Differenzial- und Integralrechnung II

Ingo Witt

SoSe 2012

Inhaltsverzeichnis

Kapitel 7

Topologische Räume

7.1 Grundlegende Begriffe

Bei topologischen Betrachtungen geht es um die Lage von Punkten im Raum zueinander, ohne dass dabei Größenverhältnisse (Längen, Winkel, etc.) eine Rolle spielen. Das heißt insbesondere, dass topologische Eigenschaften sich bei „stetigen Deformationen" nicht ändern (Gummibandgeometrie).

Definition 7.1. Ein *topologischer Raum* $(X, \mathscr{U})$ ist eine Menge X und eine Menge $\mathscr{U}$ von Teilmengen von X mit folgenden Eigenschaften:

(i) $\emptyset, X \in \mathscr{U}$,

(ii) $\mathscr{U}$ ist abgeschlossen unter endlichen Durchschnitten,

(iii) $\mathscr{U}$ ist abgeschlossen unter beliebigen Vereinigungen.

Im Allgemeinen schreiben wir X statt $(X, \mathscr{U})$, wenn aus dem Kontext klar ist, was die Menge $\mathscr{U}$ ist. Die Elemente von $\mathscr{U}$ heißen die *offenen Mengen* von X. Die Komplemente in X der Elemente von $\mathscr{U}$ heißen die *abgeschlossenen Mengen* von X. $\mathscr{U}$ heißt auch eine *Topologie* auf X.

Lemma 7.2. *Die Mengen $\emptyset, X$ sind abgeschlossen. Endliche Vereinigungen und beliebige Durchschnitte abgeschlossener Mengen von X sind abgeschlossen.*

Beispiel 7.3.

1. *Diskrete Topologie*: $\mathscr{U}$ besteht aus allen Teilmengen von X. *Triviale Topologie*: $\mathscr{U} = \{\emptyset, X\}$.

2. $X = \{a, b\}$, $\mathscr{U} = \big\{\emptyset, \{a\}, \{a, b\}\big\}$.

3. $X = \mathbb{R}$, $\mathscr{U}$ besteht aus allen offenen Teilmengen von $\mathbb{R}$ (wie im vergangenen Semester eingeführt). (Vergleiche Satz 2.56.)

4. Teilmengen von $\mathbb{R}^n$ bezüglich der induzierten Topologie (siehe unten).

5. Metrische Räume (siehe das nächste Kapitel).

Lemma 7.4. *Sei $(X, \mathscr{U})$ ein topologischer Raum, $A \subseteq X$. Dann gibt es eine größte offene Menge A°, die in A enthalten ist, und eine kleinste abgeschlossene Menge $\bar{A}$, die A enthält.*

Beweis. Es gilt $A^\circ = \bigcup\limits_{\substack{U \text{ offen}, \\ U \subseteq A}} U$, $\bar{A} = \bigcap\limits_{\substack{F \text{ abgeschlossen}, \\ F \supseteq A}} F$. □

Definition 7.5. A° heißt das *Innere*, $\bar{A}$ der *Abschluss* von A. Darüber hinaus heißt $\partial A = \bar{A} \setminus A^\circ$ der *Rand* von A.

Lemma 7.6. (i) *Eine Menge $A \subseteq X$ ist genau dann abgeschlossen, wenn die Menge A alle ihre Randpunkte enthält.*

(ii) *Sei $x \in X$. Dann ist x genau dann ein Randpunkt von A, wenn für alle $U \in \mathscr{U}$ mit $x \in U$*
$$U \cap A \neq \emptyset, \ U \cap (X \setminus A) \neq \emptyset$$
gilt.

Beweis. (i) A abgeschlossen heißt gerade $A = \bar{A}$ $(= A \cup \partial A)$.

(ii) Wir haben eine disjunkte Vereinigung
$$X = A^\circ \cup (X \setminus A)^\circ \cup \partial A.$$
□

7.2 Stetigkeit und Homöomorphie

Seien $(X, \mathscr{U})$, $(Y, \mathscr{V})$ topologische Räume, $f\colon X \to Y$ eine Abbildung.

Definition 7.7. f heißt *stetig* von X nach Y, falls für alle $V \in \mathscr{V}$ die Menge
$$f^{-1}(V) = \{x \in X \mid f(x) \in V\}$$
zu $\mathscr{U}$ gehört (d. h. falls die Urbilder offener Mengen offen sind).

Beispiel 7.8. Sei $I \subseteq \mathbb{R}$ ein Intervall, $f\colon I \to \mathbb{R}$. Dann ist f genau dann im Sinne obiger Definition stetig, wenn f im bisherigen Sinne stetig ist.
Wie in Satz 2.57 beweist man:

Satz 7.9. *Sei $f\colon X \to Y$. Dann sind äquivalent:*

(i) *f ist stetig.*

(ii) *$f^{-1}(G) \subseteq X$ ist abgeschlossen für alle abgeschlossenen Mengen $G \subseteq Y$.*

(iii) *Für alle $A \subseteq X$ gilt: $f(\bar{A}) \subseteq \overline{f(A)}$.*

Satz 7.10. *Die Komposition stetiger Abbildungen ist stetig.*

Beweis. Seien $f\colon X \to Y$ und $g\colon Y \to Z$ stetig. Dann ist $g \circ f\colon X \to Z$ stetig, da für alle offenen Mengen $W \subseteq Z$ die Menge
$$(g \circ f)^{-1}(W) = f^{-1}\big(g^{-1}(W)\big)$$
offen in X ist. □

Die folgende Definition führt den Begriff der *topologischen Äquivalenz* ein:

Definition 7.11. Zwei topologische Räume X, Y heißen *homöomorph* (topologisch äquivalent, topologisch nicht unterscheidbar), falls es eine stetige Bijektion $f\colon X \to Y$ gibt, so dass $f^{-1}\colon Y \to X$ ebenfalls stetig ist. Ein derartiges f heißt eine *Homöomorphie*.

Bemerkung. Homöomorphie ist eine Äquivalenzrelation in folgendem Sinne:

(i) $\mathrm{id}\colon X \to X$ ist eine Homöomorphie für alle X (Reflexivität).

(ii) Ist $f\colon X \to Y$ eine Homöomorphie, so auch $f^{-1}\colon Y \to X$ (Symmetrie).

(iii) Sind $f\colon X \to Y$, $g\colon Y \to Z$ Homöomorphien, so auch $g \circ f\colon X \to Z$ (Transitivität).

Beispiel 7.12. Sind die folgenden topologischen Räume homöomorph?

(a)

(b)

(c)

Zu (a): Nein, ... ist *kompakt*, ... ist es nicht.

Zu (b): Nein. Wenn wir einen Punkt entfernen, dann zerfällt ... in zwei *zusammenhängende* Mengen (...), ... jedoch nicht (...).

Zu (c): Ja. ... habe die Länge l_1, der verschlungene Knoten die Länge l_2. Wir wählen auf beiden Kurven einen beliebigen Punkt P bzw. Q und einen Umlaufsinn. Der Homöomorphismus f ordnet dann dem Punkt auf ..., der sich in Entfernung sl_1 für $s \in [0, 1)$ von P befindet, den Punkt auf ... zu, der sich in Entfernung sl_2 von Q befindet.

7.3 Konstruktion topologischer Räume

7.3.1 Die Unterraumtopologie

Sei $(X, \mathscr{U})$ ein topologischer Raum, $A \subseteq X$. Wir setzen

$$\mathscr{U}_A = \big\{ U \cap A \bigm| U \in \mathscr{U} \big\}.$$

Lemma 7.13. *$\mathscr{U}_A$ ist eine Topologie auf A.*

Beweis. Eine direkte Verifikation. □

Definition 7.14. $\mathscr{U}_A$ heißt die von $\mathscr{U}$ auf A induzierte Unterraumtopologie.

7.3.2 Die koinduzierte Topologie

Sei X eine Menge, $\mathscr{U}$, $\mathscr{V}$ seien zwei Topologien auf X. Dann heißt $\mathscr{U}$ *feiner* oder *stärker* als $\mathscr{V}$ (bzw. $\mathscr{V}$ heißt *gröber* oder *schwächer* als $\mathscr{U}$), falls jede Menge in $\mathscr{V}$ auch zu $\mathscr{U}$ gehört (d. h. wenn $\mathscr{U}$ mehr offene Mengen als $\mathscr{V}$ hat). Die diskrete Topologie ist die feinste Topologie auf X, die triviale Topologie die gröbste.

Bemerkung. Für jede Teilmenge $\mathscr{A}$ der Potenzmenge von X gibt es eine schwächste Topologie auf X, die $\mathscr{A}$ enthält (nämlich der Durchschnitt aller Topologien, die $\mathscr{A}$ enthalten). Diese heißt die *von $\mathscr{A}$ erzeugte Topologie* auf X.

Definition 7.15. Sei $(X, \mathscr{U})$ ein topologischer Raum, Y eine Menge und sei $f\colon X \to Y$ eine Abbildung. Dann heißt die stärkste Topologie $\mathscr{V}_f$ auf Y, bezüglich der f stetig ist, die von f auf Y *koinduzierte Topologie*.

Lemma 7.16. *Es gilt*

$$\mathscr{V}_f = \left\{ V \subseteq Y \;\middle|\; f^{-1}(V) \text{ ist offen in } X \right\}.$$

Beweis. Es genügt zu verifizieren, dass $\mathscr{V}_f$ in der Tat eine Topologie ist. ☐

Satz 7.17. *X, Y, Z seien topologische Räume, wobei Y die durch $f\colon X \to Y$ koinduzierte Topologie trägt. Dann ist $g\colon Y \to Z$ genau dann stetig, wenn $g \circ f\colon X \to Z$ stetig ist.*

Beweis. ($\Longrightarrow$) Offensichtlich.

($\Longleftarrow$) Sei $g \circ f\colon X \to Z$ stetig. Für eine offene Menge $W \subseteq Z$ ist $f^{-1}(g^{-1}(W))$ offen in X, d. h. $g^{-1}(W)$ ist offen in Y. Somit ist $g\colon Y \to Z$ stetig. ☐

7.3.3 Produkträume

Sei X_i, $i \in I$, eine Familie topologischer Räume.

Definition 7.18. Die *Produkttopologie* auf $X = \prod_{i \in I} X_i$ ist die von allen Mengen der Form $\prod_{i \in I} U_i$ erzeugte Topologie, wobei $U_i \subseteq X_i$ für alle i offen ist und $U_i = X_i$ *für fast alle* (d. h. mit Ausnahme von höchstens endlich vielen) i gilt.

Die Produkttopologie auf $X = \prod_{i \in I} X_i$ ist offenbar die schwächste Topologie auf X, bezüglich der alle kanonischen Projektionen $\pi_j\colon X \to X_j$ stetig sind.

Satz 7.19. *Sei Z ein topologischer Raum und $f\colon Z \to \prod_{i \in I} X_i$ eine Abbildung. Dann ist f genau dann stetig, wenn alle Abbildungen $\pi_j \circ f\colon Z \to X_j$ stetig sind.*

Beweis. ($\Longrightarrow$) Klar.

($\Longleftarrow$) Sei $U_i \subseteq X$ für $i \in I$ offen, $U_i = X_i$ für fast alle i. Dann ist

$$f^{-1}\left(\prod_{i \in I} U_i\right) = \bigcap_{i \in I} (\pi_i \circ f)^{-1}(U_i)$$

offen in Z. ☐

Beispiel 7.20. (i) $\mathbb{R}^n$ trägt die Topologie des Produktes $\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}}$.

(ii) Der Zylinder trägt die Topologie des Produkts $I \times \mathbb{S}^1$, wobei $I = [0,1]$ und $\mathbb{S}^1$ ist die Kreislinie.

7.4 Zusammenhängende topologische Räume

7.4.1 Definition und erste Eigenschaften

Definition 7.21. Ein topologischer Raum X heißt *zusammenhängend*, falls er *nicht* in der Form $X = U \cup V$ mit $U, V \subseteq X$ offen, nichtleer und disjunkt geschrieben werden kann.

Für die disjunkte Vereinigung schreiben wir auch $X = U \sqcup V$.

Beispiel 7.22. (a) Das Intervall $I = \ldots$, die Kreislinie $\mathcal{S}^1 = \ldots$ und der Zylinder $\ldots$ sind zusammenhängend.

(b) Die disjunkte Vereinigung ... zweier Intervalle ist es nicht.

Allgemeiner gilt, dass die zusammenhängenden Teilmengen von $\mathbb{R}$ gerade die leere Menge sowie die Intervalle sind.

Lemma 7.23. *Die folgenden Bedingungen sind äquivalent*:

(i) *X ist zusammenhängend.*

(ii) *Die einzigen Teilmengen von X, die sowohl offen als auch abgeschlossen sind, sind $\emptyset$ und X.*

(iii) *Jede stetige Abbildung $f\colon X \to Y$, wobei Y die diskrete Topologie trägt, ist konstant.*

Beweis. (i) $\Rightarrow$ (ii). Sei $U \subseteq X$ offen und abgeschlossen. Dann ist $X \setminus U$ ebenfalls offen und abgeschlossen. Wegen $X = U \cup (X \setminus U)$, $U \cap (X \setminus U) = \emptyset$ gilt $U = \emptyset$ oder $X \setminus U = \emptyset$, d. h. $U = X$.

(ii) $\Rightarrow$ (iii). Sei $f\colon X \to (Y, \mathcal{U}_{\text{diskret}})$ stetig. Sei $y^* = f(x^*)$ für ein $x^* \in X$. Dann ist $f^{-1}(y^*)$ offen, abgeschlossen und nichtleer. Also $f^{-1}(y^*) = X$ und $f(x) = y^*$ für alle $x \in X$, d. h. f ist konstant.

(iii) $\Rightarrow$ (i). Sei $X = U \cup V$ mit $U \cap V = \emptyset$. Dann ist die Funktion $f\colon X \to (\{0,1\}, \mathcal{U}_{\text{diskret}})$ mit

$$f(x) = \begin{cases} 0, & x \in U, \\ 1, & x \in V, \end{cases}$$

stetig, also konstant. Damit ist $U = \emptyset$ oder $V = \emptyset$. $\square$

Bemerkung. Der Beweis zeigt, dass wir in (iii) ein allgemeines Y durch die zweielementige Menge $\{0,1\}$ ersetzen können.

Satz 7.24. *Stetige Bilder zusammenhängender Mengen sind zusammenhängend.* (Vergleiche auch mit Satz 2.49.)

Beweis. Sei $f\colon X \to Y$ stetig. Weiterhin seien $V_0, V_1 \in Y$ zwei offene Mengen mit $f(X) \subseteq V_0 \cup V_1$, $f(X) \cap V_0 \cap V_1 = \emptyset$.
Sei $U_0 = f^{-1}(V_0)$, $U_1 = f^{-1}(V_1)$. Dann ist $X = U_0 \cup U_1$, $U_0 \cap U_1 = \emptyset$, U_0, U_1 sind offen. Es folgt $U_0 = \emptyset$ oder $U_1 = \emptyset$, also $f(X) \cap V_0 = \emptyset$ oder $f(X) \cap V_1 = \emptyset$.
Damit ist $f(X)$ zusammenhängend. □

Satz 7.25. *Sei X_i zusammenhängend für alle i. Dann ist $\prod_{i \in I} X_i$ zusammenhängend.*

Beweis. Diesen Beweis überlegen Sie sich selbst. □

7.4.2 Zusammenhangskomponenten

Satz 7.26. *Sei X ein topologischer Raum. Dann ist $X = \bigsqcup_i X_i$ disjunkte Vereinigung seiner (eindeutig bestimmten) Zusammenhangskomponenten X_i (d. h. jedes X_i ist zusammenhängend und zudem maximal mit dieser Eigenschaft).*

Beweis. Wir führen eine Äquivalenzrelation $\sim$ auf X ein: $x \sim y$, falls es eine zusammenhängende Teilmenge von X gibt, die x und y enthält. Um zu sehen, dass diese Relation transitiv ist, zeigen wir:

Lemma 7.27. *Seien $U, V \subseteq X$ zusammenhängend mit $U \cap V \neq \emptyset$. Dann ist $U \cup V$ zusammenhängend.*

Beweis. Sei $f\colon U \cup V \to \{0,1\}$ stetig, wobei $\{0,1\}$ die diskrete Topologie trägt. Dann ist $f\big|_U\colon U \to \{0,1\}$ stetig, also konstant, und auch $f\big|_V\colon V \to \{0,1\}$ ist stetig, also ebenfalls konstant. Wegen $U \cap V \neq \emptyset$ ist die Konstante dieselbe, also ist $f\colon U \cup V \to \{0,1\}$ konstant. □

Die Äquivalenzklassen von X bezüglich $\sim$ sind gerade die Zusammenhangskomponenten X_i von X. Offenbar gilt $X = \bigsqcup X_i$ und die X_i sind paarweise disjunkt. Weiterhin ist jedes X_i zusammenhängend (Beweis?), während aus Y zusammenhängend und $X_i \subseteq Y$ folgt, dass $X_i = Y$ gilt (Beweis?). □

Bemerkung. Die Anzahl der Zusammenhangskomponenten eines topologischen Raumes X ist offenbar eine *topologische Invariante* (d. h. diese Anzahl ist gleich für alle topologischen Räume, die homöomorph zu X sind).
Insbesondere ist ein nichtleerer topologischer Raum genau dann zusammenhängend, wenn diese Anzahl gleich 1 ist. (Aber nicht alle nichtleeren zusammenhängenden topologischen Räume sind homöomorph, z. B. ...!)

7.4.3 Wegzusammenhängende Räume

Definition 7.28. Eine (stetige) *Kurve* in X ist eine stetige Abbildung $\gamma\colon [0,1] \to X$. $\gamma(0)$ heißt der Anfangspunkt, $\gamma(1)$ der Endpunkt der Kurve γ.

Definition 7.29. Ein topologischer Raum heißt *wegzusammenhängend*, wenn es für je zwei Punkte x, $y \in X$ eine Kurve γ gibt, die diese beiden Punkte verbindet (d. h. $\gamma(0) = x$, $\gamma(1) = y$).

Beispiel 7.30. Die wegzusammenhängenden Teilmengen von $\mathbb{R}$ sind die leere Menge und die Intervalle.

Satz 7.31. *Das stetige Bild eines wegzusammenhängenden Raumes ist wegzusammenhängend.*

Beweis. Sei $f\colon X \to Y$ stetig, $y_0, y_1 \in f(X)$. Wir wählen $x_0, x_1 \in X$ mit $f(x_0) = y_0$, $f(x_1) = y_1$. Dann gibt es eine Kurve $\gamma\colon [0,1] \to X$ mit $\gamma(0) = x_0$, $\gamma(1) = x_1$. $f \circ \gamma\colon [0,1] \to f(X)$ ist dann eine Kurve in $f(X)$, die y_0 und y_1 verbindet. $\square$

Wie im Fall des Zusammenhangs lassen sich jetzt die *Wegzusammenhangskomponenten* eines topologischen Raumes X definieren. Obiger Satz zeigt uns, dass deren Anzahl wiederum eine topologische Invariante ist.

Satz 7.32. *Jeder wegzusammenhängende topologische Raum ist zusammenhängend.*

Beweis. Sei X wegzusammenhängend und $f\colon X \to \{0,1\}$ stetig, wobei $\{0,1\}$ mit der diskreten Topologie versehen ist. Sei o. B. d. A. $f(x^*) = 0$ für ein $x^* \in X$. Für einen beliebigen Punkt $x \in X$ wählen wir einen Weg $\gamma\colon [0,1] \to X$ von x^* nach x, d. h. $\gamma(0) = x^*$, $\gamma(1) = x$. Dann ist $f \circ \gamma\colon [0,1] \to \{0,1\}$ stetig. Da das Intervall $[0,1]$ zusammenhängend ist, ist $f \circ \gamma$ konstant Null, also $f(x) = f\left(\gamma(1)\right) = 0$. Damit ist f konstant. $\square$

Nicht zu überraschend ist sicherlich, dass die Umkehrung nicht gilt:

Beispiel 7.33. Sei $X \subseteq \mathbb{R}^2$,

$$X = \big\{(0,y) \bigm| |y| \le 1\big\} \cup \left\{\left(x, \sin\frac{1}{x}\right) \,\middle|\, x > 0\right\},$$

versehen mit der induzierten Topologie. Dann ist X zusammenhängend, aber nicht wegzusammenhängend.

7.5 Kompakte topologische Räume

Sei X ein topologischer Raum. Eine *offene Überdeckung* von X ist eine Familie $\{U_i\}_{i\in I}$ offener Teilmengen von X, so dass $X = \bigcup\limits_{i\in I} U_i$.

Definition 7.34. Ein topologischer Raum X heißt *kompakt*, falls jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt.

Bemerkung. Insbesondere in der algebraischen Geometrie heißen derartige Räume auch *quasikompakt*. Kompaktheit ist dann Quasikompaktheit plus die Hausdorff-Eigenschaft (siehe den nächsten Abschnitt).

Beispiel 7.35. Die kompakten Teilmengen der reellen Geraden $\mathbb{R}$ sind gerade die abgeschlossenen und beschränkten Teilmengen (Überdeckungssatz von Heine–Borel). Selbiges gilt im $\mathbb{R}^n$, $n \ge 2$, ist jedoch in allgemeinen metrischen Räumen im Allgemeinen falsch.

Satz 7.36. *Eine abgeschlossene Teilmenge eines kompakten Raumes ist kompakt.*

Beweis. Sei X kompakt, $A \subseteq X$ abgeschlossen. Seien U_i, $i \in I$, offene Mengen in X mit $A \subseteq \bigcup\limits_{i\in I} U_i$. Dann ist $X \setminus A$ offen in X und $X = (X \setminus A) \cup \bigcup\limits_{i\in I} U_i$. Folglich existiert eine endliche Teilmenge $J \subseteq I$ mit $X = (X \setminus A) \cup \bigcup\limits_{j\in J} U_j$. Dann $A \subseteq \bigcup\limits_{j\in J} U_j$. □

Satz 7.37. *Das stetige Bild eines kompakten Raumes ist kompakt.*

Beweis. Sei $f\colon X \to Y$ stetig, X kompakt. Weiterhin sei $\{V_i\}$ eine Familie offener Mengen in Y mit $f(X) \subseteq \bigcup_i V_i$. Dann sind die Mengen $f^{-1}(V_i)$ offen in X und es gilt $X = \bigcup_i f^{-1}(V_i)$. Daher gibt es endlich viele Indizes $i_1, \ldots, i_r$ mit $X = \bigcup_{k=1}^r f^{-1}(V_{i_k})$. Es folgt, dass $f(X) \subseteq \bigcup_{k=1}^r V_{i_k}$, und wir haben eine endliche Teilüberdeckung gefunden. □

Dieses Resultat zeigt, dass Kompaktheit eine topologische Eigenschaft ist.

Das folgende Ergebnis ist einer der wichtigsten Sätze der allgemeinen Topologie.

Satz 7.38 (Tychonoff)**.** *Das topologische Produkt einer beliebigen Anzahl kompakter topologischer Räume ist kompakt.*

Beweis (*für den Fall endlich vieler Faktoren*). Es genügt, den Beweis für zwei Faktoren zu führen; die Behauptung für eine beliebige Zahl endlich vieler Faktoren folgt dann per Induktion.
Seien also X, Y kompakte topologische Räume, sei $X \times Y = \bigcup\limits_i (U_i \times V_i)$ für offene Mengen $U_i \subseteq X$, $V_i \subseteq Y$. (Wieso genügt es, sich auf offene Mengen der Form $U_i \times V_i$ zu beschränken?)
Für $y \in Y$ ist die Menge $X \times \{y\} \subset X \times Y$ kompakt als homöomorphes Bild des Raumes X. (Wieso das?) Wir finden also für jedes $y \in Y$ eine endliche Überdeckung

$$X \times \{y\} \subseteq \bigcup_{r=1}^{m_y} U_r^y \times V_r^y,$$

wobei die $U_r^y \times V_r^y$ aus den $U_i \times V_i$ gewählt sind. Wir betrachten nun die Menge $V^y = \bigcap\limits_{r=1}^{m_y} V_r^y$. Diese Menge ist offen in Y, $y \in V^y$. Es gibt also endlich viele $y_1, \ldots, y_n \in Y$, so dass die $V^{y_1}, \ldots, V^{y_n}$ den Raum Y überdecken, d. h. $Y = \bigcup\limits_{s=1}^{n} V^{y_s}$. Dann gilt

$$X \times Y = \bigcup_{s=1}^{n} \bigcup_{r=1}^{m_{y_s}} U_r^{y_s} \times V_r^{y_s}.$$

Um letztere Beziehung einzusehen, wählen wir einen Punkt $(x^*, y^*) \in X \times Y$. Dann gibt es ein $s \in \{1, \ldots, n\}$ mit $y^* \in V^{y_s}$. Wegen $(x^*, y^s) \in X \times \{y^s\}$ gibt es ein $r \in \{1, \ldots, m_{y^s}\}$ mit $(x^*, y^s) \in U_r^{y_s} \times V_r^{y_s}$. Dann jedoch gilt $x^* \in U_r^{y_s}$ und $y^* \in V^{y_s} \subseteq V_r^{y_s}$, also $(x^*, y^*) \in U_r^{y_s} \times V_r^{y_s}$. □

Anhang zu Abschnitt 7.5: Die Einpunkt-Kompaktifizierung

Sei $(X, \mathscr{T})$ ein *nichtkompakter* topologischer Raum. Wir konstruieren einen kompakten topologischen Raum $\big(\hat{X}, \hat{\mathscr{T}}\big)$, in den der Raum X dicht (und homöomorph) einbettet.
Sei $\hat{X} = X \sqcup \{*\}$, wobei $*$ eine Stelle mit $* \notin X$ bezeichnet. Weiterhin sei

$$\hat{\mathscr{T}} = \mathscr{T} \cup \big\{ V \cup \{*\} \,\big|\, V \in \mathscr{T},\, X \setminus V \text{ ist kompakt} \big\}.$$

Satz 7.39. *Der so definierte Raum* $\big(\hat{X}, \hat{\mathscr{T}}\big)$ *hat die folgenden Eigenschaften*:

a) $\big(\hat{X}, \hat{\mathscr{T}}\big)$ *ist ein kompakter topologischer Raum.*

b) *Die kanonische Einbettung* $X \hookrightarrow \hat{X}$ *ist stetig, offen*[1] *und dicht.*

Beweis. Ist $\{U_i\} \subset \mathscr{T}$ eine beliebige Familie und $\{V_j\} \subset \mathscr{T}$ eine nichtleere Familie mit V_{j^*} ein Mitglied dieser Familie, wobei $X \setminus V_j$ kompakt für alle j ist, so ist

$$\Big(\bigcup\nolimits_i U_i\Big) \cup \Big(\bigcup\nolimits_j (V_j \cup \{*\})\Big) = \Big(\bigcup\nolimits_i U_i \cup \bigcup\nolimits_j V_j\Big) \cup \{*\} \in \hat{\mathscr{T}},$$

da $X \setminus \Big(\bigcup_i U_i \cup \bigcup_j V_j\Big) \subseteq X \setminus V_{j^*}$ abgeschlossen und daher kompakt ist.
Ist $\{V_j\} \subset \mathscr{T}$ eine nichleere endliche Familie, so ist

$$\Big(\bigcap\nolimits_j (V_j \cup \{*\})\Big) = \Big(\bigcap\nolimits_j V_j\Big) \cup \{*\} \in \hat{\mathscr{T}},$$

da $X \setminus \Big(\bigcap_j V_j\Big) = \bigcup_j (X \setminus V_j)$ kompakt ist. Damit ist $\hat{\mathscr{T}}$ eine Topologie auf $\hat{X}$.
Nun sei

$$\hat{X} = \Big(\bigcup\nolimits_i U_i\Big) \cup \Big(\bigcup\nolimits_j (V_j \cup \{*\})\Big)$$

für zwei Familien $\{U_i\}_{i \in I} \subset \mathscr{T}$, $\{V_j\}_{j \in J} \subset \mathscr{T}$, wobei $X \setminus V_j$ kompakt für alle j und J nichtleer mit $j^* \in J$ ist. Dann

$$X = \bigcup\nolimits_i U_i \cup \bigcup\nolimits_j V_j,$$

wobei die Menge $X \setminus V_{j^*}$ kompakt ist. Es gibt also eine endliche Teilmenge $I_0 \subseteq I$ und eine endliche Teilmenge $J_0 \subseteq J \setminus \{j^*\}$ mit

$$X \setminus V_{j^*} \subseteq \Big(\bigcup\nolimits_{i \in I_0} U_i\Big) \cup \Big(\bigcup\nolimits_{j \in J_0} V_j\Big).$$

Es folgt, dass

$$\hat{X} = \Big(\bigcup\nolimits_{i \in I_0} U_i\Big) \cup \Big(\bigcup\nolimits_{j \in J_0} (V_j \cup \{*\})\Big) \cup \big(V_{j^*} \cup \{*\}\big),$$

was die Kompaktheit des Raumes $\hat{X}$ demonstriert.
Identifizieren wir nun X mit seinem kanonischen Bild in $\hat{X}$, so trägt X die Relativtopologie, was die Stetigkeit und Offenheit der kanonischen Einbettung aufzeigt. Die Dichtheit dieser Einbettung folgt aus dem Umstand, dass $\{*\} \notin \hat{\mathscr{T}}$ gilt, dass $*$ also ein Randpunkt von X in $\hat{X}$ und somit der Abschluss von X in $\hat{X}$ der Raum $\hat{X}$ ist. $\square$

[1] Das heißt, die Bilder offener Mengen sind offen.

$\hat{X}$ heißt die *Einpunkt-Kompaktifizierung* von X.

Bemerkung. Die Einpunkt-Kompaktifizierung von X ist in dem Sinne eindeutig, dass wenn Y ein kompakter topologischer Raum und $i\colon X \to Y$ eine injektive, stetige und offene Abbildung mit der Eigenschaft ist, dass $i(X) \subset Y$ dicht und $Y \setminus i(X)$ eine einelementige Menge ist, so ist Y homöomorph zu $\hat{X}$.

Beispiel 7.40. (a) Die Einpunkt-Kompaktifizierung der Menge $\mathbb{N}$ der natürlichen Zahlen (mit der diskreten Topologie) ist homöomorph zur Menge $\{0\} \cup \{1/n \mid n \in \mathbb{N}\} \subset \mathbb{R}$.

(b) Die Einpunkt-Kompaktifizierung des $\mathbb{R}^n$ ist die n-Sphäre $\mathbb{S}^n$.

7.6 Hausdorffräume

Nach Zusammenhangs– und Kompaktheitsbegriffen diskutieren wir jetzt eine sogenannte Trennungseigenschaft, nämlich die Hausdorff-Eigenschaft oder T_2.

Definition 7.41. Ein topologischer Raum X heißt *Hausdorffsch*, falls es für alle $x, y \in X$ mit $x \neq y$ offene Mengen U, V mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$ gibt.

Beispiel 7.42. (i) Metrische Räume sind Hausdorffsch, insbesondere sind $\mathbb{R}$, $\mathbb{R}^n$ für $n \geq 2$ Hausdorffsch.

(ii) Unterräume von Hausdorffräumen sind Hausdorffsch.

(iii) Der Raum $X = \{a, b\}$ mit der Topologie $\{\emptyset, \{a\}, \{a, b\}\}$ aus Beispiel 7.3 ist nicht Hausdorffsch. Die einzige offene Menge, die b enthält, ist $\{a, b\}$, und diese Menge enthält auch a.

Satz 7.43. *Einelementige (und damit endliche) Mengen in Hausdorff-Räumen sind abgeschlossen.*

Beweis. Gemäß Definition ist $X \setminus \{x\}$ für $x \in X$ offen. □

Satz 7.44. *Sei X ein Hausdorff-Raum, $A \subseteq X$ kompakt. Dann ist A abgeschlossen.*

Beweis. Wir zeigen, dass $X \setminus A$ offen ist. Sei also $x \in X \setminus A$. Für jedes $y \in A$ wählen wir offene Mengen U_y, V_y mit $x \in U_y$, $y \in V_y$ und $U_y \cap V_y = \emptyset$. Dann ist $\{V_y\}_{y \in Y}$ eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung, d. h. es existieren $y_1, \ldots, y_m \in A$ mit $A \subseteq V_{y_1} \cup \cdots \cup V_{y_m}$. Wir setzen $U = U_{y_1} \cap \cdots \cap U_{y_m}$. Dann ist U offen, $x \in U$ und $U \cap A = \emptyset$, d. h. $U \subseteq X \setminus A$. Da $x \in X \setminus A$ beliebig gewählt war, ist $X \setminus A$ offen. □

Folgerung 7.45. *Sei $f\colon X \to Y$ bijektiv, stetig, X kompakt, Y Hausdorffsch. Dann ist f eine Homöomorphie.*

Beweis. Wir müssen zeigen, dass $f^{-1}\colon Y \to X$ stetig ist. Sei also $A \subseteq X$ abgeschlossen. Dann ist $\left(f^{-1}\right)^{-1}(A) = f(A)$ abgeschlossen in Y, da A als abgeschlossene Teilmenge eines kompakten Raumes nach Satz 7.36 kompakt und somit nach dem vorigen Resultat in Y abgeschlossen ist. □

Satz 7.46. *Sei X Hausdorfsch, A, $B \subseteq X$ kompakte Teilmengen mit $A \cap B = \emptyset$. Dann gibt es offene Mengen U, V mit $A \subseteq U$, $B \subseteq V$ und $U \cap V = \emptyset$.*

Beweis. (i) Wir behandeln zuerst den Fall, dass A eine einelementige Menge ist, d. h. $A = \{x\}$. Für jedes $y \in B$ wählen wir offene Mengen U_y, V_y mit $x \in U_y$, $y \in V_y$ und $U_y \cap V_y = \emptyset$. Dann ist $\{V_y\}_{y \in B}$ eine offene Überdeckung von B. Wir wählen eine endliche Teilüberdeckung, also $B \subseteq V_{y_1} \cup \cdots \cup V_{y_m}$. Dann sind $U = U_{y_1} \cap \cdots \cap U_{y_m}$, $V = V_{y_1} \cup \cdots \cup V_{y_m}$ offen, $x \in U$, $B \subseteq V$ und $U \cap V = \emptyset$.

(ii) Sei nun A beliebig. Für jedes $x \in A$ wählen wir gemäß (i) offene Mengen U_x, V_x mit $x \in U_x$, $B \subseteq V_x$ und $U_x \cap V_x = \emptyset$. Dann ist $\{U_x\}_{x \in A}$ eine offene Überdeckung von A. Wir wählen eine endliche Teilüberdeckung, also $A \subseteq U_{x_1} \cup \cdots \cup U_{x_m}$. Dann sind die Mengen $U = U_{x_1} \cup \cdots \cup U_{x_m}$, $V = V_{x_1} \cap \cdots \cap V_{x_m}$ offen, $A \subseteq U$, $B \subseteq V$ und $U \cap V = \emptyset$. □

Neben dem Satz von Tychonoff ist der Satz von Urysohn einer der wichtigsten Sätze der allgemeinen Topologie.

Satz 7.47 (Urysohn)**.** *Sei X ein kompakter Hausdorffraum, A, $B \subseteq X$ abgeschlossen, $A \cap B = \emptyset$. Dann existiert eine stetige Funktion $F \colon X \to [0,1]$ mit $F(x) = 0$ für alle $x \in A$ und $F(x) = 1$ für alle $x \in B$.*

Beweis. Sei I die Menge aller rationalen Zahlen der Form $k/2^j$, wobei j, $k \in \mathbb{N}$, $1 \leq k < 2^j$. Für $r \in I$ konstruieren wir induktiv offene Mengen $U_r \subseteq X$ mit

- $A \subseteq U_r$, $\overline{U}_r \subseteq X \setminus B$,
- $\overline{U}_r \subseteq U_s$ für $r < s$.

Für $j = 1$ wählen wir unter Benutzung des vorigen Satzes eine offene Menge $U_{1/2}$ mit $A \subseteq U_{1/2} \subseteq \overline{U}_{1/2} \subseteq X \setminus B$.
Für $j = 2$ wählen wir wiederum unter Benutzung des vorigen Satzes offene Mengen $U_{1/4}$, $U_{3/4}$ mit

$$A \subseteq U_{1/4} \subseteq \overline{U}_{1/4} \subseteq U_{1/2} \subseteq \overline{U}_{1/2} \subseteq U_{3/4} \subseteq \overline{U}_{3/4} \subseteq B,$$

und so weiter.
Wir definieren dann

$$F(x) = \begin{cases} \inf\{r \in I \mid x \in U_r\}, & \text{falls } x \in \bigcup_{r \in I} U_r, \\ 1, & \text{andernfalls.} \end{cases}$$

Dann ist offenbar $F(x) = 0$ für $x \in A$ und $F(x) = 1$ für $x \in B$.
Es bleibt zu zeigen, dass F stetig ist. Dazu bemerken wir, dass für $0 < r < s < 1$

$$F(x) > r \iff \exists r' > r \colon x \notin \overline{U}_{r'} \iff x \notin \bigcap_{r' > r} \overline{U}_{r'},$$

$$F(x) < s \iff \exists s' < s \colon x \in U_{s'} \iff x \in \bigcup_{s' > s} U_{s'}.$$

Also sind die Mengen

(a) $F^{-1}((r,s)) = \left(\bigcup_{s'>s} U_{s'}\right) \setminus \left(\bigcap_{r'>r} \overline{U}_{r'}\right)$,

(b) $F^{-1}([0,s)) = \bigcup_{s'>s} U_{s'}$,

(c) $F^{-1}((r,1]) = X \setminus \left(\bigcap_{r'>r} \overline{U}_{r'}\right)$

sämtlich offen in X. $\square$

Kapitel 8

Metrische Räume

8.1 Definition

Definition 8.1. Ein *metrischer Raum* (X, d) ist eine Menge X zusammen mit einer Funktion $d\colon X \times X \to [0, \infty)$ mit den folgenden Eigenschaften: Für alle x, y, $z \in X$ gilt

(i) $d(x, y) = 0 \iff x = y$,

(ii) $d(x, y) = d(y, x)$,

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

Die Funktion d heißt *Metrik* (oder *Abstandsfunktion*).

Beispiel 8.2. 1. Die durch

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

auf einer beliebigen Menge X gegebene *diskrete Metrik*.

2. $X = \mathbb{R}^n$, $1 \leq p \leq \infty$,

$$d_p(x, y) = \begin{cases} \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}, & 1 \leq p < \infty, \\ \max_{i=1,\ldots,n} |x_i - y_i|, & p = \infty, \end{cases}$$

für $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$

3. Für einen topologischen Raum Y sei $X = C_b(Y; \mathbb{R})$ die Menge aller beschränkten und stetigen Funktionen $f\colon Y \to \mathbb{R}$ mit der Metrik

$$d(f, g) = \|f - g\|_\infty = \sup_{y \in Y} |f(y) - g(y)|.$$

8.2 Die metrische Topologie

Sei (X, d) ein metrischer Raum.

Definition 8.3. Wir nennen eine Teilmenge $U \subseteq X$ offen, falls es für jedes $x \in U$ ein $r > 0$ gibt, so dass die *offene Kugel*

$$B(x, r) = \{y \in X \mid d(x, y) < r\}$$

mit Mittelpunkt x und Radius r vollständig in U enthalten ist.

Satz 8.4. *Mit diesem Begriff einer offenen Menge erhalten wir eine Topologie auf X.*

Beweis. Wir haben zu überprüfen, dass der Durchschnitt zweier offener Mengen U, V wieder offen ist. (Alle anderen Axiome einer Topologie sind offensichtlich erfüllt.)
Sei also $x \in U \cap V$ und $B(x, r) \subseteq U$, $B(x, r') \subseteq V$ mit gewissen $r > 0$, $r' > 0$. Dann $B(x, r'') \subseteq U \cap V$ mit $r'' = \min\{r, r'\}$. □

Die so definierte Topologie heißt die *durch d auf X induzierte Topologie* (oder die *metrische Topologie*).

Lemma 8.5. *Jede offene Kugel $B(x, r)$ ist offen im Sinne der obigen Definition.*

Beweis. Sei $y \in B(x, r)$, d. h. $d(x, y) < r$. Dann gilt $B(y, r') \subseteq B(x, r)$ mit $r' = r - d(x, y) > 0$, denn $d(y, z) < r'$ impliziert

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + r' = r.$$ □

Bemerkung. Die offenen Kugeln $B(x, r)$, wobei x eine dichte Teilmenge in X und r alle positiven *rationalen* Zahlen durchläuft, bilden eine Basis der durch d auf X induzierten Topologie. Insbesondere besitzt diese Topologie eine abzählbare Basis, falls es eine abzählbare dichte Teilmenge von X gibt.

Die metrische Topologie kann mit Hilfe der Konvergenz von Folgen charakterisiert werden.

Definition 8.6. (i) Eine *Folge* $\{x_n\}_{n\in\mathbb{N}}$ in X ist eine Abbildung $\mathbb{N} \to X$, $n \mapsto x_n$.
(ii) Eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset X$ heißt *konvergent* gegen eine Stelle $x^* \in X$, falls es für jedes $r > 0$ ein $n_0 = n_0(x^*, r) \in \mathbb{N}$ gibt, so dass $x_n \in B(x^*, r)$ für alle $n \geq n_0$ gilt. Wir schreiben dann $x_n \to x^*$ für $n \to \infty$.

Lemma 8.7. *Eine Folge $\{x_n\}_{n\in\mathbb{N}}$ hat höchstens einen Grenzwert.*

Beweis. Angenommen, wir hätten $x_n \to x^*$ und $x_n \to y^*$ für $n \to \infty$ mit $x^* \neq y^*$. Dann gilt für $0 < r \leq d(x^*, y^*)/2$, dass $B(x^*, r) \cap B(y^*, r) = \emptyset$, was der Definition der Folgenkonvergenz widerspricht. □

Satz 8.8. *Eine Menge A in einem metrischen Raum (X, d) ist dann und nur dann abgeschlossen, wenn die Grenzwerte aller Folgen in A, die konvergieren, in A liegen.*

Beweis. Wir zeigen tatsächlich, dass für eine beliebige Menge $B \subseteq X$ gilt:

$$x^* \in \bar{B} \iff \exists \text{ Folge } \{x_n\}_{n\in\mathbb{N}} \subset B \text{ mit } x_n \to x^* \text{ für } n \to \infty.$$

($\Longleftarrow$) Für jedes $n \in \mathbb{N}$ wählen wir ein $x_n \in B(x^*, 1/n) \cap B$. Dann gilt offenbar $x_n \to x^*$ für $n \to \infty$.

($\Longrightarrow$) Gibt es eine Folge $\{x_n\}_{n\in\mathbb{N}} \subset B$ mit $x_n \to x^*$ für $n \to \infty$, so hat jede offene Kugel $B(x^*, r)$ mit $r > 0$ einen nichtleeren Durchschnitt mit B, d. h. es gilt $x^* \in B$. □

Satz 8.9. *Eine Menge A in (X, d) ist genau dann kompakt, wenn sie abgeschlossen ist und jede Folge in A eine konvergente Teilfolge besitzt. (Äquivalent dazu ist, dass wir aus jeder Folge in A eine in A konvergente Teilfolge auswählen können.)*

Beweis. ($\Longrightarrow$) Sei $A \subseteq X$ kompakt. Dann ist A abgeschlossen, da X Hausdorffsch ist. Sei weiterhin $\{x_n\} \subset A$ eine beliebige Folge. Angenommen, $\{x_n\}$ enthält keine konvergente Teilfolge. Insbesondere kann dann kein Folgenglied unendlich oft vorkommen, so dass wir $x_m \neq x_n$ für $m \neq n$ annehmen dürfen. Jeder Punkt x_m ist ein isolierter Punkt (d. h. kein Häufungspunkt) der Menge $\{x_n \mid n \in \mathbb{N}\}$. Dann gibt es eine offene Kugel $U_m = B(x_m, r_m)$ mit $x_n \notin U_m$ für alle $n \neq m$. Man sieht leicht, dass die Menge $\{x_n \mid n \in \mathbb{N}\}$ abgeschlossen und somit die Menge $U_0 = X \setminus \{x_n \mid n \in \mathbb{N}\}$ offen in X ist. Die Mengen U_n für $n \in \mathbb{N}$ zusammen mit U_0 bilden eine offene Überdeckung von A, die keine endliche Teilüberdeckung enthält.

($\Longleftarrow$) Nun habe jede Folge in A eine in A konvergente Teilfolge. Wir zeigen zuerst:

Lemma 8.10. *Sei $\{U_i \mid i \in I\}$ eine offene Überdeckung von A. Dann gibt es ein $r > 0$, so dass es für alle $a \in A$ ein $i \in I$ mit*

$$B(a, r) \subseteq U_i$$

gibt.

Beweis. Angenommen, dies wäre nicht der Fall. Wir wählen dann eine Folge $\{x_n\} \subset A$ mit $B(x_n, 1/n) \not\subseteq U_i$ für alle i. Sodann wählen wir eine Teilfolge $\{x_{n_k}\}$ von $\{x_n\}$ mit $x_{n_k} \to x^*$ für $k \to \infty$ und einem $x^* \in A$. Es gilt $x^* \in U_{i^*}$ für ein i^* und $B(x^*, r^*) \subseteq U_{i^*}$ für ein $r^* > 0$, da U_{i^*} offen ist. Wir wählen nun ein k so groß, dass $d(x^*, x_{n_k}) \leq r^*/2$ und $1/k \leq r^*/2$ gilt. Dann haben wir für $y \in B(x_{n_k}, 1/n_k)$, dass

$$d(x^*, y) \leq d(x^*, x_{n_k}) + d(x_{n_k}, y) \leq \frac{r^*}{2} + \frac{r^*}{2} = r^*,$$

also $y \in B(x^*, r^*) \subseteq U_{i^*}$. Somit ist $B(x_{n_k}, 1/n_k) \subseteq U_{i^*}$ im Widerspruch zu $B(x_{n_k}, 1/n_k) \not\subseteq U_i$ für alle i. □

Hat nun jede Folge in A eine in A konvergente Teilfolge, so gibt es für jedes $r > 0$ endlich viele $x_1, \ldots, x_m \in A$ mit

$$A \subseteq B(x_1, r) \cup \cdots \cup B(x_m, r).$$

Andernfalls könnten wir ein $x_1 \in A$ beliebig wählen, dann ein $x_2 \in A \setminus B(x_1, r)$, dann ein $x_3 \in A \setminus (B(x_1, r) \cup B(x_2, r))$, usw., und würden auf diese Weise eine Folge $\{x_n\} \subset A$ mit $d(x_m, x_n) \geq r$ für alle $m \neq n$ erhalten, also eine Folge, aus der wir keine konvergente Teilfolge auswählen könnten. Sei nun also $\{U_i\}_{i \in I}$ eine offene Überdeckung von A. Wir wählen $r > 0$ wie in Lemma 8.10 und $x_1, \ldots, x_m \in A$ mit $A \subseteq B(x_1, r) \cup \cdots \cup B(x_m, r)$. Dann existiert für jedes $k = 1, \ldots, m$ ein $i_k \in I$ mit $B(x_k, r) \subseteq U_{i_k}$. $\{U_{i_k}\}_{k=1}^m$ ist eine endliche Teilüberdeckung von A. □

Satz 8.11. *Seien X, Y metrische Räume, $f\colon X \to Y$ eine Abbildung. Dann ist f genau dann stetig, wenn für alle Folgen $\{x_n\} \subset X$ aus $x_n \to x$ für $n \to \infty$ $f(x_n) \to f(x)$ für $n \to \infty$ folgt.*

Beweis. ($\Longrightarrow$) Sei zuerst f stetig und $\{x_n\}$ eine Folge in X mit $x_n \to x$ für $n \to \infty$. Für jedes $\epsilon > 0$ gibt es wegen der Stetigkeit von f ein $\delta > 0$ mit

$$f\left(B(x, \delta)\right) \subseteq B(f(x), \epsilon).$$

Wir wählen ein n_0, so dass $x_n \in B(x, \delta)$ für alle $n \geq n_0$. Dann ist $f(x_n) \in B(f(x), \epsilon)$ für alle $n \geq n_0$. Es folgt, dass $f(x_n) \to f(x)$ für $n \to \infty$.

($\Longleftarrow$) Nun habe f umgekehrt die in der Formulierung des Satzes genannte Eigenschaft. Sei $B \subseteq Y$ abgeschlossen. Wir wollen zeigen, dass $A = f^{-1}(B)$ ebenfalls abgeschlossen ist. Sei also $\{x_n\} \subset A$ eine konvergente Folge mit $x_n \to x$ für $n \to \infty$. Zu zeigen ist, dass $x \in A$ gilt.
Aus $x_n \to x$ für $n \to \infty$ folgt, dass $f(x_n) \to f(x)$ für $n \to \infty$. Wegen $\{f(x_n)\} \subset B$ und da B abgeschlossen ist, gilt $f(x) \in B$. Damit ist jedoch $x \in f^{-1}(B) = A$. □

8.3 Vollständige metrische Räume

Sei (X, d) ein metrischer Raum.

Definition 8.12. (i) $\{x_n\} \subset X$ heißt *Cauchyfolge*, falls

$$\forall \epsilon > 0 \ \exists n_0 = n_0(\epsilon) \ \forall m, n \geq n_0\colon d(x_m, x_n) \leq \epsilon.$$

(ii) X heißt *vollständig*, falls jede Cauchyfolge konvergiert.

Beispiel 8.13. (a) $\mathbb{R}$, $[a, b]$ sind vollständig.

(b) Vollständigkeit ist keine topologische Eigenschaft. Beispielsweise ist offene Intervall $(0, 1)$ zwar homöomorph zu $\mathbb{R}$, aber *nicht* vollständig. ($\{1/n\}_{n \in \mathbb{N}}$ ist eine Cauchyfolge in $(0, 1)$, die jedoch in $(0, 1)$ nicht konvergiert.)

Definition 8.14. Ein metrischer Raum $(\bar{X}, \bar{d})$ heißt *Vervollständigung* von (X, d), falls es eine isometrische (d. h. eine metrik-erhaltende) Abbildung

$$i\colon (X, d) \to (\bar{X}, \bar{d})$$

gibt, so dass $i(X)$ in $\bar{X}$ dicht ist (d. h. der Abschluss von $i(X)$ ist $\bar{X}$).

Theorem 8.15. *Für jeden metrischen Raum gibt es bis auf Isometrie genau eine Vervollständigung.*

Beispiel 8.16. (i) $\mathbb{Q}$ ist nicht vollständig, seine Vervollständigung ist $\mathbb{R}$.

(ii) Die Vervollständigung von $(0,1)$ ist $[0,1]$.

Beweis von Theorem 8.15. (a) (Eindeutigkeit) Seien $\bar{X}$, $\bar{X}'$ zwei Vervollständigungen. Wir können X sowohl als Unterraum von $\bar{X}$ als auch als Unterraum von $\bar{X}'$ betrachten. Dann setzt sich die identische Abbildung $X \to X$, $x \mapsto x$, in eindeutiger Weise zu einer Isometrie $f\colon \bar{X} \to \bar{X}'$ fort.

(b) (Existenz) Sei $\mathscr{X}$ der Raum aller Cauchyfolgen in X. Auf $\mathscr{X}$ betrachten wir die folgende Relation:

$$\{x_n\} \sim \{y_n\} :\Longleftrightarrow \big\{d(x_n, y_n)\big\} \text{ ist Nullfolge in } \mathbb{R}.$$

$\sim$ ist eine Äquivalenzrelation auf $\mathscr{X}$: $\sim$ ist offenbar reflexiv und symmetrisch. Die Transitivität ergibt sich wie folgt:

Sei $\{x_n\} \sim \{y_n\}$ und $\{y_n\} \sim \{z_n\}$, d. h. $\big\{d(x_n, y_n)\big\}$ und $\big\{d(y_n, z_n)\big\}$ sind Nullfolgen. Dann ist wegen

$$0 \le d(x_n, z_n) \le d(x_n, y_n) + d(y_n, z_n) \to 0 \text{ für } n \to \infty$$

auch $\big\{d(x_n, z_n)\big\}$ eine Nullfolge.

Sei $\bar{X}$ der Raum der Äquivalenzklassen von $\mathscr{X}$ bezüglich $\sim$. Auf $\bar{X}$ definieren wir die Metrik

$$\bar{d}\left([\{x_n\}], [\{y_n\}]\right) = \lim_{n\to\infty} d(x_n, y_n).$$

(i) *Zu zeigen:* $\bar{d}$ ist wohldefiniert.

(a) Seien $\{x_n\}$, $\{y_n\}$ Cauchyfolgen in X. Dann ist $\big\{d(x_n, y_n)\big\}$ wegen

$$\begin{aligned} &|d(x_m, y_m) - d(x_n, y_n)| \\ &\qquad \le |d(x_m, y_m) - d(x_n, y_n)| + |d(x_m, y_m) - d(x_n, y_n)| \\ &\qquad \le d(x_m, x_n) + d(y_m, y_n) \to 0 \end{aligned}$$

für $m, n \to \infty$ eine Cauchyfolge in $\mathbb{R}$. Also existiert $\lim\limits_{n\to\infty} d(x_n, y_n)$.

(b) (Repräsentantenunabhängigkeit) Sei $\{x'_n\}$ ein weiterer Repräsentant der Klasse $[\{x_n\}]$, d. h. $\{x_n\} \sim \{x'_n\}$. Dann gilt

$$|d(x_n, y_n) - d(x'{}_n, y_n)| \le d(x_n, x'{}_n) \to 0$$

für $n \to \infty$, d. h. die Folgen $\big\{d(x_n, y_n)\big\}$ und $\big\{d(x'{}_n, y_n)\big\}$ haben denselben Grenzwert.

(ii) Wir zeigen als Nächstes, dass der Raum $\bar{X}$ vollständig ist. Dazu wählen wir eine Cauchyfolge $\{\xi^k\}_{k\in\mathbb{N}}$ in $\bar{X}$. Wir repräsentieren jedes Glied ξ^k dieser Folge durch eine Cauchyfolge $\{x_n^k\}_{n\in\mathbb{N}}$ in X, d. h. $\xi^k = \big[\{x_n^k\}\big]$. Dann wählen wir eine Folge $\{y^k\}$ in X, von der wir zeigen werden:

(a) $\{y^k\}$ ist eine Cauchyfolge,

(b) $\{y^k\}$ repräsentiert den Grenzwert der Folge $\{\xi^k\}$.

Wir setzen $y^k = x^k_{n_k}$, wobei n_k so gewählt ist, dass $d(x^k_m, x^k_n) \leq 1/k$ für m, $n \geq n_k$ gilt.

Zu (a): Sei $\epsilon > 0$. Wir finden dann ein $k_0 \geq 1/\epsilon$, so dass

$$\lim_{n\to\infty} d(x^k_n, x^l_n) \leq \epsilon \text{ für } k,\, l \geq k_0.$$

Dann gilt für k, $l \geq k_0$

$$\begin{aligned} d(y^k, y^l) &\leq d(x^k_{n_k}, x^k_n) + d(x^k_n, x^l_n) + d(x^l_n, d^l_{n_l}) \\ &\leq \frac{1}{k} + 2\epsilon + \frac{1}{l} \leq 4\epsilon \end{aligned}$$

für ein hinreichend großes n. Also ist $\{y^k\}$ eine Cauchyfolge in X.

Zu (b): Sei $\epsilon > 0$. Indem wir k_0 wie in (a) wählen, erhalten wir für $k \geq k_0$, dass

$$\begin{aligned} d(y^m, x^k_m) &\leq d(x^m_{n_k}, x^m_n) + d(x^m_n, x^k_n) + d(x^k_n, x^k_m) \\ &\leq \frac{1}{m} + 2\epsilon + \frac{1}{k} \leq 4\epsilon \end{aligned}$$

für $m \geq k_0$ und n hinreichend groß. Insbesondere

$$\lim_{m\to\infty} d(y^m, x^k_m) \leq 4\epsilon \text{ für } k \geq k_0.$$

Folglich gilt $\lim_{k\to\infty} \xi^k = [\{y^m\}]$ in $\bar{X}$.

(iii) Die isometrische Abbildung von X nach $\bar{X}$ ergibt sich, indem jedes $x \in X$ auf die Klasse abgebildet wird, die durch die Folge konstant x repräsentiert wird. Indem wir X mit seinem Bild in $\bar{X}$ unter dieser Abbildung identifizieren, sehen wir, dass X in $\bar{X}$ dicht ist: Es gilt

$$\lim_{n\to\infty} x_n = [\{x_n\}]$$

in $\bar{X}$ für jede Cauchyfolge $\{x_n\}$ in X. □

Beispiel 8.17 (Beispiel eines vollständigen metrischen Raumes).

$$X = C([0,1], \mathbb{R}) = \big\{ f\colon [0,1] \to \mathbb{R} \;\big|\; f \text{ ist stetig} \big\}$$

mit der Metrik

$$d(f,g) = \max_{x\in[0,1]} |f(x) - g(x)|\,.$$

Behauptung. X ist vollständig.

Beweis. Sei dazu $\{f^n\}$ eine Cauchyfolge. Dann ist für jedes $x \in [0,1]$ die Folge $\{f^n(x)\}$ eine Cauchyfolge in $\mathbb{R}$. Somit konvergiert die Folge $\{f^n\}$ punktweise gegen eine Grenzfunktion f, d. h. $\lim\limits_{n\to\infty} f^n(x) = f(x)$ für alle $x \in [0,1]$. Diese Konvergenz ist tatsächlich gleichmäßig, da aus

$$\max_{x\in[0,1]} |f^m(x) - f^n(x)| \leq \epsilon$$

für ein gegebenes $\epsilon > 0$ und m, $n \geq n_0$ bei geeeignetem n_0 durch Grenzübergang $m \to \infty$ folgt, dass

$$\sup_{x\in[0,1]} |f(x) - f^m(x)| \leq \epsilon$$

für $n \geq n_0$ gilt. Damit ist die Funktion f gemäß Satz 6.7 stetig, gehört als zum Raum X.

8.4 Der Banachsche Fixpunktsatz

Sei (X,d) ein vollständiger metrischer Raum[1]. Sei ferner $T\colon X \to X$ eine Abbildung mit
$$d\,(Tx,Ty) \le \alpha\, d(x,y), \quad \forall x,\, y \in X,$$
für ein $0 \le \alpha < 1$. Eine solche Abbildung heißt *strikt kontraktiv*.

Satz 8.18 (Banach)**.** *Die strikt kontraktive Abbildung $T\colon X \to X$ in einem vollständigen metrischen Raum X besitzt genau einen Fixpunkt.*

Beweis. (a) (Eindeutigkeit) Seien $x,\, y \in X$ Fixpunkte von T. Dann gilt
$$d(x,y) = d\,(Tx,Ty) \le \alpha\, d(x,y),$$
also $(1-\alpha)\, d(x,y) \le 0$, $d(x,y) = 0$ und $x = y$.

(b) (Existenz) Zuerst erhalten wir induktiv, dass
$$d\,(T^n x, T^n y) \le \alpha^n\, d(x,y), \quad \forall x,\, y \in X.$$
Sei nun $x_0 \in X$ beliebig und $x_n = T^n x_0$ für $n \in \mathbb{N}$. Dann gilt
$$d(x_{n+1},x_n) = d\,(T^n x_1, T^n x_0) \le \alpha^n d(x_1,x_0),$$
also für $n,\, k \in \mathbb{N}$
$$\begin{aligned} d(x_{n+k},x_n) &\le d(x_{n+k},x_{n+k-1}) + \cdots + d(x_{n+2},x_{n+1}) + d(x_{n+1},x_n) \\ &\le \left(\alpha^{n+k-1} + \cdots + \alpha^{n+1} + \alpha^n\right) d(x_1,x_0) \\ &= \alpha^n\, \frac{1-\alpha^k}{1-\alpha}\, d(x_1,x_0) \le \frac{\alpha^n}{1-\alpha}\, d(x_1,x_0). \end{aligned}$$
Wegen $\frac{\alpha^n}{1-\alpha} \to 0$ für $n \to \infty$ ist $\{x_n\}$ eine Cauchyfolge in X. Wegen der Vollständigkeit von X konvergiert $\{x_n\}$ gegen ein $x^* \in X$. Indem wir in
$$d\,(Tx_{n+1},x_{n+1}) \le \alpha\, d\,(Tx_n,x_n)$$
zum Grenzwert für $n \to \infty$ übergehen, erhalten wir
$$d(Tx^*,x^*) \le \alpha\, d(Tx^*,x^*),$$
also $d(Tx^*,x^*) = 0$ und $Tx^* = x^*$. Folglich ist x^* der gesuchte Fixpunkt. □

Bemerkung. In Anwendungen ist es oft wichtig, dass dieser Beweis ein konstruktives Verfahren zur Approximation des gesuchten Fixpunktes x^* liefert: Für ein *beliebiges* x_0 in X gilt $\lim\limits_{n\to\infty} T^n x_0 = x^*$, wobei wir die Fehlerabschätzung
$$d\,(x^*, T^n x_0) \le \frac{\alpha^n}{1-\alpha}\, d(Tx_0,x_0)$$
haben.

[1] Das ist beispielsweise der Fall, wenn X (als topologischer Raum) kompakt ist.

Beispiel 8.19. Sei $X = C([0,1],\mathbb{R})$ und $k\colon [0,1]\times[0,1]\times\mathbb{R}\to\mathbb{R}$ stetig. Wir betrachten den Integraloperator $K\colon X\to X$,

$$(Kf)(x) = f(x) + \int_0^1 k(x,y,f(y))\,dy, \quad x\in[0,1].$$

Unter der Voraussetzung

$$|k(x,y,u) - k(x,y,v)| \le \alpha\,|u-v|$$

für $x,y\in[0,1]$, $u,v\in\mathbb{R}$ und ein $\alpha<1$ wollen wir zeigen, dass der Operator $K\colon X\to X$ *bijektiv* ist. Dazu wählen wir ein $h\in X$ und formulieren die Gleichung $Kf = h$ für das gesuchte f als Fixpunktgleichung $f = Tf$,

$$Tf(x) = h(x) - \int_0^1 k(x,y,f(y))\,dy, \quad x\in[0,1].$$

Wegen

$$\begin{aligned} |Tf(x) - Tg(x)| &\le \int_0^1 |k(x,y,f(y)) - k(x,y,g(y))|\,dy \\ &\le \alpha\int_0^1 |f(y)-g(y)|\,dy \le \alpha \max_{x\in[0,1]} |f(x)-g(x)| \end{aligned}$$

gilt

$$d(Tf,Tg) \le \alpha d(x,y)$$

für alle f, $g\in X$. Damit hat der Operator T genau einen Fixpunkt, d. h. bei gegebenem $h\in X$ hat die Gleichung $Kf = h$ genau eine Lösung $f\in X$.

Es gilt folgende Verallgemeinerung des Banachschen Fixpunktsatzes:

Satz 8.20. *Sei $T\colon X\to X$ eine Abbildung in einem vollständiger metrischer Raum X, so dass T^m für ein $m\in\mathbb{N}$ strikt kontraktiv ist. Dann besitzt T genau einen Fixpunkt.*

Beweis. Sei $d(T^m x, T^m y) \le \alpha d(x,y)$ für alle x, $y\in X$ und ein $0\le\alpha<1$.
T^m besitzt gemäß dem Banachschen Fixpunktsatz genau einen Fixpunkt x^*. Weiterhin besitzt T höchstens x^* als Fixpunkt, da jeder Fixpunkt von T auch Fixpunkt von T^m ist. Andererseits ist wegen

$$d(Tx^*,x^*) = d(T^{m+1}x^*, T^m x^*) \le \alpha d(Tx^*,x^*),$$

also $d(Tx^*,x^*) = 0$, x^* tatsächlich ein Fixpunkt von T. □

Kapitel 9

Differenziation von Funktionen mehrerer Veränderlicher

Im Folgenden bezeichnet $\mathbb{R}^n$ den *n-dimensionalen Euklidischen Raum* mit seiner Standardtopologie. Stellen im $\mathbb{R}^n$ bezeichnen wir mit $x = (x_1, x_2, \dots, x_n)^T$, d. h. x ist der Spaltenvektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Benötigen wir einen Abstandsbegriff im $\mathbb{R}^n$, so arbeiten wir mit der *Euklidischen Metrik*

$$|x - y| = \left(\sum_{k=1}^{n} |x_k - y_k|^2 \right)^{1/2}$$

für $x = (x_1, \dots, x_n)^T$, $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$. Insbesondere bezeichnet $|x| = \left(\sum\limits_{k=1}^{n} x_k^2 \right)^{1/2}$ den Betrag von x.

Mit den beiden vorangegangenen Kapiteln sind die Begriffe Stetigkeit einer Funktion $f \colon U \to \mathbb{R}^m$ für $U \subseteq \mathbb{R}^n$, Folgenstetigkeit, Kompaktheit, usw. klar. Insbesondere gilt im $\mathbb{R}^n$ der *Satz von Heine-Borel*, wonach eine Teilmenge von $\mathbb{R}^n$ genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist.

9.1 Partielle Ableitungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f \colon U \to \mathbb{R}$.

Definition 9.1. (i) Die *partielle Ableitung* der Funktion f nach der Veränderlichen x_k an der Stelle $x^0 = (x_1^0, \dots, x_n^0) \in U$ (falls existent) ist die Ableitung der Funktion

$$g(t) = f\left(x_1^0, \dots, x_{k-1}^0, t, x_{k+1}^0, \dots, x_n^0\right)$$

an der Stelle x_k^0. Man schreibt $\frac{\partial f}{\partial x_k}(x^0)$ oder auch $f_{x_k}(x^0)$.

(ii) Die Funktion f heißt in U partiell nach der Veränderlichen x_k differenzierbar, falls $\frac{\partial f}{\partial x_k}$ an jeder Stelle in U existiert.

Die Differenziationsregeln für Summen, Differenzen, Produkte, Quotienten und zusammengesetzte Funktionen übertragen sich direkt vom eindimensionalen auf den mehrdimensionalen Fall.

Beispiel 9.2. Im $\mathbb{R}^2$ schreiben wir auch $(x, y) = (x_1, x_2)$. Die Funktion $f(x, y) = \tan(x^2 + y)$ ist in allen Stellen (x, y) mit $x^2 + y \notin \frac{\pi}{2} + \mathbb{Z}$ sowohl nach x als auch nach y partiell differenzierbar mit

$$\frac{\partial f}{\partial x}(x, y) = \frac{2x}{\cos^2(x^2 + y)}, \qquad \frac{\partial f}{\partial y}(x, y) = \frac{1}{\cos^2(x^2 + y)}.$$

Aus der Tatsache, dass alle partiellen Ableitungen an der Stelle x^0 existieren, folgt nicht, dass f an x^0 stetig ist.

Beispiel 9.3. Die Funktion

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0), \end{cases}$$

besitzt partielle Ableitungen auf ganz $\mathbb{R}^2$, ist an $(0, 0)$ aber nicht stetig: Es gilt $f(x, 0) = f(0, y) = 0$, also $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, jedoch $f(x, x) = \frac{1}{2}$, $f(x, -x) = -\frac{1}{2}$.

9.2 Die totale Ableitung

Sei $U \subseteq \mathbb{R}^n$ offen, $f \colon U \to \mathbb{R}^m$. In Komponenten schreiben wir dann auch $f = (f_1, \dots, f_n)^T$, wobei $f_j \colon U \to \mathbb{R}$ für $j = 1, \dots, m$.

Definition 9.4. (i) f heißt an der Stelle x^0 *differenzierbar*, falls es eine Matrix $A \in M_{m \times n}(\mathbb{R})$ gibt, so dass

$$f(x^0 + h) = f(x^0) + Ah + \rho(h)$$

für alle $h \in \mathbb{R}^n$ mit $x^0 + h \in U$, wobei $\lim\limits_{h \to 0} \frac{\rho(h)}{|h|} = 0$.

(ii) f heißt in U differenzierbar, falls f an jeder Stelle von U differenzierbar ist.

Lemma 9.5. *Die Matrix A in* (i) *ist eindeutig bestimmt.*

Beweis. Sei $f(x^0 + h) = f(x^0) + Bh + \rho'(h)$ für alle $h \in \mathbb{R}^n$ mit $x^0 + h \in U$, wobei $\lim\limits_{h \to 0} \frac{\rho'(h)}{|h|} = 0$. Dann gilt $(A - B)h = -\rho(h) + \rho'(h)$, also

$$\lim_{h \to 0} (A - B) \frac{h}{|h|} = 0.$$

Daraus folgt $A - B = 0$ bzw. $A = B$. $\square$

Definition 9.6. Die Matrix A in (i) (falls existent) heißt die (*totale*) *Ableitung* von f an der Stelle x^0. Man schreibt $A = Df(x^0)$.

Satz 9.7. *Ist f an der Stelle x^0 differenzierbar, so ist f an der Stelle x^0 stetig.*

Beweis. Folgt direkt aus der Gleichung

$$f(x^0 + h) = f(x^0) + Df(x^0)h + \rho(h),$$

siehe auch Satz 3.3. □

Satz 9.8. *Ist f an der Stelle x^0 differenzierbar, so existieren die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}(x^0)$ für $j = 1, \dots, m$, $k = 1, \dots, n$ und*

$$Df = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

jeweils an der Stelle x^0, wobei $\frac{\partial f}{\partial x_k}$ der Spaltenvektor $\left(\frac{\partial f_1}{\partial x_k}, \dots, \frac{\partial f_m}{\partial x_k}\right)^T$ ist.

Beweis. Ist $A = Df(x^0)$ und $e_k = (0, \dots, 0, 1, 0, \dots, 0)^T$ mit der 1 an der k-ten Stelle, so ist Ae_k die k-te Spalte von A. Wir erhalten

$$\frac{\partial f}{\partial x_k}(x^0) = \lim_{s \to 0} \frac{f(x^0 + se_k) - f(x^0)}{s} = \lim_{s \to 0} \left(Ae_k + \frac{\rho(se_k)}{s}\right) = Ae_k. \qquad \square$$

Definition 9.9. Die Matrix $\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$ an der Stelle x^0 heißt auch die *Jacobi-Matrix* oder *Jacobische* von f an der Stelle x^0.

Umgekehrt zum vorigen Satz gilt:

Satz 9.10. *Die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}$ für $j = 1 \dots, m$, $k = 1, \dots, n$ mögen in einer kleinen Umgebung von x^0 existieren und an x^0 stetig sein. Dann ist f an der Stelle x^0 differenzierbar.*

Beweis. O. B. d. A. sei $m = 1$, $f_1 = f$. Wir definieren

$$\rho(h) = f(x^0 + h) - f(x^0) - \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x^0)h_k$$

für $h = (h_1, \dots, h_n)^T \in \mathbb{R}^n$ mit $x^0 + h \in U$. Wir setzen weiterhin $x^1 = x^0 + h_1 e_1$, $x^2 = x^1 + h_2 e_2$, …, $x^n = x^{n-1} + h_n e_n = x^0 + h$. Dann gilt

$$\begin{aligned} \rho(h) &= \sum_{k=1}^{n} \left(f(x^k) - f(x^{k-1}) - \frac{\partial f}{\partial x_k}(x^0)h_k\right) \\ &= \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}\left(x^{k-1} + \vartheta_k h_k e_k\right) - \frac{\partial f}{\partial x_k}(x^0)\right) h_k \end{aligned}$$

für gewisse $\vartheta_k \in (0, 1)$ nach dem Mittelwertsatz. Also gilt $\lim\limits_{h \to 0} \frac{\rho(h)}{|h|} = 0$, da $x^{k-1} + \vartheta_k h_k e_k \to x^0$ für $h \to 0$. □

Bemerkung. Im Fall $m = 1$ nennt man $\nabla f(x^0) = Df(x^0)$ auch den *Gradienten* von f an der Stelle x^0. Geometrisch zeigt der Gradient in die Richtung des steilsten Anstiegs von f an der Stelle x^0.

Unter der Benutzung der Matrizenschreibweise formuliert sich die *Kettenregel* sehr übersichtlich.

Satz 9.11. *Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, $f\colon U \to V$ an $x^0 \in U$ differenzierbar und $g\colon V \to \mathbb{R}^p$ an $y^0 = f(x^0) \in V$ differenzierbar. Dann ist $g \circ f\colon U \to \mathbb{R}^p$ an x^0 differenzierbar und*

$$D\left(g \circ f\right)(x^0) = Dg(y^0)Df(x^0).$$

Beweis. Wir schreiben

$$\begin{aligned} f(x^0 + h) &= f(x^0) + Df(x^0)h + \rho(h), \\ g(y^0 + k) &= g(y^0) + Dg(y^0)k + \sigma(k) \end{aligned}$$

mit $\lim\limits_{h \to 0} \frac{\rho(h)}{|h|} = 0$, $\lim\limits_{k \to 0} \frac{\sigma(k)}{|k|} = 0$. Dann gilt

$$\begin{aligned} (g \circ f)(x^0 + h) &= g\left(f(x^0) + Df(x^0)h + \rho(h)\right) \\ &= (g \circ f)(x^0) + Dg(y^0)\left(Df(x^0)h + \rho(h)\right) + \sigma\left(Df(x^0)h + \rho(h)\right) \\ &= (g \circ f)(x^0) + Dg(y^0)Df(x^0)h + \left(Dg(y^0)\rho(h) + \sigma\left(Df(x^0)h + \rho(h)\right)\right) \end{aligned}$$

Berücksichtigen wir, dass es für jedes $A \in M(m, n; \mathbb{R})$ eine Konstante $C \geq 0$ mit $|Ah| \leq C\,|h|$ für alle $h \in \mathbb{R}^n$ gibt (die kleinstmögliche Wahl für C ist $\sup\limits_{|h|=1} |Ah|$), so erhalten wir mit einer geeigneten Konstanten C, dass

$$\frac{\left|Dg(y^0)\rho(h)\right|}{|h|} \leq C\,\frac{|\rho(h)|}{|h|} \to 0 \text{ für } h \to 0.$$

Weiterhin gilt mit

$$\chi(x^0, h) = \begin{cases} \dfrac{\sigma(Df(x^0)h + \rho(h))}{|Df(x^0)h + \rho(h)|}, & Df(x^0)h + \rho(h) \neq 0, \\ 0, & Df(x^0)h + \rho(h) = 0, \end{cases}$$

dass $\sigma(Df(x^0)h + \rho(h)) = \chi(x^0, h)(Df(x^0)h + \rho(h))$ sowie mit einer geeigneten Konstanten C

$$\begin{aligned} \frac{\left|\sigma\left(Df(x^0)h + \rho(h)\right)\right|}{|h|} &= \left|\chi(x^0, h)\right| \cdot \frac{\left|Df(x^0)h + \rho(h)\right|}{|h|} \\ &\leq \left|\chi(x^0, h)\right| \cdot \left(C\,\frac{|h|}{|h|} + \frac{|\rho(h)|}{|h|}\right) \to 0 \text{ für } h \to 0 \end{aligned}$$

wegen $\chi(x^0, h) \to 0$ für $h \to 0$. $\square$

9.3 Höhere Ableitungen

Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$.

Definition 9.12. Sei f in U nach der Veränderlichen x_k partiell differenzierbar, $\frac{\partial f}{\partial x_k}\colon U \to \mathbb{R}$. Ist dann $\frac{\partial f}{\partial x_k}$ an der Stelle $x^0 \in U$ nach x_j differenzierbar, d. h. $\frac{\partial(\partial f/\partial x_k)}{\partial x_j}(x^0)$ existiert, so schreiben wir dafür $\frac{\partial^2 f}{\partial x_j \partial x_k}(x^0)$ oder auch $f_{x_j x_k}(x^0)$. Im Fall $j = k$ schreiben wir $\frac{\partial^2 f}{\partial x_k^2}$.

Induktiv definieren wir höhere Ableitungen $\dfrac{\partial^m f}{\partial x_{k_1} \dots \partial x_{k_m}}(x^0)$, wobei $k_r \in \{1, 2 \dots, n\}$ für $r = 1, \dots, m$.

Satz 9.13 (Schwarz)**.** *Es mögen die Ableitungen $\frac{\partial^2 f}{\partial x_j \partial x_k}$ und $\frac{\partial^2 f}{\partial x_k \partial x_j}$ in einer Umgebung von x^0 existieren und an x^0 stetig sein. Dann gilt*

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(x^0) = \frac{\partial^2 f}{\partial x_k \partial x_j}(x^0).$$

Beweis. Indem wir die Veränderlichen x_i mit $i \neq j, k$ als Parameter betrachten, genügt es den Fall $n = 2$ zu behandeln, $(x, y) = (x_j, x_k)$. Wir behaupten, dass

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

In der Tat ergibt eine Anwendung des Mittelwertsatzes auf die Funktion $y \mapsto f(x^0 + h, y) - f(x^0, y)$ für $h \neq 0$, $|h|$ klein, dass

$$\begin{aligned}
&\frac{f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - f(x^0, y^0 + h) + f(x^0, y^0)}{h} \\
&\qquad\qquad = f_y(x^0 + h, y^0 + \vartheta'_h h) - f_y(x^0, y^0 + \vartheta'_h h)
\end{aligned}$$

für ein $\vartheta'_h \in (0, 1)$. Nochmalige Anwendung des Mittelwertsatzes, diesmal auf $x \mapsto f_y(x, y^0 + \vartheta_h h)$, liefert

$$\begin{aligned}
&\frac{f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - f(x^0, y^0 + h) + f(x^0, y^0)}{h^2} \\
&\qquad\qquad = f_{xy}(x^0 + \vartheta_h h, y^0 + \vartheta'_h h)
\end{aligned}$$

für ein $\vartheta_h \in (0, 1)$. Im Grenzübergang $h \to 0$ folgt, dass der Grenzwert der linken Seite gleich $\frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)$ ist.

Genauso berechnet sich dieser Grenzwert zu $\frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)$, indem man zuerst den Mittelwertsatz auf $x \mapsto f(x, y^0 + h) - f(x, y^0)$ für $h \neq 0$, $|h|$ klein anwendet. $\square$

Bemerkung. Ganz analoge Resultate gelten für höhere partielle Ableitungen, vorausgesetzt, dass diese stetig sind, z. B.

$$\frac{\partial^3 f}{\partial x \partial y \partial x}(x^0) = \frac{\partial^3 f}{\partial x^2 \partial y}(x^0).$$

Analog definiert man höhere Ableitungen im Fall der totalen Ableitung. Sei dazu $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^m$. Ist f in U differenzierbar, so ist $Df\colon U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, wobei wir mit $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ den *Raum der linearen Abbildungen von* $\mathbb{R}^n$ *nach* $\mathbb{R}^m$ bezeichnen, den wir mit $M_{m\times n}(\mathbb{R})$ identifizieren.

Definition 9.14. Ist Df in $x^0 \in U$ differenzierbar, so bezeichnen wir die Ableitung mit $D^2 f(x^0)$. Analog definieren wir höhere Ableitungen $D^b f(x^0)$ für $b \geq 3$.

Bemerkung. Gemäß Definition ist $D^b f(x^0)$ eine b-lineare Abbildung[1] $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{b\text{-mal}} \to \mathbb{R}^m$, wobei wir beispielsweise im Fall $b = 2$ den Raum $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ mit dem Raum der bilinearen Abbildungen $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ identifizieren[2]. Genauer gilt

$$\begin{aligned} &D^b f(x^0 + k)(h_1, \ldots, h_b) \\ &\qquad = D^b f(x^0)(h_1, \ldots, h_b) + D^{b+1} f(x^0)(h_1, \ldots, h_b, k) + o(|k|) \end{aligned}$$

für $k \to 0$, wobei $o(|k|)$ eine (nicht weiter spezifizierte) Funktion $\varrho(k)$ mit der Eigenschaft $\lim\limits_{k\to 0} \frac{\varrho(k)}{|k|} = 0$ bezeichnet.

Lemma 9.15. *Die b-lineare Abbildung*

$$D^b f(x^0)\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{b\text{-mal}} \to \mathbb{R}^m$$

ist total-symmetrisch, d. h.

$$D^b f(x^0)(h_{\sigma(1)}, \ldots, h_{\sigma(b)}) = D^b f(x^0)(h_1, \ldots, h_b)$$

für alle Permutationen σ von $\{1, \ldots, b\}$.

Beweis. Wie im Beweis des Satzes von Schwarz zeigt man, dass

$$\begin{aligned} D^2 f(x^0)(h,k) &= \lim_{s\to 0} \frac{f(x^0 + sh + sk) - f(x^0 + sh) - f(x^0 + sk) + f(x^0)}{s^2} \\ &= D^2 f(x^0)(k,h). \end{aligned}$$

Analog für höhere Ableitungen. □

Satz 9.16. *Ist $f\colon U \to \mathbb{R}^m$ an der Stelle $x^0 \in U$ b-mal differenzierbar, so ist f an der Stelle x^0 b-mal partiell differenzierbar, und es gilt*

$$\frac{\partial^b f}{\partial x_{k_1} \ldots \partial x_{k_b}}(x^0) = \left(D^b f\right)(x^0)(e_{k_1}, \ldots, e_{k_b}),$$

wobei e_k den k-ten Einheitsvektor bezeichnet.

[1] Eine Abbildung $B\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^m$ heißt *multilinear*, wenn B linear in jedem Argument ist.

[2] We identifizieren die Abbildung $D \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ mit der bilinearen Abbildung $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ vermöge

$$B(h,k) = D(h)k, \quad \forall h,\, k \in \mathbb{R}^n,$$

wobei wir den Umstand $D(h) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ausnutzen.

Beweis. Der Fall $b = 1$ wurde in Satz 9.8 erledigt. Für $b = 2$ gilt

$$\frac{1}{t}\left(f(x^0 + se_j + te_k) - f(x^0 + se_j) - f(x^0 + te_k) + f(x^0)\right) \to \frac{\partial f}{\partial x_k}(x^0 + se_j) - \frac{\partial f}{\partial x_k}(x^0)$$

für $t \to 0$ und

$$\frac{1}{s}\left(\frac{\partial f}{\partial x_k}(x^0 + se_j) - \frac{\partial f}{\partial x_k}(x^0)\right) \to \frac{\partial^2 f}{\partial x_j \partial x_k}(x^0)$$

für $s \to 0$. Damit folgt das Ergebnis mittels der Formel im Beweis des vorangegangenen Satzes.
Höhere Ableitungen behandelt man analog. ☐

Bemerkung. Im Fall $m = 1$ heißt die symmetrische Matrix[3]

$$D^2 f(x^0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}(x^0)$$

die *Hesse-Matrix* von f an der Stelle x^0.
Umgekehrt gilt:

Satz 9.17. *Ist f b-mal partiell differenzierbar in einer Umgebung von $x^0 \in U$ und sind die b-ten partiellen Ableitungen von f an x^0 stetig, so existiert $D^b f(x^0)$.*

Beweis. Mittels Induktion nach b unter Benutzung des Satzes 9.10. ☐

9.4 Die Taylorformel

Sei $U \subseteq \mathbb{R}^n$ offen $f\colon U \to \mathbb{R}$.

Satz 9.18. *Sei f $(b+1)$-mal differenzierbar. Liegt dann die Strecke von x^0 nach $x^0 + h$ vollständig in U, so gilt*

$$f(x^0 + h) = \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!} \underbrace{(h, \ldots, h)}_{a\text{-mal}} + \frac{D^{b+1} f(x^0 + \vartheta h)}{(b+1)!} \underbrace{(h, \ldots, h)}_{(b+1)\text{-mal}}$$

für ein $\vartheta \in (0, 1)$.

[3] Eine Matrix $A \in M_n(\mathbb{R})$ heißt *symmetrisch*, falls $A = A^T$ gilt.
Wir identifizieren den Raum der *symmetrischen Bilinearformen* $a\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit dem Raum symmetrischen Matrizen A vermöge der Abbildung $a \mapsto A = \left(A_{jk}\right)_{j,k=1}^n$, wobei $A_{jk} = a(e_j, e_k)$, und wir die Bilinearform a aus A durch

$$a(h, k) = \sum_{j,k=1}^{n} A_{jk} h_j h_k$$

für $h = (h_1, \ldots, h_n)^T$, $k = (k_1, \ldots, k_n)^T \in \mathbb{R}^n$ zurückgewinnen.

Beweis. Die Funktion $g\colon [0,1] \to \mathbb{R}$, $t \mapsto f(x^0+th)$ ist $(b+1)$-mal differenzierbar, wobei

$$g^{(a)}(t) = D^a f(x^0 + th)\,\underbrace{(h,\dots,h)}_{a\text{-mal}}$$

für $a \le b+1$. Die Behauptung ergibt sich aus der Taylorformel mit Lagrangeschem Restglied gemäß Satz 6.31:

$$g(1) = \sum_{a=0}^{b} \frac{g^{(a)}(0)}{a!} + \frac{g^{(b+1)}(\vartheta)}{(b+1)!}$$

für ein $\vartheta \in (0,1)$. $\square$

Bemerkung. Möchte man vektorwertig arbeiten, also $f\colon U \to \mathbb{R}^m$ mit möglichem $m > 1$, so kann man unter der Voraussetzung, dass f $(b+1)$-mal *stetig* differenzierbar ist, den Satz 6.28 benutzen:

$$\begin{aligned} f(x^0 + h) = {} & \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!}(h,\dots,h) \\ & + \frac{1}{b!}\left[\int_0^1 (1-t)^b D^{(b+1)} f(x^0 + th)\,dt\right]\underbrace{(h,\dots,h)}_{(b+1)\text{-mal}}. \end{aligned}$$

Dabei ist $\left[\int_0^1 \dots\, dt\right]$ eine $(b+1)$-lineare Abbildung.
Benötigt man das Restglied nicht explizit, so kann man letzteres auch als

$$f(x^0 + h) = \sum_{a=0}^{b} \frac{D^a f(x^0)}{a!}(h,\dots,h) + o\left(|h|^{b+1}\right)$$

für $h \to 0$ schreiben.

Die Formel im vorigen Satz wird typischerweise anders notiert. Dazu einige *Bezeichnungen*: Statt $\frac{\partial f}{\partial x_k}$ bzw. f_{x_k} schreiben wir auch $\partial_{x_k} f$ (bzw. $\partial_k f$). Ist dann $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}_0^n$ ein sogenannter *Multiindex*, so bezeichnet $\partial^\alpha f$ (bzw. $\partial_x^\alpha f$) die partielle Ableitung $\dfrac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$, wobei $|\alpha| = \alpha_1 + \dots + \alpha_n$ die Länge des Multiindex α ist. An weiteren Bezeichnungen benötigen wir $h^\alpha = h_1^{\alpha_1} \cdots h_n^{\alpha_n}$ für $h = (h_1,\dots,h_n)^T \in \mathbb{R}^n$ sowie $\alpha! = \alpha_1! \cdots \alpha_n!$.

Satz 9.19. *Sei $f\colon U \to \mathbb{R}$ $(b+1)$-mal stetig (partiell) differenzierbar*[4]*. Die Strecke von x^0 nach $x^0 + h$ möge ganz in U liegen. Dann gilt*

$$f(x^0 + h) = \sum_{|\alpha| \le b} \frac{\partial^\alpha f(x^0)}{\alpha!} h^\alpha + \sum_{|\alpha| = b+1} \frac{\partial^\alpha f(x^0 + \vartheta h)}{\alpha!} h^\alpha$$

für ein $\vartheta \in (0,1)$.

[4]Aus den Sätzen 9.16 und 9.17 folgt, dass für f die b-malige stetige Differenzierbarkeit in U und die b-malige stetige partielle Differenzierbarkeit in U äquivalent sind.

Beweis. Es bleibt lediglich festzustellen, dass für $x \in U$ und $a \leq b+1$

$$D^a f(x) \underbrace{(h, \ldots, h)}_{a\text{-mal}} = \sum_{|\alpha|=a} \frac{a!}{\alpha!} \partial^\alpha f(x) h^\alpha$$

gilt. Dies ergibt sich für $h = (h_1, \ldots, h_n)^T \in \mathbb{R}^n$ wie folgt:

$$\begin{aligned} D^a f(x)(h, \ldots, h) &= D^a f(x) \left(\sum_{k_1=1}^n h_{k_1} e_{k_1}, \ldots, \sum_{k_a=1}^n h_{k_a} e_{k_a} \right) \\ &= \sum_{k_1, \ldots, k_a=1}^n h_{k_1} \cdots h_{k_a} D^\alpha f(x)(e_{k_1}, \ldots, e_{k_a}) \\ &= \sum_{k_1, \ldots, k_a=1}^n h_{k_1} \cdots h_{k_a} \frac{\partial^a f}{\partial x_{k_1} \ldots \partial x_{k_a}}(x) \\ &= \sum_{|\alpha|=a} \frac{a!}{\alpha!} \partial^\alpha f(x) h^\alpha, \end{aligned}$$

wobei die numerische Konstante $\frac{a!}{\alpha!}$ sich kombinatorisch[5] ergibt. □

9.5 Extremwertbestimmung

Im Weiteren sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}$. Wir möchten lokale Extrema von f bestimmen.

Definition 9.20. f nimmt an der Stelle $x^0 \in U$ ein *lokales Maximum* (*Minimum*) an, falls es ein $\delta > 0$ gibt, so dass $f(x^0) \geq f(x)$ $(f(x^0) \leq f(x))$ für alle $x \in U$ mit $\left|x - x^0\right| \leq \delta$ gilt.

Im Folgenden beweisen wir die Resultate für lokale Maxima, die Resultate für lokale Minima erhält man entsprechend (indem man z. B. f durch $-f$ ersetzt).

9.5.1 Erste Ableitungen

Satz 9.21. *Ist f an der Stelle $x^0 \in U$ differenzierbar und hat f an x^0 ein lokales Maximum (Minimum), so gilt $\nabla f(x^0) = 0$.*

Beweis. Es gilt

$$f(x^0 + h) = f(x^0) + \nabla f(x^0) h + \varrho(h)$$

für $|h|$ klein mit $\lim\limits_{h \to 0} \frac{\varrho(h)}{|h|} = 0$. Ist $\nabla f(x^0) \neq 0$, so gibt es ein $h_0 \in \mathbb{R}^n$, $|h_0| = 1$ mit $\nabla f(x^0) h_0 > 0$. Wir finden dann ein $\lambda_0 > 0$ mit $\left|\frac{\varrho(\lambda h_0)}{\lambda}\right| < \nabla f(x^0) h_0$ für $0 < \lambda \leq \lambda_0$. Es folgt

$$f(x^0 + \lambda h_0) \geq f(x^0) + \lambda \nabla f(x^0) h_0 - |\varrho(\lambda h_0)| > f(x^0)$$

für $0 < \lambda \leq \lambda_0$. Somit hat f an x^0 kein lokales Maximum. □

[5] $\binom{\alpha}{\alpha_1, \ldots, \alpha_n} = \frac{a!}{\alpha_1! \ldots \alpha_n!}$ ist die Anzahl der Permutationen von a Elementen, die in n Gruppen zu jeweils $\alpha_1, \ldots, \alpha_n$ gleichen Elementen eingeteilt sind.

Beispiel 9.22. Sei $n = 2$ und

$$f(x, y) = x^2 + y^2 + xy^2.$$

Dann ist

$$f_x(x, y) = 2x + y^2, \quad f_y(x, y) = 2(1 + x)y.$$

Folglich liegen lokale Extrema höchstens an den Stellen $x = y = 0$ und $x = -1$, $y = \pm\sqrt{2}$ vor. Ob an einer dieser Stellen tatsächlich ein lokales Extremum vorliegt, werden wir im nächsten Abschnitt entscheiden.

9.5.2 Zweite Ableitungen

Sei jetzt f in U zweimal differenzierbar und $\nabla f(x^0) = 0$. Zuerst behandeln wir eine weitere notwendige Bedingung für das Vorliegen einer Extremstelle.

Für reelle symmetrische Matrizen A, $B \in M_n(\mathbb{R})$ $(= M_{n\times n}(\mathbb{R}))$ schreiben wir $A \geq 0$, falls A nur nichtnegative Eigenwerte hat, und $A \geq B$, falls $A - B \geq 0$.[6]

Definition 9.23. Sei A eine reelle symmetrische Matrix. Dann:

- A heißt *positiv* (*negativ*) *semidefinit*, falls $A \geq 0$ ($A \leq 0$).
- A heißt *positiv* (*negativ*) *definit*, falls $A \geq cI$ ($A \leq -cI$) für ein $c > 0$. Man schreibt $A > 0$ ($A < 0$).
- A heißt *indefinit*, falls A weder positiv noch negativ semidefinit ist.[7]

Satz 9.24. *Ist f an der Stelle x^0 zweimal differenzierbar und hat f an x^0 ein lokales Maximum* (*Minimum*)*, so ist die Hesse-Matrix $D^2 f(x^0)$ negativ* (*positiv*) *semidefinit.*

Beweis. Dieser Beweis ist ähnlich dem Beweis des vorigen Satzes.
Es gilt

$$f(x^0 + h) = f(x^0) + \frac{1}{2}\, D^2 f(x^0)(h, h) + \varrho(h)$$

für $|h|$ klein, wobei $\lim\limits_{h\to 0} \frac{\varrho(h)}{|h|^2} = 0$. Ist $D^2 f(x^0) \not\leq 0$, so gibt es ein $h_0 \in \mathbb{R}^n$, $|h_0| = 1$ mit $D^2 f(x^0)(h_0, h_0) > 0$. Es gibt dann ein $\lambda_0 > 0$ mit $\left|\frac{\varrho(\lambda h_0)}{\lambda^2}\right| < \frac{1}{2}\, D^2 f(x^0)(h_0, h_0)$ für $0 < \lambda \leq \lambda_0$. Wir erhalten

$$f(x^0 + \lambda h_0) \geq f(x^0) + \frac{\lambda^2}{2}\, D^2 f(x^0)(h_0, h_0) - |\varrho(\lambda h_0)| > f(x^0)$$

[6] Eine reelle symmetrische Matrix $A \in M_n(\mathbb{R})$ hat halbeinfache reelle Eigenwerte und ist mittels einer orthogonalen Matrix $U \in O(n;\mathbb{R})$ diagonalisierbar. Das heißt, es gibt eine Matrix U mit $UU^T = I$ und $UAU^T = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$, wobei die $\lambda_j \in \mathbb{R}$ die Eigenwerte von A sind.

[7] Äquivalente Bedingungen sind:

- A ist positiv (negativ) semidefinit, falls $\langle Ah, h\rangle \geq 0$ ($\langle Ah, h\rangle \leq 0$) für alle $h \in \mathbb{R}^n$.
- A ist positiv (negativ) definit, falls $\langle Ah, h\rangle \geq c\,|h|^2$ ($\langle Ah, h\rangle \leq -c\,|h|^2$) für alle $h \in \mathbb{R}^n$ und ein $c > 0$.
- A ist indefinit, falls es h_1, $h_2 \in \mathbb{R}^n$ mit $\langle Ah_1, h_1\rangle > 0$ und $\langle Ah_2, h_2\rangle < 0$ gibt.

Hierbei ist $\langle\ ,\ \rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ das durch $\langle x, y\rangle = x^T y$ definierte *Euklidische Skalarprodukt* in $\mathbb{R}^n$. Insbesondere gilt $\langle x, x\rangle = |x|^2$ für alle $x \in \mathbb{R}^n$.

für $0 < \lambda \leq \lambda_0$. Folglich hat f kein lokales Maximum an x^0. □

Folgerung 9.25. *Ist $Df(x^0) = 0$, jedoch $D^2 f(x^0)$ indefinit, so nimmt f an x^0 keinen lokalen Extremwert an. Man sagt, dass an einer solchen Stelle ein Sattelpunkt vorliegt.*

Beispiel 9.26. In unserem obigen Beispiel gilt

$$D^2 f(x,y) = 2 \begin{pmatrix} 1 & y \\ y & 1+x \end{pmatrix}.$$

Die Eigenwerte sind

$$\lambda_{\pm} = 2 + x \pm \sqrt{x^2 + 4y^2}.$$

Folglich ist die Hesse-Matrix an $x = -1$, $y = \pm\sqrt{2}$ indefinit (die Eigenwerte sind -2, 4). Somit liegt an diesen Stellen jeweils ein Sattelpunkt vor.

Ein hinreichendes Kriterium für das Vorliegen einer Extremstelle ist die Definitheit der Hesse-Matrix.

Satz 9.27. *Ist f an der Stelle x^0 zweimal differenzierbar, ist $\nabla f(x^0) = 0$ und ist $D^2 f(x^0)$ negativ definit (positiv definit), so hat f an der Stelle x^0 ein lokales Maximum (Minimum).*

Beweis. Sei $D^2 f(x^0)(h,h) \leq -c|h|^2$ für ein $c > 0$. Wir wählen ein $\delta > 0$, so dass $|\varrho(h)| < \frac{c}{2}|h|^2$ für $|h| \leq \delta$, wobei $\varrho(h)$ den Rest in der Taylorformel bezeichnet, d. h.

$$f(x^0 + h) = f(x^0) + \frac{1}{2} D^2 f(x^0)(h,h) + \varrho(h).$$

für $|h|$ hinreichend klein. Für $|h| \leq \delta$ erhalten wir dann

$$f(x^0 + h) \leq f(x^0) - \frac{c}{2}|h|^2 + |\varrho(h)| < f(x^0).$$

Folglich ist $f(x^0)$ lokales Maximum an der Stelle x^0. □

Beispiel 9.28. In unserem Beispiel ist die Hesse-Matrix an $x = y = 0$ positiv definit (beide Eigenwerte sind 2), also haben wir an dieser Stelle ein lokales Minimum.

9.5.3 Extremwertbestimmung unter Nebenbedingungen

Sei jetzt zusätzlich eine stetig differenzierbare Abbildung $g\colon U \to \mathbb{R}^r$ für ein $r < n$ gegeben. Wir wollen Extrema der Funktion f unter der *Nebenbedingung* $g(x) = 0$ bestimmen.

Definition 9.29. Sei $x^0 \in U$ mit $g(x^0) = 0$[8]. Dann hat f an der Stelle x^0 ein *lokales Maximum (Minimum) unter der Nebenbedingung* $g = 0$, falls es ein $\delta > 0$ gibt, so dass $f(x^0) \geq f(x)$ ($f(x^0) \leq f(x)$) für alle $x \in U$ mit $g(x) = 0$ und $\left|x - x^0\right| \leq \delta$ gilt.

Wir schreiben $g = (g_1, \ldots, g_r)^T$.

[8]Das heißt, x^0 erfüllt die Nebenbedingung.

Satz 9.30. *Sei f an $x^0 \in U$ differenzierbar, $g(x^0) = 0$. Weiter habe f an $x^0 \in U$ ein lokales Maximum (Minimum) unter der Nebenbedingung $g = 0$. Die lineare Abbildung $Dg(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^r)$ habe vollen Rang*[9]*. Dann gibt es $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ mit*

$$\nabla f(x^0) = \sum_{k=1}^{r} \lambda_k \nabla g_k(x^0),$$

d. h. der Gradient von f an der Stelle x^0 ist Linearkombination der Gradienten der g_k an der Stelle x^0.

Die $\lambda_1, \ldots, \lambda_r$ heißen *Lagrange-Multiplikatoren.*

Beispiel 9.31. Wir betrachten die durch die Gleichung

$$g(x) = |x|^2 - 1 = x_1^2 + \cdots + x_n^2 + x_{n+1}^2 - 1 = 0$$

gegebene *n-Sphäre* $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ und die Höhenfunktion

$$f(x) = x_{n+1}.$$

Es gilt

$$\nabla f(x) = (\underbrace{0, \ldots, 0}_{n-\text{mal}}, 1), \quad \nabla g(x) = (2x_1, 2x_2, \ldots, 2x_{n+1}) = 2x.$$

Die Bedingung $\nabla f(x) = \lambda \nabla g(x)$ für ein $\lambda \in \mathbb{R}$, d. h.

$$(0, \ldots, 0, 1) = 2\lambda x,$$

impliziert $x_1 = \cdots = x_n = 0$ und damit $x_{n+1} = \pm 1$ wegen $|x| = 1$ (dann $\lambda = \pm 1/2$). Somit liegen mögliche Extrema der Höhenfunktion an den Stellen $(0, \ldots, 0, 1)$ (Nordpol) und $(0, \ldots, 0, -1)$ (Südpol) vor.
Tatsächlich nimmt f auf $\mathbb{S}^n$ an $(0, \ldots, 0, 1)$ ein Maximum und an $(0, \ldots, 0, -1)$ ein Minimum an.

Wir beginnen den Beweis des Satzes mit folgendem Lemma:

Lemma 9.32. *Es gibt einen Koordinatenwechsel $x \mapsto y = y(x)$ nahe x^0, so dass $y(x^0) = 0$ und $g_k(x) = y_{n-r+k}(x)$ für $k = 1, \ldots, r$.*

Wir werden dieses Ergebnis im übernächsten Abschnitt 9.7 beweisen. Dort werden wir auch sehen, dass die im Satz genannten Bedingungen *invariant unter Koordinatenwechseln* sind.

Beweis des Satzes. Wir schreiben $x = (x', x'')^T$ mit $x' = (x_1, \ldots, x_{n-r})^T$, $x'' = (x_{n-r+1}, \ldots, x_n)^T$. Obiges Lemma bedeutet, dass wir o. B. d. A. $x^0 = 0$ und $g(x) = x''$ annehmen dürfen.

[9]Wegen $r < n$ heißt das, dass die lineare Abbildung $Dg(x^0)\colon \mathbb{R}^n \to \mathbb{R}^r$ surjektiv ist bzw. die zugehörige $(r \times n)$–Matrix $\left(\frac{\partial g_k}{\partial x_j}(x^0)\right)_{\substack{k=1,\ldots,r,\\ j=1,\ldots,n}}$ den maximal möglichen Rang r hat. Das wiederum heißt, dass die r Zeilen $\nabla g_k(x^0)$ für $k = 1, \ldots, r$ dieser Matrix linear unabhängig sind.

Wir fragen uns, ob die Funktion $h(x') = f(x', 0)$ an $x' = 0$ ein lokales Extremum hat. Notwendig dafür ist $\nabla_{x'} f(0,0) = \nabla_{x'} h(0) = 0$, also

$$\begin{aligned}\nabla_x f(0,0) = (0, \nabla_{x''} f(0,0)) &= \sum_{k=1}^{r} \frac{\partial f}{\partial x_{n-r+k}}(0,0)\, e_{n-r+k} \\ &= \sum_{k=1}^{r} \lambda_k \nabla_x g_k(0,0),\end{aligned}$$

wobei $\lambda_k = \frac{\partial f}{\partial x_{n-r+k}}(0,0)$.[10] □

Beispiel 9.33. Extremstellen der Funktion $f(x,y) = x^2 + \frac{y^4}{2}$ auf der Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $a > 0$, $b > 0$. Die Nebenbedingung ist $g = 0$ mit $g(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$, also wegen

$$\nabla f(x,y) = (2x, 2y^3), \quad \nabla g(x,y) = \left(\frac{2x}{a^2}, \frac{2y}{b^2}\right)$$

ist

$$(x, y^3) = \lambda\left(\frac{x}{a^2}, \frac{y}{b^2}\right)$$

für ein $\lambda \in \mathbb{R}$ eine notwendige Bedingung für das Vorliegen einer Extremstelle. Wir unterscheiden die Fälle $x = 0$, dann $y = 0$ und schließlich $x \neq 0$, $y \neq 0$:

Fall 1: $x = 0$. Dann ist $\lambda = b^4$, $y = \pm b$ und $f(0, \pm b) = \frac{b^4}{2}$.

Fall 2: $y = 0$. Dann ist $\lambda = a^2$, $x = \pm a$ und $f(\pm a, 0) = a^2$.

Fall 3: $x \neq 0$, $y \neq 0$. Dann ist $\lambda = a^2$, $x = \pm a\sqrt{1 - \frac{a^2}{b^4}}$, $y = \pm\frac{a}{b}$ (mit voneinander unabhängig gewählten Vorzeichen) und

$$f\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right) = a^2\left(1 - \frac{a^2}{2b^4}\right).$$

Dieser Fall ist nur für $b^4 \geq a^2$ möglich.

Es gilt

$$\begin{aligned} a^2 &> a^2\left(1 - \frac{a^2}{2b^4}\right), \\ \frac{b^4}{2} &\geq a^2\left(1 - \frac{a^2}{2b^4}\right), \end{aligned}$$

wobei im zweiten Fall Gleichheit genau für $a^2 = b^4$ eintritt. Damit ergibt sich folgende Tabelle:

	Minimalstellen	Maximalstellen
$b^4 < a^2$	$(0, \pm b)$	$(\pm a, 0)$
$\frac{b^4}{2} < a^2 \leq b^4$	$\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right)$	$(\pm a, 0)$
$a^2 = \frac{b^4}{2}$	$\left(\pm\frac{a}{\sqrt{2}}, \pm\frac{b}{\sqrt{2}}\right)$	$(\pm a, 0), (0, \pm b)$
$a^2 < \frac{b^4}{2}$	$\left(\pm a\sqrt{1 - \frac{a^2}{b^4}}, \pm\frac{a}{b}\right)$	$(0, \pm b)$

[10] Die Werte der λ_k sind *nicht* invariant unter Koordinatentransformationen.

9.6 Der Satz über die implizite Funktion

Wir beginnen mit dem Satz über die lokale Umkehrfunktion.

Satz 9.34 (über die inverse Funktion)**.** *Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}^n$ stetig differenzierbar. Sei weiterhin $x^0 \in U$, $Df(x^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar. Dann gibt es offene Umgebungen $U' \subseteq U$ von x^0 und $V' = f(U')$ von $f(x^0)$, so dass f eingeschränkt auf U' invertierbar ist, d. h. $f\big|_{U'}\colon U' \to V' = f(U')$ ist bijektiv. Die Umkehrabbildung $g = \left(f\big|_{U'}\right)^{-1}\colon V' \to U'$ ist stetig differenzierbar und*

$$Dg(y) = Df(x)^{-1}, \quad y \in V'.$$

wobei $x = g(y)$.

Beweis. Indem wir f durch $Df(x^0)^{-1}\left(f(x^0 + x) - f(x^0)\right)$ ersetzen, dürfen wir $x^0 = 0$, $f(x^0) = 0$ und $Df(x^0) = I$ annehmen.

1. Wir zeigen zuerst, dass f lokal invertierbar ist. Dazu wählen wir ein $r > 0$ mit $\bar{B}(0, 2r) = \left\{x \in \mathbb{R}^n \;\middle|\; |x| \leq 2r\right\} \subseteq U$ und $|(I - Df(x))\,h| \leq |h|/2$ für $|x| \leq 2r$ und $h \in \mathbb{R}^n$.

   Sei nun $y \in \mathbb{R}^n$, $|y| \leq r$. Wir zeigen, dass die Gleichung $f(x) = y$ eine eindeutige Lösung $x \in \mathbb{R}^n$ mit $|x| \leq 2r$ besitzt. Dazu benutzen wir den Banachschen Fixpunktsatz und betrachten

   $$T_y x = x - f(x) + y.$$

   Es gilt $T_y\colon \bar{B}(0, 2r) \to \bar{B}(0, 2r)$ wegen

   $$|x - f(x)| \leq \int_0^1 |(I - Df(sx))\,x|\,ds \leq \frac{|x|}{2},$$

   also

   $$|T_y x| \leq |x - f(x)| + |y| \leq 2r.$$

   Weiterhin ist T_y eine Kontraktion auf $\bar{B}(0, 2r)$ wegen

   $$|T_y x - T_y x'| \leq \int_0^1 |(I - Df(x' + s(x - x')))\,(x - x')|\,ds \leq \frac{|x - x'|}{2}.$$

   Wegen der Vollständigkeit von $\bar{B}(0, 2r)$ folgt die Behauptung.

2. Wir setzen nun $V' = B(0, r)$ und $U' = f^{-1}(B(0, r)) \subseteq B(0, 2r)$. U' ist eine offene Umgebung der 0 wegen der Stetigkeit von f und $f\big|_{U'}\colon U' \to V'$ ist bijektiv nach Konstruktion.

3. Schließlich zeigen wir, dass $g = \left(f\big|_{U'}\right)^{-1}\colon V' \to U'$ stetig differenzierbar ist und $Dg(y) = Df(x)^{-1}$ für $x = g(y)$ gilt.

   Angenommen, wir wissen bereits, dass g differenzierbar ist. Dann folgt aus $g\,(f(x)) = x$ für $x \in U'$ und der Kettenregel, dass $Dg\,(f(x))\,Df(x) = I$ gilt, also $Dg\,(f(x)) = Df(x)^{-1}$. Da $Df\colon U' \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ stetig und, wie wir gerade gezeigt haben, auch überall invertierbar ist, folgt, dass $Dg\colon V' \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ ebenfalls stetig ist.

Somit bleibt zu zeigen, dass g differenzierbar ist. Da die Stelle $x^0 = 0$ in keiner Weise ausgezeichnet ist, genügt es zu zeigen, dass g an 0 differenzierbar ist. Dazu schreiben wir mit $f(0) = 0$ und $(Df)(0) = I$

$$f(x) = x + o(|x|)$$

für $x \to 0$, folglich mit $x = g(y)$

$$y = g(y) + o\left(|g(y)|\right)$$

beziehungsweise

$$g(y) = y + o(|y|)$$

für $y \to 0$. Es folgt, dass g an 0 differenzierbar ist und $Dg(0) = I$ gilt.

□

Wir schreiben jetzt $(x, y)^T$ mit $x \in \mathbb{R}^n$, $y \in \mathbb{R}^k$ für Stellen in $\mathbb{R}^{n+k}$.

Satz 9.35 (über die implizite Funktion). *Sei $U \subseteq \mathbb{R}^{n+k}$ offen, $f\colon U \to \mathbb{R}^k$ stetig differenzierbar, $(x^0, y^0)^T \in U$ mit $f(x^0, y^0) = 0$ und $D_y f(x^0, y^0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar. Dann gibt es offene Umgebungen $V \subseteq \mathbb{R}^n$ von x^0 und $W \subseteq \mathbb{R}^k$ von y^0 mit $V \times W \subseteq U$ und eine stetig differenzierbare Funktion $g\colon V \to W$ mit $g(x^0) = y^0$, so dass*

$$f(x, y) = 0 \text{ für } (x, y)^T \in V \times W \iff y = g(x).$$

Insbesondere gilt $f\left(x, g(x)\right) = 0$ für alle $x \in V$.

Beweis. Wir betrachten die Funktion $F\colon U \to \mathbb{R}^{n+k}$, $F(x, y) = \left(x, f(x, y)\right)^T$. Dann ist

$$DF(x, y) = \begin{pmatrix} I & 0 \\ D_x f(x, y) & D_y f(x, y) \end{pmatrix}$$

an der Stelle $(x^0, y^0)^T$ als lineare Abbildung von $\mathbb{R}^n \times \mathbb{R}^k$ nach $\mathbb{R}^n \times \mathbb{R}^k$ invertierbar. Folglich existieren nach dem Satz über die inverse Funktion offene Umgebungen $\widetilde{U} \subseteq U$ von $(x^0, y^0)^T$ und $\widetilde{V} \subseteq \mathbb{R}^{n+k}$ von $(x^0, 0)^T$ sowie die Umkehrabbildung $G\colon \widetilde{V} \to \widetilde{U}$ zu $F\big|_{\widetilde{U}}\colon \widetilde{U} \to \widetilde{V}$.
G hat notwendigerweise die Gestalt $G(x, y) = \left(x, h(x, y)\right)^T$ für eine Funktion $h\colon \widetilde{V} \to \mathbb{R}^k$ und es gilt

$$f\left(x, h(x, y)\right) = y$$

für $(x, y)^T \in \widetilde{V}$.
Nach Konstruktion gilt $f(x, y) = 0$ für $(x, y)^T \in \widetilde{U}$ genau dann, wenn $(x, 0)^T \in \widetilde{V}$ und $y = h(x, 0)$. Damit können wir jetzt $V \subseteq \left\{x \in \mathbb{R}^n \;\middle|\; (x, 0)^T \in \widetilde{V}\right\}$, $W = \left\{h(x, 0) \;\middle|\; x \in V\right\}$ und $g(x) = h(x, 0)$ für $x \in V$ setzen, wobei wir die Umgebung V von x^0 so klein wählen, dass $V \times W \subseteq U$ gilt. □

Bemerkung. Mit

$$D_x F(x, y) + D_y F(x, y) Dg(x) = 0,$$

wobei $y = g(x)$, berechnet sich die Ableitung von g zu

$$Dg(x) = -D_y F(x, y)^{-1} D_x F(x, y).$$

9.7 Differenziation in krummlinigen Koordinaten

Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f\colon U \to \mathbb{R}$ heißt von der *Klasse* C^l, wobei $l \in \mathbb{N}_0 \cup \{\infty\}$, falls f in U l-mal stetig differenzierbar ist[11]. Eine Vektorfunktion heißt von der Klasse C^l, falls jede ihrer Komponenten von der Klasse C^l ist.

Definition 9.36. Ein C^l*-Koordinatensystem in* U, wobei $l \geq 1$, sind n skalare Funktionen $y_k\colon U \to \mathbb{R}$ für $k = 1, \ldots, n$ der Klasse C^l, so dass die Abbildung

$$y = (y_1, \ldots, y_n)^T : U \to \mathbb{R}^n$$

injektiv und $Dy(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ für alle $x \in U$ invertierbar ist.

Dann ist die Abbildung $y\colon U \to \mathbb{R}^n$ bijektiv auf ihr Bild $V = y(U)$, $V \subseteq \mathbb{R}^n$ ist offen und die Umkehrabbildung $V \to U$, $y \mapsto x$, ist ebenfalls von der Klasse C^l. Man nimmt die *Werte* der Funktionen $y_1, \ldots, y_n$ an einer Stelle $x \in U$ als *Koordinaten* von x, d. h. um die Stelle x zu adressieren.

Beispiel 9.37.

1. *Euklidische Koordinaten*: Die Koordinatenfunktionen sind $x_1, \ldots, x_n$.

2. *Polarkoordinaten im* $\mathbb{R}^2$: Euklidische Koordinaten sind x, y, als U nehmen wir den $\mathbb{R}^2$ ohne einen vom Ursprung ausgehenden Strahl, z. B.

$$U = \mathbb{R}^2 \setminus \big\{(x, 0) \in \mathbb{R}^2 \;\big|\; x \leq 0\big\}.$$

Wir führen neue Koordinaten $(r, \varphi) \in \mathbb{R}_+ \times (-\pi, \pi)$ vermöge

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

ein.

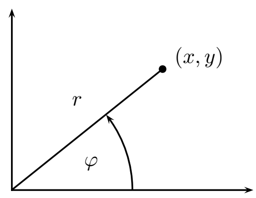

Abbildung 9.1: Polarkoordinaten

Die Abbildung

$$\begin{aligned} \mathbb{R}_+ \times (-\pi, \pi) &\to \mathbb{R}^2 \setminus \big\{(x, 0) \;\big|\; x \leq 0\big\}, \\ (r, \varphi) &\mapsto (x, y), \end{aligned}$$

ist bijektiv. Zudem gilt

$$\det \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \det \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} = r > 0.$$

[11] Für $l = 0$ ist $f\colon U \to \mathbb{R}$ stetig.

Tatsächlich ist $r^2 = x^2 + y^2$ und $\cot\varphi = \frac{\cos\varphi}{\sin\varphi} = \frac{x}{y}$ für $y \neq 0$, also

$$r = \sqrt{x^2+y^2}, \qquad \varphi = \operatorname{arccot}\frac{x}{y} \qquad \text{für } y > 0$$
$$\text{bzw. } \varphi = -\pi + \operatorname{arccot}\frac{x}{y} \qquad \text{für } y < 0$$

sowie $\varphi = 0$ für $x > 0$, $y = 0$.

3. *Zylinderkoordinaten im* $\mathbb{R}^3$: Euklidische Koordinaten sind x, y, z, wir nehmen $U = \mathbb{R}^3 \setminus \{(x,0,z) \mid x \leq 0\}$. Wir führen neue Koordinaten $(r,\varphi,z) \in \mathbb{R}_+ \times (-\pi,\pi) \times \mathbb{R}$ vermöge

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z$$

ein.

4. *Kugelkoordinaten im* $\mathbb{R}^3$: Wir nehmen U wie unter 3), die Koordinaten sind $(\varrho,\varphi,\theta) \in \mathbb{R}_+ \times (-\pi,\pi) \times (0,\pi)$ mit

$$x = \varrho\cos\varphi\sin\theta, \quad y = \varrho\sin\varphi\sin\theta, \quad z = \varrho\cos\theta.$$

Den Übergang von einem Koordinatensystem zu einem anderen bezeichnet man als *Koordinatenwechsel* (oder auch als Koordinatentransformation).

Beispiel 9.38. Koordinatenwechsel im $\mathbb{R}^2$, $(x,y) \mapsto (u,v)$

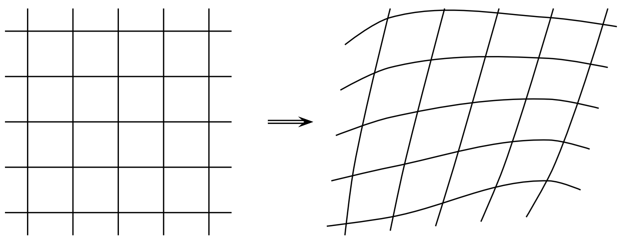

Abbildung 9.2: Übergang zu krummlinige Koordinaten

Lemma 9.39. *Sei $y\colon U \to \mathbb{R}^n$ von der Klasse C^l, $x^0 \in U$ und $Dy(x^0) \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ invertierbar. Dann kann man die $y_1,\dots,y_n$ nahe x^0 als C^l-Koordinaten nehmen.*

Beweis. Man benutze den Satz über die inverse Funktion. □

Lemma 9.40. *Seien $y_1,\dots,y_r\colon U \to \mathbb{R}$ für $r < n$ C^l-Funktionen, $l \geq 1$, so dass $Dy(x^0) \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^r)$ maximalen Rang hat (d. h. die Vektoren $\nabla y_1(x^0),\dots,\nabla y_r(x^0)$ sind linear unabhängig). Dann gibt es Funktionen*

$$y_{r+1},\dots,y_n\colon U \to \mathbb{R},$$

so dass $(y_1,\dots,y_n)$ ein C^l-Koordinatensystem nahe x^0 ist.

Beweis. Es gibt $n-r$ Indizes i_k mit $1 \le i_1 < i_2 < \cdots < i_{n-r} \le n$, so dass

$$\det\left(\nabla y(x^0)^T, e_{i_1}, \ldots, e_{i_{n-r}}\right) \neq 0.$$

Wir können dann $(y_1, \ldots, y_r, x_{i_1}, \ldots, x_{i_{n-r}})$ als C^l-Koordinatensystem nahe x^0 wählen. □

Lemma 9.41. *Seien x und y zwei Koordinatensysteme nahe x^0 mit Koordinatenwechseln $x \mapsto y = y(x)$ bzw. $y \mapsto x = x(y)$. Weiter sei f eine Funktion, die in x^0 differenzierbar ist. Dann gilt*

$$\nabla_y f(x^0) = \nabla_x f(x^0) \frac{\partial x}{\partial y}(x^0),$$

wobei $\frac{\partial x}{\partial y}$ die Jacobi-Matrix des Koordinatenwechsels $y \mapsto x$ bezeichnet.

Beweis. Wir schreiben f als Funktion von x, d. h. $f = f(x)$, und x als Funktion von y, d. h. $x = \varphi(y)$. Dann schreibt sich f als Funktion von y als $g(y) = f(\varphi(y))$. Wir erhalten mit $x^0 = \varphi(y^0)$

$$\nabla g(y^0) = \nabla f(x^0) D\varphi(y^0).$$

Das ist die Behauptung. □

Lemma 9.42. *Unter den Voraussetzungen des vorigen Lemmas sei f nahe x^0 stetig differenzierbar, $\nabla f(x^0) = 0$ (diese Bedingung ist unabhängig vom gewählten Koordinatensystem) und f sei an x^0 zweimal differenzierbar. Dann gilt*

$$D_y^2 f(x^0) = \frac{\partial x}{\partial y}(x^0)\, D_x^2 f(x^0)\, \frac{\partial x}{\partial y}(x^0)^T.$$

Insbesondere sind Definitheitsbedingungen an die Hesse-Matrix koordinateninvariant (unter der Bedingung, dass der Gradient verschwindet).

Beweis. Unter den Bezeichnungen des Beweises des vorigen Lemmas gilt

$$\frac{\partial^2 g}{\partial y_k \partial y_l}(y^0) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0) \frac{\partial \varphi_i}{\partial y_k}(y^0) \frac{\partial \varphi_j}{\partial y_l}(y^0),$$

und daraus folgt die Behauptung. □

Wir studieren schließlich das Verhalten von Differenzialausdrücken unter Koordinatentransformation.

Definition 9.43. Sei x ein C^l-Koordinatensystem in U. Ein *Differenzialausdruck* der Ordnung m in den Koordinaten $x_1, \ldots, x_n$ ist ein Ausdruck der Form

$$\sum_{|\alpha| \le m} a_\alpha(x) \partial_x^\alpha,$$

wobei $\partial_x^\alpha = \dfrac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$ wie zuvor und die $a_\alpha(x)$ für $|\alpha| \le m$ Funktionen auf U sind.

Es sei jetzt ein weiteres C^l-Koordinatensystem $y_1, \dots, y_n$ in U gegeben. Unter der Voraussetzung $m \leq l$ können wir dann einen gegebenen Differenzialausdruck $\sum\limits_{|\alpha| \leq m} a_\alpha(x) \partial_x^\alpha$ auch als

$$\sum_{|\beta| \leq m} b_\beta(y) \partial_y^\beta$$

mit gewissen resultierenden Funktionen $b_\beta(y)$ auf U schreiben. Die Frage ist, wie sich die b_β aus den a_α berechnen.
Wir betrachten dazu einige Beispiele.

Beispiel 9.44. 1. Wie sich Ableitungen erster Ordnung transformieren, haben wir bereits gesehen:

$$\frac{\partial}{\partial x_j} = \frac{\partial y_1}{\partial x_j} \frac{\partial}{\partial y_1} + \dots + \frac{\partial y_n}{\partial x_j} \frac{\partial}{\partial y_n}.$$

2. Einer der wichtigsten Differenzialoperatoren der mathematischen Physik ist der *Laplace-Operator* $\Delta = \dfrac{\partial^2}{\partial x_1^2} + \dfrac{\partial^2}{\partial x_2^2} + \dots + \dfrac{\partial^2}{\partial x_n^2}$. Wir wollen im Fall $n = 2$ diesen Operator in Polarkoordinaten schreiben.

   Aus $x = r \cos \varphi$, $y = r \sin \varphi$ folgt

$$\begin{aligned}
\frac{\partial}{\partial r} &= \frac{\partial x}{\partial r} \frac{\partial}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial}{\partial y} = \cos \varphi \frac{\partial}{\partial x} + \sin \varphi \frac{\partial}{\partial y} \\
\frac{\partial}{\partial \varphi} &= \frac{\partial x}{\partial \varphi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \varphi} \frac{\partial}{\partial y} = -r \sin \varphi \frac{\partial}{\partial x} + r \cos \varphi \frac{\partial}{\partial y}
\end{aligned}$$

   bzw.

$$\begin{pmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix} \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

$$\begin{pmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial \varphi} \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{pmatrix}$$

   oder

$$\begin{aligned}
\frac{\partial}{\partial x} &= \cos \varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial}{\partial \varphi} \\
\partial \partial y &= \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial}{\partial \varphi}.
\end{aligned}$$

   Es folgt

$$\begin{aligned}
\frac{\partial^2}{\partial x^2} &= \left(\cos \varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial}{\partial \varphi} \right) \left(\cos \varphi \frac{\partial}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial}{\partial \varphi} \right) \\
&= \cos^2 \varphi \frac{\partial^2}{\partial r^2} - \frac{1}{r} \sin \varphi \cos \varphi \frac{\partial^2}{\partial r \partial \varphi} + \frac{1}{r^2} \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi} \\
&- \frac{1}{r} \sin \varphi \cos \varphi \frac{\partial^2}{\partial r \partial \varphi} + \frac{1}{r} \sin^2 \varphi \frac{\partial}{\partial r} \\
&+ \frac{1}{r^2} \sin^2 \varphi \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r^2} \sin \varphi \cos \varphi \frac{\partial}{\partial \varphi}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2}{\partial y^2} &= \left(\sin\varphi \frac{\partial}{\partial r} + \frac{1}{r}\,\cos\varphi \frac{\partial}{\partial\varphi}\right)\left(\sin\varphi \frac{\partial}{\partial r} + \frac{1}{r}\,\cos\varphi \frac{\partial}{\partial\varphi}\right) \\
&= \sin^2\!\varphi \frac{\partial^2}{\partial r^2} + \frac{1}{r}\,\sin\varphi\cos\varphi \frac{\partial^2}{\partial r\partial\varphi} - \frac{1}{r^2}\,\sin\varphi\cos\varphi \frac{\partial}{\partial\varphi} \\
&+ \frac{1}{r}\,\sin\varphi\cos\varphi \frac{\partial^2}{\partial r\partial\varphi} + \frac{1}{r}\cos^2\!\varphi \frac{\partial}{\partial r} \\
&+ \frac{1}{r^2}\cos^2\!\varphi \frac{\partial^2}{\partial\varphi^2} - \frac{1}{r^2}\,\sin\varphi\cos\varphi \frac{\partial}{\partial\varphi},
\end{aligned}
$$

also ist

$$
\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2}
$$

der Ausdruck des Laplace-Operators in Polarkoordinaten.

Kapitel 10

Maß- und Integrationstheorie

Wir beschäftigen uns zuerst mit abstrakter Maß- und Integrationstheorie und dann im nächsten Kapitel mit dem Lebesgue-Integral.

10.1 Maßräume

Sei Ω eine Menge. Wir möchten gewissen Teilmengen von Ω ein Maß zuordnen.

Definition 10.1. Eine *σ-Algebra* $\mathscr{A}$ auf Ω ist Menge von Teilmengen von Ω mit folgenden Eigenschaften:

(i) $\Omega \in \mathscr{A}$.

(ii) Ist $A \in \mathscr{A}$, so ist auch $A^c = \Omega \setminus A \in \mathscr{A}$.

(iii) Ist $\{A_i\} \subset \mathscr{A}$ eine *abzählbare* Familie, so ist $\bigcup_i A_i \in \mathscr{A}$.

$(\Omega, \mathscr{A})$ heißt ein *messbarer Raum* und die Mengen in $\mathscr{A}$ die messbaren Teilmengen von Ω.

Offenbar gilt $\emptyset \in \mathscr{A}$, $\mathscr{A}$ ist abgeschlossen unter abzählbaren Durchschnitten (wegen $\bigcap_i A_i = (\bigcup_i A_i^c)^c$) und mit $A, B \in \mathscr{A}$ gehört auch $A \setminus B = A \cap B^c$ zu $\mathscr{A}$.

Lemma 10.2. *Ist $\mathscr{B} \subseteq \mathscr{P}(\Omega)$, so gibt es eine kleinste σ-Algebra $\sigma(\mathscr{B})$, die $\mathscr{B}$ enthält.*

Beweis. Der Durchschnitt einer beliebigen Familie von σ-Algebren ist wiederum eine σ-Algebra. □

$\sigma(\mathscr{B})$ heißt auch die von $\mathscr{B}$ *erzeugte σ-Algebra.*

Beispiel 10.3. a) $\{\emptyset, \Omega\}$ und $\mathscr{P}(\Omega) = 2^\Omega$ sind σ-Algebren.

b) Ist $A \subseteq \Omega$, so ist $\sigma(A) = \{\emptyset, A, A^c, \Omega\}$ die von $\{A\}$ erzeugte σ-Algebra.

c) Sei $X = (X, \mathscr{T})$ ein topologischer Raum. Dann ist die *Borelsche σ-Algebra* $\mathscr{B}(X) = \sigma(\mathscr{T})$ auf X die von allen offenen Mengen von X erzeugte σ-Algebra. Die Elemente von $\mathscr{B}(X)$ heißen die *Borelmengen* von X. Insbesondere gehören sämtliche abgeschlossenen Mengen dazu, aber auch F_σ-Mengen (= abzählbare Vereinigungen abgeschlossener Mengen), G_δ-Mengen (= abzählbare Durchschnitte offener Mengen), usw.

Definition 10.4. Ein *Maß* auf einem messbaren Raum $(\Omega, \mathscr{A})$ ist eine Abbildung $\mu\colon \mathscr{A} \to [0, \infty]$ mit folgenden Eigenschaften:

(i) Es gilt $\mu(\emptyset) = 0$.

(ii) (σ-Additivität) Ist $\{A_i\} \subset \mathscr{A}$ eine abzählbare Familie *paarweise disjunkter* Mengen, so gilt

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i).$$

Das Tripel $(\Omega, \mathscr{A}, \mu)$ heißt ein *Maßraum*.

Bemerkung. Die Forderung (i) kann durch die schwächere Forderung, dass es ein $A \in \mathscr{A}$ mit $\mu(A) < \infty$ gibt, ersetzt werden.[1]

Folgerung 10.5. *Ein Maß μ besitzt die folgenden Eigenschaften:*

(i) (Endliche Additivität) *Für paarweise disjunkte Mengen $A_1, A_2, \ldots, A_m \in \mathscr{A}$ gilt*

$$\mu(A_1 \cup \cdots \cup A_m) = \mu(A_1) + \cdots + \mu(A_m).$$

(ii) (Monotonie) *Für A, $B \in \mathscr{A}$ mit $A \subseteq B$ gilt*

$$\mu(A) \leq \mu(B).$$

(iii) *Ist $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}$ mit $A_1 \subseteq A_2 \subseteq \ldots$, so*

$$\mu\left(\bigcup\nolimits_i A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

(iv) *Ist $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}$ mit $A_1 \supseteq A_2 \supseteq \ldots$ und $\mu(A_1) < \infty$, so*

$$\mu\left(\bigcap\nolimits_i A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

Beweis. (i) Die Menge $\{1, 2, \ldots, m\}$ ist abzählbar.

(ii) Es gilt $\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

(iii) Sei $B_1 = A_1$ und $B_i = A_i \setminus A_{i-1}$ für $i \geq 2$. Dann sind die B_i paarweise disjunkt und $\bigcup_i A_i = \bigcup B_i$. Daher gilt

$$\mu\left(\bigcup\nolimits_i A_i\right) = \sum_{j=1}^{\infty} \mu(B_j) = \lim_{i\to\infty} \sum_{j=1}^{i} \mu(B_j) = \lim_{i\to\infty} \mu(A_i).$$

(iv) Sei $C_i = A_1 \setminus A_i$ für $i \geq 1$. Dann $C_1 \subseteq C_2 \subseteq \ldots$ und daher

$$\begin{aligned}\mu(A_1) - \mu\left(\bigcap\nolimits_i A_i\right) &= \mu\left(A_1 \setminus \left(\bigcap\nolimits_i A_i\right)\right) = \mu\left(\bigcup\nolimits_i C_i\right) = \lim_{i\to\infty} \mu(C_i)\\ &= \lim_{i\to\infty} (A_1 - \mu(A_i)) = \mu(A_1) - \lim_{i\to\infty} \mu(A_i),\end{aligned}$$

also $\mu\left(\bigcap_i A_i\right) = \lim_{i\to\infty} \mu(A_i)$. ∎

[1] Aus $\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)$ und $\mu(A) < \infty$ folgt $\mu(\emptyset) = 0$.

Bemerkung. a) Im Fall $A_1 \subseteq A_2 \subseteq \dots$ schreibt man $A_i \uparrow \bigcup_i A_i$, im Fall $A_1 \supseteq A_2 \supseteq \dots$ schreibt man $A_i \downarrow \bigcap_i A_i$.

b) Allgemeiner definiert man für eine Folge $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}$

$$\liminf_i A_i = \bigcup_i \bigcap_{j\geq i} A_j, \quad \limsup_i A_i = \bigcap_i \bigcup_{j\geq i} A_j.$$

Dann $\liminf_i A_i \in \mathscr{A}$, $\limsup_i A_i \in \mathscr{A}$ und

- $\mu(\liminf_i A_i) \leq \liminf_i \mu(A_i)$,
- $\mu(\limsup_i A_i) \geq \limsup_i \mu(A_i)$, falls $\mu(\bigcup_{i\geq j} A_j) < \infty$ für ein i.

Man schreibt $A_i \to A$, falls $A = \liminf_i A_i = \limsup_i A_i$. Dann gilt $\mu(A_i) \to \mu(A)$, falls zusätzlich $\mu(\bigcup_{i\geq j} A_j) < \infty$ für ein i.

Beispiel 10.6. a) Für $B \subseteq \Omega$ wählen wir $\mathscr{A} = \mathscr{P}(\Omega)$ und setzen für eine beliebige Menge $A \subseteq \Omega$

$$\mu_B(A) = \sharp\,(A \cap B)\,,$$

d. h. $\mu_B(A)$ ist die Anzahl der Elemente der Menge $A \cap B$. Spezielle Fälle dieses Maßes sind

- das *Zählmaß* $\mu_{\text{count}} = \mu_\Omega$,
- das in ω^0 konzentrierte *Punktmaß* $\mu_{\omega^0} = \mu_{\{\omega^0\}}$ für ein $\omega^0 \in \Omega$.

b) Sei $f\colon \Omega \to [0,\infty]$. Für $A \subseteq \Omega$ setzen wir

$$\mu_f(A) = \begin{cases} \sum_{\omega\in A} f(\omega), & \text{falls } A \cap \{f > 0\} \text{ abzählbar,} \\ \infty, & \text{andernfalls.}^2 \end{cases}$$

Das vorige Beispiel ergibt sich im Spezialfall der *Indikatorfunktion* $f = \chi_B$, wobei $\chi_B(\omega) = 1$, falls $\omega \in B$, und $\chi_B(\omega) = 0$, falls $\omega \notin B$.

c) (Später) Ist $(\Omega, \mathscr{A}, \nu)$ ein Maßraum und $f\colon (\Omega, \mathscr{A}) \to [0,\infty]$ eine messbare Funktion, so können wir

$$\mu_{f,\nu}(A) = \int_A f(\omega)\, d\nu(\omega), \quad A \in \mathscr{A},$$

setzen. Das Beispiel b) ergibt sich für $\mathscr{A} = \mathscr{P}(\Omega)$ und $\nu = \mu_{\text{count}}$.

d) Ein *Borelmaß* ist ein Maß auf der Borelschen σ-Algebra $\mathscr{B}(X)$ eines topologischen Raumes $(X, \mathscr{T})$.

Noch einige weitere Begriffe:

Definition 10.7. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum.

(i) Das Maß μ heißt *σ-endlich*, falls es eine abzählbare Familie $\{A_i\} \subset \mathscr{A}$ mit $\Omega = \bigcup_i A_i$ und $\mu(A_i) < \infty$ für alle i gibt.

(ii) Das Maß μ heißt *endlich*, falls $\mu(\Omega) < \infty$.

[2] Im Allgemeinen setzt man $\sum_{\omega\in A} f(\omega) = \sup_{B\,\subseteq\, A,\, B \text{ endlich}} \sum_{\omega\in B} f(\omega)$. Ist die Menge $A \cap \{f > 0\}$ überabzählbar, so ergibt sich Unendlich als Wert der Summe.

(iii) Das Maß μ heißt ein *Wahrscheinlichkeitsmaß* (und $(\Omega, \mathscr{A}, \mu)$ ein *Wahrscheinlichkeitsraum*), falls $\mu(\Omega) = 1$.

Beispiel 10.8. Das in Beispiel 10.6 b) konstruierte Maß μ_f ist genau dann σ-endlich, wenn f den Wert Unendlich nicht annimmt und die Menge $\{f > 0\}$ abzählbar ist.[3,4] Das Maß μ_f ist endlich, wenn (zusätzlich) $\sum_{\omega\in\Omega} f(\omega) < \infty$ gilt, und ein Wahrscheinlichkeitsmaß, wenn letztere Summe gleich Eins ist.

10.2 Konstruktion von Maßen

Für spätere Anwendungen wird es enorm wichtig sein, in geeigneter Weise Maße konstruieren zu können.

10.2.1 Äußere Maße

Sei wiederum Ω eine Menge.

Definition 10.9. Ein *äußeres Maß* auf Ω ist eine Abbildung $\mu_*\colon \mathscr{P}(\Omega) \to [0, \infty]$ mit folgenden Eigenschaften:

(i) $\mu_*(\emptyset) = 0$.

(ii) (Monotonie) $A \subseteq B$ impliziert $\mu_*(A) \leq \mu_*(B)$.

(iii) (σ-Subadditivität) Für abzählbare Familien $\{A_i\}$ von Teilmengen von Ω gilt

$$\mu_*\left(\bigcup_i A_i\right) \leq \sum_i \mu_*(A_i).$$

Zu jedem äußeren Maß gehören in natürlicher Weise eine σ-Algebra und ein Maß auf dieser.

Definition 10.10. Sei μ_* ein äußeres Maß. Eine Menge $A \subseteq \Omega$ heißt *messbar* (oder *Carathéodory-messbar*) bezüglich μ_*, falls für alle $B \subseteq \Omega$

$$\mu_*(B) = \mu_*(A \cap B) + \mu_*(A^c \cap B)$$

gilt.

Satz 10.11 (Carathéodory)**.** *Die Mengen der μ_*-messbaren Mengen bildet eine σ-Algebra $\mathscr{A}_C$ und $\mu_C = \mu_*\big|_{\mathscr{A}_C}$ ist ein Maß auf $\mathscr{A}_C$.*

Beweis. (i) Offenbar gilt $\Omega \in \mathscr{A}_C$ und mit A gehört auch A^c zu $\mathscr{A}_C$.

(ii) Als Nächstes zeigen wir, dass $\mathscr{A}_C$ abgeschlossen unter endlichen Vereinigungen und μ_* endlich additiv auf $\mathscr{A}_C$ ist.

[3]Ist die Menge $\{f > 0\}$ überabzählbar und $\{A_i\}$ eine abzählbare Überdeckung von Ω, $\Omega = \bigcup_i A_i$, so gibt es ein i^* derart, dass die Menge $A_{i^*} \cap \{f > 0\}$ überabzählbar ist. Insbesondere ist dann $\mu_f(A_{i^*}) = \infty$.

[4]Im Spezialfall $f = \chi_B$ bedeutet das gerade, dass die Menge B abzählbar ist.

Seien dazu A, $B \in \mathscr{A}_C$, $C \subseteq \Omega$. Dann gilt

$$\begin{aligned}\mu_*(C) &= \mu_*(A \cap B \cap C) + \mu_*(A^c \cap B \cap C) + \mu_*(A \cap B^c \cap C) \\ &\quad + \mu_*(A^c \cap B^c \cap C) \\ &\geq \mu_*((A \cup B) \cap C) + \mu_*((A \cup B)^c \cap C)\end{aligned}$$

wegen $A \cup B = (A \cap B) \cup (A^c \cap B) \cup (A \cap B^c)$ und $(A \cup B)^c = A^c \cap B^c$, was $A \cup B \in \mathscr{A}_C$ demonstriert.
Sei nun zusätzlich $A \cap B = \emptyset$. Dann

$$\mu_*(A \cup B) = \mu_*(A \cap (A \cup B)) + \mu_*(A^c \cap (A \cup B)) = \mu_*(A) + \mu_*(B).$$

(iii) Schließlich zeigen wir, dass $\mathscr{A}_C$ abgeschlossen unter abzählbaren Vereinigungen und μ_* σ-additiv auf $\mathscr{A}_C$ ist.
Sei dazu $\{A_i\}_{i \in \mathbb{N}} \subset \mathscr{A}_C$, $A_i \cap A_j = \emptyset$ für $\neq j$. Wir setzen $A = \bigcup_i A_i$ und $B_k = \bigcup_{i=1}^k A_i \in \mathscr{A}_C$ für $k \in \mathbb{N}$. Für $C \subseteq \Omega$ erhalten wir induktiv

$$\mu_*(B_k \cap C) = \mu_*(A_k \cap C) + \mu_*(B_{k-1} \cap C) = \ldots = \sum_{i=1}^k \mu_*(A_i \cap C)$$

und damit wegen $A^c \subseteq (B_k)^c$

$$\mu_*(C) = \mu_*(B_k \cap C) + \mu_*((B_k)^c \cap C) \geq \sum_{i=1}^k \mu_*(A_i \cap C) + \mu_*(A^c \cap C).$$

Im Grenzwert für $k \to \infty$ ergibt sich

$$\mu_*(C) \geq \sum_{i=1}^\infty \mu_*(A_i \cap C) + \mu_*(A^c \cap C) \geq \mu_*(A \cap C) + \mu_*(A^c \cap C) \geq \mu_*(C).$$

Es folgt

$$\mu_*(C) = \mu_*(A \cap C) + \mu_*(A^c \cap C),$$

also $A = \bigcup_i A_i \in \mathscr{A}_C$, sowie im Fall $C = A$

$$\mu_*(A) = \sum_{i=1}^\infty \mu_*(A_i). \qquad \square$$

Beispiel 10.12 (Das äußere Lebesgue-Maß). Für $A \subseteq \mathbb{R}^n$ setzen wir

$$\lambda_{n*}(A) = \inf \left\{ \sum_i |Q_i| \;\middle|\; \{Q_i\} \text{ ist abzählbare Familie von achsenparallelen, abgeschlossenen Würfeln mit } A \subseteq \bigcup_i Q_i \right\},$$

wobei $|Q|$ das Volumen eines n-dimensionalen Würfel $Q \subseteq \mathbb{R}^n$ bezeichnet (siehe Abbildung 10.1).
Wir werden später sehen, dass λ_{n*} ein äußeres Maß auf $\mathbb{R}^n$ ist, was uns dann zum Begriff einer Lebesgue-messbaren Menge und zum Lebesgue-Maß führt.

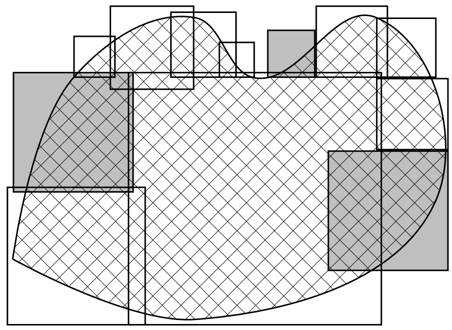

Abbildung 10.1: Überdeckung einer Volumens durch achsenparallele Würfel

10.2.2 Prämaße

Es gibt einen natürlichen Prozess, um von einem Prämaß auf einer Algebra über ein äußeres Maß schließlich zu einem Maß auf einer σ-Algebra zu gelangen. Zudem werden wir sehen, dass Prämaße aufs Engste mit der anschaulichen Vorstellung des Maßbegriffs verbunden sind.
Eine *Algebra* ist eine Menge von Teilmengen von Ω, die Ω enthält und abgeschlossen unter Komplementen und endlichen Vereinigungen ist.

Definition 10.13. Sei $\mathscr{A}_0$ eine Algebra auf Ω. Dann heißt eine Abbildung $\mu_0 \colon \mathscr{A}_0 \to [0, \infty]$ ein *Prämaß*, falls die folgenden Bedingungen erfüllt sind:

(i) $\mu_0(\emptyset) = 0$.

(ii) Ist $\{A_i\} \subset \mathscr{A}_0$ eine abzählbare Familie mit $A_i \cap A_j = \emptyset$ für $i \neq j$ und $\bigcup_i A_i \in \mathscr{A}_0$, so
$$\mu_0\Big(\bigcup_i A_i\Big) = \sum_i \mu_0(A_i).$$

Insbesondere ist ein Prämaß μ_0 endlich additiv auf $\mathscr{A}_0$.

Satz 10.14. *Sei μ_0 ein Prämaß auf $\mathscr{A}_0$ und sei*
$$\mu_*(A) = \inf\left\{ \sum_i \mu_0(A_i) \;\middle|\; \{A_i\} \subset \mathscr{A}_0 \text{ ist abzählbare Familie mit } A \subseteq \bigcup_i A_i \right\}$$
für $A \subseteq \Omega$. Dann gilt:

(i) *μ_* ist ein äußeres Maß auf Ω.*

(ii) *$\mu_*(A) = \mu_0(A)$ für alle $A \in \mathscr{A}_0$.*

(iii) *Jede Menge in $\mathscr{A}_0$ ist Carathéodory-messbar bezüglich μ_*.*

Beweis. (i) Wir zeigen zuerst, dass μ_* ein äußeres Maß ist. Dazu müssen wir die σ-Subadditivität von μ_* beweisen.

Sei also $\{A_i\}_{i\in\mathbb{N}}$ eine Familie von Teilmengen von Ω und sei $\epsilon > 0$. Wir jedes i wählen wir eine abzählbare Familie $\{A_{ij}\} \subset \mathscr{A}_0$ mit $A_i \subseteq \bigcup_j A_{ij}$ und $\mu_*(A_i) \geq \sum_j \mu_0(A_{ij}) - \epsilon/2^i$. Summation über i liefert wegen $A \subseteq \bigcup_{ij} A_{ij}$, dass

$$\sum_i \mu_*(A_i) + \epsilon \geq \sum_{i,j} \mu_0(A_{ij}) \geq \mu_*(A).$$

Da dies für alle $\epsilon > 0$ richtig ist, folgt $\mu_*(A) \leq \sum_i \mu_*(A_i)$.

(ii) Nun zeigen wir, dass μ_* das Prämaß μ_0 fortsetzt.
Sei also $A \in \mathscr{A}_0$. Da A sich selbst überdeckt, folgt $\mu_*(A) \leq \mu_0(A)$. Sei weiterhin $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}_0$ mit $A \subseteq \bigcup_i A_i$. Wir setzen $B_k = A \cap \left(A_k \setminus \bigcap_{i=1}^{k-1} A_i\right) \in \mathscr{A}_0$. Dann sind die B_i paarweise disjunkt, $B_i \subseteq A_i$ und $A = \bigcup_i B_i \in \mathscr{A}_0$. Es folgt

$$\mu_0(A) = \sum_i \mu_0(B_i) \leq \sum_i \mu_0(A_i),$$

also $\mu_0(A) \leq \mu_*(A)$.

(iii) Schließlich zeigen wir, dass alle Mengen in $\mathscr{A}_0$ messbar bezüglich μ_* sind. Seien dazu $A \in \mathscr{A}_0$, $B \subseteq \Omega$ und $\epsilon > 0$. Wir wählen eine abzählbare Familie $\{A_i\} \subset \mathscr{A}_0$ mit $\sum_i \mu_0(A_i) \leq \mu_*(B) + \epsilon$. Es folgt

$$\mu_*(B) + \epsilon \geq \sum_i \mu_0(A \cap A_i) + \sum_i \mu_0(A^c \cap A_i) \geq \mu_*(A \cap B) + \mu_*(A^c \cap B),$$

also, da $\epsilon > 0$ beliebig ist,

$$\mu_*(B) \geq \mu_*(A \cap B) + \mu_*(A^c \cap B).$$

Folglich ist A messbar. □

Satz 10.15 (Fortsetzungssatz)**.** *Sei $\mathscr{A}_0$ eine Algebra, μ_0 ein Prämaß auf $\mathscr{A}_0$ und $\mathscr{A} = \sigma(\mathscr{A}_0)$ die von $\mathscr{A}_0$ erzeugte σ-Algebra. Dann gilt:*

(i) *Es gibt ein Maß μ auf $\mathscr{A}$ mit $\mu\big|_{\mathscr{A}_0} = \mu_0$.*

(ii) *Ist das aus dem äußeren Maß μ_* des vorigen Satzes konstruierte Maß μ auf $\mathscr{A}$ σ-endlich, so ist es das einzige Maß auf $\mathscr{A}$ mit $\mu\big|_{\mathscr{A}_0} = \mu_0$.*

Bemerkung. Die Bedingung in (ii) ist äquivalent dazu, dass bereits das Prämaß μ_0 auf $\mathscr{A}_0$ σ-endlich ist.[5]

Beweis. (i) Folgt direkt aus dem vorigen Satz und den Ausführungen über äußere Maße.

(ii) (Eindeutigkeit) Sei ν ein weiteres Maß auf $\mathscr{A}$ mit $\nu\big|_{\mathscr{A}_0} = \mu_0$. Wir zeigen zuerst, dass $\mu(A) = \nu(A)$ für alle $A \in \mathscr{A}$ mit $\mu(A) < \infty$ gilt.

[5] Ist $\{A_i\}_i \subset \mathscr{A}$ eine abzählbare Familie mit $\Omega = \bigcup_i A_i$ und $\mu(A_i) < \infty$ für alle i, so gibt es für jedes i eine abzählbare Familie $\{A_{ij}\} \subset \mathscr{A}_0$ mit $A_i \subseteq \bigcup_j A_{ij}$ und $\mu(A_i) \geq \sum_j \mu_0(A_{ij}) - 1$. Dann jedoch ist $\{A_{ij}\}_{i,j} \subset \mathscr{A}_0$ eine abzählbare Familie mit $\Omega = \bigcup_{i,j} A_{ij}$ und $\mu_0(A_{ij}) < \infty$ für alle i, j.

Für $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}_0$ mit $A \subseteq \bigcup_i A_i$ gilt

$$\nu(A) \leq \sum_i \nu(A_i) = \sum_i \mu_0(A_i),$$

also $\nu(A) \leq \mu(A)$. Weiterhin gilt für eine solche Familie $\{A_i\}$ und $B = \bigcup_i A_i$, dass

$$\nu(B) = \lim_{k\to\infty} \nu\left(\bigcup_{i=1}^k A_i\right) = \lim_{k\to\infty} \mu\left(\bigcup_{i=1}^k A_i\right) = \mu(B).$$

Sei nun $\epsilon > 0$ und die Familie $\{A_i\}$ so gewählt, dass $\mu(B) \leq \mu(A) + \epsilon$. Dann folgt wegen $\mu(A) < \infty$, dass $\mu(B \setminus A) \leq \epsilon$ und

$$\mu(A) \leq \mu(B) = \nu(B) = \nu(A) + \nu(B \setminus A) \leq \nu(A) + \mu(B \setminus A) \leq \nu(A) + \epsilon,$$

also $\mu(A) \leq \nu(A)$.
Jetzt benutzen wir, dass das Maß μ auf $\mathscr{A}$ σ-endlich ist, und wählen eine paarweise disjunkte, abzählbare Familie $\{A_i\} \subset \mathscr{A}$ mit $\Omega = \bigcup_i A_i$ und $\mu(A_i) < \infty$ für alle i. Dann gilt für eine beliebige Menge $A \in \mathscr{A}$

$$\mu(A) = \sum_i \mu(A \cap A_i) = \sum_i \nu(A \cap A_i) = \nu(A),$$

was schließlich $\mu = \nu$ zeigt. □

Für spätere Referenz halten wir auch fest:

Satz 10.16 (Approximationssatz)**.** *Sei μ_0 ein Prämaß auf einer Algebra $\mathscr{A}_0$ und μ_C das oben aus dem durch μ_0 induzierten äußeren Maß konstruierte Maß auf der σ-Algebra $\mathscr{A}_C$ der Carathéodory-messbaren Mengen. Gilt $\mu_C(A) < \infty$ für ein $A \in \mathscr{A}_C$, so gibt es zu jedem $\epsilon > 0$ ein $B \in \mathscr{A}_0$ mit*[6]

$$\mu_C(A \triangle B) \leq \epsilon.$$

Beweis. Zu einem gegebenem $\epsilon > 0$ wählen wir eine Familie $\{B_i\}_{i\in\mathbb{N}} \subset \mathscr{A}_0$ paarweise disjunkter Mengen mit $A \subseteq \bigcup_i B_i$ und $\mu_C(A) \geq \sum_i \mu_0(B_i) - \frac{\epsilon}{4}$. Weiterhin sei k so groß gewählt, dass $\sum_{i>k} \mu_0(B_i) \leq \frac{\epsilon}{4}$. Wir setzen $B = \bigcup_{i=1}^k B_i \in \mathscr{A}_0$. Dann gilt

- $\mu_C(A \setminus B) \leq \sum_{i>k} \mu_0(B_i) \leq \frac{\epsilon}{4}$,
- $\mu_C(B \setminus A) = \mu_C(B) - \mu_C(A \cap B) \leq \mu_C(A) - \mu_C(A \cap B) + \frac{\epsilon}{2} = \mu_C(A \setminus B) + \frac{\epsilon}{2} \leq \frac{3\epsilon}{4}$,

also $\mu_C(A \triangle B) = \mu_C(A \setminus B) + \mu_C(B \setminus A) \leq \epsilon$. □

10.2.3 Vervollständigung von Maßen

Ein Maß μ auf $\mathscr{A}$ kann in trivialer Weise auf eine größere σ-Algebra, die alle Teilmengen von Nullmengen umfasst, fortgesetzt werden. Dabei heißt $N \in \mathscr{A}$ eine *Nullmenge*, falls $\mu(N) = 0$ gilt. Eine Idee hinter dieser Konstruktion ist, dass Nullmengen in der Integrationstheorie in dem Sinne eine untergeordnete Rolle spielen, dass ein Integral über eine Nullmenge immer Null ergibt.

[6] $A \triangle B = (A \setminus B) \sqcup (B \setminus A)$ bezeichnet die symmetrische Differenz von A, B.

Definition 10.17. Ein Maß μ auf $\mathscr{A}$ heißt *vollständig*, falls für alle Nullmengen $N \in \mathscr{A}$ und alle $B \subseteq N$ die Beziehung $B \in \mathscr{A}$ gilt.

Beispiel 10.18. Das im Abschnitt 10.2.1 aus einem äußeren Maß μ_* konstruierte Maß μ ist vollständig.

Satz 10.19. *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Dann gilt:*

(i) $\bar{\mathscr{A}} = \left\{ A \cup B \;\middle|\; A \in \mathscr{A},\, B \subseteq N \text{ für eine Nullmenge } N \in \mathscr{A} \right\}$ *ist eine σ-Algebra.*

(ii) *Sei $\bar{\mu}\colon \bar{\mathscr{A}} \to [0, \infty]$ vermöge*

$$\bar{\mu}(A \cup B) = \mu(A)$$

für $A \in \mathscr{A}$, $B \subseteq N$ für eine Nullmenge $N \in \mathscr{A}$ erklärt. Dann ist $\bar{\mu}$ wohldefiniert und ein Maß auf $\bar{\mathscr{A}}$.

(iii) $\bar{\mu}\big|_{\mathscr{A}} = \mu$.

Beweis. Eine einfache Übungsaufgabe. □

Das Maß $\bar{\mu}$ auf $\bar{\mathscr{A}}$ ist vollständig und heißt die *Vervollständigung* μ.

Satz 10.20. *Sei μ_0 ein Prämaß auf einer Algebra $\mathscr{A}_0$, μ_C das oben aus dem durch μ_0 induzierten äußeren Maß konstruierte Maß auf der σ-Algebra $\mathscr{A}_C$ der Carathéodory-messbaren Mengen, $\mathscr{A} = \sigma(\mathscr{A}_0)$ die durch $\mathscr{A}_0$ erzeugte σ-Algebra und $\mu = \mu_C\big|_{\mathscr{A}}$. Dann gilt: Ist μ σ-endlich, so ist $(\Omega, \mathscr{A}_C, \mu_C)$ die Vervollständigung von $(\Omega, \mathscr{A}, \mu)$.*

Beweis. Sei $(\Omega, \bar{\mathscr{A}}, \bar{\mu})$ die Vervollständigung von $(\Omega, \mathscr{A}, \mu)$. Da das Maß μ_C vollständig ist, gilt $\bar{\mathscr{A}} \subseteq \mathscr{A}_C$ und $\bar{\mu} = \mu_C\big|_{\bar{\mathscr{A}}}$. Wir zeigen die umgekehrte Inklusion:

a) Sei zuerst $A \in \mathscr{A}_C$, $\mu_C(A) < \infty$. Für jedes $m \in \mathbb{N}$ gibt es eine Folge $\{A_i^{(m)}\} \subseteq \mathscr{A}_0$, $A_i^{(m)}$ paarweise disjunkt, mit $A \subseteq \bigcup_i A_i^{(m)}$ und

$$\sum_i \mu_0(A_i^{(m)}) - \frac{1}{m} \le \mu_C(A) \le \sum_i \mu_0(A_i^{(m)}).$$

Wir setzen $A_m = \bigcup_i A_i^{(m)} \in \mathscr{A}$. Dann $A \subseteq A_m$ und $\mu_C(A) \le \mu(A_m) \le \mu_C(A) + \frac{1}{m}$ für alle m, also $\liminf A_m \in \mathscr{A}$, $A \subseteq \liminf A_m$ und

$$\mu_C(A) \le \mu(\liminf A_m) \le \lim \mu(A_m) = \mu_C(A),$$

folglich $\mu_C(A) = \mu(\liminf A_m)$.

Wenn $\mu_C(A) = 0$, so existiert insbesondere eine μ-Nullmenge N mit $A \subseteq N$. Nun ist $\mu_C(\liminf A_m \setminus A) = \mu(\liminf A_m) - \mu_C(A) = 0$, also existiert eine μ-Nullmenge N mit $\liminf A_m \setminus A \subseteq N$. Damit

$$A = \liminf A_m \setminus (\liminf A_m \setminus A) \in \bar{\mathscr{A}}.$$

b) Im allgemeinen Fall schreiben wir $\Omega = \bigcup_k B_k$ mit $B_k \in \mathscr{A}$ und $\mu(B_k) < \infty$ für alle k. Sei nun $A \in \mathscr{A}_C$. Gemäß a) existieren $A_k \in \mathscr{A}$, $N_k \in \mathscr{A}$ mit $A_k \subseteq A \cap B_k \subseteq A_k \cup N_k$ und $\mu(N_k) = 0$ für alle k. Dann $\bigcup_k A_k \in \mathscr{A}$, $N = \bigcup_k N_k \in \mathscr{A}$ sowie $\bigcup_k A_k \subseteq A = \bigcup_k (A \cap B_k) \subseteq \bigcup_k A_k \cup N$ und $\mu(N) = 0$. Es ergibt sich, dass $A \in \bar{\mathscr{A}}$. □

10.3 Lebesgue-Stieltjes-Maße auf $\mathbb{R}^n$

Die für diesen Kurs wichtigsten Maße sind Lebesgue-Stieltjes-Maße und als ein Spezialfall das Lebesgue-Maß auf gewissen Teilmengen des $\mathbb{R}^n$. Mit diesen Maßen beschäftigen wir uns jetzt.

10.3.1 Der eindimensionale Fall

Wir untersuchen den eindimensionalen Fall detailliert und tragen dann im mehrdimensionalen Fall nur noch die Ergebnisse zusammen.

Definition 10.21. Ein eindimensionales *Lebesgue-Stieltjes-Maß* ist ein Borelmaß μ auf $\mathbb{R}$ mit $\mu(I) < \infty$ für jedes beschränkte Intervall $I \subseteq \mathbb{R}$.

Beispiel 10.22. a) Die Punktmaße auf $\mathbb{R}$ von Beispiel 10.6 sind Lebesgue-Stieltjes-Maße, das Zählmaß auf $\mathbb{R}$ ist keines.

b) Das Lebesgue-Maß λ_1 auf $\mathbb{R}$ wird sich als das eindeutige Lebesgue-Stieltjes-Maß auf $\mathbb{R}$ herausstellen, für das $\lambda_1(I)$ ist gleich der Länge von I für jedes Intervall $I \subseteq \mathbb{R}$ ist.

Wichtige Beispiele ergeben sich wie folgt:

Lemma 10.23. *Sei $\mathscr{B}_0(\mathbb{R})$ die Menge aller endlichen, disjunkten Vereinigungen von halboffenen, nach rechts abgeschlossenen Intervallen, d. h. die Menge aller Mengen der Form*

$$\bigcup_{i=1}^{m} (a_i, b_i]$$

mit $-\infty \leq a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_m < b_m \leq \infty$. Dann gilt:

(i) *$\mathscr{B}_0(\mathbb{R})$ ist eine Algebra.*

(ii) *Die von $\mathscr{B}_0(\mathbb{R})$ erzeugte σ-Algebra ist die Borelsche σ-Algebra $\mathscr{B}(\mathbb{R})$.*

Beweis. (i) Dies ist offensichtlich.

(ii) Sei $-\infty \leq a < b \leq \infty$ und $\{b_m\} \subset \mathbb{R}$ eine Folge mit $b_m \uparrow b$. Dann folgt

$$(a, b_m] \uparrow (a, b),$$

und daher die Behauptung. □

Sei nun $F \colon \mathbb{R} \to \mathbb{R}$ eine monoton wachsende, rechtsseitig stetige [7] Funktion. Wir setzen für Elemente in $\mathscr{B}_0(\mathbb{R})$

$$\mu_0\left(\bigcup_{i=1}^{m} (a_i, b_i]\right) = \sum_{i=1}^{m} \left(F(b_i) - F(a_i)\right).$$

Lemma 10.24. *μ_0 ist wohldefiniert und endlich additiv auf $\mathscr{B}_0(\mathbb{R})$.*

[7] Das heißt, es gilt $F(x+0) = F(x)$ für alle $x \in \mathbb{R}$.

Beweis. Die Wohldefiniertheit von μ_0 folgt aus

$$F(b) - F(a) = (F(b) - F(c)) + (F(c) - F(a)),$$

d. h. $\mu((a,b]) = \mu((a,c]) + \mu((c,b])$ für $-\infty \leq a < c < b \leq \infty$. Die endliche Additivität ergibt sich direkt aus der Definition. □

Tatsächlich wird sich herausstellen, dass μ_0 ein Prämaß ist, also ein Lebesgue-Stieltjes-Maß auf $\mathbb{R}$ gemäß der Konstruktion in Abschnitt 10.2 induziert, und dass weiterhin *jedes* Lebesgue-Stieltjes-Maß auf $\mathbb{R}$ auf diese Weise entsteht.

Satz 10.25. *Die vorangegangene Konstruktion vermittelt eine Bijektion zwischen Lebesgue-Stieltjes-Maßen μ auf $\mathbb{R}$ und monoton wachsenden, rechtsseitig stetigen Funktionen $F\colon \mathbb{R} \to \mathbb{R}$, wenn man zusätzlich* (*beispielsweise*) $F(0) = 0$ *fordert.*

Beweis. ($\Rightarrow$) Sei μ ein Lebesgue-Stieltjes-Maß auf $\mathbb{R}$. Wir setzen

$$F(x) = \begin{cases} \mu\left((0,x]\right), & \text{falls } x \geq 0, \\ -\mu\left((x,0]\right), & \text{falls } x < 0. \end{cases}$$

Dann ist F monoton wachsend, rechtsseitig stetig und es gilt $F(0) = 0$ sowie

$$\mu\left((a,b]\right) = F(b) - F(a)$$

für alle $-\infty \leq a < b \leq \infty$. Um letztere Formel einzusehen, unterscheiden wir drei Fälle:

- $0 \leq a < b \leq \infty$: $\mu\left((a,b]\right) = \mu\left((0,b]\right) - \mu\left((0,a]\right) = F(b) - F(a)$.

- $-\infty \leq a < 0 \leq b \leq \infty$: $\mu\left((a,b]\right) = \mu\left((0,b]\right) + \mu\left((a,0]\right) = F(b) - F(a)$.

- $-\infty \leq a < b < 0$: $\mu\left((a,b]\right) = -\mu\left((b,0]\right) + \mu\left((a,0]\right) = F(b) - F(a)$.

Aus der Formel ergibt sich sofort die Monotonie von F: Für $-\infty \leq x < y \leq \infty$ gilt

$$F(y) = F(x) + \mu\left((x,y]\right) \geq F(x).$$

Um die rechtsseitige Stetigkeit einzusehen, sei $\{x_n\} \subset \mathbb{R}$ eine Folge mit $x_n \downarrow x$. Wir unterscheiden zwei Fälle:

- $x \geq 0$: Dann $(0,x_n] \downarrow (0,x]$, folglich

  $$F(x_n) = \mu\left((0,x_n]\right) \to \mu\left((0,x]\right) = F(x).$$

- $x < 0$: Dann $(x_n,0] \uparrow (x,0]$, folglich

  $$F(x_n) = \mu\left((x_n,0]\right) \to \mu\left((x,0]\right) = F(x).$$ □

Um die umgekehrte Richtung zu beweisen, müssen wir zeigen, dass die oben aus F konstruierte Mengenfunktion μ_0 auf $\mathscr{B}_0(\mathbb{R})$ ein Prämaß ist. Wir bereiten diesen Beweis mit zwei Hilfsaussagen vor.

Lemma 10.26. *Sei $\mathscr{A}_0$ eine Algebra, $\mu_0\colon \mathscr{A}_0 \to [0,\infty]$ endlich additiv und $\mu_0(\Omega) < \infty$. Dann ist μ_0 genau dann ein Prämaß, wenn für jede Folge $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}_0$ mit $A_i \downarrow \emptyset$ $\mu_0(A_i) \to 0$ gilt.*

Beweis. ($\Rightarrow$) Sei $\{A_i\} \subset \mathscr{A}_0$, $A_i \downarrow \emptyset$. Dann $\{A_i^c\} \subset \mathscr{A}_0$, $A_i^c \uparrow \Omega$ und damit

$$\mu_0(\Omega) - \mu_0(A_i) = \mu_0(A_i^c) \to \mu_0(\Omega),$$

d. h. $\mu_0(A_i) \to 0$.

($\Leftarrow$) Sei $\{A_i\} \subset \mathscr{A}_0$, A_i paarweise disjunkt, $A = \bigcup_i A_i \in \mathscr{A}_0$. Wir setzen $B_k = A \setminus \bigcup_{i=1}^k A_i \in \mathscr{A}_0$. Dann gilt $B_1 \supseteq B_2 \supseteq \ldots$, $\bigcap_i B_i = A \setminus \bigcup_{i=1}^\infty A_i = \emptyset$, also $B_i \downarrow \emptyset$ und $\mu_0(B_i) \to 0$. Es folgt

$$\mu_0(A) = \lim_{k\to\infty} \mu_0\left(\bigcup_{i=1}^k A_i\right) = \lim_{k\to\infty} \sum_{i=1}^k \mu_0(A_i) = \sum_{i=1}^\infty \mu_0(A_i). \qquad \square$$

Lemma 10.27. *Sei $\mathscr{A}_0$ eine Algebra, $\mu_0\colon \mathscr{A}_0 \to [0,\infty]$ endlich additiv, $\{A_i\} \subset \mathscr{A}_0$, A_i paarweise disjunkt und $\bigcup_i A_i \in \mathscr{A}_0$. Dann gilt*

$$\mu_0\left(\bigcup_i A_i\right) \geq \sum_i \mu_0(A_i).$$

Beweis. Für jedes $k \in \mathbb{N}$ gilt $\bigcup_{i=1}^\infty A_i \supseteq \bigcup_{i=1}^k A_i$, also $\mu_0\left(\bigcup_{i=1}^\infty A_i\right) \geq \mu_0\left(\bigcup_{i=1}^k A_i\right) = \sum_{i=1}^k \mu_0(A_i)$. Die Behauptung ergibt sich für $k \to \infty$. $\square$

Abschluss des Beweises von Satz 10.25. ($\Leftarrow$) Sei $F\colon \mathbb{R} \to \mathbb{R}$ monoton wachsend, rechtsseitig stetig und $\mu_0\colon \mathscr{B}_0(\mathbb{R}) \to [0,\infty]$ wie zuvor konstruiert. Um zu zeigen, dass μ_0 ein Prämaß ist, unterscheiden wir zwei Fälle:
a) $F(\infty) - F(-\infty) < \infty$. Sei $\{A_i\} \subset \mathscr{B}_0(\mathbb{R})$, $A_i \downarrow \emptyset$. Wir wählen eine Familie $\{B_i\} \subset \mathscr{B}_0(\mathbb{R})$ mit $\overline{B}_i \subseteq A_i$[8] und $\mu_0(A_i \setminus B_i) \leq \epsilon/2$ für alle i und ein gegebenes $\epsilon > 0$. Es gilt $\bigcap_i \overline{B}_i = \emptyset$, also existiert ein k mit $\bigcap_{i=1}^k B_i \subseteq \bigcap_{i=1}^k \overline{B}_i = \emptyset$.[9] Damit gilt für $m \geq k$

$$\begin{aligned}\mu_0(A_m) &= \mu_0\left(A_m \setminus \bigcap_{i=1}^m B_i\right) + \mu_0\left(\bigcap_{i=1}^m B_i\right) = \mu_0\left(A_m \setminus \bigcap_{i=1}^m B_i\right) \\ &\leq \mu_0\left(\bigcup_{i=1}^m (A_i \setminus B_i)\right) \leq \sum_{i=1}^m \mu_0(A_i \setminus B_i) \leq \epsilon,\end{aligned}$$

woraus $\mu_0(A_i) \to 0$ folgt.
b) $F(\infty) - F(-\infty) = \infty$. In diesem Fall definieren wir für $n \in \mathbb{N}$

$$F_n(x) = \begin{cases} F(n), & \text{falls } n \leq x, \\ F(x), & \text{falls } -n \leq x < n, \\ F(-n-0), & \text{falls } x < -n. \end{cases}$$

[8] Abschluss von B_i in $\overline{\mathbb{R}}$.

[9] Der Raum $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ mit der Topologie der Punktkompaktifizierung von $\mathbb{R}$ an seinen beiden Enden ist kompakt und $\{\overline{B}_i^c\}$ ist eine offene Überdeckung von $\overline{\mathbb{R}}$.

Dann ist $F_n\colon \mathbb{R} \to \mathbb{R}$ monoton wachsend, rechtsseitig stetig und $F_n(\infty) - F_n(-\infty) < \infty$. Sei μ_n das zugehörige Lebesgue-Stieltjes-Maß auf $\mathbb{R}$, das gemäß a) existiert. Nach Konstruktion gilt $\mu_n\big|_{\mathscr{B}_0(\mathbb{R})} \le \mu_0$ für alle n sowie

$$\mu_n\,(A) \to \mu_0\,(A) \quad \text{für } n \to \infty, \quad \forall A \in \mathscr{B}_0(\mathbb{R}).$$

Sei nun $\{A_i\} \subset \mathscr{B}_0(\mathbb{R})$, A_i paarweise disjunkt, $A = \bigcup_i A_i \in \mathscr{B}_0(\mathbb{R})$. Wir wollen $\mu_0\,(A) \le \sum_i \mu_0\,(A_i)$ zeigen. Gilt $\sum_i \mu_0\,(A_i) = \infty$, so bleibt nichts zu tun. Andernfalls benutzen wir

$$\mu_0\,(A) = \lim_{n\to\infty} \mu_n\,(A) = \lim_{n\to\infty} \sum_{i=1}^{\infty} \mu_n\,(A_i)$$

und erhalten

$$\mu_0\,(A) - \sum_{i=1}^{\infty} \mu_0\,(A_i) = \lim_{n\to\infty} \sum_{i=1}^{\infty} \big(\underbrace{\mu_n\,(A_i) - \mu_0\,(A_i)}_{\le 0} \big) \le 0,$$

woraus die Behauptung folgt. □

Definition 10.28. Ist μ ein Lebesgue-Stieltjes-Maß auf $\mathbb{R}$, so heißt ein zugehöriges F eine *Verteilungsfunktion* von μ. Man schreibt auch $\mu = dF$.

Bemerkung. a) Eine Verteilungsfunktion ist bis auf eine additive Konstante eindeutig bestimmt.

b) Im Fall eines Wahrscheinlichkeitsmaßes μ normiert man F häufig so, dass $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$ gilt.

Bemerkung. Sei F eine Verteilungsfunktion eines Lebesgue-Stieltjes-Maßes μ.

a) Für $-\infty \le a < b \le \infty$ gilt $\mu\,((a,b)) = F(b-0) - F(a)$, $\mu\,([a,b]) = F(b) - F(a-0)$, usw.

b) F ist genau dann an einer Stelle $a \in \mathbb{R}$ stetig (d. h. $F(a-0) = F(a)$), wenn $\mu\,(\{a\}) = 0$ gilt.

Beispiel 10.29. a) Seien $-\infty < a_1 < a_2 < \ldots < a_m < \infty$, $f_1 > 0, \ldots, f_m > 0$. Dann ist μ mit

$$\mu\,(A) = \sum_{i\colon a_i \in A} f_i, \quad \forall A \in \mathscr{B}(\mathbb{R}),$$

ein Lebesgue-Stieltjes-Maß auf $\mathbb{R}$ mit der Verteilungsfunktion

$$F(x) = \begin{cases} 0, & \text{falls } x < a_1, \\ \sum_{i=1}^{k} f_i, & \text{falls } a_k \le x < a_{k+1},\ 1 \le k < m, \\ \sum_{i=1}^{m} f_i, & \text{falls } a_m \le x. \end{cases}$$

b) Sei $f\colon \mathbb{R} \to \mathbb{R}$ auf jedem kompakten Intervall R-integrierbar, $f \ge 0$. Dann ist

$$F(x) = \int_0^x f(t)\,dt, \quad \forall x \in \mathbb{R},$$

eine stetige Verteilungsfunktion für ein Lebesgue-Stieltjes-Maß μ auf $\mathbb{R}$ mit

$$\mu\left([a,b]\right) = \mu\left((a,b)\right) = \int_a^b f(t)\,dt$$

für alle a, $b \in \mathbb{R}$, $a < b$. *Formal* haben wir hier, dass

$$dF(x) = f(x)\,dx$$

gilt, wobei dx für das Lebesgue-Maß steht. Wir werden diese Beziehung im Abschnitt über die Differenziation von Maßen näher studieren.

Definition 10.30. Das *Lebesgue-Maß* λ_1 auf $\mathbb{R}$ ist das Lebesgue-Stieltjes-Maß dx.

Bemerkung. Im Weiteren bezeichnen wir mit $\bar{\mathscr{B}}(\mathbb{R})$ die σ-Algebra der *Lebesgue-messbaren* Teilmengen von $\mathbb{R}$, die sich als die σ-Algebra aller messbaren Mengen bezüglich des äußeren Lebesgue-Maßes λ_{1*} ergibt. Aus Satz 10.20 folgt

$$\bar{\mathscr{B}}(\mathbb{R}) = \left\{A \subseteq \mathbb{R} \,\middle|\, \exists B \in \mathscr{B}(\mathbb{R})\colon \lambda_{1*}(A \triangle B) = 0\right\}.$$

Anhang zu Abschnitt 10.3.1: Eine nicht Lebesgue-messbare Teilmenge von $\mathbb{R}$

Die folgende Konstruktion benötigt das *Auswahlaxiom.* Auf dem Intervall $[0,1]$ führen wir die Äquivalenzrelation

$$x \sim y \;:\Longleftrightarrow\; x - y \in \mathbb{Q}$$

ein. Sei $[x]$ die Äquivalenzklasse, die x enthält. Das Auswahlaxiom gestattet uns die Wahl einer Menge $A \subseteq [0,1]$ mit

$$\sharp\,(A \cap [x]) = 1, \quad \forall x \in [0,1].$$

Lemma 10.31. *Die Menge A ist nicht Lebesgue-messbar.*

Beweis. Sei $\{q_k\}_{k\in\mathbb{N}}$ ein Aufzählung aller rationalen Zahlen in $[-1,1]$. Wir setzen $A_k = q_k + A$.
Nun angenommen, die Menge A ist Lebesgue-messbar. Dann ist jede der Mengen A_k Lebesgue-messbar und es gilt $\lambda_1(A_k) = \lambda_1(A)$ für alle k.[10] Weiterhin gilt $A_k \cap A_l = \emptyset$ für $k \neq l$. Wir erhalten

$$[0,1] \subseteq \bigsqcup_k A_k \subseteq [-1,2],$$

wobei man die erste Enthaltenseinsbeziehung wie folgt einsieht: Sei $x \in [0,1]$ und $y \in A \cap [x]$. Dann $x - y \in \mathbb{Q}$ und $|x-y| \leq 1$, also $x - y = q_k$ für (genau) ein k. Es folgt, dass $x \in A_k$.
Wir folgern, dass

$$1 \leq \sum_k \lambda_1(A_k) \leq 3,$$

woraus sich sowohl $\lambda_1(A) = 0$ als auch $\lambda_1(A) > 0$ ergibt; ein Widerspruch. Somit muss die Annahme, dass die Menge A Lebesgue-messbar ist, fallengelassen werden. □

[10] Es ist leicht zu sehen, dass das Lebesgue-Maß translationsinvariant ist, d. h. es gilt $\lambda_1(b+B) = \lambda_1(B)$ für alle $b \in \mathbb{R}$, $B \in \bar{\mathscr{B}}(\mathbb{R})$. Tatsächlich kann man zeigen, dass alle σ-endlichen, translationsinvarianten Borelmaße auf $\mathbb{R}$ von der Form $c\lambda_1$ für ein $c \in \mathbb{R}$, $c \geq 0$ sind.

10.3.2 Der mehrdimensionale Fall

Sei $a = (a_1, \ldots, a_n) \in \overline{\mathbb{R}}^n$, $b = (b_1, \ldots, b_n) \in \overline{\mathbb{R}}^n$ und $a < b$, d. h. es gilt $a_i < b_i$ für alle $1 \le i \le n$. Wir setzen

$$(a, b] = \bigtimes_{i=1}^{n} (a_i, b_i] = \left\{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \;\middle|\; a_i < x_i \le b_i \;\forall i\right\}.$$

Definition 10.32. Ein *Lebesgue-Stieltjes-Maß* auf $\mathbb{R}^n$ ist ein Borelmaß μ auf $\mathbb{R}^n$ mit $\mu\left((a, b]\right) < \infty$ für alle a, $b \in \mathbb{R}^n$, $a < b$.

Auch hier gibt es eine Bijektion zwischen Lebesgue-Stieltjes-Maßen auf $\mathbb{R}^n$ und Verteilungsfunktionen. Um zu sehen, welche Eigenschaften Verteilungsfunktionen charakterisieren, nehmen wir der Einfachheit halber an, dass μ ein endliches Maß ist, d. h. dass $\mu\left(\mathbb{R}^n\right) < \infty$ gilt. Wir setzen

$$F(x) = \mu\left(\bigtimes_{i=1}^{n} (-\infty, x_i]\right), \quad x \in \mathbb{R}^n.$$

Dann gilt im Fall $n = 2$

$$\begin{aligned}
\mu\left((a, b]\right) &= \mu\left((a_1, b_1] \times (a_2, b_2]\right) \\
&= \mu\left((-\infty, b_1] \times (-\infty, b_2]\right) - \mu\left((-\infty, a_1] \times (-\infty, b_2]\right) \\
&\qquad - \mu\left((-\infty, b_1] \times (-\infty, a_2]\right) + \mu\left((-\infty, a_1] \times (-\infty, a_2]\right) \\
&= F(b_1, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, a_2).
\end{aligned}$$

Ein analoges Resultat gilt auch im Fall $n \ge 3$, wobei auf der rechten Seite 2^n Summanden erscheinen.
Wir führen für $F(x) = F(x_1, \ldots, x_n)$ die Differenzenoperatoren $\Delta^{(i)}_{a_i, b_i}$,

$$\begin{aligned}
&\left(\Delta^{(i)}_{a_i, b_i} F\right)(x_1, \ldots, \widehat{x}_i, \ldots, x_n) \\
&\qquad = F(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots, x_n) - F(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n),
\end{aligned}$$

ein.[11] Weiterhin setzen wir für a, $b \in \mathbb{R}^n$, $a < b$,

$$\Delta_{a,b} = \Delta^{(1)}_{a_1, b_1} \ldots \Delta^{(n)}_{a_n, b_n}$$

Lemma 10.33. *Für $\mu\left(\mathbb{R}^n\right) < \infty$ und F wie oben definiert gilt*

$$\mu\left((a, b]\right) = \Delta_{a,b} F.$$

Beweis. Dies ist eine Umformulierung des oben Gezeigten. □

Eine Funktion $F\colon \mathbb{R}^n \to \mathbb{R}$ heißt

- *monoton wachsend*, falls $\Delta_{a,b} F \ge 0$ für alle a, $b \in \mathbb{R}^n$, $a < b$,
- *rechtsseitig stetig*, falls $F(x^k) \to F(x)$ für $k \to \infty$, wenn $x^k \downarrow x$ für $k \to \infty$, d. h. $x^k_i \downarrow x_i$ für alle $1 \le i \le n$.

[11] Die Bezeichnung $(x_1, \ldots, \widehat{x}_i, \ldots, x_n)$ bedeutet, dass der i-te Eintrag fehlt, also $(x_1, \ldots, \widehat{x}_i, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

Satz 10.34. *Es gibt eine Bijektion zwischen Lebesgue-Stieltjes-Maßen auf $\mathbb{R}^n$ und monoton wachsenden, rechtsseitig stetigen Funktionen $F\colon \mathbb{R}^n \to \mathbb{R}$, wenn man F (beispielsweise) durch $F(0,\dots,0) = 0$ normiert. Diese Beziehung ist durch*

$$\mu\left((a,b]\right) = \Delta_{a,b}F$$

für alle a, $b \in \mathbb{R}^n$, $a < b$, gegeben.

Definition 10.35. F heißt eine *Verteilungsfunktion* von μ, man schreibt $\mu = dF$.

Beispiel 10.36. Seien $F_1,\dots,F_n$ eindimensionale Verteilungsfunktionen. Dann ist $F(x_1,\dots,x_n) = F_1(x_1)F_2(x_2)\dots F_n(x_n)$ eine n-dimensionale Verteilungsfunktion. Das zugehörige Lebesgue-Stieltjes-Maß wird mit $dF_1 \dots dF_n$ bezeichnet.

Definition 10.37. Das n-dimensionale *Lebesgue-Stieltjes-Maß* λ_n auf $\mathbb{R}^n$ ist $dx_1 \dots dx_n$.

10.4 Der abstrakte Integralbegriff

10.4.1 Messbare Funktionen

Seien $(\Omega,\mathscr{A})$, $(\Omega',\mathscr{B})$ messbare Räume.

Definition 10.38. Eine Abbildung $f\colon \Omega \to \Omega'$ heißt *messbar* (genauer $(\mathscr{A},\mathscr{B})$*-messbar*, falls $f^{-1}(B) \in \mathscr{A}$ für alle $B \in \mathscr{B}$. Ist $(\Omega',\mathscr{B}) = \big(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}})\big)$, so heißt f *Borel-messbar*.

Wir schreiben $\mathscr{M}(\Omega,\Omega')$ für den Raum der messbaren Abbildungen $f\colon \Omega \to \Omega'$. Wichtige Spezialfälle sind $\mathscr{M}(\Omega,\overline{\mathbb{R}})$ und $\mathscr{M}_+(\Omega,\overline{\mathbb{R}}) = \mathscr{M}(\Omega,[0,\infty])$.
Offenbar gilt:

Lemma 10.39. *$f\colon \Omega \to \overline{\mathbb{R}}$ ist genau dann Borel-messbar genau, wenn*

$$\big\{f < a\big\} = \big\{\omega \in \Omega \bigm| f(\omega) < a\big\} \in \mathscr{A}$$

für alle $a \in \mathbb{R}$.

Beweis. Eine einfache Übungsaufgabe. □

Wir studieren zuerst elementare Eigenschaften messbarer Funktionen.

Satz 10.40. (i) *Aus $f \in \mathscr{M}(\Omega,\Omega')$, $g \in \mathscr{M}(\Omega',\Omega'')$ folgt $g \circ f \in \mathscr{M}(\Omega,\Omega'')$.*

(ii) *Sind X, Y topologische Räume, so ist $\mathscr{C}(X,Y) \subseteq \mathscr{M}(X,Y)$.*[12]

(iii) *$\mathscr{M}(\Omega,\mathbb{R})$ ist eine $\mathbb{R}$-Algebra unter den punktweisen Operationen.*

(iv) *$\mathscr{M}_+(\Omega,\overline{\mathbb{R}})$ ist ein konvexer Kegel, d. h. aus f, $g \in \mathscr{M}_+(\Omega,\overline{\mathbb{R}})$ und $c > 0$, $d > 0$ folgt $cf + dg \in \mathscr{M}_+(\Omega,\overline{\mathbb{R}})$.*

[12]Wenn nicht anders vermerkt, versehen wir einen topologischen Raum immer mit seiner Borelschen σ-Algebra.

(v) $\{f_n\} \subset \mathscr{M}(\Omega, \overline{\mathbb{R}})$ *impliziert*

$$\sup f_n,\ \inf f_n,\ \limsup f_n,\ \liminf f_n \in \mathscr{M}(\Omega, \overline{\mathbb{R}}).$$

Insbesondere folgt aus $f_n(\omega) \to f(\omega)$ für alle ω, dass $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$.

(vi) *Ist $(\Omega, \mathscr{A}, \mu)$ ein vollständiger Maßraum, $f \in \mathscr{M}(\Omega, \Omega')$ und $g\colon \Omega \to \Omega'$ habe die Eigenschaft, dass $\mu\big(\{f \neq g\}\big) = 0$, so ist g messbar.*

Beweis. (i), (ii) sind offensichtlich.

(iii) Seien f, $g \in \mathscr{M}(\Omega, \mathbb{R})$. Dann ist

$$\{f + g < a\} = \bigcup_{q \in \mathbb{Q}} \big(\{f < a - q\} \cap \{g < q\}\big)$$

für alle $a \in \mathbb{R}$ messbar, also $f + g \in \mathscr{M}(\Omega, \mathbb{R})$. Außerdem gilt $f^2 \in \mathscr{M}(\Omega, \mathbb{R})$ für $f \in \mathscr{M}(\Omega, \mathbb{R})$ wegen (i), (ii), also

$$fg = \frac{1}{2}\big((f+g)^2 - f^2 - g^2\big) \in \mathscr{M}(\Omega, \mathbb{R}).$$

(iv) Dies ist ähnlich wie (iii).

(v) Es gilt für $a \in \mathbb{R}$

$$\{\inf f_n < a\} = \bigcup_n \{f_n < a\} \in \mathscr{A},$$

also ist $\inf f_n$ messbar. Das Ergebnis in den anderen drei Fällen folgt aus

$$\begin{aligned}
\sup f_n &= -\inf(-f_n),\\
\liminf f_n &= \sup_k \inf_{n \geq k} f_n,\\
\limsup f_n &= \inf_k \sup_{n \geq k} f_n.
\end{aligned}$$

(vi) Sei $B \in \mathscr{B}$. Dann ist

$$g^{-1}(B) = \big(f^{-1}(B) \cap \{f = g\}\big) \sqcup \big(g^{-1}(B) \cap \{f \neq g\}\big)$$

messbar, da $f^{-1}(B) \cap \{f = g\} = f^{-1}(B) \cap \{f \neq g\}^c \in \mathscr{A}$ und $g^{-1}(B) \cap \{f \neq g\} \subseteq \{f \neq g\}$ sogar eine Nullmenge ist. □

Um die Einführung des Integralbegriffs vorzubereiten, zeigen wir nun, dass jede messbare Funktion punktweise durch Stufenfunktionen approximiert werden kann.

Definition 10.41. Eine *Stufenfunktion* $f\colon \Omega \to \mathbb{R}$ ist eine Borel-messbare Funktion, die nur endlich viele Werte annimmt, also von der Form

$$f = \sum_{i=1}^{m} c_i\, \chi_{A_i}$$

für ein $m \in \mathbb{N}$, $c_i \in \mathbb{R}$, $A_i \in \mathscr{A}$ ist.

$\mathscr{M}_0(\Omega, \mathbb{R})$ bezeichne den Raum der Stufenfunktionen $f\colon \Omega \to \mathbb{R}$.

Lemma 10.42. (i) *Sei* $f \in \mathscr{M}_+(\Omega, \overline{\mathbb{R}})$. *Dann existiert eine Folge* $\{f_n\} \subset \mathscr{M}_0(\Omega, \mathbb{R})$ *mit* $0 \le f_n \uparrow f$.

(ii) *Sei* $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$. *Dann existiert eine Folge* $\{f_n\} \subset \mathscr{M}_0(\Omega, \mathbb{R})$ *mit* $f_n \to f$ *und* $|f_n| \le |f|$ *für alle* n.

Beweis. (i) Wir setzen

$$f_n(\omega) = \begin{cases} \frac{k-1}{2^n}, & \text{falls } \frac{k-1}{2^n} \le f(\omega) < \frac{k}{2^n},\ k = 1, 2, \dots, n2^n, \\ n, & \text{falls } f(\omega) \ge n. \end{cases}$$

Dann ist $f_n \in \mathscr{M}_0(\Omega, \mathbb{R})$, $0 \le f_n \uparrow f$.

(ii) Wir wählen Folgen $\{g_n\}$, $\{h_n\} \subset \mathscr{M}_0(\Omega, \mathbb{R})$ mit $0 \le g_n \uparrow f^+ = \max\{f, 0\}$, $0 \le h_n \uparrow f^- = \max\{-f, 0\}$ und setzen $f_n = g_n - h_n$. □

10.4.2 Definition des Integrals

Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Wir möchten ein Integral $\int_\Omega f(\omega)\, d\mu(\omega)$ für gewisse messbare Funktionen $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ einführen .

Bemerkung (die Notation betreffend). Eigentlich müsste man

$$\int_\Omega f(\omega)\, \mu(\omega)$$

schreiben, aber die (an sich unsinnige) Bezeichnung $\int_\Omega f(\omega)\, d\mu(\omega)$ hat sich durchgesetzt. So wird beispielsweise im Fall $\Omega = \mathbb{R}^n$, $\mu = dF$ das Integral $\int_\Omega f(\omega)\, d\mu(\omega)$ dann zu $\int_{\mathbb{R}^n} f(x)\, ddF(x)$.

Definition 10.43. a) Für $f = \sum_{i=1}^m c_i\, \chi_{A_i} \in \mathscr{M}_0(\Omega, \mathbb{R})$, $f \ge 0$, setzen wir[13]

$$\int_\Omega f(\omega)\, d\mu(\omega) = \sum_{i=1}^n c_i\, \mu(A_i).$$

b) Für $f \in \mathscr{M}_+(\Omega, \overline{\mathbb{R}})$ setzen wir

$$\int_\Omega f(\omega)\, d\mu(\omega) = \sup\left\{ \int_\Omega g(\omega)\, d\mu(\omega) \,\middle|\, 0 \le g \le f,\ g \in \mathscr{M}_0(\Omega, \mathbb{R}) \right\}.$$

c) Für allgemeines $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ setzen wir

$$\int_\Omega f(\omega)\, d\mu(\omega) = \int_\Omega f^+(\omega)\, d\mu(\omega) - \int_\Omega f^-(\omega)\, d\mu(\omega),$$

falls die rechte Seite *nicht* von der Form $\infty - \infty$ ist.[14] In letzterem Fall bleibt das Integral undefiniert.

[13] Offenbar ist diese Definition unabhängig von der konkreten Darstellung von f. Man kann beispielsweise immer die A_i paarweise disjunkt und die c_i als paarweise verschieden annehmen.

[14] Das bedeutet, dass $\int_\Omega f^+\, d\mu < \infty$ oder $\int_\Omega f^-\, d\mu < \infty$ gilt.

Weiterhin setzen wir für $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ und $A \in \mathscr{A}$

$$\int_A f\, d\mu = \int_\Omega \chi_A f\, d\mu,$$

falls die rechte Seite existiert. Existiert $\int_A f\, d\mu$ und ist endlich, so heißt f *über A integrierbar*, im Fall $A = \Omega$ einfach nur integrierbar.

Für diesen Integralbegriff haben wir die folgenden elementaren Eigenschaften:

Satz 10.44. a) *Existiert $\int_\Omega f\, d\mu$ und ist $c \in \mathbb{R}$, so existiert auch $\int_\Omega cf\, d\mu$ und hat den Wert $c\int_\Omega f\, d\mu$.*

b) *Gilt $f \geq g$, so ist*

$$\int_\Omega f\, d\mu \geq \int_\Omega g\, d\mu,$$

falls beide Integrale existieren. Insbesondere impliziert $\int_\Omega g\, d\mu > -\infty$ die Existenz von $\int_\Omega f\, d\mu$, während $\int_\Omega f\, d\mu < \infty$ die Existenz von $\int_\Omega g\, d\mu$ nach sich zieht.

c) *Existiert $\int_\Omega f\, d\mu$, so gilt*

$$\left|\int_\Omega f\, d\mu\right| \leq \int_\Omega |f|\, d\mu.$$

d) *Ist $f \geq 0$, $A \in \mathscr{A}$, so ist*

$$\int_A f\, d\mu = \sup\left\{\int_A g\, d\mu \;\middle|\; 0 \leq g \leq f,\, g \in \mathscr{M}_0(\Omega, \mathbb{R})\right\}.$$

e) *Existiert $\int_\Omega f\, d\mu$, so existiert auch $\int_A f\, d\mu$ für jedes $A \in \mathscr{A}$. Ist $\int_\Omega f\, d\mu$ endlich, so ist auch $\int_A f\, d\mu$ endlich für jedes $A \in \mathscr{A}$.*

Beweis. a) • Ist $f \geq 0$ eine Stufenfunktion, so folgt die Behauptung aus der Definition.

- Für $f \in \mathscr{M}$, $f \geq 0$, $c > 0$ gilt

$$\begin{aligned}\int_\Omega cf\, d\mu &= \sup\left\{\int_\Omega g\, d\mu \;\middle|\; 0 \leq g \leq cf,\, g \in \mathscr{M}_0\right\}\\ &= c \sup\left\{\int_\Omega \frac{g}{c}\, d\mu \;\middle|\; 0 \leq \frac{g}{c} \leq f,\, \frac{g}{c} \in \mathscr{M}_0\right\} = c\int_\Omega f\, d\mu\end{aligned}$$

- Für $f \in \mathscr{M}$, $f = f^+ - f^-$, $c > 0$ gilt $(cf)^+ = cf^+$, $(cf)^- = cf^-$, also

$$\int_\Omega cf\, d\mu = c\int_\Omega f^+\, d\mu - c\int_\Omega f^-\, d\mu = c\int_\Omega f\, d\mu.$$

- Für $f \in \mathscr{M}$, $f = f^+ - f^-$, $c < 0$ gilt $(cf)^+ = -cf^-$, $(cf)^- = -cf^+$, also

$$\int_\Omega cf\, d\mu = -c\int_\Omega f^-\, d\mu + c\int_\Omega f^+\, d\mu = c\int_\Omega f\, d\mu.$$

b) • Für $f \geq 0$, $g \geq 0$, $h \in \mathscr{M}_0$, $0 \leq h \leq g$ gilt $0 \leq h \leq f$, also $\int_\Omega g\, d\mu \leq \int_\Omega f\, d\mu$.

• Für allgemeines $f, g \in \mathscr{M}$ mit $f \geq g$ gilt $f^+ \geq g^+$, $f^- \leq g^-$. Demnach impliziert $\int_\Omega g\, d\mu > -\infty$, dass

$$\int_\Omega f^-\, d\mu \leq \int_\Omega g^-\, d\mu < \infty,$$

also existiert $\int_\Omega f\, d\mu$ und

$$\begin{aligned}\int_\Omega f\, d\mu &= \int_\Omega f^+\, d\mu - \int_\Omega f^-\, d\mu \\ &\geq \int_\Omega g^+\, d\mu - \int_\Omega g^-\, d\mu = \int_\Omega g\, d\mu.\end{aligned}$$

• Der Fall $\int_\Omega f\, d\mu < \infty$ ist analog.

c) Existiert $\int_\Omega f\, d\mu$, so folgt aus $-|f| \leq f \leq |f|$, dass

$$-\int_\Omega |f|\, d\mu \leq \int_\Omega f\, d\mu \leq \int_\Omega |f|\, d\mu,$$

und damit die Behauptung.

d) Ist $g \in \mathscr{M}_0$, $0 \leq g \leq f$, so ist $\int_A g\, d\mu \leq \int_A f\, d\mu$, also

$$\int_A f\, d\mu \geq \sup\left\{\int_A g\, d\mu \,\middle|\, 0 \leq g \leq f,\, g \in \mathscr{M}_0\right\}.$$

Sei umgekehrt $h \in \mathscr{M}_0$, $0 \leq h \leq \chi_A f$. Dann ist $h = \chi_A h \leq f$ und damit

$$\int_\Omega h\, d\mu \leq \sup\left\{\int_\Omega \chi_A g\, d\mu \,\middle|\, 0 \leq g \leq f,\, g \in \mathscr{M}_0\right\}.$$

Nimmt man hier das Supremum über alle derartigen h, so ergibt sich

$$\int_A f\, d\mu \leq \sup\left\{\int_A g\, d\mu \,\middle|\, 0 \leq g \leq f,\, g \in \mathscr{M}_0\right\}.$$

e) Dies folgt aus b) und $(\chi_A f)^+ = \chi_A f^+ \leq f^+$, $(\chi_A f)^- = \chi_A f^- \leq f^-$. □

10.4.3 Grundlegende Sätze

Im Folgenden sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum.

Satz 10.45. *Sei $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $\int_\Omega f\, d\mu$ möge existieren. Sei*

$$\nu(A) = \int_A f\, d\mu, \quad A \in \mathscr{A}.$$

Dann gilt:

(i) *ν ist σ-additiv auf $\mathscr{A}$.*

(ii) *Ist $f \geq 0$, so ist ν ein Maß.*

Beweis. a) Sei $f = \sum_{i=1}^m c_i \, \chi_{A_i} \in \mathscr{M}_0(\Omega, \mathbb{R})$, $f \geq 0$. Dann ist

$$\nu(A) = \int_A f \, d\mu = \sum_{i=1}^n c_i \, \mu(A \cap A_i),$$

also ist ν ein Maß auf $\mathscr{A}$.

b) Sei $f \in \mathscr{M}_+(\Omega, \overline{\mathbb{R}})$. Wir zeigen, dass ν auch in diesem Fall σ-additiv, also ein Maß auf $\mathscr{A}$ ist.
Sei $A = \bigcup_{i=1}^\infty A_i$, $A_i \in \mathscr{A}$ paarweise disjunkt. Dann gilt für $g \in \mathscr{M}_0$, $0 \leq g \leq f$, dass

$$\int_A g \, d\mu = \sum_{i=1}^\infty \int_{A_i} g \, d\mu \leq \sum_{i=1}^\infty \int_{A_i} f \, d\mu.$$

Indem wir das Supremum über alle derartigen g nehmen, erhalten wir

$$\nu\,(A) \leq \sum_{i=1}^\infty \nu\,(A_i)\,.$$

Um die umgekehrte Abschätzung einzusehen, stellen wir fest, dass $A_i \subseteq A$, $\chi_{A_i} \leq \chi_A$ und $\nu(A_i) \leq \nu(A)$ gilt, also im Fall $\nu(A_i) = \infty$ für ein i nichts zu zeigen ist.
Sei daher $\nu(A_i) < \infty$ für alle i. Wir fixieren ein i und wählen $\epsilon > 0$. Dann existiert ein $g \in \mathscr{M}_0$, $0 \leq g \leq f$, mit

$$\int_{A_j} g \, d\mu \geq \int_{A_j} f \, d\mu - \frac{\epsilon}{i}, \quad j = 1, 2, \ldots, i.$$

Es folgt dann, dass

$$\nu(A) \geq \nu\left(\bigcup_{j=1}^i A_j\right) = \int_{\bigcup_{j=1}^i A_i} f \, d\mu \geq \sum_{j=1}^i \int_{A_j} g \, d\mu \geq \sum_{j=1}^i \int_{A_j} f \, d\mu - \epsilon,$$

also für $i \to \infty$

$$\nu(A) \geq \sum_{i=1}^\infty \nu(A_i) - \epsilon.$$

Da $\epsilon > 0$ beliebig war, folgt die Behauptung auch in diesem Fall.

c) Ist $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, so dass $\int_\Omega f \, d\mu$ existiert, dann ist $f = f^+ - f^-$ mit $\int_\Omega f^+ \, d\mu < \infty$ oder $\int_\Omega f^- \, d\mu < \infty$. Es genügt dann festzustellen, dass $\nu = \nu^+ - \nu^-$ mit Maßen ν^+, ν^-,

$$\nu^+(A) = \int_A f^+ \, d\mu, \quad \nu^-(A) = \int_A f^- \, d\mu, \quad \forall A \in \mathscr{A},$$

von denen wenigstens eines endlich ist. □

Satz 10.46 (über die monotone Konvergenz)**.** *Sei $\{f_n\} \subset \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $0 \leq f_n \uparrow f$. Dann gilt*

$$\int_\Omega f_n \, d\mu \to \int_\Omega f \, d\mu.$$

Beweis. Es gilt

$$\int_\Omega f_1\,d\mu \le \int_\Omega f_2\,d\mu \le \dots,$$

so dass $k = \lim_{n\to\infty} \int_\Omega f_n\,d\mu$ existiert. Außerdem $k \le \int_\Omega f\,d\mu$.
Sei $0 < b < 1$, $g \in \mathscr{M}_0$, $0 \le g \le f$. Wir setzen

$$A_n = \{f_n \ge bg\}.$$

Dann

$$k \ge \int_\Omega f_n\,d\mu \ge \int_{A_n} f_n\,d\mu \ge b\int_{A_n} g\,d\mu.$$

Wegen $A_n \uparrow \Omega$ ergibt sich für $n \to \infty$, dass

$$k \ge b\int_\Omega g\,d\mu$$

oder, da $0 < b < 1$ beliebig,

$$k \ge \int_\Omega g\,d\mu.$$

Nimmt man nun das Supremum über alle derartigen g, so erhält man schließlich

$$k \ge \int_\Omega f\,d\mu. \qquad \square$$

Satz 10.47 (Additivität des Integrals)**.** *Seien f, $g \in \mathscr{M}(\Omega,\overline{\mathbb{R}})$, $f+g$ sei wohldefiniert, $\int_\Omega f\,d\mu$, $\int_\Omega g\,d\mu$ mögen existieren und $\int_\Omega f\,d\mu+\int_\Omega g\,d\mu$ sei ebenfalls definiert (d. h. dieser Ausdruck ist nicht von der Form $\infty+(-\infty)$ oder $(-\infty)+\infty$). Dann existiert $\int_\Omega (f+g)\,d\mu$ und*

$$\int_\Omega (f+g)\,d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu.$$

Beweis. Sei $h = f + g$.

a) Im Fall, dass f, g nichtnegative Stufenfunktionen sind, ist die Behauptung klar.

b) Seien f, $g \in \mathscr{M}_+(\Omega,\overline{\mathbb{R}})$, $\{f_n\}$, $\{g_n\} \subset \mathscr{M}_0(\Omega,\mathbb{R})$, $0 \le f_n \uparrow f$, $0 \le g_n \uparrow g$. Dann $f_n + g_n \uparrow h$, also

$$\begin{aligned}\int_\Omega h\,d\mu &= \lim_{n\to\infty}\int_\Omega (f_n+g_n)\,d\mu = \lim_{n\to\infty}\int_\Omega f_n\,d\mu + \lim_{n\to\infty}\int_\Omega g_n\,d\mu\\ &= \int_\Omega f\,d\mu + \int_\Omega g\,d\mu.\end{aligned}$$

c) Seien nun f, $g \in \mathscr{M}(\Omega,\overline{\mathbb{R}})$, $f \ge 0$, $g \le 0$ und $h \ge 0$ (insbesondere ist g endlich). Es ergibt sich $f = h + (-g)$, also

$$\int_\Omega f\,d\mu = \int_\Omega h\,d\mu - \int_\Omega g\,d\mu$$

und da $\int_\Omega g\,d\mu > -\infty$ gelten muss[15]

$$\int_\Omega h\,d\mu = \int_\Omega f\,d\mu + \int_\Omega g\,d\mu.$$

[15] $\int_\Omega g\,d\mu = -\infty$ impliziert den Widerspruch $\int_\Omega f\,d\mu \ge -\int_\Omega g\,d\mu = \infty$.

d) Der Fall $f,\, g \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $f \geq 0$, $g \leq 0$ und $h \leq 0$ ist analog zu c).

e) Im allgemeinen Fall $f,\, g \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ führen wir die sechs Mengen

$$\begin{aligned}
A_1 &= \{f \geq 0,\, g \geq 0\}\,, \\
A_2 &= \{f \geq 0,\, g < 0,\, h \geq 0\}\,, \\
A_3 &= \{f \geq 0,\, g < 0,\, h < 0\}\,, \\
A_4 &= \{f < 0,\, g \geq 0,\, h \geq 0\}\,, \\
A_5 &= \{f < 0,\, g \geq 0,\, h < 0\}\,, \\
A_6 &= \{f < 0,\, g < 0\}
\end{aligned}$$

ein. Obige Argumente liefern

$$\int_{A_i} h\, d\mu = \int_{A_i} f\, d\mu + \int_{A_i} g\, d\mu, \quad \forall i = 1, 2, \ldots, 6.$$

Folglich gilt

$$\int_\Omega f\, d\mu + \int_\Omega g\, d\mu = \sum_{i=1}^{6} \left(\int_{A_i} f\, d\mu + \int_{A_i} g\, d\mu \right) = \sum_{i=1}^{6} \int_{A_i} h\, d\mu,$$

so dass zu zeigen bleibt, dass das Integral $\int_\Omega h\, d\mu$ existiert.
Angenommen, es gilt $\int_\Omega h^+\, d\mu = \int_\Omega h^-\, d\mu = \infty$. Dann gibt es[16] $i,\ j,\ i \neq j$ mit $\int_{A_i} h^+\, d\mu = \int_{A_j} h^-\, d\mu = \infty$, und damit $\int_{A_i} h\, d\mu = \infty$, $\int_{A_j} h\, d\mu = -\infty$. Es folgt $\int_{A_i} f\, d\mu = \infty$ oder $\int_{A_i} g\, d\mu = \infty$ bzw. $\int_\Omega f\, d\mu = \infty$ oder $\int_\Omega g\, d\mu = \infty$. Analog folgt $\int_{A_j} f\, d\mu = -\infty$ oder $\int_{A_j} g\, d\mu = -\infty$, also $\int_\Omega f\, d\mu = -\infty$ oder $\int_\Omega g\, d\mu = -\infty$. Somit erhalten wir den Widerspruch

$$\left\{ \int_\Omega f\, d\mu, \int_\Omega g\, d\mu \right\} = \{\infty, -\infty\}\,. \qquad \square$$

Folgerung 10.48. a) *Ist $\{f_n\} \subset \mathscr{M}_+$, so ist $\int_\Omega \left(\sum_n f_n\right) d\mu = \sum_n \int_\Omega f_n\, d\mu$.*

b) *Für $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ ist f genau dann integrierbar, wenn $|f|$ integrierbar ist.*

c) *Sei $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, g integrierbar, $|f| \leq g$, so ist auch f integrierbar.*

Beweis. a) Benutze $\sum_{k=1}^n f_k \uparrow \sum_{k=1}^\infty f_k$.

b) Benutze $|f| = f^+ + f^-$.

c) Folgt unmittelbar aus b). $\square$

Definition 10.49. Eine Eigenschaft, die für Stellen $\omega \in \Omega$ gelten kann oder nicht, gilt *μ-fast überall* (in Formeln schreibt man μ-f. ü.), falls eine μ-Nullmenge N existiert, so dass die Eigenschaft in $\Omega \setminus N$ gilt.

[16] Tatsächlich ist $i \in \{1, 2, 4\}$, $j \in \{3, 5, 6\}$.

Satz 10.50. *Seien f, $g \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$.*

a) *Ist $f = 0$ μ-f. ü., so ist $\int_\Omega f\, d\mu = 0$.*

b) *Ist $f = g$ μ-f. ü. und existiert das Integral $\int_\Omega f\, d\mu$, so existiert auch das Integral $\int_\Omega g\, d\mu$ und beide sind gleich.*

Beweis. a) Ist $f = \sum_{i=1}^m c_i\, \chi_{A_i}$ eine nichtnegative Stufenfunktion mit paarweise disjunkten A_i, so folgt aus $c_i > 0$ für ein i, dass $\mu(A_i) = 0$ für dieses i gilt. Insgesamt ergibt sich $\int_\Omega f\, d\mu = 0$.
Sei nun $f \in \mathscr{M}_+\mathscr{M}(\Omega, \overline{\mathbb{R}})$, $g \in \mathscr{M}_0\mathscr{M}(\Omega, \mathbb{R})$, $0 \leq g \leq f$. Dann ist $\int_\Omega g\, d\mu = 0$, also auch $\int_\Omega f\, d\mu = 0$.
Schließlich schreiben wir im allgemeinen Fall $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ $f = f^+ - f^-$ und erhalten aus $f^+ \leq |f|$, $f^- \leq |f|$, dass $f^+ = 0$ μ-f. ü., $f^- = 0$ μ-f. ü. und daher $\int_\Omega f\, d\mu = 0$.
b) Sei $A = \{f = g\}$. Dann hat $A^c = \{f \neq g\}$ die Eigenschaft $\mu(A^c) = 0$. Es folgt

$$f = f\,\chi_A + f\,\chi_{A^c}, \quad g = f\,\chi_A + g\,\chi_{A^c}.$$

Weiterhin gilt $\chi_{A^c} f = 0$ μ-f. ü., $\chi_{A^c} g = 0$ μ-f. ü., also mit a)

$$\int_\Omega f\, d\mu = \int_A f\, d\mu = \int_\Omega g\, d\mu. \qquad \square$$

Satz 10.51. *Sei $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$.*

a) *Ist f integrierbar, so ist f endlich μ-f. ü.*

b) *Ist $f \geq 0$ und $\int_\Omega f\, d\mu = 0$, so ist $f = 0$ μ-f. ü.*

Beweis. a) Sei $A = \{|f| = \infty\}$. Falls $\mu(A) > 0$, so ist $\int_\Omega |f|\, d\mu \geq \int_A |f|\, d\mu = \infty$; ein Widerspruch.

b) Sei $A = \{f > 0\}$, $A_n = \left\{f \geq \frac{1}{n}\right\}$ für $n \in \mathbb{N}$. Dann $A_n \uparrow A$. Weiterhin $0 \leq \chi_{A_n} f \leq \chi_A f = f$, also $0 = \int_{A_n} f\, d\mu \geq \frac{\mu(A_n)}{n}$. Damit ist $\mu\left(A_n\right) = 0$ für alle n und folglich $\mu\left(A\right) = 0$. $\square$

Satz 10.52 (über die monotone Konvergenz; erweiterte Version)**.** *Seien $\{f_n\} \subset \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $g \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$.*

a) *Gilt $f_n \geq g$ $\forall n$ und $\int_\Omega g\, d\mu > -\infty$, so impliziert $f_n \uparrow f$, dass*

$$\int_\Omega f_n\, d\mu \to \int_\Omega f\, d\mu.$$

b) *Gilt $f_n \leq g$ $\forall n$ und $\int_\Omega g\, d\mu < \infty$, so impliziert $f_n \downarrow f$, dass*

$$\int_\Omega f_n\, d\mu \to \int_\Omega f\, d\mu.$$

Beweis. Es genügt, a) zu beweisen.
Gilt $\int_\Omega g\, d\mu = \infty$, so ist $\int_\Omega f_n\, d\mu = \infty$ $\forall n$ und es ist nichts zu zeigen.
Sei also $\int_\Omega g\, d\mu < \infty$. Dann ist g endlich μ-f. ü. An Stellen, an denen $|g|$ unendlich ist, setzen wir g zu Null. Dann $0 \leq f_n - g \uparrow f - g$ μ-f. ü., also $\int_\Omega (f_n - g)\ d\mu \to \int_\Omega (f - g)\ d\mu$. Es verbleibt, $\int_\Omega g\, d\mu$ auf beiden Seiten zu addieren. $\square$

Eine wichtige Verallgemeinerung des vorigen Satzes ergibt sich wie folgt:

Satz 10.53 (Lemma von Fatou)**.** *Seien* $\{f_n\} \subset \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$.

a) *Gilt* $f_n \geq f$ $\forall n$ *und* $\int_\Omega f\, d\mu > -\infty$, *so ist*

$$\liminf \int_\Omega f_n\, d\mu \geq \int_\Omega \liminf f_n\, d\mu.$$

b) *Gilt* $f_n \leq f$ $\forall n$ *und* $\int_\Omega f\, d\mu < \infty$, *so ist*

$$\limsup \int_\Omega f_n\, d\mu \leq \int_\Omega \limsup f_n\, d\mu.$$

Beweis. Es genügt wiederum, a) zu zeigen.
Sei $g_n = \inf_{k \geq n} f_k$, $g = \liminf f_n$. Dann ist $f_n \geq g_n \geq f$ $\forall n$ und $g_n \uparrow g$, also

$$\int_\Omega g\, d\mu = \lim_{n\to\infty} \int_\Omega g_n\, d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n\, d\mu.$$ □

Der nächste Satz ist eines der wichtigsten Resultate der Analysis:

Satz 10.54 (Lebesguescher Satz über die majorisierte Konvergenz)**.** *Sei* $\{f_n\} \subset \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, g μ*-integrierbar,* $|f_n| \leq g$ μ-f. ü. *für alle* n. *Falls* $f_n \to f$ μ-f. ü., *so*

$$\int_\Omega f_n\, d\mu \to \int_\Omega f\, d\mu.$$

Beweis. Wegen $|f_n| \leq g$, $f \leq g$ μ-f. ü. sind f_n, f μ-integrierbar. Weiter gilt mit dem Lemma von Fatou, dass

$$\begin{aligned}\int_\Omega f\, d\mu &= \int_\Omega \lim f_n\, d\mu \leq \liminf \int_\Omega f_n\, d\mu \leq \limsup \int_\Omega f_n\, d\mu \\ &\leq \int_\Omega \lim f_n\, d\mu = \int_\Omega f\, d\mu\end{aligned}$$

und daher $\lim \int_\Omega f_n\, d\mu = \int_\Omega f\, d\mu$. □

Wir wissen bereits: Wenn $\int_\Omega f\, d\mu$, $\int_\Omega g\, d\mu$ existieren, so impliziert $f \leq g$ μ-f. ü. die Beziehung $\int_A f\, d\mu \leq \int_A g\, d\mu$ für alle $A \in \mathscr{A}$. Im σ-endlichen Fall gilt auch die Umkehrung:

Satz 10.55. *Sei* μ σ*-endlich,* $f, g \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$, $\int_\Omega f\, d\mu$, $\int_\Omega g\, d\mu$ *mögen existieren. Gilt dann* $\int_A f\, d\mu \leq \int_A g\, d\mu$ *für alle* $A \in \mathscr{A}$, *so ist* $f \leq g$ μ-f. ü.

Beweis. Es genügt den Fall μ endlich zu betrachten.

a) Für $k \in \mathbb{N}$ sei

$$A_k = \left\{ f \geq g + \frac{1}{k},\ |g| \leq k \right\}.$$

Dann ist

$$\int_{A_k} g\, d\mu \geq \int_{A_k} f\, d\mu \geq \int_{A_k} g\, d\mu + \frac{\mu(A_k)}{k},$$

also wegen

$$\left|\int_{A_k} g\,d\mu\right| \le \int_{A_k} |g|\,d\mu \le k\,\mu(A_k) < \infty$$

$\mu(A_k)/k \le 0$ und $\mu(A_k) = 0$ für alle k. Damit ist

$$\mu\left(\{f > g,\, |g| < \infty\}\right) = \mu\left(\bigcup\nolimits_k A_k\right) \le \sum_k \mu\left(A_k\right) = 0$$

und $f \le g$ μ-f. ü. auf $\{|g| < \infty\}$.

b) $f \le g$ überall auf $\{g = \infty\}$ ist trivialerweise erfüllt.

c) Für $k \in \mathbb{N}$ sei

$$B_k = \{g = -\infty,\, f \ge -k\}\,.$$

Dann ist

$$-\infty\,\mu(B_k) = \int_{B_k} g\,d\mu \ge \int_{B_k} f\,d\mu \ge -k\,\mu(B_k),$$

also $\mu\left(B_k\right) = 0$ für alle k und

$$\mu\left(\{f > g,\, g = -\infty\}\right) = \mu\left(\bigcup\nolimits_k B_k\right) = 0.$$

Folglich ist $f \le g$ μ-f. ü. auch auf $\{g = -\infty\}$. □

Die Transformationsformel

Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $(\Omega', \mathscr{B})$ ein messbarer Raum, $F\colon \Omega \to \Omega'$ $(\mathscr{A}, \mathscr{B})$-messbar. Wir definieren das *Bildmaß* $F_*\mu$ von μ unter F durch

$$\left(F_*\mu\right)(B) = \mu\left(F^{-1}(B)\right), \quad \forall B \in \mathscr{B}.$$

Offenbar ist $F_*\mu$ ein Maß auf $\mathscr{B}$.

Satz 10.56. *Für alle $g \in \mathscr{M}(\Omega', \overline{\mathbb{R}})$, $B \in \Omega'$ gilt*

$$\int_B g\,d(F_*\mu) = \int_{F^{-1}(B)} (F^*g)\,d\mu,$$

wobei diese Behauptung so zu lesen ist, dass wenn eines der beiden Integrale existiert, so auch das andere und beide Integrale gleich.

Beweis. Wegen $F^*\chi_B = \chi_{F^{-1}(B)}$ dürfen wir $B = \Omega'$ annehmen.

a) Sei zuerst $g = \sum_{i=1}^m c_i\,\chi_{B_i}$ eine Stufenfunktion, $g \ge 0$, d. h. $c_i \ge 0$, $B_i \in \mathscr{B}$. Dann ist

$$\begin{aligned}\int_{\Omega'} g\,d(F_*\mu) &= \sum_{i=1}^n c_i\,(F_*\mu)(B_i) = \sum_{i=1}^n c_i\,\mu\left(F^{-1}(B_i)\right)\\ &= \int_\Omega \sum_{i=1}^n c_i\,\chi_{F^{-1}(B_i)}\,d\mu = \int_\Omega (F^*g)\,d\mu.\end{aligned}$$

b) Sei nun $g \in \mathscr{M}_+(\Omega', \overline{\mathbb{R}})$. Wir wählen eine Folge $\{g_i\} \subset \mathscr{M}_0(\Omega', \mathbb{R})$ mit $0 \le g_i \uparrow g$. Dann sind die $F^* g_i$ Stufenfunktionen mit $0 \le F^* g_i \uparrow F^* g$. Gemäß a) gilt

$$\int_{\Omega'} g \, d(F_* \mu) = \lim_{n\to\infty} \int_{\Omega'} g_i \, d(F_* \mu) = \lim_{n\to\infty} \int_{\Omega} (F^* g_i) \, d\mu = \int_{\Omega} (F^* g_i) \, d\mu.$$

c) Der verbleibende Fall $g \in \mathscr{M}(\Omega', \overline{\mathbb{R}})$ mit $g = g^+ - g^-$, g^+, $g^- \in \mathscr{M}_+(\Omega', \overline{\mathbb{R}})$ folgt nun ebenfalls. □

10.4.4 Vergleich zwischen dem Lebesgue- und dem Riemann-Integral

Der Beweis des folgenden Satzes befindet sich in einem Anhang zu diesem Abschnitt.

Satz 10.57. *Sei $f\colon [a,b] \to \mathbb{R}$, $-\infty < a < b < \infty$. Dann gilt*

(i) *f ist genau dann Riemann-integrierbar, wenn f beschränkt und* λ_1-f. ü. *stetig ist.*

(ii) *Ist f Riemann-integrierbar, so ist f auch Lebesgue-integrierbar und beide Integrale ergeben den gleichen Wert.*

Beispiel 10.58. Seien f, $g\colon [0,1] \to \mathbb{R}$ wie folgt gegeben:

$$f(x) = \begin{cases} 1/q, & \text{falls } x = p/q,\ p,\ q \in \mathbb{Z} \text{ prim},\ q > 0, \\ 0, & \text{falls } x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

$$g(x) = \begin{cases} 1, & \text{falls } x \in [0,1] \cap \mathbb{Q}, \\ 0, & \text{falls } x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Dann ist f Riemann-integrierbar[17], g jedoch nicht[18]. Andererseits gilt $f = g = 0$ λ_1-f. ü., also sind beide Funktionen integrierbar im Lebesgueschen Sinne und

$$\int_0^1 f(x) \, dx = \int_0^1 g(x) \, dx = 0.$$

Bemerkung. Das Ergebnis dieses Satzes gilt auch in höheren Raumdimensionen, nur dass wir dort den Begriff des Riemann-Integrals nicht eingeführt haben.

Beispiel 10.59. Eine Folge $\{g_i\}$ Riemann-integrierbarer Funktionen auf $[0,1]$ mit $0 \le g_i \le 1$ und $g_i \uparrow g$, für die g nicht Riemann-integrierbar ist.

Sei $\{q_k\}$ eine Aufzählung von $[0,1] \cap \mathbb{Q}$. Wir setzen

$$g_i(x) = \begin{cases} 1, & \text{falls } x \in \{q_1, \dots, q_i\}, \\ 0, & \text{andernfalls.} \end{cases}$$

Dann leistet die Folge $\{g_i\}$ das Verlangte.

Im Fall des Riemann-Integrals hat man also, um beispielsweise die Grenzwertsätze der Integrationstheorie anwenden zu können, stets die Riemann-Integrierbarkeit der Grenzfunktion voraussetzen.

[17] f ist an allen Stellen $x \in [0,1] \setminus \mathbb{Q}$ stetig.
[18] g ist überall unstetig.

Anhang zu Abschnitt 10.4.4: Beweisskizze von Satz 10.57

10.5 Weitere Resultate der Maß- und Integrationstheorie

Wir sammeln hier weitere wichtige Sätze der allgemeinen Maß- und Integrationstheorie.

10.5.1 Der Zerlegungssatz von Hahn und Jordan

Erinnerung: Sei $(\Omega, \mathscr{A})$ ein messbarer Raum. Eine Mengenfunktion $\nu\colon \mathscr{A} \to [-\infty, \infty]$ mit $\nu(\emptyset) = 0$ heißt σ-additiv, falls sie einen der beiden Werte ∞, $-\infty$ nicht annimmt und außerdem

$$\nu\left(\bigcup_i A_i\right) = \sum_i \nu(A_i)$$

für $\{A_i\}_{i\in\mathbb{N}} \subset \mathscr{A}$, A_i paarweise disjunkt gilt.

Bemerkung. Ist $\nu\left(\bigcup_i A_i\right)$ endlich, so ist die Reihe $\sum_i \nu(A_i)$ absolut konvergent[19].

Definition 10.60. Eine derartige Mengenfunktion $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}$ heißt auch ein *signierte Maß.*

Satz 10.61. *Für ν wie oben existieren D^+, $D^- \in \mathscr{A}$ mit*

$$\nu(D^+) = \sup\left\{\nu(A) \;\middle|\; A \in \mathscr{A}\right\},$$
$$\nu(D^-) = \inf\left\{\nu(A) \;\middle|\; A \in \mathscr{A}\right\}.$$

Beispiel 10.62. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und das Integral $\int_\Omega f\, d\mu$ möge existieren. Dann ist ν mit

$$\nu(A) = \int_\Omega f\, d\mu, \quad A \in \mathscr{A},$$

ein signiertes Maß. Für D^+, D^- können wir $D^+ = \{f \geq 0\}$, $D^- = \{f < 0\}$ wählen.

Beweis des Satzes. Wir führen den Beweis für das Supremum.

Wir dürfen $\nu < \infty$ annehmen, denn andernfalls wählen wir ein $D^+ \in \mathscr{A}$ mit $\nu(D^+) = \infty$ und sind fertig.
Wir wählen eine Folge $\{A_i\} \subset \mathscr{A}$ mit $\nu(A_i) \to \sup \nu$ und setzen $A = \bigcup_i A_i \in \mathscr{A}$. Für ein gegebenes $i \in \mathbb{N}$ zerlegen wir A in 2^i disjunkte Teilmengen A_{im}, $1 \leq m \leq 2^i$, der Form

$$A_1^* \cap A_2^* \cap \ldots \cap A_i^*,$$

[19] Da diese Reihe dann unbedingt konvergent ist.

wobei entweder $A_j^* = A_j$ oder $A_j^* = A \setminus A_j$. Wir setzen weiterhin

$$B_i = \bigcup_{m\colon \nu(A_{im})\geq 0} A_{im},$$

insbesondere $B_i = \emptyset$, falls $\nu(A_{im}) < 0$ für alle m. Da A_i endliche Vereinigung einiger der A_{im} ist, gilt

$$\nu(A_i) \leq \nu(B_i).$$

Zudem ist für $i' > i$ jedes $A_{i'm'}$ entweder eine Teilmenge von A_{im} oder disjunkt davon. Damit haben wir für jedes $r \geq i$, dass $\bigcup_{k=i}^r B_k$ die Vereinigung von B_i und einer Menge E disjunkt von B_i mit $\nu(E) \geq 0$ ist. Somit gilt

$$\nu(A_i) \leq \nu(B_i) \leq \nu\left(\bigcup\nolimits_{k=i}^r B_k\right) \to \nu\left(\bigcup\nolimits_{k=i}^\infty B_k\right) \text{ für } r \to \infty.$$

Sei $D^+ = \limsup_i B_i = \bigcap_{i=1}^\infty \bigcup_{k=i}^\infty B_k$.
Dann gilt $\bigcup_{k=i}^\infty B_k \downarrow D^+$, $0 \leq \nu\left(\bigcup_{k=i}^\infty B_k\right) < \infty$ für alle i, also

$$\nu\left(\bigcup_{k=i}^\infty B_k\right) \to \nu(D^+) \text{ für } i \to \infty.$$

Es folgt

$$\sup \nu = \lim_{i\to\infty} \nu(A_i) \leq \lim_{i\to\infty} \nu\left(\bigcup\nolimits_{k=i}^\infty B_k\right) = \nu(D^+) \leq \sup \nu,$$

also $\nu(D^+) = \sup \nu$. □

Satz 10.63 (Hahn-Jordan-Zerlegung)**.** *Sei ν wie oben. Wir definieren*

$$\nu^+(A) = \sup\left\{\nu(B) \mid B \in \mathscr{A},\, B \subseteq A\right\},$$
$$\nu^-(A) = \inf\left\{\nu(B) \mid B \in \mathscr{A},\, B \subseteq A\right\}$$

für jedes $A \in \mathscr{A}$. Dann sind ν^+, ν^- Maße auf $\mathscr{A}$, wenigstens eines ist endlich und es gilt

$$\nu = \nu^+ - \nu^-.$$

Beweis. O. B. d. A. nehme ν den Wert $-\infty$ nicht an. Die Menge $D^- \in \mathscr{A}$ sei wie oben konstruiert.

a) Wir behaupten, dass $\nu(A \cap D^-) \leq 0$, $\nu(A \cap (D^-)^c) \geq 0$ für alle $A \in \mathscr{A}$.

(i) Angenommen, es wäre $\nu(A \cap D^-) > 0$ für ein $A \in \mathscr{A}$. Dann ist $\nu(D^-) = \nu(A \cap D^-) + \nu(A^c \cap D^-)$, also

$$\nu(A^c \cap D^-) = \nu(D^-) - \nu(A \cap D^-) < \nu\left(D^-\right);$$

ein Widerspruch.

(ii) Angenommen, es wäre $\nu(A \cap (D^-)^c) < 0$. Dann ist $\nu(D^- \cup (A \cap (D^-)^c)) = \nu(D^-) + \nu(A \cap (D^-)^c) < \nu(D^-)$; ein Widerspruch.

b) Wir behaupten weiterhin, dass

$$\begin{aligned}\nu^+(A) &= \nu(A \cap (D^-)^c),\\ \nu^-(A) &= -\nu(A \cap D^-)\end{aligned}$$

für $A \in \mathscr{A}$.

(i) Sei $B \in \mathscr{A}$, $B \subseteq A$. Dann ist

$$\begin{aligned}\nu(B) &= \nu(B \cap D^-) + \nu(B \cap (D^-)^c) \le \nu(B \cap (D^-)^c)\\ &\le \nu(B \cap (D^-)^c) + \nu((A \setminus B) \cap (D^-)^c) = \nu(A \cap (D^-)^c),\end{aligned}$$

also $\nu^+(A) \le \nu(A \cap (D^-)^c)$. Wegen $A \cap (D^-)^c \subseteq A$ gilt $\nu(A \cap (D^-)^c) \le \nu^+(A)$ gemäß Definition.

(ii) Sei $B \in \mathscr{A}$, $B \subseteq A$. Dann ist

$$\begin{aligned}\nu(B) &= \nu(B \cap D^-) + \nu(B \cap (D^-)^c) \ge \nu(B \cap D^-)\\ &\ge \nu(B \cap D^-) + \nu((A \setminus B) \cap D^-) = \nu(A \cap D^-),\end{aligned}$$

also $-\nu^-(A) \ge \nu(A \cap D^-)$. Die umgekehrte Richtung ergibt sich wieder aus der Definition. □

Folgerung 10.64. *Sei ν wie oben. Dann gilt*:

a) *ν ist Differenz zweier Maße, von denen wenigstens eines endlich ist.*

b) *Ist ν endlich, so ist ν beschränkt.*

c) *Es existiert eine Menge $D \in \mathscr{A}$ mit $\nu(A \cap D) \le 0$ und $\nu(A \cap D^c) \ge 0$ für alle $A \in \mathscr{A}$.*

d) *Ist E eine andere Menge in $\mathscr{A}$ mit dieser Eigenschaft, so ist $\nu^+(A) = \nu(A \cap E^c)$, $\nu^-(A) = -\nu(A \cap E)$ für alle $A \in \mathscr{A}$ sowie*

$$(\nu^+ + \nu^-)(D \vartriangle E) = 0.$$

Definition 10.65. ν^+ heißt der *nichtnegative Teil* von ν, ν^- heißt der *nichtpositive Teil* von ν und $|\nu| = \nu^+ + \nu^-$ ist die *Totalvariation.*

Bemerkung. Aus $\nu = \nu^+ - \nu^-$ folgt

$$|\nu(A)| \le |\nu|(A), \quad \forall A \in \mathscr{A}.$$

10.5.2 Differenziation von Maßen. Der Satz von Radon-Nikodym

Wir benötigen im Weiteren das Zornsche Lemma. Hier ist eine Erinnerung an die Grundtatsachen:

Sei $(X, \le)$ eine *halbgeordnete Menge*, d. h. eine Menge X versehen mit einer Relation $\le$, die reflexiv, antisymmetrisch und transitiv ist. Dann:

- $(X, \le)$ heißt *total-geordnet* (und $\le$ eine Totalordnung auf X), falls $x \le y$ oder $y \le x$ für alle x, $y \in X$.

- Ein Element $z \in X$ heißt eine *obere Schranke* einer Teilmenge $Y \subseteq X$, falls $y \leq z$ für alle $y \in Y$. [20]

- Ein Element $z \in X$ heißt ein *maximales Element* von X, falls $z \leq x$ für $x \in X$ die Beziehung $z = x$ impliziert.[21]

Satz 10.66 (Zornsches Lemma)**.** *Ist $(X, \leq)$ eine nichtleere, halbgeordnete Menge, in der jede nichtleere, total-geordnete Teilmenge eine obere Schranke besitzt, so enthält X maximale Elemente.*

Bemerkung. Unter den üblichen Axiomen der Mengenlehre (Zermelo-Fraenkel-Axiome) ist das Zornsche Lemma mit dem Auswahlaxiom äquivalent.

Sei nun $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $f\colon \Omega \to \overline{\mathbb{R}}$ messbar, $\int_\Omega f\,d\mu$ möge existieren (d. h. $\int_\Omega f^+\,d\mu < \infty$ oder $\int_\Omega f^-\,d\mu < \infty$). Dann ist $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}$ mit

$$\nu(A) = \int_A f\,d\mu, \quad A \in \mathscr{A},$$

ein signiertes Maß mit der zusätzlichen Eigenschaft, dass $\mu(A) = 0$ die Beziehung $\nu(A) = 0$ zur Folge hat.

Definition 10.67. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}$ ein signiertes Maß. Wir sagen, dass ν *absolut-stetig* bezüglich μ ist (man schreibt $\nu \ll \mu$), falls $\mu(A) = 0$ für $A \in \mathscr{A}$ die Beziehung $\nu(A) = 0$ impliziert.

Wie wir gerade sahen, ist die Absolut-Stetigkeit eine notwendige Bedingung dafür, dass ein signiertes Maß ν eine Dichte f bezüglich eines Maßes μ wie oben angegeben besitzt. Der Satz von Radon-Nikodym stellt unter einer zusätzlichen Bedingung sicher, dass die Absolut-Stetigkeit für die Existenz einer solchen Dichte f auch hinreichend ist.

Satz 10.68 (Radon-Nikodym)**.** *Sei $(\Omega, \mathscr{A}, \mu)$ ein σ-endlicher Maßraum, $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}$ ein signiertes Maß, $\nu \ll \mu$. Dann existiert ein $f \in \mathscr{M}(\Omega, \overline{\mathbb{R}})$ mit*

$$\nu\,(A) = \int_A f\,d\mu, \quad A \in \mathscr{A}.$$

Ist g eine weitere derartige Funktion, so ist $f = g$ μ-f. ü.

Definition 10.69. $f = \frac{d\nu}{d\mu}$ heißt die *Radon-Nikodym-Ableitung* von ν bezüglich μ.

Beweis des Satzes. Die Eindeutigkeit von f μ-f. ü. ist klar, wir kümmern uns um die Existenz.

a) *Seien μ, ν endliche Maße.* Sei

$$\mathscr{G} = \left\{ g \in \mathscr{M}_+(\Omega, \overline{\mathbb{R}}) \;\Big|\; \int_A g\,d\mu \leq \nu\,(A)\,, \;\; \forall A \in \mathscr{A} \right\} \Big/ \sim,$$

wobei $\sim$ die Äquivalenzrelation $g \sim h :\iff g = h$ μ-f. ü. ist. Auf $\mathscr{G}$ definieren wir (vertreterweise) eine Halbordnung durch $g \leq h :\iff g \leq h$ μ-f. ü. Sei

[20] Die Stelle z muss nicht zu Y gehören.

[21] Eine halbgeordnete Menge kann mehr als ein maximales Element besitzen.

$\mathscr{H} \subseteq \mathscr{G}$ eine nichtleere Kette (d. h. eine nichtleere, total-geordnete Teilmenge). Wir setzen

$$h^* = \sup\left\{ \int_\Omega h\, d\mu \;\middle|\; h \in \mathscr{H} \right\}.$$

Dann existiert eine Folge $\{h_i\} \subset \mathscr{H}$ mit $\int_\Omega h_i\, d\mu \uparrow h^*$. Da $\mathscr{H}$ total-geordnet ist, gilt $h_i \le h_{i+1}$ für alle i. Folglich existiert ein $g \in \mathscr{G}$ mit $h_i \uparrow g$, $\int_\Omega g\, d\mu = h^*$. Wir behaupten, dass g eine obere Schranke von $\mathscr{H}$ ist. Ist nämlich $h \in \mathscr{H}$, so sind zwei Fälle möglich:

1. Fall. $h \le h_{i^*}$ für wenigstens ein i^*. In diesem Fall gilt $h \le g$ wegen $h_i \uparrow g$.

2. Fall. $h \ge h_i$ für alle i. Dann folgt $h \ge g$, wegen $h^* \ge \int_\Omega h\, d\mu$ also $\int_\Omega h\, d\mu = \int_\Omega g\, d\mu = h^*$ und damit $h = g$ μ-f. ü.

Somit ist g tatsächlich eine obere Schranke von $\mathscr{H}$. Eine Anwendung des Zornschen Lemmas liefert ein maximales Element $f \in \mathscr{G}$. Wir werden zeigen, dass

$$\nu(A) = \int_A f\, d\mu, \quad A \in \mathscr{A},$$

gilt. Sei $\nu_1 \colon \mathscr{A} \to [0, \infty)$ durch

$$\nu_1(A) = \nu(A) - \int_A f\, d\mu, \quad A \in \mathscr{A},$$

erklärt. ν_1 ist ein Maß auf $\mathscr{A}$, $\nu_1 \ll \mu$ und $\nu\,(\Omega) < \infty$. Wir wollen zeigen, dass ν_1 identisch Null ist. Angenommen, es wäre $\nu_1(\Omega) > 0$. Wir wählen dann ein $k > 0$ mit $\mu(\Omega) - k\,\nu_1(\Omega) < 0$ und betrachten das *signierte Maß* $\mu - k\,\nu_1$. Bezüglich dieses signierten Maßes existiert ein $D \in \mathscr{A}$ mit

$$\begin{aligned} \mu\,(A \cap D) - k\,\nu_1\,(A \cap D) &\le 0, \quad A \in \mathscr{A}, \\ \mu\,(A \cap D^c) - k\,\nu_1\,(A \cap D^c) &\ge 0, \quad A \in \mathscr{A}. \end{aligned}$$

Wir behaupten, dass $\mu\,(D) > 0$ gilt. Ist $\mu\,(D) = 0$, so ist wegen der Absolut-Stetigkeit auch $\nu_1\,(D) = 0$. Wählt man speziell $A = \Omega$ in der zweiten Ungleichung, so ergibt sich mit

$$0 \le \mu(D^c) - k\,\nu_1(D^c) = \mu(\Omega) - k\,\nu_1(\Omega) < 0$$

ein Widerspruch.
Wir setzen jetzt $h = (1/k)\,\chi_D$. Ist $A \in \mathscr{A}$, so ist

$$\int_A h\, d\mu = \frac{1}{k}\,\mu(A \cap D) \le \nu_1(A \cap D) \le \nu_1(A) = \nu(A) - \int_A f\, d\mu,$$

also $\int_A (f + h)\, d\mu \le \nu\,(A)$. Es folgt, dass $f + h \in \mathscr{G}$. Jedoch ist $f + h \ge f$ und $f + h > f$ auf D mit $\mu(D) > 0$ im Widerspruch dazu, dass $f \in \mathscr{G}$ ein maximales Element von $\mathscr{G}$ ist. Also gilt tatsächlich $\nu_1 = 0$ wie gefordert.

b) *Sei μ ein endliches Maß, ν ein σ-endliches Maß.* Wir schreiben $\Omega = \bigsqcup_i A_i$ mit $A_i \in \mathscr{A}$, $\nu(A_i) < \infty$. Sei $\nu_i(A) = \nu(A \cap A_i)$ für $A \in \mathscr{A}$, $i \in \mathbb{N}$. Mit a) existiert eine messbare Funktion $f_i \ge 0$ mit $\nu_i(A) = \int_A f_i\, d\mu$, $A \in \mathscr{A}$. Setzen wir $f = \sum f_i$, so folgt

$$\nu(A) = \int_A f\, d\mu.$$

c) *μ eine endliches Maß, ν ein beliebiges Maß.* In diesem Fall betrachten wir

$$\mathscr{C} = \left\{ C \in \mathscr{A} \;\middle|\; \nu \text{ ist } \sigma\text{-endlich auf } \mathscr{A} \cap C \right\}.$$

$\mathscr{C}$ ist nichtleer wegen $\emptyset \in \mathscr{C}$. Sei $c^* = \sup \left\{ \mu(C) \;\middle|\; C \in \mathscr{C} \right\}$. Wir wählen eine Folge $\{C_i\} \subset \mathscr{C}$ mit $\mu(C_i) \uparrow c^*$. Sei $C = \bigcup_{i=1}^{\infty} C_i$. Dann ist $C \in \mathscr{C}$ und $\mu(C) = c^*$. Wegen b) existiert eine Funktion $f \geq 0$, die bezüglich $\mathscr{A} \cap C$ messbar ist, mit

$$\nu(A \cap C) = \int_{A \cap C} f \, d\mu, \quad A \in \mathscr{A}.$$

Wir wollen diese Darstellung auf das Komplement C^c fortsetzen. Sei $A \in \mathscr{A}$ beliebig. Dann gibt es zwei Fälle:

1. Fall. Ist $\mu(A \cap C^c) > 0$, so ist $\nu(A \cap C^c) = \infty$, da andernfalls $\nu(A \cap C^c) < \infty$ nämlich $C \cup (A \cap C^c) \in \mathscr{C}$ und den Widerspruch

$$c^* \geq \mu \left(C \cup (A \cap C^c) \right) = \mu (C) + \mu (A \cap C^c) > \mu (C) = c^*$$

nach sich zieht.

2. Fall. Ist $\mu(A \cap C^c) = 0$, so ist $\nu(A \cap C^c) = 0$ wegen der Absolut-Stetigkeit.

In beiden Fällen erhalten wir

$$\nu (A \cap C^c) = \int_{A \cap C^c} \infty \, d\mu$$

und wir können f durch $f \chi_C + \infty \chi_{C^c}$ auf ganz Ω fortsetzen.

d) *μ ist σ-endlich, ν ein beliebiges Maß.* Wir schreiben $\Omega = \bigsqcup_i A_i$ mit $A_i \in \mathscr{A}$, $\mu(A_i) < \infty$. Gemäß c) existieren messbare Funktionen $f_i \geq 0$ mit $\nu(A \cap A_i) = \int_{A \cap A_i} f_i \, d\mu = \int_A f_i \, d\mu$, wobei wir o. B. d. A. $f_i = 0$ auf A_i^c angenommen haben. Dann gilt

$$\nu(A) = \int_A f \, d\mu$$

mit $f = \sum_i f_i$.

e) *μ ist ein σ-endliches Maß, ν ein beliebiges signiertes Maß.* Wir benutzen den Satz von Hahn-Jordan und schreiben $\nu = \nu^+ - \nu^-$ mit Maßen ν^+, ν^- und o. B. d. A. ν^- endlich. Nach d) existieren messbare Funktionen $f^+ \geq 0$, $f^- \geq 0$ mit f^- endlich, so dass

$$\nu^+(A) = \int_A f^+ \, d\mu, \quad \nu^-(A) = \int_A f^- \, d\mu, \quad A \in \mathscr{A}.$$

Es bleibt, $f = f^+ - f^-$ zu setzen. $\square$

Folgerung 10.70. *Unter den obigen Voraussetzungen gilt*:

a) ν *ist endlich* $\iff$ f *ist integrierbar.*

b) *Ist* $|\nu|$ σ*-endlich, so ist* f *endlich* μ-f. ü.

c) ν *ist ein Maß* $\iff$ $f \geq 0$ μ-f. ü.

Definition 10.71. a) Seien μ_1, μ_2 Maße auf $\mathscr{A}$. Wir sagen, dass μ_1, μ_2 *gegenseitig singulär* sind (man schreibt $\mu_1 \perp \mu_2$), falls es ein $D \in \mathscr{A}$ mit $\mu_1(D^c) = 0$, $\mu_2(D) = 0$ gibt.

b) Seien ν_1, ν_2 signierte Maße. Dann sind ν_1, ν_2 *gegenseitig singulär*, falls $|\nu_1| \perp |\nu_2|$.

Bemerkung. Ist ν ein signiertes Maß, so gilt $\nu^+ \perp \nu^-$.

Lemma 10.72 (Borel-Cantelli)**.** *Sei* $\{A_i\} \subset \mathscr{A}$ *eine Folge,* $\sum_i \mu(A_i) < \infty$. *Dann gilt* $\mu(\limsup A_i) = 0$.

Beweis. Es gilt $\bigcup_{k \geq i} A_k \downarrow \limsup A_i$ und damit

$$\mu(\limsup A_i) \leq \mu\left(\bigcup\nolimits_{k \geq i} A_k\right) \leq \sum_{k \geq i} \mu(A_k) \to 0 \quad \text{für } i \to \infty.$$

Daraus folgt die Behauptung. □

Lemma 10.73. *Seien* μ *ein Maß,* ν, ν_1, ν_2 *signierte Maße. Dann gilt*:

a) $\nu_1 \perp \mu$, $\nu_2 \perp \mu \Longrightarrow \nu_1 + \nu_2 \perp \mu$, *falls die Summe* $\nu_1 + \nu_2$ *definiert ist.*

b) $\nu \ll \mu \iff |\nu| \ll \mu$.

c) $\nu_1 \ll \mu$, $\nu_2 \perp \mu \Longrightarrow \nu_1 \perp \nu_2$.

d) $\nu \ll \mu$, $\nu \perp \mu \Longrightarrow \nu = 0$ (*d. h.* $\nu(A) = 0$ *für alle* $A \in \mathscr{A}$).

e) *Für* ν *endlich gilt* $\nu \ll \mu \iff \lim_{\mu(A) \to 0} \nu(A) = 0$.

Beweis. Wir zeigen e), der Rest ist offensichtlich.

($\Longleftarrow$) Ist $\mu(A) = 0$, so folgt mit $A_i = A$ für alle i, dass $\nu(A_i) \to 0$, also $\nu(A) = 0$.

($\Longrightarrow$) Sei $\nu \ll \mu$. Wir nehmen an, es gilt $\limsup_{\mu(A) \to 0} |\nu|(A) > 0$. Dann existiert ein $\epsilon > 0$ und eine Folge $\{A_i\} \subset \mathscr{A}$ mit $\mu(A_i) \leq 2^{-i}$ und $|\nu|(A_i) \geq \epsilon$ für alle i. Sei $A = \limsup A_i$. Mit dem Lemma von Borel-Cantelli folgt $\mu(A) = 0$. Jedoch steht $|\nu|(A) \geq \epsilon$ im Widerspruch zur Absolut-Stetigkeit von $|\nu|$ bezüglich μ. □

Satz 10.74 (Lebesguescher Zerlegungssatz)**.** *Sei* $(\Omega, \mathscr{A}, \mu)$ σ*-endlich,* ν *ein signiertes Maß auf* $\mathscr{A}$, $|\nu|$ σ*-endlich. Dann gibt es eine eindeutige Zerlegung*

$$\nu = \nu_1 + \nu_2$$

von ν *als Summe zweier signierter Maße* ν_1, ν_2 *mit* $\nu_1 \ll \mu$, $\nu_2 \perp \mu$.

Beweis. a) *Sei ν ein σ-endliches Maß.* Wir bilden $\rho = \mu + \nu$. Dann gilt $\mu \ll \rho$, $\nu \ll \rho$. Seien $f = d\mu/d\rho$, $g = d\nu/d\rho$, d. h. $\mu(A) = \int_A f\,d\rho$, $\nu(A) = \int_A g\,d\rho$ für alle $A \in \mathscr{A}$. Sei

$$D = \{f > 0\}\,,\ D^c = \{f = 0\}$$

und

$$\nu_1(A) = \nu(A \cap D),\ \nu_2(A) = \nu(A \cap D^c), \quad A \in \mathscr{A}.$$

Dann gilt $\nu = \nu_1 + \nu_2$.

Wir behaupten, dass $\nu_1 \ll \mu$, $\nu_2 \perp \mu$ gilt.

Zu $\nu_1 \ll \mu$. Sei $\mu(A) = 0$. Dann ist $\int_A f\,d\rho = 0$, d. h. $f = 0$ ρ-f. ü. auf A. Jedoch $f > 0$ auf $A \cap D$, also $\rho(A \cap D) = 0$. Daher $\nu(A \cap D) = 0$, d. h. $\nu_1(A) = 0$.

Zu $\nu_2 \perp \mu$. Es gilt $\nu_2(D) = 0$ und

$$\mu(D^c) = \int_{\{f=0\}} f\,d\rho = 0.$$

b) *Sei ν ein σ-endliches signiertes Maß.* Wende obiges Argument getrennt auf ν^+, ν^- an.

c) *Eindeutigkeit der Zerlegung.* Sei zuerst ν endlich und $\nu = \nu_1 + \nu_2 = \nu_1' + \nu_2'$ mit $\nu_1 \ll \mu$, $\nu_1' \ll \mu$, $\nu_2 \perp \mu$, $\nu_2' \perp \mu$. Dann ist $\nu_1 - \nu_1' = \nu_2' - \nu_2$ sowohl absolutstetig als auch singulär bezüglich μ, d. h. $\nu_1 - \nu_1' = \nu_2' - \nu_2 = 0$ bzw. $\nu_1 = \nu_1'$, $\nu_2 = \nu_2'$.

Ist ν nun σ-endlich, so schreiben wir $\Omega = \bigsqcup_i A_i$ mit $|\nu|(A_i) < \infty$ für alle i. Wir erhalten die Eindeutigkeit der Zerlegung auf jedem A_i und damit auch auf Ω. $\square$

10.5.3 Produktmaße

Anschaulich ist es klar, dass $\lambda_n = \underbrace{\lambda_1 \times \ldots \times \lambda_1}_{n-\text{mal}}$ oder allgemeiner $\mu_F = \mu_{F_1} \times \ldots \times \mu_{F_n}$ gilt, wobei μ_{F_j} das Lebesgue-Stieltjes-Maß auf $\mathbb{R}^1$ mit der Verteilungsfunktion F_j und $F(x) = F_1(x_1) \ldots F_n(x_n)$ ist.
Wir wollen diese Anschauung präzisieren.

Definition 10.75. Seien $\mathscr{A}_j$ σ-Algebren auf Ω_j, $j = 1, 2, \ldots, n$.

a) Eine *messbare Box* in $\Omega = \Omega_1 \times \ldots \times \Omega_n$ ist eine Menge der Form $A_1 \times \ldots \times A_n$ mit $A_j \in \mathscr{A}_j$ für alle j.

b) Die *Produkt-σ-Algebra* $\mathscr{A} = \mathscr{A}_1 \times \ldots \times \mathscr{A}_n$ ist die von allen messbaren Boxen erzeugte σ-Algebra auf Ω.

Im Fall $A_1 = \ldots = A_n$ schreibt man auch $\mathscr{A} = (\mathscr{A}_1)^n$.

Bemerkung. Die endlichen disjunkten Vereinigungen messbarer Boxen bilden eine Algebra auf Ω.

Satz 10.76 (Produktmaße). *Sei $(\Omega_1, \mathscr{A}_1, \mu_1)$ ein σ-endlicher Maßraum, $(\Omega_2, \mathscr{A}_2)$ ein messbarer Raum. Weiterhin seien die folgenden Daten gegeben:*

(i) *Für jedes $\omega_1 \in \Omega_1$ ist $\mu_2^{\omega_1}$ ein Maß auf $\mathscr{A}_2$.*

(ii) *Für jedes $B \in \mathscr{A}_2$ ist die Abbildung $\omega_1 \mapsto \mu_2^{\omega_1}(B)$ $\mathscr{A}_1$-messbar.*

(iii) *$\mu_2^{\omega_1}$ ist bezüglich ω_1 gleichmäßig σ-endlich, d. h. es gibt eine abzählbare Zerlegung $\Omega_2 = \bigcup_k B_k$ mit $B_k \in \mathscr{A}_2$ für alle k und es existieren Konstanten C_k mit $\mu(\omega_1, B_k) \leq C_k$ für alle $\omega_1 \in \Omega_1$.*

Dann gibt es genau ein Maß μ auf $\mathscr{A} = \mathscr{A}_1 \times \mathscr{A}_2$ mit

$$\mu(A \times B) = \int_A \mu_2^{\omega_1}(B)\, d\mu_1(\omega), \quad \forall A \in \mathscr{A}_1,\ B \in \mathscr{A}_2.$$

Tatsächlich gilt für alle $C \in \mathscr{A}$

$$\mu(C) = \int_{\Omega_1} \mu_2^{\omega_1}(C(\omega_1))\, d\mu_1(\omega_1),$$

wobei $C(\omega_1) = \{\omega_2 \in \Omega_2 \mid (\omega_1, \omega_2) \in C\}$ der Schnitt von C auf Höhe ω_1 ist.[22] *Außerdem ist das Maß μ σ-endlich. Sind μ_1 und alle $\mu_2^{\omega_1}$ Wahrscheinlichkeitsmaße, so ist auch μ ein Wahrscheinlichkeitsmaß.*

Beweis. a) Sei zuerst $\mu_2^{\omega_1}$ gleichmässig endlich bezüglich ω_1, d. h. es gäbe eine Konstante $C > 0$ mit $\mu_2^{\omega_1}(\Omega_2) \leq C$ für alle $\omega_1 \in \Omega_1$.

α) Wir zeigen $C(\omega_1) \in \mathscr{A}_2$ für $C \in \mathscr{A}$, $\omega_1 \in \Omega_1$.
Sei dazu $\mathscr{C} = \{C \in \mathscr{A} \mid C(\omega_1) \in \mathscr{A}_2\}$, wobei $\omega_1 \in \Omega_1$ fixiert wurde. $\mathscr{C}$ ist eine σ-Algebra, da

$$\left(\bigcup\nolimits_i C_i\right)(\omega_1) = \bigcup\nolimits_i C_i(\omega_1), \quad C^c(\omega_1) = (C(\omega_1))^c.$$

Falls $A \in \mathscr{A}_1$, $B \in \mathscr{A}_2$, so ist $(A \times B)(\omega_1) = B$ für $\omega_1 \in A$ und $(A \times B)(\omega_1) = \emptyset$ für $\omega_1 \notin A$. Somit enthält $\mathscr{C}$ alle messbaren Boxen und folglich $\mathscr{C} = \mathscr{A}$.

β) Wir zeigen, dass die Abbildung $\omega_1 \mapsto \mu_2^{\omega_1}(C(\omega_1))$ für $C \in \mathscr{A}$ $\mathscr{A}_1$-messbar ist.
Sei $\mathscr{C}$ die Menge aller Mengen $C \in \mathscr{A}$, für die das richtig ist. Falls $C = A \times B$ mit $A \in \mathscr{A}_1$, $B \in \mathscr{A}_2$, so ist $C \in \mathscr{C}$ wegen

$$\mu_2^{\omega_1}((A \times B)(\omega_1)) = \begin{cases} \mu(\omega_1, B), & \text{falls } \omega_1 \in A, \\ 0, & \text{falls } \omega_1 \notin A, \end{cases}$$

und $\omega_1 \mapsto \mu_2^{\omega_1}((A \times B)(\omega_1)) = \chi_A(\omega_1)\, \mu_2^{\omega_1}(B)$ ist $\mathscr{A}_1$-messbar.
Ist $\{C_i\} \subset \mathscr{C}$ eine abzählbare, paarweise disjunkte Familie, so ist $\omega_1 \mapsto \mu_2^{\omega_1}((\bigcup_i C_i)(\omega_1)) = \mu_2^{\omega_1}(\bigcup_i C_i(\omega_1)) = \sum_i \mu_2^{\omega_1}(C_i(\omega_1))$ als Reihensumme nichtnegativer $\mathscr{A}_1$-messbarer Funktionen selbst $\mathscr{A}_1$-messbar.
Ist schließlich $C \in \mathscr{C}$, so ist auch $\omega_1 \mapsto \mu_2^{\omega_1}((C^c)(\omega_1)) = \mu_2^{\omega_1}(C(\omega_1)^c) = \mu_2^{\omega_1}(\Omega_2) - \mu_2^{\omega_1}(C(\omega_1))$ $\mathscr{A}_1$-messbar. Insgesamt ergibt sich $\mathscr{C} = \mathscr{A}$.

[22] Vgl. mit dem Prinzip von Cavalieri.

γ) Wir setzen jetzt

$$\mu(C) = \int_{\Omega_1} \mu_2^{\omega_1}(C(\omega_1))\, d\mu_1(\omega_1), \quad C \in \mathscr{A}.$$

Dann ist μ ein Maß auf $\mathscr{A}$. Es gilt nämlich $\mu(\emptyset) = 0$ und ist $\{C_i\} \subset \mathscr{A}$ eine abzählbare, paarweise disjunkte Familie, so ist

$$\begin{aligned}\mu\Big(\bigcup\nolimits_i C_i\Big) &= \int_{\Omega_1} \mu_2^{\omega_1}\Big(\bigcup\nolimits_i C_i(\omega_1)\Big)\, d\mu_1(\omega_1) = \sum_i \int_{\Omega_1} \mu_2^{\omega_1}(C(\omega_1))\, d\mu_1(\omega_1)\\ &= \sum_i \mu(C_i).\end{aligned}$$

Weiterhin gilt für $A \in \mathscr{A}_1$, $B \in \mathscr{A}_2$, dass

$$\begin{aligned}\mu(A \times B) &= \int_{\Omega_1} \mu_2^{\omega_1}((A \times B)(\omega_1))\, d\mu_1(\omega_1) = \int_{\Omega_1} \chi_A(\omega_1)\, \mu_2^{\omega_1}(B)\, d\mu_1(\omega_1)\\ &= \int_A \mu_2^{\omega_1}(B)\, d\mu_1(\omega_1),\end{aligned}$$

wie behauptet.

b) Ist $\mu_2^{\omega_1}$ gleichmäßig σ-endlich, $\Omega_2 = \bigsqcup_k B_k$ eine entsprechende Zerlegung von Ω_2 mit $\mu_2^{\omega_1}(B_k) \leq C_k$ für alle $\omega_1 \in \Omega_1$ und für alle k, so setzen wir

$$\mu_{2,k}^{\omega_1}(B) = \mu_2^{\omega_1}(B \cap B_k), \quad \forall B \in \mathscr{A}_2.$$

Dann ist $\mu_{2,k}^{\omega_1}$ gleichmäßig endlich, also gibt es nach a) ein Maß μ_k auf $\mathscr{A}$ mit

$$\mu_k(A \times B) = \int_A \mu_2^{\omega_1}(B \cap B_k)\, d\mu_1(\omega_1)$$

für alle $A \in \mathscr{A}_1$, $B \in \mathscr{A}_2$. Das Maß $\mu = \sum_k \mu_k$ auf $\mathscr{A}$ leistet das Verlangte.

c) *σ-Endlichkeit und Eindeutigkeit.* Wir behaupten zuerst, dass das in a), b) konstruierte Maß μ σ-endlich auf $\mathscr{A}$ ist. Ist nämlich $\Omega_2 = \bigsqcup_k B_k$ eine Zerlegung wie in b) und $\Omega_1 = \bigsqcup_i A_i$ mit $\mu_1(A_i) < \infty$ für alle i, so gilt $\Omega = \Omega_1 \times \Omega_2 = \bigsqcup_{i,k}(A_i \times B_k)$ sowie

$$\mu(A_i \times B_k) = \int_{A_i} \mu_2^{\omega_1}(B_k)\, d\mu_1(\omega_1) \leq C_k\, \mu_1(A_i) < \infty.$$

Ist nun σ ein weiteres Maß auf $\mathscr{A}$ mit

$$\sigma(A \times B) = \int_A \mu_2^{\omega_1}(B)\, d\mu_1(\omega_1) \quad \forall A \in \mathscr{A}_1,\, B \in \mathscr{A}_2,$$

so stimmt σ mit μ auf der *Algebra* der endlichen, disjunkten Vereinigungen von messbaren Boxen überein. Nach der Eindeutigkeitsaussage im Fortsetzungssatz von Carathéodory folgt damit $\sigma = \mu$ auch auf $\mathscr{A}$. $\square$

Folgerung 10.77 (Produktmaße, klassische Version). *Seien* $(\Omega_j, \mathscr{A}_j, \mu_j)$ *σ-endliche Maßräume, $j = 1, 2$. Dann gibt es genau ein Maß μ auf $\mathscr{A}_1 \times \mathscr{A}_2$ mit*

$$\mu(A \times B) = \mu_1(A)\,\mu_2(B) \quad \forall A \in \mathscr{A}_1,\, B \in \mathscr{A}_2.$$

Dieses Maß ist σ-endlich. Sind μ_1, μ_2 Wahrscheinlichkeitsmaße, so ist auch μ ein Wahrscheinlichkeitsmaß.

Beweis. Im vorigen Satz wähle man $\mu_2^{\omega_1} = \mu_2$ für alle $\omega_1 \in \Omega_1$. □

Man schreibt in dieser Situation auch $\mu = \mu_1 \times \mu_2$.

Beispiel 10.78. $\mathscr{A}_1 = \mathscr{A}_2 = \mathscr{B}(\mathbb{R})$. Dann ist $\mathscr{A}_1 \times \mathscr{A}_2 = \mathscr{B}(\mathbb{R}^2)$. Für $\mu_1 = dF_1$ und $\mu_2 = dF_2$ mit F_1, F_2 Verteilungsfunktionen auf $\mathbb{R}^1$ gilt $\mu_1 \times \mu_2 = dF_1 dF_2$.

Wir möchten jetzt Integrale über (beispielsweise) Bereichen des $\mathbb{R}^n$ als iterierte Integrale berechnen.

Satz 10.79 (Fubini). *Die obigen Annahmen seien erfüllt, $f \in \mathscr{M}(\Omega_1 \times \Omega_2, \overline{\mathbb{R}})$.*

a) *Gilt $f \geq 0$, so existiert $\int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2^{\omega_1}(\omega_2)$ und ist eine messbare Funktion von ω_1. Weiterhin gilt*

$$\int_\Omega f\, d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2^{\omega_1}(\omega_2) \right) d\mu_1(\omega_1).$$

b) *Existiert das Integral $\int_\Omega f\, d\mu$, so existiert $\int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2^{\omega_1}(\omega_2)$* μ_1-f. ü. *und definiert eine messbare Funktion, sobald letztere auf der Ausnahmemenge, auf der dieses Integral nicht existiert, zu* (*beispielsweise*) *Null gesetzt wird. Weiterhin gilt die Formel aus* a).

Beweis. a) Wir zeigen $f(\omega_1, \cdot) \in \mathscr{M}(\Omega_2, \overline{\mathbb{R}})$ für alle $\omega_1 \in \Omega_1$. Sei dazu $B \in \mathscr{B}(\overline{\mathbb{R}})$. Dann ist $\{f(\omega_1, \cdot) \in B\} = f^{-1}(B)(\omega_1) \in \mathscr{A}_2$.
Folglich existiert $\int_\Omega f(\omega_1, \omega_2)\, d\mu_2^{\omega_1}(\omega_2)$ für alle $\omega_1 \in \Omega_1$.
Sei nun $C \in \mathscr{A}$, $f = \chi_C$. Dann ist die Funktion $\omega_1 \mapsto \int_{\Omega_2} \chi_C(\omega_1, \omega_2)\, d\mu_2^{\omega_1}(\omega_2) = \int_{\Omega_2} \chi_{C(\omega_1)}(\omega_2)\, d\mu_2^{\omega_1}(\omega_2) = \mu_2^{\omega_1}(C(\omega_1))$ messbar und

$$\begin{aligned}\int_\Omega \chi_C\, d\mu = \mu(C) &= \int_{\Omega_1} \mu_2^{\omega_1}(C(\omega_1))\, d\mu_1(\omega_1) \\ &= \int_{\Omega_1} \left(\int_{\Omega_2} \chi_C(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2) \right) d\mu_1(\omega_1).\end{aligned}$$

Der Rest des Beweises ist jetzt wie üblich durch Approximation mit Stufenfunktionen.

b) O. B. d. A. sei $\int_\Omega f^-\, d\mu < \infty$. Dann gilt gemäß a)

$$\int_\Omega f^-\, d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f^-(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2) \right) d\mu_1(\omega_1) < \infty,$$

so dass $\int_{\Omega_2} f^-(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2) < \infty$ μ_1-f. ü. Insbesondere erhalten wir für $\omega_1 \in \Omega_1$ μ_1-f. ü.

$$\int_{\Omega_2} f(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2) = \int_{\Omega_2} f^+(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2) - \int_{\Omega_2} f^-(\omega_1, \omega_2)\, \mu_2^{\omega_1}(\omega_2)$$

und die Behauptung folgt mit a). □

Folgerung 10.80. *Ist $f \in \mathscr{M}(\Omega_1 \times \Omega_2, \overline{\mathbb{R}})$ und*

$$\int_{\Omega_1} \left(\int_{\Omega_2} |f(\omega_1, \omega_2)| \, \mu(\omega_1, d\omega_2) \right) d\omega_1 < \infty,$$

so ist $\int_\Omega f \, d\mu$ endlich und Fubinis Theorem ist anwendbar.

Beweis. Aus dem vorigen Satz folgt $\int_\Omega |f| \, d\mu < \infty$. $\square$

Satz 10.81 (Fubini, klassische Version)**.** *Seien $(\Omega_j, \mathscr{A}_j, \mu_j)$ σ-endliche Maßräume, $j = 1, 2$. Ist $f \in \mathscr{M}(\Omega_1 \times \Omega_2, \overline{\mathbb{R}})$ und existiert $\int_\Omega f \, d\mu$ mit $\mu = \mu_1 \times \mu_2$, so ist*

$$\int_\Omega f \, d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f \, d\mu_2 \right) d\mu_1 = \int_{\Omega_2} \left(\int_{\Omega_1} f \, d\mu_1 \right) d\mu_2.$$

Beweis. Benutze den vorigen Satz mit $\mu_2^{\omega_1} = \mu_2$ für alle $\omega_1 \in \Omega_1$. $\square$

Wir betrachten jetzt den Fall, dass mehr als zwei messbare Räume gegeben sind. Seien also $\mathscr{A}_j$ σ-Algebren auf Ω_j, $j = 1, 2, \ldots, n$. Weiter sei gegeben:

- μ_1 ist ein σ-endliches Maß auf $\mathscr{A}_1$,
- $\mu_2^{\omega_1}$ ist ein Maß auf $\mathscr{A}_2$ für jedes $\omega_1 \in \Omega_1$, die Abbildung $\omega_1 \mapsto \mu_2^{\omega_1}(A_2)$ ist $\mathscr{A}_1$-messbar für jedes $A_2 \in \mathscr{A}_2$ und $\mu_2^{\omega_1}$ ist bezüglich ω_1 gleichmäßig σ-endlich,
- $\mu_3^{\omega_1, \omega_2}$ ist ein Maß auf $\mathscr{A}_3$ für jedes $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$, die Abbildung $\Omega_1 \times \Omega_2 \to [0, \infty]$, $(\omega_1, \omega_2) \mapsto \mu_3^{\omega_1, \omega_2}(A_3)$ ist $\mathscr{A}_1 \times \mathscr{A}_2$-messbar für jedes $A_3 \in \mathscr{A}_3$ und $\mu_3^{\omega_1, \omega_2}$ ist bezüglich (ω_1, ω_2) gleichmäßig σ-endlich,

$\vdots$

- $\mu_n^{\omega_1, \ldots, \omega_{n-1}}$ ist ein Maß auf $\mathscr{A}_n$ für jedes $(\omega_1, \ldots, \omega_{n-1}) \in \Omega_1 \times \ldots \times \Omega_{n-1}$, die Abbildung $\Omega_1 \times \ldots \times \Omega_{n-1} \to [0, \infty]$, $(\omega_1, \ldots, \omega_{n-1}) \mapsto \mu_n^{\omega_1, \ldots, \omega_{n-1}}(A_n)$ ist $\mathscr{A}_1 \times \ldots \times \mathscr{A}_{n-1}$-messbar für jedes $A_n \in \mathscr{A}_n$ und $\mu_n^{\omega_1, \ldots, \omega_{n-1}}$ ist bezüglich $(\omega_1, \ldots, \omega_{n-1})$ gleichmäßig σ-endlich,

Satz 10.82. a) *Unter diesen Bedingungen gibt es genau ein Maß μ auf $\mathscr{A}_1 \times \ldots \times \mathscr{A}_n$ mit*

$$\mu(A_1 \times \ldots \times A_n) = \int_{A_1} \left(\int_{A_2} \left(\ldots \left(\int_{A_{n-1}} \left(\int_{A_n} d\mu_n^{\omega_1, \ldots, \omega_{n-1}}(\omega_n) \right) \right. \right. \right. \right.$$
$$\left. \left. \left. \left. d\mu_{n-1}^{\omega_1, \ldots, \omega_{n-2}}(\omega_{n-1}) \right) \ldots \right) d\mu_2^{\omega_1}(\omega_2) \right) d\mu_1(\omega_1).$$

b) *Ist $f \in \mathscr{M}(\Omega_1 \times \ldots \times \Omega_n, \overline{\mathbb{R}})$ derart, dass das Integral $\int_\Omega f \, d\mu$ existiert, so lässt sich dieses Integral iterativ vermöge der Formel*

$$\int_\Omega f \, d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} \left(\ldots \left(\int_{\Omega_{n-1}} \left(\int_{\Omega_n} f(\omega_1, \ldots, \omega_n) \, d\mu_n^{\omega_1, \ldots, \omega_{n-1}}(\omega_n) \right) \right. \right. \right. \right.$$
$$\left. \left. \left. \left. d\mu_{n-1}^{\omega_1, \ldots, \omega_{n-2}}(\omega_{n-1}) \right) \ldots \right) d\mu_2^{\omega_1}(\omega_2) \right) d\mu_1(\omega_1)$$

berechnen, wobei gegebenenfalls im $(n-j)$-ten für $j = n-1, \ldots, 1$ Schritt die resultierende Funktion von $\omega_1, \ldots, \omega_j$ auf einer $\widehat{\mu}_j$-Nullmenge abzuändern ist, wenn $\widehat{\mu}_j$ das durch μ_1, $\mu_2^{\omega_1}, \ldots, \mu_j^{\omega_1,\ldots,\omega_{j-1}}$ gemäß a) *bestimmte Maß auf $\mathscr{A}_1 \times \ldots \times \mathscr{A}_j$ bezeichnet.*

Beweis. Induktion über n unter Benutzung des Falls $n = 2$. □

Bemerkung. Die Formel in a) ist ein Spezialfall von b) mit $f(\omega_1, \ldots, \omega_n) = \chi_{A_1}(\omega_1) \ldots \chi_{A_n}(\omega_n)$.

Folgerung 10.83. *Sind $(\Omega_j, \mathscr{A}_j, \mu_j)$ σ-endliche Maßräume, $j = 1, \ldots, n$, so gibt es genau ein Maß μ mit*

$$\mu\left(\bigtimes_{j=1}^{n} A_j\right) = \prod_{j=1}^{n} \mu_j(A_j)$$

für alle $A_1 \in \mathscr{A}_1, \ldots, A_n \in \mathscr{A}_n$. Die Formel zur iterativen Berechnung von $\int_\Omega f\, d\mu$ für $f \in \mathscr{M}(\Omega_1 \times \ldots \times \Omega_n, \overline{\mathbb{R}})$ gilt wie im vorigen Satz angegeben.

In diesem Fall schreiben wir $\mu = \mu_1 \times \ldots \times \mu_n$.

Beispiel 10.84. Sei $n \in \mathbb{N}$, $n = k_1 + \ldots + k_m$ mit $k_i \in \mathbb{N}$. Dann gilt

$$\bar{\mathscr{B}}(\mathbb{R}^n) = \bar{\mathscr{B}}(\mathbb{R}^{k_1}) \times \ldots \times \bar{\mathscr{B}}(\mathbb{R}^{k_m})$$

und $\lambda_n = \lambda_{k_1} \times \ldots \times \lambda_{k_m}$. Eine analoge Aussage gilt für Lebesgue-Stieltjes-Maße, deren Verteilungsfunktionen F Produkte von Funktionen wenigerer Variablen sind, d. h. wenn

$$\begin{aligned} &F(x_1, \ldots, x_n) \\ &\qquad = F_1(x_1, \ldots, x_{k_1}) F_2(x_{k_1+1}, \ldots, x_{k_2}) \ldots F_m(x_{k_1+\cdots+k_{m-1}+1}, \ldots, x_n). \end{aligned}$$

Berechnung einiger Integrale

Für die folgenden Rechnungen benötigen wir einige Formeln. Wir erinnern daran, dass die *Gammafunktion* $\Gamma(x)$ durch durch

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt, \quad x > 0,$$

gegeben ist.

Satz 10.85. *Es gilt:*

(i) $\Gamma(n+1) = n!$ *für* $n \in \mathbb{N}_0$.

(ii) $\Gamma(x+1) = x\,\Gamma(x)$ *für* $x > 0$.

(iii) (Reflektionsformel) $\Gamma(x)\,\Gamma(1-x) = \dfrac{\pi}{\sin(\pi x)}$ *für* $0 < x < 1$.

(iv) (Verdopplungsformel) $\Gamma(x)\,\Gamma\left(x + \dfrac{1}{2}\right) = 2^{1-2x}\sqrt{\pi}\,\Gamma(2x)$ *für* $x > 0$.

Beweis. (i), (ii) wurden in der Vorlesung “Differenzial- und Integralrechnung I” gezeigt, einen kurzen Beweis für (iii), (iv) wird es in der Vorlesung “Funktionentheorie” geben. □

Insbesondere erhalten wir:

Folgerung 10.86. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$, $\Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4}$.

Weiterhin benötigen wir die *Betafunktion* $B(x,y)$, die durch das Integral

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt, \quad x>0,\, y>0,$$

eingeführt wird.

Lemma 10.87. *Für $x>0$, $y>0$ gilt*

$$B(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}.$$

Beweis. Mit der Substitution $t = uv$, $s = u(1-v)$ erhalten wir

$$\begin{aligned}\Gamma(x)\,\Gamma(y) &= \int_{\mathbb{R}_+^2} e^{-(t+s)} t^{x-1} s^{y-1}\,dtds \\ &= \int_{\mathbb{R}_+\times(0,1)} e^{-u} u^{x+y-1} v^{x-1}(1-v)^{y-1}\,dudv \\ &= \Gamma(x+y)\,B(x,y).\end{aligned}$$

□

Lemma 10.88. *Für $a > -1$ gilt*

$$\int_{-1}^1 \left(1-t^2\right)^a dt = \sqrt{\pi}\,\frac{\Gamma(a+1)}{\Gamma\left(a+\frac{3}{2}\right)}.$$

Beweis. Mit der Substitution $s = (1+t)/2$ und unter Benutzung der Verdopplungsformel für die Gammafunktion ergibt sich

$$\begin{aligned}\int_{-1}^1 \left(1-t^2\right)^a dt &= 2^{2a+1}\int_0^1 s^a(1-s)^a\,ds \\ &= 2^{2a+1}\,\frac{\Gamma(a+1)^2}{\Gamma(2a+2)} = \sqrt{\pi}\,\frac{\Gamma(a+1)}{\Gamma\left(a+\frac{3}{2}\right)}.\end{aligned}$$

□

Das n-dimensionale Kugelvolumen

Sei $\mathbb{B}^n = \{x \in \mathbb{R}^n \mid |x| \le 1\}$ die n-dimensionalen Kugel um 0 mit dem Radius 1 und sei ω_n deren Volumen. Aus den Saklierungseigenschaften des Lebesgue-Maßes folgt, dass $\omega_n R^n$ das Volumen der n-dimensionalen Kugel $R\,\mathbb{B}^n$ um 0 mit dem Radius $R > 0$ ist.

Satz 10.89. *Es gilt*

$$\omega_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Beispiel 10.90. Wir erhalten[23]

$$\omega_1 = \frac{\sqrt{\pi}}{\Gamma\left(\frac{3}{2}\right)} = 2, \quad \omega_2 = \frac{\pi}{\Gamma(2)} = \pi, \quad \omega_3 = \frac{\pi^{3/2}}{\Gamma\left(\frac{5}{2}\right)} = \frac{4\pi}{3},$$

was mit den aus der Schule bekannten Ergebnissen übereinstimmt.

Beweis. Für das n-dimensionale Kugelvolumen ergibt sich wegen $\lambda_n = \lambda_1 \times \lambda_{n-1}$ mit dem Satz über Produktmaße und dem vorigen Lemma

$$\begin{aligned} \omega_n = \lambda_n\left(\mathbb{B}^n\right) &= \int_{-1}^{1} \lambda_{n-1}\left(\sqrt{1-x_n^2}\,\mathbb{B}^{n-1}\right) dx_n \\ &= \omega_{n-1} \int_{-1}^{1} (1-x_n^2)^{\frac{n-1}{2}}\, dx_n = \omega_{n-1} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}. \end{aligned}$$

Die Behauptung folgt durch Induktion nach n. □

Anhang zu Abschnitt 10.5.2: Absolut-stetige Maße bezüglich des Lebesgue-Maßes

[23] Formal können wir auch $n = 0$ zulassen; $\omega_0 = 1$.

Kapitel 11

Das Lebesgue-Integral

Wir studieren jetzt genauer Eigenschaften des Lebesgue-Maßes und des Lebesgue-Integrals. Insbesondere verallgemeinern wir mit dem Satz von Gauß die Formeln für die partielle Integration vom ein- auf den mehrdimensionalen Fall.

11.1 Regularitätseigenschaften des Lebesgue-Maßes

Sei λ_{n*} das äußere Lebesgue-Maß auf $\mathbb{R}^n$, d. h.

$$\lambda_{n*}(A) = \inf\Big\{\sum_i |Q_i| \;\Big|\; \{Q_i\} \text{ abzählbare Familie abgeschlossener, achsenparalleler Würfel, } A \subseteq \bigcup_i Q_i\Big\}, \quad A \subseteq \mathbb{R}^n.$$

Wir wissen bereits, dass λ_{n*} ein äußeres Maß auf $\mathbb{R}^n$ ist.[1] Insbesondere hat λ_{n*} die folgenden Eigenschaften:

(i) $\lambda_{n*}(\emptyset) = 0$.

(ii) **(Monotonie)** Ist $A \subseteq B$, so ist $\lambda_{n*}(A) \leq \lambda_{n*}(B)$.

(iii) **(σ-Subadditivität)** Ist $\{A_i\}$ eine abzählbare Familie von Teilmengen von $\mathbb{R}^n$, so gilt $\lambda_{n*}(\bigcup_i A_i) \leq \sum_i \lambda_{n*}(A_i)$.

Weiterhin gilt:

(iv) Ist $A \subseteq \mathbb{R}^n$, so ist

$$\lambda_{n*}(A) = \inf\big\{\lambda_n(U) \;\big|\; U \subseteq \mathbb{R}^n \text{ offen, } A \subseteq U\big\}.$$

Beweis. $(\leq)$ $\lambda_{n*}(A) \leq \inf \lambda_n(U)$ folgt aus der Monotonie des äußeren Maßes.

[1] Tatsächlich haben wir dies im Fall $n = 1$ bewiesen, doch der Beweis im allgemeinen Fall ist ähnlich.

($\geq$) Sei $\epsilon > 0$. Wir wählen Würfel[2] Q_i mit $A \subseteq \bigcup_i Q_i$ und $\sum_i |Q_i| \leq \lambda_{n*}(A) + \epsilon/2$. Seien Q_{i0} offene Würfel mit $Q_{i0} \supset Q_i$ und $|Q_{i0}| \leq |Q_i| + \epsilon/2^{i+1}$. Dann ist die Menge $U = \bigcup_i Q_{i0}$ offen sowie

$$\lambda_n(U) \leq \sum_i |Q_{i0}| \leq \sum_i \left(|Q_i| + \frac{\epsilon}{2^{i+1}}\right) \leq \sum_i |Q_i| + \frac{\epsilon}{2} \leq \lambda_{n*}(A) + \epsilon.$$

Da $\epsilon > 0$ beliebig wählbar ist, erhalten wir

$$\inf \lambda_n(U) \leq \lambda_{n*}(A).$$ □

(v) Sei A, $B \subseteq \mathbb{R}^n$ mit $\inf \big\{|x - y| \;\big|\; x \in A, y \in B\big\} > 0$. Dann gilt

$$\lambda_{n*}(A \cup B) = \lambda_{n*}(A) + \lambda_{n*}(B).$$

Bemerkung. Diese Eigenschaft charakterisiert λ_{n*} als ein sogenanntes *metrisches äußeres Maß*. Es kann gezeigt werden, dass bezüglich eines metrischen äußeren Maßes auf einem beliebigen metrischen Raum (X, d) die Borelmengen Carathéodory-messbar sind, dieses äußere Maß also ein Borelmaß induziert.

Beweis. ($\leq$) Aus der σ-Subadditivität folgt $\lambda_{n*}(A \cup B) \leq \lambda_{n*}(A) \cup \lambda_{n*}(B)$.

($\geq$) Sei $\delta = \inf \big\{|x - y| \;\big|\; x \in A, y \in B\big\} > 0$. Sei weiterhin $A \cup B \subseteq \bigcup_{i \in I} Q_i$, $\sum_{i \in I} |Q_i| \leq \lambda_{n*}(A \cup B) + \epsilon$ für ein gegebenes $\epsilon > 0$. Falls notwendig, zerlegen wir jedes Q_i in achsenparallele kleinere Würfel, so dass wir annehmen dürfen, dass jeder Würfel Q_i einen Durchmesser kleiner als δ hat. Damit hat jeder Würfel Q_i einen nichtleeren Durchschnitt mit höchstens einer der Mengen A oder B. Sei $I = I_1 \cup I_2$, $I_1 \cap I_2 = \emptyset$, mit $Q_i \cap B = \emptyset$ für $i \in I_1$ und $Q_i \cap A = \emptyset$ für $i \in I_2$. Dann ist $A \subseteq \bigcup_{i \in I_1} Q_i$, $B \subseteq \bigcup_{i \in I_2} Q_i$ sowie

$$\lambda_{n*}(A) + \lambda_{n*}(B) \leq \sum_{i \in I_1} |Q_i| + \sum_{i \in I_2} |Q_i| = \sum_{i \in I} |Q_i| \leq \lambda_{n*}(A \cup B) + \epsilon,$$

woraus die Behauptung folgt. □

(vi) Sei $A = \bigcup_i Q_i$ abzählbare Vereinigung von Würfeln Q_i, wobei $Q_i \cap Q_j \subseteq \partial Q_i$ für alle $i \neq j$ gilt.[3] Dann ist A Lebesgue-messbar und

$$\lambda_n(A) = \sum_i |Q_i|.$$

Bemerkung. Jede offene Menge $U \subseteq \mathbb{R}^n$ ist eine derartige Menge A.

Beweis. Dies folgt aus der Tatsache, dass $\lambda_n(Q_i \cap Q_j) = 0$ für $i \neq j$ gilt.

Wir geben zusätzlich einen Beweis an, in dem $\lambda_n(A)$ durch $\lambda_{n*}(A)$ ersetzt wird und der die Tatsache, dass A Lebesgue-messbar ist, nicht ausnutzt.

($\leq$) Gemäß Definition ist $\lambda_{n*}(A) \leq \sum_i |Q_i|$.

[2] Im Weiteren bezeichnet Q_i stets einen abgeschlossenen, achsenparallelen Würfel.
[3] Das heißt, es gilt $Q_i \cap Q_j = \partial Q_i \cap \partial Q_j$ für alle $i \neq j$.

($\geq$) Wir zeigen die umgekehrte Abschätzung.

Seien dazu Q_{i1} Würfel mit $Q_{i1} \subset (Q_i)^\circ$, $|Q_i| \leq |Q_{i1}| + \epsilon/2^i$. Für jedes i_0 sind die Würfel $Q_{i1}, \ldots, Q_{ii_0}$ kompakt und disjunkt, haben insbesondere positiven Abstand voneinander. Damit gilt mit (v), dass

$$\lambda_{n*}\left(\bigcup\nolimits_{i=1}^{i_0} Q_{i1}\right) = \sum_{i=1}^{i_0} |Q_{i1}| \geq \sum_{i=1}^{i_0} \left(|Q_i| - \frac{\epsilon}{2^i}\right) \geq \sum_{i=1}^{i_0} |Q_i| - \epsilon.$$

Für $i_0 \to \infty$ erhalten wir

$$\lambda_{n*}(A) \geq \lambda_{n*}\left(\bigcup\nolimits_{i=1}^{\infty} Q_{i1}\right) \geq \sum_{i=1}^{\infty} |Q_i| - \epsilon$$

und damit die Behauptung. □

Im Fall der Carathéodory-Messbarkeit für das äußere Lebesgue-Maß gibt es eine einfache Charakterisierung dieser Eigenschaft:

Satz 11.1. *Eine Menge $A \subseteq \mathbb{R}^n$ ist genau dann Lebesgue-messbar, wenn es für jedes $\epsilon > 0$ eine offene Menge $U \supseteq A$ mit $\lambda_{n*}(U \setminus A) \leq \epsilon$ gibt.*

Beweis. ($\Longrightarrow$) Sei $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$, d. h. es gilt $\lambda_{n*}(B) = \lambda_{n*}(A \cap B) + \lambda_{n*}(A^c \cap B)$ für jede Menge $B \subseteq \mathbb{R}^n$. O. B. d. A. sei $\lambda_n(A) < \infty$, andernfalls zerlegen wir A disjunkt in messbare Mengen endlichen Maßes und approximieren diese getrennt. Zu $\epsilon > 0$ wählen wir eine offene Menge $U \supseteq A$ mit $\lambda_n(U) \leq \lambda_n(A) + \epsilon$. Dann ist

$$\lambda_n(U) = \lambda_n(A) + \lambda_n(U \setminus A),$$

daher

$$\lambda_{n*}(U \setminus A) = \lambda_n(U \setminus A) = \lambda_n(U) - \lambda_n(A) \leq \epsilon.$$

($\Longleftarrow$) Zu jedem $m \in \mathbb{N}$ wählen wir eine offene Menge $U_m \supseteq A$ mit $\lambda_{n*}(U_m \setminus A) \leq 1/m$. Wir setzen $B = \bigcap_m U_m \in \mathscr{B}(\mathbb{R}^n)$. Dann ist $B \supseteq A$ und $\lambda_{n*}(B \setminus A) = 0$, also wegen der Vollständigkeit des Lebesgue-Maßes $B \setminus A \in \bar{\mathscr{B}}(\mathbb{R}^n)$ und

$$A = B \setminus (B \setminus A) \in \bar{\mathscr{B}}(\mathbb{R}^n).$$ □

Das Lebesgue-Maß besitzt die folgenden sogenannten *Regularitätseigenschaften*.

Satz 11.2. *Sei $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$, $\epsilon > 0$. Dann:*

a) *Es gibt eine offene Menge $U \supseteq A$ mit $\lambda_n(U \setminus A) \leq \epsilon$.*

b) *Es gibt eine abgeschlossene Menge $F \subseteq A$ mit $\lambda_n(A \setminus F) \leq \epsilon$.*

c) *Ist $\lambda_n(A) < \infty$, so gibt es eine kompakte Menge $K \subseteq A$ mit $\lambda_n(A \setminus K) \leq \epsilon$.*

Beweis. a) Dies wurde bereits gezeigt.

b) Zu $A^c \in \bar{\mathscr{B}}(\mathbb{R}^n)$ wählen wir eine offene Menge $U \supseteq A^c$ mit $\lambda_n(U \setminus A^c) \leq \epsilon$. Es bleibt $F = U \subseteq A$ zu setzen und $U \setminus A^c = A \setminus F$ zu beobachten.

c) Sei $F \subseteq A$ abgeschlossen mit $\lambda_n(F \setminus A) \leq \epsilon/2$. Wir setzen $K_m = F \cap \{|x| \leq m\}$ für $m \in \mathbb{N}$. Die Mengen K_m sind kompakt und $A \setminus K_m \downarrow A \setminus F$, also gilt $\lambda_n(A \setminus K_m) \to \lambda_n(A \setminus F)$ wegen $\lambda_n(A) < \infty$. Somit existiert ein m^* mit $\lambda_n(A \setminus K_{m^*}) \leq \epsilon$. □

Folgerung 11.3. *Für $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$ gilt*

$$\begin{aligned} \lambda_n(A) &= \inf \left\{ \lambda_n(O) \;\middle|\; U \text{ ist offen, } U \supseteq A \right\} \\ &= \sup \left\{ \lambda_n(K) \;\middle|\; K \text{ ist kompakt, } K \subseteq A \right\}. \end{aligned}$$

11.2 Eigenschaften Lebesgue-messbarer Funktionen

In diesem Abschnitt beweisen wir zwei wichtige Sätze.

Satz 11.4 (Egorov). *Sei $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$, $\lambda_n(A) < \infty$. Seien $\{f_i\} \subset \mathscr{M}(A, \mathbb{R})$ eine Folge, $f \in \mathscr{M}(A, \mathbb{R})$ und $f_i \to f$* f. ü.[4] *Dann existiert für alle $\epsilon > 0$ eine abgeschlossene Menge $F \subseteq A$ mit $\lambda_n(A \setminus F) \le \epsilon$ und $f_i \to f$ gleichmäßig auf F.*

Beweis. Es genügt, eine Menge $B \in \bar{\mathscr{B}}(\mathbb{R}^n)$, $B \subseteq A$ mit $\lambda_n(A \setminus B) \le \epsilon$ und $f_i \to f$ gleichmäßig auf B zu finden.
O. B. d. A. gelte $f_i(x) \to f(x)$ für alle $x \in A$. Für $i, m \in \mathbb{N}$ setzen wir

$$A_{im} = \left\{ x \in A \;\middle|\; |f_j(x) - f(x)| \le \frac{1}{m} \text{ für alle } j \ge i \right\}.$$

Dann $A_{im} \subseteq A_{i+1,m}$ und $A_{im} \uparrow A$ für $i \to \infty$. Insbesondere existiert für alle m ein i_m mit

$$\lambda_n(A \setminus A_{i_m,m}) \le \frac{1}{2^m}.$$

Wir wählen m_0 mit $\sum_{m \ge m_0} \frac{1}{2^m} \le \epsilon$ und setzen $B = \bigcap_{m \ge m_0} A_{i_m,m}$. Dann gilt

$$\lambda_n(A \setminus B) \le \sum_{m \ge m_0} \lambda_n(A \setminus A_{i_m,m}) \le \epsilon$$

sowie $f_i \to f$ gleichmäßig auf B:
Sei dazu $\delta > 0$ und $m \ge \max\left\{m_0, \frac{1}{\delta}\right\}$. Dann gilt für $x \in B \subseteq A_{i_m,m}$, dass

$$|f_j(x) - f(x)| \le \frac{1}{m} \le \delta$$

für alle $j \ge i_m$, was die gleichmäßige Konvergenz beweist. □

Um den nächsten Satz vorzubereiten, zeigen wir:

Lemma 11.5. *Sei $f \in \mathscr{M}(\mathbb{R}^n, \mathbb{R})$. Dann gibt es eine Folge $\{f_i\} \subset \mathscr{M}_0(\mathbb{R}^n, \mathbb{R})$ von Stufenfunktionen mit $f_i \to f$* f. ü. *der speziellen Form*

$$f_i = \sum_k c_{ik}\, \chi_{Q_{ik}},$$

wobei jeweils die Summe endlich ist, $c_{ik} \in \mathbb{R}$ und die Q_{ik} abgeschlossene, achsenparallele Würfel sind.

[4] Eigenschaften fast überall sind in diesem Abschnitt stets bezüglich des Lebesgue-Maßes zu verstehen.

Beweis. Es genügt, das Lemma für eine charakteristische Funktion χ_A mit $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$, $\lambda_n(A) < \infty$ zu zeigen.

Zu jedem $i \in \mathbb{N}$ gibt es eine endliche Familie $\{Q_{ik}\}$ disjunkter, abgeschlossener und achsenparalleler Würfel mit $\lambda_n\left(A \triangle \bigcup_k Q_{ik}\right) \leq \frac{1}{2^i}$. Wir setzen $f_i = \sum_k \chi_{Q_{ik}}$. Dann ist $\lambda_n(\{\chi_A \neq f_i\}) = \lambda_n\left(A \triangle \bigcup_k Q_{ik}\right) \leq \frac{1}{2^i}$, also

$$\lambda_n(\limsup_{i\to\infty}\{\chi_A \neq f_i\}) = 0$$

gemäß dem Lemma von Borel-Cantelli. Jedoch $f_i \to \chi_A$ auf dem Komplement von $\limsup_{i\to\infty}\{\chi_A \neq f_i\}$, was das Lemma beweist. □

Satz 11.6 (Lusin)**.** *Sei $A \in \bar{\mathscr{B}}(\mathbb{R}^n)$, $\lambda_n(A) < \infty$, $f \in \mathscr{M}(A, \mathbb{R})$. Dann existiert für jedes $\epsilon > 0$ eine abgeschlossene Menge $F \subseteq A$ mit $\lambda_n(A \setminus F) \leq \epsilon$ und $f\big|_F$ ist stetig.*

Bemerkung. Es wird *nicht* behauptet, dass f an allen Stellen von F stetig ist. Im Allgemeinen ist dies falsch.
Sei beispielsweise $A = [0, 1]$,

$$f(x) = \begin{cases} 1, & \text{falls } x \in [0,1] \cap \mathbb{Q}, \\ 0, & \text{falls } x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Dann ist f an *allen* Stellen unstetig. Jedoch mit $B = [0,1] \setminus \mathbb{Q}$ gilt $f\big|_B = 0$. Insbesondere ist $f\big|_B$ stetig, und $\lambda_n([0,1] \setminus B) = 0$. Es verbleibt, B von "innen" durch eine abgeschlossene Menge F zu approximieren.

Beweis des Satzes. Es genügt, eine Menge $B \in \bar{\mathscr{B}}(\mathbb{R}^n)$ mit $\lambda_n(A \setminus B) \leq \epsilon$ und $f\big|_B$ ist stetig zu konstruieren.

Sei $\{f_i\}$ eine Folge von Stufenfunktionen mit $f_i \to f$ f. ü., wie diese im vorigen Lemma konstruiert wurde. Sei A_i die Menge der Unstetigkeitsstellen von f_i, $\lambda_n(A_i) = 0$. Nach dem Satz von Egorov gibt es eine abgeschlossene Menge $F \subseteq \mathbb{R}^n$ mit $F \subseteq A$, $\lambda_n(A \setminus F) \leq \epsilon$ und $f_i \to f$ gleichmäßig auf F. Wir setzen $B = F \setminus \bigcup_i A_i$. Dann ist $\lambda_n(B) = \lambda_n(F)$, insbesondere $\lambda_n(A \setminus B) \leq \epsilon$ wegen $B \subseteq F$, $f_i\big|_B$ ist stetig für alle i und $f\big|_B$ ist stetig als gleichmäßiger Grenzwert stetiger Funktionen. □

11.3 Variablenwechsel unter dem Integral

Sei $\varphi\colon [a,b] \to [c,d]$ eine bijektive C^1-Abbildung mit $\varphi'(x) \neq 0$ für alle $x \in [a,b]$. Wie wir aus der Vorlesung "Differenzial- und Integralrechnung I" wissen, ist dann für jede Riemann-integrierbare Funktion $f\colon [c,d] \to \mathbb{R}$ die Funktion $x \mapsto f(\varphi(x))\,|\varphi'(x)|$ Riemann-integrierbar auf $[a,b]$ und es gilt

$$\int_c^d f(y)\,dy = \int_a^b f(\varphi(x))\,|\varphi'(x)|\,dx.$$

Diese Formel verallgemeinert sich auf den Fall des Lebesgue-Integrals und mehr als einer Raumdimension.

Als technisches Hilfsmittel benötigen wir Zerlegungen der Eins.

Definition 11.7. Sei $K \subset \mathbb{R}^n$ kompakt und $\mathscr{U} = \{U_i\}$ eine offene Überdeckung von K, d. h. $K \subseteq \bigcup_i U_i$. Eine zu dieser Überdeckung $\mathscr{U}$ gehörige, stetige *Zerlegung der Eins* über K ist eine Familie $\{\psi_i\}$ stetiger Funktionen $\psi_i \colon \mathbb{R}^n \to [0,1]$ mit kompaktem Träger[5] mit $\psi_i = 0$ für fast alle i und den folgenden Eigenschaften:

(i) $\operatorname{supp} \psi_j \subset U_i$ für alle i,

(ii) $\sum_i \psi_j(x) = 1$ für $x \in K$.

Satz 11.8. *Für jede kompakte Menge $K \subset \mathbb{R}^n$ und jeder offenen Überdeckung $\mathscr{U}$ von K existiert eine zu $\mathscr{U}$ gehörige Zerlegung der Eins über K.*

Beweis. O. B. d. A.sei die Überdeckung $= \{U_i\}_{i=1}^I$ endlich.
Für je zwei abgeschlossene achsenparallele Würfel Q, Q' mit $Q \subseteq (Q')^\circ$ existiert eine stückweise lineare[6] Funktion $\chi_{QQ'} \colon \mathbb{R}^n \to [0,1]$ mit $\operatorname{supp} \chi_{QQ'} = Q'$ und $\chi_{QQ'}\big|_Q = 1$ (siehe Abbildung 11.4).

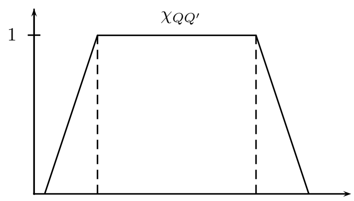

Abbildung 11.1: Eine spezielle Abschneidefunktion

Wir wählen nun für jedes $x \in K$ achsenparallele Quadrate Q_x, Q'_x mit $x \in Q_x^\circ$ und $Q_x \subset (Q'_x)^\circ \subset Q'_x \subset U_i$ für wenigstens ein $i \in \{1, \ldots, I\}$. Die offenen Mengen Q_x° für $x \in K$ überdecken K. Folglich gibt es endlich viele Stellen $x_1, x_2, \ldots, x_L \in K$ mit $K \subseteq \bigcup_{l=1}^L Q_{x_l}^\circ$. Wir setzen dann $\chi_l = \chi_{Q_{x_l} Q'_{x_l}}$ für $l = 1, \ldots, L$ und

$$\widetilde{\psi}_1 = \chi_1, \quad \widetilde{\psi}_{l+1} = (1-\chi_1)(1-\chi_2)\cdots(1-\chi_l)\psi_{l+1}$$

für $l = 1, \ldots, L-1$. Induktiv erhalten wir

$$\sum_{l=1}^{L'} \widetilde{\psi}_l = 1 - \prod_{l=1}^{L'} (1-\chi_l)$$

für $L' = 1, \ldots, L$. Insbesondere[7] folgt für $L' = L$, dass $\sum_{l=1}^L \widetilde{\psi}_l = 1$ auf K. Wählen wir nun eine Abbildung $\{1, \ldots, L\} \to \{1, \ldots, I\}$, $l \mapsto i(l)$, mit $Q'_{x_l} \subset U_{i(l)}$, so verbleibt nur

$$\psi_j = \sum_{l \colon i(l) = i} \widetilde{\psi}_l$$

[5]Ist X ein topologischer Raum und ist die Funktion $f \colon X \to \mathbb{R}$ stetig, so ist der Träger $\operatorname{supp} f = \overline{\{x \in X \mid f(x) \neq 0\}}$ von f die kleinste abgeschlossene Menge $F \subseteq X$ mit $f\big|_{X \setminus F} = 0$.

[6]Tatsächlich ist es für die folgenden Konstruktionen nur wichtig zu wissen, dass $\chi_{QQ'}$ stetig, nicht unbedingt stückweise linear ist. Wir hätten an dieser Stelle also auch den Satz von Urysohn aus der allgemeinen Topologie benutzen können.

[7]Für jedes $x \in K$ gibt es ein l mit $\chi_l(x) = 1$.

zu setzen. □

Das Transformationsverhalten des Lebesgue-Maßes unter C^1-Diffeomorphismen wird durch das folgende Resultat beschrieben:

Satz 11.9. *Seien $U, V \subseteq \mathbb{R}^n$ offene Mengen und $\varphi\colon U \to V$ eine bijektive C^1-Abbildung mit $\det D\varphi(x) \neq 0$ für alle $x \in U$. Dann gilt*

$$\varphi_* \left(|\det D\varphi(x)|\, dx\right) = dy,$$

wobei y die Variable in V bezeichnet.

Dieses Satz besagt, dass

$$\int_U f\left(\varphi(x)\right) |\det D\varphi(x)|\, dx = \int_V f(y)\, dy.$$

für jede Lebesgue-messbare Funktion $f\colon V \to \mathbb{R}$, vorausgesetzt dass eines der beiden Integrale existiert.

Folgerung 11.10. *Die Abbildung $\varphi\colon U \to V$ aus dem vorigen Satz ist genau dann* volumenerhaltend*, falls $|\det D\varphi(x)| = 1$ für alle $x \in U$ gilt.*

Beweis. Man wähle f von der Form χ_B für $B \in \bar{\mathscr{B}}(V)$. □

Bemerkung. Im Folgenden wird es wichtig sein, dass wir das Integral auf der linken Seite iterativ berechnen können. Indem wir den Integranden mit $\chi_U(x)$ multiplizieren, dürfen wir dabei annehmen, dass über ganz $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}}$ integriert wird. An dieser Stelle lässt sich dann der Satz von Fubini anwenden, beispielsweise im Fall $n = 3$ mit den parameterabhängigen Maßen $\mu_1 = dx_1$, $\mu_2^{x_1} = dx_2$ und

$$\mu_3^{x_1,x_2} = \chi_U(x_1, x_2, x_3)\, |\det D\varphi(x_1, x_2, x_3)|\, dx_3.$$

Beweis des Satzes. Wir beweisen den Satz über den Variablenwechsel unter dem Integral in fünf Schritten.

Schritt 1. Zurückführen auf stetige Integranden. Die Behauptung des Satzes ist

$$\int_{\varphi^{-1}(B)} |\det D\varphi(x)|\, dx = \int_B dy.$$

für alle $B \in \bar{\mathscr{B}}(V)$. Unter Benutzung der σ-Endlichkeit des Lebesgue-Maßes dürfen wir $\lambda_n(B) < \infty$ annehmen. Wir approximieren dann B vermöge $K_i \uparrow B$ durch eine Folge $\{K_i\}$ kompakter Teilmengen von V und unter Benutzung des Satzes von Urysohn sowie des Satzes von Lebesgue über die majorisierte Konvergenz jede der charakteristischen Funktionen χ_{K_i} durch eine stetige Funktion f_i mit kompaktem Träger in V, so dass $\left|\int_V f_i(y)\, dy - \lambda_n(K_i)\right| \leq 1/i$ gilt.
Damit genügt es zum Beweis des Satzes,

$$\int_U f\left(\varphi(x)\right) |\det D\varphi(x)|\, dx = \int_V f(y)\, dy.$$

für jede *stetige* Funktion $f\colon V \to \mathbb{R}$ mit *kompaktem* Träger zu zeigen.[8]

Schritt 2. Reduktion auf den eindimensionalen Fall. Wir benutzen hier den Satz über die implizite Funktion. Sei $x^0 \in U$ fixiert. O. B. d. A. dürfen wir $\frac{\partial \varphi_n}{\partial x_n}(x^0) \neq 0$ annehmen. Wir schreiben $x = (x', x_n)$ mit $x' = (x_1, \dots, x_{n-1})$. Wir können die Gleichung

$$\varphi_n(x', x_n) = y_n$$

in einer Umgebung von x^0 nach x_n auflösen,

$$x_n = h_n(x', y_n).$$

Wir betrachten σ_n definiert durch

$$\sigma_n(x) = (x', \varphi_n(x', x_n)).$$

σ_n ist eine bijektive C^1-Abbildung einer hinreichend kleinen Umgebung von x^0 auf eine Umgebung von $\big((x^0)', \varphi(x^0)\big)$. Es gilt

$$\left(\varphi \circ \sigma_n^{-1}(x', y_n)\right)_n = \varphi_n\left(x', h_n(x', y_n)\right) = y_n,$$

wohingegen für $1 \leq j \leq n-1$

$$\left(\varphi \circ \sigma_n^{-1}(x', y_n)\right)_j = \varphi_j\left(x', h_n(x', y_n)\right).$$

Folglich lässt $\varphi \circ \sigma_n^{-1}$ die n-te Koordinate invariant und wirkt als von y_n abhängige bijektive C^1-Abbildung in den restlichen Variablen $y_1, \dots, y_{n-1}$.
Da wir $\varphi = \left(\varphi \circ \sigma_n^{-1}\right) \circ \sigma_n$ schreiben können, wird φ schließlich in einer kleinen Umgebung von x^0 die Komposition endlich vieler derartiger σ_j für $j = 2, \dots, n$.

Schritt 3. Lokalisierung. Wir zeigen jetzt, dass der Satz gilt, falls wir für jedes $x \in \operatorname{supp} f$ eine Umgebung finden, so dass der Satz für stetige Funktionen mit kompaktem Träger in dieser Umgebung gilt.
Tatsächlich sei $\{U_i\}$ eine offene Überdeckung von $\operatorname{supp} f$, so dass der Satz wie angenommen in jeder der offenen Mengen U_i gilt, und sei $\{\psi_i\}$ eine zugehörige Zerlegung der Eins. Es folgt dann

$$\int_U (\psi_i f)(\varphi(x)) \left|\det D\varphi(x)\right| \, dx = \int_V (\psi_i f)(y) \, dy$$

für alle i. Indem wir über die i summieren und $\sum_i \psi_i(x) = 1$ für $x \in \operatorname{supp} f$ ausnutzen, erhalten wir die Behauptung für f.

Schritt 4. Kompositionen. Wir nehmen an, wir haben zwei Koordinatenwechsel $\varphi\colon U \to V$ und $\psi\colon V \to W$ und das Ergebnis sei für φ, ψ richtig. Dann gilt es auch für die Komposition $\psi \circ \varphi\colon U \to W$.
Tatsächlich gilt nach der Kettenregel

$$D(\psi \circ \varphi)(x) = (D\psi)(\varphi(x)) \, D\varphi(x)$$

[8] Insbesondere reudziert sich für derartige f das Lebesgue-Integral auf das Riemann-Integral, für das wir im eindimensionalen Fall die Transformationsformel bereits kennen!

also

$$\begin{aligned}\int_W f(z)\,dz &= \int_V f(\psi(y))\,|\det D\psi(y)|\ dy \\ &= \int_U f((\psi\circ\varphi)(x))\,|\det D\psi(\varphi(x))|\,|\det D\varphi(x)|\ dx \\ &= \int_U f((\psi\circ\varphi)(x))\,|\det D(\psi\circ\varphi)(x)|\ dx.\end{aligned}$$

Schritt 5. Der eindimensionale Fall. Wegen der bisherigen Vorbereitungen dürfen wir jetzt

$$\varphi(x) = (x_1, \dots, x_{n-1}, h(x))$$

mit einer C^1-Funktion h und $f\colon V \to \mathbb{R}$ stetig mit kompaktem Träger annehmen. Wegen

$$D\varphi(x) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} & \cdots & \frac{\partial h}{\partial x_{n-1}} & \frac{\partial h}{\partial x_n} \end{pmatrix}$$

gilt $\frac{\partial h}{\partial x_n} \neq 0$. Folglich ist die Abbildung $x_n \mapsto h(x_1, \dots, x_{n-1}, x_n)$ bei festem $(x_1, \dots, x_{n-1})$ strikt monoton (wachsend oder fallend).
In jedem Fall gilt

$$\begin{aligned}&\int_U f(\varphi(x))\,|\det D\varphi(x)|\ dx \\ &\quad= \int_U f(x_1, \dots, x_{n-1}, h(x))\left|\frac{\partial h}{\partial x_n}(x)\right|\ dx \\ &\quad= \int_{-\infty}^{\infty} \dots \left(\int_{-\infty}^{\infty} \chi_U(x)\, f(x_1, \dots, x_{n-1}, h(x))\left|\frac{\partial h}{\partial x_n}(x)\right|\ dx_n\right) \dots dx_1 \\ &\quad= \int_{-\infty}^{\infty} \dots \left(\int_{-\infty}^{\infty} \chi_V(y) f(y_1, \dots, y_n)\,dy_n\right) \dots dy_1 = \int_V f(y)\,dy.\end{aligned}$$

Das beendet den Beweis. □

Bemerkung. Hier ist eine *heuristische* Erklärung für die Gültigkeit der Transformationsformel: Jedes $D\varphi(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ lässt sich in der Form $D\varphi(x) = RU$ schreiben, wobei $R \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ mit $RR^T = I$ eine Rotation (eigentlich im Fall $\det D\varphi(x) > 0$, orientierungsumkehrend im $\det D\varphi(x) < 0$) und $U \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ mit $U = U^T$ symmetrisch ist. Die infinitesimale Rotation an der Stelle x ändert das n-dimensionale Volumen nicht[9], während U für eine infinitesimale Volumenänderung an der Stelle x durch Streckung (bzw. Stauchung) in Richtung der Hauptachsen von U um die Beträge der zugehörigen Eigenwerte verantwortlich ist.

Beispiel 11.11. Als Beispiel denken wir uns die obere offene Halbsphäre

$$\mathbb{S}^{n-1}_+ = \{(x', x_n) \in \mathbb{R}^n \mid |x'|^2 + x_n^2 = 1,\ x_n > 0\}$$

[9] Das Lebesgue-Maß ist invariant unter Verschiebungen und Rotationen.

vermöge der Gleichung $x_n = \sqrt{1-|x'|^2}$ durch $y = x'$ mit $|y| < 1$ parametrisiert. Benutzen wir zusätzlich die radiale Koordinate $r = |x|$, so erhalten wir neue Koordinaten (r, y) in $\mathbb{R}^n_+$ vermöge[10]

$$\varphi\colon \mathbb{R}_+ \times \operatorname{int}\mathbb{B}^{n-1} \to \mathbb{R}^n, \quad (r,y) \mapsto \left(ry, r\sqrt{1-|y|^2}\right).$$

Wir berechnen[11]

$$D\varphi(y) = \begin{pmatrix} y & r\,I_{n-1} \\ \sqrt{1-|y|^2} & -\frac{ry}{\sqrt{1-|y|^2}} \end{pmatrix}$$

und durch Entwicklung nach der letzten Zeile

$$\det D\varphi(y) = (-1)^{n-1} r^{n-1} \left(\sqrt{1-|y|^2} + \frac{|y|^2}{\sqrt{1-|y|^2}} \right).$$

Folglich

$$|\det D\varphi(y)| = \frac{r^{-1}}{\sqrt{1-|y|^2}}$$

und

$$\int_{\mathbb{R}^n_+} f(x)\,dx = \int_{\mathbb{R}_+ \times \operatorname{int}\mathbb{B}^{n-1}} (\varphi^* f)(r,y)\, \frac{r^{n-1}\,dr dy}{\sqrt{1-|y|^2}}$$

für jede integrierbare Funktion f. Später werden wir (r, y) als (verallgemeinerte) Polarkoordinaten interpretieren.

11.4 Untermannigfaltigkeiten des $\mathbb{R}^n$

Wir beginnen die Definition der Integration über Untermannigfaltigkeiten des $\mathbb{R}^n$ mit einer Diskussion letzteren Begriffs.

Definition 11.12. Sei $M \subseteq \mathbb{R}^n$ eine nichtleere Teilmenge, $d \in \{0, \ldots, n\}$ und $k \in \mathbb{N} \cup \{\infty\}$. Dann heißt M eine C^k-*Untermannigfaltigkeit* des $\mathbb{R}^n$ der Dimension d, falls für jedes $x^0 \in M$ eine Umgebung $U \subseteq \mathbb{R}^n$ von x^0 und eine C^k-Abbildung f einer offenen Teilmenge $V \subseteq \mathbb{R}^d$ nach $\mathbb{R}^{n-d}$ existieren, so dass

$$M \cap U = \big\{ (y, f(y)) \bigm| y \in V \big\}.$$

Hierbei ist möglicherweise die Reihenfolge der Koordinaten $(x_1, x_2, \ldots, x_n)$ im $\mathbb{R}^n$ zu ändern.

Beispiel 11.13. (i) $d = 1$: C^k-Kurven.

(ii) $d = 2$: C^k-Flächen.

(iii) $d = n - 1$: C^k-Hyperflächen. So ist $\mathbb{S}^{n-1} = \big\{ x \in \mathbb{R}^n \bigm| |x| = 1 \big\}$ eine C^∞-Hyperfläche des $\mathbb{R}^n$. Um dies zu sehen, wählen wir eine beliebige Stelle

[10] $\operatorname{int}\mathbb{B}^{n-1}$ bezeichnet das Innere von $\mathbb{B}^{n-1}$.

[11] Wir betrachten y als einen Spaltenvektor.

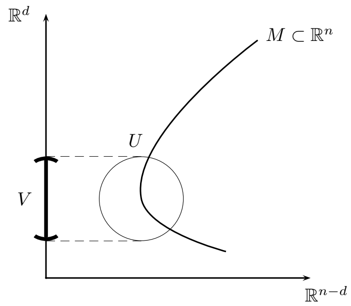

Abbildung 11.2: Zur Definition einer C^k-Untermannigfaltigkeit

$x^0 \in \mathbb{S}^{n-1}$, beispielsweise $x^0 = (0, \dots, 0, 1)$. Mit $U = \{x \in \mathbb{R}^n \mid x_n > 0\}$ gilt dann

$$\begin{aligned}\mathbb{S}^{n-1} \cap U &= \{x \in \mathbb{R}^n \mid |x| = 1,\, x_n > 0\} \\ &= \left\{\left(x', \sqrt{1 - |x'|^2}\right) \mid x' \in \mathbb{R}^{n-1},\, |x'| < 1\right\},\end{aligned}$$

wobei die Abbildung $f\colon V \to \mathbb{R}$ mit $V = \{x' \in \mathbb{R}^{n-1} \mid |x'| < 1\}$, $f(x') = \sqrt{1 - |x'|^2}$ von der Klasse C^∞ ist.

(iv) $d = n$: Offene Teilmengen im $\mathbb{R}^n$ sind C^∞-Untermannigfaltigkeiten des $\mathbb{R}^n$.

Lemma 11.14. *Ist die Abbildung $f\colon V \to \mathbb{R}^{n-d}$ wie in obiger Definition, so ist die Abbildung $h\colon V \to \mathbb{R}^n$,*

$$h(y) = (y, f(y))$$

injektiv, die Umkehrabbildung $h^{-1}\colon h(V) \to V$ ist stetig und $Dh(y) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ ist injektiv für alle $y \in V$.

Beweis. h ist offenbar injektiv und die Umkehrabbildung $h^{-1}\colon h(V) \to V$, $(y, f(y)) \mapsto y$, ist stetig als Projektion auf den ersten Eintrag. Weiter erhalten wir

$$Dh(y) = \begin{pmatrix} I_d \\ Df(y) \end{pmatrix},$$

wobei I_d die $d \times d$-Einheitsmatrix ist, so dass $Dh(y)$ für $y \in V$ injektiv ist. □

Definition 11.15. (i) Sei $D \subseteq \mathbb{R}^d$ offen, nichtleer, $d < n$. Eine C^k-Abbildung $\varphi\colon D \to \mathbb{R}^n$ mit $k \in \mathbb{N} \cup \{\infty\}$ heißt eine C^k-*Immersion*, falls $D\varphi(y) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ für alle $y \in D$ injektiv ist.

(ii) Eine C^k-Immersion $\varphi\colon D \to \mathbb{R}^n$ heißt eine C^k-*Einbettung*, falls φ injektiv und die Umkehrabbildung $\varphi^{-1}\colon \varphi(D) \to D$ stetig ist.

Beispiel 11.16. (i) Für jedes $r > 0$ ist die Abbildung

$$\phi\colon \mathbb{R} \to \mathbb{R}^2, \quad \phi(\theta) = (r\cos\theta, r\sin\theta)$$

eine C^∞-Immersion. Das Bild ist die Kreislinie $\{x \in \mathbb{R}^2 \mid |x| = r\}$ mit Mittelpunkt 0 und Radius r. Die Abbildung $\varphi\big|_{(0,2\pi)}\colon (0, 2\pi) \to \mathbb{R}$ ist eine C^∞-Einbettung.

(ii) Die Abbildung $\psi\colon \mathbb{R} \to \mathbb{R}^2$,

$$\psi(t) = (t^2 - 1, t^3 - t)$$

ist eine C^∞-Immersion. Es gilt $\psi(t) = \psi(s)$ genau dann, wenn $t = s$ oder $t, s \in \{+1, -1\}$. Die Abbildung ψ ist injektiv auf $(-\infty, 1)$, jedoch ist die Umkehrabbildung $\psi^{-1}\colon \psi((-\infty, 1)) \to (-\infty, 1)$ nicht stetig an $(0, 0)$. Andernfalls gäbe es ein $\delta > 0$ mit

$$|t + 1| = \left|t - \psi^{-1}(0, 0)\right| \leq \frac{1}{2}$$

für $t < 1$, $|\psi(t) - (0, 0)| \leq \delta$, was wegen $\lim\limits_{t \to 1-0} \psi(t) = (0, 0)$ nicht der Fall ist.

Lokal in der Nähe eines jeden Punktes ist eine Immersion jedoch eine Einbettung:

Satz 11.17. *Sei $V \subseteq \mathbb{R}^d$ offen, $d \leq n$, $k \in \mathbb{N} \cup \{\infty\}$ und sei $\varphi\colon V \to \mathbb{R}^n$ eine C^k-Abbildung mit $D\varphi(y^0)\colon \mathbb{R}^d \to \mathbb{R}^n$ injektiv für ein $y^0 \in V$. Dann gibt es eine Umgebung $D \subseteq V$ von y^0 mit:*

(i) *$\varphi(D)$ ist eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$.*

(ii) *Es gibt eine offene Umgebung $U \subseteq \mathbb{R}^n$ von $x^0 = \varphi(y^0)$ und eine bijektive C^k-Abbildung $\Phi\colon U \to \Phi(U) \subseteq \mathbb{R}^n$, so dass für alle $y \in D$*

$$\Phi \circ \varphi(y) = (y, 0) \in \mathbb{R}^d \times \mathbb{R}^{n-d}.$$

Insbesondere ist $\varphi\big|_D\colon D \to \mathbb{R}^n$ injektiv.

Beweis. (i) $D\varphi(y^0)\colon \mathbb{R}^d \to \mathbb{R}^n$ hat den Rang d, insbesondere gibt es in der Matrix

$$\begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1} & \cdots & \frac{\partial \varphi_1}{\partial y_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial y_1} & \cdots & \frac{\partial \varphi_n}{\partial y_d} \end{pmatrix} (y^0)$$

d Zeilen, die linear unabhängig sind. O. B. d. A. seien dies die ersten d Zeilen. Wir definieren

$$\begin{aligned} g &= (\varphi_1, \ldots, \varphi_d)\colon V \to \mathbb{R}^d, \\ h &= (\varphi_{d+1}, \ldots, \varphi_n)\colon V \to \mathbb{R}^{n-d}. \end{aligned}$$

Dann ist $Dg(y^0) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ invertierbar. Folglich gibt es eine offene Umgebung $D \subseteq V$ von y^0 und eine offene Umgebung $W \subseteq \mathbb{R}^d$ von $g(y^0)$, so dass $g\big|_D\colon D \to W$ bijektiv ist. Wir nehmen $w = g(y)$ als neue unabhängige Variable. Dann

$$\varphi(D) = \left\{\big(w, f(w)\big) \;\middle|\; w \in W\right\}$$

mit $W \subseteq \mathbb{R}^d$ offen und $f = h \circ g^{-1}\colon W \to \mathbb{R}^{n-d}$ eine C^k-Abbildung.

(ii) Wir definieren

$$\begin{aligned} &\psi\colon D \times \mathbb{R}^{n-d} \to W \times \mathbb{R}^{n-d}, \\ &\psi(y, z) = \varphi(y) + (0, z) = (g(y), h(y) + z)\,. \end{aligned}$$

Aus (i) folgt, dass ψ eine C^k-Inverse Φ besitzt, nämlich $\Phi(t,w) = \big(g^{-1}(w), t - f(w)\big)$. Wir setzen dann $U = W \times \mathbb{R}^{n-d}$ und erhalten

$$\Phi \circ \varphi(y) = \Phi(g(y), h(y)) = \big(g^{-1}(g(y)), h(y) - f(g(y))\big) = (y, 0)$$

für $y \in D$. □

Folgerung 11.18. *Sei $D \subseteq \mathbb{R}^d$ offen, $d \leq n$, $k \in \mathbb{N} \cup \{\infty\}$. Sei $\varphi\colon D \to \mathbb{R}^n$ eine C^k-Einbettung. Dann ist $\varphi(D)$ eine d-dimensionale C^k-Untermannigfaltigkeit von $\mathbb{R}^n$.*

Wir werden auch folgendes Resultat benötigen:

Lemma 11.19. *Sei M eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Seien $D, E \subseteq \mathbb{R}^d$ offen und $\varphi\colon D \to \mathbb{R}^n$ und $\psi\colon E \to \mathbb{R}^n$ C^k-Einbettungen mit $\varphi(D) \subseteq M$, $\psi(E) \subseteq M$ und $\varphi(D) \cap \psi(E) \neq \emptyset$. Dann sind die Mengen $\varphi(D)$, $\psi(E)$ offen in M und die Abbildung*

$$\psi^{-1} \circ \varphi\colon \varphi^{-1}(\varphi(D) \cap \psi(E)) \to \psi^{-1}(\varphi(D) \cap \psi(E))$$

ist eine C^k-Bijektion.

Beweis. Wir müssen zeigen, dass $\varphi(D)$ (und analog $\psi(E)$) offen in M ist. Sei dazu $x^0 \in \varphi(D) \subseteq M$. Gemäß (des Beweises von) (ii) des vorigen Satzes existieren eine offene Umgebung $U \subseteq \mathbb{R}^n$ von x^0 und ein C^k-Diffeomorphismus $\Phi\colon U \to \Phi(U) \subseteq \mathbb{R}^n$ mit insbesondere $M \cap U \subseteq \varphi(D)$ und der weiteren in (ii) genannten Eigenschaft. Jedoch ist $M \cap U$ eine offene Umgebung von x^0 in M, was die Offenheit der Menge $\varphi(D)$ in M zeigt. □

Satz 11.20. *Sei $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$. M ist genau dann eine d-dimensionale C^k-Untermannigfaltigkeit von M, wenn es C^k-Einbettungen $\varphi_i\colon D_i \to \mathbb{R}^n$ mit $D_i \subseteq \mathbb{R}^d$ offen, $\varphi_i(D_i) \subseteq M$ gibt, so dass die $U_i = \varphi_i(D_i)$ eine offene Überdeckung von M bilden.*

Beweis. ($\Longrightarrow$) Dies folgt direkt aus der Definition einer C^k-Untermannigfaltigkeit und Lemma 11.14.

($\Longleftarrow$) Jede der Mengen U_i ist eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Da diese Mengen M überdecken und da der Begriff einer C^k-Untermannigfaltigkeit durch eine lokale Eigenschaft definiert ist, folgt die Behauptung. □

Definition 11.21. Die $\big(U_i, \varphi_i^{-1}\big)$ wie im vorigen Satz heißen *Koordinatenumgebungen* auf M. Die vermöge φ_i^{-1} von $D_i \subseteq \mathbb{R}^d$ nach M transportierten Euklidischen Koordinaten nennt man *lokale Koordinaten* auf M (in U_i). Die Abbildungen

$$\varphi_j^{-1} \circ \varphi_i\colon \varphi_i^{-1}(U_i \cap U_j) \to \varphi_j^{-1}(U_i \cap U_j)$$

zwischen offenen Teilmengen des $\mathbb{R}^d$ (für $U_i \cap U_j \neq \emptyset$) heißen *lokale Koordinatenwechsel* auf M.

Um eine Integrationstheorie auf Untermannigfaltigkeiten aufzubauen, benötigen wir auch:

Satz 11.22. *Sei $\{U_i\}$ eine offene Überdeckung von M. Dann gibt es eine zugehörige Zerlegung der Eins, d. h. stetige Funktionen $\psi_i \colon M \to [0,1]$, von denen nur abzählbar viele verschieden Null sind, so dass $\operatorname{supp}\psi_i \subseteq U_i$ für alle i, die Familie $\{\operatorname{supp}\psi_i\}$ lokal endlich ist und*[12,13]

$$\sum_i \psi_i(x) = 1$$

für alle $x \in M$ gilt.

11.5 Integration über Untermannigfaltigkeiten

Sei M eine d-dimensionale C^1-Untermannigfaltigkeit des $\mathbb{R}^n$. Wir möchten jetzt Funktionen über M integrieren. Wir betrachten dazu zuerst das Problem lokal und denken uns M nahe $x^0 \in M$ vermöge einer C^1-Einbettung $\varphi \colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^d$ offen, $\varphi(y^0) = x^0$ für ein $y^0 \in D$ und $\varphi(D) \subseteq M$ parametrisiert. Sei weiterhin $f \colon M \to \mathbb{R}$ mit $\operatorname{supp} f \subseteq \varphi(D)$. Ist dann $f \circ \varphi$ integrierbar über D, so können wir das Integral

$$I_\varphi(f) = \int_D (f \circ \varphi)(y)\, dy$$

betrachten. Ist jedoch $\psi \colon E \to \mathbb{R}^n$ eine weitere C^1-Einbettung wie oben mit $\operatorname{supp} f \subseteq \varphi(D) \cap \psi(E)$, so erhalten wir

$$I_\psi(f) = \int_E (f \circ \psi)(z)\, dz = \int_D (f \circ \varphi)(y) \left|\det D\left(\psi^{-1} \circ \varphi\right)(y)\right| dy,$$

und letzteres ist im Allgemeinen verschieden von $I_\varphi(f)$ (außer im Spezialfall eines Koordinatenwechsels, für den $\left|\det D\left(\psi^{-1} \circ \varphi\right)(y)\right| = 1$ für alle $y \in D$ gilt). Auf diese Weise erhalten wir also *keinen invarianten* Integralbegriff.

Wir erhalten jedoch einen invarianten Integralbegriff, falls wir definieren:

Definition 11.23. Sei $\{U_i\}$ eine offene Überdeckung von M mit Koordinatenumgebungen $\left(U_i, \varphi_i^{-1}\right)$ und $\{\psi_i\}$ eine zugehörige Zerlegung der Eins. Weiter sei $A \subseteq M$.

(i) Wir nennen A *Lebesgue-messbar*, falls $\varphi_i^{-1}(A \cap U_i) \subseteq D_i$ für alle i Lebesgue-messbar ist.

(ii) Ist A Lebesgue-messbar, so setzen wir[14]

$$\lambda_d^M(A) = \sum_i \int_{\varphi_i^{-1}(A \cap U_i)} \psi_i(\varphi_i(y)) \sqrt{\det\left((D\varphi_i)^T(y)(D\varphi_i)(y)\right)}\, dy.$$

[12] Lokal endlich bedeutet, dass jede Stelle $x \in M$ eine Umgebung $V \subseteq M$ besitzt, so dass $\operatorname{supp}\psi_i \cap V$ für nur endlich viele i nichtleer ist.

[13] Für jedes $x \in M$ ist die Summe $\sum_i \psi_i(x)$ endlich.

[14] Die Bezeichnung $\operatorname{vol}_M(A)$ ist auch gebräuchlich.

Bezeichne $\bar{\mathscr{B}}(M)$ die Menge der Lebesgue-messbaren Teilmengen von M.

Lemma 11.24. *$\bar{\mathscr{B}}(M)$ ist eine σ-Algebra, die die Borelsche σ-Algebra $\mathscr{B}(M)$ enthält.*

Beweis. Dies ist offensichtlich. □

Beispiel 11.25. Wir betrachten wiederum das Beispiel der oberen offenen Halbsphäre $\mathbb{S}_+^{n-1}$. Es gilt $\mathbb{S}_+^{n-1} = \varphi(\operatorname{int}\mathbb{B}^{n-1})$, wobei

$$\varphi\colon \operatorname{int}\mathbb{B}^{n-1} \to \mathbb{R}^n, \quad y \mapsto \left(y, \sqrt{1-|y|^2}\right).$$

Wir berechnen unmittelbar, dass

$$(D\varphi)(y) = \begin{pmatrix} I_{n-1} \\ -\frac{y^T}{\sqrt{1-|y|^2}} \end{pmatrix},$$

also

$$\begin{aligned}(D\varphi)(y)^T(D\varphi)(y) &= \left(\delta_{jk} + \frac{y_j y_k}{1-|y|^2}\right)_{1\le j,\, k\le n-1} \\ &= I_{n-1} + \frac{y}{\sqrt{1-|y|^2}} \cdot \frac{y^T}{\sqrt{1-|y|^2}}\end{aligned}$$

und damit[15]

$$\sqrt{\det\left((D\varphi)(y)^T(D\varphi)(y)\right)} = \sqrt{1+\frac{|y|^2}{1-|y|^2}} = \frac{1}{\sqrt{1-|y|^2}}.$$

Das ergibt genau das Hyperflächenmaß auf $\mathbb{S}^{n-1}$, das wir bereits am Ende von Abschnitt 11.3 hergeleitet hatten.

Wir werden im Allgemeinen zeigen, dass λ_d^M ein wohldefiniertes Maß auf $\bar{\mathscr{B}}(M)$ ist.

Für eine lineare Abbildung $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ ist $A^TA \in \mathcal{L}(\mathbb{R}^d)$ und $\det\left(A^TA\right) \ge 0$ wegen $\left\langle A^TAv, v\right\rangle = \|Av\|^2 \ge 0$ für $v \in \mathbb{R}^d$.

Definition 11.26. Wir definieren $\omega\colon \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n) \to \mathbb{R}$ durch

$$\omega(A) = \sqrt{\det\left(A^TA\right)}.$$

Lemma 11.27. *Die Funktion ω hat folgende Eigenschaften:*

(a) *Es gilt $\omega(A) > 0$ genau dann, wenn A injektiv ist.*

(b) $\omega(AB) = |\det B|\,\omega(A)$ *für $B \in \mathcal{L}(\mathbb{R}^d)$,*

(c) $\omega(CA) = \omega(A)$ *für $C \in \mathrm{O}(\mathbb{R}^n)$.*

[15] Ist die Abbildung A von der Form $I + v\cdot v^T$ für ein $v \in \mathbb{R}^{n-1}$, so ist $Av = \left(1+|v|^2\right)v$ und $Aw = 0$ für jedes w mit $v^T\cdot w = 0$. Folglich $\det A = 1 + |v|^2$.

Beweis. a) Obiges Argument zeigt $\langle A^T Av, v\rangle = \|Av\|^2 > 0$ für alle $v \in \mathbb{R}^d \setminus \{0\}$ genau dann, wenn A injektiv ist.

b) Es gilt

$$\begin{aligned}\omega(AB)^2 &= \det\bigl((AB)^T(AB)\bigr) = \det\left(B^T(A^TA)B\right)\\ &= \det B^T \det(A^TA)\det B = (\det B)^2\omega(A)^2.\end{aligned}$$

c) Es gilt $C^TC = I$ und

$$\begin{aligned}\omega(CA)^2 &= \det\bigl((CA)^TCA\bigr) = \det\left(A^T(C^TC)A\right)\\ &= \det(A^TA) = \omega(A)^2.\end{aligned}$$

□

Für eine C^1-Einbettung $\varphi\colon D \to \mathbb{R}^n$, $D \subseteq \mathbb{R}^d$ offen, $\varphi(D) \subseteq M$ setzen wir jetzt für $y \in D$

$$\omega_\varphi(y) = \omega\left(D\varphi(y)\right) = \sqrt{\det\bigl(D\varphi(y)^T D\varphi(y)\bigr)}.$$

Lemma 11.28. *Seien $\varphi\colon D \to \mathbb{R}^n$, $\psi\colon E \to \mathbb{R}^n$ zwei C^1-Einbettungen wie oben mit $\varphi(D) \cap \psi(E) \neq \emptyset$. Dann gilt für $y \in \varphi^{-1}\left(\varphi(D) \cap \psi(E)\right)$*

$$\omega_\varphi(y) = \left|\det D\left(\psi^{-1} \circ \varphi\right)(y)\right| \omega_\psi(z),$$

wobei $z = \psi^{-1} \circ \varphi(y)$.

Beweis. Mit der Kettenregel gilt

$$D\varphi(y) = D\psi(z) D\left(\psi^{-1} \circ \varphi\right)(y),$$

und es bleibt, obiges Resultat mit $A = D\psi(z)$, $B = D\left(\psi^{-1} \circ \varphi\right)(y)$ anzuwenden. □

Satz 11.29. *Das oben definierte Maß λ_d^M auf $\bar{\mathscr{B}}(M)$ ist wohldefiniert.*

Beweis. Sei $\{(U_l', (\varphi_l')^{-1}\}$ eine weitere offene Überdeckung von M durch Koordinatenumgebungen, wobei $\varphi_l'\colon E_l \to \mathbb{R}^n$ mit $E_l \subseteq \mathbb{R}^d$ offen, und $\{\psi_l'\}$ eine zugehörige Zerlegung der Eins. Dann gilt für jede stetige Funktion $f\colon M \to \mathbb{R}$

$$\begin{aligned}&\sum_i \int_{D_i} (\psi_i f)(\varphi_i(y))\,\omega_{\varphi_i}(y)\,dy\\ &\qquad= \sum_i \int_{D_i} \sum_l \psi_l'(\varphi_j(y))\psi_i(\varphi_i(y))f(\varphi_i(y))\,\omega_{\varphi_i}(y)\,dy\\ &\qquad= \sum_{i,l} \int_{E_l} \psi_l'(\varphi_l'(z))\psi_i(\varphi_l'(z))f(\varphi_l'(z))\,\omega_{\varphi_i}((\varphi_i^{-1} \circ \varphi_l')(z))\\ &\qquad\qquad \times \left|\det D(\varphi_i^{-1} \circ \varphi_l')(z)\right|dz\\ &\qquad= \sum_l \int_{E_l} \sum_i \psi_i(\varphi_l'(z))(\psi_l' f)(\psi_l(z))\,\omega_{\varphi_l'}(z)\,dz\\ &\qquad= \sum_l \int_{E_l} (\psi_l' f)(z)\,\omega_{\varphi_l'}(z)\,dz,\end{aligned}$$

was die Unabhängigkeit der Definition von λ_d^M von den getroffenen Wahlen zeigt. □

Bemerkung. Statt $\int_M f\, d\lambda_d^M$ werden wir auch $\int_M f\, dM$ schreiben.

Beispiel 11.30 (*Verallgemeinerte Polarkoordinaten* auf $\mathbb{R}^n$). (i) Die Rechnungen oben zeigen, dass $d\mathbb{S}^{n-1} = \frac{dy}{\sqrt{1-|y|^2}}$ auf $\mathbb{S}^{n-1} \cap \mathbb{R}^n_+$ parametrisiert durch $(x', x_n) = (y, \sqrt{1-|y|^2})$ gilt.

Weiterhin ist für jede integrierbare Funktion f auf $\mathbb{R}^n$ die Funktion $\mathbb{S}^{n-1} \to \overline{\mathbb{R}}$, $\omega \mapsto f(r\omega)$ integrierbar bezüglich $d\mathbb{S}^{n-1}$ für $r > 0$ f. ü. und es gilt

$$\int_{\mathbb{R}^n} f(x)\, dx = \int_0^\infty r^{n-1} \left(\int_{\mathbb{S}^{n-1}} f(r\omega)\, d\mathbb{S}^{n-1}(\omega) \right) dr.$$

(ii) Für $A \subseteq \bar{\mathscr{B}}(\mathbb{S}^{n-1})$ setzen wir $\hat{A} = \{r\omega \mid 0 \le r \le 1,\, \omega \in A\} \in \bar{\mathscr{B}}(\mathbb{R}^n)$ und erhalten

$$\begin{aligned} n\, \lambda_n(\hat{A}) &= n \int_{\mathbb{R}^n} \chi_{\hat{A}}(x)\, dx \\ &= n \int_0^1 r^{n-1} \left(\int_{\mathbb{S}^{n-1}} \chi_A(\omega)\, d\omega \right) dr = \lambda_{n-1}^{\mathbb{S}^{n-1}}(A). \end{aligned}$$

Insbesondere ist $n\omega_n = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$ das Volumen der Sphere $\mathbb{S}^{n-1}$.

Es gilt folgender oft nützlicher Satz:

Satz 11.31. *Sei M eine d-dimensionale C^1-Untermannigfaltigkeit des $\mathbb{R}^n$. Weiter seien $\varphi_i \colon D_i \to \mathbb{R}^n$ abzählbar viele C^1-Einbettungen mit $\varphi_i(D_i) \subseteq M$ und*

(i) $\varphi_i(D_i) \cap \varphi_j(D_j) = \emptyset$ *für* $i \neq j$,

(ii) $\lambda_d^M \left(M \setminus \bigcup_i \varphi_i(D_i)\right) = 0$.

Dann gilt

$$\int_M f\, dM = \sum_i \int_{D_i} f(\varphi_i(y))\, \omega_{\varphi_i}(y)\, dy$$

für jede bezüglich λ_d^m-integrierbare Funktion f auf M.

Bemerkung. Später werden wir ein Beispiel kennenleren, in dem die Forderung aufgegeben wird, dass M auch entlang $M \setminus \bigcup_i \varphi_i(D_i)$ eine C^1-Mannigfaltigkeit ist, man aber immer noch intergrieren kann. Diese Bemerkung ist nützlich in Situationen, in denen M Singularitäten wie Ecken und Kanten enthält, die Menge dieser Singularitäten jedoch in einem geeigneten Sinn klein ist.

11.6 Integration einer totalen Ableitung

11.6.1 Integration eines Gradienten

Sei jetzt $\Omega \subseteq \mathbb{R}^n$ eine offene beschränkte Menge. Wir nehmen an, dass $\partial\Omega$ eine C^1-Hyperfläche im $\mathbb{R}^n$ ist und dass Ω einseitig bezüglich des Randes liegt. In mathematischen Termen bedeutet dies, dass es für jede Stelle $x^0 \in \partial\Omega$ eine offene Umgebung $U \subseteq \mathbb{R}^n$ von x^0 und eine C^1-Funktion $h \colon U \to \mathbb{R}$ mit $\nabla h \neq 0$ auf U gibt, so dass

$$\Omega \cap U = \{x \in U \mid h(x) < 0\}.$$

Bemerkung. Dann gilt weiter

$$\partial\Omega \cap U = \{x \in U \mid h(x) = 0\}, \quad U \setminus \overline{\Omega} = \{x \in U \mid h(x) > 0\}.$$

Definition 11.32. Unter diesen Bedingungen definieren wir die *äußere Einheitsnormale* $n = n(x^0)$ an $\partial\Omega$ in $x^0 \in \partial\Omega$ als

$$n = \frac{\nabla h(x^0)^T}{|\nabla h(x^0)|}.$$

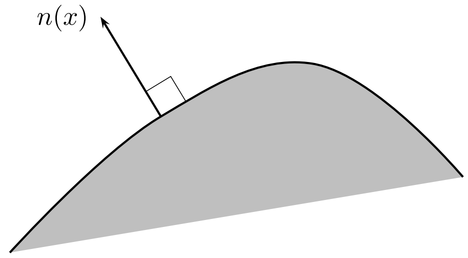

Abbildung 11.3: Die äußere Normale

Lemma 11.33. *Die äußere Einheitsnormale n ist durch folgende Bedingungen eindeutig festgelegt:*

(i) $|n| = 1$.

(ii) $x^0 + sn \notin \overline{\Omega}$ *für kleine* $s > 0$.

(iii) *Für jede C^1-Kurve $\gamma\colon (-1,1) \to \mathbb{R}^n$ mit $\gamma(0) = x^0$ und $\gamma((-1,1)) \subseteq \partial\Omega$ gilt*

$$\langle n, \gamma'(0)\rangle = 0.$$

Beweis von (iii). O. B. d. A. sei $\gamma((-1,1)) \subseteq \partial\Omega \cap U$. Dann gilt $h(\gamma(t)) = 0$, also $\nabla h(\gamma(t))\,\gamma'(t) = 0$ für alle t. An der Stelle $t = 0$ folgt $\langle n, \gamma'(0)\rangle = 0$. ☐

Satz 11.34. *Sei $\Omega \subset \mathbb{R}^n$ eine offen, beschränkte Menge mit C^1-Rand $\partial\Omega$, so dass Ω einseitig bezüglich $\partial\Omega$ liegt. Sei $f\colon \Omega \to \mathbb{R}$ eine C^1-Funktion, so dass sich f und $\nabla f\colon \Omega \to \mathbb{R}^n$ zu stetigen Funktionen auf $\overline{\Omega}$ fortsetzen lassen*[16]*. Dann gilt*

$$\int_\Omega \left(\nabla^T f\right)(x)\,dx = \int_{\partial\Omega} f(x)n(x)\,d(\partial\Omega),$$

wobei $n\colon \partial\Omega \to \mathbb{R}$ die äußere Einheitsnormale bezeichnet. In Komponenten

$$\int_\Omega \frac{\partial f}{\partial x_i}(x)\,dx = \int_{\partial\Omega} f(x)n_i(x)\,d(\partial\Omega)$$

für $i = 1, \ldots, n$.

[16]Das ist gleichbedeutend damit, dass f und ∇f gleichmäßig stetig auf Ω sind.

Bemerkung. Der Hauptsatz der Differential– und Integralrechnung, wonach

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

für eine stetig differenzierbare Funktion $f\colon [a,b] \to \mathbb{R}$ gilt, ist ein Spezialfall.

Beweis. 1. Schritt. Wir zeigen, dass es für jede Stelle $x \in \overline{\Omega}$ eine Umgebung $U \subseteq \mathbb{R}^n$ von x gibt, so dass aus $\operatorname{supp} f \subseteq U_x$ die Gültigkeit der Ausssage des Satzes für f folgt.
Fall 1. $x^0 \in \Omega$. Wir wählen ein achsenparalleles Rechteck $U = \prod_{i=1}^n (a_i, b_i)$ in Ω, das x^0 enthält. Dann gilt mit $U' = \prod_{i=1}^{n-1} (a_i, b_i)$ für $\operatorname{supp} f \subseteq U$

$$\begin{aligned}\int_\Omega \frac{\partial f}{\partial x_n}(x)\,dx &= \int_{U'} \left(\int_{a_n}^{b_n} \frac{\partial f}{\partial x_n}(x', x_n)\,dx_n \right) dx' \\ &= \int_{U'} \left(f(x', b_n) - f(x', a_n) \right) dx' = 0.\end{aligned}$$

Analog erhält man $\int_\Omega \frac{\partial f}{\partial x_i}(x)\,dx = 0$ für $i = 1, \dots, n-1$ und damit die Behauptung.

Fall 2. $x^0 \in \partial\Omega$. Wir zeigen, dass

$$\int_\Omega \nabla f(x) v\,dx = \int_{\partial\Omega} f(x) u(x) \cdot v\,d(\partial\Omega)$$

für jeden Vektor $v \in \mathbb{R}^n$ gilt. Da beide Seiten linear in v sind und sich jeder Vektor im $\mathbb{R}^n$ aus Vektoren mit $u(x^0) \cdot v > 0$ linear kombinieren lässt, genügt es, dies für alle Vektoren v mit $u(x^0) \cdot v > 0$ zu zeigen.
Sei $v \in \mathbb{R}^n$ ein derartiger Vektor. Wir koordinatisieren $\partial\Omega$ in der Nähe von x^0 vermöge einer C^1-Einbettung $\varphi\colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^{n-1}$ offen, $0 \in D$, $\varphi(0) = x^0$ und $\varphi(D) \subseteq \partial\Omega$. Wir betrachten dann

$$\Phi\colon D \times (-\epsilon, \epsilon) \to \mathbb{R}^n, \quad (y', y_n) \mapsto \varphi(y') + y_n v.$$

Wegen $\det D\Phi(0,0) = \det (D\varphi(0), v) \neq 0$ ist Φ für D und $\epsilon > 0$ hinreichend klein ein C^1-Koordinatenwechsel. Es gilt

$$(\nabla f)(\Phi(y))\, v = (\nabla f)(\varphi(y') + y_n v)\, v = \frac{\partial}{\partial y_n} \left(f(\varphi(y') + y_n v) \right).$$

Wir benötigen folgendes Resultat aus der linearen Algebra:

Lemma 11.35. *$A \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^n)$ habe den Rang $n-1$ und sei*[17] $\ker A^T = \{\lambda n \mid \lambda \in \mathbb{R}\}$ *für $n \in \mathbb{R}^n$, $|n| = 1$. Dann gilt für ein beliebiges $v \in \mathbb{R}^n$*

$$\left| \det \begin{pmatrix} A & v \end{pmatrix} \right| = |n \cdot v|\, \omega(A),$$

wobei $\omega(A)$ in Definition 11.26 *eingeführt wurde.*

[17] Beachte, dass $\dim \ker A^T = 1$ gilt.

Beweis. Wir dürfen $n = (0, \dots, 0, 1)^T$ annehmen. Dann ist $A = \binom{B}{0}$, wobei $B \in \mathcal{L}(\mathbb{R}^{n-1})$ invertierbar ist und die 0 in der letzten Zeile für den Vektor $(0, \dots, 0)$ steht, sowie $\omega(A) = |\det B|$. Mit $v = \binom{v'}{v_n}$ erhalten wir

$$\left|\det \begin{pmatrix} A & v \end{pmatrix}\right| = \left|\det \begin{pmatrix} B & v' \\ 0 & v_n \end{pmatrix}\right| = |\det B|\,|v_n| = |n \cdot v|\;\omega(A). \qquad \square$$

Indem wir dieses Lemma mit $A = D\varphi(y')$ und $n = n\,(\varphi(y'))$ anwenden, erhalten wir

$$|\det \Phi(y', 0)| = \left|\det \begin{pmatrix} D\varphi(y') & v \end{pmatrix}\right| = n(\varphi(y')) \cdot v\;\omega_\phi(y').$$

Für $\operatorname{supp} f \subseteq U = \Phi\,(D \times (-\epsilon, \epsilon))$ ergibt sich

$$\begin{aligned}
\int_\Omega \nabla f(x) v\,dx &= \int_D \int_{-\epsilon}^0 \frac{\partial}{\partial y_n}\left(f\left(\varphi(y') + y_n v\right)\right) dy_n \left|\det \begin{pmatrix} D\varphi(y') & v \end{pmatrix}\right| \,dy' \\
&= \int_D f\left(\varphi(y')\right) n(\varphi(y')) \cdot v\;\omega_\phi(y')\,dy' \\
&= \int_{\partial\Omega} f(x) n(x) \cdot v\,d(\partial\Omega).
\end{aligned}$$

2. Schritt. Sei nun $f : \Omega \to \mathbb{R}$ wie in der Formulierung des Satzes, aber ansonsten beliebig.
Sei $\{U_i\}$ ist eine endliche offene Überdeckung von $\overline{\Omega}$ mit offenen Mengen wie im 1. Schritt konstruiert. Sei $\{\chi_i\}$ eine zugehörige C^1-Zerlegung der Eins[18], d. h. die $\psi_i : \mathbb{R}^n \to [0, 1]$ sind C^1-Funktionen mit $\operatorname{supp} \psi_i \subseteq U_i$ und $\sum_i \psi_i(x) = 1$ für alle $x \in \overline{\Omega}$. Dann gilt

$$\begin{aligned}
\int_\Omega \nabla^T f(x)\,dx &= \sum_i \int_\Omega \nabla^T\left(\psi_i f\right)(x)\,dx \\
&= \sum_i \int_{\partial\Omega} \left(\psi_i f\right)(x) u(x)\,d(\partial\Omega) \\
&= \int_{\partial\Omega} f(x) u(x)\,d(\partial\Omega).
\end{aligned}$$

$\square$

11.6.2 Eine Verallgemeinerung

In Anwendungen muss man häufig über Gebiete integrieren, deren Ränder nicht C^1, aber „fast so gut“ sind (z. B. wenn Ω ein polygonales Gebiet ist). Um derartige Situationen behandeln zu können, schwächen wir die Voraussetzungen im vorigen Satz ab.

Definition 11.36. Eine Menge $S \subseteq \mathbb{R}^n$ heißt *vernachlässigbar im* $(n-1)$-*dimensionalen Sinne*, falls

$$\lambda_{n*}(S_\epsilon) = o(\epsilon) \;\text{ für } \epsilon \to +0,$$

[18] Der Beweis von Satz 11.8 läßt sich leicht dahingehend abändern, dass er eine C^1-Zerlegung der Eins produziert.

d. h. falls

$$\lim_{\epsilon \to +0} \frac{\lambda_{n*}(S_\epsilon)}{\epsilon} = 0$$

gilt. Hierbei ist $S_\epsilon = \big\{x \in \mathbb{R}^n \;\big|\; \inf\{|x-y| \mid y \in S\} \leq \epsilon\big\}$.

Beispiel 11.37. Ist S abzählbare Vereinigung von höchstens $(n-2)$-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^n$ mit endlichem Volumen, so ist S vernachlässigbar im $(n-1)$-dimensionalen Sinne.

Wir machen jetzt folgende Voraussetzungen an Ω:

(i) $\Omega \subset \mathbb{R}^n$ ist offen, beschränkt.

(ii) Es gibt abzählbar viele $(n-1)$-dimensionale Untermannigfaltigkeiten M_k des $\mathbb{R}^n$ mit $M_k \subseteq \partial\Omega$ für alle k und $M_k \cap M_l = \emptyset$ für $k \neq l$, so dass $\partial\Omega \setminus \bigcup_k M_k$ vernachlässigbar im $(n-1)$-dimensionalen Sinne ist.

(iii) Ω liegt lokal einseitig bezüglich jedes M_k.

Theorem 11.38. *Unter diesen Voraussetzungen sei $f\colon \Omega \to \mathbb{R}$ eine gleichmäßig stetige C^1-Funktion, so dass $\nabla f\colon \Omega \to \mathbb{R}^n$ integrierbar ist. Dann gilt*

$$\int_\Omega \nabla^T f(x)\, dx = \int_{\partial\Omega} f(x) n(x) d(\partial\Omega),$$

wobei die rechte Seite als $\sum_k \int_{M_k} f(x)u(x)\, dM_k$ interpretiert wird.

11.6.3 Der Satz von Gauß und Anwendungen

In den folgenden Sätzen sind Ω und alle auftretenden Funktionen hinreichend gut, damit die vorigen beiden Sätze anwendbar sind.

Satz 11.39 (Partielle Integration)**.** *Für f, $g\colon \Omega \to \mathbb{R}$ gilt*

$$\int_\Omega f(x) \frac{\partial g}{\partial x_i}(x)\, dx = \int_{\partial\Omega} f(x) g(x) n_i(x)\, d(\partial\Omega) - \int_\Omega \frac{\partial f}{\partial x_i}(x) g(x)\, dx$$

für $i = 1, \ldots, n$.

Beweis. Es gilt

$$\begin{aligned}\int_\Omega f \frac{\partial g}{\partial x_i}\, dx &= \int_\Omega \frac{\partial}{\partial x_i}(f\,g)\; dx - \int_\Omega \frac{\partial f}{\partial x_i} g\, dx \\ &= \int_{\partial\Omega} f\, g\, n_i\, d(\partial\Omega) - \int_\Omega \frac{\partial f}{\partial x_i} g\, dx.\end{aligned}$$

□

Satz 11.40 (Satz von Gauß)**.** *Sei $F\colon \Omega \to \mathbb{R}^n$ ein Vektorfeld. Dann gilt*

$$\int_\Omega \operatorname{div} F(x)\, dx = \int_{\partial\Omega} F(x) \cdot n(x)\, d(\partial\Omega).$$

Beweis. Wegen $\operatorname{div} F = \sum\limits_{i=1}^{n} \frac{\partial F_i}{\partial x_i}$ gilt

$$\begin{aligned}\int_\Omega \operatorname{div} F \, dx &= \sum_{i=1}^{n} \int_\Omega \frac{\partial F_i}{\partial x_i} \, dx = \sum_{i=1}^{n} \int_{\partial\Omega} F_i n_i \, d(\partial\Omega) \\ &= \int_{\partial\Omega} F \cdot n \, d(\partial\Omega).\end{aligned}$$

□

Satz 11.41 (1. Greensche Formel)**.** *Für $f,\, g\colon \Omega \to \mathbb{R}$ gilt*

$$\begin{aligned}\int_\Omega f(x)\,(\Delta g)\,(x)\,dx = \int_{\partial\Omega} & f(x)\frac{\partial g}{\partial n}(x)\,d(\partial\Omega) \\ & - \int_\Omega \left(\nabla^T f\right)(x) \cdot \left(\nabla^T g\right)(x)\,dx.\end{aligned}$$

Hierbei ist $\frac{\partial g}{\partial n}(x) = \nabla g(x) n$ die sogenannte Normalenableitung.

Beweis. Wegen $\Delta = \operatorname{div} \cdot \nabla^T$ gilt

$$\begin{aligned}\int_\Omega f \Delta g \, dx &= \int_\Omega \operatorname{div}\left(f \nabla^T g\right) dx - \int_\Omega \left(\nabla^T f\right) \cdot \left(\nabla^T g\right) dx \\ &= \int_{\partial\Omega} f \nabla^T g \cdot n \, d(\partial\Omega) - \int_\Omega \left(\nabla^T f\right) \cdot \left(\nabla^T g\right) dx.\end{aligned}$$

□

Satz 11.42 (2. Greensche Formel)**.** *Für $f,\, g\colon \Omega \to \mathbb{R}$ gilt*

$$\begin{aligned}&\int_\Omega \left(f(x)\,(\Delta g)\,(x) - (\Delta f)\,(x) g(x)\right) dx \\ &\qquad = \int_{\partial\Omega} \left(f(x)\frac{\partial g}{\partial n}(x) - \frac{\partial f}{\partial n}(x) g(x) \right) d(\partial\Omega).\end{aligned}$$

Beweis. Man vertausche die Rollen von f, g im vorigen Satz und subtrahiere die so erhaltenen Gleichungen. □

Kapitel 12

Gewöhnliche Differenzialgleichungen

In diesem Kapitel lernen wir die grundlegenden Existenzsätze und Lösungsmethoden für gewöhnliche Differenzialgleichungen kennen.

12.1 Grundlegende Begriffe

Definition 12.1. (i) Ein *System von M gewöhnlichen Differenzialgleichungen in N unbestimmten Funktionen* ist eine Gleichung der Form

$$F\bigl(x, u(x), u'(x), u''(x), \ldots, u^{(m)}(x)\bigr) = f(x), \quad x \in I. \tag{12.1}$$

Hierbei ist $I \subseteq \mathbb{R}$ ein Intervall, $u = \bigl(u_1, \ldots, u_N\bigr)^T$ ist die gesuchte Funktion,

$$u' = \frac{du}{dx}, \;\; u'' = \frac{d^2u}{dx^2}, \;\; \ldots, \;\; u^{(m)} = \frac{d^m u}{dx^m}$$

und die stetigen Funktionen

$$F\colon U \subseteq I \times \underbrace{\mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{m+1 \text{ Faktoren}} \to \mathbb{R}^M$$

und $f\colon I \to \mathbb{R}^M$ sind gegeben; m heißt die *Ordnung* des Systems.

(ii) Das System (12.1) heißt *linear*, falls F eine lineare Funktion in $u(x), u'(x)$, $\ldots, u^{(m)}$ ist, also

$$F\bigl(x, u, u', \ldots, u^{(m)}\bigr) \\ = a_0(x)u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x)u.$$

Hierbei sind die $a_j(x)$ für $j = 0, \ldots, m$ $M \times N$-Matrizen.

(iii) Hat F die Form

$$F\bigl(x, u, u', \ldots, u^{(m)}\bigr) = a\bigl(x, u, u', \ldots, u^{(m-1)}\bigr)\, u^{(m)} + b\bigl(x, u, u', \ldots, u^{(m-1)}\bigr),$$

so heißt das System (12.1) *quasilinear* (d. h. das System ist linear in der höchsten Ableitung).

(iv) Das allgemeine System (12.1) heißt *voll nichtlinear*.

(v) Eine *klassische Lösung* des Systems (12.1) ist eine Funktion $u\colon I \to \mathbb{R}^N$ der Klasse C^m mit $\big(x, u(x), u'(x), \ldots, u^{(m)}(x)\big)^T \in U$ für alle $x \in I$, so dass die Gleichung (12.1) punktweise erfüllt ist.

Bemerkung. In dieser Vorlesung werden wir uns nur für klassische Lösungen interessieren. Für gewöhnliche Differenzialgleichungen ist dies ein durchaus angemessenes Vorgehen, für partielle Differenzialgleichungen[1] wird der Lösungsbegriff zu schwachen Lösungen oder gar Distributionenlösungen zu verallgemeinern sein.

Beispiele für gewöhnliche Differenzialgleichungen

1. *Besselsche Differenzialgleichung*: $u'' + \frac{1}{x}\, u' + \left(1 - \frac{\nu^2}{x^2}\right) u = 0$, $\nu \in \mathbb{R}$ ist ein Parameter.
2. *Airysche Differenzialgleichung*: $u'' = xu$.
3. *Bernoullische Differenzialgleichung*: $u' = a(x)u + b(x)u^p$, $p \in \mathbb{N}$.
4. *Riccattische Differenzialgleichung*: $u' = a(x)u^2 + b(x)u + c(x)$.

12.2 Ein lokaler Existenzsatz

12.2.1 Funktionalanalytische Vorbereitungen

Definition 12.2. Sei V ein $\mathbb{R}$-(oder $\mathbb{C}$-)Vektorraum. Eine Funktion $\|\cdot\| : V \to [0, \infty)$ heißt eine *Norm auf* V, falls Folgendes für alle $u, v \in V$, $\alpha \in \mathbb{R}$ (oder $\alpha \in \mathbb{C}$) gilt:

(i) $\|u\| = 0$ impliziert $u = 0$,

(ii) $\|\alpha u\| = |\alpha|\, \|u\|$

(iii) $\|u + v\| \le \|u\| + \|v\|$.

Ein mit einer Norm versehener Vektorraum $(V, \|\cdot\|)$ heißt ein *normierter Raum*.

Beispiel 12.3. $\big(\mathbb{R}^n, |\cdot|_p\big)$ ist ein normierter Raum für $1 \le p \le \infty$, wobei

$$|x|_p = \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{1/p}, & 1 \le p < \infty, \\ \max\limits_{j=1,\ldots,n} |x_j|, & p = \infty. \end{cases}$$

Lemma 12.4. (i) *Ist* $(V, \|\cdot\|)$ *ein normierter Raum, so ist* $d\colon V \times V \to [0, \infty)$,

$$d(u, v) = \|u - v\|,$$

eine Metrik auf V.

[1] Das sind Differenzialgleichungen in mehr als einer unabhängigen Veränderlichen.

(ii) *Die Vervollständigung eines normierten Raumes als metrischer Raum ist in kanonischer Weise ein normierter Raum.*

Beweis. (i) Eine einfache Übungsaufgabe.

(ii) Ist $\{u_m\}$ eine Cauchyfolge in V, so existiert $\lim_{m\to\infty} \|u_m\|$. Ist $\{u'_m\}$ eine weitere Cauchyfolge mit $\lim_{m\to\infty} \|u_m - u'_m\| = 0$, so gilt $\lim_{m\to\infty} \|u_m\| = \lim_{m\to\infty} \|u'_m\|$. Auf diese Weise lässt sich eine Norm auf der Vervollständigung von V als

$$\|[\{u_m\}]\| = \lim_{m\to\infty} \|u_m\|$$

einführen, wobei $[\{u_m\}]$ die Äquivalenzklasse von $\{u_m\}$ bezeichnet. □

Definition 12.5. Ein normierter Raum $(V, \|\cdot\|)$, der als metrischer Raum vollständig ist, heißt ein *Banachraum.*

Definition 12.6. Sei $I \subset \mathbb{R}$ ein kompaktes Intervall, $m \in \mathbb{N}_0$, $n \in \mathbb{N}$. Wir definieren dann $C^m(I;\mathbb{R}^N)$ als den Raum aller C^m-Funktionen $u\colon I \to \mathbb{R}^N$. Anstelle von $C^0(I;\mathbb{R}^N)$ schreiben wir auch $C(I;\mathbb{R}^N)$.

Offenbar ist $C^m(I;\mathbb{R}^N)$ ein $\mathbb{R}$-Vektorraum und ein $C^m(I;\mathbb{R})$-Modul.

Lemma 12.7. *Der Raum $C^m(I;\mathbb{R}^N)$ versehen mit der Norm*

$$\|u\|_{C^m(I;\mathbb{R}^N)} = \sum_{j=0}^{m} \max_{x\in I} \left|u^{(j)}(x)\right|$$

ist ein Banachraum.

Beweis. Die Abbildung $u \mapsto \|u\|_{C^m(I;\mathbb{R}^N)}$ ist offenbar eine Norm auf $C^m(I;\mathbb{R}^N)$. Es bleibt zu zeigen, dass der so definierte normierte Raum vollständig ist.
Wir führen den Vollständigkeitsbeweis zuerst für $m = 0$. Sei also $\{u_k\} \subset C(I;\mathbb{R}^N)$ eine Cauchyfolge. Dann ist $\{u_k(x)\}$ für jedes $x \in I$ eine Cauchyfolge in $\mathbb{R}^N$, also existiert der (punktweise) Grenzwert

$$u(x) = \lim_{k\to\infty} u_k(x), \quad x \in I.$$

Als gleichmäßiger Limes stetiger Funktionen ist die Grenzfunktion u stetig, gehört also zu $C(I;\mathbb{R}^N)$. Damit $u_k \to u$ für $k \to \infty$ in $C(I;\mathbb{R}^N)$.
Für beliebiges m haben wir zuerst $u_k^{(j)} \to v_j$ für $k \to \infty$ in $C(I;\mathbb{R}^N)$ und jedes $0 \le j \le m$. Aus

$$u_k^{(j)}(x) = u_k^{(j)}(x^0) + \int_{x^0}^{x} u_k^{(j+1)}(x')\,dx', \quad x \in I,$$

für ein fixiertes $x^0 \in I$ und $0 \le j < m$ folgt durch Grenzübergang $k \to \infty$

$$v_j(x) = v_j(x^0) + \int_{x^0}^{x} v_{j+1}(x')\,dx', \quad x \in I,$$

also $v_{j+1} = v'_j$. Es ergibt sich $v_j = u^{(j)}$, wobei wir $u = v_0$ gesetzt haben, und $u_k \to u$ für $k \to \infty$ in $C^m(I;\mathbb{R}^N)$. □

12.2.2 Formulierung des Satzes und Beweis

Wir betrachten zuerst ein *Anfangswertproblem* für ein System erster Ordnung der Form

$$\begin{cases} u' = F(x,u), \quad x \in I, \\ u(x^0) = u^0. \end{cases} \tag{12.2}$$

Hierbei ist $u\colon I \to \mathbb{R}^N$ die gesuchte Funktion. Die Funktion F (aus einer Teilmenge) von $I \times \mathbb{R}^N$ nach $\mathbb{R}^N$, $x^0 \in I$ und $u^0 \in \mathbb{R}^N$ sind gegeben.

Annahme 1. Es existieren $a > 0$, $b > 0$, so dass F eine stetige Funktion auf

$$R = \left\{(x,u) \;\middle|\; \left|x - x^0\right| \le a,\ \left|u - u^0\right| \le b\right\}$$

ist. Zudem sei

$$\max_{(x,u)\in R} |F(x,u)| \le M.$$

Lemma 12.8. *Für jede C^1-Lösung $u = u(x)$ obigen Anfangswertproblems auf $\left|x - x^0\right| \le \min\{a, b/M\}$ gilt*

$$\left|u(x) - u^0\right| \le b.$$

Beweis. Sei $\mathcal{F}$ die Menge aller X mit $0 \le X \le \min\{a, b/M\}$, so dass die Abschätzung $\left|u(x) - u^0\right| \le b$ für $\left|x - x^0\right| \le X$ gilt. Offenbar ist $\mathcal{F}$ nichtleer[2] und abgeschlossen. Diese Menge ist aber auch offen in $[0, \min\{a, b/M\}]$, denn ist $X \in \mathcal{F}$ und gilt $0 \le X < \min\{a, b/M\}$, so gilt

$$\left|u(x) - u^0\right| = \left|\int_{x^0}^{x} F\left(x', u(x')\right)\, dx'\right| \le M\left|x - x^0\right| < b$$

für $\left|x - x^0\right| \le X$ und damit[3] $X < \sup \mathcal{F}$. Folglich gilt $\mathcal{F} = [0, \min\{a, b/M\}]$. □

Annahme 2. Für (x,u), $(x,v) \in R$ gilt

$$|F(x,u) - F(x,v)| \le L\,|u - v|$$

für eine Konstante $L > 0$.

Satz 12.9. *Unter diesen Voraussetzungen besitzt das Anfangswertproblem* (12.2) *eine eindeutige C^1-Lösung $u = u(x)$ für $\left|x - x^0\right| \le \min\{a, b/M\}$.*

Beweis. Existenz. Wir konstruieren die Lösung u durch sukzessive Approximation[4]. Dazu setzen wir $u_0(x) = u^0$ für $x \in I$, wobei $I = [x^0 - X, x^0 + X]$, $X = \min\{a, b/M\}$, und dann

$$\begin{cases} u_m'(x) = F\left(x, u_{m-1}(x)\right), \\ u_m(x^0) = u^0 \end{cases}$$

für $m = 1, 2, 3, \ldots$ Wir erhalten

$$u_m(x) = u^0 + \int_{x^0}^{x} F\left(x', u_{m-1}(x')\right)\, dx', \quad x \in I.$$

[2]Denn es gilt $0 \in \mathcal{F}$.
[3]Die Funktion u ist stetig.
[4]Diese Konstruktion heißt *Picard-Iteration* oder auch Picard-Lindelöf-Iteration.

Behauptung. Für alle $m \in \mathbb{N}$ gilt

$$|u_m(x) - u_{m-1}(x)| \leq \frac{ML^{m-1}|x - x^0|^m}{m!} \leq \frac{ML^{-1}}{m!}(LX)^m.$$

Beweis (durch Induktion über m). Für $m = 1$ gilt die Abschätzung wegen

$$|u_1(x) - u_0(x)| \leq \left|u_1(x) - u^0\right| \leq M\left|x - x^0\right|.$$

Der Induktionsschritt $m \to m + 1$ ergibt sich aus

$$\begin{aligned}
|u_{m+1}(x) - u_m(x)| &= \left|\int_{x^0}^{x} (F(x', u_m(x')) - F(x', u_{m-1}(x')))\, dx'\right| \\
&\leq L \int_{x^0}^{x} |u_m(x') - u_{m-1}(x')|\, |dx'| \\
&\leq \frac{ML^m}{m!} \int_{x^0}^{x} \left|x' - x^0\right|^m |dx'| = \frac{ML^m}{(m+1)!} \left|x - x^0\right|^{m+1}.
\end{aligned}$$

□

Aus $\sum\limits_{m=0}^{\infty} \frac{(LX)^m}{m!} < \infty$ folgt jetzt, dass die Folge $\{u_m\}$ gleichmäßig auf I gegen eine Grenzfunktion u konvergiert. Weiterhin ergibt sich $\left|u(x) - u^0\right| \leq b$ für $x \in I$ sowie

$$u(x) = u^0 + \int_{x^0}^{x} F(x', u(x'))\, dx', \quad x \in I.$$

Damit gehört u zu $C^1(I; \mathbb{R}^N)$ und löst das Anfangswertproblem.

Eindeutigkeit. Ist $v = v(x)$ eine weitere Lösung von (12.2), so gilt

$$u(x) - v(x) = \int_{x^0}^{x} (F(x', u(x')) - F(x', v(x')))\, dx',$$

also wegen $|u(x) - v(x)| \leq 2M\left|x - x^0\right|$ wie im Existenzbeweis

$$|u(x) - v(x)| \leq \frac{2ML^{-1}}{m!}(LX)^m$$

für alle m. Für $m \to \infty$ folgt $|u(x) - v(x)| = 0$ für alle $x \in I$, also $u = v$. □

Bemerkung. Tatsächlich wiederholte der vorangegangene Beweis den Beweis des Banachschen Fixpunktsatzes. (12.2) ist nämlich für Funktionen $u \in C(I; \mathbb{R}^N)$ mit der Integralgleichung

$$u(x) = u^0 + \int_{x^0}^{x} F(x', u(x'))\, dx', \quad x \in I,$$

äquivalent. Setzen wir also $X = C(I; \mathbb{R}^N)$ und betrachten wir den Operator $T\colon X \to X$,

$$Tu(x) = u^0 + \int_{x^0}^{x} F(x', u(x'))\, dx', \quad x \in I,$$

so zeigt der vorangegangene Beweis, dass

$$\|T^m u - T^m v\|_{C(I;\mathbb{R}^N)} \leq \frac{2ML^{-1}}{m!}(LX)^m \|u - v\|_{C(I;\mathbb{R}^N)}$$

für u, $v \in C(I; \mathbb{R}^N)$ und alle $m \in \mathbb{N}$ gilt. Für m hinreichend groß ist $\frac{2ML^{-1}}{m!}(LX)^m < 1$, wonach lediglich Satz 8.20 anzuwenden bleibt.

Ist

$$\begin{cases} u^{(m)} = F(x, u, u', \ldots, u^{(m-1)}), \quad x \in I, \\ u^{(j)}(x^0) = u^j, \quad 0 \le j \le m-1, \end{cases} \tag{12.3}$$

ein *System höherer Ordnung*, so lässt sich dieses System wie folgt auf ein System erster Ordnung reduzieren:
Wir setzen $U = \left(u, u', \ldots, u^{(m-1)}\right)^T$, $U^0 = (u^0, u^1, \ldots, u^{m-1})^T$. Dann

$$\begin{cases} U' = G(x, U), \\ U(x^0) = U^0, \end{cases} \tag{12.4}$$

wobei

$$G(x, U) = \begin{pmatrix} u_1 \\ \vdots \\ u_{m-1} \\ F(x, u_0, \ldots, u_{m-1}) \end{pmatrix}$$

mit $U = \left(u_0, \ldots, u_{m-1}\right)^T$.

Lemma 12.10. *Die Systeme* (12.3), (12.4) *sind äquivalent in dem Sinne, dass jede C^m-Lösung $u = u(x)$ von* (12.3) *eine C^1-Lösung $U = U(x)$ von* (12.4) *mittels der angegebenen Transformation ergibt. Ist umgekehrt $U = U(x)$ eine C^1-Lösung von* (12.4), *so ist die erste Komponente von U eine C^m-Lösung von* (12.3).

Beweis. Eine direkte Anwendung der Definitionen. □

12.2.3 Stetige Abhängigkeit vom Anfangswert

Unter den Voraussetzungen wie oben betrachten wir jetzt zusätzlich das Anfangswertproblem

$$\begin{cases} v' = F(x, v), \quad x \in I, \\ v(x^0) = v^0, \end{cases}$$

wobei v^0 hinreichend nahe an u^0 ist. Die (lokal eindeutige) Lösung dieses Problems sei $v = v(x)$.

Satz 12.11. *Für $\left|x - x^0\right| + \left|u^0 - v^0\right|/M \le \min\{a, b/M\}$ gilt*

$$|u(x) - v(x)| \le e^{L\left|x - x^0\right|}\left|u^0 - v^0\right|.$$

Beweis. Aus dem Existenz- und Eindeutigkeitssatz folgt, dass $v = v(x)$ auf dem angegebenen Intervall existiert.
Sei $h(x) = |u(x) - v(x)|$. Dann ist h eine C^1-Funktion[5], und wir erhalten

$$\begin{aligned} \frac{1}{2}\frac{d}{dx}\left(h^2(x)\right) &= \frac{1}{2}\frac{d}{dx}\left(|u(x) - v(x)|^2\right) \\ &= \langle u'(x) - v'(x), u(x) - v(x)\rangle \\ &= \langle F\left(x, u(x)\right) - F\left(x, v(x)\right), u(x) - v(x)\rangle \\ &\le L|u(x) - v(x)|^2 = L\,h^2(x). \end{aligned}$$

[5] Es gilt $h(x) > 0$ für alle x, falls $u^0 \ne v^0$.

Es folgt $h'(x)h(x) \le L\,h^2(x)$, $h'(x) \le L\,h(x)$ und schließlich

$$h(x) \le h(x^0)\,e^{L|x-x^0|}.$$

Letzteres ist die Behauptung. □

12.3 Lineare Differenzialgleichungen

Bei der Behandlung linearer Gleichungen ist es bequem, aller auftretenen Funktionen als komplexwertig anzunehmen, was wir ab sofort tun. Die Ergebnisse des vorigen Abschnittes übertragen sich wegen $\mathbb{C} \cong \mathbb{R}^2$ unmittelbar.

12.3.1 Allgemeines

Wir behandeln lineare Differenzialgleichungen der Form

$$\begin{cases} u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x)u = f(x), \quad x \in I, \\ u(x^0) = u^0, \quad u'(x^0) = u^1, \quad \ldots, \quad u^{(m-1)}(x^0) = u^{m-1}. \end{cases} \tag{12.5}$$

Hierbei ist $I \subseteq \mathbb{R}$ ein Intervall, $x^0 \in I$ und die rechte Seite $f \in C(I;\mathbb{C})$ bzw. die Anfangsdaten $u^0, u^1, \ldots, u^{m-1} \in \mathbb{C}$ sind gegeben. Die $a_j(x)$ sind stetige komplexwertige Funktionen auf I.

Lemma 12.12. (i) *Unter den genannten Bedingungen existiert eine eindeutige C^m-Lösung u* global *auf I.*

(ii) *Der Raum der Lösungen u zur homogenen Gleichung* (*d. h. wenn $f \equiv 0$*) *ist ein m-dimensionaler $\mathbb{C}$-Vektorraum. Die Anfangsdaten $u^0, u^1, \ldots, u^{m-1}$ sind lineare Koordinaten in diesem Lösungsraum.*

(iii) *Der Raum der Lösungen u zur inhomogenen Gleichung* (*d. h. mit einer beliebigen rechter Seite f*) *ist ein affiner Raum über $\mathbb{C}$ mit zugehörigem $\mathbb{C}$-Vektorraum wie unter* (ii) *beschrieben.*

Beweis. Dies folgt aus dem analogen Resultat für lineare Systeme erster Ordnung (siehe unten). □

Anstelle (12.5) betrachten wir auch $N \times N$-Systeme erster Ordnung der Form

$$\begin{cases} U' = A(x)U + F(x), \quad x \in I, \\ U(x^0) = U^0, \end{cases} \tag{12.6}$$

wobei $A \in C(I; M(N,\mathbb{C}))$, $F \in C(I;\mathbb{C}^N)$ und $U^0 \in \mathbb{C}^N$. Die Funktion $U \in C^1(I;\mathbb{C}^N)$ ist gesucht.

Die Gleichung (12.5) kann in ein System der Form (12.6) (mit $N = m$) transformiert werden, wobei

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_m & -a_{m-1} & -a_{m-2} & \cdots & -a_2 & -a_1 \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}$$

und $U^0 = \begin{pmatrix} u^0 \\ u^1 \\ \vdots \\ u^{m-1} \end{pmatrix}$ gilt.

Lemma 12.13. (i) *Unter den genannten Bedingungen besitzt das System* (12.6) *eine eindeutige Lösung* $U \in C^1(I;\mathbb{C}^N)$.

(ii) *Der Raum der Lösungen* U *zur homogenen Gleichung* (*d. h. wenn* $F \equiv 0$) *ist ein* N*-dimensionaler* $\mathbb{C}$*-Vektorraum. Das Anfangsdatum* $U^0 \in \mathbb{C}^N$ *parametrisiert diesen Lösungsraum linear.*

(iii) *Der Raum der Lösungen* U *zur inhomogenen Gleichung* (*d. h. mit einem beliebigen* F) *ist ein affiner Raum über* $\mathbb{C}$ *mit zugehörigem* $\mathbb{C}$*-Vektorraum wie unter* (ii) *beschrieben.*

Beweis. Lediglich (i) bedarf eines Arguments. Wir haben zu zeigen, dass für jedes kompakte Intervall $J \subseteq I$ mit $x^0 \in J$ die Lösung $U = U(x)$ auf J existiert. Dazu genügt es zu zeigen, dass für jedes $x^1 \in J$ die Lösung auf dem Intervall $[x^1 - \frac{1}{2\alpha}, x^1 + \frac{1}{2\alpha}] \cap J$ mit[6] $\alpha = \max_{x\in J} |A(x)|$ existiert.
Der Platz der Funktion $F(x,u)$ aus Abschnitt 12.2. wird von der Funktion $A(x)U + F(x)$ eingenommen. Damit kann $b > 0$ beliebig gewählt werden. Eine mögliche Wahl[7] für M ist $M = \alpha b + \beta$ mit $\beta = \max_{x\in J} \left|A(x)U^0 + F(x)\right|$. Die Behauptung folgt dann aus $\frac{b}{M} = \frac{b}{\alpha b+\beta} \to \frac{1}{\alpha}$ für $b \to \infty$. □

Wir betrachten jetzt simultan N Lösungen $U_1, \dots, U_N$ der homogenen Gleichung, d. h. es gilt

$$U_j' = A(x)U_j, \quad x \in I,$$

für $1 \le j \le N$, und fassen diese N Lösungen zu einer $N \times N$-Matrix

$$X(x) = \begin{pmatrix} U_1(x) & \cdots & U_N(x) \end{pmatrix}$$

zusammen. Dann löst $X = X(x)$ die Matrixgleichung

$$X' = A(x)X.$$

Die Lösungen $U_1, \dots, U_N$ sind genau dann linear unabhängig, wenn $\det X(x) \neq 0$ für ein $x \in I$ (und dann für alle $x \in I$) gilt. In diesem Fall heißt $X = X(x)$ eine *Fundamentalmatrix*.

Satz 12.14 (Liouville)**.** *Es gilt*

$$\det X(x) = \det X(x^0) \exp\left(\int_{x^0}^{x} \operatorname{tr} A(x')\,dx'\right).$$

[6] Die Matrixnorm ist $|A| = \max_{|u|\le 1} |Au|$ für $A \in M(N,\mathbb{C})$. Wir bemerken gleich hier für spätere Verwendung, dass $|AB| \le |A|\,|B|$ für $A, B \in M(N,\mathbb{C})$ gilt.

[7] Wegen $|A(x)U + F(x)| \le |A(x)| \left|U - U^0\right| + \left|A(x)U^0 + F(x)\right|$.

Beweis. Es gilt

$$\begin{aligned}\frac{d}{dx}\det X(x) &= \sum_{j=1}^{N}\det\begin{pmatrix}U_1 & \cdots & U_{j-1} & U_j' & U_{j+1} & \cdots & U_N\end{pmatrix}\\ &= \sum_{j=1}^{N}\det\begin{pmatrix}U_1 & \cdots & U_{j-1} & A(x)U_j & U_{j+1} & \cdots & U_N\end{pmatrix}.\end{aligned}$$

Bezeichnen wir für festes $x \in I$

$$F_x\colon \underbrace{\mathbb{C}^N \times \cdots \times \mathbb{C}^N}_{N\text{-mal}} \to \mathbb{C},$$

$$F_x(U_1,\ldots,U_N) = \sum_{j=1}^{N}\det\begin{pmatrix}U_1 & \cdots & U_{j-1} & A(x)U_j & U_{j+1} & \cdots & U_N\end{pmatrix},$$

so ist F_x eine alternierende Multilinearform. Der Raum dieser alternierenden Multilinearformen ist eindimensional, also

$$F_x(U_1,\ldots,U_N) = c_x \det\begin{pmatrix}U_1 & \cdots & U_N\end{pmatrix}$$

mit einer noch zu bestimmenden Konstanten $c_x \in \mathbb{C}$.

Für $U_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, $U_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}$, $\ldots$, $U_N = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$ ergibt sich jedoch leicht, dass $c_x = \operatorname{tr} A(x)$ gilt. Daher die Behauptung in der äquivalenten Form

$$\frac{d}{dx}\det X(x) = \operatorname{tr} A(x)\cdot\det X(x).$$ □

Für eine skalare Gleichung (12.5) wählt man m Lösungen $u_1 = u_1(x)$, $\ldots$, $u_m = u_m(x)$. Diese Lösungen sind linear unabhängig, falls die *Wronski–Determinante*

$$W(x) = \det\begin{pmatrix}u_1 & u_2 & \cdots & u_m\\ u_1' & u_2' & \cdots & u_m'\\ \vdots & \vdots & \ddots & \vdots\\ u_1^{(m-1)} & u_2^{(m-1)} & \cdots & u_m^{(m-1)}\end{pmatrix}$$

verschieden von Null ist. In diesem Fall gilt

$$W(x) = W(x^0)\exp\left(-\int_{x^0}^{x} a_1(x')\,dx'\right).$$

Definition 12.15. $X(x,x^0)$ bezeichnet die Fundamentalmatrix mit $X(x^0,x^0) = I_N$, wobei I_N die $N\times N$-Einheitsmatrix ist.

Kennt man die Fundamentalmatrizen $X(x,x^0)$, so kann man das inhomogene Anfangswertproblem lösen.

Satz 12.16. *Die Lösung des* AWP (12.6) *ist durch*

$$U(x) = X(x,x^0)U^0 + \int_{x^0}^{x} X(x,x')F(x')\,dx'$$

gegeben.

Beweis. Durch die angegebene Formel wird offenbar eine C^1-Funktion $U(x)$ definiert. Weiterhin ist $U(x^0) = X(x^0, x^0)U^0 = U^0$ und

$$\begin{aligned} U'(x) &= \frac{\partial X}{\partial x}(x, x^0)U^0 + \int_{x^0}^{x} \frac{\partial X}{\partial x}(x, x')F(x')\,dx' + X(x,x)F(x) \\ &= A(x)\left[X(x, x^0)U^0 + \int_{x^0}^{x} X(x, x')F(x')\,dx'\right] + F(x) \\ &= A(x)U(x) + F(x). \end{aligned}$$

Somit löst $U = U(x)$ das AWP (12.6). Der Eindeutigkeitssatz vervollständigt den Beweis. □

Bemerkung. Für später bemerken wir, daß die Funktion $X(x, x^0)$ auch nach dem zweiten Parameter stetig differenzierbar ist und dass

$$\frac{\partial X}{\partial x^0}(x, x^0) = -X(x, x^0)A(x^0)$$

gilt.

Beweis. Dies folgt durch Differentation der Beziehung $X(x, x^0)X(x^0, x) = I$,[8] also

$$\frac{\partial X}{\partial x^0}(x, x^0)X(x^0, x) + X(x, x^0)A(x^0)X(x^0, x) = 0$$

und damit $\frac{\partial X}{\partial x^0}(x, x^0) + X(x, x^0)A(x^0) = 0$. □

12.3.2 Lineare Differenzialgleichungen mit konstanten Koeffizienten

Wir betrachten jetzt das AWP (12.6) mit $A(x) = A \in M(N, \mathbb{C})$, also

$$\begin{cases} U' = AU + F(x), \quad x \in I \\ U(x^0) = U^0, \end{cases} \tag{12.7}$$

In diesem Fall lassen sich die Fundamentalmatrizen $X(x, x^0)$ explizit angeben. Dazu führen wir das Exponential e^A einer Matrix $A \in M(N, \mathbb{C})$ ein.

Definition 12.17. Wir definieren

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I_N + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots.$$

Die Reihe $\sum\limits_{k=0}^{\infty} \frac{A^k}{k!}$ konvergiert wegen

$$\sum_{k=0}^{\infty} \left|\frac{A^k}{k!}\right| \leq \sum_{k=0}^{\infty} \frac{|A|^k}{k!} = e^{|A|} < \infty$$

absolut.

[8] Allgemein gilt $X(x, x')X(x', x^0) = X(x, x^0)$ für alle x, x', $x^0 \in I$.

Satz 12.18. *Es gilt*

$$X(x,x^0) = e^{(x-x^0)A}.$$

Beweis. Wir haben $X(x^0,x^0) = e^{0I_N} = 1$ und

$$\begin{aligned}\frac{\partial X}{\partial x}(x,x^0) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{\infty}\frac{(x-x^0)^k}{k!}A^k\right)\\ &= \sum_{k=1}^{\infty}\frac{(x-x^0)^{k-1}}{(k-1)!}A^k = A\sum_{k=0}^{\infty}\frac{(x-x^0)^k}{k!}A^k = AX(x,x^0).\end{aligned}$$

□

Bemerkung. Gilt $\big[A(x),A(x')\big] = 0$ für alle x, $x' \in I$ in (12.6), wobei $[A,B] = AB - BA$ den *Kommutator* zweier Matrizen $A, B \in M(N,\mathbb{C})$ bezeichnet, so ist

$$X(x,x^0) = e^{\int_{x^0}^{x} A(x')\,dx'}.$$

Es bleibt e^{xA} für $A \in M(N,\mathbb{C})$ zu berechnen. Wir beginnen mit einem Jordanblock der Form

$$\begin{pmatrix} \mu & 1 & & 0\\ & \mu & \ddots & \\ & & \ddots & 1\\ 0 & & & \mu\end{pmatrix}, \quad \mu \in \mathbb{C}.$$

Satz 12.19. *Es gilt*

$$e^{xJ} = e^{\mu x}\begin{pmatrix} 1 & \frac{x}{1!} & \frac{x^2}{2!} & \cdots & \frac{x^{N-1}}{(N-1)!}\\ & 1 & \frac{x}{1!} & \cdots & \vdots\\ & & \ddots & \frac{x}{1!} & \frac{x^2}{2!}\\ & 0 & & 1 & \frac{x}{1!}\\ & & & & 1\end{pmatrix},$$

d. h. auf der j-ten Nebendiagonalen für $j = 0,\dots,N-1$ (oberhalb der Hauptdiagonalen) haben wir den konstanten Eintrag $\frac{x^j}{j!}$.

Beweis. Bezeichnen wir die rechte Seite mit $Y(x)$, so rechnet man leicht $Y(0) = I$ und $Y'(x) = JY(x)$ nach.
Alternative können wir auch $J^N = 0$ beobachten, dann

$$e^{xJ} = I_N + \frac{x}{1!}\,J + \frac{x^2}{2!}\,J^2 + \cdots + \frac{x^{N-1}}{(N-1)!}\,J^{N-1}.$$ □

Satz 12.20. *Ist $A = P\begin{pmatrix} J_1 & & 0\\ & \ddots & \\ 0 & & J_b\end{pmatrix}P^{-1}$ für $J_1 \in M(N_1,\mathbb{C}),\dots,J_b \in M(N_b,\mathbb{C})$ mit $N_1 + \cdots + N_b = N$ und $P \in \mathrm{Gl}(N,\mathbb{C})$ (dies kann beispielsweise die Jordanzerlegung von A sein), so gilt*

$$e^{xA} = P\begin{pmatrix} e^{xJ_1} & & 0\\ & \ddots & \\ 0 & & e^{xJ_b}\end{pmatrix}P^{-1}.$$

Beweis. Direktes Nachrechnen. □

Wir wollen nun ein entsprechendes Resultat für skalare Differenzialgleichungen m-ter Ordnung mit konstanten Koeffizienten formulieren. Wir betrachten also die Gleichung

$$u^{(m)} + a_1 u^{(m-1)} + \cdots + a_{m-1}u' + a_m u = f(x), \quad x \in I,$$

wobei $a_j \in \mathbb{C}$ für $1 \le j \le m$. Transformation auf ein System erster Ordnung führt auf die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_m & -a_{m-1} & -a_{m-2} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

Satz 12.21. *Es gilt*

$$\det(\mu I_m - A) = \mu^m + a_1\mu^{m-1} + \cdots + a_{m-1}\mu + a_m.$$

Beweis. Durch Induktion nach m.
Der Induktionsanfang $m = 1$. Dann ist $\det(\mu + a_1) = \mu + a_1$, d. h. die Behauptung gilt.
Der Induktionsschritt $m \to m + 1$. Wir entwickeln die Determinante nach der letzten Spalte und erhalten

$$\begin{aligned}
&\det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 1 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_2 & \mu + a_1 \end{pmatrix} \\
&= (\mu + a_1)\det\begin{pmatrix} \mu & 1 & \cdots & 0 & 0 \\ 0 & \mu & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mu & 1 \\ 0 & 0 & \cdots & 0 & \mu \end{pmatrix} \\
&\quad + \det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 1 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_3 & \mu + a_2 \end{pmatrix} \\
&\quad - \det\begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu & 0 \\ a_{m+1} & a_m & a_{m-1} & \cdots & a_3 & \mu \end{pmatrix} \\
&= (\mu + a_1)\mu^m + (\mu^m + a_2\mu^{m-1} + a_3\mu^{m-2} + \cdots + a_m\mu + a_{m+1}) - \mu^m \\
&= \mu^{m+1} + a_1\mu^m + a_2\mu^{m-1} + \cdots + a_m\mu + a_{m+1}.
\end{aligned}$$

□

Folgerung 12.22. *Es sei*

$$\mu^m + a_1\mu^{m-1} + \cdots + a_{m-1}\mu + a_m = \prod_{j=1}^{r} (\mu - \mu_j)^{m_j}$$

mit $\mu_j \neq \mu_{j'}$ für $j \neq j'$, $m_1 + \cdots + m_r = m$. Dann schreibt sich die allgemeine Lösung des homogenen Problems als

$$u(x) = \sum_{j=1}^{r} \sum_{k=0}^{m_{j-1}} c_{jk}\, e^{\mu_j x}\, \frac{x^k}{k!}$$

mit beliebigen Koeffizienten $c_{jk} \in \mathbb{C}$.

Beweis. Man führe die skalare Gleichung auf ein System erster Ordnung zurück. Dimensionsüberlegungen vervollständigen den Beweis. □

Beispiel 12.23. Die gewöhnliche Differenzialgleichung

$$u^{(3)} - 4u'' + 5u' - 2 = f(x), \quad x \in I,$$

hat das *charakteristische Polynom* $\mu^3 - 4\mu^2 + 5\mu - 2 = (\mu - 1)^2(\mu - 2)$. Ein Fundamentalsystem ist daher durch

$$e^x, \;\; xe^x, \;\; e^{2x}$$

gegeben.

12.3.3 Die Methode der Variation der Konstanten

Wir wollen jetzt eine *spezielle Lösung* des $N \times N$-Systems

$$U' = A(x)\, U + F(x), \quad x \in I, \tag{12.8}$$

finden. (Die allgemeine Lösung dieser Gleichung ist dann diese spezielle Lösung plus die allgemeine Lösung des homogenen Problems.)
Wir beginnen mit folgender Überlegung: Ist $X(x)$ eine Fundamentalmatrix, so ist $X(x)C$ für jedes $C \in \mathrm{Gl}(N, \mathbb{C})$ eine weitere Fundamentalmatrix. Umgekehrt lässt sich jede Fundamentalmatrix in dieser Form schreiben.

Beweis. Um letzteren Punkt einzusehen, nehmen wir eine weitere Fundamentalmatrix $Y(x)$ her und betrachten $X^{-1}(x)Y(x)$. Es gilt dann

$$\begin{aligned} \frac{d}{dx}\left(X^{-1}Y\right) &= -X^{-1}X'X^{-1}Y + X^{-1}Y' \\ &= X^{-1}\left[AXX^{-1}Y - AY\right] = 0. \end{aligned}$$

Folglich ist $X^{-1}(x)Y(x)$ konstant.
Um eine spezielle Lösung zu (12.8) zu finden, machen wir jetzt den Ansatz

$$U(x) = X(x)C(x)$$

mit $C \in C^1(I; \mathbb{C}^N)$. Dann gilt

$$U' = X'C + XC' = AXC + XC'.$$

Letzteres soll gleich $AU + F = AXC + F$ sein. Also benötigen wir $XC' = F$, $C' = X^{-1}F$ bzw.

$$C(x) = \int_{x^0}^{x} X^{-1}(x')F(x')\,dx'$$

für ein $x^0 \in I$. (Eine andere Wahl der Integrationskonstanten führt zu einer anderen speziellen Lösung.) □

Bemerkung. In den Bezeichnungen von Abschnitt 12.3.1. gilt

$$X(x, x') = X(x)X^{-1}(x').$$

Folglich ist die gerade konstruierte Lösung $U = U(x)$ zu (12.8) genau diejenige mit $U(x^0) = 0$.

Wir wollen nun obiges Schema auf *skalare lineare Gleichungen* spezifizieren. Sei also die Gleichung

$$u^{(m)} + a_1(x)u^{(m-1)} + \cdots + a_{m-1}(x)u' + a_m(x) = f(x), \quad x \in I \tag{12.9}$$

vorgelegt. Wir nehmen an, dass wir ein Fundamentalsystem dieser Gleichung, d. h. m linear unabhängige Lösungen $u_1, \ldots, u_m$ des homogenen Problems zu (12.9) kennen. Wir suchen dann eine Lösung $u = u(x)$ des inhomogenen Problems in der Form

$$u(x) = c_1(x)u_1(x) + \cdots + c_m(x)u_m(x)$$

mit zu bestimmenden Funktionen $c_1, \ldots, c_m$.
Gemäß unseres Schemas haben wir diese Funktionen $c_1, \ldots, c_m$ so zu wählen, dass

$$\sum_{k=1}^{m} c_k'(x)u_k^{(j)}(x) = 0, \quad 0 \le j \le m-2,$$
$$\sum_{k=1}^{m} c_k'(x)u_k^{(m-1)}(x) = f(x)$$

gilt. Dies ist ein lineares System in $c_1'(x), \ldots, c_m'(x)$ mit Koeffizientendeterminante $W(x) \neq 0$, wobei die Wronski-Determinante $W(x)$ bezüglich $u_1, \ldots, u_m$ gebildet wurde.
Bezeichnet $W_k(x)$ die Determinante, die man erhält, indem man in $W(x)$ die k-te Spalte $\begin{pmatrix} u_k \\ u_k' \\ \vdots \\ u_k^{(m-1)} \end{pmatrix}$ durch $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}$ ersetzt, so gilt (siehe die AGLA-Vorlesung)

$$c_k'(x) = \frac{W_k(x)}{W(x)}, \quad k = 1, \ldots, m,$$

also beispielsweise

$$c_k(x) = \int_{x^0}^{x} \frac{W_k(x')}{W(x')}\, dx', \quad k = 1, \ldots, m,$$

für ein $x^0 \in I$.

Beispiel 12.24. Wir betrachten die Gleichung

$$u'' - 3u' + 2u = e^{4x}.$$

Wegen $\mu^2 - 3\mu + 2 = (\mu - 1)(\mu - 2)$ ist ein Fundamentalsystem gegeben durch e^x, e^{2x}. Weiterhin gilt

$$W(x) = \det \begin{pmatrix} e^x & e^{2x} \\ e^x & 2e^{2x} \end{pmatrix} = e^{3x},$$

$$W_1(x) = \det \begin{pmatrix} 0 & e^{2x} \\ e^{4x} & 2e^{2x} \end{pmatrix} = -e^{6x},$$

$$W_2(x) = \det \begin{pmatrix} e^x & 0 \\ e^x & e^{4x} \end{pmatrix} = e^{5x}.$$

Wir erhalten $c_1'(x) = \frac{W_1(x)}{W(x)} = -e^{3x}$, $c_2'(x) = \frac{W_2(x)}{W(x)} = e^{2x}$,

$$c_1(x) = -\frac{1}{3}\, e^{3x}, \quad c_2(x) = \frac{1}{2}\, e^{2x}.$$

Eine spezielle Lösung ist

$$u(x) = c_1(x)e^x + c_2(x)e^{2x} = -\frac{1}{3}\, e^{4x} + \frac{1}{2}\, e^{4x} = \frac{1}{6}\, e^{4x}.$$

Die allgemeine Lösung ergibt sich zu

$$a_1 e^x + a_2 e^{2x} + \frac{1}{6}\, e^{4x}$$

mit beliebigen Konstanten $a_1, a_2 \in \mathbb{C}$.

12.4 Spezielle Lösungsmethoden

12.4.1 Die Methode der Separation der Variablen

Die Methode der Separation der Variablen arbeitet für skalare Gleichungen der Form

$$\begin{cases} u' = f(x)g(u), \quad x \in I, \\ u(x^0) = u^0 \end{cases} \tag{12.10}$$

mit f, g stetig.

Unter der Annahme $g(u^0) \neq 0$ ist

$$H(u) = \int_{u^0}^{u} \frac{du'}{g(u')}$$

für u nahe u^0 eine monotone, stetig differenzierbare Funktion. Folglich existiert die Umkehrfunktion H^{-1} nahe 0.

Satz 12.25. *Für $g(u^0) \neq 0$ ist die (lokal eindeutige) Lösung $u = u(x)$ zu (12.10) durch*

$$u(x) = H^{-1}\left(\int_{x^0}^{x} f(x')\,dx'\right)$$

gegeben.

Beweis. Sei $u = u(x)$ eine C^1-Lösung zu (12.10). Dann gilt

$$\int_{x^0}^{x} \frac{u'(x')\,dx'}{g\left(u(x')\right)} = \int_{x^0}^{x} f(x')\,dx',$$

also nach Variablenwechsel $v = u(x)$ für x nahe x^0

$$H\left(u(x)\right) = \int_{u^0}^{u(x)} \frac{dv}{g(v)} = \int_{x^0}^{x} f(x')\,dx'.$$

Ist umgekehrt $u = u(x)$ durch diese Formel gegeben, so ist u offenbar eine C^1-Funktion, $u(x^0) = u^0$ sowie

$$u'(x) = \frac{f(x)}{H'\left(H^{-1}\left(\int_{x^0}^{x} f(x')\,dx'\right)\right)} = \frac{f(x)}{H'\left(u(x)\right)} = f(x)g(u(x)).$$

Folglich ist u Lösung von (12.10). □

Bemerkung. Formal schreibt man $\frac{du}{g(u)} = f(x)\,dx$ und integriert, also

$$\int \frac{du}{g(u)} = \int f(x)\,dx.$$

Beispiel 12.26. (i) Wir betrachten die Gleichung

$$u' = px^{p-1}u^2$$

für $p \in \mathbb{N}$. Für Lösungen $u \not\equiv 0$ führt Separation der Variabalen auf $\int \frac{du}{u^2} = p\int x^{p-1}\,dx$, also $-\frac{1}{u} = x^p - k$ mit $k \in \mathbb{R}$ und

$$u(x) = \frac{1}{k - x^p}.$$

Die Integrationskonstante k bestimmt sich (beispielsweise) aus einer Anfangsbedingung $u(0) = u^0$ mit $u^0 \neq 0$ zu $k = 1/u^0$.

(ii) Wir betrachten die Gleichung

$$u' = u(1-u).$$

Für Lösungen $u \not\equiv 0$ und $u \not\equiv 1$ führt Separation der Variablen auf $\int \frac{du}{u(1-u)} = \int dx$, also wegen $\frac{1}{u(1-u)} = \frac{1}{u} + \frac{1}{1-u}$

$$\log\left|\frac{u}{1-u}\right| = \log|u| - \log|1-u| = x + k$$

mit $k \in \mathbb{R}$. Eine einfache algebraische Manipulation führt zur Lösung

$$u(x) = \frac{1}{1 + ae^{-x}}$$

mit $a \in \mathbb{R}$. ($a = 0$ ergibt die Lösung $u \equiv 1$.)

12.4.2 Exakte Differenzialgleichungen und integrierende Faktoren

Definitionen und Beispiele

Definition 12.27. (i) Eine Differenzialgleichung, die in der Form

$$\frac{d}{dx}F(x,u,u',\ldots,u^{(m-1)})=0$$

geschrieben werden kann [9], heißt *exakt*.

(ii) Wird eine Differenzialgleichung durch Multiplikation mit einem Faktor $G(x,u,u',\ldots,u^{(m)})$ exakt, so heißt $G(x,u,\ldots,u^{(m)})$ ein *integrierender Faktor*.

Die Lösungen einer exakten Differenzialgleichung sind *implizit* durch die Gleichungen

$$F(x,u,u',\ldots,u^{(m-1)})=k$$

gegeben, wobei k eine beliebige Konstante ist.

Beispiel 12.28. (i) Die Gleichung $u''+xu'+u=0$ kann als $\frac{d}{dx}(u'+xu)=0$ geschrieben werden und ist somit exakt. Die Lösungen sind

$$u(x)=\left(k_1\int_0^x e^{t^2/2}\,dt+k_2\right)e^{-x^2/2}.$$

(ii) Die Gleichung $u''+\frac{x+1}{x}\,u'+\frac{x-1}{x^2}\,u=0$ wird durch Multiplikation mit e^x exakt:

$$\frac{d}{dx}\left(e^xu'+\frac{e^x}{x}u\right)=e^xu''+e^x\,\frac{x+1}{x}\,u'+e^x\,\frac{x-1}{x^2}\,u=0.$$

Somit gilt $u'+\frac{u}{x}=-k_1\,e^{-x}$ mit der Lösung

$$u(x)=k_1\,\frac{x+1}{x}\,e^{-x}+\frac{k_2}{x}.$$

Der Fall $m=1$

Besonders interessant ist der Fall $m=1$.

Satz 12.29. *Die Differenzialgleichung* $F_0(x,u)u'+F_1(x,u)=0$ *ist genau dann exakt, wenn* [10]

$$\frac{\partial F_0}{\partial x}=\frac{\partial F_1}{\partial u}.$$

Beweis. ($\Longrightarrow$) Ist die Differenzialgleichung exakt, so gibt es ein $F=F(x,u)$ mit $F_0=\frac{\partial F}{\partial u}$, $F_1=\frac{\partial F}{\partial x}$ und die Behauptung folgt.

($\Longleftarrow$) Gilt umgekehrt $\frac{\partial F_0}{\partial u}=\frac{\partial F_1}{\partial x}$, so finden wir ein derartiges F, indem wir

$$F(x,u)=\int_{(x^0,u^0)}^{(x,u)}F_0(x',u')\,du'+F_1(x',u')\,dx'$$

setzen.[11] □

---

[9] F ist wenigstens von der Klasse C^1.

[10] Vorausgesetzt, diese partiellen Ableitungen existieren und sind stetig.

[11] Dies ist das Kurvenintegral zweiter Art entlang eines beliebigen stückweise C^1-Weges

Beispiel 12.30. (i) Die Gleichung $(u^2+x)u'+(u-x^2)=0$ ist exakt. Tatsächlich gilt

$$\frac{\partial}{\partial x}(u^2+x) = \frac{\partial}{\partial u}(u-x^2) = 1.$$

Die Lösungen u sind implizit durch

$$u^3 + 3xu - x^3 = k$$

für beliebiges $k \in \mathbb{R}$ gegeben.

(ii) Separable Gleichungen der Form $F_0(u)u' + F_1(x) = 0$ sind wegen $\frac{\partial F_0}{\partial x} = \frac{\partial F_1}{\partial u} = 0$ exakt.

(iii) Die Gleichung $(1+xu+u^2)+(1+xu+x^2)u' = 0$ ist nicht exakt. Ein integrierender Faktor ist e^{xu}. Man erhält

$$\frac{d}{dx}\big(e^{xu}(x+u)\big) = e^{xu}\left(1+xu+x^2\right)u' + e^{xu}\left(1+xu+u^2\right) = 0.$$

Folglich gilt

$$(x+u)\,e^{xu} = k.$$

Ein weiteres wichtiges Beispiel ist:

Lemma 12.31. *Ein integrierender Faktor für die skalare Gleichung*

$$u' + a(x)u = 0$$

erster Ordnung ist $\exp\left(\int_{x^0}^{x} a(x')\,dx'\right)$.

Beweis. Es gilt

$$\frac{d}{dx}\left(e^{\int_{x^0}^{x} a(x')\,dx'}u(x)\right) = e^{\int_{x^0}^{x} a(x')\,dx'}\big(u'(x)+a(x)u(x)\big),$$

was gerade die Behauptung ist. □

Folgerung 12.32. *Die Lösung u der inhomogenen Gleichung*

$$\begin{cases} u' + a(x)u = f(x), \\ \quad u(x^0) = u^0 \end{cases}$$

ist

$$u(x) = e^{-\int_{x^0}^{x} a(x')\,dx'}u^0 + \int_{x^0}^{x} e^{-\int_{x'}^{x} a(t)\,dt} f(x')\,dx'.$$

$\gamma\colon [a,b] \to \mathbb{R}^2$, der im (x',u')-Raum die Stellen (x^0,u^0) und (x,u) verbindet. Man erhält dieses Integral als

$$\int_a^b (F_1,F_0)\,(\gamma(t))\gamma'(t)\,dt.$$

Eine mögliche Wahl des Weges γ ist

$$\gamma(t) = \begin{cases} \left(x^0, u^0 + t(u-u^0)\right), & 0 \le t \le 1, \\ \left(x^0 + (t-1)(x-x^0), u\right), & 1 \le t \le 2 \end{cases}$$

(dann $a=0$, $b=2$).
Mit orientierter Integration werden wir uns im nächsten Semester in der Vorlesung „Differential- und Integralrechnung III" beschäftigen. Dann werden wir insbesondere sehen, dass obige Definition von F ein Ergebnis liefert, das von der Wahl des verbindenden Weges γ unabhängig ist (lokal für (x,u) nahe (x^0,u^0) immer und global unter einer dann zu besprechenden topologischen Voraussetzung).

Beweis. Mit der Rechnung aus dem vorigen Beweis gilt

$$\frac{d}{dx}\left(e^{\int_{x^0}^{x} a(x')\,dx'}u(x)\right) = e^{\int_{x^0}^{x} a(x')\,dx'}f(x).$$

Die Behauptung folgt durch Integration. $\square$

Diese Folgerung besagt aber nichts anderes, als dass das Fundamentalsystem[12] $v(x, x^0)$ mit $v(x^0, x^0) = 1$ durch

$$v(x, x^0) = e^{-\int_{x^0}^{x} a(x')\,dx'}$$

gegeben ist.

Bemerkung. Dieses Beispiel zeigt, dass das Finden eines integrierenden Faktors im Fall $m = 1$ zum Lösen der Differenzialgleichung äquivalent und damit genauso schwierig ist.

12.4.3 Autonome Differenzialgleichungen

Definition 12.33. Eine Differenzialgleichung auf $\mathbb{R}$, die invariant unter den Transformationen $x \mapsto x - k$ für $k \in \mathbb{R}$ ist, heißt *autonom*.

Autonome Differenzialgleichungen sind solche, die nicht explizit von der Variablen x abhängen. Sie lassen sich in der Ordnung um Eins reduzieren, indem man $v(x) = u'(x)$ als Funktion von u schreibt.[13] Man erhält

$$\begin{aligned}
u'(x) &= v(u),\\
u''(x) &= \frac{dv}{dx} = \frac{dv}{du}\cdot\frac{du}{dx} = v'(u)v(u),\\
u^{(3)}(x) &= \frac{d}{dx}\big(v'(u)v(u)\big) = v(u)(v')^2(u) + v^2(u)v''(u),
\end{aligned}$$

usw.

Beispiel 12.34. 1. Die Substitution $v(u) = u'(x)$ vereinfacht die Gleichung $u^{(3)} + uu' = 0$ nach Division[14] durch v zu

$$(v')^2 + vv'' + u = 0.$$

Die Lösung dieser Differenzialgleichung ist

$$v(u) = \pm\sqrt{k_1 + k_2 u - \frac{u^3}{3}}$$

mit k_1, $k_2 \in \mathbb{R}$. Es ergibt sich die Gleichung erster Ordnung

$$u' = \pm\sqrt{k_1 + k_2 u - \frac{u^3}{3}},$$

in der die Variablen separiert sind.[15]

[12] Im Fall einer skalaren linearen Gleichung erster Ordnung besteht ein Fundamentalsystem aus genau einer skalaren Funktion.

[13] Unter der Annahme $u'(x^0) \neq 0$ lässt sich die Abbildung $x \mapsto u(x)$ nahe x^0 lokal umkehren, d. h. u kann als neue unabhängige Variable eingeführt werden.

[14] Die Funktionen $u \equiv k_0$ für eine Konstante $k_0 \in \mathbb{R}$, d. h. $v \equiv 0$, sind Lösungen der ursprünglichen Gleichung.

[15] Die Lösung erfolgt in Termen der Weierstraßschen $\wp$-Funktion, die eventuell im Abschnitt über elliptische Funktionen in der Vorlesung „Funktionentheorie“ behandelt wird.

2. Die Gleichung
$$uu' = u''u^{(3)}$$
kann durch die Gleichung zweiter Ordnung
$$v(v')^3 + v^2v'v'' = u$$
ersetzt werden.

12.4.4 Eulergleichungen

Definition 12.35. Eine Differenzialgleichung auf $(0, \infty)$, die invariant unter den Transformationen $x \mapsto \frac{x}{k}$ mit $k > 0$ ist, heißt eine *Eulergleichung*[16].

Der Variablenwechsel $\mathbb{R} \to (0, \infty)$, $t \mapsto x = e^t$ mit
$$x\frac{d}{dx} = \frac{d}{dt},$$
$$x^2\frac{d}{dx^2} = \frac{d^2}{dt^2} - \frac{d}{dt},$$
usw. überführt eine solche Gleichung in eine autonome Gleichung.

Beispiel 12.36. Die Gleichung $u'' = \frac{uu'}{x}$ ist eine Eulergleichung. Indem wir $x^2u'' = u(xu')$ schreiben, erhalten wir nach dem Variablenwechsel $x = e^t$
$$u''(t) - u'(t) = u(t)u'(t).$$
Die Substitution $v(u) = u'(t)$ reduziert diese Gleichung zu
$$vv' - v = uv.$$
Es folgt $v = 0$, also $u(x) = k_0$, oder $v'(u) = u + 1$, also $v(u) = \frac{u^2}{2} + u + k_1$. Aus den Beziehungen $v(u) = u(t)$ und $x = e^t$ sehen wir schließlich, dass wir die separable Gleichung
$$u'(t) = \frac{u^2(t)}{2} + u(t) + k_1$$
bezüglich $u(x)$ lösen müssen. Die Lösung (beispielsweise) im Fall $k_1 > 1/2$ ist
$$u(t) = a_1 \tan\left(\frac{a_1t}{2} + a_2\right) - 1$$
mit a_1, $a_2 \in \mathbb{R}$, wobei $k_1 = \frac{a_1^2+1}{2}$. In den ursprünglichen x-Koordinaten lautet die Lösung
$$u(x) = a_1 \tan\left(\frac{a_1 \log x}{2} + a_2\right) - 1.$$

Eine *lineare Eulergleichung* ist von der Form
$$a_0\left(x\frac{d}{dx}\right)^m u + a_1\left(x\frac{d}{dx}\right)^{m-1} u + \cdots + a_{m-1}\left(x\frac{d}{dx}\right)u + a_m u = 0,$$
wobei $a_0, a_1, \ldots, a_m \in \mathbb{C}$, $a_0 \neq 0$.
Aus der Theorie linearer Gleichungen mit konstanten Koeffizienten erhalten wir:

[16] Im Englischen: *equidimensional*.

Satz 12.37. *Es sei*

$$a_0\mu^m + a_1\mu^{m-1} + \cdots + a_{m-1}\mu + a_m = \prod_{j=1}^{r} (\mu - \mu_j)^{m_j}$$

mit $\mu_j \neq \mu_{j'}$ *für* $j \neq j'$. *Dann ist*

$$x^{\mu_1}, x^{\mu_1}\log x, \dots, x^{\mu_1}(\log x)^{m_1-1}, \dots, x^{\mu_r}, x^{\mu_r}\log x, \dots, x^{\mu_r}(\log x)^{m_r-1}$$

ein Fundamentalsystem für obige lineare Differenzialgleichung.

Beispiel 12.38. Die Gleichung $u'' + \frac{u}{4x^2} = 0$ führt auf die Gleichung $(\mu - \frac{1}{2})^2 = 0$. Die allgemeine Lösung ist deshalb

$$u(x) = k_1\sqrt{x} + k_2\sqrt{x}\log x$$

mit k_1, $k_2 \in \mathbb{C}$.

12.5 Randwert- und Eigenwertprobleme

12.5.1 Grundlegende Eigenschaften

Wir beginnen diesen Abschnitt mit einem Beispiel.

Beispiel 12.39. Das Randwertproblem

$$\begin{cases} u'' + \mu u = 0, \\ u(0) = u(\pi) = 0, \end{cases} \tag{12.11}$$

wobei $\mu \in \mathbb{C}$ ein Parameter ist, kann eine von Null verschiedene Lösung besitzen. Tatsächlich ist die allgemeine Lösung der linearen Gleichung $u'' + \mu u = 0$ durch

$$u(x) = k_1 \sin(\sqrt{\mu}x) + k_2\cos(\sqrt{\mu}x)$$

gegeben.[17] Die Randbedingung $u(0) = 0$ erzwingt $k_2 = 0$, die Randbedingung $u(\pi) = 0$ zusammen mit $k_1 \neq 0$ impliziert $\mu = k^2$ für ein $k \in \mathbb{N}$.[18] Die Werte k^2 für $k \in \mathbb{N}$ heißen die *Eigenwerte* des Randwertproblems (12.11), die Funktionen $u_k(x) = \sin(kx)$ heißen die zugehörigen *Eigenfunktionen*.

Bemerkung. In Kapitel 6 hatten wir 2π-periodische, ungerade Funktionen nach diesen Eigenfunktionen $\sin(kx)$ für $k \in \mathbb{N}$ in *Fourierreihen* entwickelt. Diese Bemerkung zeigt, dass die Eigenfunktionen zumindestens für spezielle Randwertprobleme in Anwendungen enorm wichtig sein können.

Wir wollen im Weiteren die Frage nach Existenz von *Eigenwerten* allgemein aufgreifen. Dazu betrachten wir ein $N \times N$-System erster Ordnung der Form

$$\begin{cases} U' = A(x)\,U + \mu B(x)\,U, \quad x^0 \leq x \leq x^1, \\ M_0 U(x^0) + M_1 U(x^1) = 0, \end{cases} \tag{12.12}$$

[17] Hier ist $\sqrt{\mu}$ eine der beiden komplexen Zahlen mit $(\sqrt{\mu})^2 = \mu$.
[18] $k = 0$ ergibt die Nulllösung, $-k$ ergibt dasselbe μ wie k.

wobei $A, B \in C([x^0, x^1]; M(N; \mathbb{C}))$, $M_0, M_1 \in M(N, \mathbb{C})$, $x^0 < x^1$ und $\mu \in \mathbb{C}$ ein sogenannter *Spektralparameter* ist.
Wir machen weiterhin die Annahme, dass die lineare Abbildung

$$\mathbb{C}^{2N} \to \mathbb{C}^N, \quad (U^0, U^1) \mapsto M_0 U^0 + M_1 U^1 \tag{12.13}$$

vollen Rang N hat, d. h. surjektiv ist.

Beispiel 12.40. a) Obiges Beispiel (12.11) können wir mit $U = \binom{u}{u'}$ als

$$\begin{cases} U' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} U + \mu \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} U, \quad 0 \le x \le \pi, \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U(0) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} U(\pi) = 0 \end{cases}$$

schreiben.

b) Im Weiteren betrachten wir folgende (in Anwendungen vorkommende) Randbedingungen für die Gleichung in (12.11) detailliert (die Tabelle zeigt außerdem mögliche Wahlen für M_0, M_1):

Kürzel	Bezeichung	Randbedingung	M_0	M_1
(D)	Dirichletsche RB	$u(0) = u(\pi) = 0$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$
(N)	Neumannsche RB	$u'(0) = u'(\pi) = 0$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
(Per)	Periodische RB	$u(0) = u(\pi),\ u'(0) = u'(\pi)$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
(APer)	Antiperiodische RB	$u(0) + u(\pi) = 0,\ u'(0) + u'(\pi) = 0$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Definition 12.41. $\mu^0 \in \mathbb{C}$ heißt ein *Eigenwert* von (12.12), falls es eine Lösung $U = U(x)$ zu (12.12) mit $\mu = \mu^0$ gibt, die nicht identisch Null ist. Ein solches U heißt dann eine *Eigenfunktion* zum Eigenwert μ^0. [19]

Um Eigenwerte und Eigenfunktionen zu studieren, zeigen wir als erstes, dass die Fundamentalmatrix $Y(x, x^0; \mu)$ von (12.12) holomorph vom Parameter μ abhängt.[20]

Satz 12.42. *Die Fundamentalmatrix $Y(x, x^0; \mu)$ hat die Form*

$$\sum_{k=0}^{\infty} \mu^k Y_k(x, x^0),$$

wobei $X(x, x^0) = Y_0(x, x^0; 0)$ die Fundamentalmatrix des Systems $U' = A(x)U$ ist und die Reihe für alle $\mu \in \mathbb{C}$ gleichmäßig bezüglich $x \in [x^0, x^1]$ und absolut konvergiert.

[19] Wie in der linearen Algebra können Eigenwerte geometrische und algebraische Multiplizitäten größer 1 haben.

[20] Für uns bedeutet *Holomorphie*, dass wir die Fundamentalmatrix in einer konvergenten Potenzreihe bezüglich μ mit Konvergenzradius ∞ entwickeln können. In der Vorlesung „Funktionentheorie" werden Sie lernen, dass dies gerade die auf *ganz* $\mathbb{C}$ *komplex-differenzierbaren Funktionen*, sogenannte *ganze Funktionen* charakterisiert.

Beweis. Eine formale Manipulation der Gleichungen $Y = \sum_{k=0}^{\infty} \mu^k Y_k$ und $Y' = AY + \mu BY$ ergibt wegen $Y' = \sum_{k=0}^{\infty} \mu^k Y_k'$ die Beziehungen

$$\begin{cases} Y_0' = AY_0, \quad Y_0(x^0) = I, \\ Y_k' = AY_k + BY_{k-1}, \quad Y_k(x^0) = 0 \end{cases} \tag{12.14}$$

für $k \geq 1$. Das System (12.14) ist eindeutig lösbar, und zwar erhalten wir sukzessiv, dass $Y_0 = X$ und

$$Y_k(x) = \int_{x^0}^{x} X(x, x')B(x')Y_{k-1}(x')\,dx'$$

für $k \geq 1$ gilt.
Sei nun $I = [x^0, x^1]$. Sei ferner $a = \max_{x \in I} |A(x)|$ und $b = \max_{x \in I} |B(x)|$. Dann gilt

$$|Y_k(x)| \leq \frac{b^k (x - x_0)^k}{k!}\, e^{a(x-x^0)}, \quad \forall x \in I, \tag{12.15}$$

für alle $k \in \mathbb{N}_0$.
Beweis von (12.15) *mittels Induktion nach* k.
Der Induktionsanfang $k = 0$. Es gilt $|X(x)| = \max_{|U^0| \leq 1} \left|X(x)U^0\right|$, also müssen wir $|U(x)| \leq e^{a(x-x^0)} \left|U^0\right|$ für $x \in I$ und jede Lösung $U = U(x)$ zu

$$U' = A(x)U, \quad U(x^0) = U^0$$

zeigen. Für eine solche Lösung gilt

$$\begin{aligned} \frac{d}{dx} |U(x)|^2 &= 2 \operatorname{Re} \left\langle U'(x), \overline{U(x)} \right\rangle \\ &= 2 \operatorname{Re} \left\langle A(x)U(x), \overline{U(x)} \right\rangle \leq 2a\, |U(x)|^2, \end{aligned}$$

also $|U(x)|^2 \leq \left|U^0\right|^2 e^{2a(x-x^0)}$ und $|U(x)| \leq e^{a(x-x^0)} \left|U^0\right|$.
Der Induktionsschritt $k \to k+1$. Es gilt

$$\begin{aligned} |Y_{k+1}(x)| &\leq \int_{x^0}^{x} |X(x, x')|\, |B(x')|\, |Y_k(x')|\, dx' \\ &\leq \frac{b^{k+1}}{k!} \int_{x^0}^{x} e^{a(x-x')} (x' - x^0)^k e^{a(x'-x^0)}\, dx' \\ &= \frac{b^{k+1}}{k!}\, e^{a(x-x^0)} \int_{x^0}^{x} (x' - x^0)^k\, dx' \\ &= \frac{b^{k+1}}{k!}\, e^{a(x-x^0)} \frac{(x - x^0)^{k+1}}{k+1} = \frac{b^{k+1}(x - x^0)^{k+1}}{(k+1)!}\, e^{a(x-x^0)}. \end{aligned}$$

Aus (12.15) folgt die absolute Konvergenz der Reihe $\sum_{k=0}^{\infty} \mu^k Y_k(x, x^0)$ gleichmäßig bezüglich $x \in I$, nämlich

$$\sum_{k=0}^{\infty} \left|\mu^k Y_k(x)\right| \leq \sum_{k=0}^{\infty} |\mu|^k \frac{b^k (x - x^0)^k}{k!}\, e^{a(x-x^0)} = e^{(a+|\mu|b)(x-x^0)} < \infty. \qquad \square$$

Beispiel 12.43. Für das Beispiel (12.11) ergibt sich (mit $x^0 = 0$)[21]

$$Y(x,\mu) = \begin{cases} \begin{pmatrix} \cos(\nu x) & \frac{1}{\nu}\sin(\nu x) \\ -\nu\sin(\nu x) & \cos(\nu x) \end{pmatrix} & \text{für } \mu \neq 0, \\ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} & \text{für } \mu = 0. \end{cases}$$

wobei wir hier und im Folgenden $\mu = \nu^2$ setzen.[22]

Die Kenntnis der Fundamentalmatrix gestattet ein einfaches Kriterium für das Vorliegen von Eigenwerten:

Satz 12.44. *$\mu^0 \in \mathbb{C}$ ist dann und nur dann ein Eigenwert von* (12.12)*, wenn*

$$\Delta(\mu) = \det\big(M_0 + M_1 Y(x^1, x^0; \mu)\big)$$

an der Stelle μ^0 verschwindet.

Beweis. ($\Longrightarrow$) Sei U eine Lösung zu (12.12) mit $\mu = \mu^0$, die nicht identisch verschwindet. Sei $U^0 = U(x^0)$, $U^1 = U(x^1)$. Dann ist $U^1 = Y(x^1, x^0; \mu^0)U^0$, also

$$\big(M_0 + M_1 Y(x^1, x^0; \mu^0)\big)\, U^0 = 0. \tag{12.16}$$

Somit ist $U^0 \neq 0$ Eigenvektor zum Eigenwert 0, folglich ist $\Delta(\mu^0) = 0$.

($\Longleftarrow$) Gilt umgekehrt $\Delta(\mu^0) = 0$, so finden wir ein $U^0 \neq 0$ mit (12.16). Die gesuchte Lösung U von (12.12) ist dann $U(x) = Y(x, x^0; \mu^0)U^0$. □

Dies hat die folgende einfache Konsequenz:

Satz 12.45. *Sei $\Delta(\mu) \not\equiv 0$. Dann sind die Eigenwerte von* (12.12) *isoliert. Insbesondere können sich die Eigenwerte von* (12.12) *nicht im Endlichen häufen.*

Beweis. $\Delta(\mu)$ ist eine ganze holomorphe Funktion, d. h. $\Delta(\mu)$ ist als eine in ganz $\mathbb{C}$ konvergente Potenzreihe darstellbar. Dann folgt das Ergebnis aus dem Eindeutigkeitssatz 6.2.4 für Potenzreihen, der sinngemäß auch im Komplexen gilt. □

[21] Benutzt man die Taylorreihe für den Cosinus und den Sinus, so findet man

$$Y_k(x) = (-1)^k \begin{cases} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} & \text{für } k = 0, \\ \begin{pmatrix} \frac{x^{2k}}{(2k)!} & \frac{x^{2k+1}}{(2k+1)!} \\ \frac{x^{2k-1}}{(2k-1)!} & \frac{x^{2k}}{(2k)!} \end{pmatrix} & \text{für } k \geq 1. \end{cases}$$

[22] Der Ausdruck auf der rechten Seite ist eine gerade Funktion in ν und damit unabhängig von der Wahl der Wurzel von μ.

Beispiel 12.46. a) Die Eigenwerte und Eigenfunktionen[23] in unseren vier Beispielen sind jetzt leicht zu berechnen:

Randbed.	$\det(M_0 + M_1 Y(\pi, 0; \mu))$	Eigenwerte	Eigenfunktionen
(D)	$\begin{cases} \frac{1}{\nu}\sin(\pi\nu) & \text{für } \mu \neq 0, \\ \pi & \text{für } \mu = 0 \end{cases}$	$k^2,\ k \in \mathbb{N}$	$\sin(kx)$
(N)	$\nu\sin(\pi\nu)$	$k^2,\ k \in \mathbb{N}_0$	$\cos(kx)$
(Per)	$-2\,(1 - \cos(\pi\nu))$	$4k^2,\ k \in \mathbb{N}_0$, (EW sind doppelt für $k \geq 1$)	$1,\quad k = 0$, $\cos(2kx),\ \sin(2kx),\quad k \geq 1$
(APer)	$2\,(1 + \cos(\pi\nu))$	$(2k+1)^2,\ k \in \mathbb{N}_0$ (EW sind doppelt)	$\cos((2k+1)x),\ \sin((2k+1)x)$

b) Wählen wir hingegen die Randbedingung $u(0) + u(\pi) = 0,\ u'(0) = u'(\pi)$, also etwa $M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $M_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, so gilt $\Delta(\mu) \equiv 0$, wie man leicht verifiziert. Tatsächlich ist dann

$$u(x) = \begin{cases} \sin\left(\nu x - \frac{\pi\nu}{2}\right) & \text{für } \mu \neq 0, \\ x - \frac{\pi}{2} & \text{für } \mu = 0 \end{cases}$$

für jedes $\mu \in \mathbb{C}$ eine Eigenfunktion.

Bemerkung. Für das allgemeine Cauchyproblem zu (12.12) (mit $M_0 = I_N$, $M_1 = 0_N$) gilt offenbar $\Delta(\mu) \equiv 1$. Wir finden bestätigt, dass es hier keine Eigenwerte und Eigenfunktionen gibt, was wir auch aus der allgemeinen Theorie des Abschnitts 12.2 bereits wissen.

12.5.2 Das adjungierte Randwertproblem

Wir betrachten weiterhin das Problem (12.12). Gleichzeitig werden wir das dazu adjungierte Randwertproblem betrachten.
Wir arbeiten mit der Sesquilinearform

$$\big(U, V\big) = \big\langle U, \overline{V}\big\rangle, \quad U,\, V \in \mathbb{C}^N.$$

Die adjungierte Matrix D^* zu einer Matrix $D \in M(N; \mathbb{C})$ ist dann vermöge

$$\big(DU, V\big) = \big(U, D^* V\big), \quad U,\, V \in \mathbb{C}^N,$$

erklärt.

Definition 12.47. Das zu (12.12) *adjungierte Randwertproblem* ist von der Form

$$\begin{cases} V' = -\left(A^*(x) + \mu' B^*(x)\right) V, & x^0 \leq x \leq x^1, \\ N_0 V(x^0) + N_1 V(x^1) = 0 \end{cases} \tag{12.17}$$

mit weiter unten zu spezifizierenden Matrizen $N_0,\, N_1 \in M(N; \mathbb{C})$. Hierbei ist $\mu' \in \mathbb{C}$ der Spektralparameter.

[23] In der vierten Spalte stehen die Eigenfunktionen des ursprünglichen skalaren Problems zweiter Ordnung.

Um die Form der *adjungierten Randbedingungen* zu bestimmen, machen wir eine Reihe von Beoachtungen:

Lemma 12.48. *Die Fundamentalmatrix* $Y^*(x, x^0; \mu')$ *des adjungierten Problems* (12.17) *ist* $Y(x^0, x; \overline{\mu'})^*$.

Beweis. Dies folgt aus $Y(x^0, x^0; \overline{\mu'})^* = I^* = I$ und

$$\begin{aligned} \frac{\partial}{\partial x} Y(x^0, x; \overline{\mu'})^* &= \left(\frac{\partial}{\partial x} Y(x^0, x; \overline{\mu'}) \right)^* \\ &= \left(-Y(x^0, x; \overline{\mu'})(A(x) + \overline{\mu'} B(x)) \right)^* \\ &= -\left(A^*(x) + \mu' B^*(x) \right) Y(x^0, x; \overline{\mu'})^*, \end{aligned}$$

wobei wir Bemerkung 12.3.1 verwendeten. □

Beispiel 12.49. Für die Gleichung in (12.11) ergibt sich als adjungierte Gleichung

$$V' = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} V + \mu' \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} V, \quad 0 \le x \le \pi,$$

mit einer noch jeweils geeignet zu bestimmenden Randbedingung. Die Fundamentalmatrix (mit $x^0 = 0$) ergibt sich zu

$$Y^*(x, \mu') = \begin{cases} \begin{pmatrix} \cos(\nu' x) & \nu' \sin(\nu' x) \\ -\frac{1}{\nu'} \sin(\nu' x) & \cos(\nu' x) \end{pmatrix} & \text{für } \mu \neq 0, \\ \begin{pmatrix} 1 & 0 \\ -x & 1 \end{pmatrix} & \text{für } \mu = 0, \end{cases}$$

wobei $\mu' = \nu'^2$.

Lemma 12.50. *Ist* U *eine Lösung des Problems* (12.12) *und* V *eine Lösung des Problems* (12.17) *mit* $\mu' = \overline{\mu}$ *(beides ohne Randbedingungen), so ist die Größe*

$$\big(U(x), V(x)\big), \quad x^0 \le x \le x^1,$$

unabhängig von x.

Beweis. 1. Methode. Mit $U^0 = U(x^0)$, $V^0 = V(x^0)$ gilt für $x^0 \le x \le x^1$

$$\begin{aligned} \big(U(x), V(x)\big) &= \big(Y(x, x^0; \mu) U^0, Y^*(x^0, x; \mu) V^0\big) \\ &= \big(Y(x^0, x; \mu) Y(x, x^0; \mu) U^0, V^0\big) = \big(U^0, V^0\big). \end{aligned}$$

2. Methode. Es gilt

$$\begin{aligned} \frac{d}{dx}(U, V) &= (U', V) + (U, V') = ((A + \mu B) U, V) - (U, (A^* + \bar{\mu} B^*) V) \\ &= ((AU, V) - (U, A^* V)) + \mu \left((BU, V) - (U, B^* V) \right) = 0. \end{aligned}$$

□

Die adjungierte Randbedingung ist durch die Bedingung charakterisiert, dass die Größe $\big(U(x), V(x)\big)$ identisch verschwindet. Dazu sei

$$\beta\colon \mathbb{C}^{2N} \times \mathbb{C}^{2N} \to \mathbb{C}, \quad \left(\begin{pmatrix} U^0 \\ U^1 \end{pmatrix}, \begin{pmatrix} V^0 \\ V^1 \end{pmatrix}\right) \mapsto \left(U^0, V^0\right) - \left(U^1, V^1\right)$$

eine *nichtentartete* Sesquilinearform auf $\mathbb{C}^{2N} \cong \mathbb{C}^N \times \mathbb{C}^N$ und

$$\Pi = \left\{(U^0, U^1) \in \mathbb{C}^{2N} \;\middle|\; M_0 U^0 + M_1 U^1 = 0\right\}.$$

ein linearer Unterraum von $\mathbb{C}^{2N}$. Laut der Annahme in (12.13) ist

$$\dim \Pi = N.$$

Definition 12.51. Die *adjungierte Randbedingung* in (12.17) ist durch

$$\Pi^\perp = \left\{(V^0, V^1) \in \mathbb{C}^{2N} \;\middle|\; N_0 V^0 + N_1 V^1 = 0\right\}$$

gegeben, wobei $\Pi^\perp$ der zu Π orthogonale Unterraum bezüglich β ist.

Wir bemerken, dass $\dim \Pi^\perp = N$ gilt.

Beispiel 12.52. a) In unseren Beispielen ergeben sich die adjungierten Randbedingungen und die Funktion $\Delta^*(\mu') = \det\left(N_1 + N_0 Y^*(\pi, 0; \mu')\right)$ wie folgt:

Randbed.	N_0	N_1	$\Delta^*(\mu')$
(D)	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{cases} -\frac{1}{\nu'}\sin(\pi\nu') & \text{für } \mu' \neq 0, \\ -\pi & \text{für } \mu' = 0 \end{cases}$
(N)	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$-\nu' \sin(\pi\nu')$
(Per)	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$-2\left(1 - \cos(\pi\nu')\right)$
(APer)	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$2\left(1 + \cos(\pi\nu')\right)$

b) Im Fall der Randbedingung (D) ist die Eigenfunktion für das ursprüngliche Randwertproblem zum Eigenwert k^2, $k \in \mathbb{N}$, gleich $U(x) = \begin{pmatrix} \frac{1}{k}\sin(kx) \\ \cos(kx) \end{pmatrix}$, wohingegen die Eigenfunktion für das adjungierte Problem zum gleichen Eigenwert gleich $V(x) = \begin{pmatrix} \cos(kx) \\ -\frac{1}{k}\sin(kx) \end{pmatrix}$ ist. Tatsächlich gilt

$$(U(x), V(x)) = \frac{1}{k}\sin(kx)\cos(kx) - \cos(kx)\frac{1}{k}\sin(kx) = 0$$

für alle $0 \leq x \leq \pi$, wie in Lemma 12.50 behauptet.

Tatsächlich sind die Matrizen N_0, N_1 im Wesentlichen eindeutig bestimmt:

Lemma 12.53. *Die in der vorigen Definition angegebene Bedingung bestimmt die Matrizen N_0, $N_1 \in M(N;\mathbb{C})$ bis auf Linksmultiplikation mit einem Faktor $D \in \mathrm{Gl}(N;\mathbb{C})$.*

Beweis. $\Pi^\perp$ ist der Kern der linearen Abbildung $\mathbb{C}^{2N} \to \mathbb{C}^N$, $(V^0, V^1) \mapsto N_0V^0 + N_1V^1$, die wir mit (N_0, N_1) bezeichnen wollen. Diese Abbildung induziert eine lineare Isomorphie $\widetilde{(N_0, N_1)}\colon \mathbb{C}^{2N}/\Pi^\perp \xrightarrow{\cong} \mathbb{C}^N$. Für jedes weitere Paar von Matrizen N_0', $N_1' \in M(N;\mathbb{C})$ mit derselben Eigenschaft erhalten wir analog eine lineare Isomorphie $\widetilde{(N_0', N_1')}\colon \mathbb{C}^{2N}/\Pi^\perp \xrightarrow{\cong} \mathbb{C}^N$. Wir setzen $D = \widetilde{(N_0', N_1')}\widetilde{(N_0, N_1)}^{-1}\colon \mathbb{C}^N \xrightarrow{\cong} \mathbb{C}^N$ und identifizieren D mit einer Matrix in $\mathrm{Gl}(N;\mathbb{C})$. Dann

$$(N_0', N_1') = D \circ (N_0, N_1) = (DN_0, DN_1)$$

als lineare Abbildungen $\mathbb{C}^{2N} \to \mathbb{C}^N$ bzw. $N_j' = DN_j$ für $j = 0, 1$. □

Wir bemerken, dass die Multiplikation von N_1, N_1 mit $D \in \mathrm{Gl}(N;\mathbb{C})$ die Funktion $\Delta^*(\mu')$, nun für das Randwertproblem (12.17) definiert, nur um den konstanten Faktor $\det D \neq 0$ abändert. Insbesondere ändern sich die Nullstellen der Funktion $\Delta^*(\mu')$ dabei nicht.

Beispiel 12.54. Um die adjungierte Randbedingung in einer Klasse von Beispielen zu bestimmen, nehmen wir an, dass für ein $k \in \{0, 1, \dots, N\}$ (möglicherweise nach Multiplikation mit einer Matrix $D \in \mathrm{Gl}(N;\mathbb{C})$)

$$M_0 = \begin{pmatrix} M_0^+ & 0 \\ M_0^- & I_{N-k} \end{pmatrix}, \quad M_1 = \begin{pmatrix} I_k & M_1^+ \\ 0 & M_1^- \end{pmatrix}$$

mit $M_0^+ \in M(k;\mathbb{C})$, $M_0^- \in M(N-k, k;\mathbb{C})$, $M_1^+ \in M(k, N_k;\mathbb{C})$ und $M_1^- \in M(N_k;\mathbb{C})$ gilt. Entsprechend zerlegen wir $U^0 = \begin{pmatrix} U^0_+ \\ U^0_- \end{pmatrix}$ mit $U^0_+ \in \mathbb{C}^k$, $U^0_- \in \mathbb{C}^{N-k}$ und $U^1 = \begin{pmatrix} U^1_+ \\ U^1_- \end{pmatrix}$ mit $U^1_+ \in \mathbb{C}^k$, $U^1_- \in \mathbb{C}^{N-k}$. Die Bedingung $M_0U^0 + M_1U^1 = 0$ schreibt sich dann als

$$\begin{cases} U^0_- = -M_0^-U^0_+ - M_1^-U^1_-, \\ U^1_+ = -M_0^+U^0_+ - M_1^+U^1_-. \end{cases}$$

Wir sehen insbesondere, dass U^0_+, U^1_- frei wählbar sind, während U^0_-, U^1_+ sich wie angegeben ergeben.

Lemma 12.55. *Unter diesen Voraussetzungen ist*

$$N_0 = \begin{pmatrix} I_k & -(M_0^-)^* \\ 0 & (M_1^-)^* \end{pmatrix}, \quad N_1 = \begin{pmatrix} (M_0^+)^* & 0 \\ -(M_1^+)^* & I_{N-k} \end{pmatrix}$$

eine mögliche Wahl für N_0, N_1.

Beweis. Die Bedingung, dass $(U^0, V^0) = (U^1, V^1)$ für $(U^0, U^1) \in \Pi$, $(V^0, V^1) \in \Pi^\perp$ ist, schreibt sich als

$$\big(U^0_+, V^0_+\big) - \big(M_0^-U^0_+ + M_1^-U^1_-, V^0_-\big) = -\big(M_0^+U^0_+ + M_1^+U^1_-, V^1_+\big) + \big(U^1_-, V^1_-\big).$$

Da $U^0_+ \in \mathbb{C}^k$, $U^1_- \in \mathbb{C}^{N-k}$ beliebig sind, folgt daraus, dass

$$\begin{aligned} V^0_+ &= (M_0^-)^*V^0_- - (M_0^+)^*V^1_+, \\ V^1_- &= -(M_1^-)^*V^0_- + (M_1^+)^*V^1_+. \end{aligned}$$

Das ist gerade die Behauptung. □

Bemerkung. Die Randbedingungen (Per) und (APer) sind vom gerade behandelten Typ, die Randbedingungen (D) und (N) hingegen nicht.[24]

Die Bedeutung des adjungierten Problems wird mit folgendem Resultat klar:

Satz 12.56. *$\mu \in \mathbb{C}$ ist genau dann Eigenwert von* (12.12)*, wenn $\bar{\mu}$ Eigenwert von* (12.17) *ist.*

Beweis. Sei

$$\Sigma = \left\{(U^0, U^1) \in \mathbb{C}^{2N} \;\middle|\; U^1 = Y(x^1, x^0; \mu)U^0\right\}$$

der Raum der möglichen Randdaten für das Problem (12.12) und

$$\Sigma^* = \left\{(V^0, V^1) \in \mathbb{C}^{2N} \;\middle|\; V^1 = Y^*(x^1, x^0; \bar{\mu})V^0\right\}$$

der entsprechende Raum für das Problem (12.17) (mit $\mu' = \overline{\mu}$). Wir bemerken, dass $\dim \Sigma = \dim \Sigma^* = N$ gilt. Lemma 12.50 besagt gerade, dass $\Sigma^* = \Sigma^\perp$. Nun ist μ genau dann ein Eigenwert von (12.12), wenn $\Sigma \cap \Pi \neq \{0\}$, und das gilt genau dann, wenn $\Sigma^* \cap \Pi^\perp \neq \{0\}$, d. h. wenn $\overline{\mu}$ ein Eigenwert von (12.17) ist. □

Bemerkung. a) Tatsächlich lässt sich zeigen, dass

$$\Delta^*(\bar{\mu}) = C\,\overline{\Delta(\mu)}$$

mit einer Konstanten $C \in \mathbb{C}$, $C \neq 0$ gilt, was nochmals den vorigen Satz zur Folge hat.

b) Gleichzeitig zeigt der Beweis des vorigen Satzes auch, dass die geometrische Vielfachheit $\dim(\Sigma \cap \Pi) = \dim\left(\Sigma^* \cap \Pi^\perp\right)$ beider Eigenwerte μ und $\bar{\mu}$ übereinstimmt.

12.5.3 Existenz und Eindeutigkeit von Lösungen

Wir betrachten jetzt das *inhomogene Randwertproblem*

$$\begin{cases} U' = (A(x) + \mu B(x))\,U + F(x), \quad x^0 \leq x \leq x^1, \\ M_0 U(x^0) + M_1 U(x^1) = 0, \end{cases} \tag{12.18}$$

wobei die Annahmen an A, B, M_0, M_1, μ wie im vorigen Abschnitt sind und $F \in C([x^0, x^1]; \mathbb{C}^N)$ gilt.

Satz 12.57. a) *Das inhomogene Randwertproblem* (12.18) *besitzt genau dann eine Lösung, wenn für jede Lösung V des* (*homogenen*) *adjungierten Randwertproblems* (12.17) *mit $\mu' = \overline{\mu}$*

$$\int_{x^0}^{x^1} \big(F(x), V(x)\big)\,dx = 0$$

gilt.

[24] Auch nicht nach Ersetzung von M_0, M_1 durch DM_0, DM_1 mit $D \in \mathrm{Gl}(N; \mathbb{C})$.

b) *Insbesondere ist das inhomogene Randwertproblem* (12.18) *genau dann eindeutig lösbar, wenn das adjungierte Randwertproblem* (12.17) *nur die Nulllösung zuläßt.*

Beweis. Sei U eine Lösung zu (12.18), V eine Lösung zu (12.17) mit $\mu' = \overline{\mu}$. Dann gilt

$$\begin{aligned}\int_{x^0}^{x^1} \big(F, V\big)\, dx &= \int_{x^0}^{x^1} \big(U' - \big(A + \mu B\big), V\big)\, dx \\ &= \big(U, V\big)\Big|_{x^0}^{x^1} - \int_{x^0}^{x^1} \big(U, V' + \big(A^* + \overline{\mu} B^*\big)V\big)\, dx = 0.\end{aligned}$$

Die umgekehrte Richtung folgt aus Dimensionsgründen. □

Bemerkung. a) Ein analoges Result gilt im Fall des Problems (12.18) mit einer *inhomogenen Randbedingungen* der Form

$$M_0 U(x^0) + M_1 U(x^1) = U^2,$$

wobei $U^2 \in \mathbb{C}^N$, nur dass sich jetzt die Lösbarkeitsbedingung in a) etwa im Fall des Beispiels 12.54 als

$$\big(U^2_-, V(x^0)_-\big)_{\mathbb{C}^{N-k}} + \int_{x^0}^{x^1} \big(F(x), V(x)\big)_{\mathbb{C}^N}\, dx = \big(U^2_+, V(x^1)_+\big)_{\mathbb{C}^k} \tag{12.19}$$

für alle V wie zuvor schreibt. Hierbei bezeichnet V_+ für $V \in \mathbb{C}^N$ den Vektor aus den ersten k Komponenten und V_- den Vektor aus den letzen $N-k$ Komponenten von V.

b) Ist μ *kein Eigenwert*, so hat das Problem (12.18) mit der inhomogenen Randbedingungen (12.19) die Lösung

$$U(x) = Y(x, x^0; \mu) U^0 + \int_{x^0}^{x} Y(x, x'; \mu) F(x')\, dx',$$

wobei sich U^0 eindeutig aus der Gleichung

$$\big(M_0 + M_1 Y(x^1, x^0; \mu)\big)\, U^0 = U^2 - M_1 \int_{x^0}^{x^1} Y(x^1, x'; \mu) F(x')\, dx' \tag{12.20}$$

bestimmt. Ist μ ein Eigenwert und ist die in Satz 12.57 bzw. in a) genannte Kompatibilitätsbedingung erfüllt, so gehört der Vektor auf der rechten Seite von (12.20) immer noch zum Bild des Operators $M_0 + M_1 Y(x^1, x^0; \mu) \in \mathscr{L}(\mathbb{C}^N)$, so dass $U^0 \in \mathbb{C}^N$ immer noch geeignet gefunden werden kann, aber nicht mehr eindeutig bestimmt ist.

Beispiel 12.58. Wir betrachten das inhomogene Randwertproblem

$$\begin{cases} u'' + u = f(x), \quad 0 \le x \le \pi, \\ u(0) = 1, \;\; u(\pi) = 0 \end{cases}$$

für ein $f \in C([0, \pi]; \mathbb{C})$. Die Lösbarkeitsbedingung ist

$$\int_0^\pi \sin y\, f(y)\, dy = 1.$$

Ist diese Bedingung erfüllt, so ergibt sich die allgemeine Lösung zu

$$u(x) = k \sin x + \cos x + \int_0^x \sin(x-y) f(y)\, dy,$$

wobei der Parameter $k \in \mathbb{C}$ frei gewählt werden kann.

12.6 Asymptotische Eigenschaften von Fundamentalsystemen und Asymptotik der Eigenwerte

In diesem Abschnitt beschäftigen wir uns mit Situationen, in denen man die betrachteten Gleichungen nicht mehr in geschlossener Form lösen kann, aber immer noch für viele Anwendungen hinreichend gute Informationen über die Lösungen erhält.

12.6.1 Die Liouville-Greensche Approximation

Heuristik

Wir betrachten eine skalare homogene Gleichung zweiter Ordnung

$$\frac{d^2u}{dx^2} = f(x)u, \quad x \in I, \tag{12.21}$$

wobei f eine zweimal stetig differenzierbare, reellwertige Funktion bezeichnet, *die auf auf dem Intervall I nirgends veschwindet.* Wir möchten diese Gleichung approximativ lösen. In Regionen, in denen f in etwa konstant ist, ist eine erste Näherung

$$u(x) \sim k_1 e^{x\sqrt{f(x)}} + k_2 e^{-x\sqrt{f(x)}}$$

mit beliebigen Konstanten k_1, k_2. Diese Näherung gibt einen Hinweis auf den Charakter der Lösungen (diese sind vom exponentiellen Typ im Fall $f(x) > 0$ und vom trigonometrischen bzw. oszillierenden Typ im Fall $f(x) < 0$), doch oft ist diese Approximation nicht gut für globale Betrachtungen auf I geeignet.
Eine bessere Möglichkeit verfährt wie folgt: Wir führen eine neue Koordinate $y = y(x)$ ein und setzen

$$v = \sqrt{\frac{dy}{dx}}\, u.$$

Lemma 12.59. *Die Funktion $v = v(y)$ genügt der Differenzialgleichung*

$$\frac{d^2v}{dy^2} = \left(\left(\frac{dx}{dy}\right)^2 f(x) + \left(\frac{dx}{dy}\right)^{1/2} \frac{d^2}{dy^2}\left(\left(\frac{dx}{dy}\right)^{-1/2} \right) \right) v. \tag{12.22}$$

Beweis. Dies ist nachzurechnen. □

Bemerkung. Der Term $(dx/dy)^{1/2} \frac{d^2}{dy^2}\left((dx/dy)^{-1/2}\right)$ kann auch als

$$-\frac{1}{2}\left(\frac{d^3x/dy^3}{dx/dy} - \frac{3}{2}\left(\frac{d^2x/dy^2}{dx/dy}\right)^2 \right)$$

geschrieben werden. Letzteres ohne den Faktor $-1/2$ ist die sogenannte *Schwarzsche Ableitung*.

Der Fall $f > 0$

Wir behandeln zuerst den Fall

$$f(x) > 0, \quad \forall x \in I.$$

Unter der Annahme, dass der zweite Summand in der Klammer $(\dots)$ in (12.22) klein ist, möchten wir, dass

$$\left(\frac{dx}{dy}\right)^2 f(x) = 1, \quad x \in I,$$

gilt. Eine näherungsweise Lösung der Gleichung (12.22) ist dann

$$v(y) \sim k_1 e^{y} + k_1 e^{-y}.$$

Wir setzen also

$$y(x) = \int f^{1/2}(x)\,dx.$$

Dann schreibt sich (12.22) in der Form

$$\frac{d^2 v}{dy^2} = (1+\varphi)v \tag{12.23}$$

mit

$$\varphi = \frac{4f(x)f''(x) - 5f'^2(x)}{16 f^3(x)} = -\frac{1}{f^{3/4}}\frac{d^2}{dx^2}\left(\frac{1}{f^{1/4}}\right).$$

Definition 12.60. Die Näherungslösung

$$u(x) \sim k_1 f^{-1/4} e^{\int f^{1/2}\,dx} + k_2 f^{-1/4} e^{-\int f^{1/2}\,dx}$$

zu (12.23) heißt die *Liouville-Greensche Approximation.*

Betrachtet man stattdessen eine Gleichung der Form

$$\frac{d^2 u}{dx^2} = (f(x) + g(x))\,u,$$

so führt obige Transformation auf

$$\frac{d^2 v}{dy^2} = \left(1 + \varphi + \frac{g}{f}\right) v$$

mit φ wie oben.

Abschätzung der Fehlerterme

Seien die Annahmen wie oben erfüllt, insbesondere nehme $f(x)$ positive Werte an. Ferner sei $I = (x^0, x^1)$ ein endliches oder unendliches Intervall. Die Funktion g sei stetig auf I (nicht notwendigerweise reellwertig). Wir setzen

$$F(x) = \int \left(\frac{1}{f^{1/4}}\frac{d^2}{dx^2}\left(\frac{1}{f^{1/4}}\right) - \frac{g}{f^{1/2}}\right) dx.$$

Für eine C^1-Funktion h auf I ist die *Variation* von h durch

$$V_{x^0,x}(h) = \int_{x^0}^{x} |h'(x')|\; dx', \quad x^0 \le x \le x^1,$$

erklärt, analog für $V_{x^1,x}(h) = V_{x,x^1}(h)$.

Theorem 12.61. *Unter obigen Voraussetzungen und der weiteren Bedingung* $V_{x^j,x}(F) < \infty$ *für* $j = 0, 1$ *und alle* $x \in I$ *besitzt die Gleichung*

$$\frac{d^2u}{dx^2} = (f(x) + g(x))\, u \tag{12.24}$$

zwei C^2*-Lösungen* u_0, u_1 *mit*

$$u_0(x) = f^{-1/4}(x) \exp\left(\int f^{1/2}(x)\, dx\right)(1 + \epsilon_0(x)),$$
$$u_1(x) = f^{-1/4}(x) \exp\left(-\int f^{1/2}(x)\, dx\right)(1 + \epsilon_1(x)),$$

wobei

$$|\epsilon_j(x)| \le \exp\left(1/2 V_{x^j,x}(F)\right) - 1,$$
$$\frac{1}{2} f^{-1/2}(x) \left|\epsilon_j'(x)\right| \le \exp\left(\frac{1}{2} V_{x^j,x}(F)\right) - 1$$

für $j = 0, 1$ *gilt. Ist die Funktion* g *reellwertig, so sind diese Lösungen reellwertig.*

Beweis. Wir schreiben Gleichung (12.24) in der Form

$$\frac{d^2v}{dy^2} = (1 + \psi(y))\, v \tag{12.25}$$

mit $u = f^{-1/4}(x)v$ und

$$\psi(y) = \frac{g}{f} - \frac{1}{f^{3/4}} \frac{d^2}{dx^2}\left(\frac{1}{f^{1/4}}\right).$$

Sei $y = y^0$ für $x = x^0$ und $y = y^1$ für $x = x^1$. Dann ist ψ eine stetige Funktion im Intervall (y^0, y^1). Wir setzen $v(y) = e^y(1 + h(y))$ und erhalten

$$h''(y) + 2h'(y) - \psi(y)h(y) = \psi(y).$$

Betrachten wir in dieser Gleichung den Term $-\psi h$ als Korrekturterm, so erhalten wir

$$h(y) = \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s)\,(1 + h(s))\, ds. \tag{12.26}$$

Wir lösen diese Integralgleichung *iterativ*. Dazu nehmen wir $y^0 > -\infty$ und ψ rechtsseitig stetig an y^0 an. (Der allgemeine Fall ergibt sich analog unter Benutzung von $V_{x^0,x}(F) < \infty$.)
Wir definieren $h_0(y) = 0$ und setzen

$$h_k(y) = \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s)\,(1 + h_{k-1}(s))\, ds$$

für $k \ge 1$.
Behauptung. Für alle $k \ge 1$ gilt

$$|h_k(y) - h_{k-1}(y)| \le \frac{\Psi^k(y)}{2^k k!}, \tag{12.27}$$

wobei $\Psi(y) = \int_{y^0}^{y} |\psi(s)|\, ds$ ist.

Beweis. Wir führen den Beweis mittels Induktion über k.

Der Induktionsanfang $k = 1$. Es gilt

$$h_1(y) = \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s)\, ds,$$

also $|h_1(y)| \leq \frac{1}{2}\Psi(y)$ wegen $0 \leq 1 - e^{2(s-y)} < 1$ für $s \in [y^0, y]$.

Der Induktionsschritt $k \to k+1$. Wir haben

$$h_{k+1}(y) - h_k(y) = \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s) \left(h_k(s) - h_{k-1}(s)\right) ds, \tag{12.28}$$

also mit der Induktionsvoraussetzung

$$|h_{k+1}(y) - h_k(y)| \leq \frac{1}{2^{k+1} k!} \int_{y^0}^{y} |\psi(s)|\, \Psi^k(s)\, ds = \frac{\Psi^{k+1}(y)}{2^{k+1}(k+1)!}. \qquad \square$$

Aus (12.27) folgt, dass die Reihe

$$h(y) = \sum_{k=0}^{\infty} \left(h_{k+1}(y) - h_k(y)\right)$$

gleichmäßig auf kompakten Teilintervallen von $[x^0, x^1)$ konvergiert. Durch Summation von (12.28) finden wir, dass

$$\begin{aligned} h(y) &= h_1(y) + \sum_{k=1}^{\infty} \left(h_{k+1}(y) - h_k(y)\right) \\ &= \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s)\, ds \\ &\qquad + \sum_{k=1}^{\infty} \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s) \left(h_k(s) - h_{k-1}(s)\right) ds \\ &= \frac{1}{2} \int_{y^0}^{y} \left(1 - e^{2(s-y)}\right) \psi(s) \left(1 + h(s)\right) ds \end{aligned}$$

gilt, also genügt die Funktion h der Integralgleichung (12.26).
Wir zeigen als nächstes, dass h zweimal stetig differenzierbar ist. Dazu müssen wir zeigen, dass die Reihe

$$\sum_{k=0}^{\infty} \left(h''_{k+1}(y) - h''_k(y)\right)$$

lokal gleichmäßig konvergiert. Wir haben

$$\begin{aligned} h'_1(y) &= \int_{y^0}^{y} e^{2(s-y)} \psi(s)\, ds \\ h'_{k+1}(y) - h'_k(y) &= \int_{y^0}^{y} e^{2(s-y)} \psi(s) \left(h_k(s) - h_{k-1}(s)\right) ds. \end{aligned}$$

Zusammen mit $\left|e^{2(s-y)}\right| \leq 1$ und der Abschätzung (12.27) ergibt sich

$$\left|h'_{k+1}(y) - h'_k(y)\right| \leq \frac{\Psi^{k+1}(y)}{2^k(k+1)!}$$

für $k \geq 0$. Folglich ist $\sum\limits_{k=0}^{\infty} \left(h'_{k+1}(y) - h'_k(y)\right)$ lokal gleichmäßig konvergent.
Um die Existenz und Stetigkeit der zweiten Abbildung zu bekommen, benutzen wir die Beziehungen

$$\begin{aligned} h''_1(y) &= -2h'_1(y) + \psi(y), \\ h''_{k+1}(y) - h''_k(y) &= -2\left(h'_{k+1}(y) - h'_k(y)\right) + \psi(y)\left(h_k(y) - h_{k-1}(y)\right). \end{aligned}$$

Summation ergibt, dass h der Differenzialgleichung

$$h''(y) + 2h'(y) = (1 + \psi(y))\, h(y)$$

genügt. Zudem erhalten wir die Abschätzungen

$$|h(y)| \leq e^{\Psi(y)/2} - 1, \quad 1/2\,|h'(y)| \leq e^{\Psi(y)/2} - 1.$$

Kehren wir nun zu der ursprünglichen x-Variablen zurück, so ist $F(x) = -\int \psi(y)\,dy$ (wegen $dy = f^{1/2}(x)\,dx$), also $\Psi(y) = V_{x^0,x}(F)$. Die Behauptungen bezüglich u_0 folgen.
Die Herleitungen für u_1 sind analog. □

Bemerkung. Insbesondere erhalten wir

$$\epsilon_0(x) \to 0, \;\; f^{-1/2}(x)\epsilon'_0(x) \to 0 \text{ für } x \to x^0 + 0.$$

Analog für ϵ_1 und $x \to x^1 - 0$.

Theorem 12.61 besagt, dass wir eine Lösung u_0 mit

$$u_0(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\,dx\right) \text{ für } x \to x^0 + 0$$

finden. Eine interessante Frage im Fall, dass $\int f^{1/2}\,dx$ für $x \to x^1 - 0$ divergiert, ist, ob es eine Lösung u_2 mit

$$u_2(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\,dx\right) \text{ für } x \to x^1 - 0$$

gibt. Eine teilweise Antwort darauf gibt der folgende Satz:

Satz 12.62. *Sei zusätzlich $V_{x^0,x^1}(F) < \infty$ und $\int f^{1/2}\,dx \to \infty$ für $x \to x^1 - 0$. Dann*

$$\epsilon_0(x) \to k, \;\; f^{-1/2}(x)\epsilon'_0(x) \to 0 \text{ für } x \to x^1 - 0,$$

wobei k eine Konstante ist.

Beweis. Der Beweis benutzt $y^1 = \infty$ und die Konstruktionen des vorigen Beweises. □

Gilt jetzt also

$$\epsilon_0(x^1 - 0) \neq -1,$$

so ist u_0 eine Funktion mit den gewünschten Eigenschaften. Im Allgemeinen findet man eine solche Lösung u_2, indem man die Stelle x^0 hinreichend dicht an x^1 wählt.
Allerdings ist eine Lösung u_2 mit der Eigenschaft

$$u_2(x) \sim f^{-1/4} \exp\left(\int f^{1/2}\, dx\right) \quad \text{für } x \to x^1 - 0$$

nicht eindeutig bestimmt. Man kann stets ein beliebiges Vielfaches von u_1 addieren und erhält eine weitere Lösung mit derselben Eigenschaