathbb{R}^n$ haben wir verschiedene Normen kennengelernt. Die folgende Bemerkung zeigt, dass es keine Rolle spielt, welche spezielle Norm wir betrachten.

Bemerkung 2.1. *Sind* $||\cdot||$ *und* $||\cdot||_*$ *Normen auf* $\mathbb{R}^n$*, so gibt es* $c_1, c_2 \in \mathbb{R}^+$ *mit* $c_1||x|| \leq ||x||_* \leq c_2||x|| \quad \forall x \in \mathbb{R}^n$*. Man sagt auch: "Die Normen* $||\cdot||$ *und* $||\cdot||_*$ *sind äquivalent."*

Um Umgebung und Konvergenz zu definieren, brauchen wir uns nicht auf Vektorräume zu beschränken. Wir können auch eine Abstandsfunktion (*Metrik)* auf jeder beliebigen Menge definieren. Wir erhalten dann die Struktur eines *metrischen Raumes.*

2.1 Metrische Räume

Definition 2.2. *Sei* M *eine Menge;* $d : M \times M \to \mathbb{R}$ *heißt* <u>*Metrik*</u> *auf* M*, genau dann, wenn* $\forall x, y, z \in M$

(D1) $(x, y) \geq 0;\ d(x, y) = 0 \Leftrightarrow\ x = y$ *(Positivität)*

(D2) $d(x, y) = d(y, x)$ *(Symmetrie)*

(D3) $d(x,y) \leq d(x,z) + d(z,y)$ *(Dreieckungleichung)*

Man nennt die Funktion $d(x,y)$ *die Abstandsfunktion des metrischen Raumes* (M,d)

Beispiele:
Ist $||\cdot|| : V \to \mathbb{R}$ eine Norm auf einem Vektorraum V, dann definiert $d(x,y) := ||x-y||$ eine Metrik auf V.
Beachte: $d(y,x) = ||y-x|| = ||-(x-y)|| = |-1|||x-y|| = ||x-y|| = d(x,y)$
sowie: $d(x,z) = ||x-z|| = ||(x-y)+(y-z)|| \leq ||x-y||+||y-z|| = d(x,y)+d(y,z)$

Definition 2.3. *Sei M eine Menge mit Metrik d.*
Wir definieren für $x \in M, \epsilon > 0$ die offene ϵ-Kugel um x durch

$$K_\epsilon(x) := \{y \in M : d(x,y) < \epsilon\}$$

$A \subset M$ heißt Umgebung von $x \in M$
$\Leftrightarrow \exists \epsilon : K_\epsilon(x) \subset A$

In Worten: $A \subset M$ heißt Umgebung von einem Punkt $x \in M$, wenn es eine offene $\epsilon-$Kugel um x gibt, die in A enthalten ist.

2.2 Zahlenfolgen, Konvergenz

Definition 2.4. *Sei M eine Menge mit einer Metrik d, $(x_k)_{k\in\mathbb{N}}$ eine Folge in M*

- *$x \in M$ heißt Häufungswert der Folge $(x_k)_{k\in\mathbb{N}} \Leftrightarrow \forall$ Umgebungen U von x gilt: $\{k : x_k \in U\}$ ist unendlich.*
- *$x \in M$ heißt Grenzwert der Folge $(x_k)_{k\in\mathbb{N}}$, $x = \lim\limits_{k\to\infty} x_k$,*
  $\Leftrightarrow$ *für alle Umgebungen U von x $\exists k_0 \in \mathbb{N}$: $x_k \in U$ $\forall k \geq k_0$*
  $\Leftrightarrow$ $\forall \epsilon > 0$ $\exists k_0 \in \mathbb{N}$: $d(x_k, x) < \epsilon$ $\forall k \geq k_0$
- *Eine Folge, die einen Grenzwert besitzt, heißt konvergent.*

Anmerkung: Für eine Folge $(x_k)_{k\in\mathbb{N}}$ schreiben wir auch vereinfachend (x_k).

Beispiel 2.5. *1. $M = \mathbb{C}$ mit der von der Norm $|\cdot|$ erzeugten Metrik*

$$x_k := i^k$$
$$x_0 = 1$$
$$x_1 = i$$
$$x_2 = -1$$
$$x_3 = -i$$
$$x_4 = 1$$
$$\ldots$$

Die Folge (x_k) nimmt die Werte $i, -1, -i, 1$ an und zwar jeden der Werte unendlich oft. Daher sind $i, -1, -i, 1$ die Häufungswerte der Folge.

2. *Um die Konvergenz einer Folge (x_n) gegen x zu zeigen muss zu vorgegebenem $\epsilon > 0$ ein n_0 angegeben (berechnet) werden, so dass*

$$\forall n \geq n_0 : d(x_n, x) < \epsilon$$

Die gelingt unter Umständen durch die Rückführung des Problems auf bekannte konvergente Folgen, wie z.B.

(a) $\lim\limits_{n\to\infty} \frac{1}{n} = 0,$ *denn*
$\frac{1}{n} \to \inf\{\frac{1}{n} : n \in \mathbb{N}^+\} = 0$
Daraus folgt sofort:

(b) $\lim\limits_{n\to\infty} \frac{1}{n^p} = 0$ *(für $p \in \mathbb{Q}^+$),*
denn sei $\epsilon > 0$ vorgegeben,
damit $\frac{1}{n^p} < \epsilon$ muss $n > \frac{1}{\epsilon^{\frac{1}{p}}}$ sein.
Wir wählen also $n_0 > \frac{1}{\epsilon^{\frac{1}{p}}}$
$\Rightarrow \forall n \geq n_0 : n > \frac{1}{\epsilon^{\frac{1}{p}}} \Rightarrow \frac{1}{n^p} < \epsilon$ $\square$

(c) $\lim\limits_{n\to\infty} a^{\frac{1}{n}} = 1$ *(für $a > 0$)*

Beweis. Mit der Bernoulli'schen Ungleichung erhalten wir:

$$a = (1 + a^{\frac{1}{n}} - 1)^n \geq 1 + n(a^{\frac{1}{n}} - 1) \geq n(a^{\frac{1}{n}} - 1) \Rightarrow \; a^{\frac{1}{n}} - 1 \leq \frac{a}{n}$$

i. Falls $\underline{a \geq 1}$ ist, folgt

$$|a^{\frac{1}{n}} - 1| = a^{\frac{1}{n}} - 1 \leq \frac{a}{n}$$

Dann erhalten wir, dass für vorgegebenes $\epsilon > 0$ und $n_0 > \frac{a}{\epsilon}$ folgt:

$$|a^{\frac{1}{n}} - 1| \leq \frac{a}{n_0} < \epsilon \quad \forall n \geq n_0$$

ii. Wir führen diesen Fall auf Fall (i) zurück indem wir $1/a$ anstatt a betrachten: Falls $\underline{a < 1}$ ist gilt

$$\frac{1}{a} \geq 1 \Rightarrow \; \frac{1}{a^{\frac{1}{n}}} = \left(\frac{1}{a}\right)^{\frac{1}{n}} \to 1 \;\; \text{(wegen (i))}$$

Daher: $\forall \epsilon > 0 \; \exists$ ein $n_0 \in \mathbb{N}$ mit

$$|\frac{1}{a^{\frac{1}{n}}} - 1| < \epsilon \quad \forall n \geq n_0$$
$$\Rightarrow 1 - a^{\frac{1}{n}} < \epsilon a^{\frac{1}{n}} < \epsilon$$

Denn
$$a \leq 1 \Rightarrow \quad a^{\frac{1}{n}} \leq 1 \Rightarrow |1 - a^{\frac{1}{n}}| = 1 - a^{\frac{1}{n}} < \epsilon$$

□

(d) $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$ *(ohne Beweis)*

(e) $\lim_{n\to\infty} \frac{z^n}{n!} = 0 \quad \forall z \in \mathbb{C}$

Beweis. Es genügt, die Behauptung für $z \in \mathbb{R}^+$ zu zeigen, da $|\frac{z^n}{n!}| = \frac{|z|^n}{n!}$, $|z| \in \mathbb{R}^+$ für $z \neq 0$
Wir schreiben
$$a_n := \frac{z^n}{n!}$$
Dann gilt
$$a_n = \frac{z}{n} a_{n-1} \leq \frac{1}{2} a_{n-1} \quad \text{für} \quad n \geq n_0 \quad \text{wenn} \quad n_0 \geq 2z$$
Also gilt
$$\forall n > n_0 : \quad a_n \leq a_{n_0} \Rightarrow \quad a_n = \frac{z}{n} a_{n-1} \leq \frac{z a_{n_0}}{n}$$
Für vorgegebenes $\epsilon > 0$ wählen wir daher $n_1 \geq n_0,\ n_1 > \frac{z a_{n_0}}{\epsilon}$
Dann gilt
$$\forall n \geq n_1 : \quad a_n \leq \frac{z a_{n_0}}{n} \leq \frac{z a_{n_0}}{n_1} < \epsilon$$

□

Definition 2.6. *Sei* (M, d) *ein metrischer Raum. Eine Folge* $x_0, x_1, \ldots$ *in* (M, d) *heißt Cauchyfolge, wenn zu jedem* $\epsilon > 0$ *aus* $\mathbb{R}$ *eine natürliche Zahl* n_0 *existiert, so dass* $\forall n, m \in \mathbb{N}$

$$n, m \geq n_0 \Rightarrow \quad d(x_n, x_m) < \epsilon \qquad \textit{(Cauchy-Kriterium)}$$

Satz 2.7. *Sei* M *eine Menge mit einer Metrik* d
$(x_k)_{k\in\mathbb{N}}$ *eine konvergente Folge in* M.
Dann gilt

1. *(x_k) ist beschränkt, d.h.*

$$\forall a \in M : \{d(x_k, a) : k \in \mathbb{N}\} \subset \mathbb{R} \quad \textit{beschränkt}$$

2. *(x_k) ist eine Cauchy-Folge*

Beweis. 1. Sei x der Grenzwert von x_k
zu $\epsilon = 1$ gibt es $n_0 \in \mathbb{N}$ so dass

$$\forall k \geq n_0 : d(x_k, x) < 1$$
$$\Rightarrow \quad d(x_k, a) \overset{\text{Dreieckungleichung}}{\leq} d(x_k, x) + d(x, a) < 1 + d(x, a)$$

Also gilt $\forall k \in \mathbb{N}$

$$0 \leq d(x_k, a) \leq \max\{1 + d(x, a), d(x_0, a), d(x_1, a), ..., d(x_{n_0-1}, a)\}$$

das heißt $\{d(x_k, a) : k \in \mathbb{N}\}$ ist beschränkt.

2. Wenn $x_k \to x \Rightarrow \forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N}$
mit $d(x_k, x) < \frac{\epsilon}{2} \quad \forall k \geq n_0$
$\overset{\text{Dreieckungleichung}}{\Rightarrow} \quad \forall k, l \geq n_0 \quad d(x_k, x_l) \leq d(x_k, x) + d(x, x_l) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

□

Definition 2.8. (M, d) *heißt vollständiger metrischer Raum* $\Leftrightarrow$ *jede Cauchy-Folge in M ist konvergent.*

Die Vollständigkeit ist eine wichtige Eigenschaft, wenn wir Gleichungen in abstrakten Räumen lösen wollen.

Lemma 2.9. *Sei $x_0, x_1, ...$ eine Folge in (M, d), welche gegen $x \in M$ und $y \in M$ konvergiert. Dann ist $x = y$.*

Beweis. Wir führen einen Widerspruchbeweis.
Wäre $d(x, y) > 0$, dann existiert für $\epsilon = d(x, y)$ wegen der Konvergenz der Folge ein $n_0 = n_0(\frac{\epsilon}{2})$ aus $\mathbb{N}$ mit $d(x_n, x) < \frac{\epsilon}{2}$ für $n \geq n_0(\frac{\epsilon}{2})$. Analog existiert ein $m_0 = m_0(\frac{\epsilon}{2}) \in \mathbb{N}$ mit $d(x_n, y) < \frac{\epsilon}{2}$ für $n \geq m_0$. Aufgrund der Dreieckungleichung $d(x, y) \leq d(x, x_n) + d(x_n, y)$ und der Symmetrie $d(x, x_n) = d(x_n, x)$ folgt für alle $n \geq \max\{n_0, m_0\}$

$$d(x, y) < \epsilon$$

Wir erhalten einen Widerspruch zu der Annahme $\epsilon = d(x, y)$. Es folgt

$$x = y.$$

□

Definition 2.10. *Eine Teilfolge einer gegebenen Folge* $(x_n)_{n\in\mathbb{N}}$ *ist eine Auswahl* $(x_{n_k})_{k\in\mathbb{N}}$ *die ihrerseits auch eine Folge ist und deren Glieder allesamt in dieser Reihenfolge auch Glieder der Folge* $(x_n)_{n\in\mathbb{N}}$ *sind.*

Beispiel 2.11. *Die Folge* $x_0, x_2, x_4, x_6, ...$ *ist eine Teilfolge der Folge* $(x_n)_{n\in\mathbb{N}}$

Die geometrische Folge

Lemma 2.12. $\forall y \in \mathbb{R},\ |q| < 1,$ *konvergiert die geometrische Folge*

$$x_n = C \cdot q^n$$

gegen Null

Beweis. Sei $\epsilon > 0$ gegeben. Nach Annahme gilt $|q| < 1$ und damit $|q|^{-1} > 1$.
Somit ist $|q|^{-1} = 1 + x$ für ein $x > 0$
Die Ungleichung $d(Cq^n, 0) < \epsilon$ ist dann äquivalent zu der Ungleichung

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

Das archimedische Axiom garantiert die Existenz einer natürlichen Zahl $n_0 > \frac{C}{x\epsilon} - \frac{1}{x} = \frac{C-\epsilon}{x\epsilon}$.
Für alle $n \geq n_0$ gilt dann

$$\frac{C}{\epsilon} = \left(\frac{C}{\epsilon x} - \frac{1}{x}\right) x + 1 < n_0 x + 1 < nx + 1$$

Daraus folgt wegen der Bernoulli Ungleichung:

$$1 + nx \leq (1+x)^n$$
$$\frac{C}{\epsilon} < (1+x)^n$$

Dieses ist wie oben erhalten wurde äquivalent zu $d(Cq^n, 0) < \epsilon$. □

Folgerung 2.13. *Dies hat die folgende Konsequenz:*
Die geometrische Reihe $S_n = 1 + q + q^2 + ... + q^n$ *konvergiert für* $|q| < 1$ *und hat in diesem Fall den Grenzwert*

$$\lim_{n\to\infty} \sum_{i=0}^{n} q^i = \frac{1}{1-q}$$

Beweis. Dies folgt aus der verallgemeinerten Binomialform

$$(1-q)\cdot \underbrace{(1+q+...+q^n)}_{\sum\limits_{k=0}^{n} q^k} = 1-q^{n+1}$$

(die man durch Indukition beweisen kann).
Diese Formel zeigt, dass

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

Also

$$d\left(\sum_{k=0}^{n} q^k, \frac{1}{1-q}\right) = C\cdot |q|^n$$

für $C = |\frac{q}{1-q}|$
$\Rightarrow \quad d\left(\sum\limits_{k=0}^{n} q^k, \frac{1}{1-q}\right) < \epsilon \quad$ für $\quad n \geq n_0$ □

Satz 2.14. *Banachscher Fixpunktsatz*
Sei (M,d) *ein vollständiger metrischer Raum und* $f : M \to M$ *eine strenge Kontraktion, d.h.*

$$\exists 0 < c < 1$$
$$d(f(x), f(y)) \leq cd(x,y) \quad \forall x,y \in M$$

Dann konvergiert die Folge (x_n) *definiert durch:*

$$x_0 = a \in M \ \textit{ beliebig}$$
$$x_{n+1} = f(x_n)$$

gegen den Fixpunkt x *von* f*:* $x = f(x)$

Beweis. Wir müssen zeigen, dass (x_n) eine Cauchy-Folge ist. Dann folgt, dass (x_n) gegen ein $x \in M$ konvergiert.
Es gilt dann:

$$d(x_n, x) \to 0 \Rightarrow d(f(x), f(x)) \to 0,$$
$$\text{da} \quad d(f(x_n), f(x)) \leq cd(x_n, x)$$

Also bekommen wir:

$$\begin{cases} x_n \to x \\ f(x_n) \to f(x) \qquad \Rightarrow \quad \underline{x = f(x)} \\ x_{n+1} = f(x_n) \end{cases}$$

(x_n) ist eine Cauchy-Folge

$$d(x_{n+2}, x_{n+1}) = d(f(x_{n+1}), f(x_n)) \leq cd(x_{n+1}, x_n)$$

Durch vollständige Induktion folgt:

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0) \quad \forall n \in \mathbb{N}$$

Für $m > n$ erhalten wir mit der Dreieckungleichung durch Induktion nach m:

$$d(x_m, x_n) \leq \sum_{j=n}^{m-1} d(x_{j+1}, x_j)$$

Es folgt:

$$d(x_m, x_n) \leq \sum_{j=n}^{m-1} c^j d(x_1, x_0) = c^n d(x_1, x_0) \underbrace{\sum_{k=0}^{m-n-1} c^k}_{\leq \frac{1}{1-c}, \text{ s. Folgerung 2.13}} \leq \frac{c^n d(x_1, x_0)}{1-c}$$

Nun gibt es zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $\frac{c^n d(x_1,x_0)}{1-c} < \epsilon \ \forall n \geq n_0$
Es folgt $\forall m > n \geq n_0 : d(x_m, x_n) < \epsilon$
das heißt (x_n) ist eine Cauchy-Folge.
x ist der einzige Fixpunkt von f, denn wäre $y = f(y)$ ein weiterer Fixpunkt von f,

dann

$$d(x, y) = d(f(x), f(y)) \leq c d(x, y) = 0 \Rightarrow \quad x = y$$

□

Beispiel 2.15. $f : [0,1] \to \mathbb{R}$ *mit* $|f(x) - f(y)| \leq c|x - y|, \ c < 1$

$$f(0) = 0$$
$$x_0 = x$$
$$x_{n+1} = f(x_n)$$

Hier können wir direkt abschätzen:

$$|x_n - 0| = |f(x_{n-1}) - f(0)| \leq c|x_{n-1} - 0|$$

Mit Induktion folgt

$$|x_n| \leq c^n |x| \to 0, \ n \to \infty$$

2.3 Vollständigkeit des $\mathbb{R}^n$, folgenkompakte Mengen

Definition 2.16. *Ein metrischer Raum (M, d) heißt folgenkompakt, wenn jede Folge aus (M, d) eine konvergente Teilfolge besitzt.*

Ein folgenkompakter metrischer Raum ist automatisch vollständig, denn eine Cauchy-Folge x_n konvergiert gegen x genau dann, wenn eine Teilfolge der Cauchy-Folge gegen x konvergiert.

Definition 2.17. *Sei A eine Teilmenge eines metrischen Raumes (M, d). Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige in (M, d) konvergente Folge von Elementen aus A. Sei $x := \lim_{n\to\infty} x_n \in M$. A heißt abgeschlossen, genau dann wenn gilt: $x \in A$.*

Beispiel 2.18. *Die Intervalle*

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

sind abgeschlossene Teilmengen in $\mathbb{R}$

Satz 2.19. *Jede abgeschlossene Teilmenge A eines vollständigen metrischen Raums (M, d_M) versehen mit der eingeschränkten Metrik ist ein vollständiger metrischer Raum (A, d_M).*

Beweis. Sei eine Cauchy-Folge x_n in (A, d_M) gegeben. Dann ist per Definition x_n eine Cauchyfolge in (M, d_M). Nach Annahme $\exists x = \lim_{n\to\infty} x_n$ in (M, d_M).
Weil A abgeschlossen ist, gilt $x \in A$: Also ist per Definition $x \in A$ der Grenzwert von x_n in (A, d_M). $\square$

Definition 2.20. *Sei (M, d) ein metrischer Raum.*
$O \subset M$ heißt offen: $\Leftrightarrow$ $\forall x \in O$: O ist Umgebung von x.
$A \subset M$ heißt abgeschlossen $\Leftrightarrow$ $C_M A$ offen.
x heißt Häufungspunkt einer Menge $A \Leftrightarrow$ $\forall$ Umgebung U von x:

$$(U \setminus \{x\}) \cap A \neq \emptyset.$$

Erinnerung: $C_M A$ bezeichent das Komplement von A in M.

Bemerkung: Der Begriffe "Häufungspunkt" und "Häufungswert" sollten nicht miteinander verwechselt werden.

Satz 2.21. *Jede folgenkompakte Teilmenge A eines metrischen Raumes (M, d_M) ist beschränkt und abgeschlossen in (M, d_M).*

Beweis. Wäre A nicht beschränkt, gäbe es eine Folge (x_n) aus A mit $d_M(x_0, x_n) \geq n$ Dies liefert einen Widerspruch, denn für jede Teilfolge $(\tilde{x}_n)$ einer solchen Folge gilt insbesondere $d_M(x_0, \tilde{x}_n) \geq n$. Somit besäße (x_n) keine konvergente (und damit beschränkte) Teilfolge. Wir erhalten $\lightning$ zur Folgenkompaktheit von A.

Um zu zeigen, dass A abgeschlossen ist, betrachten wir eine beliebige Folge (x_n) aus A mit Grenzwert x in (M, d_M). Dann konvergiert aber auch jede Teilfolge $(\tilde{x}_n)$ der Folge (x_n) gegen den Grenzwert x in (M, d_M). (A, d_M) ist nach Annahme folgenkompakt. Somit existiert eine konvergente Teilfolge$(\tilde{x}_n)$ der Folge (x_n) mit einem Grenzwert $a \in A$. Aus der Eindeutigkeit des Grenzwerts folgt daher $x = a$. Somit ist $x \in A$. Also ist A eine abgeschlossene Teilmenge von (M, d_M). $\square$

Bemerkung 2.22. *Für $A \subset \mathbb{R}^n$ gilt von dem eben bewiesenen Satz auch die Umkehrung. Das gesamte Resultat ist bekannt als Satz von Bolzano-Weierstrass, den wir im Folgenden zeigen.*

Satz 2.23. *Satz von Bolzano-Weierstrass*
Sei $A \subset \mathbb{R}^n$. Folgende Aussagen sind äquivalent:

1. *A ist beschränkt und abgeschlossen*

2. *Jede Folge $(x_k)_{k\in\mathbb{N}}$ aus A besitzt eine in A konvergente Teilfolge $(x_{k_n})_{n\in\mathbb{N}}$*

Beweis. 2. $\Rightarrow$ 1. folgt aus Satz 2.21
1. $\Rightarrow$ 2. durch Induktion
Wir zeigen dies zunächst für $n = 1$
Dazu beweisen wir zuerst:

Lemma 2.24. *Jede Folge (x_k) reeller Zahlen besitzt eine monotone Teilfolge.*

Beweis. Sei $B = \{k \in \mathbb{N} : \forall l \geq k : \ x_k \geq x_l\}$
Fall 1 B unendlich
Wir zählen $B \subset \mathbb{N}$ monoton wachsend ab:

$$\begin{cases} k_0 & = \min B \\ k_{n+1} & = \min \{k \in B : k > k_n\} \end{cases}$$

Dann ist die Teilfolge (x_{kn}) von (x_k) monoton fallend.
Fall 2 B endlich oder leer:

$$\Rightarrow \ \exists k_0 \in \mathbb{N} \ \forall k \geq k_0 : \ k \notin B \quad \text{, d.h.} \quad \exists l \geq k : \ x_k < x_l$$

Damit können wir definieren:

$$k_{n+1} = \min\{l \geq k_n : \; x_{kn} < x_l\}$$

und die Teilfolge $(x_{kn})_{n\in\mathbb{N}}$ von (x_n) ist monoton wachsend. $\square$

Sei nun (x_k) eine beliebige Folge in A, (x_{kn}) eine monotone Teilfolge.
(x_{kn}) ist beschränkt, da A beschränkt ist $\Rightarrow (x_{kn})$ ist konvergent.
Wir müssen zeigen, dass $x := \lim\limits_{n\to\infty} x_{kn} \in A$
Angenommen $x \notin A \Rightarrow \; x \in \mathcal{C}A$
$\mathcal{C}A$ ist offen $\Rightarrow \; \exists$ Umgebung U von x:

$$U \subset \mathcal{C}A \Rightarrow U \cap A = \emptyset$$

Nun ist aber mit geeignetem $n_0 \in \mathbb{N}$

$$\forall k \geq k_0 : \; x_{kn} \in U : \; x_{kn} \in A \Rightarrow \; x_{kn} \in U \cap A \;\lightning$$

Wir müssen nun noch mit Induktion die Aussage (1) $\Rightarrow$ (2) für $n > 1$ beweisen.
Angenommen, sie gilt schon für ein $n \in \mathbb{N}$. Sei dann $A \subset \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ beschränkt und abgeschlossen. Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge in A.
Wir schreiben

$$x_k = (x_k^*, \xi_k) \quad x_k^* \in \mathbb{R}^n, \; \xi_k \in \mathbb{R}$$

Die Folgen $(x_k^*)_{k\in\mathbb{N}}$, $(\xi_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ bzw. $\mathbb{R}$ sind beschränkt, denn wählen wir in $\mathbb{R}^n$ die Norm

$$||x||_\infty = \max|x_j|, \;\; j = 1, 2, ..., m$$

so gilt

$$\begin{cases} ||x_k^*||_\infty & \leq ||x_k||_\infty \\ |\xi_k| & \leq ||x_k||_\infty \end{cases}$$

und $\{||x_k||_\infty, \; k \in \mathbb{N}\}$ ist nach Voraussetzung beschränkt.
Nach Induktionsvoraussetzung können wir daher eine konvergente Teilfolge:

$$(x_{k_n}^*)_{n\in\mathbb{N}} \text{ mit } \lim_{n\to\infty} x_{k_n}^* = x^* \in \mathbb{R}^n$$

auswählen. Anschließend können wir aus (ξ_{kn}) eine konvergente Teilfolge $(\xi_{knm})_{m\in\mathbb{N}}$ auswählen mit $\lim_{m\to\infty} \xi_{knm} = \xi \in \mathbb{R}$.
Damit gilt unter Beachtung des Lemmas

$$(x_{knm})_{m\in\mathbb{N}} = (x_{k_{nm}}^*, \xi_{knm})$$

ist konvergent mit

$$x := \lim_{m\to\infty} x_{knm} = (x^*, \xi^*).$$

Wie zuvor folgt aus der Abgeschlossenheit von A : $x \in A$. $\square$

Bemerkung 2.25. *Im obigen Beweis haben wir eine konvergente Teilfolge von* $((x_k^*, \xi_k))$ *gefunden, indem wir zunächst eine Teilfolge wählten, die bzgl. der ersten Komponente (hier genannt* x_k^**) konvergiert, dann haben wir aus dieser Teilfolge eine Teilfolge gewählt, die auch bezüglich der zweiten Komponente (hier genannt* ξ_k*) konvergiert. Dieses Vorgehen lässt sich verallgemeinern zum sog. Diagonalfolgenprinzip: Sei* $(f_k)_{k\in\mathbb{N}}$ *eine Folge von Funktionen mit gleichem Definitionsbereich und gleichem, folgenkompakten, Wertebereich. Seien* $(x_j)_{j\in\mathbb{N}}$ *abzählbar viele Elemente aus dem gemeinsamen Definitionsbereich der Funktionen* (f_k)*. Gesucht ist eine Teilfolge* (g_k) *von* (f_k)*, sodass* $\lim_{k\to\infty} g_k(x_j)$ *für alle* $j \in \mathbb{N}$ *konvergiert. Wir verfahren wie folgt: Sei* $(f_{1,k})_{k\in\mathbb{N}}$ *eine Teilfolge von* $(f_k)_{k\in\mathbb{N}}$ *sodass* $(f_k(x_1))_{k\in\mathbb{N}}$ *konvergiert. Sie existiert, da der Wertebereich der* (f_k) *folgenkompakt ist. Sei* $(f_{2,k})_{k\in\mathbb{N}}$ *eine Teilfolge von* $(f_{1,k})_{k\in\mathbb{N}}$*, sodass* $(f_{2,k}(x_2))_{k\in\mathbb{N}}$ *konvergiert. Sie existiert, da der Wertebereich der* (f_k) *folgenkompakt ist. Für* $(f_{2,k})_{k\in\mathbb{N}}$ *konvergiert also* $(f_{2,k}(x_1))_{k\in\mathbb{N}}$*, da* $(f_{2,k})$ *Teilfolge von* $(f_{1,k})$ *ist und* $(f_{2,k}(x_2))_{k\in\mathbb{N}}$*, nach Auswahl von* $(f_{2,k})$*. Wir fahren induktiv fort: Sei nun* $(f_{n,k})_{k\in\mathbb{N}}$ *eine Teilfolge, für die* $(f_{n,k}(x_j))_{k\in\mathbb{N}}$ *für* $j = 1 \dots n$ *konvergiert. Es kann dann eine Teilfolge* $(f_{n+1,k})_{k\in\mathbb{N}}$ *von* $(f_{n,k})_{k\in\mathbb{N}}$ *ausgewählt werden, für die* $(f_{n+1,k}(x_j))_{k\in\mathbb{N}}$ *für* $j = 1 \dots n+1$ *konvergiert. Wir betrachten die Folge* (g_j) *mit* $g_j := f_{j,j}$*. Das Element* g_j *ist das* j*-te Element derjenigen Teilfolge,* $(f_{j,k})_{k\in\mathbb{N}}$ *für die* $f_{j,k}(x_1), \dots, f_{j,k}(x_j)$ *konvergiert. Die Folge* (g_k) *wird als Cantorsche Diagonalfolge bezeichnet. Es gilt* $lim_{k\to\infty} g_k(x_m)$ *konvergiert für alle* $m \in \mathbb{N}$*. Denn es gilt für jedes feste* $m \in \mathbb{N}$*: Die Folge* $g_m, g_{m+1}, \dots$ *ist Teilfolge von* $(f_{m,k})_{k\in\mathbb{N}}$ *und konvergiert daher, wenn man als Argument* x_m *einsetzt. Dieses Konstruktionsprinzip findet häufige Anwendung, z.B. beim Beweis der Abzählbarkeit von* $\mathbb{Q}$*.*

Lemma 2.26. *Sei* (M, d) *ein metrischer Raum und* (x_k) *eine Cauchyfolge in* M*. Sei* x *Häufungswert von* $(x_k) \Rightarrow (x_k)$ *konvergiert gegen* x*.*

Beweis. Sei $\epsilon > 0$ vorgegeben. Wir wählen

$$k_0 \in \mathbb{N} \text{ mit } d(x_k, x_m) < \frac{\epsilon}{2} \quad \forall m, k > k_0,$$

dies ist möglich, weil (x_n) Cauchyfolge ist.
Wie wählen außerdem ein

$$m_0 \in \mathbb{N},\ m_0 > k \text{ mit } \quad d(x, x_{m_0}) < \frac{\epsilon}{2}.$$

Dies ist möglich, da x Häufungswert der Folge ist und somit in jeder Umgebung von x unendlich viele Folgenelemente liegen.

$$\Rightarrow \quad \forall k > m_0 \quad : \quad d(x, x_k) \underbrace{\leq}_{\text{Dreieckungleichung}} d(x, x_{m_0}) + d(x_{m_0}, x_k) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Da ϵ beliebig ist, ist x der Grenzwert der Folge. □

2.4 Rechenregeln für Grenzwerte von Folgen

Satz 2.27. *Sei V ein normierter Vektorraum über $\mathbb{K}$ ($=\mathbb{R}$ oder $\mathbb{C}$)*
$(x_k),(y_k)$ *seien konvergente Folgen in V*
$(\alpha_k),\ \ (\beta_k)\in\mathbb{K}\quad \beta_k\neq 0\ \ \forall\ \ k\quad \lim\limits_{k\to\infty}\beta_k\neq 0.$ *Dann gilt*

1. $\lim\limits_{k\to\infty} x_k\pm y_k=\lim\limits_{k\to\infty} x_k\pm\lim\limits_{k\to\infty} y_k$
2. $\lim\limits_{k\to\infty}\alpha_k x_k=(\lim\limits_{k\to\infty}\alpha_k)(\lim\limits_{k\to\infty} x_k)$
3. $\lim\limits_{k\to\infty}\frac{\alpha_k}{\beta_k}=\frac{\lim\limits_{k\to\infty}\alpha_k}{\lim\limits_{k\to\infty}\beta_k}$

Der Satz besagt, dass die algebraischen Operationen in einem Körper bzw. Vektorraum mit der Operation der Grenzwertbildung vertauschbar sind.

