Eberhard Freitag

Analysis

Teil II

Differential- und Integralrechnung
für Funktionen mehrerer reellen Veränderlichen

Self-Publishing, 2019

Eberhard Freitag
Universität Heidelberg
Mathematisches Institut
Im Neuenheimer Feld 205
69 120 Heidelberg
freitag@mathi.uni-heidelberg.de

Inhalt

1. Vorwort

Kapitel V. Funktionen auf metrischen Räumen

1. Metrische Räume

1.1 Definition. *Eine **Metrik** d auf einer Menge X ist eine Abbildung* `metR`

$$d : X \times X \longrightarrow \mathbb{R}$$

*(die sogenannte **Distanzfunktion**) mit den Eigenschaften*

a) $d(x,y) \geq 0$ *für alle* $x, y \in X$ *und* $d(x,y) = 0 \Longleftrightarrow x = y$.

b) $d(x,y) = d(y,x)$ *(Symmetrie).*

c) $d(x,z) \leq d(x,y) + d(y,z)$ *für alle* $x, y, z \in X$ *(Dreiecksungleichung).*

Ein *metrischer Raum* (X,d) ist eine Menge X, auf der eine bestimmte Metrik d ausgezeichnet ist.

Die Elemente eines metrischen Raumes nennt man häufig *Punkte*.

Beispiele metrischer Raüme.

(Der Leser möge die Axiome a)–c) jeweils nachweisen!)

1) $X = \mathbb{R}$ mit $d(x,y) = |x - y|$.

2) $X = \mathbb{R}^n$. In diesem Fall gibt es mehrere wichtige Metriken, die sämtlich Verallgemeinerungen der Betragsmetrik aus 1) sind.

   a) $d(x,y) = \max_{1 \leq \nu \leq n} |x_\nu - y_\nu|$ (Maximumsmetrik).

   b) $d(x,y) = \left(\sum_{\nu=1}^n (x_\nu - y_\nu)^2\right)^{\frac{1}{2}}$ (EUKLIDische Metrik).

Die Dreiecksungleichung ist in diesem Fall eine Folgerung aus einer weiteren Ungleichung, die wir in Paragraph 5 beweisen werden, der sogenannten CAUCHY-SCHWARZschen Ungleichung:

$$\left|\sum_{\nu=1}^n x_\nu y_\nu\right| \leq \sqrt{\sum_{\nu=1}^n x_\nu^2}\sqrt{\sum_{\nu=1}^n y_\nu^2}.$$

   c) $d(x,y) = \sum_{\nu=1}^n |x_\nu - y_\nu|$.

3) Sei X die Menge aller beschränkten Funktionen $f : D \to \mathbb{R}$, wobei D eine nicht leere Menge sei, beispielsweise ein Intervall.

Schreibweise. $X = B(D)$.

Die Norm einer Funktion f ist durch

$$\|f\| = \sup_{x \in D} |f(x)|$$

definiert (Supremum der Beträge des Wertevorrats von f). Man erhält hieraus eine Metrik

$$d(f, g) := \|f - g\|.$$

4) Sei $D = [a, b],\ a < b$, ein abgeschlossenes Intervall und sei

$$X = C(D) := \{f : D \longrightarrow \mathbb{R};\ f \text{ stetig}\}$$

die Menge der stetigen Funktionen auf D. Man erhält eine Metrik durch

$$d(f, g) = \int_a^b |f(x) - g(x)|\, dx.$$

Wir kehren nun zur allgemeinen Situation zurück und führen den fundamentalen Begriff der *Kugel* in einem metrischen Raum ein.

1.2 Definition. *Sei (X, d) ein metrischer Raum und $x_0 \in X$ ein Punkt aus* ofK
X, sowie $r > 0$ eine positive Zahl. Die Punktmenge

$$U_r(x_0) = \{x \in X;\ d(x, x_0) < r\}$$

heißt offene Kugel um x_0 vom Radius r.

Anmerkung. Es gilt stets: $x_0 \in U_r(x_0)$. Manchmal nennt man $U_r(x_0)$ auch die *r-Umgebung* von x_0.

Beispiele.

1) $\mathbb{R}^n$ mit der Maximumsmetrik: $U_r(x_0)$ ist ein achsenparalleler Würfel mit Mittelpunkt x_0 und Kantenlänge $2r$.

2) $\mathbb{R}^n$ mit der Euklidischen Metrik: $U_r(x_0)$ ist eine Euklidische Kugel mit Mittelpunkt x_0 und Radius r.

Grundtatsachen über Kugeln

1.3 Bemerkung. *Sei $U_r(x_0)$ eine Kugel in einem metrischen Raum (X, d) und sei* KugOff

$$x_1 \in U_r(X_0)$$

ein beliebiger Punkt aus dieser Kugel. Es gibt eine Zahl $\varepsilon > 0$ mit

$$U_\varepsilon(x_1) \subset U_r(x_0).$$

Beweis. Man wähle etwa $\varepsilon := r - d(x_1, x_0)$. Diese Zahl ist in der Tat positiv (weil $x_1 \in U_r(x_0)$). Aus $x \in U_\varepsilon(x_1)$ folgt

$$d(x, x_1) < \varepsilon = r - d(x_0, x_1),$$

also insbesondere

$$d(x, x_0) < r,$$

denn wegen der Dreiecksungleichung und der Symmetrie gilt

$$d(x, x_0) \leq d(x, x_1) + d(x_1, x_0) = d(x, x_1) + d(x_0, x_1). \qquad \square$$

1.4 Bemerkung. *Gegeben seien n Kugeln* nKug

$$U_{r_1}(x_1), \ldots,\ U_{r_n}(x_n)$$

in einem metrischen Raum und ein beliebiger weiterer Punkt x. Es gibt dann eine Zahl $r > 0$, so daß gilt:

$$U_{r_1}(x_1) \cup \ldots \cup U_{r_n}(x_n) \subset U_r(x).$$

Beweis. Man wähle

$$r = \max_{1 \leq \nu \leq n} \{r_\nu + d(x_\nu, x)\}. \qquad \square$$

1.5 Bemerkung. *Seien x, x' zwei verschiedene Punkte eines metrischen Raumes. Es existiert dann eine positive Zahl $\varepsilon > 0$ mit* Haus

$$U_\varepsilon(x) \cap U_\varepsilon(x') = \emptyset \qquad \textit{(Punktetrennungseigenschaft).}$$

Beweis. Man wähle etwa $\varepsilon = \frac{1}{2} d(x, x')$. $\square$

Einige topologische Begriffe in metrischen Räumen.

1.6 Definition. *Sei $x \in X$ ein Punkt eines metrischen Raumes (X, d). Eine Teilmenge $M \subset X$ heißt* ***Umgebung*** *von x, wenn es eine Zahl $\varepsilon > 0$ gibt mit* UmG

$$U_\varepsilon(x) \subset M.$$

Insbesondere ist $U_\varepsilon(x)$ selbst eine Umgebung von x. Offenbar ist der Durchschnitt von endlich vielen Umgebungen von x auch eine Umgebung von x (denn sind $M_1, \dots, M_n$ Umgebungen von x, so existieren positive Zahlen $\varepsilon_1, \dots, \varepsilon_n$ mit $U_{\varepsilon_\nu}(x) \subset M_\nu$ für $\nu = 1, \dots, n$ und es gilt

$$U_\varepsilon(x) \subset M_1 \cap \ldots \cap M_n \text{ mit } \varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\} \text{).}$$

Man kann nun Bemerkung 1.3 auch so ausdrücken:

Die Kugel $U_r(x_0)$ ist Umgebung eines jeden Punktes x, den sie enthält.

Diese Eigenschaft einer Kugel hat sich als äußerst fundamental erwiesen, so daß man ihr einen eigenen Namen gegeben hat.

1.7 Definition. *Eine Teilmenge U eines metrischen Raumes (X, d) heißt* ***offen****, wenn sie Umgebung eines jeden in ihr enthaltenen Punktes ist.* defOf

Das bedeutet also: Ist $x \in U$, so existiert ein $\varepsilon > 0$ mit

$$U_\varepsilon(x) \subset U.$$

Wie schon bemerkt: Die Kugel $U_r(x_0)$ ist offen. Dies steht im Einklang mit der bereits verwendeten Bezeichnung *offene* Kugel (Definition 1.2).

Grundeigenschaften offener Mengen

a) Die leere Menge $\emptyset$ ist offen und ebenso der ganze Raum X.

b) Sind $U_1, \dots, U_n$ (endlich viele) offene Mengen, so ist auch

$$U_1 \cap \ldots \cap U_n$$

offen.

c) Ist $(U_i)_{i \in I}$ eine Schar offener Mengen (I eine beliebige Indexmenge), so ist auch die Vereinigungsmenge

$$\bigcup_{i \in I} U_i = \{\, x \in X; \quad x \in U_i \text{ für (mindestens) ein } \ i \,\}$$

offen.

1.8 Definition. *Sei $A \subset X$ eine Teilmenge eines metrischen Raumes (X, d). Ein Punkt $x \in X$ heißt* ***Randpunkt*** *von A, wenn es in jeder Umgebung von x sowohl Punkte gibt, die in A liegen, als auch solche, die nicht in A liegen.* defRa

Also: Zu jedem $\varepsilon > 0$ existieren Punkte x', x'' mit

$$d(x, x') < \varepsilon,\ d(x, x'') < \varepsilon,\ x' \in A,\ x'' \notin A.$$

Beispiel.
$X = \mathbb{R}$ (mit der üblichen Metrik $d(x, y) = |x - y|$). Seien $a < b$ zwei Zahlen und

$$A = (a, b) \quad \text{oder} \quad (a, b] \quad \text{oder} \quad [a, b) \quad \text{oder} \quad [a, b].$$

In allen vier Fällen sind a und b die beiden einzigen Randpunkte von A.

Man sieht an diesem Beispiel, daß die Randpunkte einer Menge A dieser Menge angehören können, daß dies aber nicht sein muß. Insofern ist das *abgeschlossene* Intervall $[a, b]$ also dadurch vor den übrigen ausgezeichnet, daß es alle seine Randpunkte enthält. Dies gibt Anlaß zu folgender Definition.

1.9 Definition. *Eine Teilmenge $A \subset X$ eines metrischen Raumes (X, d) heißt* ***abgeschlossen****, wenn jeder Randpunkt von A in A enthalten ist.* defAB

Bezeichnungen.

$$\partial A = \text{Rand von } A = \text{Menge aller Randpunkte von } A,$$
$$\bar{A} = A \cup \partial A = \text{Abschluß von } A.$$

Übungsaufgaben.

1) Die Menge $\bar{A}$ ist abgeschlossen (d.h. $\bar{\bar{A}} = \bar{A}$).

2) $A \subset B \Longrightarrow \bar{A} \subset \bar{B}$.

3) Die abgeschlossene Kugel

$$\bar{U}_r(x_0) := \{x \in X;\ d(x, x_0) \le r\}$$

ist abgeschlossen.

Aus 1) und 2) folgt, daß der Abschluß der *offenen Kugel* in der abgeschlossenen Kugel enthalten ist: $\overline{U_r(x_0)} \subset \bar{U}_r(x_0)$. In vielen Fällen gilt Gleichheit, aber nicht immer.

1.10 Hilfssatz. *Eine Teilmenge $A \subset X$ eines metrischen Raumes ist genau dann abgeschlossen, wenn ihr Komplement* oKb

$$X - A := \{x \in X;\ x \notin A\}$$

offen ist.

Beweis, 1. Teil. A sei abgeschlossen. Wir zeigen, daß $X - A$ offen ist. Sei dazu $x \in X - A$. Da $x \notin A$ und da A abgeschlossen ist, kann x kein Randpunkt von A sein. Es muß daher ein $\varepsilon > 0$ geben, so daß $U_\varepsilon(x)$ nicht Punkte von A *und* von $X - A$ enthalten kann. Da nun aber x in $X - A$ liegt, muß somit

$$U_\varepsilon(x) \subset X - A$$

gelten, d.h. $X - A$ ist Umgebung von x.

2. Teil. Sei $X - A$ offen. Wir zeigen daß A abgeschlossen ist, daß also kein Punkt $x \in X - A$ ein Randpunkt von A ist. Das ist aber klar, denn es existiert $\varepsilon > 0$ mit

$$U_\varepsilon(x) \subset X - A \Longleftrightarrow U_\varepsilon(x) \cap A = \emptyset. \qquad \square$$

Abschließend noch eine weitere *Sprechweise*:

Ein Punkt $x \in M$ (M eine Teilmenge eines metrischen Raumes (X, d)) heißt *innerer* Punkt von M, wenn M Umgebung von x ist. Die Menge der inneren Punkte von M wird mit

$$M^\circ := \{\, x \in M; \quad M \text{ ist Umgebung von } x \,\}$$

bezeichnet. Offenbar ist M° offen und es gilt

$$M \text{ ist offen} \Longleftrightarrow M^\circ = M.$$

2. Konvergenz und Stetigkeit in metrischen Räumen

2.1 Definition. *Eine Folge $(x_n)_{n \in \mathbb{N}}$ von Punkten aus einem metrischen Raum (X, d) konvergiert gegen $x \in X$, wenn die Zahlenfolge $d(x_n, x)$ eine Nullfolge ist.* koFMe

Die Folge (x_n) heißt konvergent, wenn es ein $x \in X$ gibt, do daß (x_n) gegen x konvergiert.

Offenbar konvergiert die Folge (x_n) genau dann gegen x, wenn für jede Umgebung U von x gilt:

$$x_n \in U \text{ für alle } n \text{ bis auf endlich viele Ausnahmen.}$$

Der Grenzwert x ist durch die Folge (x_n) eindeutig bestimmt, denn würde (x_n) gegen zwei verschiedene Grenzwerte x, x' konvergieren, so könnte man Umgebungen U von x und U' von x' mit leerem Durchschnitt finden (Bemerkung 1.5).

Dann können aber nicht alle x_n bis auf endlich viele Ausnahmen sowohl in U als auch in U' liegen.

Bezeichnung.

$$x = \lim_{n\to\infty} x_n \quad \text{oder} \quad x_n \longrightarrow x \text{ für } n \longrightarrow \infty.$$

Beispiel.

Eine Folge von Punkten $x^{(k)} \in \mathbb{R}^n$ konvergiert bezüglich der Maximumsmetrik genau dann gegen $x \in \mathbb{R}^n$, wenn

$$\max_{1\le\nu\le n} |x_\nu^{(k)} - x_\nu| \longrightarrow 0 \text{ für } k \longrightarrow \infty$$

gilt. Dies bedeutet nichts anderes, als daß

$$\lim_{n\to\infty} x_\nu^{(k)} = x_\nu \text{ für jedes } \nu \in \{1,\dots,n\}.$$

Vergleich verschiedener Metriken.

2.2 Definition. *Zwei Metriken d, d' auf einer Menge X heißen **(streng) äquivalent**, wenn es Konstanten C, C' gibt mit* Metaeq

$$d(x,y) \le C'd'(x,y) \quad \textit{sowie} \quad d'(x,y) \le Cd(x,y).$$

Äquivalente Metriken sind in Bezug auf die Konvergenz von Folgen nicht zu unterscheiden,

$$d'(x_n,x) \longrightarrow 0 \quad \Longleftrightarrow \quad d(x_n,x) \longrightarrow 0 \quad \text{für } n \longrightarrow \infty.$$

Beispiel.

Die Maximumsmetrik und die Euklidische Metrik des $\mathbb{R}^n$ sind äquivalent:

$$\max_{1\le\nu\le n} |x_\nu - y_\nu| \le \left(\sum_{\nu=1}^n |x_\nu - y_\nu|^2\right)^{\frac{1}{2}} \le \sqrt{n} \max_{1\le\nu\le n} |x_\nu - y_\nu|.$$

Folgende beiden Aussagen für eine Folge $x^{(k)}$ im $\mathbb{R}^n$ sind also gleichbedeutend:

1) $x^{(k)}$ konvergiert *komponentenweise* gegen x, d.h. für jede Koordinate ν gilt

$$x_\nu^{(k)} \longrightarrow x_\nu \text{ für } k \longrightarrow \infty.$$

2) Der Euklidische Abstand $d(x^{(k)}, x)$ konvergiert gegen Null.

2.3 Definition. *Seien (X, d) und (Y, d') zwei metrische Räume und sei $f : X \longrightarrow Y$ eine Abbildung. Diese heißt* ***stetig*** *in einem Punkt $x_0 \in X$, wenn folgendes gilt: Ist V eine Umgebung von $y_0 = f(x_0)$ in Y, so ist $f^{-1}(V)$ eine Umgebung von x_0 in X.* metST

Übersetzen wir diese Definition in die „Sprache“ der Epsi-Deltalontik.

2.4 Bemerkung. *Die Abbildung f ist genau dann stetig in x_0, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit* epDE

$$d'(f(x), f(x_0)) < \varepsilon \text{ für alle } x \in X \text{ mit } d(x, x_0) < \delta.$$

Beweis. Wir zeigen nur eine Richtung: Sei f stetig in x_0 im Sinne der Definition 2.3 und sei $\varepsilon > 0$. Die Menge $V := U_\varepsilon(f(x_0))$ ist eine Umgebung von $f(x_0)$. Daher ist $f^{-1}(V)$ eine Umgebung von x_0 in X. Nach Definition existiert dann eine Zahl $\delta > 0$ mit

$$U_\delta(x_0) \subset f^{-1}(V) \Longrightarrow f\left(U_\delta(x_0)\right) \subset V.$$

Also gilt

$$x \in U_\delta(x_0) \Longrightarrow f(x) \in U_\varepsilon(f(x_0))$$

oder

$$d(x, x_0) < \delta \Longrightarrow d'(f(x), f(x_0)) < \varepsilon. \qquad \square$$

2.5 Bemerkung. *Eine Abbildung $f : X \to Y$ zweier metrischer Räume ist dann und nur dann stetig (d.h. stetig in jedem Punkt von X), wenn das Urbild $f^{-1}(V)$ jeder offenen Menge V aus Y offen in X ist.* StOf

Beweis. Übungsaufgabe (man benutze direkt Definition 2.3). $\square$

2.6 Hilfssatz. *Seien X, Y, Z metrische Räume und* mePS

$$f : X \longrightarrow Y, \; g : Y \longrightarrow Z$$

Abbildungen sowie x ein Punkt aus X mit der Eigenschaft

a) *f ist stetig in x,*
b) *g ist stetig in $f(x)$.*

Dann ist

$$g \circ f : X \longrightarrow Z \qquad (g \circ f(x) := g(f(x)))$$

stetig in x.

Kurz gesagt bedeutet dies: Die Zusammensetzung stetiger Funktionen ist stetig. Der Beweis ist klar und kann übergangen werden.

Zwischen Folgenkonvergenz und Stetigkeit besteht ein enger Zusammenhang.

2.7 Satz. *Seien $f : X \to Y$ eine Abbildung metrischer Räume und $x \in X$ ein Punkt. Die Abbildung f ist dann und nur dann stetig in x, wenn gilt:* meFoSt

> *Für jede Folge $x_n \in X$, die gegen x konvergiert, konvergiert die Bildfolge $f(x_n)$ gegen $f(x)$, also*
> $$x_n \longrightarrow x \Longrightarrow f(x_n) \longrightarrow f(x).$$

Beweis. Sei zunächst f stetig und die Folge x_n aus X konvergiere gegen x. Es ist zu zeigen, daß $f(x_n)$ gegen $f(x)$ konvergiert. Sei hierzu $V \subset Y$ eine Umgebung von $f(x)$. Dann ist $f^{-1}(V) \subset X$ eine Umgebung von x (weil f stetig in x ist). Daher gilt $x_n \in f^{-1}(V)$ für fast alle n (d.h. alle bis auf endlich viele Ausnahmen) und hieraus folgt

$$f(x_n) \in V \text{ für fast alle } n.$$

Nun sei f nicht stetig in $x \in X$. Wir konstruieren eine Folge (x_n) aus X, die gegen X konvergiert, ohne daß $(f(x_n))$ gegen $f(x)$ konvergiert. Da f in x unstetig ist, existiert eine Umgebung V von $f(x)$, so daß $f^{-1}(V) \subset X$ *keine* Umgebung von x ist. Zu jedem $\varepsilon > 0$ existiert daher ein Punkt aus $U_\varepsilon(x)$, der nicht in $f^{-1}(V)$ liegt. Wählt man dann speziell

$$\varepsilon = 1,\ \frac{1}{2},\ \frac{1}{3}, \ldots,$$

so erhält man durch Auswahl eines x_n für jedes n eine Folge

$$x_n \in X,\ x_n \notin f^{-1}(V),\ x_n \in U_{\frac{1}{n}}(x).$$

Diese konvergiert gegen x, denn es ist ja $d(x_n, x) < \frac{1}{n}$. Aber $f(x_n)$ konvergiert nicht gegen $f(x)$, denn sonst müßte ja $f(x_n) \in V$ für fast alle n gelten.

□

2.8 Definition. *Ein Punkt x eines metrischen Raumes X heißt* ***isoliert****, wenn es eine Zahl $\varepsilon > 0$ gibt, so daß $U_\varepsilon(x)$ nur aus dem Punkt x allein besteht.* meIS

Anders ausgedrückt: x ist genau dann isoliert, wenn $\{x\}$ eine offene Menge ist.

2.9 Definition. *Seien X und Y metrische Räume und sei $a \in X$ ein nicht isolierter Punkt. Weiter sei eine Abbildung* meGRS

$$f : X - \{a\} \longrightarrow Y \quad \textit{oder} \quad f : X \longrightarrow Y$$

gegeben. Dann sagt man: Die Funktion f besitzt den Grenzwert b für x gegen a, in Zeichen

$$f(x) \longrightarrow b \textit{ für } x \longrightarrow a \quad \textit{oder} \quad \lim_{x \to a} f(x) = b,$$

wenn die Abbildung

$$\tilde{f} : X \longrightarrow Y,\ \tilde{f}(x) = \begin{cases} f(x) & \textit{für } x \neq a \\ b & \textit{für } x = a. \end{cases}$$

in $x = a$ stetig ist. Zu jedem $\varepsilon > 0$ existiert daher ein $\delta > 0$ mit

$$d(x, a) < \delta,\ x \neq a \Longrightarrow d'(f(x), b) < \varepsilon.$$

Hierbei muß klar sein, daß b eindeutig bestimmt ist (wobei dann eingeht, daß a nicht isoliert ist). Der *Beweis* der Eindeutigkeit von b beruht auf der Punktetrennungseigenschaft Bemerkung 1.5. Es seien zwei verschiedene stetige Fortsetzungen $\tilde{f}$ und f^* von $f|(X - \{a\})$ auf X gegeben mit

$$\tilde{f}(a) = b \quad \text{und} \quad f^*(a) = b^*.$$

Wir schließen indirekt, nehmen also an, b und b^* seien verschieden. Nach Bemerkung 1.5 gibt es dann disjunkte Umgebungen V, V^* von b bzw. b^*. Wegen der Stetigkeit von $\tilde{f}$ und f^* sind die Urbilder

$$U = \tilde{f}^{-1}(V) \quad \text{sowie} \quad U^* = {f^*}^{-1}(V^*)$$

Umgebungen von a. Der Durchschnitt $U \cap U^*$ ist ebenfalls eine Umgebung von a und enthält einen Punkt

$$a^* \neq a \qquad \text{(denn } a \text{ ist nicht isoliert!).}$$

Im Widerspruch zur Disjunktheit von V und V^* gälte somit

$$\tilde{f}(a^*) = f^*(a^*) = f(a^*) \in V \cap V^*. \qquad \square$$

Die Begriffe „Konvergenz“ und „Stetigkeit“, die wir in metrischen Räumen eingeführt haben, sind sogenannte *topologische Begriffe*, d.h. sie lassen sich mit dem Umgebungsbegriff bzw. mit offenen Mengen formulieren. Die Metrik ist nur insofern von Bedeutung, als aus ihr der Umgebungsbegriff abgeleitet wurde. Was topologische Begriffsbildungen anbetrifft sind dabei äquivalente Metriken nicht zu unterscheiden. Genauer gilt

2.10 Hilfssatz. *Seien d und d' zwei äquivalente Metriken auf X. Dann ist jede Umgebung U eines Punktes a bezüglich d auch Umgebung bezüglich d' und umgekehrt. Insbesondere führen d und d' zu denselben offenen Mengen.* meAeq

Beweis. Sei

$$d'(x, y) \leq C d(x, y), \; C > 0.$$

Die Kugeln bezüglich d bzw. d' werden mit

$$U_r(a, d) \quad \text{bzw.} \quad U_r(a, d')$$

bezeichnet. Sei nun U eine Umgebung von a bezüglich d', also

$$U \supset U_r(a, d'), \; r > 0 \text{ genügend klein.}$$

Offenbar gilt

$$U_\varepsilon(a, d) \subset U_r(a, d') \text{ mit } \varepsilon := \frac{r}{C},$$

und daher ist U auch Umgebung von a bezüglich d. $\square$

Zum Schluß zählen wir noch ein paar topologische Begriffe auf (die teilweise erst noch eingeführt werden müssen):

Stetigkeit, Konvergenz, offen, abgeschlossen, innerer Punkt, isolierter Punkt, Häufungspunkt, Kompaktheit.

Nicht rein topologischer Natur sind hingegen Begriffe wie:

gleichmäßige Konvergenz, Vollständigkeit.

3. Induzierte Metrik und Produktmetrik

In der Analysis einer Variablen wurde der metrische Raum $(\mathbb{R}, d)$ mit der Metrik
$$d(x, y) = |x - y|$$
untersucht. Es wurden aber nicht nur Abbildungen
$$f : \mathbb{R} \longrightarrow \mathbb{R}$$
betrachtet, sondern allgemeiner
$$f : D \longrightarrow \mathbb{R},$$
wobei D eine Teilmenge von $\mathbb{R}$ war.

In beliebigen metrischen Räumen haben wir uns scheinbar einer Beschränkung unterworfen, indem wir nur Abbildungen studierten, die auf dem gesamten Raum definiert waren. Durch einen kleine Kunstgriff wird diese Beschränkung wieder aufgehoben. Man faßt eine Teilmenge A eines metrischen Raumes selbst wieder als einen metrischen Raum auf.

3.1 Definition. *Sei $A \subset X$ eine Teilmenge eines metrischen Raumes (X, d).* meIN
*Die **induzierte Metrik** $d' = d|A$ ist definiert durch*
$$d' : A \times A \longrightarrow \mathbb{R}, \quad d'(x, y) := d(x, y) \textit{ für alle } x, y \in A.$$

Ist B eine Teilmenge von A, so gilt offenbar
$$d|B = (d|A)|B.$$

Sind X und Y metrische Räume und ist $A \to Y$ eine Abbildung, wobei A eine Teilmenge von X sei, so heißt diese Abbildung *stetig* (in einem Punkt $a \in A$), wenn die Abbildung stetig im Sinne von Definition 2.3 ist, wobei A selbst als metrischer Raum, versehen mit der induzierten Metrik, aufzufassen ist.

Wenn A eine Teilmenge eines metrischen Raumes (X, d) ist und eine Folge $a_n \in A$ sowie $a \in A$ gegeben sind, so sieht man
$$a_n \longrightarrow a \text{ für } n \longrightarrow \infty \text{ bezüglich der Metrik } d$$
$$\Longleftrightarrow$$
$$a_n \longrightarrow a \text{ für } n \longrightarrow \infty \text{ bezüglich der induzierten Metrik } d|A.$$

Dies ist trivial, denn der Abstand von a_n und a ist ja bei Ausgangs- und induzierter Metrik derselbe. Die beiden folgenden Bemerkungen sind dann Folgerungen aus dieser trivialen Beobachtung und aus Satz 2.7.

3.2 Bemerkung. *Sei $A \subset X$ eine Teilmenge des metrischen Raumes (X, d). Die kanonische Injektion* uniIN

$$\iota : A \longrightarrow X, \;\; \iota(a) = a \text{ für } a \in A,$$

ist stetig (dabei sei A mit der induzierten Metrik versehen).

Bemerkung. *Sei $f : X \longrightarrow Y$ eine Abbildung metrischer Räume und B eine Teilmenge von Y (versehen mit der induzierten Metrik), die das Bild von f enthält, also*

$$f(X) \subset B.$$

Man kann dann die Abbildung

$$f_0 : X \longrightarrow B; \;\; f_0(x) = f(x) \text{ für } x \in X$$

betrachten. Diese ist genau dann in einem Punkt $x \in X$ stetig, wenn f stetig in X ist.

Nun stellt sich die Frage, welche Beziehung zwischen den Umgebungen eines Punktes $a \in A$ bezüglich d und bezüglich $d|A$ besteht.

Zunächst ist klar: Ist $M \subset X$ eine Umgebung von $a \in A$ bezüglich der Metrik d, so ist $M \cap A$ eine Umgebung von a bezüglich der induzierten Metrik. Hiervon gilt jedoch auch die Umkehrung.

Sei $N \subset A$ eine Umgebung von $a \in A$ bezüglich der induzierten Metrik $d|A$. Dann gibt es eine Umgebung $M \subset X$ von a bezüglich d, so daß gilt:

$$M \cap A = N.$$

Dies sieht man so: Man weiß, daß ein $\varepsilon > 0$ mit $U_\varepsilon(a, d|A) \subset N$ existiert. Man setze dann

$$M := U_\varepsilon(a, d) \cup N.$$

Halten wir noch einmal fest:

3.3 Hilfssatz. *Sei A eine Teilmenge eines metrischen Raumes (X, d) und $a \in A$. Eine Teilmenge $N \subset A$ ist genau dann Umgebung von a (bezüglich der induzierten Metrik $d|A$), wenn es eine Umgebung $M \subset X$ von a (bezüglich d) gibt, so daß gilt:* topIN

$$M \cap A = N$$

Dies überträgt sich unmittelbar auf offene Mengen.

3.4 Hilfssatz. *Sei A eine Teilmenge eines metrischen Raumes X und V eine Teilmenge von A. Dann gilt: V ist genau dann offen (bezüglich $d|A$), wenn es eine offene Teilmenge $U \subset X$ (bezüglich d) gibt mit* offIN

$$U \cap A = V.$$

Für abgeschlossene Mengen gilt Entsprechendes (*Beweis* als *Übungsaufgabe*!)

3.5 Hilfssatz. *Sei A eine Teilmenge eines metrischen Raumes X und V eine Teilmenge von A. Dann gilt: V ist genau dann abgeschlossen (bezüglich $d|A$), wenn es eine abgeschlossene Teilmenge $U \subset X$ (bezüglich d) gibt mit $U \cap A = V$.* abgIN

Natürlich ist eine Teilmenge V eines Teilraums A eines metrischen Raumes (X, d), die in A bezüglich der induzierten Metrik $d|A$ offen ist, noch lange nicht offen in X bezüglich d. Wir müssen daher immer unterscheiden, ob V in $(A, d|A)$ oder ob V in (X, d) offen sein soll: Häufig bringen wir dies so zum Ausdruck:

*Eine Teilmenge $V \subset A$ heißt **„offen in A“**, wenn sie offen in dem metrischen Raum $(A, d|A)$ ist.*

Dann braucht V noch lange nicht offen in X zu sein. Unter gewissen Voraussetzungen ist dies jedoch der Fall.

3.6 Hilfssatz. *Sei $A \subset X$ ein **offener** Teil des metrischen Raumes (X, d). Eine Teilmenge $V \subset A$ ist genau dann offen in A (bezüglich $d|A$), wenn sie offen in X (bezüglich d) ist.* spezO

Beweis. a) Sei V offen in X. Dann ist $V \cap A = A$ auch offen in X.

b) Sei nun V offen in A. Dann existiert eine in X offene Menge $U \subset X$, so daß $V = U \cap A$ gilt. Da U und A offen in X sind, ist auch ihr Durchschnitt offen in X. □

3.7 Hilfssatz. *Sei $A \subset X$ ein **abgeschlossener** Teil des metrischen Raumes (X, d). Eine Teilmenge $V \subset A$ ist genau dann abgeschlossen in A (oder: bezüglich $d|A$), wenn sie abgeschlossen in X (oder: bezüglich d) ist.* spezA

Man vergleiche Hilfssatz 3.6. □

3.8 Definition. *Es seien zwei metrische Räume (X, d') und (Y, d'') gegeben. Auf dem kartesischen Produkt* mePRo

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

*ist die sogenannte **Produktmetrik** $d := d' \times d''$ definiert durch*

$$d\left((x, y), (\tilde{x}, \tilde{y})\right) = \max\left(d'(x, \tilde{x}), d''(y, \tilde{y})\right).$$

Den (trivialen) Nachweis der Axiome für diese Metrik übergehen wir. □

Ist (a, b) ein Punkt im Produktraum, so gilt offenbar

$$U_r\left((a, b), d\right) = U_r(a, d') \times U_r(b, d'').$$

3.9 Bemerkung. *Seien X, Y metrische Räume. Eine Folge* folPR

$$(x_n, y_n) \in X \times Y \qquad n = 1, 2, \ldots$$

konvergiert genau dann (bezüglich der Produktmetrik), wenn (x_n) und (y_n) in X bzw. Y konvergiert und gegebenenfalls gilt

$$\lim_{n\to\infty} (x_n, y_n) = (\lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n).$$

Mit Hilfe von Satz 2.7 und Bemerkung 3.9 beweist man unmittelbar:

3.10 Satz. *Es seien X, Y und Z metrische Räume und* stePR

$$f : X \longrightarrow Y \times Z$$

eine Abbildung. Zerlegt man diese in ihre „zwei Komponenten“

$$f_1 : X \longrightarrow Y \quad \textit{und} \quad f_2 : Y \longrightarrow Z, \quad f(x) = (f_1(x), f_2(x)),$$

so erhält man: Die Abbildung f ist genau dann stetig in einem Punkt $x \in X$, wenn f_1 und f_2 in x stetig sind.

Als Spezialfall von Satz 3.10 ergibt sich

3.11 Bemerkung. *Seien X, Y metrische Räume. Die beiden Projektionen* proST

$$\begin{aligned} \pi_1 : X \times Y &\longrightarrow X \\ (x, y) &\longmapsto x \end{aligned} \quad \textit{und} \quad \begin{aligned} \pi_2 : X \times Y &\longrightarrow Y \\ (x, y) &\longmapsto y \end{aligned}$$

sind stetig.

Man kann den Begriff der Produktmetrik sofort auf das Produkt von n metrischen Räumen $X_1, \dots, X_n$ verallgemeinern:

$$d\left((x_1, \dots, x_n), (\tilde{x}_1, \dots, \tilde{x}_n)\right) := \max_{1 \le \nu \le n} d(x_\nu, y_\nu).$$

Die Aussagen Folgerung 3.9 – Bemerkung 3.11 übertragen sich in naheliegender Weise auf das Produkt von n metrischen Räumen.

Offenbar ist die Maximumsmetrik des $\mathbb{R}^n$ nichts anderes als eine derartige Produktmetrik. Damit erhält man beispielsweise:

Eine Abbildung

$$f : D \longrightarrow \mathbb{R}^m, \ D \subset \mathbb{R}^n, \ m, n \in \mathbb{N},$$

ist genau dann stetig, wenn die „Komponenten“

$$f_\nu : D \longrightarrow \mathbb{R} \qquad (1 \le \nu \le m)$$

dieser Abbildung stetig sind.

4. Kompaktheit

Es gibt metrische Räume X mit der Eigenschaft, daß jede stetige Funktion

$$f : X \longrightarrow \mathbb{R}$$

ein Maximum (und ein Minimum) besitzt. Das soll heißen, daß ein $a \in X$ existiert, so daß gilt:

$$f(x) \le f(a) \text{ für alle } x \in X.$$

Beispiel. Man nehme etwa ein abgeschlossenes Intervall $[a, b]$, $a < b$ (versehen mit der Metrik $d(x, y) = |x - y|$, s. Theorem II.2.6).

Es gibt auch metrische Räume, die die genannte Eigenschaft nicht besitzen, etwa das offene Intervall $(0, 1)$. Die Funktion

$$f(x) := \frac{1}{x}$$

besitzt dort kein Maximum.

Eine genaue Analyse vieler Beweise, in denen die Existenz von Maxima oder Minima eingeht, führt auf den Begriff der Kompaktheit. Um diesen formulieren zu können, benötigt man den Begriff der *Überdeckung* einer Menge X. Man versteht darunter eine Schar $(U_i)_{i \in I}$ von Teilmengen U_i von X, so daß jeder Punkt von X in mindestens einer dieser Teilmengen enthalten ist:

$$X = \bigcup_{i \in I} U_i.$$

Die Indexmenge I darf dabei beliebig sein.

Mit Hilfe des Begriffs der Potenzmenge

$$\mathcal{P}(X) = \{Y;\ Y \subset X\},$$

also der Menge aller Teilmengen von X, läßt sich der Begriff der Überdeckung folgendermaßen beschreiben:

Eine Überdeckung (I, φ) einer Menge X besteht aus

a) einer Menge I (der sogenannten Indexmenge),

b) einer Abbildung
$$\varphi : I \longrightarrow \mathcal{P}(X),\ \varphi(i) = U_i,$$
so daß gilt:
$$X = \bigcup_{i \in I} U_i.$$

(Jedes Element von X ist dann also in mindestens einem U_i enthalten).

Sei X ein metrischer Raum. Man nennt eine Überdeckung *offen*, wenn alle U_i, $i \in I$, offen sind.

4.1 Definition. *Ein metrischer Raum X heißt **kompakt**, wenn es zu **jeder** offenen Überdeckung $(U_i)_{i\in I}$ eine endliche Teilüberdeckung gibt. Es mögen also endlich viele Indizes* defKO

$$i_1, \dots, i_n \in I \qquad (n \in \mathbb{N}),$$

existieren, so daß gilt:

$$X = U_{i_1} \cup \ldots \cup U_{i_n}.$$

Dieser Begriff scheint sehr kompliziert zu sein, denn wie soll es möglich sein, über alle offenen Überdeckungen eine Aussage zu machen. Wir werden wenig später den Heine-Borelschen Satz kennenlernen (Theorem 4.8), welcher besagt, daß abgeschlossene Intervalle kompakt sind. Wir werden hieraus neue Beweise für bekannte Sätze ableiten können und mit Hilfe des Begriffs der Kompaktheit auf allgemeinere Situationen übertragen können.

Wir beginnen mit einem Beispiel für einen Raum, welcher nicht kompakt ist und zwar mit dem offenen Einheitsintervall $(0, 1)$ (aufgefaßt als metrischer Raum mit der Betragsmetrik). Dazu betrachten wir die Überdeckung

$$I := \mathbb{N} \quad \text{mit} \quad U_n := (1/n, 1).$$

Jede der Mengen U_n ist offen in $(0, 1)$. Offenbar gilt

$$(0, 1) = \bigcup_{n=1}^{\infty} (1/n, 1).$$

Allerdings existiert keine endliche Teilüberdeckung, denn zu jedem $n \in \mathbb{N}$ existiert ein x mit

$$x > 0 \quad \text{aber} \quad x < 1/n.$$

4.2 Definition. *Eine Teilmenge A eines metrischen Raumes X heißt **kompakt**, wenn A zusammen mit der induzierten Metrik ein kompakter metrischer Raum im Sinne von Definition 4.1 ist.* inKO

Diese Definition ist elegant und einfach, hat aber den Nachteil, daß man wissen muß, wie die offenen Teile von A bezüglich der induzierten Metrik $d|A$ aussehen. Man kann aber auch die Kompaktheit von A direkt mit den offenen Mengen von X beschreiben.

