Analysis 1

Ulrich Bunke

July 11, 2008

Contents

1 Einführung in die mathematische Sprache

1.1 Aussagen

Hier sind einige Beispiele von Aussagen über eine natürliche Zahl $n \in \mathbb{N} := \{1, 2, 3, \dots\}$.

1. $P_1(n) :=$ Die Zahl n ist durch 2 teilbar.
2. $P_2(n) :=$ Die Zahl n ist größer als 4.
3. $P_3(n) :=$ Die Zahl n is prim.

Die Symbolik $A :=$ ... bedeutet hier, daß das Symbol als Abkürzung für ... verwendet werden soll. Um anzugeben, daß eine Aussage P wahr oder falsch ist, schreiben wir $P = w$ oder $P = f$.

In Abhängigkeit der betrachteten Zahl n kann so eine Aussage $P_i(n)$ wahr (w) oder falsch (f) sein. Die folgende Tabelle illustriert die Werte einiger Aussagen.

1. $P_1(n) = \begin{cases} w & n = 2, 4, 6, 8, \ldots \\ f & n = 1, 3, 5, 7, \ldots \end{cases}$

2. $P_2(n) = \begin{cases} w & n = 5, 6, 7, 8, \ldots \\ f & n = 1, 2, 3, 4 \end{cases}$

3. $P_3(n) = \begin{cases} w & n = 2, 3, 5, 7, \ldots \\ f & n = 1, 4, 6, 8, 9, 10, \ldots \end{cases}$

Aussagen können negiert oder kombiniert werden:

Negation:

1. $\sim P_1(n) =$ Die Zahl n ist nicht durch 2 teilbar.
2. $\sim P_2(n) =$ Die Zahl n ist nicht größer als 4.
3. $\sim P_3(n) =$ Die Zahl n ist nicht prim.

Sind A, B zwei Aussagen, dann schreiben wir $A = B$ um zu sagen, daß die Aussagen die **gleichen Wahrheitswerte** haben. Es gilt also $A = B$ wenn entweder beide wahr oder beide falsch sind. Andernfalls gilt $A \neq B$. Ist n eine vorher fixierte Zahl (wie es verabredungsgemäß in diesem Abschnitt der Fall ist), dann bedeutet $P_1(n) = P_2(n)$ die Gleichheit der Aussagen für genau diese Zahl. Man kann aber n auch variabel annehmen, was aus dem Kontext klar werden muß. In diesem Fall würde $P_1(n) = P_2(n)$ bedeuten, daß die Aussagen für jede Wahl von n gleich sind.

Die Negation wirkt auf die Wahrheitswerte wie folgt.

P	$\sim P$
w	f
f	w

Und:

1. $P_1(n) \wedge P_2(n) =$ Die Zahl n ist durch 2 teilbar und größer als 4.
2. $P_2(n) \wedge P_3(n) =$ Die Zahl n ist größer als 4 und prim.
3. $P_2(n) \wedge P_3(m) =$ Die Zahl n ist größer als 4 und die Zahl m ist prim (*Dies ist eine Aussage über zwei natürliche Zahlen* n, m).

Die Und-Verknüfung wirkt auf die Wahrheitswerte wie folgt.

P_1	P_2	$P_1 \wedge P_2$
w	w	w
f	w	f
w	f	f
f	f	f

Zum Beispiel gilt:

$$P_1(4) \wedge P_2(4) = f \ , P_1(6) \wedge P_2(6) = w \ , P_2(7) \wedge P_3(11) = w$$

Oder:

1. $P_1(n) \vee P_2(n) =$ Die Zahl n ist durch 2 teilbar oder größer als 4.
2. $P_2(n) \vee P_3(n) =$ Die Zahl n ist größer als 4 oder prim.
3. $P_2(n) \vee P_3(m) =$ Die Zahl n ist größer als 4 und die Zahl m ist prim. (*Dies ist eine Aussage über zwei natürliche Zahlen n, m*)

Die Oder-Verknüfung wirkt auf die Wahrheitswerte wie folgt.

P_1	P_2	$P_1 \vee P_2$
w	w	w
f	w	w
w	f	w
f	f	f

Zum Beispiel gilt:

$$P_1(4) \vee P_2(4) = w \ , P_1(3) \vee P_2(3) = f \ , P_2(7) \vee P_3(11) = w$$

Wenn - dann:

1. $P_1(n) \to P_2(n) =$ Wenn die Zahl n ist durch 2 teilbar ist, dann ist sie größer als 4.
2. $P_2(n) \to P_3(n) =$ Wenn die Zahl n ist größer als 4 ist, dann ist sie prim.
3. $P_2(n) \to P_3(m) =$ Wenn die Zahl n ist größer als 4 dann ist die Zahl m prim.

Die Wenn-dann Verknüfung wirkt auf die Wahrheitswerte wie folgt.

P_1	P_2	$P_1 \to P_2$
w	w	w
f	w	w
w	f	f
f	f	w

Zum Beispiel gilt:

$$P_1(4) \to P_2(4) = f \ , P_1(3) \to P_2(3) = w \ , P_2(7) \to P_3(11) = w \ .$$

Dann und nur dann:

1. $P_1(n) \leftrightarrow P_2(n) =$ Die Zahl n ist dann und nur dann durch 2 teilbar , wenn sie größer als 4 ist.

2. $P_2(n) \leftrightarrow P_3(n) =$ Die Zahl n ist dann und nur dann größer als 4, wenn sie prim ist.

3. $P_2(n) \leftrightarrow P_3(m) =$ Die Zahl n ist dann und nur dann größer als 4, wenn die Zahl m prim ist.

Die dann und nur dann Verknüfung wirkt auf die Wahrheitswerte wie folgt.

P_1	P_2	$P_1 \leftrightarrow P_2$
w	w	w
f	w	f
w	f	f
f	f	w

Zum Beispiel gilt:

$$P_1(4) \leftrightarrow P_2(4) = f \ , P_1(3) \leftrightarrow P_2(3) = w \ , P_2(7) \leftrightarrow P_3(11) = w \ .$$

Dir Verknüpfung von Aussagen kann man iterieren.

1. $\sim P_1(n) \wedge P_1(n) =$ Die Zahl n ist durch 2 teilbar und sie ist nicht durch 2 teilbar.

2. $P_1(n) \wedge (P_2(n) \rightarrow P_3(n)) =$ Die Zahl n ist durch 2 teilbar, und wenn n größer als 4 ist, dann ist n prim.

Wir betrachten eine Kombination $C(P_1, \ldots, P_k)$ von Aussagen, zum Beispiel

$$C(P_1, P_2, P_3) := (P_1 \wedge (P_1 \rightarrow P_2)) \vee \sim (P_2 \wedge P_3) \ .$$

Die Tabelle

P_1	P_2	P_3	$C(P_1, P_2, P_3)$
w	w	w	w
w	w	f	w
w	f	w	w
w	f	f	w
f	w	w	f
f	w	f	w
f	f	w	w
f	f	f	w

heißt Wahrheitswertetabelle. Wir betrachten Kombinationen aus Aussagen als äquivalent, wenn sie die gleiche Wahrheitswertetabelle haben. Wir schreiben

$$C(P_1, P_2, P_3) \Leftrightarrow D(P_1, P_2, P_3) \ ,$$

um zu sagen, daß die Kombinationen C und D äquivalent sind. Hier sind einige Beispiele von Kombinationen, welche äquivalent sind. Es gilt

$$C(P_1, P_2, P_3) \Leftrightarrow D(P_1, P_2, P_3)$$

genau dann, wenn die Wahrheitswerte der Kombination

$$C(P_1, P_2, P_3) \leftrightarrow D(P_1, P_2, P_3)$$

durchweg wahr sind. Zur Unterscheidung, wir betrachten $C(P_1, P_2, P_3)$ und $D(P_1, P_2, P_3)$ als gleich und schreiben $C(P_1, P_2, P_3) = D(P_1, P_2, P_3)$, wenn die Symbolketten gleich wären.

1. $\sim\sim P \Leftrightarrow P$
2. $\sim (P \wedge Q) \Leftrightarrow \sim P \vee \sim Q$
3. $\sim (P \vee Q) \Leftrightarrow \sim P \wedge \sim Q$
4. $P \to Q \Leftrightarrow \sim (P \wedge \sim Q) \Leftrightarrow \sim P \vee Q$
5. $P \leftrightarrow Q \Leftrightarrow (P \to Q) \wedge (Q \to P)$
6. $P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$
7. $P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)$
8. $f \vee P \Leftrightarrow P$
9. $w \wedge P \Leftrightarrow P$

Offensichtlich gilt

$$C(P_1, P_2, P_3) \Leftrightarrow P_1 \vee \sim P_2 \vee \sim P_3 \ .$$

In der Tat kann man mit den obigen Regeln umformen:

$$\begin{aligned}
C(P_1, P_2, P_3) &\Leftrightarrow (P_1 \wedge (P_1 \to P_2)) \vee \sim (P_2 \wedge P_3) \\
&\Leftrightarrow P_1 \wedge (P_1 \to P_2)) \vee \sim P_2 \vee \sim P_3) \\
&\Leftrightarrow (P_1 \wedge \sim (P_1 \wedge \sim P_2)) \vee \sim P_2 \vee \sim P_3) \\
&\Leftrightarrow (P_1 \wedge (\sim P_1 \vee P_2)) \vee \sim P_2 \vee \sim P_3) \\
&\Leftrightarrow (P_1 \wedge \sim P_1) \vee (P_1 \wedge P_2) \vee \sim P_2 \vee \sim P_3) \\
&\Leftrightarrow (P_1 \vee \sim P_2 \vee \sim P_3) \wedge (P_2 \vee \sim P_2 \vee \sim P_3) \\
&\Leftrightarrow P_1 \vee \sim P_2 \vee \sim P_3
\end{aligned}$$

1.2 Quantoren

Bisher hatten wir Aussagen über eine fixierte natürliche Zahl gebildet. Wir erklären nun, wie man Aussagen über Mengen bildet. Als Beispiel berachten wir wieder die natürlichen Zahlen.

1. $Q_1 :=$ Für alle natürliche Zahlen n gilt $n < 3n - 1$.
2. $Q_2 :=$ Es existiert eine natürliche Zahl, die durch 3 teilbar ist.

Das Grundprinzip ist folgendes: Sei $P_1(n)$ die Aussage $n < 3n-1$ und $P_2(n)$ die Aussage $3|n$ (3 teilt n). Das ist in Symbolen

1. $Q_1 = (\forall n \in \mathbb{N} | P_1(n))$
2. $Q_2 = (\exists n \in \mathbb{N} | P_2(n))$.

Oft werden solche Aussagen abwechslungsreicher formuliert. Q_1 ist äquivalent zu:

1. Sei n eine natürliche Zahl. Dann gilt $n < 3n - 1$.
2. Jede natürliche Zahl n erfüllt $n < 3n - 1$.

Q_2 ist äquivalent zu:

1. Es gibt durch 3 teilbare natürliche Zahlen.

Die Negationen der Aussagen Q_1, Q_2 sind:

1. $\sim Q_1 =$ Nicht für alle natürlichen Zahlen n gilt $n < 3n - 1$.
2. $\sim Q_2 =$ Es existiert keine natürliche Zahl, die durch 3 teilbar ist.

In Symbolen:

1. $\sim Q_1 = \sim (\forall n \in \mathbb{N} | P_1(n))$
2. $\sim Q_2 = \sim (\exists n \in \mathbb{N} | P_2(n))$.

Beachte folgende äquivalente Umformungen:

1. $\sim Q_1 =$ Es existiert eine natürliche Zahl n so daß nicht $n < 3n - 1$ gilt.
2. $\sim Q_2 =$ Alle natürlichen Zahlen sind nicht durch 3 teilbar.

In Symbolen:

1. $\sim (\forall n \in \mathbb{N} | P_1(n)) = (\exists n \in \mathbb{N} | \sim P_1(n))$
2. $\sim (\exists n \in \mathbb{N} | P_2(n)) = (\forall n \in \mathbb{N} | \sim P_2(n))$

Wir betrachten nun Aussagen der Art

Q_1 Für alle natürliche Zahlen n gibt es eine natürliche Zahl m mit $m > n$.

Q_2 Es existiert eine natürliche Zahl n derart, daß für alle natürlichen Zahlen m gilt $n|m$..

In Symbolen:

1. $Q_1 = (\forall n \in \mathbb{N}|(\exists m \in \mathbb{N}|m > n)) = (\forall n \in \mathbb{N}\,\exists m \in \mathbb{N}|m > n)$
2. $Q_2 = (\exists n \in \mathbb{N}|(\forall m \in \mathbb{N}|n|m)) = (\exists n \in \mathbb{N}\,\forall m \in \mathbb{N}|n|m)$

Die Struktur dieser Bildung ist also wie folgt. Sei $P(n) := (\exists m \in \mathbb{N}|m > n)$ die von $n \in \mathbb{N}$ abhängige Aussage Aussage über $\mathbb{N}$: Es gibt eine natürliche Zahl m welche $m > n$ erfüllt. Dann ist

$$Q_1 = (\forall n \in \mathbb{N}|P(n)) \ .$$

Die Regeln für die Negation liefern:

$$\begin{aligned} \sim Q_1 &= \ \sim (\forall n \in \mathbb{N}|(\exists m \in \mathbb{N}|m > n)) \\ &= (\exists n \in \mathbb{N}| \sim (\exists m \in \mathbb{N}|m > n)) \\ &= (\exists n \in \mathbb{N}|(\forall m \in \mathbb{N}|m \leq n)) \\ &= (\exists n \in \mathbb{N}\,\forall m \in \mathbb{N}|m \leq n)) \end{aligned}$$

In Worten ist $\sim Q_1$ die Aussage: Es existiert eine natürliche Zahl n derart, daß für alle natürlichen Zahlen m die Relation $m < n$ gilt.

Oft kommen Aussagen der folgenden Art vor:

1. Es gibt genau eine natürliche Zahl n mit $3|n$ und $n < 5$.
2. Für jede natürliche Zahl n existiert genau eine Zahl m mit $m + 3 = n$.

In Symbolen schreiben wir

1. $(\exists! n \in \mathbb{N}|(3|n) \wedge (n < 5))$
2. $(\forall n \in \mathbb{N}\,\exists! m \in \mathbb{N}|m + 3 = n)$

Die erste Aussage ist wahr, während die zweite falsch ist.

1.3 Beweise

Die Grundlegenden Elemente der Mathematik sind die Bildung von Begriffen, die Formulierung von Aussagen über diese und ihr Beweis. In einem Beweis folgert man die Behauptung aus Aussagen, welche schon beweisen worden sind. Diese werden wieder aus schon vorher bewiesenen Aussagen gefolgert und so fort. Als Anfang stehen Aussagen zur Verfügung, die den Status von Axiomen haben und als wahr angesehen werden.

Der Begriff **folgern** erfordert hier eine Erklärung. Wir machen Aussagen über ganze Zahlen

$$P(n, m) := (0 = n + m) \ , \quad Q(n, m) := (0 = m + n) \ .$$

Wir wollen den folgenden Satz beweisen.

Satz 1.1 *Für je zwei ganze Zahlen n, m folgt aus $(n + m) = 0$ auch $(m + n) = 0$.*

Beweis: Wir müssen zeigen, daß

$$P(n, m) \to Q(n, m)$$

gilt. Als schon bewiesene Aussage nehmen wir das Kommutativitätsgesetz

$$T(n, m) := (n + m = m + n)$$

und das Axiom der Mengenlehre

$$Q(a, b, c) = ((a = b) \wedge (b = c)) \to (a = c) \ ,$$

welches insbesondere für die ganzen Zahlen gilt. Wir haben nun folgende Gleichheiten:

$$\begin{aligned}
&P(n, m) \to Q(n, m) \\
&\quad = \quad P(n, m) \wedge w \to Q(n, m) \\
&\quad = \quad (P(n, m) \wedge T(n, m)) \to Q(n, m) \\
&\quad = \quad ((0 = n + m) \wedge ((n + m = m + n)) \to 0 = n + m \\
&\quad = \quad Q(0, n + m, m + n) \\
&\quad = \quad w
\end{aligned}$$

■

Formal gesehen ist unsere Aussage eigentlich

$$A := (\forall n \in \mathbb{Z} \ , \forall m \in \mathbb{Z} | P(n, m) \to Q(n, m)) \ .$$

Das Kommutativitätsgesetz und die Transitivität der Gleichheit sind die Aussagen

$$w = (\forall n \in \mathbb{Z} \ , \forall m \in \mathbb{Z} | T(n, m)) \ , w = (\forall n \in \mathbb{Z} \ , \forall m \in \mathbb{Z} \ , \forall k \in \mathbb{Z} | Q(n, m, k)) \ ,$$

und man müßte mit diesen Ausdrücken operieren, was noch viel länger ist.

Normalerweise schreibt man den Beweis aber viel kürzer.

Satz 1.2 *Für je zwei ganze Zahlen n, m folgt aus $(n + m) = 0$ auch $(m + n) = 0$.*

Beweis: Wenn $0 = n + m$ für ganze Zahlen n, m erfüllt ist, dann folgt aus der Kommutativität $m + n = n + m$ der Addition auch $0 = m + n$. ■

Dieser kurze Satz gibt eine genügend präzise Beschreibung, wie der formale eigentliche Beweis gebaut werden kann (was man aber praktisch nie durchführt). Was **genügend präzise** ist, hängt vom Umfeld oder dem Adressaten, den man von der Richtigkeit der behaupteten Aussage überzeugen will, ab. Historisch hat sich schon oft herausgestellt, daß eine eigentlich schon akzeptierte Beweisbeschreibung doch nicht ausreichend war und einer Präzisierung bedurfte.

1.4 Mengen

Wir entwickeln hier keine Mengentheorie sondern erklären, wie man mit Mengen umgeht. Wir betrachten ein "Universum mathematischer Objekte", welche wir uns alle als Menge vorstellen, und welches wir nicht weiter erklären wollen und können. Die grundlegende Beziehung zwischen diesen Objekten ist die **Elementrelation**

$$a \in A \quad a \text{ in Element der Menge } A\ .$$

Wir nehmen mal an, daß wir die Menge $\mathbb{N}$ der natürlichen Zahlen und die einzelnen Zahlen $1, 2, \ldots$ in unserem Universum haben. Dann gilt etwa $1 \in \mathbb{N}$, $2 \in \mathbb{N}$, nicht aber $\mathbb{N} \in \mathbb{N}$.

Für eine Menge A können wir eine von x abhängige Aussage $(x \in A)$ über **alle** Objekte x unseres Universums bilden

Axiom 1.3 (Gleichheit von Mengen, Extensionalitätaxiom) *Zwei Mengen A und B sind genau dann gleich, wenn die Äquivalenz der Aussagen*

$$(x \in A) \leftrightarrow (x \in B)$$

gilt.

Wir können also eine **Menge angeben**, indem wir ihre **Elemente aufzählen**. Etwa in der Art: $\{1, 2, 3, 5, 7\}$ oder $\{2i | i \in \mathbb{N}\}$ die Menge der geraden natürlichen Zahlen. Bisher haben wir allerdings keine Information darüber, daß deratige Mengen in unserem Universum existieren, siehe weiter unten.

Eine grundlegende Relation zwischen Mengen ist die Teilmengenrelation:

$$B \subseteq A \quad B \text{ ist Teilmenge von } A\ .$$

Diese Relation gilt, wenn jedes Element von B ein Element von A ist.

Definition 1.4 $(B \subseteq A) := (x \in B) \rightarrow (x \in A)$.

Die Aussage, daß B eine Teilmenge von A, aber nicht gleich A ist, drücken wir in Symbolen durch $B \subset A$ oder $B \subsetneqq A$ aus. Wir sagen dann, daß B eine echte Teilmenge von A ist. Es gilt etwa

1. $\{1, 2, 3, 6, 7\} \subset \mathbb{N}$
2. $\mathbb{N} \subseteq \mathbb{N}$
3. $\{1, 2, 3, 4, a\} \not\subset \mathbb{N}$

Die grundlegende Konstruktion ist die Bildung von Teilmengen in der folgenden Art (wir nehmen an, daß es eine Menge $\mathbb{N}$ der natürlichen Zahlen gibt, siehe weiter unten): Sei M die Teilmenge der durch 2 teilbaren natürlichen Zahlen: In Symbolen:

$$M := \{n \in \mathbb{N} |\ 2|n\}\ .$$

Die Struktur dieser Bildung ist die folgende. Sei A eine Menge und $a \mapsto P(a)$ eine Aussage. Dann wollen wir die Menge

$$B := \{a \in A | P(a)\}$$

bilden.

Definition 1.5 *Die Menge $B := \{a \in A|P(a)\}$ ist durch*

$$(b \in B) \Leftrightarrow (b \in A) \wedge P(b) \ .$$

charakterisiert.

Genaugenommen müssen wir fordern, daß so eine Menge B in unserem Universum existiert.

Axiom 1.6 (Aussonderungsaxiom) *Für jede Aussage $a \mapsto P(a)$ existiert die Menge*

$$B := \{a \in A|P(a)\}$$

mit

$$(b \in B) \Leftrightarrow (b \in A) \wedge P(b) \ .$$

Hier ist ein weiteres Beispiel: Wir betrachten die Aussage

$$P(n,m) := ((m|n \to m = 1 \vee m = n) \wedge (n \neq 1))$$

über natürliche Zahlen n, m. Dann können wir die Aussage

$$Q(n) := \{\forall m \in \mathbb{N} | P(n,m)\}$$

bilden. Die Teilmenge

$$\{n \in \mathbb{N}|Q(n)\}$$

ist die Menge der Primzahlen.

Wir sollten uns jetzt daran erinnern, daß alle unsere Objekte Mengen sind. Damit sind die Elemente eine Menge wiederum Mengen. Ein typisches Beispiel ist

$$\{\{1,2,3,4\}, \{3,4,6,7\}, \mathbb{N}\} \ .$$

Man könnte jetzt fragen, inwiefern 1 eine Menge ist. Was sind deren Elemente? In der Tat haben wir bisher noch nicht erklärt, wie wir die natürlichen Zahlen in unserem Universum realisieren wollen. Eine vorgreifende Antwort ist

$$1 := \{\emptyset\} \ , \quad 2 := \{\emptyset, \{\emptyset\}\}, \dots \ .$$

Sei A irgend eine Menge. Dann bilden wir

$$B := \{x \in A | x \notin x\} \ .$$

Lemma 1.7 *Es gilt $B \notin A$.*

Beweis: Es gilt entweder $B \in B$ oder $B \notin B$. Im ersten Fall ergibt sich aus $B \in A$ ein Widerspruch. Im zweiten Fall ergibt sich aus $B \in A$, daß $B \in B$ gilt, auch ein Widerspruch. Also muß $B \notin A$ gelten. ■

Korollar 1.8 *Für jede Menge gibt es ein Objekt, welches nicht Element der Menge ist. es gibt also keine Menge aller Mengen.*

Mengen, die sich selbst enthalten, sind problematisch. In der axiomatischen Mengenlehre schließt man das durch das Regularitätsaxiom aus.

Axiom 1.9 (Regularität) *Jede nichtleere Menge A enthält ein Element $B \in A$ derart, daß $A \cap B = \emptyset$.*

In der Tat, sei B eine Menge mit $B \in B$. Dann bilden wir $A := \{B\}$. Das einzige Element dieser Menge ist B. Nun gilt aber $B \in A$ und $B \in B$, also $A \cap B \neq \emptyset$. Das Regularitätsaxiom schließt die Existenz von B also aus.

Bisher könnte unser Universum gar keine Mengen enthalten. Deshalb nehmen wir folgendes Axiom an.

Axiom 1.10 *Es gibt eine Menge.*

Sei jetzt f die Aussage, welche immer falsch ist.

Definition 1.11 *Wir definieren die* **leere Menge** $\emptyset$ *durch*

$$(x \in \emptyset) \leftrightarrow f \ .$$

Lemma 1.12 *Die leere Menge existiert..*

Beweis: Nach unserer Annahme gibt es eine Menge A. Wir zeigen, daß

$$\emptyset = \{x \in A | f\} \ .$$

gilt. Die rechte Seite existiert nach dem Aussonderungsaxiom.

$$\begin{aligned} (x \in \emptyset) &\Leftrightarrow f \\ &\Leftrightarrow (x \in A) \wedge f \\ &\Leftrightarrow (x \in \{x \in A | f\}) \end{aligned}$$

Im folgenden beschreiben wir einige weitere Konstruktionsprinzipien von Mengen. Mengen können auch Elemente sein. Insbesondere wollen wir folgendes.

Axiom 1.13 (Paarmengenaxiom) *Zu je zwei Mengen A, B gibt es eine Mengen $\{A, B\}$, welche genau die Mengen A und B als Elemente hat:*

$$(x \in \{A, B\}) \Leftrightarrow (x = A) \vee (x = B) \ .$$

Wir können damit insbesondere auch die Menge $\{A\} := \{A, A\}$ bilden, welche genau A als Element enthält. Beachte, daß $A \in \{A\}$, aber im allgemeinen nicht $A = \{A\}$ gilt. Zum Beispiel ist $\{\emptyset\}$ eine nicht-leere Menge. Diese Konstruktion kann iteriert werden. Allein aus der leeren Menge kann man

$$\{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$$

konstruieren. Desweiteren gibt es die Mengen

$$\{\emptyset, \{\emptyset\}\}\ , \{\emptyset, \{\emptyset, \{\emptyset\}\}\}\ , \{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}\} \ldots\ .$$

Seien A, B Mengen. Dann wollen wir die Vereinigungsmenge bilden:

Definition 1.14 *Die **Vereinigung** $C := A \cup B$ ist durch*

$$(x \in C) \Leftrightarrow (x \in A) \vee (x \in B)$$

charakterisiert.

Zum Beispiel ist

$$\mathbb{N} \cup \mathbb{N} = \mathbb{N}\ , \mathbb{N} \cup \{0\} = \mathbb{N}_0\ , \{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}\ .$$

Ist ist nicht apriori klar, daß eine solche Vereinigungsmenge existiert. Wir sichern dies durch die Annahme des folgenden Axioms.

Axiom 1.15 (Vereinigungsaxiom) *Für jede Menge Z existiert eine Menge X, welche genau die Elemente der Elemente von Z enthält. Sie wird durch*

$$(x \in X) \Leftrightarrow (\exists z \in Z | x \in z)$$

charakterisiert.

Für Mengen A, B können wir $Z := \{A, B\}$ bilden. Das Vereinigungsaxiom auf Z angewendet gibt die Existenz von $A \cup B$. Durch Iteration kann man Vereinigungen von endlich vielen Mengen bilden. das Axiom liefert aber auch die Existenz der Vereinigung einer beliebigen Menge (von Mengen), für welche wir

$$X := \bigcup_{A \in Z} A$$

schreiben.

Seien wiederum A, B Mengen. Dann wollen wir den Durchschnitt $A \cap B$ bilden.

Definition 1.16 *Der **Durchschnitt** $C := A \cap B$ ist durch*

$$(x \in C) \leftrightarrow (x \in A) \wedge (x \in B)$$

charakterisiert.

In der Tat wird die Existenz durch das Aussonderungsaxiom gesichert, da

$$A \cap B = \{x \in A | x \in B\} \ .$$

Zum Beispiel gilt

$$\{1,2,3,4\} \cap \{1,3,6,7\} = \{1,3\} \ , \mathbb{N} \cap \mathbb{N} = \mathbb{N} \ , \{\emptyset, \{\emptyset\}\} \cap \{\emptyset\} = \{\emptyset\} \ .$$

In der Tat können wir sogar den Durchschnitt einer beliebigen, auch unendlichen, Menge X von Mengen bilden.

Definition 1.17 *Wir definieren den Durchschnitt $\bigcap_{A \in X} A$ einer nichtleeren Menge von Mengen X durch*

$$(x \in \bigcap_{A \in X} A) \Leftrightarrow (\forall A \in X | x \in A) \ .$$

Um die Existenz brauchen wir uns keine Sorgen zu machen. Es gibt ja ein Element $B \in X$, und wir haben

$$\bigcap_{A \in X} A = \{b \in B | (\forall A \in X | b \in A)\} \ .$$

Es gelten die folgenden Regeln:

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

Definition 1.18 *Für zwei Mengen A, B definieren wir die Differenz*

$$A \setminus B := \{a \in A | a \notin B\}$$

Es gelten die Regeln

1. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
2. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

Für eine Menge A wollen wir die Menge $P(M)$ aller Teilmengen betrachten.

Definition 1.19 *Die **Potenzmenge** $\mathcal{P}(A)$ einer Menge A wird durch*

$$(X \in \mathcal{P}(A)) \Leftrightarrow (X \subseteq A)$$

charakterisiert.

Es ist wiederum nicht a priori klar, daß die Potenzmenge in unserem Universum existiert. Deshalb fordern wir:

Axiom 1.20 (Potenzmengenaxiom) *Für jede Menge existiert die Potenzmenge.*

Die Potenzmenge von $\{1,2\}$ ist zum Beispiel

$$\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\} \ , \quad \mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\} \ .$$

1.5 Paare, Relationen

Seien a, b zwei Objekte unseres Universums. Unter dem **geordneten Paar** (a, b) verstehen wir ein neues Objekt, dessen erste Komponente das Objekt a und dessen zweite Komponente das Objekt b ist. Zwei geordnete Paare (a, b) und (a', b') sollen dabei genau dann gleich sein, wenn $a = a'$ und $b = b'$ gilt. Eine mögliche Realisierung dieser Idee ist es,

$$(a, b) := \{\{a\}, \{a, b\}\}$$

zu setzen. Beachte, daß $(a, a) = \{\{a\}\}$ gilt. Sei P ein geordnetes Paar. Ein Objekt a ist die erste Komponente von P, wenn $\{a\} \in P$ gilt. Ein Objekt b ist die zweite Komponente, wenn $P \setminus \{\{a\}\} = \{\{a, b\}\}$ für die erste Komponente a von P oder $P \setminus \{\{b\}\} = \emptyset$ gilt.

Wir betrachten zwei Mengen A, B. Mit $A \times B$ bezeichnen wir die Menge der geordneten Paare (a, b) mit $a \in A$, $b \in B$.

Definition 1.21 *Die Menge $A \times B$ ist das* ***kartesische Produkt*** *der Mengen A und B.*

Zum Beispiel ist

$$\{1, 2\} \times \{a, b, c\} = \{(1, a), (2, a), (1, b), (2, b), (1, c), (2, c)\} .$$

Genaugenommen ergibt sich hier die Frage, warum $A \times B$ eine Menge ist. In der Tat können wir jedoch $A \times B$ als eine Teilmenge von $P(P(A \cup B))$ beschreiben. $A \times B$ is die Menge derjenigen Elemente $Q \in P(P(A \cup B))$, für welche es ein $a \in A$ und ein $b \in B$ gibt, so daß $Q = \{\{a\}, \{a, b\}\}$ gilt.

Ein geordnetes n-Tupel von Objekten $(a_1, \ldots, a_n)$ hat wohlbestimmte $1, 2, \ldots, n$-te Komponenten. Eine Realisierung ist als

$$\{\{a_1\}, \{a_1, a_2\}, \ldots, \{a_1, \ldots, a_n\}\}\} .$$

Entsprechend kann man das kartesische Produkt $A_1 \times \cdots \times A_n$ bilden, dessen Elemente die geordneten n-Tupel von Elementen mit $a_i \in A_i$ bilden. Man kann die Bildung des kartesischen Produktes auch iterieren und etwa $A_1 \times (A_2 \times A_3)$ oder $(A_1 \times A_2) \times A_3$ bilden. Diese sind in der Regel verschieden von $A_1 \times A_2 \times A_3$. Wir werden jedoch (natürliche) **Bijektionen** (eineindeutige Abbildungen)

$$A_1 \times (A_2 \times A_3) \xrightarrow{\sim} A_1 \times A_2 \times A_3 \xrightarrow{\sim} (A_1 \times A_2) \times A_3$$

durch

$$\begin{aligned} \{\{a_1\}, \{a_1, \{\{a_2\}, \{a_2, a_3\}\}\}\} &\mapsto \\ \{\{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}\} &\mapsto \\ \{\{\{a_1\}, \{a_1, a_2\}\}, \{\{\{a_1\}, \{a_1, a_2\}\}, \{\{\{a_1\}, \{a_1, a_2\}\}, a_3\}\}\} & \end{aligned}$$

konstruieren, mit welchen wir diese Bildungen vergleichen können.

Eine **binäre Relation** ist eine Aussage über geordnete Paare von Objekten. Hier sind einige Beispiele:

1. $n < m$ (über Paare natürlicher Zahlen (n, m))

2. $n = m^2$ (über Paare natürlicher Zahlen (n, m))

3. $2|(n - m)$ (über Paare natürlicher Zahlen (n, m))

4. $X \subseteq Y$ (über Paare von Mengen (X, Y))

Ist R das Symbol für die Relation (oben also $<$, $=$,$\subseteq$), dann schreibt man meist xRy, kann aber auch $R(x, y)$ verwenden.

Sei P eine Relation auf $A \times B$. Sie bestimmt eine Teilmenge

$$T_P := \{(a, b) \in A \times B | P(a, b)\} \subseteq A \times B \ .$$

Auf der anderen Seite bestimmt eine Teilmenge $T \subseteq A \times B$ eine Relation P_T durch

$$P_T(a, b) := ((a, b) \in T) \ .$$

Es gilt

$$P = P_{T_P} \ , \quad T = T_{P_T} \ .$$

Wir ändern nun die Betrachtungsweise und betrachten folgende Definition als grundlegend.

Definition 1.22 *Eine binäre Relation zwischen Elementen aus den Mengen A und B ist eine Teilmenge von $A \times B$.*

Wir haben damit eine Menge der binären Relationen, nämlich $P(A \times B)$.

Wir erklären im folgenden Eigenschaften einer binären Relation R auf einer Menge A.

Definition 1.23 *1.* **reflexiv** *: Für alle $a \in A$ gilt aRa.*

2. **symmetrisch** *: Für alle $a, b \in A$ gilt $aRb \leftrightarrow bRa$.*

3. **antisymmetrisch** *: Für alle $a, b \in A$ gilt $aRb \wedge bRa \rightarrow b = a$.*

4. **asymmetrisch***: Für alle $a, b \in A$ gilt $aRb \rightarrow \sim (bRa)$*

5. **transitiv** *: Für alle $a, b, c \in A$ gilt $aRb \wedge bRc \rightarrow aRc$*

6. **total** *: Für alle $a, b \in A$ gilt $aRb \vee bRa \vee a = b$.*

Hier sind die Eigenschaften unsere Beispiele

1. $=$: reflexiv, symmetrisch, transitiv, antisymmetrisch

2. $2|(n - m)$ (auf $\mathbb{N}$) : reflexiv, symmetrisch, transitiv

3. $<$, $\subset$: transitiv, asymmetrisch

4. $\leq, \subseteq$: transitiv, reflexiv, antisymmetrisch

Definition 1.24 *Eine* **Äquivalenzrelation** *ist eine* **reflexive, transitive und symmetrische** *Relation.*
Eine **Halbordnungsrelation** *ist eine* **reflexive, transitive und antisymmetrische** *Relation. Ein Halbordnungsrelation ist eine* **Ordnungsrelation**, *wenn sie zusätzlich* **total** *ist.*

1.6 Abbildungen

Ein grundlegender Begriff der Mathematik ist der einer **Abbildung**. Eine Abbildung $f : A \to B$ zwischen ordnet jedem $a \in A$ ein wohlbestimmtes $b \in B$ zu. Die Menge A heißt Definitinsbereich und B wird Wertebereich der Abbildung genannt. Eine Abbildung kann zum Beispiel durch eine Wertetabelle angegeben werden. Sei $A = \{1, 2, 3, 4\}$ und $B = \mathbb{N}$. Dann wird durch

a	1	2	3	4
$f(a)$	3	5	7	9

eine Abbildung $f : A \to B$ bestimmt. Mit f bezeichnen wir die Abbildung, während $f(a)$ den Wert der Abbildung auf dem Element a bezeichnet.

Wenn die Menge A viele Elemente enthält, dann ist die Angabe einer Wertetabelle unpraktisch. In diesem Fall kann man versuchen, die Vorschrift, nach der aus a das Element b gebildet wird zu beschreiben. In unserem Beispiel ist $f(a)$ das um Eins vermehrte Doppelte von a, also $f(a) := 2a + 1$.

Man darf die Abbildungsvorschrift nicht mit der Abbildung selbst verwechseln. Zum Beispiel definiert dieselbe Vorschrift auch eine Abbildung $f' : \mathbb{N} \to \mathbb{N}$. Diese ist von f verschieden, da sie einen anderen Definitionsbereich hat.

Genaugenommen wissen wir aber immer noch nicht, was eine Abbildung ist. Was bedeutet "zuordnen"? Um eine präzise Definition zu erhalten, betrachten wir den Graphen

$$\Gamma_f := \{(a, b) \in A \times B | f(a) = b\} \subseteq A \times B \ .$$

In unserem Beispiel ist $\Gamma_f = \{(1, 3), (2, 5), (3, 9), (4, 9)\}$. Der Graph ist also eine Relation. Er bestimmt die Abbildung durch: Für $a \in A$ ist $f(a)$ das durch $(a, f(a)) \in \Gamma_f$ eindeutig bestimmte Element von B. Nicht jede Relation definiert eine Abbildung. Wir nennen Relationen, welche von Abbildungen kommen, funktional.

Definition 1.25 *Eine Relation $R \subseteq A \times B$ heißt* ***funktional****, wenn für jedes $a \in A$ genau ein $b \in B$ existiert, so daß $(a, b) \in R$ gilt.*

In Symbolen:

$$R \text{ ist funktional, wenn} \qquad (\forall a \in A \exists! b \in B | (a, b) \in R) \ .$$

Wir können nun den Begriff einer Abbildung präzise machen.

Definition 1.26 *Eine Abbildung $A \to B$ von einer Menge A in eine Menge B ist eine funktionale Relation in $A \times B$.*

Im präzisen Sinn der Definition ist eine Abbildung $f : A \to B$ der Graph Γ_f. Insbesondere haben wir eine Menge der Abbildungen von $A \to B$, nämlich die Teilmenge von $P(A \times B)$ der funktionalen Relationen. In der Vorstellung sollte aber der Zuordnungsbegriff den Vorrang haben. Hier sind einige Beispiele:

1. $f(x) := 3x + 1 : \mathbb{N} \to \mathbb{N}$ entspricht $\{(n, m) \in \mathbb{N} \times \mathbb{N} | 3n + 1 = n\}$

2. Sei $b_0 \in B$. Die konstante Abbildung $A \to B$ mit dem Wert b entspricht $\{(a, b) \in A \times B | b = b_0\}$.

3. Die Identität $\mathtt{id}_A : A \to A$ entspricht $\{(a, b) \in A \times A) | a = b\}$.

Die konstante Funktion mit dem Wert $b \in B$ wird oft mit dem Symbol b bezeichnet. Es ergibt sich aus dem Kontext, ob b das Element von b oder eine konstante Funktion ist.

Meist steht das Symbol a kann für ein bestimtes Element der Menge A. Manchmal interpretiert man es aber auch als die identische Abbildung $A \to A$.

Wir werden immer nur Abbildungen zwischen zwei fixierten Mengen A, B vergleichen. Dies vermeidet folgendes Problem. Sei $B \subsetneq C$ und $R \subseteq A \times B$ eine funktionale Relation. Dann ist auch $R \subseteq A \times C$ eine funktionale Relation. Seien $f : A \to B$ und $f' : A \to C$ die entsprechenden Abbildungen. Entsprechend unserer Vereinbarung würde $f = f'$ gelten, was aber nicht erwünscht ist. Der Ausweg wäre hier, eine Abbildung f als ein Tripel $f := (A, R, B) \in \mathcal{P}(A) \times \mathcal{P}(A \times B) \times \mathcal{P}(B)$ zu verstehen. Dann wäre sicher $f = (A, R, B) \neq (A, R, C) = f'$.

Bisher haben wir Teilmengen durch das Aussonderungsaxiom, also in Form

$$\{x \in A | P(x)\}$$

beschrieben. Manchmal werden wir aber Mengen auch durch Aufzählen der Elemente angeben. Genauer, sei $f : A \to B$ eine Abbildung und P. Dann schreiben wir

$$\{f(a) | a \in A\}$$

für die Menge der Bilder unter f. Dies ist eine Schreibweise für die Teilmenge

$$\{b \in B | (\exists a \in A | b = f(a))\}$$

von B.

Abbildungen können komponiert werden. Seien $f : A \to B$ und $g : B \to C$ Abbildungen. Dann kann man eine neue Abbildung

$$h := g \circ f : A \to C$$

definieren, welche jedem Element $a \in A$ das Element $h(a) := g(f(a))$ zuordnet. Manchmal schreiben wir auch $h = g(f)$. Für eine präzise Definition der Komposition müssen wir den Graphen von h beschreiben.

Zu diesem Zweck beschreiben wir etwas allgemeiner eine Komposition von Relationen. $F \subseteq A \times B$ und $G \subseteq B \times C$.

Definition 1.27 *Wir definieren die* ***Komposition***

$$G \circ F := \{(a,c) \in A \times C | (\exists b \in B | (a,b) \in F \wedge (b,c) \in G)\} \ .$$

Ist $H \subseteq C \times D$ eine dritte Relation, dann gilt

$$H \circ (G \circ F) = (H \circ G) \circ F \ .$$

Lemma 1.28 *Wenn F und G funktional sind, dann ist $F \circ G$ funktional.*

Beweis: Sei $a \in A$. Wir müssen zeigen, daß es genau ein $c \in C$ mit $(a,c) \in F \circ G$ gibt. Da F funktional ist, gibt es genau ein $b \in B$ mit $(a,b) \in B$. Da G funktional ist, gibt es für dieses Element genau ein $c \in C$ mit $(b,c) \in G$. Damit gilt genau für dieses Element $(a,c) \in G \circ F$. ■

Definition 1.29 *Die* ***Komposition*** $h := g \circ f$ *wird durch*

$$\Gamma_h := \Gamma_g \circ \Gamma_f$$

definiert.

Manchmal ist es günstig, eine Abbildung $f : A \to B$ durch $f(a)$ zu bezeichnen. In diesem Fall interpretiert man a als identische Abbildung auf A. Wir betrachten zum Beispiel $f, g : \mathbb{N} \to \mathbb{N}$, $f(x) := (1 + x^2)$ und $g(x) := 1 + x$. Die Gleichung

$$f(x) = g(x^2)$$

beschreibt eine Relation zwischen Abbildungen, nämlich daß f und die Komposition aus g und Quadrieren gleich sind. Man könnte natürlich auch die Abbildung $q : \mathbb{N} \to \mathbb{N}$ einführen und

$$f = g \circ q$$

schreiben. Eine andere Möglichkeit ist die Aussage

$$(\forall x \in \mathbb{N} | f(x) = g(x^2)) \ .$$

Die folgenden beiden Eigenschaften einer Abbildung sind grundlegend.

Definition 1.30 *Sei $f : A \to B$ eine Abbildung.*

1. *f heißt* **injektiv**, *wenn für alle $x, y \in A$ aus $f(x) = f(y)$ die Relation $x = y$ folgt.*

2. *f heißt* **surjektiv**, *wenn für alle $b \in B$ ein $a \in A$ mit $f(a) = b$ existiert.*

3. *f heißt* **bijektiv**, *wenn sie injektiv und surjektiv ist.*

Für eine injektive Abbildung wird jedes Element von B höchstens einmal als Wert realisiert. Eine surjektive Abbildung realisiert jedes Element von B als Wert. Eine bijektive Abbildung realisiert jedes Element von B auf genau eine Weise als Wert. Es gilt folgende einfache Beobachtung.

Lemma 1.31 *Die Komposition zweier injektiver (surjektiver, bijektiver) Abbildungen ist wieder injektiv (surjektiv, bijektiv).*

Die identische Abbildung $\mathtt{id}_A$ is bijektiv. Die Abbildung $f : \{1,2,3,4\} \to \{1,2,3\}$

a	1	2	3	4
$f(a)$	2	3	1	3

ist surjektiv, aber nicht injektiv. Die Abbildung $g : \{1,2,3\} \to \{1,2,3,4\}$

a	1	2	3
$g(a)$	2	3	4

ist injektiv, aber nicht surjektiv.

Für eine Teilmenge $F \subseteq A \times B$ definieren wir die Transposition

$$F^{op} := \{(b,a) \in B \times A | (a,b) \in F\} \ .$$

Wir überzeugen uns von folgender Aussage:

Lemma 1.32 *Ist $f : A \to B$ eine bijektive Abbildung, dann ist die Relation Γ_f^{op} funktional.*

Definition 1.33 *Sei $f : A \to B$ bijektiv. Die* **Umkehrabbildung** *(inverse Abbildung) von f ist die Abbildung $f^{-1} : B \to A$ mit $\Gamma_{f^{-1}} := \Gamma_f^{op}$.*

Wir betrachten die bijektive Abbildung $f : \{1,2,3,4\} \to \{1,2,3,4\}$

a	1	2	3	4
$f(a)$	2	4	3	1

.

Die Umkehrabbildung ist

a	1	2	3	4
$f^{-1}(a)$	4	1	3	2

.

Die inverse Abbildung erfüllt

$$f \circ f^{-1} = \mathtt{id}_B \ , \quad f^{-1} \circ f = \mathtt{id}_A \ .$$

Ist $g : B \to C$ auch bijektiv, dann gilt

$$(g \circ f)^{-1} = g^{-1} \circ f^{-1} \ .$$

Dies folgt aus der Regel

$$(G \circ F)^{op} = F^{op} \circ G^{op} \ .$$

Aus der Bijektivität von f folgt die Bijektivität von f^{-1}. Es gilt

$$(f^{-1})^{-1} = f \ .$$

Sei $f : A \to B$ eine Abbildung.

Definition 1.34 *Seien $X \subseteq A$ und $Y \subseteq B$. Wir setzen*

$$f(X) := \{b \in B | (\exists a \in X | b = f(a))\} \subseteq B$$

und

$$f^{-1}(Y) := \{a \in A | f(a) \in Y\} \subseteq A \ .$$

Insbesondere ist $\mathtt{im}(f) := f(A)$ *das* **Bild** *von f.*

$f(X)$ ist die Menge der Bilder von Elementen aus X, und $f^{-1}(Y)$ ist die Menge der Elemente von A, welche nach Y abgebildet werden. Die Operation f^{-1} ist mit den mengentheoretischen Operationen verträglich.

Lemma 1.35 *Für Teilmengen $Y, Z \subseteq B$ gelten folgende leicht nachprüfbare Aussagen.*

$$\begin{aligned} f^{-1}(Y \cap Z) &= f^{-1}(Y) \cap f^{-1}(Z) \\ f^{-1}(Y \cup Z) &= f^{-1}(Y) \cup f^{-1}(Z) \\ f^{-1}(Y \setminus Z) &= f^{-1}(Y) \setminus f^{-1}(Z) \end{aligned}$$

In der anderen Richtung gilt:

Lemma 1.36 *Für Teilmengen $X, Y \subseteq A$ gilt:*

$$\begin{aligned} f(X \cap Y) &\subseteq f(X) \cap f(Y) \\ f(X \cup Y) &= f(X) \cup f(Y) \\ f(X) \setminus f(Y) &\subseteq f(X \setminus Y) \end{aligned}$$

Für injektive Abbildungen gilt in der ersten und dritten Relation die Gleichheit.

Sei I eine eine Menge.

Definition 1.37 *Eine durch I indizierte* **Familie** *(von Elementen einer Menge U) ist eine Abbildung $x : I \to U$.*

Wir schreiben Familien oft in der Form $(x_i)_{i \in I}$. oder noch kürzer, x_i, wenn die Indexmenge klar ist.

Definition 1.38 *Eine* **Folge** *in U ist eine Familie mit der Indexmenge $\mathbb{N}$ der natürlichen Zahlen.*

Definition 1.39 *Sei $(X_i)_{i\in I}$ eine Familie von Mengen. Wir definieren die Vereinigung $\bigcup_{i\in I} X_i$ und den Durchschnitt $\bigcap_{i\in I} X_i$ der Familie durch*

$$\begin{aligned}(x \in \bigcup_{i\in I} X_i) &:= (\exists i \in I | x \in X_i)\\ (x \in \bigcap_{i\in I} X_i) &:= (\forall i \in I | x \in X_i)\end{aligned}$$

Um zu sehen, daß die Vereinigung existiert, bemerken wir zuerst, daß eine Menge $V := \cup_{X\in U} X$ nach dem Vereinigungsaxiom existiert. Wir können die Vereinigung der Familie nun als Teilmenge von V beschreiben.

Definition 1.40 *Das* **kartesische Produkt** $\times_{i\in I} X_i$ *ist die Menge aller Familien $(x_i)_{i\in I}$ mit $x_i \in X_i$ für alle $i \in I$.*

Ist $I = \{1, 2\}$, so kann man $\times_{i\in I} X_i$ mit dem Produkt $X_1 \times X_2$ identifizieren. Die Familie $(x_i)_{i\in I}$ wird dabei mit dem Paar (x_1, x_2) identifiziert. Analog kann man das Produkt eine durch $\{1, 2, \ldots, n\}$ indizierten Familie als die Menge der geordneten n-Tupel betrachten. Insbesondere ist das Produkt einer endlichen Familie von Mengen genau dann leer, wenn einer der Faktoren leer ist.

Für allgemeine Familien ist dieser Fakt bisher nicht klar. In der Tat brauchen wir ein weiteres Axiom.

Axiom 1.41 (Auswahlaxiom) *Das kartesische Produkt einer nichtleeren Familie von nichtleeren Mengen ist nichtleer.*

1.7 Unendlichkeit

Verabredungsgemäß hat die leere Menge keine Elemente. Wir kennen schon die Mengen mit genau einem Element, nämlich die Einermengen $\{a\}$ für Objekte a unseres Universums. Insbesondere ist $1 := \{\emptyset\}$ eine Einermenge. Wir wollen nun die natürlichen Zahlen und den Begriff einer n-elementigen Menge definieren.

Wir definieren dazu den Nachfolger einer Menge X durch $X^+ := X \cup \{X\}$. Für den Beginn brauchen wir nur die leere Menge.

Definition 1.42 *Wir definieren*

$$\begin{aligned}1 &:= \emptyset^+ = \{\emptyset\}\\ 2 &:= 1^+ = \{\emptyset, \{\emptyset\}\}\\ 3 &:= 2^+ = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\\ &\vdots\\ n+1 &:= n^+\\ &\vdots\end{aligned}$$

Wir haben damit die einzelnen natürlichen Zahlen definiert, nicht aber die Menge $\mathbb{N}$ der natürlichen Zahlen. In der Tat brauchen wir ein weiteres Axiom, um die Existenz von $\mathbb{N}$ zu sichern-

Definition 1.43 *Eine Nachfolgermenge ist eine Menge, welche $\emptyset$ und mit jedem Element auch dessen Nachfolger enthält.*

Axiom 1.44 (Unendlichkeitsaxiom) *Es gibt eine Nachfolgermenge.*

Wie man sich nun leicht denken kann, sichert dies die Existenz der Menge der natürlichen Zahlen. Sei A eine Nachfolgermenge. Dann ist $\mathbb{N}$ die minimale Nachfolgermenge in A. Genauer ist

$$\mathbb{N} = \bigcap_{X \subseteq A | X \text{ ist Nachfolgermenge}} X \ .$$

Diese Definition hängt nicht von der speziellen Wahl von A ab. Durch Fortsetzung dieser Gedanken kann man die Arithetik der natürlichen Zahlen aufbauen.

Wir überzeugen uns jetzt, daß die natürlichen Zahlen alle paarweise verschieden sind. Dazu beweisen wir eine nützliche stärkere Aussage.

Lemma 1.45 *Ist $f : n \xrightarrow{\sim} m$ eine Bijektion, dann gilt $n = m$.*

Beweis: Wir schreiben $n = (n-1)^+$ und $m = (m-1)^+$. Sei $x := f(n-1) \in m$ und $y := f^{-1}(m-1) \in n$. Wir erhalten eine Bijektion $f' : n-1 \to m-1$ durch

$$f'(u) := \begin{cases} f(u) & u \neq y \\ x & u = y \end{cases}$$

Durch Iteration erhalten wir Bijektionen

$$f^{(k)} : n-k \xrightarrow{\sim} m-k \ .$$

Nach endlich vielen Schritten erhalten wir (falls $n \leq m$) eine Bijektion $\emptyset \xrightarrow{\sim} m-n$. Daraus folgt $m - n = \emptyset$, also $m = n$. ■

Definition 1.46 *Eine Menge X hat n* ***Elemente****, wenn es eine Bijektion $X \xrightarrow{\sim} n$ gibt. Wir schreiben in diesem Fall $\sharp X := n$ für die* ***Anzahl der Elemente*** *von X.*

Wir müssen uns von der Wohldefiniertheit der Anzahl überzeugen.

Lemma 1.47 *Wenn eine Menge X gleichzeitig m und n Elemente hat, dann gilt $n = m$.*

Beweis: Seien $f : X \xrightarrow{\sim} n$ und $g : X \xrightarrow{\sim} m$ Bijektionen. Dann ist die Komposition $f \circ g^{-1} : m \to n$ eine Bijektion, folglich $n = m$ nach Lemma 1.45. ■

Ein grundlegendes Ziel der Mengentheorie ist eine präzises Beschreibung des Unendlichen.

Definition 1.48 *Eine Menge A ist* **unendlich***, wenn es eine injektive Abbildung $f : A \to A$ mit $f(A) \neq A$ gibt. Die Menge A heißt* **endlich** *wenn sie nicht unendlich ist.*

Lemma 1.49 *1. Sei $B \subseteq A$. Wenn B unendlich ist, dann auch A.*

2. Sei $f : B \to A$ injektiv. Wenn B unendlich ist, dann auch A.

Beweis: Wir zeigen nur die erste Aussage. Sei $f : B \to B$ injektiv mit $f(B) \neq B$. Wir definieren $g : A \to A$ durch

$$g(a) := \begin{cases} f(a) & a \in B \\ a & a \in A \setminus B \end{cases} .$$

Dann ist f injektiv und $f(A) \neq A$, also A unendlich. ■
Aus Lemma 1.49 folgt: Wenn A endlich ist, dann auch B

Lemma 1.50 *Die Menge der natürlichen Zahlen $\mathbb{N}$ ist unendlich.*

Beweis: Die Abbildung $n \mapsto n+1$ von $\mathbb{N}$ nach $\mathbb{N}$ is injektiv, aber nicht surjektiv.. ■

Lemma 1.51 *1. Wenn A eine endliche Menge ist, dann gibt es ein $n \in \mathbb{N}$, so daß A eine Menge mit n Elementen ist.*

2. Ist A unendlich, dann gibt es eine injektive Abbildung $\mathbb{N} \to A$.

Beweis: Wir zeigen erst 2. Sei A unendlich und $f : A \to A$ eine injektive Abbildung mit $f(A) \neq A$. Wir wählen $a \in A \setminus f(A)$ und definieren $g : \mathbb{N} \to A$ durch $g(n) := f^n(a)$. Diese Abbildung ist injektiv. Wäre $m > n$ und $f^m(a) = f^n(a)$, dann gilt $f^{m-n}(a) = a$, also $a \in f(A)$, was nicht sein kann.

Wenn A nicht n Elemente für eine natürlich Zahl n hat, dann konstruieren wir eine injektive Abbildung $g : \mathbb{N} \to A$ induktiv wie folgt. Sei $g(k)$ für $k < m$ schon erklärt. Die Menge $A \setminus \{g(1), \ldots, g(m-1)\}$ ist nicht leer, weil A sonst $m-1$ Elemente hätte. Wir definieren $g(m)$ als ein Element dieser Differenz. Wir schließen, daß A unendlich ist. ■

Unter den unendlichen Mengen gibt es sehr verschieden große. Für den Vergleich führen wir die Relation "mächtiger" ein.

Definition 1.52 *Die Menge Y ist* **mächtiger** *als X, wenn es eine injektive Abbildung $X \to Y$ gibt. Wir schreiben $X \preceq Y$. Die Menge X ist zu Y* **gleichmächtig***, wenn es eine Bijektion zwischen X und Y gibt. Wir schreiben $X \sim Y$.*

Wir schreiben $X \prec Y$, wenn $X \preceq Y$, aber nicht $X \sim Y$ gilt.

Die Relation $\preceq$ ist offensichtlich reflexiv und transitiv, während $\sim$ eine Äquivalenzrelation ist. Aus $X \sim Y$ folgt $X \preceq Y$ und $Y \preceq X$. Es gilt das Lemma:

Satz 1.53 (Schröder-Bernstein) *Aus $X \preceq Y$ und $Y \preceq X$ folgt $X \sim Y$.*

Siehe [Hal76]. ■

Satz 1.54 (Cantor) *Für jede Menge X gilt $X \prec \mathcal{P}(X)$.*

Beweis: Durch $X \ni x \mapsto \{x\} \in \mathcal{P}(X)$ konstruieren wir eine injektive Abbildung $X \to \mathcal{P}(X)$. Also gilt $X \preceq \mathcal{P}(X)$. Wäre nun $X \sim \mathcal{P}(X)$, dann gäbe es eine Bijektion $f : X \to \mathcal{P}(X)$. Wir definieren

$$A := \{x \in X | x \notin f(x)\} .$$

Dann ist $A \in \mathcal{P}(X)$ und deshalb $A = f(a)$ für ein eindeutig bestimmtes $a \in A$. Es gilt entweder $a \in A$ oder $a \notin A$. im ersten Fall wäre $a \in f(a)$, also $a \notin A$, ein Widerspruch. Im zweiten Fall wäre $a \in A$, was wiederum ein Widerspruch ist. ■

Die kleinsten unendlichen Mengen sind die abzählbaren.

Definition 1.55 *Eine Menge X heißt* **abzählbar** *(unendlich), wenn $X \preceq \mathbb{N}$ ($X \sim \mathbb{N}$) gilt. Wenn $\mathbb{N} \prec X$ gilt, dann heißt X* **überabzählbar**

Es gibt also nichtabzählbare Mengen.

Korollar 1.56 *Die Potenzmenge der natürlichen Zahlen ist nicht abzählbar.*

2 Die rellen Zahlen

2.1 Die rellen Zahlen - Motivation

Die rationalen Zahlen $\mathbb{Q}$ bilden einen Körper, der zusätzlich geordnet ist.

Definition 2.1 *Ein Körper K heißt geordnet, wenn auf K eine Ordnungsrelation $\leq$ definiert ist, die die folgenden Eigenschaften hat:*

1. *$\leq$ ist mit der Addition verträglich: Aus $x < y$ folgt $x + z < y + z$.*

2. *$\leq$ ist mit der Multiplikation verträglich: Aus $0 < x$ und $0 < y$ folgt $0 < xy$.*

Aus diesen Eigenschaften können wir folgende einfache Folgerungen ziehen:

1. $x \leq y \Leftrightarrow -y \leq -x$: In der Tat folgt aus $x \leq y$, daß $x+(-x)+(-y) \leq x+(-x)+(-y)$, also $-y \leq -x$. Die andere Richtung zeigt man ähnlich.

2. Aus $x \geq 0$ und $y \geq z$ folgt $xy \geq xz$. In der Tat ist $y - z \geq 0$ und deshalb $x(y - z) \geq 0$, also $xy \geq xz$ (addiere xz). Umgekehrt, wenn $x \leq 0$, dann folgt aus $y \geq z$ daß $xz \leq xy$. Dann ist nämlich $-x \geq 0$. Aus $y \geq z$ folgt in diesem Fall, daß $-xy \leq -xz$, also $xz \leq xy$.

3. Es gilt $x^2 \geq 0$. In der Tat ist entweder $x \geq 0$ oder $-x \geq 0$, woraus $x^2 \geq 0$ oder $x^2 = (-x)^2 \geq 0$ folgt.

Für viele Probleme sind die rationalen Zahlen nicht ausreichend.

Lemma 2.2 *Es gibt keine rationale Zahl $q \in \mathbb{Q}$ mit $q^2 = 2$.*

Beweis: Wir nehmen an, daß es eine solche Zahl q gäbe. Sei $q = \frac{a}{b}$ für teilerfremde $a, b \in \mathbb{Z}$. Dann ist $q^2 = \frac{a^2}{b^2} = 2$, also $a^2 = 2b^2$. Wir benutzen die eindeutige Primfaktorzerlegung in $\mathbb{Z}$. Aus $2 \not| a$ würde $2 \not| a^2$ folgen, was dieser Gleichung widerspricht. Also gilt $2|a$ und damit $2^2|a^2$. Daraus würde aber $2|b^2$, also $2|b$ folgen. Das widerspricht der Teilerfremdheit von a, b. ■

Die Beschreibung von Wachstumsvorgängen, der Entladung von Kondensatoren oder des Abflusses von Stausseen führt oft auf Differentialgleichungen der Art

$$f' = f \ , \quad f(0) = 1 \ .$$

Die Veränderungsgeschwindigkeit der Größe $f(t)$ ist gleich der Größe $f(t)$ zur Zeit t selbst. Die Lösung dieser Gleichung ist $f(t) = \mathsf{e}^t$. Um diese Funktion zu bestimmen, kann man eine Potenzreihe ansetzen. Das folgende ist im Augenblick nur eine motivierende Betrachtung ohne präziesen mathematischen Gehalt. Wir setzen

$$f(t) = \sum_{n=0}^{\infty} a_n t^n \ .$$

Durch Einsetzen in die Differentialgleichung und Koeffizientenvergleich ergibt sich die Rekursion

$$\sum_{n=0}^{\infty} n a_n t^{n-1} = \sum_{n=0}^{\infty} a_n t^n \ ,$$

also

$$a_{n+1} = \frac{1}{n+1} a_n \ .$$

Mit $a_0 := 1$ erhalten wir $a_n = \frac{1}{n!}$, wobei $n! := n(n-1)\ldots 2$ ist. Damit gilt

$$\mathsf{e}^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} \ .$$

Ab jetzt werden wir wieder präziese. Genaugenommen ist e^t diejenige Zahl, für die folgendes gilt. *Für jede rationale Zahl $\epsilon > 0$ gibt es ein $N_0 \in \mathbb{N}$ derart, daß $|\mathsf{e}^t - \sum_{n=0}^{N} \frac{t^n}{n!}| < \epsilon$ für alle natürlichen Zahlen $N \geq N_0$ gilt.* Nun stellt sich leider heraus, daß es die Zahl e^t für rationale t in der Regel nicht gibt.

Lemma 2.3 e^1 *ist keine rationale Zahl.*

Beweis: Die Behauptung ist, daß keine rationale Zahl die e^1 definierende Eigenschaft besitzt. Wir nehmen das Gegenteil an. Wir zeigen zunächst, daß $2 < \mathsf{e} < 3$.

1. Wir wählen $3 \leq N_0$ für $\epsilon := \frac{1}{2}$. Dann gilt

$$\mathsf{e} \geq 2 + \frac{1}{2} + \sum_{n=3}^{N_0} \frac{1}{n!} - \epsilon > 2 \ .$$

2. Es gilt für $n \geq 2$, daß $n! \geq 2^{n-1}$ und folglich für $N \geq 4$

$$\sum_{n=0}^{N} \frac{1}{n!} < 1 + 1 + \frac{1}{2} + \frac{1}{6} + \sum_{i=4}^{N} \frac{1}{2^{i-1}} = 2 + \frac{1}{2} + \frac{1}{6} + \frac{1}{8}\frac{1-\frac{1}{2}^{N-3}}{1-\frac{1}{2}} \leq 2 + \frac{11}{12} \ .$$

   Wir wählen $4 \leq N_0$ für $\epsilon := \frac{1}{13}$. Dann gilt

$$\mathsf{e} \leq \sum_{n=0}^{N_0} \frac{1}{n!} + \frac{1}{13} \leq 2 + \frac{11}{12} + \frac{1}{13} < 3 \ .$$

Damit ist e nicht ganz.

Sei nun $\mathsf{e} := \mathsf{e}^1 = \frac{a}{b}$ für positive ganze Zahlen a, b mit $b > 1$. Wir setzen

$$x := b!(\mathsf{e} - \sum_{n=0}^{b} \frac{1}{n!})$$

und sehen, daß $x > 0$ gilt. Dann ist x eine natürliche Zahl, da der Faktor $b!$ gegen alle auftretenden Nenner gekürzt werden kann. Sei nun $N \in \mathbb{N}$ derart, daß $|\mathsf{e} - \sum_{n=0}^{N} \frac{1}{n!}| < \frac{1}{2bb!}$ gilt.

Dann ist

$$|x - \sum_{n=b+1}^{N} \frac{b!}{n!}| = |b!\mathsf{e} - \sum_{n=0}^{b} \frac{b!}{n!} - \sum_{n=b+1}^{N} \frac{b!}{n!}| = |b!\mathsf{e} - \sum_{n=0}^{N} \frac{b!}{n!}| < \frac{1}{2b} \leq \frac{1}{4} \ .$$

Für $n \geq b+1$ gilt

$$\frac{b!}{n!} = \frac{1}{n(n-1)\dots(b+1)} \leq \frac{1}{(b+1)^{n-b}}$$

und damit

$$\sum_{n=b+1}^{N} \frac{b!}{n!} \leq \sum_{n=b+1}^{N} \frac{1}{(b+1)^{n-b}} \leq \sum_{n=1}^{N-b} \frac{1}{(b+1)^n} = \frac{1}{b+1}\frac{1-\frac{1}{(b+1)^{N+1-b}}}{1-\frac{1}{b+1}} \leq \frac{1}{b} \leq \frac{1}{2} \ .$$

Wir schließen, daß $x - \frac{1}{2} < \frac{1}{4}$. Diese Ungleichung ist aber für keine natürliche Zahl erfüllt. ■

Wir betrachten diese Beispiele nun genauer.

Sei p eine positive rationale Zahl. Dann können wir die folgende Teilmenge

$$U := \{q \in \mathbb{Q} | q^2 < p \vee q \leq 0\}$$

betrachten. Aus $q \in U$ folgt $q < p+1$. In der Tat, wäre $q \geq p+1$, dann würde

$$q^2 \geq (p+1)q \geq (p+1)^2 = p^2 + 2p + 1 > p \ ,$$

also $q \notin U$ gelten. Insbesondere ist U durch $p+1$ beschränkt und nichtleer, da etwa $0 \in U$ ist. Wenn $q \in U$ und $r \leq q$ ist, dann gilt auch $r \in Q$. In der Tat gilt für $r \in U$ für $r \leq 0$ sowieso, und $r^2 \leq rq \leq q^2 < p$ für $0 \leq r \leq q$.

Wenn es eine Zahl $r \in \mathbb{Q}$ mit $r^2 = p$ gibt, dann gilt $U = \{q \in Q | q < r\}$. Wie der Fall $p = 2$ zeigt, gibt es jedoch nicht immer eine rationale Zahl q, mit welcher man U in dieser Form schreiben kann. Wir werden diese Zahl jedoch in einer Erweiterung der rationalen Zahlen finden.

Im Fall der Zahl e setzen wir

$$U := \{q \in \mathbb{Q} | (\exists N \in \mathbb{N} | q < \sum_{n=0}^{N} \frac{1}{n!}\} \ .$$

Wie wir bereits gesehen haben, gilt

$$\sum_{n=0}^{N} \frac{1}{n!} < 3$$

für alle N. Damit folgt aus $q \in U$ daß $q < 3$. Die Menge U ist wiederum beschränkt und nichtleer, da $0 \in U$. Wenn $q \in U$ und $r \leq q$ ist, dann gilt auch $r \in Q$. Mit Hilfe der hypothetischen Zahl e könnten wir

$$U = \{q \in \mathbb{Q} | q < \mathsf{e}\}$$

schreiben.

Die Mengen U in der obigen Diskussion sind sogenannte Schnitte von $\mathbb{Q}$.

Definition 2.4 *Ein* **Schnitt** *von $\mathbb{Q}$ ist eine nichtleere Teilmenge $U \subset \mathbb{Q}$ mit folgenden Eigenschaft:*

1. *U ist von oben beschränkt, d.h. es gibt eine rationale Zahl r, so daß aus $q \in U$ die Ungleichung $q < r$ folgt.*
2. *Ist $q \in U$ und $r < q$, dann gilt $r \in U$.*
3. *Für jedes $q \in U$ existiert ein $r \in U$ mit $q < r$.*

Definition 2.5 *Eine* **reelle Zahl** *ist ein Schnitt von $\mathbb{Q}$. Wir bezeichnen die Menge der reellen Zahlen mit $\mathbb{R}$.*

Im nächsten Kapitel beschreiben wir die reellen Zahlen axiomatisch. Mit Hilfe der Schnitte kann man zeigen, daß die Axiome realisiert werden können. Für die Details der Definition der arithmetischen Operationen und des Nachweises der Axiome verweisen wir auf [Rud80], Anhang zu Kap. 1.

2.2 Die Supremumseigenschaft und das Archimedische Axiom - Definition von $\mathbb{R}$

Wir betrachten eine Menge M mit einer Ordnungsrelation $\leq$. Wir sagen, daß M eine geordnete Menge ist. Zur Erinnerung (1.24): Eine Ordnungsrelation ist eine reflexive, transitive und antisymmetrische Relation. Beispiele für geordnete Mengen sind $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$. Die

Definition 2.6 *Eine Teilmenge $U \subseteq M$ heißt (von oben)* **beschränkt**, *wenn es ein $m \in M$ mit gibt, so daß $u \leq m$ für alle $u \in U$ gilt. Ein solches Element m heißt* **obere Schranke** *von U.*

Ist U von oben beschränkt und $V \subseteq U$, dann ist auch V von oben beschränkt und jede obere Schranke von U ist auch einen von V.

Sei $U \subseteq M$ eine beschränkte Teilmenge.

Definition 2.7 *Ein Element $m \in M$ heißt kleinste obere Schranke, wenn für jede obere Schranke n von U die Relation $m \leq n$ gilt. Wir schreiben $\sup U := m$ und nennen dieses Element das Supremum von U.*

Auf analoge Weise definieren wir die Begriffe "von unten beschränkt", ütere Schranke" und größte untere Schranke, welche wir mit $\inf U$ bezeichnen und das Infimum von U nennen.

Hier sind einige Beispiele (in $\mathbb{Q}$)

1. $\sup\{\frac{1}{n}|n \in \mathbb{N}\} = 1$, $\inf\{\frac{1}{n}|n \in \mathbb{N}\} = 0$
2. $\sup\{q \in \mathbb{Q}|q^2 \leq 4\} = 2$
3. $\{q \in \mathbb{Q}|q^2 \leq 2\} \subset \mathbb{Q}$ besitzt obere Schranken, zum Beispiel 3, aber kein Supremum.

Beobachte, daß in einigen Fällen $\sup U \in U$ gilt, in anderen Fällen jedoch nicht.

Definition 2.8 *Wenn $\sup U \in U$, dann ist $\sup U$ das* **Maximum** *von U welches wir auch mit $\mathtt{max}U$ bezeichnen. Analog definieren wir das* **Minimum** $\mathtt{min}U$ *von U.*

Eine endliche Teilmenge einer geordneten Menge hat immer ein Maximum und ein Minimum.

Seien $V \subseteq U \subseteq M$ Teilmengen, für welche die Suprema exitieren. Dann gilt

$$\sup V \leq \sup U \ .$$

In der Tat ist jede obere Schranke von U auch eine von V.

Seien $(M, \geq)$ und $(N, \geq)$ geordnete Mengen.

Definition 2.9 *Eine Abildung $f : M \to N$ zwischen geordneten Mengen heißt* ***monoton*** *(wachsend), wenn aus $x, y \in M$ und $x \leq y$ folgt $f(x) \leq f(y)$.*

Analog definiert man monoton fallende Abbildungen. Sei jetzt $X \subset M$ derart, daß $\sup X$ und $\sup f(X)$ existiert. Dann gilt

$$\sup f(X) \leq f(\sup X) \ .$$

In der Tat folgt aus $x \in X$, daß $x \leq \sup X$ und deshalb $f(x) \leq f(\sup X)$. Damit ist $f(\sup X)$ eine obere Schranke von $f(X)$, allerdings möglicherweise nicht die kleineste. Sei etwa $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

gegeben. Dann ist $\sup(-\infty, 0) = 0$, $\sup f((-\infty, 0)) = 0$ und $f(\sup(-\infty, 0)) = 1$.

Definition 2.10 *Eine geordnete Menge $(M, \leq)$ hat die Supremumseigenschaft, wenn für jede nichtleere und von oben beschränkte Teilmenge $U \subseteq M$ ein Supremum in M existiert.*

Definition 2.11 *Ein geordneter Körper, welcher die Supremumseigenschaft hat, ist ein Körper reeller Zahlen.*

Sei K ein geordneter Körper (2.1).

Definition 2.12 *Der Körper ist* **archimedisch** *geordnet, wenn für je zwei Elemente $x, y \in K$ mit $x > 0$ ein $n \in \mathbb{N}$ existiert, so daß $nx > y$ ist.*

Beachte, daß hier nx eine Abkürzung für die Summe $x + x + \cdots + x$ von n Summanden ist.

Die rationalen Zahlen sind archimedisch. In der Tat, sei $x = \frac{a}{b}$ mit $a, b > 0$ und $y = \frac{c}{d}$ mit $d > 0$. Dann gilt $nx > y$ genau dann, wenn $nad > bc$. Nun ist $ad \geq 1$. Wenn wir also $n > bc$ wählen, dann gilt diese Ungleichung.

Lemma 2.13 *Ein Körper reeller Zahen enthält eine Kopie der natürlichen Zahen und ist archimedisch geordnet.*

Beweis: Sei $\mathbb{R}$ ein Körper reller Zahlen. Wir definieren eine Abbildung $i : \mathbb{N} \to \mathbb{R}$ durch

$$n \mapsto i(n) := \underbrace{1 + \cdots + 1}_{n\text{-Summanden}} \ .$$

Es gilt $1 \neq 0$ und deshalb $0 < 1$ oder $1 < 0$. In der Tat kann man den zweiten Fall ausschließen, da er $0 < -1$ und damit $0 = 0^2 < (-1)^2 < 1$ implizieren würde. daraus folgt induktiv $i(n) < i(n) + 1 = i(n + 1)$. Wäre nun $n < m$ und $i(n) = i(m)$, dann würde $i(n) < i(n + 1) < \cdots < i(n + m) = i(n)$ gelten, was unmöglich ist.

Die archimedische Eigenschaft bedeutet, daß $i(\mathbb{N}) \subseteq \mathbb{R}$ unbeschränkt ist (um $nx > y$ zu erfüllen, muß man $i(n) > \frac{y}{x}$ erfüllen). In der Tat, wäre diese Menge beschränkt, dann gäbe es eine kleinste obere Schranke

$$s := \sup i(\mathbb{N}) \in \mathbb{R} \ .$$

Für alle $n \in \mathbb{N}$ gilt nun $i(n) = i(n + 1) - 1 \leq s - 1$. Also wäre auch $s - 1$ eine obere Schranke von $i(\mathbb{N})$, was aber wegen $s - 1 < s$ unmöglich ist. $\blacksquare$

Satz 2.14 *Es gibt einen Körper reeller Zahlen. Für zwei Körper reeller Zahlen $\mathbb{R}$ und $\mathbb{R}'$ gibt es genau einen Isomorphismus $\mathbb{R} \xrightarrow{\sim} \mathbb{R}'$ geordneter Körper.*

Beweis: Wir skizzieren nur die Schritte:

1. Im ersten Schritt zeigt man, daß die Menger der Schnitte $\mathbb{R}$ von $\mathbb{Q}$ (siehe 2.4) mit der Struktur eines geordneten Körpers versehen werden kann, der archimedisch ist und die Supremumseigenschaft besitzt (Details: [Rud80], Anhang zu Kap. 1). In der Tat, ist $A \subset \mathbb{R}$ nichtleer und beschränkt, dann gilt
$$\sup A = \bigcup_{U \in A} U \ .$$
Dies zeigt die Existenzaussage.

2. Sei $\mathbb{R}'$ nun ein zweiter Körper reeller Zahlen. Dann gibt es genau eine Einbettung $\mathbb{Q} \hookrightarrow \mathbb{R}'$. Diese ist automatisch mit der Ordnung verträglich. Wir sehen dann ein, daß jede Zahl $x \in \mathbb{R}'$ einen Schnitt von $\mathbb{Q}$ definiert. Auf diese Weise erhalten wir einen Körperhomomorphismus $\mathbb{R}' \to \mathbb{R}$ welcher notwendiger Weise injektiv ist. Ein Element in $\mathbb{R}$, also ein Schnitt bestimmt über die Einbettung $\mathbb{Q} \hookrightarrow \mathbb{R}'$ eine beschränkte nichtleere Teilmenge von $\mathbb{R}'$. Deren Supremum existiert und ist ein Element von $\mathbb{R}'$. Dies liefert den inversen Homomorphismus $\mathbb{R}' \to \mathbb{R}$..

3. Wir sehen ein, daß der in 2. konstruierte Isomorphismus der einzige mit der Ordnung verträgliche ist.

Für den Aufbau der Analysis fixieren wir von nun an einen Körper der reellen Zahlen. Wir werden nur die Axiome benutzen, nicht aber die konkrete Realisierung.

2.3 Potenzen, Dezimalbrüche, Überabzählbarkeit

Wir kann man nun reelle Zahlen angeben. Hier sind einige Beispiele.
Sei $n \in \mathbb{N}$.

Definition 2.15 *Für $p \in \mathbb{R}$, $p > 0$ definieren wir*
$$p^{\frac{1}{n}} := \sup\{q \in \mathbb{Q} | q^n < p\} \ .$$

In der Tat ist $0 \in \{q \in \mathbb{Q} | q^n < p\}$. Damit ist diese Menge nicht leer. Desweiteren folgt aus $0 < q$, $q^n < p$, daß $q < p + 1$ (da andernfalls $q^n \geq (p+1)^n) \geq p^n$ gelten würde). Damit ist diese Menge von oben beschränkt und besitzt ein Supremum in $\mathbb{R}$.

Lemma 2.16 *Es gilt $(p^{\frac{1}{n}})^n = p$.*

Beweis: In der Tat, wäre $(p^{\frac{1}{n}})^n < p$, gäbe es ein $0 < \epsilon \in \mathbb{Q}$ mit $(p^{\frac{1}{n}} + \epsilon)^n < p$. Um dieses zu finden beobachten wir, daß für $\epsilon < 1$

$$(p^{\frac{1}{n}} + \epsilon)^n = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} (p^{\frac{1}{n}})^i \epsilon^{n-i} \leq (p^{\frac{1}{n}})^n + \sum_{i=1}^{n} n!(p^{\frac{1}{n}})^{n-i} \epsilon$$

gilt. Wir können also

$$\epsilon := \frac{p - (p^{\frac{1}{n}})^n}{2 \sum_{i=1}^{n} n!(p^{\frac{1}{n}})^{n-i}}$$

setzen. Damit wäre aber $p^{\frac{1}{n}}$ keine obere Schranke von $\{q \in \mathbb{Q} | q^n < p\}$.

Wäre $(p^{\frac{1}{n}})^n > p$, dann existiert ein $0 < \epsilon \in \mathbb{Q}$ so daß $p^{\frac{1}{n}} - \epsilon$ immer noch eine obere Schranke von $\{q \in \mathbb{Q} | q^n < p\}$ ist. Damit kann $p^{\frac{1}{n}}$ nicht die kleinste obere Schranke dieser Menge gewesen sein. In der Tat kann ϵ wie folgt finden. Wir rechnen für $\epsilon < 1$

$$(p^{\frac{1}{n}} - \epsilon)^n = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} (p^{\frac{1}{n}})^i (-\epsilon)^{n-i} \geq (p^{\frac{1}{n}})^n - \sum_{i=1}^{n} n!(p^{\frac{1}{n}})^{n-i} \epsilon \ .$$

Wir setzen also

$$\epsilon := \frac{(p^{\frac{1}{n}})^n - p}{2 \sum_{i=1}^{n} n!(p^{\frac{1}{n}})^{n-i}} \ .$$

■

Lemma 2.17 *1. Sei $p > 0$ und $n \in \mathbb{N}$. Sei $q \in \mathbb{R}$, $q > 0$ und $q^n = p$. Dann gilt $q = p^{\frac{1}{n}}$.*

2. Seien $a, b \in \mathbb{R}$, $a, b > 0$. Dann gilt $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$.

3. Seien $n, m \in \mathbb{N}$ und $p \in \mathbb{R}$, $p > 0$. Dann gilt $(p^{\frac{1}{n}})^{\frac{1}{m}} = p^{\frac{1}{nm}}$.

Beweis: Aus $0 < x, y$ und $x < y$ folgt $x^n < y^n$. Wenn $q \in \mathbb{R}$, $q > 0$ eine der Ungleichungen $q < p^{\frac{1}{n}}$ oder $q > p^{\frac{1}{n}}$ erfüllt, dann gilt $q^n > p$ oder $q^n < p$. Wenn also $q^n = p$, dann muß $q = p^{\frac{1}{n}}$ gelten.

Die zweite Aussage folgt aus der ersten und.

$$((ab)^{\frac{1}{n}})^n = ab = (a^{\frac{1}{n}})^n (b^{\frac{1}{n}})^n = (a^{\frac{1}{n}} b^{\frac{1}{n}})^n \ .$$

Für die dritte Aussage benutzen wir $(a^n)^m = a^{nm}$. Es gilt

$$((p^{\frac{1}{n}})^{\frac{1}{m}})^{nm} = (((p^{\frac{1}{n}})^{\frac{1}{m}})^m)^n = (p^{\frac{1}{n}})^n = p \ .$$

■

Sei nun $r \in \mathbb{Q}$. Wir können $r = \frac{a}{b}$ schreiben mit $b > 0$. Für $p \in \mathbb{Q}$, $p > 0$ können wir

$$p^r := (p^{\frac{1}{b}})^a$$

definieren. Ein Problem bei dieser Definition ist, daß a, b durch r nicht eindeutig bestimmt werden, die Formel für p^r aber von der Darstellung von r als Bruch abhängt. Wir müssen zeigen, daß das nicht so ist. Dies ist der Nachweis, daß p^r **wohldefiniert** ist. Sei $r = \frac{a'}{b'}$ mit $b' > 0$ eine andere Darstellung. Dann gibt es $n, m \in \mathbb{N}$ mit $nb = mb'$ woraus $na = ma'$ folgt. Wir haben

$$(p^{\frac{1}{b}})^a = ((p^{\frac{1}{bn}})^n)^a = (p^{\frac{1}{bn}})^{na} = (p^{\frac{1}{mb'}})^{ma'} = (p^{\frac{1}{b'}})^{a'} \ .$$

Lemma 2.18 *Für $r, s \in \mathbb{Q}$ und $p \in \mathbb{R}$, $p > 0$ gilt*

$$p^{rs} = (p^r)^s \ , \quad p^{r+s} = p^r p^s \ .$$

Beweis: Wir benutzen die schon gezeigten Identitäten. Sei $r = \frac{a}{b}$ und $s = \frac{c}{d}$ für ganze Zahlen a, b, c, d mit $b, d > 0$.

$$p^{rs} = p^{\frac{ab}{cd}} = (p^{\frac{1}{bd}})^{ac} = (((p^{\frac{1}{b}})^{\frac{1}{d}})^a)^c = (((p^{\frac{1}{b}})^a)^{\frac{1}{d}})^c = (p^r)^s \ .$$

Wir können nach Hauptnennerbilden annehmen, daß $b = d$ ist.

$$p^{r+s} = (p^{\frac{1}{b}})^{a+c} = (p^{\frac{1}{b}})^a (p^{\frac{1}{b}})^c = p^r p^s \ .$$

■

Lemma 2.19 *Für $\mathbb{R} \ni x > 1$ und $r, s \in \mathbb{Q}$ mit $s < r$ gilt*

$$x^s < x^r \ .$$

Beweis: Wir schreiben $x^r = x^s x^{r-s}$. Es reicht zu zeigen, daß für $\mathbb{Q} \ni t > 0$ auch $x^t > 1$ gilt. In der Tat ist $t = \frac{a}{b}$ mit $a, b \in \mathbb{N}$ und $x^t = (x^{\frac{1}{b}})^a$. Nun ist $x^{\frac{1}{b}} > 1$, weil aus $0 < x^{\frac{1}{b}} < 1$ auch $x = (x^{\frac{1}{b}})^b < 1$ folgen würde. Damit ist aber auch $(x^{\frac{1}{b}})^a > 1$. ■

Wir werden später in der Vorlesung auch Potenzen x^r mit **reellem Exponenten** bilden. Die Idee ist für $x > 1$

$$x^r := \sup\{x^q | q \in \mathbb{Q} \wedge q \leq r\}$$

und für $x < 1$

$$x^r := \sup\{x^q | q \in \mathbb{Q} \wedge q \geq r\}$$

zu setzen. Für den Nachweis, daß diese Definition die gewünschten Eigenschaften hat, werden wir den Begriff der **Stetigkeit** essentiell verwenden. Deshalb verschieben wir diese Diskussion.

Wir zeigen nun, wie man reelle Zahlen durch **Dezimalbrüche** beschreiben kann. Der Einfachheit betrachten wir positive Zahlen. Ein Dezimalbruch ist eine Folge

$$a_0, a_1, \dots,$$

ganzer Zahlen mit $a_0 \in \mathbb{N}$ und $a_i \in \{0, \dots, 9\}$ für $i \geq 1$. Wir schreiben üblicherweise

$$193.4627365367\dots$$

für die Folge mit

$$a_0 = 193 \ , a_1 = 4 \ , a_2 = 6 \dots \ .$$

Ein Dezimalbruch (a_i) definiert eine reelle Zahl wie folgt. Wir betrachten die Zahlen

$$b_n := a_0 + \sum_{i=1}^{n} \frac{a_i}{10^i} \ .$$

Wir sehen zuerst ein, daß $a_0 \leq b_n < a_0 + 1$ für alle $n > 0$. Wir definieren den Schnitt von $\mathbb{Q}$

$$\{q \in \mathbb{Q} | (\exists n \in \mathbb{N} | q < b_n)\} \ .$$

Dieser stellt die reelle Zahl dar, welche dem Dezimalbruch (a_i) entspricht.

Sei nun x eine positive reelle Zahl. Dann setzen wir a_0 als die größte ganze Zahl mit $a_0 \leq x$. Die Folge (a_i) definieren wir nun induktiv. Seien a_i für $i = 0, \dots, n$ und damit b_n schon definiert. Dann ist $a_{n+1} \in \{0, \dots, 9\}$ die größte Zahl mit

$$b_n + \frac{a_{n+1}}{10^{n+1}} \leq x \ .$$

Beachte, daß die Darstellung reller Zahlen durch Dezimalrüche nicht ganz eindeutig ist. Zum Beispiel stellen $0.99999\dots$ und $1.00000\dots$ die gleiche reelle Zahl dar. Wir erreichen Eindeutigkeit, wenn wir Darstellungen mit Nullen am Ende vermeiden, also etwa $16.3650000000\dots$ durch $16.36499999999\dots$ ersetzen. Wir wollen solche Brüche reduziert nennen.

Lemma 2.20 *Die obigen Konstruktionen bestimmen eine Bijektion zwischen den Mengen der positiven reellen Zahlen und der reduzierten Dezimalbrüche.*

Beweis: Übungsaufgabe! ■

Wir verwenden die Dezimalbruchdarstellung im Beweis des folgenden Satzes.

Satz 2.21 *Die Menge der reellen Zahlen ist überabzählbar.*

Beweis: Es genügt zu zeigen, daß die Menge $(0, 1]$ überabzählbar ist. Äquivalent dazu ist die Überabzählbarkeit der Menge R der reduzierte Dezimalbrüche mit $a_0 = 0$. Wir nehmen die Abzählbarkeit von R an. Sei $b : \mathbb{N} \to R$ eine Bijektion. Wir schreiben $(b(n)_i)$ für den n-ten Dezimalbruch. Wir definieren einen reduzierten Dezimalbruch (a_k), indem wir $a_0 := 0$ und für jedes $k > 0$ die Ziffer a_k ungleich $b_k(k)$ oder 0 wählen. Dann ist $(a_k) \in R$, also $(a_k) = b(l)$ für ein $l \in \mathbb{N}$. Damit wäre $a_l = b(l)_l$, was aber durch unsere Wahl ausgeschlossen ist. Folglich ergibt sich ein Widerspruch. ■

2.4 Ungleichungen, Betrag

Sei $(M, \leq)$ ein geordnete Menge.

Definition 2.22 *Für $a, b \in M$ definieren wir die* **Intervalle**

$$\begin{aligned}
[a, b] &:= \{m \in M | a \leq m \wedge m \leq b\} \\
(a, b] &:= \{m \in M | a < m \wedge m \leq b\} \\
[a, b) &:= \{m \in M | a \leq m \wedge m < b\} \\
(a, b) &:= \{m \in M | a < m \wedge m < b\} \ .
\end{aligned}$$

Eine geordnete Menge kann vervollständigt werden. Wir führen dazu weiter zwei Elemente $\infty, -\infty$ ein und setzen

$$\bar{M} := M \sqcup \{\infty, -\infty\} \ .$$

Wir setzen die Ordnungsrelation von M fort auf $\bar{M}$ so daß $-\infty$ kleiner und ∞ größer als jedes Element von M ist und $-\infty < \infty$ gilt. Damit sind dann auch Intervalle $(-\infty, a), \ldots$ definiert.

Wir können diese Begriffe auf $(\mathbb{R}, \leq)$ anwenden. Die Menge $\bar{\mathbb{R}}$ ist die Menge der vervollständigten reellen Zahlen.

Definition 2.23 *Wir definieren die* **Betrags**-*Abbildung* $|.| : \mathbb{R} \to \mathbb{R}$ *durch*

$$|x| := \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases} \ .$$

Lemma 2.24 *Der Betrag hat folgende Eigenschaften:*

1. *Für alle $a \in \mathbb{R}$ gilt $|a| \geq 0$.*

2. *Für alle $a, b \in \mathbb{R}$ gilt $|ab| = |a||b|$.*

3. *Für alle $a, b \in \mathbb{R}$ gilt $|a + b| \leq |a| + |b|$.*

Beweis: Wir zeigen nur die letzte Eigenschaft durch eine Fallunterscheidung und unter Verwendung der Ungleichung $a \leq |a|$.

1. $a + b > 0 : \ |a + b| = a + b \leq |a| + |b|$

2. $a + b < 0 : \ |a + b| = -a - b \leq |a| + |b|$.

Eine nützliche Ungleichung ist

$$||a| - |b|| \leq |a - b| \ .$$

Wenn $|a| \geq |b|$ ist, dann gilt $|a| = |a - b + b| \leq |a - b| + |b|$, also $|a| - |b| \leq |a - b|$. Der Fall $|b| \geq |a|$ geht analog.

Lemma 2.25 *Sei $\mathbb{R} \ni \epsilon > 0$ und $a \in \mathbb{R}$. Die Aussage $|x - a| \leq \epsilon$ is äquivalent zu $x \in [a - \epsilon, a + \epsilon]$*

Beweis: Übungsaufgabe! ■

Definition 2.26 *Wir betrachten $a, b \in \mathbb{R}$. Das* **arithmetische Mittel** *von a, b ist durch*

$$m_{arith}(a, b) := \frac{a + b}{2}$$

definiert. Allgemeiner definiert man

$$m_{arith}(a_1, \ldots, a_n) := \frac{\sum_{i=1}^{n} a_i}{n}$$

für reelle Zahlen a_i, $i = 1, \ldots, n$.

Wenn $a < b$, dann gilt

$$m_{arith}(a, b) \in (a, b) \ .$$

Definition 2.27 *Das* **geometrische Mittel** *von positiven Zahlen $a, b \in \mathbb{R}$ ist durch*

$$m_{geom}(a, b) := (ab)^{\frac{1}{2}}$$

definiert. Allgemeiner definiert man

$$m_{geom}(a_1, \ldots, a_n) := (\prod_{i=1}^{n} a_i)^{\frac{1}{n}}$$

für positive reelle Zahlen a_i, $i = 1, \ldots, n$.

Für $0 < a < b$ gilt

$$m_{geom}(a, b) \in (a, b) \ .$$

In der Tat ist $b = ua$ für ein $u > 1$. Damit gilt

$$a < au^{\frac{1}{2}} = m_{geom}(a, b) = bu^{-\frac{1}{2}} < b \ .$$

Lemma 2.28 (Bernoulli-Ungleichung) *Sei $\mathbb{N} \ni n \geq 2$ und $\mathbb{R} \ni x > -1$, $x \neq 0$. Dann gilt*

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

Beweis: Die Binomische Formel gibt

$$(1+x)^n = 1 + nx + \sum_{i=2}^{n} \frac{n!}{i!(n-i)!} x^i \ .$$

Wir müßten zeigen, daß die Summe positiv ist. Da aber x negativ sein kann, ist das nicht offensichtlich.

Wir benutzen also eine andere Methode, namlich Induktion nach n. Sei zunächst $n = 2$. Dann gilt $(1+x)^2 = 1 + 2x + x^2 > 1 + 2x$. Sei jetzt die Bernoulli-Ungleichung für $n-1$ gezeigt. Dann gilt

$$(1+x)^{n+1} = (1+x)^n(1+x) > (1+nx)(1+x) = 1 + (n+1)x + nx^2 > 1 + (n+1)x \ ,$$

also die Bernoulli-Ungleichung für $n+1$. ■

Wir können nun das arithmetische mit dem geometrischen Mittel vergleichen.

Lemma 2.29 *Seien $a_1, \ldots, a_n$ positive reelle Zahlen. Dann gilt*

$$m_{geom}(a_1, \cdots, a_n) \leq m_{arith}(a_1, \ldots, a_n) \ .$$

Die Gleichheit gilt genau dann, wenn alle a_i gleich sind.

Beweis: Wir verwenden Induktion nach n. Der Fall $n = 1$ ist klar. Sei die Ungleichung für n gezeigt. Wir können o.B.d.A annehmen, daß $a_{n+1} \geq \mathtt{max}\{a_1, \ldots, a_n\}$ gilt (sonst umnumerieren). Wir müssen zeigen, daß

$$\prod_{i=1}^{n+1} a_i \leq \left(\frac{\sum_{i=1}^{n+1} a_i}{n+1} \right)^{n+1}$$

Wir setzen $a := m_{geom}(a_1, \ldots, a_n)$. Dann gilt $0 < a \leq a_{n+1}$, also $x := \frac{a_{n+1}-a}{(n+1)a} \geq 0$. Nun ist

$$1 + x = \frac{(n+1)a + a_{n+1} - a}{(n+1)a} = \frac{na + a_{n+1}}{(n+1)a} = \frac{a_1 + \cdots + a_{n+1}}{(n+1)a} \ .$$

Wir erhalten aus der Bernoulli-Ungleichung

$$(1+x)^{n+1} \geq 1 + (n+1)x = \frac{a_{n+1}}{a} \ .$$

Wir schließen mit der Induktionsvoraussetzung im zweiten Schritt

$$\left(\frac{\sum_{i=1}^{n+1} a_i}{n+1} \right)^{n+1} = (1+x)^{n+1} a^{n+1} \geq a^n a_{n+1} \geq \prod_{i=1}^{n+1} a_i \ .$$

Die Diskussion der Gleichheit überlassen wir als Übungsaufgabe. ■

3 Grenzwerte von Folgen

3.1 Folgen reeller Zahlen

Zur Erinnerung, eine Folge reeller Zahlen $(a_i)_{i\in\mathbb{N}}$ (oder kurz (a_i)) ist formal gesehen eine Abbildung $a : \mathbb{N} \to \mathbb{R}$, wobei a_i den Wert von a an der Stelle $i \in \mathbb{N}$ bezeichnet. Manchmal lassen wir eine Folge auch mit dem nullten Glied beginnen, betrachten also eine Abbildung $a : \mathbb{N}_0 \to \mathbb{R}$. Es gibt verschiedene Methoden Folgen anzugeben.

1. Da $\mathbb{N}$ unendlich ist, kann man sicher nicht alle Glieder **aufzählen**. Die Angabe von $1, 2, 3, 4, 5 \ldots$ legt nahe, das die Folge mit $a_i = i$ gemeint ist. Wir werden diese Methode manchmal beutzen, wenn aus dem Kontext klar wird, was gemeint ist. Formal gesehen könnte jedoch $a_6 = 0$ sein.

2. Eine Folge kann durch eine **Formel für das i-te Glied** angegeben werden. Zum Beispiel wird durch $a_i := \frac{1}{2i+1}$ eine Folge beschrieben.

3. Eine Folge kann durch eine **rekursive Vorschrift** angegeben werden. Hier ein Beispiel: $a_1 := 1$, $a_{i+1} := 2a_i$ für $i \geq 2$. In diesem Beispiel ist eine Formel für das i-te Glied leicht zu finden. Es gilt $a_i = 2^i$. Im Fall von $a_1 := 1$, $a_{i+1} := \frac{1}{1+a_i^2}$ scheint das schon fast aussichtslos zu sein.

Die Glieder der Folge $a_i = \frac{1}{i}$ nähern sich mit größer werdendem i immer mehr der Null an. Wir sagen, daß 0 der Grenzwert dieser Folge. Wir haben die Zahl e als Grenzwert der Folge $a_i := \sum_{n=0}^{i} \frac{1}{n!}$ definiert. Wir wollen nun den Begriff des Annähern präzise fassen.

Sei $(a_i)_{i=1}^{\infty}$ eine Folge reeller Zahlen und a eine reelle Zahl.

Definition 3.1 *Die Zahl a ist* ***Grenzwert*** *der Folge $(a_i)_{i=1}^{\infty}$, wenn für jede positive reelle Zahl ϵ eine natürliche Zahl N existiert, so daß für alle natürlichen Zahlen $i \geq N$ die Relation $|a_i - a| < \epsilon$ gilt. In diesem Fall schreiben wir $a = \lim_{i\to\infty} a_i$ oder $a_i \to a$ und sagen auch, daß (a_i) gegen a* ***konvergiert****.*

In Symbolen: a ist Grenzwert der Folge (a_i), wenn

$$(\forall \epsilon \in (0, \infty)\, \exists N \in \mathbb{N} | (\forall i \in N | i \geq N \to |a_i - a| < \epsilon))$$

wahr ist.

Nicht jede Folge besitzt einen Grenzwert. Sei etwa $a_i = (-1)^i$. Möge a ein Grenzwert dieser Folge sein. Dann wählen wir $\epsilon = 1$. Sei $N \in \mathbb{N}$ wie in der Definition. Dann gilt für $i \geq N$, daß

$$1 = \epsilon > |a_i - a| = |a_i - a + a_{i+1} - a_{i+1}| \geq |a_i - a_{i+1}| - |a - a_{i+1}| > 2 - \epsilon = 1\ ,$$

ein Widerspruch.

Eine Folge kann **höchstens einen Grenzwert** besitzen. Seien $a, b \in \mathbb{R}$ beides Grenzwerte von (a_i). Wir nehmen an, daß $a \neq b$ ist und wählen $\epsilon = \frac{|a-b|}{2}$. Sei $N \in \mathbb{N}$ wie in der

Definition. Dann gilt für $i \geq N$. $2\epsilon = |a-b| = |a-a_i+a_i-b| \leq |a-a_i| + |a_i-b| < 2\epsilon$, ein Widerspruch.

Sei (a_i) eine konvergente Folge. Wenn man **endlich viele Glieder ändert**, dann ist die neue Folge auch konvergent und hat den gleichen Grenzwert wie die alte.

Um zu zeigen, daß eine Folge reeller einen Grenzwert hat, kann man entsprechend der Definition vorgehen. Zuerst rät man den Grenzwert a. Dann weist man für jedes positive ϵ die Existenz eines $N \in \mathbb{N}$ mit den gewünschten Eigenschaften nach.

Oft hilft auch folgende Beobachtung. Seien (a_i) und (b_i) konvergente Folgen mit den Grenzwerten a, b und gelte $a_i \leq b_i$ für unendlich viele $i \in \mathbb{N}$. Dann gilt $a \leq b$. In der Tat, wäre $b < a$, dann wählen wir $\epsilon := \frac{a-b}{2}$. Da die Folgen konvergieren, gibt es ein $N_0 \in \mathbb{N}$ derart, daß $|a_i - a| < \epsilon$ und $|b_i - b| < \epsilon$ für alle $i \geq N_0$ gilt. Damit ist für alle $i \geq N_0$

$$a_i > a - \epsilon = b + \epsilon > b_i \ ,$$

ein Widerspruch.

Wir nehmen jetzt an, daß zusätzlich $a = b$ gilt. Sei (x_i) eine weitere Folge mit $a_i \leq x_i \leq b_i$. Dann ist (x_i) konvergent und der Grenzwert ist auch a In der Tat ist $|x_i - a| \leq \mathtt{max}\{|a_i - a|, |b_i - a|\}$.

1. Wir betrachten etwa die durch
$$a_i := \frac{1}{n}$$
definierte Folge. Wir raten $a = 0$. Der Nachweis, daß 0 Grenzwert von (a_i) ist, geht wie folgt. Sei $\epsilon > 0$ gegeben. Nach dem archimedischen Axiom gibt es $N \in \mathbb{N}$ mit $N < \frac{1}{\epsilon}$. Für $i \geq N$ gilt $\frac{1}{i} \leq \frac{1}{N} < \epsilon$, also $|a_i - 0| < \epsilon$.

2. Für $0 < r \in \mathbb{Q}$ gilt
$$\lim_{n\to\infty} \frac{1}{n^r} = 0 \ .$$
Sei $\epsilon > 0$ gegeben. Dann wählen wir $N_0 > \frac{1}{\epsilon^{\frac{1}{r}}}$. Für $n \geq N_0$ gilt $0 < \frac{1}{n^r} < \epsilon$. Das zeigt die Behauptung.

3. Sei $x \in \mathbb{R}$ mit $|x| < 1$. Dann konvergiert gilt
$$\lim_{i\to\infty} x^i = 0 \ .$$
In der Tat ist $|x^i - 0| = |x|^i$. Es gilt $|x|^{-1} = 1 + \delta$ mit $\delta > 0$. Folglich mit dr Bernoullischen ungleichung $|x|^{-i} = (1+\delta)^{-i} \leq (1+i\delta)^{-1}$. Sei jetzt $\epsilon > 0$ gegeben. Dann wählen wir (archimedisches Axiom) N_0 so groß, daß $N_0\delta > \epsilon^{-1}$. Für $i \geq N_0$ gilt dann
$$|x|^{-i} \leq (1+i\delta)^{-1} \leq i^{-1}\delta^{-1} \leq N_0^{-1}\delta^{-1} < \epsilon \ .$$

4. Für $\mathbb{R} \ni x > 0$ gilt
$$\lim_{n\to\infty} x^{\frac{1}{n}} = 1 \ .$$

Sei zunächst $x > 1$. Dann ist $a_n := x^{\frac{1}{n}} - 1 > 0$. Die binomische Formel liefert $1 + na_n \le (1 + a_n)^n = x$. Folglich gilt $0 \le a_n \le \frac{x-1}{n}$. Wir schließen $\lim_{n\to\infty} a_n = 0$. Daraus folgt die Behauptung. Wenn $0 < x < 1$, dann betrachtet man die Folge $((\frac{1}{x})^{\frac{1}{n}})$.

5. Es gilt
$$\lim_{n\to\infty} n^{\frac{1}{n}} = 1 \ .$$
Wir setzen $a_n := n^{\frac{1}{n}} - 1$. dann gilt $a_n > 0$. Mit der binomischen Formel erhalten wir (nimm nur den dritten Term)
$$n = (1 + a_n)^n \ge \frac{n(n-1)}{2} a_n^2 \ .$$
Folglich gilt $0 < a_n \le (\frac{2}{n-1})^{1/2}$. Daraus folgt die Behauptung.

6. Für $\alpha \in \mathbb{Q}$ und $x > 1$ gilt
$$\lim_{n\to\infty} \frac{n^\alpha}{x^n} = 0 \ .$$
Sei $k \in \mathbb{N}$ mit $k > \alpha$. Dann gilt für $n > 2k$, daß $n - k + 1 > \frac{n}{2}$, $(n - k) > k$, also
$$x^n = (1+(x-1))^n > \frac{n!}{(n-k)!k!}(x-1)^k = \frac{n(n-1)\dots(n-k+1)}{k!}(x-1)^k > \frac{n^k(x-1)^k}{2^k k!} \ .$$
Daraus schließen wir (für $n > 2k$,
$$0 < \frac{n^\alpha}{x^n} < \frac{2^k k!}{(x-1)^k} n^{\alpha-k} \ .$$
Nun gilt $\lim_{n\to\infty} n^{\alpha-k} = 0$. Daraus folgt die Behauptung.

Oft kennen wir den Grenzwert der betrachteten Folge nicht. Dies war etwa im Fall der e definierenden Folge der Fall. In diesem Fall sichert folgendes Lemma die Existenz eines Grenzwertes. Eine Folge reeller Zahlen ist **monoton wachsend**, wenn sie als Abbildung $\mathbb{N} \to \mathbb{R}$ monoton ist, also aus $i \le j$ folgt $a_i \le a_j$. Eine Folge reeller Zahlen ist **von oben beschränkt**, wenn die Menge $\{a_i | i \in \mathbb{N}\}$ eine obere Schranke besitzt.

Lemma 3.2 *Sei $(a_i)_{i=1}^\infty$ eine monoton wachsende von oben beschränkte Folge reeller Zahlen. Dann besitzt sie einen Grenzwert.*

Beweis: Die Menge $\{a_i | i \in \mathbb{N}\}$ ist nicht leer und nach Voraussetzung von oben beschränkt. Deshalb existiert $a := \sup\{a_i | i \in \mathbb{N}\} \in \mathbb{R}$. Wir nehmen an, daß a nicht Grenzwert der Folge ist. Dann existiert ein $\epsilon > 0$ so daß für alle $N \ge 0$ ein noch größeres $i \in \mathbb{N}$ existiert mit $|a_i - a| \ge \epsilon$. In unserem Fall ist $a \ge a_i$ so daß $a - a_i \ge \epsilon$. In anderen Worten, es gibt beliebig große $i \in \mathbb{N}$ mit $a_i \le a - \epsilon$. Sei jetzt $j \in \mathbb{N}$ gegeben. Dann gibt es sicher ein solches i mit $i \ge j$. Es folgt aus der Monotonie, daß $a_j \le a_i \le a - \epsilon$.

Also ist $a-\epsilon$ immer noch eine obere Schranke von $\{a_i|i\in\mathbb{N}\}$. Das ist ein Widerspruch. ■

Die Folge $a_i := \sum_{n=0}^{i} \frac{1}{n!}$ ist sicher monoton wachsend, da die Summanden alle positiv sind. Wir hatten weiterin 3 als eine obere Schranke gefinden. Diese beiden Beobachtungen lieferten die Existenz des Grenzwertes.

Lemma 3.3 *Sei $A \subset \mathbb{R}$ eine von oben beschränkte nichtleere Teilmenge. Dann gibt es eine (sogar monotone) Folge (a_n) mit $a_n \in A$ und $\lim_{n\to\infty} a_n = \sup A$.*

Beweis: Wir konstruieren die Folge (a_n) induktiv. Wir wählen ein Element $a_0 \in A$ aus. Sei nun a_i für $i \le n$ schon konstruiert. Wenn $a_n = \sup A$ ist, dann wählen wir $a_{n+1} := a_n$. Anderfalls ist die Menge $A \cap [\mathtt{max}\{a_n, \sup A - \frac{1}{n}\}, \sup A]$ nicht leer und wir wählen aus dieser Menge ein Element a_{n+1}. Die so konstruierte Folge ist monoton. Es gilt $|a_n - \sup A| \le \frac{1}{n}$, also $\lim_{n\to\infty} a_n = \sup A$. ■

Sei (a_i) irgend eine Folge. Dann können wir eine Folge von Mengen $A_n := \sup\{a_i|n \le i \in \mathbb{N}\}$ bilden. Die Folge (A_n) ist monoton falled bezüglicher der Relation $\subseteq$, d.h. es gilt $A_{n+1} \subseteq A_n$ für all $n \in \mathbb{N}$. Wir nehmen jetzt an, daß die Folge (a_i) beschränkt ist. Die Folge reeller Zahlen

$$b_n := \sup A_n$$

ist dann definiert, monoton fallend und von unten beschränkt. Wir definieren den **Limes superior** (Limes der kleinsten oberen Schranken) der Folge (a_i) durch

$$\limsup_{t\to\infty} a_i := \lim_{n\to\infty} b_n \ .$$

Beachte, daß $\limsup_{t\to\infty} a_i$ für jede beschränkte Folge definiert ist. Analog definiert man $\liminf_{t\to\infty} a_i$. Es gilt immer $\inf A_i \le \sup A_i$, also

$$\liminf_{t\to\infty} a_i \le \limsup_{t\to\infty} a_i \ .$$

Die Folge (a_i) konvergiert genau dann, wenn in in dieser Relation die Gleichheit gilt. In diesem Fall gilt

$$\liminf_{t\to\infty} a_i = \lim_{i\to\infty} a_i = \limsup_{t\to\infty} a_i \ .$$

Sei (a_i) eine Folge reeller Zahlen und (i_k) eine streng monoton wachsende Folge natürlicher Zahlen. Dann können wir eine neue Folge (a_{i_k}) bilden. In Termen von Abbildungen ist diese als Komposition

$$\mathbb{N} \xrightarrow{i} \mathbb{N} \xrightarrow{a} \mathbb{R}$$

definiert.

Definition 3.4 *Die Folge (a_{i_k}) nennt man eine Teilfolge von (a_i).*

Beachte, daß wir i als streng monoton annehmen.

Die Folge (a_i) **konvergiert** gegen a dann und nur dann, wenn **jede Teilfolge** gegen a konvergiert.

Die Folge $((-1)^i)$ konvergiert nicht, besitzt aber konvergente Teilfolgen $((-1)^{2i})$ und $((-1)^{2i+1})$, welche gegen 1 und -1 konvergieren.

Die rationalen Zahlen sind abzählbar. Folglich gibt es eine Folge (a_i) mit $\{a_i | i \in \mathbb{N}\} = \mathbb{Q}$. Diese Folge besitzt für jede reelle Zahl eine gegen diese Zahl konvergierende Teilfolge. Sei $x \in \mathbb{R}$ durch einen Schnitt $U \subset \mathbb{Q}$ repräsentiert. Dann ist $x = \sup U$. Die Folge finden wir mit Lemma 3.3.

Definition 3.5 *Eine Folge (a_i) reeller Zahlen heißt **beschränkt** (von oben, von unten), wenn die Menge $\{a_i | i \in \mathbb{N}\}$ diese Eigenschaften hat.*

Lemma 3.6 *Eine beschränkte Folge (a_i) reeller Zahlen besitzt eine konvergente Teilfolge.*

Beweis: Wegen der Beschränktheit der Folge gibt es ein Intervall (a, b), welches alle Folgenglieder enthält. Wir setzen $x_0 := a$ und $y_0 := b$. Wir konstruieren induktiv Folgen (x_n), (y_n) und (i_n) derart, daß (i_n) streng monoton wächst, (x_n) monoton wächst und (y_n) monoton fällt und $x_n \leq a_{i_n} \leq y_n$ gilt. Im n-ten Schritt wählen wir die obere oder untere Hälfte $[x_n, y_n]$ des Intervalls $[x_{n-1}, y_{n-1}]$ derart aus, daß dieses unendlich viele Folgenglieder enthält. Dann gilt entweder $x_n = x_{n-1}$ und $y_n < y_n - 1$ oder $x_{n-1} < x_n$ und $y_{n-1} = y_n$. Als nächstes wählen wir $i_n > i_{n-1}$ derart, daß $a_{i_n} \in [x_n, y_n]$ ist.

Die beschränkten monotonen Folgen (x_n) und (y_n) haben Grenzwerte $x \leq y$. Wegen $|y_n - x_n| \leq (b-a)2^{-n}$ gilt $x = y$. Damit konvergiert auch $a_{i_n} \to x$. ■

Mit Hilfe des Begriffs einer Cauchy-Folge erhält man ein weiteres Krtierium für Konvergenz, welches nicht die Kenntnis des Grenzwertes benutzt.

Definition 3.7 *Eine Folge reeller Zahlen (a_i) heißt **Cauchy-Folge**, wenn für alle $0 < \epsilon \in \mathbb{R}$ ein $N \in \mathbb{N}$ existiert, so daß für alle $n, m \geq N$ die Ungleichung $|a_n - a_m| < \epsilon$ gilt.*

Lemma 3.8 *Eine Folge reeller Zahlen ist dann und nur dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis: Sei (a_i) eine konvergente Folge mit dem Grenzwert a. Wir geben $\epsilon > 0$ vor und wählen N derart, daß $|a_n - a| < \frac{\epsilon}{2}$ für alle $n \geq N$. Damit gilt für $n, m \geq N$

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a_m - a| < \epsilon \ .$$

Dies zeigt die Cauchyfolgeneigenschaft.

Umgekehrt, sei (a_i) eine Cauchyfolge. Dann wählen wir $N \in \mathbb{N}$ derart, daß $|a_n - a_m| < 1$ für $n, m \geq N$ gilt. Insbesondere liegen alle Glieder a_m für $m \geq N$ im Intervall $[a_N - 1, a_N + 1]$. Damit ist die Folge beschränkt und hat eine konvergente Teilfolge (a_{i_k}) mit dem Grenzwert a. Sei nun $\epsilon > 0$ gegeben. Wir wählen $N_0 \in \mathbb{N}$ derart, daß $|a_m - a_n| < \frac{\epsilon}{2}$

und $|a_{i_n} - a| < \frac{\epsilon}{2}$ für alle $n, m \geq N_0$ gilt. Dann gilt für $m \geq N_0$ auch (beachte, daß $i_m \geq m$)

$$|a_m - a| = |a_m - a_{i_m} + a_{i_m} - a| \leq |a_m - a_{i_m}| + |a_{i_m} - a| < \epsilon \ .$$

Dies zeigt die Konvergenz. ■

Durch Kombinationen mit den Rechenoperationen können Folgen kombiniert werden. Seien (a_i) und (b_i) Folgen reeller Zahlen. Dann können wir die Folgen $(a_i + b_i)$, $(a_i b_i)$, (λa_i) für $\lambda \in \mathbb{R}$ betrachten. Wenn $b_i \neq 0$ für alle $i \in \mathbb{N}$ gilt, können wir auch $(\frac{a_i}{b_i})$ bilden.

Lemma 3.9 *Wenn (a_i) und (b_i) gegen a bzw. b konvergieren, dann konvergieren $(a_i + b_i)$ geben $a + b$, $(a_i b_i)$ gegen ab und (λa_i) gegen λa. Wenn $b_i \neq 0$ für alle $i \in \mathbb{N}$ und $b \neq 0$ ist, dann konvergiert $(\frac{a_i}{b_i})$ gegen $(\frac{a}{b})$.*

Beweis: Wir führen den Nachweis im Fall des Produktes. Alle Behauptungen werden wir später auf einfache und systematische Weise aus der Stetigkeit der Rechenoperationen schließen.

Sei $1 \geq \epsilon > 0$ gegeben. Wir wählen $N \geq 0$ derart, daß aus $i \geq N$ sowohl $|a_i - a| < \frac{\epsilon}{3(|b|+1)}$ als auch $|b_i - b| < \frac{\epsilon}{3(|a|+1)}$ folgt. Dann gilt für $i \geq N$ auch

$$\begin{aligned}
|a_i b_i - ab| &= |a_i b_i - ab_i + ab_i - ab| \\
&\leq |a_i - a||b_i| + a|b_i - b| \\
&< \frac{\epsilon |b_i|}{3(|b|+1)} + \frac{\epsilon |a|}{3(|a|+1)} \\
&\leq \frac{\epsilon(|b_i - b| + |b|)}{3(|b|+1)} + \frac{\epsilon}{3} \\
&\leq \frac{2\epsilon}{3} + \frac{\epsilon^2}{9(|a|+1)(|b|+1)} \\
&\leq \epsilon \ .
\end{aligned}$$

■

Hier sind einige Anwendungen.

1. Seien

   $$p(x) := p_r x^r + p_{r-1} x^{r-1} + \cdots + p_1 x + p_0 \ , \quad q(x) := q_r x^r + q_{r-1} x^{r-1} + \cdots + q_1 x + q_0$$

   Polynome mit reellen Koeffizienten und $q_r \neq 0$. Es gilt $q(i) = 0$ für höchstens endlich viele $i \in \mathbb{N}$. Insbesondere existiert ein $N \in \mathbb{N}$ mit $q(i) \neq 0$ für $i \geq N$. Wir betrachten eine Folge mit den Gliedern $\frac{p(i)}{q(i)}$ für $i \geq N$. Dann ist diese Folge konvergent und es gilt

   $$\lim_{i \to \infty} \frac{p(i)}{q(i)} = \frac{p_r}{q_r} \ .$$

Um dies einzusehen, schreiben wir

$$a_i = \frac{p(i)}{q(i)} = \frac{i^{-r}p(i)}{i^{-r}q(i)} \ .$$

Es gilt

$$i^{-r}p(i) = p_r + i^{-1}p_{r-1} + \dots i^{-r}p_0 \ , i^{-r}q(i) = q_r + i^{-1}q_{r-1} + \dots i^{-r}q_0 \ .$$

Die Folgen $p_{r-k}i^{-k}$ konvergieren alle gegen Null. Deshalb konvergiert deren Summe $i^{-r}p(i)$ gegen p_r. Analoges gilt für $i^{-r}q(i)$ welche gegen $q_r \neq 0$ Der Quotient konvergiert damit gegen $\frac{p_r}{q_r}$.
Also gilt etwa

$$\lim_{i\to\infty} \frac{3i^2+2i+1}{i^2+3} = 3 \ .$$

2. Sei $p(x)$ wie in 1 und $a \in \mathbb{R}$ fest. Wir betrachten die Folge mit den Gliedern

$$\frac{p(a+i^{-1})-p(a)}{i^{-1}} \ .$$

In diesem Beispiel ist der Grenzwert der Folge im Nenner 0, so daß wir nicht direkt den Satz über die Quotienten anwenden können. Wir schreiben mit Hilfe der binomischen Formel

$$p(a+i^{-1}) - p(a) = \sum_{k=0}^{r} p_k((a+i^{-1})^k - a^k = \sum_{k=1}^{r} p_k \sum_{l=1}^{k} \frac{k!}{l!(k-l)!} a^{k-l} i^{-l} \ .$$

Wir sehen daß

$$\frac{p(a+i^{-1})-p(a)}{i^{-1}} = \sum_{k=1}^{r} kp_k a^{k-1} + \sum_{l=1}^{r} c_r i^{-l}$$

für geeignete $c_l \in \mathbb{R}$. Der zweite Term ist eine Summe gegen Null konvergierender Folgen. Wir schließen

$$\lim_{i\to\infty} \frac{p(a+i^{-1})-p(a)}{i^{-1}} = kp_k a^{k-1} + (k-1)p_{k-1}a^{k-2} + \dots + 2p_2 a + p_1 \ .$$

3.2 Reihen

Sei $(a_i)_{i\in\mathbb{N}}$ eine Folge reeller Zahlen. Dann können wir eine neue Folge $(s_n)_{n\ni\mathbb{N}}$ mit dem n-ten Glied

$$s_n := \sum_{i=0}^{n} a_n$$

bilden, die n-te **Partialsumme** der Folge (a_i).

Definition 3.10 *Wenn die Folge $(s_n)_{n=0}^{\infty}$ konvergiert, dann machen wir folgende Definition:*

$$\sum_{i=0}^{\infty} a_i := \lim_{n\to\infty} s_n \ .$$

Wir sprechen dann von einer ***konvergenten Reihe*** $\sum_{i=0}^{\infty} a_i$.

Hier sind einige Beispiele:

1. Die **geometrische Reihe** $\sum_{i=0}^{\infty} x^i$ konvergiert für $|x| < 1$ gegen $\frac{1}{1-x}$. In der Tat gilt $s_n = \frac{1-x^{n+1}}{1-x}$ und $\lim_{n\to\infty} x^n = 0$.

2. Wir hatten schon gesehen, daß die Reihe

$$\mathsf{e} = \sum_{n=0}^{\infty} \frac{1}{n!}$$

konvergiert.

3. Die zum Nachweis der Konvergenz dieser Reihe angewandte Methode funktioniert allgemeiner. Sei (a_i) eine Folge nicht-negativer reeller Zahlen. Wenn es eine Zahl C gibt, so daß für alle $N \in \mathbb{N}$ $\sum_{n=0}^{N} a_n \le C$ gilt, dann konvergiert $\sum_{n=0}^{\infty} a_n$. In der Tat ist die Folge (s_n) monoton wachsend und beschränkt, folglich konvergent.

4. Eine Reihe $\sum_{n=0}^{\infty} a_n$ konvergiert genau dann, wenn die Folge der Partialsummen eine Cauchy Folge ist. Folglich konvergiert die Reihe genau dann, wenn für alle $\epsilon > 0$ ein N_0 existiert, so daß $|\sum_{i=n}^{m} a_i| < \epsilon$ gilt. Diese Aussage nennt man **Cauchy Kriterium**.

5. Seien (a_i) und (c_i) Folgen derart, daß $|a_i| < c_i$ für alle (es reicht für alle bis auf endlich viele) $i \in \mathbb{N}$ gilt und $\sum_{i=0}^{\infty} c_i$ konvergiert. Dann konvergiert auch $\sum_{i=0}^{\infty} a_i$. Die Folge (c_i) heißt **Majorante** von (a_i), und diese Aussage heißt auch **Majorantenkriterium**. Sei $\epsilon > 0$ gegeben. Wir wählen $N_0 \in \mathbb{N}$ derart, daß $\sum_{i=n}^{m} c_i < \epsilon$ für alle $n, m \ge N_0$ gilt. Für diese n, m gilt auch

$$|\sum_{i=n}^{m} a_i| \le \sum_{i=n}^{m} |a_i| \le \sum_{i=n}^{m} c_i < \epsilon \ .$$

Folglich konvergiert $\sum_{i=0}^{\infty} a_i$ nach dem Cauchy Kriterium.

6. Oft kann man die geometrische Reihe als Majorante benutzen. Wir hatte zum Beispiel beobachtet, daß $\frac{1}{n!} < 2\,2^{-n}$ ist und daraus auf die Konvergenz der e-Reihe geschlossen. Allgemeiner, sei $x \in \mathbb{R}$. Dann konvergiert die **Exponentialreihe**

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

Sei $\mathbb{N} \ni N > |x| + 1$ und $u := \frac{|x|}{N} < 1$. Dann gilt (Faktoren abzählen) $|\frac{x^n}{n!}| < \frac{|x|^N}{N!} u^{-N} u^n$. Folglich ist

$$|\frac{x^n}{n!}| \leq \frac{x^N}{N!} u^{-N} u^n \ .$$

Damit haben wir eine konvergente Majorante (geometrische Reihe) gefunden.

Wenn die Reihe $\sum_{i=0}^{\infty} a_i$ konvergiert, dann gilt offensichlich nach dem Cauchyktriterium (mit $m := n+1$), daß $\lim_{i \to \infty} a_i = 0$. Die Umkehrung stimmt nicht.

Lemma 3.11 *Sei (a_i) eine monoton fallende Folge nichtnegativer reeller Zahlen. Die Reihe $\sum_{i=1}^{\infty} a_i$ konvergiert genau dann, wenn die Reihe $\sum_{i=1}^{\infty} 2^i a_{2^i}$ konvergiert.*

Beweis: Seien $s_n := \sum_{i=1}^{n} a_i$ und $t_n := \sum_{i=0}^{n} 2^i a_{2^i}$ die Partialsummen. Die Folgen (s_n) und (t_n) sind monoton wachsend. Es gilt für $n < 2^k$ daß

$$s_n \leq \sum_{i=1}^{k} \sum_{j=2^{i-1}}^{2^i-1} a_j \leq \sum_{i=1}^{k} a_{2^{i-1}} 2^{i-1} \leq t_k \ .$$

Folglich impliziert die Beschränktkeit der Folge (t_n) die Beschränktkeit der Folge (s_n). Umgekehrt gilt für $n > 2^k$

$$s_n \geq \sum_{i=1}^{k} \sum_{j=2^{i-1}}^{2^i-1} a_j \geq \frac{1}{2} \sum_{i=1}^{k} a_{2^i} 2^i = \frac{1}{2} t_k \ .$$

Die Beschränktkeit der Folge (s_n) impliziert die Beschränktkeit der Folge (t_n). ■

1. Die Folge $(\sum_{i=1}^{n} \frac{1}{i})_{n=1}^{\infty}$ ist unbeschränkt. In anderen Worten, die **harmonische Reihe** konvergiert nicht. In der Tat ist $\sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \sum_{k=1}^{\infty} 1$ sicher nicht konvergent.
2. Durch Umkekrung des Majorantenkriteriums (wir erhalten das **Minorantenkriterium**) kann man die Nichtkonverhenz gewisser Reihen zeigen. Wenn wir die Logarithmusfunktion $\ln : (0, \infty) \to \mathbb{R}$ schon kennen, dann können wir zeigen:

   $$\sum_{n=1}^{\infty} \frac{\ln(n)}{n}$$

   konvergiert nicht. In der Tat ist $\ln(n) > 1$ für $n > 1$ und damit $\frac{1}{n} \leq \frac{\ln(n)}{n}$.
3. Die Reihe

   $$\zeta(x) := \sum_{n=1}^{\infty} \frac{1}{n^x}$$

konvergiert dann und nur dann, wenn $x > 1$ ist. In der Tat ist

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^x} = \sum_{k=0}^{\infty} 2^{(1-x)k} .$$

Wenn $x > 1$, dann ist diese Reihe konvergent. Anderfalls konvergieren nicht einmal die Glieder gegen Null.

Wir betrachten wieder eine monoton fallende Folge nichtnegativer Zahlen (a_i) mit $\lim_{i\to\infty} a_i = 0$. Dann konvergiert die **alternierende Reihe**

$$\sum_{n=0}^{\infty} (-1)^n a_n .$$

In der Tat erfüllt die Folge der Partialsummen

$$s_{2n+1} \leq s_{2n+3} \leq s_{2n+2} \leq s_{2n} .$$

Die Folge (s_{2n+1}) ist monoton wachsend und von oben durch s_0 beschränkt, hat also eine Grenzwert s_*. Sie Folge (s_{2n}) ist monoton fallend und durch von unten durch s_1 beschränkt, hat also eine Grenzwert s^*. Da auf der anderen Seite die Folge der Zahlen $|s_{2n} - s_{2n+1}| = a_{2n+1}$ gegen Null konvergiert, gilt $s_* = s^*$.

Sind $\sum_{i=0}^{\infty} a_i$ und $\sum_{i=0}^{\infty} b_i$ konvergente Reihen, dann konvergiert auch $\sum_{i=0}^{\infty} a_i + b_i$ und es gilt

$$\sum_{i=0}^{\infty} a_i + b_i = \sum_{i=0}^{\infty} a_i + \sum_{i=0}^{\infty} b_i .$$

Das folgt unmittelbar aus der entsprechenden Eigenschaft für Folgen. Der Fall eines Produktes ist nicht so einfach. Das Problem ist, daß wir beim Ausmultiplizieren der Partialsummen eine Reihenfolge der Summanden wählen müssen.

In der Tat kann das Konvergenzverhalten eine Reihe aber von der **Reihenfolge der Summation** abhängen. Es ist klar, daß man endlich viele Summanden vertauschen darf. Aber in der Regel eben nicht unendlich viele. Sei (a_i) eine Folge und $j : \mathbb{N} \to \mathbb{N}$ eine Bijektion. Dann können wir eine neue Folge durch $a_{j(i)}$ bilden. Das Konvergenzverhalten von $\sum_{i=1}^{\infty} a_{j(i)}$ kann sich aber sehr von $\sum_{i=1}^{\infty} a_i$ unterscheiden. Dazu zeigen wir folgende Behauptung.

Wir nehmen an, daß $\sum_{i=1}^{\infty} a_i$ konvergiert, nicht jedoch $\sum_{i=1}^{\infty} |a_i|$. Sei $x \in \mathbb{R}$ gegeben. Dann existiert eine Bijektion $j : \mathbb{N} \to \mathbb{N}$ derart, daß

$$\sum_{i=1}^{\infty} a_{j(i)} = x$$

gilt. Wir definieren

$$\mathbb{N}^+ := \{n \in \mathbb{N} | a_n \geq 0\} , \quad \mathbb{N}^- := \{n \in \mathbb{N} | a_n < 0\} .$$

Beide Menge sind unendlich. Die Folgen $\sum_{i\in\mathbb{N}^\pm, i\leq n} |a_i|$ sind beide unbeschränkt. Wir definieren j induktiv. Angenommen, $j(i)$ ist für $i \leq n$ schon definiert. Wenn $\sum_{i=1}^n a_{j(i)} \geq x$ ist, dann wählen wir für $j(n+1)$ das kleinste noch nicht gewählte Element aus $\mathbb{N}^-$. Anderfalls nehmen wir für $j(n+1)$ das kleinste noch nicht gewählte Element aus $\mathbb{N}^+$. Die so konstruierte Abbildung ist injektiv. Wenn sie nicht surjektiv wäre, dann gäbe es ein $N \in \mathbb{N}$ derart, daß $j(n) \in \mathbb{N}^+$ für alle $n \geq N$ oder $j(n) \in \mathbb{N}^-$ für alle $n \geq N$. Im ersten Fall wäre jedoch $\sum_{i\in\mathbb{N}^+, i\leq n} |a_i|$ beschränt, im zweiten $\sum_{i\in\mathbb{N}^-, i\leq n} |a_i|$, was jedoch nicht der Fall ist. Es gilt für genügend große n, daß $|\sum_{i=1}^n a_{j(i)} - x| < a_{j(n)}$. Wegen $\lim_{i\to\infty} a_{j(i)} = 0$ folgt $\sum_{i=1}^\infty a_{j(i)} = x$. Man kann auch eine Umordnung finden, für welche die Folge der Partialsummen $\sum_{i=1}^n a_{j(i)}$ unbeschränkt ist.

Die Annahme, daß $\sum_{i=1}^\infty |a_i|$ nicht konvergiert, ist notwendig. In der Tat gilt der **Umordnungssatz.** Die Konvergenz von $\sum_{i=1}^\infty |a_i|$ impliziert die Konvergenz von $\sum_{i=1}^\infty a_i$ nach dem Majorantenkriterium.

Definition 3.12 *Eine Reihe $\sum_{i=1}^\infty a_i$ konvergiert **absolut**, wenn $\sum_{i=1}^\infty |a_i|$ konvergiert.*

Lemma 3.13 *Sei $j : \mathbb{N} \to \mathbb{N}$ eine Bijektion und $\sum_{i=1}^\infty a_i$ absolut konvergent. Dann konvergiert $\sum_{i=1}^\infty a_{j(i)}$ und es gilt*

$$\sum_{i=1}^\infty a_i = \sum_{i=1}^\infty a_{j(i)} \ .$$

Beweis: Sei $\epsilon > 0$ vorgegeben. Das Cauchykriterium zeigt die Existenz einer Zahl $N_0 \in \mathbb{N}$ derart, daß $\sum_{i=n}^m |a_i| < \epsilon$ für $n, m \geq N_0$ gilt. Es gibt jetzt ein $N_1 \in \mathbb{N}$ derart, daß $\{1, \ldots, N_0\} \subset j(\{1, \ldots, N_1\})$ gilt. Für $k, l \geq N_1$ gilt dann $\sum_{i=k}^l |a_{j(i)}| < \epsilon$. Wir schließen also mit den Cauchy Kriterium, daß $\sum_{i=1}^\infty a_{j(i)}$ konvergiert. Desweiteren ist für $n \geq \mathtt{max}\{N_1, N_0\}$ auch $|\sum_{i=1}^n a_i - \sum_{i=1}^n a_{j(i)}| < \epsilon$. Daraus folgt $|\sum_{i=1}^\infty a_i - \sum_{i=1}^\infty a_{j(i)}| \leq \epsilon$. Da ϵ beliebig klein gewählt werden kann müssen beide Reihen den gleichen Grenzwert haben. ■

Lemma 3.14 *Seien $a := \sum_{i=0}^\infty a_i$ und $b := \sum_{i=0}^\infty b_i$ konvergente Reihen. Wenn $\sum_{i=0}^\infty a_i$ sogar absolut konvergiert, dann konvergiert $c := \sum_{i=0}^\infty c_i$ mit $c_i = \sum_{j=0}^i a_j b_{i-j}$ und es gilt $ab = c$.*

Beweis: Seien $A_n := \sum_{i=0}^n a_i$, $B_n := \sum_{i=0}^n b_i$ und $C_n := \sum_{i=0}^n c_i$ die Partialsummen. Dann gilt mit $\beta_n := B_n - b$:

$$\begin{aligned} C_n &= a_0 B_n + a_1 B_{n-1} + \cdots + a_n B_0 \\ &= a_0(b + \beta_n) + a_1(b + \beta_{n-1}) + \cdots + a_n(b + \beta_0) \\ &= A_n b + \sum_{i=0}^n a_i \beta_{n-i} \ . \end{aligned}$$

Sei $\gamma_n := \sum_{i=0}^n a_i\beta_{n-i} = \sum_{i=0}^n a_{n-i}\beta_i$. Wir müssen zeigen, daß $\lim_{n\to\infty}\gamma_n = 0$. Sei $\alpha = \sum_{i=0}^\infty |a_i|$. Sei $\epsilon > 0$ gegeben. Wir wählen N_0 so groß, daß $|\beta_n| < \frac{\epsilon}{2\alpha+1}$ für alle $n > N_0$ gilt. Dann gilt für $n > N_0$ daß $|\gamma_n| \le \sum_{i=0}^{N_0} a_{n-i}\beta_i + \epsilon$. Da $\beta_n \to 0$ gilt, erhalten wir

$$\lim_{n\to\infty} |\gamma_n| \le \epsilon \ .$$

Da ϵ beliebig klein gewählt werden kann, folgt die Behauptung. ■

1. **Eigenschaften der e-Reihe.** Wir hatten gesehen, daß die Reihe $\mathsf{e}(x) := \sum_{n=0}^\infty \frac{x^n}{n!}$ für alle $x \in \mathbb{R}$ absolut konvergiert. Die Funktion

$$\mathbb{R} \ni x \mapsto \mathsf{e}(x) \in \mathbb{R}$$

heißt **e-Funktion.** Die gebräuchliche Schreibweise als Potenz $\mathsf{e}^x := \mathsf{e}(x)$ ist nicht zufällig. In der Tat gilt nämlich $\mathsf{e}(x) = (\mathsf{e}^1)^x$. Wir zeigen zunächst, daß für $x, y \in \mathbb{R}$

$$\mathsf{e}(x+y) = \mathsf{e}(x)\mathsf{e}(y)$$

gilt. Dazu rechnen wir

$$\begin{aligned} \mathsf{e}(x)\mathsf{e}(y) &= \sum_{n=0}^\infty \frac{x^n}{n!}\sum_{m=0}^\infty \frac{y^m}{m!} \\ &= \sum_{n=0}^\infty \sum_{k=0}^n \frac{x^k y^{n-k}}{(n-k)!k!} \\ &= \sum_{n=0}^\infty \frac{1}{n!}\sum_{k=0}^n x^k y^{n-k}\frac{n!}{k!(n-k)!} \\ &= \sum_{n=0}^\infty \frac{(x+y)^n}{n!} \\ &= \mathsf{e}(x+y) \end{aligned}$$

Wir beobachten, daß $\mathsf{e}(x) > 0$ für alle $x > 0$ gilt. Daraus folgt $\mathsf{e}(x) = \frac{1}{\mathsf{e}(-x)} > 0$ für negative x. Wir sehen, daß die e-Funktion eine Abbildung

$$\mathsf{e} : \mathbb{R} \to (0,\infty)$$

ist.

Nun gilt für $n \in \mathbb{N}$, daß $\mathsf{e}(\frac{1}{n})^n = \mathsf{e}(n\frac{1}{n}) = \mathsf{e}(1) = \mathsf{e}$ und deshalb $\mathsf{e}(\frac{1}{n}) = \mathsf{e}^{\frac{1}{n}}$. Für rationales $q = \frac{a}{b}$ gilt

$$\mathsf{e}^q = (\mathsf{e}^{\frac{1}{b}})^a = \mathsf{e}(\frac{1}{b})^a = \mathsf{e}(\frac{a}{b}) = \mathsf{e}(q) \ .$$

Diese Gleichung gilt auch für reelle q. Einen Beweis erbringen wir später.

2. Für $x \in \mathbb{R}$ betrachten wie die Reihen

$$\sin(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \ , \quad \cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \ .$$

Diese Reihen konvergieren absolut für alle $x \in \mathbb{R}$. Wir werden dies mit dem sogenannten **Quotientenkriterium** einsehen.

Lemma 3.15 (Quotientenkriterium) *Eine Reihe* $\sum_{i=0}^{\infty} a_i$ *konvergiert absolut, wenn*

$$\limsup_{i\to\infty} |\frac{a_i}{a_{i+1}}| < 1$$

gilt (die Existenz der linken Seite als Voraussetzung eingeschlossen).

In unserem Fall gilt für $\sin(x)$

$$|\frac{a_{i+1}}{a_i}| = \frac{x^2}{(2i+3)(2i+2)} \ .$$

Die Folge dieser Zahlen konvergiert gegen Null. Analog argumentiert man mit $\cos(x)$. Wir zeigen nun die Identität

$$\sin(x)^2 + \cos(x)^2 = 1 \ .$$

In der Tat gilt

$$\begin{aligned}
&\sin(x)^2 + \cos(x)^2 \\
&= (\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1})^2 + (\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n})^2 \\
&= \sum_{n=0}^{\infty} x^n \left[\sum_{2i+2j+2=n} \frac{(-1)^{i+j}}{(2i+1)!(2j+1)!} + \sum_{2i+2j=n} \frac{(-1)^{i+j}}{(2i)!(2j)!} \right] \\
&= \sum_{n=0,even}^{\infty} (-1)^{\frac{n}{2}} x^n \left[- \sum_{2i+2j+2=n} \frac{1}{(2i+1)!(2j+1)!} + \sum_{2i+2j=n} \frac{1}{(2i)!(2j)!} \right]
\end{aligned}$$

Wir schreiben für gerades $n > 1$

$$\begin{aligned}
(1-1)^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (-1)^k \\
&= \sum_{2i+2j=n} \frac{n!}{(2i)!(2j)!} - \sum_{2i+2j=n+2}^{n} \frac{n!}{(2i+1)!(2j+1)!} \ .
\end{aligned}$$

Wir schließen daraus, daß in der obigen Summe alle Summanden mit $n > 0$ verschwinden, also

$$\sin(x)^2 + \cos(x)^2 = 1 \ .$$

3. Für $|x| < 1$ definieren wir die Logarithmusfunktion

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n} .$$

   Diese Reihe konvergiert absolut. In der Tat ist $|\frac{x^n}{n}| \leq |x|^n$, und damit die geometrische Reihe $\sum_{n=1}^{\infty} |x|^n$ eine konvergente Majorante.

4. Wir zeigen, daß die Reihe

$$\sum_{n=1}^{\infty} \frac{n^5}{5^n}$$

   konvergent ist. Dazu verwenden wir das **Wurzelkriterium**.

   Lemma 3.16 (Wurzelkriterium) *Eine Reihe* $\sum_{n=1}^{\infty} a_n$ *konvergiert absolut, falls*

$$\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} < 1$$

   gilt.

   In der Tat ist

$$|\frac{n^5}{5^n}|^{\frac{1}{n}} = (n^{\frac{1}{n}})^5 \frac{1}{5} .$$

   Da $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$ gilt, folgt $\lim_{n\to\infty} |\frac{n^5}{5^n}|^{\frac{1}{n}} = \frac{1}{5}$.

Wir **zeigen** nun das **Quotientenkriterium.** Aus $\limsup_{i\to\infty} |\frac{a_{i+1}}{a_i}| < 1$ folgt die Existenz einer Zahl $0 \leq c < 1$ und eines $N \in \mathbb{N}$ derart, daß $|\frac{a_{i+1}}{a_i}| \leq c$ für alle $i \geq N$ gilt. Daraus folgt

$$a_{N+2} \leq a_N c^2 \ , a_{N+3} \leq a_N c^3 \ldots , a_{N+k} \leq a_N c^k \ .$$

Die geometrische Reihe

$$a_N \sum_{i=0}^{\infty} c^i$$

ist also eine konvergente Majorante der Reihe

$$\sum_{k=0}^{\infty} a_{N+k} \ .$$

Wir **zeigen** nun das **Wurzelkriterium**. Wenn $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} < 1$, dann gibt es ein $0 < q < 1$ und $n_0 \in \mathbb{N}$ mit $|a_n|^{\frac{1}{n}} < q$ für alle $n \geq n_0$. Daraus folgt $|a_n| < q^n$. Damit ist $\sum_{n=n_0}^{\infty} q^n$ eine konvergente Majorante von $\sum_{n=n_0}^{\infty} a_n$.

4 Stetigkeit

4.1 Motivierende Argumente

Wir sagen, daß eine Abbildung f mit **Grenzwerten vertauscht**, wenn

$$f(\lim_{n\to\infty} r_n) = \lim_{n\to\infty} f(r_n)$$

für jede konvergente Folge (r_n) gilt.

Sei $0 < x \in \mathbb{R}$ fest. Wir hatten die Potenzfunktion

$$\mathbb{Q} \ni r \mapsto x^r \in \mathbb{R}$$

definiert. Das folgende Lemma besagt, daß die Potenzfunktion $x \mapsto x^r$ mit Grenzwerten vertauscht.

Lemma 4.1 *Sei (r_n) eine konvergente Folge rationaler Zahlen mit $r := \lim_{n\to\infty} r_n \in \mathbb{Q}$. Dann gilt*

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

Beweis: Ohne Beschränkung der Allgemeinheit können wir annehmen, daß $x > 1$ ist. Wenn $x < 1$, dann ersetzen wir x durch x^{-1} und r_n durch $-r_n$. Der Fall $x = 1$ ist klar.

Wir betrachten nun den Fall, daß $r_n \to 0$ gilt. Sei $\epsilon > 0$ gegeben. Es gilt $x^{\frac{1}{n}} \to 1$ und $x^{-\frac{1}{n}} \to 1$. Dann gibt es ein $n_0 \in \mathbb{N}$ derart, daß $1 - \epsilon < x^{-\frac{1}{n_0}} < 1 < x^{\frac{1}{n_0}} < 1 + \epsilon$.

Dann finden wir ein $n_1 \in \mathbb{N}$ derart, daß $-\frac{1}{n_0} < r_n < \frac{1}{n_0}$ für alle $n \geq n_1$ gilt. Für diese n gilt (Lemma 2.19) $1 - \epsilon < x^{-\frac{1}{n_0}} < x^{r_n} < x^{\frac{1}{n_0}} < 1 + \epsilon$, also $|1 - x^{r_n}| < \epsilon$. Dies zeigt $x^{r_n} \to 1 = x^0$.

Im allgemeinen Fall gilt $r_n - r \to 0$. Wir schreiben $x^{r_n} = x^{r_n - r} x^r$ und schließen, daß $x^{r_n} \to x^r$ gilt. ■

Wir erklären nun x^r für $r \in \mathbb{R}$ und $x > 0$. In der Tat können wir uns wieder auf den Fall $x \geq 1$ einschränken. Für $r \in \mathbb{R}$ wählen wir nun eine monoton wachsende Folge (r_n) mit $r_n \to r$ und definieren

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

Dieser Grenzwert existiert. In der Tat ist (x^{r_n}) monoton wachsend und $x^{r_n} < x^q$ für alle n, wobei $q \in \mathbb{Q}$ eine Zahl mit $r < q$ ist. Eine beschränkte, monoton wachsende Folge hat einen Grenzwert (Lemma 3.2).

Wir müssen die Wohldefiniertheit, also die Unabhängigkeit von der Wahl der Folge zeigen. Sei (s_n) eine beliebige Folge mit $s_n \to r$. Dann gilt $s_n - r_n \to 0$. Wir haben $x^{s_n} = x^{s_n - r_n} x^{r_n} \to x^r$, da $x^{s_n - r_n} \to 1$ und $x^{r_n} \to x^r$.

Es gilt $x^{r+s} = x^r x^s$. In der Tat, seien $(r_n), (s_n)$ Folgen mit $r_n \to r$ und $s_n \to s$. Dann gilt $r_n + s_n \to r + s$. Wir haben

$$x^{r+s} = \lim_{n\to\infty} x^{r_n + s_n} = \lim_{n\to\infty} x^{r_n} x^{s_n} = \lim_{n\to\infty} x^{r_n} \lim_{n\to\infty} x^{s_n} = x^r x^s \ .$$

Desweiteren gilt für $x > 1$ und $s < r$, daß $x^s < x^r$. Wir wählen Folgen $(r_n), (s_n)$ wie oben, wobei wir annehmen können, daß $s_n \leq r_n$ ist. In der Tat können wir sonst r_n durch $\mathtt{max}\{r_n, s_n\}$ ersetzen. Dann gilt aber $x^{s_n} \leq x^{r_n}$, woraus $x^s \leq x^r$ folgt.

Die Abbildung $r \ni \mathbb{R} \mapsto x^r \in \mathbb{R}$ vertauscht mit Grenzwerten. In der Tat, sei (r_n) eine konvergente Folge in $\mathbb{R}$. Sei $m \in \mathbb{N}$ gegeben. Dann gibt es ein n_0 derart, daß $r - \frac{1}{m} < r_n < r + \frac{1}{m}$ für $n \geq n_0$ gilt. Für diese n gilt

$$x^{-\frac{1}{m}} x^r \leq x^{r_n} \leq x^{\frac{1}{m}} x^r \ .$$

Wegen $\lim_{m\to\infty} x^{-\frac{1}{m}} = 1$ folgt

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

Wir hatten die e-Funktion $\mathsf{e}(t) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ eingeführt und gezeigt, daß $\mathsf{e}(x) = \mathsf{e}^x$ für $x \in \mathbb{Q}$ gilt.

Lemma 4.2 *Wenn die* e*-Funktion mit Grenzwerten vertauscht, dann gilt* $\mathsf{e}(x) = \mathsf{e}^x$ *für alle* $x \in \mathbb{R}$.

Beweis: Sei $x \in \mathbb{R}$ und (x_n) eine rationale Folge mit $x_n \to x$. Dann gilt

$$\mathsf{e}(x) = \mathsf{e}(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \mathsf{e}(x_n) = \lim_{n\to\infty} \mathsf{e}^{x_n} = \mathsf{e}^{\lim_{n\to\infty} x_n} = \mathsf{e}^x \ .$$

■

Lemma 4.3 *Die* e*-Funktion vertauscht mit Grenzwerten.*

Beweis: Wir überlegen zuerst, daß $\mathsf{e}(x)$ monoton wachsend ist. Zunächst sehen wir aus der Reihendarstellung, daß aus $r \geq 0$ auch $\mathsf{e}(r) \geq 1$ folgt. Nun folgt für $x \leq y$, daß $\mathsf{e}(y) - \mathsf{e}(x) = \mathsf{e}(x)(\mathsf{e}(y-x) - 1) \geq 0$.

Sei jetzt (r_n) eine beliebige Folge mit $r_n \to n$. Sei $m \in \mathbb{N}$ gegeben. Dann gibt es ein $n_0 \in \mathbb{N}$ derart, daß $r - \frac{1}{m} \leq r_n \leq r + \frac{1}{m}$ für alle $n \geq n_0$ ist. Wir schließen, daß

$$\mathsf{e}^{-\frac{1}{m}} \mathsf{e}(r) \leq \mathsf{e}(r_n) \leq \mathsf{e}(r) \mathsf{e}^{\frac{1}{m}} \ .$$

Wegen $\mathsf{e}^{\frac{1}{m}} \to 1$ schließen wir, daß

$$\lim_{n\to\infty} \mathsf{e}(r_n) = \mathsf{e}(r) = \mathsf{e}(\lim_{n\to\infty} r_n) \ .$$

■

4.2 Stetige reell-wertige Funktionen auf $\mathbb{R}$

Definition 4.4 *Eine Teilmenge $U \subseteq \mathbb{R}$ heißt **offen**, wenn für jedes $x \in U$ ein $\epsilon > 0$ existiert, so daß $(x-\epsilon, x+\epsilon) \subset U$. Eine Teilmenge $V \subseteq \mathbb{R}$ heißt **abgeschlossen**, wenn ihr Komplement $\mathbb{R} \setminus V$ offen ist.*

Offene Intervalle (a, b), (a, ∞) sind offen. Abgeschlossene Intervalle $[a, b]$ sind abgeschlossen. Intervalle $[a, b)$ (für $b \neq \infty$) sind weder offen noch abgeschlossen.

Lemma 4.5 *Die Vereinigung einer beliebigen Familie $(U_\alpha)_{\alpha \in I}$ offener Teilmengen ist offen.*

Beweis: Sei $x \in \cup_{\alpha \in I} U_\alpha$. Dann existiert ein $\alpha \in I$ mit $x \in U_\alpha$. Wir finden $\epsilon > 0$, so daß $(x-\epsilon, x+\epsilon) \subset U_\alpha$. Dann gilt auch $(x-\epsilon, x+\epsilon) \subset \cup_{\alpha \in I} U_\alpha$. ■

Eine etwas exotischere abgeschlossene Teilmenge des $\mathbb{R}$ ist die **Cantormenge** C_α für $\alpha \in (0, 1/2)$. Sei $I := [a, b]$ ein Intervall. Dann setzen wir

$$G_\alpha(I) := [a, b] \setminus ((a + \alpha(b-a), b - \alpha(b-a)) \ .$$

Ist $A \subset \mathbb{R}$ eine endliche disjunkte Vereinigung von Intervallen $A = \sqcup_k I_k$, dann bilden wir $G_\alpha(A) := \sqcup_k G_\alpha(I_k)$. Dies ist wieder eine endliche (in der Tat verdoppelte Zahl) disjunkte Vereinigung von Intervallen und es gilt $G_\alpha(A) \subseteq A$. Wir starten nun mit dem Einheitsintervall $I := [0, 1]$ und bilden

$$C_\alpha := \bigcap_{k \geq 0} G_\alpha^k(I) \ .$$

Die Cantormenge C_α is abgeschlossen, da

$$\mathbb{R} \setminus C_\alpha = \bigcup_{k \geq 0} (\mathbb{R} \setminus G_\alpha^k(I))$$

eine Vereinigung offener Mengen ist. Die Länge der Intervalle in $G^k(I)$ ist α^k.

Definition 4.6 *Eine Umgebung von $x \in \mathbb{R}$ ist eine Menge $M \subset \mathbb{R}$, für welche es eine offene Teilmenge $U \subset \mathbb{R}$ mit $x \in U \subseteq M$ gibt.*

Die Menge $(-1, 1]$ ist eine Umgebung aller Punkte von $(-1, 1)$, nicht aber von 1.

Für jeden Punkt $x \in C_\alpha$ ist C_α keine Umgebung. Sei $\epsilon > 0$ gegeben. Dann gibt es ein $k > 0$ derart, daß $\alpha^k < \epsilon$. Weiterhin gibt es ein Intervall $[u, v] \in U \subset G_\alpha^k(I)$, welches x enthält. Dann ist $(x-\epsilon, x+\epsilon) \nsubseteq G_\alpha(U)$ und damit $(x-\epsilon, x+\epsilon) \nsubseteq C_\alpha$. Es gilt etwa $\frac{u+v}{2} \notin G_\alpha(U)$, aber $|x - \frac{u+v}{2}| < \epsilon$. .

Sei $A \subset \mathbb{R}$ eine beliebige Teilmenge.

Definition 4.7 *Eine Teilmenge $V \subseteq A$ ist offen in A, wenn sie von der Form $V = A \cap U$ für eine offene Teilmenge von $\mathbb{R}$ ist. Eine Teilmenge $B \subseteq A$ ist abgeschlossen in A, wenn $A \setminus B$ offen in A ist. Sei $x \in A$. Eine Teilmenge $M \subseteq A$ ist eine Umgebung von x in A, wenn es eine Umgebung N von $x \in \mathbb{R}$ mit $M = A \cap N$ gibt.*

1. Die Menge A ist offen und abgeschlossen in A.
2. Eine beliebige Vereinigung offener Teilmengen von A ist wieder offen in A.
3. Zum Beispiel ist $(-1,1) \cap \mathbb{Q}$ eine Umgebung von 0 in $\mathbb{Q}$.
4. Die Punkte $n \in \mathbb{Z}$ sind offen und abgeschlossen in $\mathbb{Z}$.
5. Die Menge $(0,1]$ ist eine Umgebung von 1 in $(-1,1]$.

Sei $A \subset \mathbb{R}$ und $f : A \to \mathbb{R}$ eine auf A definierte Funktion. Sei $x \in A$.

Definition 4.8 (ϵ-δ-Definition) *Eine Funktion $f : A \to \mathbb{R}$ heißt (metrisch) stetig in x, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so daß aus $y \in A$ und $|y-x| < \delta$ die Ungleichung $|f(x) - f(y)| < \epsilon$ folgt.*

1. Die kanonische Einbettung $A \to \mathbb{R}$ ist (metrisch) stetig in jedem $x \in A$. In der Tat kann man für gegebenes $\epsilon > 0$ $\delta := \epsilon$ wählen.
2. Die Funktion $f(x) := x^2$ is in jedem Punkt $x_0 \in \mathbb{R}$ (metrisch) stetig. Sei $\delta > 0$ gegeben. Dann wählen wir $\delta := \mathtt{min}\{\frac{\epsilon}{3+3|x_0|}, 1\}$. Es gilt für $|x - x_0| < \delta$, daß
$$\begin{aligned} |x^2 - x_0^2| &= |x^2 - xx_0 + xx_0 - x_0^2| \\ &\leq |x^2 - xx_0| + |xx_0 - x_0^2| \\ &\leq |x||x - x_0| + |x_0||x - x_0| \\ &\leq \frac{\epsilon(|x_0| + |x - x_0|)}{3 + 3|x_0|} + \frac{\epsilon|x_0|}{3 + 3|x_0|} \\ &\leq \epsilon \, . \end{aligned}$$
3. Ein Punkt $x \in A$ heißt **isoliert**, wenn es eine Umgebung von x in A gibt, welche keine weiteren Punkte von A enthält. Zum Beispiel besteht $\mathbb{Z} \subset \mathbb{R}$ aus isolierten Punkten. Ist $x \in A$ isoliert, so ist eine beliebige Funktion $f : A \to \mathbb{R}$ in x stetig. In der Tat, sei $\epsilon > 0$ gegeben. Wegen der Isoliertheit von x können wir $\delta > 0$ derart wählen, daß $(x - \delta, x + \delta) \cap A = \{x\}$ gilt. Dann gilt $|f(x) - f(y)| < \epsilon$ für alle $y \in A$ mit $|x - y| < \delta$, da nur $y := x$ diese Bedingung erfüllt.
4. Die **Cantormenge** C_α besteht hat **keine isolierten Punkte**. Sei $[u,v] \subset G^k_\alpha(I)$. Dann sind u, v auch Anfangs und Endpunkte von Intervallen in $G^{k+1}(I)$. Insbesondere gilt $u, v \in C_\alpha$. Sei $x \in C_\alpha$ und $\epsilon > 0$ gegeben. Dann gibt es ein $k > 0$ derart, daß $\alpha^k < \epsilon$. Nun gibt es ein Intervall $[u,v] \in G^k_\alpha(I)$, welches x enthält. Dann gilt $u, v \in (x - \epsilon, x + \epsilon) \cap C_\alpha$. Zumindest einer dieser beiden Punkte muß von x verschieden sein.

Definition 4.9 (Folgenstetigkeit) *Eine Funktion $f : A \to \mathbb{R}$ heißt **folgenstetig** in $x \in A$, wenn für jede Folge (x_n) mit $x_n \in A$ und $\lim_{n\to\infty} x_n = x$ auch $(f(x_n))$ konvergiert und $\lim_{n\to\infty} f(x_n) = f(x)$ gilt.* , Die Funktion vertauscht also mit Grenzwerten bei x. Für diese Definition muß man nur Konvergenzbegriffe für Folgen in A und $\mathbb{R}$ haben.

Definition 4.10 (Topologische Definition) *Eine Funktion $f : A \to \mathbb{R}$ heißt (topologisch) stetig in $x \in A$, wenn für jede Umgebung $N \subseteq \mathbb{R}$ von $f(x)$ das Urbild $f^{-1}(N) \subseteq A$ eine Umgebung von x in A ist.*

Diese Definition benutzt die Sprache der Umgebungen.

Satz 4.11 *Die folgenden Aussagen sind äquivalent.*

1. *f ist in x (metrisch) stetig.*
2. *f ist in x folgenstetig.*
3. *f ist in x (topologisch) stetig*

Beweis: Sei f (metrisch) stetig und (x_n) eine Folge in A mit $x_n \to x$. Sei $\epsilon > 0$ gegeben. Dann wählen wir $\delta > 0$ derart, daß aus $|y - x| < \delta$ folgt $|f(y) - f(x)| < \epsilon$. Nun wählen wir $n_0 \in \mathbb{N}$ derart, daß aus $n \geq n_0$ folgt $|x_n - x| < \delta$. Nun gilt für $n \geq n_0$ auch $|f(x_n) - f(x)| < \epsilon$.

Sei f folgenstetig und $N \subset \mathbb{R}$ eine Umgebung von $f(x)$. Wir nehmen an, daß $f^{-1}(N)$ keine Umgebung von x ist. Dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in (x - \frac{1}{n}, x + \frac{1}{n}) \cap A$ mit $f(x_n \notin N)$. Folglich gilt $x_n \to x$, nicht aber $f(x_n) \to f(x)$. Dies ist aber ein Widerspruch zur Folgenstetigkeit von f in x.

Sei nun f in x (topologisch) stetig und $\epsilon > 0$ gegeben. Dann ist $N := (f(x) - \epsilon, f(x) + \epsilon)$ eine Umgebung von $f(x)$. Folglich ist $f^{-1}(N)$ eine Umgebung von x. Also gibt es ein $\delta > 0$ derart, daß $(x - \delta, x + \delta) \cap A \subset f^{-1}(N)$. Damit folgt aus $|y - x| < \delta$, daß $|f(y) - f(x)| < \epsilon$. ■

In Anbetracht dieses Satzes werden wir für reellwertige Funktionen auf Teilmengen des $\mathbb{R}$ nicht zwischen den verschiedenen Definitionen unterscheiden.

Sei $A \subset \mathbb{R}$.

Definition 4.12 *Eine Funktion $f : A \to \mathbb{R}$ heißt stetig, wenn sie in allen Punkten von A stetig ist.*

1. Die Einbettung $i : A \to \mathbb{R}$ ist stetig. Sei $x \in A$ und M eine Umgebung von $i(x) = x$. Dann ist $i^{-1}(M) = M \cap A$ nach Definition eine Umgebung von x in A. Die Einbettung ist also stetig nach der topologischen Definition.

2. Die Summe und das Produkt in x stetiger Funktionen ist stetig in x. Seien $f, g : A \to \mathbb{R}$ stetig und (x_n) eine Folge in A mit $x_n \to x$. Dann gilt

   $$\lim_{n\to\infty} f(x_n)g(x_n) = \lim_{n\to\infty} f(x_n) \lim_{n\to g(x_n)} = f(x)g(x) \ .$$

   Das Produkt folgenstetiger Funktionen ist also folgenstetig. Für die Summe argumentiert man genauso.

3. Sei $g(x)$ stetig in x und $0 \notin g(A)$. Dann ist $\frac{1}{g}$ stetig in x. In der Tat gilt für jedes Folge (x_n) in A mit $x_n \to x$, daß $\frac{1}{g(x_n)} \to \frac{1}{g(x)}$.
4. Durch Kombination dieser Aussagen sieht man, daß Polynome stetig sind.
5. Eine rationale Funktion $f : A \to \mathbb{R}$ ist eine Funktion der Form $f(x) = \frac{p(x)}{q(x)}$ mit $0 \notin g(A)$. Eine rationale Funktion ist stetig.
6. Wir hatten schon gesehen, daß für festes $x > 0$ die Potenzen $\mathbb{R} \ni r \mapsto x^r \in \mathbb{R}$ stetig sind.

Wir diskutieren nun **Unstetigkeiten**. Sei $A \subseteq \mathbb{R}$ und $f : A \to \mathbb{R}$. Wir sagen, daß f in $x \in A$ **unstetig** ist, wenn f in x nicht stetig ist.

1. Die Funktion $\theta : \mathbb{R} \to \mathbb{R}$,
$$\theta(x) := \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$
ist in 0 unstetig. In der Tat ist $(-1/2, 1/2)$ eine Umgebung von $0 = f(0)$. Das Urbild $f^{-1}((-1/2, 1/2)) = (-\infty, 0]$ ist aber keine Umgebung von 0.
2. Die Funktion $f : \mathbb{R} \to \mathbb{R}$
$$f(x) := \begin{cases} 0 & x \leq 0 \\ 0 & x > 0 \wedge x \in \mathbb{R} \setminus \mathbb{Q} \\ \frac{1}{b} & x = \frac{a}{b} \text{ mit } a, b \in \mathbb{N} \text{ teilerfremd} \end{cases}$$
ist genau in den positiven rationalen Zahlen unstetig. In der Tat ist f an allen $x \leq 0$ stetig.

   Wenn x irrational und $(x_n) \to x$ eine Folge mit $x_n \to x$ ist, dann gilt $f(x_n) \to 0$. In der Tat ist sogar $f(x_n) = 0$ für irrationale Folgenglieder. Die rationalen schreiben wir in der Form $x_n = \frac{a_n}{b_n}$ mit teilerfremden natürlichen Zahlen a_n, b_n. Dann muß $b_n \to \infty$ gelten. Folglich gilt $f(x_{n_j}) \to 0$ für die Teilfolge (x_{n_j}) der rationalen Glieder.

   Wenn jedoch x rational ist, dann gilt $f(x) \neq 0$. Wir finden jedoch eine Folge irrationaler Zahlen (x_n) mit $x_n \to x$. Es gilt $f(x_n) \to 0 \neq f(x)$.
3. Die Funktion $f : \mathbb{R} \to \mathbb{R}$,
$$f(x) := \begin{cases} 0 & x \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases}$$
ist überall unstetig. Sei $x \in \mathbb{R}$. Wenn $x \in \mathbb{Q}$ ist, dann wählen wir eine Folge (x_n) irrationaler Zahlen mit $x_n \to x$. Wenn $x \in \mathbb{R} \setminus \mathbb{Q}$, dann wählen wir eine Folge (x_n) rationaler Zahlen mit $x_n \to x$. In jedem Fall gilt $\lim_{n\to\infty} f(x_n) \neq f(x)$.

4. Sei $f : (a,b) \to \mathbb{R}$ gegeben und $x \in (a,b)$. Wir nehmen an, daß $f_{|(a,b)\setminus\{x\}}$ stetig ist. Dann können unter anderem folgende Fälle auftreten.

   (a) $f_{|[x,b)}$ ist stetig. Man spricht von einer **Sprungstelle**.

   (b) $f_{|(a,x]}$ ist stetig. Man spricht von einer **Sprungstelle**.

   (c) Es gibt eine stetige Funktion $\tilde{f} : (a,b) \to \mathbb{R}$ mit $\tilde{f}_{|(a,b)\setminus\{x\}} = f_{|(a,b)\setminus\{x\}}$. In diesem Fall spricht man von einer **hebbaren Unstetigkeit**.

   (d) Die Funktion
$$f(x) := \begin{cases} x^{-1} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
   ist auf $\mathbb{R} \setminus \{0\}$ stetig, hat aber weder eine Sprungstelle noch eine hebbare Unstetigkeit.

4.3 Zwischenwerte und Extremwerte

Wir betrachten eine stetige Funktion $f : [a,b] \to \mathbb{R}$ mit $f(a) \leq 0 \leq f(b)$.

Lemma 4.13 (Zwischenwertsatz) *Es gibt ein $\xi \in [a,b]$ mit $f(\xi) = 0$.*

Beweis: Sei $U := \{x \in [a,b] | f(x) \leq 0\}$. Dann ist $a \in U$, also U nicht leer. Folglich ist $\xi := \sup U \in [a,b]$ definiert. Wir werden zeigen, daß $f(\xi) = 0$ gilt.

Es gibt (Lemma 3.3) eine Folge (x_n) mit $x_n \in U$ und $x_n \to \xi$. Wegen der Stetigkeit von f gilt $f(x_n) \to f(\xi)$. Da $f(x_n) \leq 0$ für alle n ist, sehen wir $f(\xi) \leq 0$. Wäre nun $f(\xi) < 0$, dann wäre $f(x) < 0$ für alle x in einer Umgebung von ξ. Da in diesem Fall $\xi \neq b$ sein muß, gibt es $x \in (\xi, b]$ mit $f(x) < 0$, also $x \in U$, was aber im Widerspruch zur Definition von ξ steht. ■

Die Vorausssetzung der Stetigkeit ist notwendig, wie das Beispiel der Funktion
$$[-1,1] \mapsto \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases}$$
zeigt.

Korollar 4.14 (Zwischenwertsatz) *Sei $f : [a,b] \to \mathbb{R}$ stetig und $x \in \mathbb{R}$ zwischen $f(a)$ und $f(b)$ (also $x \in [f(a), f(b)] \cup [f(b), f(a)]$. Dann gibt es ein $\xi \in [a,b]$ mit $f(\xi) = x$.*

Beweis: Wir betrachten $g(\xi) = f(\xi) - x$ und wenden darauf (oder auf $-g$) den Zwischenwertsatz an. ■

Es ist offensichtlich, daß der Zwischenwertsatz für unstetige Funktionen nicht gelten muß.

Lemma 4.15 *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist $f([a,b])$ beschränkt. Weiterhin gibt es $\alpha, \omega \in [a,b]$ mit*

$$f(\alpha) = \inf f([a,b]) \ , \quad f(\omega) = \sup f([a,b]) \ .$$

Beweis: Die Menge $f([a,b])$ ist von oben beschränkt. Andernfalls gäbe es eine Folge (x_n), $x_n \in [a,b]$, mit $f(x_n) > n$. Diese Folge hat eine konvergente Teilfolge (x_{n_i}) mit $\lim_{i\to\infty} x_{n_i} = \xi$. Dann gilt $\lim_{i\to\infty} f(x_{n_i}) = f(\xi)$, insbesondere ist die Folge $(f(x_{n_i}))_{i\in\mathbb{N}}$ also beschränkt, ein Widerspruch.

