Analysis II für Physiker

Nicolas Ginoux

Universität Regensburg - Sommersemester 2012

7. Juli 2014

Inhaltsverzeichnis

Einführung

Dieses Skript besteht aus den Vorlesungsnotizen. Größtenteils wurden sie freundlicherweise von Frau Bonn getippt. Anmerkungen zum Skript sind bitte an Nicolas Ginoux (`nicolas.ginoux@mathematik.uni-regensburg.de`) zu schicken.

Wegen Zeitmangel wurden nicht alle Beweise in der Vorlesung durchgeführt; von Kapitel 2 aus wurden sie fast systematisch weggelassen. Die fehlenden Beweise sind allerdings in diesem Skript zu finden. Es wird empfohlen, diese Beweise zu lesen und sich selber einzuarbeiten.

Als Literatur wurden, für Kapitel 1 und Kapitel 2, hauptsächlich [3] sowie Teile aus [1, 2] verwendet. Punktweise – z.B. in den Abschnitten 2.2.2 und 2.2.3 – wurde auch [4] benutzt. Die Abschnitte 2.3 und 2.4 wurden von [10] stark inspiriert. Kapitel 3 und Anhang A lassen sich von [6], [7] und [11] ableiten. Kapitel 4 bleibt wieder sehr nahe an [4]. Für Kapitel 5 ist [5] immer noch eine sehr gute Referenz, wobei [8, 9] für die Integralsätze sehr hilfreich sind.

Herzlichen Dank an Mihaela Pilca für Korrekturen und Anregungen.

Angekündigtes Programm: In dieser Vorlesung sollen die Differential- und Integralrechnung mehrerer Variablen in dem Umfang vorgestellt werden, wie es für die Studierenden der Physik nötig ist. Dazu gehören:

- Kurven im $\mathbb{R}^n$,
- Differenzierbare Abbildungen im $\mathbb{R}^n$,
- Vektorfelder und Potentiale,
- Taylor-Entwicklung in mehreren Variablen,
- Minima und Maxima, auch mit Nebenbedingungen,
- Sätze über Umkehrfunktionen und implizite Funktionen,
- Mannigfaltigkeiten,
- Integral im $\mathbb{R}^n$,
- Transformationsformel,
- Polar- und Zylinderkoordinaten,
- Gewöhnliche Differentialgleichungen (Existenz und Eindeutigkeit von Lösungen),

- Lineare Differentialgleichungen (Systeme 1. Ordnung und eine Gleichung n-ter Ordnung),
- Differentialformen,
- Integralsätze im $\mathbb{R}^n$ (Gauß, Green, Stokes),
- Divergenz- und Rotationssatz.

Tatsächliches Programm:

1. (17.4) Vorstellung der Veranstaltung, Vorlesungsplan. Kap. 1 (Kurven): parametrisierte und regulär parametrisierte Kurven im $\mathbb{R}^n$, Umparametrisierungen, Produkt- und Kettenregeln, Orientierung, Länge, Parametrisierung nach Bogenlänge.
2. (18.4) Kap. 1, Fortsetzung: Existenz einer Umparametrisierung nach Bogenlänge, Integral einer Funktion entlang einer Kurve, Krümmung ebener Kurven, Frenet-Gleichungen.
3. (24.4) Kap. 1, Ende: Hauptsatz der Theorie der ebenen Kurven (eine ebene Kurve wird bis auf euklidische Bewegungen von ihrer Krümmungsfunktion festgelegt) mit Beweisskizze. Kap. 2 (Differenzierbare Abbildungen in mehreren Veränderlichen): Metrische Räume, offene und abgeschlossene Bälle.
4. (25.4) Kap. 2, Fortsetzung: Umgebungen, offene und abgeschlossene Teilmengen, Inneres, Abschluss und Rand einer Teilmenge, Charakterisierung des Inneren und des Abschlusses (als offene bzw. abgeschlossene Teilmenge); konvergente Folgen, Charakterisierung des Abschlusses mit Folgen; stetige Abbildungen (Definition und Charakterisierung mit Urbildern von Bällen).
5. (2.5) Kap. 2, Fortsetzung: Charakterisierung der Stetigkeit mit Urbildern offener bzw. abgeschlossener Teilmengen, die Verknüpfung zweier stetiger Abbildungen ist wieder stetig; kompakte metrische Räume: Definition, Beschränktheit, kompakte Teilmengen, Vereinigung und Durchschnitt kompakter Teilmengen, Produkte kompakter Räume, Satz von Heine-Borel (Charakterisierung der kompakten Teilmengen des $\mathbb{R}^n$), Stetigkeit und Kompaktheit (das Bild einer kompakten Teilmenge unter einer stetigen Abbildung ist wieder kompakt, jede stetige bijektive Abbildung auf einem kompakten Raum ist bereits ein Homöomorphismus), Existenz eines Maximums und eines Minimums einer stetigen Funktion auf einem kompakten Raum, Gleichmäßigstetigkeit einer stetigen Abbildung auf einem kompakten Raum.
6. (8.5) Kap. 2, Fortsetzung: vollständige metrische Räume: Cauchy-Folgen, Beispiele (jede konvergente Folge ist Cauchy, die Umkehrung gilt nicht), vollständiger metrischer Raum, Beispiele (jeder kompakte Raum ist vollständig, $\mathbb{R}^n$ ist vollständig, $\mathbb{Q}$ ist nicht vollständig), vollständige Teilmengen, Vereinigung und Durchschnitt vollständiger Teilmengen (man beachte die Bedingung $\operatorname{diam}(Y_n) \underset{n\to\infty}{\longrightarrow} 0$ für den Durchschnitt), Produkte vollständiger Räume, Banach'scher Fixpunktsatz (jede Kontraktion in einem vollständigen metrischen Raum hat einen eindeutigen Fixpunkt), Anwendungsbeispiel, Fortsetzungssatz von Tietze (jede auf einer dichten Teilmenge $D \subset X$ definierte gleichmäßig stetige Abbildung mit Werten in einem vollständigen metrischen Raum Y lässt sich eindeutig auf X zu einer stetigen Abbildung $X \to Y$ fortsetzen).

Kapitel 1

Kurven

1.1 Parametrisierung und Länge

1.1.1 Regulär parametrisierte Kurven

Definition 1.1 *Eine parametrisierte Kurve in* $\mathbb{R}^n$ *(wobei* $n \in \mathbb{N}$*) ist eine Abbildung* $c : I \longrightarrow \mathbb{R}^n$*, wobei* $I \subset \mathbb{R}$ *ein Intervall ist.*

Erinnerung: In den kanonischen Koordinaten des $\mathbb{R}^n$ kann eine Abbildung $c : I \longrightarrow \mathbb{R}^n$ in der Form

$$t \longmapsto c(t) = \begin{pmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{pmatrix}$$

geschrieben werden, wobei $c_j : I \longrightarrow \mathbb{R}$ eine reellwertige Funktion ist, für alle $1 \leq j \leq n$. Die c_j's heißen Koordinatenfunktionen von c. Es gilt

- c ist stetig $\Longleftrightarrow$ c_j ist stetig, $\forall\ 1 \leq j \leq n$.
- c ist ableitbar bzw. C^1, C^k, C^∞ $\Longleftrightarrow$ c_j ist ableitbar, C^1, C^k, C^∞, $\forall\ 1 \leq j \leq n$.
- c ist stückweise C^1 (bzw. C^k, C^∞) $\Longleftrightarrow$ c_j ist stückweise C^1 (bzw. C^k, C^∞).

Außerdem gilt: ist c k-mal ableitbar (mit $k \geq 1$), so lässt sich ihre k-te Ableitung in der Form

$$t \longmapsto c^k(t) = \begin{pmatrix} c_1^{(k)}(t) \\ \vdots \\ c_n^{(k)}(t) \end{pmatrix}$$

schreiben, wobei $c_j^{(k)} : I \longrightarrow \mathbb{R}$ die k-te Ableitung von c_j ist, für alle $1 \leq j \leq n$.

Definition 1.2 *Eine parametrisierte Kurve* $c : I \longrightarrow \mathbb{R}^n$ *heißt* regulär parametrisiert *g.d.w. sie stückweise*[1] C^1 *ist und ihre erste Ableitung* $c' : I \longrightarrow \mathbb{R}^n$ *nirgends verschwindet.*

Bemerkung 1.3 Ist $c : I \longrightarrow \mathbb{R}^n$ stückweise C^1 aber nicht C^1, so heißt "c' verschwindet nirgends": $c'(t) \neq 0$ für alle t, wo c ableitbar ist, sowie $c'(t^-) \neq 0$ und $c'(t^+) \neq 0$ an jedem t, wo c nur links- oder rechts-ableitbar ist.

[1] In der Vorlesung wurden lediglich C^1 parametrisierte Kurven betrachtet.

Bezeichnungen 1.4 Wir schreiben $\dot{c} := c'$ sowie $\ddot{c} := c''$.

Beispiele 1.5

1. Für $n = 1$ ist jede Funktion $f : I \longrightarrow \mathbb{R}$ eine parametrisierte Kurve und jede C^1 Funktion $f : I \longrightarrow \mathbb{R}$ mit $f'(t) \neq 0 \ \forall\, t \in I$ ist eine regulär parametrisierte Kurve.

2. Sei $n = 2$. Die Abbildung
$$\begin{array}{rcl} c : \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & p + tv, \end{array}$$
wobei $p, v \in \mathbb{R}^2$, ist eine C^∞ parametrisierte Kurve. Wegen $\dot{c}(t) = v \ \forall\, t \in \mathbb{R}$ gilt: c ist genau dann regulär parametrisiert, wenn $v \neq 0$ gilt.

3. Sei $n = 2$. Die Abbildung
$$\begin{array}{rcl} c : \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & \binom{\cos(t)}{\sin(t)} \end{array}$$
ist eine regulär parametrisierte Kurve, denn: c ist C^∞ und es gilt $\dot{c}(t) = \binom{-\sin(t)}{\cos(t)} \ \forall\, t \in \mathbb{R}$, insbesondere $\dot{c}(t) \neq 0$ für alle $t \in \mathbb{R}$.

3'. Für $n = 2$ ist die Abbildung
$$\begin{array}{rcl} d : \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & \binom{\cos(2t)}{\sin(2t)} \end{array}$$
ebenfalls eine C^∞ regulär parametrisierte Kurve.

4. Für $n = 2$ ist $c : \mathbb{R} \longmapsto \mathbb{R}^2$, $t \longmapsto \binom{t^3}{t^2}$, zwar eine C^∞ parametrisierte Kurve, ist aber wegen $\dot{c}(0) = 0$ keine regulär parametrisierte Kurve.

Definition 1.6 *Seien $c : I \longrightarrow \mathbb{R}^n$ und $d : J \longrightarrow \mathbb{R}^n$ regulär parametrisierte Kurven. Man sagt, dass d eine* Umparametrisierung *von c ist g.d.w. ein C^∞ Diffeomorphismus $\varphi : J \longrightarrow I$ existiert mit $d = c \circ \varphi$.*

Erinnerung: Eine Abbildung $\varphi : J \longrightarrow I$ heißt C^∞ Diffeomorphismus g.d.w.:

- φ ist C^∞,
- φ ist bijektiv,
- Die Umkehrabbildung $\varphi^{-1} : I \longrightarrow J$ von φ ist ebenfalls C^∞.

Beispiel 1.7 Sei $n = 2$, $I := J := \mathbb{R}$ und c, d wie in den Beispielen 1.5.3 und 1.5.3' oben. Dann ist d eine Umparamatrisierung von c.

Definition 1.8 *Zwei regulär parametrisierte Kurven heißen genau dann* äquivalent*, wenn die eine eine Umparametrisierung der anderen ist.*

Übungsaufgabe: Überprüfe, dass dies eine Äquivalenzrelation auf der Menge der regulär parametrisierten Kurven definiert.

VON HIER AUS BETRACHTEN WIR NUR C^∞ REGULÄR PARAMETRISIERTE KURVEN.

Bezeichnungen 1.9

- $M_{m\times n}(\mathbb{R}) :=$ Menge der $m \times n$ reellen Matrizen
- $\langle \cdot\,, \cdot \rangle :=$ kanonisches euklidisches Skalarprodukt im $\mathbb{R}^n$, d.h. $\langle x, y\rangle := \sum_{j=1}^{n} x_j y_j$ für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$.
- $|\cdot| :=$ kanonische euklidische Norm im $\mathbb{R}^n$: $|x| := \sqrt{\langle x, x\rangle} = \sqrt{\sum_{j=1}^{n} x_j^2}$ für alle $x \in \mathbb{R}^n$.

Proposition 1.10 (Rechenregeln für Kurven) *Seien $c, d : I \longrightarrow \mathbb{R}^n$ regulär parametrisierte Kurven, $\varphi : I \longrightarrow \mathbb{R}$ eine C^∞ Funktion, $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ eine C^∞ Abbildung. Dann gilt:*

(i) (Produktregel 1) $\dot{\overparen{\varphi c}} = \varphi' c + \varphi \dot{c}$.

(ii) (Produktregel 2) $\langle c, d\rangle' = \langle \dot{c}, d\rangle + \langle c, \dot{d}\rangle$.

(iii) (Kettenregel 1) $\dot{\overparen{c \circ \psi}} = (\dot{c} \circ \psi) \cdot \psi'$ *für eine C^∞ Funktion $\psi : J \longrightarrow I$ ($J \subset \mathbb{R}$ Intervall).*

(iv) (Kettenregel 2[2]) $\dot{\overparen{A \cdot c}} = \frac{dA}{dt} \cdot c + A \cdot \dot{c}$.

Beweis: In der Zentralübung. □

1.1.2 Orientierung einer Kurve

Definition 1.11 *Zwei äquivalente regulär parametrisierte Kurven $c : I \longrightarrow \mathbb{R}^n$ und $d : J \longrightarrow \mathbb{R}^n$* besitzen dieselbe Orientierung $:\Longleftrightarrow$ *$\exists$ C^∞-Diffeomorphismus $\varphi : J \longrightarrow I$ mit $\varphi' > 0$ und $c \circ \varphi = d$. Ist dies nicht der Fall, so sagt man, dass d die umgekehrte Orientierung von c hat.*

Beispiele 1.12

- Die regulär parametrisierten Kurven $c : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(t)}{\sin(t)}$ und $d : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(2t)}{\sin(2t)}$, sind äquivalent und besitzen dieselbe Orientierung.
- Die regulär parametrisierten Kurven c von oben und $\overline{c} : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(t)}{-\sin(t)}$, sind äquivalent, besitzen aber nicht dieselbe Orientierung.

[2]Sie kann auch als Produktregel gesehen werden.

1.1.3 Länge und Bogenlänge

Definition 1.13 *Das* Geschwindigkeitsvektorfeld *einer regulär parametrisierten Kurve $c : I \longrightarrow \mathbb{R}^n$ ist die Abbildung*

$$\dot{c} : I \longrightarrow \mathbb{R}^n.$$

Definition 1.14 *Die Länge einer regulär parametrisierten Kurve $c : I \longrightarrow \mathbb{R}^n$ ist $L[c] := \int_I |\dot{c}(s)|ds$.*

Proposition 1.15 *Sind c und d äquivalente Kurven, so haben sie dieselbe Länge.*

Beweis: In der Zentralübung. □

Beispiele 1.16

1. Ist $c : I \longrightarrow \mathbb{R}^n$ regulär parametrisiert mit $|\dot{c}(s)| = 1 \;\forall\; s \in I$, so gilt $L[c] = \int_I 1ds = |I|$: die Länge der Kurve stimmt mit der Länge des Intervalls überein.

2. Für $c : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto p + tv$ (mit $p \in \mathbb{R}^2$ und $v \in \mathbb{R}^2 \setminus \{0\}$) und $a < b$ in $\mathbb{R}$ gilt: $L[c_{|[a,b]}] = \int_a^b |v|ds = (b-a) \cdot |v|$.

3. Für $c : [0, 2\pi] \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(t)}{\sin(t)}$, gilt:

$$L[c] = \int_0^{2\pi} \left|\begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix}\right| ds = \int_0^{2\pi} ds = 2\pi.$$

3'. Für $c : [0, 2\pi] \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(2t)}{\sin(2t)}$, gilt:

$$L[c] = \int_0^{2\pi} \left|2\begin{pmatrix} -\sin(2s) \\ \cos(2s) \end{pmatrix}\right| ds = 2\int_0^{2\pi} ds = 4\pi.$$

Definition 1.17 *Eine regulär parametrisierte Kurve $c : I \longrightarrow \mathbb{R}^n$ heißt* nach Bogenlänge parametrisiert *g.d.w.*

$$|\dot{c}(s)| = 1$$

für alle $s \in I$ gilt.

Die Benennung kommt daher, dass im Falle $0 \in I$ gilt:

$$L[c_{|[0,t]}] = t \qquad \forall\; t \geq 0, t \in I.$$

"Die Länge wird durch den Parameter gegeben".

Proposition 1.18 *Sei $c : I \longrightarrow \mathbb{R}^n$ eine regulär parametrisierte Kurve und $t_0 \in I$ beliebig. Dann ist $c \circ \varphi$ eine nach Bogenlänge Umparametrisierung von c, wobei φ die Umkehrabbildung von $\psi : I \longrightarrow \mathbb{R}$, $t \longmapsto \int_{t_0}^t |\dot{c}(s)|ds$, ist. Außerdem unterscheiden sich zwei Umparametrisierungen nach Bogenlänge von c durch eine affine Parametertransformation der Form $t \longmapsto \pm t + r$ für ein $r \in \mathbb{R}$: ist $\tilde{c} : I' \longrightarrow \mathbb{R}^n$ eine weitere Umparametrisierung nach Bogenlänge von c, so gibt es ein eindeutiges $r \in \mathbb{R}$ s.d. entweder $\tilde{c}(t) = (c \circ \varphi)(t + r)$ $\forall\; t \in I'$ oder $\tilde{c}(t) = (c \circ \varphi)(-t + r)$ $\forall\; t \in I'$ gilt.*

Beweis: Die Abbildung ψ ist eine (C^∞) Abbildung mit $\psi'(t) = |\dot{c}(t)| > 0 \ \forall\ t \in I$, insbesondere ist ψ ein C^∞ Diffeomorphismus auf sein Bild $J := \psi(I)$, welches notwendigerweise ein Intervall ist (Mittelwertsatz). Für $d := c \circ \varphi = c \circ \psi^{-1}$ gilt dann $\forall\ t \in J$:

$$\begin{aligned} \dot{d}(t) = \dot{\widehat{c \circ \varphi}}(t) &= \dot{c}(\varphi(t)) \cdot \varphi'(t) \text{ mit } \varphi'(t) = (\psi^{-1})'(t) = \tfrac{1}{\psi'(\psi^{-1}(t))} \\ &= \tfrac{\dot{c}(\varphi(t))}{|\dot{c}(\varphi(t))|}, \end{aligned}$$

insbesondere $|\dot{d}(t)| = \left|\frac{\dot{c}(\varphi(t))}{|\dot{c}(\varphi(t))|}\right| = 1 \ \forall\ t \in J$. Somit ist $d = c \circ \varphi$ eine Umparametrisierung nach Bogenlänge von c. Ist $\tilde{c} : I' \longrightarrow \mathbb{R}^n$ eine weitere Umparametrisierung nach Bogenlänge von c, so ist $\tilde{c}$ auch Umparametrisierung nach Bogenlänge von d. Schreibe $\tilde{c} = d \circ \chi$, wobei $\chi : I' \longrightarrow J$ ein Diffeomorphismus ist. Aus

$$1 = |\dot{\tilde{c}}(t)| = |\dot{d}(\chi(t)) \cdot \chi'(t)| = \underbrace{|\dot{d}(\chi(t))|}_{=1} \cdot |\chi'(t)|$$

(für alle $t \in I'$) folgt $|\chi'(t)| = 1 \ \forall\ t \in I'$ und somit entweder $\chi' = 1$ auf I' oder $\chi' = -1$ auf I'. □

1.1.4 Integral einer Funktion entlang einer Kurve

Definition 1.19 *Sei $c : I \longrightarrow \mathbb{R}^n$ regulär parametrisierte Kurve und $f : c(I) \longrightarrow \mathbb{R}^k$ eine Abbildung, wobei $k \in \mathbb{N}$ und I ein abgeschlossenes, beschränktes Intervall ist. Angenommen, $f \circ c : I \longrightarrow \mathbb{R}^k$ sei stetig. Dann ist das* Integral von f entlang c *definiert als* $\int\limits_c f := \int_I f \circ c(s) \cdot |\dot{c}(s)| ds \in \mathbb{R}^k$.

Erinnerung: Für eine stetige Abbildung $g : I \longrightarrow \mathbb{R}^k$, $g(t) = \begin{pmatrix} g_1(t) \\ \vdots \\ g_k(t) \end{pmatrix}$ (wobei $g_j : I \longrightarrow \mathbb{R}$ ist $\int\limits_I g(s) ds := \begin{pmatrix} \int_I g_1(s) ds \\ \vdots \\ \int_I g_k(s) ds \end{pmatrix}$.

Lemma 1.20 *Ist d eine Umparametrisierung von c, so gilt*

$$\int\limits_d f = \int\limits_c f$$

für alle Funktionen $f : c(I) \longrightarrow \mathbb{R}^k$ s.d. $f \circ c$ stetig ist.

Beweis: In der Zentralübung. □

Proposition 1.21 *Sei $c : [a, b] \longrightarrow \mathbb{R}^n$ eine regulär parametrisierte Kurve mit $a < b$ in $\mathbb{R}$. Sei $f : c(I) \longrightarrow \mathbb{R}^k$ eine Abbildung mit $f \circ c : [a, b] \longrightarrow \mathbb{R}^k$ stetig. Dann gilt:*

$$\Big|\int\limits_c f\Big| \leq \int\limits_c |f| \leq \big(\max_{t \in [a,b]} |f \circ c(t)|\big) \cdot L[c].$$

Beweis: Für die 1. Ungleichung brauchen wir folgende
Behauptung: *Ist $g : [a, b] \longrightarrow \mathbb{R}^k$ stetig, so gilt*

$$\left| \int_a^b g(s)ds \right| \leq \int_a^b |g(s)|ds.$$

Beweis der Behauptung: Setze $v := \int_a^b g(s)ds$. Dann gilt

$$\begin{aligned} \left| \int_a^b g(s)ds \right|^2 &= \langle \int_a^b g(s)ds, \int_a^b g(s)ds \rangle \\ &= \langle \int_a^b g(s)ds, v \rangle. \end{aligned}$$

Wegen der Linearität des Integrals gilt

$$\langle \int_a^b g(s)ds, w \rangle = \int_a^b \langle g(s), w \rangle ds \qquad \forall\ w \in \mathbb{R}^k.$$

Aus der Cauchy-Schwarz-Ungleichung folgt

$$\int_a^b \langle g(s), w \rangle ds \leq \int_a^b |g(s)| \cdot |w| ds \qquad \forall\ w \in \mathbb{R}^k,$$

insbesondere $\left| \int_a^b g(s)ds \right|^2 \leq \int_a^b |g(s)| \cdot |v| ds = \left(\int_a^b |g(s)|ds \right) \cdot \left| \int_a^b g(s)ds \right|$. Ist $\int_a^b g(s)ds = 0$, so gilt offensichtlich die gesuchte Ungleichung. Sonst folgt aus der letzten Ungleichung oben: $\left| \int_a^b g(s)ds \right| \leq \int_a^b |g(s)|ds$, w.z.b.w. $\surd$

Aus der Behauptung folgt

$$\left| \int_c f \right| \overset{\text{DEF}}{=} \left| \int_a^b f \circ c(s) \cdot |\dot{c}(s)| ds \right| \leq \int_a^b |f \circ c(s)| \cdot |\dot{c}(s)| ds \overset{\text{DEF}}{=} \int_c |f|.$$

Außerdem gilt

$$\begin{aligned} \int_c |f| &= \int_a^b |f \circ c(s)| \cdot |\dot{c}(s)| ds \\ &\leq \int_a^b (\max_{t \in [a,b]} (|f \circ c(t)|)) \cdot |\dot{c}(s)| ds \\ &= \max_{t \in [a,b]} |f \circ c(t)| \cdot \int_a^b |\dot{c}(s)| ds \\ &= \max_{t \in [a,b]} |f \circ c(t)| \cdot L[c]. \end{aligned}$$

$\square$

1.2 Krümmung ebener Kurven

1.2.1 Definition

Definition 1.22 *Sei c eine (C^2) regulär parametrisierte Kurve im $\mathbb{R}^2$, $c : I \longrightarrow \mathbb{R}^2$. Die* Krümmung *von c ist die Abbildung*

$$\begin{aligned} \kappa : I &\longrightarrow \mathbb{R} \\ t &\longmapsto \kappa(t) := \tfrac{\det(\dot{c}(t), \ddot{c}(t))}{|\dot{c}(t)|^3}. \end{aligned}$$

Beachte, dass κ wohldefiniert (wegen $\dot{c}(t) \neq 0 \;\forall\, t \in I$) und C^∞ ist, falls c C^∞ ist.

Beispiele 1.23

1. Sei $c : I \longrightarrow \mathbb{R}^2$, $t \longmapsto p + tv$, wobei $(p, v) \in \mathbb{R}^2 \times (\mathbb{R}^2 \setminus \{0\})$. Wegen $\ddot{c} = 0$ auf I gilt $\kappa = 0$.

2. Sei $c : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(t)}{\sin(t)}$. Wegen $\dot{c}(t) = \binom{-\sin(t)}{\cos(t)}$ und $\ddot{c}(t) = \binom{-\cos(t)}{-\sin(t)}$ gilt
$$\begin{aligned}\kappa(t) &= \frac{\det\begin{pmatrix}-\sin(t) & -\cos(t)\\ \cos(t) & -\sin(t)\end{pmatrix}}{\left|\binom{-\sin(t)}{\cos(t)}\right|^3}\\ &= \frac{+\sin^2(t) + \cos^2(t)}{1}\\ &= 1\end{aligned}$$
für alle $t \in \mathbb{R}$.

3. Sei allgemeiner $c : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto p + r \cdot \binom{\cos(t)}{\sin(t)}$, wobei $p \in \mathbb{R}^2$ und $r \in]0, \infty[$ (das Bild von c ist der Kreis von Radius r um p). Dann ist $\kappa(t) = \frac{1}{r} \;\forall\, t \in \mathbb{R}$.

Proposition 1.24 *Ist $\tilde{c} = c \circ \varphi$ eine Umparametrisierung mit* derselben Orientierung *einer regulär parametrisierten Kurve im $\mathbb{R}^2$, so gilt für die Krümmung $\tilde{\kappa}$ von $\tilde{c}$ bzw. κ von c:*
$$\tilde{\kappa} = \kappa \circ \varphi.$$

Beweis: In der Zentralübung. □

Bemerkung 1.25 Hat $\tilde{c} = c \circ \varphi$ die umgekehrte Orientierung von c (d.h., $\varphi' < 0$), so gilt $\tilde{\kappa} = -\kappa \circ \varphi$.

1.2.2 Frenet-Gleichungen

Definition 1.26 *Sei $c : I \longrightarrow \mathbb{R}^2$ eine regulär parametrisierte Kurve. Das* Normalenfeld *zu c ist die Abbildung*
$$\begin{aligned} n : I &\longrightarrow \mathbb{R}^2\\ t &\longmapsto \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \cdot \frac{\dot{c}(t)}{|\dot{c}(t)|}\end{aligned}$$

Satz 1.27 (Frenet-Gleichungen) *Sei $c : I \longrightarrow \mathbb{R}^2$ eine nach Bogenlänge parametrisierte Kurve. Sei n das Normalenfeld zu c und κ die Krümmung von c. Dann gilt* $\begin{cases}\ddot{c} = \kappa n\\ \dot{n} = -\kappa \dot{c}\end{cases}$, *d.h. in Kurzform:* $\binom{\dot{c}}{n}' = \begin{pmatrix}0 & \kappa\\ -\kappa & 0\end{pmatrix} \cdot \binom{\dot{c}}{n}$.

Beweis: Zu bemerken ist zuerst, dass für jedes $t \in I$ das Paar $(\dot{c}(t), n(t))$ eine positiv-orientierte Orthonormalbasis (ONB) des $\mathbb{R}^2$ bildet: $|\dot{c}(t)| = |n(t)| = 1$, $\langle \dot{c}(t), n(t)\rangle = 0$ und $\det(\dot{c}(t), n(t)) = 1 > 0$. Nun leiten wir die Identität $|\dot{c}|^2 = 1$ nach t ab: für jedes $t \in I$ gilt
$$0 = (|\dot{c}|^2)'(t) = 2\langle \ddot{c}(t), \dot{c}(t)\rangle \qquad \text{(Produktregel 2)},$$

d.h., $\langle \ddot{c}(t), \dot{c}(t)\rangle = 0 \;\forall\, t \in I$. Da $(\dot{c}(t), n(t))$ eine ONB des $\mathbb{R}^2$ ist, existiert (für jedes t) ein $f(t) \in \mathbb{R}$ mit $\ddot{c}(t) = f(t) \cdot n(t)$. Somit gilt

$$\kappa(t) = \frac{\det(\dot{c}(t), \ddot{c}(t))}{|\dot{c}(t)|^3} = \frac{\det(\dot{c}(t), f(t)n(t))}{1} = f(t),$$

d.h., $\ddot{c} = \kappa n$ auf I. Dies zeigt die erste Gleichung. Für die zweite zerlege man $\dot{n}(t)$ in der ONB $(\dot{c}(t), n(t))$:

$$\dot{n}(t) = \langle \dot{n}(t), \dot{c}(t)\rangle \dot{c}(t) + \langle \dot{n}(t), n(t)\rangle n(t).$$

Wegen $0 = (|n|^2)'(t) = 2\langle \dot{n}(t), n(t)\rangle$ gilt $\langle \dot{n}(t), n(t)\rangle = 0$. Außerdem gilt wegen $\langle n, \dot{c}\rangle = 0$:

$$\begin{aligned} 0 &= (\langle n, \dot{c}\rangle)'(t) \\ &= \langle \dot{n}(t), \dot{c}(t)\rangle + \langle n(t), \ddot{c}(t)\rangle \\ &= \langle \dot{n}(t), \dot{c}(t)\rangle + \kappa(t)\underbrace{\langle n(t), n(t)\rangle}_{1}, \end{aligned}$$

insbesondere $\langle \dot{n}(t), \dot{c}(t)\rangle = -\kappa(t)$. Es folgt $\dot{n}(t) = -\kappa(t)\dot{c}(t)$ und die zweite Gleichung. □

Bemerkung 1.28 Insbesondere impliziert dieser Satz, dass – für nach Bogenlänge parametrisierte $c : I \longrightarrow \mathbb{R}^2$ – die Krümmung von c die „Beschleunigung" von c beschreibt: es gilt $\ddot{c}(t) = \kappa(t)n(t)$ und $\ddot{c} : I \longrightarrow \mathbb{R}^2$ ist das Beschleunigungsvektorfeld von c. Ist $\kappa(t) > 0$ für ein $t \in I$, so biegt das „Teilchen" $c(t)$ nach links; ist $\kappa(t) < 0$ für ein $t \in I$, so biegt das Teilchen $c(t)$ nach rechts; ist $\kappa(t) = 0$, so weiß man a priori nicht, in welche Richtung $c(t)$ biegt!

1.2.3 Der Hauptsatz der Theorie der ebenen Kurven

Einschub über euklidische Bewegungen:

- Eine *euklidische Bewegung* im $\mathbb{R}^2$ ist eine Abbildung $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ der Form $F(x) = A \cdot x + a$, wobei $a \in \mathbb{R}^2$ und $A \in M_{2\times 2}(\mathbb{R})$ entweder der Form $A = \left(\begin{smallmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{smallmatrix}\right)$ oder der Form $A = \left(\begin{smallmatrix}\cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{smallmatrix}\right)$ ist für ein $\theta \in \mathbb{R}$.

- Beispiele:

  1. $F(x) = A \cdot x$ mit $A = \left(\begin{smallmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{smallmatrix}\right) \;\forall\, x \in \mathbb{R}^2$. Die Abbildung F ist die Drehung um den Winkel $\theta \in R$ um den Punkt $0 \in \mathbb{R}^2$.
  2. $F(x) = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right) \cdot x \;\forall\, x \in \mathbb{R}^2$ (A ist vom 2. Typ mit $\theta = 0$). Dann ist F die orthogonale Spiegelung an der x-Achse.
  3. $F(x) = x + a$ für alle $x \in \mathbb{R}^2$ (d.h., $A = I_2$: wähle $\theta = 0$ für A vom 1. Typ). Dann ist F die Verschiebung (oder Translation) um den Vektor $a \in \mathbb{R}^2$.

- **Satz:** Die euklidischen Bewegungen im $\mathbb{R}^2$ sind genau die sogenannten *Isometrien* des $\mathbb{R}^2$, d.h. die abstandserhaltenden Abbildungen $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ (wobei „abstandserhaltend" bedeutet, dass $|F(x) - F(y)| = |x - y| \;\forall\, x, y \in \mathbb{R}^2$ gilt).

- Eine euklidische Bewegung im $\mathbb{R}^2$ heißt genau dann *orientierungserhaltend*, wenn sie der Form $F(x) = Ax + a$ ist ($a \in \mathbb{R}^2$), wobei A nur vom 1. Typ ist ($A = \left(\begin{smallmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{smallmatrix}\right)$).

Satz 1.29 (Hauptsatz der Theorie der ebenen Kurven) *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ eine C^∞ Funktion. Dann existiert eine nach Bogenlänge parametrisierte Kurve $c : I \longrightarrow \mathbb{R}^2$ mit Krümmung $\kappa = f$. Diese Kurve ist bis auf Dahinterschaltung von orientierungserhaltenden euklidischen Bewegungen eindeutig: ist $\tilde{c} : I \longrightarrow \mathbb{R}^2$ eine weitere nach Bogenlänge parametrisierte Kurve mit Krümmung $\tilde{\kappa} = f$, so existiert eine eindeutige orientierungserhaltende euklidische Bewegung F im $\mathbb{R}^2$ so, dass $\tilde{c} = F \circ c$ gilt.*

Beweisskizze:
Existenz: Fixiere $t_0 \in I$ und betrachte die Abbildung

$$\begin{array}{rcl} c : I & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & \begin{pmatrix} \int_{t_0}^t \cos(\int_{t_0}^s f(u)du)ds \\ \int_{t_0}^t \sin(\int_{t_0}^s f(u)du)ds \end{pmatrix}. \end{array}$$

Dann ist c eine C^∞ parametrisierte Kurve mit

$$\begin{array}{rrcl} * & \dot{c}(t) & = & \begin{pmatrix} \cos(\int_{t_0}^s f(u)du) \\ \sin(\int_{t_0}^s f(u)du) \end{pmatrix} \ (\Rightarrow |\dot{c}(t)| = 1 \ \forall\, t \in I) \\ * & \ddot{c}(t) & = & f(t) \begin{pmatrix} -\sin(\int_{t_0}^s f(u)du) \\ \cos(\int_{t_0}^s f(u)du) \end{pmatrix} \\ & & = & f(t) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \dot{c}(t). \end{array}$$

Daraus folgt $\kappa(t) = f(t) \ \forall\, t \in I$. Somit ist die Existenz bewiesen.

Eindeutigkeit:

- **Behauptung 1:** Es gibt ein eindeutiges $A \in M_{2\times 2}(\mathbb{R})$ der Form $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ und ein eindeutiges $a \in \mathbb{R}^2$ mit:

$$\begin{cases} A \cdot c(0) + a & = \tilde{c}(0) \\ A \cdot \dot{c}(0) & = \dot{\tilde{c}}(0). \end{cases}$$

  Beweisidee: [Bild: die Vektoren $\dot{c}(0)$ und $\dot{\tilde{c}}(0)$ haben beide Norm 1 und bilden einen Winkel $\theta \in \mathbb{R}$.] $\rightsquigarrow$ $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ erfüllt $A \cdot \dot{c}(0) = \dot{\tilde{c}}(0)$. Setze dann $a := \tilde{c}(0) - A \cdot c(0)$.

- **Behauptung 2:** Es gilt $F \circ c = \tilde{c}$ (für $F(x) := Ax + a$, $\forall\, x \in \mathbb{R}^2$).

  Beweisskizze: Beweise zuerst $\kappa_{F\circ c} = f$ (nachrechnen). Dies impliziert, dass $F \circ c$ dieselben Frenet-Gleichungen löst wie $\tilde{c}$. Nach Konstruktion von F gilt aber $\dot{\overbrace{F \circ c}}(0) = \dot{\tilde{c}}(0)$ sowie $n_{F\circ c}(0) = n_{\tilde{c}}(0)$ (A ist orientierungserhaltend). Aus dem Satz von Picard-Lindelöf (Satz 4.9) folgt dann $\dot{\overbrace{F \circ c}} = \dot{\tilde{c}}$; wegen $F \circ c(0) = \tilde{c}(0)$ folgt auch $F \circ c = \tilde{c}$.

$\square$

Kapitel 2

Differenzierbare Abbildungen in mehreren Veränderlichen

2.1 Metrische Räume

2.1.1 Metrik

Definition 2.1 *Sei X eine Menge. Eine* Metrik *auf X ist eine Abbildung $d : X \times X \longrightarrow [0, \infty[$ mit:*

i) $d(x, y) = 0 \Longleftrightarrow x = y$ *(„d trennt die Punkte")*

ii) $d(x, y) = d(y, x)$ *(d ist symmetrisch)*

iii) $d(x, z) \leq d(x, y) + d(y, z)$ *(d erfüllt die Dreiecksungleichung)*

für alle $x, y, z \in X$.

Definition 2.2 *Ein* metrischer Raum *ist ein Paar (X, d), wobei X eine Menge und d eine Metrik auf X ist.*

Beispiele 2.3

1. Sei $X := \mathbb{R}^2$ und $d : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$, $(x, y) \longmapsto |x - y|$, wobei $|\cdot|$ der Betrag ist. Dann ist (X, d) ein metrischer Raum.

2. Allgemeiner sei $X := \mathbb{R}^n$ und $d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $(x, y) \longmapsto |x - y|$, wobei $|\cdot|$ die euklidische Norm bezeichnet. Dann ist (X, d) ein metrischer Raum. Diese Metrik d heißt die *Standardmetrik* vom $\mathbb{R}^n$.

3. Jede nichtleere Menge X besitzt eine Metrik: die Abbildung

$$\begin{aligned} d^G : X \times X &\longrightarrow [0, \infty[\\ (x, y) &\longmapsto \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{falls } x \neq y, \end{cases} \end{aligned}$$

   definiert eine Metrik auf X. Diese Metrik heißt die *grobe* Metrik auf X.

4. Ist (X, d) ein metrischer Raum und $Y \subset X$ eine beliebige Teilmenge, so ist $(Y, d_{|_{Y \times Y}})$ wieder ein metrischer Raum.

Definition 2.4 *Sei (X, d) ein metrischer Raum, $x \in X$ ein Punkt und $r \in [0, \infty[$ eine nichtnegative Zahl.*

i) Der offene Ball *vom Radius r um x (kurz: der offene r-Ball um x) ist definiert als*

$$B_r(x) := \{y \in X \,|\, d(x, y) < r\}.$$

ii) Der abgeschlossene Ball *vom Radius r um x (kurz: der abgeschlossene r-Ball um x ist definiert als*

$$\overline{B}_r(x) := \{y \in X \,|\, d(x, y) \leq r\}.$$

Beispiele 2.5

1. Für $X := \mathbb{R}$, $d(x, y) := |x - y|$ ist $B_r(x) = \{y \in \mathbb{R} \,|\, |x - y| < r\}$, d.h. $B_r(x) =]x - r, x + r[$ und $\overline{B}_r(x) = [x - r, x + r]$.

2. Für $X := \mathbb{R}^n$, $d(x, y) := |x - y|$ ist $B_r(x) = \{y \in \mathbb{R}^n \,|\, |x - y| < r\}$ bzw. $\overline{B}_r(x) = \{y \in \mathbb{R}^n \,|\, |x - y| \leq r\}$ der „übliche" offene bzw. abgeschlossene r-Ball um x. Z.B. für $n = 2$ ist $B_r(x)$ die „offene" Kreisscheibe vom Radius r um x (ohne den Randkreis $\{y \in \mathbb{R}^2 \,|\, |x - y| = r\}$) und $\overline{B}_r(x)$ die „abgeschlossene" Kreisscheibe vom Radius r um x (mit dem Randkreis).

3. Ist $Y \subset X$ eine Teilmenge, wobei (X, d) ein metrischer Raum ist, so ist für jedes $x \in Y$ und jedes $r > 0$:

$$B_r^Y(x) := \{y \in Y \,|\, d(x, y) < r\} = B_r^X(x) \cap Y$$

sowie $\overline{B}_r^Y(x) = \overline{B}_r^X(x) \cap Y$.

Definition 2.6 *Sei (X, d) ein metrischer Raum.*

i) Sei $x \in X$. Eine Teilmenge U von X heißt Umgebung *von x in X g.d.w. ein $r > 0$ existiert mit $B_r(x) \subset U$.*

ii) Eine Teilmenge U von X heißt genau dann offen *in X, wenn sie eine Umgebung jedes ihrer Punkte ist, d.h. $\forall\, x \in U$, $\exists\, r > 0$ mit $B_r(x) \subset U$.*

iii) Eine Teilmenge A von X heißt genau dann abgeschlossen *in X, wenn $A^c := X \setminus A$ (Komplement von A in X) offen in X ist.*

Beispiele 2.7

1. Jedes offene Intervall ist offen in $\mathbb{R}$.

2. Allgemeiner ist jeder offene Ball eines metrischen Raumes offen in X (in der Zentralübung).

3. Jeder abgeschlossene Ball eines metrischen Raumes X ist abgeschlossen in X (in der Zentralübung).

3'. Das Intervall $[0, 1[$ ist weder offen noch abgeschlossen in $\mathbb{R}$ (mit der Standardmetrik).

4. Ist $Y \subset X$ eine Teilmenge eines metrischen Raumes X, so gilt:

- $U \subset X$ offen in $X \Longrightarrow U \cap Y$ offen in Y,
- $A \subset X$ abgeschlossen in $X \Longrightarrow A \cap Y$ abgeschlossen in Y.

Proposition 2.8 *Sei (X,d) ein metrischer Raum.*

i) $\varnothing$ und X sind offen und abgeschlossen in X.

ii) Ist $(U_i)_{i\in I}$ eine beliebige *Familie offener Teilmengen von X, so ist $\bigcup\limits_{i\in I} U_i$ (ihre Vereinigung) wieder offen in X.*

iii) Ist $(U_i)_{1\le i\le k}$ eine endliche *Familie offener Teilmengen von X, so ist $\bigcap\limits_{i=1}^{k} U_i$ (ihr Durchschnitt) wieder offen in X.*

Beweis: In der Zentralübung. □

Definition 2.9 *Sei (X,d) ein metrischer Raum und $Y \subset X$ eine Teilmenge.*

i) Das Innere *von Y in X ist die Menge*

$$\overset{\circ}{Y} := \{y \in Y \mid \exists\, r > 0 \text{ s.d. } B_r^X(y) \subset Y\}.$$

ii) Der Abschluss *von Y in X ist die Menge*

$$\overline{Y} := \{y \in X \mid \forall\, r > 0, B_r^X(y) \cap Y \neq \varnothing\}.$$

iii) Der Rand *von Y in X ist die Menge $\partial Y := \overline{Y} \setminus \overset{\circ}{Y}$.*

Beispiele 2.10

1. Für $X := \mathbb{R}$ mit der Standardmetrik und $Y :=]a,b[$ oder $[a,b[$ oder $]a,b]$ oder $[a,b]$ (mit $a < b$ reell) gilt: $\overset{\circ}{Y} =]a,b[$, $\overline{Y} = [a,b]$ und $\partial Y = \{a,b\}$.

2. Allgemeiner sei $X := \mathbb{R}^n$ und $Y := B_r(x)$ oder $\overline{B}_r(x)$ für ein $x \in X$ und ein $r > 0$. Dann ist $\overset{\circ}{Y} = B_r(x)$, $\overline{Y} = \overline{B}_r(x)$ und $\partial Y = S_r(x) := \{y \in \mathbb{R}^n \mid |x - y| = r\}$ (r-Sphäre um x in $\mathbb{R}^n$).

3. Nach Definition gilt $\overset{\circ}{\varnothing} = \varnothing = \overline{\varnothing}$ sowie $\overset{\circ}{X} = X = \overline{X}$, für jeden metrischen Raum X.

Proposition 2.11 *Sei (X,d) ein metrischer Raum und $Y \subset X$ eine Teilmenge von X. Dann gilt:*

i) Das Innere von Y in X ist die größte Teilmenge von X, die in Y enthalten ist: $\overset{\circ}{Y} \subset Y$, $\overset{\circ}{Y}$ ist offen in X und $\forall\, U \subset X$ offen in X mit $U \subset Y$ gilt $U \subset \overset{\circ}{Y}$. Insbesondere ist Y selbst genau dann offen in X, wenn $\overset{\circ}{Y} = Y$ gilt.

ii) Der Abschluss von Y in X ist die kleinste abgeschlossene Teilmenge von X, die Y enthält: $\overline{Y} \supset Y$, $\overline{Y}$ ist abgeschlossen in X und $\forall\, A \subset X$ abgeschlossen in X mit $Y \subset A$ gilt $\overline{Y} \subset A$. Insbesondere ist Y selbst genau dann abgeschlossen in X, wenn $\overline{Y} = Y$ gilt.

iii) Der Rand von Y ist abgeschlossen in X.

Beweis: In der Zentralübung oder in den Übungen. □

Bemerkungen 2.12

1. Die Begriffe Inneres, Abschluss und Rand hängen vom Raum X ab: ist z.B. $Y = [0,1] \subset X = \mathbb{R}$, so gilt bzgl. der Standardmetrik $d : (x,y) \mapsto |x-y|$: $\overset{\circ}{Y} =]0,1[$ in X aber $\overset{\circ}{Y} = Y$ in Y! Insbesondere muss darauf geachtet werden, bzgl. welcher Menge das Innere (oder der Abschluss, der Rand) betrachtet wird. In diesem Beispiel müsste man somit verschiedene Bezeichnungen, z.B. $\overset{\circ,X}{Y}$ bzw. $\overset{\circ,Y}{Y}$, verwenden.

2. Diese Begriffe hängen auch von der Metrik d ab. Z.B. gilt für $d = d^G$ aus Beispiel 2.3.3 und jedes $x \in X$: $B_1(x) = \{x\}$, $\overline{B}_1(x) = X$ und $\overset{\circ}{\overline{B_1}}(x) = X$, $\overline{B_1(x)} = \{x\}$. Ist $X = \mathbb{R}^n$ und $d' : (x,y) \longmapsto |x-y|$ die Standardmetrik, so gilt aber $B_1(x) \neq \{x\}$. Auch in diesem Beispiel müsste man die Metrik in die Bezeichnung $B_r(x)$ einbinden. In den meisten Fällen wird die Metrik aber fixiert, deshalb lassen wir das Symbol „d “weg für die Bälle.

3. Es gilt aber stets $B_r(x) \subset \overset{\circ}{\overline{B}_r(x)}$ sowie $\overline{B_r(x)} \subset \overline{B}_r(x)$ (in den Übungen oder in der Zentralübung).

Proposition 2.13 *Seien (X, d_X), (Y, d_Y) metrische Räume. Dann definiert*

$$\begin{array}{rcl} d : (X \times Y) \times (X \times Y) & \longrightarrow & [0, \infty[\\ ((x_1, y_1), (x_2, y_2)) & \longmapsto & \max(d_X(x_1, x_2), d_Y(y_1, y_2)) \end{array}$$

eine Metrik auf $X \times Y$. Sie heißt die Produktmetrik *von d_X und d_Y.*

Beweis: In der Zentralübung oder in den Übungen. □

2.1.2 Konvergente Folgen

Definition 2.14 *Sei (X,d) ein metrischer Raum und $(x_n)_{n\in\mathbb{N}}$ eine Folge aus X ($x_n \in X$ $\forall\ n \in \mathbb{N}$). Sei $\overline{x} \in X$. Wir sagen, die Folge $(x_n)_{n\in\mathbb{N}}$* konvergiert *(in X) gegen $\overline{x}$ g.d.w. $d(x_n, \overline{x}) \underset{n\to\infty}{\longrightarrow} 0$, d.h. $\forall\, \varepsilon > 0$, $\exists\ N \in \mathbb{N}$ s.d. $\forall\, n \geq N$, $d(x_n, \overline{x}) < \varepsilon$. In diesem Fall heißt $\overline{x}$* Grenzwert *der Folge $(x_n)_{n\in\mathbb{N}}$. Wir schreiben $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$ oder $\overline{x} = \lim\limits_{n\to\infty} x_n$.*

Bemerkung 2.15 Existiert ein Grenzwert für die Folge $(x_n)_{n\in\mathbb{N}}$ aus X, so ist dieser Grenzwert eindeutig: gelten nämlich $d(x_n, \overline{x}) \underset{n\to\infty}{\longrightarrow} 0$ und $d(x_n, \overline{x}') \underset{n\to\infty}{\longrightarrow} 0$, so gilt nach der Dreiecksungleichung: $d(\overline{x}, \overline{x}') \leq d(\overline{x}, x_n) + d(x_n, \overline{x}') = d(x_n, \overline{x}) + d(x_n, \overline{x}')$ für alle $n \in \mathbb{N}$. Lässt man $n \to \infty$, so bekommt man $0 \leq d(\overline{x}, \overline{x}') \leq \lim\limits_{n\to\infty} (d(x_n, \overline{x}) + d(x_n, \overline{x}')) = 0$, insbes. $d(\overline{x}, \overline{x}') = 0$. Aus der Definition einer Metrik folgt dann $\overline{x} = \overline{x}'$.

Proposition 2.16 *Sei (X,d) ein metrischer Raum, $Y \subset X$ eine Teilmenge und $\overline{y} \in X$. Dann gilt: $\overline{y} \in \overline{Y} \Longleftrightarrow \exists$ Folge $(y_n)_{n\in\mathbb{N}}$ aus Y mit $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$.*

Beweis:

"$\Rightarrow$" Gilt $\overline{y} \in \overline{Y}$, so betrachte für jedes $n \in \mathbb{N}$, $n \geq 1$, den Durchschnitt $B^X_{\frac{1}{n}}(\overline{y}) \cap Y$; wegen $\overline{y} \in \overline{Y}$ existiert (mindestens) ein $y_n \in B^X_{\frac{1}{n}}(\overline{y}) \cap Y$. Per Konstruktion bekommen wir eine Folge $(y_n)_{n\in\mathbb{N}}$ aus Y mit $d(y_n, \overline{y}) < \frac{1}{n}$ $\forall\, n \geq 1$, insbesondere $d(y_n, \overline{y}) \underset{n\to\infty}{\longrightarrow} 0$. Somit gilt $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$.

"$\Leftarrow$" Sei umgekehrt $\overline{y} \in X$ so, dass eine Folge $(y_n)_{n\in\mathbb{N}}$ aus Y existiert mit $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$. Sei $r > 0$. Dann existiert (mindestens) ein $n \in \mathbb{N}$, $n \geq 1$, mit $d(y_n, \overline{y}) < r$, also ein $y_n \in B^X_r(\overline{y}) \cap Y$. Insbesondere ist $B^X_r(\overline{y}) \cap Y \neq \varnothing$. Dies gilt für alle $r > 0$, somit ist $\overline{y} \in \overline{Y}$. □

2.1.3 Stetige Abbildungen

Definition 2.17 *Seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \longrightarrow Y$ eine Abbildung.*

i) *f heißt* stetig *in einem* Punkt *$x \in X$ g.d.w. für jede Folge $(x_n)_{n\in\mathbb{N}}$ aus X mit $x_n \underset{n\to\infty}{\longrightarrow} x$ gilt $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$.*

ii) *f heißt* stetig *auf X g.d.w. f ist stetig in jedem $x \in X$.*

Proposition 2.18 *Sei $f : X \longrightarrow Y$ eine Abbildung zwischen metrischen Räumen und $x \in X$. Dann sind folgende Aussagen äquivalent:*

i) f ist stetig in x.

ii) $\forall\, r > 0$, $\exists\, \delta > 0$ s.d. $f(B^X_\delta(x)) \subset B^Y_r(f(x))$.

iii) $\forall\, V \subset Y$ Umgebung von $f(x)$ in Y, die Teilmenge $f^{-1}(V) \subset X$ ist eine Umgebung von x in X.

Beweis:[1]
"*i) $\Rightarrow$ ii)*": Angenommen f sei stetig in x. Sei $r > 0$. Gäbe es kein $\delta > 0$ s.d. $f(B^X_\delta(x)) \subset B^Y_r(f(x))$, so betrachte insbesondere die Familie $(\delta_n = \frac{1}{n})_{n\geq 1}$; wegen $f(B^X_{\delta_n}(x)) \not\subset B^X_r(f(x))$ existiert $x_n \in B^X_{\delta_n}(x)$ mit $f(x_n) \notin B^Y_r(f(x))$, für jedes $n \in \mathbb{N}$, $n \geq 1$. Wegen $\delta_n \underset{n\to\infty}{\longrightarrow} 0$ gilt $x_n \underset{n\to\infty}{\longrightarrow} x$; wegen $f(x_n) \notin B^X_r(f(x))$ gilt $d_Y(f(x_n), f(x)) > r > 0$ $\forall\, n \geq 1$, insbesondere $f(x_n) \underset{n\to\infty}{\nrightarrow} f(x)$, Widerspruch zur Stetigkeit von f in x. Daraus folgt, dass ein solches $\delta > 0$ existiert.
"*ii) $\Rightarrow$ iii)*": Angenommen, $\forall\, r > 0$, $\exists\, \delta > 0$ s.d. $f(B^X_\delta(x)) \subset B^Y_r(f(x))$. Sei $V \subset Y$ eine beliebige Umgebung von $f(x)$ in Y. Nach Definition existiert ein $r > 0$ s.d. $B^Y_r(f(x)) \subset V$. Nach Voraussetzung existiert ein $\delta > 0$ s.d. $f(B^X_\delta(x)) \subset B^Y_r(f(x))$; insbesondere gilt

$$f^{-1}(V) \supset f^{-1}(B^Y_r(f(x)) \supset f^{-1}(f(B^X_\delta(x))) \supset B^X_\delta(x),$$

[1]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

was zeigt, dass $f^{-1}(V)$ eine Umgebung von x in X ist.
"*iii)* $\Rightarrow$ *i)*": Angenommen, $f^{-1}(V)$ ist eine Umgebung von x in X für jede Umgebung V von $f(x)$ in Y. Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus X mit $x_n \underset{n\to\infty}{\longrightarrow} x$. Sei $r > 0$ beliebig. Da $B_r^Y(f(x))$ eine Umgebung von $f(x)$ in Y ist, muss $f^{-1}(B_r^Y(f(x)))$ eine Umgebung von x in X sein, d.h., $\exists\,\delta > 0$ s.d. $B_\delta^X(x) \subset f^{-1}(B_r^Y(f(x)))$. Wegen $x_n \underset{n\to\infty}{\longrightarrow} x$ existiert ein $N \in \mathbb{N}$ mit $x_n \in B_\delta^X(x)$ für alle $n \geq N$; wegen $f(B_\delta^X(x)) \subset f(f^{-1}(B_r^Y(f(x))) \subset B_r^Y(f(x))$ gilt dann $f(x_n) \in B_r^Y(f(x))\ \forall n \geq N$, d.h., $d_Y(f(x_n), f(x)) \leq r\ \forall n \geq N$. Dies zeigt: $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$. □

Proposition 2.19 *Sei $f : X \longrightarrow Y$ eine Abbildung zwischen metrischen Räumen. Dann sind folgende Aussagen äquivalent:*

i) *f ist stetig auf X.*

ii) *$\forall\, U \subset Y$ offen in Y, die Teilmenge $f^{-1}(U) \subset X$ ist offen in X.*

iii) *$\forall\, A \subset Y$ abgeschlossen in Y, die Teilmenge $f^{-1}(A) \subset X$ ist abgeschlossen in X.*

Beweis:[2]
"*i)* $\Rightarrow$ *ii)*": Angenommen, f sei stetig auf X. Sei $U \subset Y$ offen in Y und sei $x \in f^{-1}(U)$ beliebig. Da U offen ist, ist U insbesondere eine Umgebung von $f(x) \in U$. Aus Proposition 2.18 folgt, dass dann $f^{-1}(U)$ eine Umgebung von x in X ist. Insgesamt ist $f^{-1}(U)$ eine Umgebung jedes seiner Punkte, d.h., $f^{-1}(U)$ ist offen in X.
"*ii)* $\Rightarrow$ *iii)*": Angenommen, $f^{-1}(U)$ sei offen für jede offene Teilmenge U von Y. Sei $A \subset Y$ abgeschlossen, dann ist $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ offen in X, da $Y \setminus A$ offen in Y ist. Somit ist $f^{-1}(A)$ abgeschlossen in X.
"*iii)* $\Rightarrow$ *i)*": Angenommen, $f^{-1}(A)$ sei abgeschlossen für jede abgeschlossene Teilmenge A von Y. Sei $x \in X$ und $V \subset Y$ eine Umgebung von $f(x)$ in Y, dann existiert ein $r > 0$ mit $B_r^Y(f(x)) \subset V$. Da $Y \setminus B_r^Y(f(x))$ abgeschlossen ist, ist $f^{-1}(Y \setminus B_r^Y(f(x))) = X \setminus f^{-1}(B_r^Y(f(x)))$ abgeschlossen in X, d.h., $f^{-1}(B_r^Y(f(x)))$ ist offen in X. Wegen $x \in f^{-1}(B_r^Y(f(x)))$ ist dann $f^{-1}(B_r^Y(f(x)))$ eine Umgebung von x in X, somit ist $f^{-1}(V)$ auch eine Umgebung von x in X. □

Proposition 2.20 *Seien $f : X \longrightarrow Y$ und $g : Y \longrightarrow Z$ Abbildungen zwischen metrischen Räumen und $x \in X$ ein Punkt. Ist f stetig in x und g stetig in $f(x)$, so ist $g \circ f$ stetig in x.*

Beweis:[3] Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge aus X mit $x_n \underset{n\to\infty}{\longrightarrow} x$, dann gilt $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$ (f ist stetig in x) und somit $g(f(x_n)) \underset{n\to\infty}{\longrightarrow} g(f(x))$ (g ist stetig in $f(x)$). □

Beispiele 2.21

1. Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto xy$, ist stetig auf $\mathbb{R}^2$ (wobei $\mathbb{R}^2$ die Standardmetrik trägt). Denn: konvergiert eine Folge $((x_n, y_n))_{n\in\mathbb{N}}$ gegen ein $(x, y) \in \mathbb{R}^2$, so gelten $x_n \underset{n\to\infty}{\longrightarrow} x$ und $y_n \underset{n\to\infty}{\longrightarrow} y$, insbesondere $x_n y_n \underset{n\to\infty}{\longrightarrow} xy$.

[2]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[3]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

2. Sind (X, d_X) und (Y, d_Y) metrische Räume, so sind die Projektionen

$$\begin{array}{rcl} \Pi_X : X \times Y & \longrightarrow & X \\ (x,y) & \longmapsto & x \end{array}$$

und

$$\begin{array}{rcl} \Pi_Y : X \times Y & \longrightarrow & Y \\ (x,y) & \longmapsto & y \end{array}$$

stetig. Denn analog gilt: konvergiert eine Folge $((x_n, y_n))_{n\in\mathbb{N}}$ aus $X \times Y$ gegen ein $(x, y) \in X \times Y$, so gelten $x_n \underset{n\to\infty}{\longrightarrow} x$ sowie $y_n \underset{n\to\infty}{\longrightarrow} y$ (Beweis in der Zentralübung oder in den Übungen).

3. Sei $f : X \longrightarrow Y \times Z$, $x \longmapsto (f_Y(x), f_Z(x))$, eine Abbildung, wobei $f_Y : X \longrightarrow Y$, $f_Z : X \longrightarrow Z$ Abbildungen und X, Y, Z metrische Räume sind. Dann gilt: ist f ist stetig in einem Punkt $x \in X$, so sind f_Y und f_Z stetig in x (Beweis in der Zentralübung oder in den Übungen).

4. Ist $f : X \times Y \longrightarrow Z$ eine stetige Abbildung, so sind, für jedes $x_0 \in X$ und jedes $y_0 \in Y$ die sogenannten *partiellen Abbildungen*

$$\begin{array}{rcl} X & \longrightarrow & Z \\ x & \longmapsto & f(x, y_0) \end{array}$$

und

$$\begin{array}{rcl} Y & \longrightarrow & Z \\ y & \longmapsto & f(x_0, y) \end{array}$$

stetig. Die Umkehrung gilt aber i.A. nicht! Siehe nächstes Beispiel.

5. Die Abbildung

$$\begin{array}{rcl} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} \frac{xy}{x^2+y^2} & \text{falls } (x,y) \neq (0,0) \\ 0 & \text{falls } (x,y) = (0,0) \end{cases} \end{array},$$

ist wohldefiniert, ist stetig auf $\mathbb{R}^2 \setminus \{0\}$ (argumentiere mit Folgen), ist aber nicht stetig in $0(= (0,0))$, denn: die Folge $(X_n := \binom{\frac{1}{n}}{\frac{1}{n}})_{n\geq 1}$ konvergiert gegen 0 in $\mathbb{R}^2$ (wegen $\frac{1}{n} \underset{n\to\infty}{\longrightarrow} 0$), aber $f(X_n) = \frac{1}{2} \underset{n\to\infty}{\not\rightarrow} 0$. Bemerke, dass für jedes $(x_0, y_0) \in \mathbb{R}^2$ die Abbildungen $\mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto f(x, y_0)$ sowie $\mathbb{R} \longrightarrow \mathbb{R}$, $y \longmapsto f(x_0, y)$, stetig sind!

6. Sei (X, d) ein metrischer Raum. Dann ist $d : X \times X \longrightarrow [0, \infty[$ stetig, wobei $X \times X$ die Produktmetrik trägt, denn: konvergiert eine Folge $(X_n = (x_n, y_n))_{n\in\mathbb{N}}$ gegen ein $\overline{X} = (\overline{x}, \overline{y}) \in X \times X$, so gilt (per Definition) $d_{X\times X}((x_n, y_n), (\overline{x}, \overline{y})) \underset{n\to\infty}{\longrightarrow} 0$, insbesondere $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$ und $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$. Daraus folgt

$$\begin{array}{rcl} |d(x_n, y_n) - d(\overline{x}, \overline{y})| & \underset{\text{(Dreiecksungleichung)}}{\leq} & |d(x_n, y_n) - d(\overline{x}, y_n)| + |d(\overline{x}, y_n) - d(\overline{x}, \overline{y})| \\ & \leq & d(x_n, \overline{x}) + d(y_n, \overline{y}) \underset{n\to\infty}{\longrightarrow} 0. \end{array}$$

Somit gilt $d(x_n, y_n) \underset{n\to\infty}{\longrightarrow} d(\overline{x}, \overline{y})$.

2.1.4 Kompakte metrische Räume

Definition 2.22 *Ein metrischer Raum (X, d) heißt genau dann* kompakt*, wenn jede Folge aus X eine konvergente Teilfolge besitzt, d.h. $\forall\ (x_n)_{n\in\mathbb{N}}$ mit $x_n \in X$, $\exists\, \varphi : \mathbb{N} \longrightarrow \mathbb{N} \nearrow$ (monoton wachsend) und ein $x \in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} x$.*

Beispiele 2.23

1. Jeder endliche metrische Raum (X, d) ist kompakt. Dabei beachte man, dass $\varnothing$ per Konvention kompakt ist.

2. Der Raum $\mathbb{R}$ mit der Standardmetrik ist nicht kompakt. Denn: betrachte z.B. die Folge $(x_n := n)_{n\in\mathbb{N}}$, dann ist jede Teilfolge von $(x_n)_{n\in\mathbb{N}}$ streng monoton wachsend und unbeschränkt: für jede monoton wachsende Abbildung $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ gilt $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \infty$. Insbesondere besitzt $(x_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge.

3. Allgemeiner ist $\mathbb{R}^n$ mit der Standardmetrik nicht kompakt.

Proposition 2.24 (Beschränktheit) *Ist (X, d) ein kompakter metrischer Raum, so ist (X, d)* beschränkt*, d.h., $\sup\limits_{x,y\in X} (d(x, y)) < \infty$.*

Beweis:[4] Angenommen, (X, d) wäre nicht beschränkt. Dann gäbe es Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ aus X mit $d(x_n, y_n) \underset{n\to\infty}{\longrightarrow} \infty$. Da (X, d) kompakt ist, existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_n$ und ein $\overline{x} \in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$; wiederum wegen (X, d) kompakt existiert eine Teilfolge $(y_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ von $(y_{\varphi(n)})_{n\in\mathbb{N}}$ und ein $\overline{y} \in X$ mit $y_{\varphi\circ\psi(n)} \underset{n\to\infty}{\longrightarrow} \overline{y}$. Bemerke, dass auch $x_{\varphi\circ\psi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$ gilt, da $(x_{\varphi(n)})_{n\in\mathbb{N}}$ bereits gegen $\overline{x}$ konvergiert. Daraus folgt, dass die Folge $(x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ aus $X \times X$ gegen $(\mathit{overlinex}, \overline{y})$ konvergiert, wobei $X \times X$ die Produktmetrik trägt. Da die Metrik stetig auf $X \times X$ ist (Beispiel 2.21.6), folgt $d(x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}) \underset{n\to\infty}{\longrightarrow} d(\overline{x}, \overline{y}) < \infty$, Widerspruch zu $d(x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}) \underset{n\to\infty}{\longrightarrow} \infty$. Somit muss (X, d) beschränkt sein. $\square$

Proposition 2.25 (kompakte Teilmengen) *Sei (X, d) ein metrischer Raum und $Y \subset X$ eine beliebige Teilmenge.*

i) Ist $(Y, d_{|Y\times Y})$ kompakt, so ist Y abgeschlossen in X.

ii) Ist (X, d) kompakt und $Y \subset X$ abgeschlossen in X, so ist $(Y, d_{|Y\times Y})$ kompakt.

Beweis:[5]
i) Sei $\overline{y} \in X$ so, dass eine Folge $(y_n)_{n\in\mathbb{N}}$ aus Y existiert mit $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$. Z.z. ist $\overline{y} \in Y$. Wegen $(Y, d_{|Y\times Y})$ kompakt existiert eine Teilfolge $(y_{\varphi(n)})_{n\in\mathbb{N}}$ von $(y_n)_{n\in\mathbb{N}}$ und ein $\hat{y} \in Y$ mit $y_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \hat{y}$. Da aber $(y_n)_{n\in\mathbb{N}}$ gegen $\overline{y}$ (in X) konvergiert, muss jede Teilfolge von $(y_n)_{n\in\mathbb{N}}$ gegen $\overline{y}$ konvergieren, insbesondere muss $y_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{y}$ gelten. Die Eindeutigkeit des Grenzwertes liefert $\hat{y} = \overline{y}$ gelten, woraus $\overline{y} \in Y$ folgt. Somit gilt $\overline{Y} = Y$, d.h. Y ist abgeschlossen in X (siehe Proposition 2.11).

[4] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[5] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

ii) Sei $(y_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus Y. Wegen (X,d) kompakt existiert eine Teilfolge $(y_{\varphi(n)})_{n\in\mathbb{N}}$ von $(y_n)_{n\in\mathbb{N}}$ und ein $\overline{y}\in X$ mit $y_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{y}$ (in X). Da $Y\subset X$ abgeschlossen in X ist, muss $\overline{y}$ in Y liegen (siehe Propositionen 2.16 und 2.11). Somit besitzt $(y_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge in $(Y, d_{|_{Y\times Y}})$. Insbesondere ist $(Y, d_{|_{Y\times Y}})$ kompakt. □

Proposition 2.26 (Vereinigung und Durchschnitt kompakter Teilmengen) *Sei (X,d) ein metrischer Raum.*

i) Ist $(K_i)_{i=1,\dots,p}$ eine endliche *Familie kompakter Teilmengen von X, so ist $\bigcup_{i=1}^{p} K_i$ wieder kompakt (für die induzierte Metrik).*

ii) Ist $(K_i)_{i\in I}$ eine beliebige *Familie kompakter Teilmengen von X, so ist $\bigcap_{i\in I} K_i$ wieder kompakt (für die induzierte Metrik).*

iii) Ist $(K_n)_{n\in\mathbb{N}}$ eine Folge nichtleerer *kompakter Teilmengen von X mit $K_{n+1}\subset K_n$ für alle $n\in\mathbb{N}$, so ist $\bigcap_{n\in\mathbb{N}} K_n$ kompakt und* nichtleer.

Beweis:[6]
i) Induktion über p. Der Fall $p=1$ ist trivial. Für den Induktionsschritt $p\longrightarrow p+1$ reicht es, den Fall $p=2$ zu beweisen. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge aus $K_1\cup K_2$. Dann enthält K_1 oder K_2 unendlich viele Folgenglieder: $\exists\, i\in\{1,2\}$ mit $|\{n\in\mathbb{N}\,|\,x_n\in K_i\}|=\infty$. O.B.d.A. sei $i=1$. Dann existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $x_{\varphi(n)}\in K_1$ für alle $n\in\mathbb{N}$. Wegen $(K_1, d_{|_{K_1\times K_1}})$ kompakt existiert eine Teilfolge $(x_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ von $(x_{\varphi(n)})_{n\in\mathbb{N}}$ und ein $\overline{x}_1\in K_1$ mit $x_{\varphi\circ\psi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x_1}$. Insbesondere konvergiert eine Teilfolge $(x_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ gegen ein $\overline{x_1}\in K_1\subset K_1\cup K_2$. Dies beweist die Kompaktheit von $K_1\cup K_2$.
ii) Die Teilmenge $\bigcap_{i\in I} K_i$ ist als Durchschnitt einer beliebigen Familie abgeschlossener Teilmengen abgeschlossen in X. Ist $I=\varnothing$, so ist per Definition $\bigcap_{i\in I} K_i=\varnothing$ und $\varnothing$ ist kompakt. Ist $I\neq\varnothing$, so gibt es ein $i_0\in I$, insbesondere ist $\bigcap_{i\in I} K_i\subset K_{i_0}$ auch abgeschlossen in K_{i_0}. Da K_{i_0} kompakt ist, folgt aus Proposition 2.25, dass $\bigcap_{i\in I} K_i$ kompakt ist.
iii) Wir wissen schon aus *ii*), dass $\bigcap_{n\in\mathbb{N}} K_n$ kompakt ist. Z.z. ist $\bigcap_{n\in\mathbb{N}} K_n\neq\varnothing$. Nach Voraussetzung existiert ein $x_n\in K_n$ für alle $n\in\mathbb{N}$. Die Folge $(x_n)_{n\in\mathbb{N}}$ liegt in K_0 wegen $K_{k+1}\subset K_k$ für alle k. Wegen K_0 kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ und ein $\overline{x}\in K_0$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$. Nun gilt, für jedes $p\in\mathbb{N}$: $x_q\in K_p$ für alle $q\geq p$ (wegen $K_q\subset K_p$), insbesondere auch $x_{\varphi(q)}\in K_p\ \forall\ q\geq p$. Die Folge $(x_{\varphi(q)})_{q\geq p}$ konvergiert aber gegen $\overline{x}$; da $K_p\subset K_0$ abgeschlossen ist (Proposition 2.25), muss dann $\overline{x}\in K_p$ gelten. Dies zeigt $\overline{x}\in\bigcap_{p\in\mathbb{N}} K_p$, insbesondere ist $\bigcap_{n\in\mathbb{N}} K_n\neq\varnothing$. □

Proposition 2.27 (Produkte kompakter Räume) *Sind (X,d_X) und (Y,d_Y) zwei kompakte metrische Räume, so ist $(X\times Y, d_{X\times Y})$ kompakt.*

[6] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Beweis:[7] Sei $((x_n, y_n))_{n\in\mathbb{N}}$ eine beliebige Folge aus $X \times Y$. Dann ist $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus X. Wegen (X, d) kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ und ein $\overline{x} \in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$. Wegen (Y, d_Y) kompakt existiert eine Teilfolge $(y_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ von $(y_{\varphi(n)})_{n\in\mathbb{N}}$ und ein $\overline{y} \in Y$ mit $y_{\varphi\circ\psi(n)} \underset{n\to\infty}{\longrightarrow} \overline{y}$. Daraus folgt $d_{X\times Y}((x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}), (\overline{x}, \overline{y})) = \max(d_X(x_{\varphi\circ\psi(n)}, \overline{x}),\ d_Y(y_{\varphi\circ\psi(n)}, \overline{y})) \underset{n\to\infty}{\longrightarrow} 0$, d.h., $(x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}) \underset{n\to\infty}{\longrightarrow} (\overline{x}, \overline{y})$. Somit ist $(X \times Y, d_{X\times Y})$ kompakt. □

Satz 2.28 (Heine-Borel) *Sei $\mathbb{R}^n$ mit der Standardmetrik versehen. Eine Teilmenge K von $\mathbb{R}^n$ ist* genau dann *kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis:[8] "⇒": Ist K kompakt, so ist $(K, d_{|K\times K})$ beschränkt nach Proposition 2.24. Außerdem ist K abgeschlossen in $\mathbb{R}^n$ nach Proposition 2.25.
"⇐": Sei zuerst $n = 1$. Es reicht, zu zeigen, dass jedes Intervall der Form $[a, b]$ (mit $a \leq b$ reell) kompakt ist, denn jede beschränkte Teilmenge von $\mathbb{R}$ ist in einem solchen Intervall enthalten (und Proposition 2.25 impliziert, dass jede abgeschlossene Teilmenge eines kompakten metrischen Raumes kompakt ist). Dass $[a, b]$ kompakt ist, ist aber genau die Aussage des Satzes von Bolzano-Weierstraß. Der Vollständigkeit halber beweisen wir diesen Satz jetzt. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge aus $K = [a, b]$. Setze $a_0 := a$, $b_0 := b$ und definiere induktiv das Intervall $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ als eines der beiden Intervalle $[a_n, \frac{a_n+b_n}{2}]$ und $[\frac{a_n+b_n}{2}, b_n]$, das immer noch *unendlich viele* Elemente der Folge $(x_n)_{n\in\mathbb{N}}$ enthält (mindestens eines der beiden muss diese Eigenschaft haben, sonst wäre $\mathbb{N}$ endlich). Es kann passieren, dass *beide* unendlich viele Elemente enthalten, wir wählen ggf. *eines* der beiden. Betrachte dann die reellen Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$. Aus der Konstruktion der a_n und b_n folgt, dass $(a_n)_{n\in\mathbb{N}}$ bzw. $(b_n)_{n\in\mathbb{N}}$ monoton wachsend bzw. fallend ist. Es gelten auch: $a_n \leq b_n \leq b$ und $b_n \geq a_n \geq a$ für alle $n \in \mathbb{N}$, insbesondere sind $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ *beschränkt.* Aus den grundsätzlichen Eigenschaften von $\mathbb{R}$ folgt, dass *jede monotone beschränkte reelle Folge konvergiert.* Insbesondere konvergieren $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$. Wegen

$$b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1}) = \ldots = \frac{1}{2^n}(b_0 - a_0) \underset{n\to+\infty}{\longrightarrow} 0$$

gilt $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} b_n =: l \in \mathbb{R}$, und wegen der Monotonie von $(a_n)_{n\in\mathbb{N}}$ bzw. von $(b_n)_{n\in\mathbb{N}}$ gilt auch $l \geq a_0 = a$ bzw. $l \leq b_0 = b$, d.h. $l \in [a, b]$. Was ist jetzt mit der Folge $(x_n)_{n\in\mathbb{N}}$? Nach der Konstruktion der Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ existiert, für jedes $n \in \mathbb{N}$, ein Glied $x_{\varphi(n)}$ der Folge $(x_n)_{n\in\mathbb{N}}$, welches in $[a_n, b_n]$ liegt. O.B.d.A. sei $(\varphi(n))_{n\in\mathbb{N}}$ streng monoton wachsend (denn jedes $[a_n, b_n]$ enthält unendlich viele Elemente der Folge $(x_n)_{n\in\mathbb{N}}$). Wegen

$$a_n \leq x_{\varphi(n)} \leq b_n$$

konvergiert dann die Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ gegen $l \in [a, b]$. Wir haben also bewiesen, dass jede Folge von $K = [a, b]$ eine konvergente Teilfolge besitzt, d.h., K ist kompakt.
Für ein beliebiges $n \in \mathbb{N}$, $n \geq 1$, folgt das Ergebnis aus folgendem Argument. Eine Teilmenge K des $\mathbb{R}^n$ ist nämlich genau dann beschränkt (bzgl. der Standardmetrik), wenn sie in einem Produkt von abgeschlossenen beschränkten Intervallen enthalten ist,

[7] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[8] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

d.h., wenn $K \subset \prod_{i=1}^{n}[a_i, b_i]$ für reelle Zahlen $a_i \leq b_i$, $i = 1, \ldots, n$. Diese Aussage ist eine direkte Folgerung der Ungleichungen

$$d(x, y) \leq |x - y| \leq \sqrt{n} \cdot d(x, y),$$

die für alle $x, y \in \mathbb{R}^n$ gelten, wobei $d : (x, y) \mapsto \max\limits_{1 \leq i \leq n} (|x_i - y_i|)$ die (induktiv definierte) Produktmetrik des $\mathbb{R}^n$ ist (Übungsaufgabe). Da aber jedes Intervall der Form $[a_i, b_i]$ kompakt ist (siehe den Fall $n = 1$), ist das Produkt $\prod_{i=1}^{n}[a_i, b_i]$ nach Proposition 2.27 wieder kompakt in $(\mathbb{R}^n, d)$. Wieder nach den beiden letzten Ungleichungen muss dann $\prod_{i=1}^{n}[a_i, b_i]$ kompakt bzgl. der Standardmetrik sein. Da K eine abgeschlossene Teilmenge des $\mathbb{R}^n$ - und somit auch von $\prod_{i=1}^{n}[a_i, b_i]$ - ist, liefert Proposition 2.25, dass K kompakt ist. $\square$

Beispiele 2.29

1. Jeder abgeschlossene Ball im $\mathbb{R}^n$ ist kompakt, denn per Definition und wegen der Dreiecksungleichung gilt $|x - y| \leq 2r$ für alle $x, y \in \overline{B_r}(p)$, wobei $p \in \mathbb{R}^n$ beliebig gewählt werden kann.

2. Jede r-Sphäre $S_r(p) := \{x \in \mathbb{R}^n \,,\, |x - p| = r\}$ im $\mathbb{R}^n$ ist kompakt.

Proposition 2.30 (Stetigkeit und Kompaktheit 1) *Sei $f : X \longrightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Ist $K \subset X$ kompakt, so ist $f(K) \subset Y$ kompakt.*

Beweis:[9] Sei $(y_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus $f(K)$. Nach Definition von $f(K)$ existiert, für jedes $n \in \mathbb{N}$ mindestens ein $x_n \in K$ mit $f(x_n) = y_n$. Man bekommt eine Folge $(x_n)_{n\in\mathbb{N}}$ aus K. Wegen K kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ und ein $\overline{x} \in K$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$. Wegen f stetig gilt $f(x_{\varphi(n)}) \underset{n\to\infty}{\longrightarrow} f(\overline{x})$. Wegen $f(\overline{x}) \in f(K)$ konvergiert die Teilfolge $(f(x_{\varphi(n)}))_{n\in\mathbb{N}}$ von $(f(x_n))_{n\in\mathbb{N}}$ gegen ein Element aus $f(K)$. Dies zeigt, dass $f(K)$ kompakt ist. $\square$

Proposition 2.31 (Stetigkeit und Kompaktheit 2) *Sei $f : X \longrightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Ist (X, d_X) kompakt und f bijektiv, so ist $f^{-1} : Y \longrightarrow X$ auch stetig.*

Beweis:[10] Nach Proposition 2.19 reicht es, zu beweisen, dass für jede abgeschlossene Teilmenge $A \subset X$ das Urbild $(f^{-1})^{-1}(A)$ von A unter f^{-1} wieder abgeschlossen ist. Ist $A \subset X$ abgeschlossen, so ist nach Proposition 2.25 A kompakt. Da f stetig ist, folgt aus Proposition 2.30, dass $f(A)$ kompakt ist. Insbesondere ist $(f^{-1})^{-1}(A) = f(A)$ kompakt und somit auch abgeschlossen in Y (Proposition 2.25). Dies beweist die Stetigkeit von f^{-1}. $\square$

Eine stetige bijektive Abbildung $f : X \longrightarrow Y$, deren Umkehrabbildung $f^{-1} : Y \longrightarrow X$ ebenfalls stetig ist, heißt per Definition *Homöomorphismus*.

[9] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[10] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Korollar 2.32 (Maximum und Minimum einer Funktion auf einem kompakten Raum) *Sei $f : X \longrightarrow \mathbb{R}$ eine stetige Abbildung, wobei $\mathbb{R}$ die Standardmetrik trägt und (X, d) ein beliebiger metrischer Raum ist. Ist (X, d) kompakt, so besitzt f ein Minimum und ein Maximum auf X, d.h., es gibt (mindestens) zwei Elemente $x_{\min}$, $x_{\max} \in X$ mit $f(x_{\min}) = \min_{x \in X}(f(x))$ und $f(x_{\max}) = \max_{x \in X}(f(x))$.*

Beweis:[11] Aus Proposition 2.30 folgt, dass $f(X)$ kompakt ist. Satz 2.28 impliziert dann, dass $f(X)$ abgeschlossen und beschränkt in $\mathbb{R}$ ist. Nach Analysis I besitzt jede beschränkte Teilmenge von $\mathbb{R}$ ein Infimum und ein Supremum; diese sind sogar in der Teilmenge enthalten, falls diese Teilmenge abgeschlossen ist (Übungsaufgabe). Somit existieren $x_{\min}, x_{\max} \in X$ mit $f(x_{\min}) = \mathrm{Inf}(f(x)) = \inf\{f(x) \,|\, x \in X\}$ und $f(x_{\max}) = \sup\{f(x) \,|\, x \in X\}$. $\square$

Satz 2.33 (Gleichmäßigstetigkeit auf kompakten Räumen) *Sei $f : X \longrightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Ist (X, d_X) kompakt, so ist f* gleichmäßig *stetig, d.h., $\forall\, r > 0$, $\exists\, \delta > 0$, s.d. $d_Y(f(x), f(y)) < r$ für alle $x, y \in X$ mit $d_X(x, y) < \delta$.*

Beweis: Angenommen, f wäre nicht gleichmäßig stetig. Dann existiert ein $r > 0$ so, dass für alle $\delta > 0$ (mindestens) ein Paar $(x, y) \in X \times X$ existiert mit $d_X(x, y) < \delta$ und $d_Y(f(x), f(y)) \geq r$. Wähle dann $\delta_n = \frac{1}{n}$, $n \in \mathbb{N}$, $n \geq 1$. Dann existiert $(x_n, y_n) \in X \times X$ mit $d_X(x_n, y_n) < \frac{1}{n}$ und $d_Y(f(x_n), f(y_n)) \geq r$. Wegen (X, d_X) kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ und ein $\overline{x} \in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$. Bemerke, dass wegen $d_X(x_{\varphi(n)}, y_{\varphi(n)}) \underset{n\to\infty}{\longrightarrow} 0$ auch $d_X(y_{\varphi(n)}, \overline{x}) \underset{n\to\infty}{\longrightarrow} 0$ (Dreiecksungleichung). Nun liefert die Stetigkeit von f die Konvergenz $f(x_{\varphi(n)}) \underset{n\to\infty}{\longrightarrow} f(\overline{x})$ sowie $f(y_{\varphi(n)}) \underset{n\to\infty}{\longrightarrow} f(\ overlinex)$, insbesondere (wegen der Stetigkeit von d_Y) $d_Y(f(x_{\varphi(n)}, f(y_{\varphi(n)})) \underset{n\to\infty}{\longrightarrow} d_Y(f(\overline{x}), f(\overline{x})) = 0$. Dies widerspricht aber $d_Y(f(x_{\varphi(n)}), f(y_{\varphi(n)})) \geq r > 0$ für alle $n \in \mathbb{N}$. Daraus folgt, dass f gleichmäßig stetig sein muss. $\square$

2.1.5 Vollständige metrische Räume

Definition 2.34 *Sei (X, d) ein metrischer Raum. Eine* Cauchy-Folge *in X ist eine Folge $(x_n)_{n\in\mathbb{N}}$ aus X, die folgendes erfüllt: $\forall\, \varepsilon > 0$, $\exists\, N \in \mathbb{N}$ s.d. $\forall\, p, q \geq N$, $d(x_p, x_q) \leq \varepsilon$.*

Beispiele 2.35

1. Jede konvergente Folge ist eine Cauchy-Folge. Denn: konvergiert eine Folge $(x_n)_{n\in\mathbb{N}}$ gegen ein $\overline{x} \in X$, so gilt: $\forall\, \varepsilon > 0$, $\exists\, N \in \mathbb{N}$ mit $d(x_n, \overline{x}) \leq \frac{\varepsilon}{2}$ für alle $n \geq N$. Insbesondere gilt: $\forall\, p, q \geq N$, $d(x_p, x_q) \leq d(x_p, \overline{x}) + d(\overline{x}, x_q) \leq \varepsilon$. Somit ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge.

2. Nicht jede Cauchy-Folge konvergiert: z.B. ist die induktiv definierte Folge $x_{n+1} := \frac{1}{2}(x_n + \frac{2}{x_n})$ aus $\mathbb{Q}$ mit $x_0 = 2$ zwar eine Cauchy-Folge, konvergiert aber nicht in $\mathbb{Q}$. Es gilt tatsächlich $x_n \underset{n\to\infty}{\longrightarrow} \sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$.

Definition 2.36 *Ein* vollständiger *metrischer Raum ist ein metrischer Raum, in dem alle Cauchy-Folgen konvergieren.*

[11] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Beispiele 2.37

1. Jeder kompakte metrische Raum (X,d) ist vollständig. Sei nämlich $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge aus X. Wegen (X,d) kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ und ein $\overline{x}\in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$. Sei nun $\varepsilon>0$. Wegen $(x_n)_{n\in\mathbb{N}}$ Cauchy existiert ein $N\in\mathbb{N}$ mit $d(x_p,x_q)\le\frac{\varepsilon}{2}$ für alle $p,q\ge N$; wegen $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} \overline{x}$ existiert ein $N'\in\mathbb{N}$ mit $d(x_{\varphi(n)},\overline{x})\le\frac{\varepsilon}{2}$ für alle $n\ge N'$. Nun gilt, $\forall\ n\ge\max(N,N')$: $d(x_n,\overline{x})\le d(x_n,x_{\varphi(n)})+d(x_{\varphi(n)},\overline{x})\le\varepsilon$ (beachte, dass $\varphi(n)\ge n$ immer gilt). Somit gilt $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$.

2. $\mathbb{R}^n$ mit der Standardmetrik ist vollständig. Denn: sei $(x_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge aus $\mathbb{R}^n$, dann ist die Folge $(x_k)_{k\in\mathbb{N}}$ insbesondere beschränkt, d.h. $\exists\, M\in[0,\infty[$ mit $d(x_k,x_l)\le M$ für alle $k,l\in\mathbb{N}$. Begründung: fixiere $\varepsilon>0$ (z.B. $\varepsilon=1$), dann existiert ein $N\in\mathbb{N}$ mit $d(x_p,x_q)\le\varepsilon$ für alle $p,q\ge N$. Für alle $k,l\in\mathbb{N}$ gilt dann
$$\begin{aligned} d(x_k,x_l) &\le d(x_k,x_N)+d(x_N,x_l)\\ &\le 2\max\left(\max\{d(x_j,x_N)\,|\,0\le j\le N\},\varepsilon\right)<\infty. \end{aligned}$$
Nach dem Satz von Heine-Borel (Satz 2.28) ist jede beschränkte Folge aus $\mathbb{R}^n$ in einer kompakten Teilmenge enthalten. Wie im obigen Beispiel muss dann $(x_k)_{k\in\mathbb{N}}$ – als Cauchy-Folge aus dieser kompakten Teilmenge – konvergieren.

3. $\mathbb{Q}$ (mit der von $\mathbb{R}$ induzierten Metrik) ist nach Beispiel 2.35.2 nicht vollständig.

4. Das Intervall $]0,\infty[$ ist nicht vollständig; die Folge $(x_n:=\frac{1}{n})_{n\ge1}$ ist zwar eine Cauchy-Folge aus $]0,\infty[$, konvergiert aber nicht in $]0,\infty[$ (denn: $x_n \underset{n\to\infty}{\longrightarrow} 0\notin]0,\infty[$).

Proposition 2.38 (vollständige Teilmengen) *Sei (X,d) ein metrischer Raum und $Y\subset X$ eine beliebige Teilmenge.*

i) Ist $(Y,d_{|Y\times Y})$ vollständig, so ist Y abgeschlossen.

ii) Ist (X,d) vollständig und $Y\subset X$ abgeschlossen, so ist $(Y,d_{|Y\times Y})$ vollständig.

Beweis:[12]
i) Sei $y\in X$ Grenzwert einer Folge $(y_n)_{n\in\mathbb{N}}$ aus Y. Da $(y_n)_{n\in\mathbb{N}}$ notwendigerweise eine Cauchy-Folge (aus X und aus Y) und $(Y,d_{|Y\times Y})$ vollständig ist, konvergiert $(y_n)_{n\in\mathbb{N}}$ in $(Y,d_{|Y\times Y})$, d.h., $\exists\,\overline{y}\in Y$ mit $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$. Da aber $(y_n)_{n\in\mathbb{N}}$ bereits gegen y konvergiert, muss dann $\overline{y}=y$ gelten, insbesondere $y\in Y$. Dies beweist, dass Y abgeschlossen in X ist.
ii) Sei $(y_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge aus $(Y,d_{|Y\times Y})$. Da $(y_n)_{n\in\mathbb{N}}$ insbesondere eine Cauchy-Folge aus (X,d) ist, existiert ein $\overline{x}\in X$ mit $y_n \underset{n\to\infty}{\longrightarrow} \overline{x}$. Wegen Y abgeschlossen gilt $\overline{x}\in Y$. Somit konvergiert $(y_n)_{n\in\mathbb{N}}$ in Y (gegen $\overline{x}$). Dies zeigt, dass $(Y,d_{|Y\times Y})$ vollständig ist. □

Proposition 2.39 (Vereinigung und Durchschnitt vollständiger Teilmengen) *Sei (X,d) ein metrischer Raum.*

i) Ist $(Y_i)_{i=1,\dots,p}$ eine Familie vollständiger Teilmengen von X, so ist $\bigcup\limits_{i=1}^{p} Y_i$ wieder vollständig (für die induzierte Metrik).

[12]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

ii) Ist $(Y_i)_{i\in I}$ *eine beliebige Familie vollständiger Teilmengen von* X*, so ist* $\bigcap_{i\in I} Y_i$ *wieder vollständig (für die induzierte Metrik).*

iii) Ist $(Y_n)_{n\in\mathbb{N}}$ *eine Folge nichtleerer vollständiger Teilmengen von* X *mit* $Y_{n+1} \subset Y_n$ *und* $\operatorname{diam}(Y_n) \underset{n\to\infty}{\longrightarrow} 0$*, so gilt* $\bigcap_{n\in\mathbb{N}} Y_n \neq \varnothing$ *(und ist vollständig).*

Bezeichnung: $\operatorname{diam}(X, d) := \sup_{x,y\in X} \{d(x, y)\} \in [0, \infty]$ ist der *Durchmesser* von (X, d).

Beweis:[13]
$i)$ Es reicht, den Fall $p = 2$ zu zeigen (führe dann Induktion über p durch). Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge aus $Y_1 \cup Y_2$, so existiert ein $i \in \{1, 2\}$ und eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $x_{\varphi(n)} \in Y_i$ für alle $n \in \mathbb{N}$. Da Y_i vollständig ist, konvergiert $(x_{\varphi(n)})_{n\in\mathbb{N}}$ in Y_i – somit auch in $Y_1 \cup Y_2$ – gegen ein $y \in Y_i$. Da aber $(x_n)_{n\in\mathbb{N}}$ Cauchy ist, muss sie dann gegen y konvergieren (siehe Beispiel 2.37.1). Somit ist $Y_1 \cup Y_2$ vollständig.
$ii)$ Fixiere $i_0 \in I$ (oBdA sei $I \neq \varnothing$). Dann ist $\bigcap_{i\in I} Y_i \subset Y_{i_0}$ und jedes Y_i ist abgeschlossen in X nach Proposition 2.38. Insbesondere ist $\bigcap_{i\in I} Y_i$ abgeschlossen in X, somit auch in Y_{i_0}. Wiederum nach Proposition 2.38 muss dann $\bigcap_{i\in I} Y_i$ vollständig sein.
$iii)$ Nach $ii)$ ist $\bigcap_{n\in\mathbb{N}} Y_n$ vollständig. Nach Voraussetzung existiert ein $y_n \in Y_n$ für jedes $n \in \mathbb{N}$.
Behauptung: $(y_n)_{n\in\mathbb{N}}$ *ist eine Cauchy-Folge.*
Beweis der Behauptung: Sei $\varepsilon > 0$. Wegen $\operatorname{diam}(Y_n) \underset{n\to\infty}{\longrightarrow} 0$ existiert ein $N \in \mathbb{N}$ s.d. $\operatorname{diam}(Y_n) < \varepsilon$ für alle $n \geq N$, insbesondere $d(y, z) < \varepsilon$ für alle $y, z \in Y_n$ und alle $n \geq N$. Nun gilt für alle $p, q \geq N$: $y_p, y_q \in Y_N$ (wegen $Y_p \subset Y_N$ und $Y_q \subset Y_N$) und somit $d(y_p, y_q) < \varepsilon$. Dies zeigt, dass $(y_n)_{n\in\mathbb{N}}$ Cauchy ist. ✓
Wegen Y_0 vollständig konvergiert $(y_n)_{n\in\mathbb{N}}$ gegen ein $\overline{y} \in Y_0$. Für jedes $N \in \mathbb{N}$ ist $(y_n)_{n\geq N}$ ebenfalls eine Cauchy-Folge in Y_N, somit konvergiert sie in Y_N; nach der Eindeutigkeit des Grenzwerts gilt $\overline{y} \in Y_N$. Dies gilt für alle $N \in \mathbb{N}$, woraus folgt $\overline{y} \in \bigcap_{n\in\mathbb{N}} Y_n$. Dies zeigt insbesondere $\bigcap_{n\in\mathbb{N}} Y_n \neq \varnothing$. □

Man beachte, dass der Durchschnitt $\bigcap_{n\in\mathbb{N}} Y_n$ aus Proposition 2.39.$iii)$ tatsächlich aus genau einem Element besteht (benutze $\operatorname{diam}(\bigcap_{n\in\mathbb{N}} Y_n) \leq \operatorname{diam}(Y_n) \underset{n\to\infty}{\longrightarrow} 0$).

Proposition 2.40 (Produkte vollständiger Räume) *Sind* (X, d_X) *und* (Y, d_Y) *zwei vollständige metrische Räume, so ist* $(X \times Y, d_{X\times Y})$ *vollständig.*

Beweis:[14] Ist $((x_n, y_n))_{n\in\mathbb{N}}$ eine Cauchy-Folge aus $X \times Y$, so sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ ebenfalls Cauchy-Folgen in (X, d_X) bzw. (Y, d_Y). Da (X, d_X) und (Y, d_Y) vollständig sind, existieren $\overline{x} \in X$ und $\overline{y} \in Y$ mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$ und $y_n \underset{n\to\infty}{\longrightarrow} \overline{y}$, d.h., $(x_n, y_n) \underset{n\to\infty}{\longrightarrow} (\overline{x}, \overline{y})$. Dies zeigt, dass $(X \times Y, d_{X\times Y})$ vollständig ist. □

Nun wollen wir den Zusammenhang zwischen Stetigkeit und Vollständigkeit untersuchen. Man bemerke zuerst, dass das Bild einer vollständigen Teilmenge unter einer stetigen

[13]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[14]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Abbildung nicht unbedingt vollständig ist: z.B. $f : \mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto e^x$ ist stetig und $(\mathbb{R}, |\cdot|)$ ist vollständig, jedoch ist $f(\mathbb{R}) =]0, \infty[$ nicht vollständig. Einer der wichtigsten Sätze dieses Abschnitts ist folgender:

Satz 2.41 (Banach'scher Fixpunktsatz) *Sei (X, d) ein* vollständiger *metrischer Raum und $f : X \longrightarrow X$ eine* Kontraktion, *d.h. $\exists\ k \in [0, 1[$ s.d. $d(f(x), f(y)) \leq k \cdot d(x, y)$ $\forall\, x, y \in X$. Dann hat f einen eindeutigen Fixpunkt: $\exists\ !\ x \in X$ mit $f(x) = x$. Außerdem konvergiert jede Folge der Form $x_{n+1} = f(x_n)$, $x_0 \in X$ beliebig, gegen diesen Fixpunkt x. Genauer gilt $d(x_n, x) \leq \frac{k^n}{1-k} \cdot d(x_1, x_0)$ für alle $n \in \mathbb{N}$.*

Beweis: Wir beweisen separat die Eindeutigkeit und die Existenz des Fixpunktes.
Zur Eindeutigkeit: sind x und y Fixpunkte von f, so gilt $d(x, y) = d(f(x), f(y)) \leq k \cdot d(x, y)$ nach Voraussetzung, somit $d(x, y) = 0$, d.h., $x = y$.
Zur Existenz: Sei $x_0 \in X$ beliebig und $(x_n)_{n\in\mathbb{N}}$ die durch $x_{n+1} := f(x_n)$ induktiv definierte Folge. Wir zeigen, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in (X, d) ist. Seien $n, p \in \mathbb{N}$ beliebig, dann gilt

$$d(x_{n+p}, x_n) \leq d(x_{n+p}, x_{n+p-1}) + \cdots + d(x_{n+1}, x_n) \quad \text{(Dreiecksungleichung)}$$

mit

$$\begin{aligned} d(x_{n+p}, x_{n+p-1}) &= d(f(x_{n+p-1}), f(x_{n+p-2})) \\ &\leq k \cdot d(x_{n+p-1}, x_{n+p-2}) \\ &\vdots \\ &\leq k^{p-1} \cdot d(x_{n+1}, x_n), \end{aligned}$$

somit

$$\begin{aligned} d(x_{n+p}, x_n) &\leq (k^{p-1} + k^{p-2} + \cdots + 1) d(x_{n+1}, x_n) \\ &\leq \left(\tfrac{1-k^p}{1-k}\right) \cdot k^n d(x_1, x_0) \\ &\leq \tfrac{1}{1-k} \cdot k^n \cdot d(x_1, x_0) \quad \text{wegen} \quad 0 \leq k^p < 1. \end{aligned}$$

Wegen $k \in [0, 1[$ gilt $k^n \underset{n\to\infty}{\longrightarrow} 0$, somit auch $\frac{k^n}{1-k} \cdot d(x_1, x_0) \underset{n\to\infty}{\longrightarrow} 0$. Daraus folgt, dass es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt mit $\frac{k^n}{1-k} \cdot d(x_1, x_0) < \varepsilon$ für alle $n \geq N$, insbesondere $d(x_{n+p}, x_n) < \varepsilon$ für alle $n \geq N$ und $p \in \mathbb{N}$, d.h., $d(x_p, x_q) < \varepsilon$ für alle $p, q \geq N$. Dies zeigt, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in (X, d) ist, insbesondere gegen ein $x \in X$ konvergiert. Da f aber stetig ist (gilt $u_n \underset{n\to\infty}{\longrightarrow} \overline{u}$ in X, so gilt $d(f(u_n), f(\overline{u})) \leq k \cdot d(u_n, \overline{u}) \underset{n\to\infty}{\longrightarrow} 0$, somit $f(u_n) \underset{n\to\infty}{\longrightarrow} f(\overline{u})$, für alle $\overline{u} \in X$), gilt $f(x_n) \underset{n\to\infty}{\longrightarrow} f(x)$. Wegen $x_{n+1} = f(x_n) \underset{n\to\infty}{\longrightarrow} x$, gilt $f(x) = x$, d.h., x ist Fixpunkt von f. Schließlich gilt $d(x_{n+p}, x_n) \underset{p\to\infty}{\longrightarrow} d(x, x_n)$ und wegen $d(x_{n+p}, x_n) \leq \frac{k^n}{1-k} \cdot d(x_1, x_0)$ (rechte Seite hängt nicht von p ab) gilt auch $d(x, x_n) \leq \frac{k^n}{1-k} \cdot d(x_1, x_0)$, für alle $n \in \mathbb{N}$. $\square$

Beispiel 2.42 Sei $X := [\sqrt{2}, \infty[\subset \mathbb{R}$ und $f : X \longrightarrow \mathbb{R}$, $x \longmapsto \frac{1}{2}(x + \frac{2}{x})$. Beachte, dass $f(X) = [\sqrt{2}, \infty[$ gilt (die Funktion f ist ableitbar mit $f'(x) = \frac{1}{2}(1 - \frac{2}{x^2}) \geq 0$ $\forall\ x \in [\sqrt{2}, \infty[$, $f(\sqrt{2}) = \sqrt{2}$ und $f(x) \underset{n\to\infty}{\longrightarrow} \infty$) und dass X (mit der induzierten Metrik) vollständig ist: X ist abgeschlossen in $\mathbb{R}$ und $\mathbb{R}$ ist vollständig, siehe Proposition 2.38. Außerdem gilt $f'(x) \in [0, \frac{1}{2}[$ für alle $x \in X$, insbesondere folgt aus dem *Mittelwertsatz*

$|f(x) - f(y)| \leq \sup_{t \in [x,y]} |f'(t)| \cdot |x - y| \leq \frac{1}{2} \cdot |x - y|$ für alle $x, y \in X$ mit $x < y$. Daraus folgt, dass $f : X \longrightarrow X$ eine Kontraktion ist. Aus dem Banach'schen Fixpunktsatz folgt dann, dass f einen eindeutigen Fixpunkt in X besitzt. Dieser Fixpunkt x erfüllt $f(x) = x$, d.h., $x = \sqrt{2}$. Der Banach'sche Fixpunktsatz liefert auch $x_n \underset{n\to\infty}{\longrightarrow} x$ für jede Folge $(x_n)_{n \in y}$, der Form $x_{n+1} = f(x_n)$ mit $x_0 \in X$ beliebig, siehe Beispiel 2.35.2.

Bemerkungen 2.43

1. Im Allgemeinen ist der Banach'sche Fixpunktsatz nicht so einfach anzuwenden: gegeben eine (stetige) Abbildung $f : Y \longrightarrow Y$, wobei Y ein metrischer Raum ist, sucht man eine *vollständige* Teilmenge $X \subset Y$ für die $f(X) \subset X$ gilt *und* s.d. $f_{|X}$ eine Kontraktion ist. Beides gleichzeitig zu erreichen ist aber i.A. schwierig! Siehe z.B. Beispiel 2.42 oben.

2. Keine Voraussetzung kann im Satz 2.41 weggelassen werden. Setzt man z.B. lediglich voraus, dass $d(f(x), f(y)) \leq d(x, y)$ für alle $x, y \in X$, so wird das Ergebnis falsch: wähle beispielsweise $X := \mathbb{R}$ (mit Standardmetrik) und $f(x) := x + 1$ für alle $x \in \mathbb{R}$. Dann erfüllt f die Bedingung $|f(x) - f(y)| = |x - y| \; \forall \; x, y \in \mathbb{R}$, die Abbildung f hat aber keinen Fixpunkt ($\nexists x \in \mathbb{R}$ mit $x = x + 1$).

Satz 2.44 (Fortsetzungssatz von Tietze) *Seien (X, d_X) und (Y, d_Y) metrische Räume und $D \subset X$ eine* dichte *Teilmenge von X (d.h., $\overline{D} = X$). Angenommen, (Y, d_Y) sei* vollständig. *Dann lässt sich jede* gleichmäßig stetige *Abbildung $f : D \longrightarrow Y$ eindeutig auf X fortsetzen, d.h., $\exists \,! \, \overline{f} : X \longrightarrow Y$ stetig mit $\overline{f}_{|D} = f$.*

Beweis:[15] Sei $\overline{x} \in X$. Wegen $\overline{D} = X$ existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ aus D mit $x_n \underset{n\to\infty}{\longrightarrow} x$. Wir wollen $\overline{f}(\overline{x}) := \lim_{n\to\infty} f(x_n)$ definieren. Dazu zeigen wir:

1. Die Folge $(f(x_n))_{n \in \mathbb{N}}$ konvergiert in Y (d.h., ist Cauchy in Y).

2. Der Grenzwert $\lim_{n\to\infty} f(x_n)$ hängt nicht von der Folge $(x_n)_{n \in \mathbb{N}}$ ab mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$. Insbesondere ist $\overline{f}(\overline{x})$ wohldefiniert.

3. $\overline{f}_{|D} = f$, d.h., $\overline{f}$ ist eine Fortsetzung von f auf X.

4. $\overline{f}$ ist stetig (tatsächlich gleichmäßig stetig) auf X.