Beispiel 2.28.

$$\lim_{k\to\infty}\frac{3k^2+1}{2k^2-k+1}=\lim_{k\to\infty}\frac{3+\frac{1}{k^2}}{2-\frac{1}{k}+\frac{1}{k^2}}=\frac{\lim\limits_{k\to\infty}\left(3+\frac{1}{k^2}\right)}{\lim\limits_{k\to\infty}\left(2-\frac{1}{k}+\frac{1}{k^2}\right)}=\frac{3}{2}$$

Satz 2.29. *Seiem $(a_n),(b_n)$ konvergente Folgen in $\mathbb{R}$; (x_n) sei eine konvergente Folge in einem Vektorraum. Dann gilt:*

1. $a_k\leq b_k\ \ \forall k\in\mathbb{N}\ \ \Rightarrow\ \ \lim a_k\leq\lim b_k$
2. $||x_k||\leq b_k\ \ \forall k\in\mathbb{N}\ \ \Rightarrow\ \ ||\lim x_n||\leq\lim b_k$

Beweis. 1. Sei $\epsilon>0$ vorgegeben. $\exists k_0:\forall k>k_0$:

$$\begin{aligned} b_k\leq\lim_{l\to\infty} b_l+\frac{\epsilon}{2}\ \text{ und }\ \lim_{l\to\infty} a_l+\frac{\epsilon}{2}\leq a_k\\ \Rightarrow\ \lim_{l\to\infty} a_l\leq a_k+\frac{\epsilon}{2}\leq b_k+\frac{\epsilon}{2}\leq\lim_{l\to\infty} b_l+\epsilon\end{aligned}$$

Also

$$\begin{aligned}\forall\epsilon>0:\ \lim_{l\to\infty} a_l\leq\lim_{l\to\infty} b_l+\epsilon\\ \Rightarrow\ \lim_{l\to\infty} a_l\leq\lim_{l\to\infty} b_l\end{aligned}$$

2. Wir wählen in 1. $a_k=||x_k||$
Dann müssen wir noch zeigen:

$$\lim_{k\to\infty}||x_k||=||\lim_{l\to\infty} x_l||$$

Wir haben:

$$||x_k|| - ||\lim_{l\to\infty} x_l|| \le ||x_k - \lim_{l\to\infty} x_l||$$

(s. Satz 1.48)

Nun gibt es aber $\forall \epsilon > 0$ ein k_0, so dass für alle $k \ge k_0$ die rechte (und damit auch die linke) Seite $< \epsilon$ □

Lemma 2.30. *Sei (x_k) eine Folge im $\mathbb{R}^n$. Dann gilt*

$$\lim_{k\to\infty} x_k = y \Leftrightarrow \text{ für } j = 1, ..., n \quad \lim_{k\to\infty} x_{j,k} = y_j$$

wobei

$$x_k = (x_{1,k}, ..., x_{n,k})$$
$$y = (y_1, ..., y_n)$$

Beweis. Dies beweist man sofort mit Hilfe der Definition, wenn man die Norm

$$||z||_\infty = \max\{|z_j| \;\; j = 1, 2, ..., n\} \in \mathbb{R}^n$$

verwendet. □

Wir können eine Metrik auf $\hat{\mathbb{R}} = \mathbb{R} \cup \{-\infty; \infty\}$ definieren

$$M = \mathbb{R} \cup \{-\infty; \infty\}$$
$$d(x, y) = |\xi(x) - \xi(y)|, \quad \text{wobei}$$
$$\xi(x) = \begin{cases} \frac{x}{1+|x|} & \text{für} x \in \mathbb{R} \\ \pm 1 & \text{für} x = \pm\infty \end{cases}$$

Wir können zeigen, dass

$$\lim_{k\to\infty} d(a_k, a) = 0 \quad \Rightarrow \quad \lim_{k\to\infty} |a_k - a| = 0$$

Der Konvergenzbegriff ist also unabhängig davon, welche Metrik wir verwenden.

Satz 2.31. $\hat{\mathbb{R}}$ *ist folgenkompakt*

Beweis. Wir müssen zeigen, dass jede Folge in $\hat{\mathbb{R}}$ eine konvergente Teilfolge hat. Sei also (x_k) eine beliebige Folge in $\hat{\mathbb{R}}$. Dann:

1. (x_k) hat eine beschränkte Teilfolge in $\mathbb{R}$.
   Aus dieser können wir nach dem Satz von Bolzano-Weierstrass eine konvergente Teilfolge auswählen
   oder

2. (x_k) nimmt den Wert ∞ bzw. $-\infty$ unendlich oft an. Dann können wir die entsprechende konstante (und daher konvergente) Teilfolge auswählen.
   oder

3. Sonst können wir aus (x_k) eine unbeschränkte Teilfolge in $\mathbb{R}$ (y_k) auswählen. Ist (y_k) nach oben unbeschränkt folgt

$$\Rightarrow \quad \forall l, m \in \mathbb{N}: \ \exists k_m \geq l: \ y_{k_m} \geq m$$

   Also können wir $k_0 < k_1 < ...$ finden mit $y_{k_m} \geq m \ \forall m$. Dann gilt

$$\lim_{m \to \infty} y_{k_m} = +\infty$$

   denn

$$\begin{aligned} d(y_{k_m}, \infty) &= |\xi(y_{k_m}) - \xi(\infty)| = |\frac{y_{k_m}}{1 + |y_{k_m}|} - 1| \\ &= \frac{1}{1 + y_{k_m}} < \frac{1}{m+1} \longrightarrow 0 \ (m \to \infty) \end{aligned}$$

□

Definition 2.32. *algebraische Operatoren in* $\hat{\mathbb{R}}$
Wir definieren folgende arithmetische Operationen

$$\begin{aligned} x + \infty &:= \infty + x := \infty \ \ \forall x \in \mathbb{R} \cup \{\infty\} \\ x - \infty &:= -\infty + x := -\infty \ \ \forall x \in \mathbb{R} \cup \{\infty\} \\ x \cdot \infty &:= \infty \cdot x := \begin{cases} \infty & \forall x \in \hat{\mathbb{R}}, \ \ x > 0 \\ -\infty & \forall x \in \hat{\mathbb{R}}, \ \ x < 0 \end{cases} \\ \frac{1}{\infty} &:= \frac{1}{-\infty} := 0 \end{aligned}$$

Sinnlos ist

$$\begin{aligned} \infty - \infty \\ 0 \cdot \infty \\ 0 \cdot (-\infty) \\ \frac{\infty}{\infty} \ \ \mathit{usw.} \end{aligned}$$

Damit können wir auch die Rechenregel (Satz 2.27) auf $\hat{\mathbb{R}}$ definieren.

Definition 2.33. *Sei* (x_k) *eine Folge* $\in \hat{\mathbb{R}}$, $\quad \emptyset \neq H \subset \hat{\mathbb{R}}$
Die Menge der Häufungspunkte von (x_k) *in* $\hat{\mathbb{R}}$
Dann sei

$$\underline{\lim}\ x_k := \liminf x_k := \inf H \quad \textit{Limes Inferior}$$
$$\overline{\lim}\ x_k := \limsup x_k := \sup H \quad \textit{Limes Superior}$$

Bemerkung 2.34. $x = \underline{\lim} x_k \Leftrightarrow \quad \forall \epsilon > 0$

$$\begin{cases} \textit{(1)} \quad \{k : |x - x_n| < \epsilon\} \ \textit{ist unendlich (x ist Häufungswert)} \\ \textit{und} \\ \textit{(2)} \quad \{k : x_k < x - \epsilon\} \ \textit{ist endlich (x ist der kleinste Häufungswert)} \end{cases}$$

Bemerkung 2.35. *Dass* $x = \underline{\lim} x_k$ *selbst auch Häufungswert von* x_k *ist, sehen wir folgendermaßen ein:*

Nach Definition des Infimums gibt es für alle $\epsilon > 0$ *einen Häufungswert* y *von* (x_k) *mit* $\ x \leq y < x + \epsilon$. *Sei* $\ \delta = x + \epsilon - y > 0$. *Dann ist*

$$\begin{aligned} \{k \in \mathbb{N} : \ |x_k - y| < \delta\} \ &\textit{unendlich, da } y \textit{ Häufungswert ist und} \\ \{k \in \mathbb{N} : \ |x_k - y| < \delta\} &\subset \{k \in \mathbb{N} : \ |x_k - x| < \epsilon\}, \textit{ da} \\ |x_k - x| \leq |x_k - y| + |y - x| < \delta + |y - x| &= \epsilon - |y - x| + |y - x| \end{aligned}$$

Es folgt:

$$\{k \in \mathbb{N} : \ |x_k - x| < \epsilon\} \ \textit{unendlich.}$$

Entsprechendes gilt für $\overline{\lim}$.

Beispiel 2.36.

$$\begin{aligned} x_k &= k + (-1)^k k \\ x_{2k+1} &= 0 \ \forall k, \Rightarrow 0 \ \textit{ist Häufungswert,} \\ x_{2k} = 2k \ \forall k \quad \to \quad \infty, &\Rightarrow \infty \ \textit{ist Häufungswert.} \end{aligned}$$

(x_k) *hat keine weiteren Häufungswerte.*
Also gilt

$$0 = \underline{\lim}\ x_k$$
$$\infty = \overline{\lim}\ x_k$$

Satz 2.37. *Sei* (x_k) *eine Folge* $\in \hat{\mathbb{R}}$.
Dann gilt

$$\begin{aligned} & x_k \to x \;\; in \;\; \hat{\mathbb{R}} \\ \Leftrightarrow \;\; & \underline{\lim}\; x_k = x = \overline{\lim}\; x_k \end{aligned}$$

Beweis. "$\Rightarrow$"
(x_k) konvergiert gegen $x \Rightarrow$ (x_k) hat genau einen Häufungswert x
$\Rightarrow \underline{\lim} x_k = x = \overline{\lim} x_k$
"$\Leftarrow$"

$$\underline{\lim} x_k = x = \overline{\lim} x_k \Rightarrow \;\; \forall \epsilon > 0 \begin{cases} \{k \in \mathbb{N} : |x - x_k| < \epsilon\} & \text{unendlich} \\ \{k \in \mathbb{N} : |x - x_k| \geq \epsilon\} & \text{endlich} \end{cases}$$

$$\Rightarrow \;\; \exists k_0 \in \mathbb{N}\text{:} \;\; \forall k \geq k_0 : \;\; |x - x_k| < \epsilon$$

□

2.5 Konvergente Reihen

Der Begriff der unendlichen Reihe wird auf den der Folge zurückgeführt.

Definition 2.38. *Sei* (x_k) *eine Folge in* V *(normierter Vektorraum)*
Durch die Festsetzung

$$S_n = \sum_{k=0}^{n} x_k \quad \textit{sog. Partialsummen}$$

wird eine Folge (S_n) *definiert, die man als die zu* (x_k) *gehörende unendliche Reihe bezeichnet. Ist* S_n *konvergent mit dem Grenzwert* S*, so heißt* S *die Summe der unendlichen Reihe*

$$S := \sum_{k=0}^{\infty} x_k$$

Beispiel 2.39. *1. Geometrische Reihe:*

$$S_n = \sum_{j=0}^{n} z^j$$

Wir haben bereits gesehen:

$$S_n = \frac{1 - z^{n+1}}{1 - z}$$

Für $z \in \mathbb{C}, |z| < 1$ konvergiert z^{n+1} gegen 0, d.h. die Folge (S_n) konvergiert gegen $\frac{1}{1-z}$

$$\sum_{j=0}^{\infty} z^j = \frac{1}{1-z}$$

Wir sagen: Die Reihe $\sum\limits_{j=0}^{\infty} z^j$ ist für $|z| < 1$ konvergent.

2. *Seien $S_n = \sum\limits_{j=1}^{n} \frac{1}{j}$ die Partialsummen der harmonischen Reihe $\sum\limits_{j=1}^{\infty} \frac{1}{j}$*
   Behauptung: $\lim\limits_{k \to \infty} S_n = \infty$
   Dafür schreiben wir $\sum\limits_{j=1}^{\infty} \frac{1}{j} = \infty$

   Beweis.

$$\begin{aligned} S_{2^{n+1}} &= \sum_{j=1}^{2^{n+1}} \frac{1}{j} = 1 + \frac{1}{2} + \sum_{k=1}^{n} \sum_{j=2^k+1}^{2^{k+1}} \frac{1}{j} \\ &\geq 1 + \frac{1}{2} + \sum_{k=1}^{n} \underbrace{\sum_{j=2^k+1}^{2^{k+1}} \frac{1}{2^{k+1}}}_{2^{k+1}-2^k=2^k \text{ Summanden}} \\ &= 1 + \frac{1}{2} + \frac{1}{2} n \to \infty \text{ für } n \to \infty \end{aligned}$$

$\square$

3. *Die Folge*

$$S_n = \sum_{j=0}^{n} (-1)^j = \frac{1}{2}\left(1 + (-1)^n\right)$$

   konvergiert nicht. 1 0 1 0 1 0 1 ...
   Wir sagen $\sum\limits_{j=0}^{\infty} (-1)^n$ ist divergent.

Aus den Rechenregeln für konvergente Folgen erhalten wir folgende Ergebnisse.

Satz 2.40. *Seien $\sum\limits_{j=0}^{\infty} a_j$, $\sum\limits_{j=0}^{\infty} b_j$ konvergente Reihen, $\alpha \in \mathbb{R}$*

Dann sind auch die Reihen $\sum\limits_{j=0}^{\infty}(a_j+b_j)$, $\sum\limits_{j=0}^{\infty}\alpha a_j$ *konvergent und es gilt*

$$\sum_{j=0}^{\infty}(a_j+b_j)=\sum_{j=0}^{\infty}a_j+\sum_{j=0}^{\infty}b_j$$
$$\sum_{j=0}^{\infty}\alpha a_j=\alpha\sum_{j=0}^{\infty}a_j$$

<u>Cauchy Kriterium für Partialsummen</u>
Sei V vollständig.

$$\sum_{j=0}^{\infty}a_j \text{ ist konvergent} \quad \Leftrightarrow \quad \forall \epsilon>0$$
$$\exists m_0 \;\; \forall l\geq m\geq m_0:$$
$$||\sum_{j=m}^{l}a_j||\leq\epsilon$$

(Cauchy Kriterium für Partialsummen) <u>Folgerung</u> $\sum\limits_{j=0}^{\infty}a_j$ ist konvergent
$\Rightarrow$ (a_j) ist eine Nullfolge (d.h. $a_j \underset{j\to\infty}{\to} 0$)

Beweis. Man setze in das Cauchykriterium für Folgen

$$l=m+1$$

"$\Rightarrow$"gilt bereits ohne die Annahme der Vollständigkeit □

Bemerkung 2.41. *Die Umkehrung dieser Aussage gilt nicht*

Satz 2.42. *Sei V ein normierter Vektorraum*

$$(a_j) \;\; \textit{Nullfolge in } V \Rightarrow \sum_{j=0}^{\infty}(a_j-a_{j+1})=a_0$$

Beweis. Betrachte die Partialsummen:

$$S_k=\sum_{j=0}^{k}(a_j-a_{j+1})=\sum_{j=0}^{k}a_j-\sum_{j=1}^{k+1}a_j$$
$$=a_0-a_{k+1}\Rightarrow \;\; ||S_k-a_0||=||a_{k+1}||\to 0 \quad \text{für } k\to\infty$$

□

Beispiel 2.43.

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

Satz 2.44. *Sei V ein vollständiger normierter Vektorraum.*
Sei $\sum_{j=0}^{\infty} ||a_j||$ konvergent in $\mathbb{R}$
$\Rightarrow \sum_{j=0}^{\infty} a_j$ konvergiert in V.

Beweis. Wir benutzen das Cauchykriterium.
Für die Reihe $\sum_{j=0}^{\infty} a_j$ lautet die Bedingung

$$\left|\left|\sum_{j=m}^{l} a_j\right|\right| \leq \epsilon$$

$$\left|\left|\sum_{j=m}^{l} a_j\right|\right| \leq \sum_{j=m}^{l} ||a_j||$$

und $\sum_{j=m}^{l} ||a_j|| \leq \epsilon$,weil $\sum_{j=0}^{\infty} ||a_j||$ konvergiert. □

Diese Tatsache motiviert die folgende Definition.

Definition 2.45. $\sum_{j=0}^{\infty} a_j$ *heißt* <u>*absolut konvergent*</u> *in V*
$\Leftrightarrow$ $\sum_{j=0}^{\infty} ||a_j||$ *konvergent in $\mathbb{R}$*

Wir beschäftigen uns daher im Folgenden mit notwendigen und hinreichenden Kriterien für die Konergenz reeller Reihen und erhalten daraus entsprechende Kriterien für die absolute Konvergenz von Reihen in V.

Satz 2.46. *Sei (a_j) eine Folge nichtnegativer reeller Zahlen. Dann gilt:*
$\sum_{j=0}^{\infty} a_j$ ist konvergent in $\mathbb{R} \Leftrightarrow$ Die Folge der Partialsummen ist beschränkt.
Der Beweis folgt aus der Monotonie der Partialsummenfolge.

Satz 2.47 ((Vergleichstest)). *Sei V vollständiger normierter Vektorraum, (x_j) Folge in V, (a_j) Folge in $\mathbb{R}$, $||x_j|| \leq a_j \ \forall j$.*

$$\sum_{j=0}^{\infty} a_j \ \textit{konvergent} \Rightarrow \begin{cases} \sum\limits_{j=0}^{\infty} x_j \ \textit{absolut konvergent} \\ \textit{und} \\ \left|\left|\sum\limits_{j=0}^{\infty} x_j\right|\right| \leq \sum\limits_{j=0}^{\infty} ||x_j|| \leq \sum\limits_{j=0}^{\infty} a_j \end{cases}$$

Der Beweis folgt aus dem Cauchy-Kriterium

Bemerkung 2.48. *Satz 2.47 wird auch als "Vergleichstest" bezeichnet. Der Satz erlaubt eine Konvergenzaussage wenn es möglich ist, eine vorliegende Reihe durch Abschätzung mit einer bekannten konvergenten Reihe zu vergleichen.*

Bemerkung 2.49. *Gilt $0 \leq a_j \leq ||x_j|| \ \forall j$ und ist $\sum\limits_{j=0}^{\infty} a_j$ divergent, so kann $\sum\limits_{j=0}^{\infty} x_j$ nicht absolut konvergieren.*

Beispiel 2.50.

1. $V = \mathbb{C}$, $a_j = \frac{(-1)^{j+1}}{j} z^j \ j \geq 1$, $(|z| < 1)$

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} z^j \ \textit{ist absolut konvergent,}$$

$$\textit{da} \ \ \left|\frac{(-1)^{j+1}}{j} z^j\right| \leq |z|^j$$

$$\textit{und} \ \sum_{j=1}^{\infty} |z|^j \ \textit{konvergiert (geometrische Reihe)}$$

2. *Die Reihe* $\sum\limits_{j=1}^{\infty} \frac{1}{j!}$ *konvergiert, da*

$$\frac{1}{j!} \leq \frac{1}{j(j+1)} \ \textit{für } j \geq 2 \ \textit{ und } \ \sum_{j=1}^{\infty} \frac{1}{j(j+1)} \ \textit{konvergiert.}$$

Satz 2.51. *Quotientenkriterium*
Sei (a_k) eine Folge positiver reeller Zahlen.

1. $\overline{\lim} \frac{a_{k+1}}{a_k} < 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ *ist konvergent*

2. $\exists k_0 \in \mathbb{N} : \frac{a_{k+1}}{a_k} \geq 1 \;\; \forall k \geq k_0 \Rightarrow \sum_{k=0}^{\infty} a_k = \infty$

Beweis. 1. Nach Voraussetzung
$\exists q < 1, \;\; k_0 \in \mathbb{N} \;\; \text{mit} \;\; \frac{a_{k+1}}{a_k} < q \;\; \forall k \geq k_0$
Induktion $\Rightarrow \;\; a_{k+k_0} \leq a_{k_0} \leq a_{k_0} q^k$
Vergleichstest: $\Rightarrow \;\; \sum_{k=0}^{\infty} a_{k+k_0}$ konvergiert
$\Rightarrow \;\; \sum_{k=0}^{\infty} a_k$ konvergiert (geometrische Reihe ist so genannte Majorante)

2. Ist dagegen $\frac{a_{k+1}}{a_k} \geq 1 \;\; \forall k \geq k_0$

$$\Rightarrow \;\; a_k \geq a_{k_0} > 0 \;\; \forall k \geq k_0$$
$$\Rightarrow \;\; (a_k) \;\; \text{ist keine Nullfolge}$$
$$\Rightarrow \;\; \sum_{k=0}^{\infty} a_k \;\; \text{divergiert}$$

□

Beispiel 2.52.

$$\sum_{j=0}^{\infty} \frac{z^j}{j!}$$

$$\frac{\left|\frac{z^{j+1}}{(j+1)!}\right|}{\left|\frac{z^j}{j!}\right|} = \frac{|z|}{j+1} \to 0$$

<u>Wiederholung</u>

- Eine Reihe ist eine Folge von Partialsummen
- $S_n = \sum_{k=0}^{n} a_k \;\; \Rightarrow \;\; (a_k)$ und $(S_n)_{n \in \mathbb{N}}$ sind Folgen
- Falls $\sum_{k=0}^{\infty} a_k$ konvergiert $\Rightarrow \;\; (a_k)_{k \in \mathbb{N}}$ ist eine Nullfolge
- Die Kriterien für Folgen können übertragen werden
- $\sum_{k=0}^{\infty} a_k$ heißt absolut konvergent, genau dann wenn $\sum_{k=0}^{\infty} ||a_k||$ konvergiert

Quotientenkriterium

$$\begin{aligned} &\overline{\lim}\left|\frac{a_{k+1}}{a_k}\right| < 1 \\ \Leftrightarrow \exists \theta \text{ mit } 0<\theta<1,\ \exists k_0 \in \mathbb{N}, \text{ so dass } &\left|\frac{a_{k+1}}{a_k}\right| \leq \theta\ \forall k \geq k_0 \\ &\Rightarrow \sum_{k=0}^{\infty} a_k \text{ konvergiert} \end{aligned}$$

Definition 2.53.

$$\begin{aligned} \exp:\ \mathbb{C} \to \mathbb{C} \quad \exp(z) &= \sum_{k=0}^{\infty} \frac{z^k}{k!} =: e^z \\ \sin:\ \mathbb{C} \to \mathbb{C} \quad \sin(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \\ \cos:\ \mathbb{C} \to \mathbb{C} \quad \cos(z) &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} \end{aligned}$$

$\exp, \sin, \cos$ *konvergieren für alle* $z \in \mathbb{C}$

Beweis.
- Zeige, dass $a_k = \frac{z^k}{k!}$ das Quotientenkriterium erfüllt

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{z^{k+1}}{(k+1)!}\frac{k!}{z^k}\right| = \left|\frac{z}{k+1}\right|$$
$$\text{Sei } \ k \geq 2|z| \Rightarrow \left|\frac{z}{k+1}\right| \leq \frac{1}{2}$$

- Es gilt

$$\begin{aligned} \sin(z) &= \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) \\ \cos(z) &= \frac{1}{2}\left(e^{iz} - e^{-iz}\right) \end{aligned}$$

Daher folgt die Konvergenz von $\sin(z)$ und $\cos(z)$ aus der Konvergenz von $\exp(z)$

□

Bemerkung 2.54. $\forall x \in \mathbb{R}$ *gilt*

$$\begin{aligned} \sin(x) &= \Im(exp(ix)) \\ \cos(x) &= \Re(\exp(ix)) \\ \Rightarrow \quad e^{ix} = \cos(x) + i\sin(x) \quad &\textit{Euler-Darstellung} \end{aligned}$$

Satz 2.55. *Sei (a_n) eine Folge reeller Zahlen*

1. $\overline{\lim} \sqrt[n]{|a_n|} < 1 \Rightarrow \sum_{n=0}^{\infty} a_n$ *konvergiert*

2. $\sqrt[n]{|a_n|} \geq 1$ *für unendliche $n \Rightarrow$* $\sum_{n=0}^{\infty} a_n = \infty$

Beweis. 1.

$$\begin{aligned} & \overline{\lim} \sqrt[n]{a_n} < 1 \\ \Rightarrow \quad & \exists 0 < q < 1 \ , \ n_0 \in \mathbb{N} \text{ so, dass } \sqrt[n]{a_n} < q \ \forall n \geq n_0 \\ & \Rightarrow a_n < q^n \ \forall n \geq n_0 \end{aligned}$$

$$\text{d.h. } \sum_{k=0}^{\infty} a_{k+n_0} \text{ konvergiert (geometrische Reihe ist Majorante)}$$

$$= \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n_0} a_k \text{ endlich}$$

$$\Rightarrow \sum_{k=0}^{\infty} a_k \text{ konvergiert}$$

2.

$$\begin{aligned} \sqrt[n]{a_n} \geq 1 \Rightarrow \quad & a_n \geq 1 \text{ für unendlich viele } n \in \mathbb{N} \\ & \Rightarrow (a_n)_{n \in \mathbb{N}} \text{ ist keine Nullfolge} \end{aligned}$$

□

Definition 2.56. *Sei (a_n) eine Folge, dann heißt $\sum_{n=0}^{\infty} a_n z^n$ Potenzreihe.*

Beispiel 2.57. *Wir betrachten die Potenzreihe*

$$\sum_{n=1}^{\infty} n^{\alpha} z^n \ , \ z \in \mathbb{C} \ , \ \alpha \in \mathbb{Q}$$

1. $|z| < 1$: *Die Potenzreihe konvergiert absolut, da*

$$\sqrt[n]{|n^{\alpha} z^n|} = \left(\sqrt[n]{n}\right)^{\alpha} |z| \underset{n \to \infty}{\longrightarrow} |z| < 1$$

2. $|z| \geq 1$

   (a) $\alpha \geq 0 \Rightarrow \ |n^{\alpha}| \, |z|^n \geq n^{\alpha} \geq 1$
   $\Rightarrow n^{\alpha} z^n$ *ist keine Nullfolge.*

(b)

$$\left\{ \begin{array}{ll} \alpha & < 0 \\ |z| & > 1 \end{array} \right\} \Rightarrow |n^\alpha z^n| = \left(\sqrt[n]{n^\alpha}|z|\right)^n$$

Es gilt $\lim\limits_{n\to\infty} \sqrt[n]{n^\alpha} = 1$, *d.h.* $\exists n_0 \in \mathbb{N}$ *mit*

$$\sqrt[n]{n^\alpha} \geq \frac{1}{|z|} \quad \forall n \geq n_0 \Rightarrow \sqrt[n]{n^\alpha}|z| \geq 1 \quad \forall n \geq n_0 \Rightarrow \text{ Reihe divergiert}$$

(c)

$$\left\{ \begin{array}{ll} -1 \leq \alpha & < 0 \\ |z| & = 1 \end{array} \right\} \Rightarrow |n^\alpha z^n| = |n^\alpha| \geq \frac{1}{n}$$

$$\Rightarrow \text{ Die Reihe ist nicht absolut konvergent, da } \sum \frac{1}{n} = \infty$$

(d)

$$\left\{ \begin{array}{ll} \alpha & < -1 \\ |z| & = 1 \end{array} \right\} \Rightarrow$$

$$\sum_{n=1}^{2^{k+1}-1} |n^\alpha z^n| = \sum_{l=0}^{k} \sum_{n=2^l}^{2^{l+1}-1} n^\alpha \leq \sum_{l=0}^{k} \underbrace{\sum_{n=2^l}^{2^{l+1}-1} (2^l)^\alpha}_{2^l \text{ Summanden}}$$

$$= \sum_{l=0}^{k} \left(2^{1+\alpha}\right)^l \underset{\text{geometrische Reihe}}{=} \frac{1-\left(2^{1+\alpha}\right)^{k+1}}{1-2^{1+\alpha}} \leq \frac{1}{1-2^{1+\alpha}}$$

$$\left(2^{1+\alpha}\right)^{k+1} < 1, \quad da\ \alpha < -1$$

Die Abschätzung ist unabhängig von k

$\Rightarrow$ *Die Partialsummenfolge* $\sum\limits_{n=1}^{\infty} n^\alpha$ *ist beschränkt*

$\Rightarrow$ *Die Reihe* $\sum n^\alpha z^n$ *konvergiert absolut*

Bemerkung 2.58. *zu 2.(c) : Unter den Voraussetzungen von 2 (c) gilt:* $\sum\limits_{n=\alpha}^{\infty} (-1)^n n^\alpha$ *konvergiert. Dies folgt aus dem unten stehenden Satz 2.60.*

Definition 2.59. *Sei* $(a_n)_{n\in\mathbb{N}}$ *eine reelle Folge* $\sum\limits_{k=0}^{\infty} a_k$ *heißt alternierend* $\Leftrightarrow \forall k$ *gilt* $a_{k+1}a_k < 0$

Satz 2.60. *Sei* $\sum\limits_{k=0}^{\infty} a_k$ *eine alternierende Reihe und* $(|a_k|)$ *eine monotone Nullfolge. Dann ist* $\sum a_k$ *in* $\mathbb{R}$ *konvergent*

Beweis. o.B.d.A. $a_0 > 0 \Rightarrow \ a_{2k} > 0 \ a_{2k+1} < 0$
Wir setzen $S_k = \sum\limits_{n=0}^{k} a_n, \ \ S_{2k+2} - S_{2k} = a_{2k+2} + a_{2k+1} < 0$
$\Rightarrow S_0 \geq S_2 \geq S_4 \geq ... \quad (S_{2k})$ ist monoton fallend.
Analog: $\ S_1 \leq S_3 \leq S_5 \leq ... \quad (S_{2k+1})$ ist monoton wachsend.
Wegen $S_{2k+1} - S_{2k} = a_{2k+1} < 0$ gilt: $S_{2k+1} \leq S_{2k} \ \forall k$
$\Rightarrow \quad S_1 \leq S_{2k}$ und $S_{2k+1} \leq S_0$
$\Rightarrow (S_{2k})$ ist monoton fallend und beschränkt
(S_{2k-1}) ist monoton wachsend und beschränkt
Sei $\lim\limits_{k\to\infty} S_{2k} = S \quad S' = \lim\limits_{k\to\infty} S_{2k+1}$
Es gilt $S - S' = \lim\limits_{k\to\infty} (S_{2k} - S_{2k+1}) = \lim\limits_{k\to\infty} a_{2k+1} = 0$
Sei $\epsilon > 0$, dann $\exists N_1, N_2 \in \mathbb{N}$ so dass gilt:

$$|S_{2k} - S| < \frac{\epsilon}{2} \ \forall k \geq N_1 \text{ und } |S_{2k+1} - S| < \frac{\epsilon}{2} \ \forall k \geq k \geq N_1$$
$$\text{setze } N := \max(2N_1; 2N_2 + 1)$$
$$\Rightarrow |S_n - S| < \epsilon \ \forall n \geq N$$

□

Satz 2.61. Umordnungssatz
Sei $\sum\limits_{k=0}^{\infty} a_k$ eine absolut konvergierende Reihe in einem vollständigen normierten Vektorraum. Dann gilt für jede bijektive Abbildung $\tau : \mathbb{N} \to \mathbb{N}$:

$$\sum_{k=0}^{\infty} a_{\tau(k)} = \sum_{k=0}^{\infty} a_k$$

Bemerkung 2.62. *Der Satz ist falsch, wenn die Reihe nicht absolut konvergiert.*

Beispiel 2.63. $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ *konvergiert.*
Behauptung: $\exists$ *Umordnung* τ*, so dass* $\sum\limits_{n=1}^{\infty} \frac{(-1)^{\tau(n)-1}}{\tau(n)}$ *divergiert.*
Beachte:

$$\underbrace{\frac{1}{2^k+1} + \frac{1}{2^k+3} + ... + \frac{1}{2^{k+1}-1}}_{2^{k-1} \text{ Summanden}} \geq 2^{k-1} \cdot \frac{1}{2^{k+1}} = \frac{1}{4}$$

$\Rightarrow$ *Die Umordnung*

$$
\begin{aligned}
&1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}\\
&\quad+\left(\frac{1}{5}+\frac{1}{7}\right)-\frac{1}{6}\\
&\quad+\left(\frac{1}{9}+\frac{1}{11}+\frac{1}{13}+\frac{1}{15}\right)-\frac{1}{8} \qquad >\frac{1}{4}-\frac{1}{8}=\frac{1}{8}\\
&\quad\vdots\\
&\quad+\left(\frac{1}{2^k+1}+\frac{1}{2^k+3}+\ldots+\frac{1}{2^{k+1}-1}\right)-\frac{1}{2k+2} \qquad >\frac{1}{4}-\frac{1}{8}=\frac{1}{8}
\end{aligned}
$$

konvergiert nicht.

Satz 2.64. *Cauchyprodukt für Reihen*
Seien $\sum\limits_{k=0}^{\infty} a_k$, $\sum\limits_{k=0}^{\infty} b_k$ *absokut konvergente Reihen (in* $\mathbb{R}$ *oder* $\mathbb{C}$*).*
Sei $c_m = \sum\limits_{k=0}^{m} a_k b_{m-k}$
Dann konvergiert $\sum\limits_{m=0}^{\infty} c_m = \left(\sum\limits_{k=0}^{\infty} a_k\right)\left(\sum\limits_{k=0}^{\infty} b_k\right)$

(ohne Beweis)

Beispiel 2.65. *Funktionalgleichung für* exp
$\forall x, y \in \mathbb{C}$ *gilt:*

$$\exp(x+y) = \exp(x) + \exp(y)$$

Beweis.

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

$$c_m = \sum_{k=0}^{m} \frac{x^k}{k!} \cdot \frac{y^{m-k}}{(m-k)!} = \frac{1}{m!} \sum_{k=0}^{m} \binom{m}{k} x^k y^{m-k} = \frac{1}{m!}(x+y)^m$$

$\square$

Kapitel 3

Stetige Abbildungen

3.1 Grenzwerte von Funktionen, Stetigkeit

Definition 3.1. *Seien (X, d_x) und (Y, d_y) metrische Räume. Eine Abbildung*

$$f: \ (X, d_x) \to (Y, d_y)$$

heißt stetig, wenn für jede Folge (x_n) in X gilt

$$x_n \to \xi \Rightarrow f(x_n) \to f(\xi)$$

Anders formuliert:
f führt konvergente Folgen in konvergente Folgen über und vertauscht mit Limesbildung.

$$f\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} f(x_n)$$

Ein Spezialfall
Abbildungen $f: \ (X, d_x) \to (Y, d_y)$ mit der Eigenschaft

$$d_y(f(x), f(\xi)) \leq C d_x(x, \xi)$$

mit $C > 0$ (konstant) sind stetig. Für genügend großes n gilt $d_x(x_n, \xi) < \frac{\epsilon}{C}$ wegen $x_n \to \xi$ und damit $d_y(f(x_n), f(\xi)) < \epsilon$.

Solche Abbildungen nennt man Lipschitz-stetig und C nennt man Lipschitz-Konstante Beispiele sind kontraktive Abbildungen, für diese gilt $C < 1$, bzw. $C \leq 1$.

Beispiel 3.2. 1. *Für Polynome*

$$\begin{aligned} p : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sum_{j=0}^{m} a_j x^j \end{aligned}$$

gilt für jede in $\mathbb{R}$ *konvergente Folge* (x_n)

$$\lim p(x_n) = p(\lim x_n)$$

wie wir sofort aus den Rechenregeln für konvergente Folgen schließen (Kapitel 2).

2. *Wir betrachten die Funktion*

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto [x] \quad \textit{Gaußklammer (s.S. 32)} \end{aligned}$$

Sind $(x_k), (y_k)$ *Folgen mit*

$$\lim_{k\to\infty} x_k = 1 = \lim_{k\to k} y_k$$

und $\forall k \in \mathbb{N} : (0 \leq) x_k < 1 \leq y_k (< 2)$
so gilt $\forall k$*:*

$$f(x_k) = 0 \ \ , \textit{also} \ \lim_{k\to\infty} f(x_k) = 0 \neq f\left(\lim_{k\to\infty} x_k\right) = 1$$

$$f(y_k) = 1 \ \ , \textit{also} \ \lim_{k\to\infty} f(y_k) = 1 = f\left(\lim_{k\to\infty} y_k\right)$$

Die Gaußklammer ist also nicht stetig.

3.

$$g : \mathbb{R} \to \mathbb{R}$$

$$g(x) = \begin{cases} \frac{1}{x} & \textit{für } x \neq 0 \\ 0 & \textit{für } x = 0 \end{cases}$$

Nach den Rechenregeln für konvergente Folgen gilt für Folgen (x_k) *mit*

$$\lim_{k\to\infty} x_k \neq 0 :$$

$$\lim_{k\to\infty} g(x_k) = g\left(\lim_{k\to\infty} x_k\right)$$

Ist aber

$$\lim_{k\to\infty} x_k = 0, \ \ x_k > 0 \ \ \forall k \ \textit{bzw.} \ x_k < 0 \ \ \forall k$$

so gilt

$$\lim_{k\to\infty} g(x_k) = +\infty \neq 0 = g\left(\lim_{k\to\infty} x_k\right)$$

bzw.

$$\lim_{k\to\infty} g(x_k) = -\infty \neq 0 = g\left(\lim_{k\to\infty} x_k\right)$$

Definition 3.3. *Seien* (X, d_x) , (Y, d_y) *metrische Räume.*
Eine Abbildung $f : X \to Y$ *heißt stetig im Punkt* ξ *von* X*, wenn für jede Folge* (x_n)*, die in* (X, d_x) *gegen* ξ *konvergiert, die Bildfolge* $(f(x_n))$ *in* (Y, d_y) *gegen* $f(\xi)$ *konvergiert.*

Offensichtlich ist $f : (X, d_x) \to (Y, d_y)$ genau dann stetig, wenn f in jedem Punkt ξ von X stetig ist.

Satz 3.4. *Sei* $f : (X, d_x) \to (Y, d_y)$ *stetig im Punkt* $\xi \in X$ *und sei* $g : (Y, d_y) \to (Z, d_z)$ *stetig im Punkt* $\eta = f(\xi) \in Y$.
Dann ist die Komposition

$$(g \circ f) : (X, d_x) \to (Z, d_z)$$

stetig im Punkt $\xi \in X$

Beweis. Nach Voraussetzung gilt für jede gegen ξ konvergierende Folge (x_n)

$$y_n = f(x_n) \to \eta = f(\xi)$$

Da g stetig ist, gilt wegen der Konvergenz von y_n gegen η in (Y, d_y)

$$g(y_n) = (g \circ f)(x_n) \to g(\eta) = (g \circ f)(\xi)$$

Also folgt

$$(g \circ f)(x_n) \to (g \circ f)(\xi)$$

und $g \circ f$ ist stetig im Punkt ξ. □

Folgerung 3.5. *Die Komposition stetiger Abbildungen ist wieder stetig.*

Eigenschaften stetiger Funktionen

Lemma 3.6. *Sei* $f : (X, d_x) \to (Y, d_y)$ *eine stetige Abbildung. Ist* A *in* (Y, d_y) *abgeschlossen, dann ist das Urbild* $f^{-1}(A)$ *abgeschlossen in* (X, d_x).

Beweis. Sei $x_n \to x$ eine in (X, d_x) konvergente Folge mit $x_n \in f^{-1}(A)$
zu zeigen ist $x \in f^{-1}(A)$.
Da f stetig ist, konvergiert die Bildfolge $y_n = f(x_n) \in A$ gegen $y_n \to y = f(x)$. Weil A abgeschlossen ist, folgt $y \in A$ und damit $x = f^{-1}(y) \in f^{-1}(A)$. □

Lemma 3.7. *Sei* $f : (X, d_x) \to (Y, d_y)$ *eine stetige Abbildung. Ist* (X, d_x) *folgenkompakt, dann ist auch das Bild* $f(X)$ *versehen mit der Einschränkung der Metrik* d_y *auf* $f(Y)$ *folgenkompakt.*

Beweis. Sei (y_n) eine Folge in $f(x)$.
Dann gilt $y_n = f(x_n)$ und die x_n definieren eine Folge in X.
Nach Annahme gibt es eine in (X, d_x) konvergente Teilfolge $\tilde{x}_n \to \tilde{x}$. Aus der Stetigkeit folgt, die Teilfolge $\tilde{y}_n = f(\tilde{x}_n)$ konvergiert gegen $f(\tilde{x})$ in $(f(X), d_y)$ □

Satz 3.8. *Auf einem folgenkompakten metrischen Raum (X, d_x) hat jede stetige reellwertige Funktion $f : (X, d_x) \to \mathbb{R}$ ein beschränktes Bild und das Maximum, sowie das Minimum von f werden auf X als Funktionswerte angenommen.*

Beweis. Das Bild von $f(X)$ ist folgenkompakt bezüglich der euklidischen Metrik (siehe Lemma 3.7).
Also ist $f(X)$ beschränkt und abgeschlossen.
Insbesondere gilt daher

$$\sup\left(f(X)\right) = \max\left(f(X)\right)$$
$$\inf\left(f(X)\right) = \min\left(f(X)\right)$$

Daraus folgt die Behauptung. □

Satz 3.9. *Zwischenwertsatz*
Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) \leq f(b)$. Dann gibt es für jedes $\eta \in [f(a), f(b)]$ ein $\alpha \in [a, b]$ mit

$$f(\alpha) = \eta$$

Beweis. Betrachte die (nichtleere beschränkte) Menge

$$A := \{x \in [a, b] : f(x) \leq \eta\}$$

entweder ist dann $\alpha := \sup A$ gleich b, oder es gibt per Definition ein $x \in [a, b]$ mit $x > \alpha$, i.e. $x \notin A$

$$\Rightarrow \quad f(x) > \eta$$

In beiden Fällen folgt

$$f(\alpha) \geq \eta$$

(in letzterem Fall wegen der Stetigkeit von f, da eine monoton fallende, gegen α konvergente Folge von Punkten aus A existiert, welche gegen das Supremum α von A konvergiert. Falls $\alpha = b$ gilt offensichtlich $\eta = f(\alpha) = f(b)$ und somit $f(\alpha) \geq \eta$)
Aus Stetigkeitsgründen und der Definition von A folgt

$$f(\alpha) \leq \eta.$$

Beides zusammen genommen ergibt

$$f(\alpha) = \eta.$$

□

3.2 Stetige Funktionen auf kompakten Mengen

Satz 3.10. *Das $\epsilon - \delta$ Kriterium*
Gegeben sind eine Funktion $f : (X, d_x) \to (Y, d_y)$ und ein $\xi \in X$. Dann ist f genau dann stetig in ξ, wenn zu jedem $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ existiert, so dass gilt:

$$d_x(x, \xi) < \delta \Rightarrow \ d_y\left(f(x), f(y)\right) < \epsilon$$

Veranschaulichung des $\epsilon - \delta$ Kriteriums
Betrachte die Kugel $B_\delta(x) = \{x \in X : d_x(x, \xi) < \delta\}$
Für beliebiges $\epsilon > 0$ soll es ein $\delta > 0$ geben, so dass zu jedem $x \in X$, dessen Abstand zu ξ kleiner als δ ist, der Abstand von $f(x)$ zu $f(\xi)$ kleiner als ϵ ist.
f soll also die Kugel $B_\delta(\xi)$ um ξ vom Radius $< \delta$ in die Kugel $B_\epsilon\left(f(\xi)\right)$ um $f(\xi)$ vom Radius $< \epsilon$ abbilden.

Beweis. Zunächst wollen wir zeigen, dass das $\epsilon-\delta$ Kriterium die Stetigkeit impliziert. Sei also x_n eine Folge in X mit $\lim\limits_{n\to\infty} x_n = \xi$. Also zu jedem $\tilde{\epsilon} > 0$, existiert ein $N = N(\tilde{\epsilon})$ mit $d_x(x_n, \xi) \leq \tilde{\epsilon}$ für $n \geq N$.
Sei $\epsilon > 0$ vorgegeben. Nach Annahme gilt $\exists \delta(\epsilon) > 0 :$

$$d_y(f(x_n), f(\xi)) < \epsilon \text{ für } \ d_x(x_n, \xi) < \delta = \delta(\epsilon).$$

Wählt man jetzt $\tilde{\epsilon}$ gleich δ, folgt

$$d_y(f(x_n), f(\xi)) < \epsilon$$

für $n \geq N(\delta)$. Also konvergiert $f(x_n)$ gegen $f(x)$.
Zum Beweis der Gegenrichtung nehmen wir an f sei stetig im Punkt ξ. Angenommen das $\epsilon - \delta$ Kriterium wäre falsch. Dann gilt:

$$\begin{aligned} &\exists \epsilon_0 > 0 \ \forall \delta > 0 \ \exists x \in X \\ (d_x(x, \xi) < \delta \ \text{ und } \ &d_y(f(x), f(\xi)) \geq \epsilon_0) \end{aligned}$$

Wähle nun $\delta = \frac{1}{n}$. Dann existiert ein x_n mit $d_x(x_n, \xi) < \frac{1}{n}$ und $d_y(f(x_n), f(\xi)) \geq \epsilon_0$.
Wegen $d_x(x_n, \xi) < \frac{1}{n}$ gilt $x_n \to \xi$.
Aus der Stetigkeit von f im Punkt ξ folgt $f(x_n) \to f(\xi)$ im Widerspruch zu

$$d_y(f(x), f(\xi)) \geq \epsilon_0$$

□

Folgerung
Eine Funktion $f : (X, d_x) \to (Y, d_y)$ ist genau dann stetig, wenn gilt

$$\begin{aligned} &\forall \xi \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X : \\ &d_x(x, \xi) < \delta \Rightarrow \ d_y(f(x), f(\xi)) < \epsilon \end{aligned}$$

Bemerkung 3.11. *Approximation ist ein gängiges Konzept in der Analysis. Stetigkeit kann interpretiert werden als "lokale Approximierbarkeit durch Konstanten". Sei* $f: X \to Y$ *im Punkt* x *stetig, d.h.* $\forall \epsilon > 0 \quad \exists \delta > 0: \quad d_Y(f(x), f(y)) < \epsilon$ *für alle* y *mit* $d_X(x,y) < \delta$*. Anschaulich bedeutet das, dass die Funktion* f *nahe der Stelle* x *durch die Konstante* $f(x)$ *approximiert werden kann, und dass der Fehler dieser Approximation kleiner als* ϵ *ist, wenn der Abstand von* x *kleiner als* δ *ist. Wir werden später weitere Approximationskonzepte für Funktionen kennen lernen: lokale Approximierbarkeit durch lineare Funktionen (Differenzierbarkeit), Approximierbarkeit durch Potenzreihen (Taylorreihen, analytische Funktionen) sowie Approximation von bestimmten Funktionenklassen durch Funktionen anderer Klassen z.B. durch trigonometrische Funktionen (Fourierreihenentwicklung) oder durch Polynome (Weierstrass-Approximationssatz und Taylorpolynome).*

Gleichmäßige Stetigkeit

Definition 3.12. *Eine Abbildung* $f: (X, d_x) \to (Y, d_y)$ *heißt gleichmäßig stetig auf* (X, d_x) *wenn zu jedem* $\epsilon > 0$ *ein* $\delta = \delta(\epsilon) > 0$ *existiert, so dass* $\forall \xi,\ x \in X$ *gilt*

$$d_x(x, \xi) < \delta \Rightarrow \ d_y(f(x), f(\xi)) < \epsilon$$

Bemerkung 3.13.

$$\Leftrightarrow \quad \forall \epsilon > 0 \exists \delta > 0 \ \ \forall x \in A : f(K_\delta(x) \cap A) \subset K_\epsilon(f(x))$$

Der Unterschied zur Definition der Stetigkeit auf A *ist der, dass dort die "maximale Größe* δ*"der Umgebung* $K_\delta(x)$*, für die noch* $f(K_\delta(x) \cap A) \subset K_\epsilon(f(x))$ *gilt, sowohl von* ϵ *als auch von* x *abhängt. Gleichmäßige Stetigkeit bedeutet dagegen, dass* $\forall x \in A$ *bei gegebenem* $\epsilon > 0$ *dasselbe* δ *gewählt werden kann.*

Beispiel 3.14.

$$\begin{aligned} M_1 = \mathbb{R} \setminus \{0\} \quad M_2 = \mathbb{R} \\ f: \mathbb{R} \setminus \{0\} \to \mathbb{R} \\ x \mapsto \frac{1}{x} \end{aligned}$$

1. f *ist gleichmäßig stetig auf* $A = \mathbb{R} \setminus \{]-a, a[\},\ \ a > 0$
2. f *ist nicht gleichmäßig stetig auf* $\mathbb{R} \setminus \{0\}$

Beweis.

$$\begin{aligned} |f(x) f(y)| &= \left| \frac{1}{x} - \frac{1}{y} \right| = \frac{1}{|xy|} |x - y| \\ \text{also} \quad |f(x) - f(y)| < \epsilon &\Leftrightarrow \ \ |x - y| < |xy| \epsilon \end{aligned}$$

1. Für $x, y \in A = \mathbb{R} \setminus \{-a, a\}$ gilt

$$|xy| \geq a^2$$

also

$$|x-y| < \epsilon a^2 =: \delta \Rightarrow \ |x-y| < \epsilon |xy|$$

Daher gilt

$$\forall \epsilon > 0 \ \forall x, y \in A : |x-y| < \delta := \underbrace{\epsilon a^2}_{\text{unabhängig von } x \text{ und } y} \Rightarrow \ |f(x) - f(y)| < \epsilon$$

2. Dagegen können wir $\forall \delta > 0,\ x, y \in \mathbb{R} \setminus \{0\}$ finden mit $|x-y| < \delta$ aber

$$|f(x) - f(y)| \geq 1 \Leftrightarrow \ |x-y| \geq |xy|$$

Sei $\delta > 0$ beliebig. Wähle $n \in \mathbb{N}$, sodass $\frac{\delta}{n} < 1$. Nun gilt für $|x-y| = \frac{\delta}{2n}$:

$$\begin{aligned} |xy| &\leq (|x-y| + |x|)|x| \\ &= \left(\frac{\delta}{2n} + |x|\right)|x| \leq \frac{\delta^2}{2n^2} \ \text{(für } |x| < \frac{\delta}{2n}) \\ &= \frac{\delta}{n}|x-y| \\ &\leq |x-y| \ \text{ da } \frac{\delta}{n} \leq 1. \end{aligned}$$

$\square$

Satz 3.15. <u>*Satz von Heine*</u>
Ist (X, d_x) folgenkompakt, dann gilt:
Jede stetige Funktion auf (X, d_x) ist gleichmäßig stetig

Beweis. Wäre die Aussage falsch, dann würde gelten

$$\begin{gathered} \exists \epsilon_0\colon \ \forall \delta > 0 \ \exists \xi \in X \ \exists x \in X \\ d_x(x, \xi) < \delta \ \text{ und } \ d_y(f(x), f(\xi)) \geq \epsilon_0 \end{gathered}$$

Fixiere ein solches $\epsilon_0 > 0$. $\forall n \geq 1, n \in \mathbb{N}$ und $\delta := n^{-1}$ $\exists$ $x_n, y_n \in X$ mit $d_X(\xi_n, x_n) < n^{-1}$ und $d_y(f(\xi_n), f(x_n)) \geq \epsilon_0$.

Bei sorgfältiger Auswahl von Teilfolgen kann man o.B.d.A. durch Übergang zu Teilfolgen erreichen (X ist folgenkompakt, f ist stetig)

$$\xi_n \underset{n\to\infty}{\to} \xi$$
$$x_n \underset{n\to\infty}{\to} x$$

wegen $d_x(\xi_n, x_n) < n^{-1}$ (dies gilt auch für die Teilfolgen) und

$$\begin{aligned} 0 \le d_x(x,\xi) &\le \underbrace{d_x(x,x_n)}_{<\frac{1}{3}\epsilon} + \underbrace{d_x(x_n,\xi_n)}_{<n^{-1}<\frac{1}{3}\epsilon} + \underbrace{d_x(\xi_n,\xi)}_{\frac{1}{3}\epsilon} \\ &< \epsilon \end{aligned}$$

für $n \ge N(\epsilon)$ und alle $\epsilon > 0$ folgt $d_x(x,\xi) = 0$ im Limes $n \to \infty$. Also gilt $x = \xi$ und damit folgt wegen $f(x) = f(\xi)$

$$0 < \epsilon_0 \le d_y(f(\xi_n), f(x_n)) \le \underbrace{d_y(f(\xi_n), f(\xi))}_{<\frac{1}{2}\epsilon_0} + \underbrace{f(x_n), f(x)}_{<\frac{1}{2}\epsilon_0}$$

wegen der Stetigkeit von f ist die rechte Seite $< \epsilon_0$,
falls $n \ge N_1\left(\frac{\epsilon_0}{2}\right)$ respektive $n \ge N_2\left(\frac{\epsilon_0}{2}\right)$
Wegen der Konvergenz der Folgen $f(\xi_n) \to f(\xi)$ und $f(x_n) \to f(x)$.
Dies liefert

$$0 < \epsilon_0 < \epsilon_0 \quad ↯$$

□

3.3 Reellwertige stetige Funktionen

Definition 3.16. *Für einen metrischen Raum (X, d_x) ist*

$$C(X) := \{f : X \to \mathbb{R} : f \textit{ ist stetig auf } (X, d_x)\}$$

der Raum der stetigen reellwertigen Funktionen auf X.

Lemma 3.17. *Seien $f, g \in C(X)$ und $\lambda \in \mathbb{R}$.*
Dann ist auch $f + g$, λg, fg wieder stetig auf X.

Bemerkung 3.18. *$C(X)$ bildet dann einen Ring*

Beweis. (nur für fg, die anderen Rechnungen sind trivial):
Seien also ein $\xi \in X$ und ein $\epsilon > 0$ gegeben.
Dann ist

$$\begin{aligned}|f(x)g(x) - f(\xi)g(\xi)| &\leq |f(x)g(x) - f(x)g(\xi) + f(x)g(\xi) - f(\xi)g(\xi)| \\ &\leq \underbrace{|f(x)|}_{\leq c_1}\underbrace{|g(x) - g(\xi)|}_{<\frac{\epsilon}{c_1}} + \underbrace{|g(\xi)|}_{\leq c_2}\underbrace{|f(x) - f(\xi)|}_{<\frac{\epsilon}{c_2}} < \epsilon\end{aligned}$$

für Konstanten $c_1, c_2 > 0$.
Ersteres gilt nur für $d_x(x,\xi) < \delta_1$, letzteres nur für $d_x(x,\xi) < \delta_2$.
Außerdem gilt

$$\begin{aligned}&|f(x) - f(\xi)| < 1 \text{ falls } \; d_x(x,\xi) < \delta_3 \\ &\text{Wir wählen } \; \delta = \min\{\delta_1, \delta_2, \delta_3\} \\ &\text{Falls } \; d_x(x,\xi) < \delta \text{ ist damit} \\ &|f(x)| \leq 1 + |f(\xi)| =: c_1\end{aligned}$$

Alles zusammen zeigt, dass

$$(fg)(x) = f(x)g(x)$$

stetig im Punkt ξ ist. □

Folgerung 3.19. *Polynome sind stetige Funktionen auf* $\mathbb{R}$

Lemma 3.20. *Sei* $f : (X, d_x) \to \mathbb{R}$ *stetig und* $f(x) \neq 0 \;\; \forall x \in X$.
Dann ist

$$\frac{1}{f(x)} : (X, d_x) \to \mathbb{R}$$

definiert und stetig auf (X, d_x)

Beweis. Gegeben sei ein ξ und ein $\epsilon > 0$. Dann gilt

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

Es gilt für $d(x,\xi) < \delta_1$

$$|f(x) - f(\xi)| < |\frac{1}{3}f(\xi)| =: \epsilon_1 > 0$$

aufgrund der Stetigkeit von f.
Daraus folgt

$$
\begin{aligned}
|f(x)| > \left|\frac{2}{3}f(\xi)\right| \quad & \text{wegen der unteren Dreieckungleichng} \\
& \big(|f(x) - f(\xi)| \geq ||f(\xi)| - |f(x)||\big) \\
& \Rightarrow \quad \left|\frac{1}{f(x)}\right| < \left|\frac{3}{2f(\xi)}\right| \\
& \text{Also } \left|\frac{1}{f(x)} - \frac{1}{f(\xi)}\right| = \left|\frac{1}{f(x)}\frac{f(\xi)-f(x)}{f(\xi)}\right| = \left|\frac{3}{2}\frac{f(\xi)-f(x)}{f(\xi)^2}\right| < \epsilon \\
& \text{falls } |f(x) - f(\xi)| < \frac{2}{3}|f(\xi)|^2\epsilon
\end{aligned}
$$

Dies ist erfüllt, falls $d_x(x, \xi)$ klein genug ist. □

Definition 3.21. *Seien* $f, g : (X, d_x) \to \mathbb{R}$.

$$\max(f, g)(x) = \max(f(x), g(x))$$

und analog für das Minimum.

Bemerkung 3.22. $C(X)$ *ist ein reeller Vektorraum von Funktionen.*

Satz 3.23. $\min(f, g)$ *und* $\max(f, g)$ *sind in* $C(X)$ *für* $f, g \in C(X)$

Beweis. Es geüugt, dass mit f auch $|f|$ (als Komposition stetiger Abbildungen, $|\cdot|$ ist stetig) stetig ist,
denn $\max(f, g) = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|$
und $\min(f, g) = -\max(-f, -g)$ □

3.4 Folgen stetiger Funktionen

Konvergenz von Funktionen

Definition 3.24. *Eine Folge* $(f_n)_{n\in\mathbb{N}}$ *von reellwertigen Funktionen* $f_n : (X, d_x) \to \mathbb{R}$ *heißt punktweise konvergent gegen eine Grenzfunktion* $f : (X, d_x) \to \mathbb{R}$, *wenn für jedes* $x \in X$ *gilt* $\lim\limits_{n\to\infty} f_n(x) = f(x)$.

Bemerkung 3.25. *Konvergenz von Funktionenfolgen gegen eine Grenzfunktion (bzw. Approximation einer gegebenen Funktion durch eine Folge von Funktionen) ist wichtig für die Anwendung der Analysis in der mathematischen Modellbildung und numerischen Simulation.*

Beispiel 3.26. *Die Partialsummen der Exponentialreihe stellen die Approximation der Exponentialfunktion dar, d.h.* $\forall x \in \mathbb{R}$ *gilt:*

$$\sum_{k=0}^{n} \frac{x^k}{k!} \underset{n\to\infty}{\rightarrow} \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

In diesem Fall sind die approximierenden Funktionen (als Polynome) stetig, genau so wie die Grenzfunktion e^x.

Beispiel 3.27. *Betrachte* $f_n(x) = 1 - x^n \ , \ x \in [0;1] \subset \mathbb{R}$

$$f_n(x) \underset{n\to\infty}{\rightarrow} f(x) := \begin{cases} 1 & \text{für } 0 \le x < 1 \\ 0 & \text{für } x = 1 \end{cases}$$

Wir sehen, dass die Grenzfunktion nicht stetig ist.

<u>Frage</u>: Welche Bedingungen sind zu stellen, damit für eine konvergente Folge stetiger Funktionen auch die Grenzfunktion wieder stetig ist?

Definition 3.28. <u>*Gleichmäßige Konvergenz*</u>
Eine Folge $(f_n)_{n\in\mathbb{N}}$ *von Funktionen* $f_n : (X, d_x) \to \mathbb{R}$ *heißt gleichmäßig konvergent gegen eine Grenzfunktion* $f : (X, d_x) \to \mathbb{R}$*, wenn zu jedem* $\epsilon > 0$ *ein* $N = N(\epsilon) \in \mathbb{N}$ *existiert, so dass* $\forall n \ge N$ *und alle* $x \in X$ *gilt*

$$|f_n(x) - f(x)| < \epsilon$$

Satz 3.29. <u>*Satz der gleichmäßigen Konvergenz*</u>
Konvergiert eine Folge stetiger Funktionen $f_n : X \to \mathbb{R}, \ n \in \mathbb{N}$ *gleichmäßig gegen eine Funktion* $f : X \to \mathbb{R}$*, so ist auch die Grenzfunktion* f *stetig.*

Beweis. Seien $\xi_0 \in X, \ \epsilon > 0$ gegeben. Es ist zu zeigen, dass ein $\delta_\epsilon > 0$ existiert, so dass für alle $x \in X$:

$$|x - \xi_0| < \delta_\epsilon \Rightarrow \ |f(x) - f(\xi_0)| < \epsilon$$

Wegen der gleichmäßigen Konvergenz der Folge $(f_n)_{n\in\mathbb{N}}$ gibt es ein $n \in \mathbb{N}$, so dass $\forall x \in X$ gilt

$$|f_n(x) - f(x)| < \frac{1}{3}\epsilon$$

Da f_n stetig ist, existiert weiter ein $\delta_\epsilon > 0$, so dass $\forall x \in X$ mit $|x - \xi_0| < \delta_\epsilon$ gilt

$$|f_n(x) - f_n(\xi_0)| < \frac{1}{3}\epsilon$$

Dann folgt für alle solche $x \in X$

$$|f(x) - f(\xi_0)| \leq \underbrace{|f(x) - f_n(x)|}_{<\frac{1}{3}\epsilon} + \underbrace{|f_n(x) - f_n(\xi_0)|}_{<\frac{1}{3}\epsilon} + \underbrace{|f_n(\xi_0) - f(\xi_0)|}_{<\frac{1}{3}\epsilon} < \epsilon$$

d.h. f ist stetig in ξ_0. □

Bemerkung 3.30. *Diese Aussage überträgt sich auch auf Reihen von stetigen Funktionen* $\sum_{k=1}^{\infty} f_n$*. Wenn die Partialsummen* $\sum_{k=1}^{n} f_k$ *gleichmäßig konvergieren, so ist ihr Limes* $\sum_{k=1}^{\infty} f_k$ *wieder eine stetige Funktion.*

Für beschränkte reellwertige Funktionen f definiert man

$$||f||_\infty := \sup_{\xi \in X} |f(\xi)|$$

Es gilt

- $||f||_\infty = 0 \Leftrightarrow \quad f = 0$
- $||\lambda f||_\infty = |\lambda| ||f||_\infty$
- $||f+g||_\infty \leq ||f||_\infty + ||g||_\infty$, denn $|f(\xi)+g(\xi)| \leq |f(\xi)| + |g(\xi)| \leq ||f||_\infty + ||g||_\infty$ gilt $\forall \xi \in X$.

Somit definiert $||\cdot||_\infty$ eine Norm auf dem $\mathbb{R}$-Vektorraum der beschränkten Funktionen auf X, die so genannte Supremumsnorm.

Sei (X, d_x) ein folgenkompakter metrischer Raum. Dann ist jede stetige Funktion beschränkt auf X. Damit ist $C(X)$ ein Unterraum des Raumes der beschränkten Funktionen auf X. Für stetige Funktionen $f \in C(X)$ gilt sogar

$$||f||_\infty = \max_{x \in X} |f(x)|$$

Die Supremumsnorm definiert durch

$$d(f, g) = ||f - g||_\infty$$

eine Metrik auf $C(X)$.

Lemma 3.31. *Für eine Funktionenfolge* $(f_n)_{n \in \mathbb{N}}$ *auf einem beschränkten, abgeschlossenen Intervall* $X = [a, b]$ *ist die gleichmäßige Konvergenz gegen eine Grenzfunktion* $f : [a, b] \to \mathbb{R}$ *gleichbedeutend mit*

$$||f_n - f||_\infty \underset{n \to \infty}{\longrightarrow} 0$$

Satz 3.32. *Satz der Vollständigkeit*
Sei (X, d_x) *folgenkompakt. Dann ist* $(C(X), d)$ *versehen mit der Metrik* d *der gleichmäßigen Konvergenz ein vollständiger metrischer Raum.*

Beweis. Gegeben sei eine Cauchyfolge in $C(X)$, also $f_n \in C(X)$ $n = 1, 2, ...$ mit $||f_n - f_m||_\infty < \epsilon$ für $n, m \geq N(\epsilon)$.
Wir konstruieren eine Grenzfunktion. Fixiere einen Punkt $x \in X$ und betrachte die Folge $y_n = f_n(x) \in \mathbb{R}$. Dann gilt $|y_n - y_m| \leq ||f_n - f_m||_\infty < \epsilon$, falls $n, m \geq N(\epsilon)$.
Also definiert $y_1, y_2, ...$ eine reelle Cauchyfolge. Daher existiert der Limes $y_n \to y$.
Wir setzen $f(x) := y = \lim_{n\to\infty} y_n \in \mathbb{R}$
Nun behaupten wir

$$|f(x) - f_n(x)| < \epsilon \text{ für } n > N\left(\frac{\epsilon}{2}\right) \qquad (\star)$$

und zwar für alle $x \in X$.

Sei nämlich $x \in X$ ein beliebiger Punkt. Dann gibt es ein $m_0 = m_0(\frac{1}{2}\epsilon, x)$ mit $|f(x) - f_m(x)| < \frac{\epsilon}{2}$ für $m \geq m_0(\frac{\epsilon}{2}, x)$, weil nach Konstruktion $f_m(x) \to f(x)$ gilt. Dies liefert

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < \frac{1}{2}\epsilon + \underbrace{||f_m - f_n||}_{<\frac{\epsilon}{2}} \leq \epsilon$$

falls $n, m \geq N\left(\frac{\epsilon}{2}\right)$ und $m \geq m_0(\frac{1}{2}\epsilon, x)$ gilt. Ein solches m kann man immer finden (d.h. für jede Stelle x ist ein passendes $m = m(x)$ zu wählen), so dass man jetzt unabhängig von der Wahl von x wie behauptet $|f(x) - f_n(x)| < \epsilon$ für $n \geq N\left(\frac{\epsilon}{2}\right)$ gezeigt hat.

Zum anderen behaupten wir $f \in C(X)$. Betrachte $x_1, x_2 \in X$. Dann gilt für geeignetes n und $d_x(x_1, x_2) < \delta$ (und δ geeignet)

$$|f(x_1) - f(x_2)| \leq \underbrace{|f(x_1) - f_n(x_1)|}_{\frac{1}{3}\epsilon} + \underbrace{|f_n(x_1) - f_n(x_2)|}_{\frac{1}{3}\epsilon} + \underbrace{|f_n(x_2) - f(x_2)|}_{\frac{1}{3}\epsilon}$$

Ist $d_x(x_1, x_2) < \delta = \delta(\frac{\epsilon}{3}, f_n)$, dann gilt $|f_n(x_1) - f_n(x_2)| < \frac{1}{3}\epsilon$, da f_n stetig ist und damit auch gleichmäßig stetig auf (X, d_x).
Schließlich haben wir für ein beliebiges x (auch $x = x_1, x_2$) gezeigt, dass $|f(x) - f_n(x)| < \frac{\epsilon}{3}$ für $n \geq N\left(\frac{\frac{1}{2}\epsilon}{3}\right)$. Wählt man daher ein $n \geq N$, so folgt für alle x_1, x_2 mit $d_x(x_1, x_2) < \delta$, wobei δ jetzt nur noch von ϵ abhängt, die Ungleichung

$$|f(x_1) - f(x_2)| < \epsilon.$$

Das heißt f ist stetig auf X. Die erste Beauptung $(\star)$ zeigt außerdem

$$||f - f_n|| < \epsilon \text{ für } n \geq N\left(\frac{\epsilon}{2}\right)$$

Somit konvergiert die Funktionenfolge $f_n(x)$ gleichmäßig gegen $f(x)$. Dies zeigt die Vollständigkeit von $C(X)$. □

Bemerkung 3.33. *Vollständige normierte Räume werden Banach-Räume genannt. Der Funktionenraum $C([a,b])$ ist also ein Banach-Raum. Der folgende Satz von Arzela-Ascoli bedeutet für den Funktionenraum $C([a,b])$ das gleiche, was der Satz von Bolzano- Weierstraß für $\mathbb{R}^n$ bedeutet.*

Satz 3.34. Arzela-Ascoli
Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen in $C([a,b])$, welche gleichmäßig beschränkt und gleichmäßig stetig ist. Es gelte

1. $\sup\limits_{n\in\mathbb{N}} ||f_n||_\infty < \infty$
2. $\forall \epsilon > 0 \; \exists \delta_\epsilon > 0 \; : \; \forall n \in \mathbb{N}$

$$\max_{x_1, x_2 \in [a,b], |x_1 - x_2| \leq \delta_\epsilon} |f_n(x_1) - f_n(x_2)| < \epsilon$$

Dann existiert eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$, welche gegen ein $f \in C([a,b])$ konvergiert, d.h.

$$||f_{n_k} - f||_\infty \underset{k\to\infty}{\longrightarrow} 0$$

(ohne Beweis)

Bemerkung 3.35. *Eine Menge von Funktionen die* 2. *erfüllt, heißt gleichgradig stetig auf* $[a,b]$.

Definition 3.36. *Die (reelle) Treppenfunktion zu einer endlichen Zerlegung des Intervalls $[a,b]$ durch Teilpunkte $a = x_0 < x_1 < \ldots < x_n = b$ ist definiert als*

$$f(x) = c_k \;\; , \;\; x \in (x_{k-1}, x_k)$$

In den Zerlegungspunkten x_k bleibt die Funktion unbestimmt oder wird dort dem jeweiligen Zweck entsprechend geeignet gesetzt, z.B. $f(x_{k-1}) := c_k$.

Anwendung (Approximation durch Treppenfunktionen)
Zum praktischen Rechnen mit stetigen Funktionen (z.B. auf dem Computer) ist die Definition durch Vorgabe von Werten in den überabzählbaren Punkten ihres Definitionsintervalls $[a,b] \subset \mathbb{R}$ nicht geeignet. Daher ist es von Interesse, allgemeine stetige Funktionen durch einfachere, am besten durch von endlich vielen Werten charakterisierte Funktionen zu approximieren. Kandidaten hierfür sind z.B. die Treppenfunktionen, oder Polynome.

Lemma 3.37. *Treppenapproximation*
Jede auf einem beschränkten, abgeschlossenen Intervall $[a,b]$ *definierte, stetige Funktion* $f : [a,b] \to \mathbb{R}$ *lässt sich beliebig gut durch Treppenfunktionen einschließen, d.h.* $\forall \epsilon > 0\ \exists$ *Treppenfunktionen* $\overline{c}_\epsilon, \underline{c}_\epsilon$ *(o.B.d.A. zur selben endlichen Zerlegung von* $[a,b]$ *gehörend) mit den Eigenschaften*

$$\underline{c}_\epsilon(x) \le f(x) \le \overline{c}_\epsilon(x)$$
$$|\underline{c}_\epsilon(x) - \overline{c}_\epsilon(x)| < \epsilon \quad x \in [a,b]$$

3.5 Spezielle Funktionen

Wir haben schon die Exponential-, Sinus- und Cosinusfunktion eingeführt.
Von der Exponentialfunktion wissen wir bereits, dass $f(x) = \exp(x)$ die Funktionalgleichung

$$f(x+y) = f(x)f(y)$$

löst mit der "Anfangsbedingung"

$$f(0) = 1.$$

Hieraus folgt sofort (mit Induktion)

$$f(x) = f\left(\frac{x}{n}\right)^n$$

Lemma 3.38. *Die Exponentialfunktion ist eine bijektive Abbildung von* $\mathbb{R}$ *auf* $\mathbb{R}^+$

Beweis. Die Abbildung $x \to e^x$ ist injektiv, was aus der strikten Monotonie der Funktion folgt. Zum Nachweis ihrer Surjektivität sei $a \in \mathbb{R}^+$ beliebig gegeben.
Da die Folge $(e^n)_{n\in\mathbb{N}}$ wegen $e > 1$ strikt divergiert und die Folge $(e^{-n})_{n\in\mathbb{N}}$ eine Nullfolge ist, gibt es ein $n \in \mathbb{N}$ mit

$$e^{-n} < a < e^n$$

Die Exponentialfunktion ist auf $\mathbb{R}$ und damit auch auf dem Intervall $[-n,n]$ stetig. Aus dem Zwischenwertsatz folgt dann die Existenz eines $c \in [-n,n] : e^c = a$. □

Definition 3.39. *Natürlicher Logarithmus*
Die Umkehrfunktion $f^{-1} : \mathbb{R}^+ \to \mathbb{R}$ *der reellen Exponentialfunktion ist der "natürliche Logarithmus"*$\ln(x)$

$$y := \ln(x),\ \ x \in \mathbb{R}^+ \Leftrightarrow \ \ x = e^y$$

Lemma 3.40. *Der natürliche Logarithmus*

$$\ln(\cdot) : \mathbb{R}^+ \to \mathbb{R}$$

ist stetig und streng monoton wachsend.
Es gelten die folgenden Rechenregeln:

$$\begin{aligned} \ln(xy) &= \ln(x) + \ln(y) \quad x, y \in \mathbb{R}^+ \\ \ln(y^r) &= r\ln(y) \quad y \in \mathbb{R}^+, \ r \in \mathbb{Q} \end{aligned}$$

Beweis s.u.

Beachte: Das gilt nicht für $\exp : \mathbb{C} \to \mathbb{C}$

Lemma 3.41. *Sei* $f : (a,b) \to (c,d)$ *stetig, monoton wachsend und bijektiv. Dann gilt dies auch für*

$$f^{-1} : (c,d) \to (a,b)$$

Beweis. Bijektivität klar.
Monotonie aus Lemma **??**
Stetigkeit von f^{-1}:
Sei $A \subset (a,b)$ offen. Wir müssen zeigen

$$(f^{-1})^{-1}(A) = f(A) \subset (c,d) \text{ ist offen}$$

Wähle:

$$\begin{cases} \text{zu } y \in f(A) \text{ ein } x \in A \text{ mit } y = f(x) \\ \text{und} \\ \epsilon > 0 \text{ mit } \overline{I_\epsilon(x)} \subset A \end{cases}$$

Dann gilt:

$$f(x-\epsilon) < f(x) < f(x+\epsilon)$$

Es folgt $y \in (f(x-\epsilon), f(x+\epsilon))$
und nach dem Zwischenwertsatz gilt

$$(f(x-\epsilon), f(x+\epsilon)) \subset f(A)$$

Also liegt eine ganze Umgebung von y in $f(A)$. $\square$

Beweis. (zu Lemma 3.40)

1. Nach Lemma 3.41 ist der natürliche Logarithmus stetig und streng monoton auf jedem beschränkten Intervall $[e^a, e^b] \subset \mathbb{R}^+$, welches Bild eines Intervalls $[a, b] \subset \mathbb{R}$ ist. Da man mit solchen Intervallen ganz $\mathbb{R}^+$ überdecken kann, folgen Stetigkeit und Monotonie des Logarithmus auf $\mathbb{R}^+$

2. Ausgehend von der Formel
$$e^{x+x'} = e^x e^{x'}$$
folgt für $y = e^x, \;\; y' = e^{x'}$
$$\begin{aligned} yy' &= e^{x+x'}; \\ \ln(yy') = \ln\left(e^{x+x'}\right) &= x + x' = \ln(y) + \ln(y') \end{aligned}$$

Mit $e^{\ln(x)} = x$ ergibt sich
$$e^{\ln(x^r)} = x^r = \left(e^{\ln(x)}\right)^r = e^{r\ln(x)}$$
und folglich wegen der Injektivität der Exponentialfunktion
$$\ln(x^r) = r\ln(x)$$

□

Folgerung 3.42. $\forall a \in \mathbb{R}^+$ *und durch die Festlegung*
$$a^\alpha := e^{\alpha\ln(\alpha)} > 0, \;\; \alpha \in \mathbb{R}$$
wird eine beliebige reelle a-Potenz definiert. Für diese gelten die üblichen Rechenregeln.
$$\begin{aligned} &a^{\alpha+\beta} = a^\alpha a^\beta \\ &a^{-\alpha} = (a^\alpha)^{-1} \\ &(a^\alpha)^\beta = a^{\alpha\beta} = \left(a^\beta\right)^\alpha \end{aligned}$$
Der Ausdruck 0^0 bleibt unbestimmt.
Die Exponentialfunktion $f(x) = a^x$ ist auf $\mathbb{R}$ stetig, für $a > 1$ ist sie streng monoton steigend, für $a = 1$ konstant, für $a < 1$ streng monton fallend.
Die zugehörige Umkehrfunktion ist "der Logarithmus a zur Basis a"
$$y = \log_a(x) \Leftrightarrow \;\; x = a^y$$
Es gilt:
$$\begin{aligned} &\log_a(xy) = \log_a(x) + \log_a(y) \\ &\log_a(x^\alpha) = \alpha\log_a(x) \quad \alpha \in \mathbb{R} \end{aligned}$$

Definition 3.43. *Sei V ein reeller Vektorraum*

1. *$M \subset V$ heißt konvex*
   $\Leftrightarrow$ *$\forall x, y \in M$: Die "Verbindungsstrecke"von x und y liegt ganz in M, d.h.*
$$\forall \lambda \in [0;1] : \lambda x + (1-\lambda)y \in M$$

2. *Sei $M \subset V$ konvex*
$$f : M \to \mathbb{R} \text{ heißt } \begin{cases} \text{(streng) konvex} \\ \text{(streng) konkav} \end{cases} \Leftrightarrow \forall \lambda \in (0,1) :$$
$$\begin{cases} f(\lambda x + (1-\lambda)y) \leq \text{ bzw. } < \lambda f(x) + (1-\lambda) f(y) \\ f(\lambda x + (1-\lambda)y) \geq \text{ bzw. } > \lambda f(x) + (1-\lambda) f(y) \end{cases}$$

   d.h. jede SSehneßwischen zwei Punkten des Graphen von f liegt ganz $\begin{cases} \text{oberhalb} \\ \text{unterhalb} \end{cases}$ des Graphen

Satz 3.44. *1.* exp *ist eine (streng) konvexe Funktion*

2. log *ist eine (streng) konkave Funktion*

Beweis. Für $\lambda \in (0,1)$, $x < y$ gilt:

$$\begin{aligned}
\exp(\lambda x + (1-\lambda)y) & \\
&= \exp(x + (1-\lambda)(y-x)) \\
&= \exp(x)\exp((1-\lambda)(y-x)) \\
&= \exp(x)\left(\lambda + (1-\lambda) + \sum_{j=1}^{\infty}(1-\lambda)^j \frac{(y-x)^j}{j!}\right) \\
&= \lambda\exp(x) + (1-\lambda)\exp(x)\left(1 + \sum_{j=1}^{\infty}\underbrace{(1-\lambda)^{j-1}}_{<1}\frac{(y-x)^j}{j!}\right) \\
&< \lambda\exp(x) + (1-\lambda)\exp(x)\exp(y-x) \\
&= \lambda\exp(x) + (1-\lambda)\exp(y)
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
\log\Big(\lambda x + (1-\lambda)y\Big) &= \log\Big(\lambda\exp(\log(x)) + (1-\lambda)\exp(\log(y))\Big) \\
&> \log\Big(\exp\Big(\lambda\log(x) + (1-\lambda)\log(y)\Big)\Big) \\
&= \lambda\log(x) + (1-\lambda)\log(y)
\end{aligned}$$

□

Kapitel 4

Differenzierbare Abbildungen

4.1 Ableitung

Wir betrachten normierte Vektorräume V_1, V_2

$$M \subset V_1, \;\; f : M \to V_2$$

Die Aussage, dass f in $x_0 \in M$ stetig ist, kann auch wie folgt formuliert werden:

$$f(x) = f(x_0) + r(x, x_0) \quad \forall x \in M$$

mit einem "Fehler" r, wobei

$$\lim_{x \to x_0} r(x, x_0) \to 0$$

f wird also lokal durch die konstante Funktion $x \to f(x_0)$ approximiert. Approximieren wir stattdessen durch eine lineare Funktion $x \to f(x_0) + l_{x_0}(x)$ wobei $l_{x_0}(x) : V_1 \to V_2$ eine lineare stetige Abbildung ist mit einer schärferen Forderung an den "Fehler", so gelangen wir zur Definition der Differenzierbarkeit.

Beispiel 4.1.

$$V_1 = V_2 = \mathbb{C}$$
$$\exp : \mathbb{C} \to \mathbb{C}$$
$$\sin : \mathbb{C} \to \mathbb{C} \;\; in\ x_0 \in \mathbb{C}$$

$$\begin{aligned} \exp(x) = \exp(x_0)\exp(x - x_0) &= \exp(x_0) \sum_{n=0}^{\infty} \frac{(x - x_0)^n}{n!} \\ &= \exp(x_0) + \exp(x_0)(x - x_0) + \exp(x_0)(x - x_0)\rho_1(x, x_0) \end{aligned}$$

mit $\rho_1(x, x_0) = \sum\limits_{n=1}^{\infty} \frac{(x-x_0)^n}{(n+1)!} \to 0$ *für* $x \to x_0$
Die Darstellung im Reellen zeigt, dass die approximierende lineare Funktion als Graph die "Tangenteän den Funktionsgraphen darstellt.
Beide Gleichungen haben die Gestalt:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + ||x - x_0||\rho(x, x_0)$$
$$\text{mit} \quad \rho(x, x_0) \to 0 \;\text{ für } x \to x_0$$

$$x \neq x_0$$
$$\Rightarrow \quad f'(x_0) = \lim_{\substack{x \neq x_0 \\ x \to x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

Definition 4.2. *Seien* V_1, V_2 *normierte Vektorräume (über* $\mathbb{R}$ *bzw.* $\mathbb{C}$*)*
$M \subset V_1$ *sei offen,* $x_0 \in M$
$f : M \to V_2$ *heißt differenzierbar in* x_0
$\Leftrightarrow \quad \exists$ *eine lineare, stetige Abbildung* $f'(x_0) : V_1 \to V_2$
so dass $\forall x \in M$

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + ||x - x_0||\rho(x, x_0)$$

wobei $\rho(x, x_0) \to 0$ *für* $x \to x_0$. $f'(x_0)$ *heißt* *Ableitung* *von* f *in* x_0

Bemerkung 4.3. *Zur Motivation der Forderungen an den Approximationsfehler: Für die lineare Abbildung* $x \mapsto f'(x_0)(x - x_0)$ *gilt* $f'(x_0)(x - x_0) \to 0$ *für* $x \to x_0$. *Es wird daher gefordert, dass der Approximationsfehler* $\hat{\rho}(x, x_0) := f(x) - f(x_0) - f'(x_0)(x-x_0)$ *schneller gegen Null geht als die lineare Abbildung. Daher fordert man, dass der Fehler dividiert durch den linearen Term* $||x - x_0||$ *immer noch gegen Null geht, d.h.* $\frac{\hat{\rho}(x,x_0)}{||x-x_0||} =: \rho(x, x_0) \to 0$ *für* $x \to x_0$. *Setzt man* $\hat{\rho}(x, x_0) = \rho(x, x_0)||x - x_0||$ *erhält man die Forderung aus obiger Definition. Dividiert man die lineare Abbildung hingegen durch* $||x - x_0||$ *so geht dieser Quotient nicht mehr gegen Null für* x *gegen* x_0, *sofern die lineare Abbldung von der Nullabbildung verschieden ist.*

Definition 4.4. f *heißt differenzierbar* $\Leftrightarrow$ f *ist in jedem Punkt* $x \in M$ *differenzierbar*

Beispiel 4.5. 1. $f : \mathbb{R}^m \to \mathbb{R}^n$ *sei linear. Dann gilt* $\forall x_0 \in \mathbb{R}^m$

$$f'(x_0) = f \;, \text{ denn}$$
$$f(x) = f(x_0) + \underbrace{f(x - x_0)}_{\text{lineare Abbildung in } (x - x_0)}$$

2. *Euklidisches Skalarprodukt auf* $\mathbb{R}^n$

$$\begin{aligned} f : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto < x, x > \end{aligned}$$

Es gilt

$$\begin{aligned} f'(x_0) &= 2\langle x_0, \cdot\rangle \ , \ \textit{denn} \\ f(x) &= \langle x, x\rangle = \langle x_0, x_0\rangle + 2\langle x_0, x - x_0\rangle + ||x - x_0||^2 \\ &= f(x_0) + \underbrace{2\langle x_0, x - x_0\rangle}_{\textit{lineare Abbildung in } (x - x_0)} + ||x - x_0||||x - x_0|| \end{aligned}$$

wobei $||x - x_0|| \to 0$ *für* $x \to x_0$.

3. *Determinantenfunktion: Wir schreiben für* $x \in \mathbb{R}^{(n,n)}$, $x = (x^1, ..., x^n)$ *mit 'Spaltenvektoren"*

$$x^j = \begin{pmatrix} x_1^j \\ \vdots \\ x_n^j \end{pmatrix} \in \mathbb{R}^n$$

und definieren

$$\begin{aligned} f : \mathbb{R}^{(n,n)} &\to \mathbb{R} \\ x &\mapsto \det(x^1, ..., x^n) \end{aligned}$$

f *ist differenzierbar* $\forall x_0 \in \mathbb{R}^{(n,n)}$ *mit*

$$f'(x_0) = \sum_{j=1}^{n} \det\left(x_0^1, ..., x_0^{j-1}, x^j - x_0^j, x_0^{j+1}, ..., x_0^n\right)$$

Beweis: Die Determinante ist eine alternierende Multilinearform. Daher gilt

$$\begin{aligned} f(x) &= \det\left(x^1, ..., x^n\right) \\ &= \det\left(x_0^1, x_0^2, ..., x_0^n\right) + \sum_{j=1}^{n} \det\left(x_0^1, ...x_0^{j-1}, x^j - x_0^j, x_0^{j+1}, ..., x_0^n\right) \\ &+ \tilde{\rho}(x^1, x_0^1, \ldots, x^n, x_0^n) \end{aligned}$$

$\tilde{\rho}(x^1, x_0^1, \ldots, x^n, x_0^n)$ *ist die eine Summe über Determinanten* s_k. *Für jede dieser Determinanten gibt es (mindestens)* $i, j \in \{1, \ldots n\}$, $i \neq j$ *sodass Spalte* i *gleich* $x^i - x_0^i$ *und Spalte* j *gleich* $x^j - x_0^j$ *ist.*

Wir verwenden im $\mathbb{R}^{(n,n)}$ *die Norm* $||z|| = \max_{k,l} |z_k^l|$. *Wir entwickeln jeden Summanden* s_k *nach der* i*-ten Spalte (*i *jeweils passend gewählt). Dann gilt für jeden Summanden* $|s_k| \leq \sum_{k=1}^n |x_k^i - x_{0k}^i| \cdot |\hat{\rho}_k(x^1, x_0^1, \ldots, x^n, x_0^n)| \leq \sum_{k=1}^n ||(x^1 - x_0^1, \ldots x^n - x_0^n)|| \cdot |\hat{\rho}_k(x^1, x_0^1, \ldots, x^n, x_0^n)|$, *wobei* $\hat{\rho}_k(x^1, x_0^1, \ldots, x^n, x_0^n)$ *die entsprechenden Unterdeterminanten sind. Jede dieser Unterdeterminanten enthält eine Spalte deren Komponenten eine Untermenge der Komponenten von* $x^j - x_0^j$ *sind (*j *jeweils passend gewählt). Daher geht jede dieser Unterdeterminanten gegen Null für* $x \to x_0$. *Somit ist* $|\tilde{\rho}(x^1, x_0^1, \ldots, x^n, x_0^n)| \leq ||(x^1 - x_0^1, \ldots x^n - x_0^n)|| \sum_k |q_k|$, *mit* $q_k \to 0$ *wenn* $x \to x_0$.

$$
\begin{aligned}
f(x) &= \det\left(x^1, ..., x^n\right) \\
&= \det\left(x_0^1, x_0^2, ..., x_0^n\right) + \sum_{j=1}^n \det\left(x_0^1, ...x_0^{j-1}, x^j - x_0^j, x_0^{j+1}, ..., x_0^n\right) \\
&+ \sum_{j=1}^n \left(\det\left(x_0^1, ..., x_0^{j-1}, x^j - x_0^j, x^{j+1}, ..., x^n\right) - \det\left(x^1, ..., x^{j-1}, x^j - x_0^j, x_0^{j+1}, ..., x_0^n\right)\right)
\end{aligned}
$$

Der zweite Term ist linear in $x - x_0$, *der Dritte nach dem Determinantenentwicklungssatz von der Form*

$$
\sum_{j,k=1}^n x_k^j - x_{0k}^j \rho_{j,k}(x, x_0)
$$

mit $\rho_{j,k}(x, x_0) \to 0$ *für* $x \to x_0$ $\quad$ ($\det$ *ist stetig*).
Verwenden wir im $\mathbb{R}^{(n,n)}$ *die Norm* $||z|| = \max_{j,k} |z_k^j|$, *so können wir ihn also abschätzen durch*

$$
||x - x_0|| \rho(x, x_0)
$$

mit $\rho(x, x_0) \to 0$ *für* $x \to x_0$

Satz 4.6. Komponentenweise Differentiation
Sei

$$
\begin{aligned}
f &: M \to \mathbb{R}^m \\
f(x) &= (f_1(x), ..., f_m(x))
\end{aligned}
$$

Dann gilt:
f *ist in* $x_0 \in M$ *differenzierbar*
$\Leftrightarrow \quad f_i, \;\; i = 1, ..., m$ *sind in* x_0 *differenzierbar.*
Es gilt

$$f'(x_0) = \left(f_1'(x_0), ..., f_m'(x_0)\right)$$

Beweis. Verwende die Definition und den Satz über die komponentenweise Stetigkeit. □

Definition 4.7. *Eine stetige Abbildung*

$$f : [a, b] \to \mathbb{R}^n$$
$$(a, b \in \mathbb{R}, \;\; a < b)$$

heißt <u>*Kurve*</u> *in* $\mathbb{R}^n$*.*
$f([a, b])$ *heißt* <u>*Spur der Kurve*</u>*.*

Beispiel 4.8. *Sei* $f : [a, b] \to \mathbb{R}$ *eine differenzierbare Funktion und*
$f(x) = (x, g(x))$*.*
Die Spur von f *ist der Graph der Funktion* g*. Da* $x \mapsto x$ *und* $x \mapsto g(x)$ *differenzierbar sind, ist nach Satz 4.0.72 auch* $x \mapsto (x, g(x)) = f(x)$ *differenzierbar mit*

$$f^{'}(x_0) = \left(1, g^{'}(x_0)\right)$$

$f^{'}(x_0)$ *gibt die Richtung der "Tangentenän die Kurve* f *an,* $g^{'}(x_0)$ *gibt also ihre Steigung an.*

<u>Beispiele</u> (Abbildungen $\mathbb{R} \to \mathbb{R}$)

1.

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

ist differenzierbar in jedem $x \neq 0$ mit $f^{'}(x) = \text{spur}(x)$, nicht aber in $x = 0$.

2.

$$f : \mathbb{R} \to \mathbb{R}$$
$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

ist ebenfalls nicht differenzierbar in $x = 0$ (obwohl stetig), denn für

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

gilt $\lim\limits_{n\to\infty} a_n = 0$, jedoch

$$\frac{f(a_n) - f(0)}{a_n} = \sin\left(\left(n + \frac{1}{2}\right)\pi\right) = (-1)^n$$

konvergiert nicht für $n \to \infty$.

Rechenregeln für die Ableitung

Lemma 4.9. *Linearität*
Sei $M \subset V_1,\ f, g : M \to V_2$ differenzierbar in $x_0 \in M$.
Dann gilt für $\alpha, \beta \in \mathbb{K} : \alpha f + \beta g$ ist differenzierbar in x_0 mit

$$(\alpha f + \beta g)^{'}(x_0) = \alpha f^{'}(x_0) + \beta g^{'}(x_0)$$

Beweis. Einsetzen in die Definition unter Beachtung der Eindeutigkeit. □

Lemma 4.10. *Produktregel*
Sei $V_2 = \mathbb{K},\ M \subset V_1,\ f : M \to \mathbb{K}, g : M \to \mathbb{K}$ differenzierbar in $x_0 \in M$.
Dann ist auch $fg : M \to \mathbb{K}$ mit $(fg)(x) = f(x)g(x)$ differenzierbar in x_0 mit

$$(fg)^{'}(x_0) = g(x_0)f^{'}(x_0) + f(x_0)g^{'}(x_0)$$

Beweis. Mit $\rho_i(x, x_0) \to 0$ für $x \to x_0,\ \ i = 1, 2$ ist

$$\begin{aligned} f(x) &= f(x_0) + f^{'}(x_0)(x - x_0) + ||x - x_0||_{V_1}\rho_1(x, x_0) \\ g(x) &= g(x_0) + g^{'}(x_0)(x - x_0) + ||x - x_0||_{V_1}\rho_2(x, x_0) \end{aligned}$$

Multiplikation liefert

$$\begin{aligned} f(x)g(x) = f(x_0)g(x_0) &+ (g(x_0)f^{'}(x_0) + f(x_0)g^{'}(x_0))(x - x_0) \\ &+ ||x - x_0||_{V_1}\big(\rho_1(x, x_0)(g(x_0) + g^{'}(x_0)(x - x_0)) \\ &+ \rho_2(x, x_0)(f(x_0) + f^{'}(x_0)(x - x_0))\big) + f'(x_0)(x - x_0)g'(x_0)(x - x_0) \\ &+ \rho_1(x, x_0)\rho_2(x.x_0)||x - x_0||^2_{V_1} \end{aligned}$$

Dabei gilt offenbar

$$\begin{aligned} \rho_3(x, x_0) :=& \big(\rho_1(x, x_0)(g(x_0) + g^{'}(x_0)(x - x_0)) \\ &+ \rho_2(x, x_0)(f(x_0) + f^{'}(x_0)(x - x_0))\big) \underset{x\to x_0}{\to} 0 \end{aligned}$$

und

$$f^{'}(x_0)(x-x_0)g^{'}(x_0)(x-x_0) = ||x-x_0||^2_{V_1}\rho_4(x,x_0)$$

mit

$$|\rho_4(x,x_0)| \leq \sup_{||z||_{V_1}=1}\Big(|f^{'}(x_0)(z)|\Big)\sup_{||z||_{V_1}=1}\Big(|g^{'}(x_0)(z)|\Big)$$

also

$$\begin{aligned} f(x)g(x) = f(x_0)g(x_0) &+ (g(x_0)f^{'}(x_0)+f(x_0)g^{'}(x_0))(x-x_0) \\ &+||x-x_0||_{V_1}\big(\rho_3(x.x_0)+\rho_4(x,x_0)||x-x_0||_{V_1}+\rho_1(x.x_0)\rho_2(x.x_0)||x-x_0||_{V_1}\big) \end{aligned}$$

mit

$$\rho_3(x.x_0)+\rho_4(x,x_0)||x-x_0||_{V_1}+\rho_1(x.x_0)\rho_2(x.x_0)||x-x_0||_{V_1} \underset{x\to x_0}{\longrightarrow} 0$$

□

Lemma 4.11. *Quotientenregel*
$M \subset V_1$, $V_2 = \mathbb{K}, f,g : M \to \mathbb{K}$ seien differenzierbar in $x_0 \in M, g(x) \neq 0$ (da g in x_0 stetig ist, folgt, dass $g(x) \neq 0$ $\forall x$ aus einer geeigneten Umgebung U von x.)
Dann ist auch $\frac{f}{g} : M \to \mathbb{K}$ differenzierbar in x_0 mit

$$\left(\frac{f}{g}\right)^{'}(x_0) = \frac{g(x_0)f^{'}(x_0)-f(x_0)g^{'}(x_0)}{(g(x_0))^2}$$

Beweis. Wir können $f=1$ annehmen. Der allgemeine Fall folgt dann aus der Produktregel.

$$\begin{aligned} & g(x) = g(x_0)+g^{'}(x_0)(x-x_0)+||x-x_0||_{V_1}\rho(x,x_0) \\ & \frac{1}{g(x)}-\frac{1}{g(x_0)} = -\frac{g(x)-g(x_0)}{g(x)g(x_0)} \\ & = -\frac{g^{'}(x_0)(x-x_0)+||x-x_0||_{V_1}\rho(x,x_0)}{g(x)g(x_0)} \\ & = -\frac{g^{'}(x_0)}{g(x_0)^2}(x-x_0)+\left(\frac{1}{g(x_0)^2}-\frac{1}{g(x)g(x_0)}\right)\cdot g^{'}(x_0)(x-x_0) \\ & \quad -||x-x_0||_{V_1}\rho(x,x_0)\frac{1}{g(x)g(x_0)} \end{aligned}$$

Der Betrag des zweiten Terms lässt sich abschätzen durch:

$$||x - x_0|| \qquad \underbrace{\left|\frac{1}{g(x_0)^2} - \frac{1}{g(x)g(x_0)}\right|}_{\underset{x \to x_0}{\longrightarrow} 0 \text{ (Rechenregeln für stetige Funktionen)}} \qquad \sup |g'(x_0)(z)|$$

Also haben wir gezeigt

$$\begin{aligned} \frac{1}{g(x)} = \frac{1}{g(x_0)} - \frac{g'(x_0)}{g(x_0)^2}(x - x_0) \\ + ||x - x_0||\tilde{\rho}(x, x_0) \end{aligned}$$

mit $\tilde{\rho}(x, x_0) \underset{x \to x_0}{\longrightarrow} 0$ d.h.

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$$

□

Lemma 4.12. <u>*Lipschitz-Stetigkeit differenzierbarer Funktionen*</u>
Sei $M \subset V_1$, $f : M \to V_2$ *differenzierbar in* $x_0 \Rightarrow \exists$ *Umgebung* U *um* x_0 *und* $c > 0$, *so dass* $\forall x \in U$:

$$||f(x) - f(x_0)||_{V_2} \leq c||x - x_0||_{V_1}$$

Beweis.

$$||f(x) - f(x_0)||_{V_2} \leq ||f'(x_0)(x - x_0)||_{V_2} + ||x - x_0||_{V_1}||\rho(x, x_0)||_{V_2}$$

$x \to \rho(x, x_0)$ ist stetig in x_0, also beschränkt in einer Umgebung U von x_0

$$||\rho(x, x_0)||_{V_2} \leq c_1 \ \ \forall x \in U$$

Nach dem folgenden Lemma ist ferner

$$||f'(x_0)(x - x_0)||_{V_2} \leq c_0||x - x_0||_{V_1} \ \ \forall x \in V_1$$

Es folgt

$$||f(x) - f(x_0)||_{V_2} \leq (c_0 + c_1)||x - x_0||_{V_1}$$

□

Lemma 4.13. *Sei $l : V_1 \to V_2$ linear*
l ist stetig $\Leftrightarrow$ $\exists\ c_0 > 0\ \forall\ z \in V_1$

$$||l(z)||_{V_2} \leq c_0||z||_{V_1}$$

Beweis. " $\Rightarrow$ " :
l stetig $\Rightarrow$ l stetig in 0

$$\begin{aligned}
&\Rightarrow\ \exists\ \delta > 0\text{: } \forall\ ||\xi||_{V_1} \leq \delta : ||l(\xi)||_{V_2} \leq 1 \\
&\Rightarrow\ \exists\ \delta{>}0\text{: } \forall\ z \in V_1 \setminus \{0\} : ||l(z)||_{V_2} \\
&= \frac{||z||_{V_1}}{\delta}||l\underbrace{\left(\frac{\delta}{||z||_{V_1}}z\right)}_{||\cdot||_{V_1}\leq\delta}||_{V_2} \leq \frac{1}{\delta}||z||_{V_1}
\end{aligned}$$

Wähle $c_0 = \frac{1}{\delta}$.

" $\Leftarrow$ " :

$$\begin{aligned}
&\exists c_0 : \forall z \in V_1 : ||l(z)||_{V_2} \leq c_0||z||_{V_1} \\
\Rightarrow\ &||l(x - x_0)||_{V_2} \leq c_0||x - x_0||_{V_1} < \epsilon
\end{aligned}$$

für vorgegebenes ϵ, x_0, wenn $||x - x_0||_{V_1} < \delta := \frac{\epsilon}{c_0}$
$\Rightarrow$ l ist stetig in $x_0 \Rightarrow l$ ist stetig. □

Lemma 4.14. *Kettenregel*
Seien V_1, V_2, V_3 normierte Vektorräume über $\mathbb{K}$ $(= \mathbb{R}$ oder $\mathbb{C})$ $M \subset V_1$,
Sei $f : M \to V_2$ differenziertbar in x_0, $f(M) \subset N \subset V_2$
Sei $g : N \to V_3$ differenzierbar in $f(x_0)$. Dann ist $(f \circ g) : M \to V_3$ differenzierbar in x_0 mit

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

Bemerkung 4.15. *Die Formel ist sinnvoll, denn es ist*

$$\begin{aligned}
f'(x_0) &: V_1 \to V_2 \ \text{ linear stetig} \\
\rho'(f(x_0)) &: V_2 \to V_3 \ \text{ linear stetig} \\
\Rightarrow\ g'(f(x_0))f'(x_0) &: v_1 \to V_3 \ \text{ linear stetig}
\end{aligned}$$

Beweis.

$$\begin{aligned}
(g \circ f)(x) - (g \circ f)(x_0) &= g(f(x)) - g(f(x_0)) \\
&= g'(f(x_0))(f(x) - f(x_0)) \\
&\quad + ||f(x) - f(x_0)||_{V_2}\rho_1(f(x), f(x_0)) \\
f(x) - f(x_0) &= f'(x_0)(x - x_0) + ||x - x_0||_{V_1}\rho_2(x, x_0) \\
||f'(x) - f'(x_0)||_{V_2} &\leq c||x - x_0|| \ \text{ nach Lemma}
\end{aligned}$$

Einsetzen dieser beiden Zeilen in die obige Gleichung liefert

$$\begin{aligned}(g\circ f)(x)-(g\circ f)(x_0) =& g'(f(x_0))(f'(x_0)(x-x_0))\\ &+||x-x_0||_{V_1}\underbrace{(g'(f(x_0))(\rho_2(x,x_0))+\frac{||f(x)-f(x_0)||_{V_2}}{||x-x_0||_{V_1}}\rho_1(f(x),f(x_0)))}_{=:\rho_3(x,x_0)}\\ =&(g'(f(x_0))f'(x_0))(x-x_0)\\ &+||x-x_0||_{V_1}\rho_3(x,x_0)\end{aligned}$$

Dabei gilt für $x \to x_0$

$$\rho_3(x,x_0)\to 0 \text{ , denn } \rho_2(x,x_0)\to 0$$

$g'(f(x_0))$ ist eine stetig differenzierbare lineare Abbildung.

$$\begin{aligned}&\rho_1(f(x),f(x_0=))\to 0 \ \text{ (da } f(x)\to f(x_0)\text{)}\\ &\frac{||f(x)-f(x_0)||_{V_2}}{||x-x_0||_{V_1}}\le c\end{aligned}$$

□

Beispiel 4.16.

$$(a^x)' = a^x\ln(a)$$

weil $a^x=\exp(x\ln(a))$

Lemma 4.17. *Differenzierbarkeit der Umkehrfunktion*
Sei $f:(a,b)\to\mathbb{R}$ *stetig, streng monoton und in* x_0 *differenzierbar,* $f'(x_0)\neq 0$.
Nach dem Zwischenwertsatz gibt es ein Interval $I\subset\mathbb{R}$, *so dass* $f:(a,b)\to I$ *bijektiv ist. Sei* $g:=f^{-1}:I\to(a,b)$.
Dann ist g *in* $f(x_0)$ *differenzierbar mit*

$$g'(f(x_0))=\frac{1}{f'(x_0)}$$

Beweis.

$$\begin{aligned}\frac{g(\xi)-g(f(x_0))}{\xi-f(x_0)} &= \frac{g(\xi)-x_0}{f(g(\xi))-f(x_0)}\\ &= \frac{1}{\frac{f(g(\xi))-f(x_0)}{g(\xi)-x_0}}\end{aligned}$$

Für eine geeignete Umgebung U von $f(x_0)$ gilt

$$\forall \xi \in U \setminus \{f(x_0)\} : g(\xi) \neq x_0$$
$$f(g(\xi)) = \xi \neq f(x_0)$$

Aus der Stetigkeit von g folgt

$$\lim_{\xi \to f(x_0)} g(\xi) = g(f(x_0)) = x_0$$

also

$$\lim_{\xi \to f(x_0), \xi \neq f(x_0)} \frac{f(g(\xi)) - f(x_0)}{g(\xi) - x_0} = f^{'}(x_0) \neq 0$$

Es folgt

$$\lim_{\xi \to f(x_0), \xi \neq f(x_0)} \frac{g(\xi) - g(f(x_0))}{\xi - f(x_0)} = \frac{1}{f'(x_0)}$$

d.h. g ist differenzierbar in $f(x_0)$ mit

$$g^{'}(f(x_0)) = \frac{1}{f'(x_0)}$$

□

Beispiel 4.18. *1.*

$$f = \exp : \mathbb{R} \to \mathbb{R}^+$$
$$f^{-1} = g = \log : \mathbb{R}^+ \to \mathbb{R}$$

f ist streng monoton wachsend und differenzierbar.

$$f^{'}(x) = e^x \neq 0$$

Daraus folgt mit $y = \exp(x)$

$$\log^{'}(y) = g^{'}(f(x)) = \frac{1}{f'(x)} = \frac{1}{\exp(x)} = \frac{1}{y}$$

2.

$$f = \sin : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to (-1, 1)$$
$$f^{-1} = g = \arcsin : (-1, 1) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

f ist streng monoton wachsend und differenzierbar

$$f'(x) = \cos(x) > 0$$

Damit folgt aus Lemma 4.17

$$\arcsin'(y) = g'(f(x)) = \frac{1}{f'(x)} = \frac{1}{\cos(x)} = \frac{1}{\sqrt{1-\sin^2(x)}} = \frac{1}{\sqrt{1-y^2}}$$

Definition 4.19. *Höhere Ableitungen*
Sei V_2 ein normierter $\mathbb{K}$-Vektorraum.
Sei $M \subset \mathbb{K}$ offen, $x_0 \in M$, $f : M \to V_2$
Wir definieren induktiv:
f ist 0-mal differenzierbar in x_0 $\Leftrightarrow$ f ist stetig in x_0 und

$$f^{(0)}(x_0) := f(x_0)$$

f ist m-mal differenzierbar in x_0 $\Leftrightarrow$ $\exists$ Umgebung U um x_0 : $\forall x \in U \cap M$: f ist $(m-1)$-mal differenzierbar in x_0 und die Abbildung

$$f^{(m-1)} = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{m-1} f : U \cap M \to V_2$$

ist differenzierbar in x_0. Wir setzen

$$f^{(m)}(x_0) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{m} f(x_0) := \left(f^{(m-1)}\right)'(x_0)$$

Bemerkung 4.20. *Hier wurde die lineare Abbildung $f^{(m)}(x_0): M \to V_2$ mit ihrem Wert an der Stelle 1 identifiziert. Auf diese Weise wird die Zuordnung $x \in M \mapsto f^{(m)}(x) \equiv f^{(m)}(x)(1) \in V_2$ zu einer Abbildung von M nach V_2, die dann auf Stetigkeit und Differenzierbarkeit überprüft werden kann.*

Definition 4.21. *f heißt m-mal stetig differenzierbar in M $\Leftrightarrow$ $\forall x \in M$: f ist m-mal differenzierbar in x und die Abbildung*

$$\begin{aligned} M &\to V_2 \\ x &\mapsto f^{(m)}(x) \end{aligned}$$

ist stetig auf M.

$$\begin{aligned} C^m(M, V_2) &:= \{f : M \to V_2 : f \text{ ist } m\text{-mal stetig differenzierbar}\} \\ C^\infty(M, V_2) &:= \bigcap_{m \geq 0} C^m(M, V_2) \end{aligned}$$

Beispiel 4.22. *1. Bahn eines "Massenpunktes"*

$$t \to x(t) \in \mathbb{R}^3, \;\; t \in \mathbb{R} \;\text{ Zeit, } x(t): \;\text{ Ort des Massenpunktes zur Zeit } t$$

Geschwindigkeit zur Zeit zur Zeit t:

$$x^{'}(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t)$$

Beschleunigung zur Zeit t:

$$x^{''}(t) = \frac{d^2}{dt^2}x(t) = x^{(2)}(t)$$

2. $f : K_r(x_0) \to \mathbb{C}$ *sei als Potenzreihe definiert*

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

Dann ist $f \in C^\infty(K_r(x_0))$*, denn wir wissen, dass Potenzreihen differenzierbar sind mit*

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

d.h. $f^{'}$ *ist wieder eine Potenzreihe, usw.*

Lemma 4.23. *Leibniz-Regel*
$M \subset \mathbb{K}$ *offen,* $f, g \in C^n(M, \mathbb{K})$*. Dann gilt*

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}$$

Beweis. Induktion, unter Verwendung der Produktregel □

4.2 Differenzierbare Abbildungen im $\mathbb{R}^n$

Definition 4.24. *Richtungsableitung*
Seien V_1, V_2 *normierte Vektorräume*
Sei $x_0 \in M, 0 \neq y \in V, f : M \to V_2, M \subset V_1$
f *heißt an der Stelle* x_0 *in Richtung* y *differenzierbar* $\Leftrightarrow$ *Die Abbildung*

$$\begin{aligned}[0, \epsilon] &\to V_2 \\ t &\mapsto f(x_0 + ty)\end{aligned}$$

ist in $t = 0$ differenzierbar. (ϵ so klein, dass $0 \leq t \leq \epsilon \Rightarrow \ x_0 + ty \in M$)
Die Richtungsableitung von f in Richtung y an der Stelle x_0 ist dann definiert als:

$$(\frac{\partial}{\partial y} f)(x_0) := \lim_{t\to 0} \frac{f(x_0+ty) - f(x_0)}{t} \in V_2$$

also als Ableitung (Tangentenvektor) der Kurve

$$\begin{aligned} [0,\epsilon) &\to V_2 \\ t &\mapsto f(x_0+ty) \end{aligned}$$

an der Stelle $t = 0$.

Bemerkung 4.25. *1. Ist d differenzierbar in x_0 und $y \in V_1$ mit $x_0 + ty \in M \ \ \forall t \in [0,\epsilon]$, so ist f in x_0 in y differenzierbar, und mit*

$$\begin{aligned} \gamma : [0,\epsilon] &\to M \\ \gamma(t) &= x_0 + ty \end{aligned}$$

gilt

$$\left(\frac{\partial}{\partial y} f\right)(x_0) = (f \circ \gamma)^{'}(0) = f^{'}(x_0) \circ \gamma^{'}(0) = f^{'}(x_0)(y)$$

2. Die Umkehrung gilt nicht. Differenzierbarkeit in alle Richtungen impliziert nicht die Differenzierbarkeit.
<u>Gegenbeispiel</u>:

$$\begin{aligned} f &: \mathbb{R}^2 \to \mathbb{R} \\ f(x_1,x_2) &= \begin{cases} \frac{(x_2)^3}{x_1} & \text{für } x_1 \neq 0 \\ 0 & \text{für } x_1 = 0 \end{cases} \end{aligned}$$

Für $y \neq 0$ existiert

$$\begin{cases} \lim\limits_{t\to 0} \frac{(ty_2)^3}{ty_1} = \frac{(y_2)^3}{y_1} \lim\limits_{t\to 0} t^2 = 0 & \text{für } y_1 \neq 0 \\ 0 & \text{für } y_1 = 0 \end{cases}$$

f ist aber nicht einmal stetig in $x = 0$, da z.B. $\left(\frac{1}{n^2}, \frac{1}{n}\right) \underset{n\to\infty}{\to} 0$, jedoch $f\left(\frac{1}{n^2}, \frac{1}{n}\right) = n \to \infty \neq 0 = f(0)$ gilt.
f ist also nicht differenzierbar.

Definition 4.26. *Partielle Ableitung*
Sei jetzt speziell $V_1 \subset \mathbb{R}^n, e_k := (0, ..., 0, \underbrace{1}_{k\text{-te Stelle}}, 0, ..., 0)$

f *heißt in* $\tilde{x}$ *partiell differenzierbar nach* $x_k \Leftrightarrow$ f *ist in* $\tilde{x}$ *in Richtung* e_k *differenzierbar. Wir definieren*

$$\left(\frac{\partial}{\partial k} f\right)(\tilde{x}) = \left(\frac{\partial}{\partial x_k} f\right)(\tilde{x}) := (\partial_{e_k} f)(\tilde{x})$$

$f : M \to V_2$ *heißt (stetig) partiell differenzierbar nach* $x_k \Leftrightarrow$ $\forall x \in M$ *ist* f *in* x *nach* x_k *differenzierbar und* $x \to (\frac{\partial}{\partial k} f)(x)$ *ist stetig.*

Beispiel 4.27.

$$\begin{aligned}
f : \mathbb{R}^n \setminus \{0\} &\to \mathbb{R} \\
f(x) &= \frac{1}{|x|^{n-2}} = \frac{1}{\left(\sqrt{\sum\limits_{j=1}^{n} {x_j}^2}\right)^{n-2}} \\
(\partial_j f)(x) &= -\frac{n-2}{2} \frac{1}{\left(\sqrt{\sum\limits_{j=1}^{n} {x_j}^2}\right)^{n}} \partial_j \left(\sum_{j=1}^{n} {x_j}^2\right) \\
&= -(n-2) \frac{x_j}{|x|^n} \\
\partial_k(\partial_j f)(x) &= -(n-2) \frac{\delta_{jk}}{|x|^n} + n(n-2) \frac{x_j x_k}{|x|^{n+2}}
\end{aligned}$$

wobei

$$\delta_{jk} = \begin{cases} 0 & \text{für } j \neq k \\ 1 & \text{für } j = k \end{cases}$$

Definition 4.28. *Sei* $\tilde{x} \in M \subset \mathbb{R}^n, f : M \to \mathbb{R}^m$ *sei in* $\tilde{x}$ *nach allen* $x_j, j = 1, ..., n$ *partiell differenzierbar.*
Dann heißt die Matrix

$$Df(\tilde{x}) = (\partial_i f_k(\tilde{x}) = \begin{pmatrix} \partial_1 f_1(\tilde{x}) & \dots & \partial_n f_1(\tilde{x}) \\ & & \vdots \\ \partial_1 f_m(\tilde{x}) & & \partial_n f_m(\tilde{x}) \end{pmatrix} \in \mathbb{R}^{(m,n)}$$

die Jacobi-Matrix von f in $\tilde{x}$
Für $m = 1$ heißt

$$\nabla f(\tilde{x}) = \begin{pmatrix} \partial_1 f(\tilde{x}) \\ \vdots \\ \partial_n f(\tilde{x}) \end{pmatrix} = (Df(\tilde{x}))^T \in \mathbb{R}^n$$

der Gradient von f in $\tilde{x}$

Bemerkung 4.29. 1. *Sei $m = 1, y \in \mathbb{R}^m$ (Spaltenvektor)*

$$\underbrace{\langle \nabla f(\tilde{x}), y \rangle}_{\textit{Skalarprodukt}} = \sum_{j=1}^{n} \partial_j f(\tilde{x}) y_j = \underbrace{Df(\tilde{x})y}_{\textit{Matrizenprodukt}}$$

2. *Ist f in $\tilde{x}$ differenzierbar, so gilt*

$$(\partial_j f)(\tilde{x}) = f^{'}(\tilde{x})(e_j)$$

also für $y \in \mathbb{R}^n$ (Spaltenvektor)

$$\begin{aligned}
f^{'}(\tilde{x})(y) &= f^{'}(\tilde{x}) \sum_{j=1}^{n} y_j e_j \\
&= \sum_{j=1}^{n} y_j f^{'}(\tilde{x})(e_j) = \sum_{j=1}^{n} y_j (\partial_j f)(\tilde{x}) \\
&= \begin{pmatrix} \sum\limits_{j=1}^{n} (\partial_j f_1)(\tilde{x}) y_j \\ \vdots \\ \sum\limits_{j=1}^{n} (\partial_j f_m)(\tilde{x}) y_j \end{pmatrix} \\
&= \begin{pmatrix} \partial_1 f_1(\tilde{x}) & \dots & \partial_n f_1(\tilde{x}) \\ & & \vdots \\ \partial_1 f_m(\tilde{x}) & & \partial_n f_m(\tilde{x}) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad = Df(\tilde{x})y \ \textit{Matrizenprodukt}
\end{aligned}$$

Beispiel 4.30. *Polarkoordinaten im $\mathbb{R}^n$*

$$f : \mathbb{R}^n \setminus \{0\} \to [-1,1]^{n-1} \times (0, \infty)$$

$$f_n = \sqrt{\sum_{j=1}^{n} {x_j}^2} = ||x|| := r$$

$$f_k(x) = \frac{x_k}{r} =: \theta_k \ \ (k = 1, ..., n-1)$$

Es gilt für $j = 1, ..., n$

$$\partial_j f_k(x) = \frac{\delta_{jk}}{r} - \frac{x_j x_k}{r^3} = \frac{\delta_{jk} - \theta_j\theta_k}{r} \;\; ; \; k = 1, ..., n-1$$
$$\partial_j f_n(x) = \frac{x_j}{r} = \theta_j$$

Also

$$Df(x) = \begin{pmatrix} \frac{1-\theta_1{}^2}{r} & -\frac{\theta_2\theta_1}{r} & \cdots & -\frac{\theta_{n-1}\theta_1}{r} & -\frac{\theta_n\theta_1}{r} \\ -\frac{\theta_1\theta_2}{r} & \frac{1-\theta_2{}^2}{r} & \cdots & -\frac{\theta_{n-1}\theta_2}{r} & -\frac{\theta_n\theta_2}{r} \\ \vdots & \vdots & & \vdots & \vdots \\ -\frac{\theta_1\theta_{n-1}}{r} & -\frac{\theta_2\theta_{n-1}}{r} & \cdots & \frac{1-\theta_{n-1}{}^2}{r} & -\frac{\theta_n\theta_{n-1}}{r} \\ \theta_1 & \theta_2 & \cdots & \theta_{n-1} & \theta_n \end{pmatrix}$$

Unter Berücksichtigung von $\sum\limits_{j=1}^{n} \theta_j{}^2 = 1$ *beweist man durch Induktion*

$$\det\big(Df(x)\big) = \frac{1}{r^{n-1}}$$

Satz 4.31. *Sei $M \subset \mathbb{R}^n$ offen, $f : M \to \mathbb{R}^m$ stetig partiell differenzierbar. Dann ist f stetig differenzierbar und es gilt $\forall x \in M, y \in \mathbb{R}^n$:*

$$f'(x)(y) = Df(x)y = \partial_y f(x)$$

Beweis. Es genügt, die stetige Differenzierbarkeit der Komponentenfunktionen aus deren stetiger partieller Differenzierbarkeit zu schließen. Daher können wir o.E $m = 1$ annehmen. Für differenzierbare Funktionen wurde die Formel

$$f'(x)(y) = Df(x)y = \partial_y f(x)$$

bereits bewiesen.
Für $\tilde{x} \in M$, sei δ so klein, dass

$$w_\delta := \{z \in \mathbb{R}^n : |z_j - \tilde{x}_j| < \delta, \quad \text{für } j = 1, ..., n\} \subset M$$

Dann gilt $\forall x \in w_\delta(\tilde{x})$

$$\begin{aligned} f(x) - f(\tilde{x}) &= \sum_{j=1}^{n} \left(f(\tilde{x}_1, ..., \tilde{x}_{j-1}, x_j, ..., x_n) - f(\tilde{x}_1, ..., \tilde{x}_j, x_{j+1}, ..., x_n)\right) \\ &= \sum_{j=1}^{n} \left(g_j(x_j) - g_j(\tilde{x}_j)\right) \end{aligned}$$

mit der stetig differenzierbaren Funktion

$$\begin{aligned} g_j : (\tilde{x}_j - \delta_j, \tilde{x}_j + \delta_j) &\to \mathbb{R} \\ t &\mapsto f(\tilde{x}_1, ..., \tilde{x}_{j-1}, t, \tilde{x}_{j+1}, ..., x_n) \end{aligned}$$

und δ hinreichend klein.
Es gilt

$$g_j(x_j) - g_j(\tilde{x}_j) = g_j^{'}(\xi_j)(x_j - \tilde{x}_j)$$

mit ξ_j zwischen $\tilde{x}_j$ und x_j
$\Big($das zeigen wir später: $\exists \xi \in (a,b) \overset{\text{Mittelwertsatz}}{\Rightarrow} f(b) - f(a) = (b-a)f^{'}(\xi)\Big)$

$$\begin{aligned} &= \partial_j f(\tilde{x}_1, ..., \tilde{x}_{j-1}, \xi_j, x_{j+1}, ..., x_n)(x_j - \tilde{x}_j) \\ &= \partial_j f(\tilde{x})(x_j - \tilde{x}_j) + (x_j - \tilde{x}_j)\rho_j(x, \tilde{x}) \end{aligned}$$

$$\rho_j(x, \tilde{x}) = \partial_j f(\tilde{x}_1, ..., \tilde{x}_{j-1}, \xi_j, x_{j+1}, ..., x_n) - \partial_j f(\tilde{x}) \underset{x \to \tilde{x}}{\to} 0$$

(Stetigkeit), da $\xi_j \in (x_i, \tilde{x}_i)$ gilt $\xi_j \to \tilde{x}_j$ für $x \to \tilde{x}$. Damit haben wir

$$\begin{aligned} f(x) - f(\tilde{x}) = &\sum_{j=1}^{n} \partial_j f(\tilde{x})(x_j - \tilde{x}_j) \\ &+ \sum_{j=1}^{n} (x_j - \tilde{x}_j)\rho_j(x, \tilde{x} \\ &= \langle \nabla f(\tilde{x}), x - \tilde{x} \rangle + ||x - \tilde{x}|| \underbrace{\rho(x, \tilde{x})}_{\underset{x \to \tilde{x}_0}{\to} 0} \\ &= f^{'}(\tilde{x})(x - \tilde{x}) \end{aligned}$$

$x \mapsto Df(x)$ ist nach Voraussetzung stetig. Damit haben wir ein rechnerisch anwendbares Kriterium, das es erlaubt, die Differenzierbarkeit einer Abbildung $\mathbb{R}^n \to \mathbb{R}^m$ aus der stetigen Differenzierbarkeit von Funktionen $\mathbb{R} \to \mathbb{R}$ zu schließen. □

Bemerkung 4.32. *Sei $M \subset \mathbb{R}^n, N \subset \mathbb{R}^m, M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}^n$*
f, g differenzierbar. Dann gilt für $x \in M$

$$\partial_i (g \circ f)_j(x) = \sum_{k=1}^{m} \partial_k g_j(f(x))\partial_i f_k(x)$$

Das ist die elementare Formulierung der Matrizengleichung (da $Df_{i,j} = \partial_j f_i$)

$$\begin{aligned} D(g \circ f)(x) &= Dg(f(x)) \circ Df(x) \\ &\Leftrightarrow \\ (g \circ f)^{'}(x) &= g^{'}(f(x)) \circ f^{'}(x) \end{aligned}$$

Definition 4.33. *Höhere partielle Ableitungen*
Sei $M \subset \mathbb{R}^n, f: M \to \mathbb{R}^m$ und für $k \in \mathbb{N}$ sei $j_k \in \{1, ..., n\}$
Wir definieren für $k > 1$:

1. $\frac{\partial^k}{\partial x_{j_k}, ..., \partial x_{j_1}} f(\tilde{x})$ *existiert für* $\tilde{x} \in M$
   $\Leftrightarrow$ $\exists$ *Umgebung* U *um* $\tilde{x}$: $\forall x \in U \cap M$:

$$\frac{\partial^{k-1}}{\partial x_{j_{k-1}}, ..., \partial x_{j_1}} f(x) \text{ existiert und}$$
$$x \mapsto \frac{\partial^{k-1}}{\partial x_{j_{k-1}}, ..., \partial x_{j_1}} f(x) \text{ in } \tilde{x} \text{ partiell nach } x_{j_k}$$
$$\text{differenzierbar ist.}$$

$$\begin{aligned} \frac{\partial^k}{\partial x_{j_k} ... \partial x_{j_1}} f(\tilde{x}) &:= (\partial_{j_k}, \partial_{j_{k-1}}, ..., \partial_{j_1} f)\tilde{x} \\ &:= \partial_{j_k} \left(\frac{\partial^{k-1}}{\partial x_{j-1_k} ... \partial x_{j_1}} f \right) (\tilde{x}) \end{aligned}$$

2. f *heißt k-mal stetig differenzierbar* $\Leftrightarrow$ $\forall (j_1, ..., j_k) \in \{1, ..., n\}^k$ $\forall x \in M$ $\exists \frac{\partial f}{\partial x_{j_k} ... \partial x_{j_1}} f(x)$ *und* $x \mapsto \frac{\partial^k}{\partial x_{j_k} ... \partial x_{j_1}} f(x)$ *ist stetig in* M

$$\begin{aligned} C^k(M, \mathbb{R}^m) &:= \{f : M \to \mathbb{R}^m : f \text{ ist k-mal stetig differenzierbar.}\} \\ C^\infty(M, \mathbb{R}^m) &:= \bigcap_{k \in \mathbb{N}} C^k(M, \mathbb{R}) \end{aligned}$$

3. f *heißt k-mal differenzierbar in* $\tilde{x}$
   $\Leftrightarrow$ $\exists U$ *um* $\tilde{x}$ *in* M : $f \in C^{k-1}(U, \mathbb{R}^m)$ *und alle partiellen Ableitungen k-ter Ordnung sind differenzierbar in* $\tilde{x}$.

4.3 Mittelwertsätze, Taylor-Entwicklung und Extremalbedingungen

4.3.1 Maximum und Minimum in $\mathbb{R}^1$

Definition 4.34. *Sei $D \subset \mathbb{R}^n$. Die Funktion $f : D \to \mathbb{R}$ hat in $x_0 \in D$ ein "globales Maximum"(Minimum), wenn gilt:*

$$f(x_0) \geq f(x) \;\; \forall x \in D \quad (f(x_0) \leq f(x) \;\; \forall x \in D)$$

Es handelt sich um ein "lokales Maximum (Minimum)", wenn es ein $\delta > 0$ gibt so, dass gilt:

$$f(x_0) \geq f(x) \ \ \forall x \in U_\delta(x_0) = \{x \in D \mid ||x - x_0|| < \delta\}$$
$$(f(x_0) \leq f(x) \ \ \forall x \in U_\delta(x_0))$$

Ein Extremum heißt strikt, wenn es das einzige in D bzw. $U_\delta(x_0)$ ist.

Satz 4.35. *Besitzt eine auf einem Intervall $I = (a, b)$ differenzierbare Funktion f ein lokales Extremum $x_0 \in I$, so gilt notwendigerweise*

$$f^{'}(x_0) = 0$$

Beweis. o.B.d.A. habe f in x_0 ein Minimum. Dann gilt für eine Nullfolge $(h_n)_{n\in\mathbb{N}}$ mit $h_n > 0$ und $x_0 + h_n \in U_\delta(x_0)$:

$$\frac{f(x_0 + h_n) - f(x_0)}{h_n} \geq 0$$

und für eine Nullfolge $(h_n)_{n\in\mathbb{N}}$ mit $h_n < 0$ und $x_0 + h_n \in U_\delta(x_0)$

$$\frac{f(x_0 + h_n) - f(x_0)}{h_n} \leq 0$$

Beim Grenzübergang $h_n \to 0$ folgt

$$\begin{aligned} 0 \leq& f^{'}(x_0) \leq 0 \\ \Rightarrow& f^{'}(x_0) = 0 \end{aligned}$$

Analoge Argumentation für Maxima. $\square$

Bemerkung 4.36. *Die Bedingung $f^{'}(x_0) = 0$ ist notwendig, aber nicht hinreichend.*

$$f(x) = x^3 \ \ \Rightarrow f^{'}(x) = 3x^2 \ \ \Rightarrow f^{'}(x) = 0$$

aber $x_0 = 0$ ist weder Minimum noch Maximum.

Bemerkung 4.37. *Eine stetige Funktion besitzt auf einem abgeschlossenen Intervall $I = [a, b]$ ein Maximum. Dies kann aber auch auf den Randpunkten $x = a$ oder $x = b$ liegen. Dann muss nicht $f^{'}(a) = 0$ $(f^{'}(b) = 0)$ gelten, z.B. $f(x) = x$ auf $I = [0, 1]$ hat Maximum in $x = 1$, aber $f^{'}(1) = 1 \neq 0$*

Satz 4.38. *Satz von Rolle*
Sei $f : [a, b] \to \mathbb{R}$ stetig und in (a, b) differenzierbar. Wenn $f(a) = f(b)$ gilt, dann gibt es ein $c \in (a, b)$ in dem $f^{'}(c) = 0$ ist. Insbesondere liegt zwischen zwei Nullstelle einer differenzierbaren Funktion stets eine Nullstelle der Ableitung.

Beweis. f nimmt auf $[a, b]$ ihr Maximum und ihr Minimum an (Stetigkeit). Ist f konstant, dann ist die Aussage trivialerweise rightig. Ist f nicht konstant, dann $\exists x \in (a, b)$ mit $f(x) > f(a) = f(b)$ oder $f(x) < f(a) = f(b)$. Somit wird das Maximum oder Minimum in $x_0 \in (a, b)$ angenommen.

$$\overset{\text{Satz 4.35}}{\Rightarrow} f^{'}(x_0) = 0$$

□

4.3.2 Mittelwertsätze

Satz 4.39. *Ist f im Intervall $[a, b]$ stetig und in (a, b) differenzierbar, so gibt es ein $c \in (a, b)$, so dass*

$$f^{'}(c) = \frac{f(b) - f(a)}{b - a}$$

Beweis. Wir definieren

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

$\Rightarrow$ stetig in $[a, b]$, differenzierbar in (a, b)

$$g(a) = f(a) = g(b)$$

$\overset{\text{Satz von Rolle}}{\Rightarrow}$ $\exists c \in (a, b)$ mit

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

□

Folgerung 4.40. *Sei $f : (a, b) \to \mathbb{R}$ differenzierbar. Im Fall $f^{'}(x) \geq 0$ $(f^{'}(x) \leq 0)$ für $x \in (a, b)$ ist f monoton steigend (fallend) und im Fall $f^{'}(x) > 0$ $(f^{'}(x) < 0)$ für $x \in (a, b)$ strikt monoton steigend (fallend).*
Im Fall $f^{'} \equiv 0$ auf (a, b) ist f konstant.

Beweis. Sei $f^{'}(x) > 0$ für $x \in (a, b)$. Für $x, y \in (a, b)$ mit $y > x$ gibt es nach dem Mittelwertsatz ein $\xi \in (x, y)$ mit

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

$\Rightarrow$ $f(y) > f(x)$ und f ist strikt monoton steigend.
Die anderen Fälle werden analog bewiesen. □

Folgerung 4.41. *Sei $f : (a,b) \to \mathbb{R}$ zweimal differenzierbar und es gelte $f^{'}(x_0) = 0$ für ein $x_0 \in (a,b)$. Dann hat f im Fall $f^{''}(x_0) > 0$ in x_0 ein striktes lokales Minimum und im Fall $f^{''}(x_0) < 0$ ein striktes lokales Maximum.*

Beweis. Sei $f^{''}(x_0) > 0$. Wegen

$$f^{''}(x_0) = \lim_{x \to x_0} \frac{f^{'}(x) - f^{'}(x_0)}{x - x_0} > 0$$

gibt es ein $\epsilon \in \mathbb{R}_+$ so, dass für x mit $0 < |x - x_0| < \epsilon$ gilt

$$\frac{f^{'}(x) - f'(x_0)}{x - x_0} > 0$$

Mit $f^{'}(x_0) = 0$ folgt

$$f^{'}(x) < 0 \text{ für } x \in (x_0 - \epsilon, x_0) \;\Rightarrow\; f \text{ ist monoton fallend}$$
$$f^{'}(x) > 0 \text{ für } x \in (x_0, x_0 + \epsilon) \;\Rightarrow\; f \text{ ist monoton wachsend}$$

d.h. f hat in x_0 ein striktes lokales Minimum. Der Fall $f^{''}(x_0) < 0$ wird analog behandelt. □

Bemerkung 4.42. *Die Bedingungen sind nicht notwendig.*
Zum Beispiel

$$f(x) = x^4$$

hat ein Minimum in $x_0 = 0$, aber

$$f^{''}(x) = 12x^2 \;\Rightarrow\; f^{''}(0) = 0$$

Folgerung 4.43. *Gilt für eine auf einem offenen (beschränkten oder unbeschränkten) Intervall I definierte und zwei mal differenzierbare Funktion $f : I \to \mathbb{R}$:*

$$f^{''}(x) \geq 0, \; \forall x \in I$$

so ist f konvex, d.h. für $x, y \in I$ und $\lambda \in (0,1)$ gilt

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y) \quad (*)$$

Beweis. Wegen $f^{''} \geq 0$ auf I ist $f^{'}$ monoton wachsend.
Für $x = y$ ist $(*)$ richtig. Seien $x, y \in I$, o.B.d.A $x < y$ und $\lambda \in (0,1)$. Wir setzen

$$x_\lambda := \lambda x + (1-\lambda)y \in (x\ , y).$$

Nach dem Mittelwertsatz gibt es dann $\xi \in (x, x_\lambda)$ und $\eta \in (x_\lambda, y)$ mit

$$\frac{f(x_\lambda) - f(x)}{x_\lambda - x} = f^{'}(\xi) \leq f^{'}(\eta) = \frac{f(y) - f(x_\lambda)}{y - x_\lambda}.$$

Das Relationszeichen gilt aufgrund der eben gezeigten Monotonie von f'. Es gilt

$$\begin{aligned} \underline{x_\lambda - x} &= \lambda x + (1-\lambda)y - x = \underline{(1-\lambda)(y-x)} \\ \underline{y - x_\lambda} &= y - \lambda x - (1-\lambda)y = \underline{\lambda(y-x)} \\ \Rightarrow \quad \frac{f(x_\lambda) - f(x)}{1-\lambda} &\leq \frac{f(y) - f(x_\lambda)}{\lambda} \end{aligned}$$

Durch Umformung dieser Ungleichung erhält man:

$$\begin{aligned} f(x_\lambda) &= \lambda f(x_\lambda) + (1-\lambda) f(x_\lambda) \\ &\leq \lambda f(x) + (1-\lambda) f(y). \end{aligned}$$

□

Folgerung 4.44. *Ist die Ableitung einer in $[a, b]$ stetiger und in $[a, b]$ differenzierbarer Funktion f beschränkt, also $|f^{'}(x)| \leq K \ \forall x \in (a, b)$, so ist f Lipschitz-stetig in $[a, b]$, mit Lipschitz-Konstante K.*

Beweis. Für $x, y \in [a, b], \ \ y > x$ gibt es ein $c \in (a, b)$ mit $f(x) - f(y) = f^{'}(c)(x - y)$ (Mittelwertsatz), es folgt:

$$|f(x) - f(y)| = |f^{'}(c)||x - y| \leq K|x - y|$$

□

Folgerung 4.45. *Sei f in $[a, b]$ stetig und in (a, b) differenzierbar.*
Gilt $f^{'}(x) = 0 \ \forall x \in (a, b)$, dann ist f konstant.

Beweis. Der Beweis folgt aus Folgerung 4.44 mit $K = 0$. □

Beispiel 4.46. *Seien $a_1, ..., a_n \in \mathbb{R}$ gegeben. Wir suchen $a \in \mathbb{R}$ so, dass*

$$f(a) = \sum_{k=1}^{n} (a - a_k)^2$$

minimal wird.

$$f'(a) = \sum_{k=1}^{n} 2(a - a_k) = 2\left(na - \sum_{k=1}^{n} a_k\right)$$

Die einzige Nullstelle von f ist

$$a = \frac{1}{n} \sum_{k=1}^{n} a_k$$

das arithmetische Mittel der a_k.
Aufgrund von

$$f^{''}(a) = 2n > 0$$

ist a_n ein Minimum.
$\longrightarrow$ *Methode der kleinsten Quadrate.*
Bemerkung:
In der Optimierung oder "Ausgleichsrechnung" verwendet man oft die Methode der kleinsten Quadrate. Versucht man z.B. verschiedene Messwerte a_i durch einen geeigneten Wert a derart zu repräsentieren, dass die Summe der quadratischen Abweichungen der einzelnen Messwerte zu dem Ausgleichswert a minimal ist, so gelangt man zu obiger Rechnung. Das arithmetische Mittel ist in der Hinsicht optimal, dass die Summe seiner quadrierten Abstände zu den Werten, aus denen es berechnet wurde, kleiner ist als die Summe der quadrierten Abstände dieser Werte zu irgend einer anderen reellen Zahl.

Beispiel 4.47. *Ziel: Aus dem Fermat'schen Prinzip der geometrischen Optik das Brechungsgesetz herleiten.*
Fermat'sches Prinzip: Ein Lichtstrahl von P_1 und P_2 nimmt den Weg, der die kürzeste Zeit erfordert. Wir betrachten 2 ebene, homogene Medien M_1, M_2 mit den Ausbreitungsgeschwindigkeiten v_1, v_2 für Licht.
gesucht: schnellster Weg von $P_1 = (a, h_1)$ nach $P_2 = (b, h_2)$ $\quad h_1 > 0, h_2 < 0$
Die x-Achse stellt die Trennungslinie zwischen M_1 und M_2 dar.
Annahme: Der schnellste Weg innerhalb eines Mediums ist geradlinig.
Die benötigte Zeit für den Weg von P_1 über P nach P_2 beträgt:

$$\begin{aligned} t(x) &= \frac{\sqrt{(x-a)^2 + {h_1}^2}}{v_1} + \frac{\sqrt{(x-b)^2 + {h_2}^2}}{v_2} \\ t^{'}(x) &= \frac{(x-a)}{v_1\sqrt{(x-a)^2 + {h_1}^2}} - \frac{(b-x)}{v_2\sqrt{(x-b)^2 + {h_2}^2}} \end{aligned}$$

Wegen $t^{'}(b) > 0$ und $t^{'}(a) < 0$ hat $t^{'}$ mindestens eine Nullstelle $x_0 \in (a, b)$.
$t^{'}$ wächst streng monoton, da

$$t^{''}(x) = \frac{{h_1}^2}{v_1\sqrt{(x-a)^2 + {h_1}^2}} + \frac{{h_2}^2}{v_2\sqrt{(b-x)^2 + {h_2}^2}} > 0$$

d.h. x_0 *ist die einzige Nullstelle von* t' *und es gilt*

$$\frac{\frac{x_0 - a}{\sqrt{(x_0-a)^2 + {h_1}^2}}}{\frac{b - x_0}{\sqrt{(x_0-b)^2 + {h_2}^2}}} = \frac{v_1}{v_2}$$

und somit

$$\frac{\sin(\varphi_1)}{\sin(\varphi_2)} = \frac{v_1}{v_2}$$

φ_1*: Einfallswinkel* φ_2*: Brechungswinkel*

Satz 4.48. *verallgemeinerter Mittelwertsatz*
Seien f, g *stetig in* $[a, b]$ *und differenzierbar in* (a, b) *und sei* $g'(x) \neq 0 \ \forall x \in (a, b)$, *so gibt es ein* $c \in (a, b)$ *so, dass gilt*

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)} \quad (*)$$

Beweis. Wegen $g'(x) \neq 0$ in (a, b) ist $g(a) \neq g(b)$.
Es gibt ein $c \in (a, b)$ so, dass

$$\frac{g(b) - g(a)}{b - a} = g'(c) \neq 0$$

Wir definieren

$$\begin{aligned} &F(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)) \\ &F(a) = f(a) = F(b) \\ &\overset{\text{Satz von Rolle}}{\Rightarrow} \quad \exists c \in (a, b) \text{ mit } F'(c) = 0 \end{aligned}$$

d.h.

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

und da $g'(c) \neq 0$ erhalten wir $(*)$ $\square$

4.3.3 Die Regeln von L'Hospital

Die wichtigste Anwendung von Satz 4.48 besteht in den Regeln zur Berechnung unbestimmter Ausdrücke nach L'Hospital. Daber handelt es sich um Grenzübergänge der Form

$$\lim_{x\to a}\frac{f(x)}{g(x)} \in \{\frac{0}{0},\frac{\infty}{\infty},\frac{\infty}{-\infty},\frac{-\infty}{\infty}\}$$

Satz 4.49. *Regeln von L'Hospital*
Es seien f, g zwei auf dem (beschränkten) Intervall $I = (a, b)$ differenzierbare Funktionen. Es gelte $g'(x) \neq 0,\ \forall x \in (a, b)$ und es existiere der Limes

$$\lim_{x\searrow a}\frac{f'(x)}{g'(x)} =: c \in \mathbb{R}$$

Dann gelten die folgenden Regeln:

1. *Im Fall*

$$\lim_{x\searrow a} f(x) = \lim_{x\searrow a} g(x) = 0$$

   ist $g(x) \neq 0$ in I und es gilt

$$\lim_{x\searrow a}\frac{f(x)}{g(x)} = c$$

2. *Im Fall $f(x) \to \pm\infty, g(x) \to \pm\infty$ für $x \searrow a$*
   ist $g(x) \neq 0$ für $a < x < x_* < b$ und es gilt

$$\lim_{x\searrow a}\frac{f(x)}{g(x)} = c$$

   Analog für $x \nearrow b$ und $x \to \pm\infty$

wobei

$$x \searrow x_0 \Leftrightarrow \quad x > x_0, x \to x_0$$
$$x \nearrow x_0 \Leftrightarrow \quad x < x_0, x \to x_0$$

Beweis. 1. Wir fassen f und g als Funktionen auf, die in a stetig sind und dort den Wert Null haben:

$$f(a) = g(a) = 0$$

wegen $g^{'}(x) \neq 0$ kann es dann keine weitere Nullstelle von g in I geben, d.h. $g(x) \neq 0$ in I. Nach Satz 4.48 gibt es zu jedem $x \in I$ ein $\xi \in (a,x)$ mit dem gilt

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f^{'}(\xi)}{g^{'}(\xi)}$$

Dann impliziert der Grenzübergang $x \searrow a$ auch $\xi \searrow a$ und ergibt die Gültigkeit von 1.

2. Sei $\epsilon \in \mathbb{R}^+$ beliebig. Nach Voraussetzung ist $g^{'}(x) \neq 0$ in (a,b). Wir wählen ein $\delta \in \mathbb{R}^+$ mit $a + \delta \leq x_*$ so, dass $\forall x \in (a, a+\delta)$ gilt $f(x) \neq 0$ und $g(x) \neq 0$, sowie

$$\left|\frac{f^{'}(x)}{g^{'}(x)} - c\right| < \epsilon$$

Nach Satz 4.48 gilt dann für beliebige $x, y \in (a, a+\delta)$ mit $x \neq y$ auch $g(x) \neq g(y)$ sowie

$$\left|\frac{f(x) - f(y)}{g(x) - g(y)} - c\right| < \epsilon,$$

da $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f^{'}(\xi)}{g^{'}(\xi)}$, mit einem passenden $\xi \in (a, a+\delta)$. Nun ist für beliebiges $x, y \in (a, a+\delta)$ mit $f(x) \neq f(y)$ und $g(x) \neq g(y)$:

$$\begin{aligned}\frac{f(x)}{g(x)} &= \frac{f(x) - f(y)}{g(x) - g(y)} \frac{g(x) - g(y)}{f(x) - f(y)} \frac{f(x)}{g(x)} \\ &= \frac{f(x) - f(y)}{g(x) - g(y)} \underbrace{\frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}}_{\xrightarrow[x \to a]{} 1}\end{aligned}$$

$\Rightarrow \quad \exists \delta_* \in \mathbb{R}^+$ so, dass $\forall x \in (a, a + \delta_*)$:

$$\left|\frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)}\right| < \epsilon$$

Für x mit $a < x < a + \min\{\delta, \delta_*\}$ ergibt sich damit

$$\left|\frac{f(x)}{g(x)} - c\right| < 2\epsilon$$

$\square$

Bemerkung 4.50. *Grenzprozesse für $x \to \pm\infty$, d.h.*

$$\lim_{x\to\pm\infty} \frac{f(x)}{g(x)}$$

können wir durch die Substitution $y := \frac{1}{x}$ in solche für $y \to 0$, d.h.

$$\lim_{y\to\pm 0} \frac{f\left(\frac{1}{y}\right)}{g\left(\frac{1}{y}\right)}$$

umformuliert werden.
Wir erhalten

$$\lim_{x\to\pm\infty} \frac{f(x)}{g(x)} = \lim_{x\to\pm\infty} \frac{f'(x)}{g'(x)}$$

Bemerkung 4.51. *Bei der Anwendung der Regeln von L'Hospital ist zunächst zu prüfen ob der Limes des Ableitungsquotienten überhaupt existiert.*
Zum Beispiel ist, trotz

$$\lim_{x\searrow 0} \frac{x^2 \sin\left(\frac{1}{x}\right)}{\sin(x)} = \lim_{x\searrow 0} \frac{x}{\sin(x)} x \sin\left(\frac{1}{x}\right) = 0$$

der Schluss

$$\begin{aligned}\lim_{x\searrow 0} \frac{x^2 \sin\left(\frac{1}{x}\right)}{\sin(x)} &= \lim_{x\searrow 0} \frac{2x \sin\left(\frac{1}{x}\right) - x^2 \cos\left(\frac{1}{x}\right) x^{-2}}{\cos(x)} \\ = \lim_{x\searrow 0} \frac{2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)}{\cos(x)} &= -\lim_{x\searrow 0} \cos\left(\frac{1}{x}\right)\end{aligned}$$

nicht zulässig, da der rechte Limes nicht existiert.

Beispiel 4.52. *Auf $I = (0,1)$ gilt*

1.

$$\lim_{x\searrow 0} \frac{\sin(x)}{x} = \lim_{x\searrow 0} \frac{\cos(x)}{1} = \frac{\cos(0)}{1} = 1$$

2. $I = (a,b)$ $p, q \in \mathbb{N}$

$$\lim_{x\searrow a} \frac{x^p - a^p}{x^q - a^q} = \lim_{x\searrow a} \frac{px^{p-1}}{qx^{q-1}} = \frac{pa^{p-1}}{qa^{q-1}} = \frac{p}{q} a^{p-q}$$

3. *$I = (0,1)$:*

$$\lim_{x\searrow 0} \frac{e^{2x}-1}{\ln(1+x)} = \lim_{x\searrow 0} \frac{2e^{2x}}{\frac{1}{1+x}} = 2$$

4. *Durch zweimalige Anwendung der L'Hospital'schen Regeln erhält man auf $I = (0,1)$*

$$\lim_{x\searrow 0} \left(\frac{1}{\sin(x)} - \frac{1}{x}\right) = \lim_{x\searrow 0} \frac{x-\sin(x)}{x\sin(x)}$$
$$= \lim_{x\searrow 0} \frac{1-\cos(x)}{\sin(x)+x\cos(x)} = \lim_{x\searrow 0} \frac{\sin(x)}{2\cos(x)-x\sin(x)} = 0$$

5. *$n \in \mathbb{N}$*

$$\lim_{x\to\infty} \frac{x^n}{e^x} = \lim_{x\to\infty} \frac{nx^{n-1}}{e^x} = \lim_{x\to\infty} \frac{n(n-1)x^{n-2}}{e^x} = \quad ... = \lim_{x\to\infty} \frac{n!}{e^x} = 0$$

Bemerkung 4.53. *1.*

$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} \frac{f(x)}{g(x)^{-1}}$$
$$(\textit{wenn } \lim_{x\to a} f(x) = 0,\ \lim_{x\to a} g(x) = \infty)$$

2.

$$\lim_{x\to a} f(x)^{g(x)}$$

Logarithmieren $\Rightarrow$

$$\lim_{x\to a} g(x)\ln(f(x))$$

und dann wegen Stetigkeit

$$\exp\left(\lim_{x\to a} g(x)\ln(f(x))\right)$$

Beispiel 4.54. *1.*

$$\lim_{x\searrow 0} x^x = ?$$

Logarithmieren $\Rightarrow$

$$\lim_{x\searrow 0} x\ln(x) = \lim_{x\searrow 0} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x\searrow 0} -\frac{1}{x}x^2 = 0$$

und somit

$$\lim_{x\searrow 0} x^x = e^0 = 1$$

2.

$$\lim_{x\to 1} x^{\frac{1}{(x-1)}}$$

Logarithmieren $\Rightarrow$

$$\lim_{x\to 1}\frac{\ln(x)}{x-1}=\lim_{x\to 1}\frac{1}{x}=1$$

und somit

$$\lim_{x\to 1} x^{\frac{1}{x-1}}=e^1=e$$

4.3.4 Taylor Entwicklung

Wir haben schon gesehen, dass sich gewisse Funktionen durch Potenzreihen darstellen lassen, z.B. die trigonometrischen Funktionen.
Die Exponentialfunktion

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

konvergiert absolut für all $x\in\mathbb{R}$.
Dabei gilt für einen beliebigen "Entwicklungspunkt" $x_0\in\mathbb{R}$

$$e^{x-x_0}=\sum_{k=0}^{\infty}\frac{1}{k!}(x-x_0)^k$$

$$\text{bzw. } e^x=\sum_{k=0}^{\infty}\frac{e^{x_0}}{k!}(x-x_0)^k$$

Wir wollen untersuchen unter welchen Bedingungen solch eine Potenzreihenentwicklung für eine Funktion f möglich ist und wie man diese aus f bestimmen kann.
Wir betrachten

$$p(x)=\sum_{k=0}^{n}a_kx^k$$

Mit $x=x-x_0+x_0$ wird hieraus

$$p(x)=\sum_{k=0}^{n}a_k(x-x_0+x_0)^k$$

und mit Hilfe der allgemeinen binomischen Formel

$$p(x) = \sum_{k=0}^{n} b_k (x - x_0)^k$$

mit Koeffizienten b_k.
Statt diese aus den a_k zu bestimmen, wollen wir sie direkt aus der Funktion p ableiten.
m-malige Differentiation für $0 \leq m \leq n$ ergibt:

$$p^{(m)}(x_0) = b_m m!$$

Wir finden also die Darstellung

$$p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(x_0)}{k!} (x - x_0)^k$$

Wir wollen untersuchen, in wie weit sich diese Formel auf allgemeine Funktionen übertragen lässt.

Definition 4.55. *Für eine auf dem offenen Intervall (a, b) definierte und n-mal stetig differenzierbare Funktion f heißt*

$$t_n(x_0, x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

für ein $x_0 \in (a, b)$ das "n-te Taylor-Polynom"von f um x_0.

Wir studieren den Fehler bei der Approximation von f durch das entsprechende Taylor-Polynom.

Satz 4.56. *Sei $f(\cdot)$ eine auf (a, b) definierte und $(n + 1)$-mal stetig differenzierbare Funktion und $t_n(x_0, \cdot)$ ihr n-tes Taylor-Polynom um $x_0 \in (a, b)$. Dann gibt es zu jedem $x \in (a, b)$ ein ξ zwischen x und x_0, so dass gilt*

$$f(x) = t_n(x_0, x) + \frac{f^{(n+1)}(\xi)}{(n + 1)!} (x - x_0)^{n+1}$$

$\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$ *ist das so genannte "lagrange'sche Restglied"der Taylor-Approximation.*

Beweis.

$$t_n(x_0, x_0) = f(x_0)$$

Wir definieren das "Restglied"

$$R_{n+1}(y,x) = f(x) - t_n(y,x)$$

und fassen es (für festes x) als Funktion von y auf. Wegen der $(n+1)$-maligen Differenzierbarkeit von f ist $R_{n+1}(y,x)$ mindestens einmal nach y differenzierbar und es gilt

$$\begin{aligned}
\frac{d}{dy}R_{n+1}(y,x) &= \frac{d}{dy}(f(x) - t_n(y,x)) \\
= -\frac{d}{dy}\sum_{k=0}^{n}\frac{f^{(k)}(y)}{k!}(x-y)^k \\
= -\sum_{k=0}^{n}\frac{f^{(k+1)}(y)}{k!}(x-y)^k + \sum_{k=1}^{n}\frac{f^{(k)}(y)}{(k-1)!}(x-y)^{k-1} \\
= -\frac{f^{(n+1)}(y)}{n!}(x-y)^n \qquad (\star)
\end{aligned}$$

Wir wenden Satz 4.48 an, für

$$\begin{aligned}
f(y) &:= R_{n+1}(y,x) \quad \text{und} \\
g(y) &:= (x-y)^{n+1}
\end{aligned}$$

Es gilt

$$R_{n+1}(x,x) = f(x) - t_n(x,x) = 0,$$

daher

$$\frac{R_{n+1}(y,x)}{(x-y)^{n+1}} = \frac{R_{n+1}(x,x) - R_{n+1}(y.x)}{(x-x)^{n+1} - (x-y)^{n+1}} \overset{\text{Satz } 4.48}{=} \frac{\frac{d}{dy}R_{n+1}(\xi,x)}{-(n+1)(x-\xi)^n}$$

mit einem $\xi \in (a,b)$ zwischen x und y. Mit der obigen Identität $(\star)$ für

$$\frac{d}{dy}R_{n+1}(y,x) \quad \text{für } y = \xi$$

ergibt sich schließlich

$$\frac{R_{n+1}(y,x)}{(x-y)^{n+1}} = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

□

Definition 4.57. 1. *Eine Funktion f auf einem Intervall (a, b) heißt "glattöder C^∞-Funktion,wenn sie beliebig oft differenzierbar ist, d.h. wenn für alle $k \in \mathbb{N}$ $f^{(k)}$ existiert.*
Dann:

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

2. *Konvergiert die Taylor-Reihe von f um x_0 für alle x in einer Umgebung um x_0 und gilt $f(x) = t_\infty(x_0, x)$, so heißt f "(reell) analytischin x_0.*

Satz 4.58. Taylor-Entwicklung
Sei f auf einem beschränkten Intervall (a, b) eine C^∞-Funktion mit gleichmäßig beschränkten Ableitungen, d.h.

$$\sup_{x \in (a,b)} \left| f^{(n)}(x) \right| \leq M < \infty \ , \quad \forall n \in \mathbb{N} \ \ (**)$$

Dann ist f auf (a, b) analytisch, d.h. $\forall x, x_0 \in (a, b)$ konvergiert die Taylor-Reihe von f und es gilt

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

Beweis. Aus der Restglieddarstellung in Satz 4.56 folgt mit Hilfe der Voraussetzung $(**)$

$$\begin{aligned} |f(x) - t_n(x_0, x)| &\leq \frac{\left|f^{(n+1)}(\xi)\right|}{(n+1)!}|x - x_0|^{n+1} \\ &\leq \frac{M}{(n+1)!}(b - a)^{n+1} \end{aligned}$$

zu einem beliebigen $\epsilon \in \mathbb{R}^+$ gibt es nun ein $n_\epsilon \in \mathbb{N}$, so dass $\forall n \geq n_\epsilon$ gilt

$$\frac{M}{(n+1)!}(b - a)^{n+1} < \epsilon$$

□

Bemerkung 4.59. *Eine C^∞-Funktion muss nicht analytisch sein.*
Gegenbeispiel:

$$f(x) = \begin{cases} \exp\left(-x^{-2}\right) & \textit{für } x \neq 0 \\ 0 & \textit{für} x = 0 \end{cases}$$

$$\lim_{x\to 0} \exp\left(-x^{-2}\right) = 0$$

$\Rightarrow$ *f ist in $x = 0$ und damit auf $\mathbb{R}$ stetig.*
Wir wollen die Ableitungen von f in $x_0 = 0$ bestimmen. Zunächst ist für $x \neq 0$:

$$f'(x) = 2x^{-3} \exp\left(-x^{-2}\right)$$
$$f''(x) = \left(4x^{-6} - 6x^{-4}\right) \exp\left(-x^{-2}\right)$$

$\overset{Induktion}{\Rightarrow}$ *Alle Ableitungen in $x \neq 0$ haben die Gestalt*

$$f^{(n)}(x) = p_n\left(x^{-1}\right) \exp\left(-x^{-2}\right) \quad n \geq 1$$

mit gewissen Polynomen p_n.
Alle Ableitungen sind also in $x \neq 0$ stetig. Wir substituieren

$$y := x^{-1}$$

und sehen wegen

$$\frac{y^k}{e^{y^2}} \to 0, \quad (y \to \infty), \quad k \in \mathbb{N}$$

dass sich die Ableitung in $x = 0$ stetig durch Null fortsetzen lassen,

$$f^{(n)}(x) \underset{x\to 0}{\longrightarrow} 0, \quad \forall n \in \mathbb{N}.$$

Hieraus folgt, dass f eine C^∞- Funktion ist. Ihre Taylor-Reihe in $x_0 = 0$ ist offenbar die Nullfunktion, d.h. sie ist trivialerweise für alle $x \in \mathbb{R}$ konvergent, stellt aber die Funktion f außer in $x = 0$ nirgends dar.

Anwendung der Taylor-Entwicklung
Die Taylorentwicklung (mit Restglied) einer Funktion f

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

kann unter anderem zur Berechnung von Funktionswerten $f(x)$ dienen. Die zugehörigen Taylor-Reihen sind konvergent (und stellen die Funktion dar), wenn die Restglieder für $n \to \infty$ gegen Null konvergiert.

1. Exponentialfunktion

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + R_{n+1}(x)$$

mit dem Restglied

$$R_{n+1}(x) = \frac{e^\xi}{(n+1)!} x^{n+1}$$

Die zugehörige Reihe konvergiert offenbar für alle $x \in \mathbb{R}$ und stimmt mit der bekannten Potenzreihendarstellung der Exponentialfunktion überein.

2.

$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + R_{2n+3}(x)$$

mit

$$\begin{aligned} R_{2n+3}(x) &= \frac{\sin^{(2n+3)}(\xi)}{(2n+3)!} x^{2n+3} \\ &= \frac{(-1)^{n+1}\cos(\xi)}{(2n+3)!} x^{2n+3} \end{aligned}$$

und

$$\cos(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + R_{2n+2}(x)$$

mit

$$\begin{aligned} R_{2n+2}(x) &= \frac{\cos^{(2n+2)}(\xi)}{(2n+2)!} x^{2n+2} \\ &= \frac{(-1)^{n+1}\cos(\xi)}{(2n+2)!} x^{2n+2} \end{aligned}$$

3. Logarithmus
Wir betrachten

$$f(x) = \ln(1+x) \text{ um } x_0 = 0$$

Bei Beachtung von $\ln(1) = 0$ und

$$\begin{aligned} \ln^{(k)}(1+x) \mid_{x=0} &= (-1)^{k-1} \frac{(k-1)!}{(1+x)^k} \mid_{x=0} \\ &= (-1)^{k-1}(k-1)!, \quad k \in \mathbb{N} \end{aligned}$$

erhalten wir für $-1 < x \leq 1$

$$\ln(1+x) = \sum_{k=1}^{n} \frac{\ln^{(k)}(1)}{k!} x^k + R_{n+1}$$
$$= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k + R_{n+1}$$

mit

$$R_{n+1} = \frac{\ln^{(n+1)}(1+\xi)}{(n+1)!} x^{n+1} = \frac{(-1)^n}{(n+1)(1+\xi)^{n+1}} x^{n+1}$$

Für festes $x \in (-1, 1)$ ist, da $\xi \in (0,\ x)$ bzw. $\xi \in (x,\ 0)$,

$$|R_{n+1}(x)| \leq \frac{c(x)}{(n+1)}$$

mit einer von x abhängigen Konstanten $c(x)$, so dass das Restglied für $n \to \infty$ gegen Null konvergiert. Im Grenzfall $x = 1$ ist der Zwischenwert $\xi \geq 0$ und daher

$$|R_{n+1}(x)| \leq (n+1)^{-1}$$

Das Restglied geht gegen 0, d.h. die Reihe konvergiert, aber nicht absolut (die harmonische Reihe ist divergent). Für $x = -1$ ist der Logarithmus nicht definiert. Wir finden also, dass die Taylor-Reihe des Logarithmus für $x \in (-1, 1]$ konvergiert (für $x \in (-1, 1)$ sogar absolut). Insbesondere für $x = 1$

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

4.4 Extremalbedingungen in $\mathbb{R}^n$

Definition 4.60. *Für einen Multiindex $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$ heißt*

$$|\alpha|_i = \sum_{i=1}^{n} \alpha_i$$

die Ordnung von α.
Wir schreiben kurz:

$$\alpha! = \alpha_1!...\alpha_n!$$
$$x^\alpha = x_1^{\alpha_1}...x_n^{\alpha_n} \text{ für } x \in \mathbb{R}^n$$
$$\partial^\alpha f = \partial_1^{\alpha_1}...\partial_n^{\alpha_n} f \ \ (\text{wobei } \partial_i^{\alpha_i} f = \underbrace{\partial_i...\partial_i}_{\alpha_i\text{-mal}} f), \quad \partial_i^0 f = f$$

Definition 4.61. *Mit der symbolischen Schreibweise*

$$\nabla := \begin{pmatrix} \partial_1 \\ \vdots \\ \partial_n \end{pmatrix}$$

im $\mathbb{R}^n$ liegt es nahe, für

$$y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

abzukürzen

$$(y, \nabla) := \sum_{j=1}^{n} y_j \partial_j = y_1 \partial_1 + ... + y_n \partial_n$$

Folgerung:

$$\partial_y^k = \langle y, \nabla \rangle^k = \sum_{|\alpha|=k;\alpha \in \mathbb{N}^n} \frac{k!}{\alpha!} y^\alpha \partial^\alpha$$

Satz 4.62. *Sei $M \in \mathbb{R}^n$ offen und liege die Strecke*

$$\{\tilde{x} + t(x - \tilde{x}) : t \in [0, 1]\}$$

ganz in M. Dann gilt für $f \in C^{k+1}(M)$

$$f(x) = \sum_{j=0}^{k} \frac{\langle x - \tilde{x}, \nabla \rangle^j f(\tilde{x})}{j!} + \rho_k(x, \tilde{x})$$

wobei mit geeignetem $\tau \in (0, 1)$:

$$\rho_k(x, \tilde{x}) = \frac{\langle x - \tilde{x}, \nabla \rangle^{k+1} f(\tilde{x} + \tau(x - \tilde{x}))}{(k+1)!}$$

Bemerkung 4.63. *Wir erhalten dann den expliziten Ausdruck für das Taylor-Polynom*

$$\sum_{j=0}^{k} \frac{\langle x - \tilde{x}, \nabla \rangle^j f(\tilde{x})}{j!} = \sum_{|\alpha| \leq k;\alpha \in \mathbb{N}^n} \frac{\partial^\alpha f(\tilde{x})}{\alpha!} (x - \tilde{x})^\alpha$$
$$= \sum_{\alpha_1 + ... + \alpha_n \leq k} \frac{\partial_1^{\alpha_1} ... \partial_n^{\alpha_n} f(\tilde{x})}{\alpha_1! ... \alpha_n!} (x_1 - \tilde{x}_1)^{\alpha_1} ... (x_n - \tilde{x}_n)^{\alpha_n}$$

Setzt man in einer Dimention für $\langle x - \tilde{x}, \nabla \rangle^j$ den Ausdruck $(x - \tilde{x})^j \frac{d^j}{dx^j}$ ein, so wird die Analogie der mehrdimensionalen Form zur Form der Taylorpolynome in $\mathbb{R}$ deutlich.

Satz 4.64. Mittelwertsatz
Sei $M \subset \mathbb{R}^n$ offen, $f : M \to \mathbb{R}$ sei stetig partiell differenzierbar.
Sei $\{\tilde{x} + t(x - \tilde{x}) : t \in [0,1]\} \subset M$.
Dann gilt für ein $t \in (0,1)$:

$$f(x) - f(\tilde{x}) = \langle \nabla f(\tilde{x} + \tau(x,\tilde{x})), x - \tilde{x} \rangle$$

Bemerkung 4.65. Kettenregel
Seien $M \subset \mathbb{R}^n, N \subset \mathbb{R}^m, M \xrightarrow{f} N \xrightarrow{g} \mathbb{R}^k$
Seien f, g differenzierbar. Dann gilt für $x \in M$

$$\partial_l (g \circ f)_j(x) = \sum_{i=1}^{m} \partial_i g_j(f(x)) \partial_l f_i(x)$$

Das ist die elementar aufgeschriebene Formulierung der Matrizengleichung

$$\begin{aligned} D(g \circ f)(x) &= Dg(f(x))Df(x) \\ (g \circ f)'(x) &= g'(f(x)) \circ f'(x) \end{aligned}$$

Satz 4.66. Extremwerte
Sei $M \subset \mathbb{R}^n$ offen, $f : M \to \mathbb{R}$ habe in $\tilde{x} \in M$ ein Extremum (relatives / lokales) und

1. *f ist differenzierbar in $\tilde{x}$, $\nabla f(\tilde{x}) = 0$*
2. *f ist zweimal differenzierbar in $\tilde{x}$*
   Sei $Q : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$Q(y) = \langle y, \nabla \rangle^2 f(\tilde{x}) = \sum_{j,k=1}^{n} \partial_j \partial_k f(\tilde{x}) y_j y_k$$

   Dann gilt

$$f \text{ hat in } \tilde{x} \text{ ein (relatives / lokales)} \begin{cases} \text{Minimum} \\ \text{Maximum} \end{cases}$$

$$\Rightarrow \quad Q(y) \begin{cases} \geq 0 \\ \leq 0 \end{cases} \qquad \forall y \in \mathbb{R}^n$$

Beweis. Wir reduzieren die Aussage auf den schon untersuchten eindimensionalen Fall, indem wir für $y \in \mathbb{R}^n$ die Abbildung

$$F_y : t \to f(\tilde{x} + ty)$$

in einer Umgebung von $t = 0$ betrachten.

1. F_y ist differenzierbar in $t = 0$ und hat dort ein (relatives) Extremum

$$\begin{aligned} \Rightarrow \quad F_y'(0) = 0 \Rightarrow \quad & \langle y, \nabla f(\tilde{x}) \rangle \\ = & \partial_y f(\tilde{x}) = F_y'(0) = 0 \end{aligned}$$

   Da dies $\forall y \in \mathbb{R}^n$ gilt, folgt $\nabla f(\tilde{x}) = 0$.

2. F_y ist zweimal differenzierbar in $t = 0$ und hat in $t = 0$ ein (relatives) Extremum (Minimum, für ein Maximum schließen wir analog). Für kleine $|t|$ gilt:

$$\begin{aligned} & 0 \leq F_y(t) - F_y(0) = tF_y'(\tau) \\ & \text{mit } 0 < |\tau| < |t|, \quad \tau t > 0 \\ & = t(F_y'(\tau) - \underbrace{F_y'(0)}_{=0}) \\ & = t\tau F_y''(0) + t\tau\rho(\tau) \\ & \text{mit } \lim_{\tau \to 0} \rho(\tau) = 0 \end{aligned}$$

$$\Rightarrow \quad F_y''(0) + \rho(\tau) \geq 0 \Rightarrow \quad F_y''(0) \geq 0$$

   Nun ist

$$\begin{aligned} F_y''(0) = \langle y, \nabla \rangle^2 f(\tilde{x}) = Q(y) \\ \Rightarrow Q(y) \geq 0, \end{aligned}$$

   da y beliebig war.

$\square$

Definition 4.67. *Hesse-Matrix*
Sei $M \subset \mathbb{R}^n$ offen und $f : M \to \mathbb{R}$ 2-mal differenzierbar

1. *$D^2 f(x) := (\partial_j \partial_k f(x))$ heißt Hesse-Matrix von f in x.*
   In den Bezeichnungen von oben ist für $y \in \mathbb{R}^n$:

$$Q(y) = \langle y, D^2 f(x) y \rangle$$

2. *Allgemein heißt eine reelle symmetrische $n \times n$-Matrix A bzw. die zugehörige quadratische Funktion Q mit $Q(y) =< y, Ay >$*

   - *positiv (semi-)definit $\Leftrightarrow$: $A > (\geq) 0$*

$$: \Leftrightarrow \quad \forall y \in \mathbb{R}^n \setminus \{0\} : Q(y) > (\geq) 0$$

- *negativ (semi-)definit* $\Leftrightarrow$: $A < (\leq) 0$

$$:\Leftrightarrow \quad \forall y \in \mathbb{R}^n \setminus \{0\} : Q(y) < (\leq) 0$$

- *indefinit* $\Leftrightarrow \exists y_1, y_2$:

$$\begin{aligned} Q(y_1) &> 0 \\ Q(y_2) &< 0 \end{aligned}$$

3. $x \in M$ *heißt kritischer Punkt von* $f \Leftrightarrow \nabla f(x) = 0$

Bemerkung 4.68.
Es gilt $D^2 f(x) = \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \vdots & \cdots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_n \partial x_n} f(x) \end{pmatrix}$.

Satz 4.69. *Sei* $M \subset \mathbb{R}^n$ *offen,* $f \in C^2(M)$.
Sei $\tilde{x} \in M$ *ein kritischer Punkt von* f

$$\begin{aligned} D^2 f(\tilde{x}) < 0 &\Rightarrow \quad \tilde{x} \text{ ist relatives (lokales) Maximum von } f \\ D^2 f(\tilde{x}) > 0 &\Rightarrow \quad \tilde{x} \text{ ist relatives (lokales) Minimum von } f \end{aligned}$$

Bemerkung 4.70. *Beachte, dass es sich hierbei nicht um die Umkehrung der Aussage aus Satz 4.66 handelt. Wenn an der Stelle* x *ein Maximum (Minimum) von* f *ist, so gilt* $D^2 f(x) \leq 0$ (≥ 0), *s. Satz 4.66. Umgekehrt kann man aber in kritischen Punkten von* f *nur folgern, dass es sich um ein Maximum (Minimum) handelt, wenn* $D^2 f(x) < 0$ (> 0) *ist. In kritischen Punkten mit* $D^2 f(x) = 0$ *kann z.B ein Sattelpunkt aber auch ein Extremum vorliegen. Als Beispiele betrachte man etwa im* $\mathbb{R}^1$ $f(x) = x^3$ *bzw.* $f(x) = x^4$, *jeweils in* $x = 0$. $D^2 f(x) \geq 0$ (≤ 0) *ist in kritischen Punkten* x *notwendig für die Existenz eines lokalen Extremums, nicht aber hinreichend, hinreichend ist hingegen* $D^2 f(x) > 0$ (< 0).

Beweis. Aus der Taylor-Formel folgt wegen $\nabla f(\tilde{x}) = 0$

$$f(x) - f(\tilde{x}) = \frac{1}{2} \langle x - \tilde{x}, \nabla \rangle^2 f(\tilde{x} + \tau(x - \tilde{x}))$$

mit $\tau \in (0, 1)$

$$\begin{aligned} &= \frac{1}{2} \sum_{j,k=1}^{n} \partial_j \partial_k f(\tilde{x} + \tau(x - \tilde{x}))(x_j - \tilde{x}_j)(x_k - \tilde{x}_k) \\ &= \frac{1}{2} (x - \tilde{x}, D^2 f(\tilde{x} + \tau(x - \tilde{x}))(x - \tilde{x})) \\ &= \frac{1}{2} (Q(x - \tilde{x}) + P(\tau, x, \tilde{x}) ||x - \tilde{x}||^2) \end{aligned}$$

mit

$$Q(y) = \langle y, D^2 f(\tilde{x}) y \rangle$$

und

$$P(\tau, x, \tilde{x}) \underset{x \to \tilde{x}}{\longrightarrow} 0 \text{ , da } D^2 f \text{ nach Voraussetzung stetig ist.}$$

Ist $Q > 0$, so $\exists c > 0$ mit $\inf\{Q(y) |\ ||y|| = 1\} \geq c$, daher

$$\begin{aligned} f(x) - f(\tilde{x}) =& \frac{1}{2} ||x - \tilde{x}||^2 \cdot \left(Q \left(\frac{x - \tilde{x}}{||x - \tilde{x}||} \right) + P(\tau, x, \tilde{x}) \right) \\ \geq& \frac{1}{2} ||x - \tilde{x}||^2 (c + P(\tau, x, \tilde{x})) \\ \geq& \frac{c}{4} ||x - \tilde{x}||^2 \quad \text{für } ||x - \tilde{x}|| \text{ genügend klein} \\ \geq& 0 \end{aligned}$$

Entsprechend schließt man für $Q<0$. □

Beispiel 4.71. *Sei* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ *eine reelle symmetrische Matrix und*

$$\begin{aligned} &f : \mathbb{R}^2 \to \mathbb{R} \\ &f(x) = \frac{1}{2}(x, Ax) = \frac{a{x_1}^2}{2} + b x_1 x_2 + \frac{c{x_2}^2}{2} \end{aligned}$$

Es gilt

$$\nabla f(x) = \begin{pmatrix} a x_1 + b x_2 \\ b x_1 + c x_2 \end{pmatrix} = Ax$$

und

$$D^2 f(x) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = A$$

(a) $a = c = 1$, $b = 0$*:*

$$\begin{aligned} &\nabla f(x) = 0 \Leftrightarrow \quad x = 0 \\ &D^2 f(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \textit{positiv definit} \end{aligned}$$

$\Rightarrow x = 0$ *ist Minimum.*

(b) $a = c = 0$, $b = 1$:

$$\nabla f(x) = 0 \Leftrightarrow \quad x = 0$$
$$D^2 f(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, D^2 f(0) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = 2 > 0$$
$$\left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, D^2 f(0) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = -2 < 0$$

$\Rightarrow D^2 f(0)$ *ist indefinit.*
0 *ist ein SSattelpunkt"von* f.

Wir betrachten die Funktionen

$$t \mapsto f(tu)$$
$$f(tu) = (bu_1u_2)t^2 \quad ,da\ a = c = 0.$$

mit $||u|| = 1$
Die Graphen dieser Funktionen sind die Schnitte des Graphen von f *mit Ebenen die von* $(0,0,1)$ *und* $(u_1, u_2, 0)$ *aufgespannt werden. Wandert* u *entlag des Einheitskreises, so sind dies Parabeln unterschiedlicher Krümmung, beschrieben durch* bu_1u_2*, die sowohl positiv als auch negativ sein kann.*

1. *Das Trägheitsmoment eines Systems von* l *Massepunkten* $m_1, ..., m_l$ *an den Stellen* $\xi_1, ..., \xi_l$ *im* $\mathbb{R}^2$ *in Bezug auf einen Punkt* x *ist definiert als*

$$f(x) = \sum_{k=1}^{l} m_k ||x - \xi_k||^2$$

Der Schwerpunkt ist der Punkt x*, für den* $f(x)$ *minimal ist.*

$$\nabla f(x) = 2 \sum_{k=1}^{l} m_k (x - \xi_k) \in \mathbb{R}^2 \ (da\ x\ in\ \mathbb{R}^2)$$

Im Schwerpunkt gilt also

$$\nabla f(x) = 0$$
$$\Rightarrow \quad x = \frac{\sum_{k=1}^{l} m_k \xi_k}{m}$$

mit $m = \sum_{k=1}^{l} m_k$

$$D^2 f(x) = \begin{pmatrix} 2m & 0 \\ 0 & 2m \end{pmatrix} \text{ positiv definit.}$$

Definition 4.72. *1. Sei $A \in \mathbb{R}^{n,n}$ eine Matrix,*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Sei $p \leq n$ und $i_j,\ k_j$ $(j = 1, \dots p)$ natürliche Zahlen mit $1 \leq i_1 < i_2 \dots < i_p \leq n$ sowie $1 \leq k_1 < k_2 \dots < k_p \leq n$. Wir betrachten nun die Untermatrix, die aus den Zeilen $i_1, \ \dots i_p$ von A und den Spalten $k_1, \ \dots k_p$ von A besteht. Sie hat die Form

$$\begin{pmatrix} a_{i_1k_1} & \dots & a_{i_1k_p} \\ \vdots & \dots & \vdots \\ a_{i_pk_1} & \dots & a_{i_pk_p} \end{pmatrix}.$$

Ihre Determinate bezeichnen wir durch

$$A\begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix} := \begin{vmatrix} a_{i_1k_1} & \dots & a_{i_1k_p} \\ \vdots & \dots & \vdots \\ a_{i_pk_1} & \dots & a_{i_pk_p} \end{vmatrix}.$$

Alle Determinatne dieser Form heißen Minoren p-ter Ordnung von A.

2. *Es gelten die selben Voraussetzunge wie in 1. Die Determinanten*

$$A\begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix} := \begin{vmatrix} a_{i_1k_1} & \dots & a_{i_1k_p} \\ \vdots & \dots & \vdots \\ a_{i_pk_1} & \dots & a_{i_pk_p} \end{vmatrix}$$

heißen Hauptminoren p-ter Ordnung von A, wenn gilt $i_1 = k_1, \ \dots, \ i_p = k_p$.

3. *Es gelten die selben Notationen wie in 1. Gilt $i_1 = k_1 = 1$ und sind $i_1 = k_1, i_2 = k_2, \dots i_p = k_p$ aufeinander folgende natürliche Zahlen, in dem Sinne, dass $i_2 = i_1 + 1, \ i_3 = 1_2 + 1$ etc. so heißt*

$$A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} := \begin{vmatrix} a_{11} & \dots & a_{1p} \\ \vdots & \dots & \vdots \\ a_{p1} & \dots & a_{pp} \end{vmatrix}$$

führender Hauptminor p-ter Ordnung von A (auch Hauptunterdeterminante oder Hauptabschnittsdeterminante p-ter Ordnung von A genannt).

Bemerkung 4.73. *Sei $A \in \mathbb{R}^{n,n}$ eine Matrix. Dann besitzt A genau n führende Hauptminoren. Diese sind*

$$a_1, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ldots, \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = \det(A)$$

Zu positiv-definiten Matrizen

1. Reelle symmetrische Matrizen haben reelle Eigenwerte.

   Beweis. Bezeichne $\bar{\cdot}$ die komplexe Konjugation. Dann gilt: $v \in C^n \Rightarrow \bar{v}^T v \in \mathbb{R}$ Sei v ein Eigenvektor.

   $$\begin{aligned} &\bar{\lambda}\bar{v}^T v = \overline{\lambda v}^T v = \overline{Av}^T v = \bar{v}^T \bar{A}^T v \\ =&\bar{v}^T A v = \bar{v}^T \lambda v = \lambda \bar{v}^T v \\ &\Rightarrow \quad \bar{\lambda} = \lambda \end{aligned}$$

   □

2. Kriterien für positive Definitheit reeller symmetrischer Matrizen.

   (a) $A \in \mathbb{R}^{n,n}$ positiv definit $\Leftrightarrow$ Alle Hauptunterdeterminanten (so genannte führende Hauptminoren) sind positiv, d.h.

   $$\begin{vmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \\ a_{m1} & \ldots & a_{mm} \end{vmatrix} > 0, \quad \forall m \in \{1, \ldots, n\}.$$

   (Für $n = 2$ haben wir diesen Satz bewiesen, als wir die Positivität quadratischer Formen betrachtet haben, s. Satz 1.53 und Bemerkung 1.54)

   Beispiel 4.74.

   $$A = \begin{pmatrix} 1 & 1 \\ 1 & 10 \end{pmatrix}$$

   $$|1| > 0$$

   $$\begin{vmatrix} 1 & 1 \\ 1 & 10 \end{vmatrix} > 0$$

   $\Rightarrow$ *A ist positiv definit.*

   (b) A ist positiv definit $\Leftrightarrow$ Alle Eigenwerte sind größer als 0

Definition 4.75. *Eine Matrix heißt strik diagonaldominant, falls die Beträge ihrer Diagonalelemente echt größer sind als die Summe der Beträge der jeweils verbleibenden Zeilenelemente, d.h.*

$$\sum_{j=1;i\neq j}^{n} |a_{ij}| < a_{ii} \ \ \forall i = 1, ..., n$$

Symmetrische diagonaldominante Matrizen mit positiven Diagonalemelenten sind positiv definit. Dieses Kriterium ist notwendig, aber nicht hinreichend. Es gibt also auch positiv definite Matritzen, die nicht streng diagonaldominant sind.

Beispiel 4.76. *Beispiel dafür, dass dies kein notwediges Kriterium ist:*

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

Eigenwerte:

$$(3 - 2\sqrt{2}, 3, 3 + 2\sqrt{2})$$

Wir suchen Nullstellen des charakteristischen Polynoms χ_A von A.

$$\begin{aligned} \chi_A(\lambda) =& (3-\lambda)^3 - 2 \cdot 4(3-\lambda) = 0 \Leftrightarrow \\ & (3-\lambda)(9 - 6\lambda + \lambda^2 - 8) = 0 \Leftrightarrow \\ & (3-\lambda)(1 - 6\lambda + \lambda^2) = 0 \end{aligned}$$

$$\begin{aligned} \lambda_1 &= 3 \\ \lambda_{2,3} &= \frac{6 \pm \sqrt{32}}{2} = \frac{6 \pm 4\sqrt{2}}{2} = 3 \pm 2\sqrt{2} \end{aligned}$$

A ist positiv definit, obwohl nicht streng diagonaldominant.

Beispiel 4.77.

Beispiel zur Anwendung des Kriteriums: Sei $A = \begin{pmatrix} 10 & 1 \\ 1 & 100 \end{pmatrix}$. *$A$ ist streng diagonaldominant mit positiven Diagonalelementen. A ist also positiv definit.*

Beispiel 4.78. *1.*

$$A = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 1 & -2 \\ 3 & -2 & 8 \end{pmatrix}$$

$$D_1 = 3$$
$$D_2 = \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} = 3$$
$$D_3 = \begin{vmatrix} 3 & 0 & 3 \\ 0 & 1 & -2 \\ 3 & -2 & 8 \end{vmatrix} = 3$$

$\Rightarrow$ *A ist positiv definit.*

2.

$$A = \begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{pmatrix}$$

$$D_1 = 1$$
$$D_2 = \begin{vmatrix} 1 & 4 \\ 4 & 2 \end{vmatrix} = 2 - 16 = -14$$

$$D_3 = \begin{vmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{vmatrix} = -109$$

$\Rightarrow$ *A ist indefinit.*

Satz 4.79. *Kriterien:*

- *Positive Definitheit* $\Leftrightarrow$ *Es gilt für die führenden Hauptminoren* D_i*:* $D_1 > 0, D_2 > 0, D_3 > 0, \ldots, D_n > 0$.
- *Negative Definitheit* $\Leftrightarrow$*(*$-A$ *ist positiv definit)*$\Leftrightarrow$ *Es gilt:* $D_1 < 0, D_2 > 0, D_3 < 0, \ \ldots, (-1)^n D_n > 0$.
- *Positive Semidefinitheit:* $\Leftrightarrow$ *Es gilt: Alle Hauptminoren* k*-ter Ordnung* Δ_k *sind nichtnegativ:* $\Delta_k \geq 0$.
- *Negative Semidefinitheit:* $\Leftrightarrow$ *Für alle Hauptminoren* k*-ter Ordnung* Δ_k *gilt:* $(-1)^k \Delta_k \geq 0$.

Beispiel 4.80. *Zum Beispiel ist für*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\Delta_1 = a, \ \ \Delta_1 = d$$
$$\Delta_2 = ad - bc$$

Bemerkung 4.81. *Zur Bestimmung der Semidefinitheit ist es notwendig, alle Hauptminoren zu betrachten, nicht nur die führenden Hauptminoren. Sei*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Es gilt $D_1 = 1,\ D_2 = 0,\ D_3 = 0$. *Es sind also alle führenden Hauptminoren nichtnegativ. Trotzdem ist* A *nicht positiv semidefinit, da es die Eigenwerte* $1, 0, -1$ *besitzt und es gilt* $e_3^T A e_3 = -1 < 0$, *wobei* $e_3 = (0, 0, 1)^T$. *Betrachtet man alle Hauptminoren bemerkt man, dass z.B. der Hauptminor* $\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}$ *negaitv ist. Somit ist das Kriterium für positive Semidefinitheit nicht erfüllt.*

Beispiel 4.82.

$$\begin{aligned} f(x_1, x_2) &= x_1^2 - x_2^2 - x_1 x_2 \\ \frac{\partial f}{\partial x_1} &= 2x_1 - x_2 \\ \frac{\partial f}{\partial x_2} &= -2x_2 - x_1 \end{aligned}$$

$$\det D^2 f(x) = \begin{vmatrix} 2 & -1 \\ -1 & -2 \end{vmatrix} = -5 \neq 0$$

Kritische Punkte:

$$\frac{\partial f}{\partial x_1} = 0 \Leftrightarrow x_1 = \frac{x_2}{2}$$

Eingesetzt in $\frac{\partial f}{\partial x_2} = 0$*:*

$$\begin{aligned} \frac{\partial f}{\partial x_2} = 0 &\Leftrightarrow -2x_2 - \frac{x_2}{2} = 0 \\ &\Leftrightarrow -4x_2 - x_2 = 0 \\ &\Leftrightarrow -5x_2 = 0 \end{aligned}$$

$$\Rightarrow \text{Kritischer Punkt bei } (0, 0)$$

Hesse-Matrix

$$\begin{pmatrix} 2 & -1 \\ -1 & -2 \end{pmatrix}$$

$$\begin{aligned} D_1 &= 2 \\ D_2 &= -4 - 1 = -5 \end{aligned}$$

$\Rightarrow$ *Die Matrix ist indefinit.*

Definition 4.83. *Sei $\bar{x}$ ein lokales Maximum, $\bar{x}$ ist ein globales Maximum, wenn gilt*

$$f(\bar{x}) \geq f(\tilde{x})$$

für alle lokalen Maxima $\tilde{x}$.

4.5 Differentiation und Grenzprozesse

Bemerkung 4.84. *patologische Beispiele*

1. *Eine gleichmäßig konvergente Folge differenzierbarer Funktionen mit nichtdifferenzierbarem Limes*

$$f_n(x) := \begin{cases} \frac{n}{2}x^2 + \frac{1}{2n} & \text{für } |x| \leq \frac{1}{n} \\ |x| & \text{für } |x| > \frac{1}{n} \end{cases}$$

2. *Eine Folge differenzierbarer Funktionen, die gleichmäßig gegen eine differenzierbare Funktion konvergiert, aber die Folge der Ableitungen divergiert*

$$f_n(x) := \frac{\sin(n^2 x}{n}$$

3. *Eine Folge differenzierbarer Funktionen, die gleichmäßig gegen eine differenzierbare Funktion konvergiert, die Folge der Ableitungen konvergiert ebenfalls (nicht gleichmäßig), aber nicht gegen die Ableitung der Grenzfunktion*

$$f_n(x) = x - \frac{x^n}{n}$$

Die f_n sind auf $I = [0,1]$ differenzierbar und konvergieren dort gleichmäßig.
$\Longrightarrow$ *$f(x) = x$ ist auch differenzierbar.*

$$f_n^{'}(x) = 1 - x^{n-1}$$

konvergiert in I aber in $x_0 = 1$ nicht gegen $f^{'}(1) = 1$.

Satz 4.85. *Stabilität der Differentiation*
Sei $(f_n)_{n\in\mathbb{N}}$ *eine Folge stetig differenzierbarer Funktionen auf einem (beschränkten) Intervall (offen oder abgeschlossen), welche punktweise gegen eine Funktion* f *konvergiert. Ist die Folge der Ableitungen* $\left(f_n^{'}\right)_{n\in\mathbb{N}}$ *gleichmäßig konvergent gegen eine Funktion* f^*, *so ist auch* f *differenzierbar und*

$$f^{'} = f^*,$$

das heißt

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\lim_{n\to\infty} f_n\right) = \lim_{n\to\infty} f_n^{'}.$$

Beweis. Sei $x_0 \in I$. Wir definieren auf I eine Funktion $\Delta(x)$ durch

$$\Delta(x) := \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & \text{für } x \neq x_0 \\ f^*(x_0) & \text{für } x = x_0 \end{cases}$$

Die Differenzierbarkeit von f in x_0 mit der Ableitung $f^{'}(x_0) = f^*(x_0)$ ist dann gleichbedeutend mit der Stetigkeit von $\Delta(x)$ in $x = x_0$. Für $x \in I \setminus \{x_0\}$ konvergiert

$$\Delta_n(x) := \frac{f_n(x) - f_n(x_0)}{x - x_0} \underset{n\to\infty}{\to} \Delta(x).$$

Nach dem Mittelwertsatz 4.39 gibt es nun Punkte $\xi_n \in I$ zwischen x und x_0, so dass

$$f_n^{'}(\xi_n) = \frac{f_n(x) - f_n(x_0)}{x - x_0} = \Delta_n(x).$$

Folglich ist

$$\Delta(x) - \Delta(x_0) = \Delta(x) - \Delta_n(x) + f_n^{'}(\xi_n) - f^*(x_0).$$

Sei nun ein $\epsilon \in \mathbb{R}^+$ gegeben.
Wir wählen ein $n_0 \in \mathbb{N}$ und ein $\delta \in \mathbb{R}^+$, so dass $\forall n \geq n_0$ und $x \in U_\delta(x_0)$, mit

$$U_\delta(x_0) := \{x \in I : |x - x_0| < \delta\}$$

gilt

$$\begin{aligned} \left|f_n^{'}(x) - f^*(x_0)\right| &\leq |f_n^{'}(x) - f^*(x)| \\ &\quad + |f^*(x) - f^*(x_0)| < \frac{1}{2}\epsilon. \end{aligned}$$

Hier wurde die gleichmäßige Konvergenz

$$f_n' \to f^*$$

und die Stetigkeit von f^* verwendet. Mit $x \in U_\delta(x_0)$ gilt auch $\xi_n \in U_\delta(x_0)$. Zu beliegbigem $x \in U_\delta(x_0) \setminus \{x_0\}$ können wir nun ein $m(x) \geq n_0$ finden, so dass für alle $n \geq m(x)$ gilt

$$|\Delta(x) - \Delta_n(x)| < \frac{1}{2}\epsilon.$$

Für beliebige $x \in U_\delta(x_0) \setminus \{x_0\}$ folgt dann, dass für $n \geq m(x)$ die Abschätzung gilt

$$\begin{aligned} |\Delta(x) - \Delta(x_0)| &\leq |\Delta(x) - \Delta_n(x)| + |f_n'(\xi_n) - f^*(x_0)| \\ &< \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon. \end{aligned}$$

□

Folgerung 4.86. *Seien f_k stetig differenzierbare Funktionen auf dem beschränkten Intervall I (offen und abgeschlossen) mit Ableitung f_k'.*
Wenn die Partialsummen $\sum_{k=1}^{n} f_k$ punktweise und $\sum_{k=1}^{n} f_k'$ auf I gleichmäßig konvergieren, so darf in den zugehörigen Reiehn gliedweise differenziert werden und es gilt

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=1}^{\infty} f_k = \sum_{k=1}^{\infty} f_k'.$$

Satz 4.87. *Sei $M \subset \mathbb{R}^n$, $f : M \to \mathbb{R}^m$ in $\tilde{x} \in M$ 2-mal stetig differenzierbar.*
Dann gilt $\forall j, k = 1, ..., n$

$$\partial_j \partial_k f(\tilde{x}) = \partial_k \partial_j f(\tilde{x})$$

(ohne Beweis)

Bemerkung 4.88. *Ist f k-mal differenzierbar in $\tilde{x}$ $\big(f \in C^k(M)\big)$, so kann die Reihenfolge partieller Differentiationen bis zur k-ten Ordnung beliebig vertauscht werden.*

4.6 Implizite Funktionen und Umkehrabbildungen

4.6.1 Implizite Funktionen

Expliziete Definition: $y = f(x)$
Implizite Darstellung : $F(x, y) = 0$
Beispiel
$F(x, y) = x^2 + y^2 - 1 \leftarrow$ die Punkte des Einheitskreises $\partial K(0) \in \mathbb{R}^2$
Die Gleichung läßt sich formal nach y lösen

$$y = f_{\pm}(x) = \pm\sqrt{1 - x^2}$$

Fragestellung
Sei $D = D^x \times D^y$ eine offene Menge in $\mathbb{R}^n \times \mathbb{R}^m$
$F : D \to \mathbb{R}$ eine stetige Funktion.
Wir wollen untersuchen ob die Gleichung implizit eine Funktion definiert $F(x, y) = 0$

$$x \in D^x, \quad y \in D^y$$

Wir suchen $f : D^x \to D^y$ so dass $F(x, f(x)) = 0$
Ein wichtiger Spezialfall ist $F(x, y) = g(y) - x = 0$

$$y = g^{-1}(x)$$

Satz 4.89 (Implizite Funktion)**.** *Seien $D^x \in \mathbb{R}^n$ und $D^y \in \mathbb{R}^m$ offene Mengen und $F : D^x \times D^y \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Sei $(\bar{x}, \bar{y}) \in D^x \times D^y$ ein Punkt, in dem $F(\bar{x}, \bar{y}) = 0$ gilt und die Jakobi-Matrix $D_y F(\bar{x}, \bar{y}) \in \mathbb{R}^{m\times m}$ regulär ist.*

1. *Dann gibt es eine offene Umgebung $U(\bar{x}) \times U(\bar{y}) \subset D^x \times D^y$ und eine stetige Funktion $f : U(\bar{x}) \to U(\bar{y})$, so dass $F(x, f(x)) = 0$ für $x \in U(\bar{x})$*

2. *Die Funktion f ist eindeutig bestimmt, das heißt ist $(x, y) \in U(\bar{x}) \times U(\bar{y})$ ein Punkt mit $F(x, y) = 0$ so ist $y = f(x)$*

3. *Die Funktion f ist stetig differenzierbar und ihre Jakobi-Matrix*

   $J_{f(\bar{x})} = D_x f(\bar{x}) \in \mathbb{R}^{m\times m}$ ist gegeben durch

$$J_f(\bar{x}) = -D_y F(\bar{x}, \bar{y})^{-1} D_x F(\bar{x}, \bar{y})$$

(ohne Beweis)
Bemerkung
Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ ist regulär, wenn die zugehörige lineare Abbildung bijektiv ist.

Lemma 4.90. *Für $A = (a_{ij})_{j=1}^n \in \mathbb{K}^{n\times n}$ sind die folgenden Aussagen äquivalent:*

- *A ist regulär.*
- *$Ax = b$ ist $\forall b \in \mathbb{K}^n$ eindeutig lösbar.*
- $\det(A) \neq 0$
- *$Rang(A) = n$*

4.6.2 Reguläre Abbildungen

Wir beschäftigen uns jetzt mit der Invertierbarkeit von Abbildungen $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$
Frage: Existenz von $f^{-1} : B_f \to \mathbb{R}^n$
In einer Dimension: Für stetig differenzierbare Funktionen $f : I \subset \mathbb{R} \to \mathbb{R}$ folgt aus der strikten Monotonie, wenn $f'(x) \neq 0,\ x \in D$, die Existenz von $f^{-1} : B_f \to \mathbb{R}$ welches stetig differenzierbar ist.
Beispiel: Die (offene) Teilmenge $D := \{(r, \Theta), r \in \mathbb{R}^+, \Theta \in \mathbb{R}\}$

$$x_1 = f_1(r, \Theta) := r\cos(\Theta)$$
$$x_2 = f_2(r, \Theta) := r\sin(\Theta)$$

Diese Abblidung kann nur lokal injektiv sein.
(Im Streifen $G := \{(r, \Theta), r \in \mathbb{R}^+, \Theta \in [0, 2\pi]\}$)
Für $r = 0$ tritt ein Problem auf, weil

$$(0, \Theta), \Theta \in \mathbb{R} \to (0, 0)$$

was der Injektivität wiederspräche.

Definition 4.91. *Sei $D \subset \mathbb{R}^n$ offen. Eine Abblidung $f : D \to \mathbb{R}^n$ heißt regulär in einem Punkt $\hat{x} \in D$, wenn sie in einer Umgebung $K_\delta(\hat{x}) \subset D$ von $\hat{x}$ stetig differenzierbar ist, und die Matrix $J_f(\hat{x})$ invertierbar ist. Sie heißt regulär in D, wenn sie in jedem Punkt $\hat{x} \in D$ regulär ist.*

Satz 4.92 (Umkehrabbildung)**.** *Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^n$ regulär in einem Punkt $\hat{x} \in D$. Dann gibt es eine offene Umgebung $V(\hat{x}) \subset D$ von $\hat{x}$, die von f bijektiv auf eine offene Umgebung $U(\hat{y}) \subset \mathbb{R}^n$ von $\hat{y} := f(\hat{x})$ abgebildet wird. Die Umkehrabbildung $f^{-1} : U(\hat{y}) \to V(\hat{x})$ ist ebenfalls regulär in $\hat{y}$, und es gilt:*

$$J_{f^{-1}}(\hat{y}) = (J_f(\hat{x}))^{-1}$$
$$\det J_{f^{-1}}(\hat{y}) = \frac{1}{\det J_f(\hat{x})}$$

Beweis. Sei $\hat{x} \in D$ und $y := f(\hat{x}) \in f(D)$. Wir betrachten Abbildungen $F : \mathbb{R}^n \times D \to \mathbb{R}^n$

$$F(y, x) := y - f(x)$$

Offenbar ist $F(\hat{y}, \hat{x}) = 0$
Die Jakobimatrix $D_x F(y, x) = -J_f(x)$ ist regulär in $\hat{x}$ gemäß Vorrausetzungen.
Nach dem Satz über Implizite Funktionen (angewendet mit vertauschten Rollen von x und y) gibt es (offene) Umgebungen $U(\hat{y})$ von $\hat{y}$ und $U(\hat{x})$ von $\hat{x}$ sowie eine (eindeutig bestimmte) stetig differenzierbare Funktion

$$g : U(\hat{y}) \to U(\hat{x})$$

so dass

$$0 = F(y, g(y)) = y - f(g(y)) \quad y \in U(\hat{y})$$

$\Rightarrow$ Es gibt zu jedem

$$y \in U(\hat{y})$$

genau ein

$$x = g(y) \in U(\hat{x}) \text{ mit } y = f(x)$$

Wir setzen

$$V(\hat{x}) := U(\hat{x}) \cap f^{-1}(U(\hat{y})) = \{x \in U(\hat{x}) : f(x) \in U(\hat{y})\}$$

weil

$$U(\hat{x}) \text{ und } f^{-1}(U(\hat{y}))$$

offen ist
$\Rightarrow V(\hat{x})$ offen.
Ferner wird $V(\hat{x})$ bijektiv auf $U(\hat{y})$ abgebildet.
Die Umkehrabbildung ist $f^{-1} = g$.
Wegen $J_{f \circ f^{-1}}(\cdot) = J_{id}(\cdot) = I$ folgt mit der Kettenregel

$$\begin{aligned} J_f(x) \cdot J_{f^{-1}}(f(x)) &= I \\ \to J_f(f(x))' &= (J_f(x))^{-1} \end{aligned}$$

und $\det J_{f-1}(f(x)) = \frac{1}{\det J_f(x)}$ $\square$

Folgerung 4.93. *Ist die Abbildung $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ regulär, so ist für jede offene Menge $O \subset D$, auch die Bildmenge $f(O)$ offen. Solche Abbildungen sind auch als „offen“ bezeichnet.*

Beweis. Sei $O \subset D$ offen und $y \in f(O)$ belibig mit $y = f(x) \quad x \in O$. Nach Satz 4.92 gibt es zu y die Umgebung $K_r(y) \subset f(O)$ sowie $K_s(x) \subset O$, so dass $K_r(y) \subset f(K_s(x))$ folglich ist O offen. □

Bemerkung Eine eindeutige stetige Abbildung f einer offenen Menge $f(D) \subset \mathbb{R}^n$ mit stetiger Umkehrabbildung wird „Homöomorphismus“ genannt. Sind sowohl f als auch f^{-1} stetig differenzierbar, so spricht man von einem „Diffeomorphismus“.

Beispiel $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^2$
$(x_1, x_2) = f(r, \Theta) = (r\cos(\Theta), r\sin(\Theta))$
(sogenannte Transformation der Polarkoordinaten).
Die Jakobimatrix $J_f(r,\Theta) := \begin{pmatrix} \partial_r f_1 & \partial_\Theta f_1 \\ \partial_r f_2 & \partial_\Theta f_2 \end{pmatrix} = \begin{pmatrix} \cos(\Theta) & -r\sin(\Theta) \\ \sin(\Theta) & r\cos(\Theta) \end{pmatrix}$
$\det J_f(r,\Theta) = r > 0$. Die Abbildung f ist also auf ganz $D := \mathbb{R}^+ \times \mathbb{R}$ regulär. Nach Satz 4.92 ist f also überall in D lokal umkehrbar.

$$J_{f^{-1}(x_1,x_2)} = (J_f(r,\Theta))^{-1} = \begin{pmatrix} \cos(\Theta) & \sin(\Theta) \\ -r^{-1}\sin(\Theta) & r^{-1}\cos(\Theta) \end{pmatrix}$$

Wir rechnen $r = \sqrt{x_1^2 + x_2^2}$

$$\begin{aligned} r^{-1}x_1 &= \cos(\Theta) \\ r^{-1}x_2 &= \sin(\Theta) \end{aligned}$$

$$J_{f^{-1}(x_1,x_2)} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1\sqrt{x_1^2 + x_2^2} & x_2\sqrt{x_1^2 + x_2^2} \\ -x_2 & x_1 \end{pmatrix}$$

Hier kann man f^{-1} explizit ausrechnen:

$$U := \mathbb{R}^+ \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad V := \mathbb{R}^+ \times \mathbb{R}$$

f ist bijektiv und $f^{-1}(x,y) = \left(\sqrt{x_1^2 + x_2^2}, \arctan(\frac{x_1}{x_2})\right)$
Die Abbildung f bildet $\mathbb{R}^+ \times \mathbb{R}$ auf $\mathbb{R}^2 \setminus \{0\}$ ab, sie ist aber wegen $f(r,\Theta) = f(r, \Theta + 2k\pi), k \in Z$ nicht global injektiv.

Bemerkung Zum Newton Verfahren in $\mathbb{R}^n$
Nichtlineare Gleichungen $f(x) = 0$ in $\mathbb{K}^n$.
Wenn $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar.
In $\mathbb{R}^1$ haben wir $x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}$ Wichtig ist dass dies konvergiert (Banachscher Fixpunktsatz).
In $\mathbb{R}^n$ schreiben wir: $x^{(k)} = x^{(x-1)} - J_f(x^{(k-1)})^{-1} f(x^{(k-1)}) \quad k \in \mathbb{N}$

4.7 Extremalaufgaben mit Nebenbedingungen

$f : D \to \mathbb{R}, \quad g : D \to \mathbb{R} \quad D \subset \mathbb{R}^n$
Wir suchen $\hat{x} \in D$ mit der Eigenschaft $f(\hat{x}) = \inf(f(x), x \in U(\hat{x}), g(\hat{x}) = 0)$

Satz 4.94 (Lagrange-Multiplikatoren). *Sei $D \subset \mathbb{R}^n$ offen, und $f, g : D \to \mathbb{R}$ stetig differenzierbare Abbildungen. Ferner sei $\hat{x} \in D$ ein Punkt, in dem f ein lokales Extremum unter der Nebenbedingung $g(\hat{x}) = 0$ besitzt, das heißt $f(\hat{x}) = \inf\limits_{x \in U \cap N_g} f(x)$ oder $f(\hat{x}) = \sup\limits_{x \in U \cap N_g} f(x)$, wobei*
$U(\hat{x}) \subset D$ eine Umgebung von $\hat{x}$ ist und

$$N_g := \{x \in D, g(x) = 0\}$$

Ist dann $\nabla g(\hat{x}) \neq 0$, so gibt es ein $\hat{\lambda} \in \mathbb{R}$, so dass

$$\nabla f(\hat{x}) = \hat{\lambda} \nabla g(\hat{x})$$

Der Parameter $\hat{\lambda}$ wird Lagrange-Multiplikator genannt.

Beweis. Wegen $\nabla g(\hat{x}) \neq 0$ können wir nach eventueller Umnumerierung der Koordinaten annehmen, dass $\partial_n g(\hat{x}) \neq 0$. Wir sehen

$$\begin{aligned} \hat{x} &:= (\hat{x}_1, ..., \hat{x}_n) \in \mathbb{R}^n \\ \hat{x}' &= (\hat{x}, ..., \hat{x}_{n-1}) \in \mathbb{R}^{n-1} \\ &\Rightarrow \hat{x} := (\hat{x}', \hat{x}_n) \end{aligned}$$

Satz über implizite Funktionen angewendet auf die Gleichung $F(\hat{x}', \hat{x}_n) = g(\hat{x}) = 0$ liefert die Existenz von $U(\hat{x}') \subset \mathbb{R}^{n-1}$ und $U(\hat{x}_n) \subset \mathbb{R}$ mit $U(\hat{x}') \times U(\hat{x}_n) \subset D$, sowie einer (eindeutig bestimmten) stetig differenzierbaren Funktion $\varphi : U(\hat{x}') \to U(\hat{x}_n)$, so dass $F(x', \varphi(x')) = 0$ und $N_g \cap (U(\hat{x}') \times U(\hat{x}_n)) = \{x \in U(\hat{x}') \times U(\hat{x}_n) : x_n = \varphi(x')\}$
Mit hilfe der Kettenregeln folgt aus

$$F(\underbrace{x'}_{n-1}, \underbrace{\varphi(x')}) = 0$$

$\partial_i g(\hat{x}) + \partial_n g(\hat{x}) \cdot \partial_i \varphi(\hat{x}) = 0 \quad i = 1, ..., n-1$ Da f auf N_g im Punkt $\hat{x}$ ein lokales Extremum besitzt, hat die Funktion $\tilde{f}(x') = F(x', \varphi(x'))$ auch ein lokales Extremum
$\Rightarrow 0 = \partial_i \tilde{f}(\hat{x}') = \partial_i f(\hat{x}) + \partial_n f(\hat{x}) \partial_i \varphi(\hat{x}') \quad i = 1, ..., n-1$
Definieren wir nun $\hat{\lambda} := \partial_n f(\hat{x}) \partial_n g(\hat{x})^{-1}$
beziehungsweise $\partial_n f(\hat{x}) = \hat{\lambda} \partial_n g(\hat{x})$
$\Rightarrow \partial_i f(\hat{x}) = \hat{\lambda} \partial_i g(\hat{x}) \quad i = 1..n$
$\Rightarrow \nabla f(\hat{x}) = \lambda \nabla g(\hat{x})$ $\square$

Bemerkung: Die Aussage vom Satz bedeutet, dass jedes lokale Minimum $\hat{x}$ von f unter der Bedingung $g(\hat{x}) = 0$ notwendig zu einem „Stationären Punkt“ der Lagrange Funktion korrespondiert: $\mathcal{L}(x, \lambda) = f(x) - \lambda g(x), \quad (x, \lambda) \in D \times D$
das heißt: $\nabla_{(x,\lambda)} \mathcal{L}(\hat{x}, \hat{\lambda}) = \begin{pmatrix} \nabla_x f(\hat{x}) - \hat{\lambda} \nabla_x g(\hat{x}) \\ g(\hat{x}) \end{pmatrix} = 0 \rightarrow$ „Euler-Lagrange-Formalismus”
(„induzierte“ Lösungsmethode)

Beispiel Sei $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times m}$ eine symmetrische Matrix und f die zugehörige quadratische Form

$$f := (x, Ax)_2 = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

Wir wollen extrema von f unter der Nebenbedingung $||x||_2 = 1$ finden.
Wir definieren $g(x) = ||x||_2 - 1$, $N_g = \{x \in \mathbb{R}^n, g(x) = 0\}$
Wegen $a_{ij} = a_{ji}$: $\partial_k f(x) = \sum_{i,j=1}^n a_{ij} x_j \delta_{ik} + \sum_{i,j=1}^n a_{ij} x_i \delta_{jk} = \sum_{j=1}^n a_{kj} x_j + \sum_{i=1} a_{ik} x_i = 2 \sum_{i=1}^n a_{ki} x_i, \quad k = 1, ..., n$
$\delta_{ik} = \begin{cases} 1 & i = k \\ 0 & \text{sonst} \end{cases}$

$$\nabla f(x) = 2Ax$$

Satz $\Rightarrow \exists \hat{\lambda} \in \mathbb{R} : \underset{\nabla f}{A\hat{x}} = \underset{\nabla g}{\hat{\lambda}\hat{x}}$

$$f(\hat{x}) = (\hat{x}, A\hat{x})_2 = (\hat{x}, \hat{\lambda}\hat{x}) = \hat{\lambda}$$

$\Rightarrow$ Minumum wird bei $\hat{x} =$ Eigenvektor zum kleinsten Eigenwert $\hat{\lambda}_{min}$ angenommen.
Es folgt

$$\lambda_{min} = \min_{\lambda \in \mathbb{R}^n} \frac{(x, Ax)_2}{||x||_2^2}$$

das heißt λ_{min} ist das Minimum des sogenannten „Rayley Quotienten“ $\mathbf{R(x)} := \frac{\mathbf{(x,Ax)_2}}{\mathbf{||x||_2^2}}$