Wir beweisen nur die zweite Gleichung. Sei $M = \sup f([a,b])$. Dann gibt es eine Folge (x_n) mit $x_n \in [a,b]$ derart, daß $f(x_n) \to M$ gilt. Diese Folge ist beschränkt und hat deshalb eine konvergente Teilfolge.

Indem wir die Folge durch diese Teilfolge ersetzen, können wir annehmen, daß $x_n \to \omega$ für ein $\omega \in [a,b]$. Dann gilt $f(\omega) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = M$. ■

Wir bemerken, daß dieser Beweis die spezielle Eigenschaft abgeschlossener Intervalle benutzt, daß beschränkte Folgen konvergente Teilfolgen besitzen.

Lemma 4.16 *Sei $[a,b] \subset \mathbb{R}$ und $f : [a,b] \to \mathbb{R}$ streng monoton wachsend. Dann sind folgende Aussagen äquivalent:*

1. *$f : [a,b] \to [f(a), f(b)]$ ist bijektiv.*
2. *f ist stetig.*

In beiden Fällen ist $f^{-1} : [f(a), f(b)] \to [a,b]$ auch streng monoton und stetig.

Beweis: Sei f bijektiv und in $x \in [a,b]$ nicht stetig. Sei (x_n) eine Folge mit $x_n \to x$ derart, daß $f(x_n)$ keinen Grenzwert hat. Dann gibt es Zahlen $c, d \in [f(a), f(b)]$ mit $\liminf f(x_n) < c < d < \limsup f(x_n)$. Nun ist $c = f(y)$ und $d = f(z)$ für $y, z \in [a,b]$ und es gibt beliebig große n, m derart daß $x_n < y < z < x_m$. Das ist aber wegen $x_n \to x$ unmöglich.

Wenn f stetig ist, dann ist f surjektiv nach dem Zwischenwertsatz. Wegen der Monotonie ist f auch injektiv.

In beiden Fällen existiert f^{-1} und ist streng monoton. Da f^{-1} auch bijektiv ist, ist diese Funktion stetig. ■

4.4 Abstand, metrische und topologische Räume

Wir betrachten jetzt eine Menge X. Um sagen zu können, daß eine Folge (x_n) gegen einen Punkt x konvergiert, brauchen wir eine Methode, den Abstand von x_n zu x zu messen, oder wenigstens eine Möglichkeit, von Umgebungen von x zu reden.

Definition 4.17 *Eine **Metrik** (oder Abstandsfunktion) auf X ist eine Funktion $d : X \times X \to [0, \infty)$ mit folgenden Eigenschaften:*

1. $d(x,y) = 0 \Leftrightarrow x = y$,

2. $d(x,y) = d(y,x)$,

3. $d(x,y) \leq d(x,z) + d(z,y)$,

wobei x, y, z *beliebige Punkte in* X *bezeichnen.*

Ein **metrischer Raum** ist ein Paar (X,d) aus eine Menge mit einer Abstandsfunktion.
Sei (X,d) ein metrischer Raum und $x \in X$.

Definition 4.18 *Eine Folge* (x_n) *in* X *konvergiert gegen* $x \in X$*, wenn* $d(x_n, x) \to 0$ *gilt.*

Wir können auch den Begriff einer Cauchyfolge verallgemeinern.

Definition 4.19 *Eine Folge* (x_n) *in* X *ist eine Cauchyfolge, wenn für jedes* $\epsilon > 0$ *ein* $n_0 \in \mathbb{N}$ *existiert, so daß für* $n, m \geq n_0$ *gilt* $d(x_n, x_m) < \epsilon$*.*

Hier sind einige Beispiele metrischer Räume:

1. Die Funktion $d : \mathbb{R} \times \mathbb{R} \to [0, \infty)$, $d(x,y) := |x - y|$ ist eine Abstandsfunktion. Mit dieser Abstandsfunktion ist die obige Definition der Konvergenz die uns schon bekannte.

2. Ist X ein metrischer Raum und $A \subseteq X$ ein Teilmenge, dann ist A mit dem eingeschränkten Abstand $d_{|A \times A} : A \times A \to [0, \infty)$ auch ein metrischer Raum.

3. Jede Teilmenge $A \subseteq \mathbb{R}$ ist also in natürlicher Weise ein metrischer Raum.

4. Eine konvergente Folge in einem metrischen Raum ist immer ein Cauchyfolge. Die Umkehrung gilt im allgemeinen nicht. Die Folge $(\frac{1}{n})$ ist eine Cauchyfolge in $(0, \infty)$, hat aber keinen Grenzwert, da $0 \notin (0, \infty)$.

   Definition 4.20 *Wir nennen einen metrischen Raum* ***vollständig****, wenn in ihm jede Cauchyfolge konvergiert.*

5. Sei X eine beliebige Menge. Die Funktion $d_{disc} : X \times X \to [0, \infty)$,

$$d(x,y) := \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

   ist eine Abstandsfunktion, der **diskrete Abstand**. In dem Raum (X, d_{disc}) konvergiert eine Folge genau dann gegen x, wenn $x_n = x$ für alle genügend große n ist.

6. Sei B eine endliche Menge und (T, d_T) ein metrischer Raum. Dann definieren wir auf $X := T^B$ eine Abstandsfunktion durch

$$d(f,g) = \mathtt{max}_{b\in B} d_T(f(b), g(b)) \ .$$

Wir müssen die Dreiecksungleichung zeigen. In der Tat gilt für $f, g, h \in T^B$

$$\begin{aligned} d(f,g) &= \mathtt{max}_{b\in B} d_T(f(b), g(b)) \\ &\leq \mathtt{max}_{b\in B}\{d_T(f(b), h(b)) + d_T(h(b), g(b))\} \\ &\leq \mathtt{max}_{b\in B} d_T(f(b), h(b)) + \mathtt{max}_{b\in B} d_T(f(b), h(b)) \\ &= d(f,h) + d(h,g) \ . \end{aligned}$$

Eine Folge (f_n) in (X, d) konvergiert genau dann, wenn $(f_n(b))$ für alle $b \in B$ konvergiert. In der Tat, wenn $f_n \to f$, dann gilt $\mathtt{max}_{b\in B} d_T(f_n(b), f(b)) \to 0$, also $d_T(f_n(b), f(b)) \to 0$ für alle $b \in B$.

Möge umgekehrt $d_T(f_n(b), f(b)) \to 0$ für alle $b \in B$ gelten. Für $\epsilon > 0$ finden wir eine Familie $(n_b)_{b\in B} \in \mathbb{N}$ derart, daß für alle $b \in B$ und $n \geq n_b$ gilt: $d_T(f_{b_n}(b), f(b)) < \epsilon$. Wir setzen $n_0 := \mathtt{max}\{n_b | b \in B\}$. An dieser Stelle ist die Endlichkeit von B wichtig. Dann gilt für alle $n \geq n_0$ und alle $b \in B$ auch $d_T(f_n(b), f(b)) < \epsilon$, also $d(f_n, f) = \mathtt{max}_{b\in B} d_T(f_n(b), f(b)) < \epsilon$.

7. Ein besonders wichtiges Beispiel ist der Raum $\mathbb{R}^n$ mit der Metrik

$$d_\infty(x,y) := \mathtt{max}\{|x_i - y_i| i \in \{1, \ldots, n\}\} \ .$$

8. Durch $d_{eukl}(x,y) := \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ wird die Euklidische Metrik auf $\mathbb{R}^n$ definiert. Wir müssen die Dreiecksungleichung nachweisen. Wir setzen für $x, y \in \mathbb{R}^n$

$$< x, y >:= \sum_{i=1}^n x_i y_i \qquad \|x\| :=< x, x >^{1/2} \quad .$$

Dann gilt $d_{eukl}(x,y) = \|x - y\|$. Wir zeigen zunächst die Cauchy-Schwarzsche Ungleichung

$$| < x, y > | \leq \|x\| \|y\| \ .$$

Der Fall $x = 0$ ist klar. Sei nun $x \neq 0$. Dann setzen wir $u := \frac{<x,y>}{<x,x>}$. Es gilt

$$\begin{aligned} 0 &\leq < y - ux, y - ux > \\ &= < y, y > -2u < x, y > +u^2 < x, x > \\ &= < y, y > - \frac{< x, y >^2}{< x, y >} \ . \end{aligned}$$

Daraus folgt

$$< x, y >^2 \leq < x, x >< y, y > \quad .$$

Wir schließen nun

$$\|x+y\|^2 = \|x\|^2 + 2 < x, y > + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \leq (\|x\| + \|y\|)^2 \ .$$

Daraus folgt die Dreiecksungleichung

$$d_{eukl}(x,y) = \|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\| = d_{eukl}(x,z) + d_{eukl}(z,y) \ .$$

9. Sei $C([a,b])$ die Menge der stetigen reellewertigen Funktionen auf einem beschränkten Intervall $[a,b] \subset \mathbb{R}$. Wir definieren für $f \in C([a,b])$ die **Norm** $\|f\| := \sup_{x\in[a,b]} |f(x)|$ (beachte, daß f beschränkt ist und somit das Supremum existiert) und setzen $d(f,g) := \|f-g\|$. Man sagt, daß eine Folge von Funktion (f_n) in $C([a,b])$ **gleichmäßig** gegen $f \in C([a,b])$ konvergiert, wenn sie im Raum $(C([a,b]), d)$ konvergiert. Wir müssen wieder die Dreiecksungleichung zeigen. Es gilt

$$\begin{aligned} d(f,g) &= \sup_{x\in[a,b]} |f(x) - g(x)| \\ &\leq \sup_{x\in[a,b]} \{|f(x) - h(x)| + |h(x) - g(x)|\} \\ &\leq \sup_{x\in[a,b]} |f(x) - h(x)| + \sup_{x\in[a,b]} |f(x) - h(x)| \\ &= d(f,h) + d(h,g) \ . \end{aligned}$$

10. Sei X eine Menge. Wir betrachten die Menge $X^{\mathbb{N}}$ der Folgen in X und definieren

$$d((x_n),(y_n)) := \sum_{i=1}^{\infty} 2^{-n} d_{disc}(x_n, y_n) \ .$$

Die Summe konvergiert immer, da $\sum_{n=1}^{\infty} 2^{-n}$ eine konvergente Majorante ist. Symmetrie und Nichtnegativität von d sind klar. Es gilt $d((x_n),(y_n))$ genau dann, wenn die Folgen gleich sind. Wir zeigen die Dreiecksungleichgung.

$$\begin{aligned} d((x_n),(y_n)) &= \sum_{i=1}^{\infty} 2^{-n} d_{disc}(x_n, y_n) \\ &\leq \sum_{i=1}^{\infty} 2^{-n} d_{disc}(x_n, z_n) + d_{disc}(z_n, y_n) \\ &= \sum_{i=1}^{\infty} 2^{-n} d_{disc}(x_n, z_n) + \sum_{i=1}^{\infty} 2^{-n} d_{disc}(z_n, y_n) \\ &= d_{disc}((x_n),(z_n)) + d_{disc}((z_n),(y_n)) \ . \end{aligned}$$

Wir analysieren, was $\lim_{i\to\infty}(x_{i,n}) = (x_n)$ bedeutet. Sei $\epsilon > 0$ gegeben. Damit $d((x_{i,n}),(x_n)) < \epsilon$ ist, muß $x_{i,n} = x_n$ für alle i mit $2^{-i} \geq \epsilon$ gelten. Sei i_0 derart, daß $\sum_{i=i_0}^{\infty} 2^{-i} < \epsilon$ gilt. Wenn $x_{i,n} = x_n$ für $i < i_0$ gilt, dann können die Folgenglieder $x_{i,n}$ mit $i \geq i_0$ beliebig sein. Für jedes $n \in \mathbb{N}$ muß insbesondere $x_{i,n} = x_n$ für genügend große i gelten, das ist aber nicht hinreichend.

Sei (X,d) ein metrischer Raum, $x \in X$ und $r \in \mathbb{R}$.

Definition 4.21 *Mit $B(x,r) := \{y \in X | d(x,y) < r\}$ bezeichnen wir den **Ball** um x mit dem Radius r.*

1. In $\mathbb{R}$ mit dem Abstand $d(x,y) = |x-y|$ ist $B(x,r) = (x-t, x+r)$.
2. In $(\mathbb{R}, d_{eukl})$ ist
$$B(x,r) = \{y \in \mathbb{R}^n | \sum_{i=1}^{n} (x_i - y_i)^2 < r^2\}$$
eine geometrische Kugel.
3. In (X, d_{disc}) ist $B(x,r)$ entweder gleich X für $r > 1$ oder $\{x\}$ für $r \leq 1$.

Aus der Dreiecksungleichung folgt: Wenn $y \in B(x,r)$ und $s + d(x,y) < r$, dann gilt $B(y,s) \subseteq B(x,r)$.

Sei (X,d) ein metrischer Raum.

Definition 4.22 *Wir nennen eine Menge $U \subseteq X$ **offen**, wenn für $x \in U$ ein $r > 0$ existiert mit $B(x,r) \subseteq U$. Eine Menge $A \subseteq X$ heißt **abgeschlossen**, wenn ihr Komplement $X \setminus A$ offen ist. Die Menge $\mathcal{T} := \{U \in \mathcal{P}(X) | U \text{ ist offen}\}$ heißt die von d auf X induzierte **Topologie**.*

1. X und $\emptyset$ sind offen und abgeschlossen.
2. Die Bälle $B(x,r)$ sind offen.
3. In $\mathbb{R}$ mit dem Abstand $d(x,y) = |x-y|$ sind offene Intervalle offen.
4. (X, d_{disc}) sind alle Mengen offen und abgeschlossen.
5. Die durch d_∞ und d_{eukl} auf dem $\mathbb{R}^n$ induzierten Topologien auf $\mathbb{R}^n$ stimmen überein.

Die Aussage 5. ist Spezialfall einer allgemeineren. Seien d_0, d_1 Metriken auf einer Menge X.

Definition 4.23 *Die Metriken heißen **äquivalent**, wenn es $0 < c, C \in \mathbb{R}$ gibt, so daß für alle $x, y \in X$ $cd_0(x,y) \leq d_1(x,y) \leq Cd_0(x,y)$ gilt.*

Dies ist in der Tat eine Äquivalenzrelation, da man dann auch
$$C^{-1}d_1(x,y) \leq d_0(x,y \leq c^{-1}d_1(x,y)$$
hat.

Lemma 4.24 *Wenn die Metriken d_0, d_1 auf X äquivalent sind, dann stimmen die durch sie induzierten Topologien $\mathcal{T}_0$ und $\mathcal{T}_1$ auf X überein.*

Beweis: Sei $U \in \mathcal{T}_0$. Wir müssen zeigen, daß $U \in \mathcal{T}_1$. Sei $x \in U$. Dann gibt es $r > 0$ derart, daß $B_{d_0}(x,r) \subseteq U$. Wenn für $z \in X$ die Ungleichung $d_1(x,z) < cr$ gilt, dann gilt auch $d_0(x,z) < c^{-1}d_1(x,z) < r$, also $z \in B_{d_0}(x,r)$. Wir schließen, daß $B_{d_1}(x,cr) \subseteq U$. ■

Lemma 4.25 *Die Metriken d_∞ und d_{eukl} auf $\mathbb{R}^n$ sind äquivalent.*

Beweis: In der Tat gilt offensichlich für alle $x, y \in \mathbb{R}^n$ $d_\infty(x,y) \leq d_{eukl}(x,y)$. Auf der anderen Seite gilt

$$\begin{aligned} d_{eukl}(x,y) &= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \\ &\leq \sqrt{n\mathtt{max}_{i=1}^{n}|x_i - y_i|^2} \\ &\leq \sqrt{n}\mathtt{max}_{i=1}^{n}|x_i - y_i| \\ &= \sqrt{n}d_\infty(x,y) \ . \end{aligned}$$

Sei (X,d) ein metrischer Raum und $\mathcal{T}$ die induzierte Topologie. Dann

1. $\emptyset, X \in \mathcal{T}$,

2. $\mathcal{T}$ is abgeschlossen unter endlichen Durchschnitten. In der Tat, sei $(U_i)_{i\in I}$ eine endliche Familie in $\mathcal{T}$ und $x \in \bigcap_{i\in I} U_i$. Dann gibt es $(r_i)_{i\in I}$ mit $r_i > 0$ und $B(x,r_i) \subseteq U_i$ für alle $i \in I$. Wir definieren $r := \mathtt{min}\{r_i | i \in I\} > 0$ (hier ist die Endlichkeit von I wichtig). Dann gilt $B(x,r) \subseteq B(x,r_i)$ für all $i \in I$ und damit $B(x,r) \subseteq$. Folglich ist $\bigcap_{i\in I} U_i \in \mathcal{T}$.

3. $\mathcal{T}$ ist abgeschlossen unter beliebigen Vereinigungen. In der Tat, sei $(U_i)_{i\in I}$ eine beliebige Familie in $\mathcal{T}$ und $x \in \bigcup_{i\in I} U_i$. Dann gibt es ein $i \in I$ mit $x \in U_i$. Folglich gibt es $r > 0$ mit $B(x,r) \subseteq U_i$. Damit gilt auch $B(x,r) \subseteq \bigcup_{i\in I} U_i$. Wir schließen $\bigcup_{i\in I} U_i \in \mathcal{T}$.

Wir axiomatisieren diese Eigenschaften von $\mathcal{T}$.

Definition 4.26 *Eine **Topologie** auf einer Menge X ist eine Teilmenge $\mathcal{T} \subseteq \mathcal{P}(X)$ mit:*

1. $\emptyset, X \in \mathcal{T}$,

2. $\mathcal{T}$ is abgeschlossen unter endlichen Durchschnitten.

3. $\mathcal{T}$ ist abgeschlossen unter beliebigen Vereinigungen.

*Ein **topologischer Raum** ist ein Paar $(X,\mathcal{T})$ aus einer Menge X und einer Topologie $\mathcal{T}$ auf X.*

1. Die von einem Abstand auf X induzierte Topologie ist eine Topologie in diesem Sinne.

2. Die vom diskreten Abstand auf X induzierte Topologie ist die **diskrete Topologie** $\mathcal{T}_{disc} = \mathcal{P}(X)$.

3. Ist $(X, \mathcal{T})$ ein topologischer Raum und $A \subseteq X$ eine Teilmenge. Dann ist $\mathcal{T}_A := \{U \cap A | U \in \mathcal{T}\}$ eine Topologie auf A. In der Tat ist $\emptyset = \emptyset \cap A \in \mathcal{T}_A$, $A = X \cap A \in A$. Ist $(V_i)_{i \in I}$ eine Familie in $\mathcal{T}_A$, dann gibt es eine Familie $(U_i)_{i \in I}$ in $\mathcal{T}$ mit $V_i = U_i \cap A$. Ist I endich, dann gilt

$$\bigcap_{i \in I} V_i = \bigcap_{i \in I} (U_i \cap A) = \bigcap_{i \in I} U_i \cap A \in \mathcal{T}_A \ .$$

   Für beliebige I gilt

$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} (U_i \cap A) = \bigcup_{i \in I} U_i \cap A \in \mathcal{T}_A \ .$$

   Wir nennen $\mathcal{T}_A$ die Einschränkung der Topologie $\mathcal{T}$ auf A.

4. Die Menge $\{\emptyset, X\} \subset \mathcal{P}(X)$ ist die **chaotische Topologie** auf X.

5. Sei $X = \{a, b, c\}$ und
$$\mathcal{T} := \{\emptyset, \{a\}, \{a, b\}, X\} \ .$$
   Dies ist eine Topologie. Diese kommt aber nicht von einem Abstand. In der Tat, wäre diese Topologie von einem Abstand d induziert, dann würde $0 < d(a, b) =: s$ gelten. Damit wäre $a \notin B(b, \frac{s}{2})$. Andererseits ist $B(b, \frac{s}{2})$ offen. Die einzigen offenen Mengen, welche b enthalten, sind X und $\{a, b\}$, und diese enthalten auch a.

Ein besoonders wichtiges Beispiel ist die vervollständigte Zahlengerade

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} \ .$$

Die Topologie auf $\bar{\mathbb{R}}$ wir wie folgt definiert. Eine Menge $U \subset \bar{\mathbb{R}}$ ist genau dann offen, wenn

1. $U \cap \mathbb{R}$ offen ist,

2. wenn $\infty \in U$ gilt, auch $(a, \infty) \subset U$ für ein $a \in \mathbb{R}$ gilt,

3. wenn $-\infty \in U$ gilt, auch $(-\infty, a) \subset U$ gilt.

4.5 Konvergenz im metrischen und topologischen Kontext

Wir hatten schon konvergente Folgen und Cauchyfolgen in metrischen Räumen definiert. Während Cauchyfolge ein metrischer Begriff ist, kann man konvergente Folgen im topologischen Kontext verstehen.

Lemma 4.27 *Eine konvergente Folge in einem metrischen Raum ist eine Cachyfolge.*

Die Umkehrung gilt nicht.

In einem topologischen Raum $(X, \mathcal{T})$ kann man von Umgebungen reden. Sei $(X, \mathcal{T})$ ein topologischer Raum.

Definition 4.28 *Eine Menge $M \subseteq X$ ist eine* ***Umgebung*** *von $x \in X$, wenn es eine offene Teilmenge $U \in \mathcal{T}$ mit $x \in U \subseteq M$ gibt.*

Definition 4.29 *Eine Folge (x_n) in einem topologischen Raum $(X, \mathcal{T})$* ***konvergiert (topologisch)*** *gegen $x \in X$, falls für jede Umgebung M von x ein $n_0 \in \mathbb{N}$ existiert, so daß $x_n \in M$ für alle $n \geq n_0$ gilt.*

Mit Hilfe diese Begriffs können wir etwa sagen, daß eine Folge reeller Zahlen (x_n) gegen ∞ konvergiert. Dazu betrachten wir diese Folge in der vervollständigten Zahlengeraden $\bar{\mathbb{R}}$. Explizit bedeutet $x_n \to \infty$ folgendes. Für jedes $a \in \mathbb{R}$ ist $(a, \infty) \cup \{\infty\}$ eine Umgebung von ∞. Folglich gibt für jedes $a \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ derart, daß aus $n \geq n_0$ folgt $x_n > a$. Ist V eine Umgebung von ∞. Dann gibt es ein Intervall $(a, \infty) \subset V$. Dann gilt für $n \geq n_0$, daß $a_n \in (a, \infty) \subset V$

Wir zeigen nun für metrische Räume, daß der Begriff der topologischen Konvergenz mit dem alten zusammenfällt.

Lemma 4.30 *Sei (X, d) ein metrischer Raum und $(X, \mathcal{T})$ der unterliegende topologische Raum. Eine Folge (x_n) konvergiert genau dann (metrisch) gegen x, wenn sie (topologisch) gegen x konvergiert.*

Beweis: Gelte $x_n \to x$ metrisch und sei M eine Umgebung von x. Dann gibt es ein $r > 0$ derart, daß $B(x, r) \subseteq M$. Wegen $d(x_n, d) \to 0$ finden wir ein $n_0 \in \mathbb{N}$ derart, daß für $n \geq n_0$ gilt $d(x_n, x) < r$. Für $n \geq n_0$ gilt dann $x_n \in B(x, r)$, also $x_n \in M$.

Gelte nun $x_n \to x$ topologisch. Sei $\epsilon > 0$ gegeben. Dann ist $B(x, \epsilon)$ eine Umgebung von x. Folglich finden wir ein $n_0 \in \mathbb{N}$ derart, daß für alle $n \geq n_0$ gilt $x_n \in B(x, \epsilon)$. damit folgt aus $n \geq n_0$, daß $d(x, x_n) < \epsilon$. ∎

Im allgemeinen können Topologien unerwartete Eigenschaften haben. Sei $\mathcal{T} := \{\emptyset, X\} \subset \mathcal{P}(X)$ ist die chaotische Topologie auf X. In dieser Topologie konvergiert jede Folge (x_i) gegen jeden Punkt. Insbesondere kann eine Folge **mehrere verschiedene Grenzwerte** haben.

Bei Topologien, welche von Metriken kommen, passiert das nicht. Diese haben die Hausdorffeigenschaft.

Definition 4.31 *Ein topologischer Raum X ist* ***Hausdorffsch****, wenn für alle $x, y \in X$ mit $x \neq y$ Umgebungen M, N von x, y existieren mit $M \cap N = \emptyset$.*

Zum Beispiel ist die vervollständigte Zahlengerade $\bar{\mathbb{R}}$ Hausdorffsch.

Lemma 4.32 *Der unterliegende topologische Raum eines metrischen Raumes ist* ***hausdorffsch****.*

Beweis: Sei (X, d) metrisch und $x, y \in X$ mit $x \neq y$. Dann gilt $0 < c = d(x, y)$. Wir betrachten die Umgebungen $M := B(x, \frac{c}{4})$ und $N := B(y, \frac{c}{4})$. Wenn $z \in M \cap N$ gilt, dann ist $c = d(x, y) \leq d(x, z) + d(z, y) < \frac{c}{4} + \frac{c}{4} = \frac{c}{2}$, ein Widerspruch. ■

Lemma 4.33 *In einen hausdorffschen topologischen Raum hat eine Folge* ***höchstens einen Grenzwert****.*

Beweis: Sei (x_i) eine Folge in einem hausdorffschen topologischen Raum und gelte $x_i \to x$, $x_i \to y$ für $x \neq y$. Dann gibt es Umgebungen M, N von x, y mit $M \cap N = \emptyset$. Auf der anderen Seite enthalten sowohl M also auch N alle bis auf endlich viele Folgenglieder. Das ist eine Widerspruch. ■

Ist $(X, \mathcal{T})$ hausdorffsch und $A \subseteq X$, dann ist auch $(A, \mathcal{T}_A)$ hausdorffsch. Die Hausdorffeigenschaft von X **vererbt** sich auf A.

Sei $(X, \mathcal{T})$ ein topologischer Raum.

Definition 4.34 *Eine Teilmenge $A \subset X$ heißt* ***folgenabgeschlossen****, wenn für jede Folge (a_i) in A mit $a_i \to x$ auch $x \in A$ gilt.*

Lemma 4.35 *Eine abgeschlossene Menge $A \subseteq X$ in einem topologischen Raum ist folgenabgeschlossen.*

Beweis: In der Tat, sei (a_i) eine Folge in A und $a_i \to x$. Wenn $x \notin A$, dann ist $X \setminus A$ eine Umgebung von x in X und enhält damit fast alle Folgenglieder. Das ist nicht möglich. ■

Die Umkehrung gilt nicht. Dazu betrachten wir die Menge $X := \{0, 1\}^{\mathbb{R}}$ mit folgender Topologie. Ist $f \in X$ und $B \subseteq \mathbb{R}$ endlich, so ist

$$V_B(f) := \{g \in X | f(x) = g(x) \text{ für alle } x \in B\}$$

per definitionem eine Umgebung von f. Wir setzen

$$\mathcal{T} := \{U \subseteq X | U \text{ enthält Umgebung jedes ihrer Punkte}\} .$$

Für die Konsistenz müssen wir zeigen, daß die Mengen $V_B(f)$ in $\mathcal{T}$ liegen. Sei $g \in V_B(f)$. Dann ist in der Tat $V_B(g) = V_B(f)$.

Wir weisen nun die Topologieaxiome nach. Klar ist $\emptyset, X \in \mathcal{T}$. Weiter ist $\mathcal{T}$ abgeschlossen unter beliebigen Vereinigungen. Sei $(U_i)_{i \in I}$ eine endliche Familie. Dann ist auch $\bigcap_{i \in I} U_i \in \mathcal{T}$. Sei $f \in \bigcap_{i \in I} U_i$ und $V_{B_i}(f) \subseteq U_i$. Dann ist $V_{\cup_{i \in I} B_i}(f) \subseteq \bigcap_{i \in I} U_i$. Der Raum $(X, \mathcal{T})$ ist hausdorffsch. In der Tat, sei $f \neq g$. Dann gibt es $x \in \mathbb{R}$ mit $f(x) \neq g(x)$. Dann ist $V_{\{x\}}(f) \cap V_{\{x\}}(g) = \emptyset$. In X gilt $f_n \to f$, wenn für jede endliche Teilmenge $B \subset \mathbb{R}$ für genügen große n gilt $(f_n)_{|B} = f_{|B}$. Es reicht sogar, für B einpunktige Mengen zu nehmen.

Sei $f \in X$. Wir betrachten nun die Menge

$$A := \{g \in X | \{x \in \mathbb{R} | f(x) \neq g(x)\} \text{ ist abzählbar}\} .$$

Diese Menge ist folgenabgeschlossen. In der Tat, sei (g_n) eine Folge in A mit $g_n \to g$.

Dann ist $\{g_n \neq f\}$ abzählbar für alle n. Folglich ist $C := \bigcup_{n \in \mathbb{N}} \{g_n \neq f\}$ abzählbar. Für $x \notin C$ gilt $f(x) = g_n(x) = g(x)$ für genügend große n. Damit ist $\{f \neq g\} \subseteq C$ abzählbar.

Wir zeigen nun, daß A nicht abgeschlossen ist. Sei dazu $g \in X \setminus A$. Ist $B \subset \mathbb{R}$ endlich, dann gilt $V_B(g) \cap A \neq \emptyset$. In der Tat kann man $h(x) := g(x)$ für $x \in B$ und $h(x) := f(x)$ für $x \in \mathbb{R} \setminus B$ definieren. Dann gilt $h \in V_B(g)$ und $\{h \neq f\} \subseteq B$ ist abzählbar. Damit enthält $X \setminus A$ keine Umgebung von g und ist damit nicht offen.

Lemma 4.36 *In einem metrischen Raum sind folgenabgeschlossene Mengen abgeschlossen.*

Beweis: Sei (X, d) metrisch und $A \subseteq X$ folgenabgeschlossen, aber nicht abgeschlossen. Dann gibt es einen Punkt $x \in X \setminus A$, für welchen jede seiner Umgebungen A nichttrivial schneidet. Insbesondere gibt es für jedes $n \in \mathbb{N}$ ein $a_n \in A \cap B(x, \frac{1}{n})$. Es gilt $a_n \to x$ und $x \notin A$. Widerspruch. ■

Unser Beispiel zeigt, daß es Hausdorffräume gibt, welche nicht von metrischen Räumen kommen.

Definition 4.37 *Ein topologischer Raum heißt* ***folgenkompakt****, wenn jede Folge in diesem Raum eine konvergente Teilfolge besitzt.*

Definition 4.38 *Eine* ***Überdeckung*** *eines topologischen Raumes $(X, \mathcal{T})$ ist eine Familie $(U_i)_{i \in I}$ offener Teilmengen mit $\bigcup_{i \in I} U_i = X$.*

Eine **Teilüberdeckung** von $(U_i)_{i \in I}$ ist eine Überdeckung $(U_i)_{i \in I'}$ mit einer Teilmenge $I' \subseteq I$.

Definition 4.39 *Ein topologischer Raum heißt* ***quasi-kompakt****, wenn jede Überdeckung eine endliche Teilüberdeckung besitzt. Wenn er zusätzlich Hausdorffsch ist, dann heißt er* ***kompakt****.*

Lemma 4.40 *In einem Hausdorffraum ist eine kompakte Teilmenge $A \subseteq X$ abgeschlossen.*

Beweis: Sei $x \in X \setminus A$. Wir müssen eine Umgebung von x in $X \setminus A$ konstruieren. Für jedes $a \in A$ finden wir offene Umgebungen W_a und V_a von a und x mit $W_a \cap V_a = \emptyset$. Die Familie $(W_a \cap A)_{a \in A}$ ist eine Überdeckung von A. Wir wählen eine endliche Teilüberdeckung $(W_a \cap A)_{a \in I}$ aus und bilden $U := \cap_{a \in I} U_a$. Dann gilt $U \cap A \subset U \cap \bigcup_{a \in I} W_a = \emptyset$. Weiterhin ist U offen. ■

1. Wir hatten schon gesehen, daß abgeschlossene Intervalle in $\mathbb{R}$ folgenkompakt sind.
2. Ist X folgenkompakt und $A \subseteq X$ folgenabgeschlossen, dann ist auch A folgenkompakt. In der Tat, ist (a_n) eine Folge in A. Dann hat diese eine Teilfolge, welche in X konvergiert und der Grenzwert liegt in A. Somit konvergiert (a_n) auch in A.

3. Sei (f_n) eine Folge in der Menge A im obigen Beispiel. Dann ist $C := \bigcup_{n\in\mathbb{N}} \{f_n \neq f\}$ abzählbar. Wir wählen eine Abzählung (c_n). Wir wählen nun induktiv Teilfolgen so aus, daß $f_{i,n}(c_j) = f_{i,1}(c_j)$ für alle $j \leq i$. Im Schritt $(i-1)$ nach (i) betrachten wir zwei Fälle:

   (a) Die Folge $(f_{i-1,n}(c_i))$ nimmt den Wert 0 unendlich oft an. Dann wählen wir die als Teilfolge alle Glieder mit $f_{i-1,n}(c_i) = 0$.

   (b) Andernfalls wählen wir als Teilfolge $(f_{i,n})$ alle Glieder mit $f_{i-1,n}(c_i) = 1$.

   Wir betrachten nun die Diagonalfolge $(f_{n,n})$, auch eine Teilfolge. Die Folge $(f_{n,n}(c))$ stabilisert sich für jedes $c \in C$. Deshalb konvergiert $(f_{n,n})$ gegen g. Da $\{g \neq f\} \subseteq C$ gilt und damit insbesondere abzählbar ist, ist $g \in A$. Damit ist A folgenkompakt, aber nicht abgeschlossen.

Definition 4.41 *Ist $A \subseteq X$ eine Teilmenge eines topologischen Raumes, so bezeichnet*

$$\bar{A} := \{x \in A | \textit{jede Umgebung von } x \textit{ schneidet } A\}$$

*die **abgeschlossene Hülle** von A*

und für $A \subseteq B \subseteq X$ auch $\bar{A} \subseteq B$.

Lemma 4.42 *Die abgeschlossene Hülle von A ist die kleinste abgeschlossene Teilmenge von A, welche A enthält.*

Beweis: In der Tat ist $\bar{A}$ abgeschlossen. Ist nämlich $x \notin A$, dann gibt es eine offene Umgebung von A, welche A nicht schneidet. Damit ist das Komplement von A die Vereinigung aller dieser Umgebungen über die Punkte des Komplementes und damit offen. Es gilt $A \subseteq \bar{A}$.

Sei nun $A \subseteq B$ und B abgeschlossen. Ist $x \notin B$, dann ist $X \setminus B$ eine Umgebung von X welche A nicht schneidet. Also gilt $x \notin \bar{A}$. Folglich gilt $\bar{A} \subseteq B$. ■

Lemma 4.43 *Für einen metrischen Raum sind die folgenden Aussagen äquivalent:*

1. *Der Raum ist kompakt.*
2. *Der Raum ist folgenkompakt.*

Außerdem folgt aus beiden Kompaktkeitsbedingungen die Vollständigkeit.

Beweis: Sei X ein kompakter metrischer Raum. Sei (x_n) eine Folge in einem kompakten Raum. Wir bilden

$$F_n := \overline{\{x_i | i \geq n\}}$$

und zeigen, daß es einen Punkt $x \in \bigcap_{n\geq 0} F_n$ gibt. In der Tat wäre andernfalls $(X \setminus F_n)_{n\in\mathbb{N}}$ eine Überdeckung von X, aus welcher man eine endliche Teilüberdeckung auswählen könnte. Wegen $\cdots \subseteq F_{n+1} \subseteq F_n \subseteq \dots$ wäre dann $F_n = \emptyset$ für ein $n \in \mathbb{N}$, was unmöglich ist. In jeder Umgebung $B(x, \frac{1}{i})$ gibt es ein x_{n_i} mit $n_i \geq n$ derart. Die Teilfolge (x_{n_i}) konvergiert gegen x. Damit ist der Raum folgenkompakt.

Sei nun X folgenkompakt. Wir zeigen die Vollständigkeit. Sei (x_n) eine Cauchyfolge in X. Dann besitzt diese eine konvergente Teilfolge mit Grenzwert x. Dann gilt aber auch $x_n \to x$.

Wir nehmen nun an, daß X folgenkompakt ist und schließen die Kompaktheit. Wir zeigen zuerts, daß X total beschränkt ist, also für jedes $\epsilon > 0$ durch endlich viele Bälle vom Radius $\epsilon > 0$ überdeckt werden kann. Andernfalls gäbe es eine Folge (x_n) mit der Eigenschaft, daß $\mathtt{min}_{j \leq n} d(x_{n+1}, x_j) \geq \epsilon$ für alle n gilt. Diese Folge kann keine konvergente Teilfolge enthalten.

Sei $(U_i)_{i \in I}$ eine Überdeckung. Wir nehmen an, daß diese Überdeckung keine endliche Teilüberdeckung besitzt. Wir wählen induktiv eine Folge von Kugeln $K_n := B(x_n, 2^{-n})$ derart, daß K_n nicht von einer endlichen Teilüberdeckung überdeckt wird. In der Tat kann X mit endlich vielen Kugeln vom Radius 2^{-n} überdeckt werden, und eine dieser mit nichtleerem Durchschnitt mit K_{n-1} muß die geforderte Eigenschaft haben. Sei (x_n) die Folge der Mittelpunkte. Es gilt für $m \geq n$ daß

$$d(x_n, x_m) \leq \sum_{i=n+1}^{m} d(x_{i-1}, x_i) \leq \sum_{i=n+1}^{m} 2^{1-i} \leq 2^{-n+1} .$$

Wir sehen, daß (x_n) eine Cachyfolge ist und damit einen Grenzwert x hat. Sei $x \in U_i$ für ein $i \in I$. Dann gilt für ein $\epsilon > 0$ daß $B(x, \epsilon) \subset U_i$. Es gibt nun ein $n > 0$ mit $d(x, x_n) < \frac{\epsilon}{2}$ und $2^{-n} < \frac{\epsilon}{2}$. Dann gilt aber $B(x_n, 2^{-n}) \subseteq U_i$ im Widerspruch zur Konstruktion der Folge der Bälle. ■

In der Tat ist nicht jeder kompakte Raum folgenkompakt. Hier ein Beispiel. Wir betrachten den Raum F der Folgen mit Werten im Intervall $[0, 1]$. Sei $X := \prod_{f \in F} [0, 1] = [0, 1]^F$. Nach dem Satz von Tychonov ist dieser Raum mit der Produkttopologie kompakt. Wenn wir X als Funktionenraum auffasssen, dann ist die Konvergenz die punktweise Konvergenz. Sei $\delta_n \in X$ durch $\delta_n((a_n)) := a_n$ gegeben. Die Folge (δ_n) hat keine konvergente Teilfolge. In der Tat, wenn (δ_{n_i}) eine solche Teilfolge wäre, dann wählen wir $(a_n) \in F$ derart, daß $a_{n_i} = 1/2 + (-1)^i 1/2$ gilt. Dann ist $\delta_{n_i}((a_n)) = 1/2 + (-1)^i 1/2$ sicher nicht konvergent, ein Widerspruch.

Wir betrachten den Raum $\mathbb{R}^n$ mit der von d_{eukl} induzierten Topologie. Diese Topologie ist immer gemeint, wenn nichts anderes gesagt wird.

Definition 4.44 *Eine Teilmenge eines metrischen Raumes heißt* ***beschränkt****, wenn sie in einem Ball enthalten ist.*

Lemma 4.45 *Eine Teilmenge von $\mathbb{R}^n$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*

Beweis: Sei $A \subseteq \mathbb{R}$ beschränkt. Dann gibt es ein $R \in \mathbb{R}$ derart, daß $A \subseteq B(0, R)$. Wenn $(x^1, \dots, x^n) \in A$, dann gilt $x^i < R$. Sei jetzt (x_j) eine Folge in A. Dann sind die Folgen

(x_n^i) für $i = 1, \ldots, n$ beschränkt. Wir wählen nun in n Schritten Teilfolgen aus. Im ersten Schritte wählen wir eine Teilfolge so daß $(x_{n_j}^1)$ konvergiert. Dann wählen wir aus dieser wieder eine Teilfolge so daß $(x_{n_j}^2)$ konvergiert. Nach n Schritten haben wir eine Teilfolge derart, daß $(x_{n_j}^i)$ für alle $i = 1, \ldots, n$ konvergieren. Damit konvergiert $x_j \to x$. Wenn jetzt A abgeschlossen ist, dann gilt $x \in A$.

Sei nun A kompakt. Dann ist A abgeschlossen. Wäre A nicht beschränkt, dann gäbe es eine Folge (x_n) in A mit $d(x_n, 0) \to \infty$. Diese hat dann aber keine konvergente Teilfolge, ein Widerspruch. ■

Die vervollständigte Zahlengerade ist kompakt und Folgenkompakt.

4.6 Stetigkeit im metrischen und topologischen Kontext

Wir können nun die verschiedenen Definitionen der Stetigkeit einer Abbildung $f : X \to Y$ an der Stelle x im metrischen und topologischen Kontext verallgemeinern. Seien X, Y topologische Räume.

Definition 4.46 *Die Abbildung f ist an der Stelle x **folgenstetig**, wenn für jede Folge (x_n) in X mit $x_n \to x$ auch $f(x_n) \to f(x)$ gilt.*

Definition 4.47 *Die Abbildung f ist an der Stelle x **(topologisch) stetig**, wenn für jede Umgebung M von $f(x)$ in Y das Urbild $f^{-1}(M)$ eine Umgebung von x in X ist.*

Sind (X, d_X) metrisch und $(Y, \mathcal{T}_Y)$ topologisch, dann haben wir auch eine metrische Definition

Definition 4.48 *Die Abbildung f ist an der Stelle x **(metrisch) stetig**, wenn für jede Umgebung M von $f(x)$ ein $\delta > 0$ exitsiert, so daß aus $d_X(u, x) < \delta$ folgt $f(u) \in M$.*

Im topologischen Kontext gilt:

Lemma 4.49 *Stetigkeit (topologisch) impliziert Folgenstetigkeit.*

Beweis: Sei f an der Stelle x stetig und (x_n) eine Folge in X mit $x_n \to x$. Sei M eine Umgebung von $f(x)$. Dann ist $f^{-1}(M)$ eine Umgebung von x. Deshalb gibt es ein $n_0 \in \mathbb{N}$ derart, daß aus $n \geq n_0$ folgt $x_n \in f^{-1}(M)$. Damit gilt $f(x_n) \in M$ für alle $n \geq n_0$. Folglich gilt $f(x_n) \to x$. ■

Die Umkehrung gilt im allgemeinen nicht. Hier ist wieder unser Gegenbeispiel. Wir betrachten den topologischen Raum $X = \{0, 1\}^{\mathbb{R}}$ wie oben und $A \subset X$. Sei $b \in \bar{A} \setminus A$ und $Y := A \cup \{b\}$ mit der induzierten Topologie. Wir betrachten die Funktion $\phi : Y \to \mathbb{R}$ mit $\phi(a) := 1$ für $a \in A$ und $\phi(b) := 0$. Sei (y_n) eine konvergente Folge in Y. Wenn unendlich viele Glieder der Folge in A liegen, dann hat die Teilfolge dieser Glieder einen Grenzwert $y \in A$. Dann können aber nur endlich viele Glieder gleich b sein, weil es sonst auch eine gegen $b \neq y$ konvergierende Teilfolge gäbe, was wegen der Hausdorffeigenschaft nicht möglich ist. In diesem Fall gilt $\phi(y_n) \to 1 = \phi(y)$. Wenn nur endlich viele Glieder

der Folge in A liegen, dann gilt $y_n \to b$ und $\phi(y_n) \to 0 = \phi(b)$. Wäre ϕ stetig, dann wäre die Menge $A := \phi^{-1}\{1\}$ abgeschlossen in Y, sie ist es aber nicht. Wäre nämlich $A = Y \cap E$ für eine abgeschlossene Menge $E \subseteq X$, dann wäre $b \notin E$. Dann wäre aber $X \setminus E$ eine offene Umgebung von b welche A nicht schneidet. Damit wäre $b \notin \bar{A}$.

Lemma 4.50 *Ist X ein metrischer Raum, dann stimmen die Begriffe der Folgenstetigkeit, Stetigkeit (topologisch) und Stetigkeit im Sinne der ϵ-δ-Definition überein.*

Beweis: Wir betrachten $f : X \to Y$, Y ist topologisch, und $x \in X$. Die Implikationen

$$\epsilon - \delta\text{-stetig in } x \;\Rightarrow \text{stetig in } x \Rightarrow \text{folgenstetig in } x$$

sind einfach oder schon gezeigt worden. Sei $f : X \to Y$ nun folgenstetig, nicht aber $\epsilon - \delta$-stetig. Dann finden wir eine Umgebung M von $f(x)$ derart, daß für alle $n \in \mathbb{N}$ ein $x_n \in X$ mit $d_X(x_n, x) < \delta$ und $f(x_n) \notin M$ gilt. Damit gilt $x_n \to x$, nicht aber $f(x_n) \to f(x)$, ein Widerspruch zur Folgenstetigkeit. ■

Definition 4.51 *Eine Abbildung $f : X \to Y$ zwischen topologischen Räume ist stetig, wenn sie in jedem Punkt von X stetig ist.*

Eine effektive Charakterisierung der Stetigkeit einer Abbildung $f : X \to Y$ zwischen topologischen Räumen ist:

Lemma 4.52 *$f : X \to Y$ ist genau dann stetig, wenn für jedes offene $U \subseteq Y$ das Urbild $f^{-1}(U) \subseteq X$ offen ist.*

Beweis: Sei diese Bedingung erfüllt, $x \in Y$ und $M \subset Y$ eine Umgebung von $f(x)$. Dann gibt es ein offene Umgebung $U \subset M$ von $f(x)$ und $f^{-1}(U)$ ist eine offene Umgebung con x. Folglich ist auch $f^{-1}(M)$ eine Umgebung von x.

Für die andere Richtung sei $U \subset Y$ offen. Sei $x \in f^{-1}(U)$. Dann ist U eine Umgebung von $f(x)$ und somit $f^{-1}(U)$ eine Umgebung von x. Damit gibt es ein offenes $V_x \subseteq f^{-1}(U)$ mit $x \in V_x$. Wir wählen so ein V_x für alle $x \in f^{-1}(U)$. Dann gilt

$$f^{-1}(U) = \bigcup_{x \in U} V_x \ .$$

Diese Vereinigung offener Menge ist offen. ■

Lemma 4.53 *Die Abbildungen $+, \times : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ und $^{-1} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ sind stetig.*

Beweis: Wir haben die Folgenstetigkeit schon gezeigt. ■

Seien X, Y, Z topologische Räume und $f : X \to Y$ und $g : Y \to Z$ stetig.

Lemma 4.54 *Die Komposition $g \circ f : X \to Z$ ist setig.*

Beweis: Sei $U \subseteq Z$ offen. Dann ist $g^{-1}(U) \subseteq Y$ offen. Folglich ist $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \subseteq X$ offen. ■

Weitere typische Beispiele stetiger Abbildungen

1. Sei X eine Menge und $\mathcal{T}_0 \subseteq \mathcal{T}_1$. Dann ist die identity $\mathtt{id}_X$ stetig als Abbildung $(X, \mathcal{T}_1) \to (X, \mathcal{T}_0)$.

2. Sei X eine Menge $F_b(X)$ die Menge beschränkten reellen Funktionen auf X mit der Metrik $d(f,g) := \sup_{x \in X} |f(x) - g(x)|$. Für $x \in X$ sei $\delta_x : F_b(X) \to \mathbb{R}$ die Auswerteabbildung in x. Dann ist $\delta_x : F_b(X) \to \mathbb{R}$ stetig. Wir können etwa die ϵ-δ-Definition mit $\epsilon = \delta$ verwenden. Sei $f \in F_b(X)$ und $\epsilon > 0$ gegeben. Wenn $d(g,f) < \epsilon$ gilt, dann auch $|\delta_x(f) - \delta_x(g)|$.

3. Sei (X, d) ein metrischer Raum und $A \subseteq X$ eine Teilmenge. Mit
$$d(x, A) := \inf_{a \in A} d(x, a)$$
bezeichnen wir den Abstand von x und A. Diese Funktion ist stetig. Wir benutzen die ϵ-δ-Definition mit $\delta := \frac{\epsilon}{2}$. In der Tat, sei $x \in X$ und $\epsilon > 0$ gegeben. Sei nun $y \in X$ und $d(x,y) < \frac{\epsilon}{2}$. Wir finden $a_x, a_y \in A$ derart, daß $|d(x,A) - d(x,a_x)| \le \frac{\epsilon}{2}$ und $|d(y,A) - d(y,a_y)| \le \frac{\epsilon}{2}$. Nun ist aber
$$d(x,A) \le d(x,a_y) \le d(x,y) + d(y,a_y) \le \frac{\epsilon}{2} + d(y,A) + \frac{\epsilon}{2} = d(y,A) + \epsilon \ .$$
Analog sehen wir $d(y,A) \le d(x,A) + \epsilon$, also $|d(x, A - d(y,A)| \le \epsilon$.

4. Sei l^1 die Menge aller reellen Folgen (a_i) derart, daß $\|(a_n)\| := \sum_{n=1}^{\infty} |a_n|$ konvergiert. Dann ist $d((a_n),(b_n)) := \|(a_n) - (b_n)\|$ ein Abstand. Wir betrachten die Abbildung
$$\sum : l^1 \to \mathbb{R} \ , (a_n) \mapsto \sum_{n=1}^{\infty} a_n \ .$$
Diese Abbildung ist stetig. Wir benutzen die ϵ-δ-Definition mit $\epsilon = \delta$. Sei $(a_n) \in l^1$ und $\epsilon > 0$ gegeben. Wenn $(b_n) \in l^1$ die Ungleichung $d((a_n), b(n)) < \epsilon$ erfüllt, dann gilt
$$|\Sigma(a_n) - \Sigma(b_n)| = |\sum_{n=1}^{\infty} a_n - b_n| \le \sum_{n=1}^{\infty} |a_n - b_n| = \|(a_n) - (b_n)\| = d((a_n), (b_n) \ .$$

5. Die Abildung $\mathsf{e}^{\sin(x)}$ ist als Komposition stetiger Abbildungen stetig.

Lemma 4.55 *Eine stetige folgenstetige Funktion $f : X \to \mathbb{R}$ auf einem folgenkompakten topologischen Raum $(X, \mathcal{T})$ ist beschränkt und nimmt ihre Extremwerte an.*

Beweis: Die Menge $f(X)$ ist von oben beschränkt. Andernfalls gäbe es eine Folge (x_n) in X mit $f(x_n) > n$. Diese Folge hat eine konvergente Teilfolge (x_{n_i}) mit $\lim_{i\to\infty} x_{n_i} = \xi$. Dann gilt $\lim_{i\to\infty} f(x_{n_i}) = f(\xi)$, insbesondere ist die Folge $(f(x_{n_i}))_{i\in\mathbb{N}}$ also beschränkt, ein Widerspruch.

Wir beweisen nur die zweite Gleichung. Sei $M = \sup f(X)$. Dann gibt es eine Folge (x_n) in X derart, daß $f(x_n) \to M$ gilt. Diese Folge ist beschränkt und hat deshalb eine konvergente Teilfolge.

Indem wir die Folge durch diese Teilfolge ersetzen, können wir annehmen, daß $x_n \to \omega$ für ein $\omega \in X$. Dann gilt $f(\omega) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = M$. ■

Lemma 4.56 *Ist $f : X \to Y$ eine stetige Abbildung zwischen topologischen Räumen. Wenn X kompakt ist, dann auch das Bild $f(X) \subseteq Y$.*

Beweis: Sei $(U_i)_{i\in I}$ eine Überdeckung von $f(X)$. Dann ist $(f^{-1}(U_i))_{i\in I}$ eine Überdeckung von X. Folglich gibt es eine endliche Teilmenge $I' \subseteq I$ derart, daß $X = \cup_{i\in I'} U_i$. Dann ist aber $\cup_{i\in I'} U_i$ eine endliche Teilüberdeckung von $f(X)$. ■

Korollar 4.57 *Eine stetige reellwertige Funktion auf einem kompakten Raum ist beschränkt und nimmt ihre Extremwerte an.*

Beweis: Sei $f : X \to \mathbb{R}$ settig und X kompakt. Dann ist $f(X)$ kompakt, also beschränkt und abgeschlossen. Folglich ist f beschränkt und es gilt etwa $\sup f(X) \in f(X)$, wird also als Wert realisiert. ■

Seien nun X, Y metrische Räume mit Metriken d_X und d_Y.

Definition 4.58 *Eine Abbildung $f : X \to Y$ ist gleichmäßig stetig, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so daß aus $d_X(x, y) < \delta$ gilt $d_Y(f(x), f(y)) < \epsilon$.*

Die Abbildung $x^2 : \mathbb{R} \to \mathbb{R}$ ist nicht gleichmäßig stetig. Auf der anderen Seite sind $\lambda x, |x| : \mathbb{R} \to \mathbb{R}$ gleichmäßig stetig.

Lemma 4.59 *Ist X kompakt, so ist jede stetige Funktion $X \to Y$ auch gleichmäßig stetig.*

Beweis: Möge f stetig, jedoch nicht gleichmäßig stetig sein. Dann gibt es für ein $\epsilon > 0$ Folgen $(x_n), (y_n)$ mit $d_X(x_n, y_n) < \frac{1}{n}$ und $d_Y(f(x_n), f(y_n)) \geq \epsilon$. Nach Auswahl von Teilfolgen können wir annehmen, daß $x_n \to x$ und $y_n \to x$ gilt. Nun gilt einerseits wegen der Stetigkeit von f daß

$$d_Y(f(x_n), f(x)) \to 0 \ , \quad d_Y(f(y_n), f(x)) \to 0$$

und damit

$$d_Y(f(x_n), f(x)) + d_Y(f(y_n), f(x)) \to 0 \ .$$

Andererseits gilt

$$d_Y(f(x_n), f(x)) + d_Y(f(y_n), f(x)) \geq d_Y(f(x_n), f(y_n)) \geq \epsilon \ ,$$

ein Widerspruch. ■

4.7 Grenzwerte

Sei X ein topologischer Raum und $E \subset X$ eine Teilmenge.

Definition 4.60 *Ein Punkt $x \in X$ heißt **Häufungspunkt** von E, wenn jede Umgebung von x in X einen nichtleeren Durchschnitt mit E hat.*

Der Punkt 0 ist Häufungspunkt von $[0,1]$, $(0,1)$, $\{\frac{1}{n}|n \in \mathbb{N}\}$, nicht aber von $[\frac{1}{100}, 1]$. Die Menge aller Häufungspunkte von E ist der Abschluß $\bar{E}$.

Sei Y ein topologischer Raum und $x \in X$ ein Häufungspunkt von E. Sei $f : E \to Y$ eine Abbildung.

Definition 4.61 *Ein Punkt $y \in Y$ ist **Grenzwert von f in** x, wenn für jede Umgebung M von y in Y eine Umgebung N von x in X existiert mit $f(N \cap E) \subset M$. Wir schreiben*

$$y = \lim_{e \to x} f(e) \ .$$

Der Punkt ∞ ist ein Häufungspunkt von $(a, \infty) \subset \bar{\mathbb{R}}$. Sei $f : (a, \infty) \to \mathbb{R}$ eine Funktion, dann ist der Begriff des Grenzwertes $\lim_{x\to\infty} f(x) \in \mathbb{R}$ definiert. Zum Beispiel gilt

$$\lim_{x \to \infty} \frac{1}{x} = 0 \ .$$

Wir betrachten noch einmal die Funktion $\frac{1}{x} : (0, \infty) \to \bar{\mathbb{R}}$. Der Punkt 0 ist ein Häufungspunkt des Definitionsbereichs. Es gilt

$$\lim_{x \to 0} \frac{1}{x} = \infty \ .$$

Die Funktion $\frac{1}{x} : \mathbb{R} \setminus \{0\} \to \bar{\mathbb{R}}$ hat keinen Grenzwert bei 0.

Eine typische Anwendung des Grenzwertbegriffs ist die Bildung des Differenzenquotienten. Sei $U \subset \mathbb{R}$ offen und $u \in U$. Wir betrachten eine Funktion $f : U \to \mathbb{R}$. Dann ist

$$\Delta_u(f)(x) := \frac{f(x) - f(u)}{x - u}$$

auf $U \setminus \{u\}$ definiert. Der Punkt u ist eine Häufungspunkt von U. Wir können fragen, ob der Grenzwert

$$\lim_{x \to u} \Delta_u(f)$$

existiert.

Definition 4.62 *Wenn dieser Grenzwert existiert, so heißt f an der Stelle u **differenzierbar**, und*

$$f'(u) := \lim_{x \to u} \Delta_u(f)$$

*die **Ableitung** von f an der Stelle u.*

Um zu entscheiden, ob ein Grenzwert existiert, ist folgendes Kriterium nützlich. Sei $\hat{E} := E \cup \{x\}$.

Lemma 4.63 *Es gilt* $\lim_{e \to x} f(e) = y$ *genau dann, wenn*

1. $f(x) = y$ *im Fall* $x \in E$ *gilt und*

2. $$\hat{f} : \hat{E} \to Y \ , \quad \hat{f}(z) := \begin{cases} f(z) & z \in E \setminus \{x\} \\ y & z = e \end{cases}$$

   im Punkt x *stetig ist.*

Beweis: Sei $\hat{f}$ stetig. Sei $M \subseteq Y$ eine Umgebung von y. Dann ist $N \cap E = \hat{f}^{-1}(M)$ für eine Umgebung N von x in X. Es gilt nun $f(N \cap E) \subseteq M$. Folglich gilt $\lim_{e \to x} f(x) = y$.

Sei nun $\lim_{e \to x} f(x) = y$. Sei $M \subseteq Y$ eine Umgebung von y. Dann gibt es eine Umgebung N von x in X derart, daß $\hat{f}(N \cap E \setminus \{x\}) \subseteq M$. Da auch $\hat{f}(x) = y \in M$ gilt, folgt $\hat{f}(N \cap \hat{E}) \subseteq M$. Folglich ist $\hat{f}$ an der Stelle x stetig. ■

1. Sei etwa $f : \mathbb{R} \to \mathbb{R}$ durch $f(x) = x^n$ gegeben. Dann gilt

   $$\Delta_u(f)(x) = \frac{x^n - u^n}{x - u} = x^{n-1} + x^{n-2}u + \cdots + xu^{n-2} + u^{n-1} \ .$$

   Wir sehen, daß wir die Funktion $\Delta_u(f)(x)$ durch $\Delta_u(f)(u) := nu^{n-1}$ stetig auf ganz $\mathbb{R}$ fortsetzen können. Folglich ist $x \mapsto x^n$ an der Stelle u differenzierbar und hat als Ableitung den Wert nu^{n-1}.

2. Wir betrachten $f(x) := \mathsf{e}^x$. dann gilt

   $$\Delta_u(f)(x) = \frac{\mathsf{e}^x - \mathsf{e}^u}{x - u} = \mathsf{e}^u \frac{\mathsf{e}^{u-x} - 1}{x - u} = \mathsf{e}^u \sum_{n=1}^{\infty} \frac{(x-u)^n}{n!} \ .$$

   Wir sehen, daß wir $\Delta_u(f)$ durch $\Delta_u(f)(u) := \mathsf{e}^u$ stetig auf ganz $\mathbb{R}$ fortsetzen können. Folglich ist $x \to \mathsf{e}^x$ an der Stelle u differenzierbar und hat die Ableitung e^u.

3. Sie Sinusfunktion erfüllt das Additionstheorem

   $$\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y) \ .$$

   Dieses kann man direkt aus der Potenzreighendarstellung ableiten. Es gilt

   $$\begin{aligned} \Delta_u(\sin)(x) &= \frac{\sin(x) - \sin(u)}{x - u} \\ &= \frac{\sin(x - u + u) - \sin(u)}{x - u} \\ &= \frac{\sin(x-u)\cos(u) - (\cos(x-u) - 1\sin(u))}{x - u} \end{aligned}$$

Wir beobachten nun, daß

$$\frac{\sin(x-u)}{x-u} = \sum_{n=0}^{\infty}(-1)^n \frac{(x-u)^{2n}}{(2n+1)!}$$

und

$$\frac{\cos(x-u)-1}{x-u} = \sum_{n=1}^{\infty}(-1)^n \frac{(x-u)^{2n}}{(2n)!}$$

in $u = x$ stetig fortsetzbar sind, und zwar durch 1 und 0. Folglich gilt

$$\sin'(u) = \cos(u) \ .$$

Sei $(a,b) \subset \mathbb{R}$ und $f : (a,b) \to Y$. Wir sagen, daß f einen rechtsseitigen Grenzwert in a hat, wenn der Grenzwert $\lim_{x\to a} f(x)$ existiert. Analog definieren wir den Begriff eines linksseitigen Grenzwertes in b. Wir schreiben in diesem Fall

$$f(a^+) := \lim_{x\to a} f(x) \ , \quad f(b^-) := \lim_{x\to b} f(x) \ .$$

Ist $u \in (a,b)$ und $f : (a,b) \setminus \{u\} \to X$ definiert, dann schreiben wir

$$\lim_{x\downarrow u} f(x) = f(u^+) := f_{|(u,b)}(u^+). \ , \quad \lim_{x\uparrow u} = f(u^-) := f_{|(a,u)}(u^-) \ .$$

1. Zum Beispiel existiert $f(0^+) = 0$ für $f(x) = \mathtt{e}^{-\frac{1}{x}}$.
2. Für die Vorzeichenfunktion $\mathtt{sign} : \mathbb{R}\setminus\{0\} \to \mathbb{R}$ gilt $\mathtt{sign}(0^-) = -1$ und $\mathtt{sign}(0^+) = 1$.
3. Der Grenzwert $g(0^+)$ für $g(x) = \frac{1}{x}$ auf $\mathbb{R}\setminus\{0\}$ existiert nicht.

Sei $(X, \mathcal{T})$ ein topologischer Raum, $E \subset X$ und $e \in X$ ein Häufungspunkt von E.

Lemma 4.64 *Wenn $f, g : E \to \mathbb{R}$ Grenzwerte im Punkt e haben, dann auch $f+g$ und fg und es gilt*

$$\lim_{e\to x}(f+g)(e) = \lim_{e\to x} f(e) + \lim_{e\to x} g(e) \ , \quad \lim_{e\to x}(fg)(x) = \lim_{e\to x} f(x) \ \lim_{e\to x} g(x) \ .$$

Wenn $0 \notin g(E)$ und $\lim_{x\to e} f(x) \neq 0$, dann gilt auch $\lim_{x\to e} g^{-1}(x) = g^{-1}(e)$.

Beweis: Wir argumentieren mit den stetigen Fortsetzungen $\hat{f}$ und $\hat{g}$. In der Tat gilt etwa $\widehat{f+g} = \hat{f} + \hat{g}$. Diese Funktion ist stetig, und es gilt

$$\lim_{e\to x}(f+g)(e) = \widehat{f+g}(x) = \hat{f}(x) + \hat{g}(x) = \lim_{e\to x} f(e) + \lim_{e\to x} g(e) \ .$$

. ■

4.8 Punktweise Konvergenz - gleichmäßige Konvergenz

Sei $A \subseteq \mathbb{R}$ oder allgemeiner A ein topologischer Raum. Wir betrachten Funktionen $A \to \mathbb{R}$ (oder allgemeiner mit Werten in einem topologischen Raum).

Definition 4.65 *Eine Folge von Funktionen (f_n), $f_n : A \to \mathbb{R}$* ***konvergiert punktweise*** *gegen $f : A \to \mathbb{R}$, falls für alle $a \in A$*

$$\lim_{n\to\infty} f_n(a) = f(a)$$

gilt

1. Sei $A = [0,1]$. Die Folge $f_n(x) := x^n$ konvergiert **punktweise** gegen die Funktion

$$x \mapsto \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1 \end{cases} .$$

   Beachte, daß die Folgenglieder f_n stetig sind, nicht aber die Grenzfunktion.

2. Die Folge $f_n(x) := \sum_{i=0}^{n} \frac{x^i}{i!}$ konvergiert punktweise gegen die Funktion $\mathtt{e}(x)$.

Ist der Bildraum der Funktionen $f_n : A \to X$, $f : A \to X$ ein metrischer Raum, z.B $\mathbb{R}$ mit dem Abstand $d(x,y) = |x-y|$, dann kann man von gleichmäßiger Konvergenz sprechen.

Definition 4.66 *Eine Folge von Funktionen (f_n) konvergiert* ***gleichmäßig*** *gegen f, falls*

$$\lim_{n\to\infty} \sup_{a\in A} d(f_n(a), f(a)) = 0$$

gilt. Eine Folge von Funktionen (f_n) ist eine ***gleichmäßige Cauchyfolge****, wenn für jedes $\epsilon > 0$ ein n_0 existiert, so daß aus $n, m \geq n_0$ folgt*

$$\sup_{a\in A} d(f_n(a), f_m(a)) < \epsilon .$$

Sei $x_0 \in X$ fest. Eine Abbildung $A \to X$ heiße bezüglich x_0 beschränkt, wenn die Menge

$$\{d(f(a), x_0 | a \in A\} \subseteq \mathbb{R}$$

beschränkt ist. Wenn A kompakt ist, dann ist eine stetige Funktion $f : A \to X$ bezüglich jedem Punkt $x_0 \in X$ beschränkt. In der Tat ist die Abbildung $a \mapsto d(f(a), x_0)$ stetig und damit beschränkt. Sei $C(A, (X, x_0))$ die Menge der bezüglich x_0 beschränkten Abbildungen. Dann ist

$$d(f,g) := \sup_{a\in A} d(f(a), g(a)) \tag{1}$$

eine Metrik auf $C(A, (X, x_0))$. *Die gleichmäßige Konvergenz einer Folge (f_n) in $C(A, (X, x_0))$ ist die Konvergenz im metrischen Raum $C(A, (X, x_0)), d)$.*

Die Abbildung $f(x) := x : \mathbb{R} \to \mathbb{R}$ ist aber bezüglich keinem Punkt von $\mathbb{R}$ beschränkt. Trotzdem kann man etwa zeigen, daß $(f_n(x) := x + \frac{1}{|x|+n})$ gleichmäßig gegen f konvergiert. Der Begriff der gleichmäßigen Konvergenz läßt sich also in einem allgemeineren Kontext anwenden.

Im folgenden betrachten wir Abbildung aus einem topologischen Raum A mit Werten in einem metrischen Raum (X, d).

Lemma 4.67 *Sei (f_n) eine gleichmäßig gegen f konvergente Folge stetiger Abbildungen. Dann ist f stetig.*

Beweis: Sei $a \in A$ und $x := f(a)$. Sei $\epsilon > 0$ gegeben. Wir müssen zeigen, daß $f^{-1}(B(x, \epsilon))$ eine Umgebung von a ist. Sei nun n so groß, daß

$$\sup_{b \in A} d(f_n(b), f(b)) < \frac{\epsilon}{2}$$

gilt. Dann ist $N := f_n^{-1}(B(x, \frac{\epsilon}{2}))$ eine Umgebung von a. Für $b \in N$ gilt

$$d(f(b), x) \leq d(f(b), f_n(b)) + d(f_n(b), x) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \ .$$

Damit gilt $f(N) \subseteq B(x, \epsilon)$. Folglich ist auch $f^{-1}(B(x, \epsilon))$ eine Umgebung von A, da diese Menge N enthält. ■

Lemma 4.68 *Wenn (X, d) vollständig ist, dann ist auch der Raum $C(A, (X, x_0))$ mit der Metrik (1) ist vollständig. Allgemeiner hat dann eine gleichmäßige Cauchyfolge stetiger Abbildungen einen stetigen Grenzwert.*

Beweis: Sei (f_n) eine Cauchyfolge. Dann ist für jedes $a \in A$ die Folge $(f_n(a))$ eine Cauchyfolge in X und wegen der angenommenen Vollständigkeit von X konvergent. Wir setzen

$$f(a) := \lim_{n \to \infty} f_n(a) \ .$$

Wir zeigen nun, daß f_n gleichmäßig gegen f konvergiert. Sei $\epsilon > 0$ gegeben. Dann wählen wir ein $n_0 \in \mathbb{N}$ derart, daß für alle $n, m \geq n_0$ gilt

$$\sup_{b \in A} d(f_n(b), f_m(b)) < \frac{\epsilon}{2} \ .$$

Dann gilt aber für $n \geq n_0$ und $b \in A$ auch

$$d(f_n(b), f(b)) = \lim_{m \to \infty} d(f_n(b), f_m(b)) \leq \frac{\epsilon}{2} \ ,$$

also

$$\sup_{b \in A} d(f_n(b), f(b)) < \epsilon \ .$$

Aus Lemma 4.67 folgt die Stetigkeit von f. Desweiteren gilt für alle $b \in B$ $d(f(b), x_0) \leq d(f_{n_0}(b), x_0) + \epsilon$, woraus $f \in C(A, (X, x_0))$ folgt. ■

Lemma 4.69 *Sei X ein topologischer Raum und (f_n) eine Folge von reellen Funktionen auf X. Sei weiter (a_n) eine Folge reeller Zahlen derart, daß $|f_n(x)| \leq a_n$ für alle $x \in X$ gilt und $\sum_{n=1}^{\infty} a_n$ konvergent ist. Dann konvergiert die Folge $(\sum_{i=1}^{n} f_i)$ gleichmäßig.*

Beweis: Zuerst sehen wir, daß nach dem Majorantenkriterium

$$f(x) := \sum_{n=1}^{\infty} f_n(x)$$

für alle $x \in X$ existiert. Sei $\epsilon > 0$ gegeben. Dann gibt es ein $n_0 \in \mathbb{N}$ derart, daß $\sum_{n=n_0+1}^{\infty} a_n < \epsilon$ gilt. Dann gilt aber für $n \geq n_0$ und $x \in X$, daß

$$|f(x) - f_n(x)| = |\sum_{i=n+1}^{\infty} f_n(x)| \leq \sum_{i=n+1}^{\infty} a_i \leq \sum_{n=n_0+1}^{\infty} a_n < \epsilon \ .$$

■

Wir betrachten nun eine durch eine Potenzreihe gegebene Funktion $f(x) := \sum_{n=0}^{\infty} a_n x^n$.

Lemma 4.70 *Sei $c > 0$ und die Reihe für alle $x \in [-c, c]$ absolut konvergent. Dann konvergiert die Reihe auf $[-c, c]$ gleichmäßig und $f : [-c, c] \to \mathbb{R}$ ist stetig.*

Beweis: Es gilt $|a_n x^n| \leq |a_n c^n|$ für alle $x \in [-a, a]$ und $\sum_{n=0}^{\infty} |a_n c^n| < \infty$. Folglich konvergiert $\sum_{n=0}^{\infty} a_n x^n$ auf $[-c, c]$ gleichmäßig und definiert eine stetige Funktion. ■

Die Folgen $(\sum_{i=0}^{n} \frac{x^i}{i!})$, $(\sum_{i=0}^{N} (-1)^i \frac{x^{2i+1}}{(2i+1)!}$ und $(\sum_{i=0}^{N} (-1)^i \frac{x^{2i}}{(2i)!}$ konvergieren also auf jedem beschränkten Intervall gleichmäßig gegen e^x, $\sin(x)$ und $\cos(x)$. Die Folge $\sum_{n=1}^{\infty} \frac{x^n}{n}$ konvergiert auf jedem kompakten Intervall in $(-1, 1)$ gleichmäßig gegen $\ln(1 - x)$.

Korollar 4.71 *Die Funktionen $x \mapsto \sin(x)$, $x \mapsto \cos(x)$, $x \mapsto \mathsf{e}^x$ sind auf ganz $\mathbb{R}$ stetig. Die Funktion $(-1, 1) \ni x \mapsto \ln(1 - x)$ ist stetig.*

5 Differentialrechnung

5.1 Die Ableitung

Sei $U \subseteq \mathbb{R}$ eine offene Teilmenge, $x \in U$ und $f : U \to \mathbb{R}$. Wir betrachten den Differenzenquotienten

$$\Delta_x(f)(y) := \frac{f(y) - f(x)}{y - x}$$

auf $y \in U \setminus \{x\}$. Der Punkt x ist ein Häufungspunkt dieser Menge.

Definition 5.1 *Wenn*

$$f'(x) := \lim_{y \to x} \frac{f(y) - f(x)}{y - x}$$

existiert, dann heißt f im Punkt x ***differenzierbar*** *und $f'(x)$ die* ***Ableitung*** *von f im Punkt x. Wenn f in allen Punkten von U differenzierbar ist, dann sagen wir, daß f auf U differenzierbar ist und nennen die Funktion $f' : U \to \mathbb{R}$ die Ableitung.*

Wir werden oft eine äquivalente Charakterisierung der Ableitung benutzen.

Lemma 5.2 *Die Funktion $f : U \to \mathbb{R}$ ist genau in $x \in U$ differenzierbar, wenn es eine Zahl $a \in \mathbb{R}$ und eine Funktion $r : U \to \mathbb{R}$ gibt, so daß*

1. $f(y) = f(x) + a(y - x) + r(y)(y - x)$
2. $\lim_{y \to x} r(y) = 0$

gilt. In diesem Fall ist $a = f'(x)$.

Beweis: Beide Richtungen folgen aus der Gleichung

$$r(y) = \Delta_x(f)(y) - a$$

für $y \neq x$. In der Tat, wenn r die geforderten Eigenschaften hat, dann gilt $\lim_{y \to x} \Delta_x(f)(y) = a$.

Umgekehrt, wenn die Ableitung von f existiert, dann hat die Funktion r (durch 0 nach $y = x$ fortgesetzt) die geforderten Eigenschaften. ■

Lemma 5.3 *Ist f in x differenzierbar, so ist f auch stetig.*

Beweis: Wir schreiben $f(y) = f(x) + a(y - x) + r(y)(y - x)$. Dann gilt offensichtlich $\lim_{y \to x} f(y) = f(x)$. ■

Die Umkehrung dieser Aussage gilt nicht. So ist die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := |x|$ in $x = 0$ stetig, aber nicht differenzierbar. In der Tat besitzt $\Delta_0(f)(y) = \mathtt{sign}(y)$ keinen Grenzwert in $y = 0$.

Wir betrachten eine offene Teilmenge $U \subseteq \mathbb{R}$ und $x \in U$. Sei $V \subseteq \mathbb{R}$ eine weitere offene Teilmenge, $f : U \to V$ und $g : V \to \mathbb{R}$ eine weitere Funktion.

Lemma 5.4 (Kettenregel) *Ist f in x und g in $f(x)$ differenzierbar, so ist $g \circ f$ in x differenzierbar und es gilt $(g \circ f)'(x) = g'(f(x))f'(x)$.*

Beweis: Wir schreiben

$$f(y) = f(x) + f'(x)(y-x) + r(y)(y-x) \,, \quad g(u) = g(f(x)) - g'(f(x))(u - f(x)) + s(u)(u - f(x)) \,.$$

Dann gilt

$$
\begin{aligned}
g(f(y)) &= g(f(x)) - g'(f(x))(f(y) - f(x)) + s(f(y)(f(y) - f(x))) \\
&= g(f(x)) - g'(f(x))(f(x) + f'(x)(y-x) + r(y)(y-x) - f(x)) + \\
&\quad s(f(y))(f(x) + f'(x)(y-x) + r(y)(y-x) - f(x)) \\
&= g(f(x)) + g'(f(x))f'(x)(y-x) + [r(y) + s(f(y))(f'(x) + r(y))](y-x) \; .
\end{aligned}
$$

Nun gilt $\lim_{y\to x}[r(y) + s(f(y))(f'(x) + r(y))] = 0$. Dies zeigt die Behauptung. ■

Lemma 5.5 (Produktregel) *Seien $f, g : U \to \mathbb{R}$ im Punkt $x \in U$ differenzierbar. Dann ist $f + g$ und fg im Punkt x differenzierbar und es gilt*

$$
(f+g)'(x) = f'(x) + g'(x) \; , \quad (fg)'(x) = f'(x)g(x) + f(x)g'(x) \; .
$$

Beweis: Wir zeigen nur die Produktregel. Wir schreiben

$$
f(y) = f(x) + f'(x)(y-x) + r(y)(y-x) \; , \quad g(y) = g(x) + g'(x)(y-x) + s(y)(y-x) \; .
$$

Dann gilt

$$
\begin{aligned}
(fg)(y) &= (f(x) + f'(x)(y-x) + r(y)(y-x))(g(x) + g'(x)(y-x) + s(y)(y-x)) \\
&= f(x)g(x) + [f'(x)g(x) + f(x)g'(x)](y-x) + [f(y)s(y) + r(y)g(y)](y-x) \; .
\end{aligned}
$$

Es gilt $\lim_{y\to x}[f(y)s(y) + r(y)g(y)] = 0$. ■

Lemma 5.6 *Die Funktion $\mathbb{R} \setminus \{0\} \ni x \to x^{-1}$ ist differenzierbar und es gilt $(x^{-1})' = -x^{-2}$.*

Beweis: Wir erweitern den Differenzenquotienten mit xy:

$$
\frac{y^{-1} - x^{-1}}{y - x} = \frac{x - y}{xy(y-x)} = \frac{-1}{xy} \xrightarrow{y \to x} -x^{-2} \; .
$$

■

Lemma 5.7 (Umkehrfunktion) *Sei $U \subseteq \mathbb{R}$ ein Intervall, $f : U \to V$ streng monoton, stetig und in $x \in U$ differenzierbar mit $f'(x) \neq 0$. Dann ist $f^{-1} : V \to U$ in $f(x)$ differenzierbar, und es gilt $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$.*

Beweis:

Sei (v_n) eine Folge in $V \setminus \{f(x)\}$ mit $v_n \to f(x)$. Sei $g := f^{-1}$. Wir wissen schon, daß g als Umkehrfunktion einer streng monotonen stetigen Funktion stetig ist (Lemma 4.16). Dann gilt mit $x_n := g(v_n)$ auch $x_n \to x$ und

$$\frac{g(v_n) - g(f(x))}{v_n - f(x)} = \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{\frac{f(x_n)-f(x)}{x_n - x}} \to \frac{1}{f'(x)} \ .$$

■

Hier sind einige Beispiele:

1. Wir hatten schon gesehen, daß für $n \in \mathbb{N}$ $(x^n)' = nx^{n-1}$ gilt.

2. Es gilt $\sin'(x) = \cos(x)$, $\cos'(x) = -\sin(x)$ und $(\mathsf{e}^x)' = \mathsf{e}^x$.

3. $f(x) := \mathsf{e}^{\sin(x)}$ ist auf ganz $\mathbb{R}$ differenzierbar und es gilt $f'(x) = \cos(x)\mathsf{e}^{\sin(x)}$.

4. $$f(x) := \begin{cases} x\sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$
   ist auf $\mathbb{R} \setminus \{0\}$, nicht aber in 0 differenzierbar. Es gilt für $x \neq 0$ daß
   $$f'(x) = \sin(\frac{1}{x}) - \frac{\cos(\frac{1}{x})}{x} \ .$$
   Der Differenzenquotient $\Delta_0(f)(y) = \sin(\frac{1}{y})$ hat an der Stelle 0 keinen Grenzwert.

5. Die Funktion
   $$f(x) := \begin{cases} x^2\sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$
   ist auf ganz $\mathbb{R}$ differenzierbar und hat die Ableitung (für $x \neq 0$)
   $$f'(x) = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x}) \ .$$
   Wir betrachten wieder $\Delta_0(f)(y) = y\sin(\frac{1}{y})$. In diesem Fall existiert der Grenzwert für $y \to 0$ und es gilt $f'(0) = 0$. Da $\cos(\frac{1}{x})$ für $x \to 0$ keinen Grenzwert hat, ist die Ableitung im Punkt 0 nicht stetig.

6. Die e-Funktion ist streng monoton wachsend und stetig. Damit existiert die Umkehrfunktion $\ln : (0, \infty) \to \mathbb{R}$ derart, daß $\mathsf{e}^{\ln(x)} = x$ und $\ln(\mathsf{e}^x) = x$ gilt. Die Funktion $\ln$ ist stetig und streng monoton wachsend. Da e differenzierbar ist, ist auch $\ln$ differenzierbar. Es gilt $\ln'(\mathsf{e}^x) = \frac{1}{(\mathsf{e}^x)'}$, also
   $$\ln'(x) = \frac{1}{x} \ .$$

Es gilt wegen $\mathsf{e}^{x+y} = \mathsf{e}^x \mathsf{e}^y$, daß

$$\ln(xy) = \ln(x) + \ln(y) \ .$$

Sei $r \in (0, \infty)$. Dann gilt $r = e^{\ln(r)}$ und für $x \in \mathbb{R}$, daß $r^x = (\mathsf{e}^{\ln(r)})^x = \mathsf{e}^{x \ln(r)}$. Wir schließen, daß $x \to r^x$ auf ganz $\mathbb{R}$ differenzierbar ist und die Ableitung

$$(r^x)' = \ln(r) r^x$$

hat.

7. Sei nun $r \in \mathbb{R}$. Für $x > 0$ schreiben wir wieder $x^r = \mathsf{e}^{r \ln(x)}$. Wir sehen, daß die Funktion $x^r : (0, \infty) \to \mathbb{R}$ differenzierbar ist und die Ableitung

$$(x^r)' = \frac{r}{x} x^r = r x^{r-1}$$

hat.

8. Die Funktion

$$f(x) := \begin{cases} |x|^r & x \neq 0 \\ 0 & x = 0 \end{cases}$$

ist für $r > 0$ stetig, und für $r > 1$ differenzierbar mit

$$f'(x) = \begin{cases} \mathtt{sign}(x)|x|^{r-1} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

5.2 Mittelwerte und Extremwerte

Sei X ein topologischer Raum und $f : X \to \mathbb{R}$.

Definition 5.8 *Wir sagen, daß f im Punkt $x \in X$ ein **lokales Extremum** besitzt, wenn es eine Umgebung $x \in U \subseteq X$ von x gibt mit $\sup_U f = f(x)$. Analog definieren wir den Begriff eines lokalen Minimums.*

Lemma 5.9 *Sei $U \subseteq \mathbb{R}$, $f : U \to \mathbb{R}$ in $x \in U$ differenzierbar und habe in diesem Punkt ein lokales Maximum (oder Minimum). Dann gilt $f'(x) = 0$.*

Beweis: Wir schreiben $\Delta_x(f)(y) = \frac{f(y)-f(x)}{y-x}$. Dann ist $\Delta_x(f)(y) \leq 0$ für $y < x$ und > 0 für $y > x$. Folglich gilt $\lim_{y \to x} \Delta_x(f)(y) = f'(x) = 0$. ∎

Definition 5.10 *Ist $U \subseteq \mathbb{R}$ offen und $f : U \to \mathbb{R}$ differenzierbar, dann heißt ein Punkt $x \in U$ **kritisch** (für f), wenn $f'(x) = 0$ gilt.*

Lemma 5.11 *Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gibt es ein $\xi \in (a, b)$ mit*

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

Beweis: Wir betrachten die auf $[a, b]$ stetige und auf (a, b) differenzierbare Funktion $h(x) := (f(b) - f(a))x - (b - a)f(x)$. Es gilt $h(a) = f(b)a - bf(a) = h(b)$ und $h'(x) = f(b) - f(a) - (b - a)f'(x)$. Wenn h konstant ist, dann gilt $f(x) = \frac{x(f(b)-f(a))-h}{b-a}$ und der Satz stimmt offensichtlich. Andernfalls gilt $h(t) > h(a)$ oder $h(t) < h(a)$ für ein $t \in (a, b)$. Im ersten Fall finden wir eine absolutes Maximum in einem Punkt $\xi \in (a, b)$, im zweiten Fall eine absolutes Minimum. In beiden Fällen gilt $h'(\xi) = 0$, also $f(b) - f(a) = (b - a)f'(\xi)$. ■

Lemma 5.12 *Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar.*

1. *Wenn $f'(x) \geq 0$ für alle $x \in (a, b)$ gilt, dann ist f monoton wachsend.*
2. *Wenn $f'(x) = 0$ für alle $x \in (a, b)$ gilt, dann ist f konstant.*
3. *Wenn $f'(x) \leq 0$ für alle $x \in (a, b)$ gilt, dann ist f monoton fallend.*

Beweis: Ist $x, y \in [a, b]$, $x < y$, dann gilt $f(y) - f(x) = f'(\xi)(y - x) \geq 0$ für ein $\xi \in (x, y)$. Die dritte Aussage folgt aus der ersten. Die zweite Aussage folgt, da eine Funktion, welche gleichzeitig monoton wächst und fällt, konstant sein muß. ■

Lemma 5.13 *Sei $f : (a, b) \to \mathbb{R}$ differenzierbar und $x \in (a, b)$. Wenn es ein $\epsilon > 0$ mit $f'(y) \geq 0$ für $y \in (x - \epsilon, x)$ und $f'(y) \leq 0$ für $y \in (x, x + \epsilon)$ gibt, dann ist x die Stelle eines lokalen Maximums. Diese Voraussetzung ist insbesondere dann erfüllt, wenn f' in x differenzierbar ist und $f''(x) < 0$ gilt.*

Beweis: In der Tat wächst f auf $(x - \epsilon, x)$ monoton und fällt auf $(x, x + \epsilon)$. Daraus folgt $\sup_{(x-\epsilon,x+\epsilon)} f = f(x)$.

Ist die Funktion $(a, b) \ni y \mapsto f'(y)$ im Punkt x differenzierbar, dann ist $f'(x) = 0$ wegen der Stetigkeit von f' im Punkt x. Wenn $f''(x) < 0$ ist, dann ist

$$\Delta_x(f')(y) = \frac{f'(y)}{y - x} < 0$$

in einer Umgebung von x. Folglich gibt es $\epsilon > 0$ derart, daß $f'(y) \geq 0$ für $y \in (x - \epsilon, x)$ und $f'(y) \leq 0$ für $y \in (x, x + \epsilon)$. ■

Wir halten fest:

Korollar 5.14 *1. Sei $f : (a, b) \to \mathbb{R}$ differenzierbar.*

2. *Sei $x \in (a, b)$ ein kritischer Punkte von f ($f'(x) = 0$).*
3. *Sei f' in x differenzierbar und $f''(x) < 0$ (> 0).*
Dann hat f im Punkt x ein lokales Maximum (Minimum).

1. Die Funktion $f(x) = x^2$ erfüllt $f'(0) = 0$ und $f''(0) = 2 > 0$. Sie hat im Punkt 0 ein Minimum.

2. Die Funktion $f(x) = x^3$ erfüllt $f'(0) = 0$ und $f''(0) = 0$. Sie hat am Punkt 0 keinen lokalen Extremwert.

3. Die Funktion $f(x) = x^4$ erfüllt $f'(0) = 0$ und $f''(0) = 0$. Sie hat am Punkt 0 ein lokales Minimum. Das folgt aber nicht aus dem obigen Korollar.

4. Die Funktion $f(x) = |x|$ ist im Punkt 0 nicht differenzierbar. Sie hat dort aber ein lokales Extremum.

Für Ableitungen gilt eine Art Zwischenwertsatz.

Lemma 5.15 *Sei $f : U \to \mathbb{R}$ differenzierbar, $[a, b] \subset U$ und $f'(a) < \lambda < f'(b)$. Dann existiert $x \in (a, b)$ mit $f'(x) = \lambda$.*

Beweis: Sei $g(x) = f(x) - \lambda x$. Dann ist $g'(a) < 0$ und es gibt ein $x_1 \in (a, b)$ mit $g(x_1) < g(a)$. Analog gilt $g'(b) > 0$ und es gibt $x_2 \in (a, b)$ mit $g(x_2) < g(b)$. Die stetige Funktion $g : [a, b] \to \mathbb{R}$ hat ihr absolutes Minimum also in $x \in (a, b)$. Dann gilt $g'(x) = 0$, also $f'(x) = \lambda$. ■

5.3 Die trigonometrischen Funktionen, lin. Differentialgleichungen

Wir betrachten das Anfangswertproblem für die lineare Differentialgleichung erster Ordnung mit konstanten Koeffizienten ($a \in \mathbb{R}$)

$$f' = af \ , \quad f(0) = c$$

für eine Funktion $f : \mathbb{R} \to \mathbb{R}$. Wir überzeugen uns durch Einsetzen, daß

$$f(t) := c\mathrm{e}^{at}$$

eine Lösung dieses Problems ist. In der Tat ist dies die einzige. Ist nämlich g eine Lösung, dann wäre

$$h(t) := e^{-at} g(t)$$

eine Lösung des Problems

$$h'(t) = 0 \ , \quad h(0) = c \ .$$

Daraus folgt $h \equiv c$, also $g(t) = c\mathrm{e}^{at}$.

Wir betrachten nun das folgende Anfangswertproblem für die lineare Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten ($a \in \mathbb{R} \setminus \{0\}$)

$$f'' + a^2 f = 0 \ , \quad f(0) = c \, f'(0) = d \ .$$

Wir überzeugen uns davon, daß die Funktionen $\sin(at)$ und $\cos(at)$ und damit auch jede Linearkombination

$$u \sin(at) + v \cos(at) \ , \quad u, v \in \mathbb{R}$$

dieser Differentialgleichung genügen. Die Anfangsbedingungen ergeben $v = c$ und $u = a^{-1}d$. Wir sehen, daß

$$f(t) = \frac{d}{a} \sin(at) + v \cos(at)$$

eine Lösung ist. Diese ist die einzige Lösung. Seien f_0, f_1 zwei Lösungen. Dann ist $h = f_0 - f_1$ eine Lösung von

$$h'' + a^2 h = 0 \ , \quad h(0) = c \, h'(0) = d \ .$$

Wir betrachten $E = a^2 h^2 + (h')^2$. Es gilt wegen der Differentialgleichung

$$E' = 2a^2 h h' + 2h' h'' = 2h'(a^2 h + h'') = 0 \ .$$

Aus der Anfangsbedingung folgt $E(0) = 0$ und somit $E \equiv 0$. Daraus folgt $h \equiv 0$.

Lemma 5.16 *Die Funktion* cos *hat Nullstellen auf* $[0, \infty)$.

Beweis: Möge cos keine Nullstelle haben. Dann gilt wegen Zwischenwertsatz und $\cos(0) = 1$, daß $\cos(x) > 0$ für alle $x \in \mathbb{R}$. Wegen $\sin' = \cos$ folgt, daß sin streng monoton wächst. Da $\sin \not\equiv 0$ gilt, gibt es ein $c > 0$ derart, daß $\sin(x) \geq c$ für $x \geq 1$. Wegen $\cos' = -\sin$ gilt dann $\cos'(x) \leq -c$ für alle $x \geq 1$. Wir betrachten $h(x) := \cos(x) + c(x-1) - \cos(1)$. Es gilt für $x \geq 1$, daß $h'(x) = -\sin(x) + c \leq 0$. Folglich fällt h monoton. Da $h(1) = 0$ gilt, ist $h(x) \leq 0$ für alle $x \geq 1$. Daraus folgt $-cx \geq \cos(x) - \cos(1) - c \geq -1 - \cos(1) - c$ für alle $x \geq 1$, was unmöglich ist. Also hat $\cos(x)$ eine Nullstelle. ■

Definition 5.17 *Wir definieren die Zahl*

$$\pi := 2 \inf\{x \in [0, \infty) \cos(x) = 0\} \ .$$

Es gilt also $\cos(\frac{\pi}{2}) = 0$ und dies ist die kleinste positive Nullstelle von $\cos(x)$. Wegen $\cos^2 + \sin^2 = 1$ gilt $\sin(\frac{\pi}{2}) = \pm 1$. Da $\sin(0) = 0$ und $\sin(x)$ auf $[0, \frac{\pi}{2}]$ monoton wächst (wegen $\sin'(x) = \cos(x) \geq 0$ für diese x), gilt $\sin(\frac{\pi}{2}) = 1$.

Aus dem Additionstheorem $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$ folgt mit $y = \frac{\pi}{2}$, daß

$$\sin(x + \frac{\pi}{2}) = \cos(x)$$

gilt. Wir schließen, daß $\sin(\pi) = 0$ gilt, und daß π die erste positive Nullstelle von sin ist.

Aus dem Additionstheorem $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ schließen wir mit $y = \frac{\pi}{2}$, daß

$$\cos(x + \frac{\pi}{2}) = -\sin(x) \ .$$

Folglich gilt

$$\sin(x+\pi) = -\sin(x) \ , \quad \sin(x+2\pi) = \sin(x) \ .$$

Daraus folgt nun auch

$$\cos(x+\pi) = -\cos(x) \ , \quad \cos(x+2\pi) = \cos(x) \ .$$

Die Funktionen sin und cos sind also 2π-periodisch. Die Funktion $\sin'(x) := \cos(x)$ ist auf dem Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ positiv. Damit ist die Funktion $\sin(x)$ auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend mit dem Bildbereich $[-1, 1]$.

Definition 5.18 *Die Umkehrfunktion der Abbildung* $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ *ist die Funktion* $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$.

Die Ableitung ist durch

$$\arcsin'(\sin(x)) = \frac{1}{\sin'(x)} = \frac{1}{\cos(x)}$$

gegeben. Wir setzen $\sin(x) = z$. Dann ist $\cos(x) = \sqrt{1-z^2}$ und

$$\arcsin'(z) = \frac{1}{\sqrt{1-z^2}} \ .$$

Wegen $\cos(x) = -\sin(x - \frac{\pi}{2})$ ist Umkehrfunktion von $\cos : [0, \pi] \to [-1, 1]$ durch

$$\arccos(z) = -\arcsin(z) + \frac{\pi}{2} : [-1, 1] \to [0, \pi]$$

gegeben. Es gilt

$$\arccos'(z) = -\frac{1}{\sqrt{1-z^2}} \ .$$

Wir definieren die Tangensfunktion

$$\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$$

durch

$$\tan(x) := \frac{\sin(x)}{\cos(x)} \ .$$

Es gilt

$$\tan'(x) = \frac{\cos(x)\sin'(x) - \cos'(x)\sin(x)}{\cos(x)^2} = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2} \ .$$

Wir sehen insbesondere, daß $\tan'(x) > 0$ ist. Die Tangensfunktion wächst streng monoton und ist wegen

$$\lim_{x \to \pm\frac{\pi}{2}} \tan(x) = \pm\infty$$

surjektiv. Folglich haben wir eine Umkehrfunktion

$$\arctan(x) = \tan^{-1}(x) \ .$$

Es gilt

$$\arctan'(\tan(x)) = \frac{1}{\tan'(x)} = \cos(x)^2 \ .$$

Es gilt nun

$$\tan(x)^2 = \frac{\sin(x)^2}{\cos(x)^2} = \frac{1-\cos(x)^2}{\cos(x)^2} \ .$$

Wir stellen nach $\cos(x)^2$ um und erhalten

$$\cos(x)^2 = \frac{1}{1+\tan(x)^2} \ .$$

Wir sehen, daß

$$\arctan'(z) = \frac{1}{1+z^2}$$

gilt.

5.4 Taylorformel

Sei $U \subseteq \mathbb{R}$ offen und $f : U \to \mathbb{R}$. Wenn f auf U differenzierbar ist, dann kann man die Ableitung $f' : U \to \mathbb{R}$ betrachten. Ist diese differenzierbar, dann kann man die zweite Ableitung $f'' = f^{(2)} : U \to \mathbb{R}$ bilden, usw. Sei f mindestens n mal differenzierbar.

Definition 5.19 *Das **Taylorpolynom** von f im Punkt x vom Grad n ist*

$$P(y) := \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k \ .$$

Ist f ein Polynom vom Grad $\leq n$, dann gilt $f(y) = P(y)$. In der Tat ist für $f(y) = y^l$

$$f^{(k)}(x) = \frac{l!}{(l-k)!}x^{l-k} \ .$$

Damit gilt mit der binomischen Formel

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

Im allgemeinen gilt

$$f(y) = P(y) + r(y)$$

für einen Rest r. Der Taylorsche Satz gibt eine obere Abschätzung für diesen Rest.

Satz 5.20 (Taylorformel) *Sei $[a,b] \subset U$ mit $a < x < b$. Wir machen folgende Annahmen über $f : U \to \mathbb{R}$:*

1. *Auf U ist f ist n-mal differenzierbar.*
2. *$f^{(n)}$ ist auf $[a,b]$ stetig.*
3. *$f^{(n+1)}$ existiert in (a,b).*

Dann gibt es für jedes $x \neq y \in (a,b)$ ein $\xi \in (x,y)$ (oder $\xi \in (y,x)$, falls $y < x$) so daß

$$r(y) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(y-x)^{n+1} .$$

Beweis: Wir definieren $M \in \mathbb{R}$ durch $r(y) = f(z) - P(z) = M(y-x)^{n+1}$. Dann setzen wir

$$g(z) := f(z) - P(z) - M(z-x)^{n+1} .$$

Es gilt

$$g^{(n+1)}(z) = f^{(n+1)}(z) - P^{(n+1)}(z) - (n+1)!M = f^{(n+1)}(z) - (n+1)!M .$$

Es ist zu zeigen, daß $g^{(n+1)}(z)$ eine Nullstelle in (y,x) (oder (x,y)) hat.

Wir nehmen an, daß $y < x$ ist. Es gilt für $k = 0, \ldots, n$, daß $g^{(k)}(x) = 0$. Desweiteren gilt nach der Konstruktion von M auch $g(y) = 0$. Folglich gibt es ein $\xi_1 \in (y,x)$ mit $g^{(1)}(\xi) = 0$. Wir schließen nun, daß es ein $\xi_2 \in (\xi_1, x) \subset (y,x)$ mit $g^{(2)}(\xi_2) = 0$ gibt. Wir fahren so induktiv fort bis wir ein $\xi \in (y,x)$ mit $g^{(n+1)}(\xi) = 0$ finden. ■

1. Es gilt $(\mathsf{e}^x)^{(n)}_{|x=0} = \mathsf{e}^x_{|x=0} = 1$. Folglich ist für $n \geq 1$ und geeignetes $\xi_k \in (0,x)$

   $$\mathsf{e}^x = \sum_{n=0}^{k} \frac{1}{n!}x^n + \frac{1}{(k+1)!}\mathsf{e}^{\xi_k}x^{k+1} .$$

   Für festes x gilt

   $$|\frac{1}{(k+1)!}\mathsf{e}^{\xi_k}x^{k+1}| \leq |\frac{1}{(k+1)!}\mathsf{e}^{x}x^{k+1}| \stackrel{k\to\infty}{\to} 0 .$$

   Die e-Funktion wird durch ihre Taylorreihe dargestellt. Mit dem Restglied erhalten wir sogar eine explizite Abschätzung des Fehlers. Wenn wir zum Beispiel e^{-1} berechnen wollen, dann ist

   $$|\mathsf{e}^{-1} - \sum_{n=0}^{k} \frac{(-1)^n}{n!}| \leq \frac{1}{(k+1)!} .$$

   Wollen wir die ersten drei Stellen, dann muß $(k+1)! > 10^3$ sein, also $k \geq 6$.

2. Es gilt für $k \equiv 0, 1, 2, 3 \mod 4$ daß $\sin^{(k)}(0) = 0, 1, 0, -1$. Folglich ist

$$\sin(x) = \sum_{n=0}^{k} (-1)^n \frac{1}{2n+1!} x^{2k+2} + \frac{1}{(2k+2)!} (-1)^{k+1} \cos(\xi_k) \ .$$

Wegen $|\cos(\xi_k)| \leq 1$ gilt

$$|\frac{1}{(2k+2)!} \cos(\xi_k) x^{2k+2}| \overset{k\to\infty}{\to} 0 \ .$$

Wollen wir also für $x \in (-\pi, \pi)$ den Wert von $\sin(x)$ auf drei Stellen genau, dann reicht es, k so groß zu wählen, daß $\frac{\pi^{2k+2}}{(2k+2)!} < 10^{-3}$ gilt, z.b. $k = 7$.

3. Es gilt $\ln(1-x)^{(1)} = \frac{-1}{1-x}$ und $\ln^{(n)}(1-x) = -(n-1)!(1-x)^{1-n}$. Folglich gilt

$$\ln(1-x) = -\sum_{n=1}^{k} \frac{x^n}{n} - \frac{1}{k+1} \frac{\xi_k^{k+1}}{(1-x)^{k+1}}$$

mit $\xi_k \in (0, x)$. Für $|x| \leq \frac{1}{2}$ gilt $|\frac{1}{k+1} \frac{\xi_k^{k+1}}{(1-x)^{k+1}}| \to 0$ und $\ln(1-x)$ wird durch die Taylorreihe dargestellt. Wenn wir etwa $\ln(x)$ für $x \in (\frac{1}{2}, 1 + \frac{1}{2})$ auf drei Stellen berechnen wollen, dann müssen wir $k \geq 1000$ wählen. Obwohl wir wissen, daß die Taylorreihe für $|x| < 1$ konvergiert, erhalten wir aus dem Restglied für $|x| > \frac{1}{2}$ keine brauchbare Abschätzung.

4. Wir betrachten die Funktionen

$$f_k(x) := \begin{cases} 0 & x \leq 0 \\ x^k e^{-\frac{1}{x}} & x > 0 \end{cases}$$

für $k \in \mathbb{Z}$. Diese Funktionen sind für $x \neq 0$ differenzierbar. Wir betrachten die Stelle $x = 0$. Es gilt für $y \geq 0$

$$\Delta_0(f_k)(y) = \frac{e^{-\frac{1}{y}}}{y^{1-k}} \overset{y\to 0}{\to} 0 \ .$$

Folglich gilt $f_k'(0) = 0$. Es gilt weiter

$$f'(x) = \begin{cases} 0 & x \leq 0 \\ kx^{k-1} e^{-\frac{1}{x}} + x^{k-2} e^{-\frac{1}{x}} & x > 0 \end{cases} = kf_{k-1} + f_{k-2}$$

Wir sehen, daß f_0 beliebig of differenzierbar ist. Da $f_0^{(n)} = 0$ für all $n \geq 0$ ist, verschwindet die Taylorreihe im Punkt 0 identisch.

Sei $U \subseteq \mathbb{R}$ offen.

Definition 5.21 *Eine Funktion $f : U \to \mathbb{R}$ heißt glatt, wenn sie beliebig oft differenzierbar ist.*

Definition 5.22 *Sei $U \subseteq \mathbb{R}$ und $f : U \to \mathbb{R}$ eine glatte Funktion. Die Funktion f heißt in x analytisch, wenn es eine Umgebung $x \in V \subseteq U$ von x gibt derart, daß*

$$f(y) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!}(y-x)^k$$

für all $y \in V$ gilt (auf welcher also f durch die Taylorreihe dargestellt wird.)

1. Die Funktionen $\mathtt{e}$, ln, sin , cos sind auf ihren Definitionsbereichen analytisch.

2. Die Funktion
$$f(x) := \begin{cases} 0 & x \leq 0 \\ x^k e^{-\frac{1}{x}} & x > 0 \end{cases}$$
ist im Punkt 0 nicht analytisch.

6 Integration

6.1 Definition und Existenz des Integrals

Sei $[a, b] \subset \mathbb{R}$.

Definition 6.1 *Eine Funktion $\phi : [a, b] \to \mathbb{R}$ heißt* ***einfach****, wenn es eine Folge $a = x_0 \leq x_1 \cdots \leq x_{n-1} \leq x_n = b$ (zulässige Zerlegung) gibt, so daß*

$$\phi(x) := \phi(x_{i-1}) \ , \quad \forall x \in [x_{i-1}, x_i) \ , i = 1, \ldots, n$$

gilt.

Wir beobachten, daß die Summe zweier einfacher Funktionen oder Vielfache wieder einfache Funktionen sind. In der Tat, seien $(x_i)_{i=0}^n$ und $(y_j)_{j=0}^m$ zulässige Zerlegungen für ϕ und ψ, dann ist die Folge $(z_k)_{k=0}^{n+m+1}$ der nach der Größe geordneten Zahlen x_i und y_j eine zulässige Zerlegung für $\phi + \psi$ (die gemeinsame Verfeinerung der Zerlegungen). Sei $\mathcal{E}[a, b]$ der Raum der einfachen Funktionen. $\mathcal{E}[a, b]$ ist ein reeller Vektorraum.

Definition 6.2 *Wir definieren das Integral einfacher Funktionen*

$$\int_a^b \ldots dx : \mathcal{E}[a, b] \to \mathbb{R}$$

durch

$$\int_a^b \phi(x)dx := \sum_{i=1}^{n} \phi(x_{i-1})(x_i - x_{i-1})$$

für eine zulässige Zerlegung für ϕ.

Lemma 6.3 *Das Integral ist wohldefiniert und linear. Weiterhin ist es monoton, d.h. aus $\phi \geq 0$ folgt $\int_a^b \phi(x)dx \geq 0$.*

Beweis: Man muß hier die Unabhängigkeit von der Wahl der zulässigen Zerlegung zeigen. Seien $(x_i)_{i=0}^n$ und $(y_j)_{j=0}^m$ zulässige Zerlegungen und $(z_k)_{k=0}^{n+m+1}$ die gemeinsame Verfeinerung. Dann gilt

$$\sum_{i=1}^n \phi(x_{i-1})(x_i - x_{i-1}) = \sum_{i=1}^n \phi(x_{i-1}) \sum_{k,[z_k,z_{k+1})\subset[x_i,x_{i+1})} (z_{k+1} - z_k) = \sum_{k=1}^{n+m+1} \phi(z_{k-1})(z_{k-1} - z_k) \; .$$

Seien nun $\phi, \psi \in \mathcal{E}[a,b]$. Dann können wir eine für ϕ und ψ gemeinsame zulässige Zerlegung wählen. Es gilt

$$\begin{aligned} \int_a^b (\phi + \psi)(x)dx &= \sum_{i=1}^n (\phi(x_{i-1}) + \psi(x_{i-1}))(x_i - x_{i-1}) \\ &= \sum_{i=1}^n \phi(x_{i-1})(x_i - x_{i-1}) + \sum_{i=1}^n \psi(x_{i-1})(x_i - x_{i-1}) \\ &= \int_a^b \phi(x)dx + \int_a^b \psi(x)dx \end{aligned}$$

■

Sei nun $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann gibt es $\phi, \psi \in \mathcal{E}[a,b]$ mit

$$\phi \leq f \leq \phi \; .$$

In der Tat kann man ϕ und ψ konstant mit den Werten einer unteren oder oberen Schranke wählen. Wir nennen das Paar (ϕ, ψ) eine **Zange** von f. Weiterhin gilt in dieser Situation

$$\int_a^b \phi(x)dx \leq \int_a^b \psi(x)dx \; .$$

Definition 6.4 *Wir definieren daß **Oberintegral** und **Unterintegral** von f durch*

$$\begin{aligned} \int_a^{b*} f(x)dx &:= \inf_{\psi\in\mathcal{E}[a,b], f\leq\psi} \int_a^b \psi(x)dx \\ \int_{a*}^b f(x)dx &:= \inf_{\phi\in\mathcal{E}[a,b], \phi\leq f} \int_a^b \psi(x)dx \; . \end{aligned}$$

In der Tat ist die Menge $\{\int_a^b \psi(x)dx | f \leq \psi, \psi \in \mathcal{E}[a,b]\}$ durch $\int_a^b \phi(x)dx$ für ein $\phi \in \mathcal{E}[a,b]$ mit $\phi \leq f$ von unten beschränkt. Es gilt weiter

$$\int_{a*}^b f(x)dx \leq \int_a^{b*} f(x)dx \; .$$

Definition 6.5 *Wir nennen $f : [a,b] \to \mathbb{R}$* ***riemannintegrierbar****, wenn das Oberintegral und das Unterintegral den gleichen Wert hat. In diesem Fall ist*

$$\int_a^b f(x)dx := \int_{a*}^b f(x)dx = \int_a^{b*} f(x)dx$$

das Riemannintegral von f. Mit $R[a,b]$ bezeichnen wir die Menge der Riemannintegrierbaren Funktionen auf $[a,b]$.

Aus der Monotonie und Linearität des Integrals einfacher Funktionen folgt sofort, daß einfache Funktionen Riemannintegrierbar sind und das Riemannintegral mit dem Integral einfacher Funktionen übereinstimmt. Im folgenden werden wir oft die folgende Charakterisierung riemannintegrierbarer Funktionen benutzen. Eine Zange (ϕ, ψ) von f ist eine ϵ-Zange, wenn $\int_a^b (\psi(x) - \phi(x))dx < \epsilon$ gilt.

Lemma 6.6 *Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann riemannintegrierbar, wenn es für jedes $\epsilon > 0$ eine ϵ-Zange von f gibt.*

Beweis: Es gilt $\int_{a*}^b f(x)dx - \int_a^{b_*} f(x)dx < \epsilon$ genau dann, wenn es eine δ-Zange von f für ein $\delta < \epsilon$ gibt.

Umgekehrt, wenn es ϵ-Zange für f gibt, dann ist $\int_{a*}^b f(x)dx - \int_a^{b_*} f(x)dx \leq \epsilon$.

Diese beiden Aussagen implizieren sofort die Behauptung. ■

Theorem 6.7

1. *$R[a,b]$ ist ein reeller Vektorraum (mit den üblichen Operationen)*
2. *Da Integral $\int_a^b \ldots dx : R[a,b] \to \mathbb{R}$ ist linear.*
3. *Das Integral ist monoton, d.h aus $f, g \in R[a,b]$, $f \leq g$ folgt $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.*
4. *Sei $c \in [a,b]$ und $f : [a,b] \to \mathbb{R}$. Dann sind äquivalent:*
   (a) $f \in R[a,b]$
   (b) $f_{[a,c]} \in R[a,c]$ und $f_{|[c,b]} \in R[c,b]$.

   Es gilt unter diesen Voraussetzungen

   $$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \; .$$
5. *Sei f von unten durch m und von oben durch M beschränkt und $\Phi : [m,M] \to \mathbb{R}$ stetig. Ist $f \in R[a,b]$, dann ist auch $\Phi \circ f \in R[a,b]$.*
6. *Wenn $f \in R[a,b]$, dann auch $|f| \in R[a,b]$ und es gilt*

   $$|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx \; .$$

Beweis: Seien $f, g \in R[a,b]$ und $\epsilon > 0$ gegeben. Dann finden wir $\frac{\epsilon}{2}$-Zangen (ϕ, ψ) für f und (λ, μ) für g. Dann gilt

$$\phi + \lambda \leq f + g \leq \psi + \mu$$

und

$$0 \leq \int_a^b (\psi + \mu)(x)dx - \int_a^b (\phi + \lambda)(x) < \epsilon \ .$$

Folglich ist $(\phi + \lambda, \psi + \mu)$ eine ϵ-Zange von $f + g$. Da ϵ beliebig klein gewählt werden kann, folgt $f + g \in R[a,b]$.

Wir sehen weiter, daß

$$\begin{aligned} & |\int_a^b (f+g)(x)dx - \int_a^b f(x)dx - \int_a^b g(x)dx| \\ \leq & \ |\int_a^b (\psi + \mu)(x)dx - \int_a^b \psi(x)dx - \int_a^b \mu(x)dx| + 2\epsilon \\ = & \ 2\epsilon \ . \end{aligned}$$

Daraus folgt

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx \ .$$

Ist $0 < u \in \mathbb{R}$ und $f \in R[a,b]$ dann ist auch $uf \in R[a,b]$. In der Tat, sei $(\phi, \psi) \in \mathcal{E}[a,b]$ eine $\frac{\epsilon}{u}$-Zange von f. Dann gist $(u\phi, u\psi)$ eine ϵ-Zange von uf. Da ϵ beliebig klein gewählt werden kann, gilt $uf \in R[a,b]$. Ähnlich behandelt man die Fälle $u = 0$ und $u < 0$. Man sieht nun ein, daß

$$\int_a^b (uf)(x)dx = u \int_a^b f(x)dx \ .$$

Wir haben damit gezeigt, daß $R[a,b]$ ein Vektorraum und $\int_a^b \dots dx : R[a,b] \to \mathbb{R}$ eine lineare Abbildung ist.

Sei nun $f \leq g$. Dann folgt für $\phi \in \mathcal{E}$ aus $\phi \leq f$ auch $\phi \leq g$. Damit gilt

$$\int_a^b f(x)dx = \int_a^{b*} f(x)dx \leq \int_a^{b*} g(x)dx = \int_a^b g(x) \ .$$

Das ist die Monotonie des Integrals.

Sei $c \in [a,b]$ und $\phi \in \mathcal{E}[a,b]$. Dann gilt $\phi_{|[a,c]} \in \mathcal{E}[a,c]$. Ist zusätzlich $0 \leq \phi$, dann gilt

$$\int_a^c \phi_{|[a,c]}(x)dx \leq \int_a^b \phi(x)dx \ .$$

Die erste Behauptung ist offensichtlich. Für die zweite können wir annehmen, daß c ein Punkt der für g zulässigen Zerlegung ist. Dann gilt

$$\int_a^c \phi_{|[a,c]}(x)dx = \sum_{i, x_{i-1} < c} \phi(x_{i-1})(x_i - x_{i-1}) \leq \sum_{i=1}^k \phi(x_{i-1})(x_i - x_{i-1}) = \int_a^b \phi(x)dx$$

Sei jetzt $f \in R[a,b]$. Wenn nun (ϕ,ψ) eine Zange von f ist, dann ist $(\phi_{|[a,c]},\psi_{|[a,c]})$ eine Zange von $f_{[a,c]}$. Desweiteren gilt

$$\int_a^c (\psi_{|[a,c]} - \phi_{|[a,c]}) \leq \int_a^b (\psi(x)-\phi(x))dx \ .$$

Ist also (ϕ,ψ) eine ϵ-Zange, so auch $(\phi_{|[a,c]},\psi_{|[a,c]})$. Wir schließen $f_{|[a,b]} \in R[a,b]$. Analog zeigt man $f_{[c,b]} \in R[a,b]$.

Seien nun $f_{|[a,c] \in R[a,b]}$ und $f_{|[c,b]} \in R[c,b]$. Seien (ϕ,ψ) eine $\frac{\epsilon}{2}$-Zange von $f_{[a,c]}$ und (λ,μ) eine $\frac{\epsilon}{2}$-Zange von $f_{[c,b]}$. Dann können wir eine ϵ-Zange (κ,δ) von f bilden mit

$$\kappa_{|[a,c)} := \phi \ , \kappa_{[c,b]} := \mu \ , \delta_{[|a,c)} := \psi \ , \delta_{|[c,b]} := \mu \ .$$

In der Tat gilt dann $\kappa,\delta \in \mathcal{E}[a,b]$, , $\kappa \leq f \leq \delta$ und

$$\int_a^b (\delta(x)-\kappa(x)) \leq \epsilon \ .$$

Da wir $\epsilon > 0$ beliebig wählen können, gilt $f \in R[a,b]$. Weiterhin folgt aus

$$\int_a^c \psi(x)dx + \int_c^b \mu(x)dx = \int_a^b \delta(x)dx \ ,$$

daß

$$\int_a^c f_{[|a,c]}(x)dx + \int_c^b f_{|[c,b]}(x)dx = \int_a^b f(x)dx \ .$$

Sei $f \in R[a,b]$, $m := \inf f$, $M := \sup f$ und $\Phi : [m,M] \to \mathbb{R}$ stetig.

Die Funktion Φ ist gleichmäßig stetig (Lemma 4.59). Sei $\epsilon > 0$ gegeben. Wir wählen $\epsilon > \delta > 0$ derart, daß aus $x,y \in [m,M]$ und $|y-x| < \delta$ folgt: $|\Phi(x)-\Phi(y)| < \epsilon$. Wir wählen weiter eine $\delta\epsilon$-Zange (ϕ,ψ) von f. Sei $(x_i)_{i=0}^n$ eine gemeinsame zulässige Zerlegung von $[a,b]$ für ϕ und ψ.

Dann definieren wir $\tilde{\phi},\tilde{\psi}$ so daß

$$\tilde{\phi}(x) := \inf_{z\in[x_{i-1},x_i)} f(z) \ , \quad \tilde{\psi}(x) := \sup_{z\in[x_{i-1},x_i)} f(z)$$

für $x \in [x_{i-1},x_i)$ gilt. Es folgt

$$\phi \leq \tilde{\phi} \leq f \leq \tilde{\psi} \leq \psi \ .$$

und

$$\int_a^b (\tilde{\psi}(x)-\tilde{\phi}(x)) < \delta\epsilon \ .$$

Folglich ist auch $(\tilde{\phi},\tilde{\psi})$ eine $\delta\epsilon$-Zange von f. Wir zerlegen die Indexmenge $\{1,\ldots,n\} = A \cup B$ derart, daß $i \in A$ genau dann, wenn $\tilde{\psi}(x_{i-1}) - \tilde{\phi}(x_{i-1}) < \delta$.

Wir definieren weiter eine Zange (κ, δ) von $\Phi \circ f$ durch

$$\kappa(x) := \inf_{z \in [x_{i-1}, x_i)} \Phi(f(z)) \ , \quad \delta(x) := \sup_{z \in [x_{i-1}, x_i)} \Phi(f(z))$$

für $x \in [x_{i-1}, x_i)$ gilt. Für $i \in A$ und $x, y \in [x_{i-1}, x_i)$ gilt $|f(x) - f(y)| \leq \tilde{\psi}(x_{i-1}) - \tilde{\phi}(x_{i-1}) < \delta$ und somit $|\Phi(f(x)) - \Phi(f(y))| < \epsilon$. Wir schließen, daß dann $\delta(x_{i-1}) - \kappa(x_{i-1}) \leq \epsilon$ gilt. Weiterhin gilt

$$\delta\epsilon > \sum_{i \in B} (\tilde{\psi}(x_{i-1}) - \tilde{\phi}(x_{i-1}))(x_i - x_{i-1}) \geq \delta \sum_{i \in B} (x_i - x_{i-1})$$

woraus $\sum_{i \in B}(x_i - x_{i-1}) < \epsilon$ folgt. Hieraus und der Beschränktheit von $|\delta|, |\kappa|$ durch $C := \sup_{[m,M]} |\Phi|$ schließen wir

$$\sum_{i \in B} (\delta(x_{i-1}) - \kappa(x_{i-1}))(x_i - x_{i-1}) \leq 2C \ .$$

Weiter gilt

$$\sum_{i \in A} (\delta(x_{i-1}) - \kappa(x_{i-1}))(x_i - x_{i-1}) \leq \epsilon \sum_{i \in A} (x_i - x_{i-1}) \leq \epsilon(b - a) \ .$$

Insgesamt gilt

$$\int_a^b (\delta(x) - \kappa(x)) dx \leq \epsilon(2C + (b - a)) \ .$$

Wir haben damit eine $\epsilon(2C + (b - a))$-Zange (κ, δ) von $\Phi \circ f$ gefunden. Nun war $\epsilon > 0$ beliebig. Daraus schließen wir, daß $\Phi \circ f \in R[a, b]$.

Wir können nun $\Phi(x) = |x|$ betrachten und schließen, daß mit $f \in R[a, b]$ auch $|f| \in R[a, b]$ ist. Desweiten gilt $\pm f \leq |f|$ und somit

$$\pm \int_a^b f(x) dx \leq \int_a^b |f(x)| dx \ .$$

Daraus folgt die Behauptung (6). ■

Stetige Funktionen sind riemannintegrierbar.

Satz 6.8 *Es gilt* $C([a, b]) \subset R[a, b]$.

Beweis: Wir zeigen zuerst, daß $\mathtt{id}_{[a,b]} \in R[a, b]$ ist. Dann schreiben wir $f = f \circ \mathtt{id}_{[a,b]}$ und wenden Theorem 6.7 an. Sei $\epsilon > 0$ gegeben. Dann wählen wir eine Zerlegung $(x_i)_{i=0}^n$ von $[a, b]$ derart, daß $x_i - x_{i-1} < \frac{\epsilon}{b-a}$ für alle i gilt. Wir setzen weiter

$$\phi(x) := x_{i-1} \ , \psi(x) := x_i \ , x \in [x_{i-1}, x_i) \ .$$

Dann ist (ϕ, ψ) eine Zange von $\mathtt{id}_{[a,b]}$. Weiter gilt $|\psi(x) - \phi(x)| < \frac{\epsilon}{b-a}$ und deshalb

$$\int_a^b (\psi(x) - \phi(x))dx \leq \epsilon \ .$$

Also ist (ϕ, ψ) eine ϵ-Zange von $\mathtt{id}_{[a,b]}$. Da $\epsilon > 0$ beliebig klein gewählt werden kann, gilt $\mathtt{id}_{[a,b]} \in R[a,b]$.

Definition 6.9 *Eine Funktion $f : [a,b] \to \mathbb{R}$ ist stückweise stetig, wenn es eine Zerlegung $[a,b] = \bigsqcup_{i=1}^n I_i$ in endlich viele echte Intervalle gibt, so daß $f_{|I_i}$ für alle $i = 1, \ldots, n$ stetig ist.*

Korollar 6.10 *Eine stückweise stetige Funktion ist riemannintegrierbar.*

Beweis: Das folgt aus 6.7, (4) und 6.8. ■

Das Integral ist eine lineare Abbildung $\int_a^b \ldots dx : C([a,b]) \to \mathbb{R}$. Wir versehen $C([a,b])$ mit der Metrik

$$d(f,g) := \|f - g\| := \sup_{[a,b]} |f - g| \ .$$

Lemma 6.11 *Das Integral ist eine stetige lineare Abbildung.*

Beweis: Wir verwenden die ϵ-δ-Definition mit $\delta = \frac{\epsilon}{b-a}$. Ist $d(f,g) < \delta$, dann gilt

$$\begin{aligned} |\int_a^b f(x)dx - \int_a^b g(x)dx| &= |\int_a^b (f(x) - g(x))dx| \\ &\leq \int_a^b |f(x) - g(x)|dx \\ &\leq \sup_{[a,b]} |f - g|(b-a) \\ &< \epsilon \ . \end{aligned}$$

■

Nicht jede Funktion $f : [a,b] \to \mathbb{R}$ ist riemannintegrierbar. Sei etwa $f(x) := 0$ für $x \in [a,b] \cap \mathbb{Q}$ und $f(x) := 1$ für $x \in [a,b] \setminus \mathbb{Q}$. Dann gilt für jede Zange (ϕ, ψ), daß $\phi \leq 0$ und $\psi \geq 1$ ist. Daraus folgt

$$\int_a^b (\psi(x) - \phi(x))dx \geq b - a \ .$$

Insbesondere gibt es keine ϵ-Zange für $\epsilon < b - a$.

6.2 Der Hauptsatz

Wir haben gesehen, daß stetige Funktionen riemannintegrierbar sind. Es stellt sich nun die Frage, wie man Integrale berechnen kann. Das wesentliche Hilfsmittel ist der Hauptsatz der Integralrechnung.

Das Integral der konstanten Funktion mit dem Wert λ kann man unmittelbar aus der Definition berechenen. Diese Funktion ist nämlich einfach. Es gilt $\int_a^b \lambda dx = \lambda[a,b]$.

Satz 6.12 *Sei $f : [a,b] \to \mathbb{R}$ stetig. Wir bilden die Funktion $[a,b] \ni x \mapsto F(x) := \int_a^x f(x)dx$. Dann ist F stetig, auf (a,b) differenzierbar und es gilt*

$$F'(x) = f(x) \ , \quad F(a) = 0 \ .$$

Beweis: Wir betrachten $z \in [a,b]$. Für $y > z$ (der Fall $z < y$ geht analog) gilt

$$F(y)-F(z) = \int_a^y f(x)dx - \int_a^z f(x)dx = \int_a^z f(x)dx + \int_z^y f(x)dx - \int_a^z f(x)dx = \int_z^y f(x)dx \ .$$

Sei $M := \sup_{[a,b]} |f|$. Wir schließen, daß $|F(y) - F(z)| < M|y-z|$ gilt. Daraus folgt die Stetigkeit von F.

Wir berachten nun $z \in (a,b)$. Wir schreiben

$$\begin{aligned} F(y) = F(z) + \int_z^y f(x)dx &= F(z) + \int_z^y (f(z) + (f(x) - f(y))dx \\ &= F(z) + f(z)(y-z) + \frac{\int_z^y (f(x) - f(z))dx}{y-z}(y-z)dx \ . \end{aligned}$$

Es gilt

$$\left|\frac{\int_z^y (f(x) - f(z))dx}{y-z}\right| \le \sup_{x \in [z,y]} |f(x) - f(z)| \overset{y \to z}{\to} 0$$

da f in z stetig ist. Das zeigt, daß F' im Punkt z existiert und $F'(z) = f(z)$ gilt.

Die Aussage $F(a) = 0$ ist klar. ■

Sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion.

Definition 6.13 *Eine auf $[a,b] \subset \mathbb{R}$ stetige und auf (a,b) differenzierbare Funktion $F : [a,b] \to \mathbb{R}$ heißt Stammfunktion von f, wenn auf (a,b) gilt: $F' = f$.*

Lemma 6.14 *Sind $F_0, F_1 : [a,b] \to \mathbb{R}$ Stammfunktionen von f, dann ist $F_0 - F_1$ konstant.*

Beweis: Es gilt $F_0' - F_1' = 0$. ■

Korollar 6.15 *Ist F eine Stammfunktion von f, dann gilt*

$$\int_a^b f(x)dx = F(b) - F(a) \ .$$

Beweis: Die Funktion $F(x) := \int_a^x f(x)dx$ ist eine Stammfunktion von f. Damit gilt $\int_a^b f(x)dx = F(b) - F(a)$ für diese Stammfunktion. Ist F_1 eine weitere Stammfunktion, dann gilt $F_1 = F + C$ für eine Konstante $C \in \mathbb{R}$. Damit ist

$$F_1(b) - F_1(a) = F(b) - F(a) = \int_a^b f(x)dx \ .$$

■

1. $\frac{1}{n+1}x^{n+1}$ ist eine Stammfunktion von x^n. Folglich gilt
$$\int_a^b x^n = \frac{b^{n+1} - a^{n+1}}{n+1} \ .$$

2. $\frac{\mathsf{e}^{cx}}{c}$ ist eine Stammfunktion von e^{cx}. Folglich gilt
$$\int_a^b \mathsf{e}^{cx}dx = \frac{\mathsf{e}^{bc} - \mathsf{e}^{ac}}{c} \ .$$

3. $\ln(x)$ ist eine Stammfunktion von $\frac{1}{x}$ auf $(0, \infty)$. Damit gilt etwa für $0 < a < b$, daß
$$\int_a^b \frac{1}{x}dx = \ln(b) - \ln(a) = \ln(\frac{b}{a}) \ .$$

4. $\arcsin(x)$ ist eine Stammfunktion von $\frac{1}{\sqrt{1-x^2}}$ auf $(-1, 1)$. Folglich gilt für $-1 < a \leq b < 1$
$$\int_a^b \frac{1}{\sqrt{1-x^2}}dx = \arcsin(b) - \arcsin(a) \ .$$
Insbesondere gilt
$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}dx := \lim_{a \to -1} \lim_{b \to 1} \int_a^b \frac{1}{\sqrt{1-x^2}}dx = \arcsin(1) - \arcsin(-1) = \pi \ .$$

5. $\arctan(x)$ ist eine Stammfunktion von $\frac{1}{1+x^2}$. Folglich gilt
$$\int_a^b \frac{1}{1+x^2}dx = \arctan(b) - \arctan(a) \ .$$
Es gilt
$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}dx := \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b \frac{1}{1+x^2}dx = \pi \ .$$

6.3 Uneigentliche Integrale

Sei $f : [a,b) \to \mathbb{R}$ ($b = \infty$ is zugelassen) eine Funktion derart, daß $f_{|[a,c]} \in R[a,c]$ für alle $a < c < b$ gilt. Dann können wir die Funktion $(a,b) \ni c \mapsto \int_a^c f(x)dx \in \mathbb{R}$ betrachten. Der Punkt b ist ein Häufungspunkt des Definitionsbereiches $(a,b) \subset \bar{\mathbb{R}}$ dieser Funktion.

Definition 6.16 *Wenn* $\lim_{c\to b} \int_a^x f(x)dx$ *existiert, dann heißt dieser Wert*

$$\int_a^b f(x)dx = \lim_{c\to b} \int_a^x f(x)dx$$

uneigentliche Integral von f*. Analog definiert man das uneigentliche Integral einer Funktion* $(a,b] \to \infty$*. Das uneigentliche Integral einer Funktion* $f : (a,b) \to \mathbb{R}$ *ist definiert als Summe uneigenlicher Integrale (falls diese existieren)*

$$\int_a^b f(x)dx := \int_a^c f(x)dx + \int_c^b f(x)dx$$

für ein $c \in (a,b)$*.*

Man sieht leicht ein, daß diese Definition nicht von der Wahl von c abhängt.

1. $$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}dx = \pi$$

2. $$\int_{-\infty}^{\infty} \frac{1}{1+x^2}dx = \pi$$

3. $$\int_0^{\infty} \mathsf{e}^{-x} = 1 \ .$$

   In der Tat ist $\int_0^c \mathsf{e}^{-x}dx = 1 - \mathsf{e}^{-c}$.

4. $$\int_0^{\infty} x\mathsf{e}^{-x^2} = \frac{1}{2} \ .$$

   In der Tat ist $-\frac{\mathsf{e}^{-x^2}}{2}$ eine Stammfunktion von $x\mathsf{e}^{-x^2}$.

5. Es gilt für $s < -1$

   $$\int_1^{\infty} x^s = -\frac{1}{s+1} \ .$$

   In der Tat ist $\frac{x^{s+1}}{s+1}$ eine Stammfunktion von x^s. Ähnlich sieht man für $s > -1$, daß

   $$\int_0^1 x^s dx = \frac{1}{s+1}$$

   gilt.

6.4 Partielle Integration, Substitiution

Satz 6.17 (partielle Integration) *Sei $U \subseteq \mathbb{R}$ offen, $[a,b] \subset U$ und $F, G : U \to \mathbb{R}$ stetig differenzierbar mit $f := F'$, $g := G'$. Dann gilt*

$$\int_a^b f(x)G(x)dx = FG|_a^b - \int_a^b F(x)g(x)dx \ ,$$

wobei $FG|_a^b := F(b)G(b) - F(a)G(a)$ ist.

Beweis: Wir wenden den Hauptsatz der Integralrechnung auf die Funktion FG an. Es gilt

$$(FG)' = F'G + FG' = fG + Fg \ .$$

■

1. Wir wollen $\int_0^\pi x\sin(x)dx$ berechnen. Wir setzen $G(x) := x$ und $F(x) = -\cos(x)$. Dann gilt
$$\int_0^\pi x\sin(x) = (-x\cos(x))|_0^\pi - \int_0^\pi (-\cos(x))dx \ .$$
Nun gilt $(-x\cos(x))|_0^\pi = \pi$. Ferner ist $\int_0^\pi \cos(x)dx = \sin(\pi) - \sin(0) = 0$. Wir schließen, daß
$$\int_0^\pi x\sin(x)dx = \pi$$
gilt.

2. Wir wollen $\int_0^\pi x^2\cos(x)dx$ berechnen. Wir setzen $G(x) := x^2$ und $F(x) = \sin(x)(x)$. Dann ist
$$\int_0^\pi x^2\cos(x)dx = (x^2\sin(x))|_0^\pi - \int_0^\pi 2x\sin(x)dx = -2\pi \ .$$

3. Wir wollen etwa $\int_0^a \ln(1+x)dx$ berechnen. Wir setzen $F(x) := x$ und $G(x) := \ln(1+x)$. Dann gilt
$$\int_0^a \ln(1+x)dx = (x\ln(1+x))|_0^a - \int_0^a \frac{x}{1+x}dx \ .$$
Nun ist $\frac{x}{1+x} = 1 - \frac{1}{1+x}$. Folglich gilt
$$\begin{aligned}\int_0^a \ln(1+x)dx &= a\ln(1+a) - \int_0^a (1 - \frac{1}{1+x})dx \\ &= a\ln(1+a) - a + \ln(1+a) \\ &= (1+a)\ln(1+a) - a\end{aligned}$$

Um die folgende Substitutionsregel besser formulieren zu können, setzen wir für $b \le a$

$$\int_a^b f(x)dx := -\int_b^a f(x)dx \ .$$

Sei $\phi : [a,b] \to (u,v)$ eine stetige und auf (a,b) stetig differenzierbare Bijektion.

Lemma 6.18 (Substitutionsregel) *Für eine stetige Funktion $f : (u,v) \to \mathbb{R}$ gilt*

$$\int_a^b f(\phi(x))\phi'(x)dx = \int_{\phi(a)}^{\phi(b)} f(y)dy \ .$$

Beweis: Sei F eine Stammfunktion von f. Wir betrachten die Komposition $F \circ \phi : [a,b] \to \mathbb{R}$. Auf (a,b) gilt

$$(F \circ \phi)'(x) = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x) \ .$$

Nach dem Hauptsatz der Integralrechnung gilt

$$\int_{\phi(a)}^{\phi(b)} f(x)dx = F(\phi(b)) - F(\phi(b)) = \int_a^b f(\phi(x))\phi'(x)dx \ .$$

■

1. Wir wollen $\int_0^1 \sqrt{1-x^2}dx$ berechnen. Wir setzen $\phi = \sin(u)$. Dann gilt $\phi'(u) = \cos(u)$ und

$$\int_0^1 \sqrt{1-x^2}dx = \int_0^{\frac{\pi}{2}} \sqrt{1-\sin(u)^2}\cos(u)du = \int_0^{\frac{\pi}{2}} \cos(u)^2 dx \ .$$

   Wegen $\cos(u) = \sin(\frac{\pi}{2} - u)$ sehen wir, daß

$$\int_0^{\frac{\pi}{2}} \cos(u)^2 dx = \int_0^{\frac{\pi}{2}} \sin((\frac{\pi}{2} - u))^2 dx = \int_0^{\frac{\pi}{2}} \sin(u)^2 dx$$

   gilt. Aus $1 = \cos(u)^2 + \sin(u)^2$ folgt nun $\int_0^{\frac{\pi}{2}} \cos(u)^2 dx = \frac{\pi}{4}$. Wir sehen, daß

$$\int_0^1 \sqrt{1-x^2}dx = \frac{\pi}{4} \ .$$

   In der Tat haben wir den Inhalt eines Viertelkreises ausgerechnet.

2. Wir wollen etwa

$$\int_0^\pi \sin(x)\mathsf{e}^{\cos(x)}dx$$

   berechnen. Wir setzen $\phi(x) := \cos(x)$ und beobachten, daß $-\sin(x) = \phi'(x)$ ist. Wir setzen daher $f(y) := -\mathsf{e}^y$ und er halten

$$\int_0^\pi \sin(x)\mathsf{e}^{\cos(x)}dx = \int_1^{-1} (-\mathsf{e}^y)dy = \mathsf{e} - \mathsf{e}^{-1} \ .$$

3. Wir wollen etwa
$$\int_0^\infty \frac{1}{\sqrt{1+x^2}^3} dx$$
berechnen. Wir substituieren $x := \phi(t) := \sinh(t)$. Dann ist $\phi'(t) = \cosh(t)$. Es gilt
$$\int_0^\infty \frac{1}{\sqrt{1+x^2}^3} dx = \int_0^\infty \frac{\cosh(t)}{\sqrt{1+\sinh(t)^2}^3} dx = \int_0^\infty \frac{\cosh(t)}{\sqrt{\cosh(t)^2}^3} dx = \int_0^\infty \frac{1}{\cosh(t)^2} dx \; .$$
Mit $\tanh(x) := \frac{\sinh(t)}{\cosh(t)}$ gilt
$$\tanh'(x) = \frac{\cosh(t)^2 - \sinh(t)^2}{\cosh(t)^2} = \frac{1}{\cosh(t)^2} \; .$$
Wir schließen weiter, daß
$$\int_0^\infty \frac{1}{\sqrt{1+x^2}^3} dx = \tanh(x)|_0^\infty = 1 \; .$$

6.5 Partialbruchzerlegung

Mit Hilfe der Partialbruchzerlegung kann man rationale Funktionen integrieren. Da Polynome über $\mathbb{C}$ in Linearfaktoren zerfallen, ist es günstig, komplexwertige Funktionen zu betrachten. Eine Funktion $f : U \to \mathbb{C}$ zerfällt in den Real- und Imaginärteil $\mathtt{Re}(f), \mathtt{Im}(f) : U \to \mathbb{R}$, $f = \mathtt{Re}(f) + i\mathtt{Im}(f)$. Sie ist genau dann differenzierbar, wenn $\mathtt{Re}(f)$ und $\mathtt{Im}(f)$ differenzierbar sind, und es gilt $f' = \mathtt{Re}(f)' + i\mathtt{Im}(f)'$.

Wir betrachten eine rationale Funktion
$$f(x) := \frac{p(x)}{q(x)}$$
mit teilerfremden Polynomen $p, q \in \mathbb{C}[x]$. Wenn $\deg(q) \le \deg(p)$ ist, dann schreiben wir
$$p(x) = \phi(x)q(x) + r(x)$$
für $r, \phi \in \mathbb{C}[x]$ mit $\deg(r) < \deg(q)$ und
$$f(x) = \phi(x) + \frac{r(x)}{q(x)} \; .$$
Eine Stammfunktion von ϕ is leicht zu finden, so daß wir von jetzt an annehmen, daß $\deg(p) < \deg(q)$ gilt.

Sei $\lambda \in \mathbb{C}$ eine Nullstelle der Multiplizität k von q. Wir behaupten, daß man
$$f(x) = \frac{A}{(x-\lambda)^k} + \frac{r(x)}{\tilde{q}(x)}$$

mit $\tilde{q}(x)(x-\lambda) = q(x)$ und $r \in \mathbb{C}[x]$ mit $\deg(r) < \deg(\tilde{q})$ schreiben kann. In der Tat ist mit $q_1(x) = \frac{q(x)}{(x-\lambda)^k}$

$$\frac{p(x)}{q(x)} - \frac{A}{(x-\lambda)^k} = \frac{p(x) - Aq_1(x)}{q(x)} \ .$$

Wir setzen $A := \frac{p(\lambda)}{q_1(\lambda)}$. Dann ist $\frac{p(x)-Aq_1(x)}{q(x)} = \frac{r(x)}{\tilde{q}(x)}$ für ein geeignetes $r \in \mathbb{C}[x]$ mit $\deg(r) < \deg(\tilde{q})$. Wir sehen induktiv, daß $f(x)$ in eine endliche Summe von Termen der Form

$$\frac{A}{(x-\lambda)^k}$$

entwickelt werden kann. Ist $k \geq 2$, dann ist

$$-\frac{A}{(k-1)(x-\lambda)^{k-1}}$$

eine Stammfunktion.

Ist $k = 1$ und $\lambda \in \mathbb{R}$, dann ist

$$A\ln(|x-\lambda|)$$

eine Stammfunktion.

Es bleibt der Fall $k = 1$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Wir schreiben

$$\frac{A}{x-\lambda} = \frac{Ax - A\bar{\lambda}}{x^2 - 2\mathtt{Re}(\lambda)x + |\lambda|^2} = \frac{A(x-\mathtt{Re}(\lambda)) - A(\bar{\lambda} - \mathtt{Re}(\lambda))}{(x-\mathtt{Re}(\lambda))^2 + |\lambda|^2 - \mathtt{Re}(\lambda)^2} \ .$$

Beachte, daß $|\lambda|^2 - \mathtt{Re}(\lambda)^2 > 0$ gilt.

Eine Stammfunktion ist

$$\frac{A}{2}\ln((x-\mathtt{Re}(\lambda))^2 + |\lambda|^2 - \mathtt{Re}(\lambda)^2) - A\frac{\bar{\lambda} - \mathtt{Re}(\lambda)}{\sqrt{|\lambda|^2 - \mathtt{Re}(\lambda)^2}}\arctan(\frac{x - \mathtt{Re}(\lambda)}{\sqrt{|\lambda|^2 - \mathtt{Re}(\lambda)^2}})$$

Hier ist eine Anwendung. Wir wollen

$$\int_0^\infty \frac{1}{1+2x^2+x^4}dx$$

berechnen. Die Nullstellen von $x^4 + 2x^2 + 1$ sind $\pm i$, jeweils mit doppelter Multiplizität:

$$(x+i)^2(x-i)^2 = (x^2+1)^2 = x^4 + 2x + 1 \ .$$

Wir machen den Ansatz

$$\frac{1}{1+2x^2+x^4} = \frac{A}{(x-i)^2} + \frac{Bx+C}{x^2+1} + \frac{D}{(x+i)^2} \ .$$

Nach Hauptnennerbilden bekommen wir

$$\begin{aligned}
A(x+i)^2 + (Bx+C)(x^2+1) + D(x-i)^2 &= 0 \\
Bx^3 + (A+C+D)x^2 + (2iA + B - 2iD)x + (-A+C-D) &= 1 \\
B &= 0 \\
A &= D \\
2A + C &= 0 \\
-2A + C &= 1 \\
C &= \frac{1}{2} \\
A &= \frac{-1}{4} = D
\end{aligned}$$

Als gilt

$$\frac{1}{1+2x^2+x^4} = \frac{-1}{4(x-i)^2} + \frac{-1}{4(x+i)^2} + \frac{1}{2(x^2+1)}$$

und damit ist

$$\frac{1}{4(x-i)} + \frac{1}{4(x+i)} + \frac{1}{2}\arctan(x) = \frac{x}{2(x^2+1)} + \frac{1}{2}\arctan(x)$$

eine Stammfunktion. Probe:

$$\begin{aligned}
\frac{1}{2(x^2+1)} - \frac{x^2}{(x^2+1)^2} + \frac{1}{2(x^2+1)} &= \frac{2x^2+2-2x^2}{2(x^2+1)^2} \\
&= \frac{1}{(x^2+1)^2} \\
&= \frac{1}{1+2x^2+x^4} \ .
\end{aligned}$$

Wir schließen

$$\int_0^\infty \frac{1}{1+2x^2+x^4} dx = \frac{\pi}{4} \ .$$

6.6 Bogenlänge

Sei $I := [0,1]$. Da I nicht offen ist, benutzen wir für I folgende Erweiterung des Begriffs einer differenzierbaren Funktion.

Definition 6.19 *Eine Funktion $f : I \to \mathbb{R}$ ist differenzierbar, wenn sie Einschränkung einer auf einer offenen Umgebung von I definierten differenzierbaren Funktion ist.*

Wir wollen im folgenden vektorwertige Funktionen betrachten. Eine Funktion $f : I \to \mathbb{R}^n$ kann als ein n-Tupel von Funktionen $(f_1, \ldots, f_n)$ notiert werden werden.

Definition 6.20 *Wir betrachten f als differenzierbar wenn die f_i für $i = 1, \ldots, n$ differenzierbar sind und setzen $f'(t) = (f_1'(t), \ldots, f_n'(t)) \in \mathbb{R}^n$.*

Der Begriff des Weges ist ein topologisches Konzept.

Definition 6.21 *Ein **parametrisierter Weg** in $\mathbb{R}^n$ ist eine stetige Abbildung $\gamma : [0,1] \to \mathbb{R}^n$. Mit $P\mathbb{R}^n$ bezeichnen wir den Menge der Wege in $\mathbb{R}^n$.*

Auf der Menge der Wege $P\mathbb{R}^n$ führen wir Äquivalzenzrelation $\sim$ ein, unter welcher zwei Wege als Äquivalent betrachtet werden, wenn sie durch Umparametrisierung auseinander hervorgehen.

Definition 6.22 *Es gilt $\gamma_0 \sim \gamma_1$ genau dann, wenn einen Homöomorphismus $\phi : I \to I$ gibt mit*

$$\gamma_0 = \gamma_1 \circ \phi \ .$$

Wir überlassen es als Übungsaufgabe, nachzuweisen, daß dies einen Äquivalenzrelation ist. Mit

$$\bar{P}\mathbb{R}^n := P\mathbb{R}^n / \sim$$

bezeichnen wir die Menge der unparametrisierten Wege. Im folgenden wollen wir den Begriff der Länge eines Weges untersuchen. Das topologische Konzept ist dafür zu allgemein. Der Einfachheit halber schränken wir uns hier auf differenzierbare Wege ein.

Definition 6.23 *Ein **differenzierbarer parametrisierter Weg** in $\mathbb{R}^n$ ist eine stetig differenzierbare Abbildung $\gamma : [0,1] \to \mathbb{R}^n$. Mit $P_{C^1}\mathbb{R}^n$ bezeichnen wir den Menge der differenzierbaren Wege in $\mathbb{R}^n$.*

Entsprechend definieren wir die Relation der differenzierbaren Umparametrisierung.

Definition 6.24 *Es gilt $\gamma_0 \sim_{C^1} \gamma_1$ genau dann, wenn eine stetig differenzierbare Abbildung $\phi : I \to I$ gibt mit $\phi' > 0$ und*

$$\gamma_0 = \gamma_1 \circ \phi \ .$$

Mit

$$\bar{P}_{C^1}\mathbb{R}^n := P_{C^1}\mathbb{R}^n / \sim_{C^1}$$

bezeichnen wir die Menge der unparametrisierten differenzierbaren Wege.

Sei $\gamma \in P_{C^1}\mathbb{R}^n$. Dann ist $\gamma' : I \to \in \mathbb{R}^n$ stetig. Folglich ist auch $I \ni t \mapsto \|\gamma'(t)\| \in \mathbb{R}$ stetig.

Definition 6.25 *Wir definieren die **Länge** von γ durch*

$$L(\gamma) := \int_0^1 \|\gamma'\| dt \ .$$

Die wesentliche Aussage ist, daß diese Definition nicht von der Parametrisierung abhängt.

Satz 6.26 *Gilt für $\gamma_0, \gamma_1 \in P_{C^1}\mathbb{R}^n$ die Relation $\gamma_0 \sim_{C^1} \gamma_1$, dann ist $L(\gamma_0) = L(\gamma_1)$.*

Beweis: Beweis: Sei $\gamma_0 = \gamma_1 \circ \phi$ für eine differenzierbare Abbildung $\phi : I \to I$ mit $\phi' > 0$. Dann gilt $\gamma_0'(t) = \gamma_1'(\phi(t))\phi'(t)$ und

$$\|\gamma_0'(t)\| = \|\gamma_1'(\phi(t))\|\phi'(t) \ .$$

Die Substitutionsregel für das Integral zeigt

$$\int_0^1 \|\gamma_0'(t)\|dt = \int_0^1 \|\gamma_1'((\phi(t))\|\phi(t)dt = \int_0^1 \|\gamma_1'(t)\|dt \ .$$

■

1. Sei γ eine Gerade zwischen den Punkten $p, q \in \mathbb{R}^n$, also $\gamma(t) = p + t(q - p)$. Dann gilt $\gamma'(t) = (q - p)$ und $\|\gamma')t\| = \|q - p\|$, also

   $$L(\gamma) = \|q - p\| \ .$$

   Die Länge dieser Kurve ist genau der Abstand zwischen den beiden Punkten.

2. Wir parametrisieren den Vierteleinheitskreis durch

   $$\gamma(t) = (t, \sqrt{1 - t^2}) \ .$$

   Dann gilt

   $$\gamma'(t) = (1, \frac{-t}{\sqrt{1 - t^2}})$$

   und

   $$\|\gamma'(t)\| = \sqrt{1 + \frac{t^2}{1 - t^2}} = \frac{1}{\sqrt{1 - t^2}} \ .$$

   Folglich gilt

   $$L(\gamma) = \int_0^1 \frac{1}{\sqrt{1 - t^2}}dt = \arcsin(t)|_0^1 = \frac{\pi}{2} \ .$$

   Wir haben damit verifiziert, daß der Kreisumfang durch 2π gegeben wird.

3. Der Weg $\gamma(t) = (t, t^2)$ beschreibt ein Stück der Standardparabel. Es gilt $\gamma'(t) =$

$(1,2t)$ und $\|\gamma'(t)\| = \sqrt{1+4t^2}$. Folglich

$$\begin{aligned}
L(\gamma) &= \int_0^1 \sqrt{1+4t^2}dt \\
&= \frac{1}{2}\int_0^1 \sqrt{1+(2t)^2}d(2t) \\
&= \frac{1}{2}\int_0^2 \sqrt{1+u^2}du \\
&= \frac{1}{2}\int_0^{\operatorname{arcsinh}(2)} \sqrt{1+\sinh^2(t)}d(\sinh(t)) \\
&= \frac{1}{2}\int_0^{\operatorname{arcsinh}(2)} \sqrt{1+\sinh^2(t)}\cosh(t)dt \\
&= \frac{1}{2}\int_0^{\operatorname{arcsinh}(2)} \cosh(t)^2 dt \\
&= \frac{1}{8}\int_0^{\operatorname{arcsinh}(2)} (e^{2t}+e^{-2t}+2)dt \\
&= (\frac{1}{8}\sinh(2t)+\frac{1}{4}t)|_0^{\operatorname{arcsinh}(2)} \\
&= (\frac{1}{4}\sinh(t)\cosh(t)+\frac{1}{4}t)|_0^{\operatorname{arcsinh}(2)} \\
&= (\frac{1}{4}\sinh(t)\sqrt{1+\sinh(t)^2}+\frac{1}{4}t)|_0^{\operatorname{arcsinh}(2)} \\
&= \frac{1}{2}\sqrt{5}+\frac{1}{4}\operatorname{arcsinh}(2)
\end{aligned}$$

7 Funktionalanalytische Aspekte

7.1 Funktionenräume, Vervollständigung

Seine X, Y topologische Räume. Mit $C(X, Y)$ bezeichnen wir die Menge der stetigen Abbildungen $X \to Y$. Wenn $Y = \mathbb{R}$ ist, dann schreiben wir einfach $C(X) := C(X, \mathbb{R})$. Wir werden diese Notation auch für komplexwertige Funktionen verwenden.

1. Wenn X kompakt ist, dann ist jede stetige Funktion auf X beschränkt. Für $f \in C(X)$ ist die Norm $\|f\|_\infty := \sup_X |f|$ wohldefiniert und bestimmt eine Metrik. Der metrische Raum$(C(X), \|\dots\|_\infty)$ ist vollständig (Lemma 4.68).

2. Wir betrachten jetzt $[a, b] \subset \mathbb{R}$ und die Norm

$$\|f\|_1 := \int_a^b |f(x)|dx$$

auf $C([a, b])$. Wir weisen dazu die Normeigenschaften nach.

(a) Es gilt immer $\|f\|_1 \geq 0$. Wenn $\|f\|_1 = 0$ ist, dann ist $f = 0$. In der Tat, wäre $f \neq 0$, dann gibt es $z \in (a,b)$ mit $f(z) \neq 0$. Dann finden wir $\delta, \epsilon > 0$ derart, daß $[z-\epsilon, z+\epsilon] \subset [a,b]$ und auf diesem Intervall $|f(x)| > \delta$ gilt. Damit ist aber

$$\int_a^b |f(x)dx \geq \int_{z-\epsilon}^{z+\epsilon} \delta dx = 2\epsilon\delta > 0$$

(b) Für $\lambda \in \mathbb{R}$ gilt $\|\lambda f\|_1 = |\lambda| \|f\|_1$.

(c) Schließlich gilt

$$\begin{aligned} \|f+g\|_1 &= \int_a^b |f(x)+g(x)|dx \\ &\leq \int_a^b (|f(x)|+|g(x)|)dx \\ &\leq \|f\|_1 + \|g\|_1 \ . \end{aligned}$$

Der metrische Raum $(C([a,b]), \| \dots \|_1)$ ist nicht vollständig. Als Beispiel betrachten wir $[a,b] = [0,2]$ und die Folge

$$f_n(x) := \begin{cases} x^n & x \in [0,1] \\ 1 & x \in (1,2] \end{cases}$$

Dann gilt

$$\|f_n - f_m\|_1 = \int_0^1 |x^n - x^n|dx \leq \int_0^1 (x^n + x^m)dx = \frac{1}{n+1} + \frac{1}{m+1} \ .$$

Folglich ist (f_n) eine Cauchyfolge. Wir nehmen nun an, daß $f_n \to g$ für $g \in C([0,2])$ gilt. Dann ist

$$\int_0^2 |f_n(x) - g(x)|dx \geq \int_1^2 |1 - g(x)|dx \ .$$

Da die linke Seite für $n \to \infty$ gegen 0 konvergiert, gilt $\int_1^2 |1-g(x)|dx = 0$. Daraus folgt $g(x) = 1$ für alle $x \in [1,2]$ folgt. Möge nun für ein $z \in [0,1)$ gelten $g(z) \neq 0$. Dann ist gibt es für $0 < \delta < |g(z)|$ ein $1 - z > \epsilon > 0$ derart, daß $|g(x)| > \delta$ für $x \in [z, z+\epsilon]$ gilt. Damit ist

$$\int_0^2 |g(x) - f_n(x)|dx \geq \int_z^{z+\epsilon} |g(x) - f_n(x)|dx \geq \int_{z-\epsilon}^{z+\epsilon} (|g(x)| - |f_n(x)|)dx > \epsilon\delta - \frac{1}{n+1} \ .$$

Das ist aber ein Widerspruch zum Fakt, daß die linke Seite gegen Null strebt für $n \to \infty$. Folglich gilt $g(x) = 0$ für alle $x \in [0,1)$. Damit ist aber g nicht stetig. Es gilt

$$\|f\|_1 \leq (b-a)\|f\|_\infty \ .$$

Es gibt aber keine Abschätzung in die andere Richtung. Dazu betrachten wir etwa $[a,b] = [0,1]$ und $f_n := \sqrt{n}x^n$. Es gilt $\|f_n\| = \frac{\sqrt{n}}{1+n} \to 0$ für $n \to \infty$, aber $\|f_n\|_\infty = \sqrt{n} \to \infty$.

3. Auf $C([a,b])$ können wir auch die Norm

$$\|f\|_2 := \sqrt{\int_a^b |f(x)|^2 dx}$$

betrachten. Das dies eine Norm ist, sieht am am besten dadurch ein, daß man beobachtet, daß dahinter das Skalarprodukt $< g, f >:= \int_a^b f(x)g(x)dx$ steht. Der Raum $(C([a,b]), \|\dots\|_2)$ ist ebenfalls nicht vollständig. Wir können das gleiche Beispiel verwenden wie für $\|\|_1$.

Es gilt

$$\|f\|_2 \le \sqrt{b-a}\|f\|_\infty \ .$$

Die Cauchy-Schwartzsche Ungleichung zeigt

$$\|f\|_1 =< 1, |f| > | \le \|f\|_2\|1\|_2 = \sqrt{b-a}\|f\|_2 \ .$$

Wir haben also eine Kette stetiger Identitäten

$$(C([a,b]), \|\dots\|_\infty) \to (C([a,b]), \|\dots\|_2) \to (C([a,b]), \|\dots\|_1) \ .$$

Definition 7.1 *Ein normierter Vektorraum $(E, \|.\|)$ (über $\mathbb{R}$), welcher als metrischer Raum (mit $d(x,y) := \|x-y\|$) vollständig ist, nennet man einen reellen* ***Banachraum****.*

Für kompaktes X ist $(C(X), \|.\|_\infty)$ ein Banachraum. Die Räume $(C([a,b]), \|.\|_1)$ und $(C([a,b]), \|.\|_2)$ sind keine Banachräume.

Sei (X,d) ein metrischer Raum. Dieser muß nicht notwendig vollständig sein. Im folgenden wollen wir den Begriff der Vervollständigung einführen.

Definition 7.2 *Eine* ***Vervollständigung*** *von (X, d_X) ist ein Paar $((Y, d_Y), i)$ aus einem vollständigen metrischen Raum (Y, d) und einer isometrischen Abbildung $i : X \to Y$ derart, daß für jede lipschitzstetige [1] Abbildung $f : X \to Z$ in einen vollständigen metrischen Raum Z*

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "i"'] & & Z \\ & Y \arrow[ur, dotted, "g"'] & \end{tikzcd}$$

eine eindeutige stetige Faktorisierung g existiert.

Satz 7.3 *Jeder metrische Raum besitzt eine Vervollständigung $i : X \to Y$. Dabei ist $i(X) \subset Y$ dicht. Sind $((Y, d_Y), i)$ und $((Y', d_{Y'}), i')$ zwei Vervollständigungen, dann gibt es eine eindeutige isometrische Isomorphie $j : Y \to Y'$ so daß*

$$\begin{tikzcd} & X \arrow[dl, "i"] \arrow[dr, "i'"] & \\ Y \arrow[rr, "j"] & & Y' \end{tikzcd} \quad .$$

[1] also $d_Z(f(x_0), f(x_1)) \le C d_X(x_0, x_1)$ für alle $x_0, x_1 \in X$

Beweis: Wir zeigen zuerst die Eindeutigkeitsaussage. Wir wenden die universelle Eigenschaft von $((Y,d_Y),i)$ auf $i' : X \to Y'$ an und erhalten eine eindeutige Faktorsierung j. Analog erhalten wir $j' : Y' \to Y$ aus der universellen Eigenschaft von $((Y', d'_Y), i')$ angewendet auf die Abbildung $i : Y \to X$. Die Komposition $j' \circ j$ paßt in

$$\begin{tikzcd} X \arrow[rr, "i"] \arrow[dr, "i"'] & & Y \\ & Y \arrow[ur, dotted, "j' \circ j"'] & \end{tikzcd} .$$

Da diese Faktorisierung auch durch $\mathtt{id}_Y$ geleistet wird, muß wegen der Eindeutigkeit der Faktorsierung $j' \circ j = \mathtt{id}_Y$ sein. Analog zeigt man $j \circ j' = \mathtt{id}_{Y'}$. Wir müssen uns davon überzeugen, daß j eine Isometrie ist. Dazu benutzen wir die vorerst unbewiesene Tatsache, daß das Bild von X in Y dicht ist, also $\overline{i(X)} = Y$ gilt.

Seien zwei Punkte $y_0, y_1 \in Y$ gegeben. Dann finden wir Folgen (x_{0n}) und (x_{1n}) mit $i(x_{0n}) \to y_0$ und $i(x_{1n}) \to y_1$. Es gilt weiter wegen der Stetigkeit von j und der Abstandsfunktionen

$$\begin{aligned} d_{Y'}(j(y_0), j(y_1)) &= \lim_{n\to\infty} d_{Y'}(i(j(x_{0n})), i(j(x_{1n}))) \\ &= \lim_{n\to\infty} d_{Y'}(i'(x_{0n}), i'(x_{1n})) \\ &= \lim_{n\to\infty} d_X(x_{0n}, x_{1n}) \\ &= \lim_{n\to\infty} d_Y(i(x_{0n}), i(x_{1n})) \\ &= d_Y(y_0, y_1) \end{aligned}$$

Wir zeigen nun die Dichtheit von $i(X) \subseteq Y$. In der Tat ist $\overline{i(X)}$ als abgeschlossene Teilmenge eines metrischen Raumes auch vollständig. Die universelle Eigenschaft von $((Y, d_Y), i)$ gibt eine Faktorisierung

$$\begin{tikzcd} X \arrow[rr, "i"] \arrow[dr, "i"'] & & \overline{i(X)} \\ & Y \arrow[ur, dotted, "h"'] & \end{tikzcd} .$$

und die Komposition $Y \xrightarrow{h} \overline{i(X)} \to Y$ ist die Identität. Damit ist $\overline{i(X)} = Y$.

Wir müssen nun die Existenz einer Vervollständigung zeigen. Dazu betrachten wir die Menge $\hat{Y} \subset X^{\mathbb{N}}$ der Cauchyfolgen in X. Sei $\hat{i} : X \to \hat{Y}$ die Abbildung, welche jedem Punkt x von X die konstante Folge mit dem Wert x zuordnet.

Wir definieren eine Funktion $\hat{d} : \hat{Y} \times \hat{Y} \to [0, \infty)$ durch

$$\hat{d}((x_n), (y_n)) := \lim_{n\to\infty} d_X(x_n, y_n) .$$

In der Tat ist $(d_X(x_n, y_n))$ eine Cauchyfolge. Um dies einzusehen, benutzen wir die Dreiecksungleichung

$$\begin{aligned} d_X(x_n, y_n) - d(x_m, y_m) &\le d_X(x_n, x_m) + d_X(x_m, y_m) + d_X(y_m, y_n) - d_X(x_m, y_m) \\ &= d_X(x_n, x_m) + d_X(y_m, y_n) . \end{aligned}$$

Durch Vertauschen der Rollen von n und m zeigt man auch

$$d_X(x_m, y_m) - d(x_n, y_n) \leq d_X(x_n, x_m) + d_X(y_m, y_n) \ .$$

Insgesamt ergibt ich

$$|d_X(x_n, y_n) - d(x_m, y_m)| \leq d_X(x_n, x_m) + d_X(y_m, y_n) \ .$$

Es gilt

1. $\hat{d}((x_n), (x_n)) = 0$
2. $\hat{d}((x_n), (y_n)) = \hat{d}((y_n), (x_n))$
3. $\hat{d}((x_n), (y_n)) \leq \hat{d}((x_n), (z_n)) + \hat{d}((z_n), (x_n))$. In der Tat gilt

$$\lim_{n\to\infty} d_X(x_n, y_n) \leq \lim_{n\to\infty} (d_X(x_n, z_n) + d_X(z_n, y_n)) = \hat{d}((x_n), (z_n)) + \hat{d}((z_n), (x_n)) \ .$$

Wir führen nun auf $\hat{Y}$ eine Äquivalenzrelation ein, in welcher $(x_n) \sim (y_n)$ genau dann gilt, wenn $\hat{d}((x_n), (y_n)) = 0$ ist. In der Tat ist diese Relation wegen der obigen Eigenschaften von $\hat{d}$

1. reflexiv
2. symmetrisch
3. transitiv.

Sei $Y := \hat{Y}/\sim$. Die Klasse der Folge (x_n) bezeichnen wir mit $[x_n]$. Wir definieren einen Abstand $d_Y : Y \times Y \to [0, \infty)$ durch

$$d_Y([x_n], [y_n]) := \hat{d}((x_n), (y_n)) \ .$$

Dieser ist wohldefiniert In der Tat, wenn $[x_n] = [x'_n]$, dann ist

$$\hat{d}((x_n), (y_n)) \leq \hat{d}((x_n), (x'_n)) + \hat{d}((x'_n), (y_n)) = \hat{d}((x'_n), (y_n)) \ .$$

Genauso zeigt man $\hat{d}((x'_n), (y_n)) \leq \hat{d}((x_n), (y_n))$.

In der Tat folgt nun aus $d([x_n], [y_n]) = 0$ auch $[x_n] = [y_n]$. Damit wird (Y, d_Y) ein metrischer Raum. Wir setzen $i : X \xrightarrow{\hat{i}} \hat{Y} \to Y$ und beobachten, daß i isometrisch ist.

Wir zeigen nun, daß (Y, d_Y) vollständig ist. Sei $([(x^k_n)]_k)$ eine Cauchfolge von Klassen von Cauchyfolgen. Für $k \in \mathbb{N}$ wählen wir $n(k) \in \mathbb{N}$ derart, daß $d_X(x^k_n, x^k_m) < 2^{-k}$ für alle $n, m \geq n(k)$ gilt. Dann betrachten wir die Folge (y_k) mit $y_k := x^k_{n(k)}$. Dies ist eine Cauchyfolge. In der Tat, sei $\epsilon > 0$ gegeben. Dann wählen wir k_0 derart, daß $2^{-k_0} < \epsilon/3$ und $\hat{d}((x^i_n), (x^j_n)) < \epsilon/3$ für alle $i, j \geq k_0$ gilt. Dann ist für $i, j \geq k_0$

$$d_X(y_i, y_j) \leq d_X(y_i, x^i_n) + d_X(x^i_n, x^j_n) + d_X(x^j_n, y_j) \ .$$

Wir wählen $n \geq \mathtt{max}\{n(i), n(j)\}$ so groß, daß $d_X(x_n^i, x_n^j) < \epsilon/3$ gilt. Die anderen beiden Terme sind auch durch $\epsilon/3$ beschränkt (da $d_X(x_{n(i)}^i, x_n^i) < \epsilon/3$ nach Konstruktion). Folglich gilt für $i, j \geq k_0$, daß $d_X(y_i, y_j) < \epsilon$. Damit haben wir nachgewiesen, daß (y_n) eine Cauchyfolge ist.

Als nächstes zeigen wir, daß $[(x_k^n)] \to [y_k]$ gilt für $n \to \infty$.

Wir wählen $\epsilon > 0$ und $k_0 \in \mathbb{N}$ derart, daß $2^{-k_0} < \epsilon/3$ gilt. Für $n \geq k_0$ gilt dann

$$d_X(x_k^n, y_k) = d_X(x_k^n, x_{n(k)}^k) \leq d_X(x_k^n, x_i^n) + d_Y(x_i^n, x_i^k) + d_X(x_i^k, x_{n(k)}^k) \ .$$

Wir wählen jetzt n so groß, daß $\hat{d}((x_i^n), (x_i^k)) < \epsilon/3$ für alle $k \geq n, k_0$. Dann wählen wir $k \geq n$ so groß, daß $d_X(x_k^n, x_i^n) < \epsilon/3$ für alle $i \geq k$ gilt. Dann wählen wir $i \geq n(k), k$ so groß, daß $d_Y(x_i^n, x_i^k) < \epsilon/3$ gilt. Automatisch haben wir auch $d_X(x_i^k, x_{n(k)}^k) < \epsilon/3$. Damit gilt

$$\lim_{n\to\infty} \lim_{k\to\infty} d_X(x_k^n, y_k) \leq \epsilon \ .$$

Da $\epsilon > 0$ beliebug war, gilt sogar $\lim_{n\to\infty} \lim_{k\to\infty} d_X(x_k^n, y_k) = 0$. Die Folge von Klassen von Cauchyfolgen $[(x_n^k)]$ konvergiert also gegen die Klasse $[y_k]$.

Wir weisen nun die universelle Eigenschaft nach. Sei $f : X \to Z$ eine lipschitzstetige Abbildung in einen vollständigen metrischen Raum. Ist $(x_n) \in \hat{Y}$, dann ist $(f(x_n))$ eine Cauchyfolge in Z. Diese hat einen Grenzwert, und wir setzen

$$\hat{g}((x_n)) := \lim_{n\to\infty} f(x_n) \ .$$

Wir bemerken nun, daß $\hat{g}((x_n))$ nur von der Äquivalenzklasse $[x_n]$ abhängt. In der Tat, wenn $[y_n] = [x_n]$ gilt, dann gilt $d_X(x_n, y_n) \to 0$ und damit $d_Z(f(x_n), f(y_n)) \to 0$, folglich $\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(y_n)$. Wir definieren $g : Y \to Z$ durch

$$g([x_n]) := \hat{g}((x_n)) \ .$$

Wir zeigen nun, daß $g : Y \to Z$ stetig ist. In der Tat ist für $[x_n], [y_n] \in Y$

$$d_Z(g([x_n], g([y_n])) = \lim_{n\to\infty} d_Z(f(x_n), f(y_n)) \leq C \limsup_{n\to\infty} d(x_n, y_n) = C d_Y([x_n], [y_n]) \ ,$$

wobei C die Lipschitzkonstante von f ist.

Da $i(X) \subset Y$ dicht ist, ist eine stetige Fortsetzung von f auf Y eindeutig bestimmt. Damit haben wir die Existenz und Eindeutigkeit der Faktorisierung nachgewiesen. ■

1. Die Vervollständigung von $(C([a,b]), \|\ldots\|_1)$ wird mit $L^1([a,b])$ bezeichnet. Die Norm $\|\ldots\| : C([a,b]) \to \mathbb{R}$ ist lipschitzstetig und setzt sich zu einer Norm auf $L^1([a,b])$ fort. Der metrische Raum $(L^1([a,b]), \|\ldots\|)$ ist vollständig. Folglich ist $(L^1([a,b]), \|\ldots\|)$ ein Banachraum.

2. Die Vervollständigung von $(C([a,b]), \| \dots \|_2)$ wird mit $L^2([a,b])$ bezeichnet. Für $f \in C([a,b])$ ist $< f, \cdots >: C([a,b]) \to \mathbb{R}$ lipschitzstetig und setzt sich fort zu $< f, \cdots >: L^2([a,b]) \to \mathbb{R}$. Für $g \in L^2([a,b])$ ist $< \dots, g >: C([a,b]) \to \mathbb{R}$ wiederum lipschitzstetig und hat eine Fortsetzung auf $L^2([a,b])$. Wir erhalten somit $< ., . >: L^2([a,b]) \times L^2([a,b]) \to \mathbb{R}$. Man kann zeigen, daß das wieder ein Skalarprodukt ist, welches die Ausdehnung der Norm $\|.\|_2$ von $C([a,b])$ auf $L^2([a,b])$ definiert. Das Paar $(L^2([a,b]), < ., . >)$ ist ein Beispiel eines reellen Hilbertraumes. Der normierte Raum $(L^2([a,b]), \|.\|_2)$ ist ein Banachraum.

7.2 Der Satz von Ascoli

Seien (X, d_X) und (B, d_B) metrische Räume. Mit $C(X,B)$ bezeichnen wir die Menge der stetigen Abbildungen von X nach B. Wenn X kompakt und B vollständig ist, dann wird $C(X,B)$ ein vollständiger metrischer Raum mit dem Abstand

$$d_{C(X,B)}(f,g) = \sup_X d_B(f(x), g(x))$$

(Lemma 4.68). Wenn $(B, \|.\|)$ ein Banachraum ist, dann ist $\|f\|_{C(X,B)} := \sup_X \|f(x)\|$ wieder eine Norm auf dem Vektoraum $C(X,B)$, der dann wieder ein Banachraum ist.

Definition 7.4 *Eine Teilmenge $F \subseteq C(X,B)$ heißt **gleichgradig** stetig im Punkt $x \in X$, falls für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so daß aus $f \in F$, $y \in X$ und $d_X(x,y) < \delta$ die Ungleichung $d_B(f(x), f(y))) < \epsilon$ folgt. Ist F in jedem Punkt von X gleichgradig stetig, dann ist F gleichgradig stetig.*

Eine endliche Menge stetiger Abbildung ist gleichgradig stetig.

Lemma 7.5 *Sei $F \subseteq C(X,B)$ gleichgradig stetig in x und (f_n) eine Folge in F derart, daß $f_n \to g$ punktweise für eine Funktion $g : X \to B$. Dann ist g im Punkt x stetig.*

Beweis: Sei $\epsilon > 0$ gegeben. Dann wählen wir $\delta > 0$ derart, daß aus $d_X(y,x) < \delta$ für alle $n \in \mathbb{N}$ folgt $d_B(f_n(y), f_n(x)) < \epsilon$. Sei nun $d_X(y,x) < \delta$. Dann gilt $d_B(g(x), g(y)) = \lim_{n\to\infty} d_B(f_n(x), f_n(y)) \leq \epsilon$. ■

Wir ziehen die folgende Konsequenz aus dem Beweis.

Lemma 7.6 *Sei X kompakt und $F \subseteq C(X,B)$ gleichgradig stetig. Dann ist auch $\overline{F} \subseteq C(X,B)$ gleichgradig stetig.*

Beweis: Sei $x \in X$ und $\epsilon > 0$ gegeben. Dann gibt es ein $\delta > 0$ derart, daß aus $d_X(x,y) < \delta$ für alle $f \in F$ folgt $d_B(f(x), f(y)) < \epsilon$. Ist $g \in \overline{F}$. Dann ist $g = \lim_n f_n$ für eine Folge aus F. Insbesondere gilt $f_n \to g$ punktweise. Der Beweis des Lemmas 7.5 zeigt, daß aus $d_X(x,y) < \delta$ auch $d_B(g(x), g(y)) \leq \epsilon$ folgt. ■

Lemma 7.7 *Sei $D \subseteq X$ eine dichte Teilmenge, $F \subseteq C(X,B)$ gleichgradig stetig und (B,d_B) vollständig. Sei (f_n) eine Folge in F derart, daß $f_n(x) \to \hat{g}(x)$ für alle $x \in D$ für eine Funktion $\hat{g}: D \to B$. Dann hat $\hat{g}$ eine stetige Fortsetzung $g: X \to B$ und es gilt $f_n \to g$ punktweise.*

Beweis: Wir zeigen, daß $(f_n(x))$ für jedes $x \in X$ konvergiert. Dann ist

$$g(x) := \lim_{n\to\infty} f_n(x)$$

nach Lemma 7.5 eine stetige Forsetzung von $\hat{g}$.

Da B vollständig ist, reicht es zu zeigen, daß $(f_n(x))$ eine Cauchyfolge ist. Für $\epsilon > 0$ wählen wir $\delta > 0$ derart, daß aus $d_X(y,x) < \delta$ für alle $n \in \mathbb{N}$ die Abschätzung $d_B(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ folgt. Wir finden weiter $y \in D$ mit $d_X(y,x) < \delta$, da D in X dicht ist. Wir wählen schließlich $n_0 \in \mathbb{N}$ derart, daß aus $n, m \geq n_0$ folgt $d_B(f_n(y), f_m(y)) < \frac{\epsilon}{3}$. Damit gilt für $n, m \geq n_0$

$$d_B(f_n(x), f_m(x)) < d_B(f_n(x), f_n(y)) + d_B(f_n(y), f_m(y)) + d_B(f_m(y), f_m(x)) < \epsilon\ .$$

■

Lemma 7.8 *Sei (X,d_X) kompakt, $F \subseteq C(X,B)$ gleichgradig stetig und (f_n) eine punktweise gegen $g \in C(X,B)$ konvergente Folge in F. Dann gilt $f_n \to g$ gleichmäßig.*

Beweis: Sei $\epsilon > 0$ vorgegeben. Für jedes $x \in X$ wählen $\delta_x > 0$ derart, daß aus $d_X(y,x) < \delta_x$ für alle $n \in \mathbb{N}$ die Abschätzung $d_B(f_n(x), f_n(y)) < \frac{\epsilon}{3}$ folgt. Da X kompakt ist, gibt es eine endliche Teilmenge $D \subseteq X$ derart, daß $X = \cup_{x\in D} B(x, \delta_x)$ gilt.

Wir wählen nun $n_0 \in \mathbb{N}$ derart, daß aus $n \geq n_0$ für alle $x \in D$ folgt $d_B(f_n(x), g(x)) < \frac{\epsilon}{3}$ folgt. Für $y \in X$ wählen wir nun $x \in D$ mit $y \in B(x, \delta_x)$. Dann gilt für $n \geq n_0$

$$\begin{aligned}
&d_B(f_n(y), g(y)) \\
&\quad\leq\ d_B(f_n(y), f_n(x)) + d_B(f_n(x), g(x)) + d_B(g(x), g(y)) \\
&\quad<\ \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&\quad=\ \epsilon
\end{aligned}$$

■

Der Satz von Ascoli charakterisiert kompakte Teilmengen in $C(X,B)$. Wir erinnern daran, daß eine Teilmenge eines metrischen Raumes **relativ kompakt** heißt, wenn ihr Abschluß kompakt ist.

Satz 7.9 (Ascoli) *Sei (X,d_X) eine kompakter metrischer Raum und (B,d_B) ein vollständiger metrischer Raum. Eine Teilmenge $F \subseteq C(X,B)$ ist genau dann relativ kompakt, wenn F gleichgradig stetig und für jedes $x \in X$ die Menge $\{f(x) | f \in F\} \subseteq B$ relativ kompakt ist.*

Korollar 7.10 *Sei (X, d_Y) ein kompakter metrischer Raum. Eine beschränkte gleichgradig stetige Teilmenge $F \subseteq C(X, \mathbb{R}^n)$ ist relativ kompakt.*

Beweis: Sei $C \in \mathbb{R}$ derart, daß $\|f\|_\infty < C$ für alle $f \in F$. Dann ist für jedes $x \in X$ $\{f(x)|f \in F\} \subseteq B(0, C) \subset \mathbb{R}^n$ relativ kompakt. Wenn F gleichgradig stetig ist, dann ist F relativ kompakt nach dem Satz von Ascoli. ■

Beweis: Sei F relativ kompakt. Dann ist $\bar{F}$ kompakt. Die Abbildung $C(X, B) \ni f \mapsto f(x) \in B$ ist stetig und damit $\{f(x)|f \in \bar{F}\} \subseteq B$ als stetiges Bild einer kompakten Menge kompakt. Folglich ist $\{f(x)|f \in F\}$ eine Teilmenge einer kompakten Menge und somit selbst relativ kompakt.

Wir müssen zeigen, daß F gleichgradig stetig ist. Sei $\epsilon > 0$ gegeben. Sei nun $D \subset F$ eine $\frac{\epsilon}{3}$-dichte endliche Teilmenge von F. Wir wählen $\delta > 0$ derart, daß aus $d_X(y, x) < \delta$ für alle $f \in D$ folgt $d_B(f(x), f(y)) < \frac{\epsilon}{3}$ (wir benutzen hier, daß alle stetigen Funktionen auf X gleichmäßig stetig sind). Dann gilt für $g \in F$ und geeignetes $f \in D$ mit $d_{C(X,B)}(f, g) < \frac{\epsilon}{3}$, daß

$$d_B(g(x), g(y)) \leq d_B(g(x), f(x)) + d_B(f(x), f(y)) + d_B(f(y), g(y)) < \epsilon$$

für alle $x, y \in X$ mit $d_X(x, y) < \delta$.

Umgekehrt, sei F gleichgradig stetig und $\{f(x)|f \in F\}$ für jedes x relativ kompakt. Sei $D = \{d_i|i \in \mathbb{N}\} \subseteq X$ eine abzählbare dichte Teilmenge. Sei (f_n) eine Folge in F. Die Folge $(f_n(d_i))$ hat für jedes i eine konvergente Teilfolge. Durch ein Diagonalfolgenargument wählen wir eine Teilfolge (f_{n_j}) derart aus, daß $(f_{n_j}(d_i))$ für jedes $i \in \mathbb{N}$ konvergiert. Diese Teilfolge konvergiert damit nach Lemma 7.7 punktweise gegen ein Element in $C(X, B)$. Nach Lemma 7.8 ist diese Konvergenz gleichmäßig. Wir haben also eingesehen, daß jede Folge in F eine in $C(X, B)$ konvergente Teilfolge besitzt. ■

Sei $C^1([a, b]) \subset C([a, b])$ die Teilmenge der stetig differenzierbaren Funktionen. Auf $C^1([a, b])$ betrachten wir die Norm

$$\|f\|_{C^1} := \|f\|_\infty + \|f'\|_\infty \ .$$

Lemma 7.11 $(C^1([a, b]), \| \dots \|_{C^1}$ *ist ein vollständiger metrischer Raum.*

Beweis: Sei (f_n) eine Cauchyfolge in $C^1([a, b])$. Dann sind (f_n) und (f'_n) Cauchyfolgen in $C([a, b])$. Folglich gilt $f_n \to f$ und $f'_n \to g$ in $C([a, b])$. Es bleibt zu zeigen, daß $f \in C^1([a, b])$ und $g = f'$ ist.

Sei $x \in [a, b]$. Dann schreiben wir

$$f(y) = f(x) + g(x)(y - x) + r(y)(y - x) \ .$$

Nun gilt

$$f_n(y) = f_n(x) + f'_n(x)(y - x) + r_n(y)(y - x) \ .$$

Wir werden zeigen, daß $r_n \to r$ gleichmäßig. Daraus folgt

$$\lim_{y\to x} r(y) = \lim_{y\to x}\lim_{n\to\infty} r_n(y) = \lim_{n\to\infty}\lim_{y\to x} r_n(y) = 0 \ .$$

Es ist klar, daß $r_n(y) \to r(y)$ punktweise gilt. Das Problem ist die Gleichmäßigkeit der Konvergenz. Dazu betrachten wir

$$r_n(y) - r_m(y) = \frac{f_n(y) - f_n(x) - f_m(y) + f_n(x)}{y - x} - f_n'(x) + f_m'(x) \ .$$

Wir wenden den Mittelwertsatz auf $f_n - f_m$ an und erhalten

$$r_n(y) - r_m(y) = f_n'(z) - f_m'(z) - f_n'(x) + f_m'(x)$$

für geeignetes z zwischen x, y. Folglich gilt

$$\|r_n - r_m\|_\infty \le 2\|f_n - f_m\| \ .$$

Damit ist (r_n) eine gleichmäßige Cauchyfolge. ■

Lemma 7.12 *Die Einbettung $i : C^1([a,b]) \to C([a,b])$ ist kompakt im folgenden Sinne: Für jede beschränkte Teilmenge $F \subset C^1([a,b])$ ist $i(F)$ relativ kompakt.*

Beweis: Ist $F \subset C^1([a,b])$ beschränkt, dann ist sicher $i(F)$ auch beschränkt. Sei $C := \sup_{f\in F}\|f\|_{C^1}$. Ist $\epsilon > 0$ gegeben, dann gilt für $\delta := C^{-1}\epsilon$, $f \in F$ und $|x - y| < \delta$

$$|f(x) - f(y)| < C|x - y| = \epsilon \ .$$

Sei $i(F)$ ist also gleichgradig stetig und beschränkt. Nach Korollar 7.10 ist $i(F)$ relativ kompakt. ■

8 Differentialrechnung in mehreren Veränderlichen

8.1 Ableitung als lineare Approximation

Zur Erinnerung, sei $U \subseteq \mathbb{R}$ offen, $u \in U$ und $f : U \to \mathbb{R}$ eine Funktion. Dann war f im Punkt u differenzierbar, wenn

$$f(v) = f(u) + a(v - u) + r(v)(v - u)$$

gilt für eine Zahl $a \in \mathbb{R}$ und eine Funktion $r : U \to \mathbb{R}$ mit $\lim_{v\to u} r(v) = 0$. Die Zahl a heißt dann die Ableitung und kann als Grenzwert des Differenzenquotienten

$$\Delta_u(f)(v) := \frac{f(v) - f(u)}{v - u} \ , \quad a = \lim_{v\to u}\Delta_u(f)(v)$$

bestimmt werden.

Die Zahl a bestimmt, wie schnell die Funktion f im Punkt u wächst oder fällt. Insbesondere, wenn die Funktion f dort ein Extremum hat, dann ist $a = 0$. Dies kann zur Bestimmung von Extremwerten ausgenutzt werden. Der Term $f(u)$ in $f(v) = f(u)+a(v-u)+r(v)(v-u)$ kann als Approximation von f durch eine konstante Funktion angesehen werden. Der Term $a(v-u)$ ist eine Korrektur durch eine lineare Funktion, und $r(v)(v-u)$ ist ein Restglied. Die Bedingung $r(v) \to 0$ für $v \to u$ sichert, daß dieser Rest wirklich kleiner als die anderen beiden Terme ist.

Wörtlich dieselben Betrachtungen gelten, wenn die Funktion f Werte in einem Vektorraum F hat, in welchem ein Konvergenzbegriff definiert ist, also eine Topologie. In der Tat kann unter Benutzung der Vektorraumstruktur der Differenzenquotient $\Delta_u(f)$ immer noch gebildet werden. Mit Hilfe des Konvergenzbegriffs kann man die Existenz $f'(u) := \lim_{v\to u} \Delta_u(f)(v)$ untersuchen. Natürlich ist dann $f'(u) \in F$. Für spätere Betrachtungen ist es günstig, die Ableitung f' als ein Element von $\mathtt{Hom}(\mathbb{R}, F)$ zu verstehen. Natürlich gibt es einen Isomorphismus $F \cong \mathtt{Hom}(\mathbb{R}, F)$ mit $f \mapsto \{t \mapsto tf\}$.

Ist $F = \mathbb{R}^n$, dann ist kann man f durch seine Komponenten $f_i : U \to \mathbb{R}$ beschreiben: $f = (f_1, \ldots, f_n)$. Die Ableitung $f'(u)$ existiert genau dann, wenn die Ableitungen aller Komponenten $f_i'(u)$ existiert, und es gilt $f'(u) = (f_1'(u), \ldots, f_n'(u))$.

Sehen wir uns ein komplizierteres Beispiel an. Wir betrachten die Abbildung $\Phi : \mathbb{R} \to C([0,1])$

$$\Phi : \mathbb{R} \ni x \mapsto \Phi(x) := \{t \mapsto \mathsf{e}^{tx}\} \ .$$

Die Konvergenz in $C([0,1])$ soll durch die Norm $\|f\|_\infty := \sup_{t\in[0,1]} |f(t)|$ bestimmt sein. Wir betrachten den Quotienten

$$\Delta_u(\Phi)(v)(t) = \frac{\mathsf{e}^{vt} - \mathsf{e}^{ut}}{v-u} \ .$$

In der Tat existiert

$$\lim_{u\to v} \Delta_u(\Phi)(v) = \Psi \ , \quad \Psi(t) = t\mathsf{e}^{ut} \ .$$

Wir schätzen $\|\Delta_u(\Phi)(v) - \Psi\|_\infty$ mit Hilfe der Taylorformel ab. Dazu schreiben wir

$$\mathsf{e}^{vt} = \mathsf{e}^{ut} + t\mathsf{e}^{ut}(v-u) + (v-u)^2\frac{t^2\mathsf{e}^{\xi_t t}}{2}$$

für geeignetes $\xi_t \in [u, v]$. Damit ist

$$|\Delta_u(\Phi)(v)(t) - \Psi(t)| = (v-u)\frac{t^2\mathsf{e}^{\xi_t t}}{2}$$

und es gilt

$$\lim_{v\to u} \|\Delta_u(\Phi)(v) - \Psi\|_\infty = 0 \ .$$

Für die Erklärung von Konvergenzbegriffen ist man nicht auf Normen beschränkt. Sei F ein Vektorraum und F' sein Dualraum. Wir betrachten eine Teilmenge $N \subset F'$. Diese legt einen Konvergenzbegriff fest. In der Tat, wir können dann eine Folge von Vektoren (f_i) für konvergent erklären, wenn die Folgen $(n(f_i))$ für alle $n \in N$ konvergieren. Um letztlich die Eindeutigkeit von Grenzwerten zu sichern, nehmen wir an, daß N die **Punkte trennt**, also für $f \in F$ aus $n(f) = 0$ für alle $n \in N$ folgt $f = 0$.

Ist $f : U \to F$ eine Abbildung, dann ist diese in u differenzierbar genau dann, wenn es ein Element $f' \in F$ gibt, so daß für alle $n \in N$ die Funktion $n \circ f : U \to \mathbb{R}$ in u differenzierbar ist und in u die Ableitung $n(f')$ hat. Es reicht hier nicht aus, lediglich die Differenzierbarkeit aller $n \circ f$ zu fordern.

Ein typisches Beispiel ist $F = C([0,1])$ und $N \subset F'$ ist die Menge der Punktauswertungen. Daß heißt, $N \cong [0,1]$ und dem Punkt $n \in [0,1]$ entspricht die Linearform $f \mapsto f(n)$. Die Punktauswertungen trennen die Funktionen. Der Konvergenzbegriff ist die uns schon bekannte punktweise Konvergenz.

Wir wollen nun diese Theorie auf höhere Dimensionen im Argument verallgemeinern. Damit wir von linearen Approximationen sprechen können, muß $f : U \to F$ eine Abbildung von einer Teilmenge eines Vektorraumes $U \subseteq E$ in einen weiteren Vektorraum F sein. Wir schreiben

$$f(v) = f(u) + A(v-u) + R(v)$$

für eine lineare Abbildung $A : E \to F$ und eine Abbildung $R : U \to F$. Wir müssen uns Bedingungen überlegen, welche sichern, daß die drei Terme wirklich Approximationen unterschiedlicher Ordnung beschreiben. Um die Größe von Vektoren messen zu können, verwenden wir in dieser Vorlesung Normen. Es sei aber bemerkt, daß die Theorie auch in allgemeineren Kontexten topologischer Vektorräume funktioniert.

Seien also $\|.\|_E$ und $\|.\|_F$ Normen auf E und F.

Definition 8.1 *Eine Zahl $C \in \mathbb{R}$ ist eine **Schranke** von A, wenn*

$$\|Ax\|_F \le C\|x\|_E$$

*für alle $x \in E$ gilt. Wenn A eine Schranke besitzt, dann nennen wir A **beschränkt**. Daß Infimum über alle oberen Schranken ist heißt **Norm** von A. Es gilt also*

$$\|A\|_{E\to F} := \inf\{C \in \mathbb{R} | \{\forall x \in E | \|A\|_F \le C\|x\|_E\}\} \ .$$

*Mit $B(E,F)$ bezeichen wir die Menge der **beschränkten linearen Abbildungen**.*

Für beschränktes A gilt

$$\|A\| = \sup_{0\neq x\in E} \frac{\|Ax\|_F}{\|x\|_E} \ .$$

In der Differentialrechnung werden wir annehmen, daß die lineare Approximation durch eine beschränkte Abbildung gegeben wird. Wir werden später sehen, daß Beschränktheit zur Stetigkeit gleichwertig ist. Damit sichert diese Bedingung, daß eine differenzierbare Abbildung automatisch stetig ist. Diese Eigenschaft ist für das Wohlverhalten der Theorie von entscheidender Bedeutung und wird etwa später beim Beweis der Kettenregel benutzt.

Die folgende Bedingung sichert, daß das Restglied "kleiner" als die lineare Approximation ist:

$$\lim_{v\to u} \frac{\|R(v)\|_F}{\|v-u\|_E} = 0 \ .$$

Eine nicht-triviale lineare Abbildung kann diese Bedingung nicht erfüllen. In der Tat, wäre $R(v) = R_0 + A(v-u)$. Dann muß wegen $R(u) = 0$ auch $R_0 = 0$ gelten. Wäre $A \neq 0$, dann gäbe es $x \in E$ mit $Ax \neq 0$. Es gilt dann

$$\frac{\|R(u+\lambda x)\|_F}{\|\lambda x\|_E} = \frac{\|R(u+x)\|_F}{\|x\|_E}$$

für alle $\lambda \in \mathbb{R}$. Insbesondere geht die linke Seite nicht gegen Null für $\lambda \to 0$.

Ein normierter Vektorraum $(E, \|.\|_E)$ ist insbesondere metrisch und topologisch. Man kann von Grenzwerten und offenen Mengen reden.

Definition 8.2 *Seien $(E, \|.\|_E)$ und $(F, \|.\|_F)$ normierte Vektorräume. Sei $u \in U \subseteq E$ und $f : U \to F$. Die Abbildung f ist in u* ***differenzierbar****, wenn es eine beschränkte lineare Abbildung $A \in B(E,F)$ gibt, so daß*

$$f(v) = f(u) + A(v-u) + R(v)$$

und

$$\lim_{v\to u} \frac{\|R(v)\|_F}{\|(v-u)\|_E} = 0 \ .$$

Die Abbildung $df(u) := A$ heißt ***Ableitung*** *von f im Punkt u.*

Lemma 8.3 *Die Ableitung ist wohldefiniert.*

Beweis: Wir nehmen an, daß

$$f(v) = f(u) + A'(v-u) + R'(v)$$

und

$$\lim_{v\to u} \frac{\|R'(v)\|_F}{\|(v-u)\|_E} = 0$$

für ein weiteres $A' \in B(E,F)$ und R' gilt. Dann ist

$$(A-A')(v-u) = R(v) - R'(v) \ .$$

Sei $(A-A')(x) \neq 0$ für ein $x \neq 0$. Die Abbildung $\mathbb{R} \to E$, $\lambda \mapsto u + \lambda x$, ist stetig. Damit ist $u + \lambda x \in U$ für genügend kleine λ. Hier nutzen wir aus, daß U eine Umgebung von u ist. Für diese λ ist

$$\lambda(A-A')(x) = (A-A')(\lambda x) = R(u+\lambda x) - R'(u+\lambda x) \ .$$

Wir schließen
$$(A - A')(x) = \lambda^{-1}(R(u+\lambda x) - R'(u+\lambda x)) \ .$$
Dann gilt
$$\|(A - A')(x)\|_F \leq \lambda^{-1}\|R(u+\lambda x)\|_F + \lambda^{-1}\|R'(u+\lambda x)\|_F \ .$$
Es gilt
$$\lim_{\lambda\to 0} \lambda^{-1}\|R(u+\lambda x)\|_F = \lim_{\lambda\to 0} \frac{\|R(u+\lambda x)\|_F}{\lambda\|x\|_E} = 0$$
und ein analoges Ergebnis für R'. Daraus folgt $(A - A')(x) = 0$, ein Widerspruch. ■
Hier sind einige Beispiele.

1. Sei $A \in B(E,F)$ und $f_0 \in F$. Wir betrachten $f : E \to F$, $f(e) = f_0 + Ae$. Dann gilt
$$f(e') = f(e) + A(e' - e) \ .$$
Folglich ist f differenzierbar in e und es gilt $df(e) = A$.

2. Sei $f : E \to F$ konstant. Dann ist $f(e') = f(e)$, also $df(e) = 0$.

8.2 Funktionalanalytische Grundlagen

Seien $(E, \|.\|_E)$ und $(F, \|.\|_F)$ normierte Vektorräume. Im letzten Abschnitt hatten wir den Raum $B(E,F)$ der beschränkten linearen Abbildungen eingeführt. Insbesondere war für jedes $A \in B(E,F)$ die Norm $\|A\|_{E\to F}$ definiert.

Lemma 8.4 *Der Raum $B(E,F)$ ist ein Vektorraum. Durch $A \to \|A\|_{E\to F}$ wird eine Norm auf $B(E,F)$ gegeben.*

Beweis: Es gilt
$$\|A\|_{E\to F} \geq 0$$
nach Konstruktion. Weiter ist
$$\|\lambda A\|_{E\to F} = \sup_{0\neq x\in E} \frac{\|\lambda A x\|_F}{\|x\|_E} = \lambda \sup_{0\neq x\in E} \frac{\|Ax\|_F}{\|x\|_E} = \lambda\|A\|_{E\to F} \ .$$
Insbesondere ist mit $A \in B(E,F)$ auch $\lambda A \in B(E,F)$. Für $A, A' \in B(E,F)$ und beliebiges $x \in E$ gilt
$$\|(A + A')(x)\|_F \leq \|Ax\|_F + \|A'x\|_F \leq \|A\|_{E\to F}\|x\| + \|A'\|_{E\to F}\|x\|_E \ .$$
Daraus folgt $A + A' \in B(E,F)$ und $\|A + A'\|_{E\to F} \leq \|A\|_{E\to F} + \|A'\|_{E\to F}$. Ist $\|A\|_{E\to F} = 0$, dann gilt $\|Ax\|_F = 0$ für alle $x \in E$, also $A = 0$. ■

Lemma 8.5 *Ist F vollständig (Banach), so auch $B(E,F)$.*

Beweis: Sei $(A_n) \in B(E,F)$ eine Cauchyfolge. Dann ist für jedes $e \in E$ die Folge $A_n e$ eine Cauchyfolge. In der Tat ist $\|A_n e - A_m e\|_F \leq \|A_n - A_m\|_{E\to F}\|e\|_E$. Wir definieren $A : E \to F$ durch $Ae := \lim_{n\to\infty} A_n e$. Dann gilt für $e, e' \in E$ und $\lambda \in \mathbb{R}$, daß

$$\begin{aligned} A(e+\lambda e') &= \lim_{n\to\infty} A_n(e+\lambda e') \\ &= \lim_{n\to\infty} (A_n e + \lambda A_n e') \\ &= \lim_{n\to\infty} A_n e + \lim_{n\to\infty} \lambda A_n e' \\ &= Ae + \lambda Ae' \ . \end{aligned}$$

Folglich ist A linear. Weiter gilt

$$\|A_n e\|_F \leq \sup_{n\in\mathbb{N}} \|A_n\|_{E\to F}\|e\|_F \ ,$$

also $\|A\|_{E\to F} \leq \sup_{n\in\mathbb{N}} \|A_n\|_{E\to F} < \infty$. Damit ist $A \in B(E,F)$. Wir zeigen schließlich, daß $\lim_{n\to\infty} A_n = A$. Sei $\epsilon > 0$ gegeben. Wir wählen $N \in \mathbb{N}$ derart, daß für $n, m \geq N$ gilt $\|A_n - A_m\|_{E\to F} \leq \frac{\epsilon}{2}$. Für $e \in E$ wählen wir nun $m \geq N$ so groß, daß $\|A_m e - Ae\|_F \leq \frac{\epsilon\|e\|_E}{2}$. Dann gilt

$$\begin{aligned} \|(A_n - A)e\|_F &= \|(A_n - A_m)e + (A_m - A)e\|_F \\ &\leq \|(A_n - A_m)e\|_F + \|(A_m - A)e\|_F \\ &\leq \frac{\epsilon\|e\|_E}{2} + \frac{\epsilon\|e\|_E}{2} \\ &= \epsilon\|e\|_E \ . \end{aligned}$$

Also gilt für $n \geq N$, daß $\|A_n - A\|_{E\to F} \leq \epsilon$. ■

Mit $E' := B(E,\mathbb{R})$ bezeichnen wir den Raum der stetigen linearen Abbildungen. Mit $\|.\|_{E'} := \|.\|_{E\to\mathbb{R}}$ ist dies wieder ein normierter Vektorraum.

Definition 8.6 *$(E', \|.\|_{E'})$ ist der* ***duale normierte Raum*** *von $(E, \|.\|_E)$.*

Zum Beispiel sind $\delta_0 : f \mapsto f(0)$ und $I : f \mapsto \int_{-1/2}^0 f(x)dx$ Elemente von

$$(C([-1,1]), \|.\|_\infty)' \ .$$

Es gilt

$$\|\delta_0\|_{C([-1,1])'} = 1 \ , \quad \|I\|_{C([-1,1])'} = 1/2 \ .$$

Die Maßtheorie wird eine vollständige Beschreibung von $C([-1,1])'$ liefern.

Seien $(E, \|.\|_E)$, $(F, \|\|_F)$ und $(G, \|.\|_G)$ normierte Räume, $A \in B(E,F)$ und $B \in B(F,G)$.

Lemma 8.7 *Es gilt $\|B \circ A\|_{E\to G} \leq \|A\|_{E\to F}\|B\|_{F\to G}$.*

Beweis: Wir rechnen

$$\|(B \circ A)(e)\|_G \le \|B\|_{F\to G}\|Ae\|_F \le \|B\|_{F\to G}\|A\|_{E\to F}\|e\|_E \ .$$

Folglich ist $B \circ A$ durch $\|B\|_{F\to G}\|A\|_{E\to F}$. ■
Insbesondere ist für $A \in B(E,E)$

$$\|A^n\|_{E\to E} \le \|A\|^n_{E\to E} \ .$$

Wir nehmen an, daß E und F ein Banachräume sind und betrachten $A_0 \in B(E,F)$.

Lemma 8.8 *Wenn A_0 ein Inverses $A_0^{-1} \in B(F,E)$ hat, dann existiert eine offene Umgebung $U \subset B(E,F)$ von A, so daß jedes $A \in U$ ein Inverses $A^{-1} \in B(F,E)$ besitzt. Die Abbildung $I : U \to B(F,E)$, $I : A \mapsto A^{-1}$ ist differenzierbar und besitzt die Ableitung $dI(A_0)(B) = -A_0^{-1}BA_0^{-1}$.*

Beweis: Wir betrachten die Neumann Reihe

$$N(A) := A_0^{-1}\sum_{n=0}^{\infty}(-1)^n(BA_0^{-1})^n \ ,$$

wobei $B := A - A_0$ ist. Es gilt

$$\|(BA_0^{-1})^n\|_{F\to F} \le \|BA_0^{-1}\|^n_{F\to F} \le \|B\|^n_{E\to F}\|A_0^{-1}\|^n_{E\to F} \ .$$

Sei $\epsilon := \|A_0^{-1}\|^{-1}_{F\to E}$ und $U := B(A_0,\epsilon)$. Dann konvergiert die Neumann Reihe für $A \in U$ (hier benutzen wir, daß $B(F,F)$ vollständig ist). Es gilt weiter

$$\begin{aligned}
AN(A) &= (A_0+B)A_0^{-1}\sum_{n=0}^{\infty}(-1)^n(BA_0^{-1})^n \\
&= \sum_{n=0}^{\infty}(-1)^n(BA_0^{-1})^n + BA_0\sum_{n=0}^{\infty}(-1)^n(BA_0^{-1})^n \\
&= \sum_{n=0}^{\infty}(-1)^n(BA_0^{-1})^n - \sum_{n=1}^{\infty}(-1)^n(BA_0^{-1})^n \\
&= 1
\end{aligned}$$

Analog zeigt man $N(A)A = 1$. Damit ist $N(A) = A^{-1}$. Wir schreiben nun

$$N(A) = A_0^{-1} - A_0^{-1}(A-A_0)A_0^{-1} + \sum_{n=2}^{\infty}(-1)^n(BA_0^{-1})^n \ .$$

Die Abbildung

$$B(E,F) \ni X \mapsto A_0^{-1}XA_0^{-1} \in B(E,F)$$

ist durch $\|A_0^{-1}\|_{F\to E}^2$ beschränkt. Weiter gilt

$$\begin{array}{rcl} \dfrac{\|\sum_{n=2}^\infty(-1)^n(BA_0^{-1})^n\|_{F\to F}}{\|B\|_{E\to F}} & \leq & \displaystyle\sum_{n=2}^\infty \|A_0^{-1}\|_{F\to E}^n\|B\|_{E\to F}^{n-1} \\ & \overset{\|B\|_{E\to F}\to 0}{\to} & 0\ . \end{array}$$

Definition 8.9 *Sei $GL(E) := \{A \in B(E,E) | A^{-1} \text{existiert in } B(E,E)\}$.*

Korollar 8.10 *Wenn E ein Banachraum ist, dann ist $GL(E)$ eine offene Teilmenge von $B(E,E)$.*

Wir berachten nun den Zusammenhang zwischen Beschränktheit und Stetigkeit.

Lemma 8.11 *Folgende Aussagen sind äquivalent:*

1. *$A \in B(E,F)$ ist beschränkt.*
2. *$A : E \to F$ ist linear und stetig in 0.*
3. *$A : E \to F$ ist linear und gleichmäßig stetig.*

Beweis: Sei $A \in B(E,F)$. Sei $\epsilon > 0$ gegeben. Wir setzen $\delta := \frac{\epsilon}{\|A\|_{E\to F}}$. Dann gilt für $\|x\|_E \leq \delta$, daß $\|Ax\|_F \leq \epsilon$. Folglich ist A in $0 \in E$ stetig.

Sei A linear und in $0 \in E$ stetig. Sei $\epsilon > 0$ gegeben und $\delta > 0$ so gewählt, daß aus $\|x\|_E \leq \delta$ folgt $\|Ax\|_F \leq \epsilon$. Sei nun $u \in E$. Dann folgt aus $\|x-u\|_E \leq \delta$ auch $\|Ax - Au\|_F \leq \epsilon$. Dies zeigt die gleichmäßige Stetigkeit.

Sei $A : E \to F$ linear und gleichmäßig stetig. Dann ist A insbesondere in Null stetig. Sei $\delta > 0$ so gewählt, daß aus $\|x\|_E \leq \delta$ folgt $\|Ax\|_F \leq 1$. Dann ist δ^{-1} eine Schranke für A. Folglich gilt $A \in B(E,F)$. ■

In dieser Vorlesung werden wir uns vornehmlich mit Abbildungen zwischen endlich-dimensionalen Vektorräumen beschäftigen. Der folgende Satz liefert eine bedeutende Vereinfachung der Theorie.

Satz 8.12 *Ist E endlich-dimensional, dann ist jede lineare Abbildung $A : E \to F$ stetig. In anderen Worten, $B(E,F)$ fällt mit dem Raum aller linearen Abbildungen zusammen.*

Beweis: Wir zeigen zunächst:

Lemma 8.13 *Jede Norm $\|.\|$ auf $\mathbb{R}^n$ ist zu $\|.\|_1$ äquivalent.*

Beweis: Sei $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$. Nach Definition gilt für $x = (x_1, \ldots, x_n)$

$$\|x\|_1 = \|\sum_{i=1}^n x_i e_i\|_1 := \sum_{i=1}^n |x_i| \ .$$

Sei nun $\|.\|$ eine weitere Norm. Dann gilt

$$\|x\| := \|\sum_{i=1}^n x_i e_i\| \le \sum_{i=1}^n |x_i|\|e_i\| \le \|x\|_1 C$$

mit

$$C := \mathtt{max}\{\|e_1\|, \ldots, \|e_n\|\} \ .$$

Es bleibt zu zeigen, daß für ein $D \in \mathbb{R}$ auch

$$\|x\|_1 \le D\|x\|$$

gilt. Sei $\gamma := \inf_{\{\|x\|_1 = 1\}} \|x\|$. Sei $(x^{(k)})$ eine Folge mit $\|x^{(k)}\|_1 = 1$ und $\|x^{(k)}\| \stackrel{k\to\infty}{\to} \gamma$. Dann gilt $|x_i^{(k)}| \le 1$ für alle k. Wir ersetzen die Folge $(x^{(k)})$ durch eine Teilfolge derart, daß die Folgen $(x_i^{(k)})$ reeller Zahlen konvergieren. Dann gilt $x^{(k)} = \sum_{i=1}^n x_i^{(k)} e_i \stackrel{k\to\infty}{\to} x$. Es gilt $\|x\|_1 = 1$ und damit $x \ne 0$. Die Funktion $\mathbb{R}^n \to \mathbb{R}$,

$$(y_1, \ldots, y_n) \mapsto \|\sum_{i=1}^n y_i e_i\|$$

ist stetig. Folglich ist auch

$$\gamma = \lim_{k\to\infty} \|x^{(k)}\| = \|x\| \ne 0 \ .$$

Wir setzen $D := \gamma^{-1}$. Für $y \in \mathbb{R}^n$ setzen wir $x := \frac{y}{\|y\|_1}$. Dann gilt $\gamma \le \|x\|$, also $\|y\|_1 \le D\|y\|$. ■

Dieser Satz zeigt, daß $\mathtt{id} : (\mathbb{R}^n, \|.\|_1) \to (\mathbb{R}^n, \|.\|)$ und auch die dazu inverse Abbildung beschränkt sind.

Sei nun $(E, \|.\|_E)$ ein endlich-dimensionaler normierter Verktorraum. Wir wählen eine Basis (e_i) und setzen

$$C := n\mathtt{max}\{\|Ae_i\|_F\} \ .$$

Sei $I : \mathbb{R}^n \to E$ der lineare Isomorphismus $I(x) = \sum_{i=1}^n x_i e_i$. Dann ist $A \circ I : (\mathbb{R}^n, \|.\|_1) \to (F, \|.\|_F)$ durch C beschränkt. Wir definieren die Norm $\|x\| := \|Ix\|_E$ auf $\mathbb{R}^n$. Sei $D \in \mathbb{R}$ derart, daß $\|x\|_1 \le D\|x\|$. Dann gilt

$$\|Ae\|_F \le C\|I^{-1}(e)\|_1 \le CD\|I^{-1}(e)\| = CD\|e\|_E \ .$$

Folglich ist $A : (E, \|.\|_E) \to (F, \|.\|_F)$ beschränkt. ■

8.3 Die Ableitung

In diesem Abschnitt untersuchen wir die Ableitung genauer. Insbesondere betrachten wir die Beziehung zum schon bekannten Begriff in einer Dimension.

Sei $u \in U \subseteq E$ und $f : U \to F$ wie bisher. Sei $\xi \in E$. Dann erhalten wir eine Abbildung $\gamma : \mathbb{R} \to E$ durch $\gamma(t) := u + t\xi$, die parametrisierte Gerade durch u mit Richtung ξ. Diese Abbildung ist stetig. In der Tat ist

$$\|\gamma(t) - \gamma(s)\|_E = \|t - s\xi\|_E = |(t - s)|\|\xi\|_E \ .$$

Damit ist $\gamma^{-1}(U) \subseteq \mathbb{R}$ eine offene Umgebung von $0 \in \mathbb{R}$ und $\gamma^* f := f \circ \gamma : \gamma^{-1}(U) \to \mathbb{R}$ definiert.

Definition 8.14 *f hat im Punkt u die* ***Richtungsableitung*** *$d_\xi f(u) \in F$ in Richtung ξ, wenn $\gamma^* f$ im Punkt 0 die Ableitung $d_\xi f(u) := (\gamma^* f)'(0)$ hat.*

Lemma 8.15 *Wenn f im Punkt u differenzierbar ist, dann besitzt f im Punkt u alle Richtungsableitungen und es gilt*

$$d_\xi f(u) = df(u)(\xi) \ .$$

Beweis: In der Tat gilt

$$\begin{aligned} (\gamma^* f)'(0) &= \lim_{s \to 0} \frac{df(u)(s\xi) + R(s\xi + u)}{s} \\ &= df(u)(\xi) + \lim_{s \to 0} \frac{R(s\xi + u)}{s} \\ &= df(u)(\xi) \ . \end{aligned}$$

■

Der Raum $\mathbb{R}^n$ hat die Standardbasis (e_i).

Definition 8.16 *Im Fall $E = \mathbb{R}^n$ heißt*

$$\partial_i f(u) := d_{e_i} f(u)$$

die i-te ***partielle Ableitung*** *von f im Punkt u.*

Die lineare Abbildung $df(u) : \mathbb{R}^n \to F$ ist durch die Werte $df(u)(e_i) = \partial_i f(u)$ der partiellen Ableitungen vollständig bestimmt. Um $df(u)$ zu berechnen, müssen wir folglich nur $\partial_i f(u)$ für $i = 1, \dots n$ berechnen. Ist $F \cong \mathbb{R}^m$, mit der Basis $(e_1, \dots, e_m)$, dann haben wir eine Identifikation $\mathtt{Hom}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathtt{Mat}(n, m)$ (n Spalten, m Zeilen).

Definition 8.17 *Die Matrix, welche $df(u)$ unter dieser Identifikation darstellt, bezeichnet man als* ***Jacobimatrix*** *$Jf(u) \in \mathtt{Mat}(n, m)$.*

Es gilt
$$f(x)=\sum_{j=1}^m f_j(x)e_j$$
und folglich
$$\partial_i f(u)=\sum_{j=1}^m \partial_i f_j(u)e_j \ ,$$
Daraus folgt
$$Jf(u)_{j,i}=\partial_i f_j(u) \ .$$
Allgemeiner können wir also die Jacobimatrix $Jf(u)\in \mathtt{Mat}(n,m)$ einer partiell differenzierbaren Abbildung $\mathbb{R}^n\supseteq U\xrightarrow{f}\mathbb{R}^m$ durch diese Formel bilden.

Sei etwa $f:\mathbb{R}^3\to\mathbb{R}^2$ durch $f(x,y,z)=(3x^2+2xy, z^2y-z)$ gegeben. Dann ist
$$Jf(x,y,z)=\begin{pmatrix} 6x+2y & 2x & 0\\ 0 & z^2 & 2zy-1\end{pmatrix} \ .$$
Die Richtungsableitung etwa in die Richtung $(1,1,0)$ berechnet sich dann wie folgt:
$$\begin{aligned}
d_{\begin{pmatrix}1\\1\\0\end{pmatrix}} f(x,y,z) &= Jf(x,y,z)\begin{pmatrix}1\\1\\0\end{pmatrix}\\
&= \begin{pmatrix} 6x+2y & 2x & 0\\ 0 & z^2 & 2zy-1\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}\\
&= \begin{pmatrix}8x+2x\\ z^2\end{pmatrix}
\end{aligned}$$
Es sei aber bemerkt, daß diese Rechnungen nur vorbehaltlich der Tatsache, daß F wirklich im Punkt u differenzierbar ist, gelten.

Die obige Diskussion zeigt folgende Konklusionen:

differenzierbar $\to$ *ex.alle Richtungsableitungen* $\to$ *ex. alle partiellen Ableitungen*

Die folgenden Beispiele zeigen, daß die Umkehrungen jeweils nicht gelten.

1. Die Abbildung $f:\mathbb{R}^2\to\mathbb{R}^2$,
$$f(x,y)=\begin{cases}\frac{x|y|}{x^2+y^2} & (x,y)\neq 0\\ 0 & (x,y)=0\end{cases}$$
besitzt beide partielle Ableitungen. In der Tat ist $f(x,0)=0=f(0,y)$. Es gilt aber
$$f(t(a,b))=\frac{t|t|ab}{t^2(a^2+b^2)}=\mathtt{sign}(t)\frac{ab}{(a^2+b^2)} \ .$$
Folglich besitzt f außer den Richtungsableitungen in die Koordinatenrichtungen keine weiteren.

2. Die Funktion
$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2} & (x,y) \neq 0 \\ 0 & (x,y) = 0 \end{cases}$$
besitzt alle Richtungsableitungen in 0, ist aber nicht differenzierbar. In der Tat gilt $f(t(a,b)) = t\frac{a^3-b^3}{a^2+b^2}$ und damit $d_{(a,b)}f(0) = \frac{a^3-b^3}{a^2+b^2}$. Da $(a,b) \mapsto d_{(a,b)}f(0)$ nicht mal linear ist, kann f im Punkt auch nicht differenzierbar gewesen sein.

Die Situation wird etwas besser, wenn man das Verhalten der Ableitungen auf einer Umgebung von u hinzuzieht. Ohne Beschränkung der Allgemeinheit sei diese U.

Sei $U \subseteq E$ und $f : U \to F$ gegeben.

Definition 8.18 *Die Funktion f heißt **auf U differenzierbar**, wenn sie in jedem Punkt $u \in U$ differenzierbar ist. Sie ist auf U **stetig differenzierbar**, wenn sie differenzierbar und $U \ni u \mapsto df(u) \in B(E,F)$ stetig ist.*

Definition 8.19 *Sei $E = \mathbb{R}^n$. Die Funktion heißt auf U **partiell differenzierbar**, wenn $\partial_i f(u)$ für alle $u \in U$ und $i = 1,\ldots,n$ existiert. Sie heißt **stetig partiell differenzierbar**, wenn $U \ni u \mapsto \partial_i f(u) \in F$ stetig ist für $i = 1,\ldots,n$.*

Sei $U \subset \mathbb{R}^n$.

Lemma 8.20 *Die Funktion $f : U \to F$ ist genau dann stetig differenzierbar, wenn sie stetig partiell differenzierbar ist.*

Beweis: Wir verwenden, daß für $e \in E$ die lineare Abbildung $B(E,F) \ni A \mapsto Ae \in F$ stetig ist. In der Tat gilt $\|Ae\|_F \leq \|e\|_E \|A\|_{E\to F}$. Sei f stetig differenzierbar. Dann ist $U \ni u \mapsto d_{e_i}f(u) = df(u)(e_i)$ stetig. Damit ist f stetig partiell differenzierbar.

Sei nun f stetig partiell differenzierbar. Wir schreiben unter Verwendung des eindimensionalen Mittelwertsatzes mit $y_i \in [u_i, v_i]$,

$$\begin{aligned}
f(v) &= f(v_1,\ldots,v_{n-1},u_n) + (v_n - u_n)(\partial_n f)(v_1,\ldots,v_{n-1},y_n) \\
&= f(v_1,\ldots,v_{n-2},u_{n-1},u_n) + (v_{n-1}-u_{n-1})\partial_{n-1} f(v_1,\ldots,v_{n-2},y_{n-1},u_n) \\
&\quad + (v_n - u_n)(\partial_n f)(v_1,\ldots,v_{n-1},y_n) \\
&= f(u) + \sum_{i=1}^n (v_i - u_i)\partial_i f(v_1,\ldots,v_{i-1},y_i,\ldots,y_n) \\
&= f(u) + \sum_{i=1}^n (v_i - u_i)\partial_i f(u_1,\ldots,u_n) + R(v) \\
R(v) &:= \sum_{i=1}^n (v_i - u_i)\left(\partial_i f(v_1,\ldots,v_{i-1},y_i,\ldots,y_n) - \partial_i f(u_1,\ldots,u_n)\right) .
\end{aligned}$$

Nun gilt

$$\frac{\|R(v)\|_F}{\|u-v\|_1} \leq n\mathtt{max}_i \|\partial_i f(v_1,\ldots,v_{i-1},y_i,\ldots,y_n) - \partial_i f(u_1,\ldots,u_n)\| \stackrel{v\to u}{\to} 0 .$$

In unserem Beispiel $f : \mathbb{R}^3 \to \mathbb{R}^2$ mit $f(x,y,z) = (3x^2+2xy, z^2y-z)$ sind die Einträge der Jacobimatrix

$$Jf(x,y,z) = \begin{pmatrix} 6x+2y & 2x & 0 \\ 0 & z^2 & 2zy-1 \end{pmatrix}$$

stetig. Damit ist diese Funktion auf ganz $\mathbb{R}^3$ stetig differenzierbar und Jf stellt tatsächlich die Ableitung dar.

8.4 Analytische Eigenschaften differenzierbarer Abbildungen

Seien $(E, \|\|_E)$ und $(F, \|\|_F)$ normierte Vektorräume, $U \subseteq E$ offen und $u \in U$

Lemma 8.21 *Ist $f : U \to F$ in u differenzierbar, dann ist f in u stetig.*

Beweis: Es gilt

$$\lim_{v \to u} f(v) = \lim_{v \to u} (f(u) + df(u)(v-u) + R(v)) = f(u) \ ,$$

da $df(u)$ linear und stetig ist und $R(v) \to 0$ für $v \to u$ gilt. ■

Aus der Existenz aller Richtungsableitungen in u folgt die Stetigkeit nicht. Dazu betrachten wir die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, welche überall verschwindet außer in den Punkten

$$x_n := \frac{1}{n}(\sin(\frac{1}{n}), \cos(\frac{1}{n})) \ ,$$

wo ihr Wert gleich eins sei. Da $x_n \to 0$ gilt, ist f im Punkt 0 nicht stetig. Andererseits ist für jedes $(a,b) \in \mathbb{R}^2$ die Funktion $f(t(a,b))$ auf einem kleinen Intervall um 0 konstant gleich Null. Also gilt $d_{(a,b)}f(0) = 0$.

Im folgenden wollen wir Eigenschaften von Funktionen mehrerer Veränderlicher auf schon gezeigte Eigenschaften von Funktionen einer Veränderlichen zurückführen. Dazu brauchen wir Gebiete, in welchen je zwei Punkte durch eine Stecke verbunden werden können.

Definition 8.22 *Sei E ein reeller Vektorraum. Eine Teilmenge $U \subseteq E$ heißt **konvex**, wenn für je zwei Punkte $x, y \in U$ auch $\lambda x + (1-\lambda)y \in U$ für alle $\lambda \in [0,1]$ gilt.*

Hier sind einige Beispiele:

1. E selbst ist konvex.

2. Einpunktige Mengen sind konvex.

3. Ist $\|.\|$ eine Norm auf E, dann ist der Ball $B(z,r)$ für jedes $z \in E$ und $r > 0$ konvex. Das folgt aus der Dreiecksungleichung. Sei $x, y \in B(z,r)$ und $\lambda \in [0,1]$. Dann gilt

$$\begin{aligned} \|\lambda x + (1-\lambda)y - z\| &= \|\lambda(x-z) + (1-\lambda)(y-z)\| \\ &\leq \|\lambda(x-z)\| + \|(1-\lambda)(y-z)\| \\ &= \lambda\|x-z\| + (1-\lambda)\|y-z\| \\ &< \lambda r + (1-\lambda)r \\ &= r \end{aligned}$$

4. Sind $U, V \subseteq E$ konvex, so auch $U \cap V$, nicht aber im allgemeinen $U \cup V$.

Lemma 8.23 (Mittelwertsatz) *Sei $U \subseteq E$ konvex, $f : U \to F$ auf U differenzierbar und*

$$M := \sup_{x \in U} \|f'(u)\|_{E \to F} < \infty \ .$$

Dann gilt für alle $x, y \in U$, daß $\|f(x) - f(y)\|_F \leq M\|x - y\|_E$.

Beweis: Wir betrachten $\gamma : [0,1] \to E$, $\gamma(t) = (1-t)x + ty$. Dann ist $\gamma^* f : [0,1] \to F$ definiert, $(\gamma^* f)'(t) = df(\gamma(t))(y - x)$, also $\|(\gamma^* f)'(t)\|_F \leq M\|x - y\|_E$. Daraus folgt $\|f(x) - f(y)\| = \|(\gamma^* f)(0) - (\gamma^* f)(1)\| \leq M\|x - y\|_E$. ■

Insbesondere ist eine Funktion mit verschwindender Ableitung auf einer konvexen Menge konstant.

8.5 Die Kettenregel, Summen

Wir betrachten drei normierte Räume $(E, \|.\|_E$, $(F, \|.\|_F)$ und $(G, \|.\|_G)$.

Lemma 8.24 *Die Komposition $\circ : B(F,G) \times B(E,F) \to B(E,G)$ ist wohldefiniert und stetig.*

Beweis: Sei $A \in B(E,F)$ und $B \in B(F,G)$. Dann ist

$$\|(B \circ A)(e)\|_E \leq \|B\|_{F \to G} \|Ae\|_F \leq \|B\|_{F \to G} \|A\|_{E \to F} \|e\|_E \ .$$

Folglich ist $B \circ A$ durch $\|B\|_{F \to G} \|A\|_{E \to F}$ beschränkt. Auf $B(F,G) \times B(E,F)$ betrachten wir die Metrik

$$d((A,B),(A',B')) := \|A - A'\|_{E \to F} + \|B - B'\|_{F \to G} \ .$$

Sei nun (A_i, B_i) ein Folge mit dem Grenzwert (A, B). Dann gilt $A_i \to A$ und $B_i \to B$. Wir benutzen

$$\begin{aligned} \|B_i \circ A_i - B \circ A\|_{E \to G} &= \|B_i \circ A_i - B_i \circ A + B_i \circ A - B \circ A\|_{E \to G} \\ &\leq \|B_i\|_{F \to G} \|A_i - A\|_{E \to F} + \|A\|_{E \to F} \|B_i - B\|_{F \to G} \ . \end{aligned}$$

Die Folge $\|B_i\|_{F \to G}$ ist beschränkt. Damit gilt $\|B_i \circ A_i - B \circ A\|_{E \to G} \to 0$. ■

Sei $U \subseteq E$ offen, $u \in U$, $f : U \to F$ in u differenzierbar mit der Ableitung $df(u) \in B(E,F)$. Sei weiter $V \subseteq F$ offen, $f(u) \in f(U) \subseteq V$ und $g : V \to G$ in $f(u)$ differenzierbar mit der Ableitung $dg(f(u))$.

Lemma 8.25 *Under diesen Umständen ist die Komposition $g \circ f : U \to G$ im Punkt u differenzierbar und es gilt*

$$d(g \circ f)(u) = dg(f(u)) \circ df(u) \ .$$

Beweis: Wir schreiben

$$f(v) = f(u) + df(u)(v-u) + R(v)\ , \quad g(w) = g(f(u)) - dg(f(u))(w - f(u)) + S(w)\ .$$

Dann gilt

$$\begin{aligned} g(f(v)) &= g(f(u)) - dg(f(u))(f(v) - f(u)) + S(f(v) \\ &= g(f(u)) - dg(f(u))(f(u) + df(u)(v-u) + R(v) - f(u)) + S(f(v)) \\ &= g(f(u)) + dg(f(u))df(u)(v-u) + [dg(f(u))R(v) + S(f(v))]\ . \end{aligned}$$

Nun gilt

$$\frac{\|dg(f(u))R(v) + S(f(v))\|_G}{\|v-u\|_E} \le \|dg(f(u)\|_{F\to G}\frac{\|R(v)\|_{E\to F}}{\|v-u\|_E} + \frac{\|S(f(v))\|_G}{\|v-u\|_E}$$

Nun gilt

$$\lim_{v\to u} \|dg(f(u)\|_{F\to G}\frac{\|R(v)\|_{E\to F}}{\|v-u\|_E} = 0\ .$$

Wir schreiben

$$\frac{\|S(f(v))\|_G}{\|v-u\|_E} = \begin{cases} \frac{\|S(f(v))\|_G}{\|f(v)-f(u)\|_F}\frac{\|f(v)-f(u)\|_F}{\|v-u\|_E} & f(v) \ne f(u) \\ 0 & f(v) = f(u) \end{cases}$$

Für $v \to u$ bleibt $\frac{\|f(v)-f(u)\|_F}{\|v-u\|_E}$ beschränkt, es gilt $f(v) \to f(u)$ und damit

$$\frac{\|S(f(v))\|_G}{\|v-u\|_E} \to 0\ .$$

Diese Abschätzungen zeigen, daß $g \circ f$ im Punkt u differenzierbar ist und die Ableitung $dg(f(u))df(u)$ hat. ■

Falls $E \cong \mathbb{R}^n$, $F \cong \mathbb{R}^m$ und $G \cong \mathbb{R}^k$ ist, dann werden $df(u)$ durch $Jf(u) \in \mathtt{Mat}(n,m)$ und $dg(f(u))$ durch $Jg(f(u)) \in \mathtt{Mat}(m,k)$ dargestellt. Die Komposition linearer Abbildungen wird durch das Matrixprodukt dargestellt. Folglich gilt

$$J(g \circ f)(u) = Jg(f(u))Jf(u) \in \mathtt{Mat}(n,k)\ .$$

Sei in unserem Beispiel $g : \mathbb{R}^2 \to \mathbb{R}^2$ durch $g(u,v) = (2u - v, v^2)$ gegeben. Dann ist

$$Jg(u,v) = \begin{pmatrix} 2 & -1 \\ 0 & 2v \end{pmatrix}\ .$$

Wir erhalten

$$\begin{aligned} Jg(f(u))Jf(u) &= \begin{pmatrix} 2 & -1 \\ 0 & 2v \end{pmatrix}\begin{pmatrix} 6x+2y & 2x & 0 \\ 0 & z^2 & 2zy-1 \end{pmatrix} \\ &= \begin{pmatrix} 12x+4y & 4x - z^2 & 1 - 2zy \\ 0 & 2vz^2 & 4vzy - 2v \end{pmatrix} \end{aligned}$$

wobei wir noch $v := z^2y - z$ einsetzen müssen. Die Komposition direkt ist

$$g \circ f(x, y, z) = (6x^2 + 4xy - z^2y + z, (z^2y - z)^2) \ .$$

Es gilt

$$J(g \circ f)(x, y, z) = \begin{pmatrix} 12x + 4y & 4x - z^2 & 1 - 2zy \\ 0 & 2vz^2 & 4vzy - 2v \end{pmatrix} \ .$$

Dieses Beispiel bestätigt die Kettenregel explizit.

Lemma 8.26 *Seien nun $f_0, f_1 : U \to F$ Abbildungen, welche im Punkt u differenzierbar sind. Dann ist $f_0 + f_1 : U \to F$ auch im Punkt u differenzierbar und es gilt $d(f_0 + f_1)(u) = df_0(u) + df_1(u)$.*

Beweis: Übungsaufgabe. ■

Die Elementarmatrizen $(E_{ij})_{i=1,\dots,n, j=1,\dots,m}$, bilden eine Basis von $\mathtt{Mat}(n, m)$. Hierbei ist E_{ij} die Matrix, deren einziger nichtverschwindender Eintrag eine 1 in der j-ten Spalte der i-ten Zeile ist.

Wir betrachten die Funktion $\det : \mathtt{Mat}(n, n) \to \mathbb{R}$. Zuerst bestimmen wir die Richtungsableitungen $d_{E_{ij}} \det(1)$. Für $i \neq j$ gilt $\det(1 + tE_{i,j}) \equiv 1$ und somit $d_{E_{i,j}} \det(1) = 0$. Ist $i = j$, dann ist $\det(1 + tE_{i,i}) = 1 + t$ und damit $d_{E_{i,i}} \det(1) = 1$. Wir schließen daraus, daß $d \det(1)(B) = \mathtt{Tr}(B)$ ist. Allgemeiner schreiben wir für eine invertierbare Matrix A

$$\det(A + B) = \det(A(1 + A^{-1}B) = \det(A) \det(1 + A^{-1}B) \ .$$

daraus schließen wir mit der Kettenregel, daß

$$d \det(A)(B) = \det(A)\mathtt{Tr}(A^{-1}B) \ .$$

8.6 Höhere Ableitungen

Von nun an werden wir die Indizierung der Normen weglassen, wenn durch den Eintrag klar wird, um welche Norm es sich handelt. Seien E, F normierte Räume, $U \subseteq E$ offen und $f : U \to F$. Wenn f in U differenzierbar ist, dann erhalten wir eine Abildung $df : U \to B(E, F)$, und können uns fragen, ob diese Abbildung wieder differenzierbar ist. Wenn ja, dann ist

$$d(df) : U \to B(E, B(E, F)) \ .$$

Setzen wir diese Gedanken fort, so erhalten wir den begriff einer n-ten Ableitung

$$d(\dots d(d(f) \dots) : U \to B(E, B(E, \dots, B(E, F) \dots) \ .$$

Sei $\xi_1 \in E$. Dann können wir $d_{\xi_1} f : U \to F$ bilden. Ist diese Abbildung wieder differenzierbar, dann können wir

$$d_{\xi_2}(d_{\xi_1} f) : U \to F$$

bilde, und so weiter.

Wenn f zweimal differenzierbar ist, dann gilt

$$d_{\xi_1} f(u) = df(u)(\xi) \; .$$

Als Übungsaufgabe möge man sich überlegen, daß

$$d_{\xi_2}(d_{\xi_1} f)(u) = d^2 f(u)(\xi_2)(\xi_1)$$

gilt.

Schließlich kann man versuchen, iterierte partielle Ableitungen zu bilden, also

$$\partial_i \partial_j f \;, \quad \partial_i \partial_j \partial_k f \;, \dots \; .$$

Insbesondere die Notation höherer Ableitungen wird etwas unübersichtlich, insbesondere was die Reihenfolge der Einsetzungen berifft. Glücklicherweise spielt diese in den meisten Fällen keine Rolle.

Lemma 8.27 *Seien $x, y \in E$ und existieren $d_x f, d_y f, d_x(d_y f)$ als stetige Abbildungen $U \to F$. Dann existiert auch $d_y(d_x f)$ und es gilt $d_x(d_y f) = d_y(d_x f)$.*

Beweis: Sei $u \in U$. Wir betrachten $F(s,t) := f(u + tx + sy)$ auf einer Umgebung von $(0,0) \subset \mathbb{R}^2$. Dann existieren $\partial_s F, \partial_t F, \partial_t \partial_s F$ und sind stetig. Wir setzen für geeignete $\xi, \in [0,s]$ und $\eta \in [0,t]$

$$\begin{aligned}
\Delta(s,t) &:= F(s,t) - F(0,t) - F(s,0) + F(0,0) \\
&= (F(s,t) - F(s,0)) - (F(0,t) - F(0,0)) \\
&= s(\partial_s F(\xi,t) - \partial_s F(\xi,0)) \\
&= st\partial_t \partial_s F(\xi,\eta)
\end{aligned}$$

Wir können aber auch für $\gamma \in [0,t]$

$$\begin{aligned}
\Delta(s,t) &:= F(s,t) - F(0,t) - F(s,0) + F(0,0) \\
&= (F(s,t) - F(0,t)) - (F(s,0) - F(0,0)) \\
&= t(\partial_t F(s,\gamma) - \partial_t F(0,\gamma))
\end{aligned}$$

schreiben. Wir wollen nun zeigen, daß

$$\lim_{s \to 0} \frac{\partial_t F(s,0) - \partial_t F(0,0)}{s} = \partial_t \partial_s F(0,0)$$

gilt. Sei $\epsilon > 0$ gegeben. Da $\partial_t \partial_s F$ stetig ist, finden wir ein $\delta > 0$ derart, daß für alle $s, t \in \mathbb{R}$ mit $|s| < \delta$, $|t| < \delta$ gilt

$$|\frac{\Delta(s,t)}{st} - \partial_t \partial_s F(0,0)| < \epsilon \; .$$

Wir rechnen

$$\begin{aligned}
|\frac{\Delta(s,t)}{st} - \partial_t\partial_s F(0,0)| &< \epsilon \\
\rightarrow \quad |\frac{\partial_t F(s,\gamma) - \partial_t F(0,\gamma))}{s} - \partial_t\partial_s F(0,0)| &< \epsilon \\
\rightarrow \quad \lim_{t\to 0} |\frac{\partial_t F(s,\gamma) - \partial_t F(0,\gamma))}{s} - \partial_t\partial_s F(0,0)| &\leq \epsilon \\
\rightarrow \quad |\frac{\partial_t F(s,0) - \partial_t F(0,0))}{s} - \partial_t\partial_s F(0,0)| &\leq \epsilon
\end{aligned}$$

Da $\epsilon > 0$ beliebig gewählt werden kann, gilt

$$\lim_{s\to 0} |\frac{\partial_t F(s,0) - \partial_t F(0,0))}{s} - \partial_t\partial_s F(0,0)| = 0 \ .$$

Folglich existiert $\partial_s\partial_t F(0,0)$ und es gilt $\partial_s\partial_t F(0,0) = \partial_t\partial_s F(0,0)$. ■

Ist f zweimal stetig differenzierbar, dann gilt folgende Symmetrie: $d^{(2)}f(\xi_1)(\xi_2) = d^{(2)}f(\xi_2)(\xi_1)$.

Korollar 8.28 *Ist $E \cong \mathbb{R}^n$ und f zweimal stetig partiell differenzierbar, dann gilt $\partial_i\partial_j f = \partial_j\partial_i f$. Allgemeiner, ist f mindestens k-mal stetig partiell differenzierbar, dann gilt für jede Permutation $\sigma \in \Sigma^k$, daß*

$$\partial_{i_1} \ldots \partial_{i_k} f = \partial_{i_{\sigma(1)}} \ldots \partial_{i_{\sigma(k)}} f \ .$$

Wir verwenden häufig folgende Notation:

1. Elemente $\alpha \in \mathbb{N}_0^n$ heißen Multiindizes (der Länge n). Wir schreiben $\alpha = (\alpha_1, \ldots, \alpha_n)$.
2. $|\alpha| := \sum_{i=1}^n \alpha_i$.
3. $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$
4. $\partial_\alpha f := f^{(\alpha)} := \partial_{\alpha_1} \ldots \partial_{\alpha_n} f$.

Sei $U \subseteq E$ eine Teilmenge eines Vektorraumes und $x \in U$.

Definition 8.29 *Die Menge U heißt sternförmig bezüglich u, wenn für jedes $x \in U$ auch $u + \lambda(x-u) \in U$ gilt für $\lambda \in [0,1]$.*

Konvexe Teilmengen sind sternförmig bezüglich jedes ihrer Punkte.

Satz 8.30 (Taylorformel) *Sei $U \subseteq \mathbb{R}^n$ offen sternförmig bezüglich 0 und $f : U \to \mathbb{R}$ $k+1$-mal differenzierbar. Dann existiert eine Abbildung $\eta : U \to [0,1]$ derart, daß*

$$f(x) = \sum_{\alpha\in\mathbb{N}^n, |\alpha|\leq k} \frac{f^{(\alpha)}(0)}{\alpha!} x^\alpha + \sum_{\alpha\in\mathbb{N}^n, |\alpha| = k+1} \frac{f^{(\alpha)}(\eta(x)x)}{\alpha!} x^\alpha.$$

Beweis: Sei $x \in U$. Dann existiert $\epsilon > 0$ derart, daß $\gamma : (-\epsilon, 1+\epsilon) \to \mathbb{R}^n$, $\gamma(\lambda) = \lambda x$ Werte in U hat. Die Funktion $F := \gamma^* F$ ist $k+1$-mal differenzierbar auf $(-\epsilon, 1+\epsilon)$. Die Taylorformel für F gibt

$$F(1) = \sum_{j \le k} \frac{F^{(j)}(0)}{j!} + \frac{f^{(k+1)}(\eta(x))}{(k+1)!}$$

für ein geeignetes $\eta_x \in [0,1]$. Wir berechnen nun induktiv

$$\frac{F^{(j)}(\lambda)}{j!} = \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j} \frac{f^{(\alpha)}(\lambda x)}{\alpha!} x^\alpha \ .$$

In der Tat gilt

$$F^{(0)}(\lambda) = f^{(0)}(\lambda x) \ .$$

Wenn wir die Formel schon bis $j-1$ gezeigt haben, dann gilt

$$\begin{aligned}
\frac{F^{(j)}(\lambda)}{j!} &= \frac{1}{j} \partial_\lambda \frac{F^{(j-1)}(\lambda)}{(j-1)!} \\
&= \frac{1}{j} \partial_\lambda \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j-1} \frac{f^{(\alpha)}(\lambda x)}{\alpha!} x^\alpha \\
&= \frac{1}{j} \sum_{i=1}^n \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j-1} \frac{\partial_i f^{(\alpha)}(\lambda x)}{\alpha!} x^\alpha x_i \\
&= \frac{1}{j} \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j} f^{(\alpha)}(\lambda x) x^\alpha \sum_{i=1}^n \frac{1}{\alpha_1! \dots (\alpha_i - 1)! \dots \alpha_n!} \\
&= \frac{1}{j} \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j} \frac{f^{(\alpha)}(\lambda x)}{\alpha!} x^\alpha \sum_{i=1}^n \alpha_i \\
&= \sum_{\alpha \in \mathbb{N}^n, |\alpha| = j} \frac{f^{(\alpha)}(\lambda x)}{\alpha!} x^\alpha \ .
\end{aligned}$$

■

8.7 Extremwerte

Sei $f : X \to \mathbb{R}$ eine reell-wertige Funktion auf einem topologischen Raum X. Die Funktion hat im Punkt $x \in X$ ein **lokales Maximum (Minimum)**, wenn es eine Umgebung $U \subseteq X$ von x gibt, so daß $f(x) \ge f(u)$ für alle $u \in U$ bzw. $f(x) \le f(u)$ für alle $u \in U$ gilt. Wir sprechen von einem lokalen Extremum, wenn f in x ein lokales Maximum oder lokales Minimum hat. Sei $g : Y \to X$ stetig, $y \in Y$ und $g(y) = x$. Wenn f im Punkt

x ein lokales Extremum hat, dann hat g^*f im Punkt y ein lokales Extremum. Ist U eine Umgebung von x derart, daß $f(u) \geq f(x)$ für alle $u \in U$, so ist $V := g^{-1}(U)$ eine Umgebung von y derart, daß $g^*f(v) \geq g^*f(y)$ für alle $v \in V$.

Sei nun $(E, \|.\|)$ ein normierter Raum und $U \subseteq E$ offen.

Lemma 8.31 *Sei $f : U \to \mathbb{R}$ in $u \in U$ differenzierbar und habe dort ein lokales Extremum. Dann gilt $df(u) = 0$.*

Beweis: Sei $x \in E$. Dann wählen wir $\epsilon > 0$ derart, daß $\gamma : (-\epsilon, \epsilon) \to E$, $\gamma(\lambda) = u + \lambda x$ Werte in U hat. Die Funktion $\gamma^* f$ hat in x ein lokales Extremum. Damit ist $(\gamma^* f)'(0) = df(u)(x) = 0$. Da $x \in E$ beliebig war, gilt $df(u) = 0$. ■

Sei $f : U \to \mathbb{R}$ differenzierbar.

Definition 8.32 *Ein Punkt $u \in U$ heißt **kritischer Punkt** von f, wenn $df(u) = 0$ ist.*

Wir erklären jetzt, wie man mit Hilfe der zweiten Ableitungen entscheiden kann, ob in einem kritischen Punkt ein lokales Maximum oder lokales Minimum vorliegt.

Sei $f : U \to \mathbb{R}$ zwei mal stetig differenzierbar.

Definition 8.33 *Die Abbildung $E \ni \xi \mapsto \mathtt{Hess}_u(f)(\xi) := d^{(2)} f(u)(\xi)(\xi)$ heißt **Hessesche Form** von f im Punkt u.*

Die Taylorformel für f auf einem Ball $B(u, \epsilon)$ kann in der Form

$$f(v) = f(u) + df(u)(v - u) + \frac{1}{2}\mathtt{Hess}_{u+\eta(v)(v-u)}(f)(v - u)$$

geschrieben werden.

Zur Erinnerung aus der Algebra, eine Abbildung $Q : E \to \mathbb{R}$ ist eine **quadratische Form**, wenn es eine bilineare Abbildung $\tilde{Q} : E \times E \to \mathbb{R}$ gibt mit $Q(\xi) = \tilde{Q}(\xi, \xi)$. Die Hessesche Form $\mathtt{Hess}_u(f)$ ist also eine quadratische Form.

Eine quadratische Form Q ist **positiv (negativ) definit** , wenn $Q(\xi) > 0$ ($Q(\xi) < 0$) gilt für alle $\xi \neq 0$. Läßt man das Gleichheitszeichen zu, daß spricht man von Semidefinitheit.

Sei $Q : \mathbb{R}^n \to \mathbb{R}$ eine quadratische Form auf einem endlich-dimensionalen Vektorraum. Mit $S^{n-1} := \{\xi \in \mathbb{R}^n \mid \|\xi\|_2 = 1\}$ bezeichnen wir die Einheitskugel in der euklidischen Norm.

Lemma 8.34 *Wenn Q positiv definit ist, dann existiert eine Konstante $C > 0$ derart, daß $\inf_{\xi \in S^{n-1}} Q(\xi) > 0$.*

Beweis: Eine quadratische Form ist stetig. In der Tat ist eine bilineare Abbildung $\tilde{Q} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ stetig, da sie in der Form $\tilde{Q}(\eta, \xi) = \eta^t A \xi$ für eine geeignete Matrix $A \in \mathtt{Mat}(n, n)$ gegeben ist. Damit ist auch $Q(\xi) = \xi^t A \xi$ stetig.

Die Einheitskugel $S^{n-1} = \{\xi \in \mathbb{R}^n \mid \|\xi\|_2 = 1\}$ ist abgeschlossen und beschränkt, damit kompakt. Wenn $\inf_{\xi \in S^{n-1}} Q(\xi) = 0$, dann existiert ein $\xi_0 \in S^{n-1}$ mit $Q(\xi_0) = 0$, was im Widerspruch zur positiven Definitheit von Q steht. ■

Wir nehmen jetzt an, daß $U \subset \mathbb{R}^n$ offen ist.

Satz 8.35 *Sei $f : U \to \mathbb{R}$ zwei mal stetig differenzierbar, $u \in U$ ein kritischer Punkt und $\mathtt{Hess}_u(f)$ positiv (negativ) definit. Dann hat f im Punkt u ein lokales Minimum (Maximum). Ist $\mathtt{Hess}_u(f)$ indefinit, dann hat f im Punkt u kein Extremum.*

Beweis: Die Taylorformel für f auf einem Ball um u kann in der Form

$$f(v) = f(u) + df(u)(v-u) + \frac{1}{2}\mathtt{Hess}_{u+\eta(v)(v-u)}(f)(v-u)$$

für geeignetes $\eta(v) \in [0,1]$ geschrieben werden. In der Tat ist mit $u_v := u + \eta(v)(v-u)$

$$\begin{aligned}
\frac{1}{2}\mathtt{Hess}_{u+\eta(v)(v-u)}(f)(\xi) &= \frac{1}{2}\sum_{i,j=1}^{n} \partial_i\partial_j f(u_v)\xi_i\xi_j \\
&= \frac{1}{2}\sum_{i}^{n} \partial_i\partial_j f(u_v)\xi_i^2 + \sum_{i<j} \partial_i\partial_j f(u_v)\xi_i\xi_j \\
&= \sum_{\alpha\in\mathbb{N}^n, |\alpha|=2} \frac{f^{(\alpha)}(u_v)}{\alpha!}\xi^\alpha
\end{aligned}$$

Wir nehmen an, daß $\mathtt{Hess}_u(f)$ positiv definit ist. Wir definieren

$$h(x) := \inf_{\xi\in S^{n-1}} \mathtt{Hess}_x(f)(\xi) \ .$$

Diese Funktion ist stetig. Um dies einzusehen, benutzen wir das folgende Lemma.

Lemma 8.36 *Seien X, Y topologische Räume, Y kompakt und $f : X \times Y \to \mathbb{R}$ stetig. Dann ist $X \ni x \mapsto h(x) := \inf_{y\in Y} f(x)$ stetig.*

Beweis: Die Abbildung $\inf_Y : C(Y) \to \mathbb{R}$ ist stetig. In der Tat, wenn $\|g' - g\|_\infty < \epsilon$ ist, dann gilt $|\inf_Y g' - \inf_Y g| \leq \epsilon$.

Die Funktion f bestimmt eine Abbildung $F : X \to C(Y)$ durch $F(x)(y) := f(x,y)$ und es gilt $h(x) = \inf_Y \circ F$. Es reicht also aus zu zeigen, daß F stetig ist. Wir zeigen die Stetigkeit in $x \in X$.

Sei $\epsilon > 0$ vorgegeben Da f stetig ist, existiert für jedes $y \in Y$ eine Umgebung $U_y \times V_y \subseteq X \times Y$ von (x,y) derart, daß für alle $(x',y') \in U_y \times V_y$ gilt $|f(x,y) - f(x',y')| < \epsilon$ gilt. Hierbei sind $U_x \subseteq X$ und $V_y \subseteq Y$ offen. Da Y kompakt ist, kann man eine endliche Folge $y_1, \ldots, y_n$ in Y finden, so daß $\cup_{i=1}^n V_{y_i} = Y$. Dann ist $U := \cap_{i=1}^n U_{y_i}$ eine offene Umgebung von x in X. Es gilt für alle $z \in U$, daß $\|F(z) - F(x)\|_\infty < \epsilon$. ■

Es gilt $h(u) > 0$. Wir wählen $\epsilon > 0$ derart, daß $B(u,\epsilon) \subseteq U$ und $h_{|B(u,\epsilon)} > 0$ gilt. Mit der Taylorformel erhalten wir für $v \in B(u,\epsilon)$

$$f(v) = f(u) + \frac{1}{2}\mathtt{Hess}_{u+\eta(v)(v-u)}(v-u) > f(u) \ .$$

Also hat f im Punkt u ein lokales Minimum.

Wenn $\mathtt{Hess}_u(f)$ indefinit ist, dann kann man mit analogen Argumenten $v_\pm$ beliebig nahe an u finden, so daß $\pm\mathtt{Hess}_{u+\eta(v_\pm)(v_\pm-u)}(f)(v_\pm - u) > 0$. Also hat dann f in u kein lokales Extremum. ■

Die Matrix

$$\mathtt{Hess}_x(f)_{i,j} := \partial_i\partial_f f(u)$$

heißt die Hessische Matrix. Sie stellt die Hessische Form dar durch

$$\mathtt{Hess}_x(f)(\xi) = \sum_{i,j=1}^n \mathtt{Hess}_x(f)_{i,j}\xi_i\xi_j \ .$$

Hier ist ein explizites Beispiel:

$$f : \mathbb{R}^2 \to \mathbb{R} \ , f(x,y) = x^3 + y^3 - 12x - 3y \ .$$

Dann gilt

$$\partial_x f(x,y) = 3x^2 - 12 \ , \quad \partial_y f(x,y) = 3y^2 - 3 \ .$$

Die folgenden Punkte sind kritisch:

$$(2,1), (-2,1), (2,-1), (-2,-1) \ .$$

Die Hessische Matrix ist durch

$$\mathtt{Hess}_{(x,y)} = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$

gegeben. Die Diskussion der kritischen Punkte gibt:

$$\begin{array}{ccc}
(2,1) & \begin{pmatrix} 12 & 0 \\ 0 & 6 \end{pmatrix} & \text{Minumum} \\
(-2,1) & \begin{pmatrix} -12 & 0 \\ 0 & 6 \end{pmatrix} & \text{kein Extremum} \\
(2,-1) & \begin{pmatrix} 12 & 0 \\ 0 & -6 \end{pmatrix} & \text{kein Extremum} \\
(-2,-1) & \begin{pmatrix} -12 & 0 \\ 0 & -6 \end{pmatrix} & \text{Maximum} \ .
\end{array}$$

8.8 Kontraktionen und Fixpunkte, Satz über implizite Funktionen

Sei (X, d) ein metrischer Raum, $U \subseteq X$ und $f : U \to X$.

Definition 8.37 *f heißt* ***Kontraktion*** *wenn es eine Konstante $c \in [0, 1)$ gibt, so daß $d(f(x), f(y)) \leq cd(x, y)$ für alle $x, y \in X$ gilt.*

Hier sind einige Beispiele.

1. Eine konstante Abbildung ist eine Kontraktion mit $c = 0$.

2. Die Identität ist genau dann eine Kontraktion, wenn X genau aus höchstens einem Punkt besteht.

3. Sei $(E, \|.\|)$ ein normierter Raum, $A \in B(E, E)$, $e_0 \in E$ und $\|A\| < 1$. Dann ist die Abbildung $f(e) := e_0 + Ae$ eine Kontraktion mit Konstante $\|A\|$. In der Tat ist

$$d(f(e), f(e')) = \|A(e - e')\| \leq \|A\|\|e - e'\| = \|A\|d(e, e') \ .$$

Definition 8.38 *Ein Punkt $u \in U$ heißt* ***Fixpunkt*** *von f, wenn $f(u) = u$ gilt.*

Lemma 8.39 *Eine Kontraktion hat höchstens einen Fixpunkt.*

Beweis: Sei c wie in 8.37. Seien $x, y \in U$ Fixpunkte. Dann gilt $d(x, y) = d(f(x), f(y)) \leq cd(x, y)$. Diese Gleichung kann ber nur für $d(x, y) = 0$, also $x = y$ gelten. ■

Lemma 8.40 *Sei f eine Kontraktion. Wenn X vollständig und $\overline{f(U)} \subseteq U$ gilt, so besitzt f einen Fixpunkt.*

Beweis: Wir wählen $x_0 \in U$ und setzen $x_i := f^i(x_0)$. In der Tat ist wegen der Voraussetzung an f auch $x_i \in U$.

Dann ist (x_i) eine Cauchyfolge in X. In der Tat ist

$$d(x_i, x_{i+1}) = d(f(x_{i-1}), f(x_i)) \leq cd(x_{i-1}, x_i) \ .$$

Durch Iteration dieser Ungleichung bekommen wir $d(x_i, x_{i+1}) \leq c^i d(x_0, x_1)$. Die geometrische Reihe $\sum_{i=0}^{\infty} c^i d(x_0, x_1)$ mit $c \in [0, 1)$ ist eine konvergente Majorante von $\sum_{i=0}^{\infty} d(x_i, x_{i+1})$. Sei $\epsilon > 0$ gegeben. Dann wählen wir N so groß, daß $\sum_{i=n+1}^{m} d(x_i, x_{i+1}) < \epsilon$ gilt für alle $N \leq n \leq m$. Nun ist jedoch für $N \leq n \leq m$ nach der Dreicksungleichung

$$d(x_n, x_m) \leq \sum_{i=n+1}^{m} d(x_i, x_{i+1}) \leq \epsilon \ .$$

Da X vollständig ist, gibt es den Grenzwert $\lim_{n\to\infty} x_n = x$. Da alle $x_n \in f(U)$ liegen, gilt $x \in \overline{f(U)} \subseteq U$. Nun ist

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x \ .$$

Also ist x ein Fixpunkt von f. ■

Seien E, F Banachräume, $U \subseteq E$ offen, $f : U \to F$ stetig differenzierbar, $u \in U$.

Satz 8.41 (Satz über die Umkehrfunktion) *Ist $df(u) : E \to F$ beschränkt invertierbar, dann existieren Umgebungen $x \in V \subseteq U$ und $f(x) \in W \subseteq F$ und eine stetig differenzierbare Abbildung $g : W \to V$ derart, daß $g \circ f = \mathtt{id}_V$ und $f \circ g = \mathtt{id}_W$ gilt.*

Beweis: Der Beweis besteht aus folgenden Schritten bzw. Nachweisen.

1. Konstruktion von V.
2. $f_{|V}$ ist injektiv.
3. $W := f(V)$ ist offen, $g := (f_{|V})^{-1}$
4. g ist stetig differenzierbar.

1. Sei $A := df(u)$ und $\lambda := \frac{1}{2\|A^{-1}\|}$. Wir wählen $\epsilon > 0$ so klein, daß aus $\|x - u\| < \epsilon$ folgt $\|df(u) - df(x)\| < \lambda$. Wir benutzen hier, daß $u \mapsto df(u)$ stetig ist. Wir setzen

$$V := B(u, \epsilon) \ .$$

2. Für $y \in F$ definieren wir $\phi_y : V \to E$ durch

$$\phi_y(x) := x + A^{-1}(y - f(x)) \ .$$

Es gilt $d\phi_y(x) = 1 - A^{-1}df(x)$ und deshalb für $x \in V$

$$\|d\phi_y(x)\| = \|1 - A^{-1}df(x)\| = \|A^{-1}(A - df(x))\| \leq \|A^{-1}\|\|A - df(x)\| \leq \frac{1}{2} \ .$$

Der Ball V ist konvex. Nach dem Mittelwertsatz ist

$$\|\phi_y(x) - \phi_y(x')\| \leq \frac{1}{2}\|x - x'\| \ .$$

Folglich ist ϕ_y eine Kontraktion mit Konstante $\frac{1}{2}$. Diese Kontraktion hat höchstens einen Fixpunkt in V. Seien $x, x' \in V$ mit $f(x) = f(x') = y$. Damit ist $\phi_y(x) = x$ und $\phi_y(x') = x'$. Folglich ist $x = x'$.

3. Sei $W := f(V)$ und $y \in W$, $y = f(x)$ für ein $x \in V$. Wir wählen $\delta > 0$ derart, daß $B(x,\delta) \subseteq V$ (geht, da V offen ist) und zeigen, daß $B(y,\frac{\lambda\delta}{2}) \subseteq W$ gilt. Sei $z \in B(y,\frac{\lambda\delta}{2})$. Wir müssen ein $w \in V$ mit $z = f(w)$ finden.

   Zuersteinmal gilt
$$\begin{aligned} \|\phi_z(x) - x\| &= \|A^{-1}(z-y)\| \\ &\leq \|A^{-1}\|\|z-y\| \\ &= \frac{1}{2\lambda}\|z-y\| \\ &\leq \frac{\delta}{4} . \end{aligned}$$
   und für $v \in B(x,\delta)$, daß
$$\begin{aligned} \|\phi_z(v) - x\| &\leq \|\phi_z(v) - \phi_z(x)\| + \|\phi_z(x) - x\| \\ &\leq \frac{1}{2}\|v-x\| + \frac{\delta}{4} \\ &\leq \frac{3\delta}{4} , \end{aligned}$$
   also $\phi_z(B(x,\delta)) \subseteq B(x,\frac{3\delta}{4})$ und $\overline{\phi_z(B(x,\delta))} \subseteq B(x,\delta)$. Die Kontraktion ϕ_z besitzt damit einen Fixpunkt $w \in B(x,\delta) \subseteq V$, d.h. es gilt $w + A^{-1}(z - f(w)) = w$, also $z = f(w)$.

   Wir definieren jetzt
$$g := (f_{|V})^{-1} : W \to V .$$

4. Sei $x \in V$ und $f(x) = y \in W$. Dann ist $\|df(x) - A\| \leq \lambda < \frac{1}{\|A^{-1}\|}$. Folglich existiert $df(x)^{-1}$ und $V \ni x \mapsto df(x)^{-1} \in B(F,E)$ ist stetig. Wir schreiben
$$g(z) = g(y) + df(x)^{-1}(z-y) + S(z) .$$
   Wir müssen zeigen, daß
$$\lim_{z\to y} \frac{S(z)}{\|z-y\|} = 0$$
   gilt. Wir setzen $v := g(z)$ und betrachten
$$\phi_y(v) - \phi_y(x) = v - A^{-1}(y - f(v))x - x + A^{-1}(y - f(x)) = v - x + A^{-1}(y-z) ,$$
   also $A^{-1}(y-z) = x - v + \phi_y(v) - \phi_y(x)$. Daraus folgt die Abschätzung
$$\begin{aligned} \|A^{-1}\|\|y-z\| &\geq \|x-v\| - \|\phi_y(v) - \phi_y(x)\| \\ &\geq \|x-v\| - \frac{1}{2}\|x-v\| \\ &= \frac{1}{2}\|x-v\| . \end{aligned}$$

Also gilt
$$\|y-z\| \geq \frac{1}{2\|A^{-1}\|}\|x-v\| = \lambda\|x-v\| \ . \tag{2}$$
Daraus schließen wir unter Verwendung von
$$z-y = f(v)-f(x) = df(x)(v-x)+R(v) \ ,$$
daß
$$\begin{aligned}
\frac{S(z)}{\|z-y\|} &\leq \frac{\|v-x-df(x)^{-1}(z-y)\|}{\lambda\|x-v\|} \\
&= \frac{\|v-x-df(x)^{-1}(df(x)(v-x)+R(v))\|}{\lambda\|x-v\|} \\
&= \frac{\|df(x)^{-1}(R(v))\|}{\lambda\|x-v\|} \\
&\leq \frac{\|df(x)^{-1}\|}{\lambda}\frac{\|R(v)\|}{\|x-v\|}
\end{aligned}$$
Wegen (2) gilt für $z \to y$ auch $v \to x$. Da nach Voraussetzung $\lim_{v\to x}\frac{\|R(v)\|}{\|x-v\|} = 0$ gilt, folgt $\lim_{z\to y}\frac{S(z)}{\|z-y\|} = 0$. Damit ist g im Punkt y differenzierbar und $dg(y) = df(x)^{-1}$. Insbesondere ist g in diesem Punkt stetig (das folgt auch unmittelbar aus (2)). Die Abbildung $y \to dg(y) = df(g(x))^{-1}$ ist als Komposition stetiger Abbildungen stetig.

■

Wir kommen nun zum Satz über implizte Funktionen. Wir wollen eine Gleichung der Form $F(x) = 0$ lösen, die von einem Parameter λ abhängt. Wir schreiben die Abhängigkeit als weitere Variable und erhalten $F(x,\lambda) = 0$. In guten Fällen wird die Lösung x eine Funktion des Parameters sein, daß heißt, wir suchen eine Funktion $x(\lambda)$ derart, daß $F(x(\lambda),\lambda) = 0$ für alle λ gilt. Der Satz über implizite Funktionen ist ein technisches Hilfmittel der Differentialrechnung in dieser Frage.

Seien E, F, G Banachräume. Die Menge $E\times F$ hat eine induzierte Banachraumstruktur. Der unterliegende Vektorraum ist $E\oplus F$, und als Norm kann man zum Beispiel eine der beiden äquivalenten Normen
$$\|(e,f)\| := \|e\|+\|f\| \ , \quad \text{oder} \quad \|(e,f)\| := \mathtt{max}\{\|e\|,\|f\|\}$$
nehmen. Der unterliegende topologische Raum des Banachraumes $E\times F$ ist das Produkt der topologischen Räume E und F.

Sei $U \subseteq E\times F$ offen, $p : U \to G$ stetig differenzierbar, $(e,f)\in U$ und $p(e,f) = 0$. Für $y\in F$ sei $i_y : E \to E\times F$ durch $i_y(x) := (x,y)$ gegeben. Diese Abbildung ist stetig. Sei $V := i_f^{-1}(U) \subseteq E$. Diese Menge ist offen. Wir setzen
$$d_1p(x,y) := d(i_y^*p)(x) \in B(E,G) \ .$$

Analog defineren wir für $x \in E$ die Abbildung $j_x : F \to E \times F$, $j_x(y) := (x, y)$ und setzen

$$d_2 p(x,y) := d(j_x^* p)(y) \in B(F,G) \ .$$

Satz 8.42 (Satz über implizite Funktionen) *Wenn $d_1 p(e,f)$ invertierbar ist, dann gibt es offene Umgebungen $e \in \tilde{V} \subseteq V$ und $f \in W \subseteq F$ derart, daß für jedes $y \in W$ genau ein $x \in \tilde{V}$ existiert mit $p(x,y) = 0$. Die so durch $p(g(y),y) \equiv 0$ eindeutig bestimmte Abbildung $g : W \to \tilde{V}$ ist in f stetig differenzierbar und es gilt $dg(e) = -d_1 p(e,f)^{-1} \circ d_2 p(x,y)$.*

Beweis: Sei $\Phi : U \to G \times F$ durch $\Phi(x,y) := (p(x,y), y)$ gegeben. Diese Funktion ist stetig differenzierbar und es gilt

$$d\Phi(x,y) = \begin{pmatrix} d_1 p(x,y) & d_2 p(x,y) \\ 0 & 1 \end{pmatrix} : \begin{pmatrix} E \\ F \end{pmatrix} \to \begin{pmatrix} G \\ F \end{pmatrix} \ .$$

Offensichtlich ist $d\Phi(e,f)$ invertierbar und es gilt

$$d\Phi(e,f)^{-1} = \begin{pmatrix} d_1 p(e,f)^{-1} & -d_1 p(e,f)^{-1} \circ d_2 p(e,f) \\ 0 & 1 \end{pmatrix} \ .$$

Also gibt es nach dem Satz über Umkehrfunktionen Umgebungen $(e,f) \in \tilde{U} \subseteq U$ und $(0,f) \in \tilde{W} \subseteq G \times F$ und eine Umkehrfunktion $\Psi : \tilde{W} \to \tilde{U}$ von $\Phi_{|\tilde{U}}$. Wir schreiben $\Psi = (\Psi_1, \Psi_2)$, $\Psi_1 : \tilde{W} \to E$ und $\Psi_2 : \tilde{W} \to F$. Es gilt

$$(z,y) = (\Phi \circ \Psi)(z,y) = (p(\Psi_1(z,y), \Psi_2(z,y)), \Psi_2(z,y)) \ ,$$

also

$$p(\Psi_1(z,y), \Psi_2(z,y)) = z \ , \quad \Psi_2(z,y) = y \ ,$$

woraus $p(\Psi_1(z,y), y) = z$ folgt.

Wir wählen nun Umgebungen $e \in \tilde{V} \subseteq E$ und $f \in W \subseteq F$ derart, daß $\tilde{V} \times W \subseteq \tilde{U}$. Im folgenden werden wir W noch verkleinern. Wir ersetzen W durch $W \cap h^{-1}(\tilde{W})$, wobei $h : F \to G \times F$ durch $h(y) = (0, y)$ gegeben wird. Beachte, daß $h(0,f) \in W$, so daß W wirklich eine Umgebung von f ist. Durch diese Wahl ist

$$g : W \to E \ , \quad g(y) := \Psi_1(0,y)$$

definiert. Diese Abbildung ist stetig und erfüllt $g(f) = e$. In der Tat ist $\Phi(e,f) = (p(e,f), f) = (0,f)$ und damit $\Psi_1(0,f) = (e,f)$, also $g(f) = \Psi_1(0,f) = e$. Wir ersetzen nun W durch $g^{-1}(\tilde{V})$. Durch diese Wahl erreichen wir, daß $g : W \to \tilde{V}$.

Sei jetzt $y \in W$ und $x \in \tilde{V}$ derart, daß $p(x,y) = 0$. Wegen $(x,y) \in \tilde{V} \times W \subseteq \tilde{W}$ gilt $\Phi(x,y) = (0,y)$ und deshalb $(x,y) = \Psi(0,y) = (g(x), y)$. Folglich ist $x = g(y)$ die einzige Lösung von $p(x,y) = 0$ mit $(x,y) \in \tilde{V} \times W$.

Die Abbildung g ist in f stetig differenzierbar und es gilt

$$dg(f) = d_2 \Psi_1(0,f) = -d_1 p(e,f)^{-1} \circ d_2 p(e,f) \ .$$

■

8.9 Untermannigfaltigkeiten

Seien E, F Banachräume, $U \subseteq E$ offen und $f : U \to F$ stetig differenzierbar.

Definition 8.43 *Die Funktion f ist in $u \in U$* ***regulär****, wenn $df(u) : E \to F$ surjektiv ist.*

Eine Funktion $U \to \mathbb{R}$ ist in $u \in U$ genau dann regulär, wenn u nicht kritisch ist. Wenn E endlich dimensional und $\dim(E) < \dim(F)$ gilt, dann ist f in keinem Punkt regulär.

Ab jetzt sei E endlich dimensional, $n := \dim(E)$.

Definition 8.44 *Eine Teilmenge $M \subseteq E$ heißt* ***Untermannigfaltigkeit der Kodimension*** *k, wenn es für jeden Punkt $m \in M$ eine offene Umgebung $U \subseteq E$ und eine Funktion $f : U \to \mathbb{R}^k$ gibt, so daß $\{f = 0\} = M \cap U$ gilt und f in allen Punkten von m regulär ist. Wir nennen so ein Paar (U, f) eine lokale* ***definierende Funktion*** *von M bei m.*

1. Eine offene Teilmenge $M \subseteq \mathbb{R}^n$ ist eine Untermannigfaltigkeit der Kodimension 0. Das Paar $(M, 0)$ ist eine definierende Funktion.

2. Der Punkt $M := \{0\} \subseteq \mathbb{R}^n$ ist eine Untermannigfaltigkeit der Kodimension n. Das Paar $(\mathbb{R}^n, \mathtt{id})$ ist eine definierende Funktion.

3. Die Einheitskugel $S^{n-1} \subset \mathbb{R}^n$ ist eine Untermannifgaltigkeit der Kodimension 1. Das Paar $(\mathbb{R}^n, x \mapsto f(x) := \|x\|^2)$ ist eine definieren Funktion. In der Tat ist $df(x)(\xi) = 2 < x, \xi >$, also $df(x) = 0$ genau für $x = 0$. Da dieser Punkt nicht in S^{n-1} liegt, ist f in allen Punkten von S^{n-1} regulär. Da Paar $(\mathbb{R}^n \setminus \{0\}, \|x\|)$ wäre eine andere Wahl einer definieren Funktion von S^{n-1}.

4. Der Schnitt A der E**inheitskugel** S^{n-1} mit der affinen Hyperebene $\{\sum_{i=1}^n x_i = \frac{1}{2}\}$ ist eine Untermannigfaltigkeit der Kodimension 2. Das Paar

$$(\mathbb{R}^n, f(x) = (\|x\|^2, \sum_{i=1}^n x_i - \frac{1}{2}))$$

ist eine definierende Funktion. Dazu berechnen wir

$$df(x)(\xi) = (2 < x, \xi >, \sum_{i=1}^n \xi_i) \ .$$

Es gilt $\dim \mathtt{im}(df(x)) < 2$ genau dann, wenn $x \sim (1, \ldots, 1)^t$ gilt. Die einzigen solchen Punkte mit $\|x\| = 1$ wären $x = \pm \frac{1}{\sqrt{n}}(1, \ldots, 1)^t$. Dann gilt aber $\sum_{i=1}^n x_i = \pm\sqrt{n} \neq \frac{1}{2}$, also sind diese Punkte nicht in der Hyperebene und damit nicht in A.

5. Sei $U \subset E$ offen und $f : U \to F$ eine stetig-differenzierbare Abbildung in einen k-dimensionalen Vektorraum. Dann ist der **Graph**

$$\Gamma(f) := \{(x, f(x)) \,|\, x \in U\} \subseteq E \oplus F$$

eine Untermannigfaltigkeit der Kodimension k. Wir setzen $V := \{(x,y) \,|\, x \in U\} \subseteq E \oplus F$, wählen eine Isomorphie $I : F \xrightarrow{\sim} \mathbb{R}^k$ und definieren $\phi : V \to \mathbb{R}^k$ durch $\phi(x,y) = I(y) - I(f(x))$. Es gilt $\Gamma(f) = \{\phi = 0\}$. Weiterhin ist $d\phi = (d_1\phi, d_2\phi)$ und $d_2\phi = I$, also ist $d\phi$ in allen Punkten von V surjektiv. Das Paar (V, ϕ) definiert $\Gamma(f)$.

6. Das Paar $(\mathbb{R}, x \mapsto x^3)$ ist nicht definierend für $\{0\} \subset \mathbb{R}$.

7. Die Teilmenge $\{0\} \cup \{1/n \,|\, n \in \mathbb{N}\}$ ist keine Untermannigfaltigkeit. In der Tat, wäre f bei 0 definierend, dann wäre $f'(0) = \lim_{n\to\infty} \frac{f(1/n)-f(0)}{1/n} = 0$, was im Widerspruch zur geforderten Regularität von f im Punkt 0 steht. Die Teilmenge $\{1/n \,|\, n \in \mathbb{N}\}$ ist eine Untermannigfaltigkeit. Das Paar $((0,\infty), f(x) := \sin(\frac{\pi}{x}))$ ist definierend.

8. Die Teilmenge $\{x^2 = y^3\} \subset \mathbb{R}^2$ ist keine Untermannigfaltigkeit. Jedenfalls ist $f(x,y) = x^2 - y^3$ nicht definierend, da nicht regulär im Punkt 0. Wäre g eine definierende Funktion bei 0, so wäre $g(\sqrt{y^3}, y) = 0$ für alle kleinen positiven y. Folglich, mit $x = \sqrt{y^3}$ auch $\partial_x g(x,y)\frac{3\sqrt{y}}{2} + \partial_y g(x,y) \equiv 0$. Daraus folgt $\partial_y g(0,0) = 0$.
Weiterhin wäre für kleine $x > 0$ auch $g(\pm x, x^{2/3}) \equiv 0$ woraus mit $y = x^{2/3}$ $\pm\partial_x g(\pm x, y) + \partial_y g(\pm x, y)\frac{2}{3x^{1/3}} = 0$ gelten würde. Diese beiden Gleichungen voneinander abgezogen liefern $\partial_x g(x,y) = 0$. Das Verschwinden beider partieller Ableitung steht im Widerspruch zur Regularität von g im Punkt 0.

Lemma 8.45 *Sei $M \subset E$ eine Untermannnigfaltigkeit und $0 \in M$. Dann gibt es eine Aufspaltung $E \cong T \oplus N$, offene Umgebungen $X \subseteq T$ und $Y \subseteq N$ von 0 und eine Abbildung $f : X \to Y$ derart, daß $M \cap X \times Y = \Gamma(f)$ ist.*

Beweis: Sei (U, ϕ) eine lokale definierende Funktion von M bei 0. Wir setzen $T := \ker(d\phi(0))$ und wählen ein Komplement $N \subseteq E$ derart, daß $E \cong T \oplus N$. Wir können nun ϕ als Funktion in zwei Variablen auffassen, $\phi(x) = \phi(t,n)$. Es gilt $d_1\phi(0) = 0$. Da ϕ in 0 regulär ist, muß also $d_2\phi(0)$ surjektiv sein. Aus Dimensionsgründen ist $d_2\phi(0)$ ein Isomorphismus. Nach dem Satz über implizite Funktionen existieren Umgebungen $X \subseteq T$, und $Y \subseteq N$ der Null und eine stetig differenzierbare Abbildung $f : X \to Y$ derart, daß $X \times Y \subseteq U$, und $\Gamma(f) = \{\phi = 0\} \cap X \times Y$ gilt. ■
Die in diesem Beweis konstruierte Abbildung g erfüllt noch die spezielle Bedingung $df(0) = -d_2\phi(0)^{-1}d_1\phi(0) = 0$. Dies werden wir weiter unten ausnutzen. Die Identität $df(y) = -d_2\phi(y)^{-1} \circ d_1\phi(y)$ für y nahe an 0 zeigt, daß aus der k-fachen stetigen Differenzierbarkeit der definierenden Funktion die k-fache stetige Differenzierbarkeit von f folgt.

Definition 8.46 *Eine Untermannigfaltigkeit ist von **der Klasse** C^k (**glatt** oder C^∞), wenn sie lokal durch k-fach stetig differenzierbare (glatte) Funktionen definiert werden kann.*

Eine Untermannigfaltigkeit der Klasse C^k kann also lokal als Graph einer C^k-Funktion geschrieben werden.

Sei $M \subseteq E$ eine Untermannigfaltigkeit der Kodimension k und $m_0 \in M$. Dann können wir M um m_0 verschieben: Die Teilmenge $M - m_0 := \{m - m_0 \mid m \in M\}$ ist wieder eine Untermannigfaltigkeit der Kodimension k. Ist (U, f) lokal definierend für M bei m_0, dann ist $(U - m_0, f'(x) := f(x + m_0))$ lokal definierend für $M - m_0$ bei 0.

In vielen Argumenten werden wir o.B.d.A annehmen, daß der untersuchte Punkt der Nullpunkt ist. Dies werden wir immer durch Verschieben erreichen.

Sei $M \subseteq E$ eine Untermannigfaltigkeit der Kodimension k und $m \in M$.

Lemma 8.47 *Sind (U, f) und (U', f') definierende Funktionen von M bei m, dann gilt* $\ker(df(m)) = \ker(df'(m))$.

Beweis: O.B.d.A können wir annehmen, daß $m = 0$ gilt. Wir verwenden (U, f) wie im Beweis von Lemma 8.45 und finden eine Aufspaltung $E \cong T \oplus N$ mit $T = \ker(df(0))$, $X \subseteq T$ und $Y \subseteq N$ offen und $g : X \to Y$ derart, daß $M \cap X \times Y = \Gamma(g)$ und $dg(0) = 0$. Es gilt $f'(t, g(t)) \equiv 0$, also wegen $dg(0) = 0$ auch $d_1 f'(0) = 0$. Das bedeutet $\ker(df(0)) \subseteq \ker(df'(0))$.

Durch Vertauschen der Rollen von f und f' erhält man die umgekehrte Inklusion. ■

Definition 8.48 *Sei $M \subseteq E$ eine Untermannigfaltigkeit der Kodimension k und $m \in M$. Der von der Wahl der definierenden Funktion (U, f) von M bei m unabhängige Raum* $T_m M := \ker(df(m))$ *heißt* ***Tangentialraum*** *von M an m.*

Ist $M = \Gamma(g)$ für $g : V \to N$, $V \subseteq T$, mit $dg(0) = 0$, dann ist $T \subseteq T \oplus N$ der Tangentialraum an M im Punkt 0.

Definition 8.49 *Zwei lineare Unterräume $V, W \subseteq E$ eines endlich dimensionalen Vektorraumes heißen* ***transversal*** *(wir schreiben $V \pitchfork W$), wenn* $\dim(V \cap W) = \dim(V) + \dim(W) - \dim(E)$ *gilt.*

Zwei Untermannigfaltigkeiten $M, N \subseteq E$ sind zueinander transversal, wenn für jeden Punkt $x \in M \cap N$ gilt $T_x M \pitchfork T_x N$.

Seien $M, N \subseteq E$ Untermannigfaltigkeiten der Kodimensionen m und n

Lemma 8.50 *Wenn $M \pitchfork N$, dann ist $M \cap N$ eine Untermannigfaltigkeit der Kodimension $n + m$.*

Beweis: Sei $x \in M \cap N$ und (U, f) sowie (V, g) definierend für M bzw. N bei x. Wir setzen $W := U \cap V$ und $h := (f, g) : U \to \mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m}$. Dann ist (W, h) definierend für $M \cap N$ bei x. In der Tat ist $W \cap M \cap N = \{h = 0\}$ und für $y \in W \cap M \cap N$ gilt $\dim \mathtt{im}(dh(y)) = \dim(E) - \dim\ker(h) = \dim(E) - \dim\ker(df(y)) \cap \dim\ker(dg(y)) = \dim(E) + \dim(E) - \dim\ker(df(y)) - \dim\ker(dg(y)) = n + m$. Damit ist y ein regulärer Punkt von h. ■

Sei $U \subseteq F$ eine offene Teilmenge eines m-dimensionalen Vektorraumes, $f : U \to E$ stetig differenzierbar.

Definition 8.51 *f ist eine* ***Immersion****, wenn $df(u)$ für alle $u \in U$ injektiv ist. Die Abbildung f ist eine Einbettung, wenn f eine Immersion ist und $f : U \to f(U)$ ein Homöomorphismus.*

1. Seien $e_1, \ldots, e_k \in E$. Dann ist $f : \mathbb{R}^k \to E$ genau dann eine Einbettung, wenn $(e_i)_{i=1}^k$ linear unabhängig ist.

2. Die Abbildung $f : (0,1) \cup (2,3) \to \mathbb{R}^2$ gegeben durch $f(x) = (x, \sin(x^{-1})$ für $x \in (0,1)$ und $f(x) := (0, 5 - 10(x-2))$ für $x \in (2,3)$ ist eine Immersion, aber keine Einbettung. In der Tat ist nämlich $\mathtt{im}(f)$ zusammenhängend, nicht aber der Definitionsbereich.

3. Die Abbildung $f : \mathbb{R} \to \mathbb{R}^2$, $x \mapsto f(x) := (x^2, x^3)$ ist keine Immersion, da $df(0)$ nicht injektiv ist.

4. Sei $E = F \oplus N$. Ist $g : U \to N$ stetig differenzierbar, dann ist $f : U \to F \oplus N \cong E$ gegeben durch $f(x) := (x, g(x))$ eine Einbettung mit dem Bild $\Gamma(g)$.

Lemma 8.52 *Ist $f : U \to E$ eine Einbettung, dann ist $f(U) \subseteq E$ eine Untermannigfaltigkeit. Es gilt $T_{f(u)}f(U) = \mathtt{im}df(u)$.*

Beweis: Sei $x = f(u) \in f(U)$, $u \in U$. O.B.d.A (nach verschieben) können wir annehmen, daß $x = 0$ ist. Wir setzen $T := \mathtt{im}df(u)$ und wählen ein Komplement $N \subseteq E$ derart, daß $E \cong T \oplus N$. Wir betrachten die Projektion $P : E \to T$. Das Differential $d(P \circ f)(d) = P \circ df(u)$ ist ein Isomorphismus. Folglich existieren Umgebungen $V \subseteq U$ und $W \subset T$ von u und 0 derart, daß $P \circ f : V \to W$ eine stetig differenzierbare Umkehrfunktion $h : W \to V$ besitzt. Wir finden nun eine Umgebung $X \subseteq E$ von x derart, daß $P(X) \subseteq W$ und $f^{-1}(X) \subseteq V$ ist (da f eine Einbettung ist).

Wir definieren $\phi : X \to N$ durch $\phi(y) := y - f(h(P(y)))$. In der Tat ist $P\phi(y) = P(y) - P(f(h(P(y)))) = P(y) - P(y) = 0$. Wenn $\phi(y) = 0$ ist, dann gilt $y = f(h(P(y)))$, also $y \in f(U)$. Folglich ist $f(U) \cap X = \{\phi = 0\}$.

Weiterhin ist $d\phi(y)_{|N} = \mathtt{id}_N$ und deshalb ϕ regulär in y. Damit definiert (X, ϕ) die Teilmenge $f(U)$ nahe y. ■

1. Die Spirale $f(\mathbb{R}) \subset \mathbb{R}^2$ mit $f(x) := \mathtt{e}^x(\sin(x), \cos(x))$ ist eine Untermannigfaltigkeit.

2. Das Paraboloid $g(\mathbb{R}^2) \subseteq \mathbb{R}^3$, $g(x,y) := (x, y, x^2 + y^2)$ ist eine Untermannigfaltigkeit.

3. Die Kurve $g \circ f(\mathbb{R}) \subset \mathbb{R}^3$ (f aus 1. und g aus 2.) ist eine Untermannigfaltigkeit.

Wir identifizieren $\mathtt{Mat}(n,n) := \mathtt{Mat}(n,n,\mathbb{R}) \cong \mathbb{R}^{n^2}$.

1. Die Gruppe $GL(n,\mathbb{R}) \subset \mathtt{Mat}(n,n)$ ist eine Untermannigfaltigkeit der Kodimension 0. In der Tat ist diese eine offene Teilmenge.

2. Die Gruppe $O(n) \subset GL(n,\mathbb{R})$ ist eine Untermannigfaltigkeit der Kodimension $\frac{n(n+1)}{2}$. Eine Matrix $A \in \mathtt{Mat}(n,n)$ ist genau dann in $O(n)$, wenn $AA^t = \mathtt{id}$ gilt. Also ist $O(n) = \{f = 0\}$ für $f(A) = AA^t - \mathtt{id}$. Nun ist $f(A) = f(A)^t$. Es gilt weiter $df(A)(B) = AB^t + BA^t$. Insbesondere ist $df(1)(B) = B + B^t$ immer symmetrisch. Damit kann $df(1)$ nicht surjektiv sein, wenn man f mit Werten in allen Matrizen betrachtet. Sei $Sym(n) := \{B \in \mathtt{Mat}(n,n) | B = B^t\} \subset \mathtt{Mat}(n,n)$ der lineare Teilraum der symmetrischen Matrizen. Eine symmetrische Matrix B ist durch ihre Einträge B_{ij} mit $i \leq j$ bestimmt. Folglich gilt $\dim Sym(n) = \frac{n(n+1)}{2}$. Wir betrachten nun $f : \mathtt{Mat}(n,n) \to Sym(n)$. Dann ist $df(1) : \mathtt{Mat}(n,n) \to Sym(n)$, $B \mapsto B + B^t$ sicher surjektiv. Es gilt weiter für $A \in O(n)$, daß $df(A)(BA) = B^t + B$, also für $C \in Sym(n)$ auch $df(A)(\frac{1}{2}CA) = C$. Damit ist auch $df(A)$ surjektiv. Das Paar $(\mathtt{Mat}(n,n), f : \mathtt{Mat}(n,n) \to Sym(n))$ definiert $O(n)$.

3. Die Untergruppe $SO(n) \subset O(n)$ ist eine Umtermannigfaltigkeit der Kodimension $\frac{n(n+1)}{2}$. In der Tat, die Menge $U := \{\det(A) > 0\} \subset \mathtt{Mat}(n,n)$ ist offen, und $(U, f_{|U})$ definiert $SO(n)$.

8.10 Extremwerte mit Nebenbedingungen

Sei E ein endlich-dimensionaler Vektorraum, $M \subseteq E$ eine Untermannigfaltigkeit, $U \subseteq E$ offen und $f : U \to \mathbb{R}$.

Definition 8.53 *Die Funktion f hat ein lokales **Extremum mit Nebenbedingung** M, falls $f_{|M\cap U}$ ein lokales Extremum hat.*

Satz 8.54 *Wenn f in $m \in M$ ein lokales Extremum mit Nebenbedingung M hat, dann gilt $df(m)(T_mM) = 0$.*

Beweis: Wir nehmen o.B.d.A an, daß $m = 0$. Wir wählen eine Aufspaltung $E = T_0M \oplus N$ und Umgebungen $V \subseteq T_0M$ und $W \subseteq N$ sowie eine Abbildung $g : V \to W$ derart, daß $M \cap (V \times W) = \Gamma(g)$ gilt und $dg(0) = 0$. Dann ist $\phi : V \to M \cap (V \times W)$, $\phi(x) = (x, g(x))$ eine Einbettung. Die Abbildung $\phi^* f$ hat in 0 ein lokales Extremum. Also gilt $0 = d(\phi^* f) = df(0) \circ d\phi(0)$. Die Behauptung folgt nun aus $\mathtt{im}(d\phi(0)) = T_0M$. ■

Lemma 8.55 *Sei $A : E \to F$ eine (surjektive) lineare Abbildung, $h \in E'$. Wenn $h_{|\ker(A)} = 0$, dann existiert (genau) ein $\lambda \in F'$ derart, daß $h = \lambda \circ A$. Umgekehrt, wenn $\lambda \in F'$ existiert mit $h = \lambda \circ A$, dann ist $h_{|\ker(A)} = 0$.*

Beweis: Lineare Algebra! ■

Wir können mit dieser Beobachtung die Bedingung $df(m)_{|T_mM} = 0$ umformulieren. Sei $(U, \phi : U \to F)$ eine definierende Funktion für $M \subseteq E$ bei m.

Satz 8.56 *Wenn die Funktion f in $m \in M$ ein lokales Extremum mit Nebenbedingung in M hat, dann existiert ein $\lambda \in F'$ derart, daß $\lambda \circ d\phi(m) = df(m)$.*

Die Linearform λ heißt **Langrangescher Multiplikator**. Um also die lokalen Extremwerte mit Nebenbedingung in $M = \{\phi = 0\}$ zu finden, müssen wir insbesondere das im allgemeinen nichtlineare Gleichungssystem

$$\begin{aligned} \phi(x) &= 0 \\ df(x) &= \lambda \circ d\phi(x) \end{aligned}$$

für $(x, \lambda) \in U \times F'$ lösen.

Wir brauchen zweite Ableitungen, um zu entscheiden, ob in einem Punkt $m \in M$ mit $df(m)(T_m M) = 0$ ein lokales Extremum vorliegt. Wir nehmen deshalb an, daß M von der Klasse C^2 ist. O.B.d.A sei $m = 0$. Wir wählen eine Aufspaltung $E \cong T_0 M \oplus N$ und schreiben $M \cap V \times W \cong \Gamma(g)$ als Graph einer C^2-Abbildung $g : V \to W$ für geeignete Umgebungen $V \subseteq T_0 M$ und $W \subseteq N$ der Null.

Satz 8.57 *Wenn $df(0)(T_0 M) = 0$ und die quadratische Form*

$$Q := \mathtt{Hess}_0(f) + df(0) \circ \mathtt{Hess}_0(g)$$

auf $T_0 M$ positiv (negativ) definit ist, dann hat f in m ein lokales Minimum (Maximum) mit Nebenbedingung in M.

Beweis: Sei $\phi : V \to E$, $\phi(t) = (t, g(t))$. Wir müssen zeigen, daß $\phi^* f : V \to \mathbb{R}$ in 0 ein lokales Minimum (Maximum) hat. Es gilt $d\phi^* f(t) = df(\phi(t)) \circ d\phi(t)$ und insbesondere wegen $\mathtt{im}(d\phi)(0) = T_0 M$, daß $d\phi^* f(0) = 0$. Wir berechnen nun $\mathtt{Hess}_0 \phi^* f$. Dazu leiten wir für $\xi \in T$ die Abbildung $t \mapsto df(\phi(t)) \circ d\phi(t)(\xi)$ noch einmal in die Richtung ξ ab. Mit der Produktregel erhalten wir für $t = 0$

$$d^2 f(0)(d\phi(0)(\xi))(d\phi(0)(\xi)) + df(0) \circ d^2\phi(0)(\xi)(\xi) = Q(\xi) \ .$$

Wir wenden nun auf $\phi^* f$ den entsprechenden Satz 8.35 für den Fall ohne Nebenbedingungen an. ∎

Hier ist ein Beispiel: Sei $f_1 : \mathbb{R}^3 \to \mathbb{R}$ durch

$$f_1(x, y, z) = x^2 + 2y^2 + 3z^2 - 30$$

gegeben. Dann ist $M := \{f_1 = 0\}$ ein Ellipsoid. Sei weiter

$$f_2(x, y, z) := x + 2y + 3z \ .$$

Dann ist $N := \{f_2 = 0\}$ eine Hyperebene. Wir betrachten $L := N \cap M$ und behaupten, daß L eine Untermannigfaltigkeit ist. Wir müssen zeigen, daß $N \pitchfork M$ gilt, oder alternativ, daß $f := (f_1, f_2) : \mathbb{R}^3 \to \mathbb{R}^2$ in allen Nullstellen regulär ist. Wir wählen letzteren Zugang. Es gilt

$$df(x,y,z) = \begin{pmatrix} 2x & 4y & 6z \\ 1 & 2 & 3 \end{pmatrix} .$$

Aus $\mathtt{rk}(df(x,y,z)) < 2$ folgt $x = y = z$. Der einzige solche Punkt in N ist $(x,y,z) = 0$. Dieser liegt aber nicht in M.

Wir betrachten jetzt die Höhenfunktion $h(x,y,z) := z$ und suchen deren lokale Extrema mit Nebenbedingung in L.

Es gilt $dh(x,y,z) = (1,0,0)$. Wir suchen nun alle Punkte $l \in L$ mit $(1,0,0)(T_l L) = 0$. Wir verwenden die Methode der Lagrangeschen Multiplikatoren. Wir suchen also alle diejnigen Punkte $l \in L$, für welche es eine Linearform $\lambda = (\lambda_1, \lambda_2) \in (\mathbb{R}^2)'$ derart gibt, daß

$$\lambda \circ df(l) = (1,0,0)$$

gilt. Wir müssen also die Lösbarkeit dieses Gleichungssystems studieren. Die Gleichungen lauten

$$\lambda_1 2x + \lambda_2 = 1 \ , \quad \lambda_1 4y + 2\lambda_2 = 0 \ , \quad \lambda_1 6z + 3\lambda_2 = 0 \ .$$

Die zweite Gleichung gibt $\lambda_2 = -\lambda_1 2y$. Die dritte Gleichung liefert dann $\lambda_1(6z - 6y) = 0$, also $y = z$. Die erste Gleichung liefert $(2x - 2y)\lambda_1 = 1$, also $x \neq y$. Wir suchen also diejenigen Punkte in L, welche $y = z$ und $x \neq y$ erfüllen. Wir müssen also die Gleichungen

$$x + 5y = 0 \ , \quad x^2 + 2y^2 + 3y^2 = 30$$

lösen. Die erste gibt $x = -5y$, und dies in die zweite eingesetzt, $30y^2 = 30$, also $y^2 = 1$. Wir schließen, daß $y = z = \pm 1$ und $x = \pm 5$ die kritischen Punkte sind. Potentielle Extrema liegen in $l_+ := (-5,1,1)$ und $l_- := (5,-1,-1)$ vor.

Wir bestimmen nun den Tangentialraum

$$T_{l_\pm} L := \ker \begin{pmatrix} \mp 10 & \pm 4 & \pm 6 \\ 1 & 2 & 3 \end{pmatrix} .$$

Es gilt

$$T_{l_\pm} L = \mathbb{R} \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}$$

Wir wählen als Komplement

$$N := \mathbb{R} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Wir wollen L lokal als Graph einer Funktion $g = l_\pm + (g_1, 0, g_2) : T_{l_\pm} L \to N$ darstellen. Die Funktion g ist Lösung der Gleichungen

$$\begin{aligned} (\mp 5 + g_1(t))^2 + 2(\pm 1 + 3t)^2 + 3(\pm 1 + g_2(t) - 2t)^2 - 30 &= 0 \\ \pm 5 + g_1(t) + 2(\pm 1 + 3t) + 3(\pm 1 + g_2(t) - 2t) &= 0 \ . \end{aligned}$$

Wir wissen, daß $g(0) = 0$ und $dg(0) = 0$ gilt und müssen $d^2g(0)$ berechnen. Dazu leiten wir die erste Gleichung einmal ab und erhalten

$$2(\mp 5 + g_1(t))g_1'(t) + 12t(\pm 1 + 3t) + 6(\pm 1 + g_2(t) - 2t)(g_2'(t) - 2) = 0 \ .$$

Wir leiten zum zweiten Mal ab und setzen $t = 0$ ein.

$$\mp 10 g_1''(0) + \pm 12 + 24 + \pm 6 g_2''(0) = 0 \ .$$

Die zweite Ableitung der zweiten Gleichung liefert

$$g_1'' + 3g_2'' = 0 \ .$$

Wir setzen dies in die erste Gleichung ein und erhalten

$$\pm 36 g_2''(0) + 24 + \pm 12 \ ,$$

also

$$g_2''(0) = \begin{cases} -1 & + \\ \frac{1}{3} & - \end{cases}$$

Die zu untersuchende quadratische Form ist nun durch $Q(\xi) = g_2''(0)\xi^2$ gegeben. Folglich liegt im Punkt $(-5, 1, 1)$ ein Maximum und im Punkt $(5, -1, -1)$ ein lokales Minimum mit Nebenbedingung in L vor.

In der Tat ist M und damit L kompakt. Die Funktion h muß also ein globales Maximum und Minimum auf L besitzen. Wir haben also die globalen Extremwerte gefunden.

9 Differentialgleichungen

9.1 Erste Beispiele

Wir hatten die e-Funktion als die eindeutige Lösung der Differentialgleichung

$$f' = f \ , \quad f(0) = 1$$

für eine Funktion $f : \mathbb{R} \to \mathbb{R}$ charakterisiert.

Die Funktionen

$$f_{c,d}(t) := c\sin(t) + d\cos(t) \ , \quad c, d \in \mathbb{R}$$

waren alle alle Lösungen der Differentialgleichung

$$f^{(2)} + f = 0 \ , \quad f : \mathbb{R} \to \mathbb{R} \ .$$

Insbesondere ist $f_{c,d}$ die eindeutige Lösung mit $f_{c,d}(0) = d$ und $f'_{c,d}(0) = c$.

Sei $x(t)$ die Bahnkurve eines Massenpunktes in Raum $\mathbb{R}^3$ der Masse m in einem Kraftfeld $F : \mathbb{R}^3 \to \mathbb{R}^3$. Das Newtonsche Gesetz stellt einen Beziehung zwischen der Beschleunigung $x^{(2)}$ des Massenpunktes und der auf ihn wirkenden Kraft $F \in \mathbb{R}^3$ her:

$$x^{(2)}(t) = m^{-1}F(x(t)) \ .$$

Als Beispiel betrachten wir die Bewegung in der Schwerkraft. An der Erdoberfläche ist die Anziehungskraft konstant durch $f(x) = (0, 0, mg_{Erd})$ gegeben. Wir erhalten

$$x^{(2)} = (0, 0, g_{Erd}) \ .$$

Die Bahn einer im Winkel α nach oben im Punkt 0 in der x, z-Ebene abgeschossene Kanonenkugel mit der Anfangsgeschwindigkeit v_0 erfüllt diese Gleichung mit der Anfangsbedingung

$$x(0) = 0 \ , x'(0) = v_0(\cos(\alpha), 0, \sin(\alpha)) \ .$$

Die Lösung dieser Gleichung ist

$$x(t) = (v_0 \cos(\alpha)t, 0, v_0 \sin(\alpha)t - \frac{g_{Erd}}{2}t^2) \ .$$

Die Kugel schlägt zur Zeit

$$t := \frac{2\sin(\alpha)v_0}{g_{Erd}}$$

auf. Die x-Koorrdinate des Einschlages ist

$$\frac{2\sin(\alpha)\cos(\alpha)v_0^2}{g_{Erd}} = \frac{v_0^2 \sin(2\alpha)}{g_{Erd}} \ .$$

Die maximale Schußweite wird also bei einem Winke $\alpha = \frac{\pi}{4}$ erreicht und beträgt

$$\frac{v_0^2}{g_{Erd}} \ .$$

Wir betrachten nun die Gleichung, die die Planetenbahnen bestimmt. Die Anziehungskraft der Sonne (im Ursprung) auf die Erde im Punkt x wird nach dem Gravitationsgesetz durch

$$F(x) := -\frac{\gamma m_{Sonne} m_{Erde} x}{|x|^3} \ .$$

Folglich genügt die Erdbahn $x(t)$ der Differentialgleichung

$$x^{(2)} = -\frac{\gamma m_{Sonne}}{|x|^3} x \ .$$

Wir sehen, daß die Masse der Erde nicht eingeht, sich also ein leichter Satellit nach der gleichen Regel bewegen würde. Diese Gleichung bestimmt die von Kepler gefundenen elliptischen Planetenbahnen, ist aber schwerer explizit zu lösen.

In beiden Fällen wird die Kraft durch ein Vektorfeld beschrieben:

$$F(x) = (0, 0, g_{Erd}) \ , \quad F(x) = -\frac{\gamma M_{Sonne} x}{|x|^3} \ .$$

Die Kraft wird hier durch das Newtonsche Gesetz mit der Beschleunigung in Beziehung gesetzt. Wir erhalten Differentialgleichung zweiter Ordnung. Für theoretische Betrachtungen ist es wichtig zu sehen, daß man solche Gleichgungen auch als Gleichungen erster

Ordnung schreiben kann. Dazu nehmen wir die Geschwindigkeit v als eine zweite Koordinate hinzu. Die Bahnkurve des Massenpunktes ist jetzt eine Abbildung $(x, v) : \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$. Die Bewegungsgleichung hat nun die Form

$$x' = v \ , \quad v' = m^{-1}F(x) \ ,$$

also

$$\begin{pmatrix} x' \\ v' \end{pmatrix} = \begin{pmatrix} v \\ m^{-1}F(x) \end{pmatrix} \ . \tag{3}$$

Oft is eine Reibungskraft propertional zur Geschwindigkeit, $F_{Reib} = -\epsilon v$. Mit Reibung hätte die Gleichung die Form

$$\begin{pmatrix} x' \\ v' \end{pmatrix} = \begin{pmatrix} v \\ m^{-1}F(x) - \epsilon v \end{pmatrix} \ .$$

Für die Diskussion des Verhaltens von Lösungen derartiger Gleichungen spielen Erhaltungsgrößen eine wichtige Rolle. Wir nehmen an, daß die Kraft (3) durch ein Potential $U : \mathbb{R}^3 \to \mathbb{R}$ gegeben wird:

$$F = -(\partial_x U, \partial_y U, \partial_z U) \ .$$

Im Fall des Schwerefeldes ist etwa $U(x, y, z) = mg_{Erd}z$, und im Fall des Gravitationsfeldes der Sonne ist

$$U(x) = -\frac{\gamma m_{Sonne} m_{Erde}}{|x|} \ .$$

In diesem Fall ist die Energie

$$E(x, v) : \mathbb{R}^6 \to \mathbb{R} \ , \quad E(x, v) = \frac{mv^2}{2} + U(x)$$

eine Erhaltungsgröße. In der Tat gilt

$$\frac{d}{dt}E(t) = E(x(t), v(t)) = \partial_x U v + mvv' = -Fv + vF = 0 \ .$$

Ist die Anfangsenergie $E_0 := E(x(0), v(0))$ vorgegeben, dann weiß man, daß sich die Bahnkurve auf der Hyperfläche $\{E = E_0\} \subset \mathbb{R}^6$ bewegt. Die Energie des Teilchens im Gravitationsfeldes hängt nicht von den Richtungen ab. Die Bahnkurve ist also durch

$$E_0 := \frac{v^2}{2} - \frac{\gamma m_{Sonne}}{|x|}$$

eingeschränkt. Ist $E_0 > 0$, dann ist die Energiefläche $\{E = E_0\}$ in der x-Richtung unbeschränkt. Die Geschwindigkeit

$$v_0 := \sqrt{\frac{2\gamma m_{Sonne}}{|x|}}$$

entspricht $E_0 = 0$. Startet man bei x radial von Zentrum weg mit $v > v_0$, dann entweicht die Bahn nach unendlich. Ist $v < v_0$, d.h. $E_0 < 0$, dann ist der maximale Abstand, den man bei einem solchen Start erreichen kann,

$$x_{max} = \frac{\gamma m_{Sonne}}{E_0} \; .$$

Um die Keplerbahnen herauszubekommen, muß man weitere Erhaltungsgrößen hinzuziehen, nämlich die Drehimpulse.

Ist Reibung vorhanden, dann ist E nicht mehr konstant, aber es gilt

$$\frac{d}{dt}E(x(t)) = -m\epsilon|v|^2 \le 0 \; .$$

Diese Ungleichung beschreiben den Energieverlust in Verlaufe der Zeit und ist damit eine wichtige Aussage über das Langzeitverhalten.

Wir betrachten den radioaktiven Zerfall, etwa

$${}^{238}_{92}U \xrightarrow{\alpha} {}^{234}_{90}Th \; .$$

Das empierisch bestätigtige Zerfallsgesetz besagt, daß die Anzahl der in einem Zeitintervall zerfallenden Kerne proportial zur vorhandenen Menge und der Länge des Zeitintervalls ist. Sei $x(t)$ die Anzahl der Atome zur Zeit t. Dann gilt

$$x(t + \delta t) = x(t) - \alpha x(t)\delta t \; .$$

Daus leiten wir die Differentialgleichung

$$x' = -\alpha x \; .$$

Wenn am Anfang $x(0)$ Kerne vorhanden waren, dann sind es nach der Zeit t noch $x(t) = e^{-\alpha t}x(0)$. Nach der Zeit $t_{1/2} := -\alpha^{-1} \ln \frac{1}{2}$ sind die Häfte der Kerne zerfallen. Diese Zeit heißt Halbwertszeit.

Wir betrachten die Entladung eines Kondensators C über einen Widerstand R. Die Spannung U am Kondensator ist durch seine Ladung über $UC = Q$ bestimmt. Der Entladestrom ist nach dem Ohmschen Gesetz $I := U/R$. Die Stromstärke bestimmt die Veränderung der Ladung $Q' = -I$. Folglich gilt

$$Q' = -\frac{1}{RC}Q$$

und folglich

$$Q(t) = Q(0)e^{-\frac{t}{RC}} \; ,$$

bzw

$$U(t) = U(0)e^{-\frac{t}{RC}} \; .$$

Die Größe $\frac{1}{RC}$ bestimmt also die Geschwindigkeit der Entladung.

Ersetzt man den Widerstand durch eine Spule der Induktivität L (ohne Ohmschen Widerstand), dann gilt nach dem Induktionsgesetz. $L^{-1}U = I'$. Wir erhalten

$$Q^{(2)} = -I' = -\frac{1}{LC}Q \ .$$

Wir erhalten die Schwingungsgleichung

$$Q^{(2)} + \frac{1}{LC}Q = 0 \ .$$

Die Größe

$$\omega := \frac{1}{2\pi}\sqrt{\frac{1}{LC}}$$

heißt Kreisfrequenz. Die Lösungen der Gleichung haben die Form

$$a\sin(2\pi\omega t) + b\cos(2\pi\omega t) \ .$$

Im Gegensatz zur Kondensatorentladung erhalten wir ein periodisches Verhalten.

In der Populationsdynamik ist die logistische Gleichung ein beliebtes Modell. Eine Spezies mit der Individuenzahl x vermehrt sich mit der Rate c. Dies wird durch $x'(t) = cx(t)$ beschrieben und führt auf ein exponentielles Wachstum $x(t) = e^{ct}x(0)$. Sind die Ressourcen begrenzt, kann das durch eine populationsabhängige Rate $c(x) = c(1 - A^2x^2)$ beschrieben werden. In diesem Fall wird eine kleine Population entsprechend

$$x'(t) = cx(t)(1 - A^2x^2(t))$$

erst exponentiell wachsen und sich dann der Zahl A^{-1} von unten annähern. Dies folgt aus einer qualitativen Diskussion. Für kleine t und $x(0)$ ist $x' \approx cx$ und $x(t) = e^{ct}x(0)$. Wenn sich $x(t)$ der Grenze A^{-1} nähert, dann gilt $x' = -2cA^2(x - A^{-1})$, also $x(t+s) - A^{-1} = (x(t) - A)e^{-2cA^2 s}$. Wir nehmen nun an, daß es eine weitere Art mit der Individuezahl $y(t)$ gibt, welche sich von x ernährt und sich mit eine Rate proportional zum Nahrungsangebot vermehrt und mit konstanter Rate stirbt. Dies wird durch

$$x' = cx(1 - A^2x^2) - fy \ , \quad y' = gxy - sy$$

beschrieben, also

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} cx(1 - A^2x^2) - fxy \\ gxy - sy \end{pmatrix} \ .$$

Das Kurz- und Langzeitverhalten dieses Modells werden wir später analysieren.

9.2 Vektorfelder und Integralkurven

Sei $(V, \|.\|)$ ein reeller Banachraum und $U \subseteq V$ offen.

Definition 9.1 *Ein* ***Vektorfeld*** X *auf* U *ist eine Abbildung* $X : U \to V$.

Definition 9.2 *Eine **Integralkurve** von X ist eine differenzierbare Abbildung $f : I \to U$, welche der Gleichung $f' = f^*X$ genügt, wobei $I \subseteq \mathbb{R}$ ein offenes Intervall ist.*

Hier sind einige Beispiele.

1. Sei $V := \mathbb{R}^1$ und $X(x) := 1$. Für alle $c \in \mathbb{R}$ ist $f_c : \mathbb{R} \to \mathbb{R}^1$ $f_c(t) := c + t$ eine Integralkurve. In der Tat gilt
$$f_c'(t) = 1 = X \circ f(t) = (f^*X)(t) \ .$$

2. Sei $V := \mathbb{R}^1$ und $X(x) := x$. Für alle $c \in \mathbb{R}^1$ ist $f : \mathbb{R} \to \mathbb{R}^1$, $f_c(t) := c\mathsf{e}^t$ eine Integralkurve von X. In der Tat gilt
$$f_c'(t) = c\mathsf{e}^t = f(t) = X(f(t)) = (f^*X)(t) \ .$$

3. Sei $V := \mathbb{R}^1$ und $X(x) = x^2$. Für jedes $c \in \mathbb{R}^1$ ist $f_c : I_c \to \mathbb{R}^1$, $f_c(t) := \frac{c}{1-ct}$ eine Integralkurve, wobei
$$I_c := \begin{cases} (-\infty, \frac{1}{c}) & c > 0 \\ (\frac{1}{c}, \infty) & c < 0 \\ \mathbb{R} & c = 0 \end{cases} \ .$$
In der Tat gilt
$$f'(t) = (\frac{c}{1-ct})^2 = f(t)^2 = f^*X(t) \ .$$

4. Sei $V := \mathbb{R}^2$ und $X(x,y) := (y, -x)^t$. Wir schreiben
$$D := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
und $X(v) = Dv$. Für $t \in \mathbb{R}$ setzen wir
$$A(t) := \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \ . \tag{4}$$
Für jedes $c \in \mathbb{R}^2$ ist die Kurve $f_c : \mathbb{R} \to \mathbb{R}^2$, $f_c(t) := A(t)c$, eine Integralkurve von X. In der Tat gilt
$$\begin{aligned} f_c' &= A'(t)c \\ &= \begin{pmatrix} -\sin(t) & \cos(t) \\ -\cos(t) & -\sin(t) \end{pmatrix} c \\ f^*X(t) &= X(f(t)) \\ &= DA(t)c \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} c \\ &= \begin{pmatrix} -\sin(t) & \cos(t) \\ -\cos(t) & -\sin(t) \end{pmatrix} c \end{aligned}$$

9.3 Gewöhnliche Differentialgleichungen n-ter Ordnung

Sei $U \subseteq \mathbb{R}^n$ offen und $R : U \to \mathbb{R}$ eine Abbildung.

Definition 9.3 *Sei $U \subset \mathbb{R}^{n+1}$ offen und $R : U \to \mathbb{R}$ eine Abbildung. Eine n-mal differenzierbare Funktion $h : I \to \mathbb{R}$ heißt Lösung der* ***gewöhnlichen Differentialgleichung***

$$h^{(n)} = R(t, h, h', \ldots, h^{(n-1)}) ,$$

falls für die Abbildung $I \ni t \mapsto f(t) := (t, h(t), h'(t), \ldots, h^{(n-1)}(t)) \in \mathbb{R}^{n+1}$ gilt

1. *$f(I) \subseteq U$*
2. *$h^{(n)} = f^* R$,*

wobei $I \subseteq \mathbb{R}$ ein offenes Intervall ist.

Wir betrachten das Vektorfeld auf U, welches durch $X(x) = (1, x_3, \ldots, x_{n+1}, R(x))$ gegeben wird.

Lemma 9.4 *$h : I \to \mathbb{R}$ is genau dann Lösung der gewöhnlichen Differentialgleichung*

$$h^{(n)} = R(t, h, h', \ldots, h^{(n-1)}) ,$$

wenn $f : I \to U$, $f(t) := (t, h(t), h'(t), \ldots, h^{(n-1)}(t))$ eine Integralkurve von X ist.

Beweis: Ist h Lösung der gewöhnlichen Differentialgleichung, so ist offensichtlich f eine Integralkurve. Umgekehrt, ist f eine Integralkurve von X, so setzen wir $h(t) := f_2(t)$. Es gelten die Gleichungen

$$\begin{array}{rcl} f_1(t) & = & t \\ h'(t) & = & X_2(f(t)) = f_3(t) \\ h''(t) & = & f_3'(t) = X_3(f(t)) = f_4(t) \\ & \ldots & \\ h^{(n-1)}(t) & = & f_n'(t) = X_n(f(t)) = f_{n+1}(t) \\ h^{(n)}(t) & = & f_{n+1}'(t) = X_{n+1}(t) = (f^* R)(t) \ , \end{array}$$

aus denen ersichtlich ist, daß h n-mal differenzierbar ist und die gewöhnliche Differentialgleichung erfüllt. ■

9.4 Integration Banachraumwertiger Funktionen

Weiter vorne im Kapitel 6 hatten wir das Integral reeller Funktionen über die Integration einfacher Ober- und Unterfunktionen eingeführt. Diese Methode basiert wesentlich auf der Ordnungsrelation im Wertebereich $\mathbb{R}$. Folglich funktioniert diese Methode nicht für Banachraumwertige Funktionen.

Sei $(V, \|.\|)$ ein Banachraum und $[a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall. Ein Partition P von $[a, b]$ ist eine endliche monotone Folge $a = t_0 \le t_1 \le \cdots \le t_n = b$ reeller Zahlen. Die Zahl $w(P) := \mathtt{max}\{t_i - t_{i-1} | i = 1, \ldots, n\}$ heißt Weite der Partition. Sind P und Q Partitionen, dann kann man eine gemeinsame Verfeinerung $P \sharp Q$ bilden, indem man die Vereinigung der Glieder von P und Q der Größe nach ordnet. Es gilt $w(P \sharp Q) \le \mathtt{min}\{w(P), w(Q)\}$.

Wir betrachten eine Abbildung $f : [a, b] \to V$ und bilden die Summe

$$S_P(f) := \sum_{i=1}^{n} (t_i - t_{i-1}) f(t_i) \ .$$

Definition 9.5 *Wir sagen, daß daß Integral* $\int_a^b f(t)dt$ *existiert und gleich* $x \in V$ *ist, wenn für jede Folge* (P_n) *von Partitionen von* $[a, b]$ *mit* $w(P_n) \to 0$ *gilt* $\lim_{n\to\infty} S_{P_n}(f) = x$.

Lemma 9.6 *Wenn* $f : [a, b] \to V$ *stetig ist, dann existiert* $\int_a^b f(t)dt$.

Beweis: In der Tat ist f gleichmäßig stetig, da das Intervall kompakt ist. Sei $\epsilon > 0$ gegeben. Dann finden wir ein $\delta > 0$ derart, daß aus $|t-s| \le \delta$ folgt, daß $\|f(t-f(s)\| \le \frac{\epsilon}{b-a}$ ist.

Seien nun $P = (t_0, \ldots, t_n)$ eine Partition mit $w(P) < \delta$ und $Q = (s_0 \le \cdots \le s_m)$ eine weitere Partition. Dann ist $|S_{P\sharp Q}(f) - S_P(f)| = \epsilon$. In der Tat ersetzt man in der Summe $S_{P\sharp Q}(f)$ in den Summanden mit dem Faktor $f(s_i)$ mit $t_{j-1} \le s_i \le t_j$ diesen Faktor durch $f(t_j)$. Es gilt $|s_i - t_j| < \delta$ und somit $\|f(s_i) - f(t_j)\| < \frac{\epsilon}{b-a}$. Die Summe verändert man damit höchstens durch ϵ. Durch diese Ersetzung erhält man jedoch den Wert von $S_P(f)$.

Sei nun (P_n) ein Folge von Partitionen mit $w(P_n) \to 0$. Dann ist $S_{P_n}(f)$ eine Cauchyfolge und konvergiert gegen einen Wert x. In der Tat, wenn nur $w(P_n) < \delta$ und $w(P_m) < \delta$ ist, dann ist

$$\|S_{P_n}(f) - S_{P_m}(f)\| \le \|S_{P_n}(f) - S_{P_n \sharp P_m}(f)\| + \|S_{P_n \sharp P_m}(f) - S_{P_m}(f)\| \le 2\epsilon \ .$$

Ist (Q_n) eine weitere Folge von Partitionen mit $w(Q_n) \to 0$, dann bilden wir die Folge (R_n) durch $R_{2n} := P_n$ und $R_{2n+1} := Q_n$. Da $\lim_{n\to\infty} S_{R_n}(f)$ existiert, müssen auch die beiden Teilfolgen $(S_{P_n}(f))$ und $(S_{Q_n}(f))$ den gleichen Grenzwert, nämlich x haben. ■

Lemma 9.7 *Das Integral ist eine stetige lineare Abbildung* $\int_a^b \ldots dt : C([a, b], V) \to V$. *Es gilt*

$$\| \int_a^b f(t)dt \| \le (b-a) \|f\| \ .$$

Beweis: Sei P eine Partition von $[a, b]$. Dann gilt sicherlich für $f, g \in C([a, b], V)$ und $\lambda \in \mathbb{R}$, daß

$$S_P(f + \lambda g) = S_P(f) + \lambda S_P(g) \ .$$

Wendet man dies für eine Folge von Partitionen (P_n) mit $w(P_n) \to 0$ an, so erhält im Grenzwert die Linearität des Integrals. Weiterhin gilt sicher

$$\|S_P(f)\| \leq \sum_{i=1}^{n}(t_i - t_{i-1})\|f(t_i)\| \leq \sup_{t\in[a,b]} \|f(t)\| \sum_{i=1}^{n}(t_i - t_{i-1}) = \|f\|(b-a) \ .$$

Daraus folgt im Grenzwert

$$\| \int_a^b f(t)dt\| \leq \|f\|(b-a) \ .$$

■

Sei $f : [a, b] \to V$ stetig und $c \in (a, b)$.

Lemma 9.8 *Es gilt*

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt \ .$$

Beweis: Sei P eine Partition von $[a, c]$ und Q eine Partition von $[c, b]$. Dann bilden wor die Vereinigung $P \cup Q$, eine Partition von $[a, b]$. Es gilt offensichtlich

$$S_P(f_{|[a,c]}) + S_Q(f_{|[c,b]}) = S_{P\cup Q}(f) \ .$$

Weiterhin ist $w(P \cup Q) \leq \mathtt{max}\{w(P), w(Q)\}$. Wenden wir dies auf Folgen (P_n) und (Q_n) mit $w(P_n) \to 0$ und $w(Q_n) \to 0$, dann erhalten wir im Grenzwert das gewünschte Ergebnis. ■

Sei $f : [a, b] \to V$ stetig und $c \in (a, b)$. Wir definieren $F : (a, b) \to V$ durch

$$F(t) := \begin{cases} \int_c^t f(s)ds & t > c \\ 0 & t = c \\ \int_t^c f(s)ds & t < c \end{cases} \ .$$

Lemma 9.9 (Hauptsatz der Diff. und Int.-rechnung) *Die Abbildung F ist differenzierbar und es gilt $f'(t) = f(t)$.*

Beweis: Sei O.B.d.A $t > c$. Wir bilden den Differenzenquotienten (O.B.d.A ist $h > 0$)

$$\frac{F(t+h) - F(t)}{h} := h^{-1} \int_t^{t+h} f(s)ds \ .$$

Sei $\epsilon > 0$ gegeben. Dann gibt es ein δ derart., daß aus $|t - s| < \delta$ folgt $\|f(t) - f(s)\| < \epsilon$. Wir schreiben

$$h^{-1} \int_t^{t+h} f(s)ds = f(t) + h^{-1} \int_t^{t+h} (f(s) - f(t))ds$$

(Wir benutzen hier, daß $\int_u^v g(s)ds = (v-u)x$ für eine konstante Funktion g mit dem Wert x gilt) und schätzen die Norm des zweiten Terms für $h < \delta$ durch ϵ ab. Folglich ist für $h < \delta$

$$\|\frac{F(t+h) - F(t)}{h} - f(t)\| < \epsilon \ .$$

Da $\epsilon > 0$ beliebig vorgegeben war, folgt die Behauptung. ■

9.5 Existenz und Eindeutigkeit von Integralkurven

Seien (M, d_M) und $(N.d_N)$ metrische Räume.

Definition 9.10 *Eine Abbildung $X : M \to N$ heißt* ***Lipschitz-stetig****, wenn es eine Konstante $C > 0$ gibt, so daß für alle $a, b \in M$ gilt*

$$d_N(X(a), X(b)) \leq C d_M(a, b) \ .$$

Die Zahl C heißt Lipschitzkonstante für X.

Sei $(V, \|.\|_V)$ ein reeller Banachraum und $U \subseteq V$ offen und $x \in U$.

Theorem 9.11 *Wir nehmen an, daß X Lipschitz-stetig ist. Dann existiert ein $\epsilon_0 > 0$ derart, daß es für jedes $\epsilon \in (0, \epsilon_0)$ genau eine auf $I := (-\epsilon, \epsilon)$ definierte Integralkurve $f : I \to U$ mit $f(0) = x$ gibt.*

Beweis: Die Zahl $\epsilon > 0$ wird später im Beweis festgelegt. Sei C_X ein Lipschitzkonstante von X.

Wir betrachten den Banachraum $E := C(I, V)$ der stetigen V-wertigen beschränkten Funktionen auf I mit der Norm $\|f\|_E := \sup_{t \in I} \|f(t)\|_V$. Sei $E_U := E_U(\epsilon) := \{f \in E | f(I) \subseteq U\}$. Die Teilmenge $E_U \subseteq E$ ist nicht leer. da sie zum Beispiel die konstanten Abbildungen mit Werten in U enthält. Auf E_U können wir die Abbildung

$$A := A_\epsilon : E_U \to E \ , \quad A(f)(t) := x + \int_0^t X(f(s))ds \ .$$

definieren. In der Tat ist $Af : I \to V$ beschränkt. Erstens ist nämlich $f : I \to U \to V$ beschränkt und damit $I \ni t \mapsto X(f(t)) \in V$ beschränkt ist wegen der Lipschitzstetigkeit von X. Daraus folgt die Beschränktheit von Af.

Lemma 9.12 *$f \in E_U$ ist genau dann Integralkurve von X mit $f(0) = x$, wenn $A(f) = f$ gilt.*

Beweis: Sei f eine derartige Integralkurve. Dann gilt

$$\begin{aligned} A(f)(t) &= x + \int_0^t X(f(s))ds \\ &= x + \int_0^t f'(s)ds \\ &= x + f(t) - f(0) \\ &= f(t) \ . \end{aligned}$$

Wenn $A(f) = f$ gilt, so ist offensichtlich f differenzierbar, $f(0) = x$, und

$$f'(t) = (\int_0^{\cdot} X(f(s))ds)'(t) = X(f(t)) .$$

■

Es genügt also zu zeigen, daß für ein geeignetes $\epsilon > 0$ die Abbildung A genau einen Fixpunkt in E_U besitzt.

Lemma 9.13 *Die Abbildung A ist Lipschitz-stetig mit einer Lipschitzkonstanten ϵC_X.*

Beweis: Für $f, g \in E_U$ und $t \in I$ gilt

$$\begin{aligned}
\|A(f)(t) - A(g)(t)\|_V &= \|\int_0^t [X(f(s)) - X(g(s))]ds\|_V \\
&\leq |t| \sup_{s\in[0,t]} \|X(f(s)) - X(g(s))\|_V \\
&= \epsilon C_X \sup_{s\in I} \|f(s) - g(s)\|_V \\
&= \epsilon C_X \|f - g\|_E .
\end{aligned}$$

Also

$$\|A(f) - A(g)\|_E = \sup_{t\in I} \|A(f)(t) - A(g)(t)\|_V \leq \epsilon C_X \|f - g\|_E .$$

■

Wenn wir also $\epsilon > 0$ so klein wählen, daß $\epsilon C_X < 1$, dann ist A eine Kontraktion und besitzt höchstens einen Fixpunkt.

Sei nun $f_0 \in E_U$ die konstante Abbildung mit Wert x. Sei $r > 0$ so, daß $\overline{B(x,r)} \subset U$. Wir setzen $U_1 := B(f_0, r)$. Dann gilt $U_1 \subseteq E_U$. Sei weiterhin $C_1 := \sup_{v\in B(x,r)} \|X(v)\|$. Diese Zahl ist endlich, in der Tat gilt $C_1 < rC_X + \|X(x)\|_V$.

Lemma 9.14 *Wenn $\epsilon > 0$ so klein ist, daß $\epsilon C_1 < r/2$, dann gilt $\overline{A(U_1)} \subset U_1$.*

Beweis: Sei $f \in U_1$. Dann gilt

$$\begin{aligned}
\|f - f_0\|_E &= \sup_{t\in I} \|\int_0^t X(f(s))ds\|_V \\
&\leq t \sup_{s\in[0,t]} \|X(f(s))\|_V \\
&\leq \epsilon C_1 .
\end{aligned}$$

Folglich, $A(U_1) \subset B(f_0, \epsilon C_1) \subseteq B(f_0, r/2)$ und damit $\overline{A(U_1)} \subseteq B(f_0, r)$. ■

Ist also $\epsilon_0 > 0$ so klein, daß gleichzeitig $\epsilon C_1 < r$ und $\epsilon C_X < 1$ gelten, und ist $\epsilon \in (0, \epsilon_0)$, dann besitzt A genau einen Fixpunkt f in U_1.

Damit ist die Existenzaussage bewiesen. Wir beenden nun den Beweis der Eindeutigkeit. Die Eindeutigkeit beschränkter Integralkurven ist schon gezeigt, so sie in E_U liegen. Wir müssen ausschließen, daß es noch weitere unbeschränkte Integralkurven gibt. Sei $g_1 := (-\epsilon, \epsilon)$ eine Integralkurve. Dann ist für jedes $\delta \in (0, \epsilon)$ die Zahl $\sup_{|t| \le \epsilon - \delta} \|X(g(t))\| =: C_2(\delta))$ endlich. Es gilt für $|t| \le \epsilon - \delta$, daß

$$\|g(t) - x\| \le \epsilon C_2(\delta)/2 \ .$$

Damit ist $g_{|(-\epsilon+\delta, \epsilon-\delta)}$ beschränkt und folglich geich der Einschränkung $f_{|(-\epsilon+\delta, \epsilon-\delta)}$. Da δ beliebig klein sein darf, gilt $f = g$. ■

Ein zeitabhängiges Vektorfeld auf $U \subseteq V$ ist eine Abbildung $X : T \times U \to V$ für ein offenes Intervall $T \subseteq \mathbb{R}$, welches die Null enthält. Der Existenz- und Eindeutigkeitssatz gilt auch für zeitabhängige Vektorfelder, wenn man Lipschitzstetigkeit in der Ortsvariablen voraussetzt. Der Punkt ist, daß man Lipschitzstetigkeit in der Zeit nicht annehmen muß. Ansonsten könnte man das Problem in ein autonomes verwandeln und obigen Satz 9.11 anwenden. Diese Verallgemeinerung betrifft etwa eine Gleichung der Form

$$f'(t) = t^{1/3} f(t) \ , \quad f(0) = f_0 \ .$$

Unter einer Integralkurve mit Anfang in $x \in U$ versteht man eine Kurve $f : I \to U$ mit $f(0) = x$ und $f'(t) = X(t, f(t))$.

Theorem 9.15 *Wir nehmen an, daß X stetig ist, und daß es eine Konstante C gibt derart, daß $X(t, \dots) : U \to V$ Lipschitzstetig mit Konstanten C für alle $t \in T$ ist. Dann existiert ein $\epsilon_0 > 0$ derart, daß $(-\epsilon_0, \epsilon_0) \subseteq T$ und für jedes $\epsilon \in (0, \epsilon_0)$ genau eine auf $I := (-\epsilon, \epsilon)$ definierte Integralkurve $f : I \to U$ mit $f(0) = x$ gibt.*

Beweis: Der Beweis ist wie in 9.11. Man setzt

$$Af(t) = x + \int_0^t X(s, f(s)) ds \ .$$

■

Hier ist ein Beispiel, in welchem die lokale Eindeutigkeit nicht gilt. Sei $V := \mathbb{R}^1$ und $X : \mathbb{R}^1 \to \mathbb{R}^1$ durch $X(x) := \sqrt{|x|}$ gegeben. X ist nicht Lipschitz-stetig. Die Kurven $f_i : \mathbb{R} \to \mathbb{R}^1$, $i = 0, 1$, $f_0 \equiv 0$ und

$$f_1(t) := \begin{cases} 0 & t \le 0 \\ \frac{1}{4} t^2 & t > 0 \end{cases}$$

sind beides Integralkurven mit $f(0) = 0$.

9.6 Maximale Integralkurven

Sei $U \subseteq V$ offen, $x \in U$ und $X : U \to V$ ein Vektorfeld. Sei $0 \in J \subseteq \mathbb{R}$ ein offenes Intervall. Ist $g : J \to U$ eine Integralkurve des Vektorfeldes und $I \subseteq J$ ein offenes Teilintervall mit $0 \in I$, dann ist $g_{|I} : I \to U$ auch eine Integralkurve. Wir betrachten auf der Menge der Integralkurven des Vektorfeldes X mit Anfang x die folgenden partielle Ordnung.

Definition 9.16 *Es gilt $(f : I \to U) \leq (g : J \to U)$, falls $I \subseteq J$ und $f = g_{|I}$ ist.*

Theorem 9.17 *Unter den Voraussetzungen von Satz 9.11 gibt es* ***genau eine maximale Integralkurve*** *$f : I \to \mathbb{R}$ mit $f(0) = x$.*

Beweis: Die Existenz maximaler Integralkurven folgt aus folgender Betrachtung. Ist ist $(f_\alpha : I_\alpha \to U)_{\alpha \in L}$ eine Kette von Integralkurven, dann ist $g : \cup_{\alpha \in L} I_\alpha \to U$, $g(t) := f_\alpha(t)$ für geeignetes $\alpha \in L$ so daß $t \in I_\alpha$, eine obere Schranke der Kette. Nach dem Lemma von Zorn folgt daraus die Existenz maximaler Integralkurven.

Wir zeigen nun die Eindeutigkeit. Seien $(g : J \to U)$ und $(f : I \to U)$ beides maximale Integralkurven mit $g(0) = f(0) = x$. Wir betrachten die Menge

$$E := \{t \in \mathbb{R} \mid s \in I \cap J \text{ und } g(s) = f(s) \, \forall s \in [0, t)\} \ .$$

Sei $T := \sup E$ und $S := \inf E$, wobei wir $T = \infty$ und $S = -\infty$ zulassen. Wegen der lokalen Existenz und Eindeutigkeit der Integralkurven gilt offensichtlich $T > 0 > S$.

Wir zeigen, daß $J = I = (S, T)$. Klar ist daß $(S, T) \subseteq J \cap I$. Wäre diese Inklusion echt, dann können wir etwa annehmen, daß $T \in I \cap J$. Sei $y := f(T) = g(T)$, wobei die zweite Gleichheit aus der Stetigkeit der Integralkurven folgt. Dann gilt für alle genügend kleinen $\epsilon > 0$, daß $f_T(t) := f(t + T)$ und $g_T(t) := g(t + T)$ zwei verschiedene Integralkurven von X auf $(-\epsilon, \epsilon)$ mit $g_T(0) = f_T(0) = y$ definieren. Dies steht im Widerspruch zur lokalen Eindeutigkeit nach Satz 9.11.

Auf analoge Weise führt man $S \in I \cap J$ zum Widerspruch. Wir haben damit indirekt gezeigt, daß $I \cap J = (S, T)$. Wäre $I \neq J$, dann würde $I \subset I \cup J$ oder $J \subset I \cup J$ gelten. Wir definieren $h : I \cup J \to U$ durch

$$h(t) := \begin{cases} f(t) & t \in I \\ g(t) & t \in J \end{cases} \ .$$

Dann ist nach obiger Diskussion h wohldefiniert und eine Integralkurve von X mit Anfang in x. Diese setzt aber mindestens eine von f oder g echt fort, ein Widerspruch zur Maximalität von f und g. ■

Für diesen Satz haben wir nur die lokale Existenz und Eindeutigkeit von Integralkurven benutzt. Deshalb gilt er unter etwas allgemeineren Voraussetzungen, etwa wenn das Vektorfeld durch Umwandlung eines nicht-autonomen in ein autonomes System entsteht, und das nicht-autonome nur in der Ortsrichtung Lipschitzstetig (lokal gleichmäßig in der Zeit) ist.

9.7 Vollständigkeit

Sei $U \subseteq V$ eine offene Teilmenge eines Banachraums und X ein Lipschitz-stetiges Vektorfeld auf U.

Definition 9.18 1. *Das Vektorfeld X heißt* ***vollständig****, falls für jedes $x \in U$ und die maximale Integralkurve $f : I \to U$ von X mit $f(0) = x$ gilt : $I = \mathbb{R}$.*

2. *Das Vektorfeld X heißt* ***halbvollständig****, falls für jedes $x \in U$ und die maximale Integralkurve $f : I \to U$ von X mit $f(0) = x$ gilt : $\sup I = \infty$.*

Ist $I \subset \mathbb{R}$, dann sei $I_+ := [0, \infty) \cap I$. Analog definieren I_- als den negativen Teil von I.

Satz 9.19 *Sei $f : I \to U$ eine maximale Integralkurve und $t_\infty := \sup I < \infty$. Dann ist $\overline{f(I_+)} \cap U$ nicht kompakt.*

Beweis: Wir nehmen an, daß $\overline{f(I_+)} \cap U$ kompakt sei.

Lemma 9.20 *f setzt sich stetig auf $\bar{I}_+$ fort.*

Beweis: Sei $M := \sup_{v \in \overline{f(I_+)} \cap U} \|X(v)\|_V$. Es gilt $M < \infty$.

Sei (t_n) eine Folge in I_+ mit $t_n \to t_\infty$. Sei $y \in \overline{f(I_+)} \cap U$ ein Häufungspunkt von $(f(t_n))$. Wir müssen zeigen, daß $\lim_{t \to t_\infty} f(t) = y$. Es gilt

$$\begin{aligned} \|f(t) - y\|_V &\leq \|f(t) - f(t_n)\|_V + \|f(t_n) - y)\|_V \\ &= \| \int_{t_n}^t X(f(s)) ds\|_V + \|f(t_n) - y)\|_V \\ &\leq M|t - t_n| + \|f(t_n) - y)\|_V \ . \end{aligned}$$

Für gegebenes $\epsilon > 0$ wählen wir $n \in \mathbb{N}$ derart, daß

$$M|t_\infty - t_n| + \|f(t_n) - y)\|_V < \epsilon$$

gilt. Dann gilt $\|f(t) - y\|_V < \epsilon$, falls nur $t \in (t_n, t_\infty)$. ■

Sei $g : (-\epsilon, \epsilon) \to U$ eine Integralkurve von X mit $g(0) = y$. Wir definieren

$$h : I \cup [0, t_\infty + \epsilon) \to U$$

durch $h(t) := f(t)$ für $t \in I$ und $h(t) := g(t - t_\infty)$ für $t \in [t_\infty, t_\infty + \epsilon)$. Die Abbildung h ist stetig und auf $t \neq t_\infty$ differenzierbar und dort auch Integralkurve von X. Es gilt

$$\lim_{s \downarrow t_\infty} \frac{h(s) - h(t_\infty)}{s - t_\infty} = X(y) \ .$$

Auf der anderen Seite ist für $0 < s < t < t_\infty$

$$f(t) - f(s) = \int_s^t X(f(u))du \ .$$

Wir schließen daraus

$$f(t_\infty) - f(s) = \int_s^{t_\infty} X(f(u))du$$

und folglich

$$\lim_{s \uparrow t_\infty} \frac{h(s) - h(t_\infty)}{s - t_\infty} = \lim_{s \uparrow t_\infty} \frac{\int_s^{t_\infty} X(f(u))du}{s - t_\infty} = X(y) \ .$$

Also gilt $h'(t_\infty) = X(y)$. Damit ist h eine Integralkurve von X mit Anfang in x, welche $f : I \to U$ echt fortsetzt. Das steht im Widerspruch zur Maximalität von $f : I \to U$. ■

1. Wir betrachten das konstante Vektorfeld $X : V \to V$ mit dem Wert $\xi \in V$. Dieses ist vollständig. In der Tat ist $f(t) = x + t\xi$, $t \in \mathbb{R}$, die eindeutige maximale Integralkurve mit Anfang in $x \in V$.

2. Sei $U := V \setminus \{0\}$. Die Einschränkung von X auf U ist nicht mehr vollständig. So ist etwa die maximale Integralkurve $f(t) = -\xi + t\xi$ mit Anfang in $x := -\xi$ nur noch auf $(-\infty, 1)$ definert.

3. Das Vektorfeld $X(x, y) = (1, y^2) : \mathbb{R}^2 \to \mathbb{R}^2$ ist nicht vollständig. Die Integralkurve mit Anfang in $(0, c)$ ist durch $f(t) = (t, \frac{c}{1-ct})$ gegeben. Für $c > 0$ ist diese maximal auf dem Intervall $(-\infty, c^{-1})$ definiert.

4. Schränkt man das Vektorfeld aus der vorherigen Aufgabe auf $\{y < 0\}$ ein, so ist es immer noch Halbvollständig. Die Lösung mit Anfang in (s, c) ist $f(t) := (t - s, \frac{c}{1-c(t-s)})$ und auf $(c^{-1}(1 + sc), \infty)$ definiert.

9.8 Ljapunovfunktionen

Definition 9.21 *Ein Abbildung $L : M \to N$ zwischen topologischen Räumen M und N heißt eigentlich, falls für jede kompakte Teilmenge $K \subseteq N$ das Urbild $f^{-1}(K) \subseteq M$ kompakt ist.*

Sei $\bar{\mathbb{R}}_- := [-\infty, \infty)$.

Definition 9.22 *Eine* ***Ljapunovfunktion*** *für X ist eine differenzierbare Abbildung $L : U \to \mathbb{R}$ mit folgenden Eigenschaften.*

1. *Für alle $x \in U$ gilt die Ungleichung $dL(x)(X(x)) \leq 0$.*

2. *$L : U \to \bar{\mathbb{R}}_-$ ist eigentlich.*

Theorem 9.23 *Existiert für X eine Ljapunovfunktion, so ist X halbvollständig.*

Beweis: Sei $x \in U$ und $f : I \to U$ die maximale Integralkurve mit $f(0) = x$. Dann gilt

$$\begin{aligned} (f^*L)'(t) &= dL(f(t))(f'(t)) \\ &= dL(f(t))(X(f(t))) \\ &\leq 0 \end{aligned}$$

Wir schließen, daß $L(f(t)) \leq L(x)$. Also ist $f(I_+) \subseteq L^{-1}([-\infty, L(x))$, woraus die Kompaktheit von $\overline{f(I_+)} \cap U$ folgt. Also ist $\sup I_+ = \infty$. ■

Hier sind einige Beispiele.

1. Sei $(V, <.,.>)$ ein endlich-dimensionaler euklidischer Vektorraum, $U \subset V$ offen und $L : U \to \mathbb{R}$ differenzierbar mit Lipschitz-stetiger Ableitung und so, daß $L : U \to \bar{\mathbb{R}}_-$ eigentlich ist. Wir betrachten das Vektorfeld
$$X := -\mathtt{grad}(L) = -(\partial_1 X, \ldots, \partial_n X) \ .$$
Diese Vektorfeld ist halbvollständig. In der Tat ist L eine Ljapunovfunktion für X, da $dL(x)(X(x)) = - < \mathtt{grad}(L)(x), \mathtt{grad}(L)(x) > \leq 0$.

2. Sei $V = \mathbb{R}^n$ und $L(x) = x^2$. Dann ist $X(x) = -2x$. Dieses Feld ist halbvollständig. Ind er Tat sind die Teilmengen $\{L(x) \leq R\} = \{\|x\|^2 \leq R\}$ kompakt. Man kann die Lösung hier sogar explizit finden. Es gilt für die Integralkurve mit Anfang in x:
$$f(t) = \mathtt{e}^{-2t} x \ .$$
Diese Formel zeigt sogar die Vollständigkeit.

3. Sei L wie in 1. und $Y : U \to V$ ein weiteres Lipschitz-stetiges Feld mit der Eigenschaft $< X, Y > \geq 0$. Dann ist für jede Lipschitz-stetige Funktion $\phi : U \to [0, \infty)$ das Feld $Z := \phi X + Y$ halbvollständig. L ist Ljapunovfunktion von Z.

4. Sei $V := \mathbb{R}^2$ und
$$D := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
und ϕ wie in 3. Dann ist $Z(x) := -\phi(x)x + Dx$ halbvollständig.

9.9 Hamiltonsysteme

Sei $V := \mathbb{R}^{2n}$ und $U \subseteq V$ offen. Wir betrachten die alternierende Bilinearform $\omega(x, y) := \sum_{i=1}^n x_i y_{i+n} - x_{i+n} y_i$. Beachte, daß $\omega(x, x) = 0$ für alle $x \in V$.

Sei $H : U \to \mathbb{R}$ eine differenzierbare Funktion mit Lipschitz-stetigen Ableitungen.

Definition 9.24 *Das* ***Hamiltonsche Vektorfeld*** X_H *wird durch die Gleichung*

$$dH(x)(Y) = \omega(X_H(x), Y), \quad \forall Y \in V$$

charkterisiert.

Wir schreiben $x = (q, p)$ mit $p, q \in \mathbb{R}^n$. Dann ist $\omega(x, x') =< q, p' > - < p, q' >$. Es gilt

$$X_H(q, p) = (\partial_p H(q, p), -\partial_q H(q, p)) \ .$$

In der Tat ist nämlich mit dieser Definition und $Y = (U, V) \in \mathbb{R}^n \times \mathbb{R}^n$

$$\omega(X_H, Y) =< \partial_p H, V > - < -\partial_q H, U >= dH(Y) \ .$$

.

Theorem 9.25 *Wir nehmen an, daß die Mengen* $\{H = c\}$ *kompakt sind.*

1. *Für jede Integralkurve* $f : I \to V$ *von* X_H *gilt, daß* f^*H *konstant ist.*
2. X *ist vollständig.*

Beweis: Sei $x \in U$ und $f : I \to U$ die maximale Integralkurve mit $f(0) = x$. Dann gilt

$$\begin{aligned} (f^*H)'(t) &= dH(f(t))(f'(t)) \\ &= \omega(X_H(f(t)), X_H(f(t))) \\ &= 0 \ . \end{aligned}$$

Es gilt $f(I) \subset H^{-1}(\{H(x)\})$. Damit ist $\overline{f(I)} \cap U$ kompakt und folglich $I = \mathbb{R}$. ■

1. Sei beispielsweise $V := \mathbb{R}^2$ und $H(x, y) = -\frac{1}{2}(x^2 + y^2)$. Dann ist $X_H(x, y) = (y, -x)$ daß schon in 9.2 betrachtete Beispiel.
2. In der Mechanik wird $\mathbb{R}^{2n}$ als der Phasenraum eines n-dimensionalen Systems interpretiert. Eine Punkt $(q, p) \in \mathbb{R}^{2n}$ beschreibt einen Zustand mit Ort $q \in \mathbb{R}^n$ und Impuls $p \in \mathbb{R}^n$. Der Wert der Hamiltonfunktion ist die Energie dieses Zustandes. Für eine sich im Schwerefeld über der Erdoberfläche (Konstante g) bewegende Kugel der Masse m ist beispielsweise $n = 3$ und $H = \frac{1}{2m}\|p^2\| + mgq_3$. Damit ergibt sich das Hamiltonsche Vektorfeld

$$X_H(q, p) = (\frac{p_1}{m}, \frac{p_2}{m}, \frac{p_3}{m}, 0, 0, -mg) \ .$$

3. Allgemeiner sei $F = -\mathtt{grad} U$ ein durch ein Potential gegebenes Kraftfeld. Wir setzen
$$H(q,p) = \frac{1}{2m} p^2 + U(q) \ .$$
Dann ist die Hamiltongleichung
$$q' = \frac{p}{m} \ , \quad p' = -F(x) \ .$$
Ersetzt man $x := q$ und $v. = \frac{p}{m}$, dann erhält man die Newtonsche Bewegungsgleichung
$$x' = v \ , \quad v' = -\frac{1}{m} F(x) \ .$$
Die Newtonsche Mechanik ist allgemeiner, da man im Gegensatz zur Hamiltonschen auch Reibungskräfte beschreiben kann, so daß die Energie nicht notwendig konstant ist.

9.10 Lineares Wachstum

Theorem 9.26 *Sei V endlich dimensional und $X : V \to V$ ein Lipschitz-stetiges Vektorfeld. Dann ist X ist vollständig (der Punkt ist hier, daß X auf ganz V definiert ist).*

Beweis: Es reicht (nach Ortsverschiebung und Zeitumkehr), zu zeigen, daß für die maximale Integralkurve $f : I \to V$ mit $f(0) = 0$ die Relation $\sup I = \infty$ erfüllt ist. Es gilt
$$f(t) = \int_0^t X(f(s)) ds \ .$$
Wir schließen
$$\begin{aligned} \|f(t)\|_V &\leq \int_0^t \|X(f(s))\|_V ds \\ &\leq \int_0^t (\|X(f(s)) - X(0)\|_V + \|X(0)\|_V) ds \\ &= t\|X(0)\|_V + C_X \int_0^t \|f(s)\|_V ds \ . \end{aligned}$$

Lemma 9.27 (Gronwall-Lemma) *Seien $\alpha, \beta : [0,t] \to [0,\infty]$ stetige Funktionen, und erfülle $g : [0,t) \to \mathbb{R}$ die Abschätzung*
$$g(s) \leq \alpha(s) + \int_0^s \beta(u) g(u) du \ , \quad s \in [0,t) \ .$$
Dann gilt für alle $s \in [0,t)$ daß
$$g(s) \leq \alpha(s) + \int_0^s \alpha(u)\beta(u) \mathrm{e}^{\int_u^s \beta(v) dv} du \ .$$

Beweis: Wir setzen $f(s) := \int_0^s \beta(u)g(u)du$. Diese Funktion is differenzierbar und erfüllt

$$f'(s) - \beta(s)f(s) = \beta(s)g(s) - \beta(s)f(s) \leq \beta(s)g(s) - \beta(s)g(s) + \beta(s)\alpha(s) = \beta(s)\alpha(s) \ .$$

Wir setzen nun

$$h(s) = f(s)\mathsf{e}^{-\int_0^s \beta(u)du} \ .$$

Dann gilt

$$h'(s) = [f'(s) - f(s)\beta(s)]\mathsf{e}^{-\int_0^s \beta(u)du} \leq \beta(s)\alpha(s)\mathsf{e}^{-\int_0^s \beta(u)du} \ .$$

Durch Integration folgt

$$h(s) \leq \int_0^s \beta(u)\alpha(u)\mathsf{e}^{-\int_0^u \beta(v)dv}$$

und damit

$$f(s) \leq \int_0^s \beta(u)\alpha(u)\mathsf{e}^{\int_u^s \beta(v)dv} \ .$$

Daraus folgt die Behauptung. ∎

Wir wenden das Gronwall-Lemma an mit $g(t) = \|f(t)\|_V$, $\alpha(t) = t\|X(0)\|_V$ und $\beta(t) = C_X$. Es folgt

$$\|f(t)\|_V \leq t\|X(0)\|_V + \|X(0)\|_V \int_0^t uC_X\mathsf{e}^{(t-u)C_X}du =: C(t) \ .$$

Wäre nun $\sup I =: T < \infty$, so würde

$$\sup_{t\in I_+} \|f(t)\|_V \leq C(T) \ .$$

Also wäre $\overline{f(I_+)}$ kompakt. Das steht im Widerspruch zur Maximalität. ∎

1. Sei V ein endlich dimensionaler Vektorraum und $A \in \mathtt{End}(V)$. Wir definieren das lineare Vektorfeld $X(x) := A(x)$. Dieses ist Lipschitz-stetig. In der Tat gilt

   $$\|X(x) - X(y)\|_V = \|A(x-y)\|_V \leq \|A\|_{\mathtt{End}(V)}\|x-y\|_V \ ,$$

   womit $\|A\|_{\mathtt{End}(V)}$ eine Lipschitzkonstante von X ist. Sei $f_c : \mathbb{R} \to V$ die maximale Integralkurve mit $f(0) = c \in V$. Wir definieren $\mathsf{e}^{tA} \in \mathtt{End}(V)$ für $t \in \mathbb{R}$ durch die konvergente Reihe

   $$\mathsf{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} \ .$$

   Man sieht leicht durch Ableiten dieser Summe ein, daß

   $$\frac{d}{dt}\mathsf{e}^{tA} = A\mathsf{e}^{tA} \ .$$

   Damit ist $f_c(t) = \mathsf{e}^{tA}c$.

2. Wir hatten schon gesehen, daß

   $$\exp(t\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) = \begin{pmatrix} -\sin(t) & \cos(t) \\ -\cos(t) & -\sin(t) \end{pmatrix} \ .$$

9.11 Stetige und differenzierbare Abhängigkeit von Parametern

Sei (P, d_P) ein kompakter metrischer Raum und $(V, \|.\|_V)$ ein Banachraum. Wir betrachten ein vom Parameter $p \in P$ abhängendes Vektorfeld X auf einer offenen Teilmenge $U \subseteq V$, d.h. eine stetige Abbildung $X : P \times U \to V$.

Theorem 9.28 *Sei $x \in U$. Wir nehmen an, daß es eine Konstante $C_X > 0$ derart gibt, daß $X(p, .)$ Lipschitz-stetig für alle $p \in P$ mit Konstanten C_X ist. Dann existiert ein offenes Intervall $I \subset \mathbb{R}$ mit $0 \in I$ derart, daß für alle $p \in P$ eine eindeutig bestimmte Integralkurve $f_p : I \to U$ mit $f_p(0) = x$ existiert und die Abbildung $P \times I \ni (p, t) \mapsto f_p(t) \in V$ stetig ist.*

Beweis: Wir betrachten die Teilmenge $C_U(P, V) = \{Y \in C(P, V) \,|Y(P) \subseteq U\}$ des Banachraumes $C(P, V)$.

Lemma 9.29 *$C_U(P, V) \subseteq C(P, V)$ ist eine offene Teilmenge.*

Beweis: Sei $Y \in C_U(P, V)$. Da $Y(P)$ als stetiges Bild eines kompakten Raumes kompakt ist, gibt es ein $r > 0$ derart, daß aus $d(v, Y(P)) < r$ folgt $v \in U$. Dann ist aber $B(Y, r) \subseteq C_U(P, V)$. ■

Wir definieren nun das Vektorfeld $Z : C_U(V, P) \to C(V, P)$ durch

$$Z(Y)(p) := X(p, Y(p)) \in V \ .$$

In der Tat gilt

Lemma 9.30 *$Z(Y) \in C(P, V)$.*

Beweis: Die Abbildung $p \mapsto Z(Y)(p)$ ist stetig und damit auch beschränkt. ■

Lemma 9.31 *Das Vektorfeld Z ist Lipschitz-stetig.*

Beweis: Es gilt für $Y_i \in C_U(P, V)$, $i = 0, 1$, und beliebige $p \in P$, daß

$$\|Z(Y_0)(p) - Z(Y_1)(p)\|_V = \|X(p, Y_0(p)) - X(p, Y_1(p))\|_V \leq C_X \|Y_0(p) - Y_1(p)\|_V \ .$$

Daraus folgt

$$\|Z(Y_0) - Z(Y_1)\|_{C(P,V)} \leq C_X \|Y_0 - Y_1\|_{C(P,V)} \ .$$

Folglich ist C_X eine Lipschitzkonstante von Z. ■

Sei $Y_0 : P \to U$ die konstante Funktion mit dem Wert x. Es gilt $Y_0 \in C_U(P, V)$. Eine Anwendung von Satz 9.11 auf das Vektorfeld Z liefert ein Intervall $I \subset \mathbb{R}$ mit $0 \in I$ und die Existenz einer Integralkurve $Y : I \to C_U(P, V)$ mit $Y(0) = Y_0$.

Für $p \in P$ definieren wir nun $f_p : I \to U$ durch $f_p(t) := Y(t)(p)$. Es gilt $f_p(0) = Y(0)(p) = Y_0(p) = x$ und $f_p'(t) = Y'(t)(p) = X(p, Y(t)(p)) = X(p, f_p(t))$. Damit ist

$f_p : I \to U$ die gesuchte Integralkurve und $P \times I \ni (p,t) \mapsto f_p(t) \in V$ ist stetig. ■

Wir betrachten das Vektorfeld $X(p)(t) = p^{-1}$ auf $P \times U$ mit $U := (-1,1)$, $x := 0$, und $P_0 = [1,2]$ sowie $P_1 = (0,1]$. Die Integralkurve von $X(p)$ mit Anfang in 0 ist $f_p(t) := \frac{t}{p}$. Für $p \in P_0$ sind diese auf dem Intervall $(-1,1)$ definiert. Es gibt aber kein Intervall um 0, auf welchem die Integralkurven für alle $p \in P_1$ definiert sind. Diese Beispiel unterstreicht die Bedeutung der Annahme der Kompaktkeit des Parameterraumes.

Eine Modifikation dieses Arguments gibt die differenzierbare Abhängigkeit von Parametern. Sei $P \subseteq R^N$ eine kompakte Teilmenge welche $\overline{\mathtt{int} P} = P$ erfüllt. Damm kann man von differenzierbarkeit für Funktionen auf P sprechen. Wir machen folgende zusätzliche Annahmen.

1. $X : P \times U \to V$ ist stetig differenzierbar.
2. Die Abbildung $x \mapsto dX(p,x)$ ist für alle $p \in P$ Lipschitz-stetig mit Konstanten C_X.

Theorem 9.32 *Sei $x \in U$. Es existiert ein offenes Intervall $I \subset \mathbb{R}$ mit $0 \in I$ derart, daß für alle $p \in P$ eine eindeutig bestimmte Integralkurve $f_p : I \to U$ mit $f_p(0) = x$ existiert und die Abbildung $P \times T \ni (p,t) \mapsto f_p(t) \in V$ stetig differenzierbar ist.*

Beweis: Wir betrachten den Banachraum $C^1(P,V)$ der stetig differenzierbaren Abbildungen mit der Norm $\|Y\|_1 := \sup_{p\in P} \|Y(p)\| + \sup_{p\in P} \|dY(p)\|$. Darin betrachten wir die offene Teilmenge $C^1_U(P,V) := \{f \in C^1(P,V) \mid f \in C_U(P,V)\}$. Für $Y \in C^1_U(P,V)$, $p \in P$, definieren wir wieder $Z(Y)(p) := X(p, Y(p))$.

Lemma 9.33 $Z(Y) \in C^1(P,V)$.

Beweis: Es ist mit der Kettenregel klar, daß $Z(Y)$ stetig differenzierbar ist. ■

Sei $B(Y_0, r) \subset C^1_U(P,V)$ ein Ball vom Radius $r > 0$ um die konstante Abbildung mit dem Wert x.

Lemma 9.34 *Z ist auf $B(Y_0,1)$ Lipschitz-stetig.*

Beweis: Wir benutzen die Abschätzung Lemma 9.31. Sei $\sup_{p\in P} \|dX(p,x_0)\| =: C_1$ Dann gilt für $Y_1, Y_2 \in B(Y_0,r)$

$$\begin{aligned}
&\|dZ(Y_1)(p) - dZ(Y_2)(p)\| \\
&= \|d_1X(p,Y_1(p)) + d_2X(p,Y_1(p)) \circ dY_1(p) - d_1X(p,Y_2(p)) - d_2X(p,Y_2(p)) \circ dY_2(p)\| \\
&\le C_X\|Y_1(p) - Y_2(p)\| + \|d_2X(p,Y_1(p))\|\|dY_1(p) - dY_2(p)\| \\
&\quad + \|d_2X(p,Y_1(p)) - d_2X(p,Y_2(p))\|\|dY_2(p)\| \\
&\le C_X\|Y_1 - Y_2\|_1 + \|Y_1 - Y_2\|_1(\|d_2X(p,x)\| + \|d_2X(p,Y_1(p)) - d_2X(p,x)\| \\
&\quad + \|Y_2\|_1 C_X \|Y_1 - Y_2\|_1 \\
&\le \|Y_1 - Y_2\|_1(C_X + C_1 + C_X(\|Y_1\| + \|x\|) + C_X\|Y_2\|_1) \\
&\le \|Y_1 - Y_2\|_1(C_X + C_1 + C_X(\|x_0\| + r + \|x\|) + C_X(\|x\| + r))\ .
\end{aligned}$$

■

Nach Satz 9.11 finden wir nun eine Integralkurve $Y : I \to B(Y_0, r) \subset C^1(P, V)$ von Z. Folglich existiert die Ableitung df_p (nach p) und ist stetig. Weiterhin ist $f_p'(t) = X(p, f_p(t))$ stetig. Damit hat $(p, t) \mapsto f_p(t)$ stetige partielle Ableitungen und ist damit stetig differenzierbar. ■

Wenn $(p, t) \mapsto f_p(t)$ der Differentialgleichung

$$f_p'(t) = X(p, f_p(t))) \ , f_p(0) = x$$

genügt und nach p differenzierbar ist, dann gilt

$$d_p f_p'(t) = d_1 X(p, f_p(t)) + d_2 X(p, f_p(t)) d_p f_p(t) \ , d_p f_p(0) = 0 \ .$$

Die Abbildung

$$g_p := (f_p, d_p f_p) : P \times I \to V \oplus B(\mathbb{R}^n, V)$$

erfüllt also

$$g_p' = Z(p, g_p) \ , g_p(0) = (x, 0)$$

mit

$$Z(p, (v, \phi) = (X(p, v), d_1 X(p, v) + d_2 X(p, v) \circ \phi) \ .$$

Unsere Voraussetzungen sind so gestaltet, daß das Vektorfeld $X^{(1)} Z$ die Bedingungen von Theorem 9.28 erfüllt. Will man zweifach differenzierbare Abhängigkeit der Lösung von Parametern zeigen, dann muß man die Voraussetzungen von Theorem 9.32 für $X^{(1)}$ erfüllen. Definiert man

$$X^{(k)} := (\dots (X^{(1)})^{(1)} \dots)^{(1)} \ ,$$

(k-mal) und erfüllt dieses die Voraussetzungen von 9.28, dann erhalten für eine k-fach differenzierbare Abhängikeit von Parametern. Explizit sind das folgende zusätzliche Annahmen.

1. $X : P \times U \to V$ ist k-mal stetig differenzierbar.
2. Für alle $l \leq k$ und $p \in P$ ist die Abbildung $x \mapsto d^{(l)} X(p, x)$ Lipschitz-stetig mit Konstanten C_X.

Theorem 9.35 *Sei $x \in U$. Es existiert ein offenes Intervall $I \subset \mathbb{R}$ mit $0 \in I$ derart, daß für alle $p \in P$ eine eindeutig bestimmte Integralkurve $f_p : I \to U$ mit $f_p(0) = x$ existiert und die Abbildung $P \times I \ni (p, t) \mapsto f_p(t) \in V$ k-mal stetig differenzierbar ist.*

9.12 Abhängigkeit von Anfangsbedingungen

Wir betrachten einen endlich-dimensionalen Vektorraum V und eine offene Teilmenge $U \subseteq V$. Sei $X : U \to V$ ein Lipschitz-stetiges Vektorfeld mit Lipschitzkonstante C_X. Für $x \in U$ sei $\Phi(x) : I_x \to U$, $t \mapsto \Phi_t(x)$, die eindeutige maximale Integralkurve mit $\Phi_0(x) = x$.

Theorem 9.36 *Sei $x \in U$. Dann existiert eine Umgebung $W \subset U$ von x und ein Intervall $I \subset \mathbb{R}$ mit $0 \in I$ derart, daß $I \subset \cap_{v \in W} I_v$ und $\Phi : W \times I \to V$, $(w,t) \mapsto \Phi_t(w)$, stetig ist.*

Beweis: Sei $r > 0$ so daß $B(x, 4r) \subseteq U$. Wir setzen $P := \overline{B(x,r)}$. Da wir V als endlich-dimensional annehmen, ist P kompakt. Wir definieren das parameterabhängige Vektorfeld $Y : P \times B(x, 2r) \to V$ durch $Y(p, v) = X(v - p + x)$.

Lemma 9.37 *Es gibt eine Konstante C_Y derart, daß $Y(p, .)$ Lipschitz-stetig mit Konstanten C_Y für alle $p \in P$ ist.*

Beweis: In der Tat gilt für $p \in P$ und $v, w \in B(x, 2r)$ daß

$$\|Y(p,v) - Y(p,w)\|_V = \|X(v - p + x) - X(w - p + x)\| \leq C_X \|v - w\|_V \ .$$

Wir können also $C_Y := C_X$ wählen. ■

Sei $f_p : I \to B(x, 2r)$ die Integralkurve von Y mit $f_p(0) = x$. Dann ist $\Phi_t(p) := f_p(t) + p - x$ eine Integralkurve von X mit $\Phi_0(p) = p$. Die Abbildung $B(x,r) \times I \ni (p,t) \to \Phi_t(p) \in V$ ist stetig. ■

Wenn wir annehmen, daß X Lipschitz-stetige Ableitungen bis zur Ordnung k besitzt, dann erhalten wir den folgenden Satz.

Theorem 9.38 *Sei $x \in U$. Dann existiert eine Umgebung $W \subset U$ von x und ein Intervall $I \subset \mathbb{R}$ mit $0 \in I$ derart, daß $I \subset \cap_{v \in W} I_v$ und $W \times I \ni (x,t) \mapsto \Phi_t(x) \in V$ k-mal stetig differenzierbar ist.*

9.13 Flüsse

Sei V ein topologischer Raum.

Definition 9.39 *Ein **lokaler (Halb-) Fluß** auf V ist durch folgende Daten gegeben:*

1. *Eine offene Teilmenge $D \subseteq \mathbb{R} \times X$ (oder $D \subseteq [0, \infty) \times X$ für einen Halbfluß) mit $\{0\} \times V \subset V$.*

2. *Eine stetige Abbildung $\Phi : D \to V$, $(t, v) \mapsto \Phi_t(v)$.*

Dabei müssen folgende Bedingungen erfüllt sein.

1. $\Phi_0 = \mathtt{id}_V : V \to V$

2. für jedes $x \in V$ *ist* $\{t \in \mathbb{R} | (t, x) \in D\}$ *ein (offenes) Intervall in* $\mathbb{R}$ *(oder in* $[0, \infty)$ *für einen Halbfluß).*

3. Wenn $(s, \Phi_t(x)) \in D$ *und* $(t + s, x) \in D$*, dann gilt*

$$\Phi_{t+s}(x) = \Phi_s(\Phi_t(x)) \ .$$

Man kann auch Halbflüsse in der negativen Zeitrichtung betrachten.

Auf der Menge der lokalen (Halb-) Flüsse (D, Φ) führen wir die folgende Halbordnung ein:

Definition 9.40 *Es gilt* $(D, \Phi) \leq (D', \Phi')$ *genau dann, wenn* $D \subseteq D'$ *und* $\Phi = \Phi'_{|D}$.

Satz 9.41 *Zu jedem lokalen (Halb-)Fluß gibt es einen eindeutigen größeren maximalen lokalen (Halb-) Fluß.*

Beweis: Die Existenz zeigt man mit Hilfe des Lemmas von Zorn. Sei (D_α, Φ_α) eine aufsteigende Kette lokaler Flüsse. Dann setzen wir $D := \cup_\alpha D_\alpha$ und $\Phi_t(x) := \Phi_{\alpha,t}(x)$, wobei wir α so wählen, daß $(t, x) \in D_\alpha$ gilt. Dann ist (D, Φ) ein wohldefinierter lokaler Fluß und eine obere Schranke der Kette. Nach dem Lemma von Zorn gibt es also maximale lokale Flüsse. Ananlog zeigt man die Existenz maximaler Halbflüsse.

Wir zeigen jetzt die Eindeutigkeit. Seien (E, Ψ) und (F, Γ) beides maximale (Halb-) Flüsse welche größer als (D, Φ) sind.

Lemma 9.42 *Es gilt* $\Psi_{|E \cap F} = \Gamma_{|E \cap F}$.

Beweis: Wir führen den Beweis für Halbflüsse. Wir fixieren $x \in V$. Sei T die maximale Zahl derart, daß für alle $t \in (0, T)$ mit $(t, x) \in E \cap F$ auch $\Psi_t(x) = \Gamma_t(y)$ gilt. Wir nehmen nun an, daß es ein Paar $(s, x) \in E \cap F$ gibt mit $\Psi_s(x) \neq \Gamma_s(y)$. Dann ist sicher $T \leq s$ und damit $(x, T) \in E \cap F$. Folglich gilt aus Stetigkeitsgründen auch $\Psi_T(x) = \Gamma_T(y)$. Sei nun $\epsilon_0 > 0$ derart, daß $(T + \epsilon_0, x) \in E \cap F$ und $(\epsilon_0, \Psi_T(x)) \in D$. Dann gilt für alle $\epsilon < \epsilon_0$

$$\Psi_{T+\epsilon}(x) = \Psi_\epsilon(\Psi_T(x)) = \Gamma_\epsilon(\Gamma_T(x)) = \Gamma_{T+\epsilon}(x) \ .$$

Dies steht im Widerspruch zur Maximalität von T. Folglich gibt es die Zahl s mit der behaupteten Eigenschaft nicht. ∎

.

Lemma 9.43 *Seien* (E, Ψ) *und* (F, Γ) *maximale Halbflüsse, welche* (D, Φ) *fortsetzen. Dann gilt* $(E, \Psi) = (F, \Gamma)$.

Beweis: Wir definieren einen lokalen Halbfluß (G, Δ) durch

$$G := E \cup F \ , \quad \Delta_t(x) := \begin{cases} \Psi_t(x) & (t,x) \in E \\ \Gamma_t(x) & (t,x) \in F \end{cases} \ .$$

Dies ist nach Lemma 9.42 eine wohl-definierte stetige Abbildung.

Wir nehmen nun an, daß $(t+s, x) \in G$ und $(s, \Delta_t(x)) \in G$ ist. Wir müssen zeigen, daß $\Delta_s(\Delta_t(x)) = \Delta_{s+t}(x)$ gilt.

O.b.d.A gelte $(s+t, x) \in E$ und $s \geq 0$. Wenn auch $(s, \Delta_t(x)), (t, x) \in E$ gilt, dann haben wir

$$\Delta_{s+t}(x) = \Psi_{s+t}(x) = \Psi_s(\Psi_t(x)) = \Delta_s(\Delta_t(x)) \ .$$

Seien nun $(s, \Delta_t(x)) \in F$ und $(t, x) \in E$. Dann ist

$$\Delta_{t+s}(x) = \Psi_{s+t}(x) \ , \Delta_s(\Delta_t(x)) = \Gamma_s(\Psi_t(x)) \ .$$

Wenn $\Psi_{t+s}(x) \neq \Gamma_s(\Psi_t(x))$ gilt, dann existiert ein maximales $0 \leq T < s$ derart, daß $\Psi_{u+t}(x) = \Gamma_u(\Psi_t(x)$ für alle $0 \leq u \leq T$ gilt. Insbesondere gilt $\Psi_{T+t}(x) = \Gamma_T(\Psi_t(x)$. Wir wählen $\epsilon_0 > 0$ derart, daß $(\epsilon_0, \Delta_{T+t}(x)) \in D$, $(\epsilon_0 + T, \Delta_t(x)) \in F$ und $T + \epsilon_0 < s$ gilt. Für $0 \leq \epsilon \leq \epsilon_0$ gilt

$$\Psi_{T+\epsilon+t}(x) = \Psi_\epsilon(\Psi_{T+t}(x)) = \Gamma_\epsilon(\Gamma_T(\Psi_t(x)) = \Gamma_{T+\epsilon}(\Psi_t(x)) \ .$$

Das steht aber im Widerspruch zur maximalen Wahl von T. Folglich gilt $\Psi_{t+s}(x) = \Gamma_s(\Psi_t(x))$.

■

Sei (D, Φ) ein lokaler Fluß auf V. Wir setzen $D_+ := D \cap [0, \infty) \times V$ und $D_- := D \cap (-\infty, 0] \times V$. Die Einschränkungen $(D_\pm, \Phi)$ sind lokale Halbflüsse (im Falle D_- mit negativer Zeit). Dann erhalten wir eindeutige maximale Halbflüsse $(E^\pm, \Phi^\pm)$, welche $(D_\pm, \Phi)$ erweitern. Dann definieren wir $\Psi : E := E^+ \cup E^- \to V$ durch

$$\Psi_t(x) := \Phi_t^{\mathtt{sign}(t)}(x) \ .$$

Lemma 9.44 *(E, Ψ) ist ein lokaler Fluß.*

Beweis: Wir müssen $\Psi_{s+t}(x) = \Psi_s(\Psi_t(x))$ nachrechnen. Absehen von trivialen Fällen müssen wir etwa den Fall $t > 0$ und $s < 0$ und $s + t > 0$ betrachten Es gilt für alle $n > 0$, daß

$$\Psi_s(\Psi_t(x)) = (\Phi_{s/n}^{-,n}(\Phi_{t/n}^{+,n}(x))) \ .$$

Die Menge

$$A := \{\Psi_u(x) | s \leq u \leq t\}$$

is kompakt. Deshalb gibt es $\epsilon > 0$ derart, daß $(u, y) \in D$ für alle $y \in A$ und $|u| < \epsilon$.

Für genügend große n kann man also $\Phi^{+,n}_{t/n}$ durch $\Phi_{t/n}$ ersetzen. Analog kann man $\Phi^{-}_{s/n}$ durch $\Phi^{-1}_{-s/n}$ ersetzen. Es gilt also

$$\Psi_s(\Psi_t(x)) = (\Phi^n_{s/n}(\Phi^n_{t/n}(x))) = \Psi^n_{t+s/n}(x) = \Psi^+_{t+s}(x) \ .$$

■

In der Tat ist (E, Ψ) der eindeutige maximale lokale Fluß, welcher (D, Φ) fortsetzt. Aus der Maximalität von (E, Ψ) folgt nach diesem Lemma nämlich die Maximalität der Einschränkungen $(E^{\pm}, \Psi^{\pm})$, welche aber eindeutig sind. ■

Definition 9.45 *Eine lokaler (Halb-) Fluß ist ein* ***(Halb-) Fluß*** *auf V, wenn sein Definitionsbereich $\mathbb{R} \times V$ (bzw. $[0, \infty) \times V$ ist.*

Ein Fluß auf X ist also einfach durch eine stetige Abbildung $\mathbb{R} \times V \ni (t, v) \to \Phi_t(v) \in V$ gegeben, wobei $\Phi_0 = \mathtt{id}_V$ und $\Phi_{s+t} = \Phi_s \circ \Phi_t$ für alle $t, s \in \mathbb{R}$ gilt.

Sei $U \subseteq V$ eine offene Teilmenge eines endlich-dimensionalen Vektorraumes und $X : U \to V$ ein Lipschitz-stetiges Vektorfeld. Für jedes $x \in U$ finden wir ein $r_x > 0$ derart und ein Intervall $I_x \subseteq \mathbb{R}$ derart, daß für jedes $p \in B(x, r_x)$ die Integralkurve $\Phi_t(p)$ mit Anfang in p auf I_x definiert ist. Wir betrachten nun die offene Teilmenge $D := \cup_{x \in U} I_x \times B(x, r_x) \subseteq \mathbb{R} \times U$ und die induzierte Abbildung $\Phi : D \to U$.

Lemma 9.46 $\Phi : D \to U$ *ist ein lokaler Fluß.*

Beweis: In der Tat ist Φ nach Satz 9.36 stetig. Wenn $s < t$ und $(t, x), (s, x) \in D$, dann gilt auch $(u, x) \in D$ für alle $u \in (s, t)$. Weiter, wenn $(t + s, x) \in D$ und $(s, \Phi_t(x)) \in D$ ist, dann sind $u \mapsto \Phi_u(\Phi_t(x))$ und $u \mapsto \Phi_{u+t}(x)$ beides Integralkurven mit Anfang in $\Phi_t(x)$ und damit gleich. Folglich gilt $\Phi_{s+t}(x) = \Phi_s(\Phi_t(x))$. ■

Durch Einschränkung auf $[0, \infty) \times U \cap D$ erhält man einen Halbfluß.

Definition 9.47 *Der* ***von X erzeugte lokale (Halb-) Fluß*** *ist ein maximaler lokaler (Halb)- Fluß zu dem oben konstruierten Halbfluß.*

Ist X mindestens k-mal differenzierbar und sind die Ableitungen von X bis zur Ordnung k Lipschitz-stetig, so gilt die folgende Aussage für den von X erzeugten lokalen Fluß $\Phi : D \to U$:

Lemma 9.48 *Für jedes $\Phi : D \to U$ ist k-mal stetig differenzierbar.*

Beweis: Das folgt aus Satz 9.38. ■

Umgekehrt definiert jeder lokale Fluß, für welchen $\Phi : D \to U$ differenzierbar ist, ein Vektorfeld durch

$$X(x) := \frac{d}{dt}_{|t=0} \Phi_t(x) \; .$$

Die Kurve $t \mapsto \Phi_t(x)$ ist eine Integralkurve mit Anfang in x. In der Tat gilt

$$\begin{aligned} \frac{d}{dt}\Phi_t(x) &= \frac{d}{ds}_{s=0} \Phi_{t+s}(x) \\ &= \frac{d}{ds}_{|s=0} \Phi_s(\Phi_t(x)) \\ &= X(\Phi_t(x)) \end{aligned}$$

Korollar 9.49 *Insbesondere ist also jeder differenzierbare lokale Fluß durch ein Vektorfeld erzeugt. (Halb)-Vollständige Vektorfelder entsprechen dabei (Halb-)Flüssen.*

9.14 Lineare Vektorfelder

Sei V endlich dimensional.

Definition 9.50 *Eine lineares Vektorfeld auf V ist ein Vektorfeld der Form $X(x) = Ax$ mit $A \in \mathtt{End}(V)$ (wird durch X eindeutig bestimmt).*

Aus Satz 9.26 folgt die Vollständigkeit von X. Wir hatten auch schon gesehen, daß der durch X erzeugte Fluß durch

$$\Phi_t := \exp(tA)$$

beschrieben werden kann.

Sei $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ die Komplexifizierung von V. Wir können V als die Teilmenge der reellen Vektoren auffassen. Sei $A_{\mathbb{C}} = A \otimes \mathtt{id}$ die Ausdehnung von A auf $V_{\mathbb{C}}$. Es gilt

$$\begin{aligned} \Phi_t(v) &= \exp(tA)v \\ &= \exp(tA_{\mathbb{C}})v \\ &= \mathtt{Re}(\exp(tA_{\mathbb{C}}))v \; , \end{aligned}$$

folglich

$$\Phi_t = \mathtt{Re}(\exp(tA_{\mathbb{C}})) \; .$$

Eine Konsequenz der Jordanschen Normalform ist die Existenz einer Zerlegung $A_{\mathbb{C}} = S + N$, wobei S halbeinfach ist (d.h. eine Diagonalisierung zuläßt), N nilpotent ist (d.h. $N^p = 0$ für $p >> 0$ gilt), und $[S, N] = 0$ gilt.

Lemma 9.51 *Ist W ein endlich-dimensionaler Vektorraum und sind $L, M \in \mathtt{End}(V)$ mit $[L, M] = 0$, dann gilt $\exp(L + M) = \exp(L)\exp(M)$.*

Beweis: Der Beweis geht wie im skalaren Fall. ■
Wir schreiben also

$$\Phi_t = \mathtt{Re}(\exp(tA_{\mathbb{C}})) = \mathtt{Re}(\exp(tS)\exp(tN)) \ .$$

Die Abbildung $\exp(tS)$ wird wie folgt berechnet. Wir wählen eine Basis, in welcher für die Matrix M_S von S

$$M_S = \mathtt{diag}(\lambda_1, \dots, \lambda_n)$$

gilt. Dann ist

$$M_{\exp(tS)} = \mathtt{diag}(\mathtt{e}^{t\lambda_1}, \dots, \mathtt{e}^{t\lambda_n}) \ .$$

Weiterhin ist

$$\exp(tN) = \sum_{p=0}^{\infty} \frac{t^p N^p}{p!}$$

eine endliche Summe und damit ein Polynom in t.

Im letzten Schritt multiplizieren wir die Ergebnisse und bilden den Realteil.

1. Sei $V := \mathbb{R}^2$,

   $$D := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} ,$$

   und $X(x) := Dx$. Wir haben $V_{\mathbb{C}} = \mathbb{C}^2$. $D_{\mathbb{C}}$ ist halbeinfach, und in der Basis $((1,i)^t, (1,-i)^t)$ gilt $M_D = \mathtt{diag}(i, -i)$. Also ist $M_{\exp(tD_{\mathbb{C}})} = \mathtt{diag}(\mathtt{e}^{it}, \mathtt{e}^{-it})$. Folglich

   $$\exp(tD_{\mathbb{C}}) = \begin{pmatrix} \frac{\mathtt{e}^{it}+\mathtt{e}^{it}}{2} & \frac{\mathtt{e}^{it}-\mathtt{e}-it}{2i} \\ \frac{i\mathtt{e}^{it}-i\mathtt{e}^{it}}{2} & \frac{\mathtt{e}^{it}+\mathtt{e}^{-it}}{2} \end{pmatrix} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} .$$

   Dies hatten wir schon in Gleichung (4) gesehen.

2. Das in 1. behandelte Vektorfeld kann als von der gewöhnlichen Differentialgleichung

   $$h'' = -h$$

   kommend verstanden werden. Diese Differentialgleichung beschreibt eine ungedämpfte Schwingung. Die gedämpfte Schwingung wird durch

   $$h'' = -h - 2dh'$$

   modelliert. Dazu gehört das Vektorfeld $X(x) := Ax$ mit

   $$A := \begin{pmatrix} 0 & 1 \\ -1 & -2d \end{pmatrix} .$$

   Diese Matrix ist für $d \neq 1$ halbeinfach mit den Eigenwerten

   $$\lambda_\pm := -d \pm i\sqrt{1-d^2}$$

und den Eigenvektoren

$$u_\pm := (1, \lambda_\pm) \ .$$

Damit ist

$$\exp(tA_{\mathbb{C}}) = \begin{pmatrix} \frac{\lambda_- \mathrm{e}^{t\lambda_+} - \lambda_+ \mathrm{e}^{t\lambda_-}}{\lambda_- - \lambda_+} & \frac{1}{\lambda_+ - \lambda_-}(\mathrm{e}^{t\lambda_+} - \mathrm{e}^{t\lambda_-}) \\ \frac{\lambda_+\lambda_-}{\lambda_- - \lambda_+}(\mathrm{e}^{t\lambda_+} - \mathrm{e}^{t\lambda_-}) & \frac{1}{\lambda_+ - \lambda_-}\lambda_+ \mathrm{e}^{t\lambda_+} - \lambda_- \mathrm{e}^{t\lambda_-} \end{pmatrix} \ .$$

Wenn $d > 1$, dann ist diese Matrix reell. Wenn $d < 1$, so gilt mit $\omega := \sqrt{1 - d^2}$

$$\mathtt{Re}(\exp(tA_{\mathbb{C}})) = \begin{pmatrix} \frac{d}{\omega}\mathrm{e}^{-dt}\sin(\omega t) + \mathrm{e}^{-dt}\cos(\omega t) & \frac{1}{\omega}\sin(\omega t) \\ * & * \end{pmatrix} \ ,$$

wobei sich die zweite Zeile als Ableitung der ersten ergibt.

3. Im Fall $d = 1$ ist die Matrix $A_{\mathbb{C}}$ nicht halbeinfach. Wir erhalten

$$\exp(tA_{\mathbb{C}}) = \begin{pmatrix} t\mathrm{e}^{-t} + \mathrm{e}^{-t} & t\mathrm{e}^{-t} \\ * & * \end{pmatrix}$$

(die zweite Zeile ergibt sich wieder durch Ableiten).

Die Lösungsmenge der Differentialgleichung

$$f' = Af$$

ist ein Vektorraum der Dimension n. In der Tat lösen mit f_0 und f_1 auch $f_0 + \lambda f_1$ diese Gleichung. Die Abbildung $f \mapsto f(0)$ gibt einen Isomorphismus des Lösungsraumes mit V.

Sei jetzt $R : \mathbb{R} \to V$ eine stetige Abbildung. Wir wollen das System linearer Differentialgleichungen

$$f' = Af + R \ , f(0) = f_0$$

lösen. Eine explizite Methode ist die der **Variation der der Konstanten**. Die homogene Gleichung

$$g' = Ag \ , g(0) = f_0$$

wird bekanntlich durch $g(t) := e^{tA} f_0$ gelöst. Wir machen nun den Ansatz

$$f(t) = e^{tA}C(t) \ , C(0) := f_0 \ .$$

Durch Ableiten erhalten wir

$$f'(t) = Ae^{tA}C(t) + e^{tA}C'(t) \ .$$

Setzen wir dies in die Differentialgleichung ein, dann erhalten wir die Gleichung

$$Ae^{tA}C(t) + e^{tA}C'(t) = Ae^{tA}C(t) + R(t) \ ,$$

also
$$C'(t) = e^{-tA}R(t) \ .$$
Die Lösung ist also
$$C(t) := f_0 + \int_0^t e^{-sA}R(s)ds \ .$$
Die Lösungsformel für unsere inhomogene Differentialgleichung ist damit
$$f(t) = e^{tA}f_0 + \int_0^t e^{(t-s)A}R(s)ds \ .$$

Apriori ist klar, daß der Raum der Lösungen der inhomogenen Gleichung $f' = Af + R$ ein affiner Raum über dem Raum der Lösungen der inhomogenen Gleichung ist. Man bekommet alle Lösungen der inhomogenen Gleichung aus der speziellen Lösung
$$t \mapsto \int_0^t e^{(t-s)A}R(s)ds$$
durch Addition von Lösungen der homogenen Gleichung.

Es ist etwas umständlich, eine **lineare Differentialgleichung n-ter Ordnung**
$$f^{(n)} = \sum_{i=0}^{n-1} a_i f^{(i)} + r(t)$$
in ein System erster Ordnung zu übersetzen. Im Prinzip ist das möglich und zeigt, daß wir die Lösung der homogenen Gleichung in der (komplexen) Form
$$f(t) = \sum_{j=1}^{r} \sum_{i=0}^{n_j-1} p_{j,i} t^i e^{\lambda_j t}$$
ansetzen können.

Wir bestimmen zuerst den n-dimensionalen Lösungsraum der homogenen Gleichung. Setzt man $e^{t\lambda}$ in die Differentialgleichung ein, dann ergibt sich die Bestimmungsgleichung $Q(\lambda) = \lambda^{n-1} - \sum_{i=0}^{n} a_i \lambda^i$. Die Zahlen λ_i sind also die Nullstellen des **charakteristischen Polynoms** $Q \in \mathbb{R}[\lambda]$. Es stellt sich heraus, daß die Zahlen n_j ihre Vielfachheiten sind. In der Tat, setzt man $t^f e^{t\lambda}$, dann ergibt sich $(\frac{d}{d\lambda})^f Q(\lambda) = 0$ (indem man $t^f e^{t\lambda} = (\frac{d}{d\lambda})^f e^{t\lambda}$ schreibt). Es gilt
$$n = \deg(Q) = \sum_{j=1}^{r} n_j \ .$$

Die n Funktionen $(t^f e^{t\lambda_j})_{j=1\ldots,r,0\leq f\leq n_j-1}$ liefern eine Basis der Lösungsraumes der homogenen Gleichung. Der Ansatz für die Variation der Konstanten ist in diesem Fall
$$f(t) = \sum_{j=1}^{r} \sum_{i=0}^{n_j-1} p_{j,i}(t) t^i e^{\lambda_j t}$$

und führt auf explit durch Integration lösbare Gleichungen für die Funktionen $p_{j,i}(t)$.

Wir betrachten wieder die gedämpfte Schwingungsgleichung

$$f^{(2)} = -\omega^2 f - 2\epsilon f' + r(t)$$

mit $\omega^2 > \epsilon^2$. Das charakteristische Polynom ist

$$\lambda^2 + \omega^2 + 2\epsilon\lambda = 0$$

und hat die Nullstellen

$$\lambda_\pm := -\epsilon \pm i\sqrt{\epsilon^2 - \omega^2} \ .$$

Wir setzen $\alpha := \sqrt{\epsilon^2 - \omega^2}$. Die allgemeine relle homogene Lösung ist also

$$e^{-\epsilon t}(p\sin(\alpha t) + q\cos(\alpha t)) \ .$$

Eine spezielle Lösung der inhomogene Gleichung findet man durch den Ansatz

$$f(t) = p(t)e^{-\epsilon t}\sin(\alpha t) + q(t)e^{-\epsilon t}\cos(\alpha t) \ .$$

Setzt man dies in die Gleichung ein, dann erhält man allerdings eine recht komplizierte Bestimmungsgleichung für p und q.

9.15 Lineare nicht-automome Gleichungen

Wir betrachten zuerst den eindimensionalen Fall erster Ordnung

$$f' = af \ ,$$

wobei a eine stetige Funktion in t ist. Wir probieren folgende Umschreibung

$$a = \frac{f'}{f} = (\ln(f))' \ .$$

Durch Integration ergibt sich

$$\int a(t)dt = \ln(f(t)) \ .$$

Ist A eine Stammfunktion von a, dann löst offensichtlich $f(t) = e^{A(t)}$ die Differentialgleichung. In der Tat ist

$$f'(t) = A'(t)e^{A(t)} = ae^{A(t)} = a(t)f(t) \ .$$

Der Raum der Lösungen der Differentialgleichung ist ein 1-dimensionaler Vektorraum, welcher durch $e^{A(t)}$ aufgespannt wird. Hier sind einige Beispiele

1. Die Lösung des Anfangswertproblems

   $$f' = tf \ , \quad f(0) = 1$$

   ergibt sich also durch

   $$f(t) = Ce^{\frac{t^2}{2}}, C := 1 \ .$$

2. Die Lösung des Anfangswertproblemes

$$f'(t) = \sin(t)f(t)\ , f(0) = 1$$

ist

$$f(t) = C\mathtt{e}^{-\cos(t)}\ , C = \mathtt{e}^1\ .$$

Sei nun $t \to A(t)$ eine Lipschitzstetige $n \times n$-matrixwertige Funktion. Wir betrachten das System

$$f' = Af\ .$$

Der Lösungsraum ist ein n-dimensionaler Vektorraum, $f \mapsto f(t)$ ein Isomorphismus für jedes $t \in \mathbb{R}$. Die Lösung kann in der Form $e^{\int A(t)dt}f(0)$ geschrieben werden, wenn $[A(s), A(t)] = 0$ für alle $t, s \in \mathbb{R}$ gilt. Ist diese Bedingung nicht erfüllt, dann ist die Situation komplizierter. Auf jeden Fall gibt es eine eindeute Abbildung

$$\mathbb{R} \ni t \mapsto \Phi(t) \in Gl(n) \subset \mathtt{Mat}(n, n)$$

derart, daß

$$f(t) = \Phi(t)f_0$$

das Anfangswertproblem mit Anfangswert f_0 löst. Nach Übersetzung in ein autonomes System ergibt sich das Vektorfeld $X(t, x) = (1, A(t))$. Dieses ist auf allen Teilmengen $(-A, A) \times \mathbb{R}^n$ lipschitzstetig und erzeugt einen eindeutig bestimmten maximalen lokalen Fluß, welcher auf (t, x) für alle $s \in (-A - t, A - t)$ definiert ist. Diesen kann man zu einem eindeutig bestimmten Fluß auf ganz $\mathbb{R} \times \mathbb{R}^n$ fortsetzen. Die *Lipschitzstetigkeit von A ist nicht notwendig*, da man den Beweis des Existenz-und Eindeutigkeitssatzes auf zeitabhängige lipschitzstetige Vektorfelder verallgemeinern kann und sich so die Übersetzung in ein autonomes System spart. Dies betrifft etwa eine Gleichung der Form

$$f'(t) = |t|^{1/3}f(t)\ .$$

Definition 9.52 *Die Abbildung $\Phi : \mathbb{R} \to \mathtt{Mat}(n, n)$ heißt **Fundamentalmatrix** der Gleichung $f' = Af$.*

In der Tat ist Φ selbst Lösung des Anfangswertproblemes

$$\Phi' = A\Phi\ , \quad \Phi(0) = \mathtt{id}\ .$$

Die Funktion $\det(\Phi(t))$ erfüllt die Gleichung

$$\det(\Phi(t))' = \det(\Phi(t))\mathtt{Tr}(\Phi') = \det(\Phi(t))\mathtt{Tr}(\Phi^{-1}(t)A)$$

Damit gilt wegen $\det(\Phi(0)) = 1$, daß

$$\det(\Phi(t)) = e^{\int_0^t \mathtt{Tr}(\Phi^{-1}(s)A)}\ .$$

Ein wichtiger Fall ist der periodischer Koeffizienten. Wir betrachten zum Beispiel die parametrische Schwingungsgleichung

$$f^{(2)} = (\sin(\omega t) - b^2)f \ .$$

Diese lineare Differentialgleichung 2ter Ordnung mit zeitabhängigen Koeffizienten überführen wir in ein System erster Ordnung

$$u' = \begin{pmatrix} 0 & 1 \\ \sin(\omega t) - b^2 & 0 \end{pmatrix} u \ .$$

Eine direkte explizite Lösung ist schwierig, man kann aber **Floquet-Theorie** anwenden.

Sei $A : \mathbb{R} \to \mathtt{Mat}(n,n)$ stetig und periodisch mit Periode $p > 0$ und $\Phi : \mathbb{R} \to \mathtt{Mat}(n,n)$ die Fundamentalmatrix der Gleichung $f' = Af$.

Satz 9.53 *Es gibt eine stetige p-periodische Abbildung* $Z : \mathbb{R} \to \mathtt{Mat}(n,n,\mathbb{C})$ *und eine Matrix* $R \in \mathtt{Mat}(n,n,\mathbb{C})$ *derart, daß* $\Phi(t) = Z(t)\mathtt{e}^{tR}$ *gilt.*

Beweis: Die Abbildung $\Psi(t) := \Phi(t+p)$ löst auch die Gleichung $\Psi' = A\Psi$, allerdings mit der Anfangsbedingung $\Psi(0) = \Phi(p)$. Damit gilt $\Psi(t) = \Phi(t)\Phi(p)$. Wir nehmen vorerst die Existenz einer Matrix R an mit $\mathtt{e}^{pR} = \Phi(p)$. Dann definieren wir $Z(t) := \Phi(t)\mathtt{e}^{-tR}$. Die Gleichung

$$\Phi(t) = Z(t)\mathtt{e}^{tR}$$

ist damit erfüllt. Weiterhin gilt

$$Z(t+p) = \Phi(t+p)\mathtt{e}^{-(t+p)R} = \Phi(t)\Phi(p)\mathtt{e}^{-(t+p)R} = \Phi(t)\mathtt{e}^{tR} = Z(t) \ .$$

Die Existenz von R folgt aus folgendem Lemma:

Lemma 9.54 *Ist* $C \in \mathtt{Mat}(n,n,\mathbb{C})$ *invertierbar, dann existiert* $R \in \mathtt{Mat}(n,n,\mathbb{C})$ *mit* $C = \mathtt{e}^R$.

Beweis: Wir können annehmen, daß C in Jordan-Normalform vorliegt. Es genügt weiter, die Behautung für ein Jordankästchen einzusehen, also eine Matrix der Form $C = \lambda + J$, wobei $J_{ij} = 0$ falls $j \neq i+1$, und $J_{i,i+1} = 1$. Wir definieren $R = \log(\lambda) - \sum_{i=1}^{\infty} \frac{(-1)^k J^k}{k\lambda^k}$. Diese Summe ist endlich. Dann gilt $C = \mathtt{e}^R$ aufgrund der Identität von Potenzreihen

$$1 + t = \mathtt{e}^{\log(1+t)} \ , \quad \sum_{n=0}^{\infty} \frac{1}{n!} \left[-\sum_{k=1}^{\infty} \frac{(-t)^k}{k} \right]^n \ .$$

■

Die Matrix R ist nicht eindeutig bestimmt. In der Tat gilt etwa $\mathtt{e}^{pR} = \mathtt{e}^{pR + \frac{2\pi i}{p}}$. Aber die Realteile der Eigenwerte von R sind bestimmt. Ist $f \in \mathbb{C}^n$ ein verallgemeinerter Eigenvektor von R zu Eigenwert λ, dann gilt für alle $\epsilon > 0$

$$\lim_{t\to\infty} \|\Phi(t)f\|\mathtt{e}^{-(\mathtt{Re}(\lambda)+\epsilon)t} = 0 \ , \quad \lim_{t\to\infty} \|\Phi(t)f\|\mathtt{e}^{-(\mathtt{Re}(\lambda)-\epsilon)t} = \infty \ .$$

In der Tat gilt nämlich für geeignetes $c > 0$ (wähle $c := \sup_{t\in[0,p]} \mathtt{max}\{\|Z(t)\|, \|Z(t)^{-1}\|\}$)

$$c^{-1}\|\Phi(t)f\| < \|\mathsf{e}^{tR}f\| < c\|\Phi(t)f\| \ .$$

Definition 9.55 *Die Eigenwerte von R heißen Floquet-Exponenten des Systems $f' = Af$.*

Die Floquet-Exponenten bestimmen das Wachstum der Lösungen der Gleichung $f' = Af$ für lange Zeiten. Um das Langzeitverhalten zu bestimmen, müssen wir die Gleichung $\Phi' = A\Phi$ also nur über eine Periode p lösen (z.B. mit numerischen Methoden).

Die inhomogene Gleichung

$$f' = Af + R$$

kann wieder durch Variation der Konstanten gelöst werden.

Der Ansatz $f(t) = \Phi(t)C(t)$ führt auf

$$C(t) = \int_0^t \Phi(s)^{-1}R(s)ds \ .$$

Damit ist die Lösungsformel für das Anfangswertproblem der inhomogenen Gleichung

$$f(t) = \Phi(t)f_0 + \Phi(t)\int_0^t \Phi(s)^{-1}R(s)ds \ .$$

1. Die Lösung des Anfangswertproblems

   $$f' = tf + t \ , \quad f(0) = 1$$

   ergibt sich also durch

   $$f(t) = e^{\frac{t^2}{2}} + e^{\frac{t^2}{2}}\int_0^t e^{-\frac{s^2}{2}}sds = e^{\frac{t^2}{2}} + e^{\frac{t^2}{2}}(1 - e^{-\frac{t^2}{2}}) = 2e^{\frac{t^2}{2}} - 1 \ .$$

2. Die Lösung des Anfangswertproblemes

   $$f'(t) = \sin(t)f(t) + e^{-\cos(t)} \ , f(0) = 1$$

   ist

   $$f(t) = \mathsf{e}^{1-\cos(t)} + \mathsf{e}^{1-\cos(t)}\int_0^t \mathsf{e}^{\cos(s)-1}e^{-\cos(s)}ds = \mathsf{e}^{1-\cos(t)} + t\mathsf{e}^{-\cos(t)} \ .$$

Die nicht-autonome homogene lineare Differentialgleichung n-ter Ordnung

$$f^{(n)} = \sum_{i=0}^{n-1} a_i f^i$$

mit stetigen Funktionen $t \mapsto a_i(t)$ kann auf ein nicht-autonomes System erster Ordnung zurückgeführt werden. Wir erhalten wieder einen n-dimensionalen Vektorraum von Lösungen. Die Abbildung $f \mapsto (f(0), f'(0), \ldots, f^{n-1}(0))$ liefert einen Isomorphismus des Lösungsraumes mit $\mathbb{R}^n$.

9.16 Einige spezielle explizit lösbare Gleichungen

Eine Differentialgleichung der Form

$$f'(t) = \frac{h(t)}{g(f)}$$

kann man durch **Trennung der Variablen** explizit lösen. Wir schreiben

$$g(f)f'(t) = h(t) \ .$$

Sei G eine Stammfunktion g und H eine Stammfunktion von h. Dann erfüllt eine Lösung die Bedingung $G(f(t)) = H(t) + C$. Folglich ist $f(t) = G^{-1}(H(t) + C)$. Die Konstante C wird durch die Anfangsbedingung bestimmt.

1. $$f' = a(t)f$$

   Hier ist $h(t) := a(t)$ und $g(x) = \frac{1}{x}$, also $H(t) := \int_0^t a(s)ds$ und $G(x) := \ln(|x|)$ und damit $G^{-1}(y) = \pm \mathrm{e}^y$. Wir schließen, daß

   $$f(t) = \pm \mathrm{e}^{\int_0^t a(s)ds + C} = C' \mathrm{e}^{\int_0^t a(s)ds} \ ,$$

   unsere schon bekannte Lösungsformel.

2. $$f'(t) = t^m f^n(t) \ , n > 1$$

   Hier ist $h(t) = t^m$ und $H(t) := \frac{1}{m+1} t^{m+1}$ sowie $g(x) = x^{-n}$ und $G(x) = \frac{1}{-n+1} x^{-n+1}$. Wir schließen $G^{-1}(y) = ((-n+1)y)^{\frac{1}{-n+1}}$. Die Lösungsformel ergibt

   $$f(t) = (\frac{1-n}{m+1} t^{m+1} + C)^{\frac{1}{1-n}} \ .$$

Die **Bernoullische Differentialgleichung** hat die Form

$$f' = a(t)f(t) + b(t)f(t)^\rho \ .$$

Wir substituieren

$$g(t) := f(t)^{1-\rho}$$

und erhalten

$$g'(t) = (1-\rho)f'(t)f(t)^{-\rho} = (1-\rho)(a(t)f(t)+b(t)f(t)^\rho)f(t)^{-\rho} = (1-\rho)a(t)g(t)+(1-\rho)b(t) \ .$$

Diese Gleichung ist linear. Ein Beispiel ist die die **logistische Gleichung**

$$f' = cf - f^\rho \ .$$

Wir erhalten

$$g' = (1-\rho)cg - (1-\rho)\ .$$

Die allgemeine Lösung dieser Gleichung ist

$$g(t) = Ce^{(1-\rho)ct} + \frac{\mathsf{e}^{(1-\rho)ct} - 1}{c}\ .$$

Folglich löst

$$f(t) := \left(Ce^{(1-\rho)ct} + \frac{\mathsf{e}^{(1-\rho)ct} - 1}{c}\right)^{\frac{1}{1-\rho}}$$

die logistische Gleichung. Es gilt $f(0) = C^{\frac{1}{1-\rho}}$, und damit die Relation zwischen C und der Anfangsbedingung.

Eine Differentialgleichung der Form

$$f'(t) = a(t)f(t)^2 + b(t)f(t) + c(t)$$

heißt **Riccatische Differentialgleichung**. Ist h eine Lösung dieser Gleichung, dann findet man alle weiteren Lösungen in der Form $f = h + g$, wobei g eine Lösung der Bernoullischen Differentialgleichung

$$g'(t) = [2a(t)h(t) + b(t)]g(t) + a(t)g(t)^2$$

ist. In der Tat gilt

$$f^2 - h^2 = (f-h)(f+h) = g(g+2h)\ .$$

Damit ist

$$h' + g' = a(t)(h^2 + g(t)(g(t) + 2h(t))) + b(t)(h(t) + g(t)) + c(t)\ .$$

Es folgt die Bedingung

$$g' = a(t)g(t)^2 + 2a(t)h(t)g(t) + b(t)g(t)\ .$$

Als Beispiel betrachten wir die Gleichung

$$f' = (1-t)f^2 + (2t-1)f - t\ .$$

Die Funktion $h(t) \equiv 1$ ist eine Lösung. Folglich muß man nun die Gleichung

$$g' = [2(1-t) + (2t-1)]g(t) + (1-t)g(t)^2 = g(t) + (1-t)g(t)^2$$

lösen. Wir substituieren $\phi := g^{-1}$ und erhalten

$$\phi'(t) = -\phi(t) + (t-1)\ .$$

Daraus ergibt sich

$$\phi(t) = Ce^{-t} + \int_0^t \mathsf{e}^{s-t}(t-1)dt\ .$$

10 Elemente der Qualitativen Theorie

10.1 Langzeitverhalten

Sei $\Phi : \mathbb{R} \times V \to V$ ein Fluß.

Definition 10.1 *1. Der Orbit von x unter Φ ist $\mathcal{O}_x := \Phi_{\mathbb{R}}(x)$. Wir setzen $\mathcal{O}_x^+ := \Phi_{\mathbb{R}_+}(x)$ und $\mathcal{O}^-(x) := \Phi_{\mathbb{R}_-}(x)$.*

2. *$x \in V$ ist ein* ***stationärer Punkt****, falls $\mathcal{O}_x = \{x\}$ gilt.*

3. *$\mathcal{O}_x$ ist* ***geschlossen*** *mit Periode $T \in \mathbb{R} \neq 0$, falls für ein (und damit für alle $y \in \mathcal{O}_x$) gilt $\Phi_T(y) = y$. Beachte, daß mit T auch nT, $n \in \mathbb{Z} \setminus \{0\}$, eine Periode ist. Wir sagen dann, daß x ein* ***periodischer Punkt*** *sei.*

4. *Wir definieren*

$$\begin{aligned} \omega(x) &:= \bigcap_{s \in \mathbb{R}_+} \overline{\{\Phi_{[s,\infty)}(x)} \\ \alpha(x) &:= \bigcap_{s \in \mathbb{R}_-} \overline{\{\Phi_{(-\infty,s]}(x)} \end{aligned}$$

5. *Eine Teilmenge $A \subseteq V$ heißt* ***invariant****, falls für alle $x \in A$ gilt $\mathcal{O}_x \subseteq A$.*

6. *Die Teilmenge $A \subset V$ heißt* ***anziehend****, wenn es eine Umgebung U von $\bar{A}$ derart gibt, daß für jede Umgebung V von $\bar{A}$ und Kompaktum $K \subseteq U$ ein $T > 0$ existiert, so daß $\Phi_{[T,\infty)}(K) \subseteq V$. Analog definiert man abstoßende Mengen also solche, die unter Zeitumkehr anziehend sind.*

1. Stationäre Punkte und geschlossene Orbits sind invariant.
2. Ist x ein stationärer Punkt, dann gilt $\alpha(x) = \omega(x) = \{x\}$.
3. Ist x periodisch, dann ist $\alpha(x) = \omega(x) = \mathcal{O}_x$.
4. Ist $\omega(x)$ ($\alpha(x)$) leer, dann ist $\mathcal{O}_x^+$ ($\mathcal{O}_x^-$) nicht kompakt.
5. Die Mengen $\omega(x)$ und $\alpha(x)$ sind immer abgeschlossen.

Sei V ein metrischer Raum.

Lemma 10.2 *Wenn die Menge $\mathcal{O}_x^+$ kompakt ist, dann ist $\omega(x)$ zusammenhängend.*

Beweis: Sei $\omega(x)$ kompakt und als disjunkte Vereinigung zweier kompakter Mengen C_0, C_1 zerlegt. Sei $\delta = d(C_0, C_1)$. Dann existiert eine Folge (t_n) mit

$$\lim_{k\to\infty} d(\Phi_{t_{2k}}(x), C_0) = 0 \ , \quad \lim_{k\to\infty} d(\Phi_{t_{2k+1}}(x), C_1) = 0 \ .$$

Dann existieren für alle $i > i_0$ Punkte $t_{2i} < s_i < t_{2i+1}$ mit $d(\Phi_{s_i}(x), C_0) > \delta/4$ und $d(\Phi_{s_i}(x), C_1) > \delta/4$. Wegen der Kompaktheit hat $(\Phi_{s_i}(x))_{i \geq i_0}$ einen Häufungspunkt, welcher entweder zu C_0 oder zu C_1 gehören muß. Andererseits hat dieser aber mindestens den Abstand $\delta/4$ von diesen Mengen. Widerspruch. ∎

Lemma 10.3 *Die Mengen $\omega(x)$ und $\alpha(x)$ sind invariant.*

Beweis: Sei $y \in \omega(x)$. Dann existiert eine Folge (t_n) mit $t_n \to \infty$ und $\Phi_{t_n}(x) \to y$. Sei $s \in \mathbb{R}$. Dann ist

$$\Phi_s(y) = \lim_{n\to\infty} \Phi_s(\Phi_{t_n}(x)) = \lim_{n\to\infty} \Phi_{t_n+s}(x)$$

und damit $\Phi_s(y) \in \omega(x)$. Die Invarianz von $\alpha(x)$ zeigt man analog. ∎

Sei nun V ein metrischer Raum.

Definition 10.4 1. *Ein Orbit $\mathcal{O}_x^+$ heißt stabil, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so daß aus $d(y, x) < \delta$ für alle $t \geq 0$ folgt $d(\Phi_t(y), \Phi_t(x)) < \epsilon$.*

2. *Ein Orbit $\mathcal{O}_x^+$ heißt anziehend, wenn es ein $\delta > 0$ gibt, so daß aus $d(x, y) < \delta$ folgt $\lim_{t\to\infty} d(\Phi_t(y), \Phi_t(x)) = 0$.*

Wir betrachten zum Beispiel die Differentialgleichung

$$f' = A(t)f \tag{5}$$

für eine p-periodische Funktion $A : \mathbb{R} \to \mathtt{Mat}(n, n)$. Die Übersetzung in ein autonomes System liefert das Vektorfeld auf $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$

$$X(x_0, x) = (1, A(x_0)x) \ .$$

Der Lösung $f_0(t) \equiv 0$ entspricht der Orbit $\mathcal{O}_{(0,0)}^+ = \mathbb{R}_+ \times \{0\}$ des Flusses von X. Die Fundamentalmatrix $\Phi(t)$ der Gleichung (5) hatten wir in der Form $\Phi(t) = Z(t)\mathtt{e}^{tR}$ dargestellt, wobei $Z(t)$ eine p-periodische matrixwertige Funktion war. Seien $\lambda_1, \ldots, \lambda_r$ die Floquet-Exponenten des Systems. Wenn $\mathtt{Re}(\lambda_i) \leq 0$ für alle $i = 1 \ldots, r$, dann ist jeder Orbit von X stabil. Sind alle diese Ungleichung echt, dann ist jeder Orbit anziehend. In der Tat gilt ist im ersten Fall $\|\mathtt{e}^{tR}\|$ beachränkt und es gilt $\|e^{tR}\| \to 0$ für $t \to \infty$ im zweiten Fall. In der Tat, ist seien $(u, f), (v, g) \in \mathbb{R}^{n+1}$. Dann ist $\Phi_t(u, f) = (u + t, \Phi_t\Phi_u^{-1})$ und $\Phi_t(v, g) = (v + t, \Phi_t\Phi_v^{-1}g)$. Nun ist die Differenz der ersten Komponenten durch $u + t - (v + t) = u - v$ in der Zeit konstant. Die Verktorkomponente erfültt

$$\|\Phi_t\Phi_u^{-1}f - \Phi_t\Phi_v^{-1}g\| \leq \|Z(t)\|\|e^{tR}\|\|(\Phi_u^{-1}f - \Phi_v^{-1}g)\| \ .$$

Wenn $\|\mathtt{e}^{tR}\| \to 0$, dann ist der Orbit $\mathcal{O}_{0,0}^+$ anziehend (für jeden anderen Orbit). Andernfalls kann man durch genügend kleine Wahl von $|u - v|$ und $\|f - g\|$ zumindest erreichen, daß $\|(\Phi_u^{-1}f - \Phi_v^{-1}g)\|$ kleiner als eine vorgegeben Zahl wird. Daraus folgt unmittelbar die Stabilität jedes Orbits.

10.2 1-dimensionale Vektorfelder

Sei $X : \mathbb{R} \to \mathbb{R}$ ein Lipschitz-setiges vollständiges Vektorfeld auf $\mathbb{R}$ und Φ der von X erzeugte Fluß.

Lemma 10.5 *Die stationären Punkte von Φ sind die Nullstellen von X.*

Beweis: Ist $X(x) = 0$, dann ist die konstante Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(t) := x$ die eindeutige Lösung der Differentialgleichung $f' = X(f)$ mit $f(0) = x$. Damit gilt $\Phi_t(x) = x$ für alle $t \in \mathbb{R}$ und $\mathcal{O}_x = \{x\}$.

Umgekehrt, wenn $\mathcal{O}_x = \{x\}$ ist, dann gilt $\Phi_t(x) = x$ und deshalb $X(x) = 0$ nach Ableiten. ∎

Lemma 10.6 *Sei $x \in \mathbb{R}$ und $X(x) > 0$. Entweder es gibt eine Nullstelle $y \in \mathbb{R}$ von X mit $y > x$, dann gilt $\omega(x) = \inf\{y > x | X(y) = 0\}$, oder $\omega(x) = \emptyset$. Eine analoge Aussage gilt für die α-Menge.*

Beweis: Sei $\omega(x)$ nicht leer, $y \in \omega(x)$. Wenn X keine Nullstelle oberhalb von x hätte, dann wäre $X(z) > 0$ für alle $z > x$. Für eine Folge (t_n) mit $t_n \to \infty$ gilt $y = \lim_{n\to\infty} \Phi_{t_n}(x) = y$. Wir schreiben $\Phi_{t_n} = \int_0^{t_n} X(f(s))ds$ und schließen, daß $X(y) = 0$ gelten muß. Widerspruch.

Also existiert eine Nullstelle oberhalb von x. Sei $z := \inf\{y > x | X(y) = 0\}$. Dann gilt $X(z) = 0$ aus Stetigkeitsgründen. Wenn $y > z$, dann gäbe es $t > 0$ mit $\Phi_t(x) = z$. Da aber z ein stationärer Punkt ist, gilt dann $\Phi_s(x) = z$ für alle $s \geq t$ und damit $y = z$. Widerspruch. Also gilt $y = z$. ∎

Lemma 10.7 *Sei $x \in \mathbb{R}$ eine Nullstelle von X. Wenn $X' < 0$ ist, dann ist der Orbit $\mathcal{O}_x^+$ anziehend.*

Beweis: Sei $(x - \epsilon, x + \epsilon)$ so gewählt, daß $X(y) > 0$ für $y \in (x - \epsilon, x)$ und $X(y) < 0$ für $X(y) \in (x, \epsilon)$. Für jeden Punkt $y \in (x - \epsilon, x + \epsilon)$ gilt $\lim_{t\to\infty} \Phi_t(y) = x$. Damit ist der Orbit $\mathcal{O}_x$ anziehend. ∎

Da die α und ω-Mengen von X höchstens Punkte sind, kann ein eindimensionales Vektorfeld keine periodischen Orbits haben.

10.3 2-dimensionale Vektorfelder

10.3.1 Stationäre Punkte

Sei $X : \mathbb{R}^2 \to \mathbb{R}^2$ ein Lipschitz-stetiges Vektorfeld und Φ der Fluß.

Lemma 10.8 *x ist genau dann ein stationärer Punkt, wenn $X(x) = 0$.*

Beweis: Übungsaufgabe. ■

Wir betrachten nun Beispiele für das typische Verhalten eines Flusses in der Nähe von stationären Punkten.. Wenn wir X als zweimal differenzierbar annehmen und $X(0) = 0$ gilt, dann ist

$$X(x) = Ax + O(x^2) \ .$$

Wir werden weiter hinten sehen, daß die Matrix A das Verhalten im wesentlichen bestimmt, wenn nicht gerade beide Eigenwerte rein imaginär sind.

Seien also λ_0, λ_1 die Eigenwerte. Da das charakteristische Polynom von A reell ist, gilt entweder $\mathtt{Re}(\lambda_0) = \mathtt{Re}(\lambda_1)$, oder beide Eigenwerte sind reell.

Definition 10.9 *Der stationäre Punkt* 0 *von* X *ist*

Name	*stabil*	*hyperbolisch*	*instabil*
def	$\mathtt{Re}(\lambda_0) < 0, \mathtt{Re}(\lambda_1) < 0$	$\lambda_i \in \mathbb{R}, \lambda_0\lambda_1 < 0$	$\mathtt{Re}(\lambda_0) > 0, \mathtt{Re}(\lambda_1) > 0$
A	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
e^{tA}	$\begin{pmatrix} \mathsf{e}^{-t} & 0 \\ 0 & \mathsf{e}^{-t} \end{pmatrix}$	$\begin{pmatrix} \mathsf{e}^{-t} & 0 \\ 0 & \mathsf{e}^{t} \end{pmatrix}$	$\begin{pmatrix} \mathsf{e}^{t} & 0 \\ 0 & \mathsf{e}^{t} \end{pmatrix}$
B	$\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$		$\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$
$e^{t(A+B)}$	$e^{-t}\begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$		$e^{t}\begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$

Die Eigenwerte der Matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

sind imaginär. Es gilt

$$\exp(t\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \ .$$

Das Verhalten des Flusses von $X(x) = Ax + O(x^2)$ wird wesentlich von den höheren Termen $O(x^2)$ beeinflußt.

10.3.2 Gradientenfelder

Als nächstes betrachten wir Gradientenvektorfelder. Die Argumente funktionieren in beliebiger Dimensionen. Sei $X = -\mathtt{grad}(U)$ für ein Potential $U : \mathbb{R}^n \to \mathbb{R}^n$.

Lemma 10.10 *Für* $x \in \mathbb{R}^n$ *ist entweder* $\omega(x) = \emptyset$ *oder* $\omega(x)$ *besteht aus stationäreren Punkt von* X*. Eine analoge Aussage gilt für* $\alpha(x)$*.*

Beweis: Sei $\omega(x) \neq \emptyset$ und $y \in \omega(x)$. Wenn $X(x) \neq 0$, dann ist $U(\Phi_t(x))$ streng monoton fallend. In der Tat ist $\frac{d}{dt}U(\Phi_t(x)) = -\|X(\Phi_t(x)\|^2$. Wenn $U(\Phi_t(x))$ nicht streng monoton fallend wäre, dann gäbe es ein $t > 0$ mit $X(\Phi(t)) = 0$. Dieser Punkt wäre ein stationärer Punkt und deshalb gleich x. Widerspruch. Dann gilt

$$\lim_{t\to\infty} U(\Phi_t) = U(y) = U(x) - \int_0^\infty \|X(\Phi_s(x)\|ds \ .$$

Daraus folgt $X(y) = 0$. ■

Korollar 10.11 *Ein Gradientenvektorfeld im $\mathbb{R}^n$ hat keine periodischen Orbits.*

10.3.3 periodische Orbits - ein Beispiel

Wir kommen nun auf den zwei-dimensionalen Fall zurück. Das Vektorfeld $X(x) := Ax$ mit

$$A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

hat ausschließlich periodische Orbits. Wir betrachten nun

$$X(x) = Ax + xf(\|x\|^2) \ .$$

In diesem Fall kann man den Fluß explizit berechnen. Sei $\phi_s(r)$ Lösung der Gleichung

$$\phi_s' = \phi_s f(\phi_s^2) \ , \phi_s(0) = s \ .$$

Diese kann man explizit lösen durch Separation der Variablen. Dann ist

$$\Phi_t(x) = \phi_{\|x\|}(t)\mathrm{e}^{tA}x_0 \ .$$

In der Tat gilt (wegen $\|\mathrm{e}^{tA}x\| = \|x\|$)

$$\Phi_t'(x) = \phi_{\|x\|}'(t)\mathrm{e}^{tA}x + \phi_{\|x\|}(t)A\mathrm{e}^{tA} = A\phi_{\|x\|}(t)\mathrm{e}^{tA}x + \phi_{\|x\|}(t)f(|\phi_{\|x\|}(t)|^2)\mathrm{e}^{tA}x = A\Phi_t(x) + xf(\|\Phi_t(x)\|^2) \ .$$

Die periodischen Orbits dieses Systems entsprechen den Nullstellen der Funktion f. Genauer ist $\mathcal{O}_x$ periodisch mit Periode 2π, wenn $f(\|x\|) = 0$ und $x \neq 0$ ist.

Wenn $f'(|x|) < 0$ ist, dann ist der durch x bestimmte periodische Orbit anziehend. Falls $f'(|x|) > 0$ ist, dann ist er abstoßend.

10.3.4 Lotka-Volterra Modell

Wir betrachten die Differentialgleichung

$$\begin{aligned} B' &= B(\epsilon - \alpha R) \\ R' &= R(-\gamma + \beta B) \end{aligned}$$

mit positiven Konstanten $\epsilon, \alpha, \gamma, \beta$. Der Vektor (B, R) Beschreibt die Anzahl der Individuen zweier Populationen, wobei B die beute des Räubers R ist. Es gilt für $R = 0$, daß $R' = 0$ und für $B = 0$, daß $B' = 0$. Deshalb ist die Teilmenge $U = \{R \geq 0 \ , B \geq 0\}$ invariant.

Wir berechnen nun die stationären Punkte. Zunächst einmal erhalten wir

$$(0,0) \ , (\xi, \eta) := (\frac{\gamma}{\beta}, \frac{\epsilon}{\alpha}) \ .$$

Die linearisierte Gleichung im Punkt $(0, 0)$ ist

$$B' = \epsilon B \ , \quad R' = -\gamma R \ .$$

Dieser Punkt ist hyperbolisch.

Wir linearisieren die Gleichung in (ξ, η). Dazu führen wir die Koordinaten $u = B - \xi$ $v = R - \eta$ ein und schreiben

$$u' = (u + \xi)(\epsilon - \alpha(v + \eta)) \ , \quad v' = (v + \eta)(-\gamma + \beta(u + \xi)) \ .$$

Das in $(u, v) = (0, 0)$ linearisierte System ist

$$u' = -\xi\alpha v \ , \quad v' = \eta\beta u \ .$$

Die Eigenwerte der zugrundeliegenden linearen Abbildung sind $\pm i\xi\eta\alpha\beta$.

Wir betrachten nun die Funktion

$$E := \beta B - \gamma \ln(B) + \alpha R - \epsilon \ln(R) \ .$$

Es gilt

$$\begin{aligned} E' &= \beta B' - \gamma B'/B + \alpha R' - \epsilon R'/R \\ &= \beta B' - \gamma(\epsilon - \alpha R) + \alpha R' - \epsilon(-\gamma + \beta B) \\ &= \beta(B' - \epsilon B) + \alpha(R' + \gamma R) \\ &= -\beta\alpha BR + \alpha\beta BR \\ &= 0 \end{aligned}$$

Diese Funktion ist eigentlich auf dem Inneren von U. Die Integralkurven liegen also auf den Niveaumengen $E = const$. Diese sind kompakt und damit existieren die Integralkurven mit Anfang in U für alle Zeiten. Sei $x \in \{E = c\}$ vom kritischen Punkt verschieden. Sei $\gamma_t : \mathbb{R} \to \{E = c\}$ die Integralkurve durch x. Wir möchten einsehen, daß γ periodisch ist. In der Tat ist $\gamma_{\mathbb{R}}$ präkompakt. Sei (t_n) eine Folge mit $t_n \to \infty$ und $\gamma_{t_n} \to y$. Sei γ' die Integralkurve mit Anfang in y. Dann ist γ' eine Parametrisierung von $\{E = c\}$ in der Nähe von y. Damit existieren $t_n, t_m \in \mathbb{R}$ mit $|t_n - t_m| > 10$ und $s_0, s_1 \in \mathbb{R}$ mit $|s_0|, |s_1| < 1$ so daß $\gamma_{t_n} = \gamma'_{s_0}$ und $\gamma_{t_m} = \gamma'_{s_1}$ gilt. Wir schließen, daß $\gamma_{t_n - s_0} = y = \gamma_{t_m - s_1}$. Damit ist $t_n - s_0 - t_m + s_1 \neq 0$ eine Periode von γ.

10.3.5 Die Duffing Gleichung

Sei $V : \mathbb{R} \to \mathbb{R}$ eine genügend oft differenzierbare Funktion und 0 ein lokales Minimum mit $V(0) = 0$. Dann gilt $V(x) = \frac{a}{2}x^2 + O(x^3)$ mit $a > 0$. Wir betrachten die Bewegung eines Teilchens der Masse 1 im Potential V. Die Differentialgleichung ist

$$f^{(2)} = -V'(f) \ .$$

Diese Gleichung hat die Erhaltungsgröße

$$E(f, f') := \frac{1}{2}f'^2 + V(f) \ .$$

Für kleine f ist $V'(f) \sim af$ und wir erhalten die lineare Gleichung

$$f^{(2)} = -af \ ,$$

also die eines harmonischen Oszillators. Deren Lösungen kennen wir explizit:

$$f(t) = c\sin(\omega t) + d\cos(\omega t) \ , \quad \omega = \sqrt{a} \ .$$

Die Lösungen liegen auf den Niveaukurven von

$$E(f, f') = \frac{1}{2}f'^2 + \frac{a}{2}f^2 \ ,$$

also Ellipsen. Wir wollen annehmen, daß V symmetrisch ist. Dann ist

$$V(x) = \frac{1}{2}ax^2 + \frac{b}{4}x^4 + O(x^6)$$

die nächstbeste Approximation. Wir betrachten den Fall, daß $a < 0$ und $b > 0$ ist. Läßt man die Terme höherer Ordnung weg, dann erhalten wir die ungedämpfte Duffing Gleichung

$$f^{(2)} = -af - bf^3 \ .$$

Die Lösungen liegen auf den Niveaukurven von

$$E(f, f') = \frac{1}{2}f'^2 + \frac{a}{2}f^2 + \frac{b}{4}f^4.$$

Die Gleichung

$$0 = \frac{a}{2}f^2 + \frac{b}{4}f^4$$

hat die Nullstellen $f = 0$ und $\pm\sqrt{-\frac{2a}{b}}$ falls $ab < 0$ ist. Die Nullstellen von V' liegen in 0 und $\sqrt{-\frac{a}{b}}$, wobei die letzteren Minima sind. Die Menge $E = 0$ zerlegt die Ebene in drei Bereiche, von denen zwei beschränkt sind und Potentialtöpfe genannt werden. Die Bahnkurven mit $E > 0$ umkreisen beide Töpfe. Die Bahnkurven mit $E < 0$ umkreisen den Extremwert von E im jeweiligen Topf.

Wird das Teilchen gedämpft (z.B. durch Reibung), dann führt das zu einem zusätzlichen Term

$$f^{(2)} = -ef' - af - bf^3 \ .$$

In diesem Fall wissen wir, daß für jede Lösung $E(f, f')' \leq 0$ gilt. Daraus kann man das qualitative Verhalten der Lösungen ablesen. Die kritischen Punkte sind

$$(0,0), (\sqrt{-\frac{a}{b}}, 0), (-\sqrt{-\frac{a}{b}}, 0) \ .$$

Die Bahnkurven, welche außerhalb der Töpfe starten, werden diese einige Male umkreisen um dann in der Regel (wenn sie nicht den Punkt $(0,0)$ ansteuern), in einem der Töpfe den dortigen kritischen Punkt annähern. Die Linearisierung der Gleichung im Punkt $(\sqrt{-\frac{a}{b}}, 0)$ ist mit $u + \sqrt{-\frac{a}{b}} = f$ durch $u^{(2)} = -eu' + 2au$ gegeben. Die Eigenwerte der Matrix

$$\begin{pmatrix} 0 & 1 \\ 2a & -e \end{pmatrix}$$

sind

$$-\frac{e}{2} \pm \sqrt{\frac{e^2}{4} + 2a} \ .$$

Somit ist dieser kritische Punkt wie erwartet stabil (wie auch der im anderen Topf). Kreisendes Verhalten liegt vor, wenn $\frac{e^2}{4} + 2a < 0$, also etwa für $b > 0$, $a < 0$ und kleine e.

Die Linearisierung bei $(0,0)$ ist durch die Matrix

$$\begin{pmatrix} 0 & 1 \\ -a & -e \end{pmatrix}$$

beschrieben, welche die Eigenwerte

$$-\frac{e}{2} \pm \sqrt{\frac{e^2}{4} - a}$$

hat. Diese haben unterschiedliches Vorzeichen und deshalb ist der Punkt hyperbolisch.

10.3.6 Das Poincaré-Bendixson Theorem

Wir betrachten ein zumindest halbvollständiges Vektorfeld X auf dem $\mathbb{R}^2$. Sei $\Phi : [0, \infty) \times \mathbb{R}^2 \to \mathbb{R}^2$ der Fluß von X für positive Zeiten.

Theorem 10.12 (Poincaré-Bendixson) *Sei $y \in \mathbb{R}^2$ und $\omega(y)$ nicht leer und enthalte keine Fixpunkte. Dann ist $\omega(y) = \omega(z)$ für einen periodischen Punkt z.*

Beweis: Wir skizzieren hier den Beweis, der den Jordanschen Kurvensatz benutzt.

1. Eine differenzierbare Kurve $\gamma : [a, b] \to \mathbb{R}^2$ ist **transversal** zu X, wenn $\gamma'(t)$ und $X(\gamma(t))$ linear unabhängig sind für alle $t \in [a, b]$. Sei $x \in \mathbb{R}^2$ kein stationärer Punkt von X. Dann können wir einen Vektor $N \in \mathbb{R}^2$ wählen, der von $X(x)$ linear unabhängig ist. Die Kurve $[-a, a] \mapsto x + tN \in \mathbb{R}^2$ ist für genügend kleine $a > 0$ transversal zu X. Durch jeden nicht-stationären Punkt von X gibt es also ein Transversal.

2. Sei $a < 0 < b$ und $\gamma : [a, b] \to \mathbb{R}^2$ transversal zu X. Dann existiert ein $\epsilon > 0$ derart, daß $\Psi(s, t) := \Phi_s(\gamma(t)) : (-\epsilon, \epsilon) \times (-\epsilon, \epsilon) \to \mathbb{R}^2$ ein Diffeomorphismus auf das Bild ist, welches eine offene Umgebung von $\gamma(0)$ ist. In der Tat sind $\partial_s\Psi(0, 0) = X(\gamma(0))$ und $\partial_t\Psi(0,) = \gamma'(0)$ linear unabhängig. Wir wenden nun den Satz über die Umkehrfunktion an. Insbesondere, gibt für jedes $u \in \mathbb{R}^2$ und $t \in \mathbb{R}$ mit $\Phi_t(u) \in \Psi((-\epsilon, \epsilon) \times (-\epsilon, \epsilon))$ ein t' mit $|t - t'| < \epsilon$ so daß $\Phi_{t'} \in \gamma((-\epsilon, \epsilon))$.

3. Sei $x \in \mathbb{R}^2$ kein stationärer oder periodischer Punkt. Durch geeignete Wahl der Koordinaten können wir erreichen, daß $x = (0, 0)$, $X^1(0, t) > 0$ für $t \in [-1, 1]$. Sei $0 = t_0 < t_1 < t_2 < \dots$ die (möglicherweise endliche) Folge der Zeiten, zu welchen der Orbit $\mathcal{O}_x$ die Strecke $\{0\} \times (-1, 1)$ schneidet. Es gilt also $\Phi_{t_i}(0, 0) = (0, y_i)$ mit $|y_i| < 1$. Dann ist die Folge der (y_i) streng monoton. Da x nicht periodisch ist, gilt $y_1 \neq 0$. Wir betrachten den Fall, daß $y_1 > 0$. Angenommen, die Folge (y_i) wachse nicht streng monoton. Dann gilt für ein geeignetes i, daß $y_i < y_{i+1}$ und $y_{i+2} \leq y_{i+1}$. Die Gleichheit können wir sofort ausschließen, da nämlich aus $y_{i+1} = y_{i+2}$ die Periodizität von y_{i+1} und damit von x folgen würde. Also gilt $y_{i+2} < y_{i+1}$. Die Kurve K welche aus den Segmenten $\Phi_{[t_i,t_{i+1}]}((0, y_i))$ und der Strecke zwischen $(0, y_i)$ und $(0, y_{i+1})$ zusammengesetzt wird, ist geschlossen und ohne Doppelpunkte (eine Jordankurve). Nach dem Jordanschen Kurvensatz zerfällt $\mathbb{R}^2 \setminus K$ in zwei Wegzusammenhangskomponenten C_- und C_+ (das Innere und das Äußere von K), wobei C_- die Komponente links der Strecke von $(0, y_i)$ nach $(0, y_{i+1})$ ist. Da wir für $|y| < 1$ die Ungleichung $X^1(0, y) > 0$ angenommen haben, muß schbeidet jder Orbit die Strecke von $(0, -1)$ nach $(0, 1)$ von links nach rechts. Im Punkt t_{i+1} betritt die Kurve $\Phi_t(x)$ also die Komponente C_+. Damit kreuzt die Kurve $\Phi_t(x)$ das Transversal in t_{i+2} von C_+ nach C_-. Das ist aber in dieser Richtung unmöglich.

4. Sei $\gamma : [a, b] \to \mathbb{R}$ transversal zu X, $L := \gamma([a, b])$. Dann gilt für jedes $x \in \mathbb{R}^2$, daß $|\omega(x) \cap L| \leq 1$. Sei (t_n) die monoton wachsende Folge der Zeiten mit $u_n := \Phi_{t_n}(x) \in L$. Dann ist die Folge der (u_n) nach 3. monoton in L. Ist sie unendlich, dann hat sie einen Grenzwert $\bar{z}$.

   Sei jetzt $z \in \omega(x) \cap L$. Die Betrachtung 2. zeigt, daß wenn ein Orbit $\Phi_t(x)$ zu einer bestimmten Zeit t nahe an z liegt, dann auch L in einem Punkt nahe an z für einen Zeitpunkt t' nahe an t schneiden muß. Wir finden mit 2. eine Folge von Zeiten $t_n \to \infty$ mit $\Phi_{t_n}(x) \to z$ und $\Phi_{t_n}(x) \in L$. Damit ist aber $z = z'$.

5. Unsere Annahmen implizieren die Existenz eines Punktes $z \in \omega(y)$ mit $y \neq z$. Dann gilt $\mathcal{O}_z^+ \subseteq \omega(y)$ und deshalb $\omega(z) \subseteq \omega(y)$. Sei $u \in \omega(z) \subseteq \omega(y)$. Dann gilt $X(u) \neq 0$

und deshalb existiert ein Transversal L zu X durch u. Der Orbit $\mathcal{O}_z^+$ schneidet L unendlich oft (siehe Argument weiter oben), andererseits gilt $\{u\} = \omega(y) \cap L$. Folglich fallen alle diese Schnittpunkte mit u zusammen und z ist ein periodischer Punkt. Wir sehen, daß $\omega(y)$ aus periodischen Punkten besteht.

6. Sei $z \in \omega(y)$ und $\omega(y) \backslash \omega(z) \neq \emptyset$. Da $\omega(y)$ zusammenhängend ist, existiert ein Punkt $z_1 \in \omega(z)$, welcher ein Häufungspunkt von $\omega(y) \setminus \omega(z)$ ist. Sei L ein Transversal durch z_1. Dann gibt es eine Folge von Punkten (u_i) in $\omega(y) \setminus \omega(z)$ derart, daß $\mathcal{O}_{u_i}^+ \cap L \neq \emptyset$. Da alle diese Schnittpunkte in $\omega(y)$ liegen, fallen diese Schnittpunkte mit z_1 zusammen, da $\{z_1\} = L \cap \omega(y)$. Da aber $u_i \notin \mathcal{O}_{z_1} = \omega(z_1) = \omega(z)$ gelten soll, ist das unmöglich.

■

10.4 Topologische Konjugiertheit

Seien (D, Φ) und (D', Φ') lokale Flüsse auf topologischen Räumen V und V'.

Definition 10.13 *Die beiden Flüsse heißen zueinander* ***orbitäquivalent****, falls es Homöomorphismen $h : D \to D'$ und $H : V \to V'$ derart gibt, daß*

$$\begin{array}{ccc} D & \xrightarrow{h} & D' \\ \downarrow{\scriptstyle \mathrm{pr}_2} & & \downarrow{\scriptstyle \mathrm{pr}_2} \\ V & \xrightarrow{H} & V' \end{array} \qquad \begin{array}{ccc} D & \xrightarrow{h} & D' \\ \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \Phi'} \\ V & \xrightarrow{H} & V' \end{array}$$

kommutieren und die Zeitorientierung erhalten bleibt, also us $t < s$ und $(s, x), (t, x) \in D$, $h(s, x) = (s', H(x))$ und $h(t, x) = (t', H(x))$ folgt $t' < s'$. Falls $h(t, x) = (t, H(x))$ gilt, dann heißen die Flüsse topologisch konjugiert.

Orbitäquivalenz und topologische Konjugiertheit von lokalen Flüssen sind Äquivalenzrelationen. Die Symmetrie folgt, da man die Homöomorphismen invertieren kann. Transitivität benutzt die Komposition der Homöomorphismen.

Lemma 10.14 *Seien (D, Φ) und (D', Φ') zueinander orbitäquivalent. Dann gilt*

1. *Ist x Fixpunkt von (D, Φ), dann ist $H(x)$ Fixpunkt von (D', Φ').*
2. *Liegt x auf einem geschlossenen Orbit von (D, Φ), dann liegt $H(x)$ auf einem geschlossenen Orbit von (D', Φ').*
3. *Es gilt $\omega(H(x)) = H(\omega(x))$, $\alpha(H(x)) = H(\alpha(x))$.*

Beweis: Klar ■

Seien X, X' Lipschitz-stetige Vektorfelder auf $U \subset V$, $U' \to V'$ und (D, Φ) und (D', Φ') die entsprechenden maximalen lokalen Flüsse.

Definition 10.15 *X und X' heißen zueinander orbitäquivalent oder topologisch konjugiert, wenn die entsprechenden maximalen lokalen Flüsse zueinander orbitäquivalent oder topologisch konjugiert sind.*

Orbitäquivalenz und topologische Konjugiertheit von Vektorfeldern sind Äquivalenzrelationen.

Seien A, B topologische Räume, $a \in A$. Eine **lokal bei a definierte stetige Abbildung** mit Werten in B ist ein Paar (U, f), wobei $U \subseteq A$ eine Umgebung von a und $f : U \to B$ eine stetige Abbildung ist. Zwei solche lokal bei a definierte stetige Abbildungen (U, f) und (V, g) heißen zueinander äquivalent, wenn es eine lokal bei a definierte stetige Abbildung (W, h) gibt mit $W \subset U \cap V$ und $h = f_{|W}$ und $h = g_{|W}$. Unter einem **Keim** einer stetigen Abbildung bei a mit Werten in B versteht man eine Äquivalenzklasse lokal bei a definierte stetiger Abbildungen mit Werten in B. Eine global definerte Abbildung $f : A \to B$ repräsentiert einen Keim bei a, den wir mit (f, a) notieren.

Seien jetzt V und V' Banachräume und $x \in U$, $x' \in U'$. Seien weiter (X, x) und (X', x') die Keime von Vektorfeldern X und X' in den Punkten x und x'.

Definition 10.16 *Wir sagen, daß (X, x) und (X', x') zueinander orbitäquivalent sind, falls es Umgebungen V und V' von x und x' gibt auf welcher Repräsentanten der Keime (X, x) und (X', x') definiert sind, so daß $X_{|V}$ und $X'_{|V'}$ zueinander orbitäquivalent.*

10.5 Flußboxen

Wir betrachten das konstante Vektorfeld auf $\mathbb{R}^n$, $Y(x) = (1, 0, \ldots, 0)$. Der assoziierte Fluß ist durch

$$\Psi_t(x) := x + t(1, 0, \ldots, 0)$$

gegeben.

Sei $\dim(V) = n$.

Theorem 10.17 *Sei $U \subset V$ offen und X ein Lipschitzstetiges differenzierbares Vektorfeld auf U. Sei $x \in U$ ein ein regulärer Punkt von X, d.h. $X(x) \neq 0$. Dann sind (X, x) und $(Y, 0)$ zueinander topologisch konjugiert.*

Beweis: Nach Verschiebung können wir $x = 0$ annehmen. Sei (D, Φ) der von X erzeugte maximal lokaler Fluß. Sei $N \subset V$ ein $n-1$-dimensionaler Unterraum mit $X(x) \notin N$. Wir zerlegen $V = \mathbb{R}X(x) \oplus N$ und notieren die Elemente als Paare. Wir wählen ein Intervall $I := (-\epsilon, \epsilon)$ und eine Umgebung R derart, daß $I \times R \subseteq D$ gilt. Wir betrachten nun die Abbildung $F : I \times R \to U$,

$$F(t, r) := \Phi_t(x + r)\ .$$

Im Punkt $(0, 0) \subset I \times R$ gilt $dF(0, 0)(s, r) = (sX(x), d\Phi_0(x)(r)) = (sX(x), r)$ für alle $(s, r) \in \mathbb{R} \times N$. Insbesondere ist damit $dF(0, 0)$ invertierbar. Damit existiert eine offene Umgebung $W \subset I \times R$ derart, daß $F(W)$ eine offene Umgebung von x und $F_{|W} : W \to F(W)$ diffeomorph ist. Wir können o.B.d.A durch Verkleinern von ϵ annehmen, daß

$W = I \times Q$ mit $0 \in Q \subseteq R$ gilt. Der maximale lokale Fluß von $X_{|F(W)}$ auf $D := \{(u, F(t,r)) | u + t \in I\ r \in Q\}$ definiert. Es gilt

$$\begin{aligned} \Phi_u \circ F(t,r) &= \Phi_u \circ \Phi_t(x+r) \\ &= \Phi_{u+r}(x+r) \\ &= F((t,r) + u(1,0)) \end{aligned}$$

Wir wählen eine linearen Isomorphismus $b : \mathbb{R}^{n-1} \to R$ und damit $B = (\mathtt{id}_{\mathbb{R}^1}, b) : \mathbb{R}^n \cong \mathbb{R}^1 \times \mathbb{R}^{n-1} \to \mathbb{R}^1 \times R$. Dann gilt offensichtlich auf der Umgebung $I \times b^{-1}(Q) = B^{-1}(W) \subset R^n$ von 0

$$\Phi_u \circ F \circ B(t,y) = F \circ B \circ \Psi_u(t,y)$$

für alle $u + t \in I$. Damit sind $X_{|F(W)}$ und $Y_{|B^{-1}(W)}$ durch $H : F \circ B : B^{-1}(W) \to F(W)$ und $h : D \to D'$, $h(u,x) = (u, H(x))$ zueinander topologisch konjugiert. ∎

10.6 Struktur hyperbolischer kritischer Punkte

Sei V eindlichdimensional und $A \in \mathtt{End}(V)$. Sei $\mathtt{spec}(A) \in \mathbb{C}$ die Menge der Eigenwerte von $A_{\mathbb{C}}$.

Definition 10.18 *A heißt hyperbolisch, falls $\mathtt{spec}(A_{\mathbb{C}}) \cap i\mathbb{R} = \emptyset$.*

Sei $U \subset V$ offen, und X ein Lipschitz-stetig differenzierbares Vektorfeld auf U.

Definition 10.19 *Ein Punkt $x \in U$ heißt hyperbolischer singulärer Punkt von X, falls $X(x) = 0$ und $dX(x)$ hyperbolisch ist.*

Sei $x \in U$ hyperbolischer kritischer Punkt von X und L das lineare Vektorfeld auf V, $L(v) := dX(x)(v)$.

Theorem 10.20 (Grobman und Hartman) *(X, x) und $(L, 0)$ sind orbitäquivalent.*

Der Beweis dieses Satzes wird im Verlauf von mehreren Kapiteln erbracht.

Die Voraussetzung hyperbolisch ist wichtig. Sei z.B. X auf $\mathbb{R}^2$ durch $X(x,y) := (y, -x)^+ - (x^3, y^3)^+$ gegeben. Dann ist 0 ein kritischer Punkt, aber nicht hyperbolisch. Wir haben $L(x,y) := (y, -x)^+$. Die Keime $(X, 0)$ und $(L, 0)$ sind nicht orbitäquivalent. In der Tat liegen alle Punkte von $\mathbb{R}^2$ auf geschlossenen Orbits von L. Das Feld X aber hat außer 0 keine geschlossenen Orbits.

Definition 10.21 *Für $A \in \mathtt{End}(A)$ und $\lambda \in \mathbb{C}$ ist die Multiplizität durch*

$$\mathtt{mult}(\lambda) := \lim_{n \to \infty} \dim_{\mathbb{C}} \ker(A_{\mathbb{C}} - \lambda)^n$$

definiert. Wir definieren den Index

$$\mathtt{ind}(A) := \sum_{\lambda \in \mathbb{C},\, \mathtt{Re}(\lambda) < 0} \mathtt{mult}(\lambda)$$

Seien $A, B \in \mathtt{End}(V)$ hyperbolisch und X, Y die entsprechenden linearen Vektorfelder, $X(x) := Ax$, $Y(x) := By$.

Theorem 10.22 *Wenn* $\mathtt{ind}(A) = \mathtt{ind}(B)$*, so sind* $(X, 0)$ *und* $(Y, 0)$ *orbitäquivalent.*

Sei nun X ein Vektorfeld auf $U \subset V$ und x ein hyperbolischer kritischer Punkt von X.

Definition 10.23 *Wir definieren den Index von* X *in* x *durch* $\mathtt{ind}(X, x) := \mathtt{ind}(dX(x))$.

Nimmt man die Sätze 10.20 und 10.22 zusammen, so erhält man die folgende Konsequenz.

Theorem 10.24 *Seien* X*,* X' *Lipschitz-stetig differenzierbare Vektorfelder auf* $U \subset V$*,* $U' \subset V$ *und* $x \in U$*,* $x' \in U'$ *hyperbolische kritische Punkte von* X*,* X' *mit*

$$\mathtt{ind}(X, x) = \mathtt{ind}(X', x') \ .$$

Dann sind (X, x) *und* (X', x') *orbitäquivalent.*

10.7 Stabilität

Wir betrachten das Vektorfeld

$$X_\lambda(x, y) = (\lambda - x^2, -y)$$

auf $\mathbb{R}^2$. Für $\lambda < 0$ besitzt es keine kritischen Punkte. Für $\lambda = 0$ ist $(0, 0)$ ein kritischer Punkt, während für $\lambda > 0$ die kritischen Punkte $(\sqrt{\lambda}, 0)$ und $(-\sqrt{\lambda}, 0)$ vorliegen. Hierbei ist der erste ein stabiler Fixpunkt und der zweite ein hyperbolischer Sattel. Dieses System verändert bei $\lambda = 0$ das qualitative Verhalten (Sattel-Knoten-Bifurkation). Der Fixpunkt $(0, 0)$ im Fall $\lambda = 0$ ist nicht hyperbolisch.

Stabilität formalisiert die Aussage, daß kleine Störungen des Vektorfeldes das Qualitative Verhalten des Flusses nicht ändern. Damit man von der Kleinheit der Störung reden kann, braucht man einen topologischen Raum der Vektorfelder. Eine Möglichkeit ist es, den Raum $C^1(U, V)$ zu nehmen. Wenn U konvex ist, dann sind die Elemente dieses Raumes automatisch Lipschitz-stetig. Wir machen deshalb hier diese Annahme der Einfachheit halber.

Definition 10.25 $X \in C^1(U, V)$ *heißt topologisch stabil, falls es eine Umgebung* $W \subset VF(U)$ *gibt, so daß jedes* $Y \in W$ *zu* X *orbitäquivalent ist.*

Wir können auch einen entsprechenden lokalen Begriff betrachten. Sei $x \in U$ und X stetig differenzierbar.

Definition 10.26 *Der Keim* (X, x) *heißt topologisch stabil, wenn es eine konvexe Umgebung* $U_1 \subset U$ *von* x *gibt und eine Umgebung* $W \subset C^1(U_1, V)$ *gibt, so daß für jedes* $Y \in W$ *ein* $y \in U_1$ *existiert, so daß* (X, x) *und* (Y, y) *zueindander orbitäquivalent sind.*

Das obige Beispiel $(X_0, (0,0))$ ist also nicht topologisch stabil.

Theorem 10.27 *Wenn x ein regulärer Punkt oder ein hyperbolischer Fixpunkt von X ist, dann ist (X,x) topologisch stabil.*

Beweis: Sei zuerst x regulärer Punkt von X. Sei $U_1 \subset U$ ein kleiner Ball um x. Dann ist x regulärer Punkt von allen $Y \in B(X, \frac{1}{2}\|X(x)\|) =: W$. Folglich sind (X,x) und (Y,x) zueinander orbitäquivalent.

Sei nun x ein hyperbolischer kritischer Punkt von x vom Index m. Wir zeigen, daß es Umgebungen $U_1 \subset U$ von x und $W \in C^1(U_1, V)$ gibt, so daß jedes $Y \in W$ einen kritischen Punkt $P(Y) \in U_1$ vom Index m hat.

Aus 10.24 folgt dann die Behauptung.

Wir betrachten dazu die Abbildung $B : C^1(U_1, V) \times U_2 \to V$, $B(Y,y) = Y(y)$, wobei $U_1 \subset U$ ein kleiner Ball um x ist Diese Abbildung ist stetig differenzierbar. In der Tat ist $d_1B(Y,y)(Z) = Z(y)$ und $d_2B(Y,y)(v) = dY(y)(v)$. Es gilt $B(X,x) = 0$ und $d_2B(X,x) = dX(x) : V \to V$ ist invertierbar. Damit können wir den Satz über implizite Funktionen anwenden und erhalten eine differenzierbare Abbildung $P : W \to U_1$ mit $Y(P(y)) = 0$ für alle $Y \in W$, wobei $W \subset C^1(U_1, V)$ eine geeignete Umgebung von X ist.

Wegen der Stetigkeit von $W \ni Y \mapsto dY(P(Y)) \in \mathtt{End}(V)$ folgt, daß nach eventuellem Verkleinern von W die Abbildungen $dY(P(Y))$ hyperbolisch vom Index m sind. ■

■

10.8 Beweis von Satz 10.22

Satz 10.28 *Seien $A, B \in \mathtt{End}(V)$ hyperbolisch und $\mathtt{ind}(A) = \mathtt{ind}(B) = \dim(V)$. Dann existiert ein Homöomorphismus $H : V \to V$ mit $H \circ \mathsf{e}^{tA} = \mathsf{e}^{tB} \circ H$ für alle $t \in \mathbb{R}$.*

Beweis:

Lemma 10.29 *Ist $A \in \mathtt{End}(V)$ hyperbolisch und $\mathtt{ind}(A) = \dim(V)$, so existiert ein $c < 0$ und ein Skalarprodukt $< .,. >$ auf V mit*

$$< Ax, x > \le c\|x\|^2 \ , \quad \forall x \in V \ .$$

Beweis: Wir setzen $c = \mathtt{max}\{\mathtt{Re}(\lambda) \mid \lambda \in \mathtt{spec}(A_{\mathbb{C}})\}$. Sei $(e_i)_{i=1,\dots,n}$ eine Basis von $V_{\mathbb{C}}$, bezüglich welcher $A_{\mathbb{C}}$ eine Jordansche Normalform hat. Dann gilt entweder $Ae_i = \lambda_i e_i$ oder $Ae_i = \lambda_i e_i + e_{i-1}$, wobei $\lambda_i \in \mathtt{spec}(A_{\mathbb{C}})$ ist. Sei $< .,. >_{\mathbb{C}}$ das hermitesche Skalarprodukt, bezüglich welchem die Basis (e_i) orthonormal ist. Wir definieren das

Skalarprodukt $< .,. >$ durch $< x, y >= \mathtt{Re}(< x, y >_{\mathbb{C}})$. Dann gilt für $x = \sum_i c_i e_i$

$$\begin{aligned} < Ax, x > &= \mathtt{Re}(\sum_{i,j} c_j \bar{c}_i < A_{\mathbb{C}} e_i, e_j >_{\mathbb{C}}) \\ &= \mathtt{Re}(\sum_i |c_i|^2 \bar{\lambda}_i) \\ &= \sum_i |c_i|^2 \mathtt{Re}(\lambda_i) \\ &\leq c\|x\|^2 \end{aligned}$$

■

Lemma 10.30 *Seien A und $< .,. >$ wie in Lemma 10.29. Dann existiert eine differenzierbare Funktion $t : V \setminus \{0\} \to \mathbb{R}$ derart, daß*

$$\|\mathsf{e}^{t(x)A} x\| = 1 \ , \quad \forall x \in V \ .$$

Beweis: Es ist klar,daß für jedes $x \in V \setminus \{0\}$ gilt

$$\lim_{t \to \infty} \|\mathsf{e}^{tA} x\| = 0 \ , \quad \lim_{t \to -\infty} \|\mathsf{e}^{tA} x\| = \infty \ .$$

Weiterhin ist wegen

$$\frac{d}{dt} \|\mathsf{e}^{tA} x\|^2 =< Ax, x >\leq c\|x\|^2$$

die Funktion $t \to \|x\|^2$ streng monoton fallend. Damit gibt es für jedes $x \in V \setminus \{0\}$ genau ein $t(x) \in \mathbb{R}$ mit $\|\mathsf{e}^{t(x)A} x\| = 1$. Nach dem Satz über implizite Funktionen ist $x \to t(x)$ differenzierbar. ■

Seinen nun $< .,. >_A$ und $< .,. >_B$ wie in Lemma 10.29 für A, B. Wir definieren $H : V \to V$ durch

$$H(x) := \begin{cases} \frac{\mathsf{e}^{-t(x)B} \mathsf{e}^{t(x)A} x}{\|\mathsf{e}^{t(x)A} x\|_B} \ , & x \neq 0 \\ 0 \ , & x = 0 \end{cases} \ .$$

Die Abbildung H ist stetig. In der Tat, wenn $x \to 0$, so $t(x) \to -\infty$, damit $\mathsf{e}^{-t(x)B} y \to 0$ gleichmäßig auf $S^1_{\|.\|_B}$. Wenn wir $\frac{\mathsf{e}^{t(x)A} x}{\|\mathsf{e}^{t(x)A} x\|_B} \in S^1_{\|.\|_B}$ beachten so folgt $H(x) \to 0$. Die zu H inverse Abbildung erhält man durch vertauschen der Rollen von A und B. Damit ist H ein Homöomorphismus. Es gilt

$$\begin{aligned} \mathsf{e}^{tB} \circ H(x) &= \mathsf{e}^{tB} \frac{\mathsf{e}^{-t(x)B} \mathsf{e}^{t(x)A} x}{\|\mathsf{e}^{t(x)A} x\|_B} \\ &= \frac{\mathsf{e}^{(t-t(x))B} \mathsf{e}^{(t(x)-t)A} \mathsf{e}^{tA} x}{\|\mathsf{e}^{(t(x)-t)A} \mathsf{e}^{tA} x\|_B} \\ &= H \circ \mathsf{e}^{tA}(x) \end{aligned}$$

$\blacksquare$

Wir beweisen nun den Satz 10.22. Seien A, B wie im Satz. Als Konsequenz der Jordanschen Normalform existieren A (bzw. B)-invariante Aufspaltungen $V = E^u \oplus E^s$ (bzw. $V = F^u \oplus F^s$) derart, daß $\mathtt{spec}((A_{|E^u})_{\mathbb{C}}) \subset \{\mathtt{Re} > 0\}$ und $\mathtt{spec}((A_{|E^s})_{\mathbb{C}}) \subset \{\mathtt{Re} < 0\}$ ($\mathtt{spec}((B_{|F^u})_{\mathbb{C}}) \subset \{\mathtt{Re} > 0\}$ und $\mathtt{spec}((B_{|F^s})_{\mathbb{C}}) \subset \{\mathtt{Re} < 0\}$). Nach Lemma 10.28 finden wir Homöomorphismen $H^u : E^u \to F^u$, $H^s : E^s \to F^s$ derart, daß für $H = (H^u, H^s)$ gilt $\mathrm{e}^{tB} \circ H = H \circ \mathrm{e}^{tA}$. $\blacksquare$

10.9 Beweis des Satzes 10.20

Definition 10.31 *Ein Element $A \in Gl(V)$ heißt hyperbolisch, falls $\mathtt{spec}(A_{\mathbb{C}}) \cap S^1 = \emptyset$.*

Satz 10.32 *Sei $A \in \mathtt{End}(V)$ hyperbolisch. Es existiert ein $\epsilon > 0$ derart,daß für alle Lipschitz-stetigen $\phi_1, \phi_2 \in C(V, V)$ mit Lipschitzkonstanten $< \epsilon$ ein Homöomorphismus $h : V \to V$ existiert, so daß*

$$h \circ (A + \phi_1) = (A + \phi_2) \circ h \ , \tag{6}$$

und wenn $h = 1 + u$ mit $u \in C(V, V)$, dann ist h eindeutig bestimmt.

Beweis: Als Konsequenz der Jordanzerlegung gibt es eine Konstante $a \in (0, 1)$ und eine A-invariante Aufspaltung $V = E^u \oplus E^s$ derart, daß $\|A^{-1}_{|E^u}\| \leq a$ und $\|A_{|E^s}\| \leq a$, wobei $\|.\|$ bezüglich einem Skalarprodukt gebildet wird, in welchem E^s und E^u orthogonal sind.

Wir machen den Ansatz $h = 1 + u$, $u \in C(V, V)$. Die Gleichung (6) is äquivalent zu

$$\phi_1 - \phi_2 \circ (1 + u) = A \circ u - u(A + \phi_1) \ . \tag{7}$$

Lemma 10.33 *Ist ϵ genügend klein, so ist die Abbildung $L : C(V, V) \to C(V, V)$, $L(u) := A \circ u - u \circ (A + \phi_1)$, ist stetig invertierbar. Es gilt $\|L^{-1}\| \leq \frac{\|A^{-1}\|}{1-a}$.*

Beweis: Wir schreiben $L = \tilde{A} \circ \tilde{L}$, wobei $\tilde{A}, \tilde{L} : C(V, V) \to C(V, V)$, $\tilde{A}(u) = A \circ u$, und $\tilde{L}(u) := u - A^{-1} \circ u \circ (A + \phi_1)$. Die Abbildung $\tilde{A}$ ist invertierbar mit Norm $\|\tilde{A}^{-1}\| = \|A^{-1}\|$.

Wir müssen also zeigen, daß auch $\tilde{L}$ stetig invertierbar ist und $\|\tilde{L}^{-1}\| \leq \frac{1}{1-a}$ gilt. Sei $\epsilon \leq \frac{1}{\|A^{-1}\|}$.

Wir behaupten, daß dann $A + \phi_1 : V \to V$ ein Homöomorphismus ist. Wir suchen die inverse Abbildung in der Form $A^{-1} + u$ mit $u \in C(V, V)$. Durch Auswerten von $(A + \phi_1) \circ (A^{-1} + u) = 1$ sehen wir, daß die Abbildung u der Gleichung

$$u = -A^{-1} \circ \phi_1 \circ (A^{-1} + u) =: B(u)$$

genügen muß. Die Abbildung $B : C(V, V) \to C(V, V)$ erfüllt

$$\|B(u) - B(v)\| \leq \epsilon \|A^{-1}\| \|u - v\|$$

und ist deshalb für $\epsilon \leq \frac{1}{\|A^{-1}\|}$ ein Kontraktion. Damit existiert ein eindeutiger Fixpunkt u und $A^{-1}+u$ ist ein Rechtsinverses von $A+\phi_1$. Wir suchen nun ein Linksinverses in der Form $A^{-1}+v$. Die Abbildung $v : V \to V$ muß also der Gleichung $v(A+\phi_1)+A^{-1}\circ\phi_1 = 0$ genügen. Wir komponieren von rechts mit $A^{-1}+u$ und erhalten $v = -A^{-1}\circ\phi_1\circ(A^{-1}+u)$. Damit besitzt $A+\phi_1$ auch ein Linksinverses welches dann notwendigerweise mit $A^{-1}+u$ übereinstimmen muß.

Wir machen nun die Aufspaltung $C(V,V) = C(V,E^u)\oplus C(V,E^s)$, $u = (u^u,u^s)$ und schreiben $\tilde{L}(u) = (\tilde{L}^u(u),\tilde{L}^s(u))$. Da A die Unterräume E^u, E^s erhält, hängen $\tilde{L}^u(u),\tilde{L}^s(u)$ nur von den jeweiligen Komponenten u^u, u^s ab. Wir definieren $R^u : C(V,E^u)\to C(V,E^u)$ durch $R^u(v) := -A^{-1}_{|E^u}\circ v\circ(A+\phi)$, und analog $R^s : C(V,E^s)\to C(V,E^s)$ durch $R^s(w) := -A^{-1}_{|E^s}\circ w\circ(A+\phi)$. Dann gilt $\tilde{L}(u) = (u^u+R^u(u^u), u^s+R^s(u^s))$.

Der lineare Operator R^s is invertierbar, $(R^s)^{-1}(w) = -A_{|E^s}\circ w\circ(A+\phi_1)^{-1}$. Es gilt $\|(R^s)^{-1}(w)\| \leq \|A_{|E^s}\|\|w\|$ und damit $(R^s)^{-1}\leq a$. Wir schreiben $1+R^s = R^s\circ((R^s)^{-1}+1)$. Dann gilt $(1+R^s)^{-1} = ((R^s)^{-1}+1)^{-1}\circ(R^s)^{-1}$ und damit $\|(1+R^s)^{-1}\| \leq \frac{a}{1-a}$.

Es gilt $\|R^u\|\leq a$ und damit ist $1+R^u$ invertierbar und $\|(1+R^u)^{-1}\|\leq\frac{1}{1-a}$.

Wir sehen, daß $\tilde{L}^{-1} = ((1+R^u)^{-1},(1+R^s)^{-1})$ existiert und $\|\tilde{L}^{-1}\| \leq \mathtt{max}\{\frac{1}{1-a},\frac{a}{1-a}\} = \frac{1}{1-a}$ gilt. ■

Wir lösen jetzt (7). Die Abbildung u ist Lösung dieser Gleichung, genau dann, wenn u ein Fixpunkt der Abbildung

$$S : C(V,V)\to C(V,V)\ ,\quad S(u) = \tilde{L}^{-1}(\phi_1-\phi_2\circ(A+u))$$

ist. Wegen

$$\|S(u)-S(v)\| \leq \|\tilde{L}^{-1}\|\|\phi_2(A+u)-\phi_2\circ(A+v)\| \leq \frac{\|A^1\|}{1-a}\epsilon\|u-v\|$$

ist S eine Kontraktion, wenn wir $\epsilon < \frac{1-a}{\|A^{-1}\|}$ voraussetzen. In diesem Fall hat S einen eindeutig bestimmten Fixpukt u.

Es bleibt zu zeigenm daß $h := 1+u$ ein Homöomorphismus ist. Vertauscht man die Rollen von ϕ_1 und ϕ_2, so findet man wie oben ein $v\in C(V,V)$ derart, daß

$$(1+v)\circ(A+\phi_2) = (A+\phi_1)\circ(1+v)\ .$$

Es hilt demnach

$$(1+v)\circ(1+u)\circ(A+\phi_1) = (A+\phi_1)\circ(1+v)\circ(1+u)\ .$$

Eine nochmalige Anwendung der obigen Betrachtung (Eindeutigkeitsaussage) mit $\phi_1 = \phi_2$ zeigt, daß $(1+v)\circ(1+u) = \mathtt{id}$. Analog ergibt sich $(1+u)\circ(1+v) = \mathtt{id}$. ■

Wir betrachten jetzt das Vektorfeld X wie im Satz 10.20. Wir können o.B.d.A. annehmen, daß $0\in U$ der betrachtete hyperbolische kritische Punktvon X ist. Sei $L\in\mathtt{End}(V)$ durch $L := dX(0)$ definiert. Sei $Z(x) := Lx$ das zu Z gehörige lineare Vektorfeld.

Satz 10.34 *Für jedes $\epsilon > 0$ existiert ein $R > 1$ und ein Vektorfeld Y auf V derart, daß*

1. *Y ist Lipschitz-stetig mit Konstanten C_Y.*
2. *Auf $V \setminus B(0,R)$ gilt $Z = Y$.*
3. *Auf $B(0, \frac{1}{2}R^{-1})$ gilt $X = Y$.*
4. *Sei Ψ der Fluß von Y und $\phi_t := \Psi_t - \exp(tZ)$. Dann gilt $\sup_{t\in[-2,2]} \|\phi_t\| =: M < \infty$*
5. *ϕ_1 ist Lipschitz-stetig mit Konstanten $< \epsilon$.*
6. *$D\phi_1(0) = 0$.*

Beweis: Wir schreiben $X = L + \psi$ mit $\psi : U \to V$. Dann gilt $\psi(0) = 0$, $D\psi(0) = 0$. Wir wählen nun eine Abschneidefunktion $\alpha \in C\infty(\mathbb{R})$ mit $\alpha(t) \in [0,1]$, $\alpha(t) = 1$ für $t \leq 1/2$, $\alpha(t) = 0$ für $t \geq 1$ und setzen $\sup_{t\in\mathbb{R}} |\alpha'(t)| =: K$.

Sei $\delta > 0$. Dann wählen wir R so groß, daß für alle $x \in B(0, R^{-1})$ $\|d\psi(x)\| \leq \frac{\delta}{2K}$ gilt. Wir setzen

$$\tilde{\psi}(x) := \alpha(\frac{\|x\|}{R})\psi(x) \ .$$

Dann gilt $\tilde{\psi} \in C(V,V)$, $\tilde{\psi}(0) = 0$ und $\tilde{\psi}$ ist Lipschitz-stetig mit Konstanten δ (da $\sup_x \|d\tilde{\psi}\| \leq \delta$). Wir setzen

$$Y(x) := Lx + \tilde{\psi}(x) \ .$$

Dann erfüllt Y die Bedingungen 1., 2. und 3. Die Bedingung 4. folgt aus einer Anwendung des Gronwall-Lemmas. Wir haben die Ungleichung

$$\|\Psi_t(x) - \Psi_t(y)\| = \|x - y + \int_0^t [X(\Psi_s(x)) - X(\Psi_s(y))]ds\| \leq \|x-y\| + C_Y \int_0^t \|\Psi_s(x) - \Psi_s(y)\|ds \ .$$

Daraus schließen wir

$$\|\Psi_t(x) - \Psi_t(y)\| \leq \mathsf{e}^{|t|C_Y}\|x-y\| \ . \tag{8}$$

Mit $x = \Psi_{-t}(u)$, $y = \Psi_{-t}(v)$ erhalten wir auch

$$\|\Psi_{-t}(u) - \Psi_{-t}(v)\| \geq \mathsf{e}^{-|t|C_Y}\|u-v\| \ .$$

Analog gilt

$$\|\mathsf{e}^{-tL}(u) - \mathsf{e}^{-tL}(v)\| \geq \mathsf{e}^{-|t|\|L\|}\|u-v\| \ .$$

Für $x \in V$ mit $\|x\| \geq \mathsf{e}^{2\max\{C_Y,\|L\|\}}R$ und alle $t \in [-2,2]$ gilt $\|\Psi_t(x)\| \geq R$ und $\|\mathsf{e}^{tL}(x) \geq R$. Für diese x gilt also $\phi_t(x) = 0$. Wenn $\|x\| \leq \{\mathsf{e}^{2\max\{C_Y,\|L\|\}}R$, so ist für $t \in [-2,2]$

$$\|\phi_t(x)\| \leq \|\Psi_t(x)\| + \|\mathsf{e}^{tL}(x)\| \leq \mathsf{e}^{4\max\{C_Y,\|L\|\}}R \ .$$

Folglich gilt $\sup_{t\in[-2,2]} \|\phi_t\| \leq \mathsf{e}^{4\max\{C_Y,\|L\|\}}R$.

Wie schätzen nun die Lipschitzkonstante von ϕ_1 ab.

$$\begin{aligned}\phi_t' &= \Psi_t' - (\mathsf{e}^{-tL})' \\ &= Y \circ \Psi_t - L \circ \mathsf{e}^{-tL} \\ &= L \circ \Psi_t + \tilde{\psi} \circ \Psi_t - L \circ \mathsf{e}^{-tL} \\ &= \tilde{\psi} \circ \Psi_t + L \circ \phi_t \end{aligned} \tag{9}$$

Wir integrieren :

$$\phi_t(x) = \phi_0(x) + \int_0^t [\tilde{\psi} \circ \Psi_s(x) + L \circ \phi_s(x)] ds \ .$$

Beachte, daß $\phi_0(x) = 0$. Damit folgt für alle $t \in [0,1]$

$$\begin{aligned}\|\phi_t(x) - \phi_t(y)\| &= \int_0^t [\|\tilde{\psi} \circ \Psi_s(x) - \tilde{\psi} \circ \Psi_s(y)\| + \|L\| \|\phi_s(x) - \phi_s(y)\| ds \\ &\leq \delta \int_0^t \|\Psi_s(x) - \Psi_s(y)\| ds + \|L\| \int_0^t \|\phi_s(x) - \phi_s(y)\| ds \\ &\leq \delta \mathsf{e}^{C_Y} \|x - y\| + \|L\| \int_0^t \|\phi_s(x) - \phi_s(y)\| ds\end{aligned}$$

Mit dem Gronwall-Lemma schließen wir

$$\|\phi_t(x) - \phi_t(y)\| \leq \|x - y\| \delta \mathsf{e}^{C_Y} \mathsf{e}^{\|L\|} \ .$$

Wir wählen nun $\delta \leq \epsilon \mathsf{e}^{-C_Y - \|L\|}$. Dann gilt 5.

Wir zeigen nun 6. indem wir für jedes $\rho > 0$ die Ungleichung

$$\|\phi(x)\| \leq \rho \|x\|, \quad \forall x \in V, \ \|x\| \leq r(\rho)$$

zeigen.

Es gilt $\tilde{\psi}(0) = 0$ und $d\tilde{\psi}(0) = d\psi(0) = 0$. Wir finden daher $\tilde{r}(\rho) > 0$ derart, daß für $\|x\| < r(\rho)$ gilt $\|\tilde{\psi}(x)\| \leq \eta \|x\|$ mit $\eta := \rho \mathsf{e}^{-C_Y - \|L\|}$ Für $t \in [0,1]$ gilt nun (siehe (8) $\|\Psi_t(x)\| \leq \mathsf{e}^{C_Y} \|x\|$ und damit für $\|x\| \leq \tilde{r}(\rho) \mathsf{e}^{-C_Y} =: r(\rho)$ die Ungleichung $\|\tilde{\psi} \circ \Psi_t(x)\| \leq \rho \mathsf{e}^{-\|L\|} \|x\|$. Wir erhalten mit (9)

$$\begin{aligned}\|\phi_t(x)\| &= \| \int_0^t \left(\tilde{\psi} \circ \Psi_s(x) + L \circ \phi_s(x)\right) ds \| \\ &\leq \rho \mathsf{e}^{-\|L\|} \|x\| + \|L\| \int_0^t \|\phi_s(x)\| ds \ .\end{aligned}$$

Mit dem Gronwall-Lemma schließen wir, daß für $\|x\| \leq r(\rho)$ die Ungleichung $\|\phi_1(x)\| \leq \rho \|x\|$ gilt. ■

Wir schließen jetzt den Beweis des Satzes 10.20 ab. Wir treffen die gleichen Vereinfachungen wie im Beweis von 10.34. Wir fixieren $\epsilon > 0$ und wählen Y und $R > 0$ wie dort. Dann gilt $X = Y$ auf $B(0, \frac{1}{2}R^{-1})$. Folglich sind $(X, 0)$ und $(Y, 0)$ zueinander orbitäquivalent. Es bleibt zu zeigen, daß $(Y, 0)$ und $(Z, 0)$ zueinander topologisch konjugiert sind.

Wir konstruieren einen Homöomorphismus $H : V \to V$ mit

$$H \circ \Psi_t = \mathsf{e}^{tL} \circ H \ .$$

Sei nun $A := \mathsf{e}^L \in Gl(V)$. Wir bemerken, daß A hyperbolisch ist und $\Psi_1 = A + \phi_1$ gilt.

ϵ ist eine Lipschitzkonstante von ϕ_1. Sei nun ϵ so klein, daß wir 10.32 anwenden können. Wir finden mit diesem Resultat einen Homöomorphismus $h : V \to V$ mit $h \circ \Psi_1 = \mathsf{e}^L \circ h$. Wir setzen

$$H := \int_0^1 \mathsf{e}^{-tL} \circ h \circ \Psi_t dt \ .$$

Wir zeigen, daß $H \circ \Psi_t = \mathsf{e}^{tL} \circ H$ für alle $t \in [0, 1]$ gilt. Daraus folgt diese Indentität für alle t wegen

$$H \circ \Psi_{mt} = H \circ \Psi_t^m = (\mathsf{e}^{-tL})^m \circ H = \mathsf{e}^{-mtL} \circ H \ .$$

Sei $t \in [0, 1]$. Dann gilt

$$\begin{aligned}
\mathsf{e}^{-tL} \circ H \circ \Psi_t &= = \int_0^1 \mathsf{e}^{-(s+t)L} \circ h \circ \Psi_{s+t} ds \\
&= \int_t^1 \mathsf{e}^{-(s)L} \circ h \circ \Psi_s ds + \int_0^t \mathsf{e}^{-sL} \circ \mathsf{e}^{-L} \circ h \circ \Psi_1 \circ \Psi_s ds \\
&= \int_t^1 \mathsf{e}^{-(s)L} \circ h \circ \Psi_s ds + \int_0^t \mathsf{e}^{-sL} \circ h \circ \Psi_s ds \\
&= H
\end{aligned}$$

Es bleibt zu zeigen, daß H ein Homöomorphismus ist. Es galt $h = 1 + u$ mit $u \in C(V, V)$ und damit

$$\begin{aligned}
H &= \int_0^1 \mathsf{e}^{-tL} \circ (1 + u) \circ (\mathsf{e}^{tL} + \phi_t) \\
&= 1 + U \\
U &:= \int_0^1 [\mathsf{e}^{-tL} \circ \phi_t + \mathsf{e}^{-tL} \circ u \circ (\mathsf{e}^{tL} + \phi_t)] dt \ .
\end{aligned}$$

Wegen 10.34,4., gilt $U \in C(V, V)$. Wir haben $h \circ \Psi_1 = \mathsf{e}^L \circ h$ und $H \circ \Psi_1 = \mathsf{e}^L \circ H$. Wegen der Eindeutigkeitsaussage von 10.32 muß also $h = H$ gelten, womit H ein Homöomorphismmus ist. ■

11 Integration im $\mathbb{R}^n$

11.1 Der Ausdehnungssatz von Tietze-Urysohn

Sei X metrischer Raum, $A \subset X$ und $f : A \to \mathbb{R}$ stetig.

Definition 11.1 *Eine stetige Funktion X heißt stetige Ausdehnung von f falls $F_{|A} = f$.*

Unter welchen Vorraussetzungen besitzt f stetige Ausdehnungen. Beispiele :

1. Sei $A = [0,1] \in \mathbb{R} = X$. Dann hat jede stetige Funktion $f : A \to \mathbb{R}$ eine stetige Ausdehnung, z. B.
$$F(x) := \begin{cases} f(0) & x \leq 0 \\ f(x) & x \in (0,1) \\ f(1) & x \geq 1 \end{cases} .$$

2. Sei $A = (0,1] \subset \mathbb{R} = X$ und $f(x) = \sin(\frac{1}{x})$. Diese Funktion hat keine stetige Ausdehnung, da schon $\lim_{x\to 0} f(x)$ nicht existiert.

Theorem 11.2 (Tietze-Urysohn) *Wenn A abgeschlossen und f beschränkt ist, dann existiert eine Ausdehnung F on f mit* $\sup_X F = \sup_A f =: S$ *und* $\inf_X F = \inf_A f =: I$.

Beweis: Da A abgeschlossen ist, folgt aus $x \notin A$ daß
$$d(x,A) := \inf_{a\in A} d(x,a) > 0 \ .$$
Wir definieren
$$F(x) := \begin{cases} f(x) & x \in A \\ \frac{\inf_{a\in A}(f(a)-I+1)d(x,a)}{d(x,A)} + I - 1 & x \notin A \end{cases} .$$
Es gilt $\frac{d(x,a)}{d(x,A)} \geq 1$ und damit $F(x) \geq I$ für alle $x \in X$. Auf der anderen Seite
$$\inf_{a\in A}(f(a) - I + 1)d(x,a) \leq \sup_{a\in A}(f(a) - I + 1) \inf_{a\in A} d(x,a) = (S - I + 1)d(x,A) \ .$$
Folglich, $F(x) \leq S$ für alle $x \in X$. Es beleibt zu zeigen, daß F stetig ist. Wir zerlegen dazu $X = \mathtt{int}(A) \cup (X \setminus A) \cup (A \setminus \mathtt{int}(A))$. Wir zeigen die Stetigkeit von F im Punkt $x_0 \in X$, wobei wir die entsprechenden drei Fälle unterscheiden.

$x_0 \in \mathtt{int}(A)$: F ist stetig in x_0, da $\mathtt{int}(A)$ offen ist, und F auf $\mathtt{int}(A)$ mit der stetigen Funktion f übereinstimmt.

$x_0 \in X \setminus A$: Wir zeigen, daß

1. $X \setminus A \ni x \mapsto d(x,A)^{-1}$ und

2. $X \setminus A \ni x \mapsto \inf_{a\in A}(f(a) - I - 1)d(x,a)$ stetig in x_0 sind.

Zu 1. Es gilt $d(x_0,A) \neq 0$, da A abeschlossen ist. Wir zeigen ,daß $x \mapsto d(x,A)$ stetig in x_0 ist. Sei $\epsilon > 0$ gegeben. Dann folgt aus $d(x_0,x) < \epsilon$
$$\begin{aligned} |d(x,A) - d(x_0,A)| &= |\inf_{a\in A} d(x,a) - \inf_{a\in A} d(x_0,a)| \\ &\leq |\inf_{a\in A} d(x_0,a) - \inf_{a\in A} d(x_0,a)| + d(x,x_0) \\ &< \epsilon \ . \end{aligned}$$

Nun zu 2. Sei $\epsilon > 0$ gegeben. Wir setzen $\delta := \frac{\epsilon}{S-I+2} \wedge d(x_0, A)$. Dann folgt aus $d(x, x_0) < \delta$ daß $x \in X \setminus A$. Weiterhin gilt

$$|\inf_{a\in A}(f(a) - I + 1)d(x, a) - \inf_{a\in A}(f(a) - I + 1)d(x_0, a)| \le (S - I + 1)d(x, x_0) < \epsilon \ .$$

$\underline{x_0 \in A \setminus \mathtt{int}(A)}$: Sei $\epsilon > 0$ egeben. Wir wählen $r > 0$ so daß aus $x \in A \cap B(x_0, r)$ folgt $|f(x) - f(x_0)| < \epsilon$.Sei $\delta := \frac{r}{4} \wedge \frac{r(S-I+1)}{2}$ und $C := A \cap B(x_0, r)$, $D := A \setminus C$. Ist $x \in X \setminus A$, $d(x, x_0) < \delta$, dann gilt für $a \in D$,daß $d(x, a) \ge d(x_0, a) - d(x, x_0) \ge r - \frac{r}{4} = \frac{3r}{4}$. Damit $\inf_{a\in D}(f(a) - I + 1)d(x, a) \ge \frac{3r}{4}$. Da aberzum Beispiel $(f(x_0) - I + 1)d(x, x_0) \le (S - I + 1)\delta \le \frac{r}{2}$ ist, gilt $\inf_{a\in A}(f(a) - I + 1)d(x, a) = \inf_{a\in C}(f(a) - I + 1)d(x, a)$. Es gelten gelten ferner $f(x_0) - \epsilon \le f(a) \le f(x_0) + \epsilon$ für $a \in C$ und $\inf_{a\in C} d(x, a) = d(x, A)$. Folglich ist

$$(f(x_0) - I + 1 - \epsilon)d(x, A) \le \inf_{a\in C}(f(a) - I + 1)d(x, a) \le (f(x_0) - I + 1 - \epsilon)d(x, A) \ .$$

Daraus schließen wir $|F(x) - F(x_0)| \le \epsilon$.

Ist $x \in A$ mit $d(x, x_0) < \delta$, so gilt auch $|F(x) - F(x_0)| \le |f(x) - f(x_0)| \le \epsilon$. Damit haben wir auch die Stetigkeit von F gezeigt.

Korollar 11.3 *Sei X ein metrischer Raum und $A, B \subset X$ nichtleere disjunkte abgeschlossene Teilmengen. Dann gibt es eine stetige Funktion $F \in C(X)$ mit $F(X) \subset [0, 1]$, $F_{|A} \equiv 1$ und $F_{|B} \equiv 0$.*

Beweis: Wir wenden den Satz 11.2 auf $f \in C(A \cup B)$ mit $f_{|A} \equiv 1$ und $f_{|B} \equiv 0$ an. ■

11.2 Zerlegung der Eins

Sei X topologischer Raum und $\mathcal{U} := (U_\lambda)_{\lambda\in L}$ eine Überdeckung von X durch offene Mengen.

Definition 11.4 *$\mathcal{U}$ heißt lokal endlich, falls jedes $x \in X$ eine Umgebung U besitzt, für welche die Menge $\{\lambda \in L \mid U_\lambda \cap U \ne \emptyset\}$ endlich ist.*

Sei $\mathcal{U}$ lokal endlich.

Definition 11.5 *Eine $\mathcal{U}$ untergeordnete Zerlegung der Eins ist eine Familie $\mathcal{P} := (f_\lambda)_{\lambda\in L}$ mit $f_\lambda \in C(X)$, $\mathtt{supp}(f_\lambda) \subset U_\lambda$ und $\sum_{\lambda\in L} f_\lambda(x) = 1$ für jedes $x \in X$.*

Hier ist ein Beispiel. Sei $X = \mathbb{R}$, $L = \mathbb{Z}$ und $U_\lambda = (\lambda - 2, \lambda + 2)$. Wir definieren $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := \begin{cases} 0 & x \le -1 \\ 1 - |x| & x \in [-1, 1] \\ 0 & x \ge 1 \end{cases} \ .$$

Dann setzen wir $f_\lambda(x) := f(x - \lambda)$.

Theorem 11.6 *Sei X ein metrischer Raum und $\mathcal{U}$ eine höchstens abzählbare lokal endliche offene Überdeckung von X. Dann existiert eine $\mathcal{U}$ untergeordnete Zerlegung der Eins.*

Beweis: Wir nehmen an, daß $L = \mathbb{N}$ (Falls L endlich war, so ergänzen wir $\mathcal{U}$ durch abzählbar viele leere Mengen). Wir setzen $V_0 := \emptyset$. Wir konstruieren induktiv offene Teilmengen $V_n \subset X$ mit $V_n \subset \bar{V}_n \subset U_n$ derart, daß für alle $m \in \mathbb{N}$ $\bigcup_{n<m} V_m \cup \bigcup_{n\geq m} U_n = X$. Seien V_n für $n \leq m$ schon konstruiert. Wir betrachten die offene Teilmenge $C := \bigcup_{n\leq m} V_n \cup \bigcup_{n\geq m+2} U_n$ von X. Es gilt dann $X \setminus C \subset U_{m+1}$. Ist $X \setminus C = \emptyset$ so setzen wir $V_{m+1} = \emptyset$. Andernfalls wählen wir mit Satz 11.2 eine Funktion $F \in C(X)$ mit $F_{|X\setminus C} \equiv 0$ und $F_{|X\setminus U_{m+1}} \equiv 1$. Dann definieren wir $V_{m+1} := F^{-1}((-\infty, 1/2))$. Als Urbild einer offenen Menge durch eine stetige Funktion ist diese Menge offen. Weiterhin gilt $\bar{V}_{m+1} \subset F^{-1}(-\infty, 1/2]) \subset U_{m+1}$.

Wir wählen jetzt mit Satz 11.2 für jedes $n \in \mathbb{N}$ ein $g_n \in C(X)$ mit $g_{n|\bar{V}_n} \equiv 1$ und $g_{n|X\setminus U_n} \equiv 0$. Weiter definieren wir $h_n := (g_n - \frac{1}{2}) \vee 0$. Dann gilt $\mathtt{supp} h_n \subset U_n$ und $h(x) := \sum_{n\in\mathbb{N}} h_n(x) > 0$ für jedes $x \in X$ (da dieses in mindestens einem V_m liegen muß, wo $h_m(x) = \frac{1}{2}$ gilt.

Die Funktion h ist stetig. In der Tat ist diese Funktion lokal eine endliche Summe stetiger Funktionen. Sei $x \in X$ und U eine Umgebung von x mit der Eigenschaft, daß die Menge $\{n \in \mathbb{N} \mid U_n \cap \neq \emptyset\}$ endlich ist, dann gilt

$$h_{|U} = \sum_{\{n\in\mathbb{N} \mid U_n\cap\neq\emptyset\}} h_{n|U} \ .$$

Wir definieren nun

$$f_n := \frac{h_n}{h} \ .$$

Die Familie $(f_n)_{n\in\mathbb{N}}$ ist eine $\mathcal{U}$ untergeordnete Zerlegung der Eins. ■

11.3 Iterierte Integrale

Sei X topologischer Raum und E, F Banachräume. Sei $A : E \to F$ eine stetige lineare Abbildung. Dann definieren wir für $f \in C(X, E)$ die Abbildung $A_* f : X \to F$ durch $A_* f(x) = A(f(x))$.

Lemma 11.7 *Es gilt $A_* f \in C(X, F)$. Die Abbildung A_* stetig und es gilt $\|A_*\| = \|A\|$.*

Beweis: $A_* f : X \to F$ ist als Komposition stetiger Abbildungen stetig. Die Beschränktheit folgt aus

$$\sup_{x\in X} \|A_* f(x)\| \leq \sup_{x\in X} \|A\|\|f(x)\| = \|A\|\|f\| \ .$$

Diese Ungleichung zeigt auch $\|A_*\| \leq \|A\|$. Durch Einsetzen konstanter Funktionen sieht man jedoch die Gleichheit $\|A_*\| = \|A\|$. ■

Sei $U \subset \mathbb{R}^n$ offen und $K \subset U$ kompakt. Sei E ein Banachraum. Wir definieren

$$C_K(U,E) := \{f \in C(U,E) \mid \mathtt{supp}(f) \subset K\} \ .$$

Lemma 11.8 *$C_K(U,E)$ is ein abgeschlossender Unterrraum von $C(U,E)$ und damit ein Banachraum.*

Beweis: Sei $f_i \in C_K(U,E)$ eine Folge mit $f_i \to f$. Dann gilt für $x \in U \setminus K$ daß $0 = f_i(x) \to f(x)$ und damit $f(x) = 0$. Wir sehen also, daß $\mathtt{supp}(f) \in K$ und damit $f \in C_K(U,E)$. ■

Sei $U' \subset X$ offen $K' \subset U'$ kompakt und weiterhin $U \subset U'$ und $K \subset K'$. Dann haben wir eine isometrische Abbildung $C_K(U,E) \to C_{K'}(U',E)$, welche durch fortsetzung durch Null gegeben wird.

Wir setzen ferner

$$C_c(U,E) := \{f \in C(U,E) \mid \mathtt{supp}(f) \text{ kompakt}\} \ .$$

Dieser Raum ist die Vereinigung der $C_K(U,E)$ für alle K. Er ist eine linearer Unterraum von $C(U,E)$, aber nicht notwendig abgeschlossen.

Wir schreiben nun $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}^1$ und schreiben die Punkte entsprechend in der Form $x = (x', x_n)$, $x_n \in \mathbb{R}$, $x' \in \mathbb{R}^{n-1}$. Seien $p' : \mathbb{R}^n \to \mathbb{R}^{n-1}$ und $p_n : \mathbb{R}^n \to \mathbb{R}$ durch $p(x) := x'$ und $p_n(x) := x_n$ gebeben. Für $K \subset R^n$ setzen wir $K_n := p_n(K)$, $K' := p'(K)$. Ist K kompakt, so sind es auch K_n, K'. Sei $f \in C(\mathbb{R}^n, E)$. Wir definieren $q(f)\mathbb{R}^{n-1} \to C(\mathbb{R}, E)$ durch $q(f)(x')(x_n) := f(x', x_n)$.

Lemma 11.9 *1. Sei K kompakt und $f \in C_K(\mathbb{R}^n, E)$. Dann gilt $q(f) \in C_{K'}(\mathbb{R}^{n-1}, C_{K_n}(\mathbb{R}, E))$.*

2. Die lineare Abbildung $q : C_K(\mathbb{R}^n, E) \to C_{K'}(\mathbb{R}^{n-1}, C_{K_n}(\mathbb{R}, E))$ ist stetig und efüllt $\|q\| = 1$.

Beweis: Wenn $x' \notin K'$ oder $x_n \notin K_n$, dann ist $(x', x_n) \notin K' \times K_n$ und damit wegen $K \subset K' \times K_n$ auch $x \notin K$. Deshalb gelten die Aussagen über die Träger.

Wir zeigen, daß $q(f)$ stetig ist. Aus der Kompaktheit von $\mathtt{supp}(f)$ folgt, daß f gleichmäßig stetig ist. Sei $\epsilon > 0$ gegeben. Wir wählen $\delta > 0$ derart, daß aus $\|x-y\| < \delta$ folgt $\|f(x) - f(y)\| < \epsilon$. Wenn nun $\|x' - y\prime\| < \delta$ ist, dann gilt wegen $\|(x', x_n) - (y', x_n)\| < \delta$ auch

$$\|q(f)(x') - q(f)(y'\| = \sup_{x_n} \|f(x', x_n) - f(y', x_n)\| < \delta \ .$$

Die Aussage über die Norm von q ergibt sich aus der Abschätzung

$$\begin{aligned} \|q(f)\| &= \sup_{x' \in \mathbb{R}^{n-1}} \|q(f)(x'\| \\ &= \sup_{x' \in \mathbb{R}^{n-1}} \sup_{x_n \in \mathbb{R}} \|f)(x', x_n)\| \\ &= \sup_{x \in \mathbb{R}^n} \|f\| \\ &= \|f\| \ . \end{aligned}$$

■

Wir erinnern jetzt an einige Eigenschaften des Integrals. Sei $K \subset \mathbb{R}$ kompakt. Wir betrachten die Abbildung $I : C_K(\mathbb{R}) \to \mathbb{R}$, welche durch

$$I(f) := \int_a^b f(x)dx$$

gegeben wird, wobei $a < b$ so geeignet gewählt werden, daß $K \subset [a, b]$.

Lemma 11.10 *$I :: C_K(\mathbb{R}) \to \mathbb{R}$ ist eine positive stetige Abbildung.*

Beweis: Wenn $f \geq 0$, so gilt auch $I(f) \geq 0$. Das ist die Positivität. Sei nun $K \subset [a, b]$. Dann gilt $I(f) \leq (b - a)\|f\|$. Damit ist I stetig. ■

Sei nun $K \subset \mathbb{R}^n$ kompakt. Mit Hilfe von Lemma 11.7 definieren wir $I_n : C_K(\mathbb{R}^n) \to C_{K'}(\mathbb{R}^{n-1})$ durch

$$I_n := I_* \circ q \ .$$

Wir schließen, daß I_n eine beschränkte lineare positive Abbildung ist.

Definition 11.11 *Wir definieren die beschränkte positive lineare Abbildung* $: C_K(\mathbb{R}^n) \to \mathbb{R}$ *durch als die Komposition*

$$C_K(\mathbb{R}^n) \stackrel{I_n}{\to} C_{K'}(\mathbb{R}^{n-1}) \stackrel{I_{n-1}}{\to} C_{(K')'}(\mathbb{R}^{n-2}) \stackrel{I_{n-2}}{\to} \cdots \stackrel{I_1}{\to} \mathbb{R} \ .$$

Um den Anschluß an allgemein übliche Bezeichunungen herzustellen, schreiben wir auch

$$I(f) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n)dx_n \ldots dx_1 \ .$$

Da Motivation der Definition von $I(f)$ ist, daß diese Zahl für $f \geq 0$ das Volumen der Menge $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid 0 \geq t \leq f(x)\}$ beschreiben sollte. Wir werden hier keine Unabhängige Definition dieses Volumen geben. Vielmehr stellen wir uns auf den Standpunkt, daß dieses Volumen durch $I(f)$ definert werde.

11.4 Vertauschung der Koordinaten

Sei G eine Gruppe.

Definition 11.12 *Eine Wirkung von G auf einer Menge X ist durch eine Familie von Abbildungen $(L_g)_{g\in G}$, $L_g : X \to X$, mit $L_g \circ L_h = L_{gh}$ und $L_1 = \mathtt{id}$ gegeben.*

Sei $(L_g)_{g\in G}$ eine Wirkung. Sei $\mathtt{Map}(X, Y)$ die Menge der Abbildungen von $X \to Y$. Wir definieren für $g \in G$ eine Abbildung $L_g^\sharp := L_{g^{-1}}^*$ auf $\mathtt{Map}(X, Y)$.

Lemma 11.13 *$(L^\sharp_g)_{g\in G}$ ist eine Wirkung von G auf $\mathtt{Map}(X,Y)$.*

Beweis: Es gilt $L^\sharp_1 = L^*_1 = \mathtt{id}$ und

$$L^\sharp_g L^\sharp_h = L^*_{g^{-1}} L^*_{h^{-1}} = L^*_{h^{-1}g^{-1}} = L^*_{(gh)^{-1}} = L^\sharp_{gh} \ .$$

■

Sei S_n die Gruppe der Permutationen von $[n] := \{1,\dots,n\}$. Wir identifizieren $\mathbb{R}^n \cong \mathtt{Map}([n],\mathbb{R})$. Dann definieren wir eine Wirkung von S_n auf $\mathbb{R}^n$ durch $L_s := s^\sharp$.

Wir definieren nun eine Wirkung von S_n auf $C(\mathbb{R}^n)$ durch $l_s := L^\sharp_s$. Es gilt offensichtlich $\|l_s\| = 1$.

Beachte, daß $L_s\mathtt{supp}(L^*_s f) = \mathtt{supp}(f)$ gilt t. Deshalb können wir $(l_s)_{s\in S_n}$ zu einer Wirkung von S_n auf $C_c(\mathbb{R}^n)$ einschränken.

Schließlich erhalten wir eine Wirkung $(l^\sharp_s)_{s\in S_n}$ auf $\mathtt{Map}(C_c(\mathbb{R}^n),\mathbb{R})$.

In diesem Kapitel zeigen wir, daß das iterierte Integral I invariant unter Vertauschung der Koordinaten ist.

Theorem 11.14 *Es für alle $s \in S_n$:*

$$l^\sharp_s(I) = I \ .$$

Beweis: Wir zeigen zuerst einen Spezialfall. Wir nehmen vorerst an, daß $n = 2$ und $s := (1,2) \in S_2$.

Lemma 11.15 $l^\sharp_s(I) = I$

Beweis: Sei $f \in C_c((\mathbb{R}^2)$. Wir zeigen, daß für jedes $\epsilon > 0$ gilt

$$|I(l_s(f)) - I(f)| < \epsilon \ .$$

Sei $\sup(f) \subset [a,b]\times[a,b]$ für geeignete $a < b$ und $M := (b-a)$. Die Funktion f ist gleichmäßig stetig. Wir wählen ein $\delta > 0$ derart, daß aus $\|x-y\| < \delta$ folgt $|f(x)-f(y)| < \frac{\epsilon}{2M^2}$. Wir approximieren nun Integrale durch Riemannsummen. Der Index kennzeichnet die Nummer der Variable, bezüglich welcher die Summe gebildet wird. Das Ergebnis ist eine Funktion der anderen Variable. So setzen wir $RS_2(f)(x) := \delta\sum_{m\in\mathbb{Z}} f(x,\delta m)$. Dann gilt $|I_2(f) - RS_2(f)| < \frac{\epsilon}{2M}$ und damit $|I(f) - I_1(RS_2(f))| < \epsilon$. Für $|x-x'| < \delta$ gilt $|RS_2(f)(x) - RS(f)(x'| < \frac{\epsilon}{2M}$. Daraus folgt

$$|I_1(RS_2(f)) - RS_1(RS_2(f))| < \frac{1}{2\epsilon} \ .$$

Da eine endlich Summe unabhängig von der Reihenfolge der Summanden ist, gilt offensichtlich

$$RS_1(RS_2(l_s(f))) = RS_2(RS_1(f) = RS_1(RS_2(f)) \ .$$

Daraus schließen wir

$$|I(l_s(f)) - I(f)| < \epsilon \ .$$

■

Wir folgern jetzt den Satz. Eine Permutation $s \in S_n$ heißt elementar, falls sie genau ein Paar benachbarter Elemente vertauscht, das heßt, es gibt ein $i \in \{1, \dots n-1\}$ so daß $s(k) = k$ für $k \leq i-1$ oder $k \geq i+1$, und $s(i) = i+1$, $s(i+1) = i$ gilt. Wir benutzen nun die Tatsache, daß S_n durch elementare Permutationen erzeugt wird, daß also jedes Permutation als Produkt von elementaren Permutationen geschrieben werden. Wir mussen die Gleichung $l_s^\sharp I = I$ nur für eine S_n erzeugende Menge von Permutationen zeigen. Es genügt also, diese Gleichung für elementare Permutationen einzusehen.

Sei s elementar und vertausche die Zahlen $i, i+1$. Sei $p \in S_{i+1}$ die elementare Permutation, welche $i, i+1$ vertauscht. Dann gilt

$$\begin{aligned}
l_s^\sharp I &= I \circ l_s^* \\
&= I_1 \circ I_2 \circ \dots \circ I_n \circ l_s^* \\
&= I_1 \circ I_2 \circ \dots \circ (I_i \circ I_{i+1} \circ l_p^*) \circ I_{i+2} \circ \dots \circ I_n \\
&= I_1 \circ I_2 \circ \dots \circ I_i \circ I_{i+1} \circ I_{i+2} \circ \dots \circ I_n \\
&= I \ .
\end{aligned}$$

■

11.5 Transformationsformel

Seien $U, V \subset \mathbb{R}^n$ offen.

Definition 11.16 *Ein C^1-Diffeomorphismus von U nach V ist eine stetig-differenzierbare Bijektion $\Phi : U \to V$ mit $\det(J\Phi(x)) \neq 0$ für alle $x \in U$*

Zur Erinnerung, $J\Phi$ bezeichnet die Jacobimatrix von Φ. Ist $\Phi : U \to V$ ein C^1-Diffeomorphismus, so ist nach dem Satz über die Umkehrfunktion die inverse Abbildung $\Phi^{-1} : V \to U$ auch ein C^1-Diffeomorphismus.

Ist $K \subset V$ kompakt, so ist $\Phi^{-1}(K) \subset U$ auch kompakt. Zurückziehen liefert eine Abbildung $\Phi^* : C_K(V) \to C_{\Phi^{-1}(K)}(U)$. Folglich haben wir auch $\Phi^* : C_K(\mathbb{R}^n) \to C_{\Phi^{-1}(K)}(\mathbb{R}^n)$.

Theorem 11.17 (Transformationsformel) *Seien $U, V \subset \mathbb{R}^n$ offen, $K \subset V$ kompakt und $\Phi : U \to V$ ein C^1-Diffeomorphismus. Für $f \in C_K(\mathbb{R}^n)$ gilt*

$$I(\Phi^*(f) \, | \det(J\Phi)|) = I(f) \ .$$

Beweis: Wir führen Induktion nach n durch. Sei also der Satz für alle kleineren Dimensionen bewiesen.

Wir betrachten nun die Dimension n. Wir zeigen diese Formel zuerst unter der Annahme, daß Φ spezielle Formen hat. Dann schließen wir daraus den allgemeinen Fall.

Lemma 11.18 *Der Satz stimmt, falls $Phi(x) = (x', \phi(x))^t$ für eine Funktion $\phi : U \to \mathbb{R}$.*

In diesem Fall hat die Jacobi Matrix eine Blockform

$$\begin{pmatrix} (n-1)x(n-1) & 1x(n-1) \\ (n-1)x1 & 1x1 \end{pmatrix} .$$

Es gilt

$$J\Phi = \begin{pmatrix} 1 & 0) \\ d'\phi & \partial_n\phi \end{pmatrix} . .$$

Wir sehen, daß $|\det(J\Phi)| = |\partial_n\phi|$ gilt. Die eindimensionale TRansformationsformel gibt

$$I(\Phi^*(f)\,|\det(J\Phi)|) = I_n(\Phi^* f|\partial_n\phi|) = I_n(f) .$$

■

Lemma 11.19 *Der Satz stimmt, wenn $\Phi(x) = (x_1, \psi(x))$ für ein $\psi : U \to \mathbb{R}^{n-1}$ ist.*

Beweis: In diesem Fall betrachten wir die Blockstruktur

$$\begin{pmatrix} 1x1 & 1x(n-1) \\ (n-1)x1 & (n-1)x(n-1) \end{pmatrix} .$$

$$J\Phi = \begin{pmatrix} 1 & 0 \\ \partial_1\psi & J'\psi \end{pmatrix} ,$$

wobei $J'\Psi$ die Jacobimatrix von ψ bei festgehaltener 1-ter Veränderlicher bezeichnet. Wir sehen, daß $|\det(J\Phi)\| = |\det(J'\psi)|$. Wir schreiben $I = I_1 \circ I'$. wobei I' das $n-1$-dimensionale Integral bezeichnet. Dann gilt nach Induktionsvoraussetzung

$$I(\Phi^*(f)\,|\det(J\Phi)|) = I' \circ I_1(I'(\Phi^*(f)\,|\det(J'\psi)|) = I^1 \circ I'(f) = I(f) .$$

■

Wir beweisen jetzt die Transformationsformel im allgemeinen Fall. Sei $y \in U$. Dann ist $\mathtt{rk}J\Phi(y) = n$. Folglich können wir nach umnumerieren der Koordinaten annehmen, daß $\det J'\psi(y) \neq 0$ ist. Hierbei verwenden wir die Zerlegung $\Phi(x) := (\Phi'(x), \Phi_n(x))$. Wir definieren $\Psi_1(x) := (\Phi'(x), x_n)$. Dann ist $\det(J\Psi_1)(y) = \det(J'\psi)(y) \neq 0$. Nach dem Satz über die Umkehrfunktion gibt es eine Umgebung $U_y \subset U$ von x derart, daß $\Psi_1 : U_y \to V_y$ invertierbar ist. Wir definieren nun $\Psi_2 : V_x \to U$ durch $\Psi_2 := \Phi \circ \Psi_1^{-1}$. Dann ist $\Psi_2 : V_y \to W_y := \Psi_2(V_y)$ auch invertierbar. Weiterhin gilt $\Phi = \Psi_2 \circ \Psi_1$. Offensichtlich ist $\Psi_2(z) = (z', \phi(z))$ für eine Abbildung $\phi : V_y \to \mathbb{R}$. Es gilt

$$\det J(\Phi) = \Psi_1^*[\det J(\Psi_2)] \det J(\Psi_1) .$$

Sei nun $\texttt{supp}(f) \subset V_y$, dann gilt wegen der Lemmata 11.18 und 11.19

$$\begin{aligned}I(\Phi^* f)|\det(J\Phi) &= I(\Psi_1^*(\Psi_2^* f)|\Psi_1^*[\det J(\Psi_2)]| \; |\det J(\Psi_1)| \\ &= I(\Psi_2^* f)|\det J(\Psi_2)|) \\ &= I(f) \; .\end{aligned}$$

Wir führen die obige Konstruktion für jedes $y \in U$ durch. Dann ist $\{V_y\}_{y\in U}$ eine offene Überdeckung von V. Da $K \subset V$ kompakt ist, existiert eine endliche Menge $\{y_1, \ldots, y_r\} \subset U$ derart, daß $K \subset \bigcup_{i=1}^r V_{y_i}$. Wir wählen eine $\mathcal{U}$ untergeordnete Zerlegung der Eins $(\chi_i)_{i=1,\ldots,r}$. Dann schreiben wir $f = \sum_{i=1}^r f_i$ mit $f_i := \chi_i f$. Da $\texttt{supp}(f_i) \subset V_{y_i}$ ist, gilt für alle $I = 1, \ldots, r$ $I(\Phi^* f_i)|\det(J\Phi)| = I(f_i)$. Indem wir die Summe über $i = 1, \ldots, r$ bilden, erhalten wir

$$I(\Phi^* f|\det(J\Phi)|) = I(f) \; .$$

■

11.6 Beispiele

1. Sei $p(x) = x^2 - 1$ und $f : \mathbb{R}^2 \to \mathbb{R}$ durch $f(x,y) := p(x)p(y)$ für $(x,y) \in [-1,1] \times [-1,1]$ und durch $f(x,y) := 0$ für $(x,y) \not\in [-1,1] \times [-1,1]$ gegeben. Dann ist $f \in C_{[-1,1]\times[-1,1]}(\mathbb{R}^2)$. Wir berechnen

$$\begin{aligned}I(f) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx \\ &= \int_{-1}^{1}\int_{-1}^{1} p(x)p(y)dydx \\ &= \int_{-1}^{1} |_{-1}^{1} p(x)(\frac{1}{3}x^3 - x)dx \\ &= \int_{-1}^{1} |_{-1}^{1}(x^2-1)(\frac{2}{3} - 2)dx \\ &= -\frac{2}{3}|_{-1}^{1}(\frac{1}{3}x^3 - x) \\ &= (\frac{2}{3})^2\end{aligned}$$

2. Sei $h(r) = (r-1)(r-1/2)$ für $r \in [1/2,1]$, $r = 0$ für $r \leq 1/2$ und $r \geq 1$. Wir definieren $f(x,y) = h(x^2+y^2)\sin(2\arctan(x/y))$ für $x, y \geq 0$ und $f(x,y) := 0$ sonst. Das Integral $I(f)$ berechnet man am besten in Polarkoordinaten. Sei $\Phi : (0,\infty) \times (-\pi,\pi) \to \mathbb{R}^2$ durch $\Phi(r,\alpha) := (r\cos(\alpha), r\sin(\alpha))^t$ gegeben. Φ is ein C^1-Diffeomorphismus und $\texttt{supp}(f) \subset \texttt{im}(\Phi)$. Wir berechnen

$$J\Phi(r,\alpha) = \begin{pmatrix} \cos(\alpha) & -r\sin(\alpha) \\ \sin(\alpha) & r\cos(\alpha) \end{pmatrix} .$$

Daraus folgt $\det J(\Phi)(r,\alpha) = r\cos^2(\alpha) + r\sin^2(\alpha) = r$. Es gilt

$$I(f) = I(\Phi^* f|\det J(\Phi)|)$$

Nun ist aber $\Phi^* f(r,\alpha) = (r-1)(r-1/2)\sin(2\alpha)$ für $(r,\alpha) \in (1/2,1) \times (0,\pi/2)$ und $\Phi^* f(r,\alpha) = 0$ sonst. Es ergibt sich

$$\begin{aligned} I(f) &= \int_{1/2}^{1}\int_{0}^{\pi/2} r(r-1)(r-1/2)\sin(2\alpha)d\alpha dr \\ &= |_0^{\pi/2}\frac{-1}{2}\cos(2\alpha)\int_{1/2}^{1}(r^3 - 3/2r^2 + r/2)dr \\ &= |_{1/2}^{1}(1/4r^4 - 1/2r^3 + 1/4r^2) \\ &= (1/4 - 1/2 + 1/4 - 1/64 + 1/16 - 1/16) \\ &= -1/64\ . \end{aligned}$$

References

[BF91] Martin Barner and Friedrich Flohr. *Analysis. I.* de Gruyter Lehrbuch. [de Gruyter Textbook]. Walter de Gruyter & Co., Berlin, fourth edition, 1991.

[Die85] J. Dieudonné. *Grundzüge der modernen Analysis. Band 1*, volume 8 of *Logik und Grundlagen der Mathematik [Logic and Foundations of Mathematics]*. Friedr. Vieweg & Sohn, Braunschweig, third edition, 1985. Translated from the English by Ludwig Boll and Klaus Matthes.

[Hal76] Paul R. Halmos. *Naive Mengenlehre*. Vandenhoeck & Ruprecht, Göttingen, 1976. Vierte Auflage, Aus dem Englischen übersetzt von Manfred Armbrust und Fritz Ostermann, Moderne Mathematik in elementarer Darstellung, No. 6. 1.7

[Heu91] Harro Heuser. *Lehrbuch der Analysis. Teil 1.* Mathematische Leitfäden. [Mathematical Textbooks]. B. G. Teubner, Stuttgart, ninth edition, 1991.

[Rud80] Walter Rudin. *Analysis*. Physik Verlag, Weinheim, 1980. Translated from the third English edition by Martin Lorenz and Christine Eule. 2.1, 1