4.3 Bemerkung. *Eine Teilmenge A eines metrischen Raumes ist genau dann kompakt, wenn es zu jeder Schar $(U_i)_{i\in I}$ von offenen Mengen $U_i \subset X$ mit der Eigenschaft* dezKO

$$A \subset \bigcup_{i\in I} U_i$$

bereits endlich viele Indizes $i_1, \dots, i_n$ gibt, so daß gilt:

$$A \subset U_{i_1} \cup \ldots \cup U_{i_n}.$$

Beweis. a) Sei A kompakt und

$$A \subset \bigcup_{i \in I} U_i, \ U_i \subset X \text{offen}.$$

Dann ist

$$A = \bigcup_{i \in I} (U_i \cap A).$$

Die Mengen $U_i \cap A$ sind offen in A. Weil nun A kompakt ist, gilt

$$A = (U_{i_1} \cap A) \cup \ldots \cup (U_{i_n} \cap A)$$

mit gewissen Indizes $i_1, \ldots, i_n \in I$, also

$$A \subset U_{i_1} \cup \ldots \cup U_{i_n}.$$

b) Umgekehrt zeigen wir nun, daß A kompakt ist, wenn die in Bemerkung 4.3 formulierte Eigenschaft erfüllt ist. Sei hierzu

$$A = \bigcup_{i \in I} V_i, \ V_i \text{ offen in } A.$$

Es gibt dann (in X) offene Teile $U_i \subset X$ mit $V_i = U_i \cap A$. Dann gilt offensichtlich

$$A \subset \bigcup_{i \in I} U_i$$

und daher schon

$$A \subset U_{i_1} \cup \ldots \cup U_{i_n}$$

mit gewissen Indizes $i_1, \ldots, i_n \in I$. Hieraus folgt

$$A \subset V_{i_1} \cup \ldots \cup V_{i_n}. \qquad \square$$

Man nennt eine Teilmenge A eines metrischen Raumes *beschränkt*, wenn sie in einer Kugel enthalten ist, wenn es also ein $r > 0$ und einen Punkt $a \in X$ gibt mit

$$A \subset U_r(a).$$

Die Vereinigungsmenge endlich vieler beschränkter Mengen ist wieder beschränkt (Bemerkung 1.4).

4.4 Satz. *Sei A eine kompakte Teilmenge eines metrischen Raumes X. Dann ist A beschränkt und abgeschlossen in X.* BeUAb

Beweis. Wir zeigen zunächst, daß A beschränkt ist. A ist sicherlich in der Vereinigung aller Kugeln aus X enthalten:

$$A \subset \bigcup_{r>0,\ a\in X} U_r(a).$$

Nach Bemerkung 4.3 ist A dann schon in der Vereinigung endlich vieler Kugeln enthalten und damit also beschränkt.

Nun müssen wir noch die Abgeschlossenheit von A zeigen. Dies tun wir, indem wir nachweisen, daß das Komplement $X - A$ offen ist. Ein beliebiger Punkt

$$a \in X,\ \ a \notin A,$$

muß also innerer Punkt von $X - A$ sein. Um dies einzusehen setzen wir

$$V_r := \{x \in X;\ d(x,a) > r\} \qquad (r > 0).$$

Offenbar ist

$$\bigcup_{r>0} V_r = X - \{a\},$$

wegen $a \notin A$ also insbesondere

$$A \subset \bigcup_{r>0} V_r.$$

Da A kompakt ist, gilt

$$A \subset V_{r_1} \cup \ldots \cup V_{r_n}$$

mit gewissen $r_1, \ldots, r_n$ und dann sogar

$$A \subset V_r \text{ mit } r := \min\{r_1, \ldots, r_n\}.$$

Hieraus folgt nun

$$A \cap U_r(a) = \emptyset \Longleftrightarrow U_r(a) \subset X - A,$$

so daß a also ein innerer Punkt von $X - A$ ist. □

Nun fragt es sich, inwieweit die Umkehrung von Satz 4.4 gilt.

4.5 Hilfssatz. *Sei X ein kompakter Raum. Dann ist jede abgeschlossene Teilmenge $A \subset X$ ebenfalls kompakt.* abKO

Beweis. Sei
$$A \subset \bigcup_{i \in I} U_i, \; U_i \subset X \text{ offen.}$$
Dann gilt
$$X = U \cup \bigcup_{i \in I} U_i \text{ mit } U := X - A.$$
Die Menge U ist offen, weil A nach Voraussetzung abgeschlossen ist. somit hat man eine offene Überdeckung von X. Da X kompakt ist, muß es endlich viele Indizes $i_1, \dots, i_n$ mit
$$X = U_{i_1} \cup \ldots \cup U_{i_n} \cup U$$
geben. Da U mit A leeren Durchschnitt hat, folgt hieraus
$$A \subset U_{i_1} \cup \ldots \cup U_{i_n}. \qquad \square$$

Es gibt in metrischen Räumen ein Analogon des Intervallschachtelungsprinzips. Man muß nur „abgeschlossenes Intervall“ durch „Kompaktum“ ersetzen.

4.6 Satz (Allgemeines Intervallschachtelungsprinzip). meINV
Sei X ein metrischer Raum und
$$A_0 \supset A_1 \supset A_2 \supset \ldots$$
eine absteigende Kette von nichtleeren kompakten Teilmengen. Dann ist der Durchschnitt all dieser Kompakta ebenfalls nicht leer,
$$\bigcap_{i=0}^{\infty} A_i = \big\{ a \in X; \quad a \in A_i \textit{ für alle } i \big\} \neq \emptyset.$$

Beweis (indirekt). Es sei
$$\bigcap_{i=0}^{\infty} A_i = \emptyset.$$
Dann ist offenbar
$$\bigcup_{i=0}^{\infty} (X - A_i) = X.$$
Insbesondere ist dann A_0 in der Vereinigung der offenen Mengen $X - A_i$ enthalten. Da A_0 kompakt ist, muß schon
$$A_0 \subset (X - A_{i_1}) \cup \ldots \cup (X - A_{i_n})$$

mit geeigneten Indizes $i_1, \dots, i_n$ gelten. Bezeichnet man mit

$$i := \max\{i_1, \dots, i_n\},$$

so gilt sogar

$$A_0 \subset X - A_i \Longrightarrow A_0 \cap A_i = \emptyset.$$

Dies ist aber ein Widerspruch zu

$$A_0 \cap A_i = A_i \neq \emptyset. \qquad \square$$

4.7 Satz. *Seien X, Y zwei kompakte metrische Räume. Dann ist auch $X \times Y$ (versehen mit der Produktmetrik) kompakt.* `Tych`

Beweis. Es sei

$$X \times Y = \bigcup U_i, \;\; U_i \subset X \times Y \text{ offen.}$$

Dann ist eine endliche Teilüberdeckung zu konstruieren.

1. Schritt. Es sei $b \in Y$ ein fester Punkt. Es gibt endlich viele Indizes $i_1, \dots, i_n \in I$ mit

$$X \times \{b\} \subset U_{i_1} \cup \ldots \cup U_{i_n}.$$

Dies ist klar, da $X \times \{b\}$ (versehen mit der von $X \times Y$ induzierten Metrik) ein kompakter metrischer Raum ist, genau wie X selbst.

2. Schritt. Sei $U := U_{i_1} \cup \ldots \cup U_{i_n}$. Die Menge U hängt von b ab; wir schreiben daher $U = U_b$. Diese Menge ist jedenfalls offen in $X \times Y$. Hieraus wollen wir schließen:

Es gibt $r = r(b) > 0$ mit $\quad X \times U_r(b) \subset U$.

Dies sieht man so: Da U offen ist, existiert zu jedem Punkt $x \in X$ eine Zahl $r(x) > 0$, so daß

$$U_{r(x)}(x) \times U_{r(x)}(b) \subset U$$

gilt (die Mengen $U_r(a) \times U_r(b)$ sind ja genau die Kugelumgebungen von (a, b) bezüglich der Produktmetrik). Es gilt also

$$X \times \{b\} \subset \bigcup_{x \in X} U_{r(x)}(x) \times U_{r(x)}(b) \subset U.$$

Da $X \times \{b\}$ kompakt ist, genügen schon endlich viele

$$x_1, \dots, x_n \text{ mit } r_1 := r(x_1), \dots, r_n := r(x_n),$$

so daß

$$X \times \{b\} \subset U_{r_1}(x_1) \times U_{r_1}(b) \cup \ldots \cup U_{r_n}(x_n) \times U_{r_n}(b) \subset U.$$

Setzt man
$$r := \min\{r_1, \dots, r_n\},$$
so gilt offenbar
$$X \times \{b\} \subset X \times U_r(b) \subset U.$$

3. Schritt. Wir haben bisher gezeigt, daß zu jedem $b \in Y$ eine Zahl $r > 0$ existiert, so daß $X \times U_r(b)$, $r = r(b)$, von endlich vielen der Mengen U_i überdeckt wird. Nun gilt aber
$$Y = \bigcup_{b \in Y} U_r(b).$$
Wegen der Kompaktheit von Y gibt es dann Punkte $b_1, \dots, b_k$ mit
$$Y = U_{r_1}(b_1) \cup \ldots \cup U_{r_k}(b_k)$$
und daher
$$X \times Y = \bigcup_{j=1}^{k} (X \times U_{r_j}(b_j)).$$
Jede der Mengen $X \times U_{r_j}(b_j)$ wird von endlich vielen der U_i überdeckt und gleiches gilt somit auch für $X \times Y$. □

Wir wenden uns nun speziellen Teilräumen des $\mathbb{R}^n$ zu. Fundamental ist

4.8 Theorem (Heine-Borelscher Überdeckungssatz). HeBo
Ein abgeschlossenes Intervall $D = [a, b]$, $a < b$, (versehen mit der Metrik $d(x, y) := |x - y|$) ist ein kompakter metrischer Raum.

Beweis. Sei
$$D \subset \bigcup_{i \in I} U_i, \ U_i \subset \mathbb{R} \text{ offen}$$
eine offene Überdeckung von D. Man betrachte die Menge
$$M := \left\{ x \in [a, b];\ [a, x] \text{ ist Teilmenge der Vereinigung endlich vieler der } U_i \right\}.$$
Offenbar ist $M \neq \emptyset$ ($a \in M$) und nach oben beschränkt (durch b). Man kann daher $\xi := \sup M$ betrachten. Offenbar gilt
$$x \in M \text{ und } a \le t \le x \implies t \in M,$$
d.h. M ist ein Intervall:
$$M = [a, \xi) \quad \text{oder} \quad M = [a, \xi].$$
Wir wählen nun einen Index i_0, so daß $\xi \in U_{i_0}$ gilt. Da U_{i_0} offen ist, existiert ein $\varepsilon > 0$ mit
$$(\xi - \varepsilon, \xi + \varepsilon) \subset U_{i_0}.$$

Man wähle nun irgendein x mit

$$\xi - \varepsilon < x < \xi, \ a \leq x.$$

Dann gilt $x \in M$, also

$$[a, x] \subset U_{i_1} \cup \ldots \cup U_{i_n}.$$

Es folgt

$$[a, \xi] \subset U_{i_0} \cup U_{i_1} \cup \ldots \cup U_{i_n}.$$

Also ist $\xi \in M$ und somit $M = [a, \xi]$. Es ist auch klar, daß gilt

$$\xi = b,$$

denn anderenfalls könnte man ε so klein wählen, daß noch $\xi + \varepsilon \leq b$ gilt und man hätte dann

$$[a, \xi + \varepsilon] \subset U_{i_0} \cup U_{i_1} \cup \ldots \cup U_{i_n},$$

im Widerspruch zur Definition von ξ. □

4.9 Theorem. *Eine Teilmenge $A \subset \mathbb{R}^n$ ist dann und nur dann kompakt, wenn sie beschränkt und abgeschlossen ist.* koBA

Beweis. Sei A beschränkt und abgeschlossen. Wegen der Beschränktheit von A existiert eine Zahl $r > 0$ mit

$$A \subset [-r, r]^n = \left\{ x \in \mathbb{R}^n; \quad |x_\nu| \leq r, \ 1 \leq \nu \leq n \right\}.$$

Nach Theorem 4.8 und 4.7 ist der Würfel $[-r, r]^n$ kompakt, nach 4.5 (in Verbindung mit 3.7) ist dann A ebenfalls kompakt. □

4.10 Satz. *Sei $A \subset \mathbb{R}$ ein nicht leeres Kompaktum. Dann besitzt A Maximum und Minimum.* koMaMi

Beweis. Da A beschränkt ist, existiert $a = \sup A$. Offenbar ist a Randpunkt von A. Da A abgeschlossen ist, gilt auch $a \in A$. Für das Minimum schließt man analog. □

Nun soll die Frage untersucht werden, wie sich Kompakta bei Abbildungen verhalten. Von Bedeutung ist dabei folgendes einfache

4.11 Lemma. *Sei $f : X \longrightarrow Y$ eine stetige Abbildung metrischer Räume und sei $A \subset X$ ein Kompaktum. Dann ist auch $f(A) \subset Y$ kompakt.* BiKo

Beweis. Sei

$$f(A) \subset \bigcup_{i \in I} V_i, \ V_i \subset Y \text{ offen}.$$

Es folgt

$$A \subset \bigcup_{i \in I} f^{-1}(V_i).$$

Da f stetig ist, sind die Mengen $f^{-1}(V_i)$ offen in X, so daß wegen der Kompaktheit von A gilt

$$A \subset f^{-1}(V_{i_1}) \cup \ldots \cup f^{-1}(V_{i_n}),$$

also

$$f(A) \subset V_{i_1} \cup \ldots \cup V_{i_n}. \qquad \square$$

4.12 Theorem. *Sei $f : X \longrightarrow \mathbb{R}$ eine stetige Abbildung eines nichtleeren kompakten Raumes X in $\mathbb{R}$. Dann besitzt f ein Maximum (und ein Minimum), d.h. es gibt $x_0 \in X$ mit* fuMaMi

$$f(x_0) \geq f(x) \quad (\textit{bzw. } f(x_0) \leq f(x)) \quad \textit{für alle } x \in X.$$

Beweis. Nach Lemma 4.11 ist $f(X)$ kompakt. Nun benutze man Satz 4.10. $\square$

4.13 Folgerung. *Jede stetige Funktion $f : A \to \mathbb{R}$ auf einer nicht leeren beschränkten und abgeschlossenen Menge $A \subset \mathbb{R}^n$ hat ein Maximum.* RnMaMi

Gleichmäßige Stetigkeit.

Eine Abbildung

$$f : (X, d) \longrightarrow (X', d')$$

metrischer Räume ist definitionsgemäß genau dann stetig, wenn sie in jedem Punkt $x \in X$ stetig ist, wenn also zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$d(x, y) < \delta \Longrightarrow d'\left(f(x), f(y)\right) < \varepsilon$$

existiert. Gemäß dieser Definition ist zugelassen, daß δ nicht nur von ε sondern auch von x abhängt.

4.14 Definition. *Eine Abbildung* meGlSt

$$f : (X, d) \longrightarrow (X', d')$$

*metrischer Räume heißt **gleichmäßig stetig** (vgl. Satz III.1.14), wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit*

$$d(x, y) < \delta \Longrightarrow d'\left(f(x), f(y)\right) < \varepsilon \textit{ für alle } x, y \in X.$$

Übungsaufgabe. Sei $X = \mathbb{R} - \{0\}$; $X' = \mathbb{R}$. Die Funktion

$$f(x) = \frac{1}{x}$$

ist nicht gleichmäßig stetig.

4.15 Satz. *Sei $f : X \to Y$ eine stetige Abbildung metrischer Räume. Ist X kompakt, so ist f gleichmäßig stetig.* koglSt

*Beweis *).* Sei $\varepsilon > 0$ gegeben. Nach Definition der Stetigkeit existiert zu jedem $x \in X$ ein $\delta_x > 0$ mit

$$d(x,y) < \delta_x \Longrightarrow d'\left(f(x), f(y)\right) < \frac{\varepsilon}{2}.$$

Nun wird aber X von der Gesamtheit der Kugeln

$$U_{\frac{1}{2}\delta_x}(x)$$

überdeckt, so daß wir aufgrund der Kompaktheit von X bereits mit einer endlichen Teilüberdeckung auskommen:

$$X = U_{\delta_1}(x_1) \cup \ldots \cup U_{\delta_n}(x_n), \;\; \delta_\nu := \frac{1}{2}\delta_{x_\nu}.$$

Wir setzen nun

$$\delta := \min_{1 \le \nu \le n} \delta_\nu$$

und zeigen dann

$$d(x,y) < \delta \Longrightarrow d'\left(f(x), f(y)\right) < \varepsilon.$$

Zunächst existiert ein ν mit $d(x, x_\nu) < \delta_\nu$. Hieraus folgt

$$d(y, x_\nu) \le d(x,y) + d(x, x_\nu) < \delta + \delta_\nu \le \delta_{x_\nu}.$$

Dann ist aber

$$d'\left(f(x), f(y)\right) \le d'\left(f(x), f(x_\nu)\right) + d'\left(f(x_\nu), f(y)\right) < \varepsilon. \qquad \square$$

Wir behandeln nun eine typische Anwendung des Satzes von der gleichmäßigen Stetigkeit.

4.16 Bemerkung. *Sei* IntGS

$$f : [a,b] \times [c,d] \longrightarrow \mathbb{R} \qquad (a < b, \;\; c < d)$$

eine stetige Funktion von zwei Variablen. Dann ist

$$F(x) = \int_a^b f(t,x)\,dt$$

stetig auf $[c,d]$.

**)* Es ist gut, diesen Beweis mit dem von Satz III.1.14 zu vergleichen

Beweis. Sei $\varepsilon > 0$. Da das Rechteck $[a,b] \times [c,d]$ kompakt ist, existiert nach dem Satz von der gleichmäßigen Stetigkeit ein $\delta > 0$ mit

$$|t - t'| > \delta,\ |x - x'| < \delta \Longrightarrow |f(t', x') - f(t, x)| < \frac{\varepsilon}{(b-a)}.$$

Damit ergibt sich für $|x - y| < \delta$:

$$|F(x) - F(y)| = \left| \int_a^b (f(t,x) - f(t,y))\, dt \right| \leq (b-a) \cdot \frac{\varepsilon}{(b-a)} = \varepsilon. \qquad \square$$

Man kann diese Bemerkung dazu benutzen, mehrfache Integrale zu *definieren*:

$$\int\limits_{\substack{a \leq x \leq b \\ c \leq y \leq d}} f(x,y)\, dx\, dy := \int_c^d \left[\int_a^b f(x,y)\, dx \right] dy.$$

Ist die Funktion f nicht negativ, so ist dies anschaulich das Volumen des Bereiches

$$\left\{ (x,y,z) \in \mathbb{R}^3; \quad a \leq x \leq b;\ c \leq y \leq d;\ 0 \leq z \leq f(x,y) \right\}.$$

Übungsaufgabe. Man berechne das Volumen der dreidimensionalen Kugel, indem man die Funktion

$$f(x,y) = \begin{cases} \sqrt{1 - x^2 - y^2} & \text{für } x^2 + y^2 \leq 1 \\ 0 & \text{für } x^2 + y^2 > 1 \end{cases}$$

betrachtet.

5. Gleichmäßige Konvergenz und normierte Räume

Wir erläutern zunächst den Begriff der gleichmäßigen Konvergenz auf beliebigen Mengen (vgl. Kapitel II, §3).

Sei X eine nichtleere Menge und

$$B(X) := \big\{ f : X \longrightarrow \mathbb{R}; \quad f \text{ beschränkt} \big\}$$

die Menge aller Funktionen f auf X mit beschränktem Wertevorrat $f(X)$. Man kann das Supremum

$$\|f\| = \sup_{x \in X} |f(x)|$$

des Wertevorrats von f betrachten (vgl. Definition II.3.2). Offenbar gilt

$$\|f + g\| \leq \|f\| + \|g\|.$$

Hieraus folgert man leicht, daß

$$d(f, g) := \|f - g\|$$

eine Metrik auf $B(X)$ ist.

5.1 Definition. *Sei X eine nichtleere Menge. Eine Folge von Funktionen aus $B(X)$* meGlKo

$$f_1,\, f_2,\, f_3, \ldots : X \longrightarrow \mathbb{R}$$

heißt ***gleichmäßig konvergent*** *(vgl. Definition II.3.4) gegen $f \in B(X)$, wenn sie in der Metrik*

$$d(f, g) = \|f - g\|$$

gegen f konvergiert, wenn also gilt

$$\|f_n - f\| \longrightarrow 0 \text{ für } n \longrightarrow \infty.$$

Das bedeutet: Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit der Eigenschaft

$$|f_n(x) - f(x)| < \varepsilon \text{ für alle } x \in X,\ n \geq N.$$

Man sollte den Begriff der gleichmäßigen Konvergenz mit dem der *punktweisen Konvergenz* vergleichen. Eine Folge

$$f_n : X \longrightarrow \mathbb{R}$$

konvergiert per definitionem *punktweise* gegen die Funktion

$$f : X \longrightarrow \mathbb{R},$$

wenn für alle $x \in X$ die Zahlenfolge $f_n(x)$ gegen die Zahl $f(x)$ konvergiert. Das bedeutet:

Zu jedem $x \in X$ und zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit der Eigenschaft

$$|f_n(x) - f(x)| < \varepsilon \quad \textit{für } n \geq N.$$

Nach dieser Definition kann die Schranke N nicht nur von ε, sondern auch von x abhängen. Man schreibt daher manchmal auch $N = N(\varepsilon, x)$. *Gleichmäßige Konvergenz* liegt erst dann vor, wenn man das N so finden kann, daß es von x unabhängig ist ($N = N(\varepsilon)$).

Übungsaufgabe.

$$f, g \in B(X), c \in \mathbb{R} \Longrightarrow f + g,\ f \cdot g,\ cf \in B(X).$$

Dabei ist

$$(f + g)(x) := f(x) + g(x),\ (f \cdot g)(x) := f(x)g(x) \quad \text{und} \quad (cf)(x) = cf(x).$$

5.2 Satz. *Es sei $X \neq \emptyset$ ein metrischer Raum und* `liGlSt`

$$f_1,\ f_2,\ f_3, \ldots : X \longrightarrow \mathbb{R}$$

eine Folge von stetigen beschränkten Funktionen, die gleichmäßig gegen die Funktion f konvergiert. Dann ist auch f stetig (vgl. Theorem II.3.5).

Beweis. Man hat die Ungleichung

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x) - f(x)|$$

zu benutzen. Sei $\varepsilon > 0$ gegeben. Man bestimme $n \in \mathbb{N}$ so, daß

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3} \text{ für } n \geq N \quad (\text{und } \textit{alle } x \in X)$$

gilt. Dann bestimme man $\delta > 0$ so, daß

$$|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3} \text{ für } d(x_0, x) < \delta.$$

Nutzt man nun die genannte Ungleichung für N anstelle von n aus, so ergibt sich

$$|f(x_0) - f(x)| < \varepsilon \text{ für } d(x_0, x) < \delta. \qquad \square$$

Anmerkung. Sei $X = [0, 1]$. Die Folge der stetigen und beschränkten Funktionen

$$f_n(x) := x^n$$

konvergiert punktweise gegen die *unstetige* Funktion

$$f(x) = \begin{cases} 0 & \text{für } x \neq 1 \\ 1 & \text{für } x = 1. \end{cases}$$

In Satz 5.2 ist also die Voraussetzung der gleichmäßigen Konvergenz wesentlich.

Wenn X ein metrischer Raum ist, so wird mit

$$C(X) = \{f : X \longrightarrow \mathbb{R};\ f \text{ stetig}\}$$

die Menge der stetigen Funktionen auf X bezeichnet. Wenn X nicht leer und kompakt ist, so ist jede Funktion aus $C(X)$ sogar beschränkt, genauer

$$\|f\| = \max_{x\in X} |f(x)|.$$

Man muß beachten:

$$f \text{ stetig} \Longrightarrow |f| \text{ stetig}.$$

$|f|$ ist ja die Zusammensetzung der beiden stetigen Abbildungen

$$X \xrightarrow{f} \mathbb{R} \overset{y\mapsto|y|}{\longrightarrow} \mathbb{R}.$$

5.3 Bemerkung. *Es seien $f, g : X \to \mathbb{R}$ stetige Funktionen auf einem metrischen Raum X. Dann sind auch die Funktionen* SuPrSt

$$f + g \quad \textit{und} \quad f \cdot g$$

stetig.

Insbesondere sind die Funktionen

$$\begin{aligned} \mathrm{ad} : \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R},\ \mathrm{ad}(x, y) := x + y, \\ \mathrm{m} : \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{R},\ \mathrm{m}(x, y) := x \cdot y, \end{aligned}$$

stetig.

Beweis. Sei (x_n) eine Folge in X mit dem Grenzwert $x \in X$. Dann konvergieren

$$f(x_n) \longrightarrow f(x) \quad \text{und} \quad g(x_n) \longrightarrow g(x),$$

also

$$\begin{aligned} (f + g)(x_n) &= f(x_n) + g(x_n) \longrightarrow f(x) + g(x) = (f + g)(x), \\ (f \cdot g)(x_n) &= f(x_n)g(x_n) \longrightarrow f(x)g(x) = (f \cdot g)(x), \end{aligned}$$

Die Funktionenklassen $B(X)$ und $C(X)$ sind spezielle Beispiele von Vektorräumen. Es ist in der Analysis nicht wichtig, einen abstrakten Vektorraumbegriff zu benutzen.

Sei X eine beliebige Menge. Eine Menge V von Funktionen $f : X \to \mathbb{R}$ (oder $f : X \to \mathbb{C}$) heißt *reeller* (oder *komplexer*) *Vektorraum*, falls gilt:

a) $f, g \in V \implies f + g \in V$,

b) $f \in V,\ c \in \mathbb{R}\ (\mathbb{C}) \implies cf \in V.$

Beispiele für Vektorräume.

1) Die Menge aller Funktionen

$$f : X \longrightarrow \mathbb{R}\ (\mathbb{C}).$$

Spezialfall. $X = \{1, 2, \dots, n\}$.

Eine Funktion

$$f : \{1, 2, \dots, n\} \longrightarrow \mathbb{R}$$

ist nichts anderes als ein n-Tupel von reellen Zahlen. Somit ist die Menge aller dieser Funktionen gerade der $\mathbb{R}^n$. Dieser ist also ein spezieller Vektorraum.

2) Die Menge aller beschränkten Funktionen

$$B(X) := \{f : X \longrightarrow \mathbb{R};\ f \text{ beschränkt}\}.$$

3) Die Menge aller stetigen Funktionen auf einem metrischen Raum X

$$C(X) := \{f : X \longrightarrow \mathbb{R};\ f \text{ stetig}\}.$$

4) Sei $X = [a, b];\ a < b$ ein abgeschlossenes Intervall. Dann ist auch die Menge

$$R(X) := \{f : X \longrightarrow \mathbb{R};\ f \text{ integrierbar}\}$$

der Regelfunktionen ein Vektorraum.

5.4 Definition. *Eine* ***Norm*** $\|\cdot\|$ *auf einem Vektorraum* V *ist eine Abbildung* deNO

$$V \longrightarrow \mathbb{R},\ f \longmapsto \|f\|,$$

mit den Eigenschaften

1) $\|f\| \geq 0$ *und* $\|f\| = 0 \iff f = 0.$

2) $\|cf\| = |c|\|f\|$ *für* $f \in V;\ c \in \mathbb{R}\ (\mathbb{C}).$

3) $\|f + g\| \leq \|f\| + \|g\|$ *für alle* $f, g \in V.$

Beispiel. Auf der Menge $B(X)$ hat man die *Supremumsnorm*

$$\|f\| = \sup_{x \in X} |f(x)|.$$

Die Axiome 1)–3) sind leicht nachzuprüfen.

Ein *normierter (Vektor-) Raum* $(V, \|\cdot\|)$ ist definitionsgemäß ein Vektorraum V, auf dem eine Norm $\|\cdot\|$ gegeben ist.

5.5 Bemerkung. *Ist $(V, \|\cdot\|)$ ein normierter Raum, so ist* NoMe

$$d(x,y) := \|x-y\|$$

eine Metrik auf V.

Jeder normierte Raum ist also insbesondere ein metrischer Raum.

Wichtige Beispiele für Normen erhält man aus *Skalarprodukten*.

5.6 Definition. *Ein* ***Skalarprodukt*** *auf einem (reellen oder komplexen) Vektorraum V ist eine Abbildung* deSkal

$$V \times V \longrightarrow \mathbb{R}\ (\mathbb{C}), \quad (f,g) \longmapsto \langle f,g\rangle,$$

mit den Eigenschaften

1) $\langle f_1 + f_2, g\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle$ *für* $f_1, f_2, g \in V$
   $\langle af, g\rangle = a\langle f,g\rangle$ *für* $a \in \mathbb{R}\ (\mathbb{C})$; $f, g \in V$.

2) $\langle f,g\rangle = \overline{\langle g,f\rangle}$ *(insbesondere ist $\langle f,f\rangle$ reell).*

3) $\langle f,f\rangle \geq 0$ *für alle $f \in V$, wobei das Gleichheitszeichen nur für $f = 0$ gilt.*

Beispiele für Skalarprodukte.

1) $\mathbb{R}^n$ mit $\langle x,y\rangle := \sum_{\nu=1}^{n} x_\nu y_\nu$.

2) $\mathbb{C}^n$ mit $\langle x,y\rangle := \sum_{\nu=1}^{n} x_\nu \bar{y}_\nu$.

3) Sei $C([a,b])$ die Menge der stetigen reellen) Funktionen auf dem abgeschlossenen Intervall $[a,b]$ $(a<b)$.

$$\langle f,g\rangle := \int_a^b f(x)g(x)\,dx.$$

Man kann 3) verallgemeinern. Ist etwa $p : [a,b] \longrightarrow \mathbb{R}$ eine stetige Funktion, die nur positive Werte annimmt, so definiere man:

$$\langle f,g\rangle := \int_a^b p(x)f(x)g(x)\,dx.$$

Noch allgemeiner kann man annehmen, daß die *Gewichtsfunktion p lediglich nichtnegativ* ($p(x) \geq 0$ *für alle* x) *ist und den* Wert 0 nur endlich oft annimmt. Auch in diesem Fall sind die Axiome des Skalarprodukts offenbar noch erfüllt.

Von zentraler Bedeutung ist

5.7 Satz (Cauchy-Schwarzsche Ungleichung). CSU
Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf einem Vektorraum V. Dann gilt

$$|\langle f,g\rangle| \le \|f\|\cdot\|g\| \text{ für alle } f,g\in V.$$

Dabei sei $\|f\| := {}_{+}\sqrt{\langle f,f\rangle}$.

Beweis. Ist $g=0$, so ist die Behauptung klar. Sei also $g\neq 0$. Wir betrachten

$$0 \le \|f+tg\|^2 = \langle f+tg, f+tg\rangle = \|f\|^2 + 2\mathrm{Re}t\langle f,g\rangle + |t|^2\|g\|^2.$$

Dies gilt für alle $t\in\mathbb{C}$. Wählt man speziell

$$t := -\frac{\overline{\langle f,g\rangle}}{\|g\|^2},$$

so folgt

$$0 \le \|f\|^2 - 2\frac{|\langle f,g\rangle|^2}{\|g\|^2} + \left(\frac{|\langle f,g\rangle|}{\|g\|^2}\right)\|g\|^2$$

oder

$$|\langle f,g\rangle|^2 \le \|f\|^2\cdot\|g\|^2.$$ □

5.8 Folgerung. *Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf einem Vektorraum V. Dann wird durch* SkaMe

$$\|f\| := {}_{+}\sqrt{\langle f,f\rangle}$$

eine Norm auf V erklärt. Insbesondere wird V ein metrischer Raum durch

$$d(f,g) = \|f-g\|.$$

Beweis. Es ist

$$\begin{aligned}\|f+g\|^2 &= \langle f+g,f+g\rangle = \langle f,f\rangle + \langle f,g\rangle + \langle g,f\rangle + \langle g,g\rangle\\ &\le \|f\|^2 + 2\|f\|\cdot\|g\| + \|g\|^2 = (\|f\|+\|g\|)^2.\end{aligned}$$

Hieraus folgt die Dreiecksungleichung

$$\|f+g\| \le \|f\| + \|g\|.$$

Die übrigen Normeigenschaften sind evident. □

6. Der Approximationssatz von Stone Weierstrass

Eine Funktion $h : [a,b] \to \mathbb{R},\ a < b$, auf einem abgeschlossenen Intervall heißt *Treppenfunktion*, wenn eine Partition

$$a = a_0 < a_1 < \ldots < a_n = b$$

existiert, so daß h im Inneren (a_{j-1}, a_j) eines jeden Teilintervalls konstant ist.

Mit Hilfe des Satzes von der *gleichmäßigen Stetigkeit* zeigt man, daß man zu jeder stetigen Funktion $f : [a,b] \to \mathbb{R}$ und zu jedem $\varepsilon > 0$ eine Treppenfunktion h mit der Eigenschaft

$$|f(x) - h(x)| < \varepsilon \text{ für alle } x \in [a,b]$$

finden kann. Hieraus folgt, daß jede stetige Funktion auf einem abgeschlossenen Intervall integrierbar ist (s. Satz III.1.14 und Theorem III.1.15).

Wir wollen in diesem Paragraphen ein einfaches Kriterium dafür ableiten, wann sich jede stetige Funktion $f : [a,b] \to \mathbb{R}$ durch eine vorgegebene Klasse von Funktionen h in obigem Sinne beliebig gut approximieren läßt. Wir wollen dabei allerdings nur stetige Funktionen h zulassen.

6.1 Theorem (Approximationssatz von Stone-Weierstrass, erste Variante). *Es sei X ein kompakter metrischer Raum, der mindestens zwei Punkte enthält. Weiterhin sei W eine Menge von stetigen reellwertigen Funktionen auf X mit den Eigenschaften* StWe

1) $f, g \in W \Longrightarrow f + g \in W$ *und* $cf \in W$ *für* $c \in \mathbb{R}$.
2) $f \in W \Longrightarrow |f| \in W$.
3) *(Punktetrennungseigenschaft)*
   Seien a, b zwei verschiedene Punkte von X. Dann existiert ein $f \in W$ mit $f(a) = 0$ aber $f(b) \neq 0$.

Unter diesen Voraussetzungen gilt: Ist $f \in C(X)$ eine beliebige stetige Funktion auf X und $\varepsilon > 0$, so existiert eine Funktion

$$h \in W \text{ mit } \|f - h\| \leq \varepsilon.$$

Zusatz: Insbesondere existiert dann eine Folge (h_n) von Funktionen aus W, die gleichmäßig gegen f konvergiert. (Man setze $\varepsilon := \frac{1}{n}$ und bezeichne die zugehörige Funktion mit h_n.)

Beweis. Zunächst einige Bezeichnungen:
Sei $f : X \to \mathbb{R}$ eine Funktion. Dann sei

$$f^+ : X \longrightarrow \mathbb{R},\ f^+(x) := \max\left(f(x), 0\right).$$

Offenbar gilt:

$$f^+ = \frac{1}{2}\left(|f| + f\right).$$

Insbesondere gilt

$$f \in W \Longrightarrow f^+ \in W \qquad \text{(wegen 1) und 2))}.$$

Sei $g : X \to \mathbb{R}$ eine weitere Funktion. Man definiert

$$f \vee g : X \longrightarrow \mathbb{R},\ (f \vee g)(x) = \max\left(f(x), g(x)\right),$$
$$f \wedge g : X \longrightarrow \mathbb{R},\ (f \wedge g)(x) = \min\left(f(x), g(x)\right).$$

Offenbar gilt

$$f \vee g = f + (g - f)^+$$
$$f \wedge g = f - (g - f)^+.$$

Hieraus kann man sofort folgern:

$$f, g \in W \Longrightarrow f \vee g \in W,\ f \wedge g \in W.$$

Das gleiche kann man mit endlich vielen Funktionen $f_1, \dots, f_n$ machen.

Nun soll Theorem 6.1 in mehreren Schritten bewiesen werden. Wir formulieren erst die Schritte.

1. Schritt. Seien $a, b \in X$ verschiedene Punkte und A, B zwei reelle Zahlen. Dann existiert $f \in W$ mit

$$f(a) = A,\ f(b) = B.$$

2. Schritt. Gegeben seien

$$f : X \overset{\text{stetig}}{\longrightarrow} \mathbb{R},\ a, b \in X \text{ und } \varepsilon > 0.$$

Dann gibt es eine offene Umgebung V_b von b und eine Funktion $h_{a,b} \in W$ mit

$$h_{a,b}(a) = f(a) - \frac{\varepsilon}{2} \quad \text{und} \quad h_{a,b}(x) < f(x) \text{ für alle} x \in V_b.$$

3. Schritt. Gegeben seien

$$f : X \overset{\text{stetig}}{\longrightarrow} \mathbb{R},\ a \in X \text{ und } \varepsilon > 0.$$

Dann gibt es eine offene Umgebung U_a von a und eine Funktion $h_a \in W$ mit

$$f(x) > h_a(x) \text{ für alle } x \in X$$

und

$$h_a(x) > f(x) - \varepsilon \text{ für alle } x \in U_a.$$

4. Schritt. Beweis des Theorems.

Beweis des 1. Schrittes. Nach 3) existieren Funktionen $g, h \in W$ mit

$$g(a) \neq 0,\ g(b) = 0 \quad \text{und} \quad h(a) = 0,\ h(b) \neq 0.$$

Man setze

$$f := \frac{A}{g(a)} g + \frac{B}{h(b)} h.$$

Insbesondere existiert zu jedem $a \in X$ eine Funktion $f \in W$ mit vorgegebenem Funktionswert in a.

Beweis des 2. Schrittes. Nach dem 1. Schritt existiert eine Funktion $h_{a,b} \in W$ mit

$$h_{a,b}(a) = f(a) - \frac{\varepsilon}{2} \quad \text{und} \quad h_{a,b}(b) = f(b) - \frac{\varepsilon}{2}.$$

Die Menge

$$V_b := \{x \in X;\ h_{a,b}(x) < f(x)\}$$

ist eine offene Umgebung von b (denn V_b ist das Urbild von $(0, \infty)$ bei der stetigen Abbildung $h_{a,b} - f$).

Beweis des 3. Schrittes. Sei $f \in C(X, \mathbb{R}),\ a \in X$ und $\varepsilon > 0$. Zu jedem $b \in X$ existieren $h_{a,b} \in W$ und eine offene Umgebung V_b mit den Eigenschaften, die im zweiten Schritt formuliert wurden. Es gilt

$$X = \bigcup_{b \in X} V_b.$$

Wegen der Kompaktheit von X gilt sogar

$$X = V_{b_1} \cup \ldots \cup V_{b_n} \qquad (b_1, \ldots, b_n \text{ geeignet}).$$

Wir definieren nun

$$h_a := h_{a,b_1} \wedge \ldots \wedge h_{a,b_n}.$$

Dies ist wieder eine Funktion aus W. Nach Konstruktion von $h_{a,b}$ gilt ferner

$$h_a(x) \leq h_{a,b_j}(x) < f(x) \text{ für } x \in V_{b_j}.$$

Da die V_{b_j} den ganzen Raum X überdecken, gilt:

$$h_a(x) < f(x) \text{ für } x \in X.$$

Das ist eine der Behauptungen des dritten Schrittes. Es bleibt die Umgebung U_a zu konstruieren. Wegen

$$h_a(a) = f(a) - \frac{\varepsilon}{2} > f(a) - \varepsilon$$

ist

$$U_a := \{x \in X;\ h_a(x) > f(x) - \varepsilon\}$$

wieder eine offene Umgebung von a

4. Schritt. Beweis des Theorems. Sei $f \in C(X)$. Zu jedem $a \in X$ wurde eine Funktion $h_a \in W$ konstruiert mit

$$f(x) > h_a(x) \text{ für alle } x \in X$$

aber

$$h_a(x) > f(x) - \varepsilon \text{ für alle } x \in U_a,$$

wobei U_a eine offene Umgebung von a ist. Es gilt

$$X = \bigcup_{a \in X} U_a$$

und somit aufgrund der Kompaktheit von X bereits

$$X = U_{a_1} \cup \ldots \cup U_{a_n}.$$

Man definiere nun

$$h := h_{a_1} \vee \ldots \vee h_{a_n}.$$

Offenbar liegt h in W und es gilt

$$f(x) > h(x) > f(x) - \varepsilon \text{ für alle } x \in X,$$

insbesondere also

$$|f(x) - h(x)| < \varepsilon \text{ für alle } x \Longleftrightarrow \|f - h\| < \varepsilon. \qquad \square$$

Ein Beispiel für Theorem 6.1. Sei

$$X = [a, b] \qquad (a < b)$$

und W die Menge aller stetigen Funktionen $f : [a, b] \to \mathbb{R}$, für die es Stützstellen

$$a = a_0 < a_1 < \ldots < a_n = b$$

mit, so daß f linear auf $[a_\nu, a_{\nu+1}]$ ist $(0 \leq \nu < n)$.

Die Menge W erfüllt die in Theorem 6.1 geforderten Voraussetzungen. Daher kann man jede stetige Funktion durch stückweise lineare Funktionen gleichmäßig approximieren.

6.2 Theorem (Approximationssatz von Stone-Weierstrass, zweite Variante). StWez *Es sei $X \neq \emptyset$ ein kompakter metrischer Raum und $W \subset C(X,\mathbb{R})$ eine Menge von stetigen reellwertigen Funktionen auf X mit den Eigenschaften*

1) $f, g \in W \Longrightarrow f + g \in W$ *und* $cf \in W$ *für* $c \in \mathbb{R}$.

2) $f, g \in W \Longrightarrow f \cdot g \in W$ *und die konstanten Funtionen liegen in* W.

3) W *besitzt die Punktetrennungseigenschaft (vgl. Theorem 6.1)*

Unter diesen Voraussetzungen gilt: Ist $f \in C(X)$ eine beliebige stetige Funktion auf X und $\varepsilon > 0$, so existiert eine Funktion

$$h \in W \ \textit{mit}\ \|f - h\| \leq \varepsilon.$$

Der wesentliche Unterschied zu Theorem 6.1 besteht darin, daß man die Bedingung

$$f \in W \Longrightarrow |f| \in W$$

durch die Bedingung

$$f, g \in W \Longrightarrow f \cdot g \in W$$

ersetzt hat.

Beweis von Theorem 6.2. Sei $\overline{W}$ der Abschluß von W in $C(X)$, d.h. $\overline{W}$ besteht aus allen stetigen Funktionen auf X, die sich gleichmäßig durch Funktionen aus W approximieren lassen, also

$$f \in \overline{W} \Longleftrightarrow \text{zu jedem } \varepsilon > 0 \text{ existiert } h \in W \text{ mit } \|f - h\| < \varepsilon.$$

Zunächst machen wir uns klar, daß $\overline{W}$ wieder die Eigenschaften 1)–3) von Theorem 6.2 erfüllt. Es ist zu zeigen:

$$f, g \in \overline{W} \Longrightarrow f + g,\ f \cdot g \text{ und} cf \in \overline{W}.$$

a) Wir zeigen, $f + g \in \overline{W}$. Sei $\varepsilon > 0$ vorgegeben und $f_1, g_1 \in W$ so gewählt, daß

$$\|f - f_1\| < \varepsilon \text{ und } \|g - g_1\| < \varepsilon$$

gilt. Es folgt

$$\|f + g - (f_1 + g_1)\| < \|f - f_1\| + \|g - g_1\| < 2\varepsilon.$$

b) Es soll $f \cdot g \in \overline{W}$ gezeigt werden. Sei wieder $\varepsilon > 0$ vorgegeben. Wir setzen

$$\delta := \min\left(1, \frac{\varepsilon}{1 + \|f\| + \|g\|}\right)$$

und bestimmen $f_1, g_1 \in W$ so, daß

$$\|f - f_1\| < \delta \text{ und } \|g - g_1\| < \delta$$

gilt. Es folgt

$$\begin{aligned}\|fg - f_1g_1\| &= \| - (f - f_1)(g - g_1) + f(g - g_1) + g(f - f_1)\| \\ &\le \|f - f_1\|\|g - g_1\| + \|f\|\|g - g_1\| + \|g\|\|f - f_1\| \\ &\le \delta^2 + \delta\|f\| + \delta\|g\| = \delta(\delta + \|f\| + \|g\|) \\ &\le \delta(1 + \|f\| + \|g\|) \le \varepsilon.\end{aligned}$$

Hierbei wurde neben der Dreiecksungleichung $\|f + g\| \le \|f\| + \|g\|$ auch noch die Ungleichung $\|f \cdot g\| \le \|f\| \cdot \|g\|$ benutzt, welche leicht zu verifizieren ist.

Damit besitzt also auch $\overline{W}$ die Eigenschaften 1)–3) aus Theorem 6.2. Die Punktetrennungseigenschaft ist dabei trivial, denn es gilt ja $W \subset \overline{W}$.

Jetzt zeigen wir sogar

$$f \in \overline{W} \Longrightarrow |f| \in \overline{W}.$$

Wenn wir dies gezeigt haben, so sind wir fertig, denn nach der ersten Variante Theorem 6.1 kann dann jede stetige Funktion $f \in C(X)$ gleichmäßig durch Funktionen aus $\overline{W}$ approximiert werden, d.h. ist $\varepsilon > 0$, so existiert $g \in \overline{W}$ mit

$$\|f - g\| < \varepsilon.$$

Andererseits existiert nach Definition von $\overline{W}$ eine Funktion

$$h \in W \text{ mit} \|f - h\| = \|(f - g) + (g - h)\| < 2\varepsilon.$$

Wir haben also gezeigt, daß $\overline{\overline{W}} = \overline{W}$ gilt, was offenbar in beliebigen metrischen Räumen richtig ist.

Halten wir noch einmal fest:

Die zweite Variante Theorem 6.2 des Approximationssatzes von Stone-Weierstrass ist zurückgeführt auf die erste, wenn man zeigen kann, daß für $f \in W$ die Funktion $|f|$ gleichmäßig durch Funktionen aus W approximiert werden kann.

Wir werden sogar zeigen:

Die Funktion $|f|$ läßt sich gleichmäßig durch endliche Linearkombinationen von $1, f, f^2, \ldots$ approximieren.

Der Schlüssel zum Beweis dieser Behauptung ist der folgende Hilfssatz

6.3 Hilfssatz. *Auf dem Intervall $[-1, 1]$ existiert eine Folge von Polynomen* glBet

$$p_n : [-1, 1] \longrightarrow \mathbb{R},$$

die gleichmäßig gegen die Funktion $|x|$ konvergiert.

Wir werden diese Folge (p_n) explizit angeben. Der Gedankengang ist der folgende: Man kann

$$|x| = {}_{+}\sqrt{1-(1-x^2)} \text{ für } -1 \leq x \leq 1$$

schreiben. Setzt man

$$y = 1 - x^2,$$

so variiert y zwischen 0 und 1.

Es genügt also, eine Folge von Polynomen $q_n(y)$ zu konstruieren, welche gleichmäßig gegen die Funktion

$$_{+}\sqrt{1-y}$$

konvergiert. Der zugrunde gelegte Definitionsbereich ist hierbei

$$D := \{y;\ 0 \leq y \leq 1\}.$$

(Anschließend setze man dann $p_n(x) := q_n(1-x^2)$. Dies sind dann Polynome in x).

Zur Approximation von $_{+}\sqrt{1-y}$ verwendet man die TAYLORreihe. (Diese haben wir bereits in Kapitel III, §7 im Zusammenhang mit der Taylorschen Formel behandelt. Dort haben wir uns allerdings um die Randpunkte des Konvergenzintervalls nicht gekümmert, weshalb wir das Ganze hier noch einmal aufrollen wollen.) Man ermittelt für die Taylorreihe sofort

$$\sum_{\nu=0}^{\infty} \binom{\frac{1}{2}}{\nu} (-1)^{\nu} y^{\nu},$$

wobei der verallgemeinerte Binomialkoeffizient $\binom{\alpha}{\nu}$ definiert ist durch

$$\binom{\alpha}{\nu} = \frac{\alpha(\alpha-1)\dots(\alpha-\nu+1)}{\nu!}.$$

Der Konvergenzradius dieser Reihe ist offenbar 1. Es ist aber dennoch nicht von vornherein klar, daß $_{+}\sqrt{1-y}$ durch die Reihe dargestellt wird. Jedenfalls wird durch die Reihe eine Funktion

$$f : (-1,1) \longrightarrow \mathbb{R}$$

dargestellt, die, wie man leicht nachrechnet, der Differentialgleichung

$$(1-y)f'(y) = -\alpha f(y)$$

(mit $\alpha = \frac{1}{2}$) genügt. (Gliedweise Differentiation ist im Innern des Konvergenzintervalls erlaubt!).

Hieraus folgert man leicht, daß

$$\left(\frac{f(y)}{(1-y)^\alpha}\right)' = 0$$

gilt. Also ist

$$f(y) = c \cdot (1-y)^\alpha.$$

Durch Spezialisierung ($y = 0$) ermittelt man $c = 1$.

Wir haben also gezeigt: Die Folge der Polynome

$$q_n(y) := \sum_{\nu=0}^{n} \binom{\frac{1}{2}}{\nu} (-1)^\nu y^\nu$$

konvergiert in $(-1, 1)$ gegen ${}_+\sqrt{1-y}$ und zwar ist die Konvergenz gleichmäßig in jedem Teilintervall $[-1+\varepsilon, 1-\varepsilon]$ (mit $1 > \varepsilon > 0$, ε beliebig). Mehr ist aus der allgemeinen Theorie der Potenzreihen nicht zu erhalten.

Wir wollten jedoch die Funktion ${}_+\sqrt{1-y}$, welche ja in $y = 1$ noch stetig ist, sogar gleichmäßig in $[0, 1]$ approximieren. (Der Punkt $y = 1$ entspricht gerade der Knickstelle $x = 0$ der Funktion $|x|$ und ist somit der Angelpunkt!)

Wir müssen also noch zeigen:

1) Die Folge $q_n(y)$ konvergiert auch für $y = 1$ und zwar gegen $\sqrt{1-1} = 0$.

2) Die Konvergenz in $[0, 1]$ ist gleichmäßig.

Beweis. Eine Abzählung der Vorzeichen im Binomialkoeffizienten $\binom{\alpha}{\nu}$ für $\nu \geq 1$ ergibt, daß

$$(-1)^\nu \binom{\frac{1}{2}}{\nu} = -\left|\binom{\frac{1}{2}}{\nu}\right|$$

gilt. Hieraus folgt

$$q_n(x) = 1 - \sum_{\nu=1}^{n} \left|\binom{\frac{1}{2}}{\nu}\right| y^\nu.$$

Damit ist klar, daß für $0 \leq y \leq 1$ gilt:

$$q_1(y) \geq q_2(y) \geq \dots .$$

Außerdem ist

$$\sqrt{1-y} = \lim_{n\to\infty} q_n(y) \geq 0 \text{ für } 0 \leq y < 1.$$

Hieraus folgt

$$q_n(y) \geq 0 \text{ für alle } n \text{ und } 0 \leq y < 1.$$

Die Funktionen q_n sind stetig in $y = 1$, also folgt sogar

$$q_n(y) \geq 0 \text{ für alle } n \text{ und } 0 \leq y \leq 1.$$

Die Folge $(q_n(1))$ ist also monoton fallend und nach unten (durch 0) beschränkt. Sie konvergiert daher. Das bedeutet nichts anderes, als daß die Reihe

$$\sum_{\nu=0}^{\infty}\left|\binom{\frac{1}{2}}{\nu}\right|$$

konvergiert. Für alle $y \in [0,1]$ gilt

$$\left|\binom{\frac{1}{2}}{\nu}(-1)^{\nu}y^{\nu}\right| \leq \left|\binom{\frac{1}{2}}{\nu}\right|.$$

Nach dem Majorantenkriterium konvergiert also die Folge $(q_n(y))$ daher schon gleichmäßig für $0 \leq y \leq 1$. Die Grenzfunktion

$$\lim_{n\to\infty} q_n(y)$$

ist daher stetig in $[0,1]$*). Sie stimmt mit $\sqrt{1-y}$ in $[0,1)$ überein. Da auch letztere Funktion in $y = 1$ stetig ist, müssen die beiden Funktionen auch dort übereinstimmen. Damit ist Hilfssatz 6.3 bewiesen. □

Nun ist der Approximationssatz in der zweiten Variante leicht zu beweisen. Sei etwa $f \in W$. Wir approximieren $|f|$. Sei zunächst

$$g := \frac{f}{\|f\|+1}.$$

Dann gilt offenbar

$$|g(x)| \leq 1 \text{ für alle } x \in X.$$

Nach Hilfssatz 6.3 existiert ein Polynom

$$p_n(x) = a_0 + a_1x + \ldots + a - nx^n$$

mit der Eigenschaft

$$|p_n(x) - |x|| < \varepsilon \text{ für alle } x \in [-1,1].$$

Es folgt insbesondere

$$|p_n(g(x)) - |g(x)||\varepsilon \text{ für alle } x \in X.$$

Die Funktion

$$p_n(g(x)) = a_0 + a_1g(x) + \ldots + a_ng(x)^n$$

ist eine Funktion aus W. damit ist gezeigt

$$|g| \in \overline{W} \Longrightarrow |f| = (\|f\|+1)|g| \in \overline{W}. \qquad \square$$

6.4 Folgerung (Weierstrassscher Approximationssatz). `WeApp`
Sei

$$f : [a,b] \longrightarrow \mathbb{R} \qquad (a < b)$$

stetig. Es existiert eine Folge von Polynomen, die gleichmäßig gegen f konvergiert.

---

*) Die Stetigkeit folgt auch aus dem Abelschen Grenzwertsatz

Beweis. Setze $X := [a,b]$ und $W := \{$ Polynome auf $[a,b]\}$. Dann sind die Voraussetzungen von Theorem 6.2 erfüllt. □

Es ist interessant festzustellen, daß obiger Beweis des Satzes von STONE-WEIERSTRASS Theorem 6.2 sich zusammensetzt aus allgemeinen Betrachtungen und einem extremen Spezialfall ($f(x) = |x|$), in dem einmal eine solche Approximation explizit konstruiert werden mußte.

Eine Anwendung des Weierstrassschen Approximationssatzes.

Es sei f eine stetige Funktion auf dem Rechteck

$$[a,b] \times [c,d] \qquad (a < b,\ c < d).$$

Dann gilt

$$\int_c^d \left(\int_a^b f(x,y)\,dx \right) dy = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx.$$

Beweis. Für Polynome

$$f(x,y) = \sum_{0 \le i,k \le n} a_{ik} x^i y^k$$

kann man dies leicht nachrechnen. Aufgrund des Approximationssatzes läßt sich aber jede stetige Funktion gleichmäßig durch Polynome approximieren. □

7. Konvergenzkriterien

In der Analysis ist man oft vor das Problem gestellt, von einer vorgelegten Folge $a_1, a_2, a_3, \ldots$ zu entscheiden, ob sie konvergiert oder nicht und zwar ohne den Grenzwert selbst zu kennen. Will man aber direkt die Definition der Konvergenz als Kriterium benutzen, so muß man den Grenzwert ja kennen. Es ist also wichtig, Konvergenzkriterien zu erhalten, in denen der hypothetische Grenzwert nicht auftritt.

In der Analysis einer Veränderlichen kommt man mit folgendem Kriterium aus (der Leser sollte dies nachprüfen):

Eine ***monotone*** *Folge* $a_1 \le a_2 \le a_3 \le \ldots$ *konvergiert, wenn sie* ***beschränkt*** *ist.*

Der Beweis erfolgte mit Hilfe des *Vollständigkeitsaxioms*; der Grenzwert a ist nichts anderes als das *Supremum* der Menge aller Folgenglieder.

In beliebigen metrischen Räumen hat man keine Anordnung der Punkte, man kann also von monotonen Folgen gar nicht sprechen. Man muß sich daher in metrischen Räumen neue Kriterien für die Konvergenz einfallen lassen und muß sich einen Ersatz für das Vollständigkeitsaxiom verschaffen.

7.1 Satz. *Jede Folge (x_n) in einem **kompakten** metrischen Raum X besitzt eine konvergente Teilfolge.* meTeF

Es gibt also eine Folge $\nu_1 < \nu_2 < \nu_3 < \dots$ von natürlichen Zahlen, so daß die Folge $x_{\nu_1}, x_{\nu_2}, x_{\nu_3}, \dots$ konvergiert.

Da beschränkte Mengen im $\mathbb{R}^n$ in einem Kompaktum enthalten sind gilt insbesondere der

Satz von Bolzano-Weierstrass. *Jede beschränkte Folge im $\mathbb{R}^n$ besitzt eine konvergente Teilfolge (vgl. Theorem I.4.4).*

Beweis von Satz 7.1. Sei $a \in X$. Offenbar sind folgende Aussagen gleichbedeutend:

a) Es existiert eine Teilfolge, welche gegen a konvergiert.

b) Ist U eine beliebige Umgebung von a, so gilt $x_n \in U$ für unendlich viele n.

Wir beweisen Satz 7.1 indirekt, nehmen also an, daß keine konvergente Teilfolge existiert. Zu jedem Punkt $a \in X$ existiert dann eine positive Zahl $\varepsilon = \varepsilon(a) > 0$, so daß

$$x_n \in U_\varepsilon(a) \text{ nur für endlich viele } n$$

gilt. Da X von allen Kugeln $U_\varepsilon(a)$ überdeckt wird und als kompakt vorausgesetzt wurde, muß

$$X = U_{\varepsilon(a_1)}(a_1) \cup \dots \cup U_{\varepsilon(a_m)}(a_m)$$

gelten. Hieraus ergäbe sich aber, daß die Menge der natürlichen Zahlen endlich ist. Widerspruch! □

Anmerkung. Auch die Umkehrung von Satz 7.1 ist richtig. Wenn ein metrischer Raum die Eigenschaft hat, daß jede Folge eine konvergente Teilfolge besitzt, so ist er kompakt. (Beweis in Dieudonné, *Foundations of modern analysis*, S. 56).

Häufungspunkte.

Es gibt zwei verschieden Begriffe des „Häufungspunktes“, zum Einen für Mengen und zum anderen für Folgen.:

1) Eine Teilmenge $A \subset X$ eines metrischen Raumes X besitzt den Häufungspunkt $a \in X$, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

   a) a ist nicht isolierter Punkt von $A \cup \{a\}$ (wobei $A \cup \{a\}$ mit der induzierten Metrik zu versehen ist).

   b) a ist Randpunkt der Menge $A - \{a\}$.

   c) Es existiert eine Folge (x_n)

$$x_n \in A,\ x_n \neq a \text{ für alle } a \text{ mit } \lim_{n\to\infty} x_n = a.$$

2) Eine Folge (a_n) in einem metrischen Raum hat den Häufungspunkt*) a, wenn für jede Umgebung U von a

$$a_n \in U \text{ für unendlich viele } n \in \mathbb{N}$$

gilt, wenn also eine geeignete Teilfolge von (a_n) gegen a konvergiert.

Beispiel. Die konstante Folge $a, a, a, \ldots$ hat den Häufungspunkt a. Die Menge $\{a\}$ hat keinen Häufungspunkt.

Übungsaufgabe. Eine Folge in einem metrischen Raum konvergiert genau dann, wenn sie genau einen Häufungspunkt hat. Dieser Häufungspunkt ist dann der Grenzwert der Folge.

7.2 Definition. *Man nennt eine Folge (x_n) aus einem metrischen Raum X eine* ***Cauchyfolge,*** *(vgl. Definition I.4.7), wenn zu jedem $\varepsilon > 0$ eine natürliche Zahl $N = N(\varepsilon)$ existiert, so daß gilt:* CaFo

$$d(x_n, x_m) < \varepsilon \text{ für alle } n, m \geq N.$$

Natürlich ist jede konvergente Folge eine Cauchyfolge (wegen der Dreiecksungleichung

$$d(x_n, x) + d(x, x_m) \leq d(x_n, x_m).)$$

Die Umkehrung hiervon ist nicht in beliebigen metrischen Räumen richtig.

7.3 Definition. *Ein metrischer Raum X heißt* ***vollständig****, wenn jede Cauchfolge aus X konvergiert.* meVol

Der Begriff der Cauchyfolge und daher auch der der Vollständigkeit ist nicht topologischer Natur. Man kann Cauchyfolgen nicht mit dem Umgebungsbegriff allein erklären. Dennoch ändern sich diese Begriffe nicht beim Übergang von einer Metrik zu einer äquivalenten.

7.4 Bemerkung. *Seien d, d' zwei äquivalente Metriken auf X und (x_n) eine Folge aus X. Diese ist genau dann eine Cauchyfolge bezüglich d, wenn sie eine solche bezüglich d' ist.* aeqVol

Wir versehen den $\mathbb{R}^n$ mit der Maximumsmetrik.

7.5 Satz. *Der $\mathbb{R}^n$ ist vollständig.* voRn

---

*) Manchmal spricht man bei Folgen zur begrifflichen Unterscheidung auch von „Häufungswerten“ anstelle von Häufungspunkten.

Beweis. (vgl. Definition I.4.7) Sei (x_n) ein Cauchyfolge im $\mathbb{R}^n$.

1. Schritt. (x_n) ist beschränkt. Es existiert nämlich eine natürliche Zahl N mit

$$\|x_n - x_m\| < 1 \text{ für } n, m \geq N.$$

Bis auf endlich viele Ausnahmen liegen also alle Folgenglieder in einer Kugel vom Radius 1 mit Mittelpunkt x_N.

2. Schritt. (x_n) konvergiert. Nach dem Satz von BOLZANO-WEIERSTRASS existiert eine konvergente Teilfolge (x_{ν_n}). Deren Grenzwert sei x. Zu jedem $\varepsilon > 0$ existiert eine natürliche Zahl N mit den Eigenschaften

a) $\|x_n - x_m\| < \frac{\varepsilon}{2}$ für $n, m \geq N$

b) $\|x_{\nu_n} - x\| < \frac{\varepsilon}{2}$ für $n \geq N$.

Es folgt

$$\|x - x_n\| \leq \|x - x_{\nu_n}\| + \|x_{\nu_n} - x_n\| < \varepsilon \text{ für } n \geq N. \qquad \square$$

Beispiele.

1) Jeder kompakte metrische Raum ist vollständig.
2) Ist X eine nicht leere Menge, so ist $B(X)$ mit der Supremumsnorm ein vollständiger metrischer Raum.
3) Ist $X \neq \emptyset$ ein metrischer Raum, so ist $C(X) \cap B(X)$ mit der Supremumsnorm ein vollständiger metrischer Raum.

7.6 Banachscher Fixpunktsatz. BaFix

Es sei X ein vollständiger metrischer Raum und $f : X \to X$ eine ***kontrahierende*** *Selbstabbildung. Dann existiert genau ein Fixpunkt $a \in X$, d.h. ein Punkt $a \in X$ mit*

$$f(a) = a.$$

Eine Abbildung f heißt dabei *kontrahierend*, wenn eine Zahl ρ mit $0 < \rho < 1$ existiert, so daß gilt

$$d(f(x), f(y)) \leq \rho \cdot d(x, y).$$

Beweis, 1. Existenz. Es sei $a_0 \in X$ ein beliebiger Punkt. Wir definieren

$$a_1 := f(a_0),\ a_2 := f(a_1),\ a_3 := f(a_2), \ldots,$$

allgemein

$$a_n := f(a_{n-1}) \qquad (n \geq 1).$$

Wir zeigen, daß die Folge (a_n) konvergiert. Der Grenzwert a ist dann notwendigerweise ein Fixpunkt von f, denn es gilt für beliebiges $n \in \mathbb{N}$

$$\begin{aligned} d(a, f(a)) &\leq d(a, a_{n+1}) + d(a_{n+1}, f(a)) \\ &= d(a, a_{n+1}) + d(f(a_n), f(a)) \leq d(a, a_{n+1}) + \rho\, d(a_n, a). \end{aligned}$$

Dieser Ausdruck konvergiert gegen Null für $n \to \infty$ und es folgt

$$d(a, f(a)) = 0 \Longleftrightarrow f(a) = a.$$

Um nun die *Konvergenz* der Folge (a_n) zu beweisen, zeigen wir, daß (a_n) eine Cauchyfolge ist. Zunächst ist

$$d(a_{n+1}, a_n) = d(f(a_n), f(a_{n-1})) \le \rho\, d(a_n, a_{n-1}).$$

Durch Iteration ergibt sich dann

$$d(a_{n+1}, a_n) \le \rho^n\, d(a_1, a_0).$$

Sei nun $m > n$. Mit Hilfe der Dreiecksungleichung zeigt man nun

$$\begin{aligned} d(a_m, a_n) &\le d(a_m, a_{m-1}) + d(a_{m-1}, a_{m-2}) + \ldots + d(a_{n+1}, n) \\ &\le d(a_1, a_0)(\rho^{m-1} + \ldots + \rho^n) \\ &\le \rho^n\, d(a_1, a_0)(1 + \rho + \ldots + \rho^{m-n-1}) \le \frac{\rho^n}{1-\rho}\, d(a_1, a_0). \end{aligned}$$

Nach Voraussetzung ist $0 < \rho < 1$ und somit ist die Abschätzung durch die geometrische Reihe möglich und außerem ist (ρ^n) eine Nullfolge. Es gilt daher

$$d(a_n, a_m) < \varepsilon \text{ für } m > n \ge N(\varepsilon).$$

Somit ist (a_n) eine Cauchyfolge.

2. Eindeutigkeit. Es sei b ein weiterer von a verschiedener Fixpunkt aus f. Dann gilt

$$d(a, b) = d(f(a), f(b)) \le \rho\, d(a, b).$$

Hieraus folgt $\rho \ge 1$. Widerspruch. □

Im nächsten Abschnitt geben wir eine Anwendung des Banachschen Fixpunktsatzes.

Unter den vollständigen metrischen Räumen sind besonders diejenigen wichtig, deren Metrik von einer Norm herrührt. Diese bekommen einen eigenen Namen:

7.7 Definition. *Ein **Banachraum** ist ein normierter Vektorraum, welcher vollständig bezüglich der assoziierten Metrik $d(f, g) = \|f - g\|$ ist.* BanHil

*Ein **Hilbertraum** ist ein Vektorraum zusammen mit einem Skalarprodukt, so daß der assoziierte normierte Raum $\|f\| = \sqrt{\langle f, f\rangle}$ ein Banachraum ist.*

Beispiele für Banachräume sind $B(X)$ und $B(X) \cap C(X)$, wobei im ersten Fall X eine nicht leere Menge und im zweiten Fall sogar ein metrischer Raum ist. Beispiele für Hilberträume sind schwerer zu bekommen (s. Satz ???

Erst in weiterführenden Vorlesungen über Analysis (z.B. partielle Differentialgleichungen, Funktionalanalyis) und in der mathematischen Physik wird die Theorie der Banach- und Hilberträume wirklich nutzbar gemacht. Im Rahmen dieser Einführung in die Analysis werden diese Begriffe nicht gebraucht und daher nur am Rande erwähnt.

8. Differentialgleichungen

Der Banachsche Fixpunktsatz kann benutzt werden, um die Existenz von Lösungen gewisser Differentialgleichungen zu beweisen. Unter einer gewöhnlichen Differentialgleichung versteht man eine Gleichung der Art

$$f(x, y, y', \dots, y^{(n)}) = 0.$$

Gegeben ist hierbei eine Funktion f auf dem $\mathbb{R}^{k+1}$ oder allgemeiner auf einer Teilmenge desselben. Gesucht ist eine Funktion $y(x)$ einer Veränderlicher x, definiert auf einem geeigneten Intervall, so dass obige Gleichung erfüllt ist. In vielen Fällen wird es möglich sein, diese Gleichung nach der letzten auftretenden Ableitung $y^{(n)}$ aufzulösen. Sie hat dann die Form

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}),$$

wobei die Funktion f jetzt auf dem $\mathbb{R}^n$ (oder auf einer Teilmenge) definiert ist. In der ersten Form spricht man von einer gewöhnlichen Differentialgleichung in impliziter Form, im zweiten Fall von einer in expliziter Form. Sei y eine Lösung der expliziten Form. Setzt man

$$y_0 = y,\ y_1 = y', \dots, y_{n-1} = y^{(n-1)},$$

so erhält man eine Lösung des Gleichungssystems

$$\begin{aligned} y_0' &= y_1, \\ y_1' &= y_2, \\ &\cdots \\ y_{n-2}' &= y_{n-1}, \\ y_{n-1}' &= f(x, y_0, \dots, y_{n-1}), \end{aligned}$$

und umgekehrt erhält man aus jeder Lösung des Systems eine Lösung der expliziten Differentialgleichung zurück. Diese Überlegung zeigt, dass man nur Gleichungen der Form

$$y' = f(x, y)$$

lösen muss, wenn man in Kauf nimmt, dass f und y Funktionen mit Werten in $\mathbb{R}^n$ sind. In diesem Zusammenhang vereinbaren wir, dass die Ableitung einer Funktion $h : [a, b] \to \mathbb{R}^n$ komponentenweise zu verstehen ist, $h' := (h_1', \dots, h_n')$. Dasselbe gilt für das Integral.

Im folgenden bezeichnen wir mit $\| \cdot \|$ die Maximumsnorm auf $\mathbb{R}^n$. Ein Spezialfall des *Satzes von Picard Lindelöf* besagt.

8.1 Theorem. *Sei* DGL

$$f : [a,b] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \qquad (a < b)$$

eine stetige Funktion mit folgender Eigenschaft: Es existiere eine Konstante L mit

$$\|f(x,u) - f(x,v)\| \leq L\|u - v\|, \quad x \in [a,b],\ u,v \in \mathbb{R}^n.$$

Dann existiert zu jedem Vektor $c \in \mathbb{R}^n$ *ein eindeutig bestimmtes n-Tupel* $y = (y_1, \dots, y_n)$ *differenzierbarer Funktionen auf* $[a,b]$ *mit den beiden Eigenschaften*

a) $y(a) = c$.
b) $y'(x) = f(x, y(x))$.

Man nennt $y(x)$ wegen der Bedingung a) auch die Lösung des Anfangswertproblems zum Anfangswert $y(a) = c$.

Beweis von Theorem 8.1. Wir betrachten den Raum $\mathcal{C}([a,b])$ aller stetigen Funktionen mit der Metrik,

$$d(f,b) = \|f - g\|, \quad (\|f\| = \max\{|f(x)|;\ x \in X\}).$$

Wir wissen, dass dies ein vollständiger metrischer Raum ist. Wir betrachten folgende Abbildung

$$\Phi : \mathcal{C}([a,b]) \longrightarrow \mathcal{C}([a,b]), \quad (\Phi(h))(x) := c + \int_a^x F(t, h(t))\, dt.$$

Die Lösungen der Gleichung b) mit der Eigenschaft a) sind nach dem Hauptsatz der Diffferential- und Integralrechnung genau die Fixpunkte der Abbildung Φ. Es ist daher zu zeigen, dass diese Abbildung genau einen Fixpunkt hat. Wir überprüfen die Voraussetzung des Banachschen Fixpunktsatzes: Es gilt offenbar

$$\|\Phi(h_1) - \Phi(h_2)\| \leq L(b-a)\|h_1 - h_2\|.$$

Die Voraussetzungen des Banachschen Fixpunktsatzes sind erfüllt, wenn $L(b-a) < 1$ gilt. Diese Voraussetzung ist jedoch keine ernsthafte Einschränkung. Man teilt einfach das Intervall $[a,b]$ durch eine Partition $a = a_0 < \cdots < a_m = b$ in Teilintervalle auf, so dass $L(a_j - a_{j-1}) < 1$ für $1 \leq j \leq m$ gilt. Dann löst man das Anfangswertproblem für das Intervall $[a_0, a_1]$ und geht mit dem neu gewonnen Wert $y(a_1)$ in das Anfangswertproblem für das Intervall $[a_1, a_2]$. So fortfahrend erhält man eine Lösung für das ganze Intervall und auch die Eindeutigkeit ergibt sich aus der Eindeutigkeit für die Teilintervalle. □

In der Theorie der Differentialgleichungen findet man allgemeinere Versionen des Satzes von Picard Lindelöf. Zum Beispiel ist es nicht einzusehen, dass f auf einer Teilmenge der Form $[a,b] \times \mathbb{R}^n$ definiert sein soll. Auch allgemeinere Bereiche des $\mathbb{R}^{n+1}$ sind denkbar und kommen vor. Schließlich wird die Voraussetzung an die Konstante L abgeschwächt, indem nicht wie in Theorem 8.1 gefordert wird, dass L von x unabhängig ist. Man nennt die Konstante L übrigens eine *Lipschitzkonstante.* Allgemein nennt man eine Abbildung $f : X \to Y$ metrischer Räume *Lipschitz-stetig,* falls es eine Konstante L gibt, so dass

$$d_X(f(a), f(b)) \leq L d_Y(a, b)$$

gilt. Offenbar sind Lipschitz-stetige Funktionen auch im gewöhnlichen Sinne stetig. Mann kann sich überlegen, dass stetig differenzierbare Funktionen $f : [a, b] \to \mathbb{R}$ stets Lipschitz stetig sind. Man sollte sich also Lipschitzsetigkeit als einen zwischen Stetigkeit und Differenzierbarkeit angesiedelten Begriff vorstellen.

Die Analyse des Verlaufs von Lösungen der Differentialgleichungen, konkrete Beispielklassen sowie Varianten von Theorem 8.1 muss einer speziellen Vorlesung über Differentialgleichungen überlassen bleiben.

Kapitel VI. Differentialrechnung für Funktionen mehrerer Veränderlicher

1. Partielle Ableitungen und totale Differenzierbarkeit

In diesem Paragraphen sei der Definitionsbereich $D \subset \mathbb{R}^n$ stets ein *offener* Bereich im $\mathbb{R}^n$. Zu jedem Punkt $x_0 \in D$ existiert also noch eine volle Kugel

$$U_r(x_0) \subset D;\ U_r(x_0) = \{x \in \mathbb{R}^n;\ d(x, x_0) < r\}.$$

Dabei sei d die Euklidische oder Maximumsmetrik.

Wir studieren Funktionen

$$f : D \longrightarrow \mathbb{R}$$

oder allgemeiner Abbildungen

$$f : D \longrightarrow \mathbb{R}^m.$$

Jede solche Abbildung läßt sich zerlegen in ein m-Tupel von Funktionen

$$f = (f_1, \dots, f_m), \quad f_\nu : D \longrightarrow \mathbb{R}, \quad 1 \le \nu \le m,$$

die durch die Gleichung

$$f(x) = (f_1(x), \dots, f_m(x))$$

definiert sind. Entsprechend dieser Zerlegung lassen sich viele Fragen über die Abbildung f zurückführen auf den Fall von Funktionen ($m = 1$).

Beispielsweise ist f genau dann stetig, wenn die Komponenten stetig sind (Satz V.3.10).

Schwieriger ist es, die Dimension n des Definitionsbereiches zu erniedrigen. Tatsächlich führt man viele Fragen durch geeignete *Spezialisierung* der Variablen auf Funktionen einer Veränderlichen zurück.

Ein solcher Spezialisierungsprozeß soll nun beschrieben werden. Gegeben sei also der offene Bereich D und ein fester Punkt

$$a = (a_1, \dots, a_n) \in D.$$

Man kann dann die Menge

$$D_j := \left\{ x_j \in \mathbb{R};\quad (a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n) \in D \right\}$$

betrachten.

1.1 Bemerkung. *Ist $D \subset \mathbb{R}^n$ ein offener Bereich, so ist* soezOF

$$D_j := \{ x_j \in \mathbb{R}; \quad (a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n) \in D \}$$

ein offener Teil der reellen Geraden $\mathbb{R}$.

Da D_j offen ist, existiert $\varepsilon > 0$, so daß $(-\varepsilon, \varepsilon) \subset D$. Ist $f : D \to \mathbb{R}$ eine Funktion, so erhält man durch den Spezialisierungsprozeß

$$f_j : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}; \; f_j(x_j) = f(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n)$$

Funktionen einer Veränderlichen x_j. Es kann sein, daß diese Funktion einer Veränderlicher in a_j ableitbar ist. Dies hängt natürlich nicht von der Wahl von ε ab.

1.2 Definition. *Die Funktion* defPar

$$f : D \longrightarrow \mathbb{R}, \; D \subset \mathbb{R}^n \text{offen},$$

*heißt in dem Punkt $a = (a_1, \dots, a_n)$ nach der j-ten Variablen x_j **partiell ableitbar**, wenn die Funktion*

$$f_j : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}, \quad f_j(x_j) = f(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_n),$$

im Punkt a_j differenzierbar ist.
Schreibweise: $\partial_j f(a_1, \dots, a_n) = f_j'(a_j)$.

Es ist also

$$\partial_j f(a_1, \dots, a_n) = \lim_{h \to 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h}.$$

Die Funktion heißt *in D partiell ableitbar*, wenn sie in *jedem* Punkt nach *jeder* Variablen partiell ableitbar ist. Man kann dann die partiellen Ableitungen

$$\partial_j f, \quad j = 1, \dots, n,$$

wieder als Funktionen auffassen, die in D definiert sind. Häufig verwendet man auch die Schreibweise

$$\frac{\partial f}{\partial x_j} = \partial_j f.$$

Im Fall $n = 1$ war gezeigt worden, daß aus der Ableitbarkeit einer Funktion ihre Stetigkeit folgt. Im Fall $n > 1$ ist dies (für *partielle Ableitbarkeit* anstelle von *Ableitbarkeit*) nicht der Fall, wie man durch Gegenbeispiele belegen kann.

1.3 Satz. *Die Funktion* PaSte

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \textit{ offen},$$

sei partiell ableitbar. Ist $a \in D$ *ein Punkt, so daß eine Umgebung* $U(a)$ *existiert, in der alle partiellen Ableitungen von* f *beschränkt sind,*

$$|\partial_j f(x)| \le C \text{ für alle } x \in U(a),\ j = 1, \dots, n,$$

so ist f *stetig in* a.

Diese Voraussetzung ist beispielsweise dann erfüllt, wenn die partiellen Ableitungen in a stetig sind.

Wenn also eine Funktion stetig partiell differenzierbar ist (d.h. die Ableitungen $\partial_j f$ existieren und sind stetig), so muß f stetig sein.

Beweis von 1.3. Um die Beweisidee klar hervortreten zu lassen, beschränken wir uns auf den Fall zweier Variabler und überlassen es dem geduldigen Leser, den Beweis auf n Variablen zu übertragen. Wir schreiben (x, y) und (a, b) anstelle von (x_1, x_2) und (a_1, a_2). Es gilt

$$f(x, y) - f(a, b) = (f(x, y) - f(a, y)) + (f(a, y) - f(a, b)).$$

Nach dem Mittelwertsatz der Differentialrechnung existieren Zwischenstellen

$$\xi \text{ zwischen } a \text{ und } x, \qquad \eta \text{ zwischen } b \text{ und } y$$

mit der Eigenschaft

$$f(x, y) - f(a, b) = \partial_x f(\xi, y) \cdot (x - a) + \partial_y f(a, \eta) \cdot (y - b).$$

In einer Umgebung von (a, b) gilt

$$|f(x, y) - f(a, b)| \le C\,(|x - a| + |y - b|)$$

mit einer gewissen Konstanten C. □

Höhere partielle Ableitungen.

Häufig sind auch die partiellen Ableitungen $\partial_j f$ ihrerseits wieder partiell ableitbar; man kann daher *höhere partielle Ableitungen* bilden.

Sprechweise: Eine Funktion

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \textit{ offen},$$

heißt ***k-fach stetig partiell ableitbar****, wenn die partiellen Ableitungen*

$$\partial_{j_1} \dots \partial_{j_k} f \text{ für alle } 1 \le j_1, \dots, j_k \le n$$

existieren und stetig sind.

Insbesondere ist dann f stetig. Wir bezeichnen die Menge dieser Funktionen mit $\mathcal{C}^k(D)$.

1.4 Satz. *Die Funktion* ParKo

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \textit{ offen},$$

sei zweimal stetig partiell ableitbar. Dann gilt

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

(In Worten: Partielle Ableitungen darf man vertauschen.)

Beweis. Da nur nach zwei Variablen partiell abgeleitet wird, darf man $n = 2$ annehmen. Wir schreiben (x, y) anstelle von (x_1, x_2). Sei $(a, b) \in D$ ein fest gewählter Punkt. Wir betrachten die Funktion ($x \neq a,\ y \neq b$)

$$H(x, y) := \frac{f(x, y) - f(a, y) - f(x, b) + f(a, b)}{(x - a)(y - b)}$$

und formen diese mit Hilfe des Mittelwertsatzes auf zweierlei Weise um.

1) $$H(x, y) = (y - b)^{-1} \frac{[f(x, y) - f(x, b)] - [f(a, y) - f(a, b)]}{x - a}.$$

Wendet man den Mittelwertsatz auf die Funktion

$$g(x) := f(x, y) - f(x, b) \qquad (y \text{ und } b \text{ fest})$$

an, so resultiert

$$H(x, y) = (y - b)^{-1} \left(\partial_x f(\xi, y) - \partial_x f(\xi, b)\right)$$

mit einem ξ zwischen a und x. Jetzt kann man den Mittelwertsatz auf die Funktion

$$h(y) := \partial_x f(\xi, y) - \partial_x f(\xi, b)$$

anwenden und erhält

$$H(x, y) = \partial_y \partial_x f(\xi, \eta)$$

mit ξ zwischen a und x und η zwischen b und y.

Durch Anwendung des Mittelwertsatzes auf

$$H(x, y) = (x - a)^{-1} \frac{[f(x, y) - f(a, y)] - [f(x, b) - f(a, b)]}{y - b}$$

(also indem man mit der Funktion $g^*(y) := f(x, y) - f(a, y)$ beginnt) folgt analog zu 1)

$$H(x, y) = \partial_x \partial_y f(\xi', \eta')$$

mit gewissen Zwischenstellen ξ' und η'.

Nach Voraussetzung sind aber die partiellen Ableitungen

$$\partial_x \partial_y f \text{ und} \partial_y \partial_x f$$

stetig. Vollzieht man die Grenzübergänge $x \to a$ und $y \to b$, so folgt daher

$$\lim_{(x,y)\to(a,b)} H(x,y) = \partial_x \partial_y f(a,b) = \partial_y \partial_x f(a,b).$$

Dies gilt für jeden Punkt $(a,b) \in D$. □

Wir wenden uns nun der etwas allgemeineren Situation einer Abbildung

$$f : D \longrightarrow \mathbb{R}^m, \quad D \subset \mathbb{R}^n \text{offen},$$

zu. Wir zerlegen sie in ihre Komponenten

$$f = (f_1, \dots, f_m).$$

Die sogenannte *Funktionalmatrix* (oder Jacobi-Matrix) von f faßt alle partiellen Ableitungen von $f_1, \dots, f_m$ zusammen,

$$J(f;x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} = \left(\frac{\partial f_\nu}{\partial x_\mu} \right)_{\nu\mu}$$

Diese Matrix, die für jedes $x \in D$ zu bilden ist (partielle Ableitbarkeit vorausgesetzt), hat also m Zeilen und n Spalten.

Wir untersuchen nun folgendes Problem:

Gegeben seien zwei Abbildungen

$$\begin{aligned} f : D &\longrightarrow \mathbb{R}^m, & D &\subset \mathbb{R}^n \text{offen}, \\ g : D' &\longrightarrow \mathbb{R}^p, & D' &\subset \mathbb{R}^m \text{offen}, \end{aligned}$$

die sich zusammensetzen lassen zu einer Abbildung

$$h : D \longrightarrow \mathbb{R}^p, \; h(x) = g\left(f(x)\right).$$

Der Wertevorrat $f(D)$ muß dazu im Definitionsbereich D' von g enthalten sein. Wie berechnen sich dann die partiellen Ableitungen von $h = g \circ f$ aus denen von f und g (Existenz vorausgesetzt)?

Wir erinnern daran, daß im Fall $n = m = p = 1$ die *Kettenregel*

$$(g \circ f)'(a) = g'\big(f(a)\big) f'(a)$$

gilt. Diese gilt es zu verallgemeinern.

Hierzu ist es nützlich, sich vor Augen zu halten, daß Matrizen im Zusammenhang mit linearen Abbildungen stehen. Ist

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

eine Matrix von m Zeilen und n Spalten, so ordnet man ihr eine gewisse Abbildung

$$l_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

zu. Diese werde definiert durch die Formel

$$l_A(x) = y$$

mit

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n \text{ und } y = (y_1, \dots, y_m) \in \mathbb{R}^m,$$

wobei für $i = 1, \dots, m$ gelte

$$y_i = \sum_{j=1}^{n} a_{ij} x_j .$$

Man schreibt dann meistens

$$A(x) \text{ anstatt } l_A(x).$$

Ist eine weitere lineare Abbildung

$$l_B : \mathbb{R}^m \longrightarrow \mathbb{R}^p$$

gegeben, die also durch eine Matrix

$$B = \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & & \vdots \\ b_{p1} & \dots & b_{pm} \end{pmatrix}$$

definiert wird, so kann man die zusammengesetzte Abbildung

$$l_B \circ l_A : \mathbb{R}^n \longrightarrow \mathbb{R}^p$$

betrachten. Aus der *linearen Algebra* ist bekannt, daß diese wieder durch eine Matrix beschrieben wird und zwar gilt

$$l_B \circ l_A := l_C : \mathbb{R}^n \longrightarrow \mathbb{R}^p,$$

wobei die Matrix

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{p1} & \dots & c_{pn} \end{pmatrix}$$

sich aus A und B auf folgende Weise berechnet:

$$c_{ik} = \sum_{k=1}^{m} a_{ij} b_{jk} \qquad (1 \leq i \leq p;\ 1 \leq j \leq n).$$

Halten wir noch einmal fest: Es gilt

$$B(Ax) = C(x) \quad \text{mit} \quad C = B \cdot A \text{ (Matrizenprodukt).}$$

Merkregel. c_{ij} ergibt sich als Skalarprodukt der i-ten Zeile von B mit der j-ten Spalte von A.

Schreibt man die Elemente x, y des $\mathbb{R}^n$ als Spaltenvektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} x_y \\ \vdots \\ y_n \end{pmatrix},$$

so ist $\vec{y}$ nichts anderes als das Produkt der Matrix A mit der Spalte $\vec{x}$,

$$\vec{y} = A\vec{x}.$$

Aus diesem Grunde ist es zweckmäßig, die Elemente des $\mathbb{R}^n$ in diesem Zusammenhang nicht wie gewohnt als Zeilen-, sondern als Spaltenvektoren aufzufassen. Wir tun dies gelegentlich aber nicht systematisch.

Bevor wir nun eine neue Interpretation der Funktionalmatrix geben, soll zunächst die Definition der Ableitung einer Funktion im Fall $n = m = 1$ umformuliert werden. Die Bedingung für Ableitbarkeit war, daß der Grenzwert

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

existiere, d.h.

$$\frac{f(x) - f(a)}{x - a} - f'(a) = R(x),$$

wobei $R(x) \to 0$ für $x \to a$ gilt. Definiert man

$$r(x) = (x - a)R(x),$$

so kann man dafür auch schreiben

$$f(x) = f(a) + f'(a)(x - a) + r(x),$$

wobei das Restglied $r(x)$ der Bedingung genügt:

$$\frac{r(x)}{x-a} \longrightarrow 0 \text{ für } x \longrightarrow a.$$

Gleichwertig und für uns passender ist $r(x)/|x-a| \to 0$. Genau im besagten Fall ist also f an der Stelle a ableitbar. Dies läßt sich nun formal auf den allgemeinen Fall übertragen: man ersetze lediglich $f'(a)$ durch die Jacobi-Matrix $J(f;a)$.

Da $x-a$ ein Punkt im $\mathbb{R}^n$ ist, hat $J(f;a)(x-a)$ einen wohldefinierten Sinn. Man kann jedenfalls $r(x)$ durch die Gleichung

$$f(x) = f(a) + J(f;a)(x-a) + r(x)$$

definieren und sich fragen, ob gilt:

$$\frac{r(x)}{\|x-a\|} \longrightarrow 0 \text{ für } x \longrightarrow a.$$

Dabei sei $\|\cdot\|$ die Maximumsnorm. Wie man sich leicht überlegt, kann man genauso gut eine äquivalente Norm nehmen.

1.5 Satz. *Die Abbildung* UePaTo

$$f : D \longrightarrow \mathbb{R}^m, \quad D \subset \mathbb{R}^n \textit{ offen},$$

sei für den festen Punkt $a \in D$ stetig partiell ableitbar. Definiert man dann $r(x)$ durch die Gleichung

$$f(x) = f(a) + J(f;a)(x-a) + r(x),$$

so gilt

$$\frac{r(x)}{\|x-a\|} \longrightarrow 0 \textit{ für } x \longrightarrow a.$$

(Zu jedem $\varepsilon > 0$ existiert also ein $\delta > 0$ mit

$$\|r(x)\| < \varepsilon\|x-a\| \text{ für } \|x-a\| < \delta.$$

Hierbei ist $\|\cdot\|$ eine der Standardnormen (Maximums- oder Euklidische Norm)).

Beweis. Bezeichnet man mit $f_\nu(x)$ bzw. $r_\nu(x)$ die ν-te Komponente von $f(x)$ bzw. $r(x)$ $(1 \le \nu \le m)$ und mit $\mathcal{J}_\nu(f,a)$ die ν-te Zeile von $J(f;a)$, so gilt

$$f_\nu(x) = f_\nu(a) + \mathcal{J}_\nu(f,a)(x-a) + r_\nu(x).$$

Außerdem ist $\mathcal{J}_\nu(f,a) = J(f_\nu;a)$ die Funktionalmatrix der Funktion f_ν, die nur aus einer Zeile besteht.

Damit ist klar, daß man die Behauptung nur im Fall $m = 1$ beweisen muß, denn aus

$$\frac{r_\nu(x)}{\|x-a\|} \longrightarrow 0 \text{ für } x \longrightarrow a \text{ für } \nu = 1, \dots, m$$

folgt dann

$$\frac{r(x)}{\|x-a\|} \longrightarrow 0 \text{ für } x \longrightarrow a.$$

Im Fall $m = 1$ sieht die Formel in Satz 1.5 folgendermaßen aus:

$$f(x) = f(a) + \sum_{\nu=1}^{n} \partial_\nu f(a)(x_\nu - a_\nu) + r(x).$$

Der Beweis wird wiederum besonders übersichtlich, wenn man sich auf den Fall $n = 2$ beschränkt. Wir schreiben dabei wieder

$$(x,y),\ (a,b) \text{ anstatt } (x_1,x_2),\ (a_1,a_2).$$

Dann hat man

$$f(x,y) = f(a,b) + (x-a)\partial_x f(a,b) + (y-b)\partial_y(a,b) + r(x,y).$$

Andererseits erhält man aus dem Mittelwertsatz der Differentialrechnung

$$\begin{aligned} f(x,y) - f(a,b) &= [f(x,y) - f(a,y)] + [f(a,y) - f(a,b)] \\ &= (x-a)\partial_x f(\xi,y) + (y-b)\partial_y f(a,\eta) \end{aligned}$$

mit ξ zwischen a und x und η zwischen b und y.

Ein Vergleich mit obiger Formel ergibt

$$r(x,y) = (x-a)[\partial_x f(\xi,y) - \partial_x f(a,b)] + (y-b)[\partial_y f(a,\eta) - \partial_y f(a,b)].$$

Beachtet man die Ungleichungen

$$\frac{|x-a|}{\|(x,y)-(a,b)\|} \leq 1 \text{ sowie } \frac{|x-a|}{\|(x,y)-(a,b)\|} \leq 1,$$

so ergibt sich

$$\frac{r(x,y)}{\|(x,y)-(a,b)\|} \leq |\partial_x f(\xi,y) - \partial_x f(a,b)| + |\partial_y f(a,\eta) - \partial_y f(a,b)|.$$

Für $(x,y) \to (a,b)$ konvergiert die rechte Seite aber gegen Null, denn die partiellen Ableitungen sind nach Voraussetzung stetig. □

Durch Satz 1.5 wird ein neuer Zugang zur Differentialrechnung mehrerer Veränderlicher nahegelegt.

1.6 Definition. *Eine Abbildung* TotDi

$$f : D \longrightarrow \mathbb{R}^m \quad D \subset \mathbb{R}^n \textit{ offen},$$

heißt in einem Punkt $a \in D$ ***total differenzierbar****, wenn es eine Matrix*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

gibt, so daß gilt:

$$f(x) = f(a) + A(x - a) + r(x)$$

mit

$$\frac{r(x)}{\|x - a\|} \longrightarrow 0 \textit{ für } x \longrightarrow a.$$

(Man kann mit dieser Begriffsbildung den Satz 1.5 dann einfach so ausdrücken: Jede stetig (partiell) differenzierbare Funktion ist total differenzierbar.)

Tatsächlich ist jede total differenzierbare Funktion auch partiell differenzierbar und zwar gilt:

1.7 Lemma. *Die Abbildung* $f : D \to \mathbb{R}^m$ *sei total differenzierbar in* $a \in D$ *(Bezeichnungen wie in Definition 1.6). Dann ist sie auch partiell differenzierbar in* a*, und es gilt* TofPa

$$J(f; a) = A.$$

Beweis. Wir nehmen wieder $n = 2$, $m = 1$ an. Es ist dann einfach, den Beweis auf beliebige natürliche n und m zu verallgemeinern. Die Formel lautet

$$f(x, y) = f(a, b) + \alpha(x - a) + \beta(y - b) + r(x, y), \; A = (\alpha, \beta).$$

Es folgt

$$\frac{f(x, b) - f(a, b)}{x - a} = \alpha + \frac{r(x, b)}{x - a}.$$

Durch Grenzübergang $x \to a$ erhält man

$$\partial_x f(a, b) = \alpha \qquad \text{(ebenso } \partial_y f(a, b) = \beta).$$ □

Halten wir fest:

$$\boxed{\text{stetig ableitbar} \Longrightarrow \text{total ableitbar} \Longrightarrow \text{partiell ableitbar.}}$$

Wir gelangen nun zur *Kettenregel*.

1.8 Theorem. *Gegeben seien zwei stetig differenzierbare Abbildungen* KetM

$$f : D \longrightarrow \mathbb{R}^m, \quad D \subset \mathbb{R}^n \textit{ offen,}$$
$$g : D' \longrightarrow \mathbb{R}^p, \quad D' \subset \mathbb{R}^m \textit{ offen,}$$

die sich zu einer Abbildung

$$h : D \longrightarrow \mathbb{R}^p;\ h(x) = g\,(f(x))\,.$$

zusammensetzen lassen. Dann ist auch h stetig (partiell) ableitbar und es gilt für jedes $a \in D$

$$J(g \circ f; a) = J(g; f(a)) \cdot J(f; a) \qquad \textit{(Matrizenprodukt).}$$

Beweis. Nach Voraussetzung gilt

$$f(x) = f(a) + J(f; a)(x - a) + r(x),$$
$$g(y) = g(b) + J(f; b)(y - b) + s(y),$$

wobei die Restglieder noch nach Division durch $\|x-a\|$ bzw. $\|y-b\|$ gegen Null streben (für $x \to a$ bzw. $y \to b$). Im Spezialfall $b = f(a)$ erhält man

$$g(y) = g(f(a)) + J(g; f(a)(y - f(a)) + s(y).$$

Hierin können wir speziell $y = f(x)$ eintragen. Mit der Bezeichnung $h = g \circ f$ folgt dann

$$\begin{aligned} h(x) &= h(a) + J(g; f(a))(f(x) - f(a)) + s(f(x)) \\ &= h(a) + J(g; f(a)) \cdot J(f; a)(x - a) + J(g; f(a)) \cdot r(x) + s(f(x)). \end{aligned}$$

Wenn wir zeigen können, daß die Funktion

$$\rho(x) := J(g; f(a))r(x) + s(f(x))$$

Restgliedcharakter hat (d.h.

$$\frac{\rho(x)}{\|x - a\|} \longrightarrow 0 \text{ für } x \longrightarrow a),$$

so sind wir fertig, denn nach Lemma 1.7 ist dann h in a partiell differenzierbar und hat als Funktionalmatrix gerade

$$J(g; f(a)) \cdot J(f; a).$$

Daß diese stetig von a abhängt ist dabei klar, da dies auf $J(g; f(a))$ und $J(f; a)$ zutrifft.

Abschätzung von $\dfrac{\rho(x)}{\|x - a\|}$.

1) $\dfrac{J(g; f(a)r(x)}{\|x - a\|} \longrightarrow 0$ für $x \longrightarrow a$.

Diese Aussage ist richtig, was sich sofort ergibt aus

1.9 Hilfssatz. *Sei A eine feste Matrix. Es existiert eine Konstante C mit* NoMaV

$$\|Ax\| \le C\|x\|$$

für alle Vektoren x (mit gleicher Komponentenanzahl wie die Anzahl der Spalten von A).

Sowohl bei der Maximumsnorm als auch bei der Euklidischen Norm hat man die Abschätzung

$$\Big|\sum_j a_{ij}x_j\Big| \le \sum_j |a_{ij}|\|x\| \le C_i\|x\| \text{ für jedes } i,$$

und daher

$$\|Ax\| \le C\|x\|.$$

2) $\dfrac{s(f(x))}{\|x-a\|} \longrightarrow 0$ für $x \longrightarrow a$.

Nach Voraussetzung gilt

$$\frac{s(y)}{\|y-f(a)\|} \longrightarrow 0 \text{ für } y \longrightarrow f(a);$$

das kann man auch so ausdrücken:

Die Funktion

$$l(y) := \begin{cases} \frac{\|s(y)\|}{\|y-f(a)\|} & \text{für } y \ne f(a) \\ 0 & \text{für } y = f(a) \end{cases}$$

ist stetig im Punkt $y = f(a)$. Es gilt offenbar

$$\frac{\|s(f(x))\|}{\|x-a\|} = l(f(x))\frac{\|f(x)-f(a)\|}{\|x-a\|}.$$

Wir wollen zeigen, daß dieser Ausdruck gegen Null strebt für $x \to a$. Dazu braucht man nur zu wissen, daß

$$\frac{\|f(x)-f(a)\|}{\|x-a\|} \le C$$

in einer Umgebung von a gilt. Dies sieht man mit Hilfe der Darstellung

$$f(x) - f(a) = J(f;a)(x-a) + r(x). \qquad \square$$

Andere Formulierungen der Kettenregel und Anwendungen.

Wir betrachten den Spezialfall $p = 1$:

$$f : D \longrightarrow \mathbb{R}^m,\ g : D' \longrightarrow \mathbb{R},$$
$$(D \subset \mathbb{R}^n,\ D' \subset \mathbb{R}^m \text{ beide offen},\ f(D) \subset D').$$

Dann ist die Jacobi-Matrix von g einfach eine Zeile

$$J(g;b) = (\partial_1 g, \dots, \partial_m g)(b).$$

Zerlegt man $f = (f_1, \dots, f_m)$ und $h = g \circ f = (h_1, \dots, h_p)$ in die Komponenten, so besagt die Kettenregel explizit

$$\frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(b) \frac{\partial f_k}{\partial x_j}(a) \qquad (b = f(a)).$$

Häufig schreibt man dafür einfach etwas schlampig

$$\boxed{\frac{\partial h_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k} \frac{\partial f_k}{\partial x_j}.}$$

Ein weiterer wichtiger Spezialfall ist der Fall $n = 1, p = 1$. Dann hat man also m Funktionen einer Veränderlichen

$$f_1, \dots, f_m : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R},$$

und eine Funktion von m Veränderlichen

$$g : D' \longrightarrow \mathbb{R}, \quad D' \subset \mathbb{R}^m \text{ offen.}$$

Die Funktion

$$h : D \longrightarrow \mathbb{R},\ h(x) = g(f_1(x), \dots, f_m(x)),$$

hängt wieder nur von der einen Variablen x ab. Die Kettenregel besagt jetzt:

$$\boxed{h' = \sum_{k=1}^{m} (\partial_k g) \cdot f_k'.}$$

Anwendungen der Kettenregel,

1. Richtungsableitung.

Gegeben sei eine stetig differenzierbare Funktion

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \text{ offen.}$$

Außerdem seien $a \in D$ ein fester Punkt und $\alpha \in \mathbb{R}^n$ ein Vektor. Unter der *Richtungsableitung* von f im Punkte a in Richtung α versteht man die Ableitung der Funktion

$$g(t) := f(a + t\alpha)$$

im Nullpunkt. Aus der Kettenregel folgt

$$g'(0) = \sum_{\nu=1}^{n} \partial_\nu f(a)\alpha_\nu.$$

Dies ist das Skalarprodukt des Vektors $(\partial_1 f(a), \dots, \partial_n f(a))$ mit α. Den Vektor

$$(\nabla f)(a) = (\partial_1 f(a), \dots, \partial_n f(a))$$

nennt man auch den *Gradienten* von f in a. Man erhält also

$$\boxed{g'(0) = \langle (\nabla f)(a), \alpha \rangle.}$$

2. Koordinatentransformation.

Seien $D, D' \subset \mathbb{R}^n$ offene Teile. Ein $\mathcal{C}^k$-*Diffeomorphismus* ($k \in \mathbb{N}$) ist eine bijektive Abbildung $\varphi : D \longrightarrow D'$, so daß sowohl φ als auch φ^{-1} k-fach stetig differenzierbar sind.

Sei

$$f : D' \longrightarrow \mathbb{R}$$

eine k-fach stetig differenzierbare Funktion. Dann ist auch

$$g = f \circ \varphi : D \longrightarrow \mathbb{R}$$

k-fach stetig differenzierbar und man erhält umgekehrt jede derartige Funktion $g \in \mathcal{C}^k(D)$ auf diesem Wege. Man muß lediglich $f := g \circ \varphi^{-1}$ setzen.

Sprech- und Schreibweise.

Die Funktion g entsteht aus f durch die *Koordinatentransformation* φ

$$g(y) = f(x), \quad y = \varphi(x).$$

Die partiellen Ableitungen der Funktion g erhält man aus denen von f durch Anwenden der Kettenregel

$$\partial_j f = \sum_{\nu=1}^{n} (\partial_\nu g) \frac{\partial y_\nu}{\partial x_j}$$

(genauer

$$\partial_j f(a) = \sum_{\nu=1}^{n} \partial_\nu g(\varphi(a)) \frac{\partial \varphi_\nu}{\partial x_j}(a) \;).$$

Sind f und φ sogar zweimal stetig differenzierbar, so ergibt sich

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \sum_{\nu=1}^{n} \partial_\nu g \frac{\partial^2 y_\nu}{\partial x_i \partial x_j} \sum_{\nu=1}^{n} \frac{\partial^2 g}{\partial x_i \partial y_\nu} \frac{\partial y_\nu}{\partial x_j},$$

also

$$\boxed{\frac{\partial^2 f}{\partial x_i \partial x_j} = \sum_{\nu=1}^{n} \partial_\nu g \frac{\partial^2 y_\nu}{\partial x_i \partial x_j} \sum_{\nu=1}^{n} \sum_{l=1}^{n} \frac{\partial^2 g}{\partial y_l \partial y_\nu} \frac{\partial y_l}{\partial x_i} \frac{\partial y_\nu}{\partial x_j}}$$

Polarkoordinaten

$$x = r\cos\varphi,\ \ y = r\sin\varphi.$$

Seien $D, D' \subset \mathbb{R}^2$ offene Teile, so daß durch die Abbildung

$$(r,\varphi) \mapsto (x,y) = h(r,\varphi) = (r\cos\varphi, r\sin\varphi)$$

ein Diffeomorphismus

$$h : D' \longrightarrow D$$

vermittelt wird, etwa

$$D' : = \{(r,\varphi) \in \mathbb{R}^2;\ r > 0,\ 0 < \varphi < 2\pi\}$$
$$D : = \mathbb{R}^2 - \{(x,y);\ x = 0,\ y \geq 0\}.$$

Der Laplace-Operator Δ auf D ist durch

$$\Delta f := \frac{\partial^2 f}{(\partial x)^2} + \frac{\partial^2 f}{(\partial y)^2},\ f \in \mathcal{C}^2(D),$$

erklärt. Wir wollen diesen Operator auf Polarkoordinaten umtransformieren. Darunter ist folgendes zu verstehen: Für eine Funktion $g \in \mathcal{C}^2(D')$ setzen wir

$$\Delta^* g := \Delta f \circ h \qquad (g = f \circ h).$$

Man nennt Δ^* den auf Polarkoordinaten transformierten Operator. Aus der Kettenregel ergibt sich nach einiger Rechnung:

$$\boxed{\Delta^* g = \frac{\partial^2 g}{(\partial r)^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{(\partial\varphi)^2}.}$$

2. Der Satz für implizite Funktionen

Es seien n Funktionen von n Veränderlichen gegeben, also eine Abbildung

$$f : D \longrightarrow \mathbb{R}^n,\ D \subset \mathbb{R}^n \text{ offen.}$$

Wir fragen uns nun, wann eine derartige Abbildung *umkehrbar* ist, d.h. wann eine Abbildung

$$g : f(D) \longrightarrow \mathbb{R}^n$$

existiert mit

$$f(x) = y \Longleftrightarrow x = g(y).$$

Die Gleichung

$$f(x) = y$$

ist dann bei gegebenem y aus dem Wertevorrat von f eindeutig durch ein $x \in D$ lösbar und dieses x ist gleich $g(y)$.

Nehmen wir einmal an, daß f und g stetig differenzierbar sind. Dann muß nach der Kettenregel gelten (beachte $g \circ f = \mathrm{id}$)

$$\mathcal{J}\,(g, f(a)) \cdot J(f; a) = E \qquad\qquad \text{(Einheitsmatrix)},$$

so daß also die Jacobi-Matrix $J(f; a)$ für alle $a \in D$ invertierbar sein muß.

Notwendig für die Umkehrbarkeit einer Abbildung

$$f : D \longrightarrow \mathbb{R}^n$$

(unter gewissen Differenzierbarkeitsbedingungen) ist die Invertierbarkeit der Jacobi-Matrix $J(f; a)$ für jedes $a \in D$.

Leider ist diese Bedingung nicht hinreichend, wie folgendes Beispiel zeigt:

Es sei

$$D = \mathbb{R}^2 - \{0\} = \text{punktierte Ebene}$$

und $f : D \longrightarrow \mathbb{R}^2$ definiert durch

$$(r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

Man überzeugt sich leicht davon, daß die Funktionalmatrix invertierbar ist. (Dies geschieht mit Hilfe der Funktionaldeterminante $\det J(f; a)$. Aus der linearen Algebra weiß man, daß eine quadratische Matrix genau dann invertierbar ist, wenn ihre Determinante nicht verschwindet.)

Diese Abbildung f ist aber nicht eineindeutig, denn sie hat in Abhängigkeit von φ die Periode 2π. Aber sie ist in einem eingeschränkten Sinn dennoch umkehrbar. Greift man einen beliebigen Streifen $D_0 \subset D$ der Breite 2π heraus, also

$$D_0 := \big\{\, (r, \varphi); \quad r > 0;\ \varphi_0 < \varphi \leq \varphi_0 + 2\pi \,\big\},$$

so ist die Einschränkung von f auf D_0 umkehrbar.

Kehren wir zurück zur allgemeinen Situation. Wenn die Jacobi-Matrix $J(f; a)$ für alle $a \in D$ invertierbar ist, so kann man allenfalls erwarten, daß f *lokal umkehrbar* ist, d.h. daß zu jedem Punkt $a \in D$ eine offene Umgebung $a \in D_0 \subset D$ existiert, so daß die Einschränkung $f|D_0$ von f auf D_0 umkehrbar ist. Dies läßt sich tatsächlich beweisen.

2.1 Theorem (Satz für umkehrbare Funktionen, erster Teil). SafUme
Es sei

$$f : D \longrightarrow \mathbb{R}^n, \ D \subset \mathbb{R}^n \text{ offen,}$$

eine stetig differenzierbare Abbildung mit invertierbarer Jacobi-Matrix $J(f;a)$ für jedes $a \in D$. Dann hat die Abbildung f die folgenden beiden Eigenschaften

a) *f ist offen (also: $D_0 \subset D$ offen $\Longrightarrow$ $f(D_0) \subset \mathbb{R}^n$ offen).*

b) *f ist lokal umkehrbar, d.h. zu jedem $a \in D$ existiert eine offene Umgebung D_0, $a \in D_0 \subset D$, so daß die Einschränkung $f|D_0$ umkehrbar ist.*

Wir merken noch an:

Ist f stetig differenzierbar und $a \in D$ eine fester Punkt, für den die Jacobimatrix invertierbar ist, so existiert eine offene Umgebung $a \in D' \subset D$, so daß $J(f,b)$ für alle $b \in D'$ invertierbar ist (und man kann Theorem 2.1 auf D' anstelle von D anwenden.

(Die Jacobimatrix ist genau dann invertierbar, wenn ihre Determinante von Null verschieden ist. Diese Eigenschaft überträgt sich wegen der Stetigkeit der Funktionalmatrix und damit der Funktionaldeterminante von a auf eine voll Umgebung.)

Beweis von Theorem 2.1. Offenbar genügt es, folgendes zu zeigen:

Sei $a \in D$. Es gibt positive Zahlen $R, r > 0$ mit der Eigenschaft:
Zu jedem $y_0 \in \mathbb{R}^n$ mit

$$\|y_0 - f(a)\| < R$$

existiert genau ein $x_0 \in D$ mit $\|x_0 - a\| < r$ und

$$f(x_0) = y_0.$$

(Dann ist also $U_R(f(a)) \subset f(D)$, d.h. $f(D)$ ist offen und außerdem ist f eingeschränkt auf $D \cap U_r(a)$ eineindeutig.)

Es ist im folgenden eine Vereinfachung anzunehmen, daß gilt

$$a = f(a) = 0 \quad \text{und} \quad J(f;a) = E \quad \text{Einheitsmatrix.}$$

Wir werden später sehen, daß dies keine wesentliche Einschränkung der Allgemeinheit bedeutet.

Um unser Problem dem *Banachschen Fixpunktsatz* zugänglich zu machen, schreiben wir die Gleichung $f(x) = y_0$ auf die *Fixpunktgleichung*

$$F(x) = x \text{ mit } F(x) := x - f(x) + y_0$$

um.

Im folgenden werden wir nun r und R so zu bestimmen versuchen, daß folgende Bedingungen erfüllt sind:

1) $\bar{U}_r(0) = \{x \in \mathbb{R}^n;\ \|x\| < r\} \subset D.$

2) Für jedes $y_0 \in U_R(0)$ gilt:

$$\|x\| \le r \Longrightarrow \|F(x)\| \le r.$$

3) Für $x, x' \in \bar{U}_r(0)$ gilt

$$\|F(x) - F(x')\| \le \frac{1}{2}\|x - x'\|.$$

Wenn r und R derart bestimmt sind, dann ist alles bewiesen, denn wegen 1) und 2) kann man F als Selbstabbildung von $\bar{U}_r(0)$ auffassen, die wegen 3) außerdem kontrahierend ist. Da aber $\bar{U}_r(0)$ als abgeschlossener Teil des $\mathbb{R}^n$ vollständig ist, kann man den Banachschen Fixpunktsatz Satz V.7.6 anwenden.

Konstruktion von r. Wir vergessen zunächst die Bedingung 2); diese wird dann bei der Wahl von R erfüllt. Von vornherein kann angenommen werden, daß 1) erfüllt sei (eventuelle Verkleinerung von r). Die Bedingung 3) besagt

$$\|x - f(x) - x' + f(x')\| \le \frac{1}{2}\|x - x'\| \text{ für} \|x\|, \|x'\| \le r.$$

(Diese Bedingung ist offenbar unabhängig von y_0.)

Im Fall $n = 1$ könnte man jetzt leicht mit Hilfe des Mittelwertsatzes weiterschließen. Wendet man ihn auf die Funktion

$$h(x) = x - f(x)$$

an, so würde folgen

$$|x - f(x) - x' + f(x')| = |h'(\xi)||x - x'|$$

mit einer Zwischenstelle ξ zwischen x und x'. Die Ungleichung, die wir zu realisieren haben, lautet dann

$$|h'(\xi)| \le \frac{1}{2}.$$

Nach Voraussetzung war $f'(0) = 1$, also $h'(0) = 0$. In einer genügend kleinen Umgebung von 0 ist die fragliche Ungleichung also sicher erfüllt.

Es ist naheliegend, im Fall $n > 1$ ein brauchbares Analogon des Mittelwertsatzes zu suchen.

2.2 Hilfssatz (Verallgemeinerter Mittelwertsatz der Differentialrechnung). VerMit *Es sei*

$$h : D \longrightarrow \mathbb{R}, \quad D \subset \mathbb{R}^n \text{ offen,}$$

eine stetig differenzierbare Funktion. Gegeben seien zwei Punkte $a, b \in D$, so daß auch noch ihre Verbindungsstrecke

$$a + t(b - a),\ 0 \le t \le 1,$$

ganz in D enthalten ist. Dann existiert auf dieser Verbindungsstrecke ein Punkt ξ, so daß gilt

$$h(a) - h(b) = \sum_{\nu=1}^{n} \partial_\nu h(\xi)(a_\nu - b_\nu).$$

Speziell gilt dann

$$|h(a) - h(b)| \le C \cdot \|a - b\| \text{ mit } C \ge \sum_{\nu=1}^{n} |\partial_\nu h(\xi)|.$$

Beweis. Man wendet auf die Funktion

$$h_0(t) = h(a + t(b - a))$$

den gewöhnlichen Mittelwertsatz an und erhält

$$\frac{h_0(1) - h_0(0)}{1 - 0} = h_0'(t_0);\ t_0 \in (0, 1).$$

Die *Kettenregel* liefert nun

$$h_0'(t_0) = \sum_{\nu=1}^{n} \partial_\nu h(\xi)(a_\nu - b_\nu) \text{ mit } \xi = a + t_0(a - b). \qquad \square$$

Mit Hilfe dieses verallgemeinerten Mittelwertsatzes können wir nun die Ungleichung

$$\|x - f(x) - x' + f(x')\| \le \frac{1}{2}\|x - x'\|$$

dadurch erzwingen, daß wir fordern:

$$\sum_{i,k} \left| \frac{\partial h_i}{\partial x_k}(\xi) \right| \le \frac{1}{2} \text{ für alle } \xi \text{ mit } |\xi| \le r.$$

Dabei sei wiederum

$$h(x) := x - f(x).$$

Nach Voraussetzung ist

$$J(h;0) = E - E = 0,$$

d.h. alle Ableitungen

$$\frac{\partial h_i}{\partial x_k}(0)$$

verschwinden. Wegen der Stetigkeit der partiellen Ableitungen ist die Summe ihrer Beträge in der Nähe des Nullpunkts kleiner oder gleich 1/2.

Konstruktion von R.

Es ist die Bedingung 2) zu erfüllen, d.h.

$$\|x\| \le r \text{ und } \|y_0\| \le R \Longrightarrow \|x - f(x) + y_0\| \le r.$$

Dabei muß auch r eventuell noch verkleinert werden. Nach Voraussetzung gilt

$$f(0) = 0 \text{ und } J(f;0) = E,$$

also

$$f(x) = x + r(x) \text{ mit } \frac{\|r(x)\|}{\|x\|} \longrightarrow 0 \text{ für } x \longrightarrow 0.$$

Es ist also

$$\|x - f(x) + y_0\| = \|r(x) + y_0\| \le \|r(x)\| + \|y\|_0.$$

Wir denken uns r bereits so klein gewählt, daß

$$\frac{\|r(x)\|}{\|x\|} \le \frac{1}{2} \text{ für} \|x\| \le r$$

gilt. Dann erfüllt $R = \frac{r}{2}$ das Gewünschte, denn man hat dann

$$\|x - f(x) + y_0\| \le \|r(x)\| + \|y_0\| \le \frac{r}{2} + \frac{r}{2} = r \text{ für } \|x\| \le r,\ \|y_0\| < R.$$

Als nächstes müssen wir uns von der Voraussetzung

$$f(0) = 0 \text{ und } J(f;0) = E$$

freimachen. Dies geschieht durch einen kleinen Kunstgriff.

Die Voraussetzungen von Theorem 2.1 seien gegeben. Wir betrachten nun die Abbildung

$$g(x) := J(f;a)^{-1}[f(x+a) - f(a)]$$

auf dem Definitionsbereich

$$D_0 := \{x \in \mathbb{R}^n;\ x + a \in D\}.$$

Diese hat offenbar die Eigenschaften

a) $g(0) = 0$,

b) $J(g;0) = E$.

Wir wissen daher, daß es offene Umgebungen U und V von 0 gibt, so daß die Gleichung

$$g(x) = y_0$$

für jedes $y_0 \in V$ eindeutig durch ein $x \in U$ lösbar ist.

Hieraus folgt durch Umrechnen auf f unmittelbar, daß die Gleichung

$$f(x) = y_0$$

für jedes $y_0 \in V_0$ eindeutig durch ein $x \in U_0$ lösbar ist. Dabei seien

$$\begin{aligned} U_0 &:= \{x \in \mathbb{R}^n;\ x - a \in U\}, \\ V_0 &:= \{x \in \mathbb{R}^n;\ J(f;a)^{-1}(x - f(a)) \in V\}. \end{aligned}$$

Beide Mengen sind offen.

Wir wissen jetzt also, daß die Abbildung f lokal umkehrbar ist, d.h. zu jedem $a \in D$ existiert eine offene Umgebung D_0 und eine lokale Umkehrfunktion

$$g : f(D_0) \longrightarrow \mathbb{R}^n$$

mit

$$f(x) = y,\ x = g(y) \qquad (x \in D_0,\ y \in f(D_0)).$$

2.3 Theorem (Satz für umkehrbare Funktionen, zweiter Teil). `SaUmz`
Es mögen die Voraussetzungen von Theorem 2.1 gelten; es sei also

$$f : D \longrightarrow \mathbb{R}^n,\ D \subset \mathbb{R}^n \text{ offen},$$

eine stetig differenzierbare Abbildung mit invertierbarer Jacobi-Matrix $J(f;a)$ für jedes $a \in D$. Zusätzlich sei f umkehrbar, d.h. es möge eine Abbildung

$$g : f(D) \longrightarrow \mathbb{R}^n \text{ mit } f(x) = y \iff x = g(y)$$

existieren. Dann ist auch g stetig partiell ableitbar und es gilt

$$J(g;b) = J(f;a)^{-1} \text{ für } b = f(a) \quad (\iff a = g(b)).$$

Beweis. Es gilt

$$f(x) = f(a) + J(f;a)(x-a) + r(x), \ \frac{\|r(x)\|}{\|x-a\|} \longrightarrow 0 \text{ für } x \longrightarrow a.$$

Wir machen den Ansatz

$$g(y) = g(b) + J(f;g(b))^{-1}(y-b) + s(y).$$

Abgesehen von der Stetigkeit der partiellen Ableitungen besagt unsere Behauptung nichts anderes, als

$$\frac{\|s(y)\|}{\|y-b\|} \longrightarrow 0 \text{ für } y \longrightarrow b.$$

Wir werden nun $s(y)$ durch $r(x)$ ausdrücken. Dazu setzen wir in der entsprechenden Gleichung

$$x = g(y) \text{ und } a = g(b)$$

ein und erhalten

$$y = b + J(f;g(b))(g(y) - g(b)) + r(g(y)).$$

Hieraus folgt

$$g(y) = g(b) + J(f;g(b))^{-1}(y-b) - J(f,g(b))^{-1}r(g(y)).$$

Durch Vergleich ergibt sich

$$s(y) = -\mathcal{J}^{-1}(f,g(b))^{-1}r(g(y)).$$

Wir wollen

$$\frac{\|s(y)\|}{\|y-b\|} \longrightarrow 0 \text{ für } y \longrightarrow b$$

zeigen. Dazu genügt es nach Hilfssatz 2.2 zu zeigen:

$$\frac{\|r(g(y))\|}{\|y-b\|} \longrightarrow 0 \text{ für } y \longrightarrow b.$$

Eine einfache Umformung ergibt

$$\frac{\|r(g(y))\|}{\|y-b\|} = \frac{\|r(g(y))\|}{\|g(y)-g(b)\|} \cdot \frac{\|g(y)-g(b)\|}{\|y-b\|}$$

(beachte: $y \neq b, \ g(y) \neq g(b)$).

Es bleibt jetzt noch zweierlei zu zeigen:

a) $y \to b \Longrightarrow g(y) \to g(b)$ (d.h. g ist stetig).

Hieraus folgt dann

$$\frac{\|r(g(y))\|}{\|g(y)-a\|} \longrightarrow 0 \text{ für } y \longrightarrow b \qquad (a = g(b)).$$

b) $\dfrac{\|g(y)-g(b)\|}{\|y-b\|} \le C$ in einer Umgebung von b.

Zu a). Eine Abbildung ist genau dann stetig, wenn das Urbild einer offenen Menge offen ist. Sei also $U \subset D$ offen. Es ist zu zeigen, daß $g^{-1}(U) \subset f(D)$ offen ist. Offenbar gilt

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

Nach Theorem 2.1 ist aber $f(U)$ tatsächlich offen.

Zu b). Schreiben wir $x = g(y)$, so besagt dies nichts anderes als

$$\frac{\|x-a\|}{\|f(x)-f(a)\|} \le K$$

in einer geeigneten Umgebung von a mit einer Konstanten K. Das holt man aus der Gleichung

$$f(x) = f(a) + J(f;a)(x-a) + r(x)$$

heraus und zwar durch die Umformung

$$x - a = J(f;a)^{-1}(f(x)-f(a)) - J(f;a)^{-1}r(x).$$

Anwendung der Dreiecksungleichung liefert

$$\|x-a\| \le C\|f(x)-f(a)\| + C\|r(x)\|$$

oder

$$\frac{\|f(x)-f(a)\|}{\|x-a\|} \ge \frac{1}{C} - \frac{\|r(x)\|}{\|x-a\|} \ge \frac{1}{2C}$$

in einer genügend kleinen Umgebung von a. Damit ergibt sich

$$\frac{\|x-a\|}{\|f(x)-f(a)\|} \le 2C.$$

Wir müssen nun abschließend noch zeigen, daß die partiellen Ableitungen von g stetig sind. Dies folgt aber einfach aus der Formel

$$J(g;y) = J(f;(g(y)))^{-1},$$

wenn man berücksichtigt, daß die Koeffizienten der inversen Matrix A sich aus denen von A durch rationale Operationen berechnen lassen (Cramersche Regel für das Invertieren einer Matrix.) □

2.4 Theorem (Satz für implizite Funktionen). *Sei* DafIm

$$f : D \longrightarrow \mathbb{R}^m,\ D \subset \mathbb{R}^{n+m} \text{offen},$$

eine stetig differenzierbare Abbildung. Außerdem sei ein Punkt

$$(a, b), \quad a \in \mathbb{R}^n,\ b \in \mathbb{R}^m$$

mit folgenden Eigenschaften gegeben:

a) $f(a, b) = 0$.

b) *Die Funktionalmatrix $J(f_a; b)$ der Abbildung*

$$f_a : D_a \longrightarrow \mathbb{R}^m,\ D_a := \{y \in \mathbb{R}^m; \quad (a, y) \in D\};\ f_a(y) := f(a, y)$$

ist umkehrbar.

Dann gibt es offene Umgebungen

$$a \in A \subset \mathbb{R}^n,\ b \in B \subset \mathbb{R}^m,$$

so daß für jedes $x \in A$ eine und nur eine Lösung $y \in B$ der Gleichung

$$f(x, y) = 0$$

existiert.

Zusatz. *Die Abbildung*

$$h : A \longrightarrow B,\ y = h(x),$$

ist stetig differenzierbar und es gilt

$$J(h; a) = -J(f_a; b)^{-1} J(f_b; a).$$

Dabei ist f_b analog zu f_a definiert ($f_b(x) := f(x, b)$).

Der Satz für implizite Funktionen ist allgemeiner als der Satz für umkehrbare Funktionen. Durch einen Kunstgriff kann man ihn aber auf den bereits behandelten Fall zurückführen. Da wir im folgenden davon keinen Gebrauch machen werden, geben wir lediglich die *Beweisidee* an.

Die Abbildung

$$F : D \longrightarrow \mathbb{R}^{n+m}, \quad F(x, y) := (x, f(x, y),$$

hat die Jacobi-Matrix

$$\left(\begin{array}{ccc|ccc} 1 & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & 1 & & & \\ \hline & & & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ & * & & \vdots & \ddots & \vdots \\ & & & \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{array}\right)$$

Diese ist in (a, b) nicht ausgeartet, da die Matrix

$$J(f_a; b) = \left(\frac{\partial f_\nu}{\partial y_\mu}(a, b)\right)_{\nu\mu}$$

voraussetzungsgemäß nicht ausgeartet ist. Die Voraussetzung von Theorem 2.1 wird somit von F erfüllt. Nun ist es nicht mehr sehr schwer, Theorem 2.4 mit Hilfe von Theorem 2.1 und Theorem 2.3 zu beweisen.

Beispiele. Wir betrachten die Abbildung

$$x = r\cos\varphi, \;\; y = r\sin\varphi.$$

Die Jacobi-Matrix dieser Abbildung ist

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Sie ist für $r > 0$ invertierbar, denn ihre *Determinante* ist

$$r\cos^2\varphi + r\sin^2\varphi = r$$

und somit im genannten Fall von 0 verschieden. Aus dem Satz für umkehrbare Funktionen folgt, daß die Abbildung

$$h(r, \varphi) := (r\cos\varphi, r\sin\varphi)$$

in einer geeigneten Umgebung eines beliebigen Punktes $r_0, \varphi_0)$ mit $r_0 > 0$ umkehrbar ist.

Polarkoordinaten im Raum. Sei

$$D := \{(r, \varphi, \theta) \in \mathbb{R}^3;\; r > 0\}.$$

Wir betrachten die Abbildung

$$h : D \longrightarrow \mathbb{R}^3, \quad h(r, \varphi, \theta) = (r\cos\varphi\cos\theta, r\sin\varphi\cos\theta, r\sin\theta).$$

Ihre Jacobi-Matrix ist

$$\begin{pmatrix} \cos\varphi\cos\theta & -r\sin\varphi\cos\theta & -r\cos\varphi\sin\theta \\ \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\theta & 0 & r\cos\varphi \end{pmatrix}$$

Diese Matrix hat die Determinante $r^2\cos\theta$; sie ist daher dann von Null verschieden, wenn

$$r > 0 \text{ und } \cos\theta \neq 0.$$

Jeder Punkt (r_0, θ_0) mit dieser Eigenschaft besitzt also eine Umgebung, in der h eine differenzierbare Umkehrabbildung hat. Es ist nicht schwer zu zeigen, daß die Abbildung h eine *bijektive* Abbildung von

$$D' := \left\{(r, \varphi, \theta); \quad r > 0;\ \varphi \in (0, 2\pi);\ \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)\right\}$$

auf

$$D'' := \mathbb{R}^3 - \{(x, y, z); \quad x \geq 0,\ y = 0\}$$

vermittelt. Die Umkehrabbildung $D'' \to D'$ ist nach dem eben Gesagten stetig differenzierbar.

Geometrisch ist θ der Winkel zwischen dem Vektor (x, y, z) und dem Vektor $(x, y, 0)$. φ ist der Winkel zwischen $(x, y, 0)$ und $(x, 0, 0)$ und r der Abstand von 0 und (x, y, z).

3. Extremwerte differenzierbarer Funktionen

Sei

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \text{offen},$$

eine differenzierbare Funktion, deren partielle Ableitungen verschwinden

$$\partial_1 f = \ldots = \partial_n f = 0.$$

Seien ferner a und b zwei Punkte in D.

Annahme. Die Verbindungsstrecke zwischen a und b möge ganz in D enthalten sein:

$$a + t(a - b) \in D \text{ für } 0 \leq t \leq 1.$$

Dann ist die Funktion

$$g : [0, 1] \longrightarrow \mathbb{R};\ g(t) := f(a + t(a - b))$$

definiert. Aus der Kettenregel folgt

$$g'(t) = 0.$$

Wie aus der Analysis einer Veränderlichen bekannt, folgt hieraus, daß g konstant ist, also insbesondere

$$g(a) = g(b).$$

Die offene Menge D heißt *konvex*, wenn mit je zwei Punkten $a, b \in D$ immer die Verbindungsstrecke ganz in D enthalten ist. Somit ergibt sich:

Verschwinden die partiellen Ableitungen einer Funktion $f : D \to \mathbb{R}$ und ist D konvex, so ist f konstant.

Zum Beispiel sind die Euklidischen Kugeln

$$U_r(a) = \{x \in \mathbb{R}^n;\ d(a,x) < r\}$$

konvex (hierbei sei d die Euklidische oder Maximums-Metrik).

Eine Funktion

$$f : X \longrightarrow \mathbb{R}$$

auf einem beliebigen metrischen Raum heißt *lokal konstant*, wenn es zu jedem Punkt $a \in X$ eine Umgebung $a \in U \subset X$ gibt, so daß die Einschränkung von f auf U konstant ist.

3.1 Bemerkung. *Sei* AbnfK

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \textit{ offen},$$

eine differenzierbare Funktion. Wenn die partiellen Ableitungen von f verschwinden, so ist f lokal konstant (und natürlich umgekehrt).

3.2 Definition. *Ein metrischer Raum X heißt* ***zusammenhängend****, wenn jede lokal konstante Funktion $f : X \to \mathbb{R}$ sogar konstant ist.* defZus

Beispiel. Jedes Intervall $D \subset \mathbb{R}$ ist zusammenhängend.

Beweis. Sei $f : D \to \mathbb{R}$ eine lokal konstante Funktion. Dann gilt offensichtlich $f' = 0$. Wie aus der Differentialrechnung einer Veränderlichen bekannt, folgt hieraus, daß f konstant ist.

3.3 Definition. *Ein metrischer Raum X heißt* ***bogenweise zusammenhängend****, wenn es zu je zwei Punkten $a, b \in X$ eine stetige Abbildung φ gibt, mit* bogZus

$$\varphi : [0,1] \longrightarrow X;\ \varphi(0) = a,\ \varphi(1) = b.$$

Man stelle sich $[0,1]$ als Zeitintervall vor. Dann ist $\varphi(t)$ ein Punkt, welcher sich stetig von a nach b bewegt. Das Bild von φ stelle man sich als eine (krumme) Linie vor, die a mit b verbindet. Man nennt φ auch eine Kurve, manchmal auch Bogen oder auch Weg.

3.4 Bemerkung. *Jeder bogenweise zusammenhängende Raum ist zusammenhängend.* bogZZ

Beweis. Sei $f : X \to \mathbb{R}$ eine lokal konstante Funktion auf dem metrischen Raum X. Wir müssen zeigen, daß f konstant ist, daß also $f(a) = f(b)$ für zwei vorgegebene Punkte $a, b \in X$ gilt. Nach Voraussetzung existiert eine stetige Abbildung

$$\varphi : [0,1] \longrightarrow X; \ \varphi(0) = a, \ \varphi(1) = b.$$

Man zeigt leicht mit Hilfe der Stetigkeit von φ, daß auch die Abbildung

$$g = f \circ \varphi : [0,1] \longrightarrow \mathbb{R}$$

lokal konstant ist. Da das Intervall $[0,1]$ zusammenhängend ist, muß g konstant sein. Es folgt

$$f(a) = g(0) = g(1) = f(b). \qquad \square$$

3.5 Definition. *Ein **Gebiet** im $\mathbb{R}^n$ ist eine offene, zusammenhängende Teilmenge $D \subset \mathbb{R}^n$.* defGB

Halten wir noch einmal fest:

3.6 Bemerkung. *Sei $D \subset \mathbb{R}^n$ ein Gebiet und $f : D \to \mathbb{R}$ eine differenzierbare Funktion, deren partielle Ableitungen verschwinden. Dann ist f konstant.* GebKon

Übungsaufgaben.

1) Die einzigen zusammenhängenden Teile von $\mathbb{R}$ sind die Intervalle.
2) Ein metrischer Raum X ist genau dann zusammenhängend, wenn folgendes erfüllt ist:

   Ist $X = U \cup V$; U, V offen, $U \cap V = \emptyset$, so gilt $U = X$ oder $V = X$ (und $V = \emptyset$ oder $U = \emptyset$).

Anleitung. a) Ist $f : X \to \mathbb{R}$ lokal konstant, so wähle man einen Punkt $a \in X$ und setze

$$U := \{x \in X;\ f(x) = f(a)\}, \quad V := \{x \in X;\ f(x) \neq f(a)\}.$$

b) Ist $X = U \cup V$, so betrachte man die Funktion

$$f(x) = \begin{cases} 1 & \text{für } x \in U \\ 0 & \text{für } x \in V \end{cases}$$

3) Jedes Gebiet $D \subset \mathbb{R}^n$ ist bogenweise zusammenhängend.

Anleitung. Man wähle $a \in D$ und setze

$$U := \big\{\, x \in D; \quad x \text{ ist mit } a \text{ durch eine Kurve innerhalb } D \text{ verbindbar} \,\big\}$$

und

$$V := D - U.$$

4) Man konstruiere einen metrischen Raum, welcher zwar zusammenhängend, nicht jedoch bogenweise zusammenhängend ist.

Lokale Extrema

Seien X ein metrischer Raum und $f : X \to \mathbb{R}$ eine Funktion auf X. Ein Punkt $a \in X$ heißt *lokales Maximum (Minimum)* von f, wenn es eine Umgebung $U,\ a \in U \subset X$, gibt, so daß

$$f(x) \le f(a) \qquad \bigl(f(x) \ge f(a)\bigr) \text{ für alle } x \in U$$

gilt. Der Punkt a heißt *lokales Extremum*, wenn f lokales Maximum *oder* Minimum ist.

3.7 Bemerkung. *Sei* lokEX

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \textit{offen},$$

eine differenzierbare Funktion und sei $a \in D$ ein lokales Extremum von f. Dann gilt

$$\partial_\nu f(a) = 0 \textit{ für } \nu = 1, \dots, n.$$

Der Beweis reduziert sich unmittelbar auf den bereits bekannten Fall $n = 1$. Wie auch dort ist jedoch das Verschwinden der partiellen Ableitungen keine hinreichende Bedingung für das Vorliegen eines lokalen Extremums. Solche erhält man, wenn man auch höhere Ableitungen betrachtet. Beispielsweise ergibt sich im Fall $n = 1$:

Sei

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R} \textit{ ein Intervall},$$

eine zweimal stetig differenzierbare Funktion und sei a ein Punkt mit den Eigenschaften

a) $f'(a) = 0$,
b) $f''(a) < 0$.

Dann ist a ein lokales Maximum von f.

Der Beweis ergibt sich leicht aus dem Mittelwertsatz

$$\frac{f(x) - f(a)}{x - a} = f'(\xi),\ \xi \in (a, x).$$

Nach Voraussetzung ist $f''(a)$ negativ. Aus Stetigkeitsgründen ist f'' dann in einer vollen Umgebung von a negativ. Wir können annehmen, daß $f''(x)$ für alle $x \in D$ negativ ist. Die Funktion f' ist dann streng monoton fallend, also

$$f'(x) < f'(a) = 0 \text{ für } x > a \quad \text{sowie} \quad f'(x) > f'(a) = 0 \text{ für } x < a,$$

also beispielsweise

$$\frac{f(x) - f(a)}{x - a} < 0 \text{ für } x > a.$$

Hieraus folgt

$$f(x) < f(a) \text{ für alle } x.$$

Der Beweis zeigt tatsächlich etwas mehr als wir formuliert haben, nämlich:

Sei $f : D \to \mathbb{R}$ eine zweimal stetig differenzierbar Funktion auf einem Intervall D und $a \in D$ ein Punkt mit $f'(a) = 0$. Sei D_0 ein Intervall, $a \in D_0 \subset D$, auf welchem die zweite Ableitung von f negativ ist, d.h.

$$f''(x) < 0 \textit{ für alle } x \in D_0.$$

Dann besitzt $f|D_0$ sein (absolutes) Maximum in a. Es ist also

$$f(x) < f(a) \textit{ für alle } x \in D_0.$$

Wir fassen die zweiten partiellen Ableitungen in einem Punkt $a \in D$ in einer Matrix zusammen:

$$\mathcal{H} := \mathcal{H}(f, a) := \begin{pmatrix} \partial^2_{11} f(a) & \dots & \partial^2_{1n} f(a) \\ \vdots & & \vdots \\ \partial^2_{n1} f(a) & \dots & \partial^2_{nn} f(a) \end{pmatrix} \qquad \text{(H{\small ESSE}matrix).}$$

Diese Matrix ist symmetrisch, da es auf die Reihenfolge der Differentiationen nicht ankommt.

3.8 Definition. *Eine reelle symmetrische Matrix* PosDef

$$H = \begin{pmatrix} h_{11} & \dots & h_{1n} \\ \vdots & & \vdots \\ h_{n1} & \dots & h_{nn} \end{pmatrix}, \quad h_{ij} = h_{ji} \textit{ für} 1 \le i, j \le n,$$

*heißt **positiv definit**, wenn der Ausdruck*

$$\sum_{1 \le i,j \le n} h_{ij} x_i x_j$$

für jeden von 0 verschiedenen Vektor $x \in \mathbb{R}^n$ positiv ist.

Eine einreihige Matrix $H = (h)$ ist natürlich genau dann positiv (definit), wenn die Zahl h positiv ist.

Ergänzung zu Definition 3.8. Man nennt die Matrix H *negativ definit*, wenn $-H$ positiv definit ist und *definit*, wenn sie entweder positiv oder negativ definit ist.

3.9 Satz. *Die Funktion* besLM

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \text{ offen},$$

sei zweimal stetig differenzierbar. Sei $a \in D$ ein Punkt mit den Eigenschaften

a) $\partial_j f(a) = 0$ *für* $j = 1, \dots, n$,

b) $\mathcal{H}(f, a) = \big(\partial_{ij}^2 f(a)\big)_{ij}$ *ist negativ definit.*

Dann besitzt f in a ein lokales Maximum.

Beweis. Wir betrachten für eine beliebige „Richtung" $\alpha \in \mathbb{R}^n$ die Funktion

$$g_\alpha(t) := f(a + t\alpha).$$

Sie ist in einer offenen Umgebung von $t = 0$ definiert. Aus der Kettenregel ergibt sich $g'_\alpha(a) = 0$ und

$$g''_\alpha(t) = \sum_{1 \le i,j \le n} \partial_{ij}^2 (a + t\alpha)\alpha_i \alpha_j.$$

Die Menge

$$U := \big\{\, (\alpha, t) \in \mathbb{R}^n \times \mathbb{R}; \quad a + t\alpha \in D,\ g''_\alpha(t) < 0 \,\big\}$$

ist aus Stetigkeitsgründen eine offene Teilmenge des $\mathbb{R}^{n+1}$. Es gilt

$$U \supset \mathcal{S}^n \times \{0\} \text{ mit } \mathcal{S}^n := \{\alpha \in \mathbb{R}^n;\ \|\alpha\| = 1\}.$$

Übungsaufgabe. Aus der Kompaktheit von $\mathcal{S}^n$ folgert man die Existenz einer Zahl $\varepsilon > 0$ mit der Eigenschaft

$$\|\alpha\| = 1,\ |t| < \varepsilon \Longrightarrow a + t\alpha \in D \text{ und } g''_\alpha(t) < 0.$$

Wie aus der Theorie einer Variablen bekannt, hat die Funktion

$$(-\varepsilon, \varepsilon) \longrightarrow \mathbb{R}, \quad t \longmapsto g_\alpha(t) \qquad (\|\alpha\| = 1)$$

in $t = 0$ ein absolutes Maximum. Es gilt insbesondere

$$f(x) < f(a),$$

wenn sich x in der Form

$$x = a + t\alpha,\ |t| < \varepsilon,\ \|\alpha\| = 1$$

schreiben läßt. Die Menge dieser x ist eine Umgebung von a. □

Extremwerte unter Nebenbedingungen.

Wir nehmen einmal an, daß D eine nichtleere, *beschränkte* und offene Menge im $\mathbb{R}^n$ sei. Dann ist der Abschluß $\bar{D}$ von D kompakt. Wir nehmen außerdem an, daß eine stetige Funktion $f : \bar{D} \to \mathbb{R}$ gegeben sei, welche in D zweimal stetig differenzierbar sei. Die Funktion f muß dann in $\bar{D}$ ein absolutes Maximum haben. Dann gibt es zwei Möglichkeiten. Entweder das Maximum wird im Inneren – also in D – angenommen. In diesem Fall müssen die partiellen Ableitungen von f in a (dem Maximalpunkt) verschwinden. Oder es wird auf dem Rand angenommen. Der Rand ist im Fall $n > 1$ selbst eine unendliche Punktmenge, z.B.

$$D = \{x \in \mathbb{R}^n;\ \|x\| < 1\} \Longrightarrow \partial D = \mathcal{S}^{n-1}.$$

Wie bestimmt man nun das Maximum von $f|\partial D$ oder anders gesprochen: Wie bestimmt man das Maximum der Funktion f unter der Nebenbedingung $\|x\| = 1$?

Die Methode der *Lagrange-Multiplikatoren* liefert ein allgemeines Verfahren, Maxima und Minima unter Nebenbedingungen zu untersuchen.

3.10 Satz. *Gegeben seien* LaMu

a) *eine offene Teilmenge* $D \subset \mathbb{R}^{n+m}$,

b) *eine stetig differenzierbare Funktion* $f : D \to \mathbb{R}$,

c) *eine stetig differenzierbare Abbildung*

$$g : D \longrightarrow \mathbb{R}^m,\ g = (g_1, \dots, g_m).$$

Voraussetzung. *Die Jacobi-Matrix* $J(g; x)$ *habe für jeden Punkt* $x \in D$ *den (maximal möglichen) Rang* m. *Wir setzen*

$$M := \{x \in D;\ g(x) = 0\}. \qquad \textit{(Nebenbedingungsmenge)}$$

Annahme. $a \in M$ *sei ein relatives Extremum der Einschränkung* $f|M$ *von* f *auf* M.

Behauptung. *Es existieren reelle Zahlen* $\lambda_1, \dots, \lambda_m$ *(die* ***Lagrangeschen Multiplikatoren****), so daß gilt:*

$$\partial_i f(a) = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(a) \text{ für } i = 1, \dots, n+m.$$

Die lokalen Extrema von $f|M$ sind also unter den Lösungen des Gleichungssystems

$$\boxed{\begin{aligned} g_j(x) &= 0 \text{ für } j = 1, \ldots, m \\ \partial_i f(a) &= \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(a) \text{ für } i = 1, \ldots, n+m \end{aligned}}$$

zu suchen.

Die Anzahl dieser Gleichungen ist $2m+n$. Dies ist auch die Anzahl der Unbestimmten (unbestimmt sind die Multiplikatoren und der Punkt x). Dies läßt hoffen (mehr nicht!), daß in nicht allzu entarteten Fällen dieses Gleichungssystem nur endlich viele Lösungen (x, λ) besitzt. Unter diesen sind die lokalen Extrema durch weitere Überlegungen auszusondern.

Beweis von Satz 3.10. Der Beweis erfolgt in drei Schritten.

1. Schritt. Der Spezialfall

$$g_i(x) = x_{n+i} \text{ für } i = 1, \ldots, m,$$

also

$$M = \left\{ x \in D; \quad x_{n+1} = \ldots = x_{n+m} = 0 \right\}.$$

Ist $a \in M$ ein relatives Extremum von $f|M$, so gilt

$$\partial_i f(a) = 0 \text{ für } i = 1, \ldots, n.$$

Setzt man

$$\lambda_j = \partial_{n+j} f(a) \text{ für } j = 1, \ldots, m,$$

so sind die in Satz 3.10 geforderten Gleichungen ersichtlich erfüllt.

2. Schritt. Transformationsinvarianz der Aussage von Satz 3.10.

Sei

$$\varphi : D \longrightarrow \tilde{D}, \quad D, \tilde{D} \subset \mathbb{R}^{n+m} \text{ offen},$$

ein *Diffeomorphismus*, d.h. eine bijektive Abbildung, so daß φ und φ^{-1} stetig differenzierbar sind. Wir definieren

$$\tilde{f} : \tilde{D} \longrightarrow \mathbb{R} \quad \text{und} \quad \tilde{g} : \tilde{D} \longrightarrow \mathbb{R}^m$$

durch

$$\tilde{f} \circ \varphi := f \quad \text{und} \quad \tilde{g} \circ \varphi := g$$

und setzen außerdem noch

$$\tilde{a} := \varphi(a).$$

Behauptung. Wenn die Aussage des Satzes 3.10 für die Daten (D, f, g, a) richtig ist, so ist sie es auch für $(\tilde{D}, \tilde{f}, \tilde{g}, \tilde{a})$ und umgekehrt).

Zum Beweis dieser Behauptung benutze man für die „Multiplikatorengleichung" die Matrizenschreibweise

$$J(f; a) = \lambda \cdot J(g; a) \qquad \big(\lambda = (\lambda_1, \dots, \lambda_m)\big).$$

Die Behauptung folgt dann unmittelbar aus der Kettenregel

$$J(\tilde{f}; \tilde{a}) \cdot J(\varphi; a) = J(f; a)$$

(und analog für g). Man erhält sogar dasselbe Multiplikatorensystem λ.

3. Schritt. Aus den beiden ersten Schritten folgt, daß die Aussage des Satzes für (D, f, g, a) bewiesen ist, wenn es einen Diffeomorphismus $\varphi : D \to \tilde{D}$ gibt, so daß

$$\tilde{g}_i(x) = x_{n+i} \text{ für } i = 1, \dots, m$$

gilt. Da es sich um eine lokale Aussage handelt, ist es sogar ausreichend, wenn ein derartiger Diffeomorphismus auf einer (kleinen) Umgebung $a \in U \subset D$ existiert $(\varphi : U \to \tilde{U})$. Wir werden aus dem Satz für umkehrbare Funktionen folgern, daß ein derartiger Diffeomorphismus φ (auf geeignetem U) existiert. An dieser Stelle des Beweises wird benutzt, daß die Matrix $J(g; a)$ den maximalen Rang m hat. Wie aus der linearen Algebra bekannt, bedeutet dies, daß man $J(g; a)$ zu einer quadratischen $(m + n)$-reihigen invertierbaren Matrix ergänzen kann.

Hieraus wiederum folgt, daß man eine Abbildung

$$G : D \longrightarrow \mathbb{R}^{n+m}$$

finden kann, deren letzte m Komponenten mit g übereinstimmen und deren Funktionalmatrix im Punkt a invertierbar ist. Nach dem Satz über umkehrbare Funktionen definiert G einen Diffeomorphismus

$$\varphi : U \longrightarrow \tilde{U},\ a \in U \subset D,\ \varphi(x) := G(x) \text{ für } x \in U,$$

auf einer offenen Umgebung U von a. Dieser Diffeomorphismus hat die gewünschte Eigenschaft. $\square$

Anwendungsbeispiel für Lagrange-Multiplikatoren.

Wir wollen die Extremwerte der Funktion

$$f(x) := (x_1 \cdot \ldots \cdot x_n)^2$$

auf der „Sphäre"

$$\mathcal{S} := \left\{ x; \quad x_1^2 + \ldots + x_n^2 = 1 \right\}$$

bestimmen. Die Funktion ist auf dem ganzen $\mathbb{R}^n$ definiert und stetig differenzierbar; ebenso die „Nebenbedingung“

$$g(x) = x_1^2 + \ldots + x_n^2 - 1.$$

Da $\mathcal{S}$ kompakt ist, nimmt f sein Maximum (und Minimum) in $\mathcal{S}$ an und diese sind unter den Lösungen der Gleichungen

$$\partial_i f = \lambda \cdot \partial_i g \ \text{ für } i = 1, \ldots, n$$

zu suchen, also

$$\frac{2(x_1 \cdot \ldots \cdot x_n)^2}{x_i} = \lambda \cdot 2x_i \ \text{ für } i = 1, \ldots, n$$

oder

$$(x_1 \cdot \ldots \cdot x_n)^2 = \lambda \cdot x_i^2 \ \text{ für } i = 1, \ldots, n.$$

Da wir das Maximum von f bestimmen wollen und da dieses offensichtlich von 0 verschieden ist, können wir

$$x_i \neq 0 \ \text{ für } i = 1, \ldots, n$$

annehmen. Es folgt dann $\lambda \neq 0$ und

$$x_1^2 = \ldots = x_n^2 \Longrightarrow x_i^2 = \frac{1}{n}.$$

Da die Funktion $f(x)$ sich nicht ändert, wenn man x_i durch $-x_i$ ersetzt, erhalten wir aus den bisherigen Überlegungen:

Das Maximum von f auf der Sphäre $\mathcal{S}$ wird in dem Punkt

$$\left(\sqrt{\frac{1}{n}}, \ldots, \sqrt{\frac{1}{n}}\right)$$

angenommen. Der Wert des Maximums ist

$$f\left(\sqrt{\frac{1}{n}}, \ldots, \sqrt{\frac{1}{n}}\right) = n^{-n}.$$

Wir erhalten also

$$\boxed{(x_1 \cdot \ldots \cdot x_n)^2 \leq n^{-n} \ \text{ für } x_1^2 + \ldots + x_n^2 = 1}$$

Übungsaufgabe. Man folgere aus der eben bewiesenen Ungleichung die bekannte *Ungleichung zwischen geometrischem und arithmetischem Mittel*

$$\boxed{\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leq \frac{x_1 + \ldots + x_n}{n} \quad \text{für } x_i \geq 0, \ i = 1, \ldots, n.}$$

4. Die Taylorsche Formel und analytische Funktionen

Die Taylorreihe für Funktionen mehrerer Veränderlicher erhält man durch Reduktion auf den Fall $n = 1$ (Kapitel III, §6) mit Hilfe des bereits mehrfach verwendeten Spezialisierungsprozesses

$$g(t) = f(a + t\alpha),$$

also durch Einschränken von f auf Geraden.

Zur bequemen Formulierung der Taylorschen Formel ist es zweckmäßig, den Kalkül der *Multiindizes* zu verwenden. Ein solcher Multiindex ist ein n-Tupel nicht negativer ganzer Zahlen

$$\nu = (\nu_1, \dots, \nu_n),\ \nu_j \in \mathbb{N}_0 \text{ für } 1 \le j \le n.$$

Bezeichnungen.

1) $|\nu| := \nu_1 + \nu_2 + \cdots + \nu_n$.
2) $\nu! := \nu_1! \cdots \nu_n!$.
3) $x^\nu := x_1^{\nu_1} \cdots x_n^{\nu_n}$ für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.
4) $x \cdot y := x_1 y_1 + \cdots + x_n y_n$ für $x, y \in \mathbb{R}^n$ (Skalarprodukt).
5) Ist f eine auf einem offenen Teil von $\mathbb{R}^n$ genügend oft stetig differenzierbare Funktion, so setzen wir

$$\partial_x f := \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) \quad \text{sowie} \quad \partial^\nu f := \frac{\partial^{|\nu|} f}{(\partial x_1)^{\nu_1} \dots (\partial x_n)^{\nu_n}}.$$

Beispiel.

$$\partial^{(1,2)} f = \frac{\partial^3 f}{\partial x_1 (\partial x_2)^2}.$$

Wir betrachten nun eine Funktion

$$f : D \to \mathbb{R},\ D \subset \mathbb{R}^n \text{offen},$$

welche k-fach stetig differenzierbar sei. Gegeben sei ein Punkt $a \in D$, welcher als Entwicklungspunkt dienen soll. Für einen beliebigen Richtungsvektor $\alpha \in \mathbb{R}^n$ definieren wir

$$g(t) := f(a + t\alpha).$$

Diese Funktion ist in einer offenen Umgebung von $t = 0$ k-fach stetig differenzierbar. Die k-te Ableitung von g kann durch wiederholte Anwendung der

Kettenregel gewonnen werden. Es gilt etwa ($x := a + t\alpha$):

$$\begin{aligned}
g'(t) &= \sum_{i=1}^{n} \partial_i f(x) \cdot \alpha_i, \\
g''(t) &= \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} \partial_{ji}^2 f(x) \alpha_j \Big) \alpha_i = \sum_{1 \le i,j \le n} \partial_{ij}^2 f(x) \alpha_i \alpha_j \\
&= \sum_{i=1}^{n} \partial^2 f_{ii}(x) \alpha_i^2 + 2 \sum_{1 \le i < j \le n} \partial_{ij}^2 f(x) \alpha_i \alpha_j .
\end{aligned}$$

Diesen Ausdruck kann man auch in der Form

$$g^{(2)}(t) = \sum_{|\nu|=2} \frac{2!}{\nu!} (\partial^\nu f)(x) \cdot \alpha^\nu$$

schreiben. Dabei beachte man, daß

$$|\nu| = 2 \Longrightarrow \nu! = \begin{cases} 1 & \text{für } \nu = (1,1) \\ 2 & \text{sonst.} \end{cases}$$

Durch Induktion nach k beweist man allgemein

4.1 Hilfssatz. *Unter obigen Voraussetzungen gilt* VorTay

$$g^{(k)}(t) = \sum_{|\nu|=k} \frac{k!}{\nu!} (\partial^\nu f)(x) \cdot \alpha^\nu .$$

Hierbei ist über die endlich vielen Multiindizes $\nu \in \mathbb{N}_0^n$ mit $|\nu| = k$ zu summieren.

Wir setzen nun voraus, daß f zumindest $N + 1$-mal stetig differenzierbar ist. Die Taylorreihe für die Funktion g (in einer Variablen!) lautet

$$g(t) = \sum_{k=0}^{N} \frac{g^{(k)}(0) t^k}{k!} + \frac{1}{N!} \int_0^t (t-u)^N g^{(N+1)}(u) \, du.$$

Setzen wir wieder $x := a + t\alpha$ und beachten

$$(x - a)^\nu = t^k \cdot \alpha^\nu \text{ für } |\nu| = k,$$

so folgt mittels Hilfssatz 4.1

4.2 Satz (Taylorformel). *Die Funktion* TaylF

$$f : D \to \mathbb{R},\ D \subset \mathbb{R}^n \text{ offen und konvex,}$$

sei $n+1$-mal stetig differenzierbar und es sei $a \in D$. Dann gilt für alle $x \in D$:

$$f(x) = \sum_{|\nu| \le n} \frac{(\partial^\nu f)(a)}{\nu!}(x-a)^\nu + R_n(x)$$

$$\textit{mit} \quad R_n(x) = (n+1) \sum_{|\nu|=n+1} \frac{1}{\nu!} \int_0^t (t-u)^n [(\partial^\nu f)(a+u\alpha)](a+u\alpha)^\nu \, du.$$

Wir schreiben einige Terme der Taylorschen Formel explizit nieder:

$$\begin{aligned}
&\sum_{|\nu| \le 2} \frac{(\partial^\nu f)(a)}{\nu!}(x-a)^\nu = \\
&\qquad f(a) + \quad \text{(konstanter Term)} \\
&\qquad \sum_{i=1}^n \partial_i f(a)(x_i - a_i) + \quad \text{(linearer Term)} \\
&\qquad \frac{1}{2} \sum_{1 \le i,j \le n} \partial_{ij}^2 f(a)(x_i - a_i)(x_j - a_j) \quad \text{(quadratischer Term)}
\end{aligned}$$

Taylorreihen

Wir nehmen jetzt an, daß die Funktion

$$f : D \to \mathbb{R},\ D \subset \mathbb{R}^n \text{ offen,}$$

in $a \in D$ beliebig oft stetig partiell ableitbar ist. Die unendliche Reihe

$$\sum_{\nu \in \mathbb{N}_0^n} \frac{\partial^\nu f}{\nu!}(a)(x-a)^\nu$$

heißt die *Taylorreihe* von f zum Entwicklungspunkt a. Natürlich ist wie schon im Fall $n = 1$ nicht klar, für welche x diese Reihe konvergiert und ob sie, wenn sie es tut, die Funktion f darstellt.

4.3 Definition. *Eine Funktion* defAnl

$$f : D \longrightarrow \mathbb{R},\ D \subset \mathbb{R}^n \text{ offen},$$

heißt ***analytisch****, wenn sie unendlich oft stetig partiell ableitbar ist und wenn zu jedem Punkt $a \in D$ eine Umgebung U existiert, innerhalb derer die Taylorreihe absolut konvergiert und die Funktion f darstellt:*

$$f(x) = \sum_{\nu=(\nu_1,\dots,\nu_n)} a_\nu (x-a)^\nu, \quad a_\nu = \frac{(\partial^\nu f)(a)}{\nu!}.$$

Allgemeiner heißt eine Abbildung $f : D \to \mathbb{R}^m$ analytisch, wenn alle Komponenten von f analytisch sind.

Absolute Konvergenz ist hierbei im Sinne des Summierbarkeitsbegriffes Definition I.7.2 zu verstehen (absolute Konvergenz bei irgendeiner Anordnung der Indizes).

Notwendig für die Analytizität von f ist natürlich

$$\lim_{N\to\infty} R_N(x) = 0$$

in einer Umgebung von $x = a$. Die formalen Rechenregeln für Potenzreihen und die Permanenzeigenschaften analytischer Funktionen gelten wie im Fall $n = 1$ (vgl. Kapitel IV, §6). Wir können uns daher kurz fassen und wollen auf Beweise verzichten.

Sprechweise. Eine Potenzreihe

$$P(x) := \sum_{\nu\in\mathbb{N}_0^n} a_\nu (x-a)^\nu,\ a_\nu \in \mathbb{R},$$

heißt *konvergent*, wenn es eine offene Umgebung U von a gibt, so daß die Reihe für alle $x \in U$ absolut konvergiert (also summierbar ist im Sinne von Definition I.7.2).

Die Reihe $P(x)$ definiert dann in dieser offenen Menge eine analytische Funktion. Die Taylorreihe von P und die vorgegebene Potenzreihe stimmen überein, d.h.

$$a_\nu = \frac{\partial^\nu P(a)}{\nu!}.$$

Beispiel. Die geometrische Reihe

$$\frac{1}{(1-x_1)\cdot\ldots\cdot(1-x_n)} = \sum_{\nu\in\mathbb{N}_0^n} x^\nu$$

konvergiert absolut für

$$\|x\| < 1; \ \|x\| = \max_{1 \le i \le n} |x_i|,$$

wie sich unmittelbar aus dem Fall $n = 1$ in Verbindung mit dem CAUCHYschen Multiplikationssatz ergibt.

Wie im Fall $n = 1$ beweist man: Sei $a \in \mathbb{R}^n$ ein Punkt, so daß

$$|a_\nu a^\nu| \le C \ \text{ für alle Multiindizes } \ \nu$$

mit einer geeigneten Konstanten C gilt. Dann konvergiert die Reihe

$$\sum_\nu a_\nu x^\nu$$

in dem Bereich

$$|x_i| < |a_i| \text{ für } i = 1, \dots, n.$$

Sind alle a_i von Null verschieden, so ist dieser Bereich offen, und die Potenzreihe definiert dort eine analytische Funktion.

Wir stellen abschließend die Regeln für das Rechnen mit Potenzreihen und die entsprechenden Permanenzeigenschaften analytischer Funktionen zusammen.

Sei $D \subset \mathbb{R}^n$ offen und $a \in D$ ein fester Punkt in D.

Die Summe und das Produkt zweier analytischer Funktionen sind wieder analytisch.

Die Taylorreihen von $f + g$ und $f \cdot g$ erhält man aus denen von f und g mittels der Rechenregeln

$$\sum a_\nu x^\nu + \sum b_\nu x^\nu = \sum (a_\nu + b_\nu) x^\nu$$

$$\left(\sum a_\nu x^\nu\right) \cdot \left(\sum b_\nu x^\nu\right) = \sum_\nu \left(\sum_{\alpha+\beta=\nu} a_\alpha b_\beta\right) x^\nu$$

Hierbei sind α, β und ν Multiindizes!

Sei $f : D \to \mathbb{R}^m$ eine analytische Abbildung und

$$g : D' \to \mathbb{R}, \quad D' \subset f(D) \subset \mathbb{R}^m \text{ offen},$$

eine analytische Abbildung, in welche sich f einsetzen läßt. Dann ist auch die Funktion

$$g \circ f : D \to \mathbb{R}$$

analytisch. Die zugehörige Taylorreihe erhält man aus der Taylorreihe von f in a und der von g in $f(a)$ durch Einsetzen und Umordnen.

Ein Spezialfall hiervon ist das Invertieren von Potenzreihen:

Ist $f : D \to \mathbb{R}$ *eine analytische Funktion ohne Nullstellen, so ist auch* $1/f$ *als Zusammensetzung von* f *mit der Funktion* $x \mapsto 1/x$ *analytisch.*

Sei

$$\varphi : D \to D', \quad D, D' \subset \mathbb{R}^n \text{ offen,}$$

ein Diffeomorphismus, also eine bijektive Abbildung, so daß φ und φ^{-1} stetig differenzierbar sind. Wenn φ analytisch ist, so ist auch φ^{-1} analytisch (vgl. Satz IV.6.6).

Kapitel VII. Integrationstheorie

1. Das Integral für stetige Funktionen mit kompakten Trägern

Gegeben sei ein abgeschlossener Quader Q im $\mathbb{R}^n$, welcher durch die n-Tupel

$$a = (a_1, \dots, a_n), \quad b = (b_1, \dots, b_n), \quad a_\nu \le b_\nu \quad \text{für} \quad \nu = 1, \dots, n,$$

definiert sei:

$$Q = \{x \in \mathbb{R}^n; \quad a_\nu \le x_\nu \le b_\nu\}.$$

Es sei eine stetige Funktion

$$f : Q \longrightarrow \mathbb{R}$$

gegeben. Wie versuchen, das (mehrfache) Integral von f

$$\int\limits_Q f(x)\, dx_1 \dots dx_n$$

zu definieren*). Dies soll induktiv geschehen. Zunächst kann man bei festem $x_2, \dots, x_n$ das Integral

$$\int\limits_{a_1}^{b_1} f(x)\, dx_1$$

definieren, da f als Funktion von x_1 stetig und somit integrierbar ist.

Läßt man jetzt $x_2, \dots, x_n$ wieder variieren, so kann man dieses Integral als Funktion von $x_2, \dots, x_n$ auffassen. Definitionsbereich ist der Quader im $\mathbb{R}^{n-1}$, der durch

$$a_\nu \le x_\nu \le b_\nu \quad \text{für} \quad 2 \le \nu \le n,$$

definiert ist.

*) Wir haben dies als Anwendung des Weierstraßschen Approximationssatzes im Anschluß an Folgerung V.6.4 bereits durchgeführt und rollen dies nochmals wegen seiner grundlegenden Bedeutung für den Aufbau der Integrationstheorie auf.

Jetzt will man über x_2 integrieren, d.h.

$$\int_{a_2}^{b_2}\left[\int_{a_1}^{b_1} f(x)\,dx_1\right]\,dx_2$$

betrachten, usw. und schließlich soll dann das mehrfache Integral von f über Q durch die Formel

$$\int_Q f(x)\,dx_1\dots dx_n = \int_{a_n}^{b_n}\dots\int_{a_1}^{b_1} f(x)\,dx_1\dots dx_n$$

definiert werden. Allerdings muß man wissen, daß all diese Integrationen durchführbar sind.

1.1 Hilfssatz. *Die Funktion* nachIS

$$f : Q \longrightarrow \mathbb{R}, \quad Q = \{x \in \mathbb{R}^n; \quad a_\nu \le x_\nu \le b_\nu\} \quad (a_\nu \le b_\nu)$$

sei stetig. Dann ist auch die Funktion

$$F : Q' \longrightarrow \mathbb{R}, \quad Q' = \{(x_2,\dots,x_n) \in \mathbb{R}^{n-1}; \quad a_\nu \le x_\nu \le b_\nu\},$$
$$F(x_2,\dots,x_n) = \int_{a_1}^{b1} f(x_1,x_2,\dots,x_n)\,dx_1$$

stetig.

Beweis. Dies haben wir bereits bewiesen (s. Bemerkung V.4.16). Wegen der Bedeutung für den Aufbau der Integrationstheorie gehen wir nochmal darauf ein:

Wesentliches Hilfsmittel zum Beweis ist der *Satz von der gleichmäßigen Stetigkeit*. Versuchen wir zunächst einmal, die Stetigkeit direkt zu beweisen, etwa im Punkt $(x_2^0,\dots,x_n^0)$.

Man muß zeigen:

$$\mid F(x_2,\dots,x_n) - F(x_2^0,\dots,x_n^0) \mid < \varepsilon \quad \text{für} \quad \mid x_\nu - x_\nu^0 \mid < \delta(\varepsilon) \quad (2 \le \nu \le n).$$

Dabei sei $\varepsilon > 0$ beliebig vorgegeben, also

$$\left| \int_{a_1}^{b_1} [f(x_1,x_2,\dots,x_n) - f(x_1,x_2^0,\dots,x_n^0)]\,dx_1 \right| < \varepsilon.$$

Nun gilt wegen der Stetigkeit von f tatsächlich

$$| f(x_1, x_2, \dots, x_n) - f(x_1, x_2^0, \dots, x_n^0) | < \varepsilon \text{ für } | x_\nu - x_\nu^0 | < \delta(\varepsilon) \quad (2 \le \nu \le n)$$

und man kann abschätzen:

$$| F(x_2, \dots, x_n) - F(x_2^0, \dots, x_n^0) | < \int_{a_1}^{b_1} \varepsilon \, dx_1 = (b_1 - a_1)\varepsilon$$

$$\text{für} \quad | x_\nu - x_\nu^0 | < \delta(\varepsilon) \quad (2 \le \nu \le n).$$

Das Auftreten der Konstanten $b_1 - a_1$ störte natürlich in keiner Weise, aber der Beweis funktioniert nur dann, wenn $\delta(\varepsilon)$ unabhängig von x_1 gewählt werden kann.

Dies folgt aus dem Satz von der gleichmäßigen Stetigkeit (Satz III.1.14). Damit ist folgende Definition möglich.

1.2 Definition. *Das Integral einer stetigen Funktion* `InkoQ`

$$f : Q \longrightarrow \mathbb{R}, \quad Q = \{x; \quad a_\nu \le x_\nu \le b_\nu\} \quad (a_\nu \le b_\nu),$$

auf einem kompakten Quader Q wird induktiv definiert durch

$$\int_Q f(x) \, dx_1, \dots, \, dx_n = \int_{Q'} \left[\int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) \, dx_1 \right] dx_2 \dots dx_n.$$

$$(Q' = \{(x_2, \dots, x_n) \in \mathbb{R}^{n-1}; \quad a_\nu \le x_\nu \le b_\nu\} \quad \textit{für} \quad 2 \le \nu \le n)$$

$$\int_{[a,b]} f(x) \, dx = \int_a^b f(x) \, dx \quad (a < b) \quad \textit{im Fall} \quad n = 1.$$

Schreibweise:

$$\int_Q f(x) \, dx = \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x) \, dx_1 \dots dx_n.$$

Der Aufbau der Integrationstheorie ist vergleichsweise kompliziert. Für pragmatisch denkende Anwender ist vielleicht der Hinweis nützlich, daß das Regelintegral einer Veränderlicher bereits ausreicht, um Volumenberechnungen durchzuführen. Man gehe folgendermaßen vor: Sei $A \subset \mathbb{R}^n$ eine einigermaßen anständige, sagen wir kompakte Teilmenge, welche man sich als massiven n-dimensionalen Körper vorstelle. Durchschneidet man diesen Körper mit der „Ebene" $x_n = t$, so erhält man einen $(n-1)$-dimensionalen Körper $A(t)$. Man will die Volumentheorie induktiv aufbauen, nimmt also an, daß man weiß, wie man das $(n-1)$-dimensionale Volumen von $A(t)$ berechnen kann. Induktionsbeginn sind Teilmengen von $\mathbb{R}$, welche man sich als Vereinigung endlich vieler Intervalle vorstelle. Die Funktion $h(t) = \mathrm{vol}(A(t))$ verschwindet außerhalb eines geeigneten Intervalls $[-C, C]$. In einigermaßen anständiger Situation darf man erwarten, daß sie eine Regelfunktion ist, so daß man

$$\mathrm{vol}(A) := \int_{-C}^{C} h(t)\, dt$$

definieren kann. Aber dieser naive Weg gibt keine direkte Einsicht, warum beispielsweise das Volumen gegenüber Drehungen des Körpers invariant bleibt. Dennoch: Obige Vorgehensweise reicht aus für alle praktischen Belange der Volumenbestimmung und der eine oder andere Hörer mag mit dieser Erkenntnis zufrieden sein und die Mühe des Aufbaus einer leistungsstarken Integrationstheorie als überflüssig erachten. Dies ist legitim, wenn er sich in der Mathematik von der Analysis wegorientiert und beispielsweise eine algebraische Linie bevorzugt. Tieferer Einstieg in Gebiete wie Analysis, Wahrscheinlichkeitstheorie, etc. erfordert jedoch eine leistungsstarke Integrationstheorie mit guten Grenzwertsätzen (Vertauschbarkeit von Integration mit Limesbildungen wie beispielsweise Differentiation).

Für die Integrationstheorie ist fundamental, daß man die Reihenfolge der Integrationen vertauschen darf, daß also beispielsweise im Fall $n = 2$

$$\int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2)\, dx_1 \; dx_2 = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2)\, dx_2 \; dx_1$$

gilt. Dies kann man an vielen Beispielen nachprüfen, muß aber streng bewiesen werden. Wir führen dies nochmals durch:

Gegeben sei eine Umordnung (Permutation) $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ der Zahlen $\{1, \dots, n\}$, d.h. jede Zahl 1 bis n kommt unter den σ_ν genau einmal vor.

1.3 Satz. *Gegeben sei eine stetige Funktion $f : Q \to \mathbb{R}$, Q ein kompakter Quader in $\mathbb{R}^n$ und σ eine Permutation der Variablen. Dann gilt* URdVa

$$\int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x)\, dx_1 \dots dx_n = \int_{a_{\sigma n}}^{b_{\sigma n}} \dots \int_{a_{\sigma_1}}^{b_{\sigma_1}} f(x)\, dx_{\sigma_1} \dots dx_{\sigma_n}.$$

Es kommt also auf die Reihenfolge der Integrationen nicht an.

Beweis: Die Behauptung ist trivial für Funktionen von dem speziellen Typ

$$f(x) = f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n),$$

wobei
$$f_\nu : [a_\nu\ ,\ b_\nu] \longrightarrow \mathbb{R}$$
stetige Funktionen einer Variablen sind, denn dann zerfällt das Integral
$$\int\limits_Q f(x)\,dx = \left(\int\limits_{a_1}^{b_1} f_1(x_1)\,dx_1\right)\dots\left(\int\limits_{a_n}^{b_n} f_n(x_n)\,dx_n\right)$$
in ein Produkt von Integralen.

Wir bezeichnen mit A die Menge aller stetigen Funktionen $f : Q \to \mathbb{R}$, die sich als endliche Summe von Funktionen obigen speziellen Typs schreiben lassen, also
$$f(x) = \sum_{\nu=1}^{m} f_{1\nu}(x_1)\cdots f_{n\nu}(x_n).$$
Es ist klar, daß die Behauptung auch für alle Funktionen dieser Klasse gilt.

Satz 1.3 wird dann mit Hilfe des *Approximationssatzes V.6.2 von Stone-Weierstraß* bewiesen.

Zunächst behaupten wir:

Jede stetige Funktion $f : Q \to \mathbb{R}$ ist gleichmäßiger Grenzwert einer Folge von Funktionen aus A. Dazu weisen wir die Voraussetzungen des Approximationssatzes nach. Der einzige nicht völlig triviale Punkt ist die Punktetrennung.

Seien also
$$x^0 = (x_1^0, \dots, x_n^0), \quad y^0 = (y_1^0, \dots, y_n^0)$$
zwei verschiedene Punkte aus Q, also etwa
$$x_\nu^0 \neq y_\nu^0.$$
Die Funktion f mit
$$f(x) = x_\nu - x_\nu^0$$
trennt die beiden Punkte $\big(f(x^0) = 0, \quad f(y^0) \neq 0\big)$ und liegt in A. Da man jede stetige Funktion $f : Q \to \mathbb{R}$ gleichmäßig approximieren kann durch Funktionen, für die Satz 1.3 schon bewiesen ist, muß man nur noch wissen, daß das mehrfache Integral bezüglich gleichmäßiger Konvergenz stabil ist.

1.4 Hilfssatz. *Sei $Q \subset \mathbb{R}^n$ ein kompakter Quader und* InmglS
$$f : Q \longrightarrow \mathbb{R}$$
eine stetige Funktion. Es gilt
$$\left|\int\limits_Q f(x)\,dx_1 \dots dx_n\right| \leq \|f\|(b_1 - a_1)\dots(b_n - a_n),$$
wobei $\|f\|$ die Maximumsnorm bezeichne.

Beweis. Aus den Rechenregeln für eine Veränderliche folgt unmittelbar

$$\int_Q f(x)\,dx_1 \dots dx_n \le \int_Q g(x)\,dx_1 \dots dx_n,$$

falls

$$f(x) \le g(x) \quad \text{für alle} \quad x \in Q$$

gilt. Insbesondere gilt

$$\left|\int_Q f(x)\,dx_1 \dots dx_n\right| \le \int_Q |f(x)|\,dx_1 \dots dx_n \le$$
$$\int_Q \|f\|\;dx_1 \dots dx_n = \|f\|(b_1 - a_1)\dots(b_n - a_n).$$

1.5 Folgerung. *Die Folge von stetigen Funktionen* StmvGl

$$f_k : Q \longrightarrow \mathbb{R}, \quad k = 1, 2, 3, \dots,$$

konvergiere gleichmäßig gegen f. Dann gilt

$$\int_Q f(x)\,dx_1 \dots dx_n = \lim_{k\to\infty} \int_Q f_k(x)\,dx_1 \dots dx_n.$$

Beweis:

$$\left|\int_Q f(x)\,dx_1 \dots dx_n - \int_Q f_k(x)\,dx_1 \dots dx_n\right| \le$$

$$\|f - f_k\|\;(b_1 - a_1)\dots(b_n - a_n) \longrightarrow 0 \quad \text{für} \quad k \longrightarrow \infty.$$

Man sollte sich gut vor Augen halten:

Die Stabilität des Integrals für stetige Funktionen auf kompakten Quadern ist eine Trivialität, welche auf der banalen Abschätzung in Hilfssatz 1.4 beruht.

Stetige Funktionen mit kompaktem Träger

Unter dem *Träger* einer Funktion

$$f : X \longrightarrow \mathbb{R}, \quad X \text{ ein metrischer Raum,}$$

versteht man den Abschluß der Menge aller Punkte, in denen f nicht verschwindet.

$$\text{Träger } f = \overline{\{x \in X; \quad f(x) \neq 0\}}.$$

Ein Punkt $x \in X$ gehört dann und nur dann zum Träger von f, wenn $f(x) \neq 0$ ist, oder wenn es eine Folge von Punkten $x_n \in X$ gibt mit

$$x_n \longrightarrow x \quad \text{für} \quad n \longrightarrow \infty \quad \text{und} \quad f(x_n) \neq 0 \quad \text{für alle } n.$$

Bezeichnung

Klasse der stetigen Funktionen mit kompaktem Träger auf X:

$$C_c(X) = \{f : X \longrightarrow \mathbb{R}; \quad f \text{ stetig, Träger}(f) \text{ ist } \textbf{kompakt}\}$$

Offenbar gilt

$$f, g \in C_c(X) \Longrightarrow f + g, \quad f \cdot g \text{ und } cf \in C_c(X) \quad (c \in \mathbb{R}).$$

Die konstanten Funktionen liegen nur dann in $C_c(X)$, wenn X selbst kompakt ist.

1.6 Bemerkung. *Eine stetige Funktion* FmKoT

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

gehört dann und nur dann zur Klasse $C_c(\mathbb{R}^n)$, wenn eine Zahl $R > 0$ existiert, so daß

$$f(x) = 0 \quad \textit{für} \quad \|x\| > R$$

gilt.

Beweis. 1) Wenn der Träger von f kompakt ist, so existiert eine Zahl $R > 0$, so daß

$$\text{Träger } (f) \subset U_R(0).$$

(Jedes Kompaktum ist beschränkt.)

2) Wenn ein solches R existiert, so ist der Träger von f beschränkt, außerdem ist er nach Konstruktion abgeschlossen und daher nach dem Überdeckungssatz von Heine–Borel kompakt.

Im Fall $n > 1$ erhält man stetige Funktionen mit kompaktem Träger durch

$$f(x) = f_1(x_1) \cdots f_n(x_n),$$

wobei die $f_i(x_i)$ solche in einer Veränderlichen sind.

Das Integral für stetige Funktionen mit kompaktem Träger

Es sei

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

eine stetige Funktion mit kompaktem Träger. Wir wählen $R > 0$, so daß gilt:

$$f(x) \neq 0 \Longrightarrow |x_\nu| \leq R \quad \text{für} \quad \nu = 1, \dots, n.$$

Dann definieren wir

$$\int\limits_{\mathbb{R}^n} f(x)\, dx_1 \dots dx_n \overset{\text{Def}}{=} \int\limits_Q f(x)\, dx_1 \dots dx_n,$$

wobei Q den Quader

$$-R \leq x_\nu \leq R \quad \text{für} \quad \nu = 1, \dots, n$$

bezeichne. Es ist klar, daß diese Definition von der Wahl von R nicht abhängt. R muß nur so groß sein, daß der Träger von f in Q liegt.

Das so definierte Integral für stetige Funktionen mit kompaktem Träger ist der Baustein des Lebesgue'schen Integrals, das nun mit Hilfe des Daniell-Lebesgue Prozesses gewonnen werden soll.

Warum Integrationstheorie?

Mit Hilfe der Integrationstheorie mehrerer Variablen will man u.a. mehrdimensionale Volumina von Bereichen im $\mathbb{R}^n$ berechnen.

Sei also

$$f : D \longrightarrow \mathbb{R}, \quad D \subset \mathbb{R}^n,$$

eine Funktion, die der Einfachheit halber nirgends negativ sei,

$$f(x) \geq 0 \quad \text{für alle } x \in D \quad (\text{Schreibweise } f \geq 0).$$

Gesucht ist ein Maß für das Volumen des Bereiches im $(n+1)$-dimensionalen Raum

$$\big\{(x,t) \in \mathbb{R}^{n+1}; \quad x \in D,\ 0 \leq t \leq f(x)\big\}$$

Dieses Volumen soll gerade

$$\int\limits_D f(x)\, dx_1 \cdots dx_n$$

sein.

Man könnte natürlich wie im Fall $n = 1$ versuchen, das Integral durch Approximationen mit Hilfe von „Treppen“ zu definieren. Das ist möglich, aber schwierig. Im Fall $n = 1$ sind die Definitionsbereiche D in der Regel einfache Intervalle. Diese sind sehr einfach aufzuteilen in „kleine“ Intervalle, auf denen man dann die Treppen aufbaut. Im Fall $n > 1$ kann schon der Definitionsbereich D relativ kompliziert sein. Man müßte ihn durch eine Quaderaufteilung pflastern (approximativ) und darauf die „Treppen“ aufbauen. Dieser Aufbau ist möglich. Stattdessen haben wir einen anderen Weg eingeschlagen. In sehr naheliegender Weise konnten wir direkt das Integral für stetige Funktionen definieren. Dazu machten wir einen Rückgriff auf das Regelintegral in einer Veränderlichen, das ja sehr leicht (relativ zum Fall $n \geq 1$) zu gewinnen war. Jetzt werden wir das Integral für allgemeinere Funktionentypen f dadurch gewinnen, daß wir f durch stetige Funktionen mit kompaktem Träger zu approximieren suchen.

Dabei müssen wir notgedrungen auf die gleichmäßige Konvergenz verzichten, sonst kommen wir nicht aus dem Bereich der stetigen Funktionen heraus.

Warum ist das Integral für allgemeinere Funktionstypen interessant?

Will man das n-dimensionale Volumen eines Bereiches $D \subset \mathbb{R}^n$ definieren und berechnen, so kann man folgendermaßen vorgehen.

Man betrachte die Funktion

$$f(x) = 1 \text{ für } x \in D.$$

Die Anschauung zeigt dann, daß die Definition

$$\text{Volumen } (D) = \int\limits_D 1 \; dx_1 \cdots dx_n$$

sinnvoll ist.

(Die Länge des Intervalls $[a, b]$ kann als Integral über die Funktion 1 interpretiert werden, also eindimensionales Volumen von $[a, b]$:

$$\int\limits_a^b dx = b - a,$$

oder: Die Fläche der Kreisscheibe

$$E = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R}; \quad x^2 + y^2 \leq 1 \right\}$$

kann interpretiert werden als Volumen des dreidimensionalen Zylinders

$$Z = \left\{(x, y, t) \in \mathbb{R}^3; \quad x^2 + y^2 \leq 1, \; 0 \leq t \leq 1\right\},$$

das wäre das Integral

$$\text{Volumen von } Z = \int\limits_{\substack{x^2+y^2\leq 1\\ 0\leq t\leq 1}} dx\, dy\, dt.)$$

Im folgenden werden wir nur Integrale über den $\mathbb{R}^n$ betrachten, wir nehmen also an, daß die Funktion f auf dem ganzen $\mathbb{R}^n$ definiert ist und streben von vornherein das uneigentliche Integral

$$\int\limits_{\mathbb{R}^n} f(x)\, dx_1 \cdots dx_n$$

an.

Dies ist keine Einschränkung der Allgemeinheit, wenn eine Funktion

$$f : D \longrightarrow \mathbb{R}, \quad D \subset \mathbb{R}^n$$

gegeben ist, so betrachten wir einfach $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$,

$$\tilde{f}(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

und *definieren*

$$\int\limits_{D} f(x)\, dx_1 \cdots dx_n := \int\limits_{\mathbb{R}^n} \tilde{f}(x)\, dx_1 \cdots dx_n,$$

vorausgesetzt, dass die rechte Seite schon definiert ist. Für die Volumenmessung Volumen(D) bedeutet dies folgendes: Man betrachte die sogenannte *charakteristische Funktion* von D,

$$\chi_D : \mathbb{R}^n \to \mathbb{R}, \quad \chi_D(x) = \begin{cases} 1 & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

und definiert

$$\text{Volumen}(D) = \int\limits_{\mathbb{R}^n} \chi_D(x)\, dx_1 \cdots dx_n.$$

Damit ist gezeigt, daß das Konzept des Volumens eines Bereiches $D \subset \mathbb{R}^n$ sich dem Begriffsapparat des Integrals

$$\int\limits_{\mathbb{R}^n} f(x)\, dx_1 \cdots dx_n, \quad f : \mathbb{R}^n \longrightarrow \mathbb{R},$$

unterordnet, aber man beachte, daß die Funktion χ_D nicht stetig ist (Die Randpunkte von D sind Unstetigkeitspunkte). Das Integral für stetige Funktionen mit kompaktem Träger reicht also keineswegs zur Berechnung von Volumina aus.

2. Das Bairesche Integral

In diesem Abschnitt behandeln wir den ersten Teil des Daniell-Lebesgue Prozesses, die Konstruktion des Baireschen Integrals. In §1 haben wir die Klasse $C_c(\mathbb{R}^n)$ der stetigen Funktionen mit kompaktem Träger eingeführt und darauf ein Integral definiert:

$$f \longmapsto \int\limits_{\mathbb{R}^n} f(x)\, dx_1 \cdots dx_n.$$

Wir stellen die Eigenschaften dieses Integrals, die wir im folgenden benutzen, kurz zusammen.

Dabei benutzen wir die abkürzende Schreibweise

$$I(f) := \int\limits_{\mathbb{R}^n} f(x)\, dx := \int\limits_{\mathbb{R}^n} f(x)\, dx_1 \cdots dx_n.$$

1) *Das Integral ist ein lineares Funktional, d.h.*

$$I(f+g) = I(f) + I(g), \quad I(cf) = cI(f).$$

2) *Das Integral ist „positiv“, d.h.*

$$I(f) \geq 0, \quad \text{falls} \quad f(x) \geq 0 \quad \text{für alle } x \in \mathbb{R}^n.$$

Im folgenden besteht das Problem, das Integral auf eine möglichst große Klasse von Funktionen auszudehnen.

Zunächst haben wir, das Integral für Funktionen einer Veränderlichen schon benutzend, das Integral für stetige Funktionen mit kompaktem Träger gewonnen. Allgemeinere Funktionen f versuchen wir jetzt, durch stetige Funktionen zu approximieren. *Dabei müssen wir allerdings das gelobte Land der gleichmäßigen Konvergenz verlassen, denn sonst kommen wir aus dem Bereich der stetigen Funktionen nicht heraus.*

Der Typ der Konvergenz, den wir betrachten wollen, ist die *monotone* Konvergenz. Wichtigstes Hilfsmittel für den ersten Schritt im Daniell-Lebesgue-Prozess ist folgender Satz.

2.1 Satz (Dini). *Es sei X ein kompakter metrischer Raum und* SaDin

$$f_n : X \longrightarrow \mathbb{R}, \quad n = 1, 2, 3 \ldots,$$

eine Folge von stetigen Funktionen, die monoton gegen Null fällt, d.h.

a) $f_1(x) \geq f_2(x) \geq \cdots,$

b) $\lim\limits_{n\to\infty} f_n(x) = 0 \qquad$ (*für jedes* $x \in X$).

Die Folge (f_n) konvergiert dann gleichmäßig gegen 0.

Beweis. Sei $\varepsilon > 0$ vorgegeben. Zu jedem $x \in X$ existiert eine natürliche Zahl $N(\varepsilon, x)$ mit der Eigenschaft

$$|f_n(x)| < \varepsilon \quad \text{für } n \geq N(\varepsilon, x).$$

Zu jedem x_0 existiert dann eine offene Umgebung $U(x_0)$ mit der Eigenschaft

$$|f_{N(\varepsilon,x_0)}(x)| < \varepsilon \quad \text{für } x \in U(x_0).$$

Endliche viele dieser Umgebungen überdecken X (Kompaktheit). Wir definieren

$$N = N(\varepsilon) = \max\big\{\, N(\varepsilon, x_0); \quad x_0 \in \text{ obiger endlichen Menge } \big\}.$$

Es gilt dann

$$|f_N(x)| < \varepsilon \quad \text{für alle } x.$$

Wegen der Monotonie der Folge $F_n(x)$ gilt sogar

$$|f_n(x)| < \varepsilon \quad \text{für alle } x \text{ und } n \geq N.$$

Das heißt gerade, daß f_n gleichmäßig gegen Null konvergiert. □

Der Satz von Dini gibt uns die Möglichkeit, das Integral für stetige Funktionen mit kompaktem Träger auf eine große Klasse von Funktionen auszudehnen und zwar auf solche Funktionen f, die sich *monoton* durch eine Folge von Funktionen aus $C_c(\mathbb{R}^n)$ *approximieren* lassen.

Bei der technischen Durchführung der Integrationstheorie hat es sich als zweckmäßig erwiesen, auch Funktionen zuzulassen, welche die Werte ∞ annehmen dürfen, d.h. wir erweitern die reelle Zahlengerade $\mathbb{R}$ durch Hinzufügen eines weiteren Elements, für das wir ∞ schreiben und vereinbaren die folgenden Rechenregeln auf $\mathbb{R} \cup \{\infty\}$,

$$\infty > x \ \text{ für alle } x \in \mathbb{R}.$$

Diese Erweiterung hat folgenden Vorteil. Jede Teilmenge $M \subset \mathbb{R} \cup \{\infty\}$, sofern sie nicht leer ist, besitzt nun eine obere Grenze, die wir mit $\operatorname{Sup} M$ bezeichnen wollen*).

Das Supremum $\operatorname{Sup} M$ einer nicht leeren Teilmenge $M \subset \mathbb{R} \cup \{\infty\}$ ist die *kleinste* obere Schranke von $\mathbb{R} \cup \{\infty\}$.

Ist insbesondere $M \subset \mathbb{R}$ eine in $\mathbb{R}$ nach oben beschränkte nicht leere Teilmenge, so ist $\operatorname{Sup} M = \sup M$ das gewöhnliche Supremum. Ist hingegen M durch kein Element aus $\mathbb{R}$ nach oben beschränkt, so ist $\operatorname{Sup} M = \infty$.

---

*) Wir haben eine ähnliche Konvention bereits im Zusammenhang mit der Berechnung des Konvergenzradius einer Potenzreihe verwendet, s. Kapitel I, §4

Ist $a_1, a_2, a_3, \cdots$ eine Folge von Elementen aus $\mathbb{R}$, so verstehen wir unter dem Supremum dieser Folge einfach das Supremum der Menge der Folgenglieder

$$\operatorname{Sup}(a_n) = \operatorname{Sup}\{a_1, a_2, \ldots\}.$$

Die erweiterte Zahlengerade hat die gute Eigenschaft, daß es möglich ist, die Addition $+$ auf ganz $\mathbb{R} \cup \{\infty\}$ so zu erweitern, so daß Kommutativ- und Assoziativgesetz gültig bleiben. Man definiert einfach

$$\begin{aligned} x + \infty = \infty + x &= \infty \qquad \text{für alle } x \in \mathbb{R}, \\ \infty + \infty &= \infty. \end{aligned}$$

Doch schon bei der Multiplikation gibt es Einschränkungen. Immerhin sind die Definitionen

$$\begin{aligned} x \cdot \infty &= \infty \qquad \text{für alle } x \in \mathbb{R}, \quad x > 0, \\ \infty \cdot \infty &= \infty, \end{aligned}$$

noch mit Kommutativ-, Assoziativ- und Distributivgesetzen verträglich.

Allerdings ist es nicht möglich Division und Differenzbildung auf ganz $\mathbb{R} \cup \{\infty\}$ auszudehnen, so dass diese Rechenregeln ausnahmslos erhalten bleiben.

Wir betrachten nun Funktionen, die man monoton durch stetige Funktionen mit kompaktem Träger approximieren kann.

2.2 Definition. *Eine Funktion* BaiKl

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$$

gehört der Klasse $\mathcal{B}^+$ an (Bairesche Klasse), wenn es eine Folge von stetigen Funktionen mit kompaktem Träger

$$f_\nu : \mathbb{R}^n \longrightarrow \mathbb{R}$$

mit den Eigenschaften

a) $f_1(x) \le f_2(x) \le f_3(x) \cdots$
b) $f(x) = \operatorname{Sup}_\nu f_\nu(x)$ *für alle* $x \in \mathbb{R}^n$

gibt.

Im folgenden schreiben wir einfach $f_n \uparrow f$, wenn die Eigenschaften a) und b) erfüllt sind.

(Die Bezeichnung $f_n \downarrow f$ versteht sich von selbst.)

Es ist klar, daß die Funktionen aus $\mathcal{B}^+$ den Wert ∞ annehmen können. Man kann Funktionen aus der Klasse $\mathcal{B}^+$ addieren, ohne diese Klasse zu verlassen. Die Menge $\mathcal{B}^+$ ist aber kein Vektorraum. Immerhin gilt noch

$$f \in \mathcal{B}^+, \quad C \ge 0 \implies Cf \in \mathcal{B}^+.$$

2.3 Bemerkung. *Es seien $f \in \mathcal{B}^+$ und $(f_\nu), (g_\nu)$ zwei Folgen von Funktionen aus C_c (stetige Funktionen mit kompaktem Träger) mit der Eigenschaft* StaBai

$$f_\nu \uparrow f, \quad g_\nu \uparrow f.$$

Dann gilt

$$\mathrm{Sup}_\nu \left(\int\limits_{\mathbb{R}^n} f_\nu(x)\, dx \right) = \mathrm{Sup}_\nu \left(\int\limits_{\mathbb{R}^n} g_\nu(x)\, dx \right).$$

Diese Bemerkung – wir werden sie gleich beweisen – gibt dann Anlaß zu folgender Definition.

2.4 Definition. *Es sei $f \in \mathcal{B}^+$. Das Bairesche Integral von f wird durch die Formel* BaiInt

$$I_{\mathcal{B}^+}(f) = \mathrm{Sup}_\nu \left(\int\limits_{\mathbb{R}^n} f_\nu(x)\, dx \right)$$

definiert, wobei $f_\nu \in C_c(\mathbb{R}^n)$ irgend eine Folge von Funktionen (stetig mit kompaktem Träger) ist, die f monoton wachsend approximiert, d.h. $f_\nu \uparrow f$.

Der Wert $I_{\mathcal{B}^+}(f)$ kann natürlich unendlich sein.

Die Bemerkung 2.3 besagt gerade, daß diese Definition unabhängig von der Wahl der Folge (f_ν) ist. Außerdem beachte man, daß eine Funktion $f \in C_c(\mathbb{R}^n)$ auch in $\mathcal{B}^+$ liegt, man kann sie durch die konstante Folge $f, f, f, \cdots$ monoton wachsend approximieren.

Also: Es gilt $C_c(\mathbb{R}^n) \subset \mathcal{B}^+$ und das Bairesche Integral stimmt auf C_c mit dem früher definierten Integral (§1) überein.

Dies erlaubt die Bezeichnung

$$I_{\mathcal{B}^+}(f) = I(f) = \int\limits_{\mathbb{R}^n} f(x) dx$$

auch für das Bairesche Integral.

Beweis von Bemerkung 2.3. Wir erinnern an die Bezeichnungen für die Funktionen

$$\begin{aligned} &f \vee g \text{ und } f \wedge g, \text{ die durch} \\ &(f \vee g)(x) = \max(f(x), g(x)), \\ &(f \wedge g)(x) = \min(f(x), g(x)) \end{aligned}$$

definiert sind.

Es gilt: $f, g \in C_c \Longrightarrow f \vee g$ und $f \wedge g \in C_c$.

Wir zeigen zunächst folgendes:

Sei

$$f_\nu \uparrow f, \quad f_\nu \text{ stetig mit kompaktem Träger,}$$

und sei g irgendeine stetige Funktion mit kompaktem Träger mit der Eigenschaft

$$f \geq g.$$

Dann ist

$$\operatorname{Sup} \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx \geq \int\limits_{\mathbb{R}^n} g(x)\,dx.$$

Beweis: Die Folge $g - (f_\nu \wedge g)$ fällt offenbar monoton gegen Null. Nach dem *Satz von Dini* konvergiert sie daher gleichmäßig gegen Null, und da das Integral stabil gegenüber gleichmäßiger Konvergenz ist (Folgerung 1.5), gilt

$$\lim_{\nu\to\infty} \int\limits_{\mathbb{R}^n} (f_\nu \wedge g)\,dx = \int\limits_{\mathbb{R}^n} g(x)\,dx.$$

Nun ist

$$f_\nu \wedge g \leq f_\nu \quad \text{für alle } \nu,$$

es folgt daher

$$\int\limits_{\mathbb{R}^n} g(x)\,dx \leq \operatorname{Sup}_\nu \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx.$$

Es sei nun

$$g_\nu \uparrow f, \quad g_\nu \in C_c.$$

Nach dem, was eben bewiesen wurde, gilt

$$\int\limits_{\mathbb{R}^n} g_\mu(x)\,dx \leq \operatorname{Sup}_\nu \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx \quad \text{für jedes } \mu.$$

Insbesondere gilt

$$\operatorname{Sup}_\nu \int\limits_{\mathbb{R}^n} g_\nu(x)\,dx \leq \operatorname{Sup}_\nu \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx.$$

Die umgekehrte Ungleichung gilt genauso, da man die Rollen von f_k und g_k vertauschen kann. □

Dieser Beweis zeigt in Wirklichkeit noch etwas mehr, nämlich:

2.5 Hilfssatz. *Seien $f \le g$ zwei Funktionen aus $\mathcal{B}^+$, dann gilt* MonBai

$$\int\limits_{\mathbb{R}^n} f(x)\,dx \le \int\limits_{\mathbb{R}^n} g(x)\,dx.$$

Rechenregeln für die Bairesche Klasse

Wir nennen manchmal die Funktionen aus $\mathcal{B}^+$ auch (unter-) halbstetig. Das hat seine Berechtigung in folgender Eigenschaft der Funktionen aus $\mathcal{B}^+$, die für stetige Funktionen wohlbekannt ist.

2.6 Bemerkung. *Seien $f \in \mathcal{B}^+$, C eine reelle Zahl und $a \in \mathbb{R}^n$ ein Punkt mit $f(a) > C$. Dann gilt* halbST

$$f(x) > C \quad \textit{in einer vollen Umgebung von } a.$$

Beweis: Sei

$$f_\nu \uparrow f, \quad f_\nu \in C_c.$$

Nach Definition des Supremums existiert ein Index l mit

$$f(a) \ge f_l(a) > C.$$

Für die stetige Funktion f_l stimmt aber die Behauptung und damit erst recht für f. Allgemein nennt man Funktionen mit der in Bemerkung 2.6 genannten Eigenschaft *unterhalbstetig.* Man kann umgekehrt zeigen, daß jede unterhalbstetige Funktion $f : \mathbb{R}^n \to \mathbb{R}$ mit der Eigenschaft $f(x) \ge 0$ außerhalb eines Kompaktums zu $\mathcal{B}^+$ gehört.

Als nächstes zeigen wir, daß das Integral für Funktionen aus $\mathcal{B}^+$ stabil gegenüber monotoner Approximation ist.

2.7 Hilfssatz. *Sei* StBaMo

$$f_1 \le f_2 \le f_3 \le \cdots$$

eine monotone Folge aus $\mathcal{B}^+$. Dann gilt

a) $$f = \operatorname{Sup} f_\nu \in \mathcal{B}^+,$$

b) $$\int\limits_{\mathbb{R}^n} f(x)\,dx = \operatorname{Sup} \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx.$$

Beweis. Sei
$$f_k \uparrow f, \quad f_k \in \mathcal{B}^+ \quad \text{(Bairesche Klasse).}$$
Es existiert also
$$f_{1k} \le f_{2k} \le f_{3k} \cdots \quad \text{mit } f_k = \operatorname{Sup}_i f_{ik}.$$
Wir bilden die Funktionenfolge
$$g_r = \bigvee_{i+k\le r} f_{ik} \quad \bigl(\text{also } g_r(x) = \max_{i+k\le r} (f_{ik}(x))\bigr).$$
Offenbar gilt $g_r \in C_c$.

Außerdem gelten die Ungleichungen
$$g_1 \le g_2 \le g_3 \le \cdots \le f.$$
Hieraus folgt
$$g = \operatorname{Sup} g_k \le f.$$
Beachtet man außerdem
$$f_{ik} \le g_{i+k} \le g \quad \text{für alle } i, k,$$
so folgt
$$f_k \le g \quad \text{für alle } k = 1, 2, \ldots$$
und daher
$$f = \operatorname{Sup} f_k \le g,$$
insbesondere also $f = g$.

Damit ist gezeigt:
$$g_k \uparrow f, \quad \text{also } f \in \mathcal{B}^+.$$
Außerdem gilt
$$\int\limits_{\mathbb{R}^n} f(x)\,dx = \operatorname{Sup}_k \int\limits_{\mathbb{R}^n} g_k(x)\,dx.$$
Aus den Ungleichungen
$$g_k \le f_k \quad \text{für alle } k$$
folgt (Hilfssatz 2.5)
$$\int\limits_{\mathbb{R}^n} g_k(x)\,dx \le \int\limits_{\mathbb{R}^n} f_k(x)\,dx \le \int\limits_{\mathbb{R}^n} f(x)\,dx$$
und hiermit
$$\operatorname{Sup} \int\limits_{\mathbb{R}^n} f_k(x)\,dx = \int\limits_{\mathbb{R}^n} f(x)\,dx. \qquad \square$$

2.8 Hilfssatz. *Seien $f, g \in \mathcal{B}^+$ und a, b nicht negative Zahlen, $a \geq 0,\ b \geq 0$. Dann gilt $af + bg \in \mathcal{B}^+$ und* HaGrBa

$$\int\limits_{\mathbb{R}^n} \big(af(x) + bg(x)\big)\, dx = a \int\limits_{\mathbb{R}^n} f(x)\, dx + b \int\limits_{\mathbb{R}^n} g(x)\, dx.$$

Beweis: Seien

$$f_k \uparrow f,\ g_k \uparrow g, \quad f_k, g_k \in C_c.$$

Offenbar gilt

$$af_k + bg_k \uparrow af + bg \quad \text{(beachte: } a > 0,\ b > 0!)$$

und die Behauptung ist evident. □

Hier sieht man auch den wesentlichen Nachteil des Integrals für $\mathcal{B}^+$. Leider gilt im allgemeinen nicht

$$f \in \mathcal{B}^+ \Longrightarrow -f \in \mathcal{B}^+,$$

schon deshalb nicht, weil wir $-\infty$ nicht zur erweiterten Zahlgerade hinzugefügt haben aber auch aus anderen Gründen. Zum Beispiel ist $-f$ in der Regel nicht mehr unterhalbstetig (sondern oberhalbstetig in naheliegendem Sinne).

3. Das äußere Integral

Wir kommen nun zum 2. Teil des Daniell-Lebesgue Prozesses, der Konstruktion des äußeren Integrals und darauf aufbauend der Definition des Lebesgue Integrals.

3.1 Hilfssatz. *Die Funktion* InfBai

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}, \quad f(x) = \infty \text{ für alle } x \in \mathbb{R}^n,$$

gehört der Klasse $\mathcal{B}^+$ an.

Wir deuten den Beweis nur im Fall $n = 1$ an. Es ist dann sehr einfach, ihn auf den Fall $n > 1$ zu verallgemeinern. Die Funktion f kann approximiert werden durch die Folge von „Dreiecken“

$$f_k(x) = \begin{cases} -|x| + k & \text{für } |x| \leq k, \\ 0 & \text{für } |x| \geq k. \end{cases}$$

3.2 Definition. *Das äußere Integral einer Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ ist definiert durch* DaI

$$\bar{I}(f) = \operatorname{Inf}\{I(g); \quad g \geq f,\ g \in \mathcal{B}^+\},$$

sofern dieses Infimum existiert (in $\mathbb{R} \cup \{\infty\}$). In diesem Fall sagt man, dass das äußere Integral existiert.

Wegen Hilfssatz 3.1 existiert das äußere Integral für beliebige Funktionen $f \ge 0$. (Häufig wird die reelle Zahlengerade durch ein weiteres Symbol $-\infty$ erweitert. In dieser erweiterten Zahlengerade existieren Inf und Sup für beliebige nicht leere Mengen. Allerdings ist dann die Addition nicht erklärbar, so dass das Assoziativgesetz ausnahmslos gültig ist.)

3.3 Bemerkung. *Ist $f \in \mathcal{B}^+$, so existiert das äußere Integral und stimmt mit dem Baireschen Integral überein.* BAeB

Beweis. Unter den Baireschen $h \ge f$ kommt f selbst vor. Es existiert also sogar das Minimum aller $I(h)$ und ist gleich $I(f)$. □

3.4 Hilfssatz. *Wenn das äußere Integral für f, g und $f+g$ existiert, so gilt* HaUn

$$\bar{I}(f+g) \le \bar{I}(f) + \bar{I}(g).$$

Ist $C > 0$ eine positive Konstante, so existiert das äußere Integral auch für Cf, wenn es für f existiert und es gilt $\bar{I}(Cf) = C\bar{I}(f)$.

Folgerung. *Die Funktion f nehme nur endliche Werte an. Wenn das äußere Integral für f und $-f$ existiert und endlich ist, so gilt*

$$-\bar{I}(-f) \le \bar{I}(f).$$

Beweis de Hilfssatzes. Man muss nur folgendes ausnutzen. Sind $h \ge f$ und $\tilde{h} \ge g$ Bairesche Funktionen, so ist $h + \tilde{h} \ge f + g$ auch eine Bairesche Funktion. Nimmt man das Infimum nur über diese Baireschen Funktion, so erhielte man Gleichheit. Da es oberhalb $f+g$ aber auch andere Bairesche Funktionen gibt, kann das Infimum höchstens kleiner werden. Die Folgerung erhält man aus dem Spezialfall $f + g = 0$. □

Das äußere Integral ist erfreulicherweise ordnungstreu.

3.5 Hilfssatz. *Es seien zwei Funktionen* ObUng

$$f, g : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\} \textit{ mit } f \le g$$

gegeben. Das äußere Integral für f existiere. Dann existiert auch das äußere Integral für g und es gilt $\bar{I}(f) \le \bar{I}(g)$.

Beweis. Ist $g \in \mathcal{B}^+$ eine Funktion der Baireschen Klasse mit der Eigenschaft $g \ge h$, so gilt erst recht $g \ge f$. Bei der Definition des Oberintegrals h werden also weniger Funktionen zur Konkurrenz zugelassen als bei f. Dieses wird daher höchstens größer. □

Eine Variante von Hilfssatz 3.5 besagt.

3.6 Hilfssatz. *Sei $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ eine beliebige und h eine Bairesche Funktion mit endlichem Integral. Es gelte $f + h \geq 0$. Dann existiert das äußere Integral von f und es gilt* HVb

$$\bar{I}(f) \geq -I(h).$$

Beweis. Wir müssen alle Baireschen $g \geq f$ betrachten und zeigen, dass $I(g)$ nach unten beschränkt ist. Dies folgt aus $I(g) + I(h) = I(g + h) \geq 0$. Auch die behauptete Ungleichung ergibt sich so. □

3.7 Definition. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion (ohne Unendlichkeitsstellen), so daß das äußere Integral von $-f$ existiert und endlich ist. Man nennt dann* DUi

$$\underline{I}(f) := -\bar{I}(-f)$$

das innere Integral (oder Unterintegral) von f.

(Unendlichkeitsstellen können wir hier nicht zulassen, da wir $\mathbb{R}$ nur durch ∞ nicht aber durch $-\infty$ erweitert haben.) Es gilt dann (Hilfssatz 3.4, Folgerung)

$$\underline{I}(f) \leq \bar{I}(f),$$

wenn das äußere und innere Integral von $f : \mathbb{R}^n \to \mathbb{R}$ existieren. Rechenregeln für das äußere Integral übertragen sich auf das innere Integral.

3.8 Definition. deNuF

a) *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ heißt **Nullfunktion**, wenn $\bar{I}(|f|) = 0$ gilt.*

b) *Eine Teilmenge $A \subset \mathbb{R}^n$ heißt **Nullmenge**, wenn die charakteristische Funktion χ_A*

$$\chi_A(x) = \begin{cases} 1 & \text{für } x \in A \\ 0 & \text{für } x \notin A \end{cases}$$

eine Nullfunktion ist.

Mit f ist auch h eine Nullfunktion, wenn $|h| \leq |f|$ gilt.

3.9 Hilssatz. *Sei $0 \leq f_1 \leq f_2 \leq \ldots$ eine monoton wachsende Folge von nirgends negativen Nullfunktionen, dann ist auch $\operatorname{Sup} f_\nu$ eine Nullfunktion.* HSnf

Beweis. Sei $\varepsilon > 0$. Wir wählen zu jedem ν eine Bairesche Funktion h_ν mit

$$f_\nu \leq h_\nu, \quad I(h_\nu) \leq \frac{\varepsilon}{2^\nu}.$$

Aus

$$f_\nu \leq h_\nu \leq h_1 + \cdots + h_\nu$$

folgt

$$\mathrm{Sup}(f_\nu) \leq \mathrm{Sup}(h_1 + \cdots h_\nu).$$

Auf der rechten Seite stehen nur Baireche Funktionen. Da das Bairesche Integral stabil gegenüber monotoner Konvergenz ist, folgt

$$\bar{I}(\mathrm{Sup}(f_\nu)) \leq \sum_\nu \frac{\varepsilon}{2^\nu} = \varepsilon.$$

Dies beweist den Hilfssatz. □

3.10 Satz. *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ ist dann und nur dann Nullfunktion, wenn die Menge $\{x \in \mathbb{R}^n;\ f(x) \neq 0\}$ eine Nullmenge ist.* NuMe

Beweis: Sei f eine Nullfunktion. Man wende Hilfssatz 3.9 auf die Folge $\nu|f|$ an und erhält

$$\int\limits_{\mathbb{R}^n} h(x)\,dx = 0,$$

wobei

$$h(x) = \begin{cases} \infty & \text{für } f(x) \neq 0, \\ 0 & \text{für } f(x) = 0. \end{cases}$$

Aus der Ungleichung

$$\chi_A \leq h \text{ mit } A = \{x \in \mathbb{R}^n,\ f(x) \neq 0\}$$

folgt dann, daß A eine Nullmenge ist.

Umkehrung. Sei $A = \{x \in \mathbb{R}^n;\ f(x) \neq 0\}$ eine Nullmenge. Der erste Teil dieses Beweises zeigt dann, daß h eine Nullfunktion ist und wegen $|f| \leq h$ ist dann auch f eine Nullfunktion. □

3.11 Satz. *Sei $f \geq 0$ eine nirgends negative Funktion, für die das äußere Integral endlich ist. Dann ist die Menge der Unendlichkeitsstellen von f eine Nullmenge.* SUn

Beweis. Sei χ die charakteristische Funktion der Menge der Unendlichkeitsstellen von f. Für jede natürliche Zahl gilt $f/\nu \geq \chi$, also

$$\bar{I}(f)/\nu = \bar{I}(f/\nu) \geq \bar{I}(\chi) \geq 0.$$

es folgt $\bar{I}(\chi) = 0$. □

Weiteres Beispiel einer Nullmenge

Sei U ein Untervektorraum des $\mathbb{R}^n$ mit $\dim U < n$. Dann ist

$$L = x + U \text{ für jedes } x \in \mathbb{R}^n$$

eine Nullmenge.

Übungsaufgabe. Man beweise dies wenigstens für einen achsenparallelen Unterraum, d. h.

$$U = \{(x_1, \cdots, x_n);\ x_k = \cdots = x_n = 0\} \quad (1 \leq k \leq n).$$

(Den allgemeinen Fall kann man auf diesen speziellen mittels der Transformationsformel (§9) zurückführen.)

3.12 Hilfssatz. *Sei f eine Bairesche Funktion mit endlichem Integral. Dann ist die Menge der Unendlichkeitsstellen eine Nullmenge.* HBen

Beweis. Da f Bairesch ist, existiert eine stetige Funktion mit kompaktem Träger g, so dass $f \geq g$ gilt. Man wende Hilfssatz 3.11 auf $f - g$ an. □

3.13 Satz. a) *Jede Teilmenge einer Nullmenge ist eine Nullmenge.* EigNul
b) *Jeder Punkt ist eine Nullmenge.*
c) *Sind $A_1, A_2, A_3, \cdots$ Nullmengen, so ist auch $\bigcup\limits_{\nu=1}^{\infty} A_\nu$ eine Nullmenge.*

Beweis: a) und b) sind klar. c) folgert man leicht aus Hilfssatz 3.9. □

Insbesondere ist $\mathbb{Q}$ in $\mathbb{R}$ eine Nullmenge, womit noch einmal sehr drastisch gezeigt ist, daß es irrationale Zahlen geben muß.

4. Lebesgue integrierbare Funktionen

Wir führen nun die Lebesgue integrierbaren Funktion ein.

4.1 Definition. *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt Lebesgue integrierbar, falls es zu jedem $\varepsilon > 0$ eine stetige Funktion g mit kompaktem Träger gibt, so dass $\bar{I}(|f - g|) < \varepsilon$.* DLi

Stetige Funktionen mit kompaktem Träger sind integrierbar.

4.2 Satz. *Nullfunktionen sind integrierbar* SNi

Dies folgt unmittelbar aus der Definition. Man nehme $g = 0$. □

4.3 Satz und Definition. *Ist f eine Lebesgue integrierbare Funktion, so existiert eine Folge f_ν stetiger Funktionen mit kompakten Träger, so dass $\bar{I}(|f - f_\nu|$ eine Nullfolge ist. Dann existiert* SDLi

$$I_L(f) = \lim_{\nu \to \infty} I(f_\nu)$$

und ist unabhängig von der Folge. Man nennt $I_L(f)$ das Lebesgue-Integral.

Beweis. Man wähle f_n zunächst so, dass $\bar{I}(|f - f_\nu|$ eine Nullfolge ist. Aus der Dreicksungleichung $|f_\mu - f_\nu| \leq |f - f_\nu)| + |f - f_\mu|$ folgt, dass $I(f_\nu)$ eine Cauchyfolge ist. Daher konvergiert $I(f_\nu)$. Wenn man zwei Folgen f_ν, g_ν stetiger Funktionen mit kompaktem Träger mit der Eigenschaft $\bar{I}(|f - f_n| \to 0$, $\bar{I}(|f - g_n| \to 0$, so mische man diese beidem Folgen $f_1, g_1, f_2, \ldots$ und wende den ersten Teil an. □

4.4 Hilfsatz. *Jede Bairesche Funktion f ohne Unendlichkeitsstellen und mit endlichem $I_{\mathcal{B}^+}$ ist Lebesgue integrierbar. Es gilt* HBL

$$I_{\mathcal{B}^+}(f) = I_L(f).$$

Beweis. Nach Definition des Baireschen Integrals existiert zu jedem $\varepsilon > 0$ eine stetige Funktion mit kompaktem Träger g, so dass $g \leq f$ und $0 \leq I(f) - I(g) \leq \varepsilon$. Da $-g$ eine stetige Funktion mit kompaktem Träger und somit Bairesch ist, ist auch $f - g$ Bairesch und es gilt $I(f-g) = I(f) - I(g) \leq \varepsilon$. Fa $f - g \geq 0$ ist, folgt die Integrierbarkeit von f. □

Dieser Hilfssatz erlaubt uns die Schreibweise

$$I_L(f) = I(f) = \int_{\mathbb{R}^n} f(x)dx$$

für Lebesgue integrierbare Funktionen. In diesem und im nächsten Paragraphen wird sich zeigen, daß das Lebesgue-Integral alle Eigenschaften hat, die man von einem „vernünftigen“ Integral erwartet.

Bezeichnung. Menge der integrierbaren Funktionen und Menge der Nullfunktionen

$$\begin{aligned}\mathcal{L}^1 = \mathcal{L}^1(\mathbb{R}^n) &= \{\, f : \mathbb{R}^n \longrightarrow \mathbb{R}; \quad f \text{ integrierbar} \,\},\\ \mathcal{N} = \mathcal{N}(\mathbb{R}^n) &= \{\, f : \mathbb{R}^n \longrightarrow \mathbb{R}; \quad f \text{ Nullfunktion} \,\}.\end{aligned}$$

Es gilt $\mathcal{N} \subset \mathcal{L}^1$.

4.5 Satz. *Die integrierbaren Funktionen aus $\mathcal{L}^1$ bilden einen Vektorraum und das Integral ist ein lineares Funktional, d. h. also:* intVek

$$f, g \in \mathcal{L}^1 \Longrightarrow f + g \in \mathcal{L}^1 \text{ und } cf \in \mathcal{L}^1 \text{ für } c \in \mathbb{R},$$

außerdem

$$\int_{\mathbb{R}^n} (f(x) + g(x))\, dx = \int_{\mathbb{R}^n} f(x)\, dx + \int_{\mathbb{R}^n} g(x)\, dx$$

und

$$\int_{\mathbb{R}^n} cf(x)\, dx = c \int_{\mathbb{R}^n} f(x)\, dx.$$

Außerdem gilt

$$f \geq g \Longrightarrow \int_{\mathbb{R}^n} f(x)dx \geq \int_{\mathbb{R}^n} g(x)dx.$$

Beweis. Der Beweis folgt unmittelbar aus der Definition. □

4.6 Satz. *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist genau dann Lebesgue integrierbar, wenn das äußere und innere Integral von f existieren, endlich sind und wenn $\underline{I}(f) = \bar{I}(f)$ gilt. In diesem Fall gilt* SzwC

$$I_L(f) = \bar{I}(f) = \underline{I}(f).$$

Beweis. Sei f integrierbar. Zu jedem $\varepsilon > 0$ existieren dann eine stetige Funktion mit kompaktem Träger g und eine Bairesche Funktion h mit den Eigenschaften

$$|f - g| \le h, \quad I(h) < \varepsilon.$$

Es folgt dann $f \le h + g$ und $-f \le h - g$. Aus der zweiten Ungleichung folgt $f + h - g \ge 0$. Daher (Hilfssatz 3.6) existiert das äußere Integral von f. Entsprechend folgt aus der ersten Ungleichung, dass das innere Integral existiert. Außerdem erhalten wir $\bar{I}(f) \le I(g + h)$ und $\bar{I}(-f) \le I(h - g)$. Addiert man die beiden Ungleichungen, so folgt $\bar{I}(f) + \bar{I}(-f) < 2\varepsilon$. Dies ist richtig für alle $\varepsilon > 0$, somit auch für $\varepsilon = 0$. Es folgt $\bar{I}(f) \le \underline{I}(f)$. Da die umgekehrte Ungleichung allgemein gilt, folgt die Gleichheit.

Wir beweisen die umgekehrte Richtung. Es sei als f eine Funktion ohne Unendlichkeitsstellen, so dass $\bar{I}(f)$ und $\underline{I}(f)$ existieren, endlich und gleich sind. Es existieren eine Bairesche Funktion h mit $f \le h$ und

$$I(h) - \varepsilon \le I(f) \le I(h)$$

sowie eine stetige Funktion mit kompaktem Träger $g \le h$ und

$$I(g) \le I(h) \le I(g) + \varepsilon.$$

Offenbar gilt

$$|f - g| \le (h - g) + (h - f).$$

(In den Stellen, wo h unendlich ist, gilt sogar Gleichheit.) Es folgt

$$\bar{I}(|f - g|) \le I(h - g) + \bar{I}(h - f).$$

Nun benutzen wir

$$\bar{I}(h - f) \le I(h) + \bar{I}(-f).$$

Aus $\bar{I}(f) = \underline{I}(f)$ folgt $\bar{I}(-f) = -\bar{I}(f)$. Nunmehr erhalten wir

$$\bar{I}(|f - g|) \le (I(h) - I(g)) + (I(h) - \bar{I}(f)) < 2\varepsilon.$$

Dies beweist den Satz. □

4.7 Satz. *Ist $f : \mathbb{R}^n \to \mathbb{R}$ eine integrierbare Funktion, so sind auch ihr Betrag $|f|$ sowie f^+ und f^- integrierbar. Ist $g : \mathbb{R}^n \to \mathbb{R}$ eine weitere integrierbare Funktion, so sind auch die Maxima und Minima $f \vee g$ und $f \wedge g$ integrierbar.* BeAIn

Wir erinnern an die Bezeichnungen

$$f^+(x) = \max(f(x), 0), \quad f^-(x) = -\min(f(x), 0).$$

Beweis von Satz 4.7. Es genügt zu zeigen, dass mit f auch $|f|$ integrierbar ist, da man die anderen Bildungen aus f und $|f|$ linear kombinieren kann. Die Behauptung für $|f|$ folgt aus der verschärften Dreiecksungleichung $||f| - |g|| \le |f - g|$. □

4.8 Satz. *Sei f eine integrierbare Funktion und $g : \mathbb{R}^n \to \mathbb{R}$ eine Funktion, welche nur auf einer Nullmenge von f verschieden ist. Dann ist auch g integrierbar und die Integrale von f und g stimmen überein.* UntNul

Beweis. Es gilt $I(f) - I(g) = I(f - g) = 0$. □

5. Die Grenzwertsätze

Die Stärke des Lebesgue Integrals zeigt sich in den beiden Grenzwertsätzen.

5.1 Theorem (Beppo Levi). *Gegeben sei eine monoton wachsende Folge* BepLe

$$f_1 \le f_2 \le f_3 \le \cdots$$

von integrierbaren Funktionen. Die Folge der Integrale

$$\int\limits_{\mathbb{R}^n} f_\nu(x)\,dx$$

sei beschränkt. Dann existiert eine Nullmenge S, so dass $f_\nu(x)$ außerhalb S konvergiert. Die durch

$$f(x) = \begin{cases} \lim_{\nu\to\infty} f_\nu(x) & x \notin S, \\ 0 & \textit{else} \end{cases}$$

definierte Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist integrierbar und es gilt

$$\int\limits_{\mathbb{R}^n} f(x)\,dx = \lim_{\nu\to\infty} \int\limits_{\mathbb{R}^n} f_\nu(x)\,dx.$$

Beweis. Da man die Folge f_ν durch $f_\nu - f_1$ ersetzen kann, ohne die Aussage des Satzes zu verändern, kann man ohne Beschränkung der Allgemeinheit $f_1 \geq 0$ annehmen. Sei $\varepsilon > 0$. Wir setzen

$$f_0 = 0 \quad \text{und} \quad g_\nu = f_\nu - f_{\nu-1} \qquad (\nu \geq 1).$$

Dies ist eine integrierbare und nirgends negative Funktion. Es existiert eine Bairesche Funktion h_ν mit der Eigenschaft $g_\nu \leq h_\nu$ sowie

$$I(h_\nu) \leq I(g_\nu) + \frac{\varepsilon}{2^\nu}.$$

Es gilt $g_1 + \cdots + g_k = f_k$ und somit

$$I(f_k) \leq I(h_1 + \cdots + h_k) \leq I(f_k) + \varepsilon \sum_1^k 2^{-\nu} = I(f_k) + \varepsilon.$$

Da Bairesches Integral mit monotoner Konvergenz verträglich ist, ist

$$\mathrm{Sup}_k(h_1 + \cdots + h_k)$$

Bairesch und hat endliches Bairsches Integral. Daher ist die Menge ihrer Unendlichkeitsstellen eine Nullmenge. Folgedessen konvergiert f_ν punktweise außerhalb einer Nullmenge. Wir änderen nun die Funktionen f_ν auf dieser Nullmenge ab, indem wir sie dort Null setzen. Diese Konstruktion zeigt, dass wir ohne Beschränkung der Allgemeinheit annehmen können, dass die Folge $f_\nu(x)$ für alle x konvergiert.

Aus der Ungleichung

$$0 \leq f(x) \leq \mathrm{Sup}_k(h_1 + \cdots + h_k)$$

folgt, dass inneres und äußeres Intgegral für f existieren und endlich sind. Außerdem folgt aus dieser Formel

$$\bar{I}(f) \leq I(\mathrm{Sup}_k(h_1 + \cdots + h_k)).$$

Schließlich gilt

$$I(\mathrm{Sup}_k(h_1 + \cdots + h_k)) \leq \mathrm{Sup}_k(I(f_k)) + \varepsilon \leq \underline{I}(f) + \varepsilon.$$

Beide Ungleichungen zusammen ergeben $\bar{I}(f) \leq \underline{I}(f) + \varepsilon$. Da dies für alle $\varepsilon > 0$ gilt, folgt $\bar{I}(f) \leq \underline{I}(f)$ und somit die Gleichheit. Wir haben gezeigt, dass f integrierbar ist.

Die Gleicheit

$$\int\limits_{\mathbb{R}^n} f(x)dx = \lim_{\nu\to\infty} \int\limits_{\mathbb{R}^n} f_\nu(x)dx$$

ergibt sich ebenfalls aus den obigen Ungleichungen. □

5.2 Theorem (Lebesgue'scher Grenzwertsatz). *Gegeben sei eine Folge* LGSa

$$f_k : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad k = 0, 1, 2, \dots,$$

von integrierbaren Funktionen, die ***punktweise*** *gegen eine Funktion* $f : \mathbb{R}^n \to \mathbb{R}$ *konvergiert.*

Es existiere eine Funktion $h : \mathbb{R}^n \to \mathbb{R}$ *mit den Eigenschaften*

a) $|f_k| \le h$ *für* $k = 1, 2, 3, \cdots$

b) $\bar{I}(h) < \infty$.

Dann ist auch f *integrierbar und es gilt*

$$\int_{\mathbb{R}^n} f(x)\,dx = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,dx.$$

Beweis. Aus den Ungleichungen $|f_k| \le h$ folgt $|f| \le h$. Daher ist das äußere Integral von $|f|$ endlich, $\bar{I}(|f|) < \infty$. Wir wollen den Lebesgue'schen Grenzwertsatz auf den Satz von Beppo Levi zurückführen und bilden hierzu

$$\underline{g}_k(x) = \operatorname{Inf}\{f_\nu(x),\ \nu \ge k\},$$
$$\overline{g}_k(x) = \operatorname{Sup}\{f_\nu(x),\ \nu \ge k\}.$$

Dies sind Funktionen ohne Unendlichkeitsstellen. Es gilt offenbar

$$\underline{g}_1 \le \underline{g}_2 \le \underline{g}_3 \le \cdots,$$
$$\overline{g}_1 \ge \overline{g}_2 \ge \overline{g}_3 \ge \cdots.$$

Aus $f = \lim_{k\to\infty} f_k$ folgert man leicht

$$\underline{g}_k \uparrow f, \quad \overline{g}_k \downarrow f.$$

Als nächstes wird gezeigt, daß die $\underline{g}_k$ und $\overline{g}_k$ integrierbare Funktionen sind. Dazu wird der Satz von Beppo Levi ausgenutzt. Bildet man nämlich

$$\underline{G}_{kj} = f_k \wedge f_{k+1} \wedge \cdots \wedge f_{k+j}$$

und

$$\overline{G}_{kj} = f_k \vee f_{k+1} \vee \cdots \vee f_{k+j}$$

so sind $\underline{G}_{kj}$ und $\overline{G}_{kj}$ integrierbar (Satz 4.7) und es gelten die Ungleichungen

$$-h \le \underline{G}_{kj} \le f_k \le \overline{G}_{kj} \le h.$$

Ferner gilt offensichtlich

$$\underline{G}_{kj} \downarrow \underline{g}_k \quad (j \longrightarrow \infty)$$

und

$$\overline{G}_{kj} \uparrow \overline{g}_k \quad (j \longrightarrow \infty).$$

Daher sind nach dem Satz von Beppo Levi $\underline{g}_k$ und $\overline{g}_k$ integrierbar. Wiederum aus dem Satz von Beppo Levi folgt, dass f integrierbar ist und dass die Integrale von $\underline{g}_k$ und $\overline{g}_k$ gegen das Integral von f konvergieren. Die Behauptung des Theorems folgt nun aus

$$\underline{g}_k \leq f_k \leq \overline{g}_k.$$

□

Bezeichnung. Sei f eine Funktion, deren Definitionsbereich die Menge $D \subset \mathbb{R}^n$ umfasse. Wir definieren

$$\chi_D \cdot f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

durch

$$\chi_D \cdot f(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D. \end{cases}$$

Die Funktion f heißt über D integrierbar, wenn $\chi_D \cdot f$ integrierbar ist, und man definiert

$$\int\limits_D f(x)\,dx := \int\limits_{\mathbb{R}^n} \chi_D \cdot f(x)\,dx.$$

Parameterabhängige Integrale

5.3 Satz. *Sei Y ein metrischer Raum und $f : \mathbb{R}^n \times Y \to \mathbb{R}$ eine Funktion mit folgenden Eigenschaften.* SPIs

a) *$f(x, y)$ sei für festes y integrierbar (als Funktion von x).*
b) *$f(x, y)$ sei für festes x stetig (als Funktion von y).*
c) *Es existiere eine Funktion $h : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ mit endlichem äußeren Integral und mit der Eigenschaft $|f(x, y)| \leq h(x)$ für alle x, y.*

Dann ist die Funktion $\int_{\mathbb{R}^n} f(x, y)dx$ stetig (auf Y).

Beweis. Es ist folgendes zu zeigen. Sei y_ν eine Folge, welche gegen y konvergiert, dann gilt

$$\int\limits_{\mathbb{R}^n} f(x, y_n)dx \longrightarrow \int\limits_{\mathbb{R}^n} f(x, y)dx.$$

Dies folgt aus dem Lebesgueschen Grenzwertsatz.

5.4 Satz. *Sei $I \subset \mathbb{R}$ ein Intervall und $f : \mathbb{R}^n \times I \to \mathbb{R}$ eine Funktion mit folgenden Eigenschaften.* SPds

a) *$f(x, y)$ sei für festes y integrierbar (als Funktion von x).*
b) *$f(x, y)$ sei für festes x differenzierbar (als Funktion von y).*
c) *Es existiere eine Funktion $h : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ mit endlichem äußeren Integral und mit der Eigenschaft*

$$\left|\frac{\partial}{\partial y} f(x, y)\right| \leq h(x) \text{ für alle } x, y.$$

Dann ist die Funktion $\int_{\mathbb{R}^n} f(x, y)dx$ differenzierbar und es gilt

$$\frac{d}{dy} \int\limits_{\mathbb{R}^n} f(x, y)dx = \int\limits_{\mathbb{R}^n} \frac{\partial}{\partial y} f(x, y)dx.$$

Beweis. Sei $y_\nu \to y$ eine konvergente Folge in I. Wir müssen folgendes zeigen.

$$\frac{\int f(x, y_\nu)dx - \int f(x, y)}{y_\nu - y} \longrightarrow \int \frac{\partial}{\partial y} f(x, y)dx.$$

Nach dem Mittelwertsatz der Differentialrechnung gilt

$$\frac{f(x, y_\nu)dx - f(x, y)}{y_\nu - y} = \frac{\partial}{\partial y} f(x, \xi_\nu).$$

Hierbei ist ξ_ν eine (von x abhängige) Zwischenstelle zwischen y und y_ν. Die Behauptung folgt nun aus dem Lebesgueschen Grenzwertsatz. □

Dieser Satz ist eine allgemeine Version der Leibnizschen Regel für eigentliche Integrale. Wir weisen darauf hin, dass wir bislang eine solche Regel für uneigentliche Integrale beweisen konnten.

6. Integrierbarkeitskriterien

Eine Teilmenge $A \subset \mathbb{R}^n$ heit *endlich meßbar*, wenn die charakteristische Funktion

$$\chi_A = \begin{cases} 1 & \text{für } x \in A \\ 0 & \text{sonst} \end{cases}$$

integrierbar ist und man nennt

$$v(A) \stackrel{\text{def}}{=} \int\limits_{\mathbb{R}^n} \chi_A(x)\, dx$$

das (Euklidische) Volumen von A.

6.1 Satz. *Die charakteristische Funktion* deMes

$$\chi_U : \mathbb{R}^n \to \mathbb{R}$$

einer offenen Teilmenge $U \subset \mathbb{R}^n$ gehört der Klasse $\mathcal{B}^+$ an.

Beweis, 1. Schritt. U ist ein offener Quader

$$U = \{x \in \mathbb{R}^n;\ a_\nu < x_\nu < b_\nu\} \quad (a_\nu \leq b_\nu).$$

Im Fall $n = 1$ approximiert man die charakteristische Funktion in naheliegender Weise von unten durch Trapeze. Im Fall $n > 1$ verfährt man ähnlich.

2. Schritt. U ist beliebig. Zunächst zeigen wir, daß U abzählbare Vereinigungen von offenen Quadern ist.

$$U = U_1 \cup U_2 \cup U_3 \cup \ldots, \quad U_\nu \text{ offene Quader.}$$

Man betrachte hierzu die Menge aller offenen Quader

$$Q \subset U; \quad Q = \{x; \quad a_\nu < x_\nu < b_\nu,\ 1 \leq \nu \leq n\},$$

wobei die Zahlen $a_1, \cdots, a_n,\ b_1, \cdots, b_n$ rational sind. Es ist klar, daß die Menge dieser Quader U überdeckt. Außerdem ist diese Menge abzählbar, weil die rationalen Zahlen und damit auch die n-Tupel von rationalen Zahlen abzählbar sind.

Man hat jetzt eine monotone Approximation von χ_U:

$$\chi_{U_1},\ \chi_{U_1 \cup U_2},\ \chi_{U_1 \cup U_2 \cup U_3}, \ldots. \qquad \square$$

6.2 Satz. *Jede stetige und beschränkte Funktion $f : D \to \mathbb{R}$ auf einem offenen und beschränkten Teil $D \subset \mathbb{R}^n$ ist integrierbar.* StBeIn

Beweis: Man kann annehmen, daß f nirgends negativ ist. Die Funktion

$$f \cdot \chi_D(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

gehört dann sogar der Klasse $\mathcal{B}^+(\mathbb{R}^n)$ an. Sie wird approximiert durch die Folge

$$f_k \cdot f \cdot \chi_D,$$

wobei $f_k \in C_c$ eine Folge ist, die χ_D monoton approximiert.

6.3 Folgerung. EnMes

a) *Jede beschränkte offene Teilmenge des* $\mathbb{R}^n$ *ist endlich meßbar.*
b) *Jede kompakte Teilmenge des* $\mathbb{R}^n$ *ist endlich meßbar.*

Beweis: Es ist nur noch b) zu beweisen.

Sei jetzt $K \subset \mathbb{R}^n$ kompakt. Wir werden eine Folge von offenen beschränkten Mengen

$$U_1 \supset U_2 \supset U_3 \cdots \supset K$$

konstruieren, so daß

$$\bigcap_{\nu=1}^{\infty} U_\nu = K$$

gilt.

Dann können wir wieder den Grenzwertsatz (auf die Folge χ_{U_ν}) anwenden. Wir setzen

$$U_\nu = \left\{ x \in \mathbb{R}^n; \quad \|x - y\| < \frac{1}{\nu} \text{ für mindestens ein } y \in K \right\}.$$

Es sei wieder dem Leser überlassen, die gewünschten Eigenschaften zu beweisen. □

Das Lebesgue'sche Integral ist eine Verallgemeinerung des Regelintegrals

Sei

$$f : [a, b] \to \mathbb{R}$$

eine Regelfunktion. Dann ist f auch Lebesgue-integrierbar und das Regel- und Lebesgue-Integral stimmen überein. Dies ist für Treppenfunktionen einfach zu zeigen und folgt dann allgemein aus den Grenzwertsätzen.

Allgemeiner gilt:

Sei

$$f : D \longrightarrow \mathbb{R}, \quad D \subset \mathbb{R} \text{ ein Intervall},$$

eine Funktion auf einem nicht notwendigerweise geschlossenen Intervall. Die Einschränkung von f auf jedes geschlossene Intervall sei eine Regelfunktion.

Die Funktion f ist genau dann Lebesgue-integrierbar, wenn $|f|$ uneigentlich integrierbar ist im Sinne von III, §1.

Das uneigentliche Integral und das Lebesgue-Integral stimmen dann überein.

Der Beweis ergibt sich leicht aus dem Lebesgueschen Grenzwertsatz.

Die wichtigsten Integrierbarkeitskriterien liegen in den Grenzwertsätzen. Es gibt aber auch eine einfache direkte Charakterisierung der Integrierbarkeit, welche unabhängig von ihren Anwendungen von eigenem theoretischem Interesse ist.

7. Der Satz von Fubini

In diesem Paragraphen soll Satz 1.3 über die Vertauschbarkeit der Integrationsreihenfolge auf integrierbare Funktionen verallgemeinert werden. Gegeben sei eine integrierbare Funktion

$$(x, y) \mapsto f(x, y) = f(x_1, \dots, x_n, y_1, \dots, y_m)$$

in $n + m$ Veränderlichen. Wir setzen $z = (z, y)$ und bezeichnen die Volumenelemente im

$$\mathbb{R}^{n+m} \text{ mit } dz, \quad \text{im } \mathbb{R}^n \text{ mit } dx \text{ und im } \mathbb{R}^m \text{ mit } dy.$$

Es soll eine Formel der Art

$$\int\limits_{\mathbb{R}^{n+m}} f(z)dz = \int\limits_{\mathbb{R}^m} \left[\int\limits_{\mathbb{R}^n} f(x, y)dx \right] dy = \int\limits_{\mathbb{R}^n} \left[\int\limits_{\mathbb{R}^m} f(x, y)\, dy \right] dx$$

bewiesen werden. Wir erinnern daran, dass sie für stetige Funktionen mit kompaktem Träger richtig ist.

Wir zeigen nun, dass sie auch für beliebige Bairesche Funktion f gilt. Es ist klar, dass die Funktionen

$$x \longmapsto f(x, y) \text{ bei festem } y \quad \text{und} \quad y \longmapsto f(x, y) \text{ bei festem } x$$

Bairesch sind. Inbesondere existieren das Integral

$$\int\limits_{\mathbb{R}^n} f(x, y)dx.$$

Es ist klar, dass auch diese Funktion Bairesch ist. Wir behaupten nun

$$\int\limits_{\mathbb{R}^{n+m}} f(z)dz = \int\limits_{\mathbb{R}^m} \int\limits_{\mathbb{R}^n} f(x, y)dxdy$$

Für stetige Funktionen mit kompaktem Träger haben wir diese Formel bewiesen. Sie folgt dann unmittelbar für Bairesche Funktionen. Da man die Rollen von x und y vertauschen kann, ist die Behauptung bewiesen.

Wir haben folgendes mitbewiesen. Ist $f(z)$ eine Funktion mit endlichem Baireschen Integral, so ist das Bairesche Integral $\int f(x, y)$ für alle y außerhalb einer Nullmenge endlich.

Im allgemeinen hat man das Problem, dass $f(x,y)$ bei festgehaltenem y nicht immer integrierbar zu sein braucht, wie das folgende Beispiel zeigt. Sei

$$f(x,y) = \begin{cases} 1 & \text{für } y = 0, \\ 0 & \text{else.} \end{cases}$$

Dies ist eine Nullfunktion und somit integrierbar. Aber für $y = 0$ ist sie nicht integrierbar. Daher bedarf der Satz von Fubini einer sorgfältigen Formulierung. Wir machen eine kleine Vorbemerkung, die an dieses Beispiel anknüpft. Sei $A \subset \mathbb{R}^n$ eine Nullmenge. Wir behaupten, dass dann $A \times \mathbb{R}^m$ eine Nullmenge im $\mathbb{R}^{n+m}$ ist. Zum Beweis wählen wir zu einem vorgegeben $\varepsilon > 0$ eine Bairesche Funktion $h(x)$ mit den Eigenschaften

$$\chi_A \le h, \qquad \int\limits_{\mathbb{R}^n} < \varepsilon$$

an. Die Behauptung folgt dann aus dem Satz von Fubini für die Bairesche Funktion $f(x,y) = h(x)$.

Zu einer prägnanten Formulierung des Satzes von Fubini treffen wir folgende Vereinbarung. Sei $S \subset \mathbb{R}^n$ eine Nullmenge und $f : \mathbb{R}^n - S \to \mathbb{R}$ eine Funktion, welche außerhalb f definiert ist. Wir nennen f integrierbar, falls die Funktion

$$\tilde{f}(x) = \begin{cases} f(x) & \text{falls } x \notin S, \\ 0 & \text{else} \end{cases}$$

integrierbar ist und wir schreiben dann

$$\int\limits_{\mathbb{R}^n} f(x)dx = \int\limits_{\mathbb{R}^n} \tilde{f}(x)dx.$$

7.1 Theorem (Fubini). *Gegeben sei eine integrierbare Funktion* SaFub

$$f : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}, \quad (x,y) \longmapsto f(x,y).$$

Dann existieren Nullmengen $S \subset \mathbb{R}^n$ und $T \subset \mathbb{R}^n$, so dass die beiden Funktionen

$$y \longmapsto f(x,y) \text{ für } x \notin S, \quad x \longmapsto f(x,y) \text{ für } y \notin T$$

integrierbar sind. Die Funktionen

$$x \longmapsto \int\limits_{\mathbb{R}^m} f(x,y)dy \ (x \notin S), \quad y \longmapsto \int\limits_{\mathbb{R}^n} f(x,y)dy \ (y \notin T)$$

sind ebenfalls integrierbar und es gilt die Formel

$$\int\limits_{\mathbb{R}^{n+m}} f(z)dz = \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^m} f(x,y)dydx = \int\limits_{\mathbb{R}^m} \int\limits_{\mathbb{R}^n} f(x,y)dxdy.$$

Beweis. Da f integrierbar ist, existiert eine Bairesche Funktion $h \geq |f|$ mit endlichem Integral. Wir wissen, dass $h(x, y)$ auch bei festem y Bairesch in x ist. Nach dem Satz von Fubini für Bairesche Funktionen ist die Funktionen $\int h(x, y)dx$ Bairesch und hat endliches Bairesches Integral. Sie ist nach Hilfssatz 3.12 endlich außerhalb einer Nullmenge $B \subset \mathbb{R}^m$. Aus Hilfssatz 3.6 folgt, dass die äußeren Integrale

$$\overline{\int} f(x, y)dx, \quad \overline{\int}(-f(x, y))dx,$$

für $y \notin B$ existieren. Sie sind endlich nach 3.5. Daher existiert auch das innere Integral

$$\underline{\int} f(x, y)dx$$

Wir können die Funktion f auf der Nullmenge $\mathbb{R}^n \times B$ abändern und Null einsetzen ohne die Aussage des Theorems zu verändern. Daher können wir annehmen, dass

$$\overline{\int} f(x, y)dx, \quad \underline{\int} f(x, y)dx$$

für alle $y \in \mathbb{R}^m$ existieren und endlich sind.

Als nächstes zeigen wir, dass das äußere Integral der Funktion

$$y \longmapsto \overline{\int} f(x, y)dx$$

existiert. Dazu nutzen wir $-f \leq h$ aus. Somit existiert eine Bairesche Funktion h mit endlichem Integral und $-f \leq h$ oder $f + h \geq 0$. Aus Hilfssatz 3.4 folgt

$$0 \leq \overline{\int}(f(x, y) + h(x, y))dx \leq \overline{\int} f(x, y)dx + \int h(x, y)dx.$$

Da die Funktion $\int h(x, y)dx$ Bairesch ist und endliches Integral hat, folgt die Existenz von

$$\overline{\int}\,\overline{\int} f(x, y)dxdy.$$

Aus der Monotonie des äußeren Integrals folgt

$$\overline{\int}\,\overline{\int} f(x, y)dxdy \leq \int\int h(x, y)dxdy = \int h(z)dz.$$

Bildet man das Infimum über alle h, so folgt

$$\overline{\int}\,\overline{\int} f(x,y)dxdy \le \int f(z)dz.$$

Nun können dieses Verfahren auch für $-f$ anstelle f anwenden und erhalten die Schlüsselungleichung

$$\overline{\int}\,\overline{\int} f(x,y)dxdy \le \int f(z)dz \le \underline{\int}\,\underline{\int} f(x,y)dxdy.$$

Da

$$\overline{\int} f(x,y)dx \le \int h(x,y)dx$$

durch eine Bairesche Funktion mit endlichem Integral nach oben beschränkt ist, existiert auch

$$\underline{\int}\,\overline{\int} f(x,y)dxdy.$$

Es gilt

$$\overline{\int}\,\overline{\int} f(x,y)dxdy \ge \underline{\int}\,\overline{\int} f(x,y)dxdy \ge \underline{\int}\,\underline{\int} f(x,y)dxdy.$$

Da wir die umgekehrte Ungleichung bewiesen haben, muss das Gleichheitszeichen gelten. Die beiden Funktionen

$$\overline{\int} f(x,y)dx, \quad \underline{\int} f(x,y)dx$$

haben dasselbe innere und äußere Integral und sind somit integrierbar. Ihre Differenz ist nirgends negativ und hat Integral 0. Es existiert eine Nullmenge $A \subset \mathbb{R}^n$, so dass die beiden Integrale außerhalb A übereinstimmen. Wir dürfen f auf der Nullmenge $A \times \mathbb{R}^m$ abändern und somit annehmen, dass beide Integrale gleich sind. Dann ist aber $f(x,y)$ bei festem y integrierbar und es folgt

$$\int f(z)dz = \int \int f(x,y)dxdy.$$

Da man die Rollen von x und y vertauschen kann, folgt das Theorem von Fubini. □

Eine Anwendung:

7.2 Satz. *Sei* IntVol

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f \geq 0,$$

eine integrierbare Funktion, die keine negativen Werte annimmt.

Sei

$$M = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}; \quad 0 \leq y \leq f(x)\}.$$

Dann ist M *endlich meßbar und es gilt*

$$\operatorname{vol}(M) = \int_{\mathbb{R}^n} f(x)\, dx.$$

Beweis. Das Volumen ist definiert durch die Formel

$$\operatorname{vol}(M) = \int_{\mathbb{R}^{n+1}} \chi_M(x, y)\, dz$$

Man integriere zunächst bei festem x über y und wende den Satz von Fubini an.

8. Die Transformationsformel

Wir erinnern uns daran, daß eine Funktion

$$f : D \longrightarrow \mathbb{R}, \quad D \subset \mathbb{R}^n,$$

über D integrierbar heißt, wenn die Funktion

$$\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \tilde{f}(x) = \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{für } x \notin D, \end{cases}$$

integrierbar ist, und wir setzen dann

$$\int_D f(x)\, dx = \int_{\mathbb{R}^n} \tilde{f}(x)\, dx.$$

Seien $A, B \subset \mathbb{R}^n$ offene Mengen. Unter einem *Diffeomorphismus*

$$\varphi : A \longrightarrow B$$

verstehen wir eine bijektive (=umkehrbare) stetig differenzierbare Abbildung mit nirgends verschwindender Funktionaldeterminante

$$j(\varphi, x) = \det J(\varphi; x) \neq 0 \quad \text{für alle } x \in A.$$

Nach dem Satz für implizite Funktionen ist dann auch φ^{-1} differenzierbar und es gilt

$$J(\varphi; x)^{-1} = J\left(\varphi^{-1}, \varphi(x)\right).$$

8.1 Theorem. *Es sei $u : A \to B$ ein Diffeomorphismus zwischen offenen Mengen $A, B \subset \mathbb{R}^n$. Ist $f : B \to \mathbb{R}$ eine integrierbare Funktion, so gilt* TraFoL

$$\int_B f(y)\,dy = \int_A f(u(x))\,|j(u,x)|\,dx.$$

(Insbesondere wird also behauptet, daß $f(u(x))\,|j(u,x)|$ über A integrierbar ist.)

Beweis der Transformationsformel, 1. Schritt. Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion mit kompaktem Träger. Dann gelten die Transformationsformeln

a) $$\int_{\mathbb{R}^n} f(x_1, \dots, x_{i-1}, x_i + x_j, x_{i+1}, \dots, x_n)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$$

b) $$\int_{\mathbb{R}^n} f(a_1x_1, \dots, a_nx_n)\,dx = |a_1 \cdots a_n| \int_{\mathbb{R}^n} f(x)\,dx$$

c) $$\int_{\mathbb{R}^n} f(x_1, \dots, x_n)\,dx_{\sigma(1)} \dots dx_{\sigma(n)} = \int_{\mathbb{R}^n} f(x)\,dx_1 \dots dx_n$$

d) $$\int_{\mathbb{R}^n} f(x+b)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$$

Dabei sei $b = (b_1, \dots, b_n)$ ein festes n-Tupel.

Beweis: Man benutzt die Transformationsformel im Fall $n = 1$, sowie die Tatsache, daß das Integral für stetige Funktionen mit kompaktem Träger iterativ definiert ist, beispielsweise

$$\int_{\mathbb{R}^2} f(x_1 + x_2, x_2)\,dx = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x_1 + x_2, x_2)\,dx_1 \right] dx_2.$$

Im inneren Integral macht man die Substitution $t = x_1 + x_2$ und erhält für das innere Integral

$$\int_{\mathbb{R}} f(x_1, x_2)\,dx_1.$$

2. Schritt. Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig mit kompaktem Träger und

$$A = (a_{\nu\mu}) \quad 1 \le \nu, \mu \le n$$

eine feste Matrix mit von Null verschiedener Determinante, $b \in \mathbb{R}^n$ ein fester Vektor. Dann ist

$$\int\limits_{\mathbb{R}^n} f(y)\,dy = \int\limits_{\mathbb{R}^n} f(Ax+b)\,|\det A|\;dx.$$

Beweis: Man kann $b = 0$ annehmen (1. Schritt d)). Im Spezialfall, daß die Abbildung A zu den drei Typen

a) Scherung
b) $(x_1, \dots, x_n) \to (a_1 x_1, \dots, a_n x_n)$
c) Permutation der Variablen

gehört, haben wir das im 1. Schritt erkannt.

Man muß jetzt nur aus der linearen Algebra wissen, daß sich jede lineare Abbildung A, $\det A \neq 0$, als Hintereinanderausführung von Transformationen des Typs a) – c) schreiben läßt. Dies ist nichts anderes als die Tatsache, daß sich jede Matrix A, $\det A \neq 0$, durch elementare Umformungen in die Einheitsmatrix überführen läßt.

Außerdem muß man benutzen, daß die Determinante multiplikativ ist, $\det(A_1 \cdots A_n) = \det A_1 \cdots \det A_n$.

3. Schritt. Sei $f : \mathbb{R}^n \to \mathbb{R}$ beliebig und $A = (a_{\nu,\mu})_{1\le\nu,\mu\le n}$ eine Matrix mit $\det A \neq 0$, $b \in \mathbb{R}^n$. Dann ist

$$\overline{\int} f(y)\,dy = \overline{\int} f(Ax+b)\,|\det A|\;dx.$$

Beweis: Ist $f \in \mathcal{B}^+$ und $f_k \uparrow f$ eine monotone Approximation durch Funktionen $f_k \in C_c$, so sind die Funktionen

$$g_k(x) = f_k(Ax+b)$$

ebenfalls stetig mit kompaktem Träger und approximieren monoton $g(x) = f(Ax+b)$. Damit ist die Behauptung für Funktionen $f \in \mathcal{B}^+$ zurückgeführt auf den 2. Schritt. Ist f beliebig, so folgt die Behauptung unmittelbar aus der Definition des äußeren Integrals.

Die bisherigen Überlegungen zeigen:

1) Ist f integrierbar, so gilt

$$\int\limits_{\mathbb{R}^n} f(y)\,dy = \int\limits_{\mathbb{R}^n} f(Ax+b)\,|\det A|\;dx.$$

2) Jede Teilmenge $W_0 \subset \mathbb{R}^n$ der Form

$$W_0 = a + W, \;\; W \subset \mathbb{R}^n \;\text{ ein Untervektorraum, } \dim W < n,$$

ist eine Nullmenge.

(Mit Hilfe einer linearen Transformation macht man W achsenparallel.)

4. Schritt: (A, B, u wie in 8.1.) Die Funktion $f : \mathbb{R}^n \to \mathbb{R}$ sei stetig. Sei $a \in A$ fest, $b = u(a)$. Es gelte $f(b) > 0$. Außerdem sei $q > 1$ beliebig, aber fest gewählt. Dann existiert eine Umgebung $a \in U \subset A$,

$$U = U_{\mathbb{R}}(a) = \{x;\; \|x - a\| < R\} \quad \text{(hierbei sei } \|\cdot\| \text{ die Maximumsnorm)},$$

so daß für jeden Würfel

$$W = U_r(x_0) \subset U$$

gilt:

$$q^{n+4} \int\limits_{u(W)} f(y)\, dy \geq \int\limits_{W} f(ux)\, |j(u,x)|\, dx.$$

(Es wird nicht gefordert, daß W den gleichen Mittelpunkt wie U hat!)

Beweis. Die Ungleichung wird auf den linearen Fall (3. Schritt) zurückgeführt, indem man $u(x)$ linear approximiert.

$$u(x) = u(a) + J(u,a)(x - a) + r(x).$$

Wir ersetzen u durch

$$u_0(x) = u(a) + J(u;a)(x - a)$$

und erhalten nach dem 3. Schritt

$$\int\limits_{u_0(W)} f(y)\, dy = \int\limits_{W} f(u_0 x)\, |j(u,a)|\; dx.$$

Dabei sei r so klein gewählt, daß der abgeschlossene Würfel W noch in A enthalten ist, dann ist $f(u_0 x)$ in W beschränkt und damit wegen 6.2 über W integrierbar.

Wir denken uns $\mathbb{R}$ immer so klein gewählt, daß

$$f(u_0 x) > 0 \text{ für } x \in U$$

gilt. (Beachte $u_0(a) = u(a) = b$ und $f(b) > 0$ nach Voraussetzung.)

In einer solchen Umgebung kann man

$$\frac{f(ux)}{f(u_0x)}$$

betrachten. Diese Funktion ist stetig und konvergiert gegen 1, wenn x nach a strebt. Daher gilt

$$q > \frac{f(ux)}{f(u_0x)} \quad \text{für } x \in U, \quad r \text{ genügend klein.}$$

Aus demselben Grund gilt

$$q\,|j(u,a)| \geq |j(u,x)| \quad \text{für } x \in U,$$

wenn man $\mathbb{R}$ genügend klein wählt.

Also gilt

$$q^2 \int\limits_{u_0(W)} f(y)\,dy \geq \int\limits_{W} f(ux)\,|j(u,x)|\,dx.$$

In dieser Ungleichung müßte $u(W)$ anstelle von $u_0(W)$ stehen. Wir werden daher

$$u(W) \quad \text{und} \quad u_0(W)$$

vergleichen.

Behauptung. Wählt man $\mathbb{R}$ genügend klein, so gilt $u_0(W^*) \subset u(W)$. Dabei sei

$$W^* = \left\{x; \quad \|x - x^0\| < rq^{-1}\right\}$$

der Würfel, der aus W durch Schrumpfung um den Faktor q^{-1} entsteht.

Beweis. Dies bedeutet nichts anderes als

$$v(W^*) \subset W \text{ mit } v = u^{-1}u_0,$$

also

$$\|x - x^0\| < rq^{-1} \Rightarrow \|v(x) - v(x^0)\| < r.$$

Wir schätzen $v(x) - v(x^0)$ nach dem verallgemeinerten Mittelwertsatz der Differentialrechnung ab:

$$|v_\nu(x) - v_\nu(x^0)| \leq \|x - x^0\| \sum_{\mu=1}^{n} \left|\partial_\mu v_\nu \left(\xi^{(\nu)}\right)\right|.$$

Dabei ist $\xi^{(\nu)}$ ein Punkt auf der Verbindungsstrecke zwischen x und x^0. Hieraus folgt

$$\|v(x) - v(x^0)\| \leq \|x - x^0\| < M$$

mit

$$M = \sup_{\xi \in U} \sum_{\mu=1}^{n} |\partial_\mu v_\nu(\xi)| \,.$$

Wir wollen ja U so bestimmen, daß gilt:

$$\|v(x) - v(x^0)\| < r \quad \text{falls } \|x - x^0\| < rq^{-1}.$$

Dazu benötigt man ersichtlich die Ungleichung $M \leq q$, d. h.

$$\sum_{\mu=1}^{n} |\partial_\mu v_\nu(\xi)| < q \quad \text{für alle } \xi \in U.$$

Nun beachte man, daß nach der Kettenregel für $v = u^{-1} \cdot u_0$ gilt:

$$j(v, x) = \text{ Einheitsmatrix},$$

d. h. obige Ungleichung ist im Punkt $\xi = a$ erfüllt (wegen $1 < q$). Sie gilt dann aus Stetigkeitsgründen auch in einer vollen Umgebung von a.

Damit erhalten wir nun die Ungleichung

$$q^2 \int\limits_{u(W)} f(y)\, dy \geq \int\limits_{W^*} f(u(x))\, |j(u, x)|\, dx.$$

Die Abbildung u_0 ist damit eliminiert, wir müssen allerdings noch die Integrale

$$\int\limits_{W^*} g(x)\, dx \quad \text{und} \quad \int\limits_{W} g(x)\, dx \quad \text{mit } g(x) = f(ux)\, |j(u, x)|$$

vergleichen.

Wählt man R genügend klein, so gilt

$$g(a) \cdot q > g(x) > g(a) q^{-1} \quad \text{für } x \in U \quad (\text{beachte } q > 1 \text{ und } g(a) > 0).$$

Hieraus folgt

$$\int\limits_{W^*} g(x)\, dx \geq g(a) q^{-1} = \text{vol}\, (W_{rq^{-1}}) = g(a) q^{-1} \cdot q^{-n} (2r)^n;$$

andererseits ist

$$\int\limits_{W} g(x)\, dx \leq g(a) \cdot q \cdot \text{vol}\, (W) = g(a) q (2r)^n.$$

Aus den beiden Ungleichungen folgt

$$q^{n+2} \int\limits_{W^*} g(x)\,dx \geq \int\limits_{W} g(x)\,dx.$$

Wir erhalten die gewünschte Ungleichung

$$q^{n+4} \int\limits_{u(W)} f(y)\,dy \geq \int\limits_{W} f(ux)\,|j(u,x)|\,dx$$

für $W \subset U = U_R(a),\ R$ genügend klein.

5. Schritt, Konstruktion einer geeigneten Würfelüberdeckung.

Jede offene Teilmenge $U \subset \mathbb{R}^n$ läßt sich als abzählbare Vereinigung von abgeschlossenen Würfeln schreiben

$$U = \overline{W}_1 \cup \overline{W}_2 \cup \overline{W}_3 \cup \cdots,$$

wobei die Würfel W_ν offen und paarweise disjunkt sind

$$\overline{W}_\nu = \{x;\ \|x-a\| \leq \varepsilon, \quad W_\nu = \{x;\ \|x-a\| < \varepsilon.$$

Beweis. Wir betrachten die Menge aller Würfel

$$W = \left\{x;\ \frac{a_n u}{2^r} < x_\nu < \frac{a_n u + 1}{2^r}, \quad \nu = 1, \cdots, n\right\},$$

wobei r alle natürlichen und a_ν alle ganzen Zahlen durchläuft.

Die Menge dieser Würfel ist abzählbar. (Die Menge $\mathbb{N} \times \mathbb{Z}^n$ ist abzählbar, weil allgemein das kartesische Produkt von abzählbaren Mengen abzählbar ist.) Man überlegt sich nun (dies im einzelnen durchzuführen sei dem Leser überlassen):

a) Sind W und W' zwei der beschriebenen Würfelmengen, so gilt

$$W \cap W' \neq \emptyset \Rightarrow W \subset W' \text{ oder } W' \subset W.$$

b) Jeder Punkt $a \in \mathbb{R}^n$ ist enthalten in einem der Würfel W, wobei noch r beliebig groß gewählt werden kann (und daher die Kantenlänge beliebig klein).

Aus a) und b) konstruiert man nun leicht eine Würfelaufteilung der gewünschten Art.

6. Schritt. Die Funktion

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f \geq 0,$$

sei stetig mit kompaktem Träger in B. Es gilt

$$\int_B f(y)\,dy \geq \int_A f(u(x))\,|j(u,x)|\,dx.$$

Beweis indirekt: Die Ungleichung sei falsch. Dann kann jedenfalls f nicht identisch Null sein, aus Stetigkeitsgründen sind beide Integrale nicht Null und man kann eine Zahl $q > 1$ finden, so daß sogar

$$q^{n+4} \int_B f(y)\,dy < \int_A f(u(x))\,|j(u,x)|\,dx$$

gilt.

Wir betrachten nun die Menge

$$A_0 = \{x \in \mathbb{R}^n; \quad f(u(x)) \neq 0\} \subset A.$$

Diese Menge ist offen, da f stetig ist. Nach dem 5. Schritt existiert eine Überdeckung

$$A_0 = \overline{W}_1 \cup \overline{W}_2 \cdots .$$

wobei die Würfel W_ν offen und paarweise disjunkt sind.

Behauptung. Für mindestens einen dieser Würfel, nennen wir ihn $W^{(1)}$, gilt die Ungleichung

$$q^{n+4} \int_{u(W^{(1)})} f(y)\,dy < \int_{W^{(1)}} f(u(x))\,|j(u,x)|\,dx.$$

Beweis. Würde für alle Würfel W die Ungleichung "$\geq$" gelten, so könnte man mit Hilfe des Lebesgueschen Grenzwertsatzes sogar

$$\begin{aligned} q^{n+4} \int_B f(y)\,dy &\geq \sum_{\nu=1} q^{n+4} \int_{u(W_\nu)} f(y)\,dy \\ &\geq \sum_{\nu=1} \int_{W_\nu} f(u(x))\,|j(u,x)|\,dx \\ &= \int_A f(u(x))\,|j(u,x)|\,dx \end{aligned}$$

schließen. (Man beachte, daß das Integral über den Rand eines Würfels Null ergibt.)

Die Existenz eines Würfels $W^{(1)}$ mit

$$q^{n+4} \int\limits_{u(W^{(1)})} f(y)\,dy < \int\limits_{W^{(1)}} f(u(x))\,|j(u,x)|\,dx$$

ist damit gesichert. Den Würfel $W^{(1)}$ kann man durch Halbieren der Kantenlänge in 2^n Würfel aufteilen. Mit der gleichen Schlußweise folgt die Existenz eines Würfels $W^{(2)}$ mit der obigen Ungleichung.

So fortfahrend erhält man eine Folge von Würfeln

$$W^{(1)} \supset W^{(2)} \supset W^{(3)} \supset \dots$$

mit den Eigenschaften

a) Kantenlänge von $W^{(k)} \to 0$ für $k \to \infty$.

b) $q^{n+4} \int\limits_{u(W^{(k)})} f(y)\,dy < \int\limits_{W^{(k)}} f(u(x))\,|j(u,x)|\,dx.$

Nach dem verallgemeinerten Intervallschachtelungsprinzip existiert ein Punkt

$$a \in \overline{W^{(1)}} \cap \overline{W^{(2)}} \cap \overline{W^{(3)}} \dots .$$

Zu diesem a betrachten wir die im 4. Schritt konstruierte Umgebung U. Wählt man k genügend groß, so gilt

$$W^{(k)} \subset U$$

und wir haben einen Widerspruch zwischen den Ungleichungen des 4. und 6. Schritts erhalten.

7. Schritt, Beweis von Theorem 8.1 für stetige Funktionen mit kompaktem Träger f.

Man kann annehmen, daß $f \geq 0$ gilt. Dies liegt an der Möglichkeit des Aufspaltens

$$f = f^+ - f^-$$

mit

$$f^+ = f \vee 0, \quad f^- = (-f)^+.$$

Nach dem 6. Schritt gilt die Ungleichung

$$\int\limits_B f(y)\,dy \geq \int\limits_A f(u(x))\,|j(u,x)|\,dx.$$

Wendet man diese Ungleichung an auf

A anstelle von B
B anstelle von A

$$u^{-1} \text{ anstelle von } u$$
$$f(u(x))\,|j(u,x)| \text{ anstelle von } f,$$

so resultiert die umgekehrte Gleichung.

8. Schritt, Beweis von 9.1. Sei

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f \geq 0.$$

eine Bairesche Funktion. Dann ist auch $f\chi_B$ Bairesch. Man wähle

$$f_k \uparrow f$$

und

$$g_k \uparrow \chi_B; \quad g_k \geq 0,$$

wobei f_k und g_k stetig mit kompaktem Träger sei. Dann gilt

$$f_k \cdot g_k \uparrow f \cdot \chi_B.$$

Die Gleichung

$$\int\limits_B f(y)\,dy = \int\limits_A F(u(x))\,|j(u,x)|\,dx$$

folgt nun aus dem 8. Schritt. Jetzt folgt die Transformationsformel unmittelbar für das äußere Integral und dann erst recht für das Integral.

9. Meßbarkeit

Man kann jede Funktion $f : \mathbb{R}^n \to \mathbb{R}$ als Limes einer Folge von beschränkten Funktionen mit kompaktem Träger schreiben, nämlich

$$f = \lim f_k$$

mit

$$f_k(x) = \begin{cases} f(x), & \text{falls } \|x\| \leq k \text{ und } |f(x)| \leq k, \\ 0, & \text{sonst.} \end{cases}$$

9.1 Definition. *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt meßbar, falls die Funktionen f_k, $k = 1, 2, \ldots$, alle integrierbar sind.* defMes

Es gilt offenbar

$$f_k = (\varphi_k \wedge f) \vee (-\varphi_k),$$

wenn φ_k die Funktion

$$\varphi_k(x) = \begin{cases} k & \text{für } \|x\| \leq k \\ 0 & \text{sonst} \end{cases}$$

bezeichnet. Da φ_k integrierbar ist, folgt

9.2 Bemerkung. *Jede integrierbare Funktion ist meßbar.* IntfMe

Diese und die folgende Bemerkung folgen unmittelbar aus dem Lebesgueschen Grenzwertsatz.

9.3 Bemerkung. *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine meßbare Funktion. Es existiere eine integrierbare Funktion h mit der Eigenschaft* MesfIn

$$|f(x)| \leq |h(x)| \quad \textit{für alle } x.$$

Dann ist auch f integrierbar.

Die meßbaren Funktionen haben alle wünschenswerten Stabilitätseigenschaften.

9.4 Satz. StabMe

1) *Summe und Produkt von meßbaren Funktionen sind meßbar, konstante Funktionen sind meßbar.*
2) *Seien f, g meßbare Funktionen. Dann sind auch $f \wedge g$, $f \vee g$, insbesondere $|f|$ meßbar.*
3) *Sei f_n eine punktweise konvergente Folge meßbarer Funktionen. Dann ist auch ihr Grenzwert meßbar.*
4) *Man darf eine meßbare Funktion auf einer Nullmenge abändern, ohne ihre Meßbarkeit zu verlieren.*

Beweis. Es genügt zu zeigen, daß das Produkt meßbarer Funktionen meßbar ist, da alle anderen Stabilitätseigenschaften aus entsprechenden Eigenschaften integrierbarer Funktionen folgen.

Wegen der Formel

$$4(f \cdot g) = (f + g)^2 - (f - g)^2$$

braucht man nur zu beweisen, daß mit f auch f^2 meßbar ist. Da man f durch f_k ersetzen darf, genügt es zu zeigen:

Das Quadrat einer *beschränkten* integrierbaren Funktion ist integrierbar. Dies folgt aus der Charakterisierung in Satz ??? Verwendung der Formel ($f^2 - h^2 = (f - h)(f + h)$).

Als Anwendung des Begriffs der meßbaren Funktion dient:

9.5 Ergänzung zum Satz von Fubini. *Sei $f : \mathbb{R}^{n+m} \to \mathbb{R}$ eine meßbare Funktion, so daß* VarFub

$$\bar{\int} \bar{\int} |f(x, y)|\, dy \; dx$$

endlich ist. Dann ist f integrierbar und es gilt folgedessen der Satz von Fubini 7.1.

Der Beweis ist einfach und wird übergangen.

10. Räume integrierbarer Funktionen

Wir bezeichen mit $\mathcal{N}$ die Menge aller Nullfunktionen $f : \mathbb{R}^n \to \mathbb{R}$. Dies ist ein Untervektorraum von $\mathcal{L}^1(\mathbb{R}^n)$. Da es auf Nullfunktionen meist nicht ankommt, ist es sinnvoll, diese aus der Betrachtung zu eliminieren und sie zu Null zu machen. Dies geschieht durch das Konstrukt des Faktorvektorraums

$$L^1(\mathbb{R}^n) := \mathcal{L}^1(\mathbb{R}^n)/\mathcal{N}.$$

Diesen kann man konkret wie folgt beschreiben. Man führt eine Äquivalenzrelation in $\mathcal{L}^1(\mathbb{R}^n)$ ein. Man nennt zwei Funktionen f, g äquivalent ($f \sim g$), wenn $f - g$ eine Nullfunktion ist. Die Forderungen an eine Äquivalenzrelation (Reflexivität, Symmetrie und Transitivität) sind offensichtlich. Ist $f \in \mathcal{L}^1(\mathbb{R}^n)$, so bezeichnet man die Äquivalenzklasse mit

$$[f] := \{\, g;\ g \sim f \,\}.$$

Desinitionsgemäß ist $L^1(\mathbb{R}^n)$ die Menge all dieser Äquivalenzklassen. Man überlegt sich leicht, dass die Definitionen

$$[f] + [g] := [f + g], \quad [Cf] := C[f]$$

unabhängig von der Wahl des Repräsentanten sind und macht so $L^1(\mathbb{R}^n)$ zu einem Vektorraum. Die 1-Norm einer intergierbaren Funktion f ist durch

$$\|f\|_1 := \int\limits_{\mathbb{R}}^{n} |f(x)|\, dx$$

defniert. Diese hängt nur von der Äquivalenzklasse von f ab, man kann also $\|[f]\| := \|f\|$ definieren. Der Vorteil der Konstruktion von L^1 zeigt sich nun in: *Durch* $\|\cdot\|$ *wird* $L^1(\mathbb{R}^n)$ *zu einem normierten Raum.*

Dies bedeutet bekanntlich
a) $\|Ca\|_1 = |C| \cdot \|a\|_1 \qquad (C \in \mathbb{R})$,
b) $\|a + b\|_1 \le \|a\|_1 + \|b\|_1$,
c) $\|a\|_1 \ge 0$ und $= 0$ nur für $a = 0$.

Der Vorteil der Konstruktion des Faktorraums L^1 ist, dass die Definitheit in c) gilt. In $\mathcal{L}^1$ wäre diese falsch. Die große Stärke des Lebesgue-Integrals ist:

10.1 Satz. *Der Raum $L^1(\mathbb{R}^n)$ zusammen mir der Eins-Norm ist ein **Banachraum**.* LB

Beweis. Die Aussage bedeutet folgendes:

Sei (f_n) eine Folge integrierbarer Funktionen, so dass zu jedem $\varepsilon > 0$ ein N exisiert mit $\|f_n - f_m\| \le \varepsilon$ für $n, m \ge N$. Dann esistiert eine integrierbare Funktion f, so dass $\|f - f_n\|_1$ eine Nullfolge ist.

Mit Hilfe der Dreiecksungleichung sieht man leicht: Es genügt zu zeigen, dass eine Teilfolge $g_n = f_{\nu_n}$ existiert, so dass $\|f - g_n\|_1$ gegen Null konvergiert (betrachte $f - f_n + f_n - g_n$). Man kann sich leicht überlegen, dass eine Teilfolge mit der Eigenschaft

$$\|g_{n+1} - g_n\| \le 2^{-n}$$

existiert. Dazu betrachtet man eine Folge $N_1 < N_2 < \dots$, so dass

$$\|g_m - g_n\| \le 2^{-k} \quad \text{für} \quad n, m \ge N_k.$$

Die Teilfolge $g_k = f_{N_k}$ hat dann die gewünschte Eigenschaft. Jetzt beachte man

$$g_n = g_1 + (g_2 - g_1) + \cdots + (g_n - g_1)$$

und definiere

$$G_n := |g_1| + |g_2 - g_1| + \cdots + |g_n - g_1|.$$

Die Folge der G_n ist monoton wachsend. Die Folge ihrer Integrale ist beschränkt (durch die geometrische Reihe). Nach dem Satz von Beppo Levi ist ihr Supremum eine integrierbare Funktion. Deren Unendlichkeitsstellen müssen eine Nullmenge bilden. Daher konvergiert G_n außerhalb einer Nullmenge. Die Folge g_n konvergiert außerhalb dieser Nullmenge nach dem Majorantenkriterium ebenfalls. Aus dem Lebesgueschen Grenzwertsatz folgt, dass die Grenzfunktion f integrierbar ist und es folgt auch $\|g_n - f\|_1 \to 0$.

Wir haben bei diesem Beweis mitgezeigt:

10.2 Satz. *Aus der Konvergenzaussage $\|f - f_n\| \to 0$ in $\mathcal{L}^1(\mathbb{R}^n)$ folgt: Es gibt eine Teilfolge, welcher wenigstens außerhalb einer geeigneten Nullmenge punktweise gegen f konvergiert.* STkp

Sprechweise. Anstelle $\|f - f_\nu\| \to 0$ sagt man auch, dass die Folge (f_ν) *im Mittel* gegen f konvergiert.

Wir geben noch einige Varianten des Raumes $\mathcal{L}^1$ an. Sei $p \ge 1$ eine reelle Zahl. Wir definieren

$$\mathcal{L}^p(\mathbb{R}^n) := \big\{\, f : \mathbb{R}^n \longrightarrow \mathbb{R} \text{ messbar}; \quad |f|^p \text{ integrierbar} \,\big\}.$$

Im Fall $p = 1$ erhält man natürlich den ursprünglichen $\mathcal{L}^1$. Die Bedingung „messbar“ benötigt man für den Beweis, das $\mathcal{L}^p$ ein Vektorraum ist. Wir definieren für $f \in \mathcal{L}^p(\mathbb{R}^n)$ die sogenannte p-Norm durch

$$\|f\|_p := \sqrt[p]{\int_{\mathbb{R}}^{n} |f(x)|^p \, dx}.$$

Es gilt

a) $\|cf\| = |c| \ \ \|f\|_p$,

b) $\|f\|_p \geq 0$,

c) Mit f und g ist auch $f + g$ in $\mathcal{L}^p(\mathbb{R}^n)$ enthalten und es gilt

d) $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Die Ungleichung d) heißt Minkowskische Ungleichung. Sie folgt aus der Hölderschen Ungleichung (s. Kapitel IV §5).

Man kann ergänzend noch $\mathcal{L}^\infty$ definieren. Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt im wesentlichen beschränkt, wenn es eine Konstante C gibt, so dass $|f(x)| \leq C$ außerhalb einer geeigneten (von C abhängigen) Nullmenge gibt. Das Infimum aller wesentlichen Schranken bezeichnet man mit $\|f\|_\infty$. Man setzt

$$\mathcal{L}^\infty(\mathbb{R}^n) := \big\{\, f : \mathbb{R}^n \longrightarrow \mathbb{R} \text{ messbar}; \quad f \text{ im wesentlichen beschränkt} \,\big\}.$$

Die Menge der Nullfunktionen $\mathcal{N}$ ist ein Unterraum von $\mathcal{L}^p(\mathbb{R}^n)$. Man kann also wieder

$$L^p(\mathbb{R}^n) := \mathcal{L}^p(\mathbb{R}^n)/\mathcal{N}$$

definieren und die p-Norm drückt wieder auf diesen Raum Raum durch.

10.3 Satz. *Der Raum $L^p(\mathbb{R}^n)$ versehen mit der p-Norm $\|f\|_p$ ist ein Banachraum für $1 \leq p \leq \infty$.* LpR

Der Beweis geht analog wie der von Satz 10.1. □

Ein besonders wichtiger Fall ist der Fall $p = 2$. In diesem Fall kann man ein Skalarprodukt

$$\langle f, g \rangle := \int_{\mathbb{R}^n} f(x)g(x) \, dx$$

definieren. Man kann sich leicht überlegen, dass das Integral für alle $f, g \in \mathcal{L}^2(\mathbb{R}^n)$ existiert. Es gilt

$$\|f\|_2 = \sqrt{\langle f, g \rangle}.$$

Die Minkowskische Ungleichung folgt in diesem Fall aus der Cauchy-Schwarzschen Ungleichung.

10.4 Satz. *Der Raum $L^2(\mathbb{R}^n)$ mit dem angegebenen Skalarprodukt ist ein Hilbertraum.* LzH

Die Räume L^p wurden als reelle Vektorräume definiert. Es gibt aber auch komplexe Varianten. Man nennt hierzu eine Funktion $f : \mathbb{R}^n \to \mathbb{C}$ integrierbar (meßbar), wenn dies für Real- und Imaginärteil getrennt richtig ist. Wir verzichten auf weitere Einzelheiten.

Übrigens: Wenn f und g stetige Funktionen mit kompaktem Träger sind, so gilt $[f] = [g] \Longrightarrow f = g$. Es gilt also $C_c(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$.

Index