Im Anschluss zeigen wir dann, dass jede weitere stetige Fortsetzung von f auf X mit $\overline{f}$ übereinstimmen muss.
Ad 1: Wegen (Y, d_Y) vollständig reicht es, zu zeigen, dass $(f(x_n))_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Sei $\varepsilon > 0$. Wegen f gleichmäßig stetig existiert ein $\delta > 0$ so, dass $d_Y(f(x), f(x')) \leq \varepsilon$ gilt für alle $x, x' \in D$ mit $d_X(x, x') \leq \delta$. Da $(x_n)_{n \in \mathbb{N}}$ gegen $\overline{x}$ konvergiert, ist sie Cauchy in (X, d_X), insbesondere existiert ein $N \in \mathbb{N}$ mit $d_X(x_p, x_q) \leq \delta$ für alle $p, q \geq N$. Es folgt $d_Y(f(x_p), f(x_q)) \leq \varepsilon$ für alle $p, q \geq N$. Dies zeigt, dass $(f(x_n))_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Insbesondere ist $\lim_{n\to\infty} f(x_n) \in Y$ wohldefiniert. $\checkmark$
Ad 2: Ist $(x'_n)_{n \in \mathbb{N}}$ eine weitere Folge aus D mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$, so definiere die Folge $(y_n)_{n \in \mathbb{N}}$ aus D durch $y_n := x_{\frac{n}{2}}$ falls n gerade und $y_n := x'_{\frac{n-1}{2}}$ falls n ungerade. Dann gilt $y_n \underset{n\to\infty}{\longrightarrow} \overline{x}$,

[15] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

insbesondere ist $(f(y_n))_{n\in\mathbb{N}}$ eine Cauchy-Folge in (Y, d_Y) nach 1, somit existiert ein $\overline{y} \in Y$ mit $f(y_n) \underset{n\to\infty}{\longrightarrow} \overline{y}$. Daher gelten auch $f(x_n) = f(y_{2n}) \underset{n\to\infty}{\longrightarrow} \overline{y}$ und $f(x'_n) = f(y_{2n+1}) \underset{n\to\infty}{\longrightarrow} \overline{y}$, insbesondere $\lim\limits_{n\to\infty} f(x_n) = \overline{x} = \overline{y} = \lim\limits_{n\to\infty} f(x'_n)$. Dies zeigt, dass $\lim\limits_{n\to\infty} f(x_n)$ unabhängig von der Folge $(x_n)_{n\in\mathbb{N}}$ aus D mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$ ist. $\surd$

Ad 3: Ist $\overline{x} \in D$, so kann $(x_n := \overline{x})_{n\in\mathbb{N}}$ als Folge gewählt werden mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$. In dem Fall gilt $\overline{f}(\overline{x}) := \lim\limits_{n\to\infty} f(x_n) = \lim\limits_{n\to\infty} f(\overline{x}) = f(\overline{x})$. Daraus folgt $\overline{f}_{|D} = f$. $\surd$

Ad 4: Wir zeigen, dass $\overline{f}$ gleichmäßig stetig ist. Sei $\varepsilon > 0$. Dann existiert ein $\delta > 0$ mit $d_Y(f(x), f(x')) \leq \frac{\varepsilon}{3}$ für alle $x, x' \in D$ mit $d_X(x, x') \leq \delta$. Seien nun $\overline{x}, \overline{x}' \in X$ gegeben mit $d_X(\overline{x}, \overline{x}') \leq \frac{\delta}{3}$. Dann existieren Folgen $(x_n)_{n\in\mathbb{N}}$ und $(x'_n)_{n\in\mathbb{N}}$ aus D mit $x_n \underset{n\to\infty}{\longrightarrow} \overline{x}$ sowie $x'_n \underset{n\to\infty}{\longrightarrow} \overline{x}'$. Nach Konstrution von $\overline{f}$ gelten $f(x_n) \underset{n\to\infty}{\longrightarrow} \overline{f}(\overline{x})$ und $f(x'_n) \underset{n\to\infty}{\longrightarrow} \overline{f}(\overline{x}')$. Daher existiert ein $N \in \mathbb{N}$ mit $d_X(x_n, \overline{x}) \leq \frac{\delta}{3}$, $d_X(x'_n, \overline{x}') \leq \frac{\delta}{3}$, $d_Y(\overline{f}(\overline{x}), f(x_n)) \leq \frac{\varepsilon}{3}$ und $d_Y(\overline{f}(\ overlinex'), f(x'_n)) \leq \frac{\varepsilon}{3}$ für alle $n \geq N$. Für ein beliebiges $n \geq N$ gilt dann:

$$\begin{array}{rcl} d_Y(\overline{f}(\overline{x}), \overline{f}(\overline{x}')) & \leq & d_Y(\overline{f}(\overline{x}), f(x_n)) + d_Y(f(x_n), f(x'_n)) + d_Y(f(x'_n), \overline{f}(\overline{x}')) \\ & \leq & \frac{\varepsilon}{3} + d_Y(f(x_n), f(x'_n)) + \frac{\varepsilon}{3}, \\ \text{mit} \quad d_X(x_n, x'_n) & \leq & d_X(x_n, \overline{x}) + d_X(\overline{x}, \overline{x}') + d_X(\overline{x}', x'_n) \\ & \leq & \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta, \end{array}$$

insbesondere $d_Y(f(x_n), f(x'_n)) \leq \frac{\varepsilon}{3}$ und somit $d_Y(\overline{f}(\overline{x}), \overline{f}(\overline{x}')) \leq \varepsilon$. Dies zeigt, dass $\overline{f}$ gleichmäßig stetig ist. $\surd$

Die Eindeutigkeit der Fortsetzung beruht auf folgender

Behauptung: Seien $f : X \longrightarrow Y$ und $g : X \longrightarrow Y$ stetige Abbildungen. Dann ist $\{x \in X \mid f(x) = g(x)\}$ abgeschlossen in (X, d_X). Daraus folgt: $D \subset \{x \in X \mid f(x) = g(x)\}$ für eine Teilmenge D von $X \Longrightarrow \overline{D} \subset \{x \in X \mid f(x) = g(x)\}$.

Beweis der Behauptung: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge aus X mit $f(x_n) = g(x_n)$ für alle $n \in \mathbb{N}$. Angenommen, $x_n \underset{n\to\infty}{\longrightarrow} \overline{x} \in X$. Dann gilt wegen f und g stetig:

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

somit ist $\overline{x} \in \{x \in X \mid f(x) = g(x)\}$. Dies zeigt, dass $\{x \in X \mid f(x) = g(x)\}$ abgeschlossen in (X, d_X) ist. $\surd$

Ist nun $\overline{f}' : X \longrightarrow Y$ stetig mit $\overline{f}'_{|D} = f$, so gilt nach der letzten Behauptung $X = \overline{D} \subset \{x \in X \mid \overline{f}(x) = \overline{f}'(x)\}$, d.h., $\overline{f} = \overline{f}'$. $\square$

2.1.6 Zusammenhängende metrische Räume

Definition 2.45 *Ein metrischer Raum (X, d) heißt genau dann* zusammenhängend, *wenn die einzigen Teilmengen von X, die gleichzeitig offen und abgeschlossen in (X, d) sind, $\varnothing$ und X sind.*

Eine Teilmenge Y eines metrischen Raumes (X, d) heißt genau dann zusammenhängend, wenn $(Y, d_{|Y\times Y})$ zusammenhängend ist.

Proposition 2.46 *Sei (X, d) ein metrischer Raum. Dann sind folgende Aussagen äquivalent:*

i) *(X, d) ist zusammenhängend.*

ii) *X lässt sich nicht schreiben als disjunkte Vereinigung nichtleerer offener Teilmengen von X.*

iii) *X lässt sich nicht schreiben als disjunkte Vereinigung nichtleerer abgeschlossener Teilmengen von X.*

iv) *Jede stetige Abbildung $f : X \longrightarrow \{0, 1\}$ ist konstant, wobei $\{0, 1\}$ die grobe Metrik trägt (die sowieso die einzige Metrik – bis auf positives Vielfaches – auf $\{0, 1\}$ ist).*

Beweis:[16]
i) $\Rightarrow$ *ii)*: Sind U_1, U_2 offene Teilmengen von X mit $X = U_1 \cup U_2$ und $U_1 \cap U_2 = \varnothing$, so sind U_1 und U_2 offen und abgeschlossen (wegen $U_1^c = U_2$) in X. Wegen (X, d) zusammenhängend muss entweder ($U_1 = \varnothing$ und $U_2 = X$) oder ($U_1 = X$ und $U_2 = \varnothing$) gelten. Daraus folgt, dass U_1 und U_2 nicht gleichzeitig nichtleer sein können.
ii) $\Rightarrow$ *iii)*: Sind A_1, A_2 abgeschlossene Teilmengen von X mit $X = A_1 \cup A_2$ und $A_1 \cap A_2 = \varnothing$, so gilt wegen $A_1^c = A_2$ auch $X = A_2^c \cup A_1^c$. Da A_1^c und A_2^c abgeschlossen sind mit $A_1^c \cap A_2^c = (A_1 \cup A_2)^c = \varnothing$, muss entweder ($A_1^c = \varnothing$ (d.h., $A_1 = X$) und $A_2^c = X$ (d.h., $A_2 = \varnothing$)) oder ($A_1^c = X, A_2^c = \varnothing$) – d.h., ($A_1 = \varnothing, A_2 = X$) – gelten. Daraus folgt, dass A_1 und A_2 nicht gleichzeitig nichtleer sein können.
iii) $\Rightarrow$ *iv)*: Sei $f : X \longrightarrow \{0, 1\}$ stetig. Dann gilt insbesondere $X = f^{-1}(\{0, 1\}) = f^{-1}(\{0\}) \cup f^{-1}(\{1\})$ mit $f^{-1}(\{0\}) \cap f^{-1}(\{1\}) = \varnothing$. Wegen $\{i\} \subset \{0, 1\}$ abgeschlossen (z.B. ist $\{i\} = \overline{B_{\frac{1}{2}}}(i)$ in $(\{0, 1\}, d^G)$) und f stetig sind $f^{-1}(\{0\})$ und $f^{-1}(\{1\})$ abgeschlossen in X. Es folgt entweder $f^{-1}(\{0\}) = \varnothing$, d.h., f ist konstant gleich 1 auf X, oder $f^{-1}(\{1\}) = \varnothing$, d.h., f ist konstant gleich 0 auf X. In beiden Fällen ist f konstant auf X, was zu beweisen war.
iv) $\Rightarrow$ *i)*: Sei $U \subset X$ offene und abgeschlossene Teilmenge von X. Dann ist $U^c = X \setminus U$ ebenfalls offen und abgeschlossen in X mit $X = U \cup U^c$ und $U \cap U^c = \varnothing$. Ist U nichtleer und nicht gleich X, so ist U^c nichtleer und nicht gleich X; in diesem Fall definiert die Abbildung $f : X \longrightarrow \{0, 1\}$, $x \in U \longmapsto 0$, $x \in U^c \longmapsto 1$, eine stetige Abbildung (wegen $f^{-1}(\{0\}) = U$ und $f^{-1}(\{1\}) = U^c$ offen in X, wobei $\{0\}$ und $\{1\}$ offen in $\{0, 1\}$ sind), die aber nicht konstant ist, Widerspruch. Somit muss entweder $U = \varnothing$ oder $U = X$ gelten, d.h., X muss zusammenhängend sein. □

Beispiele 2.47

1. Eine Teilmenge I von $\mathbb{R}$ (mit der Standardmetrik) ist genau dann zusammenhängend, wenn I ein Intervall ist. Denn: Ist I kein Intervall, so existieren ein x mit $x \notin I$ und $y, z \in I$ mit $y < x < z$. Die Teilmengen $I_1 :=]-\infty, x[\cap I$ und $I_2 :=]x, +\infty[\cap I$ sind dann offen, nichtleer (denn es gelten $y \in I_1$ und $z \in I_2$) und disjunkt in I mit $I = I_1 \cup I_2$. Aus Proposition 2.46 folgt, dass I nicht zusammenhängend ist. Ist umgekehrt I ein Intervall, so betrachte eine beliebige stetige Abbildung $f : I \longrightarrow \{0, 1\}$. Angenommen, f sei nicht konstant (also surjektiv). Dann existieren $x_0, y_0 \in I$ mit $f(x_0) = 0$ und $f(y_0) = 1$. O.B.d.A. sei $x_0 < y_0$ (sonst ersetze f durch $\tilde{f} := 1 - f$). Konstruiere induktiv die Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ durch: falls $f(\frac{x_n+y_n}{2}) = 0$ setze man $x_{n+1} := \frac{x_n+y_n}{2}$ und $y_{n+1} := y_n$, sonst setze man $x_{n+1} := x_n$

[16] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

und $y_{n+1} := \frac{x_n+y_n}{2}$. Dann sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ beschränkte monotone Folgen mit $|x_n - y_n| \le \frac{1}{2^n}|x_0 - y_0|$ (*Übungsaufgabe*), somit konvergieren sie beide gegen einen und denselben Punkt $l \in \mathbb{R}$. Da I ein Intervall ist und $x_n \le l \le y_n$ für alle $n \in \mathbb{N}$ gilt, gilt $l \in I$. Da f stetig ist, gilt $\lim\limits_{n\to\infty} f(x_n) = f(l) = \lim\limits_{n\to\infty} f(y_n)$, d.h. $0 = f(l) = 1$, Widerspruch. Mit Proposition 2.46 folgt, dass I zusammenhängend ist.

2. Die Teilmenge $Y = [0,1] \cup [2,3]$ von $\mathbb{R}$ (mit der Standardmetrik) ist nicht zusammenhängend, denn $[0,1]$ und $[2,3]$ sind disjunkte nichtleere abgeschlossene Teilmengen von Y mit $Y = [0,1] \cup [2,3]$.

3. Die Teilmenge $\mathbb{Q}$ von $\mathbb{R}$ (mit der Standardmetrik) ist nicht zusammenhängend, denn es gibt zwischen je zwei rationalen Zahlen eine irrationale Zahl.

Proposition 2.48 (Vereinigung zusammenhängender Teilmengen) *Sei $(Y_i)_{i\in I}$ eine beliebige Familie zusammenhängender Teilmengen eines metrischen Raumes (X,d) mit $\bigcap\limits_{i\in I} Y_i \neq \varnothing$. Dann ist $\bigcup\limits_{i\in I} Y_i$ zusammenhängend.*

Beweis:[17] Betrachte $x \in \bigcap\limits_{i\in I} Y_i$ (ein solches x existiert wegen der Voraussetzung $\bigcap\limits_{i\in I} Y_i \neq \varnothing$) und sei Ω eine Teilmenge von $\bigcup\limits_{i\in I} Y_i$, die gleichzeitig offen und abgeschlossen in $\bigcup\limits_{i\in I} Y_i$ ist. Dann ist, für jedes $i \in I$, die Teilmenge $\Omega \cap Y_i$ offen und abgeschlossen und Y_i. Ist Ω nichtleer, so trifft Ω mindestens eines der Y_i's, sei es Y_{i_0}, und da Y_{i_0} zusammenhängend ist, gilt $\Omega \cap Y_{i_0} = Y_{i_0}$. Insbesondere gilt $x \in \Omega$ und somit $\Omega \cap Y_i \neq \varnothing$ für alle $i \in I$ nach der Wahl von x. Daraus folgt $\Omega \cap Y_i = Y_i$ für alle $i \in I$, d.h., $\Omega = \bigcup\limits_{i\in I} Y_i$. □

Bemerkungen 2.49

1. Die Voraussetzung $\bigcap\limits_{i\in I} Y_i \neq \varnothing$ bei Proposition 2.48 ist wichtig. Z.B. gilt für $X := \mathbb{R}$, $Y_1 := [0,1]$ und $Y_2 := [2,3]$: Y_1 und Y_2 sind zusammenhängend, $Y_1 \cup Y_2$ aber nicht, siehe Beispiel 2.47.2.

2. Die Teilmenge $\bigcap\limits_{i\in I} Y_i$ muss nicht zusammenhängend sein. Z.B. gilt für $Y_1 := \{x \in \mathbb{R}^2 \setminus \{0\} \,|\, x_2 \ge 0\}$ und $Y_2 := \{x \in \mathbb{R}^2 \setminus \{0\} \,|\, x_2 \le 0\}$: Y_1, Y_2 sind zusammenhängend, $Y_1 \cap Y_2$ aber nicht.

Proposition 2.50 (Abschluss einer zusammenhängenden Teilmenge) *Sei Y eine zusammenhängende Teilmenge eines metrischen Raumes (X,d). Dann ist $\overline{Y}$ zusammenhängend.*

Beweis:[18] Sei Ω eine offene und abgeschlossene Teilmenge von $\overline{Y}$, dann ist $\Omega \cap Y$ auch offen und abgeschlossen in Y, so mit gilt entweder $\Omega \cap Y = \varnothing$ oder $\Omega \cap Y = Y$. Im ersten Fall muss dann $\Omega = \varnothing$ gelten (benuzte dazu die Definition 2.9), im zweiten Fall muss $Y \subset \Omega$ gelten, woraus mit Proposition 2.11 $\Omega = \overline{Y}$ folgt (denn Ω ist abgeschlossen in $\overline{Y}$). □

[17] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[18] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Bemerkung 2.51 Die entsprechende Aussage für das Innere gilt nicht. Betrachte z.B. für Y die Vereinigung zweier sich in genau einem Punkt berührender abgeschlossener Kreisscheiben in $X := \mathbb{R}^2$. Dann ist Y zwar zusammenhängend, das Innere $\overset{\circ}{Y}$ von Y in X aber nicht: $\overset{\circ}{Y}$ ist die Vereinigung der entsprechenden *offenen* Kreisscheiben und diese sind disjunkt, insbesondere ist $\overset{\circ}{Y}$ nicht zusammenhängend.

Proposition 2.52 (Stetigkeit und Zusammenhang) *Sei $f : X \longrightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Ist $Z \subset X$ zusammenhängend, so ist $f(Z)$ zusammenhängend.*

Beweis:[19] Sei Ω offene und abgeschlossene Teilmenge von $f(Z)$, dann ist wegen der Stetigkeit von f die Teilmenge $f^{-1}(\Omega)$ offen und abgeschlossen in Z. Da Z zusammenhängend ist, gilt entweder $f^{-1}(\Omega) = \varnothing$ oder $f^{-1}(\Omega) = Z$. Im ersten Fall gilt dann $\Omega = \varnothing$ und im zweiten Fall gilt $\Omega = Z$. $\square$

Beispiele 2.53

1. Sei $C \subset \mathbb{R}^n$ eine nichtleere *konvexe* Teilmenge, d.h. für alle $x, y \in C$ gilt $[x, y] \subset C$, wobei $[x, y] := \{(1-t)x + ty \,|\, t \in [0, 1]\}$ die abgeschlossene Geradenstrecke zwischen x und y bezeichnet. Dann ist C zusammenhängend. Denn: für jedes $x \in C$ gilt $C = \bigcup\limits_{y \in C} [x, y]$ und jedes $[x, y]$ ist zusammenhängend (nach Proposition 2.52) mit $\bigcap\limits_{y \in C} [x, y] = \{x\} \neq \varnothing$ (wende dann Proposition 2.48 an).

2. Sei $S \subset \mathbb{R}^n$ eine nichtleere *sternförmige* Teilmenge, d.h., es gibt ein $x \in \mathbb{R}^n$ so, dass $[x, y] \subset S$ für alle $y \in S$. Bemerke, dass jede konvexe Teilmenge sternförmig ist, dass aber nicht jede sternförmige Teilmenge konvex ist: z.B.
$$S = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \,|\, x_1 \geq 0 \quad \text{oder} \quad (x_1 < 0 \quad \text{und} \quad x_2 \geq 0) \right\}$$
ist sternförmig aber nicht konvex. Wie für konvexe Teilmengen kann dennoch bewiesen werden, dass jede nichtleere sternförmige Teilmenge des $\mathbb{R}^n$ zusammenhängend ist (der Beweis ist wortwörtlich der gleiche, wähle dabei $x \in S$ mit $[x, y] \subset S$ für alle $y \in S$) .

3. Sei $f : X \longrightarrow Y$ eine stetige Abbildung wobei X zusammenhängend ist. Dann ist $\mathrm{Graph}(f) := \{(x, f(x)) | x \in X\} \subset X \times Y$ zusammenhängend. Denn: die Abbildung $X \longrightarrow X \times Y$, $x \longmapsto (x, f(x))$, ist stetig (*Übungsaufgabe*), insbesondere ist ihr Bild – d.h., Graph(f) – nach Proposition 2.52 zusammenhängend. Aus Proposition 2.50 folgt insbesondere, dass $\overline{\mathrm{Graph}(f)}$ (der Abschluss von Graph(f) in $X \times Y$) ebenfalls zusammenhängend ist. Als Anwendung dieser Beobachtung betrachten wir den Fall $X :=]0, \infty[$ und $f : X \longrightarrow \mathbb{R}$, $x \mapsto \sin(\frac{1}{x})$. Die Abbildung f ist stetig und X ist zusammenhängend, insbesondere ist nach der letzten Überlegung $\overline{\mathrm{Graph}(f)}$ zusammenhängend. Nun ist zu bemerken, dass $\overline{\mathrm{Graph}(f)} = \mathrm{Graph}(f) \bigcup \big(\{0\} \times [-1, 1]\big)$ nicht auf offensichtliche Weise zusammenhängend ist!

[19] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Korollar 2.54 (Zwischenwertsatz) *Sei $f : X \longrightarrow \mathbb{R}$ eine stetige Funktion auf einem zusammenhängenden metrischen Raum (X, d). Gilt $\{a, b\} \subset f(X)$ für reelle Zahlen $a < b$, so gilt $[a, b] \subset f(X)$, d.h. für jedes $c \in [a, b]$ existiert (mindestens) ein $x \in X$ mit $f(x) = c$.*

Beweis: Nach Proposition 2.52 ist $f(X)$ zusammenhängend, d.h., nach Beispiel 2.47.1, ein Intervall. Daraus folgt die Behauptung. □

Proposition 2.55 (Produkte zusammenhängender Räume) *Sind (X, d_X) und (Y, d_Y) zusammenhängende metrische Räume, so ist $(X \times Y, d_{X\times Y})$ zusammenhängend.*

Beweis:[20] Sei $f : X \times Y \longrightarrow \{0, 1\}$ eine beliebige stetige Abbildung. Dann ist, für jedes $x \in X$, die Abbildung $f_x : Y \longrightarrow \{0, 1\}$, $y \mapsto f(x, y)$ auch stetig, siehe Beispiel 2.21.4. Da Y zusammenhängend ist, muss dann f_x konstant sein. Wir bezeichnen diese Konstante weiterhin mit f_x. Die Abbildung $X \longrightarrow \{0, 1\}$, $x \mapsto f_x$, ist dann auch stetig auf X (sie stimmt mit der Abbildung $x \longmapsto f(x, y)$ überein, wobei $y \in Y$ beliebig gewählt werden kann). Da X zusammenhängend ist, muss dann diese Abbildung konstant sein. Insgesamt folgt, dass f konstant ist. Proposition 2.46 impliziert, dass $X\times Y$ zusammenhängend ist. □

Definition 2.56 *Sei (X, d) ein metrischer Raum und $x \in X$ ein Punkt. Die* Zusammenhangskomponente *von x in X ist die Vereinigung aller zusammenhängenden Teilmengen von X, die x enthalten.*

Beachte, dass $\{x\}$ wohl eine x enthaltende zusammenhängende Teilmenge von X ist; insbesondere ist die Zusammenhangskomponente von x in X wohldefiniert.

Proposition 2.57 *Sei (X, d) ein metrischer Raum und $x \in X$ ein Punkt. Dann gilt:*

i) Die Zusammenhangskomponente von x in X ist die größte zusammenhängende Teilmenge von X, die x enthält.

ii) Die Zusammenhangskomponente von x in X ist abgeschlossen in X.

iii) Die Beziehung

$$x \sim y :\Longleftrightarrow y \text{ liegt in der Zusammenhangskomponente von } x \text{ in } X,$$

definiert eine Äquivalenzrelation auf X. Insbesondere lässt sich jeder metrische Raum als disjunkte Vereinigung (abgeschlossener) zusammenhängender Teilmengen schreiben.

Beweis:[21] Aus Proposition 2.48 folgt, dass die Zusammenhangskomponente von x in X selber zusammenhängend ist, somit ist sie die größte zusammenhängende Teilmenge von X, die x enthält. Da der Abschluss einer zusammenhängenden Teilmenge zusammenhängend ist (Proposition 2.50), muss dann nach 1. die Zusammenhangskomponente von x in X abgeschlossen in X sein. Die letzte Aussage wird als Übungsaufgabe gelassen. □

Diese zusammenhängenden Teilmengen heißen *Zusammenhangskomponenten von X*. Beachte, dass sie nicht unbedingt offen sind: z.B. $\{0\}$ ist eine Zusammenhangskomponente des metrischen Raumes $X := \{0\} \cup \{\frac{1}{n} \,|\, n \in \mathbb{N},\, n \geq 1\}$, ist aber nicht offen in X.

[20] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

[21] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

2.2 Differenzierbarkeit

2.2.1 Differentialabbildung und partielle Ableitungen

Von hier aus bezeichne $\left(e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} i \right)_{i=1,\ldots,n}$ die kanonische Basis des $\mathbb{R}^n$.

Definition 2.58 *Sei $U \subset \mathbb{R}^n$ eine nichtleere offene Teilmenge des $\mathbb{R}^n$ und $f : U \longrightarrow \mathbb{R}^p$ eine Abbidlung. Sei $x \in U$ ein Punkt.*

i) Die Abbildung f heißt genau dann differenzierbar *in x, wenn eine lineare Abbildung $L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, ein $r > 0$ und eine Abbildung $\varepsilon : B_r(0) \longrightarrow \mathbb{R}^p$ so existieren, dass*

$$f(x+h) = f(x) + L(h) + |h| \cdot \varepsilon(h)$$

für alle $h \in B_r(0)$ gilt, wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$.

ii) Die Abbildung f besitzt genau dann eine Richtungsableitung *an x entlang $h \in \mathbb{R}^n$, wenn $\lim\limits_{t\to 0^+} \frac{f(x+th)-f(x)}{t} \in \mathbb{R}^p$ existiert.*

iii) Die Abbildung f besitzt genau dann eine partielle Ableitung *an x in Richtung x_i, $1 \leq i \leq n$, wenn die Abbildung $t \longmapsto f(x+te_i) = f(x_1, \ldots, x_i + t, \ldots, x_n)$ in $t = 0$ ableitbar ist.*

Bemerkungen 2.59

1. Ist f differenzierbar in $x \in U$, so ist $L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ linear mit $f(x+h) = f(x) + L(h) + |h| \cdot \varepsilon(h)$ für alle $h \in B_r(0)$, $r > 0$ hinreichend klein, eindeutig. Denn: sei $L' : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ und $\varepsilon' : B_{r'}(0) \longrightarrow \mathbb{R}^p$ mit L' linear und $f(x+h) = f(x) + L'(h) + |h| \cdot \varepsilon'(h)$ für alle $h \in B_{r'}(0)$ sowie $\varepsilon'(h) \underset{|h|\to 0}{\longrightarrow} 0$. Wähle $r'' := \min(r, r') > 0$. Dann gilt für alle $h \in B_{r''}(0)$: $f(x)+L(h)+|h|\varepsilon(h) = f(x+h) = f(x)+L'(h)+|h|\cdot\varepsilon'(h)$, insbesondere $(L'-L)(h) = |h| \cdot (\varepsilon(h) - \varepsilon'(h))$. Ersetzt man h durch th mit $t \in]0,1]$, so bekommt man wegen der Linearität von $L' - L$:
$$t(L'-L)(h) = (L'-L)(th) = |th| \cdot (\varepsilon(th) - \varepsilon'(th)) = t \cdot |h| \cdot (\varepsilon(th) - \varepsilon'(th)),$$
insbesondere gilt $(L'-L)(h) = |h| \cdot (\varepsilon(th) - \varepsilon'(th))$. Wegen $\varepsilon(th) - \varepsilon'(th) \underset{t\to 0}{\longrightarrow} 0$ gilt $(L'-L)(h) = 0$, d.h., $L'(h) = L(h)$. Dies gilt für alle $h \in B_{r''}(0)$ und, wegen der Linearität von L und L', für alle $h \in \mathbb{R}^n$. Insgesamt folgt $L = L'$.

2. Für $n = 1$ gilt: f ist genau dann differenzierbar in x, wenn f ableitbar in x ist. Denn: ist f differenzierbar in x, so existiert nach Voraussetzung ein $L : \mathbb{R} \longrightarrow \mathbb{R}^p$ linear und ein $\varepsilon :]-r, r[\longrightarrow \mathbb{R}^p$ (für ein $r > 0$) mit $f(x+h) = f(x) + L(h) + |h|\varepsilon(h)$ für alle $h \in]-r, r[$, wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$. Für alle $h \in]-r, r[\setminus\{0\}$ gilt insbesondere:

$\frac{f(x+h)-f(x)}{h} = \frac{L(h)+|h|\varepsilon(h)}{h} = L(1) + \frac{|h|}{|h|}\varepsilon(h) \underset{h\to 0}{\longrightarrow} L(1)$. Daraus folgt, dass f ableitbar in x ist mit $f'(x) = L(1)$. Ist umgekehrt f ableitbar in x, so definiert man $L : \mathbb{R} \longrightarrow \mathbb{R}^p$ durch $L(h) := h \cdot f'(x)$ und $\varepsilon :]-r, r[\longrightarrow \mathbb{R}^p$ durch $\varepsilon(h) := \frac{f(x+h)-f(x)-L(h)}{|h|}$ für alle $h \in]-r, r[, h \neq 0$ sowie $varepsilon(0) := 0$. Dann gilt $f(x+h) = f(x)+L(h)+|h|\varepsilon(h)$ für alle $h \in]-r, r[$ und $\varepsilon(h) = \frac{f(x+h)-f(x)}{|h|} - f'(x) \cdot \frac{h}{|h|} \underset{h\to 0}{\longrightarrow} 0$ nach Definition der Ableitung. Da L offensichtlich linear ist, folgt, dass f differenzierbar in x ist.

Definition 2.60 *Sei $U \subset \mathbb{R}^n$ eine nichtleere offene Teilmenge.*

1. *Ist $f : U \longrightarrow \mathbb{R}^p$ differenzierbar in $x \in U$, so heißt L wie in* Definition 2.58 *das* Differential *(oder die* Differentialabbildung*) von f in x. Wir bezeichnen sie mit $d_x f$.*

2. *Hat $f : U \longrightarrow \mathbb{R}^p$ eine partielle Ableitung an x in Richtung x_i, $1 \leq i \leq n$, so heißt die Ableitung von $t \longmapsto f(x_1, \ldots, x_i + t, \ldots, x_n)$ an der Stelle $t = 0$ die* erste partielle Ableitung *von f an x in Richtung x_i. Wir bezeichnen sie mit $\frac{\partial f}{\partial x_i}(x)$.*

Eine Abbildung $f : U \longrightarrow \mathbb{R}^p$ heißt genau dann differenzierbar auf U, wenn sie in jedem $x \in U$ differenzierbar ist.

Proposition 2.61 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine Abbildung, wobei $U \subset \mathbb{R}^p$ offen ist und $x \in U$ ein Punkt.*

i) Ist f differenzierbar in x, so hat f Richtungsableitungen in allen Richtungen an x und es gilt $d_x f(h) = \lim\limits_{t\to 0^+} \frac{f(x+th)-f(x)}{t}$ für alle $h \in \mathbb{R}^n$.

ii) Ist f differenzierbar in x, so hat f partielle Ableitungen in allen Richtungen x_i an x und es gilt: $\forall\ h = \sum\limits_{i=1}^{n} h_i e_i \in \mathbb{R}^n$,

$$d_x f(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x).$$

iii) Ist f differenzierbar in x, so ist f stetig in x.

Beweis:[22] *i)* Nach Definition gilt $f(x+h) = f(x) + d_x f(h) + |h| \cdot \varepsilon(h)$ für alle $h \in B_r(0)$ ($r > 0$ hinreichend klein), wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$. Daraus folgt, für alle $h \in B_r(0)$ und $t \in]0, 1]$:

$$\begin{array}{rcl} \frac{f(x+th)-f(x)}{t} & = & \frac{1}{t}(d_x f(th) + |th| \cdot \varepsilon(th)) \\ & = & \frac{t}{t}(d_x f(h) + |h| \cdot \varepsilon(th)) \\ & & \underset{t\to 0^+}{\longrightarrow} d_x f(h) \quad \text{wegen} \quad \varepsilon(th) \underset{t\to 0^+}{\longrightarrow} 0. \end{array}$$

ii) Im Spezialfall $h = \pm e_i$, $1 \leq i \leq n$, erhalten wir

$$d_x f(e_i) = \lim_{t\to 0^+} \frac{f(x+te_i) - f(x)}{t} \quad \text{und} \quad d_x(f)(-e_i) = \lim_{t\to 0^+} \frac{f(x-te_i) - f(x)}{t},$$

[22]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

d.h. $d_xf(e_i) = \lim\limits_{t\to 0^-} \frac{f(x+te_i)-f(x)}{t}$. Somit existiert $\lim\limits_{t\to 0} \frac{f(x+te_i)-f(x)}{t}$, d.h. $t \longmapsto f(x+te_i)$ ist in $t = 0$ ableitbar, d.h., f hat eine partielle Ableitung an x in Richtung x_i und es gilt $d_xf(e_i) = \frac{\partial f}{\partial x_i}(x)$. Insbesondere folgt, für alle $h = \sum\limits_{i=1}^n h_ie_i \in \mathbb{R}^n$:

$$d_xf(h) = d_xf(\sum_{i=1}^n h_ie_i) = \sum_{i=1}^n h_i d_xf(e_i) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x).$$

iii) Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge aus U mit $x_k \underset{k\to\infty}{\longrightarrow} x$. Sei $r > 0$ s.d. $f(x+h) = f(x)+d_xf(h)+|h|\cdot \varepsilon(h)$ für alle $h \in B_r(0)$, wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$. Dann existiert ein $N \in \mathbb{N}$ so, dass $|x_k - x| < r$ für alle $k \geq N$ gilt. Für jedes solche k gilt dann

$$f(x_k) - f(x) = d_xf(x_k - x) + |x_k - x| \cdot \varepsilon(x_k - x).$$

Wegen $|x_k - x| \underset{k\to\infty}{\longrightarrow} 0$ gilt $\varepsilon(x_k - x) \underset{k\to\infty}{\longrightarrow} 0$ und somit auch $|x_k - x| \cdot \varepsilon(x_k - x) \underset{k\to\infty}{\longrightarrow} 0$. Außerdem gilt $d_xf(x_k - x) \underset{k\to\infty}{\longrightarrow} 0$, denn: $d_xf(x_k - x) = \sum\limits_{i=1}^n (x_k - x)_i \frac{\partial f}{\partial x_i}(x)$ nach *ii)* und für jedes $1 \leq i \leq n$ gilt $|(x_k - x)_i| \leq |x_k - x| \underset{k\to\infty}{\longrightarrow} 0$, somit $|\sum\limits_{i=1}^n (x_k - x)_i \frac{\partial f}{\partial x_i}(x)| \leq \sum\limits_{i=1}^n |(x_k - x)_i| \cdot |\frac{\partial f}{\partial x_i}(x)| \underset{k\to\infty}{\longrightarrow} 0$. Daraus folgt $|f(x_k) - f(x)| \underset{k\to\infty}{\longrightarrow} 0$. Dies zeigt, dass f stetig in x ist. □

Bemerkungen 2.62

1. Die Umkehrung der Aussagen aus Proposition 2.61 gilt nicht. Sei z.B. $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{falls } (x,y) \neq (0,0) \\ 0 & \text{falls } (x,y) = (0,0) \end{cases}$. Dann hat f partielle Ableitungen an $0 \in \mathbb{R}^2$ in Richtungen x und y: es gilt $f(0,t) = \begin{cases} \frac{0\cdot t}{0+t^2} = 0 & \text{falls } t \neq 0 \\ 0 & \text{sonst.} \end{cases}$, d.h., $f(0,t) = 0$ für alle $t \in \mathbb{R}$, somit gilt $\frac{\partial f}{\partial y}(0,0) = 0$. Analog ist $\frac{\partial f}{\partial x}(0,0) = 0$. Dennoch ist f nicht stetig in $(0,0) \in \mathbb{R}^2$ (siehe Beispiel 2.21.5), insbesondere nicht differenzierbar in $(0,0)$. Ein anderes Beispiel ist $f(x,y) := \begin{cases} 1 & \text{falls } y = x^2 \text{ und } x > 0 \\ 0 & \text{sonst.} \end{cases}$. Man kann zeigen, dass f Richtungsableitungen an $0 = (0,0) \in \mathbb{R}^2$ in allen Richtungen hat, die alle verschwinden; dennoch ist f auch nicht stetig in 0.

2. Im Fall $p = 1$, d.h., $f : U \longrightarrow \mathbb{R}$, bezeichnet

$$\mathrm{grad}(f)(x) := (\nabla f)(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^n$$

den *Gradienten* von f (oder das *Gradientenfeld* von f) an der Stelle $x \in U$. Dieser Vektor ist wohldefiniert, sobald f partielle Ableitungen in allen Richtungen x_i, $1 \leq i \leq n$, hat. Ist f zusätzlich differenzierbar in $x \in U$, so gilt nach Proposition 2.61:

$$d_xf(h) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) = \langle \nabla f(x), h \rangle$$

für alle $h \in \mathbb{R}^n$, wobei $\langle \cdot, \cdot \rangle$ das kanonische euklidische Skalarprodukt im $\mathbb{R}^n$ bezeichnet.

Proposition 2.63 (Kettenregel) *Seien $f : U \longrightarrow \mathbb{R}^p$, wobei $U \subset \mathbb{R}^n$ offen ist, und $g : V \longrightarrow \mathbb{R}^q$, wobei $V \subset \mathbb{R}^p$ offen ist, Abbildungen mit $f(U) \subset V$. Ist f differenzierbar in $x \in U$ und g differenzierbar in $f(x) \in V$, so ist $g \circ f : U \longrightarrow \mathbb{R}^q$ differenzierbar in $x \in U$ und es gilt*

$$d_x(g \circ f) = d_{f(x)} g \circ d_x f.$$

Beweis:[23] Nach Voraussetzung existieren $r, r' > 0$ mit $B_r(x) \subset U$, $B_{r'}(f(x)) \subset V$ und Abbildungen $\varepsilon : B_r(x) \longrightarrow \mathbb{R}^p$ und $\varepsilon' : B_{r'}(f(x) \longrightarrow \mathbb{R}^q$ so, dass für alle $h \in B_r(0)$ und $h' \in B_{r'}(0)$

$$\begin{aligned} f(x+h) &= f(x) + d_x f(h) + |h| \cdot \varepsilon(h) \\ g(f(x)+h') &= g(f(x)) + d_{f(x)} g(h') + |h'| \cdot \varepsilon'(h') \end{aligned}$$

gelten, wobei $d_x f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ sowie $d_{f(x)} g : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ linear sind und $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$ sowie $\varepsilon'(h') \underset{|h'|\to 0}{\longrightarrow} 0$. Da f stetig in x ist (Proposition 2.61), existiert ein $r'' \in]0, r[$ s.d. $|f(x+h)| - f(x)| < r'$ für alle $h \in B_{r''}(0)$. Für alle $h \in B_{r''}(0)$ gilt dann:

$$\begin{aligned} g \circ f(x+h) &= g(f(x+h)) \\ &= g(f(x) + \underbrace{d_x f(h) + |h| \cdot \varepsilon(h)}_{h'}) \\ &= g(f(x)) + d_{f(x)} g(h') + |h'| \cdot \varepsilon'(h') \\ &= g \circ f(x) + d_{f(x)} g(d_x f(h)) + d_{f(x)} g(|h| \cdot \varepsilon(h)) + |h'| \cdot \varepsilon'(h'), \end{aligned}$$

wo wir im letzten Schritt die Linearität von $d_{f(x)} g$ benutzt haben. Wieder wegen der Linearität von $d_{f(x)} g$ gilt $d_{f(x)} g(|h| \cdot \varepsilon(h)) + |h'| \cdot \varepsilon'(h') = |h| \cdot (d_{f(x)} g(\varepsilon(h)) + \frac{|h'|}{|h|} \varepsilon'(h'))$ für alle $h \in B_{r''}(0) \setminus \{0\}$. Aus $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$ folgt $d_{f(x)} g(\varepsilon(h)) \underset{|h|\to 0}{\longrightarrow} 0$ (siehe z.B. Einschub über Matrixnormen). Außerdem ist

$$\frac{|h'|}{|h|} = \frac{|d_x f(h) + |h| \cdot \varepsilon(h)|}{|h|} = |d_x f(\frac{h}{|h|}) + \varepsilon(h)|,$$

insbesondere $\frac{|h'|}{|h|} \leq \|d_x f\| \cdot \underbrace{|\frac{h}{|h|}|}_{1} + |\varepsilon(h)| \leq \|d_x f\| + 1$ für alle $h \in B_{r''}(0)$ hinreichend klein (d.h., $|h| < \tilde{r}$ für ein hinreichend kleines $\tilde{r} \in]0, r''[$). Wegen $|h'| \leq (\|d_x f\| + 1) \cdot |h|$ gilt insbesondere $|h'| \underset{|h|\to 0}{\longrightarrow} 0$ und somit $\varepsilon'(h') \underset{|h'|\to 0}{\longrightarrow} 0$. Insgesamt bekommen wir $g \circ f(x+h) = g \circ f(x) + d_{f(x)} g \circ d_x f(h) + |h| \cdot \epsilon(h)$ für alle $h \in B_{\tilde{r}}(0)$, wobei $\epsilon(h) := d_{f(x)} g(\varepsilon(h)) + \frac{|d_x f(x) + |h| \varepsilon(h)|}{|h|} \cdot \varepsilon'(d_x f(h) + |h| \varepsilon(h))$ erfüllt: $|\epsilon(h)| \leq |d_{f(x)} g(\varepsilon(h))| + (\|d_x f\| + 1)$ $varepsilon'(d_x f(h) + |h| \varepsilon(h)) \underset{|h|\to 0}{\longrightarrow} 0$.

Dies beweist, dass $g \circ f$ differenzierbar in $x \in U$ ist mit Differential

$$d_x(g \circ f) = d_{f(x)} g \circ d_x f.$$

□

Einschub über Matrixnormen:

[23] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

- Sei $A \in M_{p\times n}(\mathbb{R})$ eine Matrix. Man definiert

$$\|A\| := \sup_{x\in\mathbb{R}^n\setminus\{0\}} \left(\frac{|A\cdot x|}{|x|}\right)$$

(Norm – auch *Spektralnorm* genannt – der Matrix A). Folgende Eigenschaften können überprüft werden: $\|A\| \in [0,\infty[$, $\|A\| = 0 \Longleftrightarrow A = 0$, $\|A+B\| \leq \|A\|+\|B\|$ und $\|\lambda\cdot A\| = |\lambda|\cdot\|A\|$ für alle $A, B \in M_{p\times n}(\mathbb{R})$ und $\lambda \in \mathbb{R}$. Somit definiert $\|\cdot\|$ eine sogenannte Norm auf dem Vektorraum $M_{p\times n}(\mathbb{R})$. Aus diesen Eigenschaften folgt, dass $d(A,B) := \|A-B\|$ eine Metrik auf $M_{p\times n}(\mathbb{R})$ definiert.

- Nach Definition von $\|\cdot\|$ gelten:

$$\begin{array}{lll} * & |A\cdot x| \leq \|A\|\cdot|x| & \text{für alle } x \in \mathbb{R}^n, \\ * & \|A\cdot B\| \leq \|A\|\cdot\|B\| & \text{für alle } A \in M_{p\times n}(\mathbb{R}) \text{ und } B \in M_{n\times q}(\mathbb{R}). \end{array}$$

Beispiele 2.64

1. Ist $f : U \longrightarrow \mathbb{R}^p$ konstant, so ist f differenzierbar auf U und es gilt $d_xf = 0$ für alle $x \in U$. Denn: für $r > 0$ mit $B_r(x) \subset U$ gilt, $\forall\ h \in B_r(0)$,

$$f(x+h) - f(x) = 0 = 0\cdot h + |h|\cdot 0,$$

d.h., mit $L = 0$ und $\varepsilon(h) := 0$ bekommt man $f(x+h) = f(x) + L(h) + |h|\cdot\varepsilon(h)$ und $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$.

2. Ist $f := A_{|_U} : U \longrightarrow \mathbb{R}^p$, wobei $A : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ linear ist, so ist f differenzierbar auf U und es gilt $d_xf = A$ für alle $x \in U$. Denn: für $r > 0$ mit $B_r(x) \subset U$ gilt, für alle $h \in B_r(0)$,

$$\begin{aligned} f(x+h) - f(x) &= A(x+h) - A(x) \\ &= A(x) + A(h) - A(x) \\ &= A(h), \end{aligned}$$

wobei A bereits linear ist. Setzt man $\varepsilon(h) := 0$ für alle $h \in B_r(0)$, so gilt $f(x+h) = f(x) + A(h) + |h|\cdot\varepsilon(h)$ für alle $h \in B_r(0)$, mit $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$. Dies zeigt, dass f differenzierbar in $x \in U$ ist mit $d_xf = A$.

3. Seien $f : U \longrightarrow \mathbb{R}^p$ und $g : U \longrightarrow \mathbb{R}^{p'}$ Abbildungen und $x \in U$ ein Punkt. Betrachte die Abbildung

$$\begin{aligned} (f,g) :\ & U \longrightarrow \mathbb{R}^p \times \mathbb{R}^{p'} = \mathbb{R}^{p+p'} \\ & x \longmapsto (f(x), g(x)) \end{aligned}$$

Behauptung: *Die Abbildung (f,g) ist differenzierbar in $x \Longleftrightarrow f$ und g sind differenzierbar in x. Ggf. gilt $d_x(f,g)(h) = (d_xf(h), d_xg(h))$ für alle $h \in \mathbb{R}^n$.*

Beweis: "$\Longrightarrow$": Sind f und g differenzierbar in $x \in U$, so definiere die Abbildungen $\mathrm{pr}_1 : \mathbb{R}^{p+p'} \longrightarrow \mathbb{R}^p$, $x = (x_1,\ldots,x_{p+p'}) \longmapsto (x_1,\ldots,x_p)$ und $\mathrm{pr}_2 : \mathbb{R}^{p+p'} \longrightarrow \mathbb{R}^{p'}$, $x = (x_1,\ldots,x_{p+p'}) \longmapsto (x_{p+1},\ldots,x_{p+p'})$. Die Abbildung pr_1 ist die kanonische Projektion auf die p ersten Koordinaten, die Abbildung pr_2 ist die kanonische Projektion auf die p' letzten Koordinaten. Beide sind lineare Abbildungen. Nach dem Beispiel 2.64.2 und der Kettenregel sind dann $\mathrm{pr}_1 \circ (f,g)$ und $\mathrm{pr}_2 \circ (f,g)$ differenzierbar in

x. Wegen $\mathrm{pr}_1 \circ (f,g) = f$ und $\mathrm{pr}_2 \circ (f,g) = g$ folgt, dass f und g differenzierbar in x sind.

"$\Longleftarrow$": Sind f und g differenzierbar in x, so schreibe

$$\begin{array}{rrcl} & f(x+h) & = & f(x) + d_x f(h) + |h| \cdot \varepsilon(h) \\ \text{und} & g(x+h) & = & g(x) + d_x g(h) + |h| \cdot \varepsilon'(h), \end{array}$$

wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$ und $\varepsilon'(h) \underset{|h|\to 0}{\longrightarrow} 0$. Daraus folgt

$$\begin{aligned} (f,g)(x+h) &= (f(x+h), g(x+h)) \\ &= (f(x) + d_x f(h) + |h| \cdot \varepsilon(h), g(x) + d_x g(h) + |h| \cdot \varepsilon'(h)) \\ &= (f(x), g(x)) + (d_x f(h), d_x g(h)) + (|h| \cdot \varepsilon(h), |h| \cdot \varepsilon'(h)) \\ &= (f(x), g(x)) + (d_x f(h), d_x g(h)) + |h| \cdot (\varepsilon(h), \varepsilon'(h)), \end{aligned}$$

wobei $h \longmapsto (d_x f(h), d_x g(h))$, $\mathbb{R}^n \longrightarrow \mathbb{R}^p \times \mathbb{R}^{p'}$, linear ist und $|(\varepsilon(h), \varepsilon'(h))| \underset{|h|\to 0}{\longrightarrow} 0$ (benutze $|(\varepsilon(h), \varepsilon'(h))| \leq \sqrt{2} \cdot \max(|\varepsilon(h)|, |\varepsilon'(h)|)$). Dies zeigt, dass (f,g) differenzierbar in x ist mit Differential $h \longmapsto (d_x f(h), d_x g(h))$. $\checkmark$

4. (Produktregel) Seien $f, g : U \longrightarrow \mathbb{R}^p$ Abbildungen und $b : \mathbb{R}^p \times \mathbb{R}^p \longrightarrow \mathbb{R}^q$ eine *bilineare* Abbildung, d.h. für alle $x, y \in \mathbb{R}^p$ sind $\mathbb{R}^p \longrightarrow \mathbb{R}^q$, $z \longmapsto b(x,z)$ und $\mathbb{R}^p \longrightarrow \mathbb{R}^q$, $z \longmapsto b(z,y)$, linear. Sind f und g differenzierbar in $x \in U$, so ist die Abbildung $b(f,g) : U \longrightarrow \mathbb{R}^q$, $x \longmapsto b(f(x), g(x))$, differenzierbar in x und es gilt

$$d_x(b(f,g))(h) = b(d_x f(h), g(x)) + b(f(x), d_x g(h))$$

für alle $h \in \mathbb{R}^n$. Denn: man schreibt $b(f,g)$ als $b \circ (f,g)$, wobei $(f,g) : U \longrightarrow \mathbb{R}^p \times \mathbb{R}^p$, $x \longmapsto (f(x), g(x))$. Da f und g differenzierbar in x sind, ist (f,g) differenzierbar in x nach Beispiel 2.64.3 oben.

Behauptung: *Die Abbildung b ist differenzierbar auf $\mathbb{R}^p \times \mathbb{R}^p$ und es gilt, für alle $x, y \in \mathbb{R}^p$:*

$$d_{(x,y)} b(h, h') = b(h, y) + b(x, h')$$

für alle $h, h' \in \mathbb{R}^p$.

Beweis: Wir berechnen für alle $x, y, h, h' \in \mathbb{R}^p$:

$$\begin{aligned} b(x+h, y+h') &= b(x+h, y) + b(x+h, h') \\ &= b(x,y) + b(h,y) + b(x,h') + b(h,h'). \end{aligned}$$

Nun ist die Abbildung $(h,h') \longmapsto b(h,y) + b(x,h')$, $\mathbb{R}^p \times \mathbb{R}^p \longrightarrow \mathbb{R}^q$, linear und es gilt $\frac{b(h,h')}{|(h,h')|} \underset{|(h,h')|\to 0}{\longrightarrow} 0$: schreibe nämlich $h = \sum_{i=1}^p h_i e_i$, $h' = \sum_{i=1}^p h'_i e_i$, dann ist $b(h,h') = \sum_{i,j=1}^p h_i h'_j b(e_i, e_j)$, insbesondere

$$\begin{aligned} |b(h,h')| &\leq \sum_{i,j=1}^p |h_i h'_j b(e_i, e_j)| \\ &\leq \sum_{i,j=1}^p |h_i| \cdot |h'_j| \cdot |b(e_i, e_j)| \\ &\leq |h| \cdot |h'| \cdot \underbrace{\sum_{i,j=1}^p |b(e_i, e_j)|}_{=:M}, \end{aligned}$$

wobei $|h_i| \leq |h|$ und $|h'_i| \leq |h'|$ benutzt wurden; es gilt, für alle $(h, h') \neq 0$:

$$\frac{|b(h,h')|}{|(h,h')|} \leq \frac{|h|\cdot|h'|}{|(h,h')|}\cdot M \leq \frac{|(h,h')|^2}{|(h,h')|}\cdot M = M\cdot|(h,h')| \underset{|(h,h')|\to 0}{\longrightarrow} 0.$$

Hierbei haben wir $|h| \leq |(h,h')|$ und $|h'| \leq |(h,h')|$ verwendet. Daraus folgt $b(x+h, y+h') = b(x,y)+b(h,y)+b(x,h')+|(h,h')|\cdot\varepsilon(h,h')$, wobei $\varepsilon(h,h') := \frac{b(h,h')}{|(h,h')|} \underset{|(h,h')|\to 0}{\longrightarrow} 0$. Dies beweist die Behauptung. $\checkmark$

Aus der Behauptung und der Kettenregel folgt, dass $b(f,g)$ differenzierbar in x ist mit

$$\begin{aligned} d_x(b(f,g))(h) &= (d_{(f(x),g(x))}b)\circ(d_x(f,g)(h)) \\ &= d_{(f(x),g(x))}b\circ(d_xf(h), d_xg(h)) \\ &= b(d_xf(h), g(x)) + b(f(x), d_xg(h)) \end{aligned}$$

für alle $h \in \mathbb{R}^n$, was zu beweisen war.

4'. Spezialfälle vom Beispiel 2.64.4 sind u.a. folgende:

- Sei $b = \langle\cdot\,,\cdot\rangle : \mathbb{R}^p\times\mathbb{R}^p \longrightarrow \mathbb{R}$ das kanonische euklidische Skalarprodukt auf $\mathbb{R}^p$. Sind f und g differenzierbar in $x \in U$, so ist $\langle f, g\rangle : U \longrightarrow \mathbb{R}$ differenzierbar in x und es gilt für alle $h \in \mathbb{R}^n$:

  $$d_x\langle f,g\rangle(h) = \langle d_xf(h), g(x)\rangle + \langle f(x), d_xg(h)\rangle.$$

  Im Fall $n = 1$ ist diese Identität äquivalent zur Produktregel für Kurven, siehe Proposition 1.10. Im noch spezielleren Fall $(n = 1, p = 1)$ ist diese Identität äquivalent zur üblichen Produktregel $(fg)' = f'\cdot g + f\cdot g'$ aus der Analysis I.
- Sei $b : \mathbb{R}^3\times\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ das *Kreuzprodukt* auf $\mathbb{R}^3$, also $b(x,y) := x\times y := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$ für alle $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ und $y = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ aus $\mathbb{R}^3$. Beachte, dass b wohl bilinear ist. Sind nun $f : U \longrightarrow \mathbb{R}^3$ und $g : U \longrightarrow \mathbb{R}^3$ differenzierbar in $x \in U$, so ist $f\times g : U \longrightarrow \mathbb{R}^3$ differenzierbar in x mit $d_x(f\times g)(h) = d_xf(h)\times g(x) + f(x)\times d_xg(h)$ für alle $h \in \mathbb{R}^n$.

5. (Linearität des Differentials) Eine weitere Anwendung der obigen Beispiele ist folgende:
Behauptung: *Sind $f : U \longrightarrow \mathbb{R}^n$ und $g : U \longrightarrow \mathbb{R}^n$ differenzierbar in $x \in U$, so ist für alle $\lambda, \mu \in \mathbb{R}$ die Abbildung $\lambda f + \mu g : U \longrightarrow \mathbb{R}^n$ differenzierbar in x und es gilt, für alle $h \in \mathbb{R}^n$: $d_x(\lambda f + \mu g)(h) = \lambda d_xf(h) + \mu d_xg(h)$.*
Beweis: Zu beachten sind: $\mathbb{R}^p \longrightarrow \mathbb{R}^p$, $x \overset{\lambda\cdot\mathrm{Id}}{\longmapsto} \lambda\cdot x$ ist linear und $\mathbb{R}^p\times\mathbb{R}^p \overset{b}{\longrightarrow} \mathbb{R}^p$, $(x,y) \longmapsto x+y$ ist linear. Nach den Beispielen 2.64.2 und 2.64.3 gilt: $\lambda f + \mu g = b\circ(\lambda\mathrm{Id}, \mu\mathrm{Id})\circ(f,g)$ ist als Nacheinanderausführung von differenzierbaren Abbildungen (in x) wieder differenzierbar in x und es gilt, für alle $h \in \mathbb{R}^n$:

$$\begin{aligned} d_x(\lambda f + \mu g)(h) &= d_{(\lambda f(x),\mu g(x))}b\circ(d_x(\lambda f)(h), d_x(\mu g)(h)) \\ &= b\circ(\lambda d_xf(h), \mu d_xg(h)) \\ &= \lambda d_xf(h) + \mu d_xg(h). \end{aligned}$$

$\checkmark$

Definition 2.65 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine Abbildung mit $U \subset \mathbb{R}^n$ offen, und $x \in U$ ein Punkt. Hat f partielle Ableitungen in allen Richtungen x_i, $1 \leq i \leq n$, in x, so heißt die Matrix*

$$J_f(x) := \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$$

die Jacobi-Matrix *von f an der Stelle x. Hierbei bezeichnen die f_i's die Koordinatenfunktionen von f, d.h., $f(y) = (f_1(y), \ldots, f_p(y))$ für alle $y \in U$.*

Bemerkungen 2.66

1. Bemerke: ist f zusätzlich differenzierbar in x, so gilt nach Proposition 2.61:

$$d_x f(h) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(x) = \sum_{i=1}^{p}\sum_{j=1}^{n} h_j \frac{\partial f_i}{\partial x_j}(x) e_i,$$

d.h., $d_x f(h) = J_f(x) \cdot h$.

2. Ist $f : U \longrightarrow V$ differenzierbar in $x \in U$ und ist $g : V \longrightarrow \mathbb{R}^q$ differenzierbar in $f(x) \in V$, so gilt für die Jacobi-Matrizen:

$$J_{g\circ f}(x) = J_g(f(x)) \cdot J_f(x).$$

Für den Beweis wird die Kettenregel angewendet, unter Berücksichtigung, dass die Nacheinanderausführung linearer Abbildungen dem Produkt der entsprechenden Matrizen entspricht (in den entsprechenden kanonischen Basen von $\mathbb{R}^n$, $\mathbb{R}^p$ und $\mathbb{R}^q$).

3. In der Physik werden die Bezeichnungen bei der Anwendung der Kettenregel möglichst intuitiv gestaltet – wenn auch manchmal etwas irreführend! Dazu geben wir zwei Beispiele an:
   - "*Für eine Funktion $f = f(x, y)$ und eine weitere Funktion $z = z(x, y)$ gelten: $\frac{\partial f(x,y)}{\partial y} = \frac{\partial f(x,z)}{\partial z} \cdot \frac{\partial z(x,y)}{\partial y}$ und $\frac{\partial f(x,y)}{\partial x} = \frac{\partial f(x,z)}{\partial x} + \frac{\partial f(x,z)}{\partial z} \cdot \frac{\partial z(x,y)}{\partial x}$.*"
   Dabei wird folgendes gemeint: sei $f : U \longrightarrow \mathbb{R}$ eine Abbildung, $(x, y) \in U$ ein Punkt, wobei $U \subset \mathbb{R}^2$ offen (und nichtleer) ist. Angenommen, f sei differenzierbar auf U. Sei $z : U \longrightarrow \mathbb{R}$ eine in (x, y) differenzierbare Funktion. Wir betrachten die Abbildungen $h : U \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (x, z(x, y))$ und $g := f \circ h : U \longrightarrow \mathbb{R}$, d.h., $g(x, y) = f(x, z(x, y))$. Dann ist g differenzierbar in $(x, y) \in U$ und es gelten: $\frac{\partial g}{\partial y}(x, y) = \frac{\partial f}{\partial x_2}(x, z(x, y)) \cdot \frac{\partial z}{\partial y}(x, y)$ und $\frac{\partial g}{\partial x}(x, y) = \frac{\partial f}{\partial x_1}(x, z(x, y)) + \frac{\partial f}{\partial x_2}(x, z(x, y)) \cdot \frac{\partial z}{\partial x}(x, y)$.
   Begründung: nach der Kettenregel gilt $J_g(x, y) = J_f(h(x, y)) \cdot J_h(x, y)$, wobei

$$\begin{aligned} J_f(x', y') &= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x', y') & \frac{\partial f}{\partial x_2}(x', y') \end{pmatrix} \\ \text{und}\quad J_h(x', y') &= \begin{pmatrix} \frac{\partial x}{\partial x}(x', y') & \frac{\partial x}{\partial y}(x', y') \\ \frac{\partial z}{\partial x}(x', y') & \frac{\partial z}{\partial y}(x', y') \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\partial z}{\partial x}(x', y') & \frac{\partial z}{\partial y}(x', y') \end{pmatrix}, \end{aligned}$$

für alle $(x', y') \in U$. Daraus folgt

$$\begin{aligned} J_g(x, y) &= \left(\frac{\partial f}{\partial x_1}(h(x, y)) \quad \frac{\partial f}{\partial x_2}(h(x, y))\right) \cdot \begin{pmatrix} 1 & 0 \\ \frac{\partial z}{\partial x}(x, y) & \frac{\partial z}{\partial y}(x, y) \end{pmatrix} \\ &= \left(\frac{\partial f}{\partial x_1}(h(x, y)) + \frac{\partial f}{\partial x_2}(h(x, y)) \cdot \frac{\partial z}{\partial x}(x, y) \quad \frac{\partial f}{\partial x_2}(h(x, y)) \cdot \frac{\partial z}{\partial y}(x, y)\right). \end{aligned}$$

Dies liefert die beiden Identitäten.

- „$\frac{\partial f(x,y)}{\partial x} = -\frac{\partial f}{\partial y}(x,y) \cdot \frac{\partial y}{\partial x}(x,f)$". Dabei wird folgendes gemeint: angenommen, es gibt eine Funktion $z = z(x)$ so, dass z differenzierbar (also ableitbar) ist und $x \longmapsto f(x, z(x))$ *konstant* ist für alle x in einem offenen Intervall. Dann gilt für alle x aus diesem Intervall: $\frac{\partial f}{\partial x}(x, z(x)) = -\frac{\partial f}{\partial y}(x, z(x)) \cdot \underbrace{\frac{\partial z}{\partial x}}_{z'(x)}(x)$.

  Begründung: Wir leiten die konstante Funktion $x \longmapsto f(x, z(x))$ nach x ab und wenden die Kettenregel an: $0 = \frac{d}{dx}(f(x, z(x))) = \left(\frac{\partial f}{\partial x}(x, z(x)) \quad \frac{\partial f}{\partial y}(x, z(x))\right) \cdot \binom{1}{z'(x)} = \frac{\partial f}{\partial x}(x, z(x)) + \frac{\partial f}{\partial y}(x, z(x)) \cdot z'(x)$. Dies liefert die Identität. Dieser Situation werden wir beim Satz über implizite Funktionen (Korollar 2.78) begegnen.

Definition 2.67 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine Abbildung, wobei $U \subset \mathbb{R}^n$ offen ist. Die Abbildung f heißt genau dann* stetig differenzierbar *(kurz: C^1) auf U, wenn f differenzierbar auf U ist und die Abbildung $J_f : U \longrightarrow M_{p\times n}(\mathbb{R})$, $x \longmapsto J_f(x)$, stetig ist.*

Beispiele 2.68

1. Eine Abbildung $c : I \longrightarrow \mathbb{R}^p$, wobei $I \subset \mathbb{R}$ ein offenes Intervall ist, ist genau dann C^1 im Sinne von Definition 2.67, wenn sie im üblichen Sinne C^1 ist (d.h., c ist ableitbar und $\dot{c} : I \longrightarrow \mathbb{R}^p$ ist stetig).

2. Jede affin-lineare Abbildung ist C^1: ist f der Form $f(x) = A\cdot x + b$, mit $A \in M_{p\times n}(\mathbb{R})$ und $b \in \mathbb{R}^p$, so ist f differenzierbar auf $\mathbb{R}^n$ und es gilt $d_x f = A$ für alle $x \in \mathbb{R}^n$; insbesondere ist J_f konstant, und somit auch stetig auf $\mathbb{R}^n$.

2.2.2 Der Mittelwertsatz

Satz 2.69 (Mittelwertsatz) *Sei $f : U \longrightarrow \mathbb{R}^p$ eine auf einer (nichtleeren) offenen Teilmenge*[24] *U vom $\mathbb{R}^n$ differenzierbare Abbildung. Seien $x, y \in U$, so, dass $[x, y] \subset U$ gilt. Dann gilt*

$$|f(x) - f(y)| \leq \sup_{z\in[x,y]} (\|J_f(z)\|) \cdot |x - y|.$$

Beweis:[25] O.B.d.A. sei $x \neq y$. Definiere die Abbildung $\hat{f} : [0, 1] \longrightarrow \mathbb{R}^p$, $t \longmapsto f((1-t)x + ty)$.

Behauptung 1: *Die Abbildung $\hat{f}$ ist ableitbar auf $[0, 1]$ mit $\hat{f}'(t) = d_{(1-t)x+ty} f(y - x)$ für alle $t \in [0, 1]$.*

Beweis: Die Funktion $h : [0, 1] \longrightarrow U$, $t \longmapsto (1 - t)x + ty$ ist offensichtlich ableitbar mit $h'(t) = y - x$ für alle $t \in [0, 1]$. Nach der Kettenregel ist daher $f \circ h = \hat{f}$ auch ableitbar (oder differenzierbar, es ist äquivalent) mit $\hat{f}'(t) = d_t\hat{f}(1) = d_{h(t)}f \circ d_t h(1) = d_{h(t)}f(h'(t)) = d_{(1-t)x+ty}f(y - x)$, für alle $t \in [0, 1]$. $\checkmark$

Wir setzen $M := \sup_{z\in[x,y]} (\|J_f(z)\|)$, wobei $\|\cdot\|$ die Matrixnorm (siehe Einschub) ist. Ist

[24] In der Vorlesung wurde angenommen, dass U konvex ist. Dies hat aber keinen Einfluss auf das Ergebnis, da die Voraussetzung $[x, y] \subset U$ für alle $x, y \in U$ dann automatisch erfüllt ist.

[25] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

$M = \infty$, so ist die Ungleichung offensichtlich erfüllt. Sei also $M < \infty$. Für ein beliebiges $\varepsilon > 0$ definieren wir

$$A := \{t \in [0,1],\ |\hat{f}(t) - \hat{f}(0)| \leq (M+\varepsilon) \cdot |x-y| \cdot t + \varepsilon\} \subset [0,1].$$

Wir zeigen, dass $1 \in A$ gilt.
Behauptung 2: *Die Teilmenge A ist abgeschlossen in $[0,1]$, nichtleer und es gilt $\beta := \sup(A) > 0$.*
Beweis: Die Abbildung $g : [0,1] \longrightarrow \mathbb{R}$, $t \longmapsto |\hat{f}(t) - \hat{f}(0)| - (M+\varepsilon) \cdot |x-y| \cdot t - \varepsilon$ ist wegen $\hat{f}$ stetig offenbar selbst stetig und es gilt $A = \{t \in [0,1],\ g(t) \leq 0\} = g^{-1}([0,\infty[)$. Da $[0,\infty[$ abgeschlossen in $\mathbb{R}$ ist, folgt aus Proposition 2.19, dass A abgeschlossen in $[0,1]$ ist. Insbesondere ist $\beta \in A$.
Bemerke nun, dass $0 \in A$ gilt, denn $|\hat{f}(0) - \hat{f}(0)| = 0 \leq \varepsilon$ ist erfüllt. Es gibt sogar ein $\eta > 0$ so, dass $\eta \in A$ gilt, denn: die obige Funktion g ist stetig mit $g(0) = 0$, insbesondere existiert ein $\eta > 0$ so, dass $|g(t)| \leq \varepsilon$ für alle $t \in [0,\eta]$. Dies beweist $A \neq \varnothing$ sowie $\sup(A) \geq \eta > 0$. $\surd$

Behauptung 3: *Es gilt $\beta = 1$.*
Beweis: Angenommen, es gelte $\beta < 1$. Wegen $|\lim\limits_{t\to\beta} \frac{\hat{f}(t)-\hat{f}(\beta)}{t-\beta}| = |\hat{f}'(\beta)|$ und, nach der Behauptung 1,

$$\begin{aligned}
|\hat{f}'(\beta)| &= |d_{(1-t)x+ty} f(y-x)| \\
&= |J_f(\underbrace{(1-t)x+ty}_{\in[x,y]}) \cdot (y-x)| \\
&\leq |J_f((1-t)x+ty)| \cdot |y-x| \\
&\leq \sup_{z\in[x,y]} \|J_f(z)\| \cdot |x-y| \\
&= M \cdot |x-y| \\
&< (M+\varepsilon) \cdot |x-y|
\end{aligned}$$

existiert ein $\delta > 0$ so, dass für alle $t \in]\beta-\delta,\beta[\cup]\beta,\beta+\delta[$ gilt: $|\frac{\hat{f}(t)-\hat{f}(\beta)}{t-\beta}| \leq (M+\varepsilon) \cdot |y-x|$, insbesondere $|\hat{f}(t) - \hat{f}(\beta)| \leq (M+\varepsilon) \cdot |x-y| \cdot (t-\beta)$ für alle $t \in [\beta,\beta+\delta[$ (für $t = \beta$ ist diese Ungleichung offensichtlich erfüllt). Daraus folgt, für alle $t \in [\beta,\beta+\delta[$:

$$\begin{aligned}
|\hat{f}(t) - \hat{f}(0)| &\leq |\hat{f}(t) - \hat{f}(\beta)| + |\hat{f}(\beta) - \hat{f}(0)| \qquad \text{(Dreiecksungleichung)} \\
&\leq (M+\varepsilon) \cdot |x-y| \cdot (t-\beta) + (M+\varepsilon) \cdot |x-y| \cdot \beta + \varepsilon \\
&= (M+\varepsilon) \cdot |x-y| \cdot t + \varepsilon,
\end{aligned}$$

insbesondere $t \in A$. Dies widerspricht aber $\beta = \sup(A)$. Daraus folgt $\beta = 1$. $\surd$

Die Behauptung 3 liefert das Ergebnis, denn $\sup(A) = 1$ bedeutet, dass $|f(y) - f(x)| = |\hat{f}(1) - \hat{f}(0)| \leq (M+\varepsilon) \cdot |x-y| \cdot 1 + \varepsilon$. Dies gilt aber für alle $\varepsilon > 0$. Lässt man ε gegen 0 laufen, so bekommt man $|f(y) - f(x)| \leq M \cdot |x-y|$, was zu beweisen war. $\square$

Bemerkungen 2.70

1. Im Fall $n = 1$, d.h. $U \subset \mathbb{R}$ offen, ist die Ungleichung von Satz 2.69 äquivalent zu:

   $$|f(b) - f(a)| \leq (\sup_{t\in[a,b]} |f'(t)|) \cdot (b-a)$$

   für alle $a < b$ mit $[a,b] \subset U$. Denn: es gilt $d_t f(h) = hf'(t)$ für alle $h \in \mathbb{R}$.

2. Man beachte, dass im Fall $p \geq 2$ folgende Aussage falsch ist: *"Für $n = 1$ und $a, b \in U$ mit $a < b$ und $[a, b] \subset U$ gilt*

$$f(b) - f(a) = (b - a) \cdot f'(c) \quad \text{für ein} \quad c \in]a, b[."$$

   Im Fall $p = 1$ wäre die Aussage natürlich richtig (Mittelwertsatz der Analysis I). Für $p = 2$ kann man z.B. $U = \mathbb{R}$, $f : U \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{\cos(t)}{\sin(t)}$ betrachten. Die Abbildung f ist differenzierbar auf U und es gilt, für $a = 0$ und $b = 2\pi$ einerseits $f(b) - f(a) = \binom{0}{0} \in \mathbb{R}^2$ und andererseits $f'(t) = \binom{-\sin(t)}{\cos(t)} \neq 0$ für alle $t \in \mathbb{R}^2$, insbesondere ist $(b - a)f'(c) = 2\pi \cdot f'(c)$ für alle $c \in \mathbb{R}$ und daher kann $f(b) - f(a) = (b - a)f'(c)$ für kein c gelten.

Korollar 2.71 *Sei $U \subset \mathbb{R}^n$ eine nichtleere zusammenhängende offene Teilmenge und $f : U \longrightarrow \mathbb{R}^p$ eine differenzierbare Abbildung. Gilt $df = 0$ auf U (d.h., $d_x f = 0$ für alle $x \in U$), so ist f konstant.*

Beweis:[26] Der Beweis beruht auf folgenden Behauptungen.
Behauptung 1: *Die Abbildung f ist lokal konstant auf U, d.h., für alle $x \in U$ existiert ein $r > 0$ s.d. $f_{|B_r(x)}$ konstant ist.*
Beweis: Sei $x \in U$ beliebig und $r > 0$ mit $B_r(x) \subset U$. Für jedes $y \in B_r(x)$ ist $[x, y] \subset B_r(x)$ (jeder Ball im $\mathbb{R}^n$ ist konvex) und Satz 2.69 liefert $|f(y) - f(x)| \leq \sup\limits_{z \in [x,y]} \|J_f(z)\| \cdot |y - x|$. Nach Voraussetzung ist $J_f(z) = 0$ für alle $z \in U$, somit gilt $|f(y) - f(x)| = 0$, d.h., $f(y) = f(x)$. Dies zeigt, dass $f_{|B_r(x)}$ konstant gleich $f(x)$ ist. $\checkmark$
Behauptung 2: *Sei $f : X \longrightarrow Y$ eine lokal konstante Abbildung zwischen metrischen Räumen. Ist X zusammenhängend, so ist f konstant.*
Beweis: Für einen beliebigen Punkt $x \in X$ definiere

$$C_x := \{x' \in X \mid f(x') = f(x)\}.$$

Bemerke, dass wegen $x \in C_x$ die Teilmenge C_x nichtleer ist. Wir zeigen, dass C_x offen in X ist. Sei $x' \in C_x$. Da f lokal konstant ist, existiert ein $r > 0$ s.d. $f_{|B_r(x')}$ konstant ist; insbesondere gilt $f(z) = f(x') = f(x)$ für alle $z \in B_r(x')$, d.h., $B_r(x') \subset C_x$. Dies zeigt, dass C_x offen in X ist. Nun definiere die Beziehung $\sim$ auf X durch: für $x, x' \in X$ gilt $x \sim x'$ g.d.w. $f(x) = f(x')$. Diese Beziehung ist offensichtlich eine Äquivalenzrelation auf X. Die Äquivalenzklasse eines Punktes $x \in X$ ist genau C_x. Da X die disjunkte Vereinigung seiner Äquivalenzklassen bzgl. $\sim$ ist, muss jede Äquivalenzklasse C_x als Komplement einer Vereinigung offener Teilmengen auch abgeschlossen sein. Mit X zusammenhängend folgt, dass es nur eine Äquivalenzklasse bzgl. $\sim$ gibt, d.h., $X = C_x$ für ein $x \in X$, d.h., f ist konstant. $\checkmark$
Der Beweis folgt direkt aus den beiden Behauptungen. $\square$

Satz 2.72 (Konvergenz einer Folge differenzierbarer Funktionen) *Sei $U \subset \mathbb{R}^n$ eine nichtleere konvexe offene Teilmenge und $(f_k)_{k \in \mathbb{N}}$ eine Folge von differenzierbaren Abbildungen $U \longrightarrow \mathbb{R}^p$. Man nehme an:*

i) Ein $u \in U$ existiert s.d. die Folge $(f_k(u))_{k \in \mathbb{N}}$ konvergiert im $\mathbb{R}^p$;

[26] Dieser Beweis wurde in der Vorlesung nur teilweise durchgeführt.

ii) Eine Abbildung $g : U \longrightarrow M_{p\times n}(\mathbb{R})$ existiert s.d. die Folge $(J_{f_k})_{k\in\mathbb{N}}$ gleichmäßig auf *U gegen g konvergiert, d.h., $\sup_{x\in U} \|J_{f_k}(x) - g(x)\| \underset{k\to\infty}{\longrightarrow} 0$.*

Dann existiert eine Abbildung $f : U \longrightarrow \mathbb{R}^p$ mit:

a) Für jede beschränkte Teilmenge $Y \subset U$ konvergiert die Folge $(f_{k|_Y})_{k\in\mathbb{N}}$ gleichmäßig gegen $f_{|_Y}$;

b) Die Abbildung f ist differenzierbar auf U mit $J_f(x) = g(x)$ für alle $x \in U$.

Beweis:[27] Bemerke, dass wegen U konvex $[x, y] \subset U$ für alle $x, y \in U$ gilt. Für jedes $x \in U$ und alle $m, q \in \mathbb{N}$ gilt, nach Satz 2.69 (angewendet auf die Abbildung $f_m - f_q$):

$$|(f_m - f_q)(x) - (f_m - f_q)(u)| \le \sup_{z\in[x,u]} \big(\|J_{f_m}(z) - J_{f_q}(z)\|\big) \cdot |x - u|.$$

Sei $\varepsilon > 0$. Da die Folge $(f_k(u))_{k\in\mathbb{N}}$ konvergiert, ist sie eine Cauchy-Folge (Beispiel 2.35.1), insbesondere existiert ein $N \in \mathbb{N}$ mit $|f_m(u) - f_q(u)| \le \varepsilon$ für alle $m, q \ge N$. Wegen der *gleichmäßigen* Konvergenz von $(J_{f_k})_{k\in\mathbb{N}}$ gegen g auf U existiert auch ein $N' \in \mathbb{N}$ mit $\|J_{f_k}(z) - g(z)\| \le \varepsilon$ für alle $k \ge N'$ und *alle* $z \in U$, insbesondere

$$\|J_{f_m}(z) - J_{f_q}(z)\| \le \|J_{f_m}(z) - g(z)\| + \|g(z) - J_{f_q}(z)\| \le 2\varepsilon$$

für alle $m, q \ge N'$ und alle $z \in U$. Daher existiert ein $N'' \in \mathbb{N}$ s.d. $|f_m(u) - f_q(u)| \le \varepsilon$ und $\|J_{f_m}(z) - J_{f_q}(z)\| \le \varepsilon$ für alle $z \in U$ und $m, q \ge N''$, insbesondere gilt $\sup_{z\in[x,u]} \big(\|J_{f_m}(z) - J_{f_q}(z)\|\big) \le \varepsilon$ für alle $m, q \ge N''$. Aus der Dreiecksungleichung folgt, für alle $m, q \ge N''$:

$$\begin{aligned} |f_m(x) - f_q(x)| &\le |f_m(x) - f_q(x) - (f_m(u) - f_q(u))| + |f_m(u) - f_q(u)| \\ &\le \sup_{z\in[x,u]} \big(\|J_{f_m}(z) - J_{f_q}(z)\|\big) \cdot |x - u| + |f_m(u) - f_q(u)| \\ &\le \varepsilon \cdot |x - u| + \varepsilon. \end{aligned} \tag{2.1}$$

Dies zeigt, dass für alle $x \in U$ die Folge $(f_k(x))_{k\in\mathbb{N}}$ eine Cauchy-Folge im $\mathbb{R}^p$ ist. Da $\mathbb{R}^p$ mit der Standardmetrik vollständig ist (Beispiel 2.37.2), konvergiert die Folge $(f_k(x))_{k\in\mathbb{N}}$ gegen ein Element aus $\mathbb{R}^p$, welches wir mit $f(x)$ bezeichnen. Somit bekommen wir eine Abbildung $f : U \longrightarrow \mathbb{R}^p$ s.d. die Folge $(f_k)_{k\in\mathbb{N}}$ *punktweise* gegen f konvergiert, d.h., $f_k(x) \underset{k\to\infty}{\longrightarrow} f(x)$ für alle $x \in U$. Tatsächlich konvergiert $(f_k)_{k\in\mathbb{N}}$ *gleichmäßig* auf jeder beschränkten Teilmenge von U gegen f: ist nämlich $B \subset U$ eine beschränkte Teilmenge, so existiert ein $M \in [0, \infty[$ mit $|b - b'| \le M$ für alle $b, b' \in B$; legen wir ein Element $b \in B$ fest, so folgt insbesondere $|x - u| \le |x - b| + |b - u| \le M + |b - u|$ für alle $x \in B$. Aus (2.1) folgt, für alle $m, q \ge N''$ wie oben und alle $x \in U$:

$$|f_m(x) - f_q(x)| \le \varepsilon \cdot |x - u| + \varepsilon \le \varepsilon \cdot (M + |b - u|) + \varepsilon.$$

Lässt man q gegen ∞ laufen, so bekommt man $|f_m(x) - f(x)| \le \varepsilon \cdot (M + |b - u| + 1)$. Dabei ist zu beachten, dass das obige N'' *unabhängig* von x gewählt worden ist. Daraus folgt $\sup_{x\in B} (|f_m(x) - f(x)|) \underset{m\to\infty}{\longrightarrow} 0$, d.h., die Folge $(f_{m|_B})_{m\in\mathbb{N}}$ konvergiert gleichmäßig gegen $f_{|_B}$.

[27] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Nun ist der (lokal) gleichmäßige Limes einer Folge von stetigen Abbildungen stets stetig. Der Vollständigkeit halber beweisen wir diese Tatsache in der kommenden
Behauptung: *Die Abbildung f ist stetig.*
Beweis: Sei $x \in U$ und betrachte, für ein beliebiges $r > 0$, die beschränkte Umgebung $B := B_r^U(x) \subset U$ von x in U (wir brauchen an dieser Stelle die Tatsache nicht, dass U offen im $\mathbb{R}^n$ ist). Sei $\varepsilon > 0$. Da $(f_{m|_B})_{m\in\mathbb{N}}$ gleichmäßig gegen $f_{|_B}$ konvergiert (siehe oben), existiert ein $N \in \mathbb{N}$ mit $|f_m(y) - f(y)| \leq \varepsilon$ für alle $m \geq N$ und *alle* $y \in B$. Da außerdem f_N als differenzierbare Abbildung stetig ist (siehe Proposition 2.61), ist $f_{N|_B}$ auch stetig, insbesondere existiert ein $r' > 0$ mit $|f_N(y) - f_N(x)| \leq \varepsilon$ für alle $y \in B$ mit $|y - x| \leq r'$. Mit der Dreiecksungleichung folgt

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| \\ &\leq 3\varepsilon, \end{aligned}$$

d.h., $(f_{|_B})^{-1}(B_{3\varepsilon}(f(x))) \supset B_{r'}(x) \cap B$. Proposition 2.18 liefert, dass $f_{|_B}$ stetig in x ist; da B eine offene Umgebung von x in U ist, liefert dies, dass f selbst stetig in x ist (argumentiere mit Folgen). Da $x \in U$ beliebig gewählt werden konnte, folgt, dass f stetig auf U ist. $\checkmark$

Es bleibt, zu zeigen, dass f differenzierbar ist mit g als punktweise Jacobi-Matrix. Sei $x \in U$ und $r > 0$ mit $B_r(x) \subset U$. Für alle $h \in B_r(0)$ und $m \in \mathbb{N}$ gilt

$$\begin{aligned} |f(x+h) - f(x) - g(x)\cdot h| &\leq |f(x+h) - f(x) - (f_m(x+h) - f_m(x))| \\ &\quad + |f_m(x+h) - f_m(x) - J_{f_m}(x)\cdot h| \\ &\quad + |J_{f_m}(x)\cdot h - g(x)\cdot h|. \end{aligned} \tag{2.2}$$

Sei $\varepsilon > 0$. Für den ersten Summanden auf der rechten Seite von (2.2) gilt

$$|f(x+h) - f(x) - (f_m(x+h) - f_m(x))| = \lim_{q\to\infty} |f_q(x+h) - f_q(x) - (f_m(x+h) - f_m(x))|$$

mit, für alle $q \in \mathbb{N}$:

$$\begin{aligned} |f_q(x+h) - f_q(x) - (f_m(x+h) - f_m(x))| &= |(f_q - f_m)(x+h) - (f_q - f_m)(x)| \\ &\overset{(\text{Satz2.69})}{\leq} \sup_{z\in[x,x+h]} \left(\|J_{f_q}(z) - J_{f_m}(z)\|\right)\cdot|h|. \end{aligned}$$

Wie oben finden wir ein $N \in \mathbb{N}$ s.d. $\sup_{z\in U}\left(\|J_{f_q}(z) - J_{f_m}(z)\|\right) \leq \varepsilon$ für alle $m, q \geq N$, insbesondere $|f_q(x+h) - f_q(x) - (f_m(x+h) - f_m(x))| \leq \varepsilon\cdot|h|$ für alle $m, q \geq N$ und $h \in B_r(0)$. Lässt man q gegen ∞ laufen, so bekommt man

$$|f(x+h) - f(x) - (f_m(x+h) - f_m(x))| \leq \varepsilon\cdot|h|, \tag{2.3}$$

für alle $m \geq N$ und $h \in B_r(0)$. Für den dritten Summanden auf der rechten Seite von (2.2) gilt

$$|J_{f_m}(x)\cdot h - g(x)\cdot h| = |(J_{f_m}(x) - g(x))\cdot h| \leq \|J_{f_m}(x) - g(x)\|\cdot|h| \leq \varepsilon\cdot|h|$$

für alle $m \geq N$, wobei die letzte Ungleichung aus $\|J_{f_m}(x) - g(x)\| = \lim_{q\to\infty} \|J_{f_m}(x) - J_{f_q}(x)\| \leq \varepsilon$ folgt. Nun *fixieren* wir ein $m \geq N$ und betrachten den zweiten Summanden auf der rechten Seite von (2.2). Da f_m differenzierbar in x ist, existiert eine Abbildung $\varepsilon_m : B_r(0) \longrightarrow \mathbb{R}^p$ mit $\varepsilon_m(h) \underset{|h|\to 0}{\longrightarrow} 0$ so, dass $f_m(x+h) = f_m(x) + J_{f_m}(x) \cdot h + |h| \cdot \varepsilon_m(h)$ für alle $h \in B_r(0)$ gilt. Wegen $\varepsilon_m(h) \underset{|h|\to 0}{\longrightarrow} 0$ existiert ein $r' \in]0, r[$ s.d. $|\varepsilon_m(h)| \leq \varepsilon$ für alle $h \in B_{r'}(0)$ gilt, insbesondere

$$|f_m(x+h) - f_m(x) - J_{f_m}(x) \cdot h| = ||h| \cdot \varepsilon_m(h)| = |h| \cdot |\varepsilon_m(h)| \leq \varepsilon \cdot |h|$$

für alle $h \in B_{r'}(0)$. Insgesamt folgt für (2.2) und für alle $h \in B_{r'}(0)$:

$$|f(x+h) - f(x) - g(x) \cdot h| \leq 3\varepsilon \cdot |h|,$$

d.h., $\frac{1}{|h|} \cdot |f(x+h) - f(x) - g(x) \cdot h| \leq 3\varepsilon$ für alle $h \in B_{r'}(0) \setminus \{0\}$. Dies beweist $\frac{1}{|h|} \cdot |f(x+h) - f(x) - g(x) \cdot h \underset{|h|\to 0}{\longrightarrow} 0$, d.h., dass f differenzierbar in x ist mit $J_f(x) = g(x)$. Dies gilt für alle $x \in U$. Der Beweis von Satz 2.72 ist damit zu Ende. □

Satz 2.73 (Zusammenhang zwischen Differenzierbarkeit und partieller Differenzierbarkeit) *Sei $f : U \longrightarrow \mathbb{R}^p$ eine auf einer (nichtleeren) offenen Teilmenge $U \subset \mathbb{R}^n$ definierten Abbildung. Dann sind folgende Aussagen äquivalent:*

i) Die Abbildung f ist C^1 auf U.

ii) Alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ existieren und sind stetig auf U.

Beweis:[28]
"$i) \Longrightarrow ii)$": Nach Proposition 2.61 besitzt f partielle Ableitungen in allen Richtungen und es gilt $\frac{\partial f}{\partial x_i}(x) = d_x f(e_i) = J_f(x) \cdot e_i$, für alle $x \in U$ und $1 \leq i \leq n$. Da $x \mapsto J_f(x)$ nach Voraussetzung stetig ist, ist $x \mapsto J_f(x) \cdot e_i$ ebenfalls stetig (nutze $|J_f(x) \cdot e_i - J_f(y) \cdot e_i| \leq \|J_f(x) - J_f(y)\|$), somit ist $\frac{\partial f}{\partial x_i} : U \longrightarrow \mathbb{R}^p$ stetig, für jedes $1 \leq i \leq n$.
"$ii) \Longrightarrow i)$": Sei $x \in U$ beliebig. Angenommen, die Abbildungen $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq n$, seien stetig in x. Sei $r > 0$ mit $B_r(x) \subset U$ und $h \in B_r(0)$ beliebig. Dann gilt

$$\begin{aligned}
f(x+h) - f(x) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x) &= f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) \\
&\quad - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) \\
&= f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, x_2 + h_2, \ldots, x_n + h_n) \\
&\quad - h_1 \frac{\partial f}{\partial x_1}(x_1, \ldots, x_n) \\
&\quad + f(x_1, x_2 + h_2, \ldots, x_n + h_n) - f(x_1, x_2, x_3 + h_3, \ldots, x_n + h_n) \\
&\quad - h_2 \frac{\partial f}{\partial x_2}(x_1, \ldots, x_n) \\
&\quad \vdots \\
&\quad + f(x_1, x_2, \ldots, x_{n-1}, x_n + h_n) - f(x_1, \ldots, x_n) \\
&\quad - h_n \frac{\partial f}{\partial x_n}(x_1, \ldots, x_n).
\end{aligned}$$

[28] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Mit der Dreiecksungleichung folgt

$$|f(x+h) - f(x) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x| \quad \leq$$

$$\sum_{j=1}^{n} |f(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - f(x_1, \ldots, x_j, \ldots, x_n + h_n) - h_j \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)|.$$

Nun ist zu bemerken, dass für jedes $j \in \{1, \ldots, n\}$ die Abbildung

$$p_j : h_j \mapsto f(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - h_j \frac{\partial f}{\partial x_1}(x_1, \ldots, x_j, \ldots, x_n)$$

nach Voraussetzung ableitbar auf einem kleinen Intervall um $0 \in \mathbb{R}$ ist mit Ableitung gegeben durch $p_j'(h_j) = \frac{\partial f}{\partial x_j}(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)$. Nach dem Mittelwertsatz angewendet auf p_j gilt $|p_j(h_j) - p_j(0)| \leq \sup\limits_{z \in [0, h_j]} \left(|p_j'(z)|\right) \cdot |h_j|$, d.h.,

$$\begin{aligned}
&|f(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - f(x_1, \ldots, x_j, \ldots, x_n + h_n) - h_j \frac{\partial f}{\partial x_1}(x_1, \ldots, x_j, \ldots, x_n)| \\
&\qquad \leq \sup_{z \in [0, h_j]} \left(|\frac{\partial f}{\partial x_j}(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)| \right) \cdot |h_j|.
\end{aligned}$$

Wegen $|h_j| \leq |h|$ für alle j folgt

$$\begin{aligned}
&|f(x+h) - f(x) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x)| \\
&\quad \leq \sum_{j=1}^{n} \sup_{z \in [0, h_j]} \left(|\frac{\partial f}{\partial x_j}(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)| \right) \cdot |h_j| \\
&\quad \leq |h| \cdot \sum_{j=1}^{n} \sup_{z \in [0, h_j]} \left(|\frac{\partial f}{\partial x_j}(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n) - \frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)| \right).
\end{aligned}$$

Sei $\varepsilon > 0$. Da jedes $\frac{\partial f}{\partial x_i}$ stetig in x ist, existiert ein $\eta > 0$ s.d. $|\frac{\partial f}{\partial x_i}(y) - \frac{\partial f}{\partial x_i}(x)| \leq \varepsilon$ für alle $y \in B_\eta(x)$. Es folgt, für alle $h \in B_\eta(0) \setminus \{0\}$:

$$\frac{1}{|h|} \cdot |f(x+h) - f(x) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x)| \leq n\varepsilon.$$

Dies beweist $\frac{1}{|h|} \cdot |f(x+h) - f(x) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x)| \underset{|h| \to 0}{\longrightarrow} 0$. Somit ist f differenzierbar mit $d_x f(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x)$ für alle $h \in \mathbb{R}^n$. Dies gilt nun für alle $x \in U$, da die $\frac{\partial f}{\partial x_i}$'s stetig auf ganz U sind. Außerdem liefert dies auch die Stetigkeit von $J_f : U \longrightarrow M_{p \times n}(\mathbb{R})$ (siehe Beispiel 2.21.3). Insgesamt folgt, dass f C^1 auf U ist. □

2.2.3 Der Umkehrsatz und der Satz über implizite Funktionen

Definition 2.74 *Seien U und V (nichtleere) offene Teilmengen des $\mathbb{R}^n$.*

1. *Eine Abbildung $\varphi : U \longrightarrow V$ heißt genau dann* Diffeomorphismus *zwischen U und V, wenn:*

   i) *φ ist differenzierbar auf U,*

   ii) *$\varphi : U \longrightarrow V$ ist bijektiv,*

   iii) *$\varphi^{-1} : V \longrightarrow U$ (Umkehrabbildung von φ) ist differenzierbar auf V.*

2. *Eine Abbildung $\varphi : U \longrightarrow V$ heißt genau dann ein C^1-Diffeomorphismus zwischen U und V, wenn φ ein Diffeomorphismus zwischen U und V ist und φ sowie φ^{-1} C^1-Abbildungen auf U bzw. V sind.*

Proposition 2.75 *Sei $\varphi : U \longrightarrow V$ ein Homöomorphismus ($\varphi : U \longrightarrow V$ ist bijektiv, φ und φ^{-1} sind stetig) zwischen offenen Teilmengen des $\mathbb{R}^n$ und $x \in U$ ein Punkt. Dann gilt:*

i) *Ist φ differenzierbar in $x \in U$, so ist φ^{-1} genau dann differenzierbar in $\varphi(x) = y \in V$, wenn $d_x\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ invertierbar ist. Ggf. gilt*

$$d_y\varphi^{-1} = (d_x\varphi)^{-1} = (d_{\varphi^{-1}(y)}\varphi)^{-1}.$$

ii) *Ist φ C^1 auf U, so ist φ genau dann ein C^1-Diffeomorphismus auf U, wenn für alle $z \in U$ die Abbildung $d_z\varphi$ invertierbar ist.*

Beweis:[29] Ad *i*): Ist φ^{-1} differenzierbar in y, so gilt nach der Kettenregel angewendet auf $\mathrm{Id} = \varphi^{-1} \circ \varphi$:

$$\mathrm{Id} = d_x\mathrm{Id} = d_{\varphi(x)}(\varphi^{-1}) \circ d_x\varphi.$$

Daraus folgt, dass $d_x\varphi$ invertierbar ist mit $(d_x\varphi)^{-1} = d_y(\varphi^{-1})$.
Umgekehrt sei $d_x\varphi$ invertierbar. Sei $r, r' > 0$ mit $B_r(x) \subset U$ und $B_{r'}(y) \subset V$. Da φ^{-1} nach Voraussetzung stetig ist, kann o.B.d.A. angenommen werden, dass $\varphi^{-1}(B_{r'}(y)) \subset B_r(x)$ gilt (sonst verkleinere r'). Da φ differenzierbar in x ist, gilt für alle $h \in B_r(x)$:

$$\varphi(x+h) = \varphi(x) + d_x\varphi(h) + |h| \cdot \varepsilon(h),$$

wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow}$. Sei nun $h' \in B_{r'}(0)$. Dann gilt für $x = \varphi^{-1}(y)$ und $h := \varphi^{-1}(y+h') - \varphi^{-1}(y)$:

$$\begin{aligned}\varphi(\underbrace{\varphi^{-1}(y) + h}_{\varphi^{-1}(y+h')}) &= \varphi(\varphi^{-1}(y)) + d_x\varphi(\varphi^{-1}(y+h') - \varphi^{-1}(y)) \\ &\quad + |\varphi^{-1}(y+h') - \varphi^{-1}(y)| \cdot \varepsilon(\varphi^{-1}(y+h') - \varphi^{-1}(y)),\end{aligned}$$

d.h., $d_x\varphi(\varphi^{-1}(y+h')-\varphi^{-1}(y)) = y+h'-y-|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\cdot\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y))$.
Da nach Voraussetzung $d_x\varphi$ invertierbar ist, folgt

$$\begin{aligned}\varphi^{-1}(y+h') - \varphi^{-1}(y) &= (d_x\varphi)^{-1}(h') \\ &\quad -(d_x\varphi)^{-1}\left(|\varphi^{-1}(y+h') - \varphi^{-1}(y)| \cdot \varepsilon(\varphi^{-1}(y+h') - \varphi^{-1}(y))\right).\end{aligned}$$

[29]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Zu zeigen ist nun $\frac{1}{|h'|}\cdot|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\cdot\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y))|\underset{|h'|\to 0}{\longrightarrow}0$, denn dann folgt, dass φ^{-1} differenzierbar in y ist (mit Differential gleich $(d_x\varphi)^{-1}$). Die letzte Identität liefert aber, für alle $h'\in B_{r'}(0)$:

$$\begin{aligned}|\varphi^{-1}(y+h')-\varphi^{-1}(y)| \;\le\; & |(d_x\varphi)^{-1}(h')| \\ & +|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\cdot|(d_x\varphi)^{-1}(\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y)))|.\end{aligned}$$

Sei $\epsilon>0$ mit $\epsilon<\frac{1}{2}$ gegeben. Wegen $\varphi^{-1}(y+h')-\varphi^{-1}(y)\underset{|h'|\to 0}{\longrightarrow}0$ gilt $\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y))\underset{|h'|\to 0}{\longrightarrow}0$ und somit auch $(d_x\varphi)^{-1}(\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y)))\underset{|h'|\to 0}{\longrightarrow}0$, daher existiert ein $\eta>0$ mit $|(d_x\varphi)^{-1}(\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y)))|\le\epsilon$ für alle $h'\in B_\eta(0)$. Es folgt, für alle h' mit $|h'|\le\min(\eta,r')$:

$$|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\le\|(d_x\varphi)^{-1}\|\cdot|h'|+\epsilon\cdot|\varphi^{-1}(y+h')-\varphi^{-1}(y)|,$$

insbesondere $|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\le\frac{\|(d_x\varphi)^{-1}\|}{1-\epsilon}\cdot|h'|\le 2\|(d_x\varphi)^{-1}\|\cdot|h'|$. Dies zeigt $\frac{1}{|h'|}\cdot|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\le 2\|(d_x\varphi)^{-1}\|$ für alle $h\neq 0$ hinreichend klein. Aus $\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y))\underset{|h'|\to 0}{\longrightarrow}0$ folgt schließlich $\frac{1}{|h'|}\cdot|\varphi^{-1}(y+h')-\varphi^{-1}(y)|\cdot\varepsilon(\varphi^{-1}(y+h')-\varphi^{-1}(y))|\underset{|h'|\to 0}{\longrightarrow}0$, was zu beweisen war.

Ad *ii*): Ist φ ein C^1-Diffeomorphismus auf U, so ist φ^{-1} differenzierbar auf V, insbesondere liefert *i*), dass $d_z\varphi$ für alle $z\in U$ invertierbar ist.

Sei umgekehrt der Homöomorphismus φ eine C^1-Abbildung mit invertierbarem Differential in jedem Punkt. Nach *i*) is dann φ^{-1} ebenfalls differenzierbar in jedem Punkt. Es bleibt lediglich, zu zeigen, dass die Abbildung $J_{\varphi^{-1}}:V\longrightarrow M_{n\times n}(\mathbb{R})$ stetig ist. Nach *i*) gilt aber $J_{\varphi^{-1}}(y)=(J_\varphi(\varphi^{-1}(y)))^{-1}$ für alle $y\in V$. Da J_φ und φ^{-1} stetig sind, ist $y\mapsto J_\varphi(\varphi^{-1}(y))$ stetig.

Behauptung: *Die Abbildung* $\mathrm{Inv}:\mathrm{GL}(n,\mathbb{R})\longrightarrow\mathrm{GL}(n,\mathbb{R})$, $A\mapsto A^{-1}$, *ist stetig.*

Beweis: Zur Erinnerung ist $\mathrm{GL}(n,\mathbb{R}):=\{A\in M_{n\times n}(\mathbb{R})\mid\det(A)\neq 0\}$ die lineare Gruppe, d.h., die Gruppe der invertierbaren reellen $n\times n$-Matrizen. Es ist erstens zu bemerken, dass $\mathrm{GL}(n,\mathbb{R})$ offen in $M_{n\times n}(\mathbb{R})$ ist. Betrachte zunächst die Identitätsmatrix $I_n\in\mathrm{GL}(n,\mathbb{R})$. Für alle $H\in M_{n\times n}(\mathbb{R})$ mit $\|H\|<1$ betrachten wir die Folge $\left(u_k:=\sum_{j=0}^k H^j\right)_{k\in\mathbb{N}}$, wobei $H^0:=I_n$. Wir behaupten, dass diese Folge eine Cauchy-Folge in $M_{n\times n}(\mathbb{R})$ ist. Denn: für alle $p,q\in\mathbb{N}$ mit $p<q$ gilt

$$\|u_p-u_q\|=\|\sum_{j=p+1}^{q}H^j\|=\|H^{p+1}\cdot\sum_{j=0}^{q-p-1}H^j\|\le\|H\|^{p+1}\cdot\sum_{j=0}^{q-p-1}\|H\|^j.$$

Wegen $\|H\|<1$ ist (siehe Analysis I für die Berechnung der Summe einer geometrischen Reihe) $\sum_{j=0}^{q-p-1}\|H\|^j\le\sum_{j=0}^{\infty}\|H\|^j=\frac{1}{1-\|H\|}<\infty$. Daraus folgt, für alle $p<q$ in $\mathbb{N}$:

$$\|u_p-u_q\|\le\frac{\|H\|^{p+1}}{1-\|H\|}.$$

Sei $\epsilon>0$. Wegen $\|H\|^{p+1}\underset{p\to\infty}{\longrightarrow}0$ existiert ein $N\in\mathbb{N}$ mit $\frac{\|H\|^{p+1}}{1-\|H\|}\le\epsilon$ für alle $p\ge N$, insbesondere $\|u_p-u_q\|\le\epsilon$ für alle $p,q\ge N$ (vertausche p und q falls $p\ge q$). Dies zeigt, dass $(u_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $M_{n\times n}(\mathbb{R})$ ist. An dieser Stelle akzeptieren wir, dass $M_{n\times n}(\mathbb{R})$ bzgl. der Metrik $d:(A,B)\mapsto\|A-B\|$ *vollständig* ist (nutze die Vollständigkeit des $\mathbb{R}^n$ und argumentiere mit den Spalten der betrachteten Matrizen). Diese Tatsache impliziert, dass $(u_k)_{k\in\mathbb{N}}$ gegen ein $u \in M_{n\times n}(\mathbb{R})$ konvergiert. Diesen Grenzwert bezeichnen wir mit $u =: \sum_{k=0}^{\infty} H^k$. Nun ist folgende Identität zu bemerken:

$$(I_n - H)\cdot u_k = (I_n - H)\cdot \sum_{j=0}^{k} H^j = \sum_{j=0}^{k} H^j - \sum_{j=0}^{k} H^{j+1} = I - H^{k+1}.$$

Wegen $\|H^{k+1}\| \le \|H\|^{k+1} \xrightarrow[k\to\infty]{} 0$ gilt $H^{k+1} \xrightarrow[k\to\infty]{} 0$. Aus $u_k \xrightarrow[k\to\infty]{} u$ folgt $(I_n - H)\cdot u = I_n$, insbesondere ist $I_n - H$ invertierbar mit Inversem $u = \sum_{k=0}^{\infty} H^k$. Insgesamt folgt, dass jedes $A \in M_{n\times n}(\mathbb{R})$ mit $\|I_n - A\| < 1$ invertierbar ist: die Matrix $H := I_n - A$ hat nämlich Matrixnorm $\|H\| < 1$, daher ist $I_n - H = A$ invertierbar mit Inversem $\sum_{k=0}^{\infty}(I_n - A)^k$. Dies heißt wiederum, dass im metrischen Raum $M_{n\times n}(\mathbb{R})$ die Inklusion $B_1(I_n) \subset \mathrm{GL}(n,\mathbb{R})$ gilt. Ist nun $A \in \mathrm{GL}(n,\mathbb{R})$ beliebig, so ist $A + H = A(I_n + A^{-1}H)$ für alle $H \in M_{n\times n}(\mathbb{R})$. Gilt insbesondere $\|A^{-1}H\| < 1$, so ist nach der vorangegangenen Überlegung $I_n + A^{-1}H$ invertierbar, daher auch $A + H = A(I_n + A^{-1}H)$. Wegen $\|A^{-1}H\| \le \|A^{-1}\|\cdot\|H\|$ können wir sicherstellen, dass $A + H$ invertierbar ist, sobald $\|H\| < \frac{1}{\|A^{-1}\|}$ gilt. Dies beweist die Inklusion $B_{\frac{1}{\|A^{-1}\|}}(A) \subset \mathrm{GL}(n,\mathbb{R})$. Daraus folgt, dass $\mathrm{GL}(n,\mathbb{R})$ offen in $M_{n\times n}(\mathbb{R})$ ist.
Sei nun $A \in \mathrm{GL}(n,\mathbb{R})$ und $H \in M_{n\times n}(\mathbb{R})$ mit $\|H\| < \frac{1}{\|A^{-1}\|}$. Dann ist

$$\begin{aligned}(A+H)^{-1} - A^{-1} &= \big(A(I_n + A^{-1}H)\big)^{-1} - A^{-1}\\ &= (I_n + A^{-1}H)^{-1}\cdot A^{-1} - A^{-1}\\ &= \big((I_n + A^{-1}H)^{-1} - I_n\big)\cdot A^{-1},\end{aligned}$$

mit $(I_n + A^{-1}H)^{-1} = \sum_{k=0}^{\infty}(-1)^k(A^{-1}H)^k$, insbesondere

$$(I_n + A^{-1}H)^{-1} - I_n = \sum_{k=1}^{\infty}(-1)^k(A^{-1}H)^k = -A^{-1}H\cdot\sum_{k=0}^{\infty}(-1)^k(A^{-1}H)^k.$$

Wie oben ist $\|\sum_{k=0}^{\infty}(-1)^k(A^{-1}H)^k\| \le \frac{1}{1-\|A^{-1}H\|} < \infty$, deshalb gilt

$$-A^{-1}H\cdot\sum_{k=0}^{\infty}(-1)^k(A^{-1}H)^k \xrightarrow[H\to 0]{} 0.$$

Dies liefert $(A+H)^{-1} - A^{-1} \xrightarrow[H\to 0]{} 0$, was zu beweisen war. $\surd$
Da $y \mapsto J_\varphi(\varphi^{-1}(y))$ stetig ist, ist nach der Behauptung und Proposition 2.20 die Abbildung $y \mapsto \mathrm{Inv}\circ J_\varphi(\varphi^{-1}(y))$ stetig, d.h., $J_{\varphi^{-1}}$ ist stetig. $\square$

Satz 2.76 (Umkehrsatz) *Sei $f : W \longrightarrow \mathbb{R}^n$ eine auf einer offenen Teilmenge $W \subset \mathbb{R}^n$ definierte C^1-Abbildung. Angenommen, d_uf sei invertierbar in einem Punkt $u \in W$. Dann existieren offene Umgebungen U von u in W und V von $f(u)$ im $\mathbb{R}^n$ so, dass $f_{|U} : U \longrightarrow V$ ein C^1-Diffeomorphismus ist.*

Man sagt, dass dann f *lokal um u* ein C^1-Diffeomorphismus ist.

Beweis:[30] O.B.d.A. seien $u = 0$, $f(u) = 0$ und $d_uf = \mathrm{Id}$ (sonst betrachte die Abbildung $\widehat{f} : -u + U \longrightarrow \mathbb{R}^n$, $x \mapsto (d_uf)^{-1}(f(u+x) - f(u))$ und beobachte, dass die Verschiebung

[30] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

um einen konstanten Vektor und die Verknüpfung mit einer konstanten invertierbaren Matrix C^1-Diffeomorphismen sind). Wir wollen die Gleichung $y = f(x)$ eindeutig lösen können, zumindest für x und y nahe genug an $0 \in \mathbb{R}^n$. Die Idee ist folgende: fasse diese Gleichung als Fixpunktgleichung auf und wende dann den Banach'schen Fixpunktsatz an. Nämlich ist

$$y = f(x) \iff y - f(x) = 0 \iff x + y - f(x) = x.$$

Betrachte nun, für $y \in \mathbb{R}^n$ *a priori* beliebig, die Abbildung $\varphi_y : U \longrightarrow \mathbb{R}^n$, $x \mapsto x - f(x) + y$. Da f C^1 ist, ist φ_y auch C^1 mit $d_x\varphi_y = \mathrm{Id} - d_x f$ für alle $x \in U$. Da nach Voraussetzung $d_0\varphi_y = \mathrm{Id} - \mathrm{Id} = 0$ und $x \mapsto d_x f$ stetig ist, existiert ein $r > 0$ – *unabhängig* von y – mit $\|J_{\varphi_y}(x)\| \leq \frac{1}{2}$ für alle $x \in \overline{B}_r(0)$. Aus dem Mittelwertsatz (Satz 2.69) folgt, für alle $x, x' \in \overline{B}_r(0)$:

$$|\varphi_y(x) - \varphi_y(x')| \leq \sup_{z \in [x,x']} \left(\|J_{\varphi_y}(z)\|\right) \cdot |x - x'| \leq \frac{1}{2} \cdot |x - x'|. \tag{2.4}$$

Diese Ungleichung liefert bereits die Tatsache, dass $\varphi_{y|_{\overline{B}_r(0)}}$ eine Kontraktion ist. Gilt auch $\varphi_y(\overline{B}_r(0)) \subset \overline{B}_r(0)$? Wegen $\varphi_y(0) = 0 - 0 + y = y$ gilt, für alle $x \in \overline{B}_r(0)$:

$$|\varphi_y(x)| \leq |\varphi_y(x) - \varphi_y(0)| + |\varphi_y(0)| \overset{(2.4)}{\leq} \frac{1}{2} \cdot |x| + |y| \leq \frac{r}{2} + |y|.$$

Erfüllt y die Ungleichung $|y| \leq \frac{r}{2}$, so gilt $|\varphi_y(x)| \leq r$ für alle $x \in \overline{B}_r(0)$, d.h., $\varphi_y(x) \in \overline{B}_r(0)$. Für ein $y \in \overline{B}_{\frac{r}{2}}(0)$ ist also $\varphi_{y|_{\overline{B}_r(0)}} : \overline{B}_r(0) \to \overline{B}_r(0)$ eine Kontraktion. Da $\overline{B}_r(0)$ als abgeschlossene Teilmenge des vollständigen metrischen Raumes $\mathbb{R}^n$ vollständig ist (siehe Beispiel 2.37.2 und Proposition 2.38), folgt aus dem Banach'schen Fixpunktsatz (Satz 2.41) die Existenz und Eindeutigkeit eines Fixpunktes für $\varphi_{y|_{\overline{B}_r(0)}}$, d.h., es gibt ein eindeutiges $x \in \overline{B}_r(0)$ mit $\varphi_y(x) = x$. Nach der obigen Bemerkung heißt dies: für jedes $y \in \overline{B}_{\frac{r}{2}}(0) =: V$ existiert ein eindeutiges $x \in \overline{B}_r(0)$ mit $f(x) = y$. Bemerke, dass dann $x \in \overline{B}_r(0) \cap f^{-1}(\overline{B}_{\frac{r}{2}}(0)) =: U$. Wir haben also bewiesen, dass $f_{|_U} : U \longrightarrow V$ *bijektiv* ist. Die obige Ungleichung (2.4) liefert auch, für alle $x, x' \in U$:

$$\begin{aligned} |x - x'| &\leq |x - x' - (f(x) - f(x'))| + |f(x) - f(x')| \\ &= |\varphi_y(x) - \varphi_y(x')| + |f(x) - f(x')| \\ &\overset{(2.4)}{\leq} \frac{1}{2} \cdot |x - x'| + |f(x) - f(x')|, \end{aligned}$$

inbesondere $|x - x'| \leq 2|f(x) - f(x')|$. Daraus folgt $|f^{-1}(y) - f^{-1}(y')| \leq 2|y - y'|$ für alle $y, y' \in V$. Dies liefert direkt die Stetigkeit von $f^{-1} : V \longrightarrow U$. Deshalb ist $f_{|_U} : U \longrightarrow V$ ein Homöomorphismus. Es bleibt, zu zeigen, dass $d_x f$ in jedem $x \in U$ invertierbar ist. Für alle $x \in U$ gilt aber $\|J_{\varphi_y}(x)\| = \|I_n - J_f(x)\| \leq \frac{1}{2} < 1$ nach Konstruktion von r, insbesondere ist $I_n - (I_n - J_f(x)) = J_f(x)$ invertierbar (siehe Beweis von Proposition 2.75). Proposition 2.75 liefert schließlich, dass $f_{|_U} : U \longrightarrow V$ ein C^1-Diffeomorphismus ist.
□

Beispiele 2.77

1. Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \longrightarrow \mathbb{R}$ eine C^1-Funktion. Sei $u \in I$. Ist $f'(u) \neq 0$, so folgt aus dem Umkehrsatz, dass f lokal um u ein C^1-Diffeomorphismus ist.

Man beachte, dass f i.A. kein globaler Diffeomorphismus ist (f muss nicht bijektiv sein).

2. Sei $W := \mathbb{R}^2$ und $f : W \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (e^x \cos(y), e^x \sin(y))$. Dann hat f offenbar erste partielle Ableitungen nach x und y und es gilt, für alle $(x, y) \in \mathbb{R}^2$:

$$\begin{aligned} J_f(x,y) &= \begin{pmatrix} \frac{\partial(e^x\cos(y))}{\partial x} & \frac{\partial(e^x\cos(y))}{\partial y} \\ \frac{\partial(e^x\sin(y))}{\partial x} & \frac{\partial(e^x\sin(y))}{\partial y} \end{pmatrix} = \begin{pmatrix} e^x\cos(y) & -e^x\sin(y) \\ e^x\sin(y) & e^x\cos(y) \end{pmatrix} \\ &= e^x \cdot \begin{pmatrix} \cos(y) & -\sin(y) \\ \sin(y) & \cos(y) \end{pmatrix}. \end{aligned}$$

Da $J_f : W \longrightarrow M_{2\times 2}(\mathbb{R})$ insbesondere stetig ist (alle Koeffizienten der Matrix $J_f(x, y)$ sind stetige Funktionen von (x, y)), folgt aus Satz 2.73, dass f C^1 ist. Außerdem gilt $\det(J_f(x, y)) = e^{2x} \cdot \det \begin{pmatrix} \cos(y) & -\sin(y) \\ \sin(y) & \cos(y) \end{pmatrix} = e^{2x} \cdot (\cos^2(y) + \sin^2(y)) = e^{2x}$. Insbesondere ist $\det(J_f(x, y)) \neq 0$, daher ist $J_f(x, y)$ invertierbar für alle $(x, y) \in U$, d.h., $d_{(x,y)}f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ist invertierbar für alle $(x, y) \in W$. Aus dem Umkehrsatz folgt, dass f ein lokaler C^1-Diffeomorphismus ist. Dabei muss man nicht einmal lokal die Umkehrabbildung von f suchen! Man bemerke, dass in diesem Beispiel f auch nicht bijektiv ist.

3. (Polarkoordinaten) Sei $W :=]0, \infty[\times] - \pi, \pi[$ und

$$\begin{aligned} f : W &\longrightarrow \mathbb{R}^2 \\ (r, \varphi) &\longmapsto \begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi) \end{pmatrix}. \end{aligned}$$

Es gilt: $f(W) = \mathbb{R}^2 \setminus \{(x_0) \in \mathbb{R}^2 \,|\, x \leq 0\}$ und $f : W \longrightarrow f(W)$ ist ein Homöomorphismus (siehe Aufgabe 1 im Blatt 5). Desweiteren hat f erste partielle Ableitungen auf W und es gilt: $J_f(r, \varphi) = \begin{pmatrix} \frac{\partial(r\cos\varphi))}{\partial r} & \frac{\partial(r\cos(\varphi))}{\partial \varphi} \\ \frac{\partial(r\sin(\varphi))}{\partial r} & \frac{\partial(r\sin(\varphi))}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}$ für alle $(r, \varphi) \in W$. Da $(r, \varphi) \longmapsto J_f(r, \varphi)$ stetig auf W ist, ist f C^1 auf W (wiederum nach Satz 2.73). Wegen $\det(J_f(r, \varphi)) = \det \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} = r\cos^2(\varphi) + r\sin^2(\varphi) = r > 0$ ist $J_f(r, \varphi)$, d.h. $d_{(r,\varphi)}f$, invertierbar für jedes $(r, \varphi) \in W$. Aus Proposition 2.75 folgt, dass f ein C^1-Diffeomorphismus ist.

4. (Zylinderkoordinaten) Sei $W :=]0, \infty[\times] - \pi, \pi[\times \mathbb{R}$ und

$$\begin{aligned} f : W &\longrightarrow \mathbb{R}^3 \\ (r, \varphi, z) &\longmapsto \begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi) \\ z \end{pmatrix}. \end{aligned}$$

Wie im letzten Beispiel kann bewiesen werden, dass $f(W) = \mathbb{R}^3 \setminus \{(x, 0, z) \,|\, x \leq 0$ und $z \in \mathbb{R}\}$ gilt und dass $f : W \longrightarrow f(W)$ ein Homöomorphismus ist. Außerdem hat f erste partielle Ableitungen auf W mit

$$\begin{aligned} J_f(r, \varphi, z) &= \begin{pmatrix} \frac{\partial(r\cos\varphi)}{\partial r} & \frac{\partial(r\cos\varphi)}{\partial \varphi} & \frac{\partial(r\cos\varphi)}{\partial z} \\ \frac{\partial(r\sin\varphi)}{\partial r} & \frac{\partial(r\sin\varphi)}{\partial \varphi} & \frac{\partial(r\sin\varphi)}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial z} \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

für alle $(r, \varphi, z) \in W$. Dies zeigt einerseits, dass f C^1 ist (Satz 2.73), andererseits dass $\det(J_f(r, \varphi, z)) = r > 0$ für alle $(r, \varphi, z) \in W$. Proposition 2.75 liefert wiederum, dass f ein C^1-Diffeomorphismus ist.

5. (Kugelkoordinaten) Sei $W :=]0, \infty[\times] - \pi, \pi[\times] - \frac{\pi}{2}, \frac{\pi}{2}[$ und

$$\begin{aligned} f : W &\longrightarrow \mathbb{R}^3 \\ (r, \theta, \varphi) &\longmapsto \begin{pmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ r\sin(\varphi) \end{pmatrix}. \end{aligned}$$

Es kann wiederum bewiesen werden, dass $f(W) = \mathbb{R}^3 \setminus \{(x, 0, z) | x \leq 0 \text{ und } z \in \mathbb{R}\}$ gilt und dass $f : W \longrightarrow f(W)$ ein Homöomorphismus ist. Die Abbildung f hat erste partielle Ableitungen auf W mit

$$\begin{pmatrix} \frac{\partial(r\cos\theta\cos\varphi)}{\partial r} & \frac{\partial(r\cos\theta\cos\varphi)}{\partial\theta} & \frac{\partial(r\cos\theta\cos\varphi)}{\partial\varphi} \\ \frac{\partial(r\sin\theta\cos\varphi)}{\partial r} & \frac{\partial(r\sin\theta\cos\varphi)}{\partial\theta} & \frac{\partial(r\sin\theta\cos\varphi)}{\partial\varphi} \\ \frac{\partial(r\sin\varphi)}{\partial r} & \frac{\partial(r\sin\varphi)}{\partial\theta} & \frac{\partial(r\sin\varphi)}{\partial\varphi} \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi & -r\sin\theta\cos\varphi & -r\cos\theta\sin\varphi \\ \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\varphi & 0 & r\cos\varphi \end{pmatrix},$$

insbesondere ist f C^1 auf W mit

$$\begin{aligned} \det(J_f(r, \theta, \varphi)) &= r^2\cos^2(\theta)\cos(\varphi) + r^2\sin^2(\theta)\cos(\varphi) \\ &= r^2\cos(\varphi) > 0 \quad \text{für alle} \quad r > 0 \quad \text{und} \quad \varphi \in] - \tfrac{\pi}{2}, \tfrac{\pi}{2}[. \end{aligned}$$

Daraus folgt, dass $J_f(r, \theta, \varphi)$, d.h., $d_{(r,\theta,\varphi)}f$, invertierbar für alle $(r, \theta, \varphi) \in W$ ist. Proposition 2.75 liefert, dass f ein C^1-Diffeomorphismus ist.

Korollar 2.78 (Satz über implizite Funktionen) *Sei $W \subset \mathbb{R}^n \times \mathbb{R}^p$ eine nichtleere offene Teilmenge und $f : W \longrightarrow \mathbb{R}^p$, $(x, y) \mapsto f(x, y)$, eine C^1-Abbildung, wobei x im $\mathbb{R}^n$ und y im $\mathbb{R}^p$ laufen. Angenommen, es gibt ein $w = (a, b) \in W$ (wobei $(a, b) \in \mathbb{R}^n \times \mathbb{R}^p$) mit $f(w) = 0$ und so, dass die Matrix $\left(\frac{\partial f_i}{\partial y_j}(w)\right)_{1 \leq i,j \leq p}$* invertierbar *ist. Dann existieren offene Umgebungen U von a im $\mathbb{R}^n$ und V von $(a, b) = w$ in W sowie eine C^1-Abbildung $g : U \longrightarrow \mathbb{R}^p$ mit:*

$$\begin{cases} (x, y) \in V \\ f(x, y) = 0 \end{cases} \Longleftrightarrow \begin{cases} x \in U \\ y = g(x) \end{cases}.$$

Beweis:[31] Betrachte die Abbildung

$$\begin{aligned} \widehat{f} : W &\longrightarrow \mathbb{R}^n \times \mathbb{R}^p \\ (x, y) &\longmapsto (x, f(x, y)). \end{aligned}$$

Da f C^1 ist, ist $\widehat{f}$ auch C^1 mit

$$J_{\widehat{f}}(x, y) = \begin{pmatrix} \frac{\partial x_i}{\partial x_j} & \frac{\partial x_i}{\partial y_j} \\ \frac{\partial f_i}{\partial x_j} & \frac{\partial f_i}{\partial y_j} \end{pmatrix}(x, y) = \begin{pmatrix} I_n & 0 \\ \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{pmatrix},$$

[31] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

wobei wir folgende Bezeichnungen verwenden:

$$\left(\frac{\partial f}{\partial x}(x,g(x))\right) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x,g(x)) & \dots & \frac{\partial f_1}{\partial x_n}(x,g(x)) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(x,g(x)) & \dots & \frac{\partial f_p}{\partial x_n}(x,g(x)) \end{pmatrix}$$

und

$$\left(\frac{\partial f}{\partial y}(x,g(x))\right) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x,g(x)) & \dots & \frac{\partial f_1}{\partial y_p}(x,g(x)) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial y_1}(x,g(x)) & \dots & \frac{\partial f_p}{\partial y_p}(x,g(x)) \end{pmatrix}.$$

Aus dem Umkehrsatz (Satz 2.76) folgt die Existenz einer offene Umgebung V' von (a,b) in W und einer Umgebung $\widehat{V}'$ von $\widehat{f}(a,b) = (a,0)$ in $\mathbb{R}^n \times \mathbb{R}^p$ s.d. $\widehat{f}_{|V'} : V' \longrightarrow \widehat{V}'$ ein C^1-Diffeomorphismus ist. Bemerke nun: es gibt Umgebungen U von a im $\mathbb{R}^n$ und U' von 0 im $\mathbb{R}^p$ mit $U \times U' \subset \widehat{V}'$. Denn: ein $r > 0$ eixstiert mit $B_r(a,0) \subset \widehat{V}'$ (wegen $\widehat{V}'$ offen); wähle z.B. $U := B_{\frac{r}{2}}(a) \subset \mathbb{R}^n$ und $U' := B_{\frac{r}{2}}(0) \subset \mathbb{R}^p$ und überprüfe, dass $U \times U' \subset \widehat{V}'$ wohl erfüllt ist. Setze nun $V := (\widehat{f}_{|V'})^{-1}(U \times U')$ (bemerke, dass V offen in V' und somit in W ist, da $\widehat{f}$ stetig ist). Nach Einschränkung ist $\widehat{f}_{|V} : V \longrightarrow U \times U'$ wieder ein C^1-Diffeomorphismus. Betrachte die Abbildung

$$\begin{aligned} g : U &\longrightarrow \mathbb{R}^p \\ x &\longmapsto \mathrm{pr}_2 \circ (\widehat{f}_{|V})^{-1} \circ \iota_1, \end{aligned}$$

wobei $\mathrm{pr}_2 : \mathbb{R}^n \times \mathbb{R}^p \longrightarrow \mathbb{R}^p$ die kanonische Projektion auf die p letzten Koordinaten und $\iota_1 : U \longrightarrow U \times U'$, $x \mapsto (x,0)$ ist. Inbesondere ist $(\widehat{f}_{|V})^{-1}(x,0) = (x,g(x))$ für alle $x \in U$. Ist nun $(x,y) \in V$ mit $f(x,y) = 0$, so ist $\widehat{f}(x,y) = (x,0)$, insbesondere gilt $x \in U$ und $(x,y) = (\widehat{f}_{|V})^{-1}(\widehat{f}(x,y)) = (\widehat{f}_{|V})^{-1}(x,0) = (x,g(x))$ und daher ist $y = g(x)$; also $x \in U$ und $y = g(x)$. Umgekehrt sei $x \in U$ und $y := g(x)$. Wegen $(x,y) = (x,g(x)) = (\widehat{f}_{|V})^{-1}(x,0)$ gilt $(x,y) \in V$; außerdem ist $\widehat{f}((\widehat{f}_{|V})^{-1}(x,0)) = (x,0) = (x,f(x,y))$, insbesondere gilt $f(x,y) = 0$. Dies schließt den Beweis ab. □

Bemerkungen 2.79

1. Unter den Voraussetzungen von Korollar 2.78 kann die Gleichung $f(x,y) = 0$ *lokal nach y aufgelöst werden*. Dies geht i.A. aber nur lokal, siehe z.B. Beispiel 2.80 unten. Es ist auch zu bemerken, dass die Funktion i.A. nicht *explizit* gegeben ist.

2. Es kann allerdings folgende Formel für g hergeleitet werden: nämlich für alle $x \in U$ gilt (nach Korollar 2.78) $f(x,g(x)) = 0$, somit – da f und g differenzierbar sind – nach der Kettenregel (Proposition 2.63):

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x,g(x)) & \dots & \frac{\partial f_1}{\partial x_n}(x,g(x)) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(x,g(x)) & \dots & \frac{\partial f_p}{\partial x_n}(x,g(x)) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x,g(x)) & \dots & \frac{\partial f_1}{\partial y_p}(x,g(x)) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial y_1}(x,g(x)) & \dots & \frac{\partial f_p}{\partial y_p}(x,g(x)) \end{pmatrix} \cdot J_g(x) = 0.$$

   Mit den obigen Bezeichnungen bekommen wir folgende Formel für g:

$$J_g(x) = -\left(\frac{\partial f}{\partial y}(x,g(x))\right)^{-1} \cdot \left(\frac{\partial f}{\partial x}(x,g(x))\right).$$

   Dies ist eine sogenannte *partielle Differentialgleichung*, welche g erfüllt.

Beispiel 2.80 Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto 2x^2y - \cos(y) + 1$. Die Abbildung f ist offensichtlich C^1 mit $\frac{\partial f}{\partial y}(x, y) = 2x^2 + \sin(y)$. Für $a = 1$ und $b = 0$ ist insbesondere $\frac{\partial f}{\partial y}(a, b) = 2 \neq 0$. Der Satz über implizite Funktionen impliziert, dass auf einer Umgebung von (a, b) die Gleichung $f(x, y) = f(a, b) = 0$ nach y aufgelöst werden kann, d.h., es gibt eine auf einer offenen Umgebung U von a in $\mathbb{R}$ definierte C^1-Funktion g s.d. $f(x, g(x)) = 0$ für alle $x \in U$. Nach der Bemerkung 2.79.2 ist g durch folgende Identität gegeben:

$$g'(x) = -\frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))} = -\frac{4xg(x)}{2x^2 + \sin(g(x))}$$

für alle $x \in U$. Insbesondere löst g eine (nichtlineare) *Differentialgleichung erster Ordnung*, siehe Kapitel 4.

2.2.4 Höhere Ableitungen und Taylor-Formel

Definition 2.81 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine auf einer (nichtleeren) offenen Teilmenge des $\mathbb{R}^n$ definierte Abbildung. Man sagt, dass f* partielle Ableitungen zweiter Ordnung *an einem Punkt $x \in U$ besitzt, wenn f erste partielle Ableitungen in allen Richtungen auf einer Umgebung von x hat und diese erste partielle Ableitungen an x in allen Richtungen besitzen. Ggf. definiert man*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) := \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(x) \qquad \textit{für alle} \quad 1 \leq i, j \leq n.$$

Im Fall $i = j$ wird $\frac{\partial^2 f}{\partial x_i \partial x_i}(x)$ oft mit $\frac{\partial^2 f}{\partial x_i^2}(x)$ bezeichnet.

Beispiel 2.82 Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto 2x^2y - \cos(y)$. Dann hat f erste partielle Ableitungen an allen Punkten $(x, y) \in \mathbb{R}^2$: es gilt $\frac{\partial f}{\partial x}(x, y) = 4xy$ und $\frac{\partial f}{\partial y}(x, y) = 2x^2 + \sin(y)$. Da $(x, y) \longmapsto 4xy$ und $(x, y) \longmapsto 2x^2 + \sin(y)$ offensichtlich erste partielle Ableitungen besitzen auf $\mathbb{R}^2$, hat f partielle Ableitungen zweiter Ordnung an jedem $(x, y) \in \mathbb{R}^2$. Es gelten: $\frac{\partial^2 f}{\partial x^2}(x, y) = 4y$, $\frac{\partial^2 f}{\partial y \partial x}(x, y) = 4x$, $\frac{\partial^2 f}{\partial x \partial y}(x, y) = 4x$ und $\frac{\partial^2 f}{\partial y^2}(x, y) = \cos(y)$. Dabei bemerken wir, dass $\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y)$ gilt. Dies ist aber kein Zufall.

Satz 2.83 (Satz von Schwarz) *Sei $f : U \longrightarrow \mathbb{R}^p$ eine auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ definierte Abbildung. Hat f partielle Ableitungen zweiter Ordnung in einer Umgebung eines Punktes $x \in U$,* welche in x stetig sind, *so gilt, für alle $1 \leq i, j \leq n$:*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x).$$

Beweis:[32] Wir schreiben den Beweis unter der leicht stärkeren Voraussetzung, dass $\frac{\partial^2 f}{\partial x_i \partial x_j}$ stetig *um* x (d.h., auf einer offenen Umgebung von x in U) sind. Seien $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ (sonst ist nichts zu zeigen). Da die anderen Veränderlichen x_k, $k \neq i, j$, keine Rolle spielen, kann o.B.d.A. angenommen werden, dass $n = 2$, d.h. $x_1 = x$ und $x_2 = y$.

[32] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Sei also $f : U \longrightarrow \mathbb{R}^p$, $(x,y) \longmapsto f(x,y)$ wobei $U \subset \mathbb{R}^2$ offen ist. Nach Voraussetzung hat $\frac{\partial f}{\partial y}$ eine partielle Ableitung in x-Richtung an (x,y). Nach Definition gilt:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x \partial y}(x,y) &= \lim_{h\to 0} \frac{\frac{\partial f}{\partial y}(x+h,y) - \frac{\partial f}{\partial y}(x,y)}{h} \\
&= \lim_{h\to 0} \frac{1}{h}\left(\lim_{k\to 0} \frac{f(x+h,y+k)-f(x+h,y)}{k} - \lim_{k\to 0}\frac{f(x,y+k)-f(x,y)}{k}\right) \\
&= \lim_{h\to 0}\left(\lim_{k\to 0}\left(\frac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)}{hk}\right)\right).
\end{aligned}$$

Analog hat (nach Voraussetzung) $\frac{\partial f}{\partial x}$ eine partielle Ableitung in y-Richtung an (x,y) und es gilt:

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \lim_{k\to 0}\left(\lim_{h\to 0}\left(\frac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)}{hk}\right)\right).$$

Es ist daher zu beweisen, dass $\lim\limits_{k\to 0}$ mit $\lim\limits_{h\to 0}$ vertauscht werden kann. Betrachte, für ein festes $h > 0$ die Abbildung $z \longmapsto f(x+h,z)$. Nach Voraussetzung ist diese Abbildung zweimal ableitbar auf einem kleinen offenen Intervall um y und ihre zweite Ableitung ist stetig nahe y. Nach der Taylor-Formel mit integralem Rest gilt

$$f(x+h,y+k) - f(x+h,y) = k\frac{\partial f}{\partial y}(x+h,y) + k^2\int_0^1 (1-t)\frac{\partial^2 f}{\partial y^2}(x+h,y+tk)dt$$

sowie

$$f(x,y+k) - f(x,y) = k\frac{\partial f}{\partial y}(x,y) + k^2\int_0^1 (1-t)\frac{\partial^2 f}{\partial y^2}(x,y+tk)dt.$$

Analog gilt $f(x+h,y+k)-f(x,y+k) = h\cdot\frac{\partial f}{\partial x}(x,y+k) + h^2\int_0^1(1-t)\frac{\partial^2 f}{\partial x^2}(x+th,y+k)dt$ sowie $f(x+h,y)-f(x,y) = h\cdot\frac{\partial f}{\partial x}(x,y) + h^2\int_0^1(1-t)\frac{\partial^2 f}{\partial x^2}(x+th,y)dt$. Desweiteren sind $z \longmapsto \frac{\partial f}{\partial y}(x,y)$ und $z \longmapsto \frac{\partial f}{\partial x}(x,z)$ ableitbar in x bzw. z und es gilt $\frac{\partial f}{\partial y}(x+h,y) = \frac{\partial f}{\partial y}(x,y) + h\frac{\partial^2 f}{\partial x\partial y}(x,y) + h\varepsilon_1(h)$ mit $\varepsilon_1(h) \underset{h\to 0}{\longrightarrow} 0$ sowie $\frac{\partial f}{\partial x}(x,y+k) = \frac{\partial f}{\partial x}(x,y) + k\cdot\frac{\partial^2 f}{\partial y\partial x}(x,y) + k\varepsilon_2(k)$ mit $\varepsilon_2(k) \underset{k\to 0}{\longrightarrow} 0$. Es folgt einerseits

$$\begin{aligned}
f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y) &= k\left(\frac{\partial f}{\partial y}(x,y) + h\frac{\partial^2 f}{\partial x\partial y}(x,y) + h\varepsilon_1(h)\right) \\
+k^2\int_0^1(1-t)\frac{\partial^2 f}{\partial y^2}(x+h,y+tk)dt - k\cdot\frac{\partial f}{\partial y}(x,y) &- k^2\int_0^1(1-t)\frac{\partial^2 f}{\partial y^2}(x,y+tk)dt \\
&= hk\left(\frac{\partial^2 f}{\partial x\partial y}(x,y) + \varepsilon_1(h)\right) \\
+k^2\int_0^1(1-t)&\left(\frac{\partial^2 f}{\partial y^2}(x+h,y+tk) - \frac{\partial^2 f}{\partial y^2}(x,y+tk)\right)dt
\end{aligned}$$

und andererseits

$$\begin{aligned}
f(x+h,y+k)-f(x+h,y)-f(x,y+k)-f(x,y) &= h\left(\frac{\partial f}{\partial x}(x,y)+k\frac{\partial^2 f}{\partial y\partial x}(x,y)+k\varepsilon_2(k)\right)\\
+h^2\int_0^1(1-t)\frac{\partial^2 f}{\partial x^2}(x+th,y+k)dt-h\cdot\frac{\partial f}{\partial x}(x,y) &- h^2\int_0^1(1-t)\frac{\partial^2 f}{\partial x^2}(x+th,y)dt\\
&= hk\left(\frac{\partial^2 f}{\partial y\partial x}(x,y)+\varepsilon_2(k)\right)\\
+h^2\int_0^1(1-t)&\left(\frac{\partial^2 f}{\partial x^2}(x+th,y+tk)-\frac{\partial^2 f}{\partial x^2}(x+th,y)\right)dt.
\end{aligned}$$

Vergleicht man beide Identitäten, so bekommt man

$$\begin{aligned}
hk\left(\frac{\partial^2 f}{\partial x\partial y}(x,y)-\frac{\partial^2 f}{\partial y\partial x}(x,y)+\varepsilon_1(h)-\varepsilon_2(k)\right) &=\\
k^2\int_0^1(1-t)\left(\frac{\partial^2 f}{\partial y^2}(x,y+tk)-\frac{\partial^2 f}{\partial y^2}(x+h,y+tk)\right)dt-h^2\int_0^1(1-t)&\left(\frac{\partial^2 f}{\partial x^2}(x+th,y)-\frac{\partial^2 f}{\partial x^2}(x+th,y+k)\right)dt.
\end{aligned}$$

Ersetzt man nun h durch λh und k durch λk (wobei $\lambda\in]0,1[$), so bekommt man

$$\begin{aligned}
\lambda^2 hk\left(\frac{\partial^2 f}{\partial x\partial y}(x,y)-\frac{\partial^2 f}{\partial y\partial x}(x,y)+\varepsilon_1(\lambda h)-\varepsilon_2(\lambda k)\right) &=\\
\lambda^2k^2\int_0^1(1-t)&\left(\frac{\partial^2 f}{\partial y^2}(x,y+t\lambda k)-\frac{\partial^2 f}{\partial y^2}(x+h,y+t\lambda k)\right)dt\\
-\lambda^2h^2\int_0^1(1-t)&\left(\frac{\partial^2 f}{\partial x^2}(x+t\lambda h,y)-\frac{\partial^2 f}{\partial x^2}(x+t\lambda h,y+k)\right)dt.
\end{aligned}$$

Mit der Transformation $s=t\lambda$ und durch Kürzung der λ^2-Faktoren auf beiden Seiten erhält man

$$\begin{aligned}
hk\left(\tfrac{\partial^2 f}{\partial x\partial y}(x,y)-\tfrac{\partial^2 f}{\partial y\partial x}(x,y)+\varepsilon_1(\lambda h)-\varepsilon_2(\lambda k)\right) &= \tfrac{k^2}{\lambda}\int_0^\lambda(1-\tfrac{s}{\lambda})\tfrac{\partial^2 f}{\partial y^2}(x,y+sk)ds\\
&- \tfrac{h^2}{\lambda}\int_0^\lambda(1-\tfrac{s}{\lambda})\tfrac{\partial^2 f}{\partial x^2}(x+sh,y)ds\\
&= \tfrac{k^2}{\lambda^2}\int_0^\lambda(\lambda-s)\tfrac{\partial^2 f}{\partial y^2}(x,y+sk)ds\\
&- \tfrac{h^2}{\lambda^2}\int_0^\lambda(\lambda-s)\tfrac{\partial^2 f}{\partial x^2}(x+sh,y)ds.
\end{aligned}$$

Sei jetzt $\epsilon>0$. Da $\frac{\partial^2 f}{\partial y^2}$ und $\frac{\partial^2 f}{\partial x^2}$ stetig in (x,y) sind, existieren ein $M\in[0,\infty[$ und ein $\eta>0$ s.d. $\left|\frac{\partial^2 f}{\partial x^2}(x+h,y+k)-\frac{\partial^2 f}{\partial x^2}(x,y)\right|\le\epsilon$ und $\left|\frac{\partial^2 f}{\partial y^2}(x+h,y+k)-\frac{\partial^2 f}{\partial y^2}(x,y)\right|\le\epsilon$ für alle h, k mit $|h| \leq \eta$ und $|k| \leq \eta$. Für solche h und k gilt

$$\begin{aligned}
&|\frac{k^2}{\lambda^2}\int_0^\lambda (\lambda - s)\left(\frac{\partial^2 f}{\partial y^2}(x, y + sk) - \frac{\partial^2 f}{\partial y^2}(x + h, y + sk)\right) ds \\
&\qquad - \frac{h^2}{\lambda^2}\int_0^\lambda (\lambda - s)\left(\frac{\partial^2 f}{\partial x^2}(x + sh, y) - \frac{\partial^2 f}{\partial x^2}(x + sh, y + k)\right) ds| \\
&\leq \frac{k^2}{\lambda^2}\int_0^\lambda (\lambda - s)|\left(\frac{\partial^2 f}{\partial y^2}(x, y + sk) - \frac{\partial^2 f}{\partial y^2}(x + h, y + sk)\right)|ds \\
&\qquad + \frac{h^2}{\lambda^2}\int_0^\lambda (\lambda - s)|\left(\frac{\partial^2 f}{\partial x^2}(x + sh, y) - \frac{\partial^2 f}{\partial x^2}(x + sh, y + k)\right)|ds \\
&\leq \frac{\epsilon(k^2 + h^2)}{\lambda^2}\int_0^\lambda (\lambda - s) ds \\
&\leq \frac{(h^2 + k^2)\epsilon}{2} \quad \text{für alle} \quad \lambda \in]0, 1[.
\end{aligned}$$

Wegen $\varepsilon_1(\lambda h) \underset{\lambda \to 0}{\longrightarrow} 0$ und $\varepsilon_2(\lambda k) \underset{\lambda \to 0}{\longrightarrow} 0$ existiert ein $\lambda > 0$ mit $|\varepsilon_1(\lambda h) - \varepsilon_2(\lambda k)| \leq \epsilon$. Daraus folgt

$$\begin{aligned}
\left|hk\left(\tfrac{\partial^2 f}{\partial x \partial y}(x, y) - \tfrac{\partial^2 f}{\partial y \partial x}(x, y)\right)\right| &\leq \left|hk\left(\tfrac{\partial^2 f}{\partial x \partial y}(x, y) - \tfrac{\partial^2 f}{\partial y \partial x}(x, y) + \varepsilon_1(\lambda h) - \varepsilon_2(\lambda k)\right)\right| \\
&\qquad + |\varepsilon_2(\lambda h) - \varepsilon_2(\lambda k)| \cdot |hk| \\
&\leq \tfrac{(h^2 + k^2)\epsilon}{2} + \epsilon|hk|,
\end{aligned}$$

d.h., $\frac{(h^2+k^2)\epsilon}{2} + |hk| \cdot \left(\epsilon - \left|\frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(x, y)\right|\right) \geq 0$. Die Diskriminante des Polynoms zweiten Grades $\frac{\epsilon x^2}{2} + \left(\epsilon - \left(\left|\frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(x, y)\right|\right)\right) x + \frac{\epsilon}{2}$ ist gleich $(\epsilon - \left|\frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(x, y)\right|)^2 - \frac{4\epsilon^2}{4} = \left|\frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(x, y)\right|^2 - 2\epsilon\left|\frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial y \partial x}(x, y)\right|$. Die Bedingung, dass diese Diskriminante nichtpositiv ist, liefert $\left|\frac{\partial^2 f}{\partial y \partial x}(x, y) - \frac{partial^2 f}{\partial x \partial y}(x, y)\right| \leq 2\epsilon$. Da $\epsilon > 0$ beliebig klein gewählt werden konnte, folgt $\left|\frac{\partial^2 f}{\partial y \partial x}(x, y) - \frac{\partial^2 f}{\partial x \partial y}(x, y)\right| = 0$, was zu beweisen war. $\square$

Bemerkung 2.84 Die Voraussetzung *“f hat partielle Ableitungen zweiter Ordnung um x, welche in x stetig sind”* ist wichtig, wie das folgende Beispiel zeigt. Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto \begin{cases} \frac{xy(x^2-4y^2)}{x^2+y^2} & \text{falls} \quad (x, y) \neq (0, 0) \\ 0 & \text{falls} \quad (x, y) = (0, 0). \end{cases}$ Dann hat f erste und zweite partielle Ableitungen auf $\mathbb{R}^2$. Es gilt $\frac{\partial f}{\partial x}(x, y) = \frac{y(2x^4+7x^2y^2-4y^4)}{(x^2+y^2)^2}$ und $\frac{\partial f}{\partial y}(x, y) = \frac{x(x^4-13x^2y^2-4y^4)}{(x^2+y^2)^2}$, insbesondere $\frac{\partial f}{\partial y}(x, y) \underset{(x,y)\to(0,0)}{\longrightarrow} 0 = \frac{\partial f}{\partial y}(0, 0)$ und $\frac{\partial f}{\partial x}(x, y) \underset{(x,y)\to(0,0)}{\longrightarrow} 0 = \frac{\partial f}{\partial x}(0, 0)$. Wegen $\frac{\partial f}{\partial y}(x, 0) = \frac{x^5}{x^4} = x$ gilt $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$. Wegen $\frac{\partial f}{\partial x}(0, y) = \frac{-4y^5}{y^4} = -4y$ gilt $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -4$, insbesondere $\frac{\partial^2 f}{\partial x \partial y}(0, 0) \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0)$. Tatsächlich kann direkt nachgewiesen werden, dass weder $\frac{\partial^2 f}{\partial x \partial y}$ noch $\frac{\partial^2 f}{\partial y \partial x}$ stetig in $(0, 0)$ ist.

Definition 2.85 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine Abbildung, wobei $U \subset \mathbb{R}^n$ offen ist, und $k \in \mathbb{N}$, $k \geq 1$.*

i) Man sagt, dass f partielle Ableitungen k-ter Ordnung *auf U besitzt, wenn f partielle Ableitungen $(k-1)$-ter Ordnung auf U hat, welche in allen Richtungen erste partielle Ableitungen besitzen.*

ii) Die Abbildung f heißt C^k (oder k-mal stetig differenzierbar*) auf U, wenn sie partielle Ableitungen k-ter Ordnung besitzt, die stetig auf U sind.*

iii) Die Abbildung f heißt C^∞ (oder unendlich oft differenzierbar*) auf U, wenn sie C^k für alle $k \geq 1$ ist.*[33]

Per Konvention heißt f C^0 auf U, wenn f stetig auf U ist.

Bezeichnungen 2.86

1. Sei $k \in \mathbb{N}$ und $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ mit $\alpha_1 + \cdots + \alpha_n = k$. Induktiv wird $\frac{\partial^k f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}$ durch $\frac{\partial^k f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} := \frac{\partial}{\partial x_i} \left(\frac{\partial^{k-1} f}{\partial x_i^{\alpha_i - 1} \ldots \partial x_n^{\alpha_n}} \right)$ definiert, wobei $i := \min\{j \in \{1, \ldots, n\} \mid \alpha_j \neq 0\}$. Per Konvention ist $\frac{\partial^0 f}{\partial x_i^0} = f$ für alle $1 \leq i \leq n$.

2. Für $\alpha \in \mathbb{N}^n$ definiert man $|\alpha| := \sum_{i=1}^{n} \alpha_i \in \mathbb{N}$, $\alpha! := \alpha_1! \cdot \cdots \cdot \alpha_n! \in \mathbb{N}$ und $\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} = \frac{\partial^{\alpha_1 + \cdots + \alpha_n} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}$ für eine Abbildung $f : U \longrightarrow \mathbb{R}^p$. Per Konvention ist also $\frac{\partial^0 f}{\partial x^0} := f$. Man definiert $h^\alpha := h_1^{\alpha_1} \cdot \cdots \cdot h_n^{\alpha_n} \in \mathbb{R}$ für alle $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$.

3. Für $\alpha, \beta \in \mathbb{N}^n$ definiert man $\alpha + \beta := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$ und $\alpha - \beta := (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n)$. Per Definition gilt $\alpha \leq \beta :\Longleftrightarrow \alpha_i \leq \beta_i$ für alle $1 \leq i \leq n$.

4. Das Symbol "$o(r)$" bezeichnet eine Funktion $o : \mathbb{R} \longrightarrow \mathbb{R}^p$, welche $\frac{|o(r)|}{|r|} \underset{|r| \to 0}{\longrightarrow} 0$ erfüllt. Das Symbol "$O(r)$" bezeichnet eine Funktion $O : \mathbb{R} \longrightarrow \mathbb{R}^p$, für die ein $M \in [0, \infty[$ und ein $\delta > 0$ existieren mit $|O(r)| \leq M \cdot |r|$ für alle $|r| \leq \delta$.

Satz 2.87 (Taylor-Formeln) *Sei $f : U \longrightarrow \mathbb{R}^p$ eine Abbildung, wobei $U \subset \mathbb{R}^n$ offen ist. Sei $x \in U$ und $h \in \mathbb{R}^n$ so, dass $[x, x+h] \subset U$. Dann gilt*

i) Ist f C^k auf U, so gilt

$$f(x+h) = \sum_{|\alpha| \leq k} \frac{h^\alpha}{\alpha!} \cdot \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) + o(|h|^k).$$

ii) Ist f C^{k+1} auf U, so gilt

$$f(x+h) = \sum_{|\alpha| \leq k} \frac{h^\alpha}{\alpha!} \cdot \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) + \underbrace{\sum_{|\alpha| = k+1} \frac{h^\alpha}{\alpha!} \int_0^1 \frac{(1-t)^k}{k!} \frac{\partial^{k+1} f}{\partial x^\alpha}(x + th) dt}_{O(|h|^{k+1})}.$$

[33] Mit dem Satz 2.73 ist dies noch äquivalent dazu, dass f partielle Ableitungen beliebiger Ordnung und in beliebigen Richtungen auf U besitzt.

Die Taylor-Formel aus Satz 2.87.i) kann folgendermaßen umgeschrieben werden:

$$f(x+h) = \sum_{|\alpha|\leq k} \frac{h^\alpha}{\alpha!} \cdot \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) + |h|^k \cdot \varepsilon(h),$$

wobei $\varepsilon(h) \underset{|h|\to 0}{\longrightarrow} 0$.

Beweisskizze von Satz 2.87:[34]
i) Induktion über k. Für $k=0$ (d.h., f stetig) ist zu zeigen, dass $f(x+h) - f(x) \underset{|h|\to 0}{\longrightarrow} 0$, was aber aus der Stetigkeit von f in x folgt. Für den Induktionsschritt $k \longrightarrow k+1$ betrachte man die Abbildung $\varphi : h \longmapsto f(x+h) - \sum\limits_{|\alpha|\leq k+1} \frac{h^\alpha}{\alpha!} \cdot \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x)$, welche auf einer (offenen) Umgebung von $0 \in \mathbb{R}^n$ definiert ist; man wende die Taylor-Formel bis zur k-ten Ordnung auf $J_f : U \longrightarrow M_{p\times n}(\mathbb{R})$ zusammen mit dem Satz von Schwarz (Satz 2.83) an und bekomme $|\varphi(h)| = o(|h|^{k+1})$, was zu beweisen ist.
ii) Definiere die Abbildung $\hat{f} : [0,1] \longrightarrow \mathbb{R}^p$, $t \longmapsto f(x+th)$. Dann ist $\hat{f}$ C^{k+1} mit $\hat{f}'(t) = \sum\limits_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x+th)$ und, allgemeiner, $\hat{f}^{(\ell)}(t) = \sum\limits_{|\alpha|=\ell} h^\alpha \frac{\partial f}{\partial x^\alpha}(x+th)$, für alle $1 \leq \ell \leq k+1$. Die Abbildung $v : [0,1] \longrightarrow \mathbb{R}^p$, $t \longmapsto \sum\limits_{\ell=0}^k \frac{(1-t)^\ell}{\ell!} \hat{f}^{(\ell)}(t)$ ist dann C^1 auf $[0,1]$ mit $v(1) = \hat{f}(1)$ und $v(0) = \sum\limits_{\ell=0}^k \frac{\hat{f}^{(\ell)}(0)}{\ell!}$. Die Formel $v(1) - v(0) = \int_0^1 v'(t)dt$ zusammen mit $v'(t) = \frac{(1-t)^k}{k!} \hat{f}^{(k+1)}(t)$ liefert

$$\hat{f}(1) - \sum_{\ell=0}^k \frac{\hat{f}^{(\ell)}(0)}{\ell!} = \int_0^1 \frac{(1-t)^k}{k!} \hat{f}^{(k+1)}(0)dt.$$

Ersetzt man $\hat{f}$ durch $t \mapsto f(x+th)$, so bekommt man die gesuchte Identität. ☐

Proposition 2.88 (Rechenregeln für höhere Ableitungen) *Sei $f : U \longrightarrow \mathbb{R}^p$ eine C^k-Abbildung für ein $k \in \mathbb{N} \cup \{\infty\}$.*

i) Ist $g : U \longrightarrow \mathbb{R}^p$ C^k, so ist $\lambda f + \mu g : U \longrightarrow \mathbb{R}^p$ ebenfalls C^k für alle $\lambda, \mu \in \mathbb{R}$. Ggf. gilt $\frac{\partial^{|\alpha|}(\lambda f + \mu g)}{\partial x^\alpha} = \lambda \frac{\partial^{|\alpha|} f}{\partial x^\alpha} + \mu \frac{\partial^{|\alpha|} g}{\partial x^\alpha}$ für alle $\alpha \in \mathbb{N}^n$ mit $|\alpha| \leq k$.

ii) (Leibniz-Regel) Ist $g : U \longrightarrow \mathbb{R}^p$ C^k, so ist $\langle f, g\rangle : U \longrightarrow \mathbb{R}$, $x \longmapsto \langle f(x), g(x)\rangle$ ebenfalls C^k mit $\frac{\partial^{|\alpha|}\langle f,g\rangle}{\partial x^\alpha} = \sum\limits_{\substack{\beta\in\mathbb{N}^n \\ \beta\leq\alpha}} \binom{\alpha}{\beta} \langle \frac{\partial^{|\beta|} f}{\partial x^\beta}, \frac{\partial^{|\alpha-\beta|} g}{\partial x^\beta} \rangle$, wobei $\binom{\alpha}{\beta} := \prod\limits_{j=1}^n \binom{\alpha_j}{\beta_j} := \prod\limits_{j=1}^n \frac{\alpha_j!}{\alpha_j!(\alpha_j-\beta_j)!} \in \mathbb{N}$.

iii) Ist $g : V \longrightarrow \mathbb{R}^q$ C^k, wobei $V \subset \mathbb{R}^p$ offen ist mit $f(U) \subset V$, so ist $g \circ f$ C^k.

Beweisskizze:[35] Der Beweis erfolgt per Induktion über k bzw. $|\alpha|$. ☐

[34] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[35] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

2.3 Lokale Extrema

2.3.1 Freie Extrema

Definition 2.89 *Sei $f : U \longrightarrow \mathbb{R}$ eine Abbildung, wobei $U \subset \mathbb{R}^n$ offen ist.*

i) Ein Punkt $u \in U$ heißt Minimum *(bzw.* Maximum*) der Funktion f g.d.w. $f(u) \leq f(x)$ (bzw. $f(u) \geq f(x)$) für alle $x \in U$ gilt.*

ii) Ein Punkt $u \in U$ heißt lokales Minimum *(bzw.* lokales Maximum*) der Funktion f g.d.w. eine Umgebung V von u in U so existiert, dass u ein Minimum (bzw. Maximum) von $f_{|V}$ ist, d.h., wenn $f(u) \leq f(x)$ (bzw. $f(x) \geq f(u)$) für alle $x \in V$ gilt.*

iii) Ein Punkt $u \in U$ heißt (freies) Extremum *von f, wenn u ein Minimum oder ein Maximum von f ist. Entsprechend heißt u lokales (freies) Extremum von f, wenn u ein lokales Minimum oder ein lokales Maximum von f ist.*

Definition 2.90

i) Sei $f : U \longrightarrow \mathbb{R}$ eine C^2-Funktion wobei $U \subset \mathbb{R}^n$ offen ist und $x \in U$ ein Punkt. Die Hesse-Matrix[36] *von f in x ist die $n \times n$-Matrix*

$$H_f(x) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \leq i,j \leq n}$$

ii) Eine Matrix $A \in M_{n\times n}(\mathbb{R})$ heißt genau dann positiv-semidefinit*, wenn $\langle A \cdot h, h \rangle \geq 0$ für alle $h \in \mathbb{R}^n$ gilt. Sie heißt genau dann* positiv-definit*, wenn $\langle A \cdot h, h \rangle > 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$ gilt.*

iii) Entsprechend heißt $A \in M_{n\times n}(\mathbb{R})$ genau dann negativ-semidefinit *(bzw.* negativ-definit*), wenn $-A$ positiv-semidefinit (bzw. positiv-definit) ist.*

Beachte, dass nach dem Satz von Schwarz die Hesse-Matrix einer C^2-Funktion symmetrisch ist.

Satz 2.91 (notwendige Bedingungen für die Existenz eines lokalen Extremums) *Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge des $\mathbb{R}^n$ definierte Abbildung.*

i) Ist f C^1 und $u \in U$ ein lokales Extremum von f, so gilt $J_f(u) = 0$, d.h., $(\nabla f)(u) = 0$.

ii) Ist f C^2 und $u \in U$ ein lokales Minimum (bzw. Maximum) von f, so gilt $(\nabla f)(u) = 0$ und die Hesse-Matrix von f in u ist positiv-semidefinit *(bzw. negativ-semidefinit).*

Beweis:[37] *i)* OBdA sei u ein lokales Minimum von f. Sei $h \in \mathbb{R}^n$ beliebig und betrachte die Funktion $\hat{f} : t \longmapsto f(u + th)$. Beachte, dass wegen U offen ein $\eta > 0$ existiert mit $\{u + th \,|\, t \in] - \eta, \eta[\} \subset U$, insbesondere ist $\hat{f}$ auf $] - \eta, \eta[$ wohldefiniert. Da f C^1 ist, ist $\hat{f}$ C^1. Nach Voraussetzung existiert eine Umgebung V von u in U s.d. $f(u) \leq f(x)$ für

[36] Tatsächlich kann die Hesse-Matrix $H_f(x)$ definiert werden, sobald alle partiellen Ableitungen zweiter Ordnung in x existieren.

[37] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

alle $x \in V$. Bis auf Verkleinerung von $\eta > 0$ kann angenommen werden, dass $u + th \in V$ für alle $t \in]-\eta, \eta[$ gilt. Es folgt $\hat{f}(t) = f(u+th) \geq f(u) = \hat{f}(0)$ für alle $t \in]-\eta, \eta[$. Daraus folgt einerseits $\hat{f}'(0) = \lim\limits_{t \to 0^+} \frac{\hat{f}(t)-\hat{f}(0)}{t} \geq 0$, andererseits $\hat{f}'(0) = \lim\limits_{t \to 0^-} \frac{\hat{f}(t)-\hat{f}(0)}{t} \leq 0$, somit ist $\hat{f}'(0) = 0$. Nach der Kettenregel ist aber $\hat{f}'(0) = d_u f(h) = J_f(u) \cdot h$. Daraus folgt $J_f(u) \cdot h = 0$ für alle $h \in \mathbb{R}^n$, d.h., $J_f(u) = 0$.
ii) Sei $u \in U$ ein lokales Minimum von f. Wir wissen nach *i)*, dass bereits $J_f(u) = 0$ gelten muss. Sei nun V eine Umgebung von u in V mit $f(x) \geq f(u)$ für alle $x \in V$. Da f C^2 ist, können wir die Taylor-Formel bis zur zweiten Ordnung (siehe Satz 2.87.*i)*) anwenden: es gilt, für alle $h \in \mathbb{R}^n$ mit $[u, u+h] \subset V$,

$$\begin{aligned} f(u+h) - f(u) &= \underbrace{J_f(u) \cdot h}_{0} + \sum_{i,j=1}^{n} \frac{h_i h_j}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(u) + o(|h|^2) \\ &= (h_1, \ldots, h_n) \cdot \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(u) \right)_{1 \leq i,j \leq n} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + o(|h|^2) \\ &= \langle H_f(u) \cdot h, h \rangle + o(|h|^2). \end{aligned}$$

Schreibe $o(|h|^2) = |h|^2 \varepsilon(h)$ für eine Funktion $\varepsilon : \mathbb{R}^n \longrightarrow \mathbb{R}$ mit $\varepsilon(h) \underset{|h| \to 0}{\longrightarrow} 0$. Ersetzt man h durch th, wobei $t \in]0, 1[$, so bekommt man

$$\begin{aligned} f(u+th) - f(u) &= \langle H_f(u) \cdot th, th \rangle + |th|^2 \cdot \varepsilon(th) \\ &= t^2 (\langle H_f(u) \cdot h, h \rangle + |h|^2 \cdot \varepsilon(th)), \end{aligned}$$

insbesondere ist $\langle H_f(u) \cdot h, h \rangle = \frac{f(u+th)-f(u)}{t^2} - |h|^2 \varepsilon(th)$ für alle $t \in]0, 1[$. Lässt man t gegen 0 laufen, so bekommt man

$$\langle H_f(u) \cdot h, h \rangle = \lim_{t \to 0^+} \frac{f(u+th) - f(u)}{t^2} \geq 0.$$

Wegen U offen gilt dies für alle hinreichend kleine h's und damit (wegen $\langle H_f(u) \cdot \lambda h, \lambda h \rangle = \lambda^2 \langle H_f(u) \cdot h, h \rangle$ für alle $\lambda \in \mathbb{R}$) für alle $h \in \mathbb{R}^n$. □

Satz 2.92 (hinreichende Bedingungen für die Existenz eines lokalen Extremums) *Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge des $\mathbb{R}^n$ definierte C^2-Abbildung und $u \in U$ ein Punkt. Angenommen, es sei $(\nabla f)(u) = 0$ und $H_f(u)$* positiv-definit *(bzw. negativ-definit). Dann ist u ein lokales Minimum (bzw. Maximum) von f.*

Beweis:[38] Der Beweis beruht auf der Taylor-Formel und folgender Beobachtung.
Behauptung: *Es gibt ein $a \in]0, \infty[$ mit $\langle H_f(u) \cdot h, h \rangle \geq a \cdot |h|^2$ für alle $h \in \mathbb{R}^n$.*
Beweis: Betrachte die Funktion $g : S_1(0) \longrightarrow \mathbb{R}$, $h \longmapsto \langle H_f(u) \cdot h, h \rangle$, wobei $S_1(0) := \{x \in \mathbb{R}^n | \, |x| = 1\} \subset \mathbb{R}^n$. Diese Funktion g ist als quadratische Form offenbar stetig. Da $S_1(0)$ kompakt ist (siehe Beispiel 2.29.2), besitzt g nach Korollar 2.32 die Funktion ein Minimum auf $S_1(0)$, d.h., es gibt ein $\overline{h} \in S_1(0)$ mit $g(\overline{h}) \leq g(h)$ für alle $h \in S_1(0)$. Setze $a := g(\overline{h})$ und bemerke, dass wegen $H_f(u)$ positiv-definit $a > 0$ ist. Nun gilt, für alle $h \in \mathbb{R}^n \setminus \{0\}$:

$$\langle H_f(u) \cdot h, h \rangle = |h|^2 \cdot \langle H_f(u) \cdot \frac{h}{|h|}, \frac{h}{|h|} \rangle = |h|^2 g\left(\frac{h}{|h|}\right) \geq |h|^2 a.$$

[38] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

und somit $\langle H_f(u)\cdot h, h\rangle \geq a\cdot |h|^2$ für alle $h\in\mathbb{R}^n$ (trivial für $h=0$). $\checkmark$
Wir wenden die Taylor-Formel bis zur zweiten Ordnung in der Nähe von u an (möglich wegen f C^2): für alle $h\in\mathbb{R}^n$ mit $[u,u+h]\subset U$ gilt

$$f(u+h)-f(u)=\langle H_f(u)\cdot h,h\rangle + o(|h|^2)=\langle H_f(u)\cdot h,h\rangle+|h|^2\varepsilon(h),$$

wobei $\varepsilon(h)\underset{|h|\to 0}{\longrightarrow} 0$. Wegen $\varepsilon(h)\underset{|h|\to 0}{\longrightarrow} 0$ existiert ein $r>0$ mit $B_r(u)\subset U$ und $|\varepsilon(h)|\leq\frac{a}{2}$ für alle $h\in B_r(0)$. Es folgt, für alle $h\in B_r(u)$:

$$\begin{aligned} f(u+h)-f(u) &\geq a\cdot|h|^2+|h|^2\varepsilon(h) \qquad \text{(nach der Behauptung)}\\ &\geq \tfrac{a}{2}\cdot|h|^2, \end{aligned}$$

insbesondere $f(u+h)-f(u)\geq 0$ für alle $h\in B_r(0)$. D.h., u ist ein Minimum von $f_{|B_r(u)}$. $\square$

Beispiele 2.93

1. Sei $f:\mathbb{R}^2\longrightarrow\mathbb{R}$, $(x,y)\longmapsto x^3+x^2+y^2$. Dann ist f C^∞ (insbesondere C^2) mit $(\nabla f)(x,y)=\begin{pmatrix}\frac{\partial f}{\partial x}(x,y)\\ \frac{\partial f}{\partial y}(x,y)\end{pmatrix}=\begin{pmatrix}3x^2+2x\\ 2y\end{pmatrix}$. Insbesondere gilt

$$\begin{aligned}(\nabla f)(x,y)=0 &\iff \left\{\begin{aligned} 3x^2+2x&=0\\ 2y&=0\end{aligned}\right.\\ &\iff \left\{\begin{array}{l} x=0 \text{ oder } x=-\frac{2}{3}\\ y=0\end{array}\right.\\ &\iff (x,y)=(0,0) \text{ oder } (x,y)=(-\frac{2}{3},0).\end{aligned}$$

Außerdem gilt $H_f(x,y)=\begin{pmatrix}\frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x\partial y}(x,y)\\ \frac{\partial^2 f}{\partial y\partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y)\end{pmatrix}=\begin{pmatrix}6x+2 & 0\\ 0 & 2\end{pmatrix}$. Für $u=(0,0)$ ist $(\nabla f)(u)=0$ und $H_f(u)=\begin{pmatrix}2&0\\0&2\end{pmatrix}$ ist positiv-definit:

$$\langle H_f(u)\cdot h,h\rangle=\langle\begin{pmatrix}2&0\\0&2\end{pmatrix}\cdot\begin{pmatrix}h_1\\h_2\end{pmatrix},\begin{pmatrix}h_1\\h_2\end{pmatrix}\rangle=\langle\begin{pmatrix}2h_1\\2h_2\end{pmatrix},\begin{pmatrix}h_1\\h_2\end{pmatrix}\rangle=2h_1^2+2h_2^2\geq 0$$

mit $\langle H_f(u)\cdot h,h\rangle=0\iff h_1=0=h_2$, d.h., $\langle H_f(u)\cdot h,h\rangle=0\iff h=0$. Aus Satz 2.92 folgt, dass $u=(0,0)$ ein lokales Minimum von f ist. Beachte aber, dass $u=(0,0)$ kein globales Minimum von f ist! Es gilt nämlich $f(\mathbb{R}^2)=\mathbb{R}$.

2. Für dieselbe Funktion f betrachten wir nun den Punkt $u:=(-\frac{2}{3},0)$. Nach 1. gilt $(\nabla f)(u)=0$ und $H_f(u)=\begin{pmatrix}-2&0\\0&2\end{pmatrix}$. Insbesondere ist $H_f(u)$ nicht mehr positiv-definit – und auch nicht negativ-definit: gilt z.B. $\langle H_f(u)\cdot e_1,e_1\rangle=-2<0$ und $\langle H_f(u)\cdot e_2,e_2\rangle=2>0$. Tatsächlich hat die Funktion $x\longmapsto f(x,0)=x^3+x^2$ ein lokales Maximum in $x=-\frac{2}{3}$, hingegen hat die Funktion $y\longmapsto f(-\frac{2}{3},y)=\frac{4}{27}+y^2$ ein lokales Minimum in $y=0$. Der Punkt $(-\frac{2}{3},0)$ ist daher weder ein lokales Minimum noch ein lokales Maximum von f. Der Punkt $(-\frac{2}{3},0)$ ist ein sogenannter *Sattelpunkt* von f.

2.3.2 Extrema unter Nebenbedingungen

Definition 2.94 *Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge des $\mathbb{R}^n$ definierte Abbildung. Für ein $k \in \mathbb{N} \setminus \{0\}$ seien Funktionen $g_1, \ldots, g_k$, $g_j : U \longrightarrow \mathbb{R}$, definiert. Man setze $M := \{x \in U | g_j(x) = 0$ für alle $1 \leq j \leq k\}$. Man nennt einen Punkt $u \in M$ lokales Minimum von f* unter den Nebenbedingungen $g_1, \ldots, g_k$, *wenn eine Umgebung V von u in U so existiert, dass u ein Minimum von $f_{|V \cap M}$ ist, d.h., wenn $f(u) \leq f(x)$ für alle $x \in V \cap M$ gilt. Entsprechend wird ein lokales Minimum (bzw. Extremum) unter den Nebenbedingungen $g_1, \ldots, g_k$ definiert. Falls $V = U$ gewählt werden kann, spricht man von globalem Minimum (bzw. Maximum, Extremum) von f unter den Nebenbedingungen $g_1, \ldots, g_k$.*

Beispiel 2.95 Für eine symmetrische Matrix $A \in M_{n\times n}(\mathbb{R})$ sei $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ die durch $f(x) := \frac{\langle A\cdot x,x\rangle}{2}$ definierte Funktion. Sei $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x \longmapsto \frac{|x|^2-1}{2}$. Dann ist $M := \{x \in \mathbb{R}^n, g(x) = 0\} = \{x \in \mathbb{R}^n, |x| = 1\} = S_1(0)$ die 1-Sphäre um den Punkt $0 \in \mathbb{R}^n$. Da f auf $\mathbb{R}^n$ stetig ist (f ist sogar C^∞), ist $f_{|M}$ auch stetig; weil M kompakt ist (Beispiel 2.29.2) besitzt $f_{|M}$ (mindestens) ein Minimum und ein Maximum auf M. Daraus folgt, dass f mindestens zwei Extrema unter der Nebenbedingung g hat.

Satz 2.96 (Lagrange-Multiplikatoren) *Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge des $\mathbb{R}^n$ definierte C^1-Abbildung und $g_1, \ldots, g_k : U \longrightarrow \mathbb{R}$ C^1-Funktionen, wobei $k < n$. Sei $M := \{x \in U \,|\, g_j(x) = 0$ für alle $1 \leq j \leq k\} \subset U$. Man nehme an, dass für jedes $x \in M$ die Vektoren $\nabla g_1(x), \ldots, \nabla g_k(x)$* linear unabhängig *seien. Ist $u \in M$ ein lokales Extremum von f unter den Nebenbedingungen $g_1, \ldots, g_k$, so existieren eindeutige $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit*

$$(\nabla f)(u) = \sum_{j=1}^{k} \lambda_j \cdot (\nabla g_j)(u).$$

Die Koeffizienten λ_j heißen dann die Lagrange-Multiplikatoren *von f an der Stelle u.*

Beweis:[39] Sei $c :]-\eta, \eta[\longrightarrow \mathbb{R}^n$ eine C^1 parametrisierte Kurve mit $c(t) \in M$ für alle $t \in]-\eta, \eta[$ (wobei $\eta > 0$ beliebig klein gewählt werden kann) und $c(0) = u$. Dann gilt nach Voraussetzung $f \circ c(t) \geq f \circ c(0)$ für alle $t \in]-\eta', \eta'[$ für ein $\eta' \in]0, \eta[$. Mit der Kettenregel folgt $(f \circ c)'(0) = 0$, d.h., $\langle (\nabla f)(c(0)), \dot{c}(0)\rangle = 0$, d.h., $\langle \nabla f(u), \dot{c}(0)\rangle = 0$. Andererseits gilt wegen $c(t) \in M$ auch $g_j \circ c(t) = 0$ für alle $t \in]-\eta, \eta[$ und $1 \leq j \leq k$; wieder mit der Kettenregel bekommen wir $\langle \nabla g_j(u), \dot{c}(0)\rangle = 0$ für alle $1 \leq j \leq k$. Dies zeigt, dass $\dot{c}(0)$ orthogonal auf die Vektoren $\nabla g_1(u), \ldots, \nabla g_k(u)$ steht, insbesondere auf jede Linearkombination davon. Es gilt also $\dot{c}(0) \in \mathrm{Span}\{\nabla g_j(u), 1 \leq j \leq k\}^\perp := \{X \in \mathbb{R}^n, \langle X, Z\rangle = 0$ für alle $Z \in \mathrm{Span}\{\nabla g_j(u), 1 \leq j \leq k\}\}$. Gelingt es uns, zu zeigen, dass jeder Vektor $X \in \mathrm{Span}\{\nabla g_j(u),\, 1 \leq j \leq k\}^\perp$ der Form $X = \dot{c}(0)$ ist für eine C^1 parametrisierte Kurve $c :]-\eta, \eta[\longrightarrow M$ mit $c(0) = u$ (und $\eta > 0$ hinreichend klein), so folgt aus der ersten Beobachtung $\langle \nabla f(u), X\rangle = 0$ für alle $X \in \mathrm{Span}\{\nabla g_j(u),\, 1 \leq j \leq k\}^\perp$, d.h., $\nabla f(u) \in (\mathrm{Span}\{\nabla g_j(u), 1 \leq j \leq k\}^\perp)^\perp = \mathrm{Span}\{\nabla g_j(u),\, 1 \leq j \leq k\}$. Dies impliziert die Existenz und Eindeutigkeit der gesuchten Koeffizienten $\lambda_1, \ldots, \lambda_k$.
Behauptung: *Sei $X \in \mathrm{Span}\{\nabla g_j(u), 1 \leq j \leq k\}^\perp$. Dann existieren ein $\eta > 0$ und eine C^1-Abbildung $c :]-\eta, \eta[\longrightarrow U$ mit $c(t) \in M$ für alle $t \in]-\eta, \eta[$, $c(0) = u$ und $\dot{c}(0) = X$.*

[39] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Beweis: Nach Voraussetzung sind die Vektoren $\nabla g_1(u), \ldots, \nabla g_k(u)$ aus $\mathbb{R}^n$ linear unabhängig. Nach Linearalgebra I können diese Vektoren daher zu einer Basis

$$(\nabla g_1(u), \ldots, \nabla g_k(u), V_{k+1}, \ldots, V_n)$$

von $\mathbb{R}^n$ ergänzt werden. Betrachte die Abbildung $\hat{g} : U \longrightarrow \mathbb{R}^n$, $x \longmapsto \sum_{j=1}^{k} g_j(x)e_j + \sum_{j=k+1}^{n} x_j V_j$, wobei $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Diese Abbildung ist wegen g_j C^1 für alle j wohl C^1 und die Matrix der linearen Abbildung $d_u\hat{g}$ in den Basen $(e_1, \ldots, e_n)$ und $(e_1, \ldots, e_k, V_{k+1}, \ldots, V_n)$ von $\mathbb{R}^n$ ist gegeben durch

$$\widehat{J_{\hat{g}}}(u) = \begin{pmatrix} \frac{\partial g_1(u)}{\partial x_1}(u) & & \cdots & & \frac{\partial g_1}{\partial x_n}(u) \\ \vdots & & & & \vdots \\ \frac{\partial g_k(u)}{\partial x_1}(u) & & \cdots & & \frac{\partial g_k}{\partial x_n}(u) \\ 0 \;\ldots & 0 & 1 & 0 & \\ \vdots & \vdots & & \ddots & \\ 0 \;\ldots & 0 & 0 & 1 & \end{pmatrix} = \left(\begin{array}{ccc} & \nabla g_1(u) & \\ \hline & \vdots & \\ \hline & \nabla g_k(u) & \\ \hline 0 & 0 & I_{n-k} \end{array}\right)$$

Wegen $\mathrm{rg}(\{\nabla g_1(u), \ldots, \nabla g_k(u)\}) = k$ hat die Matrix $\widehat{J_{\hat{g}}}(u)$ vollen Rang n, d.h., sie ist invertierbar. Aus dem Umkehrsatz (Satz 2.76) folgt die Existenz von offenen Umgebungen V von u in U und W von $\hat{g}(u)$ im $\mathbb{R}^n$ s.d. $\hat{g}_{|V} : V \longrightarrow W$ ein C^1-Diffeomorphismus ist. Wir bezeichnen diesen Diffeomorphismus mit φ und bemerken, dass nach Definition von $\hat{g}$ die Abbildung $g \circ \varphi^{-1} : W \longrightarrow \mathbb{R}$ durch $\sum_{i=1}^{k} y_i e_i + \sum_{i=k+1}^{n} y_i V_i \longmapsto y_j$ gegeben ist, für alle $1 \leq j \leq k$.
Sei nun $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ mit $X \in \mathrm{Span}\{\nabla g_j(u), 1 \leq j \leq k\}^{\perp}$, d.h., $\langle X, \nabla g_j(u)\rangle = 0$ für alle $1 \leq j \leq k$. Sei $Y := J_\varphi(u) \cdot X \in \mathbb{R}^n$ und betrachte die Kurve $c : t \longmapsto \varphi^{-1}(\varphi(u) + tY)$ für $|t|$ klein genug s.d. $\varphi(u) + tY \in W$ gilt (es gibt ein $\eta > 0$ mit $\varphi(u) + tY \in W$ für alle $t \in]-\eta, \eta[$, da W offen im $\mathbb{R}^n$ ist). Dann ist c eine C^1-Abbildung mit $c(0) = \varphi^{-1}(\varphi(u)) = u$, $\dot{c}(0) = J_{\varphi^{-1}}(\varphi(u)) \cdot Y = J_{\varphi^{-1}}(\varphi(u)) \cdot J_\varphi(u) \cdot X = X$ (wende die Kettenregel an). Außerdem gilt, nach Definition von φ, $Y = J_\varphi(u) \cdot X = \sum_{j=1}^{k} \underbrace{\langle \nabla g_j(u), X\rangle}_{0} e_j + \sum_{j=k+1}^{n} X_j V_j = \sum_{j=k+1}^{n} X_j V_j$.
Daraus folgt $g_j \circ c(t) = g_j \circ \varphi^{-1}(\varphi(u) + tY) = (\ varphi(u) + tY)_j = 0$ wegen $\varphi(u)_j = g_j \circ \varphi^{-1}(\varphi(u)) = g_j(u) = 0$ nach Voraussetzung und $Y_j = 0$ nach der obigen Rechnung. Wir haben damit gezeigt, dass $c(t) \in M$ für alle $t \in]-\eta, \eta[$ mit $c(0) = u$ und $\dot{c}(0) = X$, was zu beweisen war. □

Beispiel 2.97 Sei wie im Beispiel 2.95 eine symmetrische Matrix $A \in M_{n\times n}(\mathbb{R})$ gegeben und $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ durch $f(x) := \frac{\langle Ax, x\rangle}{2}$ definiert. Sei $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $g(x) := \frac{|x|^2-1}{2}$ und $M := g^{-1}(\{0\})$. Nach dem Beispiel 2.95 besitzt f mindestens zwei Extrema unter der Nebenbedingung g. Nach Satz 2.96 – anwendbar wegen f und g C^1 – muss dann $(\nabla f)(u) = \lambda(\nabla g)(u)$ (für ein $\lambda \in \mathbb{R}$) bei jedem lokalen Extremum u von f unter der Nebenbedingung g gelten. Es gilt aber $(\nabla f)(u) = Au$ und $(\nabla g)(u) = u$. Beachte insbesondere, dass $\nabla g(u) \neq 0$ für alle $u \in M$. Nun erfüllt ein $u \in M$ genau dann die Gleichung $Au = \lambda u$ für ein $\lambda \in \mathbb{R}$, wenn u ein *Eigenvektor* der Matrix A ist; ggf. ist λ der dazu gehörige *Eigenwert*.

Da A aber symmetrisch ist, ist A in einer Orthonormalbasis des $\mathbb{R}^n$ diagonalisierbar, d.h., es gibt Vektoren $v_1, \ldots, v_n$ im $\mathbb{R}^n$ und reelle Zahlen $\lambda_1, \ldots, \lambda_n$ mit $Av_i = \lambda_i v_i$ für alle $1 \leq i \leq n$ und $\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$. Es gilt nun, für alle $x = \sum_{i=1}^n \alpha_i v_i \in \mathbb{R}^n$: $Ax = \sum_{i=1}^n \alpha_i A \cdot v_i = \sum_{i=1}^n \alpha_i \lambda_i v_i$ und daher

$$\langle Ax, x \rangle = \sum_{i,j=1}^n \alpha_i \lambda_i \alpha_j \langle v_i, v_j \rangle = \sum_{i=1}^n \lambda_i \alpha_i^2.$$

Bis auf Umnummerierung kann $\lambda_1 \leq \cdots \leq \lambda_n$ angenommen werden. Dann gilt $\lambda_1 \cdot \sum_{i=1}^n \alpha_i^2 \leq \sum_{i=1}^n \lambda_i \alpha_i^2 \leq \lambda_n \cdot \sum_{i=1}^n \alpha_i^2$, insbesondere ist, für jedes $x = \sum_{i=1}^n \alpha_i v_i$ mit $\sum_{i=1}^n \alpha_i^2 = 1$ (d.h., $g(x) = 0$):

$$\frac{\lambda_1}{2} = \frac{\lambda_1}{2} \sum_{i=1}^n \alpha_i^2 \leq f(x) \leq \frac{\lambda_n}{2} \sum_{i=1}^n \alpha_i^2 = \frac{\lambda_n}{2}.$$

Bemerke, dass diese Ungleichungen u.a. folgendes implizieren: eine symmetrische Matrix A ist genau dann positiv-semidefinit (bzw. positiv-definit), wenn alle Eigenwerte von A nichtnegativ (bzw. positiv) sind. Es folgt auch, dass die Punkte v_1 und $-v_1$ (globale) Minima der Funktion f unter der Nebenbedingung g und die Punkte v_n und $-v_n$ (globale) Maxima der Funktion f unter der Nebenbedingung g sind. Für ein $i \in \{1, \ldots, n\}$ mit $\lambda_1 < \lambda_i < \lambda_n$ ist aber v_i kein lokales Extremum von f unter der Nebenbedingung g, obwohl die Gleichung $(\nabla f)(v_i) = \lambda_i (\nabla g)(v_i)$ erfüllt ist.

Zum Schluss behandeln wir die sogenannten *Euler-Lagrange-Gleichungen*, die in manchen physikalischen Problemen auftreten. Sei $I := [a, b] \subset \mathbb{R}$ ein Intervall und, für ein $n \in \mathbb{N}$, $n \geq 1$

$$\begin{array}{rcl} L : I \times \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (t, y, p) & \longmapsto & L(t, y, p) \end{array}$$

eine C^2-Abbildung. In der Literatur wird L häufig *Lagrange-Funktional* oder *Lagrange-Funktion* genannt. Für Konstanten $\alpha, \beta \in \mathbb{R}^n$ betrachte man

$$M := \{\varphi \in C^2([a, b], \mathbb{R}^n) \mid \varphi(a) = \alpha \text{ und } \varphi(b) = \beta\}.$$

Man beachte, dass M keine Teilmenge eines $\mathbb{R}^k$ ist! Tatsächlich ist M ein sogenannter unendlich-dimensionaler affiner Raum. Man wolle jetzt die Abbildung

$$\begin{array}{rcl} S : M & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & \int_a^b L(t, \varphi(t), \dot{\varphi}(t)) dt \end{array}$$

auf (lokale) Extrema untersuchen. Zuerst bemerkt man, dass S wohldefiniert ist: die Abbildungen $t \longmapsto \varphi(t)$ und $t \longmapsto \dot{\varphi}(t)$ sind nach Voraussetzung stetig, insbesondere ist $t \longmapsto L(t, \varphi(t), \dot{\varphi}(t))$ auch stetig und daher Riemann-integrierbar auf $[a, b]$.

Satz 2.98 (Euler-Lagrange-Gleichungen) *Ist $\varphi \in M$ ein Minimum von S (d.h., gilt $S(\varphi) \leq S(\psi)$ für alle $\psi \in M$), so gilt, für alle $j \in \{1, \ldots, n\}$:*

$$\frac{\partial L}{\partial y_j}(t, \varphi(t), \dot{\varphi}(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial p_j}(t, \varphi(t), \dot{\varphi}(t))\right) = 0$$

für alle $t \in [a, b]$.

Beweis:[40] Sei $M_0 := \{h \in C^2([a,b],\mathbb{R}^n) \mid h(a) = h(b) = 0\}$. Dann ist M_0 ein (unendlich-dimensionaler) reeller Vektorraum und es gilt: $\psi \in M \iff \psi - \varphi \in M_0$. Sei nun $h = (h_1, \dots, h_n) : [a,b] \longrightarrow \mathbb{R}^n$ eine C^2-Abbildung, wobei $h_1, \dots, h_n$ C^2-Funktionen sind $[a,b] \longrightarrow \mathbb{R}$ mit $h_j(a) = h_j(b) = 0$ für alle $1 \le j \le n$. Insbesondere ist $h \in M_0$. Man betrachte die Funktion

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ r &\longmapsto S(\varphi + rh), \end{aligned}$$

d.h., $f(r) := \int_a^b L(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t) dt$. Da L eine C^2-Abbildung ist, ist, für jedes $t \in [a,b]$, die Funktion $r \longmapsto L(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t))$ ebenfalls C^2 mit Ableitung gleich

$$\frac{\partial L}{\partial t}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t)) \cdot \underbrace{\frac{\partial t}{\partial r}}_{0} + \sum_{j=1}^n \frac{\partial L}{\partial y_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + rh'(t)) \cdot h_j(t)$$

$+\sum_{j=1}^n \frac{\partial L}{\partial p_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t)) \cdot h_j'(t)$. Da diese Funktion stetig von (t,r) abhängt, ist nach Kapitel 3 die Funktion f ableitbar mit Ableitung

$$f'(r) = \int_a^b \sum_{j=1}^n \frac{\partial L}{\partial y_j}(t, \varphi(t){+}rh(t), \dot\varphi(t){+}r\dot h(t)){\cdot}h_j(t){+}\frac{\partial L}{\partial p_j}(t, \varphi(t){+}rh(t), \dot\varphi(t){+}r\dot h(t))h_j'(t)dt.$$

Eine partielle Integration bei den letzten Summanden (welche wegen L C^2 auch möglich ist) liefert

$$\begin{aligned} \int_a^b \frac{\partial L}{\partial p_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t))h_j'(t)dt &= \left[\frac{\partial L}{\partial p_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t))h_j(t)\right]_a^b \\ &\quad - \int_a^b \frac{d}{dt}\left(\frac{\partial L}{\partial p_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t))\right) h_j(t)dt \end{aligned}$$

für alle $1 \le j \le n$. Es folgt

$$f'(r) = \sum_{j=1}^n \int_a^b \left(\frac{\partial L}{\partial y_j}(t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t)) - \frac{d}{dt}(\frac{\partial L}{\partial p_j}((t, \varphi(t) + rh(t), \dot\varphi(t) + r\dot h(t))))\right) h_j(t)dt.$$

Nun gilt wegen der Voraussetzung $S(\varphi) \le S(\psi)$ für alle $\psi \in M$:

$$f(r) = S(\underbrace{\varphi + rh}_{\in M}) \ge S(\varphi) = f(0)$$

für alle $r \in \mathbb{R}$. Da f ableitbar ist, folgt aus Satz 2.91:

$$f'(0) = 0,$$

d.h., $\sum_{j=1}^n \int_a^b \underbrace{\left(\frac{\partial L}{\partial y_j}(t, \varphi(t), \dot\varphi(t)) - \frac{d}{dt}(\frac{\partial L}{\partial p_j}(t, \varphi(t), \dot\varphi(t)))\right)}_{q_j(t)} \cdot h_j(t)dt = 0.$

Dies muss für alle $h \in M_0$ gelten, also alle $h_j \in C^2([a,b],\mathbb{R})$ mit $h_j(a) = h_j(b) = 0$.

[40] Dieser Beweis wurde in der Vorlesung nur teilweise durchgeführt.

Gilt nun $q_j \neq 0$ für ein $j \in \{1, \ldots, n\}$, so existiert ein $t_0 \in [a,b]$ mit $q_j(t_0) \neq 0$. Da die Funktion q_j stetig ist, können wir oBdA annehmen, dass $t_0 \in]a,b[$ gilt und dass ein $\varepsilon > 0$ und ein $\eta > 0$ so existieren, dass entweder $q_j(t) \geq \varepsilon$ mit $a \leq t_0 - \eta \leq t \leq t_0 + \eta \leq b$, oder $q_j(t) \leq -\varepsilon$ für alle $t \in]t_0 - \eta, t_0 + \eta[\subset]a,b[$. Nun existiert eine C^2-Funktion $h_j : [a,b] \longrightarrow \mathbb{R}$ mit $h_j \geq 0$, $h_j(t) = 0$ für $t \leq t_0 - \eta$ oder $t \geq t_0 + \eta$ und $h_j(t_0) > 0$.
Beachte, dass insbesondere $h_j(a) = 0 = h_j(b)$. Wähle nun $h : [a,b] \longrightarrow \mathbb{R}^n$ mit $h_i := 0$ für alle $i \neq j$ und h_j wie oben. Für dieses h gilt

$$\int_a^b q_j(t) h_j(t) dt = 0,$$

d.h.,

$$\begin{aligned} 0 &= \int_{t_0-\eta}^{t_0+\eta} q_j(t) h_j(t) dt \qquad \text{(wegen } h_j(t) = 0 \text{ für alle } t \notin]t_0 - \eta, t_0 + \eta]) \\ &= \begin{cases} \geq \varepsilon \int_{t_0-\eta}^{t_0+\eta} h_j(t) dt \text{ falls } q_j(t) \geq \varepsilon & \forall\, t \in]t_0 - \eta, t_0 + \eta[\\ \leq -\varepsilon \int_{t_0-\eta}^{t_0+\eta} h_j(t) dt \text{ falls } q_j(t) \leq -\varepsilon & \forall\, t \in]t_0 - \eta, t_0 + \eta[. \end{cases} \end{aligned}$$

Da $\int_{t_0-\eta}^{t_0+\eta} h_j(t) dt > 0$, bekommen wir einen Widerspruch. Dies zeigt also $q_j = 0$ für alle $1 \leq j \leq n$, was zu beweisen war. □

2.4 Vektorfelder und Potentiale

Definition 2.99 *Sei $U \subset \mathbb{R}^n$ eine (nichtleere) offene Teilmenge.*

i) Ein Vektorfeld *ist eine Abbildung $v : U \longrightarrow \mathbb{R}^n$.*

ii) Ein Potential *eines Vektorfeldes $v : U \longrightarrow \mathbb{R}^n$ ist eine differenzierbare Funktion $P : U \longrightarrow \mathbb{R}$ mit $\nabla P = v$, d.h., $\frac{\partial P}{\partial x_i}(x) = v_i(x)$ für alle $x \in U$ und $1 \leq i \leq n$, wobei $v(x) = (v_1(x), \ldots, v_n(x))$.*

Definition 2.100 *Sei $v : U \longrightarrow \mathbb{R}^n$ ein auf einer offenen Teilmenge des $\mathbb{R}^n$ definiertes stetiges Vektorfeld und $c : [a,b] \longrightarrow U$ eine stückweise C^1-Abbildung, d.h., c ist C^0 und es gibt eine Teilung $a = t_0 < t_1 < \cdots < t_k = b$ des Intervalls $[a,b]$ mit $c_{|[t_i,t_{i+1}]}$ C^1 (die Abbildung c hat bei jedem $t_i \in]a,b[$ Rechts- und Links-Ableitungen, die nicht unbedingt gleich sind) für alle $i = 0, 1, \ldots, k-1$. Das* Kurvenintegral *von v längs c wird definiert durch*

$$\int_c v := \int_a^b \langle v(c(s)), \dot{c}(s) \rangle ds \in \mathbb{R}.$$

Man beachte, dass wegen der Transformationsformel der Analysis I das Kurvenintegral invariant unter orientierungserhaltender Umparametrisierung ist: ist $\varphi : [a', b'] \longrightarrow [a,b]$ ein C^1-Diffeomorphismus mit $\varphi' > 0$, so gilt $\int_c v = \int_{c \circ \varphi} v$ für jedes stetige Vektorfeld v auf U.

Satz 2.101 (Potential und Kurvenintegral) *Sei $v : U \longrightarrow \mathbb{R}^n$ ein auf einer offenen Teilmenge des $\mathbb{R}^n$ definiertes stetiges Vektorfeld. Dann sind folgende Aussagen äquivalent:*

i) Das Vektorfeld v besitzt ein Potential.

ii) Für jede stückweise C^1 geschlossene Kurve $c : [a, b] \longrightarrow U$ gilt $\int_c v = 0$. Hierbei heißt c geschlossen, *wenn $c(a) = c(b)$ gilt.*

Der Beweis von Satz 2.101 beruht auf folgendem Lemma.

Lemma 2.102 *Sei $v : U \longrightarrow \mathbb{R}^n$ ein stetiges Vektorfeld mit Potential P, d.h., $\nabla P = v$. Dann gilt*

$$\int_c v = P(c(b)) - P(c(a))$$

für jede stückweise C^1 Kurve $c : [a, b] \longrightarrow U$.

Beweis von Lemma 2.102: Nach Definition ist

$$\int_c v = \int_a^b \langle v(c(s)), \dot{c}(s)\rangle ds = \int_a^b \langle \nabla P(c(s)), \dot{c}(s)\rangle ds = \int_a^b d_{c(s)}P(\dot{c}(s))ds.$$

Nach der Kettenregel ist $d_{c(s)}P(\dot{c}(s)) = \overset{\cdot}{\widehat{P \circ}}\, c(s)$. Dies liefert die Identität. □

Beweis von Satz 2.101[41]:
i) $\Longrightarrow$ *ii)*: Ist $v = \nabla P$ für eine differenzierbare Funktion $P : U \longrightarrow \mathbb{R}$, so gilt nach Lemma 2.102:

$$\int_c v = P(c(b)) - P(c(a)) = 0$$

für jede stückweise C^1-Kurve $c : [a, b] \longrightarrow U$ mit $c(a) = c(b)$.
ii) $\Longrightarrow$ *i)*: OBdA sei U zusammenhängend (sonst argumentiere auf jeder Zusammenhangskomponente von U). Sei $x_0 \in U$ ein fester Punkt.
Behauptung 1: *Für jedes $x \in U$ existiert eine stückweise C^1-Kurve $c_x : [0, 1] \longrightarrow U$ mit $c_x(0) = x_0$ und $c_x(1) = x$.*
Beweis: Sei $W := \{y \in U \mid \exists\, c_y : [0, 1] \longrightarrow U$ stückweise C^1 mit $c_y(0) = x_0, c_y(1) = y\}$. Offensichtlich ist $x_0 \in W$ (wähle $c_{x_0}(t) := x_0$ für alle $t \in [0, 1]$). Wir zeigen, dass W offen ist. Sei $y \in W$, dann existiert eine stückweise C^1-Kurve $c_y : [0, 1] \longrightarrow U$ mit $c_y(0) = x_0$ und $c_y(1) = y$. Wegen U offen existiert ein $r > 0$ mit $B_r(y) \subset U$. Für jedes $x \in B_r(y)$ ist dann $c_x : [0, 1] \longrightarrow U$, $[0, \frac{1}{2}] \ni t \longmapsto c_y(2t), [\frac{1}{2}, 1] \ni t \longmapsto (1 - (2t - 1))y + (2t - 1)x = 2(1-t)y + (2t-1)x$, eine stückweise C^1-Kurve in U mit $c_x(0) = x_0$ und $c_x(1) = x$. Daraus folgt $x \in W$ für alle $x \in B_r(y)$, d.h. $B_r(y) \subset U$. Dies zeigt, dass W offen in U ist.
Nun betrachte folgende Beziehung "$\sim$" auf U: für $x, y \in U$ gilt
$x \sim y :\Longleftrightarrow \exists\, c : [0, 1] \longrightarrow U$ stückweise C^1 mit $c(0) = x$ und $c(1) = y$.
Dann definiert "$\sim$" eine Äquivalenzrelation auf U (für die Transitivität setzt man Kurvenstücke zusammen, siehe oben). Wie oben kann bewiesen werden, dass jede Äquivalenzklasse offen in U ist. Insbesondere ist jede Äquivalenzklasse als Komplement der Vereinigung von offenen Teilmengen auch abgeschlossen. Insbesondere ist die Äquivalenzklasse W von x_0 offen, abgeschlossen und nichtleer ($x_0 \in W$). Da U zusammenhängend ist, folgt $W = U$, was zu beweisen war. $\surd$

Nun definieren wir die Funktion $P : U \longrightarrow \mathbb{R}$ durch

$$P(x) := \int_{c_x} v,$$

[41] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

wobei $c_x : [0,1] \longrightarrow U$ stückweise C^1 mit $c_x(0) = x_0$ und $c_x(1) = x$ ist.
Behauptung 2: *Die Abbildung P ist wohldefiniert, d.h., sind c_x und $\tilde{c}_x$ zwei Kurven wie oben, so gilt $\int_{c_x} v = \int_{\tilde{c}_x} v$.*
Beweis: Die Abbildung $d : [0,2] \longrightarrow U$, $t \longmapsto \begin{cases} c_x(t) \text{ für } t \in [0,1] \\ \tilde{c}_x(2-t) \text{ für } t \in [1,2] \end{cases}$ ist eine stückweise C^1-Kurve mit $d(0) = c_x(0) = x_0 = \tilde{c}_x(0) = d(2)$, insbesondere gilt nach Voraussetzung $\int_d v = 0$. Nach Definition gilt aber

$$\begin{aligned} \int_d v = \int_0^2 \langle v(d(t)), \dot{d}(t)\rangle dt &= \int_0^1 \langle v(c_x(t)), \dot{c}_x(t)\rangle dt + \int_1^2 \langle v(\tilde{c}_x(2-t)), -\dot{\tilde{c}}_x(2-t)\rangle dt \\ &= \int_{c_x} v - \int_0^1 \langle v(\tilde{c}_x(s)), \dot{\tilde{c}}_x(s)\rangle ds \quad (\text{setze } s := 2-t) \\ &= \int_{c_x} v - \int_{\tilde{c}_x} v. \end{aligned}$$

Daraus folgt $\int_{c_x} v = \int_{\tilde{c}_x} v$. $\checkmark$

Behauptung 3: *Die Abbildung P ist differenzierbar und es gilt $\nabla P = v$ auf U.*
Beweis: Sei $x \in U$ und $r > 0$ mit $B_r(x) \subset U$. Sei $h \in B_r(0)$. Sei $c_x : [0,1] \longrightarrow U$ stückweise C^1 mit $c_x(0) = x_0$ und $c_x(1) = x$. Sei $c_{x+h} : [0,2] \longrightarrow U$ die durch

$$t \longmapsto \begin{cases} c_x(t) \text{ für } t \in [0,1] \\ x + (t-1)h \text{ für } t \in [1,2] \end{cases}$$

definierte stückweise C^1-Kurve (beachte, dass $c_{x+h}(0) = x_0$ und $c_{x+h}(2) = x + h$ gelten). Da nach Behauptung P wohldefiniert ist, gilt

$$\begin{aligned} P(x+h) - P(x) = \int_{c_{x+h}} v - \int_{c_x} v &= \int_0^1 \langle v(c_x(t)), \dot{c}_x(t)\rangle dt + \int_1^2 \langle v(x + (t-1)h), h\rangle dt \\ &\quad - \int_0^1 \langle v(c_x(t)), \dot{c}_x(t)\rangle dt \\ &= \int_0^1 \langle v(x+sh), h\rangle ds \quad (\text{mit } s := t-1) \\ &= \int_0^1 \langle v(x), h\rangle ds + \int_0^1 \langle v(x+sh) - v(x), h\rangle ds \\ &= \langle v(x), h\rangle + \int_0^1 \langle v(x+sh) - v(x), h\rangle ds. \end{aligned}$$

Die Abbildung $h \longmapsto \langle v(x), h\rangle$ ist offensichtlich linear. Da außerdem v stetig in x ist, existiert für jedes $\epsilon > 0$ ein $\eta > 0$ mit $|v(x + \overline{h}) - v(x)| \le \epsilon$ für alle $\overline{h} \in B_\eta(0)$. Für alle $h \in B_\eta(0)$ gilt somit

$$\begin{aligned} \left| \int_0^1 \langle v(x+sh) - v(x), h\rangle ds \right| &\le \int_0^1 \left| \langle v(x+sh) - v(x), h\rangle \right| ds \\ &\le \int_0^1 \left| v(x+sh) - v(x) \right| \cdot |h| ds \qquad (\text{Cauchy-Schwarz}) \\ &\le \epsilon \cdot |h|. \end{aligned}$$

Dies beweist $\frac{1}{|h|} \cdot \int_0^1 \langle v(x+sh) - v(x), h\rangle ds \underset{|h| \to 0}{\longrightarrow} 0$ und die Behauptung. $\checkmark$

Der Beweis des Satzes ist damit abgeschlossen. $\square$

Man beachte, dass unter den Voraussetzungen von Satz 2.101 das Potential P notwendigerweise C^1 ist (wegen v stetig).

Beispiele 2.103

1. Sei $v : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $(x,y) \longmapsto (x,y)$. Sei $c : [0,1] \longrightarrow \mathbb{R}^2$, $t \longmapsto \binom{c_1(t)}{c_s(t)}$, eine beliebige stückweise C^1 geschlossene Kurve. Dann gilt
$$\begin{array}{rcl}\int_c v &=& \int_0^1 \langle v(c(s)), \dot{c}(s)\rangle ds \\ &=& \int_0^1 \langle \binom{c_1(s)}{c_2(s)}, \binom{\dot{c}_1(s)}{\dot{c}_2(s)} \rangle ds \\ &=& \int_0^1 c_1(s)\dot{c}_1(s) + c_2(s)\dot{c}_2(s) ds \\ &=& \frac{1}{2}\int_0^1 (c_1^2)'(s) + (c_2^2)'(s) ds \\ &=& \frac{1}{2}[c_1^2 + c_2^2]_0^1 \\ &=& 0 \qquad \text{wegen } c_1(0) = c_1(1) \text{ und } c_2(0) = c_2(1). \end{array}$$
Satz 2.101 liefert, dass v ein Potential besitzt. Tatsächlich kann ein solches Potential mit dem direkten Ansatz
$$\begin{cases} \frac{\partial P}{\partial x}(x,y) &= x \\ \\ \frac{\partial P}{\partial y}(x,y) &= y \end{cases}$$
gefunden werden: z.B. $P(x,y) = \frac{x^2+y^2}{2}$ + Konstante.

2. Sei allgemeiner $v : U \longrightarrow \mathbb{R}^n$ der Form $v(x_1, \ldots, x_n) := (v_1(x_1), \ldots, v_n(x_n))$, wobei $U = I_1 \times \cdots \times I_n$ das Produkt von offenen Intervallen ist und $v_i : I_i \longrightarrow \mathbb{R}$ eine stetige Funktion ist. Sei, für jedes $1 \leq i \leq n$, eine Stammfunktion V_i von v_i gewählt (d.h., $V_i : I_i \longrightarrow \mathbb{R}$ mit $V_i' = v_i$). Dann ist $P : U \longrightarrow \mathbb{R}$, $x \longmapsto \sum_{i=1}^n V_i(x_i)$ ein Potential für v.

3. Sei $v : \mathbb{R}^2 \setminus \{0\} \longrightarrow \mathbb{R}^2$, $(x,y) \longmapsto \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$. Sei $c : [0,1] \longrightarrow \mathbb{R}^2 \setminus \{0\}$, $t \longmapsto \binom{\cos(2\pi t)}{\sin(2\pi t)}$. Dann ist c eine C^∞ geschlossene Kurve in $\mathbb{R}^2 \setminus \{0\}$ mit
$$\int_c v = \int_0^1 \langle \begin{pmatrix} \frac{-\sin(2\pi s)}{\cos^2(2\pi s)+\sin^2(2\pi s)} \\ \frac{\cos(2\pi s)}{\cos^2(2\pi s)+\sin^2(2\pi s)} \end{pmatrix}, \begin{pmatrix} -2\pi\sin(2\pi s) \\ 2\pi\cos(2\pi s) \end{pmatrix} \rangle ds = 2\pi \int_0^1 \underbrace{\sin^2(2\pi s) + \cos^2(2\pi s)}_{1} ds = 2\pi.$$
Wegen $2\pi \neq 0$ impliziert Satz 2.101, dass v *kein* Potential besitzt.

Satz 2.104 (Potential und partielle Ableitungen) *Sei $v : U \longrightarrow \mathbb{R}^n$, $x \longmapsto (v_1(x), \ldots, v_n(x))$ ein auf einer offenen Teilmenge des $\mathbb{R}^n$ definiertes C^1-Vektorfeld.*

i) Besitzt v ein Potential, so gilt $\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ auf U, für alle $1 \leq i, j \leq n$.

ii) Ist U sternförmig bzgl. *$u \in U$, so gilt die Umkehrung: jedes C^1-Vektorfeld $v : U \longrightarrow \mathbb{R}^n$ mit $\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ für alle $1 \leq i, j \leq n$ besitzt ein Potential P. Ein Potential für P ist ggf. gegeben durch*
$$P(x) := \int_0^1 \langle v(u + t(x-u)), x - u\rangle dt \qquad \textit{für alle } x \in U.$$

Beweis:[42] *i)* Sei $v = \nabla P$ für eine differenzierbare Funktion $P : U \longrightarrow \mathbb{R}$. Beachte, dass wegen v C^1 die Funktion P C^2 auf U ist. Dann folgt aus dem Satz von Schwarz (Satz 2.83) und $v_i = \frac{\partial P}{\partial x_i}$:
$$\frac{\partial v_i}{\partial x_j} = \frac{\partial}{\partial x_j}(\frac{\partial P}{\partial x_i}) = \frac{\partial^2 P}{\partial x_j \partial x_i} = \frac{\partial^2 P}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} v_j = \frac{\partial v_j}{\partial x_i}.$$

[42] Dieser Beweis wurde in der Vorlesung nur teilweise durchgeführt.

ii) Nach Voraussetzung gilt $[u, x] = \{(1-t)u + tx \,|\, t \in [0,1]\} \subset U$ für jedes $x \in U$. Betrachte die Funktion $P : U \longrightarrow \mathbb{R}$, $x \longmapsto \int_0^1 \langle v(u + t(x-u)), x - u\rangle dt$, d.h., $P(x) = \sum_{i=1}^n (x_i - u_i) \cdot \int_0^1 v_i(u + t(x-u))dt$. Wir zeigen, dass P ein Potential für v ist. Zuerst ist $P(x)$ wegen $[u, x] \subset U$ für alle $x \in U$ wohldefiniert. Fixiere $i \in \{1, \ldots, n\}$. Da v_i nach Voraussetzung C^1 ist, ist die Abbildung $x \longmapsto \int_0^1 v_i(u + t(x-u))dt$ ebenfalls C^1 (siehe Kapitel 3) mit

$$\begin{aligned}\tfrac{\partial}{\partial x_j}\left(\int_0^1 v_i(u + t(x-u))dt\right) &= \int_0^1 \tfrac{\partial}{\partial x_j}(v_i(u + t(x-u))dt \\ &= \int_0^1 t \cdot \tfrac{\partial v_i}{\partial x_j}(u + t(x-u))dt.\end{aligned}$$

Daraus folgt, dass P eine C^1-Funktion ist mit

$$\begin{aligned}\tfrac{\partial P}{\partial x_i}(x) &= \sum_{j=1}^n \tfrac{\partial}{\partial x_i}(x_j - u_j) \cdot \int_0^1 v_j(u + t(x-u)dt + (x_j - u_j) \cdot \int_0^1 t \cdot \tfrac{\partial v_j}{\partial x_i}(u + t(x-u))dt \\ &= \int_0^1 v_i(u + t(x-u))dt + \sum_{j=1}^n (x_j - u_j) \int_0^1 t \cdot \tfrac{\partial v_j}{\partial x_i}(u + t(x-u))dt.\end{aligned}$$

Nach Voraussetzung und Anwendung der Kettenregel gilt, für jedes $1 \leq i \leq n$:

$$\begin{aligned}\sum_{j=1}^n (x_j - u_j) \int_0^1 t \cdot \tfrac{\partial v_j}{\partial x_i}(u + t(x-u))dt &= \sum_{j=1}^n (x_j - u_j) \int_0^1 t \cdot \tfrac{\partial v_i}{\partial x_j}(u + t(x-u))dt \\ &= \int_0^1 t \cdot \sum_{j=1}^n (x_j - u_j) \tfrac{\partial v_i}{\partial x_j}(u + t(x-u))dt \\ &= \int_0^1 t \cdot \tfrac{\partial}{\partial t}(v_i(u + t(x-u)))dt.\end{aligned}$$

Eine partielle Integration liefert

$$\begin{aligned}\int_0^1 t \cdot \tfrac{\partial}{\partial t}(v_i(u + t(x-u)))dt &= [t \cdot v_i(u + t(x-u))]_0^1 - \int_0^1 v_i(u + t(x-u))dt \\ &= v_i(x) - \int_0^1 v_i(u + t(x-u))dt.\end{aligned}$$

Es folgt $\frac{\partial P}{\partial x_i}(x) = v_i(x)$. Da v_i C^1 ist für jedes $1 \leq i \leq n$, ist insbesondere P differenzierbar auf U. □

Beispiele 2.105

1. Sei $v : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (-y, x)$. Dann ist v C^1 mit $\frac{\partial v_1}{\partial y}(x, y) = -1$ und $\frac{\partial v_2}{\partial x}(x, y) = 1$, insbesondere $\frac{\partial v_1}{\partial y} \neq \frac{\partial v_2}{\partial x}$ auf $\mathbb{R}^2$. Satz 2.104 impliziert, dass v *kein* Potential besitzt.

2. Sei $v : \mathbb{R}^2 \setminus \{0\} \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2})$. Dann ist v C^1 auf $\mathbb{R}^2 \setminus \{0\}$ mit $\frac{\partial v_2}{\partial x}(x, y) = \frac{x^2+y^2-x(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$ und $\frac{\partial v_1}{\partial x}(x, y) = \frac{-(x^2+y^2)+y(2y)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$. Insbesondere gilt $\frac{\partial v_2}{\partial x} = \frac{\partial v_1}{\partial y}$ auf $\mathbb{R}^2 \setminus \{0\}$. Beachte aber, dass wegen $\mathbb{R}^2 \setminus \{0\}$ nicht sternförmig der Satz 2.104 nicht angewendet werden kann; tatsächlich hat v nach Beispiel 2.103.3 kein Potential.

3. Sei $U := I_1 \times I_2$ das Produkt zweier offener Intervalle und $v : U \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (f_1(x, y), f_2(x, y))$, wobei f_1, f_2 C^1-Funktionen von U nach $\mathbb{R}$ sind. Gilt $\frac{\partial f_1}{\partial y} = \frac{\partial f_2}{\partial x}$ auf U, so folgt aus Satz 2.104 die Existenz eines Potentials für v. Ein Ansatz, um ein

solches Potential zu bestimmen, ist folgender: Sei $F_1 : U \longrightarrow \mathbb{R}$ mit $\frac{\partial F_1}{\partial x} = f_1$ (z.B. fixiere $a \in I_1$ und setze $F_1(x,y) := \int_a^x f_1(t,y)dt$) und suche das Potential P in der Form $P(x,y) = F_1(x,y) + G(x,y)$ für eine unbekannte Funktion G. Es müssen $\frac{\partial P}{\partial x} = f_1$ und $\frac{\partial P}{\partial y} = f_2$ gelten. Wegen der Definition von F_1 ist $\frac{\partial P}{\partial x} = f_1 + \frac{\partial G}{\partial x}$; dies ist gleich f_1, sobald G nicht von x abhängt, also $G : I_2 \longrightarrow \mathbb{R}$. Die zweite Bedingung liefert $\frac{\partial F_1}{\partial y} + G' = f_2$, d.h., $G' = f_2 - \frac{\partial F_1}{\partial y}$. Dabei ist zu beachten, dass wegen $\frac{\partial f_2}{\partial x} = \frac{\partial f_1}{\partial y}$ die Funktion $f_2 - \frac{\partial F_1}{\partial y}$ unabhängig von x ist, denn: $\frac{\partial f_2}{\partial x} - \frac{\partial^2 F_1}{\partial x \partial y} = \frac{\partial f_1}{\partial y} - \frac{\partial^2 F_1}{\partial y \partial x} = \frac{\partial f_1}{\partial y} - \frac{\partial f_1}{\partial y} = 0$ auf U. Insbesondere löst $P = F_1 + G$, wobei G eine Stammfunktion von $f_2 - \frac{\partial F_1}{\partial y}$ ist, beide Gleichungen $\frac{\partial P}{\partial x} = f_1$ und $\frac{\partial P}{\partial y} = f_2$ auf U.

Definition 2.106 *Sei $v : U \longrightarrow \mathbb{R}^n$ ein auf einer offenen Teilmenge des $\mathbb{R}^n$ definiertes differenzierbares Vektorfeld.*

i) Die Divergenz *von v ist die Funktion* $\mathrm{div}(v) : U \longrightarrow \mathbb{R}$, $\mathrm{div}(v) := \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i}$, *wobei $v(x) = (v_1(x), \ldots, v_n(x))$ für alle $x \in U$. Das Vektorfeld v heißt genau dann* quellenfrei, *wenn* $\mathrm{div}(v) = 0$ *auf U gilt.*

ii) Für $n = 3$ ist die Rotation *von v das Vektorfeld* $\mathrm{rot}(v) : U \longrightarrow \mathbb{R}^3$, $\mathrm{rot}(v) := \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}$. *Das Vektorfeld v heißt genau dann* wirbelfrei, *wenn* $\mathrm{rot}(v) = 0$ *auf U gilt.*

Definition 2.107 *Sei $f : U \longrightarrow \mathbb{R}^p$ eine auf einer offenen Teilmenge des $\mathbb{R}^n$ definierte C^2-Abbildung. Man definiere*

$$\begin{aligned} \Delta f : U &\longrightarrow \mathbb{R}^p \\ x &\longmapsto \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}. \end{aligned}$$

Der Operator $\Delta : C^2(U, \mathbb{R}^p) \longrightarrow C^0(U, \mathbb{R}^p)$ heißt der Laplace-Operator *auf U.*
Hierbei ist $C^k(U, \mathbb{R}^p) := \{f : U \longrightarrow \mathbb{R}^p,\ f\ C^k\}$.

Bezeichnungen 2.108

1. Man bezeichne mit ∇ den Operator $\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$.

2. Formal können Gradient, Divergenz, Rotation und Laplace-Operator folgendermaßen beschrieben werden:

   - Für jede differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ ist $\mathrm{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \nabla f$.
   - Für jedes differenzierbare Vektorfeld $v : U \longrightarrow \mathbb{R}^n$ ist $\mathrm{div}(v) = \langle \nabla, v \rangle$.
   - Für jedes differenzierbare Vektorfeld $v : U \longrightarrow \mathbb{R}^3$ ist $\mathrm{rot}(v) = \nabla \times v$.
   - Für jede C^2-Funktion $f : U \longrightarrow \mathbb{R}$ ist $\Delta f = \langle \nabla, \nabla \rangle f$.

Korollar 2.109 *Sei $v : U \longrightarrow \mathbb{R}^3$ ein auf einer offenen Teilmenge des $\mathbb{R}^3$ definiertes C^1-Vektorfeld.*

i) Besitzt v ein Potential, so gilt $\mathrm{rot}(v) = 0$.

ii) Gilt $\mathrm{rot}(v) = 0$ und ist U sternförmig, so hat v ein Potential.

Beweis: Es gilt $\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ für alle $1 \leq i, j \leq 3$ g.d.w. $\mathrm{rot}(v) = 0$. Wende Satz 2.104 an. □

Proposition 2.110 *Sei $U \subset \mathbb{R}^3$ eine (nichtleere) offene Teilmenge. Dann gilt:*

i) $\mathrm{rot}(\mathrm{grad}(f)) = 0$ *für jede C^2-Funktion $f : U \longrightarrow \mathbb{R}$.*

ii) $\mathrm{div}(\mathrm{rot}(v)) = 0$ *für jedes C^2-Vektorfeld $v : U \longrightarrow \mathbb{R}^3$.*

iii) $\mathrm{div}(\mathrm{grad}(f)) = \Delta f$ *für jede C^2-Funktion $f : U \longrightarrow \mathbb{R}$.*

Beweis:[43]

i) Wende den Satz von Schwarz an.

ii) Wende ebenfalls den Satz von Schwarz an.

iii) Die Identität folgt direkt aus den Definitionen. □

Beispiele 2.111

1. Sei $v : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $x \longmapsto x$. Dann ist v C^∞ mit
$$\mathrm{div}(v) = \sum_{i=1}^{3} \frac{\partial x_i}{\partial x_i} = 3 \text{ und } \mathrm{rot}(v) = \begin{pmatrix} \frac{\partial x_3}{\partial x_2} - \frac{\partial x_2}{\partial x_3} \\ \frac{\partial x_1}{\partial x_3} - \frac{\partial x_3}{\partial x_1} \\ \frac{\partial x_2}{\partial x_1} - \frac{\partial x_1}{\partial x_2} \end{pmatrix} = 0.$$

2. Sei $v : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $x \longmapsto \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}$. Dann ist v C^∞ mit
$$\mathrm{div}(v) = -\frac{\partial x_2}{\partial x_1} + \frac{\partial x_1}{\partial x_2} = 0 \text{ und } \mathrm{rot}(v)(x) = \begin{pmatrix} -\frac{\partial x_1}{\partial x_3} \\ -\frac{\partial x_2}{\partial x_3} \\ \frac{\partial x_1}{\partial x_1} + \frac{\partial x_2}{\partial x_2} \end{pmatrix}(x) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = 2e_3 \text{ für}$$
alle $x \in \mathbb{R}^3$.

[43] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Kapitel 3

Integralrechnung in mehreren Veränderlichen

3.1 Mess- und Maßräume

3.1.1 Messräume

Definition 3.1 *Eine Menge X heißt genau dann* abzählbar*, wenn entweder X endlich ist oder eine bijektive Abbildung $X \longrightarrow \mathbb{N}$ existiert.*

Proposition 3.2

i) Jede Teilmenge einer abzählbaren Menge ist abzählbar.

ii) Das Produkt endlich vieler abzählbarer Mengen ist abzählbar.

iii) Ist $(Y_n)_{n\in\mathbb{N}}$ eine abzählbare Familie von abzählbaren Teilmengen einer Menge X, so ist $\bigcup\limits_{n\in\mathbb{N}} Y_n$ abzählbar.

Beweis:[1] *i*) Sei $Y \subset X$ mit X abzählbar. Ist Y nicht endlich, so ist X auch nicht endlich, ggf. betrachte eine bijektive Abbildung $\varphi : X \longrightarrow \mathbb{N}$. Die Teilmenge $\varphi(Y)$ von $\mathbb{N}$ ist dann unendlich und kann daher in der Form $\varphi(Y) = \{a_k \,|\, k \in \mathbb{N}\}$ geschrieben werden, für eine eindeutige monoton wachsende Folge $(a_k)_{k\in\mathbb{N}}$ aus $\mathbb{N}$. Nun ist die Verknüpfung
$\begin{array}{rcl} Y \xrightarrow{\varphi|_Y} \varphi(Y) & \longrightarrow & \mathbb{N} \\ a_k & \longmapsto & k \end{array}$ bijektiv.
ii) Zu zeigen ist lediglich, dass $\mathbb{N}\times\mathbb{N}$ abzählbar ist. Dazu betrachte die Abbildung $\varphi : \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{N}$, $(n_1,n_2) \longmapsto \frac{(n_1+n_2)(n_1+n_2+1)}{2} + n_1$. Man zeige direkt, dass φ bijektiv ist.
iii) Schreibe $Y_n = \{y_k^{(n)} \,|\, k \in \mathbb{N}\}$ für alle $n \in \mathbb{N}$ (möglich wegen Y_n abzählbar) und betrachte die Abbildung $\varphi : \mathbb{N}\times\mathbb{N} \longrightarrow \bigcup\limits_{n\in\mathbb{N}} Y_n$, $(n,k) \longmapsto y_k^{(n)}$. Nach Konstruktion ist φ surjektiv. Da nach *ii*) die Menge $\mathbb{N}\times\mathbb{N}$ abzählbar ist, existiert eine bijektive Abbildung $\psi : \mathbb{N} \longrightarrow \mathbb{N}\times\mathbb{N}$. Daraus folgt, dass $\varphi\circ\psi : \mathbb{N} \longrightarrow \bigcup\limits_{n\in\mathbb{N}} Y_n$ surjektiv ist. Nun ist die Abbildung $\chi : \bigcup\limits_{n\in\mathbb{N}} Y_n \longrightarrow \mathbb{N}$, $y \longmapsto \min((\varphi\circ\psi)^{-1}(\{y\}))$, injektiv. Wie oben leitet man daraus her, dass $\bigcup\limits_{n\in\mathbb{N}} Y_n$ abzählbar ist. $\square$

[1] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Beispiele 3.3

1. Die Mengen $\mathbb{N}, \mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar. Denn: $\mathbb{Z}$ lässt sich als Vereinigung zweier abzählbarer Teilmengen schreiben und die Abbildung $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \longrightarrow \mathbb{Q}$, $(p,q) \longmapsto \frac{p}{q}$, ist surjektiv.

2. Die Menge $\{0,1\}^{\mathbb{N}} := \{u : \mathbb{N} \longrightarrow \{0,1\}\}$ ist nicht abzählbar. Denn: wäre $\{0,1\}^{\mathbb{N}}$ abzählbar, so gäbe es eine Folge $(u_n)_{n\in\mathbb{N}}$ aus $\{0,1\}^{\mathbb{N}}$ mit $\{0,1\}^{\mathbb{N}} = \{u_n \,|\, n \in \mathbb{N}\}$. Nun definiere die Abbildung $v : \mathbb{N} \longrightarrow \{0,1\}$, $n \longmapsto v(n)$ mit $v(n) \neq u_n(n)$ für alle $n \in \mathbb{N}$ (möglich wegen $u_n(n) = 0$ oder 1). Dann ist $v \in \{0,1\}^{\mathbb{N}}$ mit $v \neq u_n$ für alle $n \in \mathbb{N}$, Widerspruch. Dieses sogenannte *Cantor'sche Diagonalverfahren* zeigt, dass $\{0,1\}^{\mathbb{N}}$ nicht abzählbar ist.

3. Das Intervall $[0,1]$ ist nicht abzählbar. Denn: die Abbildung
$$\begin{array}{rcl} \Phi : \{0,1\}^{\mathbb{N}} & \longrightarrow & [0,1] \\ u_n & \longmapsto & \sum\limits_{j=1}^{\infty} \frac{u_n(j)}{2^j} \end{array}$$
definiert eine bijektive Abbildung (siehe dyadische Darstellung reeller Zahlen). Da $\{0,1\}^{\mathbb{N}}$ nicht abzählbar ist, ist $[0,1]$ auch nicht abzählbar.

4. Insbesondere ist kein Intervall der Form $]a,b]$ oder $[a,b]$ oder $[a,b[$ oder $]a,b[$ mit $a < b$ abzählbar. Daraus folgt, dass $\mathbb{R}^k$ nicht abzählbar ist für alle $k \in \mathbb{N}$, $k \geq 1$.

Definition 3.4 *Sei X eine Menge und $\mathcal{P}(X) := \{Y \subset X\}$ die Potenzmenge von X. Eine* σ-Algebra *auf X ist eine Teilmenge $\mathcal{A} \subset \mathcal{P}(X)$ mit:*

1. $\varnothing \in \mathcal{A}$

2. $A \in \mathcal{A} \Longrightarrow A^c := X \setminus A \in \mathcal{A}$

3. $(A_n)_{n\in\mathbb{N}}$ *Folge aus* $\mathcal{A} \Longrightarrow \bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Ein Messraum *ist ein Paar $(X, \mathcal{A})$, wobei $\mathcal{A}$ eine σ-Algebra auf X ist.*

Beispiele 3.5 *Sei X eine Menge.*

1. $\mathcal{A} := \{\varnothing, X\}$ ist eine σ-Algebra auf X.

2. $\mathcal{A} := \mathcal{P}(X)$ ist eine σ-Algebra auf X.

3. $\mathcal{A} := \mathcal{P}_{\mathrm{abz}}(X) := \{A \subset X | A \text{ oder } A^c \text{ ist abzählbar}\}$ ist eine σ-Algebra auf X.

Proposition 3.6 *Sei $\mathcal{T} \subset \mathcal{P}(X)$ eine Teilmenge von $\mathcal{P}(X)$ für eine Menge X. Dann ist $\mathcal{A}\langle\mathcal{T}\rangle := \bigcap\limits_{\substack{\mathcal{A}\ \sigma-\text{Algebra} \\ \mathcal{T}\subset\mathcal{A}}} \mathcal{A}$ eine σ-Algebra auf X. Die σ-Algebra $\mathcal{A}\langle\mathcal{T}\rangle$ ist die kleinste σ-Algebra auf X, die $\mathcal{T}$ enthält. Man nennt sie die von $\mathcal{T}$* erzeugte σ-Algebra.

Beweis[2]: Man beweise direkt (mit der Definition von σ-Algebren) die Tatsache, dass der Durchschnitt einer beliebigen Familie von σ-Algebren wieder eine σ-Algebra ist. Der Rest ist leicht zu überprüfen. □

[2]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Definition 3.7 *Sei (X, d) ein metrischer Raum. Sei $\mathcal{O} := \{U \subset X \mid U \text{ offen}\}$. Die von $\mathcal{O}$ erzeugte σ-Algebra $\mathcal{B}(X) := \mathcal{A}\langle\mathcal{O}\rangle$ heißt die* Borel'sche σ-Algebra *des metrischen Raumes (X, d).*

Bemerkungen 3.8

1. Sei $\mathcal{I} := \{I \subset \mathbb{R} \mid I =]a, b[\text{ für } a < b \text{ reell}\}$. Dann ist $\mathcal{A}\langle\mathcal{I}\rangle = \mathcal{B}(\mathbb{R})$. Denn: jede offene Teilmenge von $\mathbb{R}$ lässt sich als abzählbare Vereinigung offener Intervalle dieser Form schreiben (Übungsaufgabe). Analog ist $\mathcal{A}\langle\mathcal{I}'\rangle = \mathcal{B}(\mathbb{R})$, wobei $\mathcal{I}' := \{I \subset \mathbb{R} \mid I = [a, b[\text{ für } a < b \text{ reell}\}$ (und dies, obwohl $[a, b[$ keine offene Teilmenge von $\mathbb{R}$ ist).

2. Allgemeiner sei $\mathbb{R}^n$ mit der Standardmetrik versehen und $\mathcal{W} := \{U \subset \mathbb{R}^n \mid U =]a_1, b_1[\times \cdots \times]a_n, b_n[\text{ für } a_j < b_j \text{ reell und alle } 1 \leq j \leq n\}$. Dann gilt $\mathcal{A}\langle\mathcal{W}\rangle = \mathcal{B}(\mathbb{R}^n)$.

3. Wir erwähnen einen möglichen Ansatz, welcher helfen kann, eine gewisse Eigenschaft E für alle Elemente einer durch eine Teilmenge $\mathcal{T} \subset \mathcal{P}(X)$ erzeugte σ-Algebra $\mathcal{A}$ nachzuweisen. Der Ansatz ist zweistufig: als Erstes zeige man, dass alle Elemente von $\mathcal{T}$ die Eigenschaft E erfüllen; der zweite Schritt besteht darin, die Teilmenge
$$\mathcal{E}_X := \{A \subset X \mid A \text{ erfüllt die Eigenschaft } E\}$$
von $\mathcal{P}(X)$ zu betrachten und zu zeigen, dass $\mathcal{E}_X$ eine σ-Algebra auf X ist. Gelingt uns, diese beiden Aussagen zu beweisen, so gilt dann automatisch wegen $\mathcal{T} \subset \mathcal{E}_X$ (nach dem ersten Schritt) auch $\mathcal{A} = \mathcal{A}\langle\mathcal{T}\rangle \subset \mathcal{E}_X$, da $\mathcal{E}_X$ nach dem zweiten Schritt eine σ-Algebra ist und somit die kleinste $\mathcal{T}$ enthaltende σ-Algebra enthält, siehe Proposition 3.6. Dies zeigt dann, dass alle Elemente von $\mathcal{A}$ die Eigenschaft E erfüllen. Dieser Ansatz heißt *Prinzip der guten Mengen* und wird in den kommenden Abschnitten häufig auftreten.

3.1.2 Maße

Bezeichnungen 3.9

1. Wir bezeichnen $[0, \infty] := [0, \infty[\cup \{\infty\}$ und setzen das additive und das multiplikative Gesetz von $[0, \infty[$ auf $[0, \infty]$ folgendermaßen fort: setze $a + \infty := \infty =: \infty + a$ für alle $a \in [0, \infty]$ und $a \cdot \infty := \infty \cdot a := \begin{cases} 0 & \text{falls } a = 0 \\ \infty & \text{falls } a \neq 0. \end{cases}$

2. Kommutativität, Assoziativität und Distributivität bleiben erhalten. Man beachte: "$a + b = a + c \Longrightarrow b = c$" gilt nur falls $a \in [0, \infty[$; "$a \cdot b = a \cdot c \Longrightarrow b = c$" gilt nur falls $a \in]0, \infty[$.

3. Die natürliche Ordnungsrelation "$\leq$" wird auf $[0, \infty]$ durch $a \leq \infty$ für alle $a \in [0, \infty]$ und $\infty \not\leq a$ für alle $a \in [0, \infty[$ fortgesetzt. Dann gilt, wie in $[0, \infty[$: $a \leq b \Longrightarrow a + c \leq b + c$ und $a \cdot c \leq b \cdot c$ für alle $a, b, c \in [0, \infty]$.

4. Ist $(u_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge aus $[0, \infty]$ (d.h., $u_n \in [0, \infty]$ und $u_n \leq u_{n+1}$ für alle $n \in \mathbb{N}$), so setzt man
$$\lim_{n\to\infty} u_n := \begin{cases} \lim\limits_{n\to\infty} u_n \in [0, \infty[& \text{falls } (u_n)_{n\in\mathbb{N}} \text{ beschränkt in } \mathbb{R} \\ \infty & \text{sonst.} \end{cases}$$

Per Definition gilt also: $\lim_{n\to\infty} u_n = \infty \iff u_n \underset{n\to\infty}{\longrightarrow} \infty$ (im üblichen Sinne) für jede Folge $(u_n)_{n\in\mathbb{N}}$ aus $[0,\infty[$. Analog definieren wir, für eine beliebige Folge $(x_n)_{n\in\mathbb{N}}$ aus $[0,\infty]$, die Summe $\sum\limits_{n=0}^{\infty} x_n := \lim\limits_{n\to\infty} (\sum\limits_{k=0}^{n} x_k) \in [0,\infty]$.

Definition 3.10 *Sei $\mathcal{A}$ eine σ-Algebra auf einer Menge X. Ein (positives)* Maß *auf $\mathcal{A}$ ist eine Abbildung $\mu : \mathcal{A} \longrightarrow [0,\infty]$ mit:*

i) $\mu(\varnothing) = 0$ *und*

ii) $\mu(\bigcup\limits_{n\in\mathbb{N}} A_n) = \sum\limits_{n=0}^{\infty} \mu(A_n)$ *für jede* disjunkte abzählbare *Familie $(A_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{A}$ (d.h. $A_n \cap A_m = \varnothing$ für alle $n \neq m$). Diese Eigenschaft heißt* σ-Additivität.

Ein Maßraum *ist ein Tripel $(X,\mathcal{A},\mu)$, wobei $(X,\mathcal{A})$ ein Messraum und μ ein Maß auf $\mathcal{A}$ ist.*

Beispiele 3.11

1. Sei X eine Menge und $\mathcal{A} := \mathcal{P}(X)$ (siehe Beispiel 3.5.2). Man definiere $\mu : \mathcal{P}(X) \longrightarrow [0,\infty],\ A \longmapsto \begin{cases} |A| & \text{falls } A \text{ endlich,} \\ \infty & \text{sonst} \end{cases}$, wobei $|A|$ die Anzahl der Elemente in A bezeichnet. Dann ist μ ein Maß auf $\mathcal{P}(X)$, welches *Zählmaß* genannt wird.

2. Sei X eine nichtleere Menge und $x \in X$ ein Punkt. Dann definiert $\delta_x : \mathcal{P}(X) \longrightarrow [0,\infty],\ A \longmapsto \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \notin A \end{cases}$ ein Maß auf $\mathcal{P}(X)$. Dieses Maß wird *Dirac-Maß* genannt.

Proposition 3.12 *Sei $(X,\mathcal{A},\mu)$ ein Maßraum.*

1. *Seien $A,B \in \mathcal{A}$, dann gilt $\mu(A\cup B) + \mu(A\cap B) = \mu(A) + \mu(B)$. Insbesondere gilt $\mu(A) \leq \mu(B)$ falls $A \subset B$.*

2. *Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathcal{A}$. Dann gilt:*

   i) $\mu(\bigcup\limits_{n\in\mathbb{N}} A_n) \leq \sum\limits_{n\in\mathbb{N}} \mu(A_n)$.

   ii) Ist $(A_n)_{n\in\mathbb{N}}$ monoton wachsend (d.h., $A_n \subset A_{n+1}$ für alle n), so gilt $\mu(\bigcup\limits_{n\in\mathbb{N}} A_n) = \lim\limits_{n\to\infty} \mu(A_n)$.

   iii) Ist $(A_n)_{n\in\mathbb{N}}$ monoton fallend (d.h., $A_n \supset A_{n+1}$ für alle n) und gilt $\mu(A_0) < \infty$, *so gilt* $\mu(\bigcap\limits_{n\in\mathbb{N}} A_n) = \lim\limits_{n\to\infty} \mu(A_n)$.

Beweis:[3] 1. Bemerke zuerst, dass für jede endliche disjunkte Familie $(A_i)_{i=1,\ldots,k}$ von Elementen von $\mathcal{A}$ die Identität $\mu(\bigcup\limits_{i=1}^{k} A_i) = \sum\limits_{i=1}^{k} \mu(A_i)$ erfüllt ist (setze $A_j := \varnothing$ für alle $j \geq k+1$ und benutze die σ-Additivität von μ). Außerdem gilt, für alle $A,B \in \mathcal{A} : A\cap B = (A^c \cup B^c)^c$ mit A^c, $B^c \in \mathcal{A}$, somit $A \cap B \in \mathcal{A}$. Analog ist $A \setminus B := A \cap B^c \in \mathcal{A}$. Wegen $A\cup B = (A\setminus B) \cup (A\cap B) \cup (B \setminus A)$ mit $A\setminus B$, $A \cap B$, $B\setminus A \in \mathcal{A}$ und disjunkt gilt

[3] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

$\mu(A \cup B) = \mu(A \setminus B) + \mu(A \cap B) + \mu(B \setminus A)$. Wegen $A = (A \setminus B) \dot\cup (A \cap B)$ (disjunkte Vereinigung) gilt $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$ und analog $\mu(B) = \mu(B \setminus A) + \mu(A \cap B)$. Daraus folgt

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A \setminus B) + \mu(A \cap B) + \mu(B \setminus A) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Dies beweist die erste Identität. Gilt nun $A \subset B$, so gilt für $A' := A$ und $B' := B \setminus A \in \mathcal{A}$: $\mu(A' \cup B') = \mu(B)$, $\mu(A' \cap B') = \mu(\varnothing) = 0$, $\mu(B') = \mu(B \setminus A) \geq 0$, somit $\mu(B) = \mu(A) + \mu(B') \geq \mu(A)$.

2. *i*) Setze, für alle $n \in \mathbb{N}$: $\overline{A}_n := A_n \setminus (\bigcup_{i=0}^{n-1} A_i)$ (und $\overline{A}_0 := A_0$). Dann ist $(\overline{A}_n)_{n\in\mathbb{N}}$ eine abzählbare disjunkte Familie von Elementen von $\mathcal{A}$ mit $\bigcup_{n\in\mathbb{N}} \overline{A}_n = \bigcup_{n\in\mathbb{N}} A_n$. Insbesondere gilt $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \mu(\bigcup_{n\in\mathbb{N}} \overline{A}_n) = \sum_{n=0}^{\infty} \mu(\overline{A}_n)$. Wegen $\overline{A}_n \subset A_n$ gilt nach 1: $\mu(\overline{A}_n) \leq \mu(A_n)$ für alle $n \in \mathbb{N}$.

ii) Wie vorher setzt man $\overline{A}_n := A_n \setminus (\bigcup_{i=0}^{n-1} A_i)$ für alle $n \in \mathbb{N}$. Beachte, dass $\overline{A}_n = A_n \setminus A_{n-1}$ (wegen $(A_n)_{n\in\mathbb{N}}$ monoton wachsend) gilt. Da $(\overline{A}_n)_{n\in\mathbb{N}}$ eine disjunkte abzählbare Familie von Elementen von $\mathcal{A}$ mit $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \overline{A}_n$ ist, gilt

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \mu(\bigcup_{n\in\mathbb{N}} \overline{A}_n) = \sum_{n=0}^{\infty}(\overline{A}_n).$$

Nun ist aber $\overline{A}_0 \cup \cdots \cup \overline{A}_n = A_n$, insbesondere $\sum_{i=0}^{n} \mu(\overline{A}_i) = \mu(A_n)$; daraus folgt

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty}(\sum_{i=0}^{n} \mu(\overline{A}_i)) = \lim_{n\to\infty} \mu(A_n).$$

iii) Bemerke zuerst, dass $(\mu(A_n))_{n\in\mathbb{N}}$ als monoton fallende und von unten beschränkte Folge konvergiert (in $[0, \infty]$). Wegen $\bigcap_{n\in\mathbb{N}} A_n \subset A_k$ gilt $\mu(\bigcap_{n\in\mathbb{N}} A_n) \leq \mu(A_k)$ für alle $k \in \mathbb{N}$ und somit $\mu(\bigcap_{n\in\mathbb{N}} A_n) \leq \lim_{n\to\infty} \mu(A_n)$. Nun betrachte $\hat{A}_n := A_0 \setminus A_n$ für alle $n \in \mathbb{N}$. Dann ist $(\hat{A}_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge aus X mit $\bigcup_{n\in\mathbb{N}} \hat{A}_n = A_0 \setminus (\bigcap_{n\in\mathbb{N}} A_n)$, insbesondere gilt $\mu(A_0) = \mu(\bigcup_{n\in\mathbb{N}} \hat{A}_n) + \mu(\bigcap_{n\in\mathbb{N}} A_n)$. Nach *ii*) gilt somit $\mu(A_0) = \lim_{n\to\infty} \mu(\hat{A}_n) + \mu(\bigcap_{n\in\mathbb{N}} A_n)$. Wiederum gilt $\mu(A_0) = \mu(\hat{A}_n) + \mu(A_n)$. Da $\mu(A_0) < \infty$ ist, gilt dann $\mu(\hat{A}_n), \mu(A_n) < \infty$ für alle n mit $\mu(A_n) = \mu(A_0) - \mu(\hat{A}_n)$, somit $\lim_{n\to\infty} \mu(A_n) = \mu(A_0) - \lim_{n\to\infty} \mu(\hat{A}_n) = \mu(\bigcap_{n\in\mathbb{N}} A_n)$. □

Bemerkungen 3.13

1. Die Voraussetzung $\mu(A_0) < \infty$ in Proposition 3.12.*iii*) kann durch "$\mu(A_p) < \infty$ für ein $p \in \mathbb{N}$" ersetzt werden (ersetze dann $(A_n)_{n\in\mathbb{N}}$ durch $(A_{n+p})_{n\in\mathbb{N}}$).

2. Diese Voraussetzung ist wichtig. Betrachte z.B. $X := \mathbb{N}$, $\mathcal{A} := \mathcal{P}(X)$ und $\mu :=$ Zählmaß (Beispiel 3.11.1). Sei $(A_n := \{n, n+1, \dots\} = \{k \,|\, k \geq n\})_{n\in\mathbb{N}}$. Dann

ist $(A_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge von Elementen von $\mathcal{A}$ mit $\bigcap\limits_{n\in\mathbb{N}} A_n = \varnothing$, insbesondere $\mu(\bigcap\limits_{n\in\mathbb{N}} A_n) = \mu(\varnothing) = 0$. Dennoch ist $\mu(A_n) = \infty$ für alle $n \in \mathbb{N}$, daher gilt $\lim\limits_{n\to\infty} \mu(A_n) > \mu(\bigcap\limits_{n\in\mathbb{N}} A_n)$.

Definition 3.14 *Sei $(X, \mathcal{A})$ ein Messraum.*

i) Ein Maß μ auf $\mathcal{A}$ heißt genau dann endlich*, wenn $\mu(X) < \infty$ gilt.*

ii) Ein Maß μ auf $\mathcal{A}$ heißt genau dann σ-endlich*, wenn eine Folge $(A_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{A}$ existiert mit $\bigcup\limits_{n\in\mathbb{N}} A_n = X$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$.*

Ein *Wahrscheinlichkeitsraum* ist ein Maßraum $(X, \mathcal{A}, \mu)$ mit $\mu(X) = 1$.

3.1.3 Das Lebesgue-Maß auf $\mathbb{R}$

Satz 3.15 *Es gibt ein eindeutiges Maß μ auf $\mathcal{B}(\mathbb{R})$ mit*

$$\mu([a, b[) = b - a$$

für alle $a, b \in \mathbb{R}$ mit $a \leq b$.

Die Konstruktion dieses Maßes wird in den Anhang A verschoben.

Definition 3.16 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum.*

i) Eine μ-Nullmenge *ist eine Teilmenge N von X, so dass ein $N' \in \mathcal{A}$ existiert mit $N \subset N'$ und $\mu(N') = 0$.*

ii) Der Maßraum $(X, \mathcal{A}, \mu)$ heißt genau dann vollständig*, wenn jede μ-Nullmenge ein Element von $\mathcal{A}$ ist.*

Beispiele 3.17

1. Sei $\mathcal{B}(\mathbb{R})$ mit dem Maß μ aus Satz 3.15 versehen. Dann gilt, für jedes $a \in \mathbb{R}$: $\mu(\{a\}) = 0$. Denn: $\{a\} = \bigcap\limits_{\substack{n\in\mathbb{N}\\ n\geq 1}} [a, a+\frac{1}{n}[$ mit $\mu([a, a+1[) = 1 < \infty$ und $[a, a+\frac{1}{n+1}[\subset [a, a+\frac{1}{n}[$ für alle $n \geq 1$, somit folgt aus Proposition 3.12: $\mu(\{a\}) = \lim\limits_{n\to\infty} \mu([a, a+\frac{1}{n}[) = \lim\limits_{n\to\infty} \frac{1}{n} = 0$.

2. Daraus folgt insbesondere, dass jede abzählbare Teilmenge N von $\mathbb{R}$ eine μ-Nullmenge ist. Denn: schreibe $N = \{a_n, n \in \mathbb{N}\}$ (wobei N eventuell endlich ist), dann gilt $\mu(N) \leq \sum\limits_{n=0}^{\infty} \mu(\{a_n\})$ nach Proposition 3.12 und somit $\mu(N) \leq 0$ nach dem letzten Beispiel, d.h., $\mu(N) = 0$.

3. Nicht jede μ-Nullmenge ist abzählbar. Betrachte z.B. die sogenannte *Cantormenge* $C \subset [0, 1]$, welche folgendermaßen definiert wird. Definiere induktiv eine Folge $(C_n)_{n\in\mathbb{N}}$ von Teilmengen von $[0, 1]$ durch: $C_0 := [0, 1]$ und C_{n+1} entsteht durch das Entfernen des (offenen) mittleren Drittels jedes Intervalls aus C_n. Z.B. ist $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ usw. Setze dann $C := \bigcap\limits_{n\in\mathbb{N}} C_n$. Dann ist $C \in \mathcal{B}(\mathbb{R})$ mit $\mu(C) = 0$, dennoch ist C *nicht* abzählbar (Übungsaufgabe).

4. Jedes auf $\mathcal{P}(X)$ definierte Maß ist selbstverständlich vollständig, da jede Teilmenge von X ein Element von $\mathcal{P}(X)$ ist. Insbesondere sind das Zählmaß und das Dirac-Maß auf $\mathcal{P}(X)$ vollständig.

Proposition 3.18 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Man definiere*

$$\hat{\mathcal{A}} := \{A \cup N \mid A \in \mathcal{A}, N\ \mu - \mathit{Nullmenge}\}$$
$$\hat{\mu} : \hat{\mathcal{A}} \longrightarrow [0, \infty],\ A \cup N \longmapsto \mu(A).$$

Dann ist $\hat{\mathcal{A}}$ eine σ-Algebra, die $\mathcal{A}$ enthält und $\hat{\mu}$ ein vollständiges Maß auf $\hat{\mathcal{A}}$, welches μ fortsetzt (d.h., $\hat{\mu}_{|\mathcal{A}} = \mu$). Außerdem ist $\hat{\mu}$ die minimale vollständige Fortsetzung von μ: ist $(\tilde{\mathcal{A}}, \tilde{\mu})$ ein weiteres Paar mit σ-Algebra $\tilde{\mathcal{A}}$ auf X und vollständigem Maß $\tilde{\mu}$ auf $\tilde{\mathcal{A}}$ so dass $\mathcal{A} \subset \tilde{\mathcal{A}}$ und $\tilde{\mu}_{|\mathcal{A}} = \mu$ gelten, so gelten $\hat{\mathcal{A}} \subset \tilde{\mathcal{A}}$ und $\tilde{\mu}_{|\hat{\mathcal{A}}} = \hat{\mu}$.

Beweis:[4] Wir zeigen zuerst, dass $\hat{\mathcal{A}}$ eine σ-Algebra ist. Da $\emptyset$ eine μ-Nullmenge ist, ist $\mathcal{A} \subset \hat{\mathcal{A}}$ und somit gilt $\emptyset \in \hat{\mathcal{A}}$. Analog ist jede μ-Nullmenge ein Element von $\hat{\mathcal{A}}$. Sei $\hat{A} \in \hat{\mathcal{A}}$, dann existieren $A \in \mathcal{A}$ und eine μ-Nullmenge $N \subset X$ so, dass $\hat{A} = A \cup N$. Nach Definition einer μ-Nullmenge existiert ein $N' \in \mathcal{A}$ mit $N \subset N'$ und $\mu(N') = 0$. Es folgt $\hat{A}^c = A^c \cap N^c = (A^c \cap N'^c) \cup (A \cap (N^c \setminus N'^c))$ mit $A^c \cap N'^c \in \mathcal{A}$ und $A \cap (N^c \setminus N'^c) = A \cap (N' \setminus N) \subset N' \setminus N \subset N'$, wobei $N' \in \mathcal{A}$ Maß $\mu(N') = 0$ hat; das heißt, $A \cap (N^c \setminus N'^c)$ ist eine μ-Nullmenge. Damit ist $\hat{A}^c \in \hat{\mathcal{A}}$. Sei nun $(\hat{A}_n)_{n\in\mathbb{N}}$ eine Folge von Elementen von $\hat{\mathcal{A}}$. Dann existiert, für jedes $n \in \mathbb{N}$, ein $A_n \in \mathcal{A}$ und eine μ-Nullmenge N_n mit $\hat{A}_n = A_n \cup N_n$. Es folgt $\cup_{n\in\mathbb{N}}\hat{A}_n = (\cup_{n\in\mathbb{N}}A_n) \cup (\cup_{n\in\mathbb{N}}N_n)$, mit $\cup_{n\in\mathbb{N}}A_n \in \mathcal{A}$ (weil $\mathcal{A}$ eine σ-Algebra ist); außerdem is $\cup_{n\in\mathbb{N}}N_n$ eine μ-Nullmenge, denn: für jedes $n \in \mathbb{N}$ existiert ein $N_n' \in \mathcal{A}$ mit $N_n \subset N_n'$ und $\mu(N_n') = 0$, insbesondere ist $\cup_{n\in\mathbb{N}}N_n \subset \cup_{n\in\mathbb{N}}N_n'$, wobei $\cup_{n\in\mathbb{N}}N_n' \in \mathcal{A}$ und $\mu(\cup_{n\in\mathbb{N}}N_n') \leq \sum_{n=0}^{\infty}\mu(N_n') = 0$, d.h., $\mu(\cup_{n\in\mathbb{N}}N_n') = 0$. Dies beweist $\cup_{n\in\mathbb{N}}\hat{A}_n \in \hat{\mathcal{A}}$. Insgesamt bekommen wir, dass $\hat{\mathcal{A}}$ eine $\mathcal{A}$ und alle μ-Nullmengen enthaltende σ-Algebra auf X ist.
Als Nächstes zeigen wir, dass $\hat{\mu}$ ein wohldefiniertes Maß auf $\hat{\mathcal{A}}$ ist. Sei $\hat{A} \in \hat{\mathcal{A}}$. Um zu zeigen, dass $\hat{\mu}(\hat{A})$ wohldefiniert ist, müssen wir zeigen, dass $\mu(A_1) = \mu(A_2)$ gilt, sobald $\hat{A} = A_1 \cup N_1 = A_2 \cup N_2$ mit $A_i \in \mathcal{A}$ und μ-Nullmenge N_i, für $i = 1, 2$. Wie vorher existieren $N_1', N_2' \in \mathcal{A}$ mit $N_i \subset N_i'$ und $\mu(N_i') = 0$, für $i = 1, 2$. Wegen $A_1 \subset A_2 \cup N_2 \subset A_2 \cup N_2'$ mit $A_2 \cup N_2' \in \mathcal{A}$ folgt $\mu(A_1) \leq \mu(A_2 \cup N_2') \leq \mu(A_2) + \mu(N_2') = \mu(A_2)$, somit $\mu(A_1) \leq \mu(A_2)$. Analog ist $\mu(A_2) \leq \mu(A_1)$, woraus $\mu(A_1) = \mu(A_2)$ folgt. Damit ist $\hat{\mu}$ wohldefiniert auf $\hat{\mathcal{A}}$. Dass $\hat{\mu}$ ein Maß auf $\hat{\mathcal{A}}$ ist, ist eine direkte Folgerung daraus, dass μ ein Maß auf $\mathcal{A}$ ist. Zuerst gilt offensichtlich $\hat{\mu}(\emptyset) = 0$ (schreibe $\emptyset = \emptyset \cup \emptyset$). Ist $(\hat{A}_n)_{n\in\mathbb{N}}$ eine Folge von disjunkten Elementen von $\hat{\mathcal{A}}$, so existieren, für alle $n \in \mathbb{N}$ ein $A_n \in \mathcal{A}$ und eine μ-Nullmenge N_n mit $\hat{A}_n = A_n \cup N_n$; da $(A_n)_{n\in\mathbb{N}}$ dann auch eine Folge von disjunkten Elementen von $\mathcal{A}$ und $\cup_{n\in\mathbb{N}}N_n$ eine μ-Nullmenge sind (siehe oben), gilt

$$\hat{\mu}(\cup_{n\in\mathbb{N}}\hat{A}_n) = \hat{\mu}(\cup_{n\in\mathbb{N}}A_n \cup \cup_{n\in\mathbb{N}}N_n) = \mu(\cup_{n\in\mathbb{N}}A_n) = \sum_{n=0}^{\infty}\mu(A_n) = \sum_{n=0}^{\infty}\hat{\mu}(\hat{A}_n).$$

Dies beweist, dass $\hat{\mu}$ ein Maß auf $\hat{\mathcal{A}}$ ist.
Der Maßraum $(X, \hat{A}, \hat{\mu})$ ist vollständig: sei nämlich $\hat{N}$ eine $\hat{\mu}$-Nullmenge, dann existieren $A \in \mathcal{A}$ und eine μ-Nullmenge N mit $\hat{N} \subset A \cup N$ und $\hat{\mu}(A \cup N) = 0$, d.h., $\mu(A) = 0$. Insbesondere ist A – und somit $\hat{N} \cap A$ sowie $\hat{N} \cap N$ – eine μ-Nullmenge, daher ist $\hat{N} =$

[4] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

$(\hat{N} \cap A) \cup (\hat{N} \cap N)$ als Vereinigung zweier μ-Nullmengen wieder eine μ-Nullmenge. Da jede μ-Nullmenge in $\hat{\mathcal{A}}$ liegt, folgt $\hat{N} \in \hat{\mathcal{A}}$.
Als Letztes zeigen wir, dass $\hat{\mu}$ die minimale vollständige Fortsetzung von μ ist. Sei $(\tilde{\mathcal{A}}, \tilde{\mu})$ ein weiteres Paar mit σ-Algebra $\tilde{\mathcal{A}}$ auf X und vollständigem Maß $\tilde{\mu}$ auf $\tilde{\mathcal{A}}$ so dass $\mathcal{A} \subset \tilde{\mathcal{A}}$ und $\tilde{\mu}_{|\mathcal{A}} = \mu$ gelten. Dann ist wegen $\tilde{\mu}_{|\mathcal{A}} = \mu$ jede μ-Nullmenge auch eine $\tilde{\mu}$-Nullmenge (gilt $N \subset N'$ mit $N' \in \mathcal{A}$ und $\mu(N') = 0$, so gilt auch $N' \in \tilde{\mathcal{A}}$ mit $\tilde{\mu}(N') = \mu(N') = 0$); da $\tilde{\mu}$ vollständig ist, muss jede μ-Nullmenge zu $\tilde{\mathcal{A}}$ gehören. Dies zeigt $A \cup N \in \tilde{\mathcal{A}}$ für alle $A \in \mathcal{A}$ und μ-Nullmenge N, d.h., $\hat{\mathcal{A}} \subset \tilde{\mathcal{A}}$. Außerdem gilt, für jedes $\hat{A} = A \cup N \in \hat{\mathcal{A}}$ mit $A \in \mathcal{A}$ und μ-Nullmenge N:

$$\tilde{\mu}(\hat{A}) = \tilde{\mu}(A \cup N) \leq \tilde{\mu}(A) + \tilde{\mu}(N) = \tilde{\mu}(A) = \mu(A) = \hat{\mu}(\hat{A}).$$

Zusammem mit $\tilde{\mu}(A \cup N) \geq \tilde{\mu}(A) = \mu(A) = \hat{\mu}(\hat{A})$ folgt $\tilde{\mu}(\hat{A}) = \hat{\mu}(\hat{A})$, d.h., $\tilde{\mu}_{|_{\hat{\mathcal{A}}}} = \hat{\mu}$. □

Das zum Maß μ gehörige Maß $\hat{\mu}$ aus Proposition 3.18 heißt *Vervollständigung* des Maßes μ.

Definition 3.19 *Das* Lebesgue-Maß *auf* $\mathbb{R}$ *ist die Vervollständigung des Maßes* μ *aus* Satz 3.15. *Es wird üblicherweise* λ *(oder* λ_1*) bezeichnet und die zugehörige* σ*-Algebra mit* $\mathcal{L}$ *(oder* $\mathcal{L}_1$*).*

Bemerke, dass insbesondere $\mathcal{B}(\mathbb{R}) \subset \mathcal{L}$ und $\lambda_{|\mathcal{B}(\mathbb{R})} = \mu$ gelten, daher auch $\lambda([a, b[) = b - a$ für alle $a, b \in \mathbb{R}$. Bemerke auch, dass λ ein σ-endliches (aber kein endliches) Maß ist.

3.2 Integration auf einem Maßraum

3.2.1 Messbare Funktionen

Definition 3.20 *Seien* $(X, \mathcal{A})$ *und* $(Y, \mathcal{B})$ *Messräume. Eine Abbildung* $f : X \longrightarrow Y$ *heißt genau dann* messbar*, wenn* $f^{-1}(B) \in \mathcal{A}$ *für alle* $B \in \mathcal{B}$ *gilt.*

Beispiele 3.21

1. Für eine Menge X und $\mathcal{A} := \mathcal{P}(X)$ ist jede Abbildung $f : X \longrightarrow Y$ messbar, unabhängig von der σ-Algebra $\mathcal{B}$ auf Y.

2. Für metrische Räume (X, d_X) und (Y, d_Y) ist jede stetige Abbildung $f : X \longrightarrow Y$ messbar bzgl. $\mathcal{A} := \mathcal{B}(X)$ und $\mathcal{B} := \mathcal{B}(Y)$. Betrachte nämlich $\mathcal{B}' := \{B \subset Y \,|\, f^{-1}(B) \in \mathcal{A}\}$. Dann ist es leicht, zu überprüfen, dass $\mathcal{B}'$ eine σ-Algebra auf Y ist, welche alle offenen Teilmengen von Y enthält (denn: f stetig $\Longrightarrow f^{-1}(U)$ ist offen in X für jede offene Teilmenge U von Y, siehe Proposition 2.19). Nach Konstruktion der erzeugten σ-Algebra $\mathcal{B}(Y)$ muss dann $\mathcal{B}(Y) \subset \mathcal{B}'$ gelten, siehe das Prinzip der guten Mengen (Bemerkung 3.8.3) und Proposition 3.6. Insbesondere gilt $f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(Y)$, was zu beweisen war.

3. Für eine Teilmenge A von X definiert man die *charakteristische Funktion* von A durch $\chi_A : X \longrightarrow \{0, 1\}$, $x \longmapsto \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases}$. Für eine gegebene σ-Algebra $\mathcal{A}$ auf X gilt: die Abbildung χ_A ist genau dann messbar (bzgl. $\mathcal{A}$ und $\mathcal{B} := \mathcal{P}(\{0, 1\})$, wenn $A \in \mathcal{A}$ gilt. Denn: $\chi_A^{-1}(\{1\}) = A$.

Proposition 3.22 *Seien $(X,\mathcal{A})$, $(Y,\mathcal{B})$ und $(Z,\mathcal{C})$ messbare Räume sowie $f : X \longrightarrow Y$ und $g : Y \longrightarrow Z$ Abbildungen. Sind f und g messbar, so ist $g \circ f$ messbar.*

Beweis: Sei $C \in \mathcal{C}$, dann gilt $(g \circ f)^{-1}(C) = f^{-1}(\underbrace{g^{-1}(C)}_{\in \mathcal{B}}) \in \mathcal{A}$. □

Nun interessieren wir uns für die messbaren reellwertigen Funktionen, wobei $\mathbb{R}$ die Borel'sche σ-Algebra trägt. Wie im Beispiel 3.21.2 ist zu bemerken, dass $f : X \longrightarrow \mathbb{R}$ messbar ist, sobald $f^{-1}(B) \in \mathcal{A}$ für alle Elemente B eines Erzeugersystems $\mathcal{T}$ von $\mathcal{B}(\mathbb{R})$ gilt, also: $\Big(f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{T}$, wobei $\mathcal{A}\langle\mathcal{T}\rangle \supset \mathcal{B}(\mathbb{R}) \Longrightarrow f$ ist messbar$\Big)$.

Proposition 3.23 *Sei $(X,\mathcal{A})$ ein Messraum und $f, g : X \longrightarrow \mathbb{R}$ messbare Abbildung. Dann sind, für alle $\alpha, \beta \in \mathbb{R}$, die Funktionen $\alpha f + \beta g$, $f \cdot g$, $|f|$, $\max(f,g)$ und $\min(f,g)$ messbar.*

Beweis:[5] Als Erstes zeigen wir folgende
Behauptung: *Die Abbildung $(f,g) : X \longrightarrow \mathbb{R}^2$, $x \longmapsto (f(x), g(x))$, ist messbar.*
Beweis: für $a_1, b_1, a_2, b_2 \in \mathbb{R}$ mit $a_i \le b_i \ \forall\ i = 1, 2$ gilt

$$(f,g)^{-1}(]a_1, b_1[\times]a_2, b_2[) = f^{-1}(]a_1, b_1[) \cap g^{-1}(]a_2, b_2[)$$

und $f^{-1}(]a_1, b_1[) \in \mathcal{A}$, $g^{-1}(]a_2, b_2[) \in \mathcal{A}$, somit $(f,g)^{-1}(]a_1, b_1[\times]a_2, b_2[) \in \mathcal{A}$. Nun wird $\mathcal{B}(\mathbb{R}^2)$ durch Teilmengen der Form $]a_1, b_1[\times]a_2, b_2[$ erzeugt, siehe Bemerkung 3.8.2. Daraus folgt $(f,g)^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{R}^2)$, d.h., (f,g) ist messbar. √
Nun betrachte die Abbildungen $\mathbb{R} \xrightarrow{m_\alpha} \mathbb{R}$, $x \longmapsto \alpha x$ und $\mathbb{R} \times \mathbb{R} \xrightarrow{+} \mathbb{R}$, $(x,y) \longmapsto x + y$. Beide Abbildungen sind stetig, insbesondere messbar, siehe Beispiel 3.21.2. Mit der Proposition 3.22 folgt, dass $\alpha f + \beta g$ als Verknüpfung messbarer Abbildungen wieder messbar ist. Analog sind $|\cdot| : \mathbb{R} \longrightarrow \mathbb{R}$ und $\mathbb{R} \times \mathbb{R} \xrightarrow{\cdot} \mathbb{R}$, $(x,y) \longmapsto x \cdot y$, stetig, daher messbar, somit sind $|f|$ und $f \cdot g$ messbar. Es bleibt, zu bemerken, dass die Abbildungen $\max : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto \max(x,y)$ und $\min : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto \min(x,y)$, stetig – und somit messbar – sind (argumentiere mit Folgen), woraus folgt, dass $\max(f,g)$ und $\min(f,g)$ messbar sind. □

Proposition 3.24 *Sei $f : X \longrightarrow \mathbb{R}^p$, $x \longmapsto (f_1(x), \ldots, f_p(x))$, wobei $f_j : X \longrightarrow \mathbb{R}$ und $(X,\mathcal{A})$ ein Messraum ist. Dann ist f genau dann messbar, wenn $f_1, \ldots, f_p$ messbar sind.*

Beweis:[6] Da $\mathcal{B}(\mathbb{R}^p)$ durch Teilmengen der Form $]a_1, b_1[\times \cdots \times]a_p, b_p[$ (mit $a_j \le b_j$ reell für alle $1 \le j \le p$) erzeugt wird (siehe Bemerkung 3.8.2), ist f genau dann messbar, wenn $f^{-1}(]a_1, b_1[\times \cdots \times]a_p, b_p[) \in \mathcal{A}$ für alle $a_j \le b_j$ gilt. Wegen $f^{-1}(]a_1, b_1[\times \cdots \times]a_j, b_j[) = \bigcap\limits_{j=1}^{p} f_j^{-1}(]a_j, b_j[)$ ist f messbar, sobald $f_1, \ldots, f_p$ messbar sind. Dies zeigt die eine Richtung. Für die andere Richtung ist lediglich zu bemerken, dass die Projektionen $\mathbb{R}^p \xrightarrow{\pi_j} \mathbb{R}$, $x = (x_1, \ldots, x_p) \longmapsto x_j$, stetig und somit messbar sind, insbesondere ist $f_j = \pi_j \circ f$ nach Proposition 3.22messbar, sobald f messbar ist, für alle $1 \le j \le p$. □

[5]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[6]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Proposition 3.25 (Punktweiser Limes einer Folge messbarer Funktionen)
Sei $(X, \mathcal{A})$ ein Messraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer Funktionen $f_n : X \longrightarrow \mathbb{R}$. Angenommen, die Folge $(f_n(x))_{n\in\mathbb{N}}$ konvergiere gegen ein Element $f(x) \in \mathbb{R}$ für alle $x \in X$. Dann ist $f : X \longrightarrow \mathbb{R}$ messbar.

Beweis[7]: Da $\mathcal{B}(\mathbb{R})$ durch Teilmengen von $\mathbb{R}$ der Form $[a, b]$, mit $a \leq b$ reell, erzeugt wird (Bemerkung 3.8.1), reicht es, zu zeigen, dass $f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{R})$ der Form $B = [a, b]$ gilt. Seien nun $a, b \in \mathbb{R}$ mit $a \leq b$.
Behauptung: *Es gilt* $f^{-1}([a, b]) = \bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\} \Big)$.
Beweis: Wir zeigen beide Inklusionen. Sei $x \in f^{-1}([a, b])$, d.h. $f(x) \in [a, b]$. Sei $m \in \mathbb{N}\setminus\{0\}$. Dann gilt $[a - \frac{1}{m}, b + \frac{1}{m}] \supset [a, b]$. Wegen $a - \frac{1}{m} < a$ und $b + \frac{1}{m} > b$ ist $[a - \frac{1}{m}, b + \frac{1}{m}]$ eine Umgebung von $f(x)$ in $\mathbb{R}$. Wegen $f_k(x) \underset{k\to\infty}{\longrightarrow} f(x)$ existiert dann ein $n \in \mathbb{N}$ so, dass $f_k(x) \in [a-\frac{1}{m}, b+\frac{1}{m}]$ für alle $k \geq n$. D.h., es gibt ein $n \in \mathbb{N}$ mit $x \in f_k^{-1}([a-\frac{1}{m}, b+\frac{1}{m}])$ für alle $k \geq n$, d.h., $x \in \bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a-\frac{1}{m}, b+\frac{1}{m}]) \big\}$. Da dies für alle $m \in \mathbb{N}\setminus\{0\}$ gilt, bekommen wir $x \in \bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a-\frac{1}{m}, b+\frac{1}{m}]) \big\} \Big)$. Dies zeigt die Inklusion $f^{-1}([a, b]) \subset \bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\} \Big)$. Sei umgekehrt $x \in \bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\} \Big)$. Für jedes $m \in \mathbb{N} \setminus \{0\}$ ist $x \in \bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\}$, insbesondere existiert ein $n \in \mathbb{N}$ mit $x \in \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\}$, was nichts anderes als $f_k(x) \in [a - \frac{1}{m}, b + \frac{1}{m}]$ für alle $k \geq n$ heißt. Wegen $[a - \frac{1}{m}, b + \frac{1}{m}]$ abgeschlossen in $\mathbb{R}$ folgt insbesondere $f(x) = \lim\limits_{k\to\infty} f_k(x) \in [a - \frac{1}{m}, b + \frac{1}{m}]$. Dies gilt für alle $m \in \mathbb{N} \setminus \{0\}$, insbesondere muss $f(x) \in \bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} [a - \frac{1}{m}, b + \frac{1}{m}] = [a, b]$ gelten, d.h., $x \in f^{-1}([a, b])$. Dies zeigt die Inklusion $\bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\} \Big) \subset f^{-1}([a, b])$ und schließt den Beweis der Behauptung ab. $\surd$
Da jedes f_k messbar ist, gehört für jedes $m \in \mathbb{N} \setminus \{0\}$ die Teilmenge $f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}])$ von X zu $\mathcal{A}$. Daraus folgt, dass $\bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}])$ als abzählbarer Durchschnitt von Elementen von $\mathcal{A}$ ebenfalls zu $\mathcal{A}$ gehört. Analog ist $\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\}$ als abzählbare Vereinigung von Elementen von $\mathcal{A}$ wieder ein Element von $\mathcal{A}$. Schließlich ist $\bigcap\limits_{m\in\mathbb{N}\setminus\{0\}} \Big(\bigcup\limits_{n\in\mathbb{N}} \big\{ \bigcap\limits_{k\geq n} f_k^{-1}([a - \frac{1}{m}, b + \frac{1}{m}]) \big\} \Big)$ als abzählbarer Durchschnitt von Elementen von $\mathcal{A}$ auch in $\mathcal{A}$. Aus der Behauptung folgt dann $f^{-1}([a, b]) \in \mathcal{A}$, was zu beweisen war. $\square$

Bezeichnungen 3.26

1. Wir bezeichnen mit $\overline{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ und versehen $\overline{\mathbb{R}}$ mit der Metrik $\overline{d} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \longrightarrow [0, \infty[$, $(x, y) \longmapsto |\varphi(x) - \varphi(y)|$, wobei $\varphi : \overline{\mathbb{R}} \longrightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$, $x \longmapsto \begin{cases} -\frac{\pi}{2} & \text{falls } x = -\infty \\ \arctan(x) & \text{falls } x \in \mathbb{R} \\ \frac{\pi}{2} & \text{falls } x = \infty \end{cases}$. Beachte, dass $\overline{d}$ wohl eine Metrik auf $\mathbb{R}$ definiert, deren Einschränkung auf $\mathbb{R}$ mit der Standardmetrik aber nicht übereinstimmt.

[7] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Dennoch definieren $\overline{d}_{|\mathbb{R}\times\mathbb{R}}$ und die Standardmetrik dieselben offenen Teilmengen (denn $\varphi_{|\mathbb{R}} : \mathbb{R} \longrightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$ ist ein Homöomorphismus mit $\varphi(t) \underset{t\to-\infty}{\longrightarrow} -\frac{\pi}{2}$ und $\varphi(t) \underset{t\to\infty}{\longrightarrow} \frac{\pi}{2}$).

2. Mit dieser Definition gilt $\mathcal{B}(\mathbb{R}) \subset \mathcal{B}(\overline{\mathbb{R}})$. Wenn wir $[0,\infty]$ als Teilmenge von $\overline{\mathbb{R}}$ ansehen, gilt auch $\mathcal{B}([0,\infty[) \subset \mathcal{B}([0,\infty]) \subset \mathcal{B}(\overline{\mathbb{R}})$, wobei $[0,\infty]$ die von $\overline{d}$ induzierte Metrik trägt.

3. Die natürliche Ordnungsrelation $\leq$ wird auf $\overline{\mathbb{R}}$ durch $-\infty \leq x \leq \infty$ für alle $x \in \overline{\mathbb{R}}$, $x \not\leq -\infty$ für alle $x \in]-\infty,\infty]$ und $x \not\geq \infty$ für alle $x \in [-\infty,\infty[$, fortgesetzt. Mit dieser Definition besitzt jede nichtleere Teilmenge $A \subset \overline{\mathbb{R}}$ ein Infimum und ein Supremum: setze $\inf(A) := \begin{cases} \inf(A) & \text{falls } A \subset \mathbb{R} \text{ und von unten beschränkt} \\ \infty & \text{falls } A = \{\infty\} \\ -\infty & \text{sonst.} \end{cases}$ Definiere $\sup(A) \in \overline{\mathbb{R}}$ analog. Insbesondere hat jede monoton steigende (bzw. fallende) Folge $(u_n)_{n\in\mathbb{N}}$ aus $\overline{\mathbb{R}}$ einen Grenzwert In $\overline{\mathbb{R}}$: es gilt $\lim\limits_{n\to\infty} u_n = \sup\limits_{n\in\mathbb{N}}\{u_n\} \in \overline{\mathbb{R}}$ (bzw. $\lim\limits_{n\to\infty} u_n = \inf\{u_n, n \in \mathbb{N}\} \in \overline{\mathbb{R}}$). Für eine reelle Folge stimmt dieser Begriff mit dem üblichen Konvergenzbegriff überein.

4. Sei eine Folge $(u_n)_{n\in\mathbb{N}}$ aus $\overline{\mathbb{R}}$ gegeben: dann ist die Folge durch $v_n := \sup\{u_k, k \geq n\} \in \overline{\mathbb{R}}$ definierte Folge $(v_n)_{n\in\mathbb{N}}$ monoton fallend, insbesondere existiert $\lim\limits_{n\to\infty} v_n \in \overline{\mathbb{R}}$. Wir bezeichnen diesen Grenzwert mit $\overline{\lim\limits_{n\to\infty}}\, u_n$. Analog ist die durch $w_n := \inf\{u_k, k \geq n\}$ definierte Folge $(w_n)_{n\in\mathbb{N}}$ monoton wachsend, insbesondere existiert der Grenzwert $\lim\limits_{n\to\infty} w_n \in \overline{\mathbb{R}}$, welchen wir mit $\underline{\lim\limits_{n\to\infty}}\, u_n$ bezeichnen. Mit den obigen Bezeichnungen konvergiert eine Folge $(u_n)_{n\in\mathbb{N}}$ im metrischen Raum $(\overline{\mathbb{R}}, \overline{d})$ g.d.w. $\underline{\lim\limits_{n\to\infty}}\, u_n = \overline{\lim\limits_{n\to\infty}}\, u_n$, ggf. ist $\lim\limits_{n\to\infty} u_n = \underline{\lim\limits_{n\to\infty}}\, u_n = \overline{\lim\limits_{n\to\infty}}\, u_n \in \mathbb{R}$.

Proposition 3.27 *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Abbildungen $f_n : X \longrightarrow \overline{\mathbb{R}}$, wobei $\overline{\mathbb{R}}$ die σ-Algebra $\mathcal{B}(\overline{\mathbb{R}})$ trägt. Dann gilt:*

i) Die Abbildungen $\inf\limits_{n\in\mathbb{N}}(f_n) : X \longrightarrow \overline{\mathbb{R}}$, $x \longmapsto \inf\limits_{n\in\mathbb{N}}\{f_n(x)\}$ und $\sup\limits_{n\in\mathbb{N}}(f_n) : X \longrightarrow \overline{\mathbb{R}}$, $x \longmapsto \sup\limits_{n\in\mathbb{N}}\{f_n(x)\}$, sind messbar.

ii) Die Abbildungen $\underline{\lim\limits_{n\to\infty}}\, f_n : X \longrightarrow \overline{\mathbb{R}}$, $x \longmapsto \underline{\lim\limits_{n\to\infty}}\, f_n(x)$ und $\overline{\lim\limits_{n\to\infty}}\, f_n : X \longrightarrow \overline{\mathbb{R}}$, $x \longmapsto \overline{\lim\limits_{n\to\infty}}\, f_n(x)$, sind messbar. Insbesondere ist $\lim\limits_{n\to\infty} f_n : X \longrightarrow \overline{\mathbb{R}}$, $x \longmapsto \lim\limits_{n\to\infty} f_n(x)$, messbar, sobald $\lim\limits_{n\to\infty} f_n(x) \in \overline{\mathbb{R}}$ für alle $x \in X$ existiert.

Beweis:[8] *i)* Für jedes $a \in \mathbb{R}$ gilt $(\inf\limits_{n\in\mathbb{N}}(f_n))^{-1}([a,\infty]) = \{x \in X, f_n(x) \geq a$ für alle $n \in \mathbb{N}\} = \bigcap\limits_{n\in\mathbb{N}} f_n^{-1}([a,\infty]) \in \mathcal{A}$, da alle f_n's messbar sind. Da $\mathcal{B}(\overline{\mathbb{R}})$ durch Elemente der Form $[a,\infty]$ erzeugt wird (*Übungsaufgabe*), muss nach dem Prinzip der guten Mengen $(\inf\limits_{n\in\mathbb{N}}(f_n))^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\overline{R})$ gelten. Für $\sup\limits_{n\in\mathbb{N}}(f_n)$ ist der Beweis analog

[8] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

(betrachte dann Erzeuger von $\mathcal{B}(\overline{R})$ der Form $[-\infty, a]$, mit $a \in \mathbb{R}$).
ii) Nach *i*) ist, für jedes $n \in \mathbb{N}$, die Abbildung $\inf_{k\geq n}(f_k) : X \longrightarrow \overline{\mathbb{R}}$ messbar; wiederum nach *i*) ist dann

$$\underline{\lim}_{n\to\infty} f_n = \sup_{n\in\mathbb{N}}(\inf_{k\geq n}(f_k)) : X \longrightarrow \overline{\mathbb{R}}$$

auch messbar. Für $\overline{\lim}_{n\to\infty} f_n$ ist der Beweis analog. □

3.2.2 Treppenfunktionen

Definition 3.28 *Sei $(X, \mathcal{A})$ ein Messraum. Eine* Treppenfunktion *auf X ist eine Abbildung $s : X \longrightarrow \mathbb{R}$, die messbar ist und nur endlich viele Werte hat.*

Proposition 3.29 *Sei $s : X \longrightarrow \mathbb{R}$ eine Treppenfunktion auf einem Messraum X. Dann existieren eindeutige $n \geq 1$, nichtleere $A_1, \ldots, A_n \in \mathcal{A}$ und reelle Zahlen $\alpha_1 < \cdots < \alpha_n$ so, dass $\bigcup_{i=1}^n A_i = X$, $A_i \cap A_j = \varnothing$ für alle $1 \leq i \neq j \leq n$ und $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$, wobei χ_{A_i} die charakteristische Funktion von A_i ist. Diese Darstellung heißt die* kanonische Darstellung *der Treppenfunktion s.*

Beweis[9]: Nach Voraussetzung ist $s(X)$ eine endliche Teilmenge von $\mathbb{R}$, insbesondere existieren reelle Zahlen $\alpha_1, \ldots, \alpha_n$ mit $s(X) = \{\alpha_1, \ldots, \alpha_n\}$. O.B.d.A. seien $\alpha_1, \ldots, \alpha_n$ monoton wachsend angeordnet, d.h., $\alpha_1 < \ldots < \alpha_n$. Setze $A_j := s^{-1}(\{\alpha_j\}) \subset X$ und bemerke, dass $A_j \in \mathcal{A}$ gilt, weil s nach Voraussetzung messbar ist; außerdem gilt $A_j \cap A_k = \varnothing$ wegen $\alpha_j \neq \alpha_k$, für alle $j \neq k$ und $X = s^{-1}(s(X)) = \bigcup_{j=1}^n s^{-1}(\{\alpha_j\}) = \bigcup_{j=1}^n A_j$. Sei $x \in X$, dann existiert ein eindeutiges $i \in \{1, \ldots, n\}$ mit $x \in A_i$. Es folgt

$$\begin{aligned} s(x) &= \alpha_i \\ &= \alpha_i \underbrace{\chi_{A_i}(x)}_{1} + \sum_{\substack{j=1 \\ j\neq i}}^n \alpha_j \underbrace{\chi_{A_j}(x)}_{0} \\ &= \sum_{j=1}^n \alpha_j \chi_{A_j}(x) \end{aligned}$$

und damit $s = \sum_{j=1}^n \alpha_j \chi_{A_j}$. Dies zeigt die Existenz einer solchen Darstellung.
Für die Eindeutigkeit nehmen wir an, dass es zwei solche Darstellungen gibt, also $s = \sum_{j=1}^n \alpha_j \chi_{A_j}$ und $s = \sum_{l=1}^m \beta_l \chi_{B_l}$, wobei $\alpha_1 < \ldots < \alpha_n$ und $\beta_1 < \ldots < \beta_m$ (sonst könnte man z.B. eines der A_i's in zwei Teilmengen teilen, $A_i = C_1 \cup C_2$ mit $C_1, C_2 \in \mathcal{A}$ und $C_1 \cap C_2 = \varnothing$, ggf. ist $\chi_{A_i} = \chi_{C_1} + \chi_{C_2}$). Dann gilt $s(X) = \{\alpha_1, \ldots, \alpha_n\}$ und $s(X) = \{\beta_1, \ldots, \beta_m\}$, insbesondere muss $m = n$ und $\alpha_j = \beta_j$ für alle $1 \leq j \leq n$ gelten. Mit der Darstellung $s = \sum_{j=1}^n \alpha_j \chi_{A_j}$ ist aber $A_j = s^{-1}(\{\alpha_j\})$ für alle $1 \leq j \leq n$, somit gilt $A_j = s^{-1}(\{\alpha_j\}) = s^{-1}(\{\beta_j\}) = B_j$ für alle $1 \leq j \leq n$. Dies beweist die Eindeutigkeit der Darstellung und die Proposition 3.29. □

Proposition 3.30 *Sei $f : X \longrightarrow [0, \infty]$ eine messbare Abbildung auf einem Messraum X. Dann existiert eine monoton wachsende Folge $(s_n)_{n\in\mathbb{N}}$ von Treppenfunktionen $s_n : X \longrightarrow [0, \infty[$ mit $s_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ für alle $x \in X$.*

[9] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Beweis:[10] Für jedes $n \in \mathbb{N}$ definiere $G_n := f^{-1}([n, \infty])$ und beachte, dass wegen f messbar $G_n \in \mathcal{A}$ gilt. Setze dann

$$s_n := \sum_{i=0}^{2^n - 1} \frac{i}{2^n} \chi_{f^{-1}([\frac{i}{2^n}, \frac{i+1}{2^n}[)} + n\chi_{G_n}, \qquad s_n : X \longrightarrow [0, \infty[.$$

Dann ist s_n eine nichtnegative Treppenfunktion (s_n ist messbar, da f messbar ist) auf X.
Behauptung 1: *Die Folge $(s_n)_{n \in \mathbb{N}}$ ist monoton wachsend, d.h., es gilt $s_n \leq s_{n+1}$ auf X, für alle $n \in \mathbb{N}$.*
Beweis: Sei $x \in X$ beliebig. Ist $x \notin G_n$, so existiert ein eindeutiges $i \in \{0, \ldots, 2^n - 1\}$ mit $x \in f^{-1}([\frac{i}{2^n}, \frac{i+1}{2^n}[)$. Nach Definition ist dann $s_n(x) = \frac{i}{2^n}$. Dann gibt es zwei Fälle: entweder ist $x \in f^{-1}([\frac{i}{2^n}, \frac{2i+1}{2^{n+1}}[) = f^{-1}([\frac{2i}{2^{n+1}}, \frac{2i+1}{2^{n+1}}[)$, ggf. ist $s_{n+1}(x) = \frac{i}{2^n} = s_n(x)$, oder $x \in f^{-1}([\frac{2i+1}{2^{n+1}}, \frac{i+1}{2^n}[) = f^{-1}([\frac{2i+1}{2^{n+1}}, \frac{2i+2}{2^{n+1}}[)$, ggf. ist $s_{n+1}(x) = \frac{2i+1}{2^{n+1}} > s_n(x)$. In beiden Fällen gilt wohl $s_{n+1}(x) \geq s_n(x)$.
Ist nun $x \in G_n$, so ist nach Definition $s_n(x) = n$. Da gibt es wieder zwei Fälle: entweder gilt $x \in G_{n+1}$, ggf. ist $s_{n+1}(x) = n + 1 > s_n(x)$, oder $x \notin G_{n+1}$, ggf. ist $x \in f^{-1}([\frac{i}{2^{n+1}}, \frac{i+1}{2^{n+1}}[)$, wobei jetzt $n \leq \frac{i}{2^{n+1}}$, insbesondere ist in diesem letzten Fall $s_{n+1}(x) = \frac{i}{2^{n+1}} \geq s_n(x)$. Wiederum gilt $s_{n+1}(x) \geq s_n(x)$. Dies beweist die Behauptung 1. $\surd$
Behauptung 2: *Für jedes $x \in X$ gilt $s_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$.*
Beweis: Sei $x \in X$ beliebig. Falls $f(x) = \infty$ gilt nach Definition $s_n(x) = n$ für alle $n \in \mathbb{N}$, somit $s_n(x) \underset{n\to\infty}{\longrightarrow} \infty = f(x)$ (im üblichen Sinne für reelle Folgen oder – es ist äquivalent – bezüglich der Metrik $\overline{d}$ aus den Bezeichnungen 3.26). Falls $f(x) < \infty$ gilt, für alle $n \geq [f(x)] + 1$ (wobei $[f(x)] \in \mathbb{N}$ die größte ganze Zahl bezeichnet, die kleiner gleich $f(x)$ ist): es gibt ein eindeutiges $i \in \{0, \ldots, 2^n - 1\}$ mit $x \in f^{-1}([\frac{i}{2^n}, \frac{i+1}{2^n}[)$, somit ist $f(x) \in [\frac{i}{2^n}, \frac{i+1}{2^n}[= [s_n(x), s_n(x) + \frac{1}{2^n}[$. Daraus folgt $|s_n(x) - f(x)| \leq \frac{1}{2^n}$ und damit $s_n(x) \underset{n\to\infty}{\longrightarrow} 0$, d.h., $s_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$. Dies zeigt die Behauptung 2. $\surd$
Dies beweist die Proposition 3.30. $\square$

Definition 3.31 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum.*

i) Für eine Treppenfunktion $s : X \longrightarrow [0, \infty[$ definieren wir

$$\int_X s d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i) \in [0, \infty],$$

wobei $s = \sum\limits_{i=1}^{n} \alpha_i \chi_{A_i}$ die kanonische Darstellung von s ist.

ii) Für eine messbare Abbildung $f : X \longrightarrow [0, \infty]$ definieren wir

$$\int_X f d\mu := \sup\left\{ \int_x s d\mu,\ s : X \longrightarrow [0, \infty[\text{ Treppenfunktion mit } s \leq f \right\} \in [0, \infty].$$

iii) Für $A \in \mathcal{A}$ definieren wir

$$\int_A f d\mu := \int_A f_{|A} d\mu_{|A} \in [0, \infty].$$

wobei A die σ-Algebra $\mathcal{A}_{|A} := \{B \cap A, B \in \mathcal{A}\}$ trägt und $\mu_{|A} : \mathcal{A}_{|A} \longrightarrow [0, \infty]$, $\mathcal{B} \longmapsto \mu(B)$ ist.

[10] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Bemerkungen 3.32

1. Für $f, g : X \longrightarrow [0,\infty]$ messbar mit $f \leq g$ gilt offensichtlich $\int_X f d\mu \leq \int_X g d\mu$.
2. Für $f : X \longrightarrow [0,\infty]$ messbar und $\alpha \in [0,\infty[$ gilt $\int_X \alpha f d\mu = \alpha \int_X f d\mu$.
3. Für $f : X \longrightarrow [0,\infty]$ messbar und $A \in \mathcal{A}$ gilt $\int_A f d\mu = \int_X f \chi_A d\mu$.
4. Für $f : X \longrightarrow [0,\infty]$ messbar und $A \in \mathcal{A}$ mit $\mu(A) = 0$ gilt $\int_A f d\mu = 0$.
5. Für $f : X \longrightarrow [0,\infty]$ messbar und $A, B \in \mathcal{A}$ mit $A \subset B$ gilt $\int_A f d\mu \leq \int_B f d\mu$.

3.2.3 Der Satz über monotone Konvergenz

Satz 3.33 (Satz über monotone Konvergenz) *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von messbaren Funktionen $f_n : X \longrightarrow [0,\infty]$ (d.h., $f_n(x) \leq f_{n+1}(x)$ für alle $n \in \mathbb{N}$ und $x \in X$). Dann gilt $\int_X \lim\limits_{n\to\infty} f_n d\mu = \lim\limits_{n\to\infty} \int_X f_n d\mu$.*

Beweis:[11] Beachte, dass nach Proposition 3.27 die Funktion $f := \lim\limits_{n\to\infty} f_n : X \longrightarrow [0,\infty]$ wohldefiniert und messbar ist. Wegen $f_n \leq f$ für alle $n \in \mathbb{N}$ gilt nach Bemerkung 3.32.1 auch $\int_X f_n d\mu \leq \int_X f d\mu$; da $(\int_X f_n d\mu)_{n\in\mathbb{N}}$ monoton wachsend ist, existiert $\lim\limits_{n\to\infty} \int_X f_n d\mu \in [0,\infty]$ und es gilt $\lim\limits_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu$. Wir wollen jetzt die andere Ungleichung zeigen. Sei $s : X \longrightarrow [0,\infty]$ eine Treppenfunktion mit $s \leq f$. Sei $\lambda \in]0,1[$. Betrachte, für alle $n \in \mathbb{N}$, die Teilmenge $A_n := \{x \in X, f_n(x) \geq \lambda s(x)\} \subset X$. Wegen f_n und λs messbar ist $A_n = (f_n - \lambda s)^{-1}([0,\infty]) \in \mathcal{A}$. Außerdem gilt $A_n \subset A_{n+1}$ für alle $n \in N$ (wegen $f_n \leq f_{n+1}$) und $\bigcup\limits_{n\in\mathbb{N}} A_n = X$, denn: für jedes $x \in X$ gilt entweder $f(x) > \lambda s(x)$ (ggf. ist wegen $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ auch $f_n(x) \geq \lambda s(x)$ für $n \in \mathbb{N}$ hinreichend groß) oder $f(x) = \lambda s(x)$ (ggf. gilt $f(x) = s(x) = \lambda s(x)$ und, wegen $\lambda \in]0,1[$, $f(x) = s(x) = 0$; dann ist $f_n(x) = 0 = \lambda s(x)$ für alle $n \in \mathbb{N}$).
Behauptung: *Es gilt $\int_{A_n} s d\mu \underset{n\to\infty}{\longrightarrow} \int_X s d\mu$.*
Beweis: Nach Proposition 3.12.2.*ii*) reicht es, zu zeigen, dass $\mathcal{A} \longrightarrow [0,\infty]$, $A \longmapsto \int_A s d\mu$ ein Maß ist. Offenbar ist $\int_\varnothing s d\mu = 0$ (siehe Bemerkung 3.32.4). Sei $(A_n)_{n\in\mathbb{N}}$ eine disjunkte Familie von Elementen von $\mathcal{A}$. Stelle s in seine kanonische Darstellung $s = \sum\limits_{i=1}^{k} \alpha_i \chi_{B_i}$ dar, wobei $B_i \in \mathcal{A}$, $\alpha_i \in [0,\infty[$ und $\bigcup\limits_{i=1}^{k} B_i = X$, $B_i \cap B_j = \varnothing$ für alle $i \neq j$. Dann ist $s_{|\bigcup\limits_{n\in\mathbb{N}} A_i} = \sum\limits_{i=1}^{k} \alpha_i \chi_{B_i \cap (\bigcup\limits_{n\in\mathbb{N}} A_n)}$, somit ist $\int_{\bigcup\limits_{n\in\mathbb{N}} A_n} s d\mu = \sum\limits_{i=1}^{k} \alpha_i \mu(B_i \cap (\bigcup\limits_{n\in\mathbb{N}} A_n))$. Da aber $B_i \cap (\bigcup\limits_{n\in\mathbb{N}} A_n) = \bigcup\limits_{n\in\mathbb{N}} (B_i \cap A_n)$ mit $B_i \cap A_n \in \mathcal{A}$ und $(B_i \cap A_n) \cap (B_j \cap A_n) = \varnothing$ für alle $i \neq j$, gilt $\mu(B_i \cap (\bigcup\limits_{n\in\mathbb{N}} A_n)) = \sum\limits_{n=0}^{\infty} \mu(B_i \cap A_n)$, für alle $1 \leq i \leq k$. Da alle $\mu(B_i \cap A_n)$ nichtnegativ sind, kann folgende Umordnung der Summanden durchgeführt werden:

$$\sum_{i=1}^{k} \alpha_i \mu(B_i \cap (\bigcup_{n\in\mathbb{N}} A_n)) = \sum_{i=1}^{k} \alpha_i \sum_{n=0}^{\infty} \mu(B_i \cap A_n) = \sum_{n=0}^{\infty} \sum_{i=1}^{k} \alpha_i \mu(B_i \cap A_n).$$

[11] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Nach Definition gilt aber $\sum_{i=1}^{k} \alpha_i \mu(B_i \cap A_n) = \int_{A_n} s d\mu$, insbesondere ist $\sum_{i=1}^{k} \alpha_i \mu(B_i \cap (\bigcup_{n\in\mathbb{N}} A_n)) = \sum_{n=0}^{\infty} \int_{A_n} s d\mu$. Dies zeigt $\int_{\bigcup_{n\in\mathbb{N}} A_n} s d\mu = \sum_{n=0}^{\infty} \int_{A_n} s d\mu$ und die Behauptung. ✓

Nun ist $s_{|A_n}$ eine Treppenfunktion mit $s_{|A_n} \leq \frac{1}{\lambda} f_{n|A_n}$, insbesondere gilt nach Definition des Integrals $\int_{A_n} s_{|A_n} d\mu_{|A_n} = \int_{A_n} s d\mu \leq \int_{A_n} \frac{1}{\lambda} \cdot f_n d\mu \leq \frac{1}{\lambda} \int_X f_n d\mu$. Lässt man n gegen ∞ laufen, so bekommt man nach der Behauptung

$$\int_X s d\mu = \lim_{n\to\infty} \int_{A_n} s d\mu \leq \frac{1}{\lambda} \lim_{n\to\infty} \int_X f_n d\mu.$$

Dies gilt für alle Treppenfunktionen $s : X \longrightarrow [0, \infty[$ mit $s \leq f$. Bilden wir das Supremum auf allen Treppenfunktionen, so bekommen wir $\int_X f d\mu \leq \frac{1}{\lambda} \lim_{n\to\infty} \int_X f_n d\mu$. Lässt man schließlich λ gegen 1 laufen, so erhalten wir $\int_X f d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu$, was zu beweisen war.
□

Korollar 3.34 *Seien $f, g : X \longrightarrow [0, \infty]$ messbare Funktionen auf einem Maßraum $(X, \mathcal{A}, \mu)$. Dann gilt $\int_X f + g d\mu = \int_X f d\mu + \int_X g d\mu$.*

Beweis:[12] Nach Proposition 3.30 existieren monoton wachsende Folgen $(s_n)_{n\in\mathbb{N}}$ und $(\tau_n)_{n\in\mathbb{N}}$ von nichtnegativen Treppenfunktionen auf X mit $s_n(x) \longrightarrow f(x)$ und $\tau_n(x) \longrightarrow g(x)$, für alle $x \in X$. Da $(s_n + \tau_n)_{n\in\mathbb{N}}$ ebenfalls eine monoton wachsende Folge von Treppenfunktionen ist mit $s_n(x) + \tau_n(x) \underset{n\to\infty}{\longrightarrow} (f + g)(x)$ für alle $x \in X$, folgt aus dem Satz 3.33 (mehrfach angewendet):

$$\begin{aligned}
\int_X f + g d\mu &= \lim_{n\to\infty} \int_X s_n + \tau_n d\mu \\
&= \lim_{n\to\infty} \left(\int_X s_n d\mu + \int_X \tau_n d\mu\right) \text{ mit } \begin{array}{ccc} \int_X s_n d\mu & \underset{n\to\infty}{\longrightarrow} & \int_X f d\mu \text{ und} \\ \int_X \tau_n d\mu & \underset{n\to\infty}{\longrightarrow} & \int_X g d\mu \end{array} \\
&= \int_X f d\mu + \int_X g d\mu.
\end{aligned}$$

Dies beweist die gesuchte Identität. Dabei haben wir die Tatsachen verwendet, dass für zwei Treppenfunktionen $s, s' : X \longrightarrow [0, \infty[$ die Summe $s + s' : X \longrightarrow [0, \infty[$ ebenfalls eine Treppenfunktion ist mit $\int_X s + s' d\mu = \int_X s d\mu + \int_X s' d\mu$ (*Übungsaufgabe*). □

Korollar 3.35 *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Abbildungen $f_n : X \longrightarrow [0, \infty]$ auf einem Maßraum $(X, \mathcal{A}, \mu)$. Dann gilt $\sum_{n=0}^{\infty} \int_X f_n d\mu = \int_X \sum_{n=0}^{\infty} f_n d\mu \in [0, \infty]$.*

Beweis:[13] Betrachte die Folge $g_n := \sum_{k=0}^{n} f_k : X \longrightarrow [0, \infty]$. Dann ist $(g_n)_{n\in\mathbb{N}}$ eine monoton wachsende (wegen $f_n \geq 0$ für alle n) Folge von messbaren Abbildungen auf X. Satz 3.33 liefert die Identität. □

[12] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[13] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Korollar 3.36 (Fatou-Lemma) *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Abbildungen $f_n : X \longrightarrow [0,\infty]$ auf einem Maßraum $(X,\mathcal{A},\mu)$. Dann gilt*

$$\int_X \varliminf_{n\to\infty} f_n d\mu \leq \varliminf_{n\to\infty} \int_X f_n d\mu.$$

Beweis:[14] Setze $g_n := \inf\{f_k, k \geq n\} : X \longrightarrow [0,\infty]$. Dann ist $(g_n)_{n\in\mathbb{N}}$ nach Proposition 3.27 eine Folge von messbaren Funktionen. Außerdem ist $(g_n)_{n\in\mathbb{N}}$ nach Definition monoton wachsend. Satz 3.33 liefert $\int_X \lim_{n\to\infty} g_n d\mu = \lim_{n\to\infty} \int_X g_n d\mu$. Nach Definition ist aber $\lim_{n\to\infty} g_n = \varliminf_{n\to\infty} f_n$ und es gilt $g_n \leq f_n$ für alle $n \in \mathbb{N}$, insbesondere ist $\int_X g_n d\mu \leq \int_X f_n d\mu$ für alle $n \in \mathbb{N}$. Daraus folgt

$$\int_X \varliminf_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_X g_n d\mu = \varliminf_{n\to\infty} \int_X g_n d\mu \leq \varliminf_{n\to\infty} \int_X f_n d\mu,$$

was zu beweisen war. □

Bemerkung 3.37 I.A. gilt die Gleichheit in der Ungleichung von Korollar 3.36 nicht: z.B. betrachte $X := \mathbb{R}$, $\mathcal{A} := \mathcal{L}$, $\mu := \lambda$ (Lebesgue-Maß) und $f_n := \chi_{[n,\infty[} : X \longrightarrow \{0,1\}$. Wegen $[n,\infty[\in \mathcal{B}(\mathbb{R})$ ist f_n messbar. Wegen $n \underset{n\to\infty}{\longrightarrow} \infty$ (!) gilt $f_n(x) \underset{n\to\infty}{\longrightarrow} 0$, für alle $x \in X$. Insbesondere ist $\varliminf_{n\to\infty} f_n = 0$ und somit $\int_X \varliminf_{n\to\infty} f_n d\mu = 0$. Dennoch ist $\int_X f_n d\mu = \mu([n,\infty[) = \infty$ für alle $n \in \mathbb{N}$, insbesondere ist $\int_X \varliminf_{n\to\infty} f_n d\mu = 0 < \infty = \varliminf_{n\to\infty} \int_X f_n d\mu$.

Proposition 3.38 *Sei $f : X \longrightarrow [0,\infty]$ eine messbare Funktion auf einem Maßraum $(X,\mathcal{A},\mu)$. Dann definiert $\nu : \mathcal{A} \longrightarrow [0,\infty]$, $A \longmapsto \int_A f d\mu$, ein Maß auf $\mathcal{A}$. Außerdem gilt $\int_X g d\nu = \int_X fg d\mu$ für jede messbare Funktion $g : X \longrightarrow [0,\infty]$.*

Beweis:[15] Nach Definition des Integrals einer nichtnegativen messbaren Abbildung gilt $\nu(\varnothing) = \int_\varnothing f d\mu = 0$. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von Elementen von $\mathcal{A}$ mit $A_n \cap A_m = \varnothing$ für alle $n \neq m$. Dann gilt

$$\begin{aligned}
\nu(\bigcup_{n\in\mathbb{N}} A_n) &= \int_{\bigcup_{n\in\mathbb{N}} A_n} f d\mu \\
&= \int_X f\chi_{\cup_{n\in\mathbb{N}} A_n} d\mu \qquad \text{nach Bemerkung 3.32.3} \\
&= \int_X f \cdot \sum_{n=0}^{\infty} \chi_{A_n} d\mu \\
&= \int_X \sum_{n=0}^{\infty} f\chi_{A_n} d\mu \\
&= \sum_{n=0}^{\infty} \int_X f\chi_{A_n} d\mu \qquad \text{nach Korollar 3.35} \\
&= \sum_{n=0}^{\infty} \int_{A_n} f d\mu \qquad \text{nach Bemerkung 3.32.3} \\
&= \sum_{n=0}^{\infty} \nu(A_n).
\end{aligned}$$

[14] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.
[15] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Dies zeigt, dass $\nu : \mathcal{A} \longrightarrow [0, \infty]$ ein Maß ist.
Die letzte Aussage zeige man erstens für eine nichtnegative Treppenfunktion $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ (mit $\alpha_1, \ldots, \alpha_n \in [0, \infty[$). Nämlich gilt

$$\begin{aligned}
\int_X s d\nu &= \int_X \sum_{i=1}^{n} \alpha_i \chi_{A_i} d\nu \\
&= \sum_{i=1}^{n} \alpha_i \int_X \chi_{A_i} d\nu \\
&= \sum_{i=1}^{n} \alpha_i \nu(A_i) \qquad \text{nach Definition des Integrals} \\
&= \sum_{i=1}^{n} \alpha_i \int_{A_i} f d\mu \qquad \text{nach Definition von } \nu \\
&= \sum_{i=1}^{n} \alpha_i \int_X f \chi_{A_i} d\mu \qquad \text{nach Bemerkung 3.32.3} \\
&= \int_X \sum_{i=1}^{n} \alpha_i f \chi_{A_i} d\mu \\
&= \int_X f s d\mu. \qquad (3.1)
\end{aligned}$$

Sei nun $g : X \longrightarrow [0, \infty]$ eine beliebige messbare Abbildung. Dann existiert nach Proposition 3.30 eine monoton wachsende Folge $(s_n)_{n \in \mathbb{N}}$ von Treppenfunktionen $s_n : X \longrightarrow [0, \infty[$ mit $s_n(x) \underset{n \to \infty}{\longrightarrow} g(x)$ für alle $x \in X$. Beachte, dass die Folge $(f s_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge von messbaren Abbildungen $X \longrightarrow [0, \infty]$ ist, die punktweise gegen die messbare Abbildung $fg : X \longrightarrow [0, \infty]$ konvergiert. Aus dem Satz über monotone Konvergenz und dem letzten Schritt folgt dann

$$\begin{aligned}
\int_X g d\nu &= \int_X \lim_{n \to \infty} s_n d\nu \\
&= \lim_{n \to \infty} \int_X s_n d\nu \qquad \text{nach Satz 3.33} \\
&\overset{(3.1)}{=} \lim_{n \to \infty} \int_X f s_n d\mu \\
&= \int_X \lim_{n \to \infty} f s_n d\mu \qquad \text{nach Satz 3.33} \\
&= \int_X f g d\mu.
\end{aligned}$$

Dies beweist die letzte Identität und die Proposition 3.38. □

3.2.4 Das Integral reell- oder vektorwertiger Funktionen

Definition 3.39 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Eine Abbildung $f : X \longrightarrow \mathbb{R}$ heißt genau dann* μ-integrierbar*, wenn f messbar ist und $\int_X |f| d\mu < \infty$ ist.*

Beachte dazu, dass $|f| = |\cdot| \circ f : X \longrightarrow [0, \infty[$ als Verknüpfung zweier messbarer Funktionen wieder messbar ist.

Bezeichnungen 3.40 Für $f : X \longrightarrow \mathbb{R}$ bezeichnen $f_+ := \max(f, 0) : X \longrightarrow [0, \infty[$ und $f_- := \max(-f, 0) : X \longrightarrow [0, \infty[$. Dann ist $f = f_+ - f_-$ und $|f| = f_+ + f_-$. Beachte: ist f messbar, so sind nach Proposition 3.23 f_+ und f_- messbar. Desweiteren gilt wegen $f_+ \leq |f|$ und $f_- \leq |f|$ auch $\int_X f_+ d\mu \leq \int_X |f| d\mu$ und $\int_X f_- d\mu \leq \int_X f d\mu$. Insbesondere gilt $\int_X f_\pm d\mu < \infty$, falls $\int_X |f| d\mu < \infty$ ist.

Definition 3.41 *Sei $f : X \longrightarrow \mathbb{R}$ eine μ-integrierbare Funktion auf einem Maßraum $(X, \mathcal{A}, \mu)$. Das* Integral *von f auf X wird durch*

$$\int_X f d\mu := \int_X f_+ d\mu - \int_X f_- d\mu = \int_X \max(f, 0) d\mu - \int_X \max(-f, 0) d\mu \in \mathbb{R}$$

definiert.

Proposition 3.42 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum.*

i) Für alle μ-integrierbaren Funktionen $f, g : X \longrightarrow \mathbb{R}$ und $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ integrierbar mit $\int_X \alpha f + \beta g d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$.

ii) Für jede μ-integrierbare Funktion $f : X \longrightarrow \mathbb{R}$ gilt $\left| \int_X f d\mu \right| \leq \int_X |f| d\mu$.

Beweis:[16] *i)* Nach Proposition 3.23 ist $\alpha f + \beta g$ messbar mit $|\alpha f + \beta g| \leq |\alpha| \cdot |f| + |\beta| \cdot |g|$, insbesondere $\int_X |\alpha f + |g| d\mu \leq \int_X |\alpha| \cdot |f| + |\beta| \cdot |g| d\mu = |\alpha| \cdot \int_X |f| d\mu + |\beta| \cdot \int_X |g| d\mu < \infty$. Daraus folgt, dass $\alpha f + \beta g$ μ-integrierbar ist.
ii) Nach Definition gilt

$$|\int_X f d\mu| = |\int_X f_+ d\mu - \int_X f_- d\mu| \leq \int_X f_+ d\mu + \int_X f_- d\mu = \int_X f_+ + f_- d\mu = \int_X |f| d\mu.$$

Dies beweist die Proposition. □

Definition 3.43 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $p \in \mathbb{N} \setminus \{0\}$.*

*i) Eine Abbildung $f : X \longrightarrow \mathbb{R}^p$, $x \longmapsto (f_1(x), \ldots, f_p(x))$, heißt genau dann μ-*integrierbar, *wenn $f_j : X \longrightarrow \mathbb{R}$ μ-integrierbar ist für alle $1 \leq j \leq p$. Ggf. definiert man $\int_X f d\mu := (\int_X f_1 d\mu, \ldots, \int_X f_p d\mu) \in \mathbb{R}^p$.*

ii) Analog ist eine Abbildung $f : X \longrightarrow \mathbb{C}$, $x \longmapsto \mathrm{Re}(f(x)) + i\mathrm{Im}(f(x))$, genau dann μ-integrierbar, wenn $\mathrm{Re}(f) : X \longrightarrow \mathbb{R}$ und $\mathrm{Im}(f) : X \longrightarrow \mathbb{R}$ μ-integrierbar sind. Ggf. definiert man $\int_X f d\mu := \int_X \mathrm{Re}(f) d\mu + i \int_X \mathrm{Im}(f) d\mu \in \mathbb{C}$.

Proposition 3.44 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $p \in \mathbb{N} \setminus \{0\}$. Dann gilt*

1. *Eine messbare Abbildung $f : X \longrightarrow \mathbb{R}^p$ ist genau dann μ-integrierbar, wenn $\int_X |f| d\mu < \infty$, d.h., wenn $|f| : X \longrightarrow \mathbb{R}$, $x \longmapsto |f(x)|$, μ-integrierbar ist. Analog ist eine messbare Abbildung $f : X \longrightarrow \mathbb{C}$ genau dann μ-integrierbar, wenn $\int_X |f| d\mu < \infty$ gilt.*

[16] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

2. *Für alle μ-integrierbaren Funktionen $f, g : X \longrightarrow \mathbb{R}^p$ und $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g : X \longrightarrow \mathbb{R}^p$ μ-integrierbar mit $\int_X \alpha f + \beta g d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$. Analog für $f, g : X \longrightarrow \mathbb{C}$.*

3. *Für jede μ-integrierbare Funktion $f : X \longrightarrow \mathbb{R}^p$ (oder $f : X \longrightarrow \mathbb{C}$) gilt $\left|\int_X f d\mu\right| \leq \int_X |f| d\mu \in [0, \infty[$.*

Beweis:[17] Wir beweisen beide Aussagen für $f : X \longrightarrow \mathbb{R}^p$. Der Fall $f : X \longrightarrow \mathbb{C}$ lässt sich daraus herleiten.

1. Es ist lediglich zu beachten, dass für alle $1 \leq j \leq p$ folgendes gilt:

$$|f_j| \leq |f| \leq \sqrt{p} \cdot \max_{1 \leq i \leq p}(|f_i|) \leq \sqrt{p} \cdot \left(\sum_{i=1}^{p} |f_i| \right).$$

2. Für jedes $j \in \{1, \ldots, p\}$ ist $\alpha f_j + \mu g_j : X \longrightarrow \mathbb{R}$ μ-integrierbar. Die Identität folgt direkt aus der entsprechenden Eigenschaft für reellwertige Funktionen, siehe Proposition 3.42.

3. Der Beweis geht analog wie der von Proposition 1.21, wobei i.W. lediglich die Linearität des Integrals genutzt wird. $\square$

3.2.5 Der Satz über majorisierte Konvergenz

Satz 3.45 (Satz über majorisierte Konvergenz) *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_{n \in \mathbb{N}}$ eine Folge von messbaren Funktionen $f_n : X \longrightarrow \mathbb{R}$. Wir nehmen an, dass*

a) für jedes $x \in X$ der Grenzwert $\lim\limits_{n \to \infty} f_n(x) \in \mathbb{R}$ existiert,

b) eine μ-integrierbare Funktion $g : X \longrightarrow \mathbb{R}$ existiert mit $|f_n| \leq g$ für alle $n \in \mathbb{N}$.

Dann sind die Funktionen f_n und $f : X \longrightarrow \mathbb{R}$, $x \longmapsto \lim\limits_{n \to \infty} f_n(x)$, μ-integrierbar und es gilt $\int_X |f_n - f| d\mu \underset{n \to \infty}{\longrightarrow} 0$, insbesondere $\int_X \lim\limits_{n \to \infty} f_n d\mu = \lim\limits_{n \to \infty} \int_X f_n d\mu$.

Beweis[18]: Sei $n \in \mathbb{N}$ beliebig. Wegen $|f_n| \leq g$ gilt $\int_X |f_n| d\mu \leq \int_X g d\mu$; da g μ-integrierbar ist, gilt $\int_X g d\mu < \infty$ und somit auch $\int_X |f_n| d\mu < \infty$, d.h., f_n ist μ-integrierbar. Beachte auch, dass für jedes $x \in X$ die Ungleichung $|f_n(x)| \leq g(x)$ durch $n \longrightarrow \infty$ die Ungleichung $|f(x)| \leq g(x)$ liefert, d.h., es gilt $|f| \leq g$ auf X. Insbesondere ist f μ-integrierbar. Außerdem gilt $|f_n - f| \leq |f_n| + |f| \leq 2g$ und somit $0 \leq 2g - |f_n - f| \leq 2g$ auf X. Bemerke auch, dass wegen $\int_X g d\mu < \infty$ die Funktion $|f_n - f|$ ebenfalls μ-integrierbar ist, für alle $n \in \mathbb{N}$ und dass $\overline{\lim}_{n \to \infty} \int_X |f_n - f| d\mu \leq 2 \int_X g d\mu < \infty$. Nun wenden wir das Fatou-Lemma auf die Folge $(2g - |f_n - f|)_{n \in \mathbb{N}}$ von nichtnegativen messbaren Funktionen an: wegen $|f_n - f|(x) \underset{n \to \infty}{\longrightarrow} 0$ für alle $x \in X$ gilt $\underline{\lim}_{n \to \infty}(2g - |f_n - f|) = 2g + \underline{\lim}_{n \to \infty}(-|f_n - f|) = 2g - \overline{\lim}_{n \to \infty} |f_n - f| = 2g - \lim_{n \to \infty} |f_n - f| = 2g$ auf X; Korollar 3.36 liefert $\int_X \underline{\lim}_{n \to \infty}(2g - |f_n - f|) d\mu \leq \underline{\lim}_{n \to \infty} \int_X (2g - |f_n - f|) d\mu$, d.h.,

$$\int_X 2g d\mu \leq \underline{\lim}_{n \to \infty} \left(\int_X 2g d\mu - \int_X |f_n - f| d\mu\right) = \int_X 2g d\mu - \overline{\lim}_{n \to \infty} \int_X |f_n - f| d\mu,$$

[17] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

[18] Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

so dass $\overline{\lim}_{n\to\infty} \int_X |f_n - f| d\mu \leq 0$. Wegen $\int_X |f_n - f| d\mu \geq 0$ für alle n muss dann $\underline{\lim}_{n\to\infty} \int_X |f_n - f| d\mu \geq 0$ und somit auch $\lim_{n\to\infty} \int_X |f_n - f| d\mu = 0$ gelten. Proposition 3.42.*ii*) liefert schließlich $|\int_X (f_n - f) d\mu| \leq \int_X |f_n - f| d\mu$; wegen $\int_X |f_n - f| d\mu \underset{n\to\infty}{\longrightarrow} 0$ folgt $|\int_X (f_n - f) d\mu| \underset{n\to\infty}{\longrightarrow} 0$, d.h., $\int_X f_n d\mu \underset{n\to\infty}{\longrightarrow} \int_X f d\mu$. □

Korollar 3.46 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Funktionen $f_n : X \longrightarrow [0, \infty[$. Angenommen, die Folge $(f_n)_{n\in\mathbb{N}}$ sei monoton fallend (d.h., $f_{n+1}(x) \leq f_n(x)$ für alle $n \in \mathbb{N}$ und $x \in X$) und $\int_X f_0 d\mu < \infty$. Dann ist jedes f_n μ-integrierbar, $\lim\limits_{n\to\infty} f_n : X \longrightarrow [0, \infty[$ ist μ-integrierbar und es gilt $\int_X \lim\limits_{n\to\infty} f_n d\mu = \lim\limits_{n\to\infty} \int_X f_n d\mu$.*

Beweis[19]: Betrachte die durch $g_n := f_0 - f_n$ (wohldefinierte) Folge $(g_n)_{n\in\mathbb{N}}$. Nach Voraussetzung ist für alle $n \in \mathbb{N}$ die Funktion $g_n : X \longrightarrow [0, \infty[$ eine nichtnegative messbare Funktion mit $g_n \leq g_{n+1}$ (wegen $f_n \geq f_{n+1}$). Desweiteren konvergiert $(g_n)_{n\in\mathbb{N}}$ punktweise gegen $f_0 - f$, wobei $f : X \longrightarrow [0, \infty[$, $x \longmapsto \lim\limits_{n\to\infty} f_n(x)$ (dieser Grenzwert existiert wegen $(f_n(x))_{n\in\mathbb{N}}$ monoton fallend und nichtnegativ). Bemerke, dass f als punktweiser Limes einer Folge von messbaren Funktionen auch messbar ist, siehe Proposition 3.25. Wegen $f_0 - f_n \geq 0$ für alle $n \in \mathbb{N}$ gilt auch $f_0 - f \geq 0$, d.h., $f_0 \geq f$ auf X. Die Identität $\int_X f_0 d\mu = \int_X f_0 - f d\mu + \int_X f d\mu$ zusammen mit $\int_X f_0 d\mu < \infty$ liefert $\int_X f_0 - f d\mu < \infty$ sowie $\int_X f d\mu < \infty$, d.h., $f_0 - f$ und f sind μ-integrierbar; analog sind $f_0 - f_n$ und f_n μ-integrierbar für alle $n \in \mathbb{N}$. Der Satz über monotone Konvergenz (Satz 3.33) impliziert dann

$$\lim_{n\to\infty} \int_X g_n d\mu = \int_X f_0 - f d\mu,$$

d.h. $\int_X f_0 d\mu - \lim_{n\to\infty} \int_X f_n d\mu = \int_X f_0 d\mu - \int_X f d\mu$. Daraus folgt $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$, was zu beweisen war. □

Definition 3.47 *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Man sagt, dass eine gewisse Eigenschaft (E)* μ-fast überall *(kurz: μ-f.ü.) gilt, wenn (E) auf $N^c \subset X$ für eine μ-Nullmenge N erfüllt ist.*

Z.B. gilt $f = g$ μ-f.ü. für zwei Funktionen $f, g : X \longrightarrow \mathbb{R}$, wenn eine μ-Nullmenge $N \subset X$ existiert mit $f(x) = g(x)$ für alle $x \notin N$. Ein anderes Beispiel ist: sei $(f_n)_n$ eine Folge von Funktionen $f_n : X \longrightarrow \mathbb{R}$ und $f : X \longrightarrow \mathbb{R}$ eine weitere Funktion, per Definition konvergiert $(f_n)_n$ punktweise μ-f.ü. gegen f, wenn eine μ-Nullmenge $N \subset X$ so existiert, dass $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ für alle $x \notin N$.

Korollar 3.48 (Satz über majorisierte Konvergenz "μ-f.ü.") *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge von messbaren Abbildungen $f_n : X \longrightarrow \mathbb{R}$. Wir nehmen an, dass*

a) der punktweise Limes $\lim\limits_{n\to\infty} f_n$ μ-f.ü. existiert,

b) eine μ-integrierbare Funktion $g : X \longrightarrow \mathbb{R}$ existiert mit

$$|f_n| \leq g \ \ \mu\text{-f.ü.}, \quad \text{und für alle } n \in \mathbb{N}.$$

[19]Dieser Beweis wurde in der Vorlesung nicht durchgeführt.

Dann ist f_n μ-integrierbar für alle $n \in \mathbb{N}$ und es existiert eine μ-integrierbare Funktion $f : X \longrightarrow \mathbb{R}$ mit $\lim_{n\to\infty} f_n = f$ μ-f.ü. Außerdem gilt $\int_X |f_n - f| d\mu \underset{n\to\infty}{\longrightarrow} 0$ und insbesondere $\int_X f_n d\mu \underset{n\to\infty}{\longrightarrow} \int_X f d\mu$.

Der Beweis von Korollar 3.48 beruht auf Satz 3.45 und auf der Tatsache, dass $\int_X f d\mu = \int_X g d\mu$ gilt, falls f und g μ-integrierbare Funktionen sind mit $f = g$ μ-f.ü.

Bemerkung 3.49 Satz 3.45 und Korollar 3.48 gelten ebenfalls für Funktionen mit Werten im $\mathbb{R}^p$ bzw. in $\mathbb{C}$: man ersetze lediglich den Betrag durch die Norm in der Bedingung *b)*.

3.2.6 Vergleich mit dem Riemann-Integral

Erinnerung: Eine beschränkte Funktion $f : [a, b] \longrightarrow \mathbb{R}$ heißt genau dann *Riemann-integrierbar*, wenn für jedes $\varepsilon > 0$ eine Unterteilung $a = t_0 < t_1 < \cdots < t_k = b$ des Intervalls $[a, b]$ so existiert, dass

$$(0 \leq) \sum_{i=0}^{k-1} (t_{i+1} - t_i)\left(\sup_{[t_i,t_{i+1}]} (f) - \inf_{[t_i,t_{i+1}]} (f)\right) < \varepsilon$$

Ggf. definiert man das Riemann-Integral von f durch

$$\begin{aligned} \int_a^b f(x)dx &:= \sup_{\substack{a=t_0<\cdots<t_k=b \\ k\in\mathbb{N}}} \left(\sum_{i=0}^{k-1} (t_{i+1} - t_i) \cdot \inf_{[t_i,t_{i+1}]} (f) \right) \\ &= \inf_{\substack{a=t_0<\cdots<t_k=b \\ k\in\mathbb{N}}} \left(\sum_{i=0}^{k-1} (t_{i+1} - t_i) \cdot \sup_{[t_i,t_{i+1}]} (f) \right). \end{aligned}$$

Satz 3.50 *Sei $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt und Riemann-integrierbar. Dann ist f auch λ-integrierbar (wobei λ das Lebesgue-Maß auf $\mathbb{R}$ ist) auf $[a, b]$ und es gilt*

$$\int_{[a,b]} f d\lambda = \int_a^b f(x)dx.$$

Beweis: Nach Voraussetzung existiert, für jedes $n \in \mathbb{N} \setminus \{0\}$, eine Unterteilung $a = t_{0,n} < \ldots < t_{k,n} = b$ des Intervalles $[a, b]$ so, dass

$$\sum_{i=0}^{k-1} \left(\sup_{[t_{i,n},t_{i+1,n}]} (f) - \inf_{[t_{i,n},t_{i+1,n}]} (f)\right)(t_{i+1,n} - t_{i,n}) \leq \frac{1}{n}$$

gilt. Hierbei hängt k auch von n ab. Wir setzen

$$\begin{aligned} s_n &:= \sum_{i=0}^{k-1} \inf_{[t_{i,n},t_{i+1,n}]} (f) \cdot \chi_{[t_{i,n},t_{i+1,n}[} + f(b) \cdot \chi_{\{b\}} \\ S_n &:= \sum_{i=0}^{k-1} \sup_{[t_{i,n},t_{i+1,n}]} (f) \cdot \chi_{[t_{i,n},t_{i+1,n}[} + f(b) \cdot \chi_{\{b\}}, \end{aligned}$$

für alle $n \in \mathbb{N} \setminus \{0\}$. Dann sind s_n und S_n (wegen f beschränkt) wohldefinierte Treppenfunktionen auf $[a, b]$ mit $s_n \leq f \leq S_n$ auf $[a, b]$ für alle $n \in \mathbb{N} \setminus \{0\}$. Die Funktionen

$s := \sup_{n\geq 1} (s_n)$ und $S := \inf_{n\geq 1} (S_n)$ sind daher wohldefiniert und, nach Proposition 3.27, messbar, wobei $[a,b]$ die σ-Algebra $\mathcal{L}_{|[a,b]}$ trägt. Nach Konstruktion gilt

$$\inf_{[a,b]}(f) \leq s_n \leq s \leq f \leq S \leq S_n \leq \sup_{[a,b]}(f).$$

Da beschränkte messbare Funktionen auf beschränkten Intervallen λ-integrierbar sind, sind s_n, s, S_n, S λ-integrierbar auf $[a,b]$ und es gilt

$$\int_{[a,b]} s_n d\lambda \leq \int_{[a,b]} s d\lambda \leq \int_{[a,b]} S d\lambda \leq \int_{[a,b]} S_n d\lambda \tag{3.2}$$

für alle $n \in \mathbb{N} \setminus \{0\}$. Wegen

$$0 \leq \int_{[a,b]} S_n d\lambda - \int_{[a,b]} s_n d\lambda = \sum_{i=0}^{k-1} (\sup_{[t_{i,n},t_{i+1,n}]} (f) - \inf_{[t_{i,n},t_{i+1,n}]} (f))(t_{i+1,n} - t_{i,n}) \leq \frac{1}{n}$$

gilt $\int_{[a,b]} S_n d\lambda - \int_{[a,b]} s_n d\lambda \underset{n\to\infty}{\longrightarrow} 0$. Dies impliziert $\int_{[a,b]} S - sd\lambda = \int_{[a,b]} Sd\lambda - \int_{[a,b]} sd\lambda = 0$.

Behauptung 1: *Sei $g : X \longrightarrow [0,\infty]$ eine messbare Abbildung auf einem Maßraum $(X, \mathcal{A}, \mu)$ mit $\int_X g d\mu = 0$. Dann gilt $g = 0$ μ-f.ü.*

Beweis: Sei $Y := g^{-1}(]0,\infty]) \subset X$. Zu zeigen ist, dass Y eine μ-Nullmenge ist. Nach Voraussetzung ist $Y \in \mathcal{A}$ (wegen g messbar). Angenommen, $\mu(Y) > 0$. Wegen $]0,\infty] = \bigcup_{n\geq 1}]\frac{1}{n},\infty]$ ist $Y = \bigcup_{n\geq 1} Y_n$ mit $Y_n := g^{-1}(]\frac{1}{n},\infty]) \in \mathcal{A}$. Außerdem ist $Y_n \subset Y_{n+1}$ für alle $n \geq 1$. Proposition 3.12 liefert $\lim_{n\to\infty} \mu(Y_n) = \mu(Y) > 0$, insbesondere existiert mindestens ein $n \geq 1$ mit $\mu(Y_n) > 0$. Dies aber impliziert

$$\int_X g d\mu = \int_{Y_n} g d\mu + \underbrace{\int_{Y_n^c} g d\mu}_{\geq 0} \geq \int_{Y_n} g d\mu \geq \int_{Y_n} \frac{1}{n} d\mu = \frac{\mu(Y_n)}{n} > 0,$$

Widerspruch zur Annahme $\int_X g d\mu = 0$. Dies zeigt $\mu(Y) = 0$, was zu beweisen war. $\checkmark$

Die Funktion $S - s : [a,b] \longrightarrow [0,\infty[$ ist messbar mit $\int_{[a,b]} S - sd\lambda = 0$, insbesondere gilt nach der Behauptung 1 $S - s = 0$ λ-f.ü. Wegen $s \leq f \leq S$ gilt dann $s = f = S$ λ-f.ü.

Behauptung 2: *Sei $g : X \longrightarrow \mathbb{R}$ eine Abbildung auf einem vollständigen Maßraum $(X, \mathcal{A}, \mu)$, die μ-f.ü. einer messbaren Abbildung $h : X \longrightarrow \mathbb{R}$ gleicht. Dann ist g ebenfalls messbar.*

Beweis: Nach Voraussetzung gilt $g = h$ auf $X \setminus N$, wobei $N \subset X$ eine μ-Nullmenge ist. Da μ vollständig ist, gilt $N \in \mathcal{A}$. Ist nun $B \in \mathcal{B}(\mathbb{R})$, so gilt $g^{-1}(B) = g^{-1}_{|_{N^c}}(B) \cup g^{-1}_{|_N}(B) = h^{-1}_{|_{N^c}}(B) \cup g^{-1}_{|_N}(B)$, wobei $h^{-1}_{|_{N^c}}(B) \in \mathcal{A}$ wegen h messbar und $g^{-1}_{|_N}(B) \in \mathcal{A}$ wegen N μ-Nullmenge. Daraus folgt $g^{-1}(B) \in \mathcal{A}$ und die Behauptung. $\checkmark$

Da das Lebesgue-Maß *vollständig* ist, impliziert die Behauptung 2, dass f messbar ist und daher λ-integrierbar auf $[a,b]$ (wegen f beschränkt) mit $\int_{[a,b]} sd\lambda = \int_{[a,b]} fd\lambda = \int_{[a,b]} Sd\lambda$. Schließlich bleibt, zu bemerken, dass $\lim_{n\to\infty} \int_{[a,b]} s_n d\lambda = \int_a^b f(x)dx = \lim_{n\to\infty} \int_{[a,b]} S_n d\lambda$ nach Konstruktion von s_n und S_n gilt. Die Kette von Ungleichungen (3.2) liefert $\lim_{n\to\infty} \int_{[a,b]} s_n d\lambda = \int_{[a,b]} sd\lambda$, somit $\int_{[a,b]} fd\lambda = \int_a^b f(x)dx$. Dies beweist den Satz 3.50. $\square$

Beispiele 3.51

1. Sei $X := [0,1]$, $\mathcal{A} := \mathcal{L}_{|[0,1]}$ und $\mu := \lambda_{|[0,1]}$ (Lebesgue-Maß). Betrachte die Folge $f_n : X \longrightarrow \mathbb{R}$, $x \longmapsto x^n$. Jede Funktion f_n ist natürlich messbar (da stetig) und es gilt $f_{n+1}(x) \leq f_n(x)$ für alle $n \in \mathbb{N}$ und $x \in [0,1]$. Da f_0 die konstante Funktion 1 ist und $\int_{[0,1]} 1 d\lambda = \lambda([0,1]) = 1 < \infty$, folgt aus Korollar 3.46:
$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \lim_{n\to\infty} f_n d\mu,$$
wobei $\lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{falls } x \in [0,1[\\ 1 & \text{falls } x = 1 \end{cases}$, also $\lim_{n\to\infty} f_n = \chi_{\{1\}}$. Wegen $\lambda(\{1\}) = 0$ folgt $\int_0^1 f_n(x)dx = \int_{[0,1]} f_n d\lambda \underset{n\to\infty}{\longrightarrow} 0$. Tatsächlich ist $\int_0^1 f_n(x)dx = \left[\frac{x^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1} \underset{n\to\infty}{\longrightarrow} 0$. Beachte aber, dass $(f_n)_{n\in\mathbb{N}}$ *nicht gleichmäßig* gegen $\chi_{\{1\}} = \lim_{n\to\infty} f_n$ konvergiert.

2. Sei nochmal $(X, \mathcal{A}, \mu) := ([0,1], \mathcal{L}_{|[0,1]}, \lambda_{|[0,1]})$ und betrachte $f_n : X \longrightarrow \mathbb{R}$, $x \longmapsto \begin{cases} \sin(\frac{1}{nx}) & \text{für } x \in]0,1] \\ 0 & \text{für } x = 0 \end{cases}$, für alle $n \in \mathbb{N}$, $n \geq 1$. Beachte: f_n ist zwar messbar auf X (*Übungsaufgabe*) aber nicht stetig ($\lim_{x\to 0^+} f_n(x)$ existiert nicht). Es gilt allerdings: für jedes $x \in X$ ist $\lim_{n\to\infty} f_n(x) = 0$ (wegen $\sin(x) \underset{x\to 0}{\longrightarrow} \sin(0) = 0$) und für die μ-integrierbare Funktion $g : X \longrightarrow \mathbb{R}$, $x \longmapsto 1$ gilt $|f| \leq g$ auf X. Satz 3.45 impliziert, dass
$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \lim_{n\to\infty} f_n d\mu,$$
d.h., $\int_{[0,1]} f_n d\lambda \underset{n\to\infty}{\longrightarrow} 0$. In dem Fall konvergiert $(f_n)_{n\in\mathbb{N}}$ nicht gleichmäßig gegen 0 auf $[0,1]$; außerdem wäre eine Stammfunktion von f_n nur schwierig zu bestimmen!

Nun vergleichen wir das Lebesgue-Integral für Funktionen $f : I \longrightarrow \mathbb{R}$ (wobei $I \subset \mathbb{R}$ Intervall) mit dem sogenannten *uneigentlichen Riemann-Integral*.
Erinnerung: Sei $I := [a,b[\subset \mathbb{R}$ ein halboffenes Intervall mit $a \in \mathbb{R}$ und $b \in]a,\infty]$. Sei $f : I \longrightarrow \mathbb{R}$ eine Funktion. Man sagt, dass das Integral $\int_a^b f(x)dx$ *konvergiert*, wenn für jedes $c \in]a,b[$ die Funktion $f_{|[a,c]} : [a,c] \longrightarrow \mathbb{R}$ beschränkt und Riemann-integrierbar ist und der Grenzwert $\lim_{c\to b} \int_a^c f(x)dx$ existiert in $\mathbb{R}$. Ggf. definiert man
$$\int_a^b f(x)dx := \lim_{c\to b} \int_a^c f(x)dx \in \mathbb{R}.$$
Analog für eine Funktion $f :]a,b] \longrightarrow \mathbb{R}$ (mit $b \in \mathbb{R}$ und $a \in [-\infty, b[$). Für eine Funktion $f :]a,b[\longrightarrow \mathbb{R}$ (mit $a, b \in \overline{\mathbb{R}}, a < b$) konvergiert das Integral $\int_a^b f(x)dx$, wenn für ein (und dann jedes) $c \in]a,b[$ die Integrale $\int_a^c f(x)dx$ und $\int_c^b f(x)dx$ konvergieren. Ggf. setzt man $\int_a^b f(x)dx := \int_a^c f(x)dx + \int_c^b f(x)dx$. Das Integral $\int_a^b f(x)dx$ heißt genau dann *absolut konvergent*, wenn das Integral $\int_a^b |f(x)|dx$ konvergiert. Ggf. konvergiert ebenfalls $\int_a^b f(x)dx$ und es gilt $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$, siehe Analysis I.

Proposition 3.52 *Sei $f : [a,b[\longrightarrow \mathbb{R}$ eine Funktion, wobei $a \in \mathbb{R}$ und $b \in]a,\infty]$. Angenommen, für jedes $c \in]a,b[$ sei $f_{|[a,c]} : [a,c] \longrightarrow \mathbb{R}$ beschränkt und Riemann-integrierbar. Dann sind folgende Aussagen äquivalent:*

i) Die Funktion f ist λ-integrierbar auf $[a,b[$.

ii) Das Integral $\int_a^b f(x)dx$ konvergiert absolut.

Ggf. gilt $\int_{[a,b[} f d\lambda = \int_a^b f(x)dx$.

Beweis: Nach Voraussetzung ist, für jedes $c \in [a,b[$, die Funktion $f_{|_{[a,c]}}$ Riemann-integrierbar. Daraus folgt insbesondere, dass f messbar ist, denn: ist $B \in \mathcal{B}(\mathbb{R})$ beliebig, so gilt $f^{-1}(B) = \bigcup\limits_{n\in\mathbb{N}} f^{-1}_{|_{[a,c_n]}}(B)$, wobei $(c_n)_{n\in\mathbb{N}}$ eine Folge aus $[a,b[$ ist mit $c_n \underset{n\to\infty}{\longrightarrow} b$; wegen $f^{-1}_{|_{[a,c_n]}}(B) \in \mathcal{L}$ gilt dann $f^{-1}(B) \in \mathcal{L}$. Außerdem ist, nach Analysis I, die Funktion $|f|_{|_{[a,c]}}$ ebenfalls Riemann-integrierbar. Satz 3.50 liefert, dass $|f|_{|_{[a,c]}}$ λ-integrierbar ist mit $\int_{[a,c]} |f| d\lambda = \int_a^c |f(x)|dx$. Ist nun $(c_n)_{n\in\mathbb{N}}$ eine beliebige monoton wachsende Folge aus $[a,b[$ mit $c_n \underset{n\to\infty}{\longrightarrow} b$, so folgt aus dem Satz über monotone Konvergenz (Satz 3.33)

$$\int_{[a,b[} |f| d\lambda = \lim_{n\to\infty} \int_{[a,c_n]} |f| d\lambda,$$

denn $\int_{[a,c_n]} |f| d\lambda = \int_{[a,b[} |f| \cdot \chi_{[a,c_n]} d\lambda$ und die Folge $(|f| \cdot \chi_{[a,c_n]})_{n\in\mathbb{N}}$ ist eine monotone wachsende Folge nichtnegativer messbarer Funktionen, welche punktweise gegen $|f|$ konvergiert (auf $[a,b[$). Insgesamt gilt $\int_{[a,b[} |f| d\lambda = \lim\limits_{n\to\infty} \int_a^{c_n} |f(x)|dx$, insbesondere ist $\int_{[a,b[} |f| d\lambda$ genau dann endlich, wenn die Folge $(\int_a^{c_n} |f(x)|dx)_{n\in\mathbb{N}}$ einen (reellen) Grenzwert besitzt, für alle Folgen $(c_n)_{n\in\mathbb{N}}$ wie oben. Das heißt, die Funktion f ist genau dann λ-integrierbar auf $[a,b[$, wenn das Integral $\int_a^b |f(x)|dx$ konvergiert, d.h., wenn das Integral $\int_a^b f(x)dx$ *absolut* konvergiert. Dies beweist die Äquivalenz zwischen $i)$ und $ii)$. Ist $i)$ oder $ii)$ erfüllt, so gilt $\int_{[a,b[} |f| d\lambda = \lim\limits_{n\to\infty} \int_a^{c_n} |f(x)|dx = \int_a^b |f(x)|dx$ und der Satz über majorisierte Konvergenz (Satz 3.45) liefert

$$\int_{[a,b[} f d\lambda = \lim_{n\to\infty} \int_{[a,c_n]} f d\lambda = \lim_{n\to\infty} \int_a^{c_n} f(x)dx = \int_a^b f(x)dx,$$

wobei genutzt wurde, dass das Integral $\int_a^b f(x)dx$ automatisch konvergiert, wenn sie absolut konvergiert (siehe Analysis I). $\square$

Bemerkung 3.53 Es kann passieren, dass das Integral $\int_a^b f(x)dx$ konvergiert, dass aber f nicht λ-integrierbar auf $[a,b[$ ist. Betrachte z.B. $f : [0,\infty[\longrightarrow \mathbb{R}, x \longmapsto \begin{cases} 1 & \text{für } x = 0 \\ \frac{\sin(x)}{x} & \text{für } x > 0 \end{cases}$. Dann konvergiert das Integral $\int_0^\infty f(x)dx$ (siehe Analysis I); dieses Integral konvergiert aber nicht absolut wegen $\int_0^t |f(x)|dx \underset{t\to\infty}{\longrightarrow} \infty$, somit ist nach Proposition 3.52 die Funktion f nicht λ-integrierbar auf $[0,\infty[$.

3.2.7 Parameterabhängige Integrale

Satz 3.54 *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ eine Funktion. Angenommen, für alle $a,b \in I$ mit $a \leq b$ sei $f_{|[a,b]} : [a,b] \longrightarrow \mathbb{R}$ λ-integrierbar. Fixiere $x_0 \in I$. Dann gilt:*

i) Die Funktion $F : I \longrightarrow \mathbb{R},\ x \longmapsto \begin{cases} \int_{[x_0,x]} f d\lambda & \textit{falls } x_0 \leq x \\ -\int_{[x,x_0]} f d\lambda & \textit{falls } x_0 \geq x \end{cases}$, *ist wohldefiniert und stetig.*

ii) Ist f *zusätzlich stetig in einem* $x \in I$, *so ist* F *ableitbar in* x *mit* $F'(x) = f(x)$.

Beweis:
$i)$ Bemerke, dass nach Voraussetzung die Funktion f auf jedem kompakten Intervall λ-integrierbar ist, insbesondere auf $[x_0, x]$ (falls $x_0 \leq x$) bzw. $[x, x_0]$ (falls $x_0 \geq x$), für alle $x \in I$; das heißt, F ist wohldefiniert auf I. Sei $x \in I$ und $(x_n)_{n\geq 1}$ eine beliebige Folge aus I mit $x_n \underset{n\to\infty}{\longrightarrow} x$. Zu zeigen ist $F(x_n) \underset{n\to\infty}{\longrightarrow} F(x)$. Setze $I_n := [x_0, x_n]$ falls $x_n \geq x_0$ und $I_n := [x_n, x_0]$ falls $x_n \leq x_0$. Dann gilt für alle $n \in \mathbb{N}$

$$F(x_n) = \begin{cases} \int_{I_n} f d\lambda & = \int_{\mathbb{R}} f \cdot \chi_{I_n} d\lambda & \text{falls } x_n \geq x_0 \\ -\int_{I_n} f d\lambda & = -\int_{\mathbb{R}} f \cdot \chi_{I_n} d\lambda & \text{falls } x_n \leq x_0 \end{cases}.$$

Fixiere $\varepsilon > 0$ mit $[x - \varepsilon, x + \varepsilon] \subset I$ (möglich wegen I offen in $\mathbb{R}$). Wegen $x_n \underset{n\to\infty}{\longrightarrow} x$ existiert ein $N \in \mathbb{N}$ mit $x_n \in]x - \varepsilon, x + \varepsilon[$ für alle $n \geq N$. Betrachte die Funktionenfolge $(f_n := f \cdot \chi_{I_n})_{n\geq 1}$. Jede Funktion $f_n : \mathbb{R} \longrightarrow \mathbb{R}$ ist als Produkt messbarer Funktionen messbar. Außerdem gilt $|f_n| \leq |f \cdot \chi_J|$ für alle $n \geq N$, wobei $J := [x_0, x] \cup [x - \varepsilon, x + \varepsilon]$ falls $x_0 \leq x$ und $J := [x, x_0] \cup [x - \varepsilon, x + \varepsilon]$ falls $x_0 \geq x$. Da f auf dem kompakten Intervall J λ-integrierbar ist, ist $|f \cdot \chi_J|$ λ-integrierbar auf $\mathbb{R}$. Wegen $x_n \underset{n\to\infty}{\longrightarrow} x$ konvergiert die Folge $(\chi_{I_n})_{n\geq N}$ punktweise auf $\mathbb{R} \setminus \{x\}$ gegen $\chi_{[x_0,x[}$ falls $x_0 \leq x$ bzw. gegen $\chi_{]x,x_0]}$ falls $x_0 \geq x$. Insbesondere konvergiert die Folge $(f \cdot \chi_{I_n})_{n\geq N}$ punktweise auf $\mathbb{R} \setminus \{x\}$ gegen $f \cdot \chi_{[x_0,x[}$ falls $x_0 \leq x$ bzw. gegen $f \cdot \chi_{]x,x_0]}$ falls $x_0 \geq x$. Nun folgt aus dem Satz über majorisierte Konvergenz μ-fast-überall (Korollar 3.48), dass

$$\int_{\mathbb{R}} f \cdot \chi_{I_n} d\lambda \underset{n\to\infty}{\longrightarrow} \begin{cases} \int_{\mathbb{R}} f \cdot \chi_{[x_0,x[} d\lambda & \text{falls } x_0 \leq x \\ \int_{\mathbb{R}} f \cdot \chi_{]x,x_0]} d\lambda & \text{falls } x_0 \geq x \end{cases}$$

gilt. Da jeder Punkt aus $\mathbb{R}$ eine λ-Nullmenge ist, folgt in allen Fällen $F(x_n) \underset{n\to\infty}{\longrightarrow} F(x)$. Dies zeigt die Stetigkeit von F auf I und somit $i)$.
$ii)$ Angenommen, f sei in einem $x \in I$ ableitbar. Sei $h \in \mathbb{R}$ mit $x + h \in I$. Dann gilt

$$\begin{aligned} F(x+h) - F(x) &= \int_{[x,x+h]} f d\lambda \\ &= \int_{[x,x+h]} f - f(x) d\lambda + \int_{[x,x+h]} f(x) d\lambda \\ &= h f(x) + \int_{[x,x+h]} f - f(x) d\lambda \end{aligned}$$

im Fall $h \geq 0$ und analog $F(x + h) - F(x) = hf(x) - \int_{[x+h,x]} f - f(x) d\lambda$ falls $h \leq 0$. Beachte, dass nach Voraussetzung $f - f(x)$ auf jedem kompakten Intervall λ-integrierbar ist, da jede konstante Funktion auf jedem kompakten Intervall λ-integrierbar ist. Da f stetig in x ist, existiert zu jedem $\varepsilon > 0$ ein $\eta > 0$ so, dass $|f(y) - f(x)| \leq \varepsilon$ für alle $y \in]x - \eta, x + \eta[\cap I$. Insbesondere folgt, für alle $|h| < \eta$,

$$\left|\int_{[x,x+h]} f - f(x) d\lambda\right| \overset{\text{Prop.3.42}}{\leq} \int_{[x,x+h]} |f - f(x)| d\lambda \leq h \cdot \varepsilon$$

falls $h \geq 0$ und analog $|\int_{[x+h,x]} f - f(x)d\lambda| \leq (-h) \cdot \varepsilon$ falls $h \leq 0$. In beiden Fällen gilt $|F(x+h) - f(x) - h \cdot f(x)| \leq |h| \cdot \varepsilon$ für alle h mit $|h| < \eta$. Dies zeigt, dass F differenzierbar - d.h., ableitbar - in x ist mit Ableitung $f(x)$. $\square$

Bemerke, dass insbesondere F ableitbar – und sogar C^1 – ist, sobald f stetig auf I ist. Allgemeiner ist F C^{k+1} auf I, sobald f C^k auf I ist.

Satz 3.55 (stetige Abhängigkeit vom Parameter) *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum, (Y, d) ein metrischer Raum und $f : X \times Y \longrightarrow \mathbb{R}^p$ eine Abbildung. Wir nehmen an:*

a) Für alle $y \in Y$ sei die Abbildung $f_y : X \longrightarrow \mathbb{R}^p$, $x \longmapsto f(x, y)$, messbar;

b) Für μ-fast alle $x \in X$ sei die Abbildung $Y \longrightarrow \mathbb{R}^p$, $y \longmapsto f(x, y)$, stetig auf Y;

c) Eine μ-integrierbare Funktion $g : X \longrightarrow \mathbb{R}$ existiere mit $|f_y| \leq g$ μ-f.ü. auf X und für alle $y \in Y$.

Dann ist, für alle $y \in Y$, die Funktion $f_y : X \longrightarrow \mathbb{R}^p$, μ-integrierbar und die Abbildung $F : Y \longrightarrow \mathbb{R}^p$, $y \longmapsto \int_X f_y d\mu$, ist stetig auf (Y, d).

Beweis: Für ein beliebiges $y \in Y$ ist nach Voraussetzung $f_y : X \longrightarrow \mathbb{R}$ messbar mit $|f_y| \leq g$ auf dem Komplement einer μ-Nullmenge, insbesondere gilt $\int_X |f_y| d\mu \leq \int_X g d\mu < \infty$. Somit ist f_y μ-integrierbar auf X; daher ist $F(y)$ wohldefiniert. Sei nun $(y_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus Y mit $y_n \underset{n\to\infty}{\longrightarrow} y$, d.h., $d(y_n, y) \underset{n\to\infty}{\longrightarrow} 0$. Betrachte die Funktionenfolge $(g_n)_{n\in\mathbb{N}}$, wobei $g_n(x) := f(x, y_n) - f(x, y)$ für alle $x \in X$. Nach Definition ist $g_n = f_{y_n} - f_y$ als Differenz zweier messbarer Funktionen wieder messbar auf X mit $|g_n| \leq |f_{y_n}| + |f_y| \leq 2g$ μ-fast-überall auf X, wobei $2g : X \longrightarrow [0, \infty[$ μ-integrierbar ist. Nach Voraussetzung konvergiert g_n auf dem Komplement einer μ-Nullmenge punktweise gegen 0 (wegen der Stetigkeit von f in der zweiten Variablen). Der Satz über majorisierte Konvergenz μ-fast-überall (Korollar 3.48) liefert $\int_X g_n d\mu \underset{n\to\infty}{\longrightarrow} 0$, d.h., $F(y_n) - F(y) \underset{n\to\infty}{\longrightarrow} 0$. Dies zeigt die Stetigkeit von F an der Stelle $y \in Y$. Da y beliebig gewählt werden konnte, folgt die Stetigkeit von F auf Y. $\square$

Satz 3.56 (Differentiation unter dem Integral) *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $I \subset \mathbb{R}$ ein offenes Intervall und $f : X \times I \longrightarrow \mathbb{R}^p$ eine Abbildung. Wir nehmen an:*

a) Für alle $t \in I$ sei $f_t : X \longrightarrow \mathbb{R}^p$, $x \longmapsto f(x, t)$, μ-integrierbar;

b) Für μ-fast alle $x \in X$ sei die Abbildung $I \longrightarrow \mathbb{R}^p$, $t \longmapsto f(x, t)$, ableitbar;

c) Eine μ-integrierbare Funktion $g : X \longrightarrow \mathbb{R}$ existiere mit $\left|\frac{\partial f}{\partial t}(x, t)\right| \leq g(x)$ für μ-fast alle $x \in X$ und alle $t \in I$, wobei $\frac{\partial f}{\partial t}$ die t-Ableitung von f bezeichnet.

Dann ist $F : I \longrightarrow \mathbb{R}^p$, $t \longmapsto \int_X f_t d\mu$, ableitbar mit $F'(t) = \int_X \frac{\partial f}{\partial t} d\mu$, für alle $t \in I$.

Beweis: Für ein beliebiges $t \in I$ sei $(h_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathbb{R}\setminus\{0\}$ mit $h_n \underset{n\to\infty}{\longrightarrow} 0$. Betrachte

$$\frac{F(t + h_n) - F(t)}{h_n} = \frac{1}{h_n} \int_X (f(x, t + h_n) - f(x, t))d\mu(x) = \int_X \underbrace{\frac{f(x, t + h_n) - f(x, t)}{h_n}}_{=:g_n(x)} d\mu(x).$$

Jede Funktion $g_n : X \longrightarrow \mathbb{R}$ ist nach Voraussetzung μ-integrierbar - insbesondere messbar - auf X mit $g_n(x) \underset{n\to\infty}{\longrightarrow} \frac{\partial f}{\partial t}(x,t)$ für μ-fast alle $x \in X$. Desweiteren folgt aus dem Mittelwertsatz (Satz 2.69)

$$|g_n(x)| \leq \sup_{s\in[t,t+h_n]} (|\frac{\partial f}{\partial t}(x,s)|),$$

für μ-fast alle $x \in X$, insbesondere auch $|g_n(x)| \leq g(x)$ für μ-fast alle $x \in X$ nach Voraussetzung. Da g μ-integrierbar ist, kann der Satz über majorisierte Konvergenz μ-fast-überall (Korollar 3.48) angewendet werden und er liefert $\int_X g_n d\mu \underset{n\to\infty}{\longrightarrow} \int_X \frac{\partial f}{\partial t}(x,t)d\mu(x)$. Da die Folge $(h_n)_{n\in\mathbb{N}}$ beliebig gewählt werden konnte, folgt, dass F ableitbar ist mit Ableitung $F'(t) = \int_X \frac{\partial f}{\partial t}(x,t)d\mu(x)$ für alle $t \in I$, was zu beweisen war. □

Korollar 3.57 *Sei $f : [a,b] \times I \longrightarrow \mathbb{R}^p$ eine C^1-Abbildung, wobei $a \leq b$ reell sind und $I \subset \mathbb{R}$ ein offenes Intervall ist. Dann ist die Funktion $F : t \longmapsto \int_a^b f(x,t)dx$, $I \longrightarrow \mathbb{R}^p$ C^1 auf I mit Ableitung*

$$F'(t) = \int_a^b \frac{\partial f}{\partial t}(x,t)dx \quad \textit{für alle } t \in I.$$

Beweis: Für jedes $t \in I$ ist $f_t = f(\cdot,t) : [a,b] \longrightarrow \mathbb{R}$ als stetige Funktion auf einem kompakten Intervall λ-integrierbar – daher auch Riemann-integrierbar nach Satz 3.50. Nach Voraussetzung ist, für jedes $x \in [a,b]$, die Funktion $I \xrightarrow{f(x,\cdot)} \mathbb{R}$ ableitbar, mit Ableitung $\frac{\partial f}{\partial t}(x,\cdot)$. Für jedes kompakte Intervall $[c,d] \subset I$ ist außerdem die Funktion $\frac{\partial f}{\partial t} : [a,b] \times [c,d] \longrightarrow \mathbb{R}$ stetig und somit beschränkt (siehe Korollar 2.32), d.h., es gibt ein $M \in [0,\infty[$ mit $|\frac{\partial f}{\partial t}(x,t)| \leq M$ für alle $(x,t) \in [a,b] \times [c,d]$, insbesondere auch für alle $(x,t) \in [a,b] \times]c,d[$. Da jede konstante Funktion auf einem kompakten Intervall λ-integrierbar ist, kann Satz 3.56 auf $f : [a,b] \times]c,d[\longrightarrow \mathbb{R}$ angewendet werden, welcher liefert, dass $t \overset{F}{\longmapsto} \int_{[a,b]} f(x,t)d\lambda(x)$ ableitbar auf $]c,d[$ ist mit Ableitung $F'(t) = \int_{[a,b]} \frac{\partial f}{\partial t}(x,t)d\lambda(x)$, d.h., $F'(t) = \int_a^b \frac{\partial f}{\partial t}(x,t)dx$ (nach Satz 3.50) für alle $t \in]c,d[$. Da $[c,d] \subset I$ beliebig gewählt werden konnte, folgt die Behauptung. □

3.3 Das Produktmaß

3.3.1 Produkte von Maßräumen

Definition 3.58 *Seien $(X,\mathcal{A})$ und $(Y,\mathcal{B})$ Messräume. Die* Produkt σ-Algebra *von $\mathcal{A}$ und $\mathcal{B}$ ist die von $\mathcal{A} \times \mathcal{B} := \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\} \subset \mathcal{P}(X \times Y)$ erzeugte σ-Algebra auf $X \times Y$. Sie wird mit $\mathcal{A} \otimes \mathcal{B}$ bezeichnet.*

Beachte, dass $\mathcal{A} \times \mathcal{B}$ keine σ-Algebra ist *(Bild)*.

Beispiel 3.59 Sei wie vorher $\mathcal{B}(\mathbb{R}^n)$ die Borel'sche σ-Algebra von $\mathbb{R}^n$ (welche nach Definition durch die offenen Teilmengen des $\mathbb{R}^n$ erzeugt wird). Dann gilt $\mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^l) = \mathcal{B}(\mathbb{R}^{k+l})$ für alle $k,l \in \mathbb{N}$. Denn: wir wissen schon, dass $\mathcal{B}(\mathbb{R}^k)$ durch Teilmengen der Form $\prod_{i=1}^k]a_i,b_i[$ mit $a_i \leq b_i$ reell erzeugt wird, für jedes $k \in \mathbb{N}$ (siehe Bemerkung 3.8.2). Da das Produkt einer Teilmenge dieser Form wieder eine Teilmenge dieser Form ist, gilt

$\mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^l) \subset \mathcal{B}(\mathbb{R}^{k+l})$ nach Definition der Produkt-σ-Algebra; wegen $\prod_{i=1}^{k+l}]a_i, b_i[= \big(\prod_{i=1}^{k}]a_i, b_i[\big) \times \big(\prod_{i=k+1}^{k+l}]a_i, b_i[\big) \in \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^l)$ gilt dann auch $\mathcal{B}(\mathbb{R}^{k+l}) \subset \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^l)$.

Proposition 3.60 *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Maßräume mit μ und ν* σ-endlich. *Dann existiert ein eindeutiges Maß $\mu \otimes \nu$ auf $\mathcal{A} \otimes \mathcal{B}$ mit*

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$$

für alle $(A, B) \in \mathcal{A} \times \mathcal{B}$. Außerdem ist $\mu \otimes \nu$ σ-endlich.

Der Beweis von Proposition 3.60 wird in den Anhang A verschoben.

Definition 3.61 *Das Maß $\mu \otimes \nu$ aus der* Proposition 3.60 *heißt das* Produktmaß *von μ und ν.*

Die Konstruktion des Produktmaßes braucht folgendes Lemma.

Lemma 3.62 *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Maßräume mit μ und ν* σ-endlich. *Sei $E \in \mathcal{A} \otimes \mathcal{B}$ beliebig. Dann gilt:*

i) Für alle $x \in X$ ist $E_x := \{y \in Y \mid (x, y) \in E\} \subset Y$ Element von $\mathcal{B}$; analog ist für jedes $y \in Y$ die Teilmenge $E_y := \{x \in X \mid (x, y) \in E\} \subset X$ Element von $\mathcal{A}$.

ii) Die Abbildungen $X \longrightarrow [0, \infty]$, $x \longmapsto \nu(E_x)$ und $Y \longrightarrow [0, \infty]$, $y \longmapsto \mu(E_y)$, sind messbar.

Lemma 3.62 wird im Anhang A bewiesen, siehe Lemma A.16.

Mit den Bezeichnungen von Lemma 3.62 wird das Produktmaß von μ und ν folgendermaßen definiert, siehe Satz A.17:

$$(\mu \otimes \nu)(E) := \int_X \nu(E_x) d\mu(x)$$

für alle $E \in \mathcal{A} \otimes \mathcal{B}$. Die Eindeutigkeit dieses Produktmaßes zeigt insbesondere, dass $\mu \otimes \nu$ auch durch $(\mu \otimes \nu)(E) = \int_Y \mu(E_y) d\nu(y)$ für alle $E \in \mathcal{A} \otimes \mathcal{B}$ gegeben ist. Dies führt zum Satz von Fubini.

3.3.2 Der Satz von Fubini

Satz 3.63 (Satz von Fubini) *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Maßräume mit μ und ν* σ-endlich.

1. (Version für nichtnegative Funktionen) *Sei $f : X \times Y \longrightarrow [0, \infty]$ eine messbare Abbildung. Dann gilt:*

   i) Für jedes $(x, y) \in X \times Y$ sind $f_y : X \longrightarrow [0, \infty]$, $z \longmapsto f(z, y)$ und $f_x : Y \longrightarrow [0, \infty]$, $w \longmapsto f(x, w)$, messbar.

   ii) Die Abbildungen $Y \longrightarrow [0, \infty]$, $y \longmapsto \int_X f_y d\mu$ und $X \longrightarrow [0, \infty]$, $x \longmapsto \int_Y f_x d\nu$, sind messbar und es gilt

   $$\int_Y \Big(\int_X f_y d\mu\Big) d\nu(y) = \int_{X \times Y} f d\mu \otimes \nu = \int_X \Big(\int_Y f_x d\nu\Big) d\mu(x).$$

2. (Version für integrierbare Funktionen) *Sei* $f : X \times Y \longrightarrow \mathbb{R}$ *(bzw.* $f : X \times Y \longrightarrow \mathbb{C}$ *oder* $f : X \times Y \longrightarrow \mathbb{R}^p$*) eine* $\mu \otimes \nu$-integrierbare *Abbildung. Dann gilt:*

   i) Für alle $(x, y) \in X \times Y$ *sind (in den obigen Bezeichnungen) die Abbildungen* f_y *und* f_x *messbar. Außerdem ist die Funktion* f_y *μ-integrierbar für ν-fast alle* $y \in Y$ *und* f_x *ist ν-integrierbar für μ-fast alle* $x \in X$.

   ii) Die ν-fast überall definierte Abbildung $y \longmapsto \int_X f_y d\mu$ *ist ν-fast überall gleich einer ν-integrierbaren Abbildung* $Y \longrightarrow \mathbb{R}$ *(und analog für* f_x*) und es gilt*

$$\int_Y (\int_X f_y d\mu) d\nu(y) = \int_{X\times Y} f d\mu \otimes \nu = \int_X (\int_Y f_x d\nu) d\mu(x).$$

Zusammengefasst heißt der zweite Fall im Satz 3.63: die Integrale $\int_Y(\int_X f_y d\mu)d\nu(y)$ und $\int_X(\int_Y f_x d\nu)d\mu(x)$ existieren und sind beide gleich $\int_{X\times Y} f d\mu \otimes \nu$. Hierbei deutet die Bezeichnung $d\nu(y)$ (bzw. $d\mu(x)$) darauf hin, dass die Integrationsvariable y (bzw. x) ist.

Beweis von Satz 3.63: Der Beweis beruht größtenteils auf der Konstruktion des Produktmaßes sowie auf dem Satz über monotone Konvergenz. Sei zuerst $y \in Y$ beliebig. Dann ist f_y gleich der Verknüpfung $f \circ \iota_y$, wobei $\iota_y : X \longrightarrow X \times Y$, $x \mapsto (x, y)$. Die Abbildung ι_y ist aber messbar, denn: für jedes $E \in \mathcal{A} \otimes \mathcal{B}$ gilt $\iota_y^{-1}(E) = \{x \in X \,|\, \iota_y(x) \in E\} = \{x \in X \,|\, (x, y) \in E\} = E_y$ mit $E_y \in \mathcal{A}$ nach Lemma A.16.*i*). Da f in beiden Versionen messbar vorausgesetzt wird, ist f_y als Verknüpfung zweier messbarer Funktionen wieder messbar (Proposition 3.22). Analog ist f_x für jedes $x \in X$ messbar auf Y.
1. Da für jedes $x \in X$ die Abbildung $f_x : Y \longrightarrow [0, \infty]$ messbar und nichtnegativ ist, existiert $\int_X f_x d\nu \in [0, \infty]$. Nun wollen wir zeigen, dass $x \mapsto \int_X f_x d\nu$ wieder messbar ist und die Integrale aus Teil 1 gleich sind. Wir betrachten zuerst den Fall, wo $f = \chi_E$ die charakteristische Funktion eines $E \in \mathcal{A}\otimes\mathcal{B}$ ist (insbesondere ist $f : X\times Y \to \{0, 1\}$ messbar, siehe Beispiel 3.21.3). Wegen $f_x = (\chi_E)_x = \chi_{E_x}$ (elementar) gilt $\int_Y f_x d\nu = \int_Y \chi_{E_x} d\nu = \nu(E_x)$. Da aber ν σ-endlich ist, folgt aus Lemma A.16.*ii*), dass $x \mapsto \nu(E_x)$ messbar ist. Analog ist wegen μ σ-endlich die Abbildung $y \mapsto \mu(E_y) = \int_X f_y d\mu$ messbar. Desweiteren gilt, nach der Konstruktion und der Eindeutigkeit des Produktmaßes,

$$\int_X \nu(E_x) d\mu(x) = (\mu \otimes \nu)(E) = \int_Y \mu(E_y) d\nu(y),$$

d.h., $\int_X(\int_Y f_x d\nu)d\mu(x) = \int_{X\times Y} f d\mu \otimes \nu = \int_Y(\int_X f_y d\mu)d\nu(y)$. Dies beweist die Aussage in diesem Spezialfall. Ist f eine nichtnegative Treppenfunktionen, so gelten die Aussagen ebenfalls für f, da f eine (nichtnegative) endliche Linearkombination von charakteristischen Funktionen ist. Ist nun $f : X \times Y \longrightarrow [0, \infty]$ eine beliebige nichtnegative messbare Abbildung, so existiert nach Proposition 3.30 eine monoton wachsende Folge $(s_n)_{n\in\mathbb{N}}$ von nichtnegativen Treppenfunktionen $X\times Y \to [0, \infty]$ mit $s_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ für alle $x \in X$. Sei $x \in X$ beliebig. Bemerke, dass dann $f_x : Y \longrightarrow [0, \infty]$ der punktweise Limes der monoton wachsenden Folge $((s_n)_x)_{n\in\mathbb{N}}$, wobei $(s_n)_x : Y \longrightarrow [0, \infty]$ wieder eine nichtnegative Treppenfunktion ist (wegen $(\chi_E)_x = \chi_{E_x}$, siehe oben). Der Satz über monotone Konvergenz (Satz 3.33) liefert

$$\int_Y f_x d\nu = \int_Y \lim_{n\to\infty} (s_n)_x d\nu = \lim_{n\to\infty} \int_Y (s_n)_x d\nu,$$

wobei $x \mapsto \int_Y (s_n)_x d\nu$ messbar ist, siehe Spezialfall. Insbesondere ist $x \mapsto \int_Y f_x d\nu$ als punktweiser Grenzwert einer Folge von messbaren Abbildungen wieder messbar (Proposition 3.25). Analog ist $y \mapsto \int_X f_y d\mu$ messbar auf Y. Desweiteren folgt aus dem Spezialfall und dem Satz über monotone Konvergenz

$$\begin{aligned}\int_X(\int_Y f_x d\nu)d\mu(x) &= \int_X(\lim_{n\to\infty}\int_Y (s_n)_x d\nu)d\mu(x)\\ &= \lim_{n\to\infty}\int_X(\int_Y (s_n)_x d\nu)d\mu(x)\\ &= \lim_{n\to\infty}\int_{X\times Y} s_n d\mu\otimes\nu\\ &= \int_{X\times Y}\lim_{n\to\infty} s_n d\mu\otimes\nu\\ &= \int_{X\times Y} f d\mu\otimes\nu\end{aligned}$$

und analog $\int_Y(\int_X f_y d\mu)d\nu(y) = \int_{X\times Y} f d\mu\otimes\nu$. Dies beweist 1.
2. Wir betrachten lediglich den Fall, wo f reellwertig ist; dann können alle Aussagen auf die reellwertigen Komponentenfunktionen von $f : X\times Y\to\mathbb{C}$ oder $f : X\times Y\to\mathbb{R}^p$ angewendet werden, was die Behauptungen in diesen Fällen liefert.
Wir wissen schon, dass $f_x : Y\longrightarrow\mathbb{R}$ und $f_y : X\longrightarrow\mathbb{R}$ messbar sind für alle $(x,y)\in X\times Y$. Betrachte die messbaren Abbildungen $f_+, f_- : X\times Y\longrightarrow[0,\infty[$. Bemerke, dass $(f_\pm)_x = (f_x)_\pm$ (bzw. $(f_\pm)_y = (f_y)_\pm$) für alle $x\in X$ (bzw. für alle $y\in Y$) gilt. Da f_+, f_- messbar und nichtnegativ sind, sind nach Teil 1 die Abbildungen $x\mapsto\int_Y(f_x)_\pm d\nu$ und $y\mapsto\int_X(f_y)_\pm d\mu$ messbar mit

$$\int_X(\int_Y(f_x)_+d\nu)d\mu(x)+\int_X(\int_Y(f_x)_-d\nu)d\mu(x) = \int_{X\times Y} f_++f_-d\mu\otimes\nu = \int_{X\times Y}|f|d\mu\otimes\nu<\infty,$$

was $\int_X(\int_Y(f_x)_\pm d\nu)d\mu(x)<\infty$ liefert. Dies impliziert einerseits, dass $\int_Y(f_x)_\pm d\nu<\infty$ für μ-fast alle $x\in X$ gelten, siehe Behauptung 1 im Beweis von Satz 3.50; d.h., die Abbildungen $(f_x)_\pm$ sind für alle x im Komplement einer (o.B.d.A.) gemeinsamen μ-Nullmenge $N\in\mathcal{A}$ ν-integrierbar auf Y. Andererseits liefert $\int_X(\int_Y(f_x)_\pm d\nu)d\mu(x)<\infty$ auch, dass die Abbildungen $x\mapsto\int_Y(f_x)_\pm d\nu$, $X\setminus N\to\mathbb{R}$, μ-integrierbar auf $X\setminus N$ sind. Daraus folgt erstens, dass $f_x = (f_x)_+-(f_x)_-$ für μ-fast alle $x\in X$ integrierbar auf Y mit $\int_Y f_x d\nu = \int_Y(f_x)_+d\nu-\int_Y(f_x)_-d\nu$ ist; zweitens, dass die Abbildung $X\setminus N\longrightarrow\mathbb{R}$, $x\mapsto\int_Y f_x d\nu$, als Differenz der integrierbaren Funktionen $x\mapsto\int_Y(f_x)_\pm d\nu : X\setminus N\longrightarrow\mathbb{R}$ wieder integrierbar auf $X\setminus N$ ist und somit *μ-fast überall gleich einer μ-integrierbaren* Abbildung $X\to\mathbb{R}$. Insbesondere haben die Integrale $\int_X(\int_Y(f_x)_\pm d\nu)d\mu(x), \int_X(\int_Y f_x d\nu)d\mu(x)\in\mathbb{R}$ wohl Sinn und es folgt wieder aus Teil 1, dass

$$\begin{aligned}\int_X(\int_Y f_x d\nu)d\mu(x) &= \int_X(\int_Y(f_x)_+d\nu-\int_Y(f_x)_-d\nu)d\mu\\ &= \int_X(\int_Y(f_x)_+d\nu)d\mu-\int_X(\int_Y(f_x)_+d\mu)d\nu\\ &= \int_{X\times Y} f_+d\mu\otimes\nu-\int_{X\times Y} f_-d\mu\otimes\nu\\ &= \int_{X\times Y} f d\mu\otimes\nu.\end{aligned}$$

Die enstprechenden Aussagen und Identitäten für f_y sind analog. Dies beweist den Teil 2 und somit den Satz 3.63. $\square$

Bemerkungen 3.64

1. Der Satz von Fubini (Teil 1) impliziert insbesondere, dass eine messbare Funktion $f : X \times Y \longrightarrow \mathbb{R}$ (bzw. $f : X \times Y \longrightarrow \mathbb{C}$ oder $f : X \times Y \longrightarrow \mathbb{R}^p$) genau dann $\mu \otimes \nu$-integrierbar ist, wenn $\int_Y (\int_X |f_y| d\mu) d\nu(y)$ oder $\int_X (\int_Y |f_x| d\nu) d\mu(x)$ endlich ist; denn beide Integrale sind gleich $\int_{X\times Y} |f| d\mu \otimes \nu$.

2. Im Fall, wo nicht mehr auf ganz $X \times Y$ sondern auf einer Teilmenge $E \in \mathcal{A} \otimes \mathcal{B}$ integriert wird, nehmen die Identitäten im Satz von Fubini folgende Gestalt an:

$$\int_E f d\mu \otimes \nu = \int_{E_1} (\int_{E_x} f(x, y) d\nu(y)) d\mu(x) = \int_{E_2} (\int_{E_y} f(x, y) d\mu(x)) d\nu(y)$$

   für jede $\mu \otimes \nu$-integrierbare Funktion f auf E, wobei $E_1 := \{x \in X \mid E_x \neq \varnothing\}$ und $E_2 := \{y \in Y \mid E_y \neq \varnothing\}$.

3.3.3 Das Lebesgue-Maß auf $\mathbb{R}^n$

Definition 3.65 *Sei* $\mu : \mathcal{B}(\mathbb{R}) \longrightarrow [0, \infty]$ *das Borel-Lebesgue-Maß aus* Satz 3.15. *Für* $k \geq 1$ *wird das* k-dimensionale Borel-Lebesgue-Maß $\mu_k : \mathcal{B}(\mathbb{R}^k) \longrightarrow [0, \infty]$ *induktiv definiert durch*

$$\mu_1 := \mu, \qquad \mu_k := \mu_{k-1} \otimes \mu_1 \quad \text{für } k \geq 2.$$

Bemerke, dass nach dem Beispiel 3.59 der Definitionsbereich von μ_k wohl $\mathcal{B}(\mathbb{R}^k)$ ist.

Definition 3.66 *Für* $k \geq 1$ *wird das* k-dimensionale-Lebesgue-Maß λ_k *definiert als die Vervollständigung des* k*-dimensionalen-Borel-Lebesgue-Maßes* μ_k. *Die zugehörige* σ*-Algebra wird mit* $\mathcal{L}_k$ *bezeichnet.*

Insbesondere ist λ_1 das Lebesgue-Maß auf $\mathbb{R}$, siehe Definition 3.19. Nach Definition der Vervollständigung gelten $\mathcal{B}(\mathbb{R}^k) \subset \mathcal{L}_k$ sowie $\lambda_{k|_{\mathcal{B}(\mathbb{R}^k)}} = \mu_k$, für alle $k \in \mathbb{N} \setminus \{0\}$. Man beachte allerdings, dass $\mathcal{L}_k \otimes \mathcal{L}_l \neq \mathcal{L}_{k+l}$ gilt, insbesondere ist $\lambda_k \otimes \lambda_l \neq \lambda_{k+l}$.

Definition 3.67 *Das* n-dimensionale Volumen *eines Elements* $A \in \mathcal{L}_n$ *wird definiert als* $\lambda_n(A) \in [0, \infty]$.

Bemerkungen 3.68

1. Für $n = 2$ wird dieses n-dimensionale Volumen *Flächeninhalt* genannt. Für $n = 3$ stellt das n-dimensionale Volumen wohl das "übliche" Volumen dar: beispielsweise ist für $A = \prod_{j=1}^3 [a_j, b_j]$ (mit $a_j \leq b_j$ reell) $\lambda_3(A) = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3)$ das Volumen des Quaders A, welcher Seitenlängen $b_1 - a_1$, $b_2 - a_2$ und $b_3 - a_3$ hat.

2. Für eine λ_n-integrierbare (oder lediglich nichtnegative messbare) Funktion $f : A \longrightarrow \mathbb{R}$ (wobei $A \in \mathcal{L}_n$) wird $\int_A f d\lambda_n$ auch $\int_A f(x_1, \ldots, x_n) dx_1 \ldots dx_n$ bezeichnet. Dies wird auch für $\mathbb{C}$- oder $\mathbb{R}^p$-wertige λ_n-integrierbare Funktionen verwendet. Im Falle, dass $A = \prod_{j=1}^n I_j$ ein Produkt von Intervallen ist, wird dieses Integral auch mit $\int_{I_1} \cdots \int_{I_n} f(x_1, \ldots, x_n) dx_1 \ldots dx_n$ bezeichnet.

3. Im Fall, wo $f : A \longrightarrow \mathbb{R}$ (oder mit Werten in $\mathbb{C}$ bzw. $\mathbb{R}^p$) eine *stetige* Funktion und, für alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$, die "Schnitte"

$$A_{x_1,\ldots,\widehat{x_j},\ldots,x_n} := \{x \in \mathbb{R} \,|\, (x_1, \ldots, x, \ldots, x_n) \in A\}, \qquad 1 \leq j \leq n-1,$$

Intervalle sind, kann das Integral $\int_A |f| d\lambda_n$ – und auch $\int_A f d\lambda_n$ falls $\int_A |f| d\lambda_n < \infty$ – mit Hilfe des eindimensionalen Riemann-Integrals berechnet werden: nach dem Satz von Fubini (mehrfach angewendet) ist, mit den obigen Bezeichnungen,

$$\int_A |f| d\lambda_n = \int_{A_1} \ldots \int_{A_{x_1,\ldots,x_{n-1}}} |f(x_1, \ldots, x_n)| d\lambda_1(x_1) \ldots d\lambda_1(x_n).$$

Da f stetig ist, ist insbesondere jede partielle Abbildung $x_n \mapsto f(x_1, \ldots, x_n)$ stetig, insbesondere Riemann-integrierbar auf jedem in $A_{x_1,\ldots,x_{n-1}}$ enthaltenen kompakten Intervall. Aus Proposition 3.52 folgt, dass $\int_{A_{x_1,\ldots,x_n}} |f| d\lambda$ genau dann endlich ist, wenn das (möglicherweise uneigentliche) Integral $\int_{A_{x_1,\ldots,x_n}} f(x_1, \ldots, x_n) dx_n$ absolut konvergiert und ggf. sind die beiden Integrale gleich; induktiv wird somit die λ_n-Integrierbarkeit von $|f|$ auf A durch die Riemann-Integrierbarkeit der sukzessiv durch Integration nach $x_n, \ldots, x_1$ gebildeten Funktionen überprüft. In dem Fall liefert auch der Satz von Fubini, dass in diesem Fall (wo f λ_n-integrierbar – oder lediglich messbar und nichtnegativ – ist) die Reihenfolge der Integrationsverfahren nach den sukzessiven Variablen (beim Berechnen von $\int_A f d\lambda_n$) beliebig geändert werden kann.

3.4 Die Transformationsformel

In diesem Abschnitt wollen wir die Transformationsformel der Analysis I für Funktionen mehrerer Veränderlichen verallgemeinern.

Proposition 3.69 *Sei $L \in \mathcal{L}_n$ und $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ der Form $f(x) = A \cdot x + b$, wobei $A \in M_{n\times n}(\mathbb{R})$ und $b \in \mathbb{R}^n$. Dann ist $f(L) \in \mathcal{L}_n$ und es gilt*

$$\lambda_n(f(L)) = |\det(A)| \cdot \lambda_n(L).$$

Insbesondere gilt $\lambda_n(f(L)) = \lambda_n(L)$, sobald f eine Isometrie ist, d.h., wenn A eine Orthogonalmatrix ist.

Beweis: Wegen der Eindeutigkeit der Vervollständigung auf $\mathcal{L}_n$ (siehe Proposition 3.18) reicht es, die Identität für μ_n zu beweisen. Nach der Translationsinvarianz des Borel-Lebesgue-Maßes (Proposition A.18) können wir $b = 0$ annehmen. Wir können auch $\mathrm{rg}(A) = n$ annehmen: ist nämlich A nicht invertierbar, so ist das Bild von A in einem $n-1$-dimensionalen Untervektorraum des $\mathbb{R}^n$ enthalten; da jeder $n-1$-dimensionale Untervektorraum des $\mathbb{R}^n$ Maß Null hat (bzgl. μ_n), gilt insbesondere $\mu_n(A(\mathbb{R}^n)) = 0$ und die Identität ist wegen $\det(A) = 0$ bewiesen. Von hier aus sei also A invertierbar. Die *Polarzerlegung* aus der Linearalgebra I liefert die Existenz einer Orthogonalmatrix $Q \in M_{n\times n}(\mathbb{R})$ und einer positiv-definiten symmetrischen Matrix $S \in M_{n\times n}(\mathbb{R})$ mit $A = QS$. Da μ_n unvariant unter Isometrien ist (Proposition A.19) und $|\det(Q)| = 1$ für jede Orthogonalmatrix Q gilt, können wir o.B.d.A. $Q = \mathrm{I}_n$ voraussetzen, d.h., $A = S$. Da nun A symmetrisch und positiv-definit ist, existiert eine *Orthogonalmatrix* $P \in M_{n\times n}(\mathbb{R})$ mit

$P^{-1}AP = D := \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$, wobei $\alpha_1, \ldots, \alpha_n \in]0, \infty[$ die Eigenwerte von A sind.

Nun ist

$$\mu_n(D \cdot [0,1[^n) = \mu_n(\prod_{i=1}^{n} [0, \alpha_i[) = \prod_{i=1}^{n} \alpha_i = \det(D) = \det(A).$$

Da $B \mapsto \mu_n(D \cdot B)$ wieder ein translationsinvariantes Maß definiert, folgt $\mu_n(D \cdot B) = \det(A)\mu_n(B)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$ (nach Proposition A.18). Zusammen mit Proposition A.19 folgt

$$\begin{aligned} \mu_n(A \cdot B) &= \mu_n(PDP^{-1}B) \\ &= \mu_n(DP^{-1}B) \qquad \text{wegen } P \text{ orthogonal} \\ &= \det(A) \cdot \mu_n(P^{-1}B) \\ &= \det(A) \cdot \mu_n(B) \qquad \text{wegen } P \text{ orthogonal,} \end{aligned}$$

was Proposition 3.69 beweist. □

Satz 3.70 (Transformationsformel) *Sei $f : V \longrightarrow [0, \infty]$ (bzw. mit Werten in $\mathbb{C}$ oder $\mathbb{R}^p$) eine auf einer offenen Teilmenge V des $\mathbb{R}^n$ definierte messbare (bzw. λ_n-integrierbare) Abbildung. Ist $\varphi : U \longrightarrow V$ ein C^1-Diffeomorphismus, so ist $f \circ \varphi \cdot |\det(J_\varphi)| : U \longrightarrow \mathbb{R}$ ebenfalls messbar (bzw. λ_n-integrierbar) und es gilt*

$$\int_V f d\lambda_n = \int_U f \circ \varphi \cdot |\det(J_\varphi)| d\lambda_n. \tag{3.3}$$

Der Beweis dieses Satzes würde den Rahmen dieser Vorlesung sprengen, siehe z.B. [6, Abschn. V.4] oder [7, Kap. 10].

Bemerkung 3.71 Für $n = 1$, für ein Intervall $U =]a, b[$ und eine Riemann-integrierbare Funktion $f : \varphi(U) \longrightarrow \mathbb{R}$, wobei $\varphi :]a, b[\longrightarrow \varphi(]a, b[)$ ein C^1-Diffeomorphismus ist, gilt nach der Transformationsformel aus der Analysis I: $\int_{\varphi(a)}^{\varphi(b)} f(x)dx = \int_a^b f(\varphi(x))\varphi'(x)dx$. Dabei muss darauf aufgepasst werden, dass kein Betrag in dieser Formel auftaucht. Jedoch stimmt diese Formel mit der Transformationsformel (3.3) überein. Ist nämlich φ orientierungserhaltend, d.h., gilt $\varphi' > 0$ auf $]a, b[$, so ist $|\varphi'| = \varphi'$ und es gilt $\int_{\varphi(a)}^{\varphi(b)} f(x)dx = \int_{]\varphi(a), \varphi(b)[} f(x)dx$; ist φ orientierungsumkehrend, d.h., gilt $\varphi' < 0$ auf $]a, b[$, so ist $|\varphi'| = -\varphi'$ und es gilt $\int_{\varphi(a)}^{\varphi(b)} f(x)dx = -\int_{]\varphi(b), \varphi(a)[} f(x)dx$ (wegen $\varphi' < 0$ ist φ monoton fallend, somit gilt $\varphi(b) < \varphi(a)$ und daher $\varphi(]a, b[) =]\varphi(b), \varphi(a)[$).

Korollar 3.72 *Sei $f : V \longrightarrow \mathbb{R}$ (oder mit Werten in $\mathbb{C}$ bzw. $\mathbb{R}^p$) eine auf einer offenen Teilmenge V des $\mathbb{R}^n$ definierte λ_n-integrierbare Abbildung.*

i) (Transformationsformel für Polarkoordinaten)) *Für $n = 2$ und V der Form $V = \{(r\cos(\varphi), r\sin(\varphi)) \,|\, r \in]r_1, r_2[, \ \varphi \in]\varphi_1, \varphi_2[\}$, wobei $0 \leq r_1 \leq r_2 \leq \infty$ und $-\pi \leq \varphi_1 \leq \varphi_2 \leq \pi$, gilt*

$$\int_V f d\lambda_2 = \int_{]r_1, r_2[} \int_{]\varphi_1, \varphi_2[} f(r\cos(\varphi), r\sin(\varphi)) r dr d\varphi.$$

ii) (Transformationsformel für Zylinderkoordinaten)) *Für* $n = 3$ *und* V *der Form* $V = \{(r\cos(\varphi), r\sin(\varphi), z) \mid r \in]r_1, r_2[, \varphi \in]\varphi_1, \varphi_2[, z \in]z_1, z_2[\}$, *wobei* $0 \le r_1 \le r_2 \le \infty$, $-\pi \le \varphi_1 \le \varphi_2 \le \pi$ *und* $-\infty \le z_1 \le z_2 \le \infty$, *gilt*

$$\int_V f d\lambda_3 = \int_{]r_1,r_2[} \int_{]\varphi_1,\varphi_2[} \int_{]z_1,z_2[} f(r\cos(\varphi), r\sin(\varphi), z)\, r dr d\varphi dz.$$

iii) (Transformationsformel für Kugelkoordinaten)) *Für* $n = 3$ *und* V *der Form* $V = \{(r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi)) \mid r \in]r_1, r_2[, \theta \in]\theta_1, \theta_2[, \varphi \in]\varphi_1, \varphi_2[\}$, *wobei* $0 \le r_1 \le r_2 \le \infty$, $-\pi \le \theta_1 \le \theta_2 \le \pi$ *und* $-\frac{\pi}{2} \le \varphi_1 \le \varphi_2 \le \frac{\pi}{2}$, *gilt*

$$\int_V f d\lambda_3 = \int_{]r_1,r_2[} \int_{]\theta_1,\theta_2[} \int_{]\varphi_1,\varphi_2[} f(r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi)) r^2 \cos(\varphi)\, dr d\theta d\varphi.$$

Beweis: Im ersten Fall gilt

$$\det(J_\Phi)_{(r,\varphi)} = \det \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} = r,$$

wobei $\Phi(r, \varphi) = (r\cos(\varphi), r\sin(\varphi))$. Satz 3.70 liefert i). Analog gilt, für die Transformation $\Phi(r, \varphi, z) := (r\cos(\varphi), r\sin(\varphi), z)$,

$$\det(J_\Phi)_{(r,\varphi,z)} = \det \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) & 0 \\ \sin(\varphi) & r\cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix} = r,$$

was ii) beweist. Für $\Phi(r, \theta, \varphi) := (r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi))$ gilt

$$\det(J_\Phi)_{(r,\theta,\varphi)} = \det \begin{pmatrix} \cos(\theta)\cos(\varphi) & -r\sin(\theta)\cos(\varphi) & -r\cos(\theta)\sin(\varphi) \\ \sin(\theta)\cos(\varphi) & r\cos(\theta)\cos(\varphi) & -r\sin(\theta)\sin(\varphi) \\ \sin(\varphi) & 0 & r\cos(\varphi) \end{pmatrix} = r^2\cos(\varphi),$$

was iii) impliziert. □

Beispiele 3.73

1. Betrachte die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto e^{-x^2-y^2}$. Dann ist f eine nichtnegative stetige (und daher messbare) Funktion auf $\mathbb{R}^2$. Die Transformationsformel für Polarkoordinaten (Korollar 3.72.i)) und der Satz von Fubini liefern

$$\begin{aligned} \int_{\mathbb{R}^2} f d\lambda_2 &= \int_{\mathbb{R}^2 \setminus \{(x,0) \mid x \le 0\}} f d\lambda_2 \qquad \text{wegen } \lambda_2(\{(x,0) \mid x \le 0\}) = 0 \\ &= \int_{]0,\infty[} \int_{]-\pi,\pi[} f(r\cos(\varphi), r\sin(\varphi)) r dr d\varphi \\ &= \int_{-\pi}^{\pi} (\int_0^\infty e^{-r^2\cos^2(\varphi) - r^2\sin^2(\varphi)} r dr) d\varphi \\ &= \int_{-\pi}^{\pi} (\int_0^\infty r e^{-r^2} dr) d\varphi \\ &= 2\pi \int_0^\infty r e^{-r^2} dr \\ &= 2\pi \left[-\frac{e^{-r^2}}{2} \right]_0^\infty \\ &= \pi \qquad \text{wegen } e^{-r^2} \underset{r\to\infty}{\longrightarrow} 0. \end{aligned}$$

Das Integral $\int_0^\infty e^{-x^2}dx$ kann daraus hergeleitet werden: es gilt nämlich, wieder nach dem Satz von Fubini:

$$(\int_0^\infty e^{-x^2}dx)^2 = \int_0^\infty e^{-x^2}dx \cdot \int_0^\infty e^{-y^2}dy = \int_{]0,\infty[\times[0,\infty[} e^{-x^2-y^2}dxdy.$$

Wegen der Symmetrien $f(x,y) = f(-x,y) = f(x,-y)$ für alle $(x,y) \in \mathbb{R}^2$ gilt nach der Transformationsformel $\int_{]0,\infty[\times[0,\infty[} e^{-x^2-y^2}dxdy = \frac{1}{4}\int_{\mathbb{R}^2} e^{-x^2-y^2}dxdy$, so dass $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.

2. Sei allgemeiner $V \subset \mathbb{R}^2$ der Form $V = \{(r\cos(\varphi), r\sin(\varphi)) \,|\, r \in]r_1, r_2[,\ \varphi \in]\varphi_1, \varphi_2[\}$ und f eine sogenannte *rotationssymmetrische* Funktion auf V, d.h., die Abbildung $(r,\varphi) \mapsto f(r\cos(\varphi), r\sin(\varphi))$ hänge *nicht von* φ ab. Wir bezeichnen diese Abbildung mit $h :]r_1, r_2[\longrightarrow \mathbb{R}$, also $h(r) := f(r\cos(\varphi), r\sin(\varphi)$ für alle $r \in]r_1, r_2[$, wobei φ beliebig gewählt werden kann. Dann folgt aus Korollar 3.72.i):

$$\int_V f d\lambda_2 = \int_{]r_1,r_2[}\int_{]\varphi_1,\varphi_2[} f(r\cos(\varphi), r\sin(\varphi))rdrd\varphi = (\varphi_2 - \varphi_1)\int_{]r_1,r_2[} h(r)rdr.$$

Insbesondere bekommen wir das Integral $\int_V f d\lambda_2$, sobald wir eine Stammfunktion von $r \mapsto rh(r)$ bestimmen können.

3. Analog sei $V \subset \mathbb{R}^3$ der Form $V = \{(r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi)) \,|\, r \in]r_1, r_2[,\ \theta \in]\theta_1, \theta_2[,\ \varphi \in]\varphi_1, \varphi_2[\}$ und f eine sogenannte *kugelsymmetrische* Funktion auf V, d.h., die Abbildung $(r,\theta,\varphi) \mapsto f(r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi))$ hänge weder von θ noch von φ ab. Bezeichne mit $h :]r_1, r_2[\longrightarrow \mathbb{R}$ die durch $h(r) := f(r\cos(\theta)\cos(\varphi), r\sin(\theta)\cos(\varphi), r\sin(\varphi))$ definierte Funktion, wobei θ und φ beliebig sind. Dann folgt aus Korollar 3.72.iii):

$$\begin{aligned}\int_V f d\lambda_3 &= \int_{]r_1,r_2[}\int_{]\theta_1,\theta_2[}\int_{]\varphi_1,\varphi_2[} h(r)r^2\cos(\varphi)\,drd\theta d\varphi \\ &= (\theta_2 - \theta_1)\cdot(\sin(\varphi_2) - \sin(\varphi_1))\cdot\int_{]r_1,r_2[} h(r)r^2dr.\end{aligned}$$

Insbesondere bekommen wir das Integral $\int_V f d\lambda_3$, sobald wir eine Stammfunktion von $r \mapsto r^2h(r)$ bestimmen können.

4. Die Vertauschbarkeit der Integrationsverfahren nach den sukzessiven Variablen beim Satz von Fubini ist nicht immer möglich! Betrachte z.B. die Abbildung $f :]0,1[^2 \longrightarrow \mathbb{R}$, $(x,y) \mapsto \frac{x^2-y^2}{(x^2+y^2)^2}$. Die Abbildung f ist wohldefiniert und stetig auf $]0,1[^2$, nicht aber (fast-)überall positiv: tatsächlich gilt $f(x,y) = -f(y,x)$ für alle $(x,y) \in]0,1[^2$. Wir berechnen das Integral $\int_0^1(\int_0^1 f(x,y)dy)dx$, d.h., wir bestimmen die Funktion $x \mapsto \int_0^1 f(x,y)dy$ und integrieren sie auf dem Intervall $]0,1[$. Für ein gegebenes $x \in]0,1[$ ist die Funktion $y \mapsto f(x,y)$ stetig und beschränkt – daher Riemann-

integrierbar – und es gilt

$$\begin{aligned}
\int_0^1 f(x,y)dy &= \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}dy \\
&= \int_0^1 \frac{x^2-y^2-x^2+x^2}{(x^2+y^2)^2}dy \\
&= \int_0^1 \frac{2x^2}{(x^2+y^2)^2} - \frac{1}{x^2+y^2}dy \\
&= 2x^2\int_0^1 \frac{dy}{(x^2+y^2)^2} - \int_0^1 \frac{dy}{x^2+y^2} \\
&= \frac{2}{x^2}\int_0^1 \frac{dy}{(1+\frac{y^2}{x^2})^2} - \frac{1}{x^2}\int_0^1 \frac{dy}{1+\frac{y^2}{x^2}} \\
&= \frac{2}{x^2}\int_0^{\frac{1}{x}} \frac{xdu}{(1+u^2)^2} - \frac{1}{x^2}\int_0^{\frac{1}{x}} \frac{xdu}{1+u^2} \qquad \text{(Substitution } u := \frac{y}{x}) \\
&= \frac{2}{x}\int_0^{\frac{1}{x}} \frac{du}{(1+u^2)^2} - \frac{1}{x}\int_0^{\frac{1}{x}} \frac{du}{1+u^2} \\
&= \frac{2}{x}\int_0^{\frac{1}{x}} \Big(\frac{1}{2(1+u^2)} - \frac{1}{4}(\frac{1}{(u-i)^2}+\frac{1}{(u+i)^2})\Big)du - \frac{1}{x}\left[\arctan(u)\right]_0^{\frac{1}{x}} \\
&= \frac{1}{x}\arctan(\frac{1}{x}) + \frac{1}{2x}\left[\frac{1}{u-i}+\frac{1}{u+i}\right]_0^{\frac{1}{x}} - \frac{1}{x}\arctan(\frac{1}{x}) \\
&= \frac{1}{2}(\frac{1}{1-ix}+\frac{1}{1+ix}) \\
&= \frac{1}{1+x^2}.
\end{aligned}$$

Dabei haben wir eine *Partialbruchzerlegung* für $\frac{1}{(1+u^2)^2}$ benutzt. Da die Funktion $x \mapsto \frac{1}{1+x^2}$ stetig auf $[0,1]$ ist, ist sie Riemann-integrierbar und es gilt

$$\begin{aligned}
\int_0^1 (\int_0^1 f(x,y)dy)dx &= \int_0^1 \frac{dx}{1+x^2} \\
&= \left[\arctan(x)\right]_0^1 \\
&= \frac{\pi}{4}.
\end{aligned}$$

Nun ist das Integral $\int_0^1(\int_0^1 f(x,y)dx)dy$ durch Umbennung der Variablen x und y gleich $\int_0^1(\int_0^1 f(y,x)dy)dx$. Mit der obigen Bemerkung ist aber $\int_0^1(\int_0^1 f(y,x)dy)dx = -\int_0^1(\int_0^1 f(x,y)dy)dx = -\frac{\pi}{4}$. Damit folgt $\int_0^1(\int_0^1 f(x,y)dx)dy \neq \int_0^1(\int_0^1 f(x,y)dy)dx$. Dies zeigt insbesondere, dass f nicht λ_2-integrierbar auf $]0,1[^2$ sein kann!

Kapitel 4

Gewöhnliche Differentialgleichungen

4.1 Differentialgleichungen und Cauchy-Probleme

4.1.1 Differentialgleichungen erster Ordnung

Definition 4.1

i) Eine Differentialgleichung erster Ordnung *im $\mathbb{R}^n$ ist eine Gleichung der Form*

$$y' = f(t, y), \tag{4.1}$$

wobei $f : I \times U \longrightarrow \mathbb{R}^n$ eine Abbildung, $I \subset \mathbb{R}$ ein Intervall und $U \subset \mathbb{R}^n$ eine offene Teilmenge ist.

ii) Eine Lösung *der Differentialgleichung* (4.1) *ist eine auf einem Intervall $J \subset I$ definierte ableitbare Abbildung $y : J \longrightarrow U$ mit $y'(t) = f(t, y(t))$ für alle $t \in J$. Diese Lösung heißt genau dann* global*, wenn $J = I$ ist.*

iii) Ein Cauchy-Problem *(oder auch* Anfangswertproblem*) erster Ordnung im $\mathbb{R}^n$ ist ein System der Form*

$$\begin{cases} y' &= f(t, y) \\ y(t_0) &= y_0, \end{cases} \tag{4.2}$$

wobei $y' = f(t, y)$ eine Differentialgleichung erster Ordnung im $\mathbb{R}^n$ ist und $(t_0, y_0) \in I \times U$. Der Punkt y_0 heißt ggf. das Cauchy-Datum *(oder der* Anfangswert*) des Cauchy-Problems* (4.2)*. Eine Lösung des Cauchy-Problems* (4.2) *ist eine Lösung $y : J \longrightarrow U$ der Differentialgleichung $y' = f(t, y)$ mit $y(t_0) = y_0$, wobei J den Punkt t_0 enthält.*

Bemerkung 4.2 In der Literatur wird manchmal, im Fall $n > 1$, von *Systemen* von Differentialgleichungen die Rede. Dabei wird die Differentialgleichung (4.1) durch n sogenannte *skalare* Differentialgleichungen ersetzt: man schreibt $f(t, x) = (f_1(t, x), \ldots, f_n(t, x))$ mit $f_j : I \times U \longrightarrow \mathbb{R}$ und $y = (y_1, \ldots, y_n)$, wobei jedes y_j nun eine reellwertige unbekannte Funktion ist; ggf. ist (4.1) äquivalent zu

$$\begin{cases} y_1' &= f_1(t, y_1, \ldots, y_n) \\ \vdots & \vdots \quad \vdots \\ y_n' &= f_n(t, y_1, \ldots, y_n). \end{cases}$$

Man beachte, dass man jede Differentialgleichung $y_j' = f_j(t, y_1, \ldots, y_n)$ i.A. nicht unabhängig von den anderen lösen kann.

Beispiele 4.3

1. Wir fangen mit der einfachen Differentialgleichung $y' = y$ im $\mathbb{R}^n$ an. Dann ist die Abbildung $(t,x) \overset{f}{\longmapsto} x$ auf $\mathbb{R}\times\mathbb{R}^n$ definiert. Für jedes $y_0 \in \mathbb{R}^n$ definiert $y(t) := e^t \cdot y_0$ die einzige Lösung des entsprechenden Cauchy-Problems (4.2). Diese Lösung ist auf $\mathbb{R}$ definiert und somit global.

2. Betrachte nun die Differentialgleichung $y' = y^2$ im $\mathbb{R}$. Dann ist die Lösung dieser Differentialgleichung mit $y(0) = 1$ gegeben durch $y(t) = \frac{1}{1-t}$. Man beachte, dass y nicht auf $\mathbb{R}$ definiert ist, obwohl die Funktion $(t,x) \overset{f}{\longmapsto} x^2$ auf $\mathbb{R}\times\mathbb{R}$ definiert ist. Wegen $\frac{1}{1-t} \underset{t\to 1^-}{\longrightarrow} \infty$ kann y sogar nicht auf $\mathbb{R}$ zu einer stetigen Funktion fortgesetzt werden.

3. Das letzte Beispiel zeigt auch, dass die globale Lösbarkeit des Cauchy-Problems (4.2) i.A. vom Cauchy-Datum y_0 abhängt. Wählt man z.B. $y_0 = 0$, so ist die Lösung des Cauchy-Problems $y' = y^2$ mit $y(0) = y_0$ die Nullabbildung, welche auf $\mathbb{R}$ wohldefiniert ist.

4.1.2 Differentialgleichungen höherer Ordnung

Definition 4.4 *Sei $k \in \mathbb{N}\setminus\{0\}$.*

i) Eine Differentialgleichung k-ter Ordnung *im $\mathbb{R}^n$ ist eine Gleichung der Form*

$$y^{(k)} = f(t, y, y', \ldots, y^{(k-1)}), \tag{4.3}$$

wobei $f : I \times U_1 \times \ldots \times U_k \longrightarrow \mathbb{R}^n$ eine Abbildung, $I \subset \mathbb{R}$ ein Intervall und $U_1, \ldots, U_k \subset \mathbb{R}^n$ offene Teilmengen sind. Eine Lösung von (4.3) ist eine auf einem Intervall $J \subset I$ definierte k-mal ableitbare Abbildung $y : J \longrightarrow \mathbb{R}^n$ mit $(y(t), y'(t), \ldots, y^{(k-1)}(t)) \in U_1 \times \ldots \times U_k$ und $y^{(k)}(t) = f(t, y(t), y'(t), \ldots, y^{(k-1)}(t))$ für alle $t \in J$.

ii) Ein Cauchy-Problem k-ter Ordnung *im $\mathbb{R}^n$ ist ein System der Form*

$$\begin{cases} y^{(k)} & = f(t, y, y', \ldots, y^{(k-1)}) \\ y(t_0) & = y_0 \\ \vdots & \vdots \ \vdots \\ y^{(k-1)}(t_0) & = y_0^{(k-1)}, \end{cases} \tag{4.4}$$

wobei $y^{(k)} = f(t, y, y', \ldots, y^{(k-1)})$ eine Differentialgleichung k-ter Ordnung im $\mathbb{R}^n$ ist und $(t_0, y_0, \ldots, y_0^{(k-1)}) \in I \times U_1 \times \ldots \times U_k$.

In der Definition 4.4 bezeichnet $y^{(j)}$ die j-te Ableitung der Funktion y; hingegen ist $y_0^{(j)}$ lediglich ein (beliebiger) Punkt aus $\mathbb{R}^n$. Die Bezeichnung soll nur darauf hindeuten, dass $y_0^{(j)}$ das Cauchy-Datum für die Funktion $y^{(j)}$ ist.

Eine Differentialgleichung k-ter Ordnung kann immer auf eine Differentialgleichung erster Ordnung zurückgeführt werden, indem man die $k-1$ ersten Ableitungen als zusätzliche Variablen auffasst:

Proposition 4.5 *Sei $f : I \times U_1 \times \ldots \times U_k \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $U_1, \ldots, U_k \subset \mathbb{R}^n$ offene Teilmengen sind. Dann löst eine auf einem Intervall $J \subset I$ definierte k-mal ableitbare Abbildung $y : J \longrightarrow \mathbb{R}^n$ genau dann die Differentialgleichung (4.3), wenn die Abbildung $Y := (y, y', \ldots, y^{(k-1)}) : J \longrightarrow \mathbb{R}^{nk} = \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{k \text{ mal}}$ folgende Differentialgleichung erster Ordnung löst:*

$$\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(k-1)} \end{pmatrix}' = \begin{pmatrix} y' \\ y'' \\ \vdots \\ f(t, y, y', \ldots, y^{(k-1)}) \end{pmatrix}.$$

Beweis: Der Beweis ist wegen $(y^{(l)})' = y^{(l+1)}$ trivial. □

Anders ausgedrückt: wir setzen $Y := (Y_0, \ldots, Y_{k-1})$, wobei jedes Y_j eine $\mathbb{R}^n$-wertige unbekannte Funktion ist, und lösen die Differentialgleichung erster Ordnung

$$Y' = F(t, Y),$$

wobei $F : I \times U_1 \times \ldots \times U_k \longrightarrow \mathbb{R}^{nk}$, $F(t, Y) := \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{k-1} \\ f(t, Y_0, Y_1, \ldots, Y_{k-1}) \end{pmatrix}$. Die Komponente Y_0 ist dann eine Lösung der Differentialgleichung (4.3).

Analog ist das Cauchy-Problem (4.4) äquivalent zum Cauchy-Problem

$$\begin{cases} Y' &= F(t, Y) \\ Y(t_0) &= (y_0, \ldots, y_0^{(k-1)}), \end{cases}$$

welches wohl der Form (4.2) ist.

Aus diesen Beobachtungen folgt, dass das Studium Differentialgleichungen k-ter Ordnung auf das der Differentialgleichungen erster Ordnung reduziert werden kann.

Beispiele 4.6

1. Die Differentialgleichung zweiter Ordnung $y'' = y$ im $\mathbb{R}^n$ besitzt globale Lösungen. Tatsächlich gibt es zu jedem Paar $(y_0, y_0') \in (\mathbb{R}^n)^2 = \mathbb{R}^{2n}$ von Cauchy-Daten eine eindeutige Lösung y dieser Differentialgleichung mit $y(0) = y_0$ und $y'(0) = y_0'$, nämlich $y : \mathbb{R} \longrightarrow \mathbb{R}^n$, $y(t) := \cosh(t) \cdot y_0 + \sinh(t) \cdot y_0'$. Wie oben kann diese Differentialgleichung als Differentialgleichung erster Ordnung aufgefasst werden: die entsprechende Differentialgleichung lautet $\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} y' \\ y \end{pmatrix}$, d.h.,

$$Y' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot Y.$$

   Im Abschnitt 4.2 lernen wir, wie wir diese lineare Differentialgleichung lösen können.

2. Die newtonsche Bewegungsgleichung eines Teilchens im Raum lautet

$$m\ddot{x} = F(x),$$

wobei x die Position des Teilchens (aufgefasst als Kurve im $\mathbb{R}^3$) und F die Kraft (aufgefasst als Vektorfeld im $\mathbb{R}^3$) bezeichnen. Dies ist wohl eine Differentialgleichung zweiter Ordnung im $\mathbb{R}^3$. Im Fall einer affin-linearen Funktion F, d.h. $F(x) = A \cdot x + b$ mit $A \in M_{3\times 3}(\mathbb{R})$ und $b \in \mathbb{R}^3$, hat diese Gleichung eine globale Lösung (auf $\mathbb{R}$), siehe Abschnitt 4.2 unten.

4.1.3 Existenz und Eindeutigkeit von Lösungen

Im Allgemeinen kann eine gegebene Differentialgleichung *nicht* explizit gelöst werden. Die *Existenz* einer Lösung ist aber für viele mit Differentialgleichungen verbundene physikalische Probleme von großer Bedeutung. Außerdem kann man sich fragen, *wieviele* Lösungen es geben kann, d.h., ob eine gewisse Lösung *eindeutig* ist. In diesem Abschnitt besprechen wir folgende Fragen:

1. Hat eine gegebene Differentialgleichung immer (mindestens) eine Lösung?
2. Wenn ja, wieviele Lösungen gibt es? Ist das entsprechende Cauchy-Problem immer lösbar und ggf. wie hängt die Lösung vom Cauchy-Datum ab?
3. Wenn nicht, gibt es zumindest eine *lokale* Lösung, also eine Lösung, die auf einem hinreichend kleinen Intervall definiert ist?
4. Im Fall, wo die Funktion f von einem weiteren Parameter abhängt: wie hängt die/eine Lösung vom Parameter ab?

Zuerst erinnern wir an den folgenden Begriff:

Definition 4.7 *Sei U eine offene Teilmenge des $\mathbb{R}^n$.*

i) Eine Abbildung $f : I \times U \longrightarrow \mathbb{R}^n$, $(t, x) \longmapsto f(t, x)$, heißt genau dann Lipschitz-stetig in $x \in U$, *wenn ein $L \in [0, \infty[$ so existiert, dass*

$$|f(t, x) - f(t, y)| \leq L \cdot |x - y|$$

für alle *$t \in I$ und $x, y \in U$ gilt.*

ii) Eine Abbildung $f : I \times U \longrightarrow \mathbb{R}^n$, $(t, x) \longmapsto f(t, x)$, heißt genau dann lokal Lipschitz-stetig in $x \in U$, *wenn zu jedem $(\hat{t}, \hat{x}) \in I \times U$ Umgebungen $I_{\hat{t}}$ von $\hat{t}$ in I und $U_{\hat{x}}$ von $\hat{x}$ in U so existieren, dass $f_{|_{I_{\hat{t}} \times U_{\hat{x}}}}$ Lipschitz-stetig in $x \in U_{\hat{x}}$ ist.*

Die lokale Lipschitz-Stetigkeit von f erlaubt es, die *Eindeutigkeit* der Lösungen des Cauchy-Problems (4.2) zu untersuchen.

Proposition 4.8 *Sei $f : I \times U \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall und $U \subset \mathbb{R}^n$ eine offene Teilmenge ist. Angenommen, f sei* stetig auf $I \times U$[1] *und* lokal Lipschitz-stetig in $x \in U$. *Sei $(t_0, y_0) \in I \times U$ beliebig. Dann gilt: sind y und z zwei auf einem t_0 enthaltenden Intervall $J \subset I$ definierte Lösungen des Cauchy-Problems* (4.2) *mit Cauchy-Datum y_0, so gilt $y = z$.*

[1]Diese Annahme fehlte in der Vorlesung.

Beweis: Man betrachtet $G := \{t \in J \,|\, y(t) = z(t)\} \subset J$ und zeigt, dass G eine nichtleere offene und abgeschlossene Teilmenge von J ist; wegen J zusammenhängend (siehe Beispiel 2.47.1) folgt dann $G = J$, was zu zeigen ist. Nach Voraussetzung gilt wegen $y(t_0) = y_0 = z(t_0)$ bereits $t_0 \in J$. Da y und z als ableitbare Funktionen automatisch stetig sind, ist G abgeschlossen in J. Sei nun $\hat{t} \in G$ und $\hat{y} := y(\hat{t}) = z(\hat{t}) \in U$. Da nach Voraussetzung f lokal Lipschitz-stetig in $x \in U$ ist, existiert eine offene Umgebung von $(\hat{t}, \hat{y})$ der Form $I_{\hat{t}} \times U_{\hat{y}}$ in $I \times U$ (wobei $I_{\hat{t}}$ ein offenes Intervall ist) und eine Konstante $L \in [0, \infty[$ mit $|f(t,x) - f(t,x')| \leq L \cdot |x - x'|$ für alle $t \in I_{\hat{t}}$ und $x, x' \in U_{\hat{y}}$. Da y und z stetig sind, kann bis auf Verkleinerung von $I_{\hat{t}}$ angenommen werden, dass $y(t), z(t) \in U_{\hat{y}}$ für alle $t \in I_{\hat{t}} \cap J$ gilt. Wir schreiben jetzt die Lösungen y und z der Differentialgleichung $y' = f(t,y)$ als Lösungen einer Integralgleichung um: da nämlich f stetig ist, ist $s \mapsto f(s, y(s))$ auf jedem kompakten Intervall in J stetig und daher Riemann-integrierbar, insbesondere gilt nach Integration der Differentialgleichung zwischen $\hat{t}$ und einem beliebigen $t \in I_{\hat{t}} \cap J$:

$$y(t) - y(\hat{t}) = \int_{\hat{t}}^{t} y'(s)ds = \int_{\hat{t}}^{t} f(s, y(s))ds$$

und analog $z(t) - z(\hat{t}) = \int_{\hat{t}}^{t} f(s, z(s))ds$. Wegen $y(\hat{t}) = z(\hat{t})$ folgt dann $y(t) - z(t) = \int_{\hat{t}}^{t} f(s, y(s)) - f(s, z(s))ds$ und, mit $y(s), z(s) \in U_{\hat{y}}$ für alle $t \in I_{\hat{t}} \cap J$,

$$\begin{aligned} |y(t) - z(t)| &\leq \int_{\hat{t}}^{t} |f(s, y(s)) - f(s, z(s))|ds \\ &\leq L \int_{\hat{t}}^{t} |y(s) - z(s)|ds \end{aligned}$$

für alle $t \in I_{\hat{t}} \cap J$ mit $\hat{t} \leq t$ (bzw. $|y(t) - z(t)| \leq L \int_{t}^{\hat{t}} |y(s) - z(s)|ds$ falls $t \leq \hat{t}$). Die Funktion $t \mapsto g(t) := \int_{\hat{t}}^{t} |y(s) - z(s)|ds$, $I_{\hat{t}} \cap J \cap [\hat{t}, \infty[\to \mathbb{R}$, ist damit ableitbar (siehe Satz 3.54) mit $g'(t) \leq Lg(t)$ für alle $t \in I_{\hat{t}} \cap J \cap [\hat{t}, \infty[$. Multipliziert man beide Seiten mit dem Exponentiel $t \mapsto e^{-Lt}$, so bekommt man[2]

$$e^{-Lt} g'(t) - Le^{-Lt} g(t) \leq 0,$$

d.h., $(e^{-Lt} g)'(t) \leq 0$ für alle $t \in I_{\hat{t}} \cap J \cap [\hat{t}, \infty[$. Nach Integration zwischen $\hat{t}$ und t bekommt man $e^{-Lt} g(t) - e^{-L\hat{t}} g(\hat{t}) \leq 0$, wobei $g(\hat{t}) = \int_{\hat{t}}^{\hat{t}} |y(s) - z(s)|ds = 0$, somit gilt $g(t) \leq 0$ für alle $t \in I_{\hat{t}} \cap J \cap [\hat{t}, \infty[$. Da aber per Definition $g(t) \geq 0$, muss $g(t) = 0$ für alle $t \in I_{\hat{t}} \cap J \cap [\hat{t}, \infty[$ gelten und damit wegen $|y(s) - z(s)| \geq 0$ für alle s auch $|y(s) - z(s)| = 0$ für alle $s \in [\hat{t}, t]$ und insgesamt für alle $s \in I_{\hat{t}} \cap J \cap [\hat{t}, \infty[$. Das Argument für $s \in I_{\hat{t}} \cap J \cap]-\infty, \hat{t}]$ ist analog. Insgesamt folgt $y(s) = z(s)$ für alle $s \in I_{\hat{t}} \cap J$, d.h., $I_{\hat{t}} \cap J \subset G$. Dies beweist, dass G offen in J ist und damit die Proposition. □

Die Existenz einer globalen Lösung des Cauchy-Problems (4.2) gilt allerdings nicht immer, wie wir bereits im Beispiel 4.3.2 gesehen haben. Dennoch gibt es immer eine *lokale* Lösung, sobald f *stetig* und *lokal Lipschitz-stetig* in x ist:

Satz 4.9 (Satz von Picard-Lindelöf) *Sei $f : I \times U \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall und $U \subset \mathbb{R}^n$ eine offene Teilmenge ist. Angenommen, f sei* stetig auf $I \times U$ *und* lokal Lipschitz-stetig in $x \in U$. *Dann gilt:*

[2]Dieser Trick zum "Lösen" einer linearen Differentialungleichung wird manchmal in der Literatur *Lemma von Grönwall* genannt.

i) Für jedes $(t_0, y_0) \in I \times U$ existieren $\varepsilon > 0$ und $r > 0$ so, dass für alle $\hat{y_0} \in B_r(y_0)$ das Cauchy-Problem (4.2) mit Cauchy-Daten $(t_0, \hat{y_0})$ eine eindeutige Lösung auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ besitzt.

ii) Diese Lösung ist C^{k+1}, sobald f C^k ist ($k \in \mathbb{N}$).

iii) Diese Lösung hängt stetig von $\hat{y_0}$ ab: die Abbildung $B_r(y_0) \longrightarrow C^0([t_0-\varepsilon, t_0+\varepsilon], U)$, $\hat{y_0} \longmapsto y$, wobei y die Lösung des Cauchy-Problems (4.2) mit Cauchy-Daten $(t_0, \hat{y_0})$ ist, ist stetig.

Letzteres bedeutet: $\sup\limits_{t\in[t_0-\varepsilon,t_0+\varepsilon]} (|\overline{y}(t) - \hat{y}(t)|) \underset{y_0\to\hat{y_0}}{\longrightarrow} 0$, wobei $\overline{y}$ (bzw. $\hat{y}$) die zu den Cauchy-Daten $(t_0, \overline{y_0})$ (bzw. $(t_0, \hat{y_0})$) gehörige Lösung des Cauchy-Problems (4.2) ist, für alle $(t_0, \overline{y_0})$, $(t_0, \hat{y_0}) \in I \times U$.

Beweis von Satz 4.9: Der Beweis beruht auf dem Banach'schen Fixpunktsatz.
i) Nach Voraussetzung existieren $L \in [0, \infty[$, $\eta > 0$ und $R > 0$ so, dass $|f(t,x) - f(t,x')| \leq L \cdot |x - x'|$ für alle $t \in [t_0 - \eta, t_0 + \eta] \subset I$ und alle $x, x' \in \overline{B}_R(y_0) \subset U$ gilt. Setze $r := \frac{R}{2}$ und

$$\varepsilon := \min\Big(\eta, \frac{1}{2L}, \frac{R}{4 \cdot \max\limits_{\substack{s\in[t_0-\eta,t_0+\eta] \\ v\in\overline{B}_{\frac{R}{2}}(y_0)}} (|f(s,v)|)}\Big).$$

Erstens ist zu bemerken, dass ε wohldefiniert und positiv ist: ist $L = 0$, so ist $\frac{1}{2L} = \infty$, welches zur Definition von ε nicht beiträgt (wegen $\eta < \infty$); desweiteren ist f als stetige Funktion auf der *kompakten* Teilmenge $[t_0 - \eta, t_0 + \eta] \times \overline{B}_{\frac{R}{2}}(y_0)$ *beschränkt*, insbesondere ist $\max\limits_{\substack{s\in[t_0-\eta,t_0+\eta] \\ v\in\overline{B}_{\frac{R}{2}}(y_0)}} (|f(s,v)|) < \infty$; da alle drei Zahlen in der Definition von ε positiv sind, ist ε positiv. Nun betrachten wir den Raum $X := C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n) := \{g : [t_0 - \varepsilon, t_0 + \varepsilon] \longrightarrow \mathbb{R}^n \,|\, g \text{ stetig}\}$ mit der Metrik $d(g_1, g_2) := \max\limits_{s\in[t_0-\varepsilon,t_0+\varepsilon]} (|g_1(s) - g_2(s)|)$. Beachte, dass wegen $[t_0 - \varepsilon, t_0 + \varepsilon]$ kompakt die Metrik d wohldefiniert ist. Da der gleichmäßige Limes einer Folge von stetigen Funktionen wieder eine stetige Funktion ist, ist (X, d) vollständig (*Übungsaufgabe*). Wir fixieren $\hat{y_0} \in \overline{B}_r(y_0)$ und bezeichnen die konstante Funktion $t \mapsto \hat{y_0}$, $[t_0 - \varepsilon, t_0 + \varepsilon] \to \mathbb{R}^n$, mit $\underline{\hat{y_0}}$. Sei $\overline{B}_r(\underline{\hat{y_0}}) := \{g \in X \,|\, d(g, \underline{\hat{y_0}}) \leq r\} \subset X$ der abgeschlossene r-Ball um $\underline{\hat{y_0}}$ in X. Wir definieren

$$\begin{aligned} F : \overline{B}_r(\underline{\hat{y_0}}) &\longrightarrow X \\ y &\longmapsto \left(t \mapsto \hat{y_0} + \int_{t_0}^{t} f(s, y(s))ds\right). \end{aligned}$$

Die Abbildung F ist wohldefiniert, weil $s \mapsto f(s, y(s))$ wohldefiniert (nach Definition von r und wegen $\varepsilon \leq \eta$) und stetig – somit integrierbar – auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ ist, daher ist $t \mapsto \int_{t_0}^{t} f(s, y(s))ds$ wohldefiniert und stetig, siehe Satz 3.54. Nun wollen wir zeigen, dass $F(\overline{B}_r(\underline{\hat{y_0}})) \subset \overline{B}_r(\underline{\hat{y_0}})$ gilt und dass $F^k : \overline{B}_r(\underline{\hat{y_0}}) \to \overline{B}_r(\underline{\hat{y_0}})$ für ein hinreichend großes $k \in \mathbb{N}$

eine Kontraktion ist. Für alle $y \in \overline{B}_r(\underline{\hat{y}_0})$ und alle $t \in [t_0, t_0+\varepsilon]$ gilt

$$\begin{aligned}
|F(y)(t) - \hat{y}_0| &= |\int_{t_0}^t f(s,y(s))ds| \\
&\leq \int_{t_0}^t |f(s,y(s))|ds \\
&\leq \int_{t_0}^t |f(s,y(s)) - f(s,\hat{y}_0)| + |f(s,\hat{y}_0)|ds \\
&\leq L\int_{t_0}^t \underbrace{|y(s)-\hat{y}_0|}_{\leq r} ds + \int_{t_0}^t |f(s,\hat{y}_0)|ds \\
&\leq L\cdot r\cdot(t-t_0) + (t-t_0)\cdot \max_{\substack{s\in[t_0-\eta,t_0+\eta] \\ v\in \overline{B}_{\frac{R}{2}}(y_0)}} (|f(s,v)|) \\
&\leq \varepsilon\cdot(Lr + \max_{\substack{s\in[t_0-\eta,t_0+\eta] \\ v\in \overline{B}_{\frac{R}{2}}(y_0)}} (|f(s,v)|)) \\
&\leq \frac{r}{2} + \frac{r}{2} \qquad \text{nach Definition von } \varepsilon
\end{aligned}$$

und analog $|F(y)(t) - \hat{y}_0| \leq r$ für alle $t \in [t_0-\varepsilon, t_0]$. Damit gilt $d(F(y), \underline{\hat{y}_0}) \leq r$, d.h., $F(y) \in \overline{B}_r(\underline{\hat{y}_0})$. Dies zeigt $F(\overline{B}_r(\underline{\hat{y}_0})) \subset \overline{B}_r(\underline{\hat{y}_0})$. Sind nun $y, z \in \overline{B}_r(\underline{\hat{y}_0})$, so gilt, für alle $t \in [t_0, t_0+\varepsilon]$,

$$\begin{aligned}
|F(y)(t) - F(z)(t)| &\leq \int_{t_0}^t |f(s,y(s)) - f(s,z(s))|ds \\
&\leq L\cdot\int_{t_0}^t |y(s)-z(s)|ds, \qquad (4.5)
\end{aligned}$$

so dass

$$|F(y)(t) - F(z)(t)| \leq L\cdot d(y,z)\cdot(t-t_0) \qquad (4.6)$$

für alle $t \in [t_0, t_0+\varepsilon]$. Analog gilt $|F(y)(t) - F(z)(t)| \leq L\cdot d(y,z)\cdot(t_0-t)$ für alle $t \in [t_0-\varepsilon, t_0]$. Für die Abbildung $F^2 := F\circ F : \overline{B}_r(\underline{\hat{y}_0}) \to \overline{B}_r(\underline{\hat{y}_0})$ folgern wir

$$\begin{aligned}
|F^2(y)(t) - F^2(z)(t)| &\overset{(4.5)}{\leq} L\cdot\int_{t_0}^t |F(y)(s) - F(z)(s)|ds \\
&\overset{(4.6)}{\leq} L^2\cdot d(y,z)\cdot\int_{t_0}^t s - t_0 dt \\
&\leq \frac{L^2\cdot(t-t_0)^2}{2}\cdot d(y,z)
\end{aligned}$$

für alle $t \in [t_0, t_0+\varepsilon]$ (und analog für $t \in [t_0-\varepsilon, t_0]$). Induktiv über $k \in \mathbb{N}$ und unter Benutzung von (4.5) und (4.6) zeigen wir für $F^k := F\circ F^{k-1}$:

$$|F^k(y)(t) - F^k(z)(t)| \leq \frac{(L\cdot|t-t_0|)^k}{k!}\cdot d(y,z)$$

für alle $t \in [t_0-\varepsilon, t_0+\varepsilon]$ (verwende lediglich $\int_0^t s^k ds = \frac{t^{k+1}}{k+1}$). Daraus folgt insbesondere $d(F^k(y), F^k(z)) \leq \frac{(L\cdot\varepsilon)^k}{k!}\cdot d(y,z)$ für alle $y, z \in \overline{B}_r(\underline{\hat{y}_0})$. Wegen $\frac{(L\cdot\varepsilon)^k}{k!} \underset{k\to\infty}{\longrightarrow} 0$ existiert

ein $k \geq 1$ mit $d(F^k(y), F^k(z)) \leq \frac{1}{2} \cdot d(y,z)$ für alle $y, z \in \overline{B}_r(\underline{\hat{y}_0})$. Insbesondere ist für dieses k die Abbildung $F^k : \overline{B}_r(\underline{\hat{y}_0}) \to \overline{B}_r(\underline{\hat{y}_0})$ eine Kontraktion auf $\overline{B}_r(\underline{\hat{y}_0})$. Da $\overline{B}_r(\underline{\hat{y}_0})$ im vollständigen metrischen Raum (X, d) abgeschlossen ist, ist $\overline{B}_r(\underline{\hat{y}_0})$ versehen mit der von d induzierten Metrik *vollständig*, siehe Proposition 2.38. Nun liefert der Banach'sche Fixpunktsatz (Satz 2.41), dass F^k einen eindeutigen Fixpunkt besitzt, d.h., es gibt ein eindeutiges $y \in \overline{B}_r(\underline{\hat{y}_0})$ mit $F^k(y) = y$. Dieses y ist notwendigerweise ein Fixpunkt von F selber, denn: es gilt $F^k(F(y)) = F^{k+1}(y) = F(F^k(y)) = F(y)$, d.h., $F(y)$ ist auch ein Fixpunkt von F^k, somit gilt $F(y) = y$ wegen der Eindeutigkeit des Fixpunktes von F^k. Außerdem ist y der einzige Fixpunkt von F, denn: ist z ein weiterer Fixpunkt von F (also $F(z) = z$), so gilt induktiv $F^l(z) = z$ für alle $l \geq 1$, insbesondere $F^k(z) = z$, was mit der Eindeutigkeit des Fixpunktes von F^k wiederum $z = y$ impliziert.
Es bleibt, zu bemerken, dass ein solcher Fixpunkt y das Cauchy-Problem $y' = f(t, y)$, $y(t_0) = \hat{y}_0$, auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ löst. Nämlich ist $F(y) = y$ dazu äquivalent, dass $y(t) = \hat{y}_0 + \int_{t_0}^t f(s, y(s))ds$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ gilt. Nach Konstruktion von ε und y gilt $(t, y(t)) \in [t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B}_r(\hat{y}_0) \subset I \times U$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$. Da f und y nach Konstruktion stetig sind, ist $t \mapsto \int_{t_0}^t f(s, y(s))ds$ ableitbar auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ (Satz 3.54), insbesondere ist y ableitbar mit $y'(t) = f(t, y(t))$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$, wobei $y(t_0) = \hat{y}_0 + 0 = \hat{y}_0$ gilt. Dies beweist, dass y das Cauchy-Problem $y' = f(t, y)$, $y(t_0) = \hat{y}_0$, auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ löst. Dass y die einzige Lösung des Cauchy-Problems $y' = f(t, y)$, $y(t_0) = \hat{y}_0$, auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ ist, folgt unmittelbar aus Proposition 4.8. Bemerke schließlich, dass ε unabhängig von $\hat{y}_0 \in \overline{B}_r(y_0)$ ist.
ii) Wir führen eine Induktion über k durch. Für $k = 0$ ist per Definition jede Lösung $y : J \longrightarrow U$ der Differentialgleichung $y' = f(t, y)$ auf einem Intervall $J \subset I$ ableitbar mit $y'(t) = f(t, y(t))$ für alle $t \in J$; da $t \mapsto f(t, y(t))$ als Verknüpfung stetiger Funktionen wieder stetig ist, ist $y' : J \longrightarrow \mathbb{R}^n$ stetig, somit ist y C^1 auf J. Angenommen, die Behauptung gelte für $k \in \mathbb{N}$. Sei f eine C^{k+1} Abbildung und $y : J \longrightarrow U$ eine beliebige Lösung der Differentialgleichung $y' = f(t, y)$ (auf einem Intervall $J \subset I$). Da f bereits C^k ist, ist y nach Induktionsvoraussetzung C^{k+1} mit Ableitung $t \mapsto f(t, y(t))$; da f C^{k+1} ist, ist $t \mapsto f(t, y(t))$ als Verknüpfung zweier C^{k+1} Funktionen wieder C^{k+1}, insbesondere ist $y' : J \longrightarrow \mathbb{R}^n$ C^{k+1} und somit ist $y : J \longrightarrow U$ C^{k+2}. Dies beweist den Induktionsschritt $k \to k + 1$ und somit die Behauptung.
iii) Seien $y, z \in C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ die Lösungen der Differentialgleichung $y' = f(t, y)$ mit Cauchy-Datum $y(t_0) = \hat{y}_0$ bzw. $z(t_0) = \hat{z}_0$, wobei $\hat{y}_0, \hat{z}_0 \in \overline{B}_r(y_0)$ beliebig sind und ε, r wie in Teil *i)* konstruiert werden (ε, r hängen lediglich von t_0, y_0 und f ab). Dann gilt, nach Integration, $y(t) = \hat{y}_0 + \int_{t_0}^t f(s, y(s))ds$ bzw. $z(t) = \hat{z}_0 + \int_{t_0}^t f(s, z(s))ds$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$. Wie oben kann o.B.d.A. angenommen werden, dass ein $L \in [0, \infty[$ existiert mit $|f(t, x) - f(t, x')| \leq L \cdot |x - x'|$ für alle $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ und $x, x' \in \overline{B}_{2r}(y_0)$. Daraus folgt, für alle $t \in [t_0, t_0 + \varepsilon]$,

$$\begin{aligned} |y(t) - z(t)| &= |\hat{y}_0 - \hat{z}_0 + \int_{t_0}^t f(s, y(s)) - f(s, z(s))ds| \\ &\leq |\hat{y}_0 - \hat{z}_0| + \int_{t_0}^t |f(s, y(s)) - f(s, z(s))|ds \\ &\leq |\hat{y}_0 - \hat{z}_0| + L \int_{t_0}^t |y(s) - z(s)|ds. \end{aligned}$$

Die Funktion $[t_0, t_0 + \varepsilon] \xrightarrow{g} \mathbb{R}$, $t \mapsto \int_{t_0}^t |y(s) - z(s)|ds$, ist ableitbar und erfüllt dann $g'(t) \leq |\hat{y}_0 - \hat{z}_0| + L \cdot g(t)$ für alle $t \in [t_0, t_0 + \varepsilon]$, wobei $g(t_0) = 0$ gilt. Im Fall $L > 0$

multiplizieren wir wie vorhin beide Seiten mit e^{-Lt} und bekommen $(e^{-Lt}g)'(t) \leq |\hat{y}_0 - \hat{z}_0| \cdot e^{-Lt}$ für alle $t \in [t_0, t_0 + \varepsilon]$, somit nach Integration $e^{-Lt}g(t) \leq |\hat{y}_0 - \hat{z}_0| \cdot \frac{e^{-Lt_0} - e^{-Lt}}{L}$, d.h., $g(t) \leq |\hat{y}_0 - \hat{z}_0| \cdot \frac{e^{L(t-t_0)}-1}{L}$. Setzen wir diese letzte Ungleichung in die obige ein, so erhalten wir

$$g'(t) = |y(t) - z(t)| \leq |\hat{y}_0 - \hat{z}_0| + L \cdot |\hat{y}_0 - \hat{z}_0| \cdot \frac{e^{L(t-t_0)} - 1}{L} = |\hat{y}_0 - \hat{z}_0| \cdot e^{L(t-t_0)}$$

für alle $t \in [t_0, t_0 + \varepsilon]$. Für $t \leq t_0$ bekommen wir auf analoge Weise die Ungleichung $|y(t) - z(t)| \leq |\hat{y}_0 - \hat{z}_0| \cdot e^{L(t_0-t)}$ für alle $t \in [t_0 - \varepsilon, t_0]$. Insgesamt folgt

$$d(y, z) = \max_{t \in [t_0 - \varepsilon, t_0 + \varepsilon]} (|y(t) - z(t)|) \leq |\hat{y}_0 - \hat{z}_0| \cdot e^{L \cdot \varepsilon}.$$

Beachte, dass diese Ungleichung im Fall $L = 0$ wegen $y(t) - z(t) = \hat{y}_0 - \hat{z}_0$ (für alle t) immer noch gilt . Daraus folgt, dass die Abbildung $\hat{y}_0 \mapsto y$, $\overline{B}_r(y_0) \longrightarrow C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$, Lipschitz-stetig – und insbesondere stetig – ist. Dies beweist *iii*) und somit den Satz 4.9. □

Satz 4.9 wird in der Literatur manchmal *Satz von Cauchy-Lipschitz* genannt. Die globale Eindeutigkeitsaussage aus Proposition 4.8 erlaubt es, die lokalen Lösungen aus Satz 4.9 auf ein *maximales* Intervall zu erweitern:

Korollar 4.10 *Sei $f : I \times U \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall und $U \subset \mathbb{R}^n$ eine offene Teilmenge ist. Angenommen, f sei* stetig auf $I \times U$ *und* lokal Lipschitz-stetig in $x \in U$. *Dann existiert für jedes $(t_0, y_0) \in I \times U$ eine auf einem* maximalen *offenen Intervall um t_0 definierte Lösung des Cauchy-Problems* (4.2).

Beweis: Für ein beliebiges Paar $(t_0, y_0) \in I \times U$ sei

$$\mathcal{J} := \{J \subset I \,|\, J \text{ Intervall}, t_0 \in J \text{ und } \exists\, y : J \longrightarrow U \text{ Lösung von } y' = f(t, y), y(t_0) = y_0\}$$

und $I_{t_0,y_0} := \bigcup_{J \in \mathcal{J}} J \subset I$ (beachte, dass dieses I_{t_0,y_0} mit dem oben definierten I_{t_0} nichts zu tun hat). Nach Satz 4.9.*i*) existiert ein $\varepsilon > 0$ mit $]t_0 - \varepsilon, t_0 + \varepsilon[\in \mathcal{J}$, insbesondere ist $\mathcal{J}$ nicht leer. Nach Proposition 2.48 ist I_{t_0,y_0} als zusammenhängende Teilmenge von $\mathbb{R}$ ein Intervall. Definiere nun $\widetilde{y} : I_{t_0,y_0} \longrightarrow U$, $J \ni t \mapsto y(t)$, wobei $y : J \longrightarrow U$ das Cauchy-Problem $y' = f(t, y)$, $y(t_0) = y_0$, auf J löst. Zu bemerken ist, dass $\widetilde{y}$ wohldefiniert ist, denn: ist $\overline{y} : \overline{J} \longrightarrow U$ eine Lösung des Cauchy-Problems $y' = f(t, y)$, $y(t_0) = y_0$, auf $\overline{J}$, wobei $\overline{J} \ni t_0$, so lösen y und $\overline{y}$ dasselbe Cauchy-Problem auf dem Intervall $J \cap \overline{J} \ni t_0$, insbesondere folgt aus Proposition 4.8, dass $y(s) = \overline{y}(s)$ für alle $s \in J \cap \overline{J}$ gilt, insbesondere $y(t) = \overline{y}(t)$. Nach Konstruktion von $\widetilde{y}$ löst $\widetilde{y}$ das Cauchy-Problem $y' = f(t, y)$, $y(t_0) = y_0$, auf I_{t_0,y_0}. Nach Proposition 4.8 ist $\widetilde{y}$ die einzige Lösung dieses Cauchy-Problems und nach Definition von I_{t_0,y_0} ist $\widetilde{y}$ eine maximale Lösung dieses Cauchy-Problems: ist $\widehat{y} : \widehat{I} \longrightarrow U$ eine Lösung desselben Cauchy-Problems auf einem Intervall $\widehat{I} \ni t_0$, wobei $\widehat{I} \supset I_{t_0,y_0}$ gilt, so ist per Definition $\widehat{I} \subset I_{t_0,y_0}$, insbesondere gilt $\widehat{I} = I_{t_0,y_0}$ und – mit Proposition 4.8 – $\widehat{y} = \widetilde{y}$.
Es bleibt lediglich, zu zeigen, dass I_{t_0,y_0} offen in I (insbesondere in $\mathbb{R}$) ist. Ist aber $\hat{t}_0 \in I_{t_0,y_0}$, so setze $\hat{y}_0 := \widetilde{y}(t_0) \in U$ und betrachte das Cauchy-Problem $y' = f(t, y)$, $y(\hat{t}_0) = \hat{y}_0$. Nach Satz 4.9.*i*) existiert ein $\varepsilon > 0$ so, dass eine Lösung $\hat{y} :]\hat{t}_0 - \varepsilon, \hat{t}_0 + \varepsilon[\longrightarrow U$ dieses Cauchy-Problems existiert. Da $\widetilde{y}$ bereits dasselbe Cauchy-Problem auf I_{t_0,y_0} löst, gilt $\widetilde{y}(t) = \hat{y}(t)$ auf $I_{t_0,y_0} \cap]\hat{t}_0 - \varepsilon, \hat{t}_0 + \varepsilon[$ (Proposition 4.8), insbesondere ist $\widetilde{y}$ auf

$I_{t_0,y_0} \cup]\hat{t}_0 - \varepsilon, \hat{t}_0 + \varepsilon[$ definiert, d.h., $]\hat{t}_0 - \varepsilon, \hat{t}_0 + \varepsilon[\subset I_{t_0,y_0}$. Dies beweist, dass I_{t_0,y_0} offen ist. Damit ist das Korollar 4.10 gezeigt. $\square$

Anders ausgedrückt: es gibt ein "größtes" offenes Intervall I_{t_0,y_0} mit $t_0 \in I_{t_0,y_0}$ und auf dem eine Lösung des Cauchy-Problems (4.2) existiert; hierbei ist "größtes" so zu verstehen, dass die (eindeutige) Lösung $y : I_{t_0,y_0} \longrightarrow U$ von (4.2) auf kein größeres Intervall fortsetzbar ist. Man beachte, dass dieses Intervall nicht immer ganz I ist, d.h., die Lösung y muss nicht global sein, siehe Beispiel 4.3.2. Schaut man genauer hin, warum eine Lösung eventuell nicht global ist, so stellt man folgendes fest *(ohne Beweis)*: für das maximale Definitionsintervall $]t_{\min}, t_{\max}[$ einer gegebenen Lösung y von (4.2) mit $-\infty < t_{\min} < t_{\max} < \infty$ gilt

- entweder $|y(t)| \underset{t \to t_{\max}}{\longrightarrow} \infty$
- oder $\lim\limits_{t \to t_{\max}} y(t)$ existiert, liegt aber nicht in U

und analog bei $t_{\min}$. *(Bild)*

Ein Kriterium für die Existenz einer *globalen* Lösung ist folgender:

Satz 4.11 *Sei $f : I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ eine stetige Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall ist. Angenommen, es gäbe eine* stetige *Funktion $L : I \longrightarrow [0, \infty[$ so, dass*

$$|f(t,x) - f(t,y)| \leq L(t) \cdot |x - y|$$

für alle *$t \in I$ und $x, y \in \mathbb{R}^n$ gilt. Dann ist das maximale Definitionsintervall jeder Lösung des Cauchy-Problems* (4.2) *gleich I.*

Beweis: Der Beweis folgt dem Beweis von Satz 4.9, indem wir den Banach'schen Fixpunktsatz auf *jedem* (und nicht nur hinreichend kleinen) kompakten Intervall anwendet. Sei $(t_0, y_0) \in I \times \mathbb{R}^n$ beliebig.
Behauptung: *Für jedes $(t_0, y_0) \in I \times \mathbb{R}^n$ und jedes $\varepsilon > 0$ mit $[t_0 - \varepsilon, t_0 + \varepsilon] \subset I$ existiert eine eindeutige Lösung $y : [t_0 - \varepsilon, t_0 + \varepsilon] \longrightarrow \mathbb{R}^n$ des Cauchy-Problems $y' = f(t, y)$, $y(t_0) = y_0$.*
Beweis der Behauptung: Die Eindeutigkeit ist nach Proposition 4.8 gesichert. Für die Existenz betrachten wir – wie im Beweis von Satz 4.9 – die Abbildung

$$\begin{array}{rcl} F : C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n) & \longrightarrow & C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n) \\ y & \longmapsto & \left(t \mapsto y_0 + \displaystyle\int_{t_0}^{t} f(s, y(s)) ds \right). \end{array}$$

Diese Abbildung ist wohldefiniert (da braucht man nicht auf $y(s) \in U$ aufpassen, da hier $U = \mathbb{R}^n$) und es gilt, genau wie vorhin,

$$\begin{array}{rcll} |F(y)(t) - F(z)(t)| & \leq & \displaystyle\int_{t_0}^{t} |f(s, y(s)) - f(s, z(s))| ds & \\ & \leq & \displaystyle\int_{t_0}^{t} L(s) \cdot |y(s) - z(s)| ds & \text{nach Voraussetzung} \\ & \leq & \widehat{L} \cdot \displaystyle\int_{t_0}^{t} |y(s) - z(s)| ds, & \end{array}$$

für alle $t \in [t_0, t_0 + \varepsilon]$ und $y, z \in X := C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$, wobei $\widehat{L} := \max_{t \in [t_0 - \varepsilon, t_0 + \varepsilon]} (L(t))$ *endlich* ist, da L auf dem kompakten Intervall $[t_0 - \varepsilon, t_0 + \varepsilon]$ *stetig* ist, siehe Korollar 2.32. Daraus folgt, auch ähnlich wie im Beweis von Satz 4.9,

$$d(F^k(y), F^k(z)) \leq \frac{\widehat{L}^k \varepsilon^k}{k!} d(y, z)$$

für alle $y, z \in X$ und $k \in \mathbb{N} \setminus \{0\}$, wobei $d(g_1, g_2) := \max_{s \in [t_0 - \varepsilon, t_0 + \varepsilon]} (|g_1(s) - g_2(s)|)$. Das wortwörtliche Übertragen des Arguments im Beweis von Satz 4.9 liefert, dass F einen eindeutigen Fixpunkt im vollständigen metrischen Raum (X, d) hat, welcher das Cauchy-Problem $y' = f(t, y)$, $y(t_0) = y_0$, löst. $\checkmark$

Die Behauptung liefert unmittelbar das Ergebnis. Wäre nämlich z.B. $b := \sup(I_{t_0, y_0}) < \sup(I)$, wobei $I_{t_0, y_0} \subset I$ das maximale Existenzintervall der Lösung $\widetilde{y}$ des Cauchy-Problems $y' = f(t, y)$, $y(t_0) = y_0$, ist (siehe Korollar 4.10), so würde man, für ein hinreichend kleines $\varepsilon > 0$ mit $[b - \varepsilon, b + \varepsilon] \subset I$ und $\inf(I_{t_0, y_0}) < b - \varepsilon$, das Intervall $[b - \varepsilon, b + \varepsilon]$ betrachten. Dann würde die Behauptung die Existenz und Eindeutigkeit einer Lösung $\widehat{y} : [b - \varepsilon, b + \varepsilon] \longrightarrow \mathbb{R}^n$ des Cauchy-Problems $y' = f(t, y)$, $y(b - \varepsilon) = \widetilde{y}(b - \varepsilon)$, liefern; da $\widetilde{y}$ bereits eine Lösung desselben Problems auf $[b - \varepsilon, b[$ ist, würde $\widetilde{y}(t) = \widehat{y}(t)$ für alle $t \in [b - \varepsilon, b[$ gelten und somit würde sich $\widetilde{y}$ auf das strikt größere Intervall $I_{t_0, y_0} \cup [b, b + \varepsilon]$ forsetzen lassen, Widerspruch zur Maximalität von $\widetilde{y}$. Deshalb muss $I_{t_0, y_0} = I$ gelten, was zu beweisen war. $\square$

Der Mittelwertsatz (Satz 2.69) liefert insbesondere folgendes Ergebnis:

Korollar 4.12 *Sei $f : I \times U \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall und $U \subset \mathbb{R}^n$ eine offene Teilmenge ist. Angenommen, f sei C^1. Dann existiert für jedes $(t_0, y_0) \in I \times U$ eine auf einem maximalen offenen Intervall um t_0 definierte Lösung des Cauchy-Problems* (4.2).

Beweis: Ist f C^1, so ist f differenzierbar und somit stetig auf $I \times U$ (siehe Proposition 2.61.*iii*)). Außerdem folgt aus dem Mittelwertsatz (Satz 2.69), für alle (t, x) und $\eta, r > 0$ mit $[t - \eta, t + \eta] \times \overline{B}_r(x) \subset I \times U$ und alle $s \in [t - \eta, t + \eta]$ sowie $y, y' \in \overline{B}_r(x)$:

$$|f(s, y) - f(s, y')| \leq \sup_{\substack{\tau \in [t - \eta, t + \eta] \\ z \in \overline{B}_r(x)}} (\|M_f(\tau, z)\|) \cdot |y - y'|,$$

wobei $M_f(\tau, z) := \left(\frac{\partial f_i}{\partial x_j}(\tau, z)\right)_{1 \leq i,j \leq n}$ die Jacobi-Matrix der Abbildung $x \mapsto f(\tau, x)$ ist. Da aber $(\tau, x) \mapsto J_f(\tau, x)$ nach Voraussetzung stetig auf $I \times U$ ist, ist $(\tau, z) \mapsto M_f(\tau, z)$ auch stetig (die Matrix $M_f(\tau, z)$ besteht aus den n letzten Spalten von $J_f(\tau, x)$), somit gilt $\sup_{\substack{\tau \in [t - \eta, t + \eta] \\ z \in \overline{B}_r(x)}} (\|M_f(\tau, z)\|) < \infty$ (das Produkt $[t - \eta, t + \eta] \times \overline{B}_r(x)$ ist kompakt). Dies zeigt, dass f lokal Lipschitz-stetig in $x \in U$ ist. Korollar 4.10 liefert das Ergebnis. $\square$

Als Letztes besprechen wir sogenannte *parameterabhängige* Differentialgleichungen erster Ordnung, d.h., den Fall, wo die Funktion f von einem zusätzlichen Parameter stetig abhängt. In diesem Fall hängt die Lösung des Cauchy-Problems (4.2) ebenfalls stetig vom Parameter ab, sobald die Funktion f lokal Lipschitz-stetig in $x \in U$ *gleichmäßig* bezüglich des Parameters ist, wie folgender Satz behauptet:

Satz 4.13 *Sei $f : I \times U \times P \longrightarrow \mathbb{R}^n$ eine Abbildung, wobei $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Teilmenge und P ein metrischer Raum (mit Metrik d) ist. Angenommen, f sei* stetig *auf $I \times U \times P$ (mit Produktmetrik),* beschränkt *auf jeder Teilmenge der Form $K \times P$, wobei $K \subset I \times U$ kompakt ist*[3] *und jeder Punkt $(\hat{t}, \hat{y}) \in I \times U$ besitze eine Umgebung der Form $I_{\hat{t}} \times U_{\hat{t}}$ (mit $I_{\hat{t}} \subset \mathbb{R}$ Intervall und $U_{\hat{t}} \subset \mathbb{R}^n$ offen) so, dass eine Konstante $L_{\hat{t},\hat{x}} \in [0, \infty[$ existiert mit*

$$|f(t, x, p) - f(t, x', p)| \leq L_{\hat{t},\hat{x}} \cdot |x - x'|$$

für alle $t \in I_{\hat{t}}$, $x, x' \in U_{\hat{t}}$ und alle *$p \in P$. Sei $(t_0, y_0) \in I \times U$ beliebig.*
Dann existieren $\varepsilon > 0$ und $r > 0$ so, dass für alle $\hat{y}_0 \in B_r(y_0)$ und $p \in P$ das Cauchy-Problem

$$\begin{cases} y' &= f(t, y, p) \\ y(t_0) &= \hat{y}_0 \end{cases} \tag{4.7}$$

mit Cauchy-Daten $(t_0, \hat{y}_0)$ eine eindeutige Lösung auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ besitzt. Außerdem hängt diese Lösung stetig von p ab: die Abbildung $P \longrightarrow C^0([t_0 - \varepsilon, t_0 + \varepsilon], U)$, $p \longmapsto y$, wobei y die Lösung des Cauchy-Problems (4.7) *mit Cauchy-Daten (t_0, y_0) ist, ist stetig.*

Beweis: Die Eindeutigkeit einer Lösung (für ein festes $p \in P$) folgt bereits aus Proposition 4.8, weil f insbesondere lokal Lipschitz-stetig in $x \in U$ ist (für ein festes $p \in P$). Die Existenz einer Lösung läuft analog wie im Beweis von Satz 4.9: für ein beliebiges $(t_0, y_0) \in I \times U$ wähle man $\eta, r > 0$ mit $|f(t, x, p) - f(t, x', p)| \leq L \cdot |x - x'|$ für alle $t \in [t_0 - \eta, t_0 + \eta]$, $x, x' \in \overline{B}_{2r}(y_0)$ und *alle* $p \in P$, wobei $L \in [0, \infty[$ eine geeignete Lipschitz-Konstante ist; nach Voraussetzung ist es möglich. Dann setze man

$$\varepsilon := \min\left(\eta, \frac{1}{2L}, \frac{r}{2 \cdot \max\limits_{\substack{(s,v) \in [t_0 - \eta, t_0 + \eta] \times \overline{B}_r(y_0) \\ p \in P}} (|f(s, v, p)|)}\right)$$

und bemerke, dass wegen der Kompaktheit von $[t_0 - \eta, t_0 + \eta] \times \overline{B}_r(y_0)$ die Abbildung $(s, v, p) \mapsto f(s, v, p)$ nach Voraussetzung auf $[t_0 - \eta, t_0 + \eta] \times \overline{B}_r(y_0) \times P$ *beschränkt* ist, insbesondere ist ε wohldefiniert, positiv und *unabhängig* von $p \in P$. Für ein beliebiges $p \in P$ und ein beliebiges $\hat{y}_0 \in \overline{B}_r(y_0)$ betrachte man dann die Abbildung

$$\begin{aligned} F_p : \overline{B}_r(\underline{\hat{y}_0}) &\longrightarrow X \\ y &\longmapsto \left(t \mapsto \hat{y}_0 + \int_{t_0}^t f(s, y(s), p) ds\right), \end{aligned}$$

wobei, wie vorher, $X := C^0([t_0 - \varepsilon, t_0 + \varepsilon], \mathbb{R}^n)$ mit Supremumsmetrik d und $\overline{B}_r(\underline{\hat{y}_0}) := \{y \in X,\ |y(t) - \hat{y}_0| \leq r\ \forall\, t \in [t_0 - \varepsilon, t_0 + \varepsilon]\} \subset X$ der abgeschlossene r-Ball um die Konstante Funktion $t \mapsto \hat{y}_0$ in (X, d) ist. Wie im Beweis von Satz 4.9 zeigt man, dass $F_p(\overline{B}_r(\underline{\hat{y}_0})) \subset \overline{B}_r(\underline{\hat{y}_0})$ gilt. Dabei geht – wie oben erwähnt wurde – die wesentliche Voraussetzung ein, dass f auf jedem Produkt $K \times P$ beschränkt ist. Wie im Beweis von Satz 4.9 ist dann F_p^k für hinreichend hohes $k \in \mathbb{N}$ eine Kontraktion im vollständigen metrischen Raum $(\overline{B}_r(\underline{\hat{y}_0}), d)$, insbesondere hat F_p (nach Satz 2.41) einen eindeutigen Fixpunkt y, welcher dann (wie im Beweis von Satz 4.9) das Cauchy-Problem $y' = f(t, y, p)$, $y(t_0) = \underline{\hat{y}_0}$, auf $[t_0 - \varepsilon, t_0 + \varepsilon]$ löst. Dabei ist zu bemerken, dass ε weder von $p \in P$ noch von $\hat{y}_0 \in \overline{B}_r(y_0)$

[3] Diese Voraussetzung fehlte in der Vorlesung.

abhängt. Dies zeigt die erste Aussage.
Nun wollen wir die Abbildung $P \longrightarrow C^0([t_0-\varepsilon, t_0+\varepsilon], \mathbb{R}^n)$, $p \mapsto y$, wobei $y : [t_0-\varepsilon, t_0+\varepsilon] \longrightarrow \mathbb{R}^n$ die Lösung des Cauchy-Problems $y' = f(t,y,p)$, $y(t_0) = y_0$, auf $[t_0-\varepsilon, t_0+\varepsilon]$ ist, auf Stetigkeit untersuchen. O.B.d.A. kann man annehmen, dass eine Konstante $L \in [0,\infty[$ so existiert, dass $|f(t,x,p)-f(t,x',p)| \leq L\cdot|x-x'|$ für alle $t \in [t_0-\varepsilon, t_0+\varepsilon]$, $x, x' \in \overline{B}_{2r}(y_0)$ und alle $p \in P$ gilt, wobei r auch wie oben gewählt wird. Wir fixieren $p \in P$ und bezeichnen mit y die Lösung des entsprechenden Cauchy-Problems auf $[t_0-\varepsilon, t_0+\varepsilon]$. Sei $\delta > 0$ beliebig. Da die Funktion $[t_0-\varepsilon, t_0+\varepsilon] \times P \longrightarrow \mathbb{R}^n$, $(s,p) \mapsto f(s,y(s),p)$, stetig und $[t_0-\varepsilon, t_0+\varepsilon]$ kompakt ist, ist diese Funktion *gleichmäßig stetig* in $t \in [t_0-\varepsilon, t_0+\varepsilon]$, d.h., für alle $p \in P$ und $\epsilon > 0$ existiert ein $\rho > 0$ so, dass $|f(s,y(s),p') - f(s,y(s),p)| \leq \epsilon$ für alle $p' \in B_\rho(p)$ und *alle* $s \in [t_0-\varepsilon, t_0+\varepsilon] \times P \longrightarrow \mathbb{R}^n$ gilt, siehe Proposition B.4. Hier setzt man $\epsilon := \frac{\delta}{\varepsilon e^{L\varepsilon}}$ und wählt das entsprechende $\rho > 0$ mit der obigen Eigenschaft. Für ein beliebiges $p' \in B_\rho(p)$ mit zugeordneter Lösung z des Cauchy-Problems $z' = f(t,z,p')$, $z(t_0) = y_0$, auf $[t_0-\varepsilon, t_0+\varepsilon]$, folgt dann, für alle $t \in [t_0, t_0+\varepsilon]$:

$$\begin{aligned}
|y(t)-z(t)| &= \Big|\int_{t_0}^{t} f(s,y(s),p) - f(s,z(s),p')ds\Big| \\
&\leq \int_{t_0}^{t} |f(s,y(s),p) - f(s,z(s),p')|ds \\
&\leq \int_{t_0}^{t} |f(s,y(s),p) - f(s,y(s),p')|ds + \int_{t_0}^{t} |f(s,y(s),p') - f(s,z(s),p')|ds \\
&\leq \frac{\delta}{\varepsilon e^{L\varepsilon}} \cdot \varepsilon + L \cdot \int_{t_0}^{t} |y(s)-z(s)|ds,
\end{aligned}$$

wobei $d(y, \underline{y_0}) \leq r$ und $d(z, \underline{y_0}) \leq r$ verwendet wurden. Die Abbildung $[t_0, t_0+\varepsilon] \xrightarrow{g} \mathbb{R}$, $t \mapsto \int_{t_0}^{t} |y(s)-z(s)|ds$, ist dann ableitbar mit $g'(t) \leq \delta e^{-L\varepsilon} + L\cdot g(t)$ für alle $t \in [t_0, t_0+\varepsilon]$. Der gleiche Trick *à la Grönwall* wie im Beweis von Satz 4.9 liefert $g'(t) \leq \delta e^{L(t-t_0-\varepsilon)}$ für alle $t \in [t_0, t_0+\varepsilon]$, insbesondere $\max\limits_{t\in[t_0,t_0+\varepsilon]} |y(t)-z(t)| \leq \delta$. Ein analoges Argument auf $[t_0-\varepsilon, t_0]$ impliziert ebenfalls $\max\limits_{t\in[t_0-\varepsilon,t_0]} |y(t)-z(t)| \leq \delta$. Insgesamt bekommen wir $\max\limits_{t\in[t_0-\varepsilon,t_0+\varepsilon]} |y(t)-z(t)| \leq \delta$, d.h., $d(y,z) \leq \delta$. Zusammengefasst haben wir, für jedes $p \in P$ und jedes $\delta > 0$ ein $\rho > 0$ so finden können, dass $d(y,z) \leq \delta$ für alle $p' \in B_\rho(p)$ gilt. Dies beweist die stetige Abhängigkeit vom Parameter und damit den Satz 4.13. □

4.2 Lineare Differentialgleichungen

4.2.1 Globale Lösbarkeit und Raum der Lösungen

Definition 4.14

i) Eine lineare *Differentialgleichung erster Ordnung ist eine Differentialgleichung erster Ordnung der Form*

$$y' = A(t)\cdot y + b(t),$$

wobei $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ auf einem Intervall $I \subset \mathbb{R}$ definierte Abbildungen sind.

ii) Allgemeiner heißt, für $k \in \mathbb{N} \setminus \{0\}$, eine Differentialgleichung k-ter Ordnung $y^{(k)} = f(t, y, y', \ldots, y^{(k-1)})$ genau dann linear, *wenn für jedes $t \in I$ die Abbildung $U_1 \times \ldots \times U_k \longrightarrow \mathbb{R}^n$, $(y_0, \ldots, y_{k-1}) \longmapsto f(t, y_0, \ldots, y_{k-1})$,* affin-linear *ist.*

Eine erste bemerkenswerte Eigenschaft der linearen Differentialgleichungen ist die *globale* Existenz und Eindeutigkeit der Lösungen. Satz 4.11 liefert nämlich folgendes

Korollar 4.15 *Seien $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ auf einem offenen Intervall $I \subset \mathbb{R}$ definierte* stetige *Abbildungen. Sei $(t_0, y_0) \in I \times U$ beliebig. Dann besitzt das Cauchy-Problem*

$$\begin{cases} y' &= A(t) \cdot y + b(t) \\ y(t_0) &= y_0 \end{cases}$$

eine eindeutige globale *Lösung $y : I \longrightarrow \mathbb{R}^n$.*

Beweis: Sei $f : I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $(t, x) \mapsto A(t) \cdot x + b(t)$, dann ist per Definition $y' = A(t) \cdot y + b(t)$ äquivalent zu $y' = f(t, y)$. Die Abbildung ist wegen $A : I \to M_{n\times n}(\mathbb{R})$ und $b : I \to \mathbb{R}^n$ stetig auch stetig. Desweiteren gilt, für alle $(t, x, x') \in I \times \mathbb{R}^n \times \mathbb{R}^n$,

$$|f(t, x) - f(t, x')| = |A(t) \cdot (x - x')| \leq \|A(t)\| \cdot |x - x'|,$$

wobei $\|A(t)\|$ die Matrixnorm bezeichnet. Da A stetig ist, ist $L(t) := \|A(t)\|$, $I \longrightarrow \mathbb{R}$, stetig. Satz 4.11 liefert die Behauptung. □

Definition 4.16 *Sei eine lineare Differentialgleichung erster Ordnung im $\mathbb{R}^n$ der Form $y' = A(t) \cdot y + b(t)$ gegeben, mit $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein Intervall ist.*

i) Die homogene Gleichung *zu dieser Differentialgleichung ist die Differentialgleichung $y' = A(t) \cdot y$.*

ii) Im Fall $b \neq 0$ heißt die Differentialgleichung $y' = A(t) \cdot y + b(t)$ inhomogen.

Korollar 4.17 *Sei eine lineare Differentialgleichung erster Ordnung im $\mathbb{R}^n$ der Form $y' = A(t) \cdot y + b(t)$ gegeben, mit $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein Intervall ist.*

i) Ist $y_{\mathrm{Inh}} : I \longrightarrow \mathbb{R}^n$ eine gegebene Lösung der inhomogenen Gleichung $y' = A(t) \cdot y + b(t)$, so ist jede Lösung y von $y' = A(t) \cdot y + b(t)$ der Form $y = y_H + y_{\mathrm{Inh}}$, wobei y_H eine beliebige *Lösung der zugehörigen homogenen Gleichung ist.*

ii) Der Raum der Lösungen der homogenen Gleichung $y' = A(t) \cdot y$ ist ein n-dimensionaler reeller Vektorraum. Genauer gilt, für ein beliebiges $t_0 \in I$: die Abbildung

$$\begin{array}{ccc} \{y : I \longrightarrow \mathbb{R}^n \text{ ableitbar mit } y'(t) = A(t) \cdot y(t) \text{ für alle } t \in I\} & \longrightarrow & \mathbb{R}^n \\ y & \longmapsto & y(t_0) \end{array}$$

ist ein Isomorphismus.

Beweis:
i) Ist $y : I \longrightarrow \mathbb{R}^n$ eine beliebige ableitbare Abbildung, so gilt:

$$\begin{aligned} y'(t) = A(t)\cdot y(t) + b(t) \quad &\Longleftrightarrow \quad y'(t) - y'_{\text{Inh}}(t) = A(t)\cdot y(t) + b(t) - A(t)\cdot y_{\text{Inh}}(t) - b(t) \\ &\Longleftrightarrow \quad (y - y_{\text{Inh}})'(t) = A(t)\cdot (y - y_{\text{Inh}})(t), \end{aligned}$$

wobei die Gleichheit jeweils für alle $t \in I$ gemeint ist. Das heißt, y löst genau dann die Differentialgleichung $y' = A(t)\cdot y + b(t)$ auf I, wenn $y_H := y - y_{\text{Inh}} : I \longrightarrow \mathbb{R}^n$ die homogene Differentialgleichung $y' = A(t)\cdot y$ löst. Dies ist eben die Aussage i).
ii) Sei $L := \{y : I \longrightarrow \mathbb{R}^n$ ableitbar mit $y'(t) = A(t)\cdot y(t)$ für alle $t \in I\} \subset C^1(I, \mathbb{R}^n)$ der Raum aller Lösungen der homogenen Differentialgleichung $y' = A(t)\cdot y$ auf I (jede Lösung ist automatisch C^1, siehe Satz 4.9.ii)). Sind $y, z \in L$, so gilt, für alle $\alpha, \beta \in \mathbb{R}$ und $t \in I$,

$$\begin{aligned} (\alpha y + \beta z)'(t) &= \alpha y'(t) + \beta z'(t) \\ &= \alpha A(t)\cdot y(t) + \beta A(t)\cdot z(t) \\ &= A(t)\cdot(\alpha y(t) + \beta z(t)). \end{aligned}$$

Dies beweist, dass L ein reeller Untervektorraum von $C^1(I, \mathbb{R}^n)$ ist. Außerdem ist die Abbildung $E_0 : L \longrightarrow \mathbb{R}^n$, $y \mapsto y(t_0)$, wohldefiniert und offensichtlich linear. Die Abbildung $F_0 : \mathbb{R}^n \longrightarrow L$, $y_0 \mapsto y$, wobei $y \in L$ die Lösung des Cauchy-Problems $y' = A(t)\cdot y$, $y(t_0) = y_0$, ist, ist nach Korollar 4.15 ebenfalls wohldefiniert (diese Lösung existiert und ist eindeutig) mit $E_0 \circ F_0 = \text{id}_{\mathbb{R}^n}$ (nach Definition eines Cauchy-Problems) und $F_0 \circ E_0 = \text{id}_L$ (nach der Eindeutigkeit der Lösungen eines Cauchy-Problems, siehe Proposition 4.8). Dies beweist, das E_0 ein reeller Vektorraumisomorphismus ist. Damit ist ii) gezeigt. □

Definition 4.18 *Die* Resolvente *zu einer linearen Differentialgleichung erster Ordnung $y' = A(t)\cdot y + b(t)$ ist die Abbildung $R : I \times I \longrightarrow M_{n\times n}(\mathbb{R})$ mit: für alle $t_0 \in \mathbb{R}^n$ ist $t \longmapsto R(t, t_0)$ die Lösung der linearen Differentialgleichung*

$$\begin{cases} y' = A(t)\cdot y \\ y(t_0) = \mathrm{I}_n \end{cases}$$

in $M_{n\times n}(\mathbb{R})$, wobei I_n die Identitätsmatrix im $\mathbb{R}^n$ ist.

Bemerkungen 4.19

1. Nach Korollar 4.15 ist R wohldefiniert mit $R(t_0, t_0) = \mathrm{I}_n$. Desweiteren gilt, für alle t, s, u in I: $R(t, s)\cdot R(s, u) = R(t, u)$. Denn: legen wir s und u fest, so ist $t \overset{F}{\longmapsto} R(t, s)\cdot R(s, u)$ ableitbar mit

   $$F'(t) = \frac{d}{dt}(R(t, s))\cdot R(s, u) = A(t)\cdot R(t, s)\cdot R(s, u) = A(t)\cdot F(t) \quad \text{für alle } t \in I$$

   wobei $F(s) = R(s, s)\cdot R(s, u) = R(s, u)$. Da aber $t \longmapsto R(t, u)$ dieselbe Differentialgleichung löst mit $R(s, u) = F(s)$, gilt $R(t, u) = F(t)$ für alle $t \in I$ nach Korollar 4.15.

2. Wegen $R(t_0, t_0) = I_n$ ist insbesondere $R(t, t_0)$ eine invertierbare Matrix, für alle t, t_0 mit $R(t, t_0)^{-1} = R(t_0, t)$.

Proposition 4.20 (Variation der Konstanten) *Sei eine lineare Differentialgleichung erster Ordnung im* $\mathbb{R}^n$ *der Form* $y'(t) = A(t) \cdot y + b(t)$ *gegeben, mit stetigen Abbildungen* $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ *und* $b : I \longrightarrow \mathbb{R}^n$*, wobei* $I \subset \mathbb{R}$ *ein Intervall ist. Sei* $R : I \times I \longrightarrow M_{n\times n}(\mathbb{R})$ *die Resolvente dieser Differentialgleichung. Seien* $t_0 \in I$ *und* $y_0 \in \mathbb{R}^n$*. Dann ist die Lösung* $y : I \longrightarrow \mathbb{R}^n$ *des Cauchy-Problems* $y' = A(t) \cdot y + b(t)$*,* $y(t_0) = y_0$*, gegeben durch:*

$$y(t) = R(t, t_0) \cdot y_0 + \int_{t_0}^{t} R(t, s)b(s)ds.$$

Beweis: Es ist lediglich, zu überprüfen, dass die angegebene Funktion y das Cauchy-Problem $y' = A(t) \cdot y$, $y(t_0) = y_0$, löst; eine andere Methode, welche die Formel liefert, wird unter dem Beweis beschrieben (man führt die Differentialgleichung auf eine Integralgleichung zurück). Sei zuerst $g : I \times I \longrightarrow \mathbb{R}^n$, $(x, y) \mapsto \int_{t_0}^{x} R(y, s) \cdot b(s)ds$. Für jedes feste $y \in I$ ist $s \mapsto R(y, s) \cdot b(s)$ als Produkt von stetigen Funktionen stetig auf I, insbesondere ist $x \mapsto g(x, y)$ ableitbar mit Ableitung $\frac{\partial g}{\partial x}(x, y) = R(y, x) \cdot b(x)$, siehe Satz 3.54. Für die Ableitbarkeit nach y fixieren wir ein $x \in I$, ein $\hat{y} \in I$ und ein $\eta > 0$ mit $[\hat{y} - \eta, \hat{y} + \eta] \subset I$. Für jedes $s \in I$ ist $y \mapsto R(y, s) \cdot b(s)$ ableitbar auf I mit Ableitung $\frac{\partial R}{\partial y}(y, s)\cdot b(s) = A(y)\cdot R(y, s)\cdot b(s)$ nach Definition der Resolvente R. Desweiteren existiert, wegen der Stetigkeit von $(s, y) \mapsto A(y) \cdot R(y, s) \cdot b(s)$ auf $I \times I$ und der Kompaktheit von $[t_0, x] \times [\hat{y} - \eta, \hat{y} + \eta]$, eine Konstante $C \in [0, \infty[$ mit $|A(y) \cdot R(y, s) \cdot b(s)| \leq C$ für alle $(s, y) \in [t_0, x] \times [\hat{y} - \eta, \hat{y} + \eta]$. Da jede stetige (insbesondere jede konstante) Funktion auf einem kompakten Intervall integrierbar ist, folgt aus Satz 3.56, dass $y \mapsto g(x, y)$ ableitbar ist mit Ableitung

$$\frac{\partial g}{\partial y}(x, y) = \int_{t_0}^{x} \frac{\partial R}{\partial y}(y, s) \cdot b(s)ds = \int_{t_0}^{x} A(y) \cdot R(y, s) \cdot b(s)ds = A(y) \cdot \int_{t_0}^{x} R(y, s) \cdot b(s)ds,$$

wobei die Linearität des Integrals im letzten Schritt verwendet wurde. Da beide Funktionen $(x, y) \mapsto R(y, x) \cdot b(x)$ und $(x, y) \mapsto A(y) \cdot \int_{t_0}^{x} R(y, s) \cdot b(s)ds$ stetig auf $I \times I$ sind (nutze Satz 3.54 und Satz 3.55), ist nach Satz 2.73 die Funktion g C^1 auf $I \times I$. Mit der Kettenregel (Proposition 2.63) folgt, dass die Abbildung $I \xrightarrow{h} \mathbb{R}^n$, $t \mapsto g(t, t)$, ableitbar auf I ist mit

$$h'(t) = \frac{\partial g}{\partial x}(t, t) + \frac{\partial g}{\partial y}(t, t) = \underbrace{R(t, t)}_{\mathrm{I}_n} \cdot b(t) + A(t) \cdot \int_{t_0}^{t} R(t, s) \cdot b(s)ds,$$

für alle $t \in I$. Daraus folgt, dass die angegebene Funktion y ableitbar ist mit

$$\begin{aligned}
y'(t) &= \frac{d}{dt}\left(R(t, t_0) \cdot y_0\right) + h'(t) \\
&= A(t) \cdot R(t, t_0) \cdot y_0 + b(t) + A(t) \cdot \int_{t_0}^{t} R(t, s) \cdot b(s)ds \\
&= A(t) \cdot \left(R(t, t_0) \cdot y_0 + \int_{t_0}^{t} R(t, s) \cdot b(s)ds\right) + b(t) \\
&= A(t) \cdot y(t) + b(t),
\end{aligned}$$

für alle $t \in I$. Schließlich ist $y(t_0) = \underbrace{R(t_0, t_0)}_{\mathrm{I}_n} \cdot y_0 + \underbrace{\int_{t_0}^{t_0} R(t, s) \cdot b(s)ds}_{0} = y_0$. Damit ist y die Lösung des Cauchy-Problems $y' = A(t) \cdot y$, $y(t_0) = y_0$ und die Proposition 4.20 ist bewiesen. □

Die Bezeichnung "Variation der Konstanten" kommt daher, dass $t \mapsto R(t, t_0) \cdot y_0$ die homogene Gleichung $y' = A(t) \cdot y$ löst mit $y(t_0) = y_0$; der Ansatz, um die obige Lösung der inhomogenen Gleichung zu bekommen, besteht darin, die *Konstante* y_0 *durch eine unbekannte Funktion* $\hat{y}$ zu ersetzen, d.h., die Lösung y in der Form $y(t) = R(t, t_0) \cdot \hat{y}(t)$ zu suchen. Setzt man diese Funktion y in die Gleichung ein, so bekommt man

$$\begin{aligned}
y' = A(t) \cdot y + b(t) \quad &\Longleftrightarrow \quad \frac{d}{dt}\Big(R(t, t_0) \cdot \hat{y}(t)\Big) = A(t) \cdot R(t, t_0) \cdot \hat{y}(t) + b(t) \\
&\Longleftrightarrow \quad \Big(\frac{d}{dt}R(t, t_0)\Big) \cdot \hat{y}(t) + R(t, t_0) \cdot \hat{y}'(t) = A(t) \cdot R(t, t_0) \cdot \hat{y}(t) + b(t) \\
&\Longleftrightarrow \quad A(t) \cdot R(t, t_0) \cdot \hat{y}(t) + R(t, t_0) \cdot \hat{y}'(t) = A(t) \cdot R(t, t_0) \cdot \hat{y}(t) + b(t) \\
&\Longleftrightarrow \quad R(t, t_0) \cdot \hat{y}'(t) = b(t) \\
&\Longleftrightarrow \quad \hat{y}'(t) = R(t, t_0)^{-1} \cdot b(t) \qquad (R(t, t_0) \text{ ist invertierbar}) \\
&\Longleftrightarrow \quad \hat{y}'(t) = R(t_0, t) \cdot b(t) \qquad (R(t, t_0)^{-1} = R(t_0, t)) \\
&\Longleftrightarrow \quad \hat{y}(t) - \hat{y}(t_0) = \int_{t_0}^{t} R(t_0, s) b(s) ds \qquad (\text{nach Integration}) \\
&\Longleftrightarrow \quad y(t) - R(t, t_0)\hat{y}(t_0) = R(t, t_0) \cdot \int_{t_0}^{t} R(t_0, s) b(s) ds \\
&\Longleftrightarrow \quad y(t) - R(t, t_0)\hat{y}(t_0) = \int_{t_0}^{t} R(t, t_0) \cdot R(t_0, s) b(s) ds \\
&\Longleftrightarrow \quad y(t) - R(t, t_0)\hat{y}(t_0) = \int_{t_0}^{t} R(t, s) b(s) ds.
\end{aligned}$$

In den letzten beiden Schritten haben wir die Linearität von $R(t, t_0)$ (als Transformation des $\mathbb{R}^n$) sowie die Identität $R(t, t_0) \cdot R(t_0, s) = R(t, s)$ benutzt. Es bleibt, zu bemerken, dass die Bedingung $y(t_0) = y_0$ die Bedingung $\hat{y}(t_0) = y_0$ verlangt (wegen $R(t_0, t_0) = \mathrm{I}_n$). Damit bekommen wir die obige Gestalt der Lösung y.

4.2.2 Lineare Differentialgleichungen mit konstanten Koeffizienten

Definition 4.21 *Sei* $n \in \mathbb{N} \setminus \{0\}$.

i) Eine lineare Differentialgleichung erster Ordnung mit konstanten Koeffizienten *im* $\mathbb{R}^n$ *ist eine Differentialgleichung der Form*

$$y' = A \cdot y + b(t),$$

wobei $A \in M_{n\times n}(\mathbb{R})$ *eine (konstante) Matrix und* $b : I \longrightarrow \mathbb{R}^n$ *eine auf einem Intervall* $I \subset \mathbb{R}$ *definierte Abbildung ist.*

ii) Allgemeiner heißt, für $k \in \mathbb{N} \setminus \{0\}$, *eine Differentialgleichung* k*-ter Ordnung* $y^{(k)} = f(t, y, y', \ldots, y^{(k-1)})$ *genau dann linear mit konstanten Koeffizienten, wenn die zugehörige Differentialgleichung erster Ordnung für* $(y, y', \ldots, y^{(k-1)})$ *linear mit konstanten Koeffizienten ist.*

Zum Lösen linearer Differentialgleichungen mit konstanten Koeffizienten spielt der folgende Begriff eine zentrale Rolle:

Definition 4.22 *Das* Exponential *einer Matrix $A \in M_{n\times n}(\mathbb{R})$ (oder $A \in M_{n\times n}(\mathbb{C})$) ist die Matrix*

$$\exp(A) = e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} \in M_{n\times n}(\mathbb{R}),$$

wobei $A^0 := \mathrm{I}_n$.

Bemerkung 4.23 Die Matrix e^A ist wohldefiniert, weil die Folge $\left(u_m := \sum\limits_{k=0}^{m} \frac{A^k}{k!}\right)_{m\in\mathbb{N}}$ eine Cauchy-Folge im vollständigen metrischen Raum $(M_{n\times n}(\mathbb{R}), \|\cdot\|)$ (oder $M_{n\times n}(\mathbb{C})$) ist, wobei $\|\cdot\|$ die Matrixnorm bezeichnet.

Lemma 4.24 *Seien $A, B \in M_{n\times n}(\mathbb{R})$ (oder $M_{n\times n}(\mathbb{C})$). Dann gilt*

i) Für die Nullmatrix 0_n gilt $e^{0_n} = \mathrm{I}_n$ und $e^{A+B} = e^A \cdot e^B$, sobald $AB = BA$.

ii) Die Abbildung $t \longmapsto e^{tA}$ ist ableitbar und es gilt $\frac{d}{dt}(e^{tA}) = A \cdot e^{tA}$.

iii) Für jede invertierbare Matrix $P \in M_{n\times n}(\mathbb{R})$ gilt $e^{P^{-1}AP} = P^{-1} \cdot e^A \cdot P$.

Beweis:
i) Wegen $0_n^m = 0_n$ für alle $m \geq 1$ ist $e^{0_n} = \frac{\mathrm{I}_n}{0!} + 0_n = \mathrm{I}_n$. Für beliebige $A, B \in M_{n\times n}(\mathbb{R})$ mit $AB = BA$ gilt, nach der binomischen Formel – welche hier wegen $AB = BA$ anwendbar ist –

$$\frac{(A+B)^m}{m!} = \frac{1}{m!}\sum_{k=0}^{m} \binom{m}{k} A^k \cdot B^{m-k} = \sum_{k=0}^{m} \frac{A^k}{k!} \cdot \frac{B^{m-k}}{(m-k)!}$$

für alle $m \in \mathbb{N}$. Zu bemerken ist nun, dass für ein beliebiges $X \in M_{n\times n}(\mathbb{R})$ die Reihe $\sum_m \frac{X^m}{m!}$ nicht nur konvergent sondern auch *absolut* konvergent ist: es gilt $\sum_{m=0}^{p} \frac{\|X^m\|}{m!} \leq \sum_{m=0}^{p} \frac{\|X\|^m}{m!} \leq \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} < \infty$ für alle $p \in \mathbb{N}$. Dies impliziert, dass das sogenannte Cauchy-Produkt der Reihen $\sum_m \frac{A^m}{m!}$ und $\sum_m \frac{B^m}{m!}$, welches formal durch $\sum_m (\sum_{k=0}^{m} \frac{A^k}{k!} \cdot \frac{B^{m-k}}{(m-k)!})$ definiert ist, wohldefiniert und (absolut) konvergent ist mit $\sum_{m=0}^{\infty}(\sum_{k=0}^{m} \frac{A^k}{k!} \cdot \frac{B^{m-k}}{(m-k)!}) = \left(\sum_{m=0}^{\infty} \frac{A^m}{m!}\right) \cdot \left(\sum_{m=0}^{\infty} \frac{B^m}{m!}\right)$. Dies liefert die Identität $e^{A+B} = e^A \cdot e^B$.
ii) Für jedes $m \geq 1$ ist $t \mapsto \frac{t^m A^m}{m!}$ offensichtlich ableitbar auf $\mathbb{R}$ mit Ableitung $mt^{m-1} \cdot \frac{A^m}{m!} = t^{m-1}\frac{A^m}{(m-1)!}$. Ist $T \in]0, \infty[$ beliebig, so gilt

$$\|t^{m-1}\frac{A^m}{(m-1)!}\| \leq |t|^{m-1} \cdot \|A\| \cdot \frac{\|A\|^{m-1}}{(m-1)!} \leq T^{m-1} \cdot \|A\| \cdot \frac{\|A\|^{m-1}}{(m-1)!}$$

für alle $t \in]-T, T[$ und $m \geq 1$. Wegen

$$\sum_{m=1}^{\infty} T^{m-1} \cdot \|A\| \cdot \frac{\|A\|^{m-1}}{(m-1)!} = T^{m-1} \cdot \|A\| \cdot \sum_{m=0}^{\infty} \frac{\|A\|^m}{(m)!} = T^{m-1} \cdot \|A\| \cdot e^{\|A\|} < \infty$$

liefert der Satz über Differentiation unter dem Integral (Satz 3.56), dass die Abbildung $t \mapsto e^{tA} = \sum_{m=0}^{\infty} \frac{t^m A^m}{m!}$ auf $]-T, T[$ ableitbar ist mit Ableitung

$$\frac{d}{dt}\left(e^{tA}\right) = \sum_{m=0}^{\infty} \frac{d}{dt}\left(\frac{t^m A^m}{m!}\right) = \sum_{m=1}^{\infty} t^{m-1}\frac{A^m}{(m-1)!} = A \cdot \sum_{m=1}^{\infty} t^{m-1}\frac{A^{m-1}}{(m-1)!} = A \cdot e^{tA}$$

für alle $t \in]-T,T[$ (*Übungsaufgabe: warum kann* Satz 3.56 *angewendet werden? Hinweis: betrachten Sie das Zählmaß auf* $\mathbb{N}$). Im letzten Schritt wurde auch verwendet, dass die Matrixmultiplikation *stetig* ist. Dies gilt für alle $T \in]0,\infty[$, somit ist die Behauptung bewiesen.
iii) Die Formel folgt unmittelbar aus $(P^{-1}AP)^m = P^{-1}A^mP$ für alle $m \in \mathbb{N}$ (verwende auch die Stetigkeit der Matrixmultiplikation). □

Im Fall einer linearen Differentialgleichung mit konstanten Koeffizienten lässt sich die Resolvente durch die obige Exponentialfunktion elementar ausdrücken:

Satz 4.25 *Sei eine lineare Differentialgleichung mit konstanten Koeffizienten erster Ordnung im $\mathbb{R}^n$ der Form $y' = A \cdot y(t) + b(t)$ gegeben, wobei $A \in M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ eine auf einem offenen Intervall I definierte stetige Abbildung ist.*
Dann ist die Resolvente dieser Differentialgleichung gegeben durch

$$R : I \times I \longrightarrow M_{n\times n}(\mathbb{R}), \qquad (t,t_0) \longmapsto e^{(t-t_0)A}.$$

Insbesondere ist die Lösung y des Cauchy-Problems $y' = A \cdot y + b(t)$, $y(t_0) = y_0$, gegeben durch

$$y(t) = e^{(t-t_0)A} \cdot y_0 + \int_{t_0}^{t} e^{(t-s)A} b(s)ds$$

für alle $t \in I$.

Beweis: Nach Lemma 4.24.*i*) ist, für alle $t \in \mathbb{R}$, $e^{(t-t_0)A} = e^{tA} \cdot e^{-t_0A}$ wegen $t_0A \cdot tA = tA \cdot t_0A$. Lemma 4.24.*ii*) liefert, dass $t \mapsto e^{tA} \cdot e^{-t_0A}$ ableitbar auf $\mathbb{R}$ ist mit Ableitung $A \cdot e^{tA} \cdot e^{-t_0A} = A \cdot e^{(t-t_0)A}$. Insbesondere ist $I \xrightarrow{H} M_{n\times n}(\mathbb{R})$, $t \mapsto e^{(t-t_0)A}$, ableitbar mit $H'(t) = A \cdot e^{(t-t_0)A} = A \cdot H(t)$ für alle $t \in I$ sowie $H(t_0) = e^{0_n} = \mathrm{I}_n$ (wieder nach Lemma 4.24.*i*)). Nach der Eindeutigkeit von Lösungen eines Cauchy-Problems (Proposition 4.8) gilt dann $H(t) = R(t,t_0)$ für alle $t \in I$. Die zweite Identität folgt aus dieser Tatsache zusammen mit Proposition 4.20. □

Bemerke, dass die Lösung $t \mapsto e^{(t-t_0)A} \cdot y_0$ des homogenen Cauchy-Problems $y' = A \cdot y$, $y(t_0) = y_0$, auf ganz $\mathbb{R}$ definiert ist. Nun wollen wir diese Lösung möglichst explizit hinschreiben. Dazu müssen wir das Exponential einer Matrix ebenfalls möglichst explizit berechnen. Dafür sind die Diagonalisierungs- und Trigonalisierungsverfahren aus der linearen Algebra sehr nützlich.

Erinnerung:
• Eine Matrix $A \in M_{n\times n}(\mathbb{R})$ heißt genau dann *diagonalisierbar*, wenn eine invertierbare Matrix $P \in M_{n\times n}(\mathbb{R})$ so existiert, dass $P^{-1}AP$ eine Diagonalmatrix ist, d.h., $P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Die reellen Zahlen $\lambda_1, \ldots, \lambda_n$ heißen dann die *Eigenwerte* der Matrix A. Analoges gilt für ein $A \in M_{n\times n}(\mathbb{C})$.
• Eine Matrix $A \in M_{n\times n}(\mathbb{R})$ heißt genau dann *trigonalisierbar*, wenn eine invertierbare Matrix $P \in M_{n\times n}(\mathbb{R})$ so existiert, dass $P^{-1}AP$ eine obere Dreiecksmatrix ist, d.h., $P^{-1}AP = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$ (alle Koeffizienten unterhalb der Diagonale verschwinden).

Analoges gilt für ein $A \in M_{n\times n}(\mathbb{C})$.
• Nicht jede reelle oder komplexe Matrix ist diagonalisierbar; eine notwendige Bedingung dafür ist, dass das sogenannte *charakteristische Polynom*

$$\chi_A(\lambda) := \det(A - \lambda \cdot \mathrm{I}_n)$$

von A in *Linearfaktoren* zerfällt; diese Bedingung reicht aber nicht. Ist z.B. A symmetrisch (d.h., ${}^tA = A$) oder hat A nur einfache Eigenwerte, so ist A diagonalisierbar.
• Eine Matrix A ist genau dann trigonalisierbar, wenn ihr charakteristisches Polynom in Linearfaktoren zerfällt. Insbesondere ist nach dem Satz von Gauß (*jedes komplexe Polynom vom Grad mindestens* 1 *zerfällt in Linearfaktoren*) jede komplexe Matrix trigonalisierbar.
• Für eine trigonalisierbare (insbesondere für eine komplexe) Matrix A kann eine obere Dreiecksform besonderer Art für A hingeschrieben werden. Zerlege nämlich $\chi_A(\lambda)$ in Linearfaktoren: es gibt $k \in \mathbb{N} \setminus \{0\}$, $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ (oder $\mathbb{C}$) paarweise verschieden und $m_1, \ldots, m_k \in \mathbb{N} \setminus \{0\}$ mit $\chi_A(\lambda) = (-1)^n \prod_{j=1}^{k} (\lambda - \lambda_j)^{m_j}$ und $\sum_{j=1}^{k} m_j = n$ (wegen $d^o\chi_A(\lambda) = n$). Setze $E_j := \ker((A - \lambda_j \cdot \mathrm{I}_n)^{m_j}) \subset \mathbb{R}^n$, dann gilt: $\mathbb{R}^n = \bigoplus_{j=1}^{k} E_j$, $\dim(E_j) = m_j$ (algebraische Vielfachheit der Wurzel λ_j im Polynom $\chi_A(\lambda)$) und $A \cdot E_j \subset E_j$. Außerdem existiert, für jedes $1 \le j \le k$, eine Basis von E_j, in der $A_{|E_j}$ der Form $\begin{pmatrix} \lambda_j & * & 0 \\ & \ddots & * \\ 0 & & \lambda_j \end{pmatrix}$, wobei jeder Koeffizient über der Diagonale entweder 1 oder 0 ist. Insgesamt existiert eine invertierbare Matrix $P \in M_{n\times n}(\mathbb{R})$ mit

$$P^{-1}AP = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & * & 0 \\ & \ddots & * \\ 0 & & \lambda_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \lambda_2 & * & 0 \\ & \ddots & * \\ 0 & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_k & * & 0 \\ & \ddots & * \\ 0 & & \lambda_k \end{matrix}} \end{pmatrix}.$$

Dies ist die sogenannte *Jordan'sche Normalform* von A und $\lambda_1, \ldots, \lambda_k$ sind die Eigenwerte von A. Für eine diagonale Matrix $A = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$ gilt $e^A = \begin{pmatrix} e^{\mu_1} & & 0 \\ & \ddots & \\ 0 & & e^{\mu_n} \end{pmatrix}$, insbesondere $e^{tA} = \begin{pmatrix} e^{t\mu_1} & & 0 \\ & \ddots & \\ 0 & & e^{t\mu_n} \end{pmatrix}$. Für eine obere Dreiecksmatrix der Form $A =$

$\begin{pmatrix} \mu & & * \\ & \ddots & \\ 0 & & \mu \end{pmatrix}$ schreiben wir $A = \begin{pmatrix} \mu & & 0 \\ & \ddots & \\ 0 & & \mu \end{pmatrix} + \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} = \mu \cdot \mathrm{I}_n + N$. Wegen $N^n = 0$ und $\mu \mathrm{I}_n \cdot N = N \cdot \mu \mathrm{I}_n$ gilt $e^A = e^{\mu \cdot \mathrm{I}_n} \cdot e^N = e^\mu \cdot \mathrm{I}_n \cdot \sum\limits_{k=0}^{n-1} \frac{N^k}{k!} = e^\mu \sum\limits_{k=0}^{n-1} \frac{N^k}{k!}$. Insbesondere ist $e^{tA} = e^{t\mu} \cdot \sum\limits_{k=0}^{n-1} \frac{t^k N^k}{k!}$. Bemerke hierbei, dass $\sum_{k=0}^{n-1} \frac{t^k N^k}{k!}$ ein Polynom in der Variablen t (mit Koeffizienten in $M_{n\times n}(\mathbb{R})$) vom Grad höchstens $n-1$ ist.
Für die obige Zerlegung von $P^{-1}AP$ gilt daher

$$P^{-1} e^{tA} P = e^{tP^{-1}AP} = \begin{pmatrix} e^{t\lambda_1}\hat{A}_1(t) & & & 0 \\ & e^{t\lambda_2}\hat{A}_2(t) & & \\ & & \ddots & \\ 0 & & & e^{t\lambda_k}\hat{A}_k(t) \end{pmatrix}$$

wobei $\hat{A}_j(t) \in M_{m_j \times m_j}(\mathbb{R})$ und der Form $\hat{A}_j(t) = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$, wobei alle Koeffizienten von $\hat{A}_j(t)$ oberhalb der Diagonale Polynome vom Grad höchstens $m_j - 1$ in t sind.

Korollar 4.26 *Sei eine lineare Differentialgleichung erster Ordnung mit konstanten Koeffizienten der Form $y' = A \cdot y + b(t)$, wobei $A \in M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ eine auf einem offenen Intervall I definierte stetige Abbildung ist. Angenommen, A sei* trigonalisierbar. *Dann ist jede Lösung der homogenen Gleichung $y' = A \cdot y$ der Form*

$$y(t) = \sum_{j=1}^{k} e^{\lambda_j t} p_j(t),$$

wobei $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte von A sind und $p_j(t) = \sum\limits_{l=0}^{m_j-1} t^l c_{jl}$ Polynome vom Grad $m_j - 1$ (wobei m_j die algebraische Vielfachheit des Eigenwertes λ_j ist) mit Koeffizienten $c_{jl} \in \ker((A - \lambda_j \cdot \mathrm{I}_n)^{m_j})$.

Beweis: Nach Satz 4.25 ist jede Lösung y der homogenen Differentialgleichung $y' = A \cdot y$ der Form $y(t) = e^{tA} \cdot y_0$, wobei $y_0 \in \mathbb{R}^n$ ein konstanter Vektor ist. Wegen $\mathbb{R}^n = \bigoplus\limits_{1 \le j \le k} E_j$ (mit $E_j := \ker((A - \lambda_j \cdot \mathrm{I}_n)^{m_j})$) lässt sich aber jeder Vektor $v \in \mathbb{R}^n$ eindeutig in $v = \sum_{j=1}^k V_j$ zerlegen, wobei $V_j \in E_j$ für alle $1 \le j \le k$ gilt. Ist nun $y_0 \in \mathbb{R}^n$ beliebig, so zerlege entsprechend $y_0 = \sum_{j=1}^k Y_j$ mit $Y_j \in E_j$ für alle $1 \leq j \leq k$. Dann gilt

$$\begin{aligned} e^{tA} \cdot y_0 &= \sum_{j=1}^k e^{tA} \cdot Y_j \\ &= \sum_{j=1}^k e^{t\lambda_j \mathrm{I}_n} \cdot e^{t(A-\lambda_j \mathrm{I}_n)} \cdot Y_j \qquad \text{nach Lemma 4.24.}i) \\ &= \sum_{j=1}^k e^{t\lambda_j} \cdot e^{t(A-\lambda_j \mathrm{I}_n)} \cdot Y_j \\ &= \sum_{j=1}^k e^{t\lambda_j} \cdot \left(\sum_{l=0}^{m_j-1} \frac{t^l (A-\lambda_j \mathrm{I}_n)^l}{l!} \right) \cdot Y_j \qquad \text{wegen } (A-\lambda_j \mathrm{I}_n)^{m_j} \cdot Y_j = 0 \\ &= \sum_{j=1}^k e^{t\lambda_j} \cdot p_j(t), \end{aligned}$$

wobei $p_j(t) := \sum_{l=0}^{m_j-1} t^l \frac{(A-\lambda_j \mathrm{I}_n)^l}{l!} \cdot Y_j$ für alle $1 \leq j \leq k$ und $t \in \mathbb{R}$. Es bleibt lediglich, zu bemerken, dass $c_{jl} := \frac{1}{l!}(A-\lambda_j \mathrm{I}_n)^l \cdot Y_j \in E_j$ wegen $A \cdot E_j \subset E_j$ gilt, für alle $0 \leq j \leq m_j - 1$ und $1 \leq j \leq k$. $\square$

Die Methode in diesem Fall lautet daher: berechne $\chi_A(\lambda)$, zerlege es in Linearfaktoren $\chi_A(\lambda) = (-1)^n \prod_{j=1}^k (\lambda - \lambda_j)^{m_j}$ (bekomme insbesondere die Eigenwerte λ_j und ihre Vielfachheiten m_j), bestimme die Untervektorräume $E_j := \ker((A - \lambda_j \cdot \mathrm{I}_n)^{m_j})$ (löse ein lineares Gleichungssystem) und schreibe dann eine beliebige Lösung in der Form $\sum_{j=1}^k e^{\lambda_j t} p_j(t)$, wobei jedes p_j ein Polynom vom Grad $m_j - 1$ ist mit beliebigen Koeffizienten in E_j; füge diese Funktion in die homogene Gleichung $y' = A \cdot y$ ein und schreibe die entsprechenden linearen Gleichungen hin, welche die Koeffizienten erfüllen müssen. Für jedes $j \in \{1, \dots, k\}$ müssen für den E_j-Anteil am Ende genau m_j Koeffizienten unbestimmt bleiben.

Bemerkung 4.27 Ist $A \in M_{n \times n}(\mathbb{R})$ *beliebig*, so betrachte A als *komplexe* Matrix. Dann ist A trigonalisierbar wie oben, wobei $\lambda_1, \dots, \lambda_k$ *komplexe* Zahlen und $E_j := \ker((A - \lambda_j \cdot \mathrm{I}_n)^{m_j}) \subset \mathbb{C}^n$ *komplexe* Unterräume sind. Ggf. ist jede Funktion der Form $t \longmapsto e^{\lambda_j t} p_j(t)$, $I \longrightarrow \mathbb{C}^n$, eine Lösung der homogenen Gleichung $y' = A \cdot y$, ist aber *komplexwertig*. Die Lösungen der "reellen" Gleichung bekommt man durch Bildung des Real- und Imaginärteils dieser Lösungen. Genauer: Schreibt man – im Fall λ_j *nicht* reell – $\lambda_j = \mathrm{Re}(\lambda_j) + i\mathrm{Im}(\lambda_j)$, so ist

$$e^{\lambda_j t} = e^{\mathrm{Re}(\lambda_j)t}(\cos(\mathrm{Im}(\lambda_j)t) + \sin(\mathrm{Im}(\lambda_j)t)).$$

Eine Lösung der (reellen) Differentialgleichung bekommt man dann durch den Ansatz $y(t) = e^{\mathrm{Re}(\lambda_j)t}(\cos(\mathrm{Im}(\lambda_j)t)p_j(t) + \sin(\mathrm{Im}(\lambda_j)t)q_j(t)$, wobei p_j, q_j $\mathbb{R}^n$-wertige Polynome vom Grad höchstens $m_j - 1$ sind.

Beispiele 4.28

1. Betrachte die homogene Differentialgleichung
$$y' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot y$$
im $\mathbb{R}^2$. Wir wollen die obige Methode auf $A := \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ anwenden. Es gilt $\chi_A(\lambda) = \det(A - \lambda \cdot I_2) = \left(\begin{smallmatrix} -\lambda & 1 \\ 1 & -\lambda \end{smallmatrix}\right) = \lambda^2 - 1$, d.h., $\chi_A(\lambda) = (\lambda - 1)(\lambda + 1)$. Da A zwei verschiedene reelle Eigenwerte hat, ist A diagonalisierbar (A ist sowieso symmetrisch). Eine kurze Rechnung liefert $A \cdot (e_1 + e_2) = e_1 + e_2$ und $A(e_1 - e_2) = -(e_1 - e_2)$. Somit sind $e_1 + e_2$ und $e_1 - e_2$ Eigenvektoren für A (zu den Eigenwerten 1 bzw. -1). Die Basiswechselmatrix zwischen $\{e_1, e_2\}$ und $\{e_1 + e_2, e_1 - e_2\}$ ist $P = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, mit Inversem $P^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Wegen $P^{-1}AP = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ gilt $P^{-1}e^{tA}P = e^{tP^{-1}AP} = \left(\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t} \end{smallmatrix}\right)$, somit
$$\begin{aligned} e^{tA} &= P \left(\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t} \end{smallmatrix}\right) P^{-1} \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t} \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} e^t & e^t \\ e^{-t} & -e^{-t} \end{smallmatrix}\right) \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} e^t + e^{-t} & e^t - e^{-t} \\ e^t - e^{-t} & e^t + e^{-t} \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{smallmatrix}\right). \end{aligned}$$
Die Lösung des Cauchy-Problems $y' = A \cdot y$ mit $y(t_0) = y_0$ (mit $t_0 \in \mathbb{R}$ und $y_0 = (y_{01}, y_{02}) \in \mathbb{R}^2$) ist daher gegeben durch
$$\begin{aligned} y(t) &= e^{(t-t_0)A} \cdot y_0 = \left(\begin{smallmatrix} \cosh(t-t_0) & \sinh(t-t_0) \\ \sinh(t-t_0) & \cosh(t-t_0) \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} y_{01} \\ y_{02} \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} \cosh(t-t_0)y_{01} + \sinh(t-t_0)y_{02} \\ \sinh(t-t_0)y_{01} + \cosh(t-t_0)y_{02} \end{smallmatrix}\right). \end{aligned}$$
Daraus folgt insbesondere, dass die Lösung z der homogenen Differentialgleichung zweiter Ordnung $z'' = z$ in $\mathbb{R}$ mit Cauchy-Daten $z_0, z_0' \in \mathbb{R}$ an der Stelle $t_0 \in \mathbb{R}$ gegeben ist durch
$$z(t) = \cosh(t - t_0)z_0 + \sinh(t - t_0)z_0',$$
siehe Beispiel 4.6.1.

2. Betrachte die homogene Differentialgleichung erster Ordnung $y' = A \cdot y$ wobei $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in M_{2\times 2}(\mathbb{R})$ ist. Es gilt jetzt $\chi_A(\lambda) = \det(A - \lambda I_2) = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$. Wiederum hat A zwei Eigenwerte, die aber jetzt *rein imaginär* sind. Eine Basis von $\mathbb{C}^2$, welche durch Eigenvektoren von A gegeben ist, ist $\{e_1 + ie_2, e_1 - ie_2\}$ (wegen $A(e_1 + ie_2) = -e_2 + ie_1 = i(e_1 + ie_2)$ und $A(e_1 - ie_2) = -e_2 - ie_1 = -i(e_1 - ie_2)$). Die Basiswechselmatrix P zwischen der kanonischen Basis $\{e_1, e_2\}$ und der Basis $\{e_1 + ie_2, e_1 - ie_2\}$ von $\mathbb{C}^2$ ist gegeben durch $P = \left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right)$, mit Inversem $P^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & -i \\ 1 & i \end{smallmatrix}\right)$. Wegen $P^{-1}AP = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ gilt, wie vorher,
$$\begin{aligned} e^{tA} &= P \left(\begin{smallmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{smallmatrix}\right) P^{-1} \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 1 & -i \\ 1 & i \end{smallmatrix}\right) \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} e^{it} & -ie^{it} \\ e^{-it} & ie^{-it} \end{smallmatrix}\right) \\ &= \tfrac{1}{2}\left(\begin{smallmatrix} e^{it} + e^{-it} & -ie^{it} + ie^{-it} \\ ie^{it} - ie^{-it} & e^{it} + e^{-it} \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{smallmatrix}\right). \end{aligned}$$

Die Lösung des Cauchy-Problems $y' = A \cdot y$, $y(0) = y_0$ (wobei $y_0 = \begin{pmatrix} y_{01} \\ y_{02} \end{pmatrix} \in \mathbb{R}^2$) ist daher gegeben durch

$$y(t) = e^{tA} \cdot y_0 = \begin{pmatrix} \cos(t)y_{01} + \sin(t)y_{02} \\ -\sin(t)y_{01} + \cos(t)y_{02} \end{pmatrix}.$$

3. Betrachte die homogene Differentialgleichung erster Ordnung $y' = A \cdot y$, wobei $A = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ ist. Es gilt $\chi_A(\lambda) = (\lambda+1)^2$, insbesondere besitzt A lediglich -1 als Eigenwert. Diesmal ist aber $A - (-I_2) = A + I_2 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \neq 0$. Genauer gilt, für ein $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$: $(A + I_2) \cdot x = 0 \iff x_1 + x_2 = 0$, d.h., $\ker(A + I_2) = \mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Wählen wir $\{e_1 - e_2, e_1\}$ als Basis der $\mathbb{R}^2$, so ist die Basiswechselmatrix P zwischen der kanonischen Basis $\{e_1, e_2\}$ und der Basis $\{e_1 - e_2, e_1\}$ gegeben durch $P = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, mit Inversem $P^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$. Außerdem ist $P^{-1}AP = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ (nachrechnen). Daraus folgt

$$\begin{aligned} e^{tA} &= P \cdot \exp\begin{pmatrix} -t & t \\ 0 & -t \end{pmatrix} \cdot P^{-1} = P \exp\begin{pmatrix} -t & 0 \\ 0 & -t \end{pmatrix} \cdot \exp\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} \cdot P^{-1} \\ &= P \begin{pmatrix} e^{-t} & \\ 0 & e^{-t} \end{pmatrix} \cdot \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \cdot P^{-1} \\ &= P \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} P^{-1} \\ &= e^{-t} \begin{pmatrix} t+1 & t \\ -t & 1-t \end{pmatrix} \qquad \text{(nachrechnen).} \end{aligned}$$

Die Lösung y des Cauchy-Problems $y' = A \cdot y$, $y(0) = y_0$ (mit $y_0 = \begin{pmatrix} y_{01} \\ y_{02} \end{pmatrix} \in \mathbb{R}^2$) ist daher gegeben durch

$$y(t) = e^{tA} \cdot y_0 = e^{-t} \begin{pmatrix} (t+1)y_{01} + ty_{02} \\ -ty_{01} + (1-t)y_{02} \end{pmatrix}.$$

4. Betrachte die homogene Differentialgleichung zweiter Ordnung $y'' - 2y' + 2y = 0$ in $\mathbb{R}$. Diese Gleichung ist äquivalent zur folgenden (homogenen) Differentialgleichung erster Ordnung im $\mathbb{R}^2$:

$$\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} y' \\ 2y' - 2y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} y \\ y' \end{pmatrix}.$$

Setze $A := \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$. Dann ist $\chi_A(\lambda) = \lambda^2 - 2\lambda + 2$, d.h., $\chi_A(\lambda) = (\lambda-(1-i))(\lambda-(1+i))$. Wiederum hat A zwei konjugierte komplexe Eigenwerte $1+i$ und $1-i$. Man rechnet nach, dass eine Basis von $\mathbb{C}^2$ durch Eigenvektoren von A durch $\{e_1+(1+i)e_2, e_1+(1-i)e_2\}$ gegeben ist, wobei $A(e_1+(1+i)e_2) = (1+i)(e_1+(1+i)e_2)$ und $A(e_1 + (1-i)e_2) = (1-i)(e_1 + (1-i)e_2)$. Die Basiswechselmatrix P zwischen $\{e_1, e_2\}$ und $\{e_1 + (1+i)e_2, e_1 + (1-i)e_2\}$ ist $P = \begin{pmatrix} 1 & 1 \\ 1+i & 1-i \end{pmatrix}$, mit Inversem $P^{-1} = \frac{1}{2}\begin{pmatrix} 1+i & -i \\ 1-i & i \end{pmatrix}$ (nachrechnen). Wie vorher ist

$$\begin{aligned} e^{tA} &= P \begin{pmatrix} e^{(1+i)t} & 0 \\ 0 & e^{(1-i)t} \end{pmatrix} P^{-1} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1+i & 1-i \end{pmatrix} \cdot \begin{pmatrix} e^{(1+i)t} & 0 \\ 0 & e^{(1-i)t} \end{pmatrix} \cdot \begin{pmatrix} (1+i) & -i \\ 1-i & i \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2\mathrm{Re}((1+i)e^{(1+i)t}) & 2\mathrm{Re}(ie^{(1-i)t}) \\ 2\mathrm{Re}(2ie^{(1+i)t}) & 2\mathrm{Re}(i(1-i)e^{(1-i)t}) \end{pmatrix} \\ &= \begin{pmatrix} e^t(\cos(t)-\sin(t)) & e^t\sin(t) \\ -2e^t\sin(t) & e^t(\cos(t)+\sin(t)) \end{pmatrix}. \end{aligned}$$

Insbesondere ist die Lösung y des Cauchy-Problems $y'' - 2y' + 2y = 0$, $y(0) = y_0$ und $y'(0) = y_0'$ (wobei $y_0, y_0' \in \mathbb{R}$), gegeben durch

$$y(t) = e^t(\cos(t) - \sin(t))y_0 + e^t \sin(t) y_0'.$$

Ein möglicher Ansatz, um die Lösung zu finden, ist, wie vorhin beschrieben: suche y in der Form $y(t) = a\cos(t)e^t + b\sin(t)e^t$, wobei $a, b \in \mathbb{R}$ die Unbekannten sind.

Bemerkungen 4.29

1. Die Formel $R(t,t_0) = e^{(t-t_0)A}$ für die Resolvente R der linearen Differentialgleichung erster Ordnung mit konstanten Koeffizienten (im $\mathbb{R}^n$) $y' = A \cdot y + b(t)$ kann für eine Differentialgleichung der Form $y' = A(t) \cdot y + b(t)$ (wobei $A : I \longrightarrow M_{n\times n}(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ auf einem offenen Intervall I definierte stetige Abbildungen sind) unter einer bestimmten Bedingung an A verallgemeinert werden. Nämlich gilt für die Resolvente $R : I \times I \longrightarrow M_{n\times n}(\mathbb{R})$ der Differentialgleichung $y' = A(t) \cdot y$:

$$R(t,t_0) = \exp(\int_{t_0}^t A(s)ds) \quad \text{für alle } (t,t_0) \in I \times I,$$

sobald $\int_{t_0}^t A(s)ds \cdot A(t) = A(t) \cdot \int_{t_0}^t A(s)ds$ für alle $t \in I$ (und für ein festes $t_0 \in I$) gilt. Denn damit gilt $\frac{d}{dt}\Big(\exp(\int_{t_0}^t A(s)ds)\Big) = \frac{d}{dt}(\int_{t_0}^t A(s)ds) \cdot \exp(\int_{t_0}^t A(s)ds) = A(t) \cdot \exp(\int_{t_0}^t A(s)ds) = A(t) \cdot R(t,t_0)$ und $\exp(\int_{t_0}^{t_0} A(s)ds) = \exp(0_n) = \mathrm{I}_n$.
Ein Anwendungsbeispiel dieser Bemerkung ist die explizite Formel für eine nach Bogenlänge parametrisierte Kurve $c : I \longrightarrow \mathbb{R}^2$ mit vorgegebener Krümmung $\kappa : I \longrightarrow \mathbb{R}$. Nämlich erfüllt das Paar $(\dot{c}, n)$ (wobei $n : I \longrightarrow \mathbb{R}^2$ das Einheitsnormalenfeld zu c ist) die Frenet-Gleichungen $\left\{\begin{array}{l} \ddot{c} = \kappa n \\ \dot{n} = -\kappa \dot{c} \end{array}\right.$, d.h., $\binom{\dot{c}}{n}' = \begin{pmatrix} 0 & \kappa \\ -\kappa & 0 \end{pmatrix} \cdot \binom{\dot{c}}{n}$ (siehe Satz 1.27).
Die Abbildung $A : I \longrightarrow M_{n\times n}(\mathbb{R})$, $t \longmapsto \begin{pmatrix} 0 & 0 & \kappa(t) & 0 \\ 0 & 0 & 0 & \kappa(t) \\ -\kappa(t) & 0 & 0 & 0 \\ 0 & -\kappa(t) & 0 & 0 \end{pmatrix}$, ist nach Voraussetzung stetig mit $\int_{t_0}^t A(s)ds = \begin{pmatrix} 0 & 0 & \int_{t_0}^t \kappa(s)ds & 0 \\ 0 & 0 & 0 & \int_{t_0}^t \kappa(s)ds \\ -\int_{t_0}^t \kappa(s)ds & 0 & 0 & 0 \\ 0 & -\int_{t_0}^t \kappa(s)ds & 0 & 0 \end{pmatrix}$.
Man kann direkt nachweisen, dass $\int_{t_0}^t A(s)ds \cdot A(t) = A(t) \cdot \int_{t_0}^t A(s)ds$ gilt: benutze dazu $A(t) = \kappa(t) \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$ und somit

$$\int_{t_0}^t A(s)ds = \int_{t_0}^t \kappa(s)ds \cdot \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Daraus folgt, dass die Lösung $y = \binom{\dot{c}}{n}$ des Cauchy-Problems $y' = A(t) \cdot y$, $y(t_0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, gegeben ist durch

$$y(t) = \exp(\int_{t_0}^{t} A(s)ds) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Man rechnet nach, dass dies genau die Formel für $\dot{c}$ liefert – und somit, nach einer zusätzlichen Integration nach t, auch c, siehe Beweis von Satz 1.29.

2. Im Beispiel 4.28.4 stellt man fest, dass für die Differentialgleichung $y'' - 2y' + 2y = 0$ das charakteristische Polynom $\chi_A(\lambda)$ der dazu assoziierten Matrix $A = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}$ der Form $\chi_A(\lambda) = \lambda^2 - 2\lambda + 2$ ist. Dies ist kein Zufall, wie wir im nächsten Satz sehen.

Satz 4.30 *Für $k \in \mathbb{N}$ sei eine homogene lineare Differentialgleichung k-ter Ordnung mit konstanten Koeffizienten in $\mathbb{R}$ gegeben, also*

$$y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)},$$

wobei $a_0, \ldots, a_{k-1} \in \mathbb{R}$. Dann hat die dazu assoziierte Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{k-1} \end{pmatrix}$$

das charakteristische Polynom $\chi_A(\lambda) = (-1)^k \cdot (\lambda^k - \sum_{j=0}^{k-1} a_j \lambda^j)$. Ist $\chi_A(\lambda) = (-1)^k \prod_{j=1}^{m} (\lambda - \lambda_j)^{m_j}$ die Zerlegung von $\chi_A(\lambda)$ in Linearfaktoren auf $\mathbb{C}$ (mit $\lambda_j \in \mathbb{C}$ und $m_j \in \mathbb{N} \setminus \{0\}$, für alle $1 \leq j \leq m$), so ist jede Lösung y der obigen Differentialgleichung der Form

$$y(t) = \sum_{j=1}^{m} e^{\mathrm{Re}(\lambda_j)t} \cdot (\cos(\mathrm{Im}(\lambda_j)t) p_j(t) + \sin(\mathrm{Im}(\lambda_j)t) q_j(t)),$$

wobei $p_j(t), q_j(t)$ Polynome vom Grad höchstens $m_j - 1$ mit Koeffizienten in $\mathbb{R}$ sind. Desweiteren ist jede Funktion dieser Form eine Lösung der Differentialgleichung[4] *und der Raum aller Lösungen ist ein k-dimensionaler reeller Vektorraum.*

Beweis: Bemerke erstens, dass die dazu assoziierte Matrix A durch Reduktion der Differentialgleichung k-ter Ordnung $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$ auf eine Differentialgleichung erster Ordnung (mit Koeffizienten im $\mathbb{R}^k$) entsteht: es ist nämlich $y^{(k)} =$

[4]Hier sollte man folgendes beachten: im Fall $\mathrm{Im}(\lambda_j) \neq 0$ ist $\overline{\lambda_j}$ automatisch ein Eigenwert mit derselben Vielfachheit von $\chi_A(\lambda)$, da die Koeffizienten dieses Polynoms sowieso reell sind; ggf. sollte m_j für die Vielfachheit *eines* der beiden Eigenwerte λ_j, $\overline{\lambda_j}$ stehen.

$a_0y + a_1y' + \cdots + a_{k-1}y^{(k-1)}$ genau dann, wenn

$$\begin{pmatrix} y \\ y' \\ \vdots \\ y^{(k-1)} \end{pmatrix}' = \begin{pmatrix} y' \\ y'' \\ \vdots \\ a_0y + a_1y' + \cdots + a_{k-1}y^{(k-1)} \end{pmatrix} = A \cdot \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(k-1)} \end{pmatrix}$$

für alle $t \in \mathbb{R}$ gilt, siehe Proposition 4.5. Die Formel für das charakteristische Polynom von A zeige man durch Induktion über k: für $k = 1$ ist $A = (a_0)$ und daher $\chi_A(\lambda) = a_0 - \lambda = (-1)(\lambda - a_0)$ wohl von der gewünschten Form. Angenommen, die Aussage gelte für $k-1 \geq 1$, dann würde man nach Entwicklung der Determinante nach der ersten Spalte folgendes bekommen:

$$\begin{aligned}
\chi_A(\lambda) &= \det \begin{pmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & -\lambda & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{k-1}-\lambda \end{pmatrix} \\
&= -\lambda \det \begin{pmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & -\lambda & 1 \\ a_1 & a_2 & a_3 & \cdots & a_{k-1}-\lambda \end{pmatrix} + (-1)^{k+1}a_0 \underbrace{\det \begin{pmatrix} 1 & & & 0 \\ -\lambda & 1 & & \\ & \ddots & \ddots & \\ 0 & & -\lambda & 1 \end{pmatrix}}_{=1} \\
&= -\lambda(-1)^{k-1}(\lambda^{k-1} - \sum_{j=0}^{k-2} a_{j+1}\lambda^j) + (-1)^{k+1}a_0 \\
&= (-1)^k(\lambda^k - \sum_{j=0}^{k-2} a_{j+1}\lambda^{j+1} - a_0) \\
&= (-1)^k(\lambda^k - \sum_{j=0}^{k-1} a_j\lambda^j),
\end{aligned}$$

was den Induktionsschritt $k-1 \longrightarrow k$ beweist. Korollar 4.26 liefert nun, dass jede $\mathbb{C}^k$-*wertige* Lösung $Y : \mathbb{R} \longrightarrow \mathbb{C}^k$ der homogenen Differentialgleichung $Y' = A \cdot Y$ der Form $Y(t) = \sum_{j=1}^m e^{\lambda_j t} P_j(t)$ (für alle $t \in \mathbb{R}$) ist, wobei $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ die paarweise verschiedenen *komplexen* Eigenwerte von A sind und $P_j(t) = \sum_{l=0}^{m_j-1} t^l C_{jl}$ Polynome vom Grad $m_j - 1$ (wobei $m_j \geq 1$ die algebraische Vielfachheit von m_j ist) mit Koeffizienten $C_{jl} \in \ker((A - \lambda_j \mathrm{I}_k)^{m_j}) \subset \mathbb{C}^k$ sind, wobei $0 \leq l \leq m_j - 1$ und $1 \leq j \leq m$. Die erste Koordinatenfunktion von Y in der kanonischen Basis des $\mathbb{C}^k$ (und nicht bezüglich der Zerlegung $\mathbb{C}^k = \bigoplus\limits_{1 \leq j \leq m} E_j$) ist nach Proposition 4.5 eine Lösung der Differentialgleichung $y^{(k)} = a_0y + a_1y' + \cdots + a_{k-1}y^{(k-1)}$. Diese Lösung hat die gleiche Form wie Y, d.h., sie lässt sich schreiben als $\sum_{j=1}^m e^{\lambda_j t} p_j(t)$, wo diesmal p_j ein Polynom vom Grad $m_j - 1$ mit *komplexen* Koeffizienten ist, für alle $1 \leq j \leq m$ (das Koeffizient von t^l in p_j ist die erste kanonische Koordinate von C_{jl}).

Es bleibt, zu beweisen, dass jede Funktion der Form $t \mapsto \sum_{j=1}^m e^{\lambda_j t} p_j(t)$, wobei p_j ein *beliebiges* komplexes Polynom vom Grad $m_j - 1$ ist für alle $1 \leq j \leq m$, eine Lösung der Differentialgleichung $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$ ist. Zuerst bemerken wir, dass die Abbildung $y \mapsto (y, y', \ldots, y^{(k-1)})$ einen *Isomorphismus* zwischen dem Raum aller Lösungen von $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$ und dem Raum aller Lösungen von $Y' = A \cdot Y$ definiert. Diese Abbildung ist nach Proposition 4.5 erstens wohldefiniert und sie ist offenbar linear. Desweiteren liefert, wegen der Form der Matrix A, jede Lösung Y von $Y' = A \cdot Y$ durch das Extrahieren der ersten Komponentenfunktion eine Lösung y von $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$; diese letztere Abbildung ist die Umkehrabbildung von $y \mapsto (y, y', \ldots, y^{(k-1)})$, wie man leicht sieht. Diese Bemerkung impliziert (zusammen mit Korollar 4.17, welches auch im komplexen Fall gilt), dass der Raum aller Lösungen von $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$ genau k-dimensional ist, also komplex k-dimensional für komplexwertige Lösungen und reell k-dimensional für reellwertige Lösungen. Schließlich ist es eine Übungsaufgabe, zu beweisen, dass die Funktionen $t \mapsto t^l e^{\lambda_j t}$, $0 \leq l \leq m_j - 1$, $1 \leq j \leq m$, als Vektoren aus beispielsweise $C^0(\mathbb{R}, \mathbb{C}^k)$, *linear unabhängig* sind. Da die Abbildung $\{y : \mathbb{R} \to \mathbb{C} \,|\, y \text{ löst } y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}\} \longrightarrow \mathrm{Span}_{\mathbb{C}}(t \mapsto t^l e^{\lambda_j t} \,|\, 0 \leq l \leq m_j - 1,\, 1 \leq j \leq m)$, $y \mapsto \sum_{j=1}^m e^{\lambda_j t} p_j(t)$ mit $y(t) = \sum_{j=1}^m e^{\lambda_j t} p_j(t)$ wie oben, linear und injektiv ist, ist sie aus Dimensionsgründen dann auch automatisch surjektiv. Das heißt, dass für *beliebige* komplexe Polynome $p_j(t)$ vom Grad $m_j - 1$ die Funktion $t \mapsto \sum_{j=1}^m e^{\lambda_j t} p_j(t)$ die Differentialgleichung $y^{(k)} = a_0 y + a_1 y' + \cdots + a_{k-1} y^{(k-1)}$ löst. Durch Bildung des Real- und Imaginärteiles bekommt man die letzte Behauptung des Satzes 4.30. Damit ist der Beweis zu Ende. □

Beispiel 4.31 Betrachte die lineare Differentialgleichung k-ter Ordnung $y^{(k)} = y$ im $\mathbb{R}$, wobei $k \in \mathbb{N} \setminus \{0\}$. Nach dem Satz 4.30 ist für die dazu assoziierte Matrix A das charakteristische Polynom $\chi_A(\lambda)$ gegeben durch: $\chi_A(\lambda) = (-1)^k(\lambda^k - 1)$ wegen $a_0 = 1$ und $a_j = 0$ für alle $j \geq 1$. Die Gleichung $\lambda^k = 1$ hat in $\mathbb{C}$ genau k Lösungen, die sogenannten *k-ten Wurzeln der Eins*, nämlich $\{1, e^{\frac{2i\pi}{k}}, e^{\frac{4i\pi}{k}}, \ldots, e^{\frac{2i(k-1)\pi}{k}}\} = \{e^{\frac{2i\ell\pi}{k}}, 0 \leq \ell \leq k-1\} \subset \mathbb{C}$. Der Satz 4.30 liefert, dass die Lösungen y von $y^{(k)} = y$ die Funktionen der Form

$$y(t) = \sum_{k=0}^{k-1} e^{\cos(\frac{2\ell\pi}{k})t} (\cos(\sin(\frac{2\ell\pi}{k})t) \cdot \alpha_\ell + \sin(\sin(\frac{2\ell\pi}{k})t) \cdot \beta\ell)$$

sind, für $\alpha_\ell, \beta_\ell \in \mathbb{R}$ beliebig, $0 \leq \ell \leq k-1$.[5]

4.3 Andere Typen von Differentialgleichungen

• $y' = f(t)$ (d.h., f hängt nicht von y ab). Ggf. ist jede Lösung der Form $y(t) = \int_{t_0}^t f(s)ds + Kste$

• *Autonome* Differentialgleichungen $y' = f(y)$ (f hängt nicht von t ab). Ggf. schreibe formal falls $y(t) \in \mathbb{R}$: $\frac{dy}{dt} = f(y) \iff \frac{dy}{f(y)} = dt \iff \int^y \frac{dx}{f(x)} = t + Kste$ (man bekommt t als Funktion von y). Diese Technik heißt *Trennung der Variablen*.

[5]Tatsächlich ist eine *Basis* des Raumes aller Lösungen der Differentialgleichung gegeben durch:

- falls k gerade: $\{e^t, e^{-t}, e^{\cos(\frac{2\ell\pi}{k})t}\cos(\sin(\frac{2\ell\pi}{k})t), e^{\cos(\frac{2\ell\pi}{k})t}\sin(\sin(\frac{2\ell\pi}{k})t), 1 \leq \ell \leq \frac{k}{2} - 1\}$.
- falls k ungerade: $\{e^t, e^{\cos(\frac{2\ell\pi}{k})t}\cos(\sin(\frac{2\ell\pi}{k})t), e^{\cos(\frac{2\ell\pi}{k})t}\sin(\sin(\frac{2\ell\pi}{k})t), 1 \leq \ell \leq \frac{k-1}{2}\}$.

• Allgemeiner kann diese Technik auf sogenannte *separierbare* Differentialgleichungen angewendet werden: $y' = f(y)g(t)$, ggf. ist

$$\frac{dy}{dt} = f(y)g(t) \Longleftrightarrow \frac{dy}{f(y)} = g(t)dt \Longleftrightarrow \int^y \frac{dx}{f(x)} = \int^t g(s)ds + Kste.$$

• *Exakte* Differentialgleichungen (im $\mathbb{R}$): $X_1(t,y) + X_2(t,y)y' = 0$, wobei $X : U \longrightarrow \mathbb{R}^2$ ein auf einer offenen Teilmenge $U \subset \mathbb{R}^2$ definiertes stetiges Vektorfeld ist. Hat X ein *Potential* P auf U, so ist y genau dann eine Lösung dieser Differentialgleichung, wenn

$$\frac{\partial P}{\partial t}(t, y(t)) + \frac{\partial P}{\partial x}(t, y(t))y'(t) = 0$$

für alle t gilt, d.h., wenn $\frac{d}{dt}(P(t, y(t)) = 0$ für alle t im Definitionsbereich gilt (Kettenregel), d.h., wenn $t \mapsto P(t, y(t)) = Kste$: die Kurve $t \longmapsto (t, y(t))$ muss eine *Niveaulinie* der Funktion P sein, siehe dazu den Satz über implizite Funktionen (Korollar 2.78).

• *Bernoullische* Differentialgleichung

$$y' = f_1(t)y + f_2(t)y^\alpha, \quad \text{wobei } \alpha \in \mathbb{R} \setminus \{1\}.$$

Ansatz für eine Lösung y: schreibe $z := y^{1-\alpha}$ und übersetze diese Gleichung in eine Gleichung für z. Bekomme folgende lineare Differentialgleichung $z' = (1-\alpha)(f_1(t)z + f_2(t))$.

• *Riccatische* Differentialgleichung:

$$y' = f_1(t)y + f_2(t)y^2 + f_3(t).$$

Ist $\hat{y}$ eine Lösung dieser Differentialgleichung, so ist jede Lösung y der Form $y = \hat{y} + z$, wobei z folgende Bernoullische Gleichung löst: $z' = f_2(t)z^2 + (2\hat{y}(t)f_2(t) + f_1(t))z$.

Kapitel 5

Untermannigfaltigkeiten und Differentialformen

5.1 Untermannigfaltigkeiten des $\mathbb{R}^n$

5.1.1 Definitionen

Definition 5.1 *Für $m, n \in \mathbb{N}$ mit $m \leq n$ heißt eine Teilmenge $M^m \subset \mathbb{R}^n$ genau dann* m-dimensionale Untermannigfaltigkeit *des $\mathbb{R}^n$, wenn für jedes $x \in M^m$ offene Teilmengen U und V von $\mathbb{R}^n$ mit $x \in U$ und ein C^∞-Diffeomorphismus $\varphi : U \longrightarrow V$ so existieren, dass*

$$\varphi(U \cap M) = V \cap (\mathbb{R}^m \times \{0_{n-m}\}).$$

gilt. Ggf. heißt m die Dimension *der Untermannigfaltigkeit M^m.*

Hierbei bezeichnet $0_{n-m} = (0, \ldots, 0) \in \mathbb{R}^{n-m}$ den Nullvektor aus $\mathbb{R}^{n-m}$. Insbesondere ist $\mathbb{R}^m \times \{0_{n-m}\} = \{x = (x_1, \ldots, x_m, 0, \ldots, 0) \in \mathbb{R}^n \,|\, (x_1, \ldots, x_m) \in \mathbb{R}^m\} \subset \mathbb{R}^n$.

Für $m = 2$ spricht man von *Fläche*. Bevor wir Beispiele von Untermannigfaltigkeiten vorstellen, geben wir eine äquivalente Definition von Untermannigfaltigkeiten an.

Proposition 5.2 *Für $m, n \in \mathbb{N}$ mit $m \leq n$ sei $M^m \subset \mathbb{R}^n$ eine Teilmenge. Dann sind äquivalent:*

i) M^m ist eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.

ii) Für jedes $x \in M^m$ existiert eine offene Umgebung U von x im $\mathbb{R}^n$ und eine C^∞-Abbildung $g : U \longrightarrow \mathbb{R}^{n-m}$ mit $d_y g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv für alle $y \in U \cap M$ und $U \cap M = g^{-1}(\{0\})$.

Zur Erinnerung ist $d_y g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ genau dann surjektiv, wenn die Jacobi-Matrix $J_g(y) \in M_{(n-m)\times n}(\mathbb{R})$ vollen Rang $n - m$ hat.

Beweis von Proposition 5.2:
i) $\Longrightarrow$ *ii)*: Sei M^m eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und $x \in M$ ein beliebiger Punkt. Nach Definition existiert eine offene Umgebung U von x im $\mathbb{R}^n$ und ein C^∞ Diffeomorphismus $\varphi : U \longrightarrow \varphi(U)$ auf eine offene Teilmenge $\varphi(U)$ von $\mathbb{R}^n$ mit $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Definiere

$$g : U \longrightarrow \mathbb{R}^{n-m}, \qquad y \longmapsto (\varphi_{m+1}(y), \ldots, \varphi_n(y)),$$

wobei $\varphi_j : U \longrightarrow \mathbb{R}$ die j-te Koordinatenfunktion von φ bezeichnet. Da φ C^∞ ist, sind alle Koordinatenfunktionen von φ C^∞ und daher ist g ebenfalls C^∞ auf U. Desweiteren gilt, für alle $y \in U$,

$$J_g(y) = \begin{pmatrix} {}^t(\nabla\varphi_{m+1}(y)) \\ \hline \vdots \\ \hline {}^t(\nabla\varphi_n(y)) \end{pmatrix} \in M_{(n-m)\times n}(\mathbb{R}).$$

Da φ ein Diffeomorphismus ist, ist $J_\varphi(y)$ invertierbar, was dazu äquivalent ist, dass die Vektoren $(\nabla\varphi_1)(y), \ldots, (\nabla\varphi_n)(y)$ linear unabhängig sind. Insbesondere sind die Vektoren $(\nabla\varphi_{m+1})(y), \ldots, (\nabla\varphi_n)(y)$ linear unabhängig, d.h., $J_g(y)$ hat vollen Rang $n-m$, d.h., $d_y g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ ist surjektiv. Außerdem gilt

$$\begin{aligned} g^{-1}(\{0\}) &= \{y \in U \,|\, g(y) = 0\} \quad \text{wobei hier } 0 = 0_{n-m} \\ &= \{y \in U \,|\, \varphi_{m+1}(y) = \ldots = \varphi_n(y) = 0\} \\ &= \{y \in U \,|\, \varphi(y) \in \mathbb{R}^m \times \{0_{n-m}\}\} \\ &= \varphi^{-1}(\mathbb{R}^m \times \{0_{n-m}\}) \\ &= \varphi^{-1}(\varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})) \\ &= \varphi^{-1}(\varphi(U \cap M)) \\ &= U \cap M, \end{aligned}$$

was zu zeigen war.
ii) $\Longrightarrow$ *i)*: Sei $x \in M$ ein beliebiger Punkt. Nach Voraussetzung existiert eine offene Umgebung U von x im $\mathbb{R}^n$ und eine C^∞ Abbildung $g : U \longrightarrow \mathbb{R}^{n-m}$ so, dass $d_y g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv ist und $U \cap M = g^{-1}(\{0\})$. Aus der Identität $J_g(x) = \begin{pmatrix} {}^t(\nabla g_1(x)) \\ \hline \vdots \\ \hline {}^t(\nabla g_{n-m}(x)) \end{pmatrix}$ und der Surjektivität von $d_x g$ leitet man her, dass die Vektoren $(\nabla g_1)(x), \ldots, (\nabla g_{n-m})(x)$ im $\mathbb{R}^n$ linear unabhängig sind. Nun erweitern wir diese Vektoren zu einer Basis von $\mathbb{R}^n$: es gibt Vektoren $v_1, \ldots, v_m \in \mathbb{R}^n$ so, dass $\{v_1, \ldots, v_m, (\nabla g_1)(x), \ldots, (\nabla g_{n-m})(x)\}$ eine Basis von $\mathbb{R}^n$ ist. Wir betrachten die Abbildung

$$\varphi : U \longrightarrow \mathbb{R}^n \qquad y \longmapsto \sum_{j=1}^m \langle y, v_j\rangle e_j + g(y),$$

wobei wir $\mathbb{R}^{n-m}$ mit $\{0_m\} \times \mathbb{R}^{n-m} \subset \mathbb{R}^n$ identifizieren, insbesondere wird $g(y) \in \mathbb{R}^{n-m}$ mit $(\underbrace{0, \ldots, 0}_{m}, g_1(y), \ldots, g_{n-m}(y)) \in \mathbb{R}^n$ identifiziert, für alle $y \in U$. Dann ist wegen g C^∞ und $y \mapsto \langle y, v_j\rangle$ linear (insbesondere C^∞) die Abbildung φ C^∞ mit $d_y\varphi(X) = \sum_{j=1}^m \langle X, v_j\rangle e_j +$

$d_yg(X)$ für alle $y \in U$ und $X \in \mathbb{R}^n$, insbesondere ist

$$J_\varphi(y) = \begin{pmatrix} {}^t v_1 \\ \vdots \\ {}^t v_m \\ {}^t(\nabla g_1)(y) \\ \vdots \\ {}^t(\nabla g_{n-m})(y) \end{pmatrix}$$

für alle $y \in U$. Für $y = x$ ist nach Wahl der Vektoren $v_1, \ldots, v_m$ die Matrix $J_\varphi(x)$ invertierbar. Der Umkehrsatz (Satz 2.76) liefert die Existenz einer offenen Umgebung V von x in U (daher auch offen im $\mathbb{R}^n$) so, dass $\varphi(V)$ eine offene Umgebung von $\varphi(x)$ im $\mathbb{R}^n$ ist und $\varphi_{|_V} : V \longrightarrow \varphi(V)$ ein Diffeomorphismus ist. Bemerke, dass $\varphi_{|_V} : V \longrightarrow \varphi(V)$ ein C^∞ Diffeomorphismus ist, da φ C^∞ ist[1]. Nun gilt

$$\begin{aligned} y \in V \cap M &\iff y \in V \text{ und } y \in U \cap M \\ &\iff y \in V \text{ und } g(y) = 0 \\ &\iff y \in V \text{ und } \varphi(y) = \sum_{j=1}^m \langle y, v_j \rangle e_j \\ &\iff y \in V \text{ und } \varphi(y) \in \mathbb{R}^m \times \{0_{n-m}\} \\ &\iff y \in \varphi_{|_V}^{-1}(\varphi(V) \cap (\mathbb{R}^m \times \{0_{n-m}\})), \end{aligned}$$

d.h., $\varphi_{|_V}(V \cap M) = \varphi(V) \cap (\mathbb{R}^m \times \{0_{n-m}\}) = \varphi_{|_V}(V) \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Somit erfüllt das Paar $(V, \varphi_{|_V})$ die Eigenschaften der Definition 5.1. Dies gilt für alle $x \in U$, somit ist M^m eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. $\square$

Beispiele 5.3

1. Jede offene Teilmenge M von $\mathbb{R}^n$ ist eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$: für jedes $x \in M$ wähle man $U = V = M$ und $\varphi = id_M$.

2. Jeder m-dimensionale affine Unterraum M des $\mathbb{R}^n$ ist eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Denn: schreibe $M = x_0 + E$, wobei $x_0 \in \mathbb{R}^n$ ein Punkt und E ein m-dimensionaler Untervektorraum des $\mathbb{R}^n$ ist. Fixiere eine Basis $(v_1, \ldots, v_m)$ von E_0 und erweitere sie zu einer Basis $(v_1, \ldots, v_n)$ des $\mathbb{R}^n$. Die Abbildung $g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$, $x_0 + \sum\limits_{i=1}^n \lambda_i v_i \longmapsto (\lambda_{m+1}, \ldots, \lambda_n)$, ist dann C^∞ mit $d_yg : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv für alle $y \in \mathbb{R}^n$ (denn g ist eine affin-lineare Abbildung, deren linearer Anteil die Abbildung $\sum_{i=1}^n \lambda_i v_i \mapsto (\lambda_{m+1}, \ldots, \lambda_n)$ ist, welche surjektiv ist) und $g^{-1}(\{0\}) = \{x_0 + \sum\limits_{i=1}^n \lambda_i v_i \,|\, \lambda_{m+1} = \cdots = \lambda_n = 0\} = x_0 + E = M$. Proposition 5.2 impliziert, dass M eine Untermannigfaltigkeit der Dimension m des $\mathbb{R}^n$ ist. Analog ist jede offene Teilmenge eines m-dimensionalen Unterraums eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.

[1]Dies folgt aus der Tatsache, dass die Abbildung $A \mapsto A^{-1}$, $\mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$, C^∞ ist, siehe Beweis von Satz 2.76.

3. Allgemeiner ist der Durchschnitt $U \cap M^m$, wobei $U \subset \mathbb{R}^n$ offen und $M^m \subset \mathbb{R}^n$ m-dimensionale Untermannigfaltigkeit, eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.

4. Sei $U \subset \mathbb{R}^m$ offen und $h : U \longrightarrow \mathbb{R}^{n-m}$ eine C^∞-Abbildung. Dann ist $\mathrm{Graph}(h) := \{(x, h(x)) \,|\, x \in U\} \subset \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Denn: betrachte die Abbildung

$$\begin{array}{rcl} g : U \times \mathbb{R}^{n-m} & \longrightarrow & \mathbb{R}^{n-m} \\ (x,y) & \longmapsto & y - h(x), \qquad (\text{wobei } x \in U, y \in \mathbb{R}^{n-m}). \end{array}$$

Wegen h C^∞ ist auch g C^∞ mit: $d_{(x,y)}g(X,Y) = Y - d_x h(X)$ für alle $X \in \mathbb{R}^m$ und $Y \in \mathbb{R}^{n-m}$. Wegen $d_{(x,y)}g(0,Y) = Y$ für alle $Y \in \mathbb{R}^{n-m}$ ist insbesondere $d_{(x,y)}g : \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv für alle $(x,y) \in U \times \mathbb{R}^{n-m}$. Außerdem gilt

$$\begin{aligned} g^{-1}(\{0\}) &= \{(x,y) \in U \times \mathbb{R}^{n-m} \,|\, g(x,y) = 0\} \\ &= \{(x,y) \in U \times \mathbb{R}^{n-m} \,|\, y - h(x) = 0\} \\ &= \{(x, h(x)) \in U \times \mathbb{R}^{n-m}\} \\ &= \mathrm{Graph}(h). \end{aligned}$$

Aus Proposition 5.2 folgt, dass $\mathrm{Graph}(h)$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ ist.

5. Sei $I \subset \mathbb{R}$ ein offenes Intervall und $c : I \longrightarrow \mathbb{R}^n$ eine regulär parametrisierte Kurve *ohne Selbstdurchschnitte*, d.h., c ist injektiv und $c : I \longrightarrow c(I)$ soll ein Homöomorphismus sein, wobei $c(I)$ die vom $\mathbb{R}^n$ induzierte Metrik trägt. Dann ist $c(I) =: M$ eine eindimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Denn: fixiere $t_0 \in I$ beliebig, erweitere $\dot{c}(t_0)$ zu einer Basis $(\dot{c}(t_0), v_2, \ldots, v_n)$ des $\mathbb{R}^n$ (möglich wegen $\dot{c}(t_0) \neq 0$) und betrachte die Abbildung $\psi : I \times \mathbb{R}^{n-1} \longrightarrow \mathbb{R}^n$, $\psi(t, t_2, \ldots, t_n) := c(t) + \sum_{j=2}^{n} t_j v_j$. Dann ist ψ C^∞ auf $I \times \mathbb{R}^{n-1}$ mit

$$J_\psi(t_0, 0, \ldots, 0) = \left(\dot{c}(t_0)\Big|v_2\Big|\ldots\Big|v_n\right) \in M_{n\times n}(\mathbb{R}).$$

Nach Konstruktion der Vektoren $v_2, \ldots, v_n$ hat $J_\psi(t_0, 0, \ldots, 0\}$ vollen Rang, insbesondere ist $J_\psi(t_0, 0, \ldots, 0)$ invertierbar. Aus dem Umkehrsatz (Satz 2.76) folgt, dass offene Umgebungen V von $(t_0, 0, \ldots, 0)$ und U von $\psi(t_0, 0, \ldots, 0)$ im $\mathbb{R}^n$ so existieren, dass $\psi : V \longrightarrow U$ ein C^∞-Diffeomorphismus ist. Die Abbildung $\varphi := \psi^{-1} : U \longrightarrow V$ ist dann ebenfalls ein C^∞-Diffeomorphismus mit

$$\varphi \circ c(t) = \psi^{-1} \circ c(t) = \psi^{-1}(c(t) + \sum_{j=2}^{n} 0 \cdot v_j) = (t, 0, \ldots, 0)$$

für alle $t \in I$ mit $c(t) \in U$ (und dies ist eine offene Umgebung von t_0 in I). Dies zeigt $\varphi(U \cap c(I)) = \varphi(U) \cap (\mathbb{R} \times \{0_{n-1}\})$. Somit ist $c(I)$ eine eindimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.

Bemerkung 5.4 Für eine auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ definierte differenzierbare Abbildung $g : U \longrightarrow \mathbb{R}^{n-m}$ und ein $x \in U$ ist $d_x g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ genau dann surjektiv, wenn die Jacobi-Matrix $J_g(x)$ vollen Rang hat, d.h., wenn die Vektoren $\nabla g_1(x), \ldots, \nabla g_{n-m}(x)$

linear unabhängig sind, wobei $g = (g_1, \ldots, g_{n-m})$. Denn: $J_g(x) = \begin{pmatrix} {}^t\nabla g_1(x) \\ \hline \vdots \\ \hline {}^t\nabla g_{n-m}(x) \end{pmatrix}$. Insbesondere ist, beim Satz über die Lagrange-Multiplikatoren (Satz 2.96), die Teilmenge $M := \{x \in U \,|\, g_j(x) = 0 \text{ für alle } 1 \leq j \leq k\}$ eine $n-k$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$[2].

Definition 5.5 *Sei $U \subset \mathbb{R}^n$ offen und $g : U \longrightarrow \mathbb{R}^k$ eine differenzierbare Abbildung, wobei $k \in \{0, 1, \ldots, n\}$. Man nennt $y \in \mathbb{R}^k$ genau dann* regulären Wert *von g, wenn $y \in g(U)$ und $d_x g : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ surjektiv für jedes $x \in g^{-1}(\{y\})$ ist.*

Eine unmittelbare Folgerung von Proposition 5.2 ist folgendes Korollar:

Korollar 5.6 (Satz über den regulären Wert) *Sei $g : U \longrightarrow \mathbb{R}^k$ eine auf einer offenen Teilmenge U des $\mathbb{R}^n$ definierte C^∞-Abbildung. Ist $y \in \mathbb{R}^k$ ein regulärer Wert von g, so ist $M := g^{-1}(\{y\})$ eine $n-k$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.*

Beweis: Die Abbildung $\overline{g} : U \longrightarrow \mathbb{R}^k$, $x \mapsto g(x) - y$, ist C^∞ mit $\overline{g}^{-1}(\{0\}) = M$. Außerdem ist, nach Definition eines regulären Werts, die Abbildung $d_x\overline{g} = d_x g : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ surjektiv für alle $x \in M$. Wende nun Proposition 5.2 an, wobei der Definitionsbereich U von $\overline{g}$ selbst als offene Umgebung jedes Punktes von $x \in M$ gewählt werden kann (und mit $\overline{g}$ als Funktion s.d. $M = U \cap M = \overline{g}^{-1}(\{0\})$). $\square$

Ggf. nennt man M eine *Niveaumenge* von g; für $n - k = 2$ (bzw. $n - k = 1$) wird M *Niveaufläche* (bzw. *Niveaulinie*) genannt.

Beispiele 5.7

1. Für ein beliebiges $x_0 \in \mathbb{R}^n$ betrachte die Abbildung $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x \longmapsto |x - x_0|^2$. Die Funktion ist C^∞ mit $d_x g(X) = 2\langle x - x_0, X\rangle$ für alle $X \in \mathbb{R}^n$. Insbesondere ist $\nabla g(x) = 2(x - x_0) \neq 0$ für alle $x \in \mathbb{R}^n \setminus \{x_0\}$. D.h., jedes $r \in]0, \infty[$ ist regulärer Wert von g. Korollar 5.6 liefert, dass für jedes $r \in]0, \infty[$ die Teilmenge $\{x \in \mathbb{R}^n \,|\, g(x) = r^2\} = S_r(x_0)$ eine $n-1$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ ist. Tatsächlich ist $S_r(x_0)$ die *Sphäre* vom Radius r und mit Zentrum x_0. Um anzudeuten, dass $S_r(x_0)$ Dimension $n-1$ hat, schreibt man $S_r(x_0) = S_r^{n-1}(x_0)$. Im Spezialfall $x_0 = 0 \in \mathbb{R}^n$ und $r = 1 \in \mathbb{R}$ schreibt man meistens $S_1^{n-1}(0) = S^{n-1}$.

2. Sei $g : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $(x, y, z) \longmapsto x^2 + y^2$. Dann ist g C^∞ mit $\nabla g(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix}$ für alle $(x, y, z) \in \mathbb{R}^3$. Insbesondere ist $\nabla g(x, y, z) \neq 0$ für alle $(x, y, z) \in \mathbb{R}^3$ mit $x^2 + y^2 \neq 0$. Daraus folgt, dass jedes $r \in]0, \infty[$ regulärer Wert von g ist. Korollar 5.6 liefert wiederum, dass für jedes $r \in]0, \infty[$ die Teilmenge $Z := \{x, y, z) \in \mathbb{R}^3 \,|\, g(x, y, z) = r^2\}$ eine zweidimensionale Untermannigfaltigkeit des $\mathbb{R}^3$ ist. Tatsächlich ist Z der *Zylinder* vom Radius r um die z-Achse.

[2]vorausgesetzt, die Funktionen $g_1, \ldots, g_k$ seien C^∞.

3. Sei $g : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $(x, y, z) \longmapsto x^2 + y^2 - z^2$ und $U := \mathbb{R}^3 \setminus \{0\}$. Dann ist g C^∞ mit $\nabla g(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ -2z \end{pmatrix}$ für alle $(x, y, z) \in \mathbb{R}^3$. Insbesondere ist $\nabla g(x, y, z) \neq 0$ für alle $(x, y, z) \in \mathbb{R}^3 \setminus \{0\}$. Korollar 5.6 liefert wiederum, dass für jedes $r \in \mathbb{R}$ die Teilmenge $g_{|_U}^{-1}(\{r\})$ eine zweidimensionale Untermannigfaltigkeit des $\mathbb{R}^3$ ist. Tatsächlich ist $g_{|_U}^{-1}(\{r\})$ ein *einschaliges Hyperboloid* für $r > 0$, ein *Doppelkegel* (ohne Spitze!) für $r = 0$ und ein *zweischaliges Hyperboloid* für $r < 0$.

4. Beachte aber, dass $0 \in \mathbb{R}$ kein regulärer Wert der Funktion g im letzten Beispiel ist, da $(\nabla g)(0, 0, 0) = 0$ und $(0, 0, 0) \in g^{-1}(\{0\})$ gilt. Insbesondere kann der Satz über den regulären Wert (Korollar 5.6) nicht angewendet werden. In unserem Fall ist $g^{-1}(\{0\})$ tatsächlich keine zweidimensionale Untermannigfaltigkeit des $\mathbb{R}^3$.

5. Ein weiteres Beispiel für eine Teilmenge M, welche keine eindimensionale Untermannigfaltigkeit des $\mathbb{R}^2$ ist, ist $M := \mathbb{R} \cdot (e_1 + e_2) \cup \mathbb{R}(e_1 - e_2) \subset \mathbb{R}^2$.

5.1.2 Der Tangentialraum

Definition 5.8 *Sei $M^m \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und $x \in M^m$ ein beliebiger Punkt. Der* Tangentialraum *von M^m an x ist*

$$T_xM := \{\dot{c}(0) \,|\, \exists\, \varepsilon > 0 \text{ und } c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n C^\infty \text{ mit } c(0) = x \text{ und } c(t) \in M \text{ für alle } t \in]-\varepsilon, \varepsilon[\} \subset \mathbb{R}^n.$$

Die Elemente von T_xM heißen Tangentialvektoren *am M^m an der Stelle x.*

Proposition 5.9 *Sei $M^m \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und $x \in M^m$ ein beliebiger Punkt. Sei $U \subset \mathbb{R}^n$ eine offene Umgebung von x so, dass eine offene Teilmenge V von $\mathbb{R}^n$ und ein C^∞-Diffeomorphismus $\varphi : U \longrightarrow V$ existiert mit $\varphi(U \cap M) = V \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Dann gilt*

$$T_xM = (d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\}) = \{(d_x\varphi)^{-1}(v, \underbrace{0, \ldots, 0}_{n-m}) \,|\, v \in \mathbb{R}^m\}.$$

Insbesondere ist T_xM ein m-dimensionaler Untervektorraum von $\mathbb{R}^n$.

Beweis: Wir beweisen beide Inklusionen auf elementare Weise. Sei $X \in T_xM$, dann existiert eine auf einem hinreichend kleinen offenen Intervall $]-\varepsilon, \varepsilon[$ definierte C^∞ Kurve $c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n$ mit $c(0) = x$, $\dot{c}(0) = X$ und $c(t) \in M$ für alle $t \in]-\varepsilon, \varepsilon[$. Wegen $c(0) = x \in U$ und der Stetigkeit von c (sowie wegen U offen im $\mathbb{R}^n$) kann bis auf Verkleinerung von ε angenommen werden, dass $c(]-\varepsilon, \varepsilon[) \subset U$ gilt. Nun gilt, wegen $c(t) \in U \cap M$, auch $\varphi \circ c(t) \in \mathbb{R}^m \times \{0_{n-m}\}$ für alle $t \in]-\varepsilon, \varepsilon[$. Insbesondere folgt mit der Kettenregel $d_x\varphi(\dot{c}(0)) = (\varphi \circ c)'(0) \in \mathbb{R}^m \times \{0_{n-m}\}$ (der Untervektorraum $\mathbb{R}^m \times \{0_{n-m}\}$ ist abgeschlossen im $\mathbb{R}^n$). Daraus folgt

$$X = \dot{c}(0) = (d_x\varphi)^{-1} \circ d_x\varphi(\dot{c}(0)) = (d_x\varphi)^{-1}((\varphi \circ c)'(0)) \in (d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\}).$$

Dies beweist $T_xM \subset (d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\})$.
Sei umgekehrt $X \in (d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\})$, dann existiert ein $Y \in \mathbb{R}^m \times \{0_{n-m}\}$ mit $X = (d_x\varphi)^{-1}(Y)$. Sei $\hat{x} := \varphi(x) \in \mathbb{R}^m \times \{0_{n-m}\}$ und definiere $c(t) := \varphi^{-1}(\hat{x} + tY)$; bemerke, dass wegen $\varphi(U)$ offen im $\mathbb{R}^n$ ein $\varepsilon > 0$ existiert mit $\hat{x} + tY \in \varphi(U)$ für alle

$t \in]-\varepsilon, \varepsilon[$, insbesondere ist c auf $]-\varepsilon, \varepsilon[$ wohldefiniert. Per Konstruktion ist c C^∞ mit $c(0) = \varphi^{-1}(\hat{x}) = x$ und $c(t) \in M$ (wegen $\hat{x} + tY \in \mathbb{R}^m \cap \{0_{n-m}\}$). Desweiteren folgt aus der Kettenregel

$$\dot{c}(0) = d_{\hat{x}}(\varphi^{-1})(\frac{d}{dt}_{|t=0}(\hat{x} + tY)) = d_{\hat{x}}(\varphi^{-1})(Y) = (d_x\varphi)^{-1}(Y) = X,$$

insbesondere gilt $X \in T_xM$. Wir bekommen $(d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\}) \subset T_xM$ und damit $T_xM = (d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\})$. Es bleibt, zu bemerken, dass $(d_x\varphi)^{-1}(\mathbb{R}^m \times \{0_{n-m}\})$ als Bild des m-dimensionalen Untervektorraumes $\mathbb{R}^m \times \{0_{n-m}\}$ des $\mathbb{R}^n$ unter dem Isomorphismus $(d_x\varphi)^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ ebenfalls ein m-dimensionaler Untervektorraum des $\mathbb{R}^n$ ist. □

Im Fall, wo M lokal als Niveaumenge einer Funktion beschrieben wird, kann der Tangentialraum als Lösungsraum einer linearen Gleichung (bzw. eines linearen Gleichungssystems) aufgefasst werden:

Proposition 5.10 *Sei $M^m \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und $x \in M^m$ ein beliebiger Punkt. Sei $U \subset \mathbb{R}^n$ eine offene Umgebung von x im $\mathbb{R}^n$ so, dass eine C^∞-Abbildung $g : U \longrightarrow \mathbb{R}^{n-m}$ mit $U \cap M = g^{-1}(\{0\})$ und $d_yg : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv für alle $y \in U \cap M$ existiert. Dann gilt*

$$T_xM = \ker(d_xg) = \{X \in \mathbb{R}^n \,|\, d_xg(X) = 0\}.$$

Beweis: Ist $X \in T_xM$, so existiert eine auf einem hinreichend kleinen offenen Intervall $]-\varepsilon, \varepsilon[$ definierte C^∞ Kurve $c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n$ mit $c(0) = x$, $c(t) \in M$ für alle $t \in]-\varepsilon, \varepsilon[$ sowie $\dot{c}(0) = X$. O.B.d.A. sei $c(t) \in U$ für alle $t \in]-\varepsilon, \varepsilon[$ (sonst verkleinere ε). Wegen $c(t) \in U \cap M$ gilt dann $g(c(t)) = 0$ für alle $t \in]-\varepsilon, \varepsilon[$, insbesondere gilt nach Ableitung und Anwendung der Kettenregel

$$0 = \frac{d}{dt}_{|t=0}(g(c(t))) = d_{c(0)}g(\dot{c}(0)) = d_xg(X),$$

d.h., $X \in \ker(d_xg)$. Dies zeigt $T_xM \subset \ker(d_xg)$. Wegen d_xg linear und surjektiv ist aber $\ker(d_xg)$ ein Untervektorraum des $\mathbb{R}^n$ mit $\dim(\ker(d_xg)) = n - (n-m) = m$. Da T_xM nach Proposition 5.9 auch m-dimensional ist, folgt $T_xM = \ker(d_xg)$. □

Bemerkungen 5.11

1. I.A. hat T_xM keine kanonische Basis! Siehe Beispiele unten.

2. Schreibt man $g = (g_1, \ldots, g_{n-m}) : U \longrightarrow \mathbb{R}^{n-m}$, so gilt, für jedes $x \in U$:

$$\begin{aligned}
\ker(d_xg) &= \{X \in \mathbb{R}^n \,|\, d_xg(X) = 0\} \text{ mit } d_xg(X) = (d_xg_1(X), \ldots, d_xg_{n-m}(X)) \\
&= \{X \in \mathbb{R}^n \,|\, d_xg_j(X) = 0 \text{ für alle } 1 \leq j \leq n-m\} \\
&= \{X \in \mathbb{R}^n \,|\, \langle \nabla g_j(x), X\rangle = 0 \text{ für alle } 1 \leq j \leq n-m\} \text{ wegen } d_xg_j(x) = \langle \nabla g_j(x), X\rangle \\
&= \bigcap_{1\leq j\leq n-m} \{X \in \mathbb{R}^n \,|\, \langle \nabla g_j(x), X\rangle = 0\} \\
&= \bigcap_{1\leq j\leq n-m} (\nabla g_j(x))^\perp.
\end{aligned}$$

Hierbei bezeichnet, für jede Teilmenge $T \subset \mathbb{R}^n$, $T^\perp := \{X \in \mathbb{R}^n \,|\, \langle X, t\rangle = 0 \ \forall\, t \in T\}$. Insbesondere ist, in der Proposition 5.10, $T_xM = \bigcap_{1\leq j\leq m} (\nabla g_j(x))^\perp$ für alle $x \in U \cap M$. Im Spezialfall $n - m = 1$ (d.h., $m = n-1$), ist $T_xM = (\nabla g(x))^\perp$: das Gradientenfeld von g steht orthogonal auf T_xM, d.h., auf die Niveaumenge $U \cap M$.

Beispiele 5.12

1. Sei $M := U$ offene Teilmenge des $\mathbb{R}^n$ und $x \in M$. Dann gilt $T_xM = \mathbb{R}^n$, denn: für die Wahl $U := M$, $V := U$ und $\varphi := id$ ist $T_xM = (d_x\varphi)^{-1}(\mathbb{R}^n) = \mathbb{R}^n$ nach Proposition 5.9. In diesem Fall hängt T_xM nicht von x ab.

2. Sei $M := x_0 + E$ ein m-dimensionaler affiner Unterraum des $\mathbb{R}^n$, wobei $x_0 \in \mathbb{R}^n$ ein Punkt und $E \subset \mathbb{R}^n$ ein m-dimensionaler Untervektorraum ist. Dann gilt $T_xM = E$ für alle $x \in M$. Denn: betrachte die im Beispiel 5.3.2 konstruierte Abbildung $g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$, $x_0 + \sum_{i=1}^n \lambda_i v_i \longmapsto (\lambda_{m+1}, \ldots, \lambda_n)$. Wir wissen schon, dass $d_xg : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ surjektiv ist für alle $x \in \mathbb{R}^{n-m}$ mit $g^{-1}(\{0\}) = M$. Proposition 5.10 liefert
$$\begin{aligned} T_xM = \ker(d_xg) &= \{X = \sum_{i=1}^n \lambda_i v_i \,|\, (\lambda_{m+1}, \ldots, \lambda_n) = 0_{n-m}\} \\ &= \{X = \sum_{i=1}^n \lambda_i v_i \,|\, \lambda_{m+1} = \cdots = \lambda_n = 0\} \\ &= \{\sum_{i=1}^m \lambda_i v_i \,|\, \lambda_i \in \mathbb{R} \text{ für alle } 1 \leq i \leq m\} \\ &= E. \end{aligned}$$
In diesem Fall hängt T_xM auch nicht von x ab.

3. Sei M eine regulär parametrisierte Kurve ohne Selbstdurchschnitte wie im Beispiel 5.3.5. Dann ist für jedes $t_0 \in I$, der Tangentialraum an $c(t_0)$ von $M := c(I)$ gegeben durch $T_{c(t_0)}M = \mathbb{R} \cdot \dot{c}(t_0)$: der Tangentialraum ist die zur üblichen Tangente gehörige Gerade im $\mathbb{R}^n$.

4. Sei $M := S^{n-1} = \{x \in \mathbb{R}^n,\ |x| = 1\}$ die $n-1$-dimensionale Einheitssphäre um den Ursprung im $\mathbb{R}^n$. Wir haben schon gesehen, dass $M = g^{-1}(\{1\})$, wobei $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $x \longmapsto |x|^2$. Nach Proposition 5.10 gilt dann wegen $\nabla g(x) = 2x$
$$\begin{aligned} T_xM &= (\nabla g(x))^\perp = \{X \in \mathbb{R}^n \,|\, \langle \nabla g(x), X\rangle = 0\} \\ &= x^\perp \end{aligned}$$
für alle $x \in S^{n-1}$.

5. Sei $M := Z := \{(x, y, z) \in \mathbb{R}^3 \,|\, x^2 + y^2 = 1\}$ der Zylinder vom Radius 1 um die z-Achse im $\mathbb{R}^3$. Wir wissen schon (Beispiel 5.7.2), dass $M = g^{-1}(\{1\})$ ist, wobei $g : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $(x, y, z) \longmapsto x^2 + y^2$. Wegen $(\nabla g)(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix}$ ist, nach Proposition 5.10, der Tangentialraum $T_{(x,y,z)}M$ an $(x, y, z) \in M$ von M gegeben durch
$$\begin{aligned} T_{(x,y,z)}M &= \ker(d_{(x,y,z)}g) = (\nabla g(x, y, z))^\perp = \{X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \in \mathbb{R}^3 \,|\, 2xX_1 + 2yX_2 = 0\} \\ &= \{X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \in \mathbb{R}^3 \,|\, \exists\, \lambda \in \mathbb{R} \text{ s.d. } (X_1, X_2) = \lambda(-y, x)\} \\ &= \begin{pmatrix} -\lambda y \\ \lambda x \\ X_3 \end{pmatrix} \,|\, \lambda, X_3 \in \mathbb{R}\}. \end{aligned}$$

Hier haben wir $(x,y)^\perp = \mathbb{R}\cdot(-y,x)$ im $\mathbb{R}^2$ benutzt.

6. Sei $M := \{(x,y,z)\in\mathbb{R}^3\setminus\{0\}\,|\,x^2+y^2-z^2=0\}$ der Doppelkegel ohne Spitze aus dem Beispiel 5.7.3. Dann ist $M = g^{-1}(\{0\})$, wobei $g:\mathbb{R}^3\setminus\{0\}\longrightarrow\mathbb{R}$, $(x,y,z)\longmapsto x^2+y^2-z^2$. Wegen $(\nabla g)(x,y,z) = (2x,2y,-2z)$ ist, nach Proposition 5.10.:

$$\begin{aligned} T_{(x,y,z)}M &= \{X=(X_1,X_2,X_3)\in\mathbb{R}^3\,|\,2xX_1+2yX_2-2zX_3=0\} \\ &= \{X=(X_1,X_2,X_3)\in\mathbb{R}^3\,|\,xX_1+yX_2-zX_3=0\}. \end{aligned}$$

Beachte, dass in diesem Fall $(x,y,z)\in T_{(x,y,z)}M$ gilt.

5.1.3 Differentialabbildung und Vektorfelder

Definition 5.13 *Seien M^m und N^p Untermannigfaltigkeiten des $\mathbb{R}^n$ bzw. $\mathbb{R}^q$. Eine stetige Abbildung $f: M^m\longrightarrow N^p$ heißt genau dann* differenzierbar *(bzw. C^k, wobei $k\in\mathbb{N}\cup\{\infty\}$, $k\geq 1$), wenn für alle C^∞-Diffeomorphismen $\varphi: U\longrightarrow\varphi(U)$, $\psi: V\longrightarrow\psi(V)$ mit $U,\varphi(U)$ offen im $\mathbb{R}^n$, $\varphi(U\cap M)=\varphi(U)\cap(\mathbb{R}^m\times\{0_{n-m}\})$ (bzw. $V,\psi(V)$ offen im $\mathbb{R}^q$, $\psi(V\cap N)=\psi(V)\cap(\mathbb{R}^p\times\{0_{q-p}\})$) und $f(U\cap M)\subset V\cap N$ die Abbildung*

$$\psi\circ f\circ\varphi^{-1}:\varphi(U\cap M)\longrightarrow\psi(V\cap N)$$

differenzierbar (bzw. C^k) ist.

Bemerkungen 5.14

1. Per Konvention ist eine Abbildung $f: M^m\longrightarrow N^p$ genau dann C^0, wenn f im üblichen Sinne stetig ist, wobei M^m und N^p die vom $\mathbb{R}^n$ bzw. $\mathbb{R}^q$ induzierten Metriken tragen.

2. Für M^m offen im $\mathbb{R}^m$ stimmt der obige Begriff von Differenzierbarkeit (bzw. C^k-Regularität) mit dem üblichen Begriff von Differenzierbarkeit (bzw. C^k-Regularität) aus dem Kapitel 2 überein, wobei f als Abbildung von M nach $\mathbb{R}^q$ aufgefasst wird.

3. Ist $f: M^m\longrightarrow N^p$ die Einschränkung auf M^m einer differenzierbaren (bzw. C^k) Abbildung $\tilde f: U\longrightarrow\mathbb{R}^q$, wobei $U\subset\mathbb{R}^n$ offen ist, so ist $f=\tilde f_{|_M}: M\longrightarrow N$ differenzierbar (bzw. C^k).

Definition 5.15 *Sei $f: M^m\longrightarrow N^p$ eine differenzierbare Abbildung zwischen Untermannigfaltigkeiten. Die* Differentialabbildung *(oder* Tangentialabbildung*) von f an der Stelle $x\in M^m$ ist die Abbildung*

$$\begin{aligned} d_xf: T_xM &\longrightarrow T_{f(x)}N \\ X=\dot c(0) &\longmapsto d_xf(X) := \dot{\widehat{f\circ c}}(0), \end{aligned}$$

wobei $c:]-\varepsilon,\varepsilon[\longrightarrow\mathbb{R}^n$ C^∞-Kurve mit $c(0)=x$ und $c(]-\varepsilon,\varepsilon[)\subset M$ ist.

Proposition 5.16 *Sei $f: M^m\longrightarrow N^p$ eine differenzierbare Abbildung zwischen Untermannigfaltigkeiten und $x\in M^m$ ein Punkt. Sei U eine offene Umgebung von x im $\mathbb{R}^n$ so, dass ein C^∞-Diffeomorphismus $\varphi: U\longrightarrow\varphi(U)$ existiert mit $\varphi(U)$ offen im $\mathbb{R}^n$ und $\varphi(U\cap M)=\varphi(U)\cap(\mathbb{R}^m\times\{0_{n-m}\})$. Dann gilt, für jeden Tangentialvektor $X\in T_xM$:*

$$d_xf(X) = d_{\varphi(x)}(f\circ\varphi^{-1})(d_x\varphi(X)).$$

Insbesondere ist $d_xf: T_xM\longrightarrow T_{f(x)}N$ linear.

Beweis: Sei $X \in T_xM$ und $c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n$ eine auf einem hinreichend kleinen offenen Intervall $]-\varepsilon, \varepsilon[$ definierte C^∞ Kurve mit $c(0) = x$, $c(t) \in M$ für alle $t \in]-\varepsilon, \varepsilon[$ sowie $\dot{c}(0) = X$. O.B.d.A. sei $c(]-\varepsilon, \varepsilon[) \subset U$ (sonst verkleinere ε). Setze $\hat{c}(t) := \varphi \circ c(t)$ für alle $t \in]-\varepsilon, \varepsilon[$. Dann ist $\hat{c}$ eine C^∞ Kurve $]-\varepsilon, \varepsilon[\longrightarrow \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$ mit $\dot{\hat{c}}(0) = d_x\varphi(\dot{c}(0)) = d_x\varphi(X)$. Daraus folgt, per Definition der Differentialabbildung und mit der Kettenregel,

$$\begin{aligned} d_xf(X) &= (f \circ c)'(0) \\ &= (f \circ \varphi^{-1} \circ \varphi \circ c)'(0) \\ &= (f \circ \varphi^{-1} \circ \hat{c})'(0) \\ &= d_{\hat{c}(0)}(f \circ \varphi^{-1})(\dot{\hat{c}}(0)) \\ &= d_{\varphi(x)}(f \circ \varphi^{-1})(d_x\varphi(X)), \end{aligned}$$

was zu beweisen war. □

Beispiele 5.17

1. Für M^m offen im $\mathbb{R}^m$ wissen wir schon, dass $T_xM = \mathbb{R}^m$ für alle $x \in M^m$ gilt. Die Differentialabbildung $d_xf : T_xM \longrightarrow T_{f(x)}N$ stimmt dann mit dem üblichen Differential $d_xf : \mathbb{R}^m \longrightarrow \mathbb{R}^q$ überein (insbesondere muss $d_xf(\mathbb{R}^m) \subset T_{f(x)}N$ gelten).

2. Ist $f : M^m \longrightarrow N^p$ die Einschränkung auf M^m einer differenzierbaren Abbildung $\tilde{f} : U \longrightarrow \mathbb{R}^q$, wobei $U \subset \mathbb{R}^n$ offen ist, so gilt $d_xf = d_x\tilde{f}_{|T_xM} : T_xM \longrightarrow T_{f(x)}N$, für alle $x \in M^m$. Denn: die Kettenregel liefert $d_xf(\dot{c}(0)) = \frac{d}{dt}(\tilde{f} \circ c)_{t=0} = d_x\tilde{f}(\dot{c}(0))$ für jede C^∞-Kurve $c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n$ mit $c(]-\varepsilon, \varepsilon[) \subset M^m$ und $c(0) = x$.

3. Sei $f : S^2 \longrightarrow \mathbb{R}$, $x = (x_1, x_2, x_3) \longmapsto x_3$. Dann ist $f = \tilde{f}_{|S^2}$, wobei $\tilde{f} : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $x = (x_1, x_2, x_3) \longmapsto x_3$. Da $\tilde{f}$ C^∞ ist, ist $\tilde{f}_{|S^2} = f$ ebenfalls C^∞ und es gilt, für alle $x \in S^2$ und $X = (X_1, X_2, X_3) \in \mathbb{R}$:

$$\begin{aligned} d_xf(X) &= d_x\tilde{f}(X) \\ &= X_3 \quad \text{da } \tilde{f} \text{ linear ist.} \end{aligned}$$

Beachte insbesondere, dass $d_{e_3}f = 0$ sowie $d_{-e_3}f = 0$ gelten, denn $T_{e_3}S^2 = T_{-e_3}S^2 = \{X = (X_1, X_2, X_3) \in \mathbb{R}^3 \,|\, X_3 = 0\}$.

Proposition 5.18 *Sei M^m eine Untermannigfaltigkeit des $\mathbb{R}^n$.*

i) Sind f und g differenzierbare (bzw. C^k) Abbildungen $M^m \longrightarrow \mathbb{R}^p$, so ist $\lambda f + \mu g : M^m \longrightarrow \mathbb{R}^p$ für alle $\lambda, \mu \in \mathbb{R}$ differenzierbar (bzw. C^k) und es gilt $d_x(\lambda f + \mu g) = \lambda d_xf + \mu d_xg : T_xM \longrightarrow \mathbb{R}^p$, für alle $x \in M^m$.

ii) Sind f und g differenzierbare (bzw. C^k) Abbildungen $M^m \longrightarrow \mathbb{R}^p$, so ist $\langle f, g\rangle : M^m \longrightarrow \mathbb{R}^p$, $x \longmapsto \langle f(x), g(x)\rangle$ auch differenzierbar (bzw. C^k) und es gilt (Produktregel):

$$d_x(\langle f, g\rangle)(X) = \langle d_xf(X), g(x)\rangle + \langle f(x), d_xg(X)\rangle$$

für alle $x \in M^m$ und $X \in T_xM$.

iii) Sind $M^m \xrightarrow{f} N^p$ *und* $N^p \xrightarrow{g} Q^d$ *differenzierbare (bzw.* C^k*) Abbildungen zwischen Untermannigfaltigkeiten, so ist* $g \circ f : M^m \longrightarrow Q^d$ *ebenfalls differenzierbar (bzw.* C^k*) und es gilt (Kettenregel)*

$$d_x(g \circ f) = d_{f(x)}g \circ d_x f$$

für alle $x \in M^m$*.*

Beweis: Die Proposition folgt unmittelbar aus den entsprechenden Eigenschaften für auf offenen Teilmengen des $\mathbb{R}^n$ definierte differenzierbare Abbildungen, siehe Propositionen 2.63 und Beispiele 2.64. □

Definition 5.19 *Sei* M^m *eine* m*-dimensionale Untermannigfaltigkeit des* $\mathbb{R}^n$*.*

i) Ein Tangentialvektorfeld *auf* M^m *ist eine Abbildung* $X : M^m \longrightarrow \mathbb{R}^n$ *mit* $X(x) \in T_xM$ *für alle* $x \in M^m$*.*

ii) Ein Normalenvektorfeld *auf* M^m *ist eine Abbildung* $X : M^m \longrightarrow \mathbb{R}^n$ *mit* $X(x) \perp T_xM$ *für alle* $x \in M^m$*.*

Entsprechend heißt X *genau dann differenzierbar (bzw.* C^k*), wenn* X *als Abbildung zwischen der Untermannigfaltigkeit* M^m *und* $\mathbb{R}^n$ *differenzierbar (bzw.* C^k*) ist.*

Bemerkung 5.20 Für eine offene Teilmenge M^m des $\mathbb{R}^m$ stimmt der Begriff von Tangentialvektorfeld mit dem Begriff des Abschnitts 2.4 (Definition 2.99) überein: in diesem Fall ist ein Vektorfeld eine Abbildung $X : U \longrightarrow \mathbb{R}^m$, da $T_xM = \mathbb{R}^m$ für alle $x \in M$ gilt.

Beispiele 5.21

1. Sei $M := S^1 = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$ der Einheitskreis um den Ursprung im $\mathbb{R}^2$. Nach dem Beispiel 5.7.1 ist S^1 eine eindimensionale Untermannigfaltigkeit des $\mathbb{R}^2$ mit $T_{(x,y)}S^1 = \{(z, t) \in \mathbb{R}^2 \,|\, xz + yt = 0\} = \mathbb{R} \cdot (-y, x)$ für alle $(x, y) \in S^1$. Insbesondere ist $X : S^1 \longrightarrow \mathbb{R}^2$, $(x, y) \longmapsto (-y, x)$, ein Tangentialvektorfeld auf S^1.

2. Sei $M := S^{n-1} \subset \mathbb{R}^n$ die Einheitssphäre um den Ursprung im $\mathbb{R}^n$, wobei $n \in \mathbb{N} \setminus \{0, 1\}$. Dann definiert

$$X : S^{n-1} \longrightarrow \mathbb{R}^n,\; x \longmapsto e_n - \langle e_n, x\rangle x = \begin{pmatrix} -x_n x_1 \\ -x_n x_2 \\ \vdots \\ -x_n x_{n-1} \\ 1 - x_n^2 \end{pmatrix}$$

ein Tangentialvektorfeld auf S^{n-1}. Beachte insbesondere, dass $X(e_n) = X(-e_n) = 0 \in \mathbb{R}^n$ gilt.

3. Sei wiederum $M := S^{n-1} \subset \mathbb{R}^n$ mit $n \geq 2$. Dann definiert $X : S^{n-1} \longrightarrow \mathbb{R}^n$, $x \longmapsto x$, ein Normalenvektorfeld auf S^{n-1}. Beachte, dass $X(x) = \frac{1}{2}(\nabla g)(x)$ für alle $x \in S^{n-1}$ gilt, wobei $g(x) := |x|^2$ für alle $x \in \mathbb{R}^n$.

4. Allgemeiner ist nach der Bemerkung 5.11.2, jedes Vektorfeld der Form $X(x) := \lambda(x)(\nabla g)(x)$ auf einer Untermannigfaltigkeit M^m der Form $M^m = g^{-1}(\{0\})$ für $g : \mathbb{R}^{m+1} \longrightarrow \mathbb{R}$ mit $0 \in \mathbb{R}$ als regulärem Wert und für eine beliebige Abbildung $\lambda : M^m \longrightarrow \mathbb{R}$, ein Normalenvektorfeld auf M^m.

5. Noch allgemeiner ist, für $M^m := g^{-1}(\{0_{n-m}\})$ mit $g : U \longrightarrow \mathbb{R}^{n-m}$, $U \subset \mathbb{R}^n$ offen, $0 \in \mathbb{R}^{n-m}$ regulärer Wert von g, jedes Vektorfeld der Form $X(x) := \sum\limits_{j=1}^{n-m} \lambda_j(x)(\nabla g_j)(x)$, $x \in M^m$, $g = (g_1, \ldots, g_{n-m})$, $\lambda_j : M^m \longrightarrow \mathbb{R}$, ein Normalenvektorfeld auf M^m, siehe Bemerkung 5.11.2.

Wie im $\mathbb{R}^n$ ist jede differenzierbare Abbildung mit verschwindedem Differential immer lokal konstant:

Proposition 5.22 *Sei $f : M^m \longrightarrow N^p$ eine differenzierbare Abbildung zwischen Untermannigfaltigkeiten. Angenommen, M^m sei* zusammenhängend *(bzgl. der induzierten Metrik) und es gelte $d_x f = 0$ für alle $x \in M^m$. Dann ist f konstant auf M^m.*

Beweis: Für ein $c \in f(M) \subset N$ betrachte man $W := f^{-1}(\{c\}) = \{x \in M \mid f(x) = c\} \subset M$. Da f per Definition stetig ist, ist W eine abgeschlossene Teimenge von M und wegen $c \in f(M)$ ist $W \neq \varnothing$. Ist $x \in W$, so existiert eine offene Umgebung U von x im $\mathbb{R}^n$ und ein C^∞ Diffeomorphismus $\varphi : U \longrightarrow \varphi(U)$ auf eine offene Teilmenge $\varphi(U)$ von $\mathbb{R}^n$ mit $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Nach eventueller Verkleinerung von U kann angenommen werden, dass $U \cap M$ zusammenhängend ist (dies ist der Fall, wenn z.B. $\varphi(U) \subset \mathbb{R}^n$ ein offener Ball ist). Dann ist die Abbildung $f \circ \varphi^{-1} : \varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\}) \longrightarrow \mathbb{R}^q$ differenzierbar mit $d_y(f \circ \varphi^{-1}) = d_{\varphi^{-1}(y)} f \circ d_y(\varphi^{-1}) = 0$ für alle $y \in \varphi(U \cap M)$ (verwende Proposition 5.16), wobei $N^p \subset \mathbb{R}^q$. Korollar 2.71 liefert, dass dann $f \circ \varphi^{-1}$ konstant auf $\varphi(U \cap M)$ ist. Daraus folgt, dass f auf $U \cap M$ konstant ist, insbesondere gilt $U \cap M \subset W$. Dies beweist, dass W auch offen in M ist. Da M zusammenhängend ist, folgt $W = M$, was zu beweisen war. □

Lokale Extrema lassen sich auch analog wie im $\mathbb{R}^n$ charakterisieren:

Proposition 5.23 *Sei $f : M^m \longrightarrow \mathbb{R}$ eine auf einer Untermannigfaltigkeit differenzierbare reellwertige Funktion. Angenommen, ein $u \in M$ sei ein* lokales Extremum *von f, d.h., es gäbe eine offene Umgebung U von x im $\mathbb{R}^n$ mit entweder $f(u) \leq f(x)$ für alle $x \in U \cap M$ (falls u lokales Minimum) oder $f(u) \geq f(x)$ für alle $x \in U \cap M$ (falls u lokales Maximum). Dann gilt $d_u f = 0$, d.h., $d_u f(X) = 0$ für alle $X \in T_u M$.*

Beweis: Wir behandeln den Fall, wo u ein lokales Minimum von f auf M ist (ersetzt man f durch $-f$, so ergibt sich der andere). Nach Voraussetzung existiert eine offene Umgebung U von u im $\mathbb{R}^n$ mit $f(u) \leq f(x)$ für alle $x \in U \cap M$. Sei $c :]-\varepsilon, \varepsilon[\longrightarrow \mathbb{R}^n$ eine beliebige C^∞ Kurve mit $c(0) = u$ und $c(t) \in M$ für alle $t \in]-\varepsilon, \varepsilon[$. O.B.d.A. sei $c(]-\varepsilon, \varepsilon[) \subset U$. Dann gilt $f(u) = (f \circ c)(0) \leq (f \circ c)(t)$ für alle $t \in]-\varepsilon, \varepsilon[$. Daraus folgt $(f \circ c)'(0) = 0$, d.h., per Definition der Differentialabbildung, $d_u f(\dot{c}(0)) = 0$ (vgl. Satz 2.91.i), wo allerdings die Voraussetzung "f differenzierbar" reicht). Dies gilt für alle Kurven wie oben, somit gilt $d_u f = 0$. □

Korollar 5.24 (Lagrange-Multiplikatoren, zweite Version) *Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge U von $\mathbb{R}^n$ definierte differenzierbare Funktion. Sei $M^m \subset U$ eine Untermannigfaltigkeit des $\mathbb{R}^n$, welche in U enthalten ist. Ist $u \in M^m$ ein lokales Extremum von $f_{|M}$, so gilt $(\nabla f)(u) \in (T_uM)^\perp$.*

Beweis: Nach Proposition 5.23 gilt $d_u(f_{|M}) = 0$. Da aber $d_u(f_{|M}) = d_uf|_{|T_uM}$ gilt (siehe Beispiel 5.17.2), folgt $d_uf(X) = 0$ für alle $X \in T_uM$, d.h., $\langle(\nabla f)(u), X\rangle = 0$ für alle $X \in T_uM$, d.h., $(\nabla f)(u) \in (T_uM)^\perp$. □

Ist $M^m = g^{-1}(\{0\}$ das Urbild eines regulären Werts, so gilt

$$\begin{aligned}(T_uM)^\perp &= \ker(d_ug)^\perp \\ &= \big((\nabla g_1)(u)^\perp \cap \ldots \cap (\nabla g_{n-m})(u)^\perp\big)^\perp \\ &= \mathrm{Span}((\nabla g_1)(u), \ldots, (\nabla g_{n-m})(u)).\end{aligned}$$

Insbesondere impliziert Korollar 5.24 den Satz 2.96.

5.2 Differentialformen

5.2.1 Multilineare Algebra

In diesem Abschnitt sei V ein endlichdimensionaler reeller Vektorraum (später: $V = \mathbb{R}^n$ oder $V = T_xM$).

Definition 5.25

1. *Eine* Linearform *auf V ist eine lineare Abbildung $\theta : V \longrightarrow \mathbb{R}$.*
2. *Für $k \in \mathbb{N}\backslash\{0\}$ ist eine k-Multilinearform auf V eine Abbildung $\omega : \underbrace{V \times \cdots \times V}_{k-mal} \longrightarrow \mathbb{R}$, welche in jedem Eintrag linear ist: für alle $v_1, \ldots, v_{k-1} \in V$ und alle $j \in \{1, \ldots, k\}$ ist $V \longrightarrow \mathbb{R}$, $v \longmapsto \omega(v_1, \ldots, v, \ldots, v_{k-1})$, wobei v an der j-ten Stelle eingesetzt wird, linear.*
3. *Eine k-Multilinearform $\omega : V \times \cdots \times V \longrightarrow \mathbb{R}$ heißt genau* alternierend, *wenn $\omega(v_1, \ldots, v_k) = 0$ für alle $v_1, \ldots, v_k \in V$ mit $v_i = v_j$ für mindestens ein Paar $(i, j) \in \{1, \ldots, k\}^2$ mit $i \neq j$.*

Bezeichnungen 5.26 Die Menge aller Linearformen auf V bezeichnen wir mit V^*; beachte, dass V^* die natürliche Struktur eines reellen Vektorraumes trägt. Die Menge aller alternierenden k-Multilinearformen auf V bezeichnen wir mit Λ^kV^*. Per Konvention ist $\Lambda^0V^* := \mathbb{R}$. Per Definition ist $\Lambda^1V^* = V^*$.

Bemerkungen 5.27

1. Ebenso wie V^* trägt Λ^kV^* die natürliche Struktur eines reellen Vektorraumes (jede Linearkombination alternierender k-Multilinearformen ist wieder eine alternierende k-Multilinearform). Ist V n-dimensional, so gilt $\Lambda^kV^* = \{0\}$ für alle $k \in \mathbb{N}$ mit $k > n$. Denn: wählt man eine Basis $\{u_1, \ldots, u_n\}$ von V, so lässt sich jeder Vektor

von V in dieser Basis zerlegen. Sind $v_1, \dots, v_k$ beliebige Vektoren, so zerlege sie in der Basis $\{u_1, \dots, u_k\}$:

$$v_j = \sum_{i=1}^{n} \lambda_{ij} u_i \qquad \text{mit } \lambda_{ij} \in \mathbb{R}.$$

Ist $\omega \in \Lambda^k V^*$, so gilt

$$\begin{aligned} \omega(v_1, \dots, v_k) &= \omega(\sum_{i_1=1}^{n} \lambda_{i_1 1} u_{i_1}, \dots, \sum_{i_k=1}^{n} \lambda_{i_k k} u_{i_k}) \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \lambda_{i_1 1} \dots \lambda_{i_k k} \omega(u_{i_1}, \dots, u_{i_k}) \quad \text{wegen } \omega \text{ multilinear.} \end{aligned}$$

Wegen $k > n$ taucht einer der Vektoren $u_1, \dots, u_n$ mindestens zweimal in jedem Tupel $(u_{i_1}, \dots, u_{i_k})$ auf; da ω alternierend ist, folgt $\omega(u_{1_1}, \dots, u_{i_k}) = 0$ und damit $\omega(v_1, \dots, v_k) = 0$. Da $v_1, \dots, v_k$ beliebig gewählt werden konnten, folgt $\omega = 0$, was zu beweisen war.

2. Eine Multilinearform ω auf V heißt genau dann *schiefsymmetrisch*, wenn für alle $v_1, \dots, v_k \in V$ und alle $1 \leq i < j \leq k$ gilt:

$$\omega(v_1, \dots, \underset{i}{v_i}, \dots, \underset{j}{v_j}, \dots, v_k) = -\omega(v_1, \dots, \underset{i}{v_j}, \dots, \underset{j}{v_i}, \dots, v_k).$$

Es ist zu bemerken: ω ist genau dann schiefsymmetrisch, wenn ω alternierend ist. Denn: ist ω schiefsymmetrisch, so gilt

$$\omega(v_1, \dots, \underset{i}{v}, \dots, \underset{j}{v}, \dots, v_{k-2}) = -\omega(v_1, \dots, \underset{i}{v}, \dots, \underset{j}{v}, \dots, v_{k-2}),$$

insbesondere $\omega(v_1, \dots, v, \dots, v, \dots, v_{k-2}) = 0$. Sei umgekehrt ω alternierend, dann gilt für alle $v_1, \dots, v_k \in V$: $\omega(v_1, \dots, v_i \underset{i}{+} v_j, \dots, v_i \underset{j}{+} v_j, \dots, v_k) = 0$, daher

$$\begin{aligned} &\omega(v_1, \dots, v_i, \dots, v_i, \dots v_k) + \omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \\ &+\omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k) + \omega(v_1, \dots, v_j, \dots, v_j, \dots v_k) = 0 \end{aligned}$$

Wegen alternierend gilt $\omega(v_1, \dots, v_i, \dots, v_i, \dots, v_k) = \omega(v_1, \dots, v_j, \dots, v_j, \dots, v_k) = 0$. Es folgt $\omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k) = -\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k)$, was zu beweisen war.

3. Ist $\{u_1, \dots, u_n\}$ eine Basis von V (mit $n = \dim(V)$), so ist die sogenannte *duale Basis* $\{u_1^*, \dots, u_n^*\}$ eine Basis von V^*. Dabei wird jedes $u_i^* \in V^*$ in der Basis $\{u_1, \dots, u_n\}$ wie folgt definiert:

$$u_i^*(u_j) := \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases}.$$

Ist $\theta \in V^*$ eine Linearform, so ist $\theta = \sum_{i=1}^{n} \theta(u_i) \cdot u_i^*$ die Darstellung von θ in der Basis $\{u_1^*, \dots, u_n^*\}$ von V^*. Im Spezialfall $V = \mathbb{R}^n$ wird die zur kanonischen Basis $\{e_1, \dots, e_n\}$ assoziierte duale Basis $\{e_1^*, \dots, e_n^*\}$ oft mit $dx_i = e_i^*$ bezeichnet. Insbesondere gilt $\theta = \sum_{i=1}^{n} \theta(e_i) dx_i$ für jedes $\theta \in (\mathbb{R}^n)^*$. Beachte auch: $e_i^*(v) = v_i$ für alle $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ und dx_i ist wirklich das Differential der linearen Abbildung $\mathbb{R}^n \longrightarrow \mathbb{R}$, $x = (x_1, \dots, x_n) \longmapsto x_i$.

Beispiele 5.28

1. Die Determinante auf $\mathbb{R}^n$ ist eine alternierende n-Multilinearform auf $\mathbb{R}^n$. Ursprünglich ist die Determinante auf dem Raum $M_{n\times n}(\mathbb{R})$ definiert. Eine quadratische Matrix kann aber als Tupel von n Spaltenvektoren aufgefasst werden.

2. Sei $f : U \longrightarrow \mathbb{R}$ eine auf einer offenen Teilmenge U des $\mathbb{R}^n$ definierte differenzierbare Funktion. Dann ist in jedem Punkt $x \in U$ das Differential $d_x f : \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Linearform auf dem $\mathbb{R}^n$. Wegen $\frac{\partial f}{\partial x_i}(x) = d_x f(e_i)$ gilt insbesondere $d_x f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) dx_i$.

Definition 5.29 *Für $k, \ell \in \mathbb{N}$ seien $\alpha \in \Lambda^k V^*$ und $\beta \in \Lambda^\ell V^*$ alternierende Multilinearformen. Das* äußere Produkt *von α mit β ist die alternierende Multilinearform $\alpha \wedge \beta \in \Lambda^{k+\ell} V^*$, welche definiert ist durch*

$$(\alpha \wedge \beta)(v_1, \ldots, v_{k+\ell}) := \sum_{\substack{\sigma \in \mathcal{S}_{k+\ell} \\ \sigma(1)<\cdots<\sigma(k) \\ \sigma(k+1)<\cdots<\sigma(k+\ell)}} \varepsilon(\sigma)\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)})$$

für alle $v_1, \ldots, v_{k+\ell} \in V$. Hierbei bezeichnet $\mathcal{S}_{k+\ell}$ die Gruppe aller bijektiven Abbildungen $\{1, \ldots, k+\ell\} \longrightarrow \{1, \ldots, k+\ell\}$ und $\varepsilon(\sigma) \in \{-1, 1\}$ ist die sogenannte Signatur eines $\sigma \in \mathcal{S}_{k+\ell}$.

Beachte: für $k = 0$ ist ein $\alpha \in \Lambda^0 V^* = \mathbb{R}$ eine reelle Zahl, ggf. ist $\alpha \wedge \beta = \alpha \cdot \beta$ die Multiplikation mit β.

Proposition 5.30 *Sei $k, \ell \in \mathbb{N} \setminus \{0\}$, $\alpha \in \Lambda^k V^*$, $\beta \in \Lambda^\ell V^*$ und $\theta_1, \ldots, \theta_k \in V^*$. Dann gilt:*

i) Das äußere Produkt ist assoziativ: für jedes $\gamma \in \Lambda^m V^*$ gilt

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

ii) Das äußere Produkt ist antikommutativ:

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha.$$

iii) Für alle $v_1, \ldots, v_k \in V$ gilt $(\theta_1 \wedge \cdots \wedge \theta_k)(v_1, \ldots, v_k) = \det((\theta_i(v_j))_{1\le i,j\le k})$.

Beispiele 5.31

1. Für zwei Linearformen $\theta_1, \theta_2 \in V^*$ gilt insbesondere

$$(\theta_1 \wedge \theta_2)(v_1, v_2) = \begin{vmatrix} \theta_1(v_1) & \theta_1(v_2) \\ \theta_2(v_1) & \theta_2(v_2) \end{vmatrix} = \theta_1(v_1)\theta_2(v_2) - \theta_2(v_1)\theta_1(v_2).$$

Daraus folgt $\theta \wedge \theta = 0$ für alle $\theta \in V^*$, insbesondere $\theta_1 \wedge \theta_2 = -\theta_2 \wedge \theta_1$.

2. Sei $V := \mathbb{R}^n$, $\theta_i = e_i^*$ für $1 \leq i \leq 3$ und $v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ und

$v_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \\ 5 \\ 2 \end{pmatrix}$. Dann gilt $(e_1^* \wedge e_2^* \wedge e_3^*)(v_1, v_2, v_3) = \det((\theta_i(v_j))_{1 \leq i,j \leq 3})$, d.h.,

$$(e_1^* \wedge e_2^* \wedge e_3^*)(v_1, v_2, v_3) = \begin{vmatrix} 1 & -1 & 3 \\ 2 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} - 2 \begin{vmatrix} -1 & 3 \\ 1 & 1 \end{vmatrix} = 10.$$

Korollar 5.32 *Sei $\{u_1, \ldots, u_n\}$ eine Basis von V. Dann ist*

$$\{u_{i_1}^* \wedge \cdots \wedge u_{i_k}^* \,|\, 1 \leq i_1 < \cdots < i_k \leq n\}$$

eine Basis von $\Lambda^k V^*$ für jedes $k \in \{1, \ldots, n\}$. Außerdem lässt sich jedes $\alpha \in \Lambda^k V^*$ wie folgt darstellen:

$$\alpha = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \alpha(u_{i_1}, \ldots, u_{i_k}) u_{i_1}^* \wedge \cdots \wedge u_{i_k}^*.$$

Insbesondere ist $\Lambda^k V^*$ endlichdimensional mit $\dim(\Lambda^k V^*) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$ für alle $k \in \{0, 1, \ldots, n\}$ (und $\Lambda^k V^* = \{0\}$ für alle $k > n$).

Beispiel 5.33 Für $V = \mathbb{R}^n$ gilt nach Proposition 5.30:

$$\det = e_1^* \wedge \cdots \wedge e_n^*.$$

Denn: $\det(v_1, \ldots, v_n) = \begin{vmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \ldots & v_{nn} \end{vmatrix} = \det((e_i^*(v_j))_{1 \leq i,j \leq n}) = (e_1^* \wedge \cdots \wedge e_n^*)(v_1, \ldots, v_n)$

für alle $v_1, \ldots, v_n \in \mathbb{R}^n$, wobei $v_j = \sum_{i=1}^{n} v_{ij} e_i$ (mit $v_{ij} \in \mathbb{R}$).

5.2.2 Differentialformen auf $\mathbb{R}^n$

Definition 5.34 *Sei $k \in \mathbb{N}$ und $U \subset \mathbb{R}^n$ eine offene Teilmenge. Eine* Differentialform *vom Grad k (kurz: k-Form) auf U ist eine Abbildung $\omega : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$. Eine k-Form ω heißt differenzierbar bzw. C^ℓ (mit $\ell \in \mathbb{N} \cup \{\infty\}$), wenn ω als Abbildung von U nach $\Lambda^k(\mathbb{R}^n)^* \cong \mathbb{R}^k$ differenzierbar bzw. C^ℓ ist.*

Bemerkung 5.35 Insbesondere lässt sich jede k-Form $\omega : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ auf eindeutige Weise wie folgt darstellen:

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \omega(e_{i_1}, \ldots, e_{i_k}) e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*,$$

wobei $\omega(e_{i_1}, \ldots, e_{i_k}) : U \longrightarrow \mathbb{R}$, $x \mapsto \omega_x(e_{i_1}, \ldots, e_{i_k})$, eine Funktion ist, für alle $1 \leq i_1 < \cdots < i_k \leq n$. Ggf. ist ω genau dann differenzierbar bzw. C^ℓ, wenn alle Funktionen $\omega(e_{i_1}, \ldots, e_{i_k})$ differenzierbar bzw. C^ℓ sind. Beachte, dass für $k = 0$ eine 0-Form nichts anderes als eine reellwertige Funktion ist; sie ist dann differenzierbar bzw. C^ℓ als 0-Form g.d.w. sie als Funktion differenzierbar bzw. C^ℓ ist.

Definition 5.36 *Für $k, \ell \in \mathbb{N}$ seien $\alpha : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ und $\beta : U \longrightarrow \Lambda^\ell(\mathbb{R}^n)^*$. Das* äußere Produkt *von α mit β wird definiert als*

$$\begin{array}{rcl}\alpha \wedge \beta : U & \longrightarrow & \Lambda^{k+\ell}(\mathbb{R}^n)^* \\ x & \longmapsto & \alpha_x \wedge \beta_x.\end{array}$$

Für $k = 0$ ist $\alpha : U \longrightarrow \mathbb{R}$ eine Funktion und $\alpha \wedge \beta = \alpha \cdot \beta$ ist die Multiplikation mit der Funktion α.

Definition 5.37 *Für $k \in \mathbb{N}$ sei $\omega : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ eine differenzierbare k-Form auf $U \subset \mathbb{R}^n$. Das* äußere Differential *von ω ist die $k+1$-Form $d\omega : U \longrightarrow \Lambda^{k+1}(\mathbb{R}^n)^*$,*

$$d\omega := \sum_{\substack{1 \leq i \leq n \\ 1 \leq i_1 < \cdots < i_k \leq n}} \frac{\partial(\omega(e_{i_1}, \ldots, e_{i_k}))}{\partial x_i} dx_i \wedge e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*,$$

wobei $\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \omega(e_{i_1}, \ldots, e_{i_k}) e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*$ die punktweise Darstellung von ω in der Basis $\{e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* \,|\, 1 \leq i_1 < \cdots < i_k \leq n\}$ von $\Lambda^k(\mathbb{R}^n)^*$ ist.

Proposition 5.38 *Das äußere Differential erfüllt folgende Eigenschaften:*

i) Für jede differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ ist $df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$ das übliche Differential von f.

ii) Es gilt $d^2\omega := d(d\omega) = 0$ für alle differenzierbaren k-Formen ω.

iii) (Produktregel) Es gilt $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta$ für alle differenzierbaren k-Formen α und ℓ-Formen β.

Bemerke, dass die Formel aus Proposition 5.38.*iii)* $d(f\beta) = df \wedge \beta + fd\beta$ lautet, falls $\alpha = f : U \longrightarrow \mathbb{R}$ eine differenzierbare Funktion ist.

Beispiele 5.39

1. Sei $\omega : \mathbb{R}^2 \longrightarrow \Lambda^1(\mathbb{R}^2)^*$, $\omega_{(x,y)} := xdx + ydy$ (kurz: $\omega = xdx + ydy$). Da $(x, y) \longmapsto x$ und $(x, y) \longmapsto y$ C^∞-Funktionen auf $\mathbb{R}^2$ sind, ist ω C^∞ mit
$$\begin{array}{rcl} d\omega & = & d(xdx) + d(ydy) \\ & = & dx \wedge dx + dy \wedge dy \\ & = & 0. \end{array}$$

2. Sei $\omega := -ydx + xdy$ auf $\mathbb{R}^2$, also $\omega : \mathbb{R}^2 \longrightarrow \Lambda^1(\mathbb{R}^2)^* = (\mathbb{R}^2)^*$, $\omega_{(x,y)} := -ydx + xdy$. Dann gilt
$$\begin{array}{rcl} d\omega & = & d(-ydx) + d(xdy) \\ & = & -dy \wedge dx + dx \wedge dy \\ & = & 2dx \wedge dy. \end{array}$$

3. Sei $\omega := -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$ auf $\mathbb{R}^2 \setminus \{0\}$. Dann gilt
$$\begin{array}{rcl} d\omega & = & -\frac{\partial}{\partial y}(\frac{y}{x^2+y^2})dy \wedge dx + \frac{\partial}{\partial x}(\frac{x}{x^2+y^2})dx \wedge dy \\ & = & \frac{y^2-x^2}{(x^2+y^2)^2}dy \wedge dx + \frac{y^2-x^2}{(x^2+y^2)}dx \wedge dy \\ & = & 0. \end{array}$$

4. Allgemein sei $\omega = \sum_{i=1}^{n} \omega_i dx_i$ eine beliebige 1-Form auf U, also $\omega : U \longrightarrow \Lambda^1(\mathbb{R}^n)^*$ mit $\omega_i : U \longrightarrow \mathbb{R}$ für alle $1 \le i \le n$, wobei $U \subset \mathbb{R}^n$ offen ist. Ist ω differenzierbar, so sind $\omega_1, \ldots, \omega_n$ differenzierbar und es gilt

$$\begin{aligned} d\omega &= \sum_{i=1}^{n} d(\omega_i) \wedge dx_i \\ &= \sum_{i,j=1}^{n} \frac{\partial \omega_i}{\partial x_j} dx_j \wedge dx_i \quad \text{mit } dx_j \wedge dx_i = -dx_i \wedge dx_j \\ &= \sum_{1 \le i < j \le n} \left(\frac{\partial \omega_j}{\partial x_i} - \frac{\partial \omega_i}{\partial x_j}\right) dx_i \wedge dx_j. \end{aligned}$$

Insbesondere ist $d\omega = 0$ g.d.w. $\frac{\partial \omega_j}{\partial x_i} = \frac{\partial \omega_i}{\partial x_i}$ für alle $1 \le i, j \le n$ (für $i = j$ ist die Bedingung leer). Man vergleiche diese Identitäten mit dem Satz 2.104.

Definition 5.40 *Für $k \in \mathbb{N}$ sei $\omega : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ eine C^∞ k-Form auf U.*

i) Die k-Form ω heißt genau dann geschlossen, *wenn $d\omega = 0$ auf U gilt.*

ii) Im Fall $k \ge 1$ heißt ω genau dann exakt, *wenn eine C^∞ $k-1$-Form α auf U existiert mit $d\alpha = \omega$ auf U.*

Bemerkungen 5.41

1. Nach Proposition 5.38 ist jede exakte Form geschlossen: ist $\omega = d\alpha$, so gilt $d\omega = d(d\alpha) = 0$.

2. Für eine C^∞ 1-Form ω auf U der Form $\omega = \sum_{i=1}^{n} \omega_i dx_i$ (mit $\omega_i \in C^\infty(U, \mathbb{R})$) ist nach Beispiel 5.39.4 ω genau dann geschlossen, wenn $\frac{\partial \omega_i}{\partial x_j} = \frac{\partial \omega_j}{\partial x_i}$ für alle $1 \le i, j \le n$. Desweiteren ist ω genau dann exakt, wenn eine C^∞ Funktion $P : U \longrightarrow \mathbb{R}$ so existiert, dass $dP = \omega$ gilt, d.h., wenn

$$\sum_{i=1}^{n} \frac{\partial P}{\partial x_i} dx_i = \sum_{i=1}^{n} \omega_i dx_i$$

gilt, d.h., wenn $(\omega_1, \ldots, \omega_n) = \nabla P$ gilt: somit ist ω genau dann exakt, wenn das *Vektorfeld* $U \longrightarrow \mathbb{R}^n$, $x \longmapsto (\omega_1(x), \ldots, \omega_n(x))$, ein *Potential* besitzt.

3. Insbesondere ist nicht jede geschlossene Form exakt. Betrachte z.B. $\omega = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$ auf $\mathbb{R}^2 \setminus \{0\}$. Nach Beispiel 5.39.3 gilt $d\omega = 0$, d.h., ω ist geschlossen auf $\mathbb{R}^2 \setminus \{0\}$. Nach Beispiel 2.103.3 hat das Vektorfeld $(x, y) \longmapsto \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ kein Potential auf $\mathbb{R}^2 \setminus \{0\}$, d.h., ω ist nicht exakt.

Satz 5.42 (Poincaré-Lemma) *Sei $\omega : U \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ eine auf einer offenen Teilmenge U des $\mathbb{R}^n$ definierte C^∞ k-Form, wobei $k \in \mathbb{N}$, $k \ge 1$. Angenommen, U sei* sternförmig *bzgl. eines seiner Punkte. Dann gilt: ist ω geschlossen auf U, so ist ω exakt auf U.*

Definition 5.43 *Sei $\varphi : U \longrightarrow V$ eine C^∞-Abbildung zwischen offenen Teilmengen des $\mathbb{R}^n$. Sei $\omega : V \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ eine C^∞ k-Form auf V, wobei $k \in \mathbb{N}$. Dann heißt*

$$\begin{array}{rcl} \varphi^*\omega : U & \longrightarrow & \Lambda^k(\mathbb{R}^n)^* \\ x & \longmapsto & ((\varphi^*\omega)_x : (X_1, \ldots, X_k) \longmapsto \omega_{\varphi(x)}(d_x\varphi(X_1), \ldots, d_x\varphi(X_k))) \end{array}$$

die durch φ zurückgezogene *k-Form auf U. Für $k = 0$ ist $\varphi^*\omega = \omega \circ \varphi$ die Verknüpfung mit φ.*

Proposition 5.44 *Sei $\varphi : U \longrightarrow V$ eine C^∞-Abbildung zwischen offenen Teilmengen des $\mathbb{R}^n$. Dann gilt:*

i) Für jede C^∞ k-Form ω auf V ist $\varphi^*\omega$ eine C^∞ k-Form auf U. Außerdem ist die Abbildung $\omega \longmapsto \varphi^*\omega$ linear.

ii) Es gilt $\varphi^*(d\omega) = d(\varphi^*\omega)$ für jede C^∞ k-Form ω auf V.

ii') $\varphi^*(\alpha \wedge \beta) = (\varphi^*\alpha) \wedge (\varphi^*\beta)$ für alle Differentialformen α und β auf U.

iii) Ist $\psi : \tilde{U} \longrightarrow U$ eine weitere C^∞-Abbildung, so gilt $(\varphi \circ \psi)^*\omega = \psi^*(\varphi^*\omega)$.

5.2.3 Differentialformen auf Untermannigfaltigkeiten

Für eine Untermannigfaltigkeit M des $\mathbb{R}^n$ und ein $x \in M$ ist $T_xM \subset \mathbb{R}^n$ ein Untervektorraum des $\mathbb{R}^n$. Für jeden Untervektorraum E von $\mathbb{R}^n$ gilt aber $E \oplus E^\perp = \mathbb{R}^n$, wobei $E^\perp := \{v \in \mathbb{R}^n \,|\, \langle v, X\rangle = 0 \text{ für alle } X \in E\}$ das orthogonale Komplement zu E im $\mathbb{R}^n$ ist. Insbesondere ist die Abbildung

$$\begin{array}{rcl} \mathrm{pr}_x : \mathbb{R}^n & \longrightarrow & T_xM \\ X & \longmapsto & X_T, \end{array}$$

wobei $X = X_T + X_\perp$ mit $X_T \in T_xM$, $X_\perp \in T_x^\perp M$ die Zerlegung von X bzgl. $\mathbb{R}^n = T_xM \oplus T_xM^\perp$ ist, wohldefiniert und linear. Die Abbildung pr_x heißt *Orthogonalprojektion* auf den Untervektorraum T_xM. Diese Abbildung ist surjektiv. Dies impliziert, dass die Abbildung

$$\begin{array}{rcl} \mathrm{pr}_x^* : T_x^*M & \longrightarrow & (\mathbb{R}^n)^* \\ \theta & \longmapsto & \theta \circ \mathrm{pr}_x \end{array}$$

linear und *injektiv* ist. Hierbei bezeichnet $T_x^*M := (T_xM)^*$ den Dualraum von T_xM. Insbesondere ist $\mathrm{pr}_x^* : T_x^*M \longrightarrow \mathrm{pr}_x^*(T_x^*M) \subset (\mathbb{R}^n)^*$ ein Isomorphismus.
Fazit: Die Abbildung pr_x^* erlaubt es, den Vektorraum T_x^*M mit dem Untervektorraum $\mathrm{pr}_x^*(T_x^*M) \subset (\mathbb{R}^n)^*$ zu *identifizieren*. Von nun an *sehen* wir T_x^*M als Untervektorraum von $(\mathbb{R}^n)^*$.

Definition 5.45 *Sei M^m eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und $k \in \mathbb{N}$. Eine k-*Form *(oder* Differentialform vom Grad *k) auf M^m ist eine Abbildung $\omega : M^m \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ mit $\omega_x \in \Lambda^kT_x^*M$ für alle $x \in M^m$, wobei $\Lambda^kT_x^*M := \Lambda^k(T_xM)^*$ als Untervektorraum von $\Lambda^k(\mathbb{R}^n)^*$ aufgefasst wird (siehe oben). Entsprechend heißt ω genau dann differenzierbar bzw. C^ℓ (mit $\ell \in \mathbb{N} \cup \{\infty\}$), wenn $\omega : M^m \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ als Abbildung zwischen Untermannigfaltigkeiten differenzierbar bzw. C^ℓ ist.*

Beispiel 5.46 Ist $M^m \subset U$, wobei $U \subset \mathbb{R}^n$ offen ist, so kann jede k-Form ω auf U auf M^m folgendermaßen eingeschränkt werden: für jedes $x \in M$ definiert $T_xM \times \cdots \times T_xM \longrightarrow \mathbb{R}$, $(X_1, \ldots, X_k) \longmapsto \omega_x(X_1, \ldots, X_k)$, eine alternierende k-Multilinearform, also ein Element aus $\Lambda^k T_x^* M$. Diese eingeschränkte k-Form bezeichnen wir weiterhin mit $\omega_{|_M}$. Es ist relativ elementar, nachzuweisen, dass $\omega_{|_M}$ differenzierbar bzw. C^ℓ auf M ist, sobald ω differenzierbar bzw. C^ℓ auf U ist.

Definition 5.47 *Sei M^m eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ und α bzw. β eine k-Form bzw. eine ℓ-Form auf M^m. Das* äußere Produkt *von α mit β ist die $k+\ell$-Form $\alpha \wedge \beta : M^m \longrightarrow \Lambda^{k+\ell}(\mathbb{R}^n)^*$, $x \longmapsto \alpha_x \wedge \beta_x$.*

Das äußere Produkt auf Differentialformen erfüllt die gleichen Eigenschaften wie das äußere Produkt auf alternierenden Multilinearformen (u.a. Assoziativität, Anti-Kommutativität).

Definition 5.48 *Sei ω eine k-Form auf einer Untermannigfaltigkeit des $\mathbb{R}^q$ und $\varphi : M^m \longrightarrow N^p$ eine differenzierbare Abbildung. Man definiert*

$$\begin{array}{rcl} \varphi^*\omega : M^m & \longrightarrow & \Lambda^k(\mathbb{R}^n)^* \\ x & \longmapsto & ((\varphi^*\omega)_x : (X_1, \ldots, X_k) \longmapsto \omega_{\varphi(x)}(d_x\varphi(X_1), \ldots, d_x\varphi(X_k)) \end{array}$$

für alle $X_1, \ldots, X_k \in T_xM$. Insbesondere ist $\varphi^*\omega$ eine k-Form auf M^m.

Genau wie für Differentialformen auf $\mathbb{R}^n$ gilt $(\varphi \circ \psi)^*\omega = \psi^*(\varphi^*\omega)$ für alle differenzierbaren Abbildungen $Q^q \xrightarrow{\psi} M^m$ und $M^m \xrightarrow{\varphi} N^p$.

Definition 5.49 *Sei ω eine differenzierbare k-Form auf einer Untermannigfaltigkeit M^m des $\mathbb{R}^n$. Das* äußere Differential *von ω ist die $k+1$-Form $d\omega$ auf M^m, die wie folgt definiert ist: sei $U \subset \mathbb{R}^n$ offen mit $U \cap M \neq \varnothing$ und so, dass ein C^∞-Diffeomorphismus $\varphi : U \longrightarrow \varphi(U)$ (mit $\varphi(U) \subset \mathbb{R}^n$ offen) existiert mit $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Dann ist*

$$d\omega_{|_{U\cap M}} := \varphi^* d((\varphi^{-1})^* \omega_{|_{\mathbb{R}^m \times \{0_{n-m}\}}}).$$

Es kann bewiesen werden, dass $d\omega$ wohldefiniert ist: sind $U \xrightarrow{\varphi} \varphi(U)$ und $V \xrightarrow{\psi} \psi(V)$ zwei solche Diffeomorphismen, so gilt $\varphi^* d((\varphi^{-1})^*\omega) = \psi^* d((\psi^{-1})^*\omega)$ auf $U \cap V \cap M$.

Proposition 5.50 *Sei ω eine differenzierbare k-Form auf einer Untermannigfaltigkeit M^m des $\mathbb{R}^n$. Dann gilt:*

i) $\varphi^*(d\omega) = d(\varphi^*\omega)$ für alle C^∞-Abbildungen $\varphi : N \longrightarrow M$, wobei N^p eine Untermannigfaltigkeit des $\mathbb{R}^q$ ist.

ii) $d(\omega \wedge \omega') = d\omega \wedge \omega' + (-1)^k \omega \wedge d\omega'$ für alle differenzierbaren ℓ-Formen ω' auf M^m.

iii) $\varphi^*(\omega \wedge \omega') = \varphi^*\omega \wedge \varphi^*\omega'$ für alle differenzierbaren ℓ-Formen ω' auf M und C^∞ Abbildungen $\varphi : N \longrightarrow M$.

Bevor wir das Integral von Differentialformen auf Untermannigfaltigkeiten einführen, besprechen wir, wie man mit Differentialformen *lokal* rechnet.
Sei M^m eine Untermannigfaltigkeit des $\mathbb{R}^n$ und $U \xrightarrow{\varphi} \varphi(U)$ ein C^∞-Diffeomorphismus zwischen offenen Teilmengen des $\mathbb{R}^n$, wobei $U \cap M \neq \varnothing$ und $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$. Eine solche Abbildung $U \xrightarrow{\varphi} \varphi(U)$ nennt man *lokale Karte* von M^m. Beachte, dass $U \cap M$ als offene Teilmenge einer m-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^n$ ebenfalls eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ ist, die zur offenen Teilmenge $\varphi(U) \cap (\mathbb{R}^n \times \{0_{n-m}\})$ von $\mathbb{R}^m$ diffeomorph ist (durch $\varphi : U \cap M \longrightarrow \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$).
Nun ziehen wir die "kanonischen" Vektorfelder $x \longmapsto e_i$ und 1-Formen $x \longmapsto dx_i$ auf $\varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$ zurück auf $U \cap M$ durch φ: wir setzen, für $1 \leq i \leq m$ und $x \in U\ capM$,

$$\begin{aligned}\left(\frac{\partial}{\partial x_i}\right)_x &:= (d_x\varphi)^{-1}(e_i) \in T_xM \\ (dx_i)_x &:= (\varphi^* dx_i)_x := dx_i \circ d_x\varphi \in T_x^*M.\end{aligned}$$

Da $\varphi : U \cap M \longrightarrow \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$ ein Diffeomorphismus ist, bilden die Vektoren $\left(\frac{\partial}{\partial x_1}\right)_x, \ldots, \left(\frac{\partial}{\partial x_m}\right)_x$ eine Basis des Tangentialraums T_xM und die Linearformen $(dx_1)_x, \ldots, (dx_m)_x$ bilden eine Basis von T_x^*M. Außerdem hängen $\frac{\partial}{\partial x_i}$ und dx_i C^∞ von x ab, somit sind $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}$ (bzw. $dx_1, \ldots, dx_m$) C^∞ Tangentialvektorfelder (bzw. C^∞ 1-Formen) auf $U \cap M$.
Bzgl. dieser Vektorfelder und Differentialformen (die nur auf $U \cap M$ definiert sind und von φ abhängen) können jetzt alle Vektorfelder und Differentialformen auf $U \cap M$ dargestellt werden und dabei alle Rechenregeln für Differentialformen auf $\mathbb{R}^m$ verwendet werden. Z.B. lässt sich jede k-Form ω auf $U \cap M$ wie folgt darstellen:

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq m} \omega\left(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}\right) dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

wobei $\omega\left(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}\right) : U \cap M \longrightarrow \mathbb{R}$ Funktionen sind. Im Spezialfall $\omega = df$ für eine differenzierbare Funktion $f : U \cap M \longrightarrow \mathbb{R}$ schreiben wir $df\left(\frac{\partial}{\partial x_i}\right) =: \frac{\partial f}{\partial x_i}$, somit gilt $df = \sum_{i=1}^m \frac{\partial f}{\partial x_i} dx_i$ auf $U \cap M$.
Auf $U \cap M$ berechnet man das äußere Produkt und das äußere Differential wie im $\mathbb{R}^m$: ist z.B. ω wie oben, so ist

$$d\omega = \sum_{i=1}^m \sum_{1 \leq i_1 < \cdots < i_k \leq m} \frac{\partial\left(\omega\left(\frac{\partial}{\partial x_{i_1}}, \ldots, \frac{\partial}{\partial x_{i_k}}\right)\right)}{\partial x_i} dx_i \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

5.3 Integration von Differentialformen auf Untermannigfaltigkeiten

5.3.1 Orientierbarkeit und Orientierung einer Untermannigfaltigkeit

Definition 5.51 *Sei $M^m \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.*

i) Die Untermannigfaltigkeit M^m heißt genau dann orientierbar, *wenn eine Familie $(U_i, \varphi_i)_{i\in I}$ von lokalen Karten von M so existiert, dass $M \subset \bigcup_{i\in I} U_i$ und $\det(d_x(\varphi_i \circ \varphi_j^{-1})) > 0$ für alle $x \in \varphi_j(U_i \cap U_j) \cap (\mathbb{R}^m \times \{0_{n-m}\})$ und alle $i, j \in I$ mit $U_i \cap U_j \cap M \neq \varnothing$.*

ii) Ist M^m orientierbar, so ist eine Orientierung *von M^m eine "maximale" Familie von lokalen Karten $(U_i, \varphi_i)_{i\in I}$ mit den obigen Eigenschaften.*

Diese lokalen Karten $(U_i, \varphi_i)_{i\in I}$ heißen ggf. orientierte lokale Karten *von M^m.*

Hierbei heißt die Familie *maximal*, wenn jede lokale Karte (U, φ) von M^m, welche $\det(d_x(\varphi \circ \varphi_i^{-1})) > 0$ für alle $x \in \varphi_i(U \cap U_i)$ und alle $i \in I$ mit $U \cap U_i \neq \varnothing$ erfüllt, bereits in der Familie liegt.

Definition 5.52 *Sei $M^m \subset \mathbb{R}^n$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Eine* Volumenform *auf M^m ist eine nirgendsverschwindende C^∞ m-Form auf M^m.*

Proposition 5.53 *Eine m-dimensionale Untermannigfaltigkeit M^m des $\mathbb{R}^n$ ist genau dann orientierbar, wenn sie eine Volumenform besitzt. Ggf. legt jede Volumenform eine Orientierung von M^m fest.*

Ist ω eine Volumenform auf M^m, so wird eine Orientierung auf M^m folgendermaßen festgelegt: sei (U, φ) eine lokale Karte von M^m, seien $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}$ die dazu gehörigen Vektorfelder auf $U \cap M$; dann gehört (U, φ) zur maximalen Familie $(U_i, \varphi_i)_{i\in I}$ aus Definition 5.51 g.d.w.

$$\omega\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right) > 0$$

auf $U \cap M$ gilt.

5.3.2 Integral einer Differentialform

Definition 5.54 *Sei $\omega : M^m \longrightarrow \Lambda^k(\mathbb{R}^n)^*$ eine auf einer Untermannigfaltigkeit M^m vom $\mathbb{R}^n$ definierte stetige k-Form. Der* Träger *von ω ist*

$$\mathrm{supp}(\omega) := \{x \in M^m \,|\, \omega_x \neq 0\} \subset M^m.$$

Per Definition ist $\mathrm{supp}(\omega)$ abgeschlossen in M^m (wegen ω stetig) und es gilt $\omega_x = 0$ für alle $x \in M^m \setminus \mathrm{supp}(\omega)$.

Definition 5.55 *Sei ω eine auf einer orientierten Untermannigfaltigkeit M^m vom $\mathbb{R}^n$ definierte stetige m-Form. Angenommen, $\mathrm{supp}(\omega)$ sei* kompakt *und in einem* Kartenbereich *U enthalten, wobei $U \xrightarrow{\varphi} \varphi(U)$ eine* orientierte *lokale Karte von M^m ist. Dann ist das Integral von ω auf M^m wie folgt definiert:*

$$\int_M \omega := \int_{\varphi(U)} (\varphi^{-1})^*\omega(e_1, \ldots, e_m) dx_1 \ldots dx_m \in \mathbb{R},$$

wobei $(\varphi^{-1})^*\omega = (\varphi^{-1})^*\omega(e_1, \ldots, e_m) dx_1 \wedge \cdots \wedge dx_m$ mit $(\varphi^{-1})^*\omega(e_1, \ldots, e_m) : \varphi(U) \longrightarrow \mathbb{R}$ stetig (mit kompaktem Träger in $\varphi(U)$).

Man soll natürlich nachweisen, dass diese Definition nicht von der Wahl der orientierten lokalen Karte von M^m abhängt; dies liegt an der Transformationsformel (Satz 3.70). Beachte auch, dass $\int_M \omega$ wegen $\mathrm{supp}(\omega)$ kompakt wohldefiniert ist. Die Definition 5.55 betrifft nur die m-Formen mit Träger in einem Kartenbereich. Für den allgemeinen Fall, wo der Träger von ω immer noch kompakt aber nicht mehr in einem Kartenbereich enthalten ist, brauchen wir folgendes Verfahren:

Definition 5.56 *Sei $M^m \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $(U_i)_{i\in I}$ eine Familie von offenen Teilmengen von M^m mit $\bigcup\limits_{i\in I} U_i = M^m$. Eine* Teilung der Eins *(oder Zerlegung der Eins) zur Familie $(U_i)_{i\in I}$ ist eine Familie $(\chi_i)_{i\in I}$ (mit derselben Indexmenge I) von C^∞ Funktion $\chi_i : M \longrightarrow [0,1]$ mit:*

i) Für alle $i \in I$ ist $\mathrm{supp}(\chi_i)$ kompakt und enthalten in U_i,

ii) Für alle $x \in M$ existiert eine offene Umgebung U von x in M s.d. $\{i \in I \,|\, \mathrm{supp}(\chi_i) \cap U \neq \varnothing\}$ endlich ist,

iii) $\sum\limits_{i\in I} \chi_i = 1$, d.h., $\sum\limits_{i\in I} \chi_i(x) = 1$ für alle $x \in M$.

Beachte, dass die Summe $\sum\limits_{i\in I} \chi_i(x)$ wegen *ii)* endlich (es gibt nur endlich viele i's aus I mit $\chi_i(x) \neq 0$) und somit wohldefiniert ist. Wir akzeptieren folgendes Ergebnis ohne Beweis.

Proposition 5.57 *Ist $(U_i)_{i=1,\ldots,p}$ eine endliche Überdeckung durch offene Teilmengen einer kompakten Untermannigfaltigkeit, so existiert eine Teilung der Eins zur Familie $(U_i)_{i=1,\ldots,p}$.*

Nun sind wir in der Lage, das Integral einer m-Form mit kompaktem Träger auf einer Untermannigfaltigkeit M^m des $\mathbb{R}^n$ zu definieren.

Definition 5.58 *Sei ω eine auf einer orientierten Untermannigfaltigkeit M^m vom $\mathbb{R}^n$ definierte stetige m-Form. Angenommen, $\mathrm{supp}(\omega)$ sei kompakt. Sei $(U_i)_{i=1,\ldots,p}$ eine endliche Familie von Kartenbereichen mit $\bigcup\limits_{i=1}^{p} U_i \supset \mathrm{supp}(\omega)$, wobei $U_i \xrightarrow{\varphi_i} \varphi_i(U_i)$ orientierte lokale Karten von M^m sind. Sei $(\chi_i)_{i=1,\ldots,p}$ eine zugeordnete Teilung der Eins. Dann ist das Integral von ω auf M^m definiert durch*

$$\int_M \omega := \sum_{i=1}^{p} \int_M \chi_i \cdot \omega \in \mathbb{R},$$

wobei $\int_M \chi_i \cdot \omega$ wie in Definition 5.55 *erklärt ist.*

Beachte, dass $\mathrm{supp}(\chi_i \cdot \omega) \subset \mathrm{supp}(\chi_i) \subset U_i$ und ist kompakt, deshalb kann die Definition 5.55 für $\chi_i{\cdot}\omega$ verwendet werden. Dazu sollte man auch überprüfen (Beweis davon wird weggelassen), dass diese Definition unabhängig von der Teilung der Eins $(\chi_i)_{i=1,\ldots,p}$ ist und dass eine endliche Familie von Kartengebieten existiert, welche die kompakte Teilmenge $\mathrm{supp}(\omega)$ überdeckt.

Beispiel 5.59 Sei M^m eine Kurve ohne Selbstdurchschnitte im $\mathbb{R}^n$ wie im Beispiel 5.3.5, sei $c :]a, b[\longrightarrow M^m$ eine Parametrisierung von M^m, wobei $]a, b[\subset \mathbb{R}$ ein offenes Intervall ist. Sei ω eine stetige 1-Form auf M^m. Dann lässt sich $\int_M \omega$ als Kurvenintegral umschreiben: es gilt

$$\int_M \omega = \int_a^b \omega_{c(t)}(\dot{c}(t))dt.$$

Man vergleiche diese Identität mit der Definition des Kurvenintegrals für Vektorfelder (Definition 2.100). Bemerke auch, dass das Integral $\int_M \omega$ das Vorzeichen wechselt, wenn die Orientierung (d.h., den Durchlaufsinn) von M wechselt.

Bevor wir den Satz von Stokes und seine Korollare vorstellen, brauchen wir den Begriff von Untermannigfaltigkeiten mit Rand und von induzierter Orientierung auf diesem Rand.

5.3.3 Untermannigfaltigkeiten mit Rand

Definition 5.60 *Eine m-dimensionale* Untermannigfaltigkeit mit Rand *ist eine Teilmenge M^m von $\mathbb{R}^n$ so, dass jedes $x \in M^m$ eine offene Umgebung U in $\mathbb{R}^n$ besitzt, auf welcher ein Diffeomorphismus $\varphi : U \longrightarrow \varphi(U)$ auf eine offene Teilmenge $\varphi(U)$ von $\mathbb{R}^n$ so existiert, dass entweder $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\})$ oder $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m_+ \times \{0_{n-m}\})$ gilt, wobei $\mathbb{R}^m_+ := \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m \,|\, x_1 \geq 0\} \subset \mathbb{R}^m$.*
Die Menge der Punkte $x \in M^m$, welche ausschließlich offene Umgebungen der zweiten Art (mit $\varphi(U) \cap M) = \varphi(U) \cap (\mathbb{R}^m_+ \times \{0_{n-m}\})$) zulassen, wird der Rand *von M^m genannt und mit ∂M bezeichnet.*

Es kann bewiesen werden, dass ∂M eine $m-1$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ ist. Z.B. ist $M^n := \overline{B}_1(0) := \{x \in \mathbb{R}^n, \, |x| \leq 1\} \subset \mathbb{R}^n$ eine n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ mit $\partial M = S^{n-1} = \{x \in \mathbb{R}^n, \, |x| = 1\}$.

Definition 5.60 verallgemeinert Definition 5.1: eine Untermannigfaltigkeit im Sinne von Definition 5.1 ist lediglich eine Untermannigfaltigkeit im Sinne von Definition 5.60 mit Rand $\partial M = \varnothing$.

Definition 5.61 *Sei M^m eine orientierte Untermannigfaltigkeit mit Rand (d.h., eine Familie $(U_i, \varphi_i)_{i \in I}$ von lokalen Karten wie in* Definition 5.60 *existiere mit* $\det(d_x(\varphi_i \circ \varphi_j^{-1})) >$ *0 für alle $x \in \varphi_j(U_i \cap U_j) \cap (\mathbb{R}^m \times \{0\})$ oder $x \in \varphi_j(U_i \cap U_j) \cap (\mathbb{R}^m_+ \times \{0\})$ und alle $i, j \in I$ mit $U_i \cap U_j \cap M \neq \varnothing$). Die* induzierte Orientierung *auf ∂M wird wie folgt definiert: sei α eine Volumenform auf M^m, welche die Orientierung von M^m definiert, dann ist eine Basis $(v_1, \ldots, v_{m-1})$ von $T_x\partial M$ genau dann orientiert, wenn $(\nu_x, v_1, \ldots, v_{m-1})$ eine orientierte Basis von T_xM ist, d.h., wenn $\alpha_x(\nu_x, v_1, \ldots, v_{m-1}) > 0$, wobei ν_x das sogenannte äußere Einheitsnormalenfeld von ∂M in M ist.*

Z.B. ist, für $M^2 \subset \mathbb{R}^2$, die induzierte Orientierung auf ∂M derjenige *Durchlaufsinn* von ∂M, für den M immer *links* von ∂M steht.

5.4 Integralsätze

Satz 5.62 (Stokes) *Sei ω eine C^1 $m-1$-Form mit kompaktem Träger auf einer orientierten m-dimensionalen Untermannigfaltigkeit M^m mit Rand vom $\mathbb{R}^n$. Sei ∂M mit der* *wie in* Definition 5.61 *definierten induzierten Orientierung versehen. Dann gilt*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Der Kern des Beweises dieses Satzes ist die berühmte Formel

$$\int_a^b f'(x)dx = f(b) - f(a)$$

aus der Analysis I, wobei $f : [a, b] \longrightarrow \mathbb{R}$ eine beliebige C^1 Funktion ist.

Es gibt zahlreiche Spezialfälle dieses Satzes. Zuerst gilt $\int_M d\omega = 0$, sobald der Rand ∂M leer ist, d.h., wenn M^m eine Untermannigfaltigkeit im Sinne der Definition 5.1 ist (das Integral irgendeiner Form auf der leeren Menge verschwindet).

Korollar 5.63 (Green) *Sei $\omega = Pdx + Qdy$ eine C^1 1-Form mit kompaktem Träger auf einer Fläche mit Rand $M^2 \subset \mathbb{R}^2$. Sei M^2 mit der kanonischen Orientierung vom $\mathbb{R}^2$ und ∂M mit der induzierten Orientierung versehen. Dann gilt*

$$\int_M \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy = \int_{\partial M} Pdx + Qdy.$$

Dieses Korollar kann – im Fall von Kurven ohne Selbstdurchschnitte – für die Charakterisierung der Vektorfelder mit Potential genutzt werden, siehe Sätze 2.101 und 2.104.

Korollar 5.64 (Gauß'scher Integralsatz) *Sei X ein C^1 Tangentialvektorfeld mit kompaktem Träger auf einer Untermannigfaltigkeit $M^m \subset \mathbb{R}^m$ mit Rand. Sei $\nu : \partial M^m \longrightarrow \mathbb{R}^m$ das zugehörige äußere Einheitsnormalenfeld. Dann gilt*

$$\int_{M^m} \mathrm{div}(X) dx_1 \ldots dx_m = \int_{\partial M^m} \langle X, \nu \rangle d\sigma,$$

wobei ∂M die induzierte Orientierung trägt und $d\sigma$ folgende Volumenform auf ∂M ist: für $x \in \partial M$ ist $(d\sigma)_x = v_1^* \wedge v_2^* \wedge \cdots \wedge v_{m-1}^*$ für jede beliebige orientierte Orthonormalbasis $(v_1, \ldots, v_{m-1})$ von $T_x\partial M$.

Ist z.B. X quellenfrei, so impliziert Korollar 5.64, dass der *Fluss* von X durch den Rand ∂M verschwindet.

Korollar 5.65 (Rotationssatz) *Sei M^2 eine orientierbare 2-dimensionale Untermannigfaltigkeit mit Rand vom $\mathbb{R}^3$, sei $\nu : M^2 \longrightarrow \mathbb{R}^3$ ein stetiges Einheitsnormalenfeld auf M^2. Die Fläche M^2 trage die von ν induzierte Orientierung (für die eine Basis $\{v_1, v_2\}$ von T_xM genau dann orientiert ist, wenn $\{\nu_x, v_1, v_2\}$ eine orientierte Basis vom $\mathbb{R}^3$ ist) und der Rand ∂M trage die von M^2 induzierte Orientierung. Sei $T : \partial M \longrightarrow \mathbb{R}^3$ das Einheitstangential-Vektorfeld längs ∂M, welches die Orientierung von ∂M definiert. Dann gilt, für jedes auf einer offenen Umgebung von M^2 definierte C^1 Vektorfeld X mit kompaktem Träger:*

$$\int_{M^2} \langle \mathrm{rot}(X), \nu \rangle d\sigma = \int_{\partial M^2} \langle X, T \rangle,$$

wobei $d\sigma$ wie im Korollar 5.64 *definiert wird (und $\int_{\partial M^2} \langle X, T \rangle$ das Integral der Funktion $\langle X, T \rangle$ längs der Kurve ∂M^2 bezeichnet, siehe* Definition 1.19*).*

Anhang A

Konstruktion des Lebesgue-Maßes

A.1 Lebesgue-Maß auf $\mathbb{R}$

Die Konstruktion des Lebesgue-Maßes auf $\mathbb{R}$ erfolgt in drei Schritten. Als Erstes definieren wir das Maß der Intervalle.

A.1.1 Lebesgue-Maß auf Intervallen

Wir definieren die Mengen

$$\mathcal{J} := \{]-\infty, a[,\ [b, c[,\ [d, \infty[\ \mid a, b, c, d \in \mathbb{R}\} \subset \mathcal{P}(\mathbb{R})$$

und

$$\mathcal{R} := \{\text{endl. Vereinigungen von Elementen von } \mathcal{J}\} \subset \mathcal{P}(\mathbb{R}).$$

Beachte insbesondere, dass $\varnothing$ und $\mathbb{R}$ Elemente von $\mathcal{R}$ sind und dass $\mathcal{R}$ stabil unter endlicher Vereinigung (per Definition von $\mathcal{J}$) und Komplementbildung (denn das Komplement in $\mathbb{R}$ jedes Elements von $\mathcal{J}$ ist wieder die Vereinigung von höchstens zwei Elementen von $\mathcal{J}$ und der Durchschnitt zweier Elemente von $\mathcal{J}$ ist ein Element von $\mathcal{J}$). Man sagt, dass $\mathcal{R}$ eine *Mengenalgebra* auf $\mathbb{R}$ ist. Die Menge $\mathcal{R}$ ist aber keine σ-Algebra auf $\mathbb{R}$, denn z.B. $\bigcup\limits_{n\in\mathbb{N}} [n, n+\frac{1}{2}[$ ist eine abzählbare Vereinigung von Elementen von $\mathcal{R}$, ist aber kein Element von $\mathcal{R}$. Bemerke auch, dass jedes Element A von $\mathcal{R}$ sich auf eindeutige Weise als endliche disjunkte Vereinigung von Elementen von $\mathcal{J}$ schreiben lässt: die entsprechenden Elemente (welche Intervalle sind) sind genau die Zusammenhangskomponenten von A. Für ein $I \in \mathcal{J}$ definieren wir die Länge $|I| \in [0, \infty]$ von I durch $|I| := \infty$ falls $I =]-\infty, a[$ oder $I = [d, \infty[$, $|I| := c - b$ falls $I = [b, c[$ mit $b \leq c$ und $|I| := 0$ falls $I = \varnothing$.

Definition A.1 *Das Lebesgue-Maß auf $\mathcal{R}$ ist die Abbildung*

$$\begin{aligned} \nu : \mathcal{R} &\longrightarrow [0, \infty] \\ A &\longmapsto \sum_{k=1}^{m} |I_k|, \end{aligned}$$

wobei $A = \dot{\bigcup}_{k=1}^{m} I_k$ die Zerlegung von A in endlich vielen Elementen von $\mathcal{J}$ bezeichnet.

Per Definition ist $\nu(\varnothing) = 0$. Offenbar erfüllt ν die Monotonie-Eigenschaft $\nu(A) \leq \nu(B)$ für alle $A, B \in \mathcal{R}$ mit $A \subset B$ (dies folgt aus der entsprechenden Eigenschaft für Intervalle).

Es ist auch elementar, nachzuweisen, dass für alle $A, B \in \mathcal{R}$ die Ungleichung $\nu(A \cup B) \leq \nu(A) + \nu(B)$ mit Gleichheit im Fall $A \cap B = \varnothing$ gilt (betrachte den Fall wo $A, B \in \mathcal{J}$ mit $\nu(A), \nu(B) < \infty$ und unterscheide zwischen zwei Fällen). Diese letzte Eigenschaft ist schon eine Eigenschaft eines Maßes. Eine viel stärkere Aussage kann tatsächlich erreicht werden:

Proposition A.2 *Es gilt* $\nu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=0}^{\infty} \nu(A_n)$ *für jede Folge* $(A_n)_{n\in\mathbb{N}}$ *von Elementen von* $\mathcal{R}$ *mit* $A_n \cap A_m = \varnothing$ *für alle* $n \neq m$ *mit* $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R}$.

Beweis: Das Ergebnis scheint offensichtlich zu sein ("ordne die Intervalle an und addiere die Längen zusammen"), ist es aber nicht: die Intervalle, welche sich zu $\bigcup_{n\in\mathbb{N}} A_n$ zusammensetzen, müssen nicht angeordnet werden – und können es manchmal auch nicht! Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von Elementen von $\mathcal{R}$ mit $A_n \cap A_m = \varnothing$ für alle $n \neq m$ und so, dass $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R}$. Da $A := \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R}$ gilt, hat A nur endlich viele Zusammenhangskomponenten, die alle Elemente von $\mathcal{J}$ sind. Wenn wir das Ergebnis für $A \in \mathcal{J}$ zeigen, bekommen wir das Ergebnis für ein beliebiges A durch Addition endlich vieler Reihen und eventueller Umordnung der Terme (die Umordnung kann wohl durchgeführt werden, weil alle Summanden nichtnegativ sind). Wir nehmen also an, dass $A \in \mathcal{J}$. Analog können wir o.B.d.A. annehmen, dass jedes A_n ein Element von $\mathcal{J}$ ist.
Da für alle $n \in \mathbb{N}$ die Inklusion $\cup_{k=0}^{n} A_k \subset A$ erfüllt ist, folgt aus den vor Proposition A.2 erwähnten Eigenschaften $\nu(A) \geq \nu(\cup_{k=0}^{n} A_k) = \sum_{k=0}^{n} \nu(A_k)$ (die A_k's sind paarweise disjunkt). Dies gilt für alle $n \in \mathbb{N}$; mit $n \longrightarrow \infty$ bekommen wir dann die Ungleichung $\nu(A) \geq \sum_{n=0}^{\infty} \nu(A_n)$. Für die andere Ungleichung fangen mit dem Fall an, wo $A = [b, c[$ für $b < c$ reell. Dann ist jedes A_n der Form $A_n = [b_n, c_n[$ mit $b \leq b_n \leq c_n \leq c$, für alle $n \in \mathbb{N}$ (beachte: $b_n = c_n$ g.d.w. $A_n = \varnothing$). Fixiere $\varepsilon > 0$. Wir vergrößern jedes Intervall $A_n = [b_n, c_n[$ zu $]b_n - \frac{\varepsilon}{2^n}, c_n[$ und gleichzeitig verkleinern $A = [b, c[$ zu $[b, c - \varepsilon]$. Wegen $A = \bigcup_{n\in\mathbb{N}} A_n$ gilt offenbar $[b, c-\varepsilon] \subset \bigcup_{n\in\mathbb{N}}]b_n - \frac{\varepsilon}{2^n}, c_n[$. Da aber $[b, c-\varepsilon]$ *kompakt* ist, existieren nach Satz B.1 endliche viele $n_1, \ldots, n_p \in \mathbb{N}$ mit $[b, c-\varepsilon] \subset \bigcup_{j=1}^{p}]b_{n_j} - \frac{\varepsilon}{2^{n_j}}, c_{n_j}[$, insbesondere $[b, c-\varepsilon[\subset \bigcup_{j=1}^{p} [b_{n_j} - \frac{\varepsilon}{2^{n_j}}, c_{n_j}[$. Nach den vor Proposition A.2 erwähnten Eigenschaften von ν gilt dann $\nu([b, c-\varepsilon[) \leq \sum_{j=1}^{p} \nu([b_{n_j} - \frac{\varepsilon}{2^{n_j}}, c_{n_j}[)$, d.h., $c - \varepsilon - b \leq \sum_{j=1}^{p} c_{n_j} - b_{n_j} + \frac{\varepsilon}{2^{n_j}}$. Daraus folgt

$$\begin{aligned} c - b - \varepsilon &\leq \sum_{n=1}^{\infty} c_n - b_n + \frac{\varepsilon}{2^n} \\ &= \sum_{n=1}^{\infty} c_n - b_n + 2\varepsilon \\ &= \sum_{n=1}^{\infty} \nu(A_n) + 2\varepsilon. \end{aligned}$$

Lassen wir ε gegen 0 laufen, so bekommen wir $c - b \leq \sum_{n=1}^{\infty} \nu(A_n)$, d.h., $\nu(A) \leq \sum_{n=1}^{\infty} \nu(A_n)$. In Fall, wo $A = [d, \infty[$ (und somit $\nu(A) = \infty$) ist, ist lediglich $\sum_{n=0}^{\infty} \nu(A_n) = \infty$ zu zeigen. Für jedes $k \in \mathbb{N}$ gilt aber $[d, d+k[= A \cap [d, d+k[= \cup_{n\in\mathbb{N}} A_n \cap [d, d+k[$ mit $A_n \cap [d, d+k[\in \mathcal{J}$, somit folgt aus dem ersten Schritt (angewendet auf $[d, d+k[$) und der

Monotonie-Eigenschaft von ν:

$$\begin{aligned} k &= \nu([d, d+k[) \\ &\leq \sum_{n=0}^{\infty} \nu(A_n \cap [d, d+k[) \\ &\leq \sum_{n=0}^{\infty} \nu(A_n). \end{aligned}$$

Lassen wir $k \longrightarrow \infty$, so bekommen wir $\sum_{n=0}^{\infty} \nu(A_n) = \infty = \nu(A)$. Der Fall, wo $A =]-\infty, a]$, ist analog. Insgesamt gilt $\nu(A) = \sum_{n=1}^{\infty} \nu(A_n)$, was zu beweisen war. $\square$

Proposition A.2 impliziert, dass ν genau die Eigenschaften eines Maßes erfüllt, mit dem Unterschied, dass $\mathcal{R}$ keine σ-Algebra ist (siehe oben). Die Fortsetzung von ν^* auf eine σ-Algebra erfolgt über einem abstrakten Weg: wir setzen ν erstens auf *ganz* $\mathcal{P}(\mathbb{R})$ zu einem sogenannten *äußeren Maß* fort, welches allerdings kein Maß ist; dann zeigen wir, dass dieses äußere Maß auf eine geeignete σ-Algebra eingeschränkt werden kann, die $\mathcal{R}$ immer noch enthält und auf der das äußere Maß zu einem Maß wird. Dies ist der Gegenstand des nächsten Abschnitts.

A.1.2 Äußere Maße und Fortsetzung von Maßen

Definition A.3 *Ein* äußeres Maß *auf einer Menge X ist eine Abbildung $\nu^* : \mathcal{P}(X) \longrightarrow [0, \infty]$ mit*

i) $\nu^*(\varnothing) = 0$,

ii) $\nu^*(A) \leq \nu^*(B)$ *für alle* $A, B \in \mathcal{P}(X)$ *mit* $A \subset B$ *(ν^* ist monoton).*

iii) $\nu^*(\bigcup\limits_{n\in\mathbb{N}} B_n) \leq \sum_{n=0}^{\infty} \nu(B_n)$ *für jede Folge* $(B_n)_{n\in\mathbb{N}}$ *aus* $\mathcal{P}(X)$ *(ν^* ist σ-subadditiv).*

Die Bezeichnung "äußeres Maß" kommt daher, dass folgendes Verfahren, wo Teilmengen von X "von außen" gemessen werden, ein äußeres Maß liefert:

Proposition A.4 *Sei X eine Menge und $\mathcal{R} \subset \mathcal{P}(X)$ eine unter endlicher Vereinigung und Komplementbildung stabile Teilmenge mit $\varnothing \in \mathcal{R}$ (eine solche Teilmenge heißt* Mengenalgebra *auf X, siehe oben). Sei $\nu : \mathcal{R} \longrightarrow [0, \infty]$ eine Abbildung mit $\nu(\varnothing) = 0$. Dann definiert*

$$\begin{aligned} \nu^* : \mathcal{P}(X) &\longrightarrow [0, \infty] \\ A &\longmapsto \inf \left\{ \sum_{n=0}^{\infty} \nu(A_n) \,|\, (A_n)_{n\in\mathbb{N}} \text{ Folge aus } \mathcal{R} \text{ mit } A \subset \cup_{n\in\mathbb{N}} A_n \right\}. \end{aligned}$$

ein äußeres Maß auf X.

Beweis: Beachte zuerst, dass ν^* wegen $X = \varnothing^c \in \mathcal{R}$ wohldefiniert ist. Die Eigenschaften *i)* und *ii)* aus der Definition A.3 sind trivialerweise erfüllt. Sei $(B_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathcal{P}(X)$. Gilt $\nu^*(B_n) = \infty$ für mindestens ein $n \in \mathbb{N}$, so ist *iii)* natürlich erfüllt. Sei also $\nu^*(B_n) < \infty$ für alle $n \in \mathbb{N}$. Fixiere $\varepsilon > 0$. Nach Definition von ν^* existiert, für jedes $n \in \mathbb{N}$, eine Folge $(A_{n,k})_{k\in\mathbb{N}}$ von Elementen von $\mathcal{R}$ mit $B_n \subset \cup_{k\in\mathbb{N}} A_{n,k}$ und $\sum_{k=0}^{\infty} \nu(A_{n,k}) \leq$

$\nu^*(B_n)+\frac{\varepsilon}{2^n}$ (wegen $\nu^*(B_n)<\infty$). Die Familie $(A_{n,k})_{n,k\in\mathbb{N}}$ ist dann eine abzählbare Familie von Elementen von $\mathcal{R}$ mit $\cup_{n\in\mathbb{N}}B_n\subset\cup_{n,k\in\mathbb{N}}A_{n,k}$, insbesondere gilt

$$\begin{aligned}\nu^*(\cup_{n\in\mathbb{N}}B_n) &\leq \sum_{n,k=0}^{\infty}\nu(A_{n,k})\\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\nu(A_{n,k})\\ &\leq \sum_{n=0}^{\infty}\nu^*(B_n)+\frac{\varepsilon}{2^n}\\ &= \sum_{n=0}^{\infty}\nu^*(B_n)+2\varepsilon.\end{aligned}$$

Lässt man $\varepsilon\longrightarrow 0$, so bekommt man $\nu^*(\cup_{n\in\mathbb{N}}B_n)\leq\sum_{n=0}^{\infty}\nu^*(B_n)$, was zu beweisen war. □

Man beachte, dass die in Proposition A.4 definierte Abbildung ν^* i.A. kein Maß auf $\mathcal{P}(X)$ ist. Betrachte z.B. die Menge $\mathcal{R}$ aus dem Abschnitt A.1.1 und die Abbildung ν aus der Definition A.1. Für $A:=\mathbb{Q}\cap[0,1]$ und $B:=(\mathbb{R}\setminus\mathbb{Q})\cap[0,1]$ gilt $\nu^*(A)=1=\nu^*(B)$ (die Teilmengen A und B liegen beide dicht in $[0,1]$), dennoch gilt $\nu^*([0,1])=\nu^*(A\cup B)=1\neq\nu^*(A)+\nu^*(B)$ und dies, obwohl $A\cap B=\varnothing$ gilt.

Allerdings ist die Einschränkung eines äußeren Maßes auf eine geeignete Teilmenge der Potenzmenge immer ein Maß:

Proposition A.5 *Sei X eine Menge und $\nu^*:\mathcal{P}(X)\longrightarrow[0,\infty]$ ein äußeres Maß. Dann definiert*

$$\mathcal{A}_{\nu^*}:=\Big\{A\in\mathcal{P}(X)\,|\,\nu^*(B)=\nu^*(B\cap A)+\nu^*(B\cap A^c)\ \ \forall\,B\in\mathcal{P}(X)\Big\}$$

eine σ-Algebra auf X und die Einschränkung $\nu^*_{|\mathcal{A}_{\nu^*}}$ ist ein Maß auf $\mathcal{A}_{\nu^*}$.

Beweis: Wegen $\nu^*(B)=\underbrace{\nu^*(\varnothing)}_{0}+\nu^*(B)=\nu^*(B\cap\varnothing)+\nu^*(B\cap X)$ für alle $B\subset X$ gilt $\varnothing\in\mathcal{A}_{\nu^*}$. Ist $A\in\mathcal{A}_{\nu^*}$, so ist nach Definition $A^c=X\setminus A\in\mathcal{A}_{\nu^*}$ (denn $(A^c)^c=A$).
Behauptung 1: *Für alle $A_1,A_2\in\mathcal{A}_{\nu^*}$ gilt $A_1\cup A_2\in\mathcal{A}_{\nu^*}$.*
Beweis: Sei $B\in\mathcal{P}(X)$ beliebig. Wegen $A_1\in\mathcal{A}_{\nu^*}$ gilt für die Teilmenge $B\cap(A_1\cup A_2)\subset X$:

$$\begin{aligned}\nu^*(B\cap(A_1\cup A_2)) &= \nu^*(B\cap(A_1\cup A_2)\cap A_1)+\nu^*(B\cap(A_1\cup A_2)\cap A_1^c)\\ &= \nu^*(B\cap A_1)+\nu^*(B\cap A_2\cap A_1^c),\end{aligned}$$

somit

$$\begin{aligned}\nu^*(B\cap(A_1\cup A_2))+\nu^*(B\cap(A_1\cup A_2)^c) &= \nu^*(B\cap(A_1\cup A_2))+\nu^*(B\cap A_1^c\cap A_2^c)\\ &= \nu^*(B\cap A_1)\\ &\quad+\nu^*(B\cap A_2\cap A_1^c)+\nu^*(B\cap A_1^c\cap A_2^c)\\ &= \nu^*(B\cap A_1)+\nu^*(B\cap A_1^c)\ \text{ wegen } A_2\in\mathcal{A}_{\nu^*}\\ &= \nu^*(B).\end{aligned}$$

Dies beweist $A_1 \cup A_2 \in \mathcal{A}_{\nu^*}$. $\surd$
Behauptung 2: *Seien $A_1, \ldots, A_k \in \mathcal{A}_{\nu^*}$ mit $A_i \cap A_j = \varnothing$ für alle $1 \leq i \neq j \leq k$, dann gilt $\nu^*(B \cap (\cup_{j=1}^k A_j)) = \sum_{j=1}^k \nu^*(B \cap A_j)$ für alle $B \subset X$.*
Beweis: Wir fangen mit dem Fall $k = 2$ an. Sei $B \in \mathcal{P}(X)$ beliebig. Wegen $A_1 \in \mathcal{A}_{\nu^*}$ gilt, für die Teilmenge $B \cap (A_1 \cup A_2) \subset X$:

$$\begin{aligned} \nu^*(B \cap (A_1 \cup A_2)) &= \nu^*(B \cap (A_1 \cup A_2) \cap A_1) + \nu^*(B \cap (A_1 \cup A_2) \cap A_1^c) \\ &= \nu^*(B \cap A_1) + \nu^*(B \cap A_2) \ \text{ wegen } A_2 \subset A_1^c. \end{aligned}$$

Dies beweist die Behauptung für $k = 2$. Für höhere k's führen wir eine Induktion über k durch, wobei genutzt wird, dass die Behauptung für $k = 2$ gilt, sobald $A_1 \in \mathcal{A}_{\nu^*}$ *oder* $A_2 \in \mathcal{A}_{\nu^*}$ gilt. $\surd$
Behauptung 3: *Sei $(A_n)_{n\in\mathbb{N}}$ eine Familie von Elementen von $\mathcal{A}_{\nu^*}$ mit $A_i \cap A_j = \varnothing$ für alle $i \neq j$. Dann gilt $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}_{\nu^*}$.*
Beweis: Sei $B \in \mathcal{P}(X)$ beliebig. Da $B = \big(B \cap (\cup_{n\in\mathbb{N}} A_n)\big) \cup \big(B \cap (\cup_{n\in\mathbb{N}} A_n)^c\big)$ und ν^* ein äußeres Maß ist, gilt

$$\nu^*(B) = \nu^*\Big(\big(B\cap(\cup_{n\in\mathbb{N}}A_n)\big)\cup\big(B\cap(\cup_{n\in\mathbb{N}}A_n)^c\big)\Big) \leq \nu^*\big(B\cap(\cup_{n\in\mathbb{N}}A_n)\big)+\nu^*\big(B\cap(\cup_{n\in\mathbb{N}}A_n)^c\big).$$

Fixiere $n \in \mathbb{N}$ und betrachte die (disjunkte) endliche Familie $(A_j)_{j=0,\ldots,n}$. Nach der Behauptung 1 ist $\cup_{j=0}^n A_j \in \mathcal{A}_{\nu^*}$. Es folgt

$$\begin{aligned} \nu^*(B) &= \nu^*(B \cap (\cup_{j=0}^n A_j)) + \nu^*(B \cap (\cup_{j=1}^k A_j)^c) \\ &\overset{\text{(Beh.2)}}{=} \sum_{j=0}^n \nu^*(B \cap A_j) + \nu^*(B \cap (\cup_{j=1}^k A_j)^c) \\ &\geq \sum_{j=0}^n \nu^*(B \cap A_j) + \nu^*(B \cap (\cup_{n\in\mathbb{N}} A_n)^c) \ \text{ wegen } \cup_{j=0}^n A_j \subset \cup_{n\in\mathbb{N}} A_n. \end{aligned}$$

Lässt man n gegen unendlich laufen, so bekommt man $\nu^*(B) \geq \sum_{n=0}^\infty \nu^*(B\cap A_n)+\nu^*(B\cap (\cup_{n\in\mathbb{N}} A_n)^c)$. Da ν^* ein äußeres Maß ist, folgt schließlich

$$\nu^*(B) \geq \nu^*(B \cap (\cup_{n\in\mathbb{N}} A_n)) + \nu^*(B \cap (\cup_{n\in\mathbb{N}} A_n)^c).$$

Insgesamt erhalten wir $\nu^*(B) = \nu^*(B \cap (\cup_{n\in\mathbb{N}} A_n)) + \nu^*(B \cap (\cup_{n\in\mathbb{N}} A_n)^c)$. Dies zeigt $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}_{\nu^*}$. $\surd$
Ist nun $(A_n)_{n\in\mathbb{N}}$ eine beliebige abzählbare Familie von Elementen von $\mathcal{A}_{\nu^*}$, so definiere $A'_n := A_n \setminus (\cup_{j=0}^{n-1} A_j)$ für alle $j \in \mathbb{N}$ (mit $A'_0 := A_0$). Dann ist $(A'_n)_{n\in\mathbb{N}}$ eine disjunkte Familie von Elementen von $\mathcal{A}_{\nu^*}$ (nach der Behauptung 1) mit $\cup_{n\in\mathbb{N}} A'_n = \cup_{n\in\mathbb{N}} A_n$. Nach der Behauptung 3 ist aber $\cup_{n\in\mathbb{N}} A'_n \in \mathcal{A}_{\nu^*}$, somit ist $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}_{\nu^*}$. Daraus folgt, dass $\mathcal{A}_{\nu^*}$ eine σ-Algebra ist. $\square$

Nun legen wir die Propositionen A.4 und A.5 zusammen:

Satz A.6 *Sei X eine Menge und $\mathcal{R} \subset \mathcal{P}(X)$ eine unter endlicher Vereinigung und Komplementbildung stabile Teilmenge mit $\varnothing \in \mathcal{R}$. Sei $\nu : \mathcal{R} \longrightarrow [0, \infty]$ ein Maß auf $\mathcal{R}$, d.h., $\nu(\varnothing) = 0$ und $\nu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n=0}^\infty \nu(A_n)$ für jede disjunkte Familie $(A_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{R}$ mit $\cup_{n\in\mathbb{N}} A_n \in \mathcal{R}$. Sei $\nu^* : \mathcal{P}(X) \longrightarrow [0, \infty]$ das zu ν gehörige äußere Maß (siehe* Proposition A.4*). Dann gilt:*

a) Für die in Proposition A.5 *definierte Menge* $\mathcal{A}_{\nu^*}$ *gilt* $\mathcal{R} \subset \mathcal{A}_{\nu^*}$ *und* $\nu^*_{|\mathcal{R}} = \nu$.

b) Sei ν σ*-endlich, d.h., eine Familie* $(A_n)_{n\in\mathbb{N}}$ *von Elementen von* $\mathcal{R}$ *existiere mit* $X = \cup_{n\in\mathbb{N}} A_n$ *und* $\nu(A_n) < \infty$ *für alle* $n \in \mathbb{N}$. *Ist* ν' *ein Maß auf einer* σ*-Algebra* $\mathcal{A}$ *auf* X *mit* $\mathcal{R} \subset \mathcal{A} \subset \mathcal{A}_{\nu^*}$ *und* $\nu'|_{\mathcal{R}} = \nu$, *so gilt* $\nu^*|_{\mathcal{A}} = \nu'$.

Beweis: a) Sei $A \in \mathcal{R}$ und $B \subset X$ eine beliebige Teilmenge. Wir wollen zeigen, dass $\nu^*(B) \geq \nu^*(B \cap A) + \nu^*(B \cap A^c)$ gilt (die andere Ungleichung folgt aus der Tatsache, dass ν^* ein äußeres Maß ist). Ist $(A_n)_{n\in\mathbb{N}}$ eine beliebige abzählbare Familie von Elementen von $\mathcal{R}$ mit $B \subset \cup_{n\in\mathbb{N}} A_n$, so gilt wegen $A_n = (A_n \cap A) \cup (A_n \cap A^c)$ (mit $A_n \cap A, A_n \cap A^c \in \mathcal{R}$) die Identität $\nu(A_n) = \nu(A_n \cap A) + \nu(A_n \cap A^c)$ und somit $\sum_{n=0}^{\infty} \nu(A_n) = \sum_{n=0}^{\infty} \nu(A_n \cap A) + \nu(A_n \cap A^c)$. Wegen $\cup_{n\in\mathbb{N}} A_n \cap A \supset B \cap A$ und $\cup_{n\in\mathbb{N}} A_n \cap A^c \supset B \cap A^c$ gelten $\sum_{n=0}^{\infty} \nu(A_n \cap A) \geq \nu^*(B \cap A)$ und $\sum_{n=0}^{\infty} \nu(A_n \cap A^c) \geq \nu^*(B \cap A^c)$. Daraus folgt $\sum_{n=0}^{\infty} \nu(A_n) \geq \nu^*(B \cap A) + \nu^*(B \cap A^c)$. Bilden wir das Infimum über alle Familien $(A_n)_{n\in\mathbb{N}}$, so bekommen wir $\nu^*(B) \geq \nu^*(B \cap A) + \nu^*(B \cap A^c)$. Dies beweist $A \in \mathcal{A}_{\nu^*}$.
Sei $A \in \mathcal{R}$. Die Familie $(A_0 := A, A_n := \varnothing \ \forall n \geq 1)$ ist eine abzählbare Familie von Elementen von $\mathcal{R}$ mit $A \subset \cup_{n\in\mathbb{N}} A_n$, insbesondere gilt nach Definition des äußeren Maßes $\nu^*(A) \leq \sum_{n=0}^{\infty} \nu(A_n) = \nu(A)$. Ist andererseits $(A_n)_{n\in\mathbb{N}}$ eine beliebige abzählbare Familie von Elementen von $\mathcal{R}$ mit $A \subset \cup_{n\in\mathbb{N}} A_n$, so gilt $A = A \cap \cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} A \cap A_n$ mit $A \cap A_n \in \mathcal{R}$, insbesondere folgt aus der σ-Additivität $\nu(A) = \sum_{n=0}^{\infty} \nu(A \cap A_n)$. Da ν monoton ist, folgt $\nu(A) \leq \sum_{n=0}^{\infty} \nu(A_n)$. Dies zeigt $\nu(A) \leq \nu^*(A)$ und damit $\nu(A) = \nu^*(A)$.
b) Sei $A \in \mathcal{A}$. Wir zeigen beide Ungleichungen $\nu'(A) \leq \nu^*(A)$ und $\nu'(A) \geq \nu^*(A)$ separat. Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von Elementen von $\mathcal{R}$ mit $A \subset \cup_{n\in\mathbb{N}} A_n$, so gilt

$$\sum_{n=0}^{\infty} \nu(A_n) = \sum_{n=0}^{\infty} \nu'(A_n) \geq \nu'(\cup_{n\in\mathbb{N}} A_n) \geq \nu'(A).$$

Daraus folgt $\nu'(A) \leq \nu^*(A)$.
Da ν nach Voraussetzung σ-endlich ist, existiert eine Folge $(B_n)_{n\in\mathbb{N}}$ mit $B_n \cap B_m = \varnothing$ für alle $n \neq m$, $\cup_{n\in\mathbb{N}} B_n = X$ und $\nu(B_n) < \infty$ (nehme die Folge $(A_n)_{n\in\mathbb{N}}$ aus Satz A.6.b) und setze $B_n := A_n \setminus (\cup_{j=0}^{n-1} A_j) \in \mathcal{R}$). Wegen $A \in \mathcal{A} \subset \mathcal{A}_{\nu^*}$ gilt $\nu^*(B_n) = \nu^*(B_n \cap A) + \nu^*(B_n \cap A^c)$. Wegen $B_n \in \mathcal{R} \subset \mathcal{A}$ gilt $\nu'(B_n) = \nu'(B_n \cap A) + \nu'(B_n \cap A^c)$. Aus $\nu^*(B_n) = \nu(B_n) = \nu'(B_n)$ folgt $\nu^*(B_n \cap A) + \nu^*(B_n \cap A^c) = \nu'(B_n \cap A) + \nu'(B_n \cap A^c)$. Da $\nu(B_n) < \infty$ ist, sind alle Summanden der letzten Gleichung endlich. Es folgt die Identität $\nu^*(B_n \cap A) - \nu'(B_n \cap A) = \nu'(B_n \cap A^c) - \nu^*(B_n \cap A^c)$, deren rechte Seite nichtpositiv ist nach dem ersten Schritt. Damit ist $\nu^*(B_n \cap A) - \nu'(B_n \cap A) \leq 0$ und die σ-Additivität von ν' und ν^* liefert

$$\begin{aligned} \nu^*(A) &= \sum_{n=0}^{\infty} \nu^*(A \cap B_n) \\ &\leq \sum_{n=0}^{\infty} \nu'(A \cap B_n) \\ &= \nu'(A). \end{aligned}$$

Damit bekommen wir $\nu'(A) = \nu^*(A)$, was zu beweisen war. □

Nun sind wir in der Lage, das "Maß" ν aus dem Abschnitt A.1.1 auf eine σ-Algebra fortzusetzen.

Korollar A.7 *Die Abbildung* $\nu : \mathcal{R} \longrightarrow [0, \infty]$ *aus der* Definition A.1 *lässt sich auf eindeutige Weise zu einem (σ-endlichen) Maß* $\nu^* : \mathcal{A}_{\nu^*} \longrightarrow [0, \infty]$ *fortsetzen, wobei* $\mathcal{A}_{\nu^*}$ *und* ν^* *wie in* Propositionen A.5 *und* A.4 *definiert sind. Außerdem gilt* $\mathcal{B}(\mathbb{R}) = \mathcal{A}\langle\mathcal{R}\rangle \subset \mathcal{A}_{\nu^*}$ *und die Einschränkung* $\nu^*_{|\mathcal{B}(\mathbb{R})} : \mathcal{B}(\mathbb{R}) \longrightarrow [0, \infty]$ *ist ebenfalls das einzige (σ-endliche) Maß auf* $\mathcal{B}(\mathbb{R})$ *mit* $\nu^*([a, b[) = b - a$ *für alle* $a \leq b$ *reell.*

Beweis: Nach Proposition A.2 erfüllt ν die Eigenschaften eines Maßes auf der Mengenalgebra $\mathcal{R}$. Satz A.6.a) liefert, dass die Einschränkung des zu ν gehörigen äußeren Maßes ν^* auf $\mathcal{A}_{\nu^*}$ ein Maß ist mit $\mathcal{R} \subset \mathcal{A}_{\nu^*}$ und $\nu^*_{|\mathcal{R}} = \nu$, insbesondere gilt $\mathcal{A}\langle\mathcal{R}\rangle \subset \mathcal{A}_{\nu^*}$ nach Proposition 3.6. Dies beweist die Existenz der Fortsetzung auf $\mathcal{A}_{\nu^*}$. Wir bezeichnen $\nu^*_{|\mathcal{A}_{\nu^*}}$ weiterhin mit ν^*. Beachte, dass jedes offene Intervall zu $\mathcal{A}\langle\mathcal{R}\rangle$ gehört (trivial, denn z.B. $]a, b[= \bigcup\limits_{n \in \mathbb{N}\setminus\{0\}} [a + \frac{1}{n}, b[)$, insbesondere ist $\mathcal{B}(\mathbb{R}) \subset \mathcal{A}\langle\mathcal{R}\rangle$; umgekehrt ist jedes Element aus $\mathcal{J}$ – und somit aus $\mathcal{R}$ – auch ein Element aus $\mathcal{B}(\mathbb{R})$ (z.B. $[a, b[= \bigcap\limits_{n \in \mathbb{N}\setminus\{0\}}]a - \frac{1}{n}, b[)$, somit ist $\mathcal{A}\langle\mathcal{R}\rangle \subset \mathcal{B}(\mathbb{R})$. Dies beweist $\mathcal{B}(\mathbb{R}) = \mathcal{A}\langle\mathcal{R}\rangle$. Da $\mathbb{R} = \cup_{n \in \mathbb{N}}[-n, n[$ mit $[-n, n[\in \mathcal{R}$ und $\nu([-n, n[) = 2n < \infty$, sind ν – und somit auch ν^* und $\nu^*_{|\mathcal{B}(\mathbb{R})}$ – σ-endlich. Die Eindeutigkeit von ν^* und $\nu^*_{|\mathcal{B}(\mathbb{R})}$ folgt aus Satz A.6.b) zusammen mit der Tatsache, dass jedes auf $\mathcal{R}$ definierte Maß τ mit $\tau([a, b[) = b - a$ für alle $a \leq b$ reell automatisch $\tau(] - \infty, a[) = \tau([d, \infty[) = \infty$ erfüllt, für alle $a, d \in \mathbb{R}$ und somit stimmt mit ν auf $\mathcal{R}$ überein. $\square$

Korollar A.7 beweist insbesondere die Existenz und die Eindeutigkeit des Maßes μ aus dem Satz 3.15: in den obigen Bezeichnungen ist $\mu = \nu^*_{|\mathcal{B}(\mathbb{R})}$. Wir behaupten nun, dass das Maß ν^* aus dem Korollar A.7 mit dem Lebesgue-Maß aus der Definition 3.19 übereinstimmt. Dafür muss die Vollständigkeit von ν^* untersucht werden.

A.1.3 Vollständigkeit der Fortsetzung

In diesem Abschnitt vergleichen wir das Fortsetzungsverfahren für Maße aus dem Abschnitt A.1.2 mit dem der Vervollständigung (Proposition 3.18).

Proposition A.8 *Sei* $(X, \mathcal{R}, \nu)$ *wie im* Satz A.6 *mit σ-endlichem* $\nu : \mathcal{R} \to [0, \infty]$. *Dann ist* $(X, \mathcal{A}_{\nu^*}, \nu^*_{|\mathcal{A}_{\nu^*}})$ *ein vollständiger Maßraum und es gilt* $\mathcal{A}_{\nu^*} = \widehat{\mathcal{A}\langle\mathcal{R}\rangle}$ *sowie* $\nu^*_{|\mathcal{A}_{\nu^*}} = \widehat{\nu}$.

Beweis: Sei $B \subset X$ eine ν^*-Nullmenge, dann existiert ein $N \in \mathcal{A}_{\nu^*}$ mit $B \subset N$ und $\nu^*(N) = 0$. Da ν^* ein äußeres Maß auf $\mathcal{P}(X)$ ist, folgt $\nu^*(B) \leq \nu^*(N) = 0$, d.h., $\nu^*(B) = 0$. Für alle $E \subset X$ ist $E \cap B \subset B$, somit gilt auch $\nu^*(E \cap B) = 0$ und daher $\nu^*(E) \geq \nu^*(E \cap B^c) = \nu^*(E \cap B) + \nu^*(E \cap B^c)$. Da die andere Ungleichung $\nu^*(E) \leq \nu^*(E \cap B) + \nu^*(E \cap B^c)$ für ein äußeres Maß immer erfüllt ist, folgt $\nu^*(E) = \nu^*(E \cap B) + \nu^*(E \cap B^c)$. Dies zeigt $B \in \mathcal{A}_{\nu^*}$ und die Vollständigkeit von $(X, \mathcal{A}_{\nu^*}, \nu^*_{|\mathcal{A}_{\nu^*}})$.
Wie im Beweis von Korollar A.7 ist zu bemerken, dass wegen $\mathcal{R} \subset \mathcal{A}_{\nu^*}$ auch $\mathcal{A}\langle\mathcal{R}\rangle \subset \mathcal{A}_{\nu^*}$ gilt. Proposition 3.18 liefert, dass $\nu^*_{|\mathcal{A}_{\nu^*}}$ eine Fortsetzung von $\widehat{\nu}$ ist, d.h., $\widehat{\mathcal{A}\langle\mathcal{R}\rangle} \subset \mathcal{A}_{\nu^*}$ und $\nu^*_{|\widehat{\mathcal{A}\langle\mathcal{R}\rangle}} = \widehat{\nu}$. Es bleibt, $\mathcal{A}_{\nu^*} \subset \widehat{\mathcal{A}\langle\mathcal{R}\rangle}$ zu beweisen. Dafür ist folgende Behauptung nutzlich.
Behauptung: *Für* $B \subset X$ *gilt* $B \in \mathcal{A}_{\nu^*}$ *g.d.w.* $A, C \in \mathcal{A}\langle\mathcal{R}\rangle$ *existieren mit* $A \subset B \subset C$ *und* $\nu^*(C \setminus A) = 0$.
Beweis: Sei $B \subset X$ so, dass $A, C \in \mathcal{A}\langle\mathcal{R}\rangle$ existieren mit $A \subset B \subset C$ und $\nu^*(C \setminus A) = 0$.

Sei $E \subset X$ beliebig. Wegen $A, C \in \mathcal{A}\langle\mathcal{R}\rangle \subset \mathcal{A}_{\nu^*}$ gelten $\nu^*(E) = \nu^*(E \cap A) + \nu^*(E \cap A^c)$ sowie $\nu^*(E) = \nu^*(E \cap C) + \nu^*(E \cap C^c)$. Andererseits gilt

$$\begin{aligned}\nu^*(E \cap C) &= \nu^*(E \cap \underbrace{C \cap A}_{=A}) + \nu^*(E \cap C \cap A^c) \\ &= \nu^*(E \cap A) + \underbrace{\nu^*(E \cap (C \setminus A))}_{\leq \nu^*(C\setminus A)=0} \\ &= \nu^*(E \cap A)\end{aligned}$$

und analog $\nu^*(E \cap A^c) = \nu^*(E \cap C^c) + \nu^*(E \cap (C \setminus A)) = \nu^*(E \cap C^c)$. Daraus folgt

$$\begin{aligned}\nu^*(E) &\leq \nu^*(E \cap B) + \nu^*(E \cap B^c) \\ &\leq \nu^*(E \cap C) + \nu^*(E \cap A^c) \\ &= \nu^*(E \cap C) + \nu^*(E \cap C^c) \\ &= \nu^*(E),\end{aligned}$$

woraus $\nu^*(E) = \nu^*(E \cap B) + \nu^*(E \cap B^c)$ folgt. Dies beweist $B \in \mathcal{A}_{\nu^*}$. Sei umgekehrt $B \in \mathcal{A}_{\nu^*}$. Wir nehmen zuerst an, dass $\nu^*(B) < \infty$ ist. Dann existiert, für jedes $k \in \mathbb{N}$, eine abzählbare Familie $(C_{n,k})_{n\in\mathbb{N}}$ von Elementen von $\mathcal{R}$ mit $B \subset \cup_{n\in\mathbb{N}} C_{n,k}$ und $\sum_{n=0}^{\infty} \nu(C_{n,k}) \leq \nu^*(B) + \frac{1}{2^k}$. Die Teilmenge $C := \cap_{k\in\mathbb{N}}(\cup_{n\in\mathbb{N}} C_{n,k})$ von X ist dann ein Element von $\mathcal{A}\langle\mathcal{R}\rangle$ mit $B \subset C$ und

$$\nu^*(B) \leq \nu^*(C) \leq \nu^*(\cup_{n\in\mathbb{N}} C_{n,k}) \leq \sum_{n=0}^{\infty} \nu(C_{n,k}) \leq \nu^*(B) + \frac{1}{2^k}$$

für alle $k \in \mathbb{N}$, somit $\nu^*(C) = \nu^*(B)$. Beachte, dass wegen $B \in \mathcal{A}_{\nu^*}$ insbesondere $\nu^*(C) = \nu^*(C \cap B) + \nu^*(C \cap B^c) = \nu^*(B) + \nu^*(C \setminus B)$ gilt; mit $\nu^*(C) = \nu^*(B) < \infty$ folgt $\nu^*(C \setminus B) = 0$. Analog existiert für die Teilmenge $C \setminus B \subset X$ (welche Maß $0 < \infty$ hat) ein Element $D \in \mathcal{A}\langle\mathcal{R}\rangle$ mit $C \setminus B \subset D$ und $\nu^*(C \setminus B) = \nu^*(D) = 0$. Setze nun $A := C \cap D^c$, dann ist $A \in \mathcal{A}\langle\mathcal{R}\rangle$ mit $A \subset C \cap (C \setminus B)^c = C \cap B = B$ und $\nu^*(C \setminus A) = \nu^*(C \cap (C^c \cup D)) \leq \nu^*(D) = 0$, damit $\nu^*(C \setminus A) = 0$. Dies beweist die andere Inklusion in diesem Fall. Im Fall, wo $\nu^*(B)$ beliebig ist, zerlegen wir $X = \cup_{n\in\mathbb{N}} B_n$ mit $B_n \in \mathcal{R}$, $B_n \cap B_m = \varnothing$ für alle $m \neq n$ und $\nu(B_n) < \infty$ für alle $n \in \mathbb{N}$ (möglich da ν σ-endlich). Insbesondere gilt $B \cap B_n \in \mathcal{A}_{\nu^*}$ mit $\nu^*(B \cap B_n) \leq \nu^*(B_n) < \infty$ für alle $n \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$ finden wir nach dem ersten Fall Elemente $A_n, C_n \in \mathcal{A}\langle\mathcal{R}\rangle$ mit $A_n \subset B \cap B_n \subset C_n$ und $\nu^*(C_n \setminus A_n) = 0$. Setze $A := \cup_{n\in\mathbb{N}} A_n$ und $C := \cap_{n\in\mathbb{N}} C_n$, dann gilt $A, C \in \mathcal{A}\langle\mathcal{R}\rangle$ mit $A \subset \cup_{n\in\mathbb{N}} B \cap B_n = B \subset \cap_{n\in\mathbb{N}} C_n = C$ und $\nu^*(C \setminus A) = \nu^*(\cap_{n\in\mathbb{N}} C_n \cap \cap_{k\in\mathbb{N}} A_k^c) \leq \nu^*(C_0 \setminus A_0) = 0$, d.h., $\nu^*(C \setminus A) = 0$. Dies zeigt die Behauptung. $\surd$

Sei nun $B \in \mathcal{A}_{\nu^*}$, dann existieren nach der Behauptung $A, C \in \mathcal{A}\langle\mathcal{R}\rangle$ mit $A \subset B \subset C$ und $\nu^*(C \setminus A) = 0$. Es folgt $B = A \cup (B \setminus A)$ mit $A \in \mathcal{A}\langle\mathcal{R}\rangle$, $B \setminus A \subset C \setminus A$ und $\nu^*(C \setminus A) = 0$, d.h., B ist die Vereinigung eines Elements aus $\mathcal{A}\langle\mathcal{R}\rangle$ und einer $\nu^*_{|\mathcal{A}\langle\mathcal{R}\rangle}$-Nullmenge. Daraus folgt $B \in \widehat{\mathcal{A}\langle\mathcal{R}\rangle}$. Dies zeigt $\mathcal{A}_{\nu^*} = \widehat{\mathcal{A}\langle\mathcal{R}\rangle}$ und schließt den Beweis der Proposition A.8 ab. $\square$

Proposition A.8 impliziert, dass das in der Definition 3.19 definierte Lebesgue-Maß λ mit dem Maß ν^* aus dem Korollar A.7 übereinstimmt.

A.2 Lebesgue-Maß auf $\mathbb{R}^k$

Wie bereits im Abschnitt 3.3.3 erwähnt wird, entsteht das Lesbesgue-Maß auf $\mathbb{R}^k$ durch das Vervollständigen eines Produktmaßes, siehe Definition 3.66. Dafür braucht man die Konstruktion des Produktmaßes zweier σ-endlicher Maße. Eine gute Referenz für diese Konstruktion ist [6].

A.2.1 Existenz und Eindeutigkeit des Produktmaßes

In diesem Abschnitt wollen wir Proposition 3.60 beweisen. Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Maßräume mit μ und ν *σ-endlich*. WOHLDEFINIERTHEIT VON ρ ZEIGEN... Wir müssen zeigen, dass die Abbildung $\mathcal{A} \times \mathcal{B} \longrightarrow [0, \infty]$, $A \times B \longmapsto \mu(A) \cdot \nu(B)$, sich auf $\mathcal{A} \otimes \mathcal{B}$ zu einem eindeutigen Maß fortsetzen lässt. Wir besprechen zuerst die Eindeutigkeit der Fortsetzung. Dafür führen wir den Begriff einer *monotonen Klasse* ein:

Definition A.9 *Eine* monotone Klasse *auf einer Menge X ist eine Teilmenge $\mathcal{M} \subset \mathcal{P}(X)$, welche folgende Monotonieeigenschaften erfüllt:*

i) Für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{M}$ mit $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

ii) Für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{M}$ mit $A_n \supset A_{n+1}$ $\forall n \in \mathbb{N}$ gilt $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

Zum Beispiel ist jede σ-Algebra eine monotone Klasse; nicht jede monotone Klasse ist aber eine σ-Algebra (wähle z.B. $\mathcal{M} := \{X\} \subset \mathcal{P}(X)$). Der Begriff einer monotonen Klasse kommt daher, dass die Menge der Elemente einer gegebenen σ-Algebra, wo zwei endliche Maße übereinstimmen, eine monotone Klasse ist: sind $\mu_1, \mu_2 : \mathcal{A} \longrightarrow [0, \infty]$ zwei endliche Maße, so ist

$$\{A \in \mathcal{A} \,|\, \mu_1(A) = \mu_2(A)\}$$

eine monotone Klasse (*Übungsaufgabe; benutze dabei* Proposition 3.12).

Wie bei σ-Algebren existiert der Begriff einer *erzeugten* monotonen Klasse:

Proposition A.10 *Für eine Menge X sei $\mathcal{S} \subset \mathcal{P}(X)$ eine beliebige Teilmenge. Dann ist*

$$\mathcal{M}\langle\mathcal{S}\rangle := \bigcap_{\substack{\mathcal{M} \text{ monot. Kl.} \\ \mathcal{S}\subset\mathcal{M}}} \mathcal{M}$$

eine monotone Klasse auf X, die $\mathcal{S}$ enthält. Diese monotone Klasse ist die kleinste monotone Klasse auf X, die $\mathcal{S}$ enthält. Sie wird die von $\mathcal{S}$ erzeugte monotone Klasse genannt.

Beweis: Per Definition enthält $\mathcal{M}\langle\mathcal{S}\rangle$ die Menge $\mathcal{S}$. Sei $(A_n)_{n\in\mathbb{N}}$ eine beliebige Folge von Elementen von $\mathcal{M}\langle\mathcal{S}\rangle$. Für jede monotone Klasse $\mathcal{M}$ auf X mit $\mathcal{S} \subset \mathcal{M}$ ist insbesondere $(A_n)_{n\in\mathbb{N}}$ eine Folge von Elementen von $\mathcal{M}$. Falls $(A_n)_{n\in\mathbb{N}}$ monoton wachsend ist (d.h., $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}$), ist $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$ nach Definition einer monotonen Klasse. Dies gilt für jede monotone Klasse $\mathcal{M}$ auf X mit $\mathcal{S} \subset \mathcal{M}$, d.h., $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}\langle\mathcal{S}\rangle$. Der Fall einer monoton fallenden Folge ist analog. Dass $\mathcal{M}\langle\mathcal{S}\rangle \subset \mathcal{M}$ für jede monotone Klasse $\mathcal{M}$ auf X mit $\mathcal{S} \subset \mathcal{M}$ gilt, ist per Definition von $\mathcal{M}\langle\mathcal{S}\rangle$ klar. Damit ist $\mathcal{M}\langle\mathcal{S}\rangle$ die kleinste $\mathcal{S}$

enthaltende monotone Klasse auf X. $\square$

Da jede σ-Algebra eine monotone Klasse ist, gilt insbesondere $\mathcal{M}\langle\mathcal{S}\rangle \subset \mathcal{A}\langle\mathcal{S}\rangle$, wobei $\mathcal{A}\langle\mathcal{S}\rangle$ die von $\mathcal{S}$ erzeugte σ-Algebra bezeichnet. Im Fall, wo die Teilmenge $\mathcal{S} \subset \mathcal{P}(X)$ eine Mengenalgebra ist, ist die davon erzeugte monotone Klasse eine σ-Algebra:

Proposition A.11 *Für eine Menge X sei $\mathcal{S} \subset \mathcal{P}(X)$ eine Mengenalgebra über X, d.h., es gelten $\varnothing \in \mathcal{S}$, $A^c := X \setminus A \in \mathcal{S}$ für alle $A \in \mathcal{S}$ und $A \cup B \in \mathcal{S}$ für alle $A, B \in \mathcal{S}$. Dann ist $\mathcal{M}\langle\mathcal{S}\rangle$ eine σ-Algebra auf X. Insbesondere gilt $\mathcal{M}\langle\mathcal{S}\rangle = \mathcal{A}\langle\mathcal{S}\rangle$.*

Beweis: Wir zeigen zuerst, dass $\mathcal{M}\langle\mathcal{S}\rangle$ eine Mengenalgebra ist. Daraus folgt unmittelbar, dass $\mathcal{M}\langle\mathcal{S}\rangle$ eine σ-Algebra ist. Nach Definition gilt $\varnothing \in \mathcal{S} \subset \mathcal{M}\langle\mathcal{S}\rangle$.
Behauptung 1: *Für alle $A \in \mathcal{M}\langle\mathcal{S}\rangle$ ist $A^c \in \mathcal{M}\langle\mathcal{S}\rangle$.*
Beweis der Behauptung: Wir betrachten die Menge

$$\mathcal{N} := \{A \in \mathcal{P}(X) \,|\, A^c \in \mathcal{M}\langle\mathcal{S}\rangle\}$$

und zeigen, dass $\mathcal{N}$ eine monotone Klasse auf X ist, die $\mathcal{S}$ enthält; somit enthält sie automatisch $\mathcal{M}\langle\mathcal{S}\rangle$ und die Behauptung ist bewiesen.
Ist $A \in \mathcal{S}$, so ist nach Definition einer Mengenalgebra $A^c \in \mathcal{S}$ und somit $A^c \in \mathcal{M}\langle\mathcal{S}\rangle$; dies beweist $\mathcal{S} \subset \mathcal{N}$. Sei nun $(A_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Elementen von $\mathcal{N}$. Dann ist $(\bigcup_{n\in\mathbb{N}} A_n)^c = \bigcap_{n\in\mathbb{N}} A_n^c$, wobei $(A_n^c)_{n\in\mathbb{N}}$ eine monoton fallende Folge von Teilmengen von X ist. Für alle $n \in \mathbb{N}$ ist aber $A_n^c \in \mathcal{M}\langle\mathcal{S}\rangle$ (wegen $A_n \in \mathcal{N}$), insbesondere ist $(A_n^c)_{n\in\mathbb{N}}$ eine monoton fallende Folge von Elementen von $\mathcal{M}\langle\mathcal{S}\rangle$. Da $\mathcal{M}\langle\mathcal{S}\rangle$ eine monotone Klasse ist, folgt $\bigcap_{n\in\mathbb{N}} A_n^c \in \mathcal{M}\langle\mathcal{S}\rangle$. Der Fall einer monoton wachsenden Folge von Elementen von $\mathcal{N}$ ist analog. Damit ist $\mathcal{N}$ eine $\mathcal{S}$ enthaltende monotone Klasse auf X und die Behauptung ist bewiesen. $\surd$
Behauptung 2: *Für alle $A, B \in \mathcal{M}\langle\mathcal{S}\rangle$ gilt $A \cup B \in \mathcal{M}\langle\mathcal{S}\rangle$.*
Beweis der Behauptung: Sei zuerst $B \in \mathcal{S}$. Wie vorher betrachten wir die Menge

$$\mathcal{U}_B := \{A \in \mathcal{P}(X) \,|\, A \cup B \in \mathcal{M}\langle\mathcal{S}\rangle\}$$

und zeigen, dass $\mathcal{U}_B$ eine $\mathcal{S}$ enthaltende monotone Klasse ist. Ist zuerst $A \in \mathcal{S}$, so gilt $A \cup B \in \mathcal{S} \subset \mathcal{M}\langle\mathcal{S}\rangle$ (weil $\mathcal{S}$ eine Mengenalgebra ist), insbesondere gilt $A \in \mathcal{U}_B$. Sei nun $(A_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Elementen von $\mathcal{U}_B$. Dann ist $(\bigcup_{n\in\mathbb{N}} A_n) \cup B = \bigcup_{n\in\mathbb{N}}(A_n \cup B)$, wobei $(A_n \cup B)_{n\in\mathbb{N}}$ eine monoton wachsende Folge von Teilmengen von $\mathcal{P}(X)$ ist. Da aber $A_n \in \mathcal{U}_B$ ist, gilt $A_n \cup B \in \mathcal{M}\langle\mathcal{S}\rangle$, insbesondere $\bigcup_{n\in\mathbb{N}}(A_n \cup B) \in \mathcal{M}\langle\mathcal{S}\rangle$, da $\mathcal{M}\langle\mathcal{S}\rangle$ eine monotone Klasse ist. Dies zeigt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{U}_B$. Der Fall einer monoton fallenden Folge von Elementen von $\mathcal{U}_B$ ist analog. Damit ist $\mathcal{U}_B$ eine $\mathcal{S}$ enthaltende monotone Klasse auf X, insbesondere enthält sie auch $\mathcal{M}\langle\mathcal{S}\rangle$. D.h., wir erhalten damit, dass $A \cup B \in \mathcal{M}\langle\mathcal{S}\rangle$ für alle $A \in \mathcal{M}\langle\mathcal{S}\rangle$ und alle $B \in \mathcal{S}$.
Sei nun $B \in \mathcal{M}\langle\mathcal{S}\rangle$ beliebig. Wir betrachten die oben definierte Menge $\mathcal{U}_B$ und zeigen wiederum, dass $\mathcal{U}_B$ eine $\mathcal{S}$ enthaltende monotone Klasse ist. Nach dem obigen Fall ist $\mathcal{S} \subset \mathcal{U}_B$. Der Beweis davon, dass $\mathcal{U}_B$ eine monotone Klasse ist, ist derselbe. Dies zeigt $\mathcal{M}\langle\mathcal{S}\rangle \subset \mathcal{U}_B$. Insgesamt erhalten wir $A \cup B \in \mathcal{M}\langle\mathcal{S}\rangle$ für alle $A, B \in \mathcal{M}\langle\mathcal{S}\rangle$, was zu beweisen war. $\surd$
Nach den Behauptungen 1 und 2 (sowie wegen $\varnothing \in \mathcal{S} \subset \mathcal{M}\langle\mathcal{S}\rangle$) ist $\mathcal{M}\langle\mathcal{S}\rangle$ eine Mengenalgebra. Es bleibt noch, zu zeigen, dass $\mathcal{M}\langle\mathcal{S}\rangle$ eine σ-Algebra ist.
Behauptung 3: *Die monotone Klasse $\mathcal{M}\langle\mathcal{S}\rangle$ ist eine σ-Algebra.*

Beweis der Behauptung: Es ist lediglich, zu zeigen, dass $\mathcal{M}\langle\mathcal{S}\rangle$ unter abzählbarer Vereinigung stabil ist. Sei $(A_n)_{n\in\mathbb{N}}$ eine beliebige Folge von Elementen von $\mathcal{M}\langle\mathcal{S}\rangle$. Dann gilt $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} E_n$, wobei $E_n := \bigcup_{k=0}^n A_k$. Da aber $\mathcal{M}\langle\mathcal{S}\rangle$ eine Mengenalgebra ist, gilt $E_n \in \mathcal{M}\langle\mathcal{S}\rangle$. Außerdem ist die Folge $(E_n)_{n\in\mathbb{N}}$ offenbar monoton wachsend (es gilt $E_n \subset E_{n+1}$ für alle $n \in \mathbb{N}$). Da $\mathcal{M}\langle\mathcal{S}\rangle$ eine monotone Klasse ist, folgt $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} E_n \in \mathcal{M}\langle\mathcal{S}\rangle$. Dies beweist die Behauptung. $\surd$
Da jetzt $\mathcal{M}\langle\mathcal{S}\rangle$ eine σ-Algebra ist, die $\mathcal{S}$ enthält, gilt $\mathcal{A}\langle\mathcal{S}\rangle \subset \mathcal{M}\langle\mathcal{S}\rangle$. Die andere Inklusion haben wir bereits bewiesen, siehe oben. Damit gilt $\mathcal{A}\langle\mathcal{S}\rangle = \mathcal{M}\langle\mathcal{S}\rangle$ und die Proposition A.11 ist bewiesen. □

Proposition A.11 erlaubt es, die Gleichheit zweier endlicher Maße zu beweisen, die bereits auf einer echt kleineren Teilmenge der σ-Algebra übereinstimmen. Dafür brauchen wir den Begriff von Mengenhalbalgebra:

Definition A.12 *Eine* Mengenhalbalgebra *auf einer Menge X ist eine Teilmenge $\mathcal{H} \subset \mathcal{P}(X)$, welche folgende Eigenschaften besitzt:*

i) Für alle $A, B \in \mathcal{H}$ gilt $A \cap B \in \mathcal{H}$.

ii) Für jedes $A \in \mathcal{H}$ existiert eine endliche paarweise disjunkte *Familie $(A_i)_{i=0}^k$ aus $\mathcal{H}$ (wobei $k \in \mathbb{N}$) mit $A^c = \bigcup_{i=0}^k A_i$.*

Ein einfaches Beispiel einer Mengenhalbalgebra ist die Teilmenge $\mathcal{A} \times \mathcal{B} := \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$ von $\mathcal{P}(X \times Y)$, wobei $\mathcal{A}$ bzw. $\mathcal{B}$ eine σ-Algebra auf X bzw. Y ist. Nämlich gilt $(A\times B)\cap(C\times D) = (A\cap C)\times(B\cap D)$ und $(A\times B)^c = (A^c\times B) \dot\cup (A\times B^c) \dot\cup (A^c\times B^c)$. Jede Mengenalgebra ist offensichtlich eine Mengenhalbalgebra, die Umkehrung gilt aber nicht (wähle z.B. $\mathcal{A}\times\mathcal{B}$ wie oben oder – noch einfacher – $X := \{0, 1\}$ und $\mathcal{H} := \{\varnothing, \{0\}, \{1\}\}$). Bemerke, dass $\varnothing \in \mathcal{H}$ gilt, sobald $\mathcal{H}$ nichtleer ist, denn: ist $A \in \mathcal{H}$, so ist $A^c = \bigcup_{i=0}^k A_i$ für eine endliche paarweise disjunkte Familie $(A_i)_{i=0}^k$ aus $\mathcal{H}$, insbesondere ist $\varnothing = A \cap A_0 \in \mathcal{H}$.

Proposition A.13 *Sei $\mathcal{H}$ eine nichtleere Mengenhalbalgebra auf einer Menge X. Dann ist die kleinste Mengenalgebra auf X, die $\mathcal{H}$ enthält, gleich der Menge $\mathcal{C}\langle\mathcal{H}\rangle$ der endlichen paarweise disjunkten Vereinigungen von Elementen von $\mathcal{H}$:*

$$\mathcal{C}\langle\mathcal{H}\rangle := \left\{ A \in \mathcal{P}(X) \mid A = \bigcup_{i=0}^{k} A_i \text{ wobei } k \in \mathbb{N}, A_i \in \mathcal{H}, A_i \cap A_j = \varnothing \;\forall\, 0 \le i \ne j \le k \right\}.$$

Beweis: Wir zeigen zuerst, dass $\mathcal{C}\langle\mathcal{H}\rangle$ eine Mengenalgebra auf X ist, die $\mathcal{H}$ enthält. Per Definition gilt $\mathcal{H} \subset \mathcal{C}\langle\mathcal{H}\rangle$ (wähle $k = 0$). Wegen $\varnothing \in \mathcal{H}$ (siehe oben) gilt also $\varnothing \in \mathcal{C}\langle\mathcal{H}\rangle$. Sind $A, B \in \mathcal{C}\langle\mathcal{H}\rangle$, so existieren paarweise disjunkte Elemente $A_0, \ldots, A_k \in \mathcal{H}$ bzw. $B_0, \ldots, B_l \in \mathcal{H}$ mit $A = \bigcup_{i=0}^k A_i$ bzw. $B = \bigcup_{j=0}^l B_j$. Damit gilt $A \cap B = (\bigcup_{i=0}^k A_i) \cap (\bigcup_{j=0}^l B_j) = \bigcup_{\substack{0\le i\le k\\ 0\le j\le l}} A_i \cap B_j$, wobei die Teilmengen $A_i \cap B_j$ paarweise disjunkt sind und in $\mathcal{H}$ liegen (da $\mathcal{H}$ eine Mengenhalbalgebra ist); daraus folgt $A \cap B \in \mathcal{C}\langle\mathcal{H}\rangle$. Sei nun $A = \bigcup_{i=0}^k A_i \in \mathcal{C}\langle\mathcal{H}\rangle$, wobei $A_0, \ldots, A_k \in \mathcal{H}$ paarweise disjunkt sind. Per Induktion über k zeigen wir, dass dann $A^c \in \mathcal{C}\langle\mathcal{H}\rangle$ gilt. Für $k = 0$ ist dies die Definition einer Mengenhalbalgebra. Angenommen, die Behauptung wäre für $k \in \mathbb{N}$ richtig. Ist $A = \bigcup_{i=0}^{k+1} A_i$, so gilt $A^c = \left((\bigcup_{i=0}^k A_i) \cup A_{k+1}\right)^c = (\bigcup_{i=0}^k A_i)^c \cap A_{k+1}^c$, wobei $(\bigcup_{i=0}^k A_i)^c = \bigcup_{j=0}^m B_j$ mit

paarweise disjunkten $B_0, \ldots, B_m \in \mathcal{H}$ nach Induktionsvoraussetzung. Wegen $A_{k+1} \in \mathcal{H}$ existieren paarweise disjunkte $C_0, \ldots, C_n \in \mathcal{H}$ mit $A_{k+1}^c = \bigcup_{p=0}^n C_p$. Insbesondere gilt $A^c = \bigcup_{\substack{0 \leq j \leq m \\ 0 \leq p \leq n}} B_j \cap C_p$, wobei die Teilmengen $B_j \cap C_p$ paarweise disjunkt sind und in $\mathcal{H}$ liegen. Dies beweist die Behauptung für $k+1$ und somit für alle $k \in \mathbb{N}$. Insgesamt erhalten wir $A^c \in \mathcal{C}\langle\mathcal{H}\rangle$. Die zwei letzten Eigenschaften liefern $A \cup B \in \mathcal{C}\langle\mathcal{H}\rangle$ für alle $A, B \in \mathcal{C}\langle\mathcal{H}\rangle$, denn: $A \cup B = (A^c \cap B^c)^c$ mit $A^c, B^c \in \mathcal{C}\langle\mathcal{H}\rangle$, somit $A^c \cap B^c \in \mathcal{C}\langle\mathcal{H}\rangle$ und daher $(A^c \cap B^c)^c \in \mathcal{C}\langle\mathcal{H}\rangle$. Dies beweist, dass $\mathcal{C}\langle\mathcal{H}\rangle$ eine Mengenalgebra auf X ist.
Dass $\mathcal{C}\langle\mathcal{H}\rangle$ die kleinste $\mathcal{H}$ enthaltende Mengenalgebra auf X ist, ist offensichtlich: ist $\mathcal{C}$ eine beliebige Mengenalgebra auf X, die $\mathcal{H}$ enthält, so enthält sie alle endlichen Vereinigungen von Elementen von $\mathcal{H}$, insbesondere enthält sie $\mathcal{C}\langle\mathcal{H}\rangle$. □

Wir kommen zum Hauptsatz, der es erlaubt, die Eindeutigkeit von Maßen zu beweisen.

Satz A.14 *Seien $(X, \mathcal{A})$ ein Messraum und $\mu_1, \mu_2 : \mathcal{A} \longrightarrow [0, \infty]$ zwei Maße.*

i) Angenommen, μ_1 und μ_2 seien endlich *mit $\mu_{1|_\mathcal{H}} = \mu_{2|_\mathcal{H}}$ für eine nichtleere Mengenhalbalgebra $\mathcal{H} \subset \mathcal{A}$. Dann gilt $\mu_{1|_{\mathcal{A}\langle\mathcal{H}\rangle}} = \mu_{2|_{\mathcal{A}\langle\mathcal{H}\rangle}}$, wobei $\mathcal{A}\langle\mathcal{H}\rangle$ die von $\mathcal{H}$ erzeugte σ-Algebra (auf X) bezeichnet.*

ii) Angenommen, es gäbe eine Mengenhalbalgebra $\mathcal{H} \subset \mathcal{A}$ mit $\mu_{1|_\mathcal{H}} = \mu_{2|_\mathcal{H}}$ sowie eine Folge $(H_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{H}$ mit $X = \bigcup_{n\in\mathbb{N}} H_n$ und $\mu_1(H_n)(= \mu_2(H_n)) < \infty$ für alle $n \in \mathbb{N}$ (insbesondere sind μ_1, μ_2 σ-endlich*). Dann gilt $\mu_{1|_{\mathcal{A}\langle\mathcal{H}\rangle}} = \mu_{2|_{\mathcal{A}\langle\mathcal{H}\rangle}}$.*

Beweis: Der Beweis folgt auf elementare Weise aus den vorigen Ergebnissen.
i) Gilt $\mu_{1|_\mathcal{H}} = \mu_{2|_\mathcal{H}}$, so gilt $\mu_1(\bigcup_{i=0}^k A_i) = \sum_{i=0}^k \mu_1(A_i) = \sum_{i=0}^k \mu_2(A_i) = \mu_2(\bigcup_{i=0}^k A_i)$ für jede endliche paarweise disjunkte Familie $(A_i)_{i=0}^k$ von Elementen von $\mathcal{H}$ (denn μ_1 und μ_2 sind Maße). Mit Proposition A.13 folgt $\mu_1(A) = \mu_2(A)$ für alle $A \in \mathcal{C}\langle\mathcal{H}\rangle$, d.h., $\mu_{1|_{\mathcal{C}\langle\mathcal{H}\rangle}} = \mu_{2|_{\mathcal{C}\langle\mathcal{H}\rangle}}$. Da aber μ_1 und μ_2 endlich sind, ist die Menge $\{A \in \mathcal{A} \,|\, \mu_1(A) = \mu_2(A)\}$ eine monotone Klasse auf X (siehe oben); wenn diese monotone Klasse $\mathcal{C}\langle\mathcal{H}\rangle$ enthält, enthält sie dann auch die davon erzeugte monotone Klasse nach Proposition A.10, d.h., $\mu_{1|_{\mathcal{M}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle}} = \mu_{2|_{\mathcal{M}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle}}$. Da aber $\mathcal{C}\langle\mathcal{H}\rangle$ eine Mengenalgebra ist, gilt $\mathcal{M}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle = \mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle$ nach Proposition A.11, insbesondere auch $\mu_{1|_{\mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle}} = \mu_{2|_{\mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle}}$. Es bleibt, zu bemerken, dass $\mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle = \mathcal{A}\langle\mathcal{H}\rangle$ gilt. Die Inklusion $\mathcal{A}\langle\mathcal{H}\rangle \subset \mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle$ folgt unmittelbar aus $\mathcal{H} \subset \mathcal{C}\langle\mathcal{H}\rangle$; die andere Inklusion $\mathcal{A}\langle\mathcal{C}\langle\mathcal{H}\rangle\rangle \subset \mathcal{A}\langle\mathcal{H}\rangle$ folgt aus $\mathcal{C}\langle\mathcal{H}\rangle \subset \mathcal{A}\langle\mathcal{H}\rangle$ (jede σ-Algebra ist bereits eine Mengenalgebra). Insgesamt erhalten wir $\mu_{1|_{\mathcal{A}\langle\mathcal{H}\rangle}} = \mu_{2|_{\mathcal{A}\langle\mathcal{H}\rangle}}$, was zu beweisen war.
ii) Der Beweis basiert auf Teil *i*) sowie auf folgenden Behauptungen.
Behauptung 1: *Es kann angenommen werden, dass die Familie $(H_n)_{n\in\mathbb{N}}$ paarweise disjunkt ist.*
Beweis der Behauptung: Sei $(H_n)_{n\in\mathbb{N}}$ wie in den Voraussetzungen (beachte insbesondere, dass $\mathcal{H}$ nicht leer sein kann). Man setzt $\overline{H}_0 := H_0$ und betrachtet die induktiv definierte Folge $\overline{H}_{n+1} := H_{n+1} \setminus (\bigcup_{i=0}^n \overline{H}_i)$, $n \in \mathbb{N}$. Per Konstruktion gilt $\overline{H}_i \cap \overline{H}_j = \varnothing$ für alle $i, j \in \mathbb{N}$ mit $i \neq j$. Es ist außerdem elementar, zu überprüfen, dass $\bigcup_{i=0}^n \overline{H}_i = \bigcup_{i=0}^n H_i$ für alle $n \in \mathbb{N}$ gilt, insbesondere auch $\bigcup_{n\in\mathbb{N}} \overline{H}_n = \bigcup_{n\in\mathbb{N}} H_n = X$. Bemerke, dass $\overline{H}_n \in \mathcal{C}\langle\mathcal{H}\rangle$ für alle $n \in \mathbb{N}$ gilt: für $n = 0$ ist es klar ($\mathcal{H} \subset \mathcal{C}\langle\mathcal{H}\rangle$); gilt $\overline{H}_n \in \mathcal{C}\langle\mathcal{H}\rangle$ für ein $n \in \mathbb{N}$, so gilt $\overline{H}_{n+1} = \underbrace{H_{n+1}}_{\in\mathcal{H}} \cap \underbrace{(\bigcup_{i=0}^n \overline{H}_i)^c}_{\in\mathcal{C}\langle\mathcal{H}\rangle} \in \mathcal{C}\langle\mathcal{H}\rangle$, insbesondere liefert diese Induktion (über n)

$\overline{H}_n \in \mathcal{C}\langle\mathcal{H}\rangle$ für alle $n \in \mathbb{N}$. Proposition A.13 liefert die Existenz, für jedes $n \in \mathbb{N}$, einer natürlichen Zahl $k(n) \geq 1$ und einer paarweise disjunkten endlichen Familie $(A_i^{(n)})_{i=0}^{k(n)}$ von Elementen von $\mathcal{H}$ mit $\overline{H}_n = \bigcup_{i=0}^{k(n)} A_i^{(n)}$. Schließlich gilt wegen $\mu_1(H_n) = \mu_2(H_n) < \infty$ und $\overline{H}_n \subset H_n$ (per Definition von H_n) auch $\mu_i(\overline{H}_n) \leq \mu_i(H_n) < \infty$ für $i = 1, 2$ und somit $\mu_1(A_i^{(n)}) = \mu_2(A_i^{(n)}) < \infty$ für alle $i \in \{0, \ldots, k(n)\}$ (beachte, dass $A_i^{(n)} \in \mathcal{H} \subset \mathcal{A}$ für alle i, n gilt). Daraus folgt $X = \bigcup_{n\in\mathbb{N}} \bigcup_{i=0}^{k(n)} A_i^{(n)}$, wobei $A_i^{(n)} \cap A_j^{(m)} = \varnothing$ für alle $(i, n) \neq (j, m)$ und mit $\mu_1(A_i^{(n)}) = \mu_2(A_i^{(n)}) < \infty$; da eine abzählbare Vereinigung von endlichen Mengen wieder abzählbar ist (Proposition 3.2), folgt die Behauptung 1. $\surd$

Von hier aus nehmen wir also an, dass die Familie $(H_n)_{n\in\mathbb{N}}$ paarweise disjunkt sei, also $H_n \cap H_m = \varnothing$ für alle $n \neq m$. Für jedes $n \in \mathbb{N}$ definieren wir $\mathcal{A}_n := \{A \cap H_n \,|\, A \in \mathcal{A}\} \subset \mathcal{P}(H_n)$.

Behauptung 2: *Die Menge $\mathcal{A}_n$ ist eine σ-Algebra auf H_n und es gilt $\mathcal{A}_n = \{A \in \mathcal{A} \,|\, A \subset H_n\}$.*

Beweis der Behauptung: Die Tatsache, dass $\mathcal{A}_n$ eine σ-Algebra auf H_n ist, ist elementar zu beweisen (*Übungsaufgabe*). Für die letzte Identität ist die Inklusion $\mathcal{A}_n \supset \{A \in \mathcal{A} \,|\, A \subset H_n\}$ trivial (wegen $A = A \cap H_n$ falls $A \subset H_n$) während die andere Inklusion $\mathcal{A}_n \subset \{A \in \mathcal{A} \,|\, A \subset H_n\}$ aus $H_n \in \mathcal{A}$ folgt. $\surd$

Die Maße $\mu_{1|_{\mathcal{A}_n}}, \mu_{2|_{\mathcal{A}_n}}$ sind nun wegen $\mu_1(H_n) = \mu_2(H_n) < \infty$ endliche Maße auf $\mathcal{A}_n$ mit $\mu_1(H) = \mu_2(H)$ für jedes $H \in \mathcal{H}$ mit $H \subset H_n$. Analog wie in der Behauptung 2 ist es elementar, $\{H \in \mathcal{H} \,|\, H \subset H_n\} = \mathcal{H} \cap H_n := \{H \cap H_n \,|\, H \in \mathcal{H}\}$ zu beweisen (benutze $H \cap H_n \in \mathcal{H}$); desweiteren ist $\mathcal{H} \cap H_n$ eine Mengenhalbalgebra auf H_n (auch elementar, wird als Übungsaufgabe gelassen) und ist auch nicht leer. Teil $i)$ aus Satz A.14 impliziert dann $\mu_{1|_{\mathcal{A}\langle\mathcal{H}\cap H_n\rangle}} = \mu_{2|_{\mathcal{A}\langle\mathcal{H}\cap H_n\rangle}}$, wobei $\mathcal{A}\langle\mathcal{H} \cap H_n\rangle$ die von der Mengenhalbalgebra $\mathcal{H} \cap H_n$ erzeugte σ-Algebra auf H_n ist.

Behauptung 3: *Es gilt $\mathcal{A}\langle\mathcal{H} \cap H_n\rangle = \mathcal{A}\langle\mathcal{H}\rangle \cap H_n := \{A \cap H_n \,|\, A \in \mathcal{A}\langle\mathcal{H}\rangle\}$.*

Beweis der Behauptung: Die Inklusion $\mathcal{A}\langle\mathcal{H} \cap H_n\rangle \subset \mathcal{A}\langle\mathcal{H}\rangle \cap H_n$ folgt aus $\mathcal{H} \cap H_n \subset \mathcal{A}\langle\mathcal{H}\rangle \cap H_n$ (wegen $\mathcal{H} \subset \mathcal{A}\langle\mathcal{H}\rangle$ und per Definition von $\mathcal{A}\langle\mathcal{H}\rangle \cap H_n$) und der Tatsache, dass $\mathcal{A}\langle\mathcal{H}\rangle \cap H_n$ eine σ-Algebra auf H_n ist (*Übungsaufgabe*). Für die andere Inklusion betrachtet man die Menge $\mathcal{Q}_n := \{A \subset X \,|\, A \cap H_n \in \mathcal{A}\langle\mathcal{H} \cap H_n\rangle\} \subset \mathcal{P}(X)$ und zeigt, dass $\mathcal{Q}_n$ eine σ-Algebra auf X ist, die $\mathcal{H}$ enthält (*Übungsaufgabe*), somit enthält $\mathcal{Q}_n$ auch $\mathcal{A}\langle\mathcal{H}\rangle$, d.h., es gilt $A \cap H_n \in \mathcal{A}\langle\mathcal{H} \cap H_n\rangle$ für alle $A \in \mathcal{A}\langle\mathcal{H}\rangle$, was zu beweisen war. $\surd$

Sei nun $A \in \mathcal{A}\langle\mathcal{H}\rangle$ beliebig, dann gilt $A = \bigcup_{n\in\mathbb{N}}(A \cap H_n)$, wobei $A \cap H_n \in \mathcal{A}\langle\mathcal{H}\rangle \cap H_n = \mathcal{A}\langle\mathcal{H} \cap H_n\rangle$ (nach Behauptung 3) und $A \cap H_n \in \mathcal{A}$. Nach Voraussetzung ist die Familie $(A \cap H_n)_{n\in\mathbb{N}}$ paarweise disjunkt (siehe oben), somit folgt aus der σ-Additivität von μ_1 und μ_2:

$$
\begin{aligned}
\mu_1(A) &= \sum_{n=0}^{\infty} \mu_1(A \cap H_n) \\
&= \sum_{n=0}^{\infty} \mu_2(A \cap H_n) \\
&= \mu_2(A),
\end{aligned}
$$

was zu beweisen war. $\square$

Insbesondere bekommen wir die *Eindeutigkeit* des Produktmaßes aus Proposition 3.60:

Korollar A.15 *Seien* $(X, \mathcal{A}, \mu)$ *und* $(Y, \mathcal{B}, \nu)$ *Maßräume mit* μ *und* ν σ-endlich. *Dann existiert* höchstens ein *Maß* ρ *auf* $\mathcal{A} \otimes \mathcal{B}$ *mit*

$$\rho(A \times B) = \mu(A)\nu(B) \qquad \text{für alle } (A, B) \in \mathcal{A} \times \mathcal{B}.$$

Beweis: Sind ρ und ρ' zwei Maße auf $\mathcal{A} \otimes \mathcal{B}$ mit $\rho(A \times B) = \mu(A)\nu(B) = \rho'(A \times B)$ für alle $(A, B) \in \mathcal{A} \times \mathcal{B}$, so gilt insbesondere $\rho_{|\mathcal{A}\times\mathcal{B}} = \rho'_{|\mathcal{A}\times\mathcal{B}}$, wobei $\mathcal{A} \times \mathcal{B}$ eine Mengenhalbalgebra ist, siehe oben. Da außerdem μ und ν σ-endlich sind, existieren abzählbare Familien $(A_n)_{n\in\mathbb{N}}$ aus $\mathcal{A}$ und $(B_m)_{m\in\mathbb{N}}$ aus $\mathcal{B}$ mit $\bigcup_{n\in\mathbb{N}} A_n = X$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$ (bzw. $\bigcup_{m\in\mathbb{N}} B_m = X$ und $\nu(B_m) < \infty$ für alle $m \in \mathbb{N}$). Es folgt $X \times Y = \bigcup_{m,n\in\mathbb{N}} A_n \cup B_m$ mit $A_n \times B_m \in \mathcal{A} \times \mathcal{B}$ und $\rho(A_n \times B_m) = \rho'(A_n \times B_m) = \mu(A_n)\nu(B_m) < \infty$, wobei die Familie $(A_n \times B_m)_{m,n\in\mathbb{N}}$ abzählbar ist (Proposition 3.2). Da per Definition die von $\mathcal{A} \times \mathcal{B}$ erzeugte σ-Algebra auf $X \times Y$ die σ-Algebra $\mathcal{A} \otimes \mathcal{B}$ ist, liefert Satz A.14.*ii*) die Identität $\rho_{|\mathcal{A}\otimes\mathcal{B}} = \rho'_{|\mathcal{A}\otimes\mathcal{B}}$, d.h., $\rho = \rho'$. $\square$

Nun wollen wir die Existenz eines (notwendigerweise eindeutigen) Produktmaßes zeigen. Dafür müssen wir Lemma 3.62 beweisen.

Lemma A.16 *Seien* $(X, \mathcal{A}, \mu)$ *und* $(Y, \mathcal{B}, \nu)$ *Maßräume und* $E \in \mathcal{A} \otimes \mathcal{B}$. *Dann gilt:*

i) Für alle $x \in X$ *ist* $E_x := \{y \in Y \,|\, (x, y) \in E\} \subset Y$ *Element von* $\mathcal{B}$*; analog ist für jedes* $y \in Y$ *die Teilmenge* $E_y := \{x \in X \,|\, (x, y) \in E\} \subset X$ *Element von* $\mathcal{A}$.

ii) Ist ν σ*-endlich, so ist die Abbildung* $X \longrightarrow [0, \infty]$, $x \longmapsto \nu(E_x)$ *messbar. Analog gilt die Aussage für* $Y \longrightarrow [0, \infty]$, $y \longmapsto \mu(E_y)$*, falls* μ σ*-endlich ist.*

Beweis: Beide Aussagen folgen wiederum aus dem Prinzip der guten Mengen.
i) Man fixiere $x \in X$ und betrachte $\mathcal{K}_x := \{E \subset X \times Y \,|\, E_x \in \mathcal{B}\} \subset \mathcal{P}(X \times Y)$. Wenn wir zeigen, dass $\mathcal{K}_x$ eine $\mathcal{A} \times \mathcal{B}$ enthaltende σ-Algebra auf $X \times Y$ ist, folgt insbesondere $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{K}_x$, was die Behauptung liefert (da $x \in X$ beliebig gewählt wurde). Dass $\mathcal{K}_x$ eine σ-Algebra auf $X \times Y$ ist, ist aber klar (benutze lediglich, dass $\mathcal{B}$ eine σ-Algebra ist sowie $(E^c)_x = (E_x)^c$ und $(\bigcup_{n\in\mathbb{N}} E_n)_x = \bigcup_{n\in\mathbb{N}} (E_n)_x$, was auch elementar zu zeigen ist). Ist $E = A \times B \in \mathcal{A} \times \mathcal{B}$, so gilt $E_x = B \in \mathcal{B}$ falls $x \in A$ und $E_x = \varnothing \in \mathcal{B}$ falls $x \in A^c$, insbesondere immer $E_x \in \mathcal{B}$, d.h., $E \in \mathcal{K}_x$. Dies beweist $\mathcal{A} \times \mathcal{B} \subset \mathcal{K}_x$ und somit $E_x \in \mathcal{B}$ für alle $x \in X$. Bemerke, dass wir dabei die Tatsache nicht verwendet haben, dass $\mathcal{A}$ eine σ-Algebra ist. Der Fall von E_y (mit $y \in Y$) ist analog (nutze diesmal nur, dass $\mathcal{A}$ eine σ-Algebra ist). Dies zeigt *i*).
ii) Nach *i*) ist diese Abbildung erstmal wohldefiniert. Wir fangen mit dem Fall an, wo ν *endlich* ist. Betrachte

$$\mathcal{E} := \{E \subset X \times Y \,|\, E_x \in \mathcal{B} \text{ für alle } x \in X \text{ und } x \mapsto \nu(E_x) \text{ ist messbar}\} \subset \mathcal{P}(X \times Y).$$

Wir zeigen, dass $\mathcal{E}$ eine $\mathcal{C}\langle\mathcal{A} \times \mathcal{B}\rangle$ enthaltende monotone Klasse auf $X \times Y$ ist; daraus folgt $\mathcal{E} \supset \mathcal{M}\langle\mathcal{C}\langle\mathcal{A} \otimes \mathcal{B}\rangle\rangle$, d.h., $\mathcal{E} \supset \mathcal{A}\langle\mathcal{C}\langle\mathcal{A} \times \mathcal{B}\rangle\rangle$ (Proposition A.11), was zusammen mit $\mathcal{A}\langle\mathcal{C}\langle\mathcal{A} \times \mathcal{B}\rangle\rangle = \mathcal{A}\langle\mathcal{A} \times \mathcal{B}\rangle$ ("$\supset$" wegen $\mathcal{C}\langle\mathcal{A} \times \mathcal{B}\rangle \supset \mathcal{A} \times \mathcal{B}$, "$\subset$" wegen $\mathcal{C}\langle\mathcal{A} \times \mathcal{B}\rangle \subset \mathcal{A}\langle\mathcal{A} \times \mathcal{B}\rangle$, siehe Beweis von Satz A.14.*i*)) und $\mathcal{A}\langle\mathcal{A} \times \mathcal{B}\rangle = \mathcal{A} \otimes \mathcal{B}$ (per Definition) die Behauptung liefert. Ist zuerst $E = A \times B \in \mathcal{A} \times \mathcal{B}$, so gilt $E_x \in \mathcal{B}$ für alle $x \in X$ (siehe oben) sowie $\nu(E_x) = \nu(B)$ falls $x \in A$ und $\nu(E_x) = \nu(\varnothing) = 0$ falls $x \in A^c$, d.h., die Abbildung $x \mapsto \nu(E_x)$ stimmt mit $\nu(B) \cdot \chi_A : X \longrightarrow [0, \infty]$ überein, wobei $\chi_A : X \longrightarrow \{0, 1\}$ die charakteristische Funktion von A ist. Wegen $A \in \mathcal{A}$ ist aber

χ_A – und somit auch $\nu(B) \cdot \chi_A$ – messbar, siehe Beispiel 3.21.3. Dies zeigt $E \in \mathcal{E}$ und somit $\mathcal{A} \times \mathcal{B} \subset \mathcal{E}$. Sind $E, F \in \mathcal{E}$ disjunkt, so gilt $(E \cup F)_x = E_x \cup F_x$ für alle $x \in X$, wobei $E_x \cap F_x = \varnothing$ (wegen $E \cap F = \varnothing$) und $E_x \cup F_x \in \mathcal{B}$ (wegen $\mathcal{B}$ σ-Algebra); außerdem ist $x \mapsto \nu(E_x \cup F_x) = \nu(E_x) + \nu(F_x)$ als Summe zweier messbarer Funktionen wieder messbar, somit gilt $E \cup F \in \mathcal{E}$. Dies zeigt insbesondere, dass $\mathcal{E}$ alle endlichen paarweise disjunkten Vereinigungen von Elementen der (wohl nichtleeren) Mengenhalbalgebra $\mathcal{A} \times \mathcal{B}$ enthält und somit auch $\mathcal{C}\langle \mathcal{A} \times \mathcal{B} \rangle$ nach Proposition A.13. Nun beweisen wir, dass $\mathcal{E}$ eine monotone Klasse auf $X \times Y$ ist. Sei $(E_n)_{n\in\mathbb{N}}$ eine beliebige monoton wachsende Folge von Elementen von $\mathcal{E}$. Dann ist $(\bigcup_{n\in\mathbb{N}} E_n)_x = \bigcup_{n\in\mathbb{N}} (E_n)_x \in \mathcal{B}$ (da $\mathcal{B}$ eine σ-Algebra ist) für alle $x \in X$. Da für jedes $x \in X$ die Folge $((E_n)_x)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{B}$ monoton wachsend ist, gilt $\nu(\bigcup_{n\in\mathbb{N}} (E_n)_x) = \lim\limits_{n\to\infty} \nu((E_n)_x)$ nach Proposition 3.12. Wegen $E_n \in \mathcal{E}$ ist aber jede Abbildung $x \mapsto \nu((E_n)_x)$ messbar, insbesondere liefert Proposition 3.25, dass $x \mapsto \nu(\bigcup_{n\in\mathbb{N}} (E_n)_x)$ als punktweiser Grenzwert einer Folge von messbaren Funktionen wieder messbar ist. Dies zeigt somit $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{E}$. Für eine monoton fallende Folge $(E_n)_{n\in\mathbb{N}}$ von Elementen von $\mathcal{E}$ ist wiederum $(\bigcap_{n\in\mathbb{N}} E_n)_x = \bigcap_{n\in\mathbb{N}} (E_n)_x \in \mathcal{B}$ für alle $x \in X$. Wegen $(E_{n+1})_x \subset (E_n)_x$ für alle $n \in \mathbb{N}, x \in X$ und weil ν *endlich* ist, gilt $\nu(\bigcap_{n\in\mathbb{N}} (E_n)_x) = \lim\limits_{n\to\infty} \nu((E_n)_x)$, wieder nach Proposition 3.12. Proposition 3.25 impliziert, dass $x \mapsto \nu(\bigcap_{n\in\mathbb{N}} (E_n)_x)$ dann als punktweiser Grenzwert einer Folge von messbaren Funktionen wieder messbar ist, somit gilt $\bigcap_{n\in\mathbb{N}} E_n \in \mathcal{E}$. Für ein endliches Maß ν ist somit $\mathcal{E}$ eine $\mathcal{C}\langle \mathcal{A} \times \mathcal{B} \rangle$ enthaltende monotone Klasse auf $X \times Y$ ist, was zu beweisen war.
Im Falle, wo ν lediglich σ-endlich ist, schreiben wir $Y = \bigcup_{n\in\mathbb{N}} Y_n$ mit $Y_n \in \mathcal{B}$ und $\nu(Y_n) < \infty$. O.B.d.A. sei die Familie $(Y_n)_{n\in\mathbb{N}}$ monoton wachsend (sonst ersetze Y_n durch $\overline{Y}_n := \bigcup_{k=0}^n Y_k \in \mathcal{B}$ für alle $n \in \mathbb{N}$). Dann ist, für jedes $n \in \mathbb{N}$, die Abbildung $\nu_n : \mathcal{B} \longrightarrow [0, \infty]$, $B \mapsto \nu(B \cap Y_n)$, ein Maß auf $\mathcal{B}$ (*Übungsaufgabe*) mit $\nu_n(Y) = \nu(Y_n) < \infty$. Sei nun $E \in \mathcal{A} \otimes \mathcal{B}$ beliebig. Nach dem obigen Fall ist dann, für jedes $n \in \mathbb{N}$, die Abbildung $x \mapsto \nu_n(E_x)$ messbar. Außerdem gilt $E_x = \bigcup_{n\in\mathbb{N}} E_x \cap Y_n$ für alle $x \in X$, wobei die Folge $(E_x \cap Y_n)_{n\in\mathbb{N}}$ aus $\mathcal{B}$ monoton wachsend ist. Proposition 3.12 liefert $\nu(E_x) = \lim\limits_{n\to\infty} \nu(E_x \cap Y_n) = \lim\limits_{n\to\infty} \nu_n(E_x)$ für alle $x \in X$. Wir erhalten $x \mapsto \nu(E_x)$ somit als punktweisen Grenzwert einer Folge von messbaren Funktionen, insbesondere impliziert Proposition 3.25, dass $x \mapsto \nu(E_x)$ messbar ist. Dies beweist Lemma A.16. $\square$

Beachte, dass Lemma A.16 Lemma 3.62 beweist. Nun sind wir in der Lage, die Existenz eines (des) Produktmaßes zu zeigen.

Satz A.17 *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Maßräume mit ν* σ-endlich. *Dann definiert*

$$\rho : \mathcal{A} \otimes \mathcal{B} \longrightarrow [0, \infty], \qquad E \longmapsto \int_X \nu(E_x) d\mu(x),$$

ein Maß auf $\mathcal{A} \otimes \mathcal{B}$ mit $\rho(A \times B) = \mu(A)\nu(B)$ für alle $A \times B \in \mathcal{A} \times \mathcal{B}$. Desweiteren ist ρ σ-endlich, sobald μ auch σ-endlich ist.

Beweis: Nach Lemma A.16 ist ρ wohldefiniert, da für alle $E \in \mathcal{A} \otimes \mathcal{B}$ die Abbildung $x \mapsto \nu(E_x)$ wohldefiniert, messbar und nichtnegativ ist. Wegen $\varnothing_x = \varnothing$ für alle $x \in X$ ist $x \mapsto \nu(E_x)$ die Nullabbildung und somit gilt $\rho(\varnothing) = \int_X 0 d\mu = 0$. Sei $(E_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Elementen von $\mathcal{A} \otimes \mathcal{B}$, dann gilt

$$\begin{aligned}
\rho(\bigcup_{n\in\mathbb{N}} E_n) &= \int_X \nu((\bigcup_{n\in\mathbb{N}} E_n)_x) d\mu(x) \\
&= \int_X \nu(\bigcup_{n\in\mathbb{N}} (E_n)_x) d\mu(x) \qquad \text{mit } (E_n)_x \cap (E_m)_x = \varnothing \;\; \forall\, n \neq m \\
&= \int_X \sum_{n=0}^{\infty} \nu((E_n)_x) d\mu(x) \\
&= \sum_{n=0}^{\infty} \int_X \nu((E_n)_x) d\mu(x) \qquad \text{nach Korollar 3.35} \\
&= \sum_{n=0}^{\infty} \rho(E_n).
\end{aligned}$$

Dies beweist, dass ρ ein Maß auf $\mathcal{A} \otimes \mathcal{B}$ ist. Ist $E = A \times B \in \mathcal{A} \times \mathcal{B}$, so ist $x \mapsto \nu(E_x)$ die Abbildung $\nu(B) \cdot \chi_A$ (siehe Beweis von Lemma A.16), insbesondere gilt

$$\rho(A \times B) = \int_X \nu(B) \cdot \chi_A d\mu = \nu(B) \int_X \chi_A d\mu = \nu(B)\mu(A).$$

Ist nun μ auch σ-endlich, so existieren abzählbare Familien $(A_n)_{n\in\mathbb{N}}$ aus $\mathcal{A}$ und $(B_m)_{m\in\mathbb{N}}$ aus $\mathcal{B}$ mit $\bigcup_{n\in\mathbb{N}} A_n = X$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$ (bzw. $\bigcup_{m\in\mathbb{N}} B_m = X$ und $\nu(B_m) < \infty$ für alle $m \in \mathbb{N}$). Es folgt $X \times Y = \bigcup_{m,n\in\mathbb{N}} A_n \cup B_m$ mit $A_n \times B_m \in \mathcal{A} \times \mathcal{B}$ und $\rho(A_n \times B_m) = \mu(A_n)\nu(B_m) < \infty$, wobei die Familie $(A_n \times B_m)_{m,n\in\mathbb{N}}$ abzählbar ist (Proposition 3.2). Dies zeigt, dass ρ dann ebenfalls σ-endlich ist. □

Mit Satz A.17 und Korollar A.15 erhalten wir die Existenz und Eindeutigkeit des Produktmaßes und somit Proposition 3.60.

A.2.2 Charakterisierung des Lebesgue-Maßes

Das n-dimensionale Lebesgue-Maß besitzt bestimmte Eigenschaften, die u.a. zur Transformationsformel führen. Erstens kann dieses Maß durch seine *Translationsinvarianz* und das Maß des Einheitsquaders charakterisiert werden.

Proposition A.18 *Für ein beliebiges $n \in \mathbb{N}$ ist das n-dimensionale Borel-Lebesgue-Maß μ_n das einzige Maß auf $\mathcal{B}(\mathbb{R}^n)$ mit:*

i) μ_n ist translationsinvariant, d.h., $\mu_n(x + B) = \mu_n(B)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$ und $x \in \mathbb{R}^n$,

ii) $\mu_n([0, 1[^n) = 1$.

Beweis: Nach Konstruktion des Maßes μ_n gilt $\mu_n(\prod_{i=1}^n [a_i, b_i[) = \prod_{i=1}^n (b_i - a_i)$, insbesondere $\mu_n([0, 1[^n) = 1$. Um zu beweisen, dass μ_n translationsinvariant ist, fixieren wir ein $x \in \mathbb{R}^n$ und betrachten die Abbildung $\widetilde{\mu}_n : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$, $\widetilde{\mu}_n(B) := \mu_n(x+B)$. Dann ist $\widetilde{\mu}_n$ wohldefiniert (die Abbildung $\mathbb{R}^n \xrightarrow{x+\mathrm{Id}} \mathbb{R}^n$ ist ein Homöomorphismus, insbesondere gilt

$(x+\mathrm{Id})(B) \in \mathcal{B}(\mathbb{R}^n)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$)) und ist ein Maß auf $\mathcal{B}(\mathbb{R}^n)$ (*Übungsaufgabe*). Desweiteren gilt, für alle $a_i \leq b_i \in \mathbb{R}$, $1 \leq i \leq n$,

$$\widetilde{\mu}_n(\prod_{i=1}^n [a_i,b_i[) = \mu_n(x+\prod_{i=1}^n [a_i,b_i[) = \mu_n(\prod_{i=1}^n [x_i+a_i,x_i+b_i[) = \prod_{i=1}^n (b_i-a_i) = \mu_n(\prod_{i=1}^n [a_i,b_i[),$$

wobei $x=(x_1,\ldots,x_n)$. Da $\mathcal{B}(\mathbb{R}^n)$ durch Produktmengen der Form $\prod_{i=1}^n [a_i,b_i[$ erzeugt wird, folgt aus Satz A.14 die Identität $\widetilde{\mu}_n = \mu_n$. Insbesondere gilt $\mu_n(x+B)=\mu_n(B)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$, d.h., μ_n ist translationsinvariant.
Nun beweisen wir die Eindeutigkeit. Sei $\mu : \mathcal{B}(\mathbb{R}^n) \longrightarrow [0,\infty]$ ein beliebiges Maß, welches $i)$ und $ii)$ erfüllt. Wie vorher reicht es, $\mu(\prod_{i=1}^n [a_i,b_i[) = \prod_{i=1}^n (b_i-a_i)$ für alle $a_i \leq b_i \in \mathbb{R}$, $1 \leq i \leq n$, zu beweisen, um $\mu=\mu_n$ zu zeigen. Seien $2n$ beliebige reelle Zahlen $a_i \leq b_i$, $1 \leq i \leq n$. Da μ translationsinvariant ist, können wir o.B.d.A. annehmen, dass $a_1 = \ldots = a_n = 0$ gilt (sonst ersetze $\prod_{i=1}^n [a_i,b_i[$ durch $-(a_1,\ldots,a_n)+\prod_{i=1}^n [a_i,b_i[= \prod_{i=1}^n [0,b_i-a_i[$). Im Fall, wo $b_1,\ldots,b_n \in \mathbb{Q}$, schreiben wir $b_i = \frac{p_i}{q_i}$ mit $p_i \in \mathbb{N}$ und $q_i \in \mathbb{N}\setminus\{0\}$, $1 \leq i \leq n$. Dann gilt $\prod_{i=1}^n [0,b_i[= \bigcup_{j_1=1}^{p_1} \ldots \bigcup_{j_n=1}^{p_n} [\frac{j_1-1}{q_1},\frac{j_1}{q_1}[\times \ldots \times [\frac{j_n-1}{q_n},\frac{j_n}{q_n}[$, wobei die Vereinigung auf der rechten Seite disjunkt ist. Nach der Translationsinvarianz von μ gilt dann $\mu(\prod_{i=1}^n [0,b_i[) = p_1 \cdot \ldots \cdot p_n \cdot \mu([0,\frac{1}{q_1}[\times \ldots \times [0,\frac{1}{q_n}[)$. Nun können wir auch $[0,1[^n$ als disjunkte Vereinigung schreiben: es gilt $[0,1[^n = \bigcup_{j_1=1}^{q_1} \ldots \bigcup_{j_n=1}^{q_n} [\frac{j_1-1}{q_1},\frac{j_1}{q_1}[\times \ldots \times [\frac{j_n-1}{q_n},\frac{j_n}{q_n}[$, somit (wieder mit der Translationsinvarianz von μ) gilt $1 = q_1 \cdot \ldots \cdot q_n \cdot \mu([0,\frac{1}{q_1}[\times \ldots \times [0,\frac{1}{q_n}[)$. Daraus folgt $\mu([0,\frac{1}{q_1}[\times \ldots \times [0,\frac{1}{q_n}[) = \frac{1}{q_1 \ldots q_n}$ und somit auch $\mu(\prod_{i=1}^n [0,b_i[) = \prod_{i=1}^n \frac{p_i}{q_i} = \prod_{i=1}^n b_i$, was zu beweisen war. Im allgemeinen Fall, wo $b_1,\ldots,b_n \in \mathbb{R}$, existieren n monoton wachsende rationale Folgen $(u_{k,i})_{k\in\mathbb{N}}$ mit $u_{k,i} \underset{k\to\infty}{\longrightarrow} b_i$ für alle $1 \leq i \leq n$. Dann ist $\prod_{i=1}^n [0,b_i[= \bigcup_{k_1=0}^{\infty} \ldots \bigcup_{k_n=0}^{\infty} [0,u_{k_1,1}[\times \ldots \times [0,u_{k_n,n}[$. Da die Vereinigung über k_1 monoton wachsend ist, folgt aus Proposition 3.12, dass $\mu(\prod_{i=1}^n [0,b_i[) = \lim_{k_1\to\infty} \mu([0,u_{k_1,1}[\times \bigcup_{k_2=0}^{\infty} \ldots \bigcup_{k_n=0}^{\infty} [0,u_{k_2,2}[\times \ldots \times [0,u_{k_n,n}[)$. Sukzessiv auf $k_1,k_2,\ldots,k_n$ angewendet liefert dieses Argument, zusammen mit der Diskussion des ersten Falles,

$$\begin{aligned}
\mu(\prod_{i=1}^n [0,b_i[) &= \lim_{k_1\to\infty} \ldots \lim_{k_n\to\infty} \mu([0,u_{k_1,1}[\times \ldots \times [0,u_{k_n,n}[) \\
&= \lim_{k_1\to\infty} \ldots \lim_{k_n\to\infty} u_{k_1,1} \cdot \ldots \cdot u_{k_n,n} \\
&= \lim_{k_1\to\infty} \ldots \lim_{k_{n-1}\to\infty} u_{k_1,1} \cdot \ldots \cdot u_{k_{n-1},n-1} \cdot b_n \\
&\vdots \quad \vdots \\
&= b_1 \cdot \ldots \cdot b_n,
\end{aligned}$$

was zu beweisen war. □

Das Lebesgue-Maß bleibt auch erhalten unter Isometrien.

Proposition A.19 *Für ein beliebiges $n \in \mathbb{N}$ ist das n-dimensionale Borel-Lebesgue-Maß μ_n invariant unter Isometrien, d.h., es gilt $\mu_n(F(B)) = \mu_n(B)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$ und alle Isometrien F des $\mathbb{R}^n$.*

Beweis: Jede Isometrie F des $\mathbb{R}^n$ ist der Form $F(x) = A\cdot x + b$, wobei $A \in M_{n\times n}(\mathbb{R})$ eine Orthogonalmatrix und $b \in \mathbb{R}^n$ ein Vektor sind. Da μ_n nach Proposition A.18 translationsinvariant ist, kann $b=0$ angenommen werden. Da außerdem jede Isometrie als

Nacheinanderausführung von endlich vielen Orthogonalspiegelungen geschrieben werden kann (siehe Linearalgebra I oder Mathematische Methoden der Physik), können wir auch annehmen, dass A die Matrix einer Orthogonalspiegelung ist. Nun betrachten wir die Abbildung $\widetilde{\mu}_n : \mathcal{B}(\mathbb{R}^n) \longrightarrow [0, \infty]$, $B \mapsto \mu_n(A \cdot B)$. Bemerke, dass $\widetilde{\mu}_n$ wohldefiniert ist, weil $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ ein Homöomorphismus ist (somit gilt $A \cdot B \in \mathcal{B}(\mathbb{R}^n)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$). Dann ist $\widetilde{\mu}_n$ ein Maß auf $\mathcal{B}(\mathbb{R}^n)$ (*Übungsaufgabe*) und ist wegen der Translationsinvarianz von μ_n translationsinvariant: für alle $x \in \mathbb{R}^n$ und $B \in \mathcal{B}(\mathbb{R}^n)$ gilt $\widetilde{\mu}_n(x + B) = \mu_n(A \cdot (x + B)) = \mu_n(A \cdot x + A \cdot B) = \mu_n(A \cdot B) = \widetilde{\mu}_n(B)$. Aus Proposition A.18 folgt, dass $\widetilde{\mu}_n$ proportional zu μ_n ist: für $\alpha := \widetilde{\mu}_n([0, 1[^n)$ gilt $\widetilde{\mu}_n = \alpha \cdot \mu_n$, denn $\frac{1}{\alpha}\widetilde{\mu}_n$ translationsinvariant mit $\frac{1}{\alpha}\widetilde{\mu}_n([0, 1[^n) = 1$ (bemerke, dass $\alpha \in]0, \infty[$ wegen A Homöomorphismus und $A([0, 1[^n) \subset [0, \sqrt{n}[^n)$. Schließlich gilt $\alpha = 1$ wegen $A^2 = \mathrm{I}_n$: es gilt $\mu_n(B) = \mu_n(A^2 \cdot B) = \mu_n(A \cdot (A \cdot B)) = \alpha \cdot \mu_n(A \cdot B) = \alpha^2 \mu_n(B)$ für alle $B \in \mathcal{B}(\mathbb{R}^n)$, somit $\alpha^2 = 1$ (mit z.B. $B := [0, 1[^n$). Dies beweist die Proposition. □

Bemerke, dass sowohl Proposition A.18 als Proposition A.19 für λ_n statt μ_n gelten: für jede Isometrie F des $\mathbb{R}^n$ definiert $L \mapsto \lambda_n(F(L))$ ein *vollständiges* Maß auf $\mathcal{L}_n$ (*Übungsaufgabe*); wenn dieses Maß mit μ_n auf $\mathcal{B}(\mathbb{R}^n)$ übereinstimmt, dann stimmt es automatisch mit λ_n auf $\mathcal{L}_n$ überein, siehe Proposition 3.18. Dabei ist zu beachten, dass $N \subset \mathbb{R}^n$ genau dann eine μ_n-Nullmenge ist, wenn $F(N)$ eine μ_n-Nullmenge ist (somit gilt $F(L) \in \mathcal{L}_n$ für alle $L \in \mathcal{L}_n$).

Anhang B

Mehr über Kompaktheit

In diesem Abschnitt sammeln wir einige wichtige Eigenschaften kompakter metrischer Räume, die wir im Laufe der Vorlesung gebraucht haben aber nicht vorstellen konnten.

B.1 Kompaktheit und offene Überdeckungen

Nach Definition ist ein metrischer Raum genau dann kompakt, wenn jede Folge eine konvergente Teilfolge besitzt. Es gibt eine äquivalente Definition in Termen von *offenen Überdeckungen*.

Satz B.1 *Ein metrischer Raum (X, d) ist genau dann kompakt, wenn* jede *offene Überdeckung von X eine* endliche *Teilüberdeckung besitzt, d.h., wenn für jede Familie $(U_i)_{i\in I}$ von offenen Teilmengen von X mit $X = \bigcup\limits_{i\in I} U_i$ eine endliche Teilmenge $\{i_1, \ldots, i_k\} \subset I$ existiert mit $X = \bigcup\limits_{1\leq j\leq k} U_{i_j}$.*

Beweis: Angenommen, jede offene Überdeckung von X besitze eine endliche Teilüberdeckung. Zuerst beweisen wir eine Eigenschaft von X, die das Analogon von Proposition 2.26.*iii*) ist.
Behauptung: *Sei (X, d) ein metrischer Raum so, dass jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt. Sei $(K_n)_{n\in\mathbb{N}}$ eine beliebige monoton fallende (d.h., $K_{n+1} \subset K_n$ für alle $n \in \mathbb{N}$) Folge von* nichtleeren *abgeschlossenen Teilmengen von X. Dann ist $\bigcap\limits_{n\in\mathbb{N}} K_n$ eine* nichtleere *Teilmenge von X.*
Beweis der Behauptung: Die Teilmenge $\bigcap\limits_{n\in\mathbb{N}} K_n$ von X ist abgeschlossen als (unendlicher) Durchschnitt abgeschlossener Teilmengen von X. Angenommen, es gelte $\bigcap\limits_{n\in\mathbb{N}} K_n = \varnothing$. Wegen $X \setminus (\bigcap\limits_{n\in\mathbb{N}} K_n) = \bigcup\limits_{n\in\mathbb{N}} (X \setminus K_n)$ gilt dann

$$\bigcup_{n\in\mathbb{N}} (X \setminus K_n) = X \setminus \varnothing = X,$$

mit $X \setminus K_n$ offen (wegen K_n abgeschlossen), d.h., die Menge $\{X \setminus K_n,\, n \in \mathbb{N}\}$ bildet eine offene Überdeckung von X. Nach Voraussetzung existieren endlich viele $K_{n_1}, \ldots, K_{n_p}$ s.d.

$$X = \bigcup_{i=1}^{p}(X \setminus K_{n_i}).$$

Setze $m := \max(n_1, \ldots, n_p)$, dann gilt $X \setminus K_{n_i} \subset X \setminus K_m$ (weil $(K_n)_{n\in\mathbb{N}}$ monoton fallend ist), und somit $X = \bigcup_{i=1}^{p}(X \setminus K_{n_i}) = X \setminus K_m$, d.h. $K_m = \varnothing$, Widerspruch zur Annahme $K_n \neq \varnothing$ für alle $n \in \mathbb{N}$. Es folgt $\bigcap\limits_{n\in\mathbb{N}} K_n \neq \varnothing$. $\surd$

Sei nun $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge aus X. Dann ist, für alle $N \in \mathbb{N}$, die Teilmenge $K_N := \overline{\{x_n,\, n \geq N\}} \subset X$ abgeschlossen in X. Es gilt offenbar $K_{n+1} \subset K_n$ und $K_n \neq \varnothing$ für alle $n \in \mathbb{N}$. Somit ist $(K_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge abgeschlossener Teilmengen von X. Aus der letzten Behauptung folgt, dass $\bigcap\limits_{n\in\mathbb{N}} K_n \neq \varnothing$, d.h. es existiert ein $x \in \bigcap\limits_{n\in\mathbb{N}} K_n$.
Konstruiere nun eine streng monoton wachsende Abbildung $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ wie folgt: wähle $\varphi(0) \in \mathbb{N}$ s.d. $x_{\varphi(0)} \in B_1(x) \cap \{x_p,\, p \in \mathbb{N}\}$, und konstruiere induktiv $\varphi(n)$ durch:

$$x_{\varphi(n)} \in B_{\frac{1}{n+1}}(x) \cap \{x_p,\, p \geq \varphi(n-1)+1\}$$

(wegen $x \in K_p$ für alle $p \in \mathbb{N}$ ist der Durchschnitt auf der rechten Seite nicht leer). Die so definierte Abbildung φ ist offensichtlich streng monoton wachsend, und die Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ erfüllt, für jedes $n \in \mathbb{N}$:

$$d(x_{\varphi(n)}, x) < \frac{1}{n+1},$$

insbesondere $d(x_{\varphi(n)}, x) \underset{n\to+\infty}{\longrightarrow} 0$, d.h., $x_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} x$. Daraus folgt, dass $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge in X besitzt. Damit ist (X, d) kompakt.
Umgekehrt setzen wir voraus, dass (X, d) kompakt sei.

Lemma B.2 *Sei (X, d) ein kompakter metrischer Raum. Sei $\mathcal{U}$ eine offene Überdeckung von X. Dann existiert ein $\varepsilon > 0$ so, dass jeder offene Ball vom Radius ε in (mindestens) einem Element von $\mathcal{U}$ enthalten ist.*

Beweis von Lemma B.2: Angenommen, dies gelte nicht. Für jedes $n \in \mathbb{N}$, $n \geq 1$, würde dann ein $x_n \in X$ so existieren, dass $B_{\frac{1}{n}}(x_n)$ in keinem Element von $\mathcal{U}$ enthalten wäre. Die Folge $(x_n)_{n\in\mathbb{N}}$ würde eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ besitzen, die gegen ein $x \in X$ konvergieren würde. Da $\mathcal{U}$ eine Überdeckung von X ist, würde ein $U \in \mathcal{U}$ existieren mit $x \in U$, und wegen U offen würde auch ein $r > 0$ so existieren, dass $B_r(x) \subset U$. Sei dann $N \in \mathbb{N} \setminus \{0\}$ s.d. $\frac{1}{\varphi(N)} < \frac{r}{2}$ und $d(x_{\varphi(N)}, x) < \frac{r}{2}$. Aus der Dreiecksungleichung würde damit

$$B_{\frac{1}{\varphi(N)}}(x_{\varphi(N)}) \subset B_r(x) \subset U$$

gelten, Widerspruch zur Konstruktion der Folge $(x_n)_{n\in\mathbb{N}}$. $\surd$

Lemma B.3 *Sei (X, d) ein kompakter metrischer Raum. Dann existiert es für jedes $\varepsilon > 0$ eine* endliche *Überdeckung von X durch offene Bälle vom Radius ε.*

Beweis von Lemma B.3: Angenommen, dies gelte nicht, d.h., es würde ein $\varepsilon > 0$ so geben, dass jede endliche Familie von offenen Bällen vom Radius ε die Menge X *nicht* überdeckt. Daher würde eine Folge $(x_n)_{n\in\mathbb{N}}$ von X existieren mit $x_{n+1} \notin \bigcup\limits_{0\leq i\leq n} B_\varepsilon(x_i)$ für alle $n \in \mathbb{N}$.
Diese Folge würde eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ besizten, die gegen ein $x \in X$ konvergieren würde. Für ein festes $r \in]0, \frac{\varepsilon}{2}[$ würde dann ein $N \in \mathbb{N}$ existieren mit $d(x_{\varphi(n)}, x) < r$ für alle $n \geq N$, und somit (nach der Dreiecksungleichung)

$$d(x_{\varphi(p)}, x_{\varphi(q)}) < \varepsilon$$

für alle $p, q \geq N$, Widerspruch zur Konstruktion der Folge $(x_n)_{n\in\mathbb{N}}$. $\checkmark$

(*Ende des Beweises von* Satz B.1) Sei $\mathcal{U}$ eine beliebige offene Überdeckung von X. Nach Lemma B.2 existiert ein $\varepsilon > 0$ s.d. jeder Ball in X von Radius ε in einem Element von $\mathcal{U}$ enthalten ist. Für dieses ε existiert nach Lemma B.3 eine endliche Familie $B_1, \ldots, B_p$ von Bällen von Radius ε, die X überdecken ($X = \bigcup\limits_{1\leq i\leq p} B_i$). Da jedes B_i in einem Element U_i von $\mathcal{U}$ enthalten ist, folgt $X = \bigcup\limits_{1\leq i\leq p} U_i$. Somit existiert eine endliche Teilmenge von $\mathcal{U}$, die X immer noch überdeckt. Daraus folgt, dass jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt. $\square$

B.2 Gleichmäßigstetigkeit

Wir haben schon gesehen, dass jede auf einem kompakten metrischen Raum definierte stetige Abbildung automatisch *gleichmäßig* stetig ist (Proposition 2.33). Wir beweisen nun eine leicht stärkere Aussage, in der Situation, wo die Abbildung auf einem Produkt von metrischen Räumen definiert ist, von dem nur einer kompakt vorausgesetzt wird.

Proposition B.4 *Sei $f : X \times Y \longrightarrow Z$ eine stetige Abbildung, wobei (X, d_X), (Y, d_Y), (Z, d_Z) metrische Räume sind und $X \times Y$ mit der Produktmetrik versehen wird. Ist der Raum (X, d_X) kompakt, so ist f* gleichmäßig stetig in *$x \in X$, d.h., für alle $y \in Y$ und $\varepsilon > 0$ existiert ein $\eta > 0$ so, dass $d_Z(f(x, y'), f(x, y)) < \varepsilon$ für alle $y' \in B_\eta(y) \subset Y$ und* alle *$x \in X$.*

Beweis: Wie bei der Proposition 2.33 führen wir einen Widerspruchsbeweis durch. Angenommen, dies gelte nicht, dann würde ein $\varepsilon > 0$ und ein $y \in Y$ so existieren, dass es für jedes $\eta > 0$ ein $y' \in B_\eta(y)$ und ein $x \in X$ gibt mit $d_Z(f(x, y'), f(x, y)) \geq \varepsilon$. Insbesondere existieren, für jedes η der Form $\frac{1}{n}$ (wobei $n \in \mathbb{N} \setminus \{0\}$), Elemente $y_n \in B_{\frac{1}{n}}(y)$ und $x_n \in X$ mit $d_Z(f(x_n, y_n), f(x_n, y)) \geq \varepsilon$. Nach Konstruktion gilt $y_n \underset{n\to\infty}{\longrightarrow} y$. Wegen (X, d) kompakt existiert eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\geq 1}$ und ein $x \in X$ mit $x_{\varphi(n)} \underset{n\to\infty}{\longrightarrow} x$. Da f stetig ist, folgt

$$d_Z(f(x_{\varphi(n)}, y_{\varphi(n)}), f(x_{\varphi(n)}, y)) \underset{n\to\infty}{\longrightarrow} d_Z(f(x, y), f(x, y)) = 0,$$

Widerspruch zu $d_Z(f(x_{\varphi(n)}, y_{\varphi(n)}), f(x_{\varphi(n)}, y)) \geq \varepsilon > 0$ für alle $n \in \mathbb{N}$. $\square$

Literaturverzeichnis

[1] T. Bröcker, *Analysis II*, 2. Auflage, Spektrum Akademischer Verlag, 1995.

[2] T. Bröcker, *Analysis III*, BI-Wissenschaftsverlag, 1992.

[3] T. Bröcker, Teilskript zur Vorlesung "Analysis II für Physiker", erhaltbar unter `http://www.mathematik.uni-regensburg.de/broecker/kapv-vi.pdf`.

[4] H. Cartan, *Differential calculus*, Hermann, Paris; Houghton Mifflin Co., Boston, Mass., 1971.

[5] H. Cartan, *Differential forms*, Houghton Mifflin Co., Boston, Mass. 1970.

[6] J. Elstrodt, *Maß- und Integrationstheorie*, 7. Auflage, Springer, Berlin, 2011.

[7] A. Gramain, *Intégration*, Hermann, Paris, 1988.

[8] K. Jänich, *Mathematik 2. Geschrieben für Physiker*, Springer, 2002.

[9] K. Jänich, *Vektoranalysis*, 5. Auflage, Springer, 2005.

[10] H. Kerner, W. von Wahl, *Mathematik für Physiker*, 2. Auflage, Springer, 2007.

[11] D. Werner, *Einführung in die höhere Analysis*, 2. Auflage, Springer, 2009.

Bezeichnungen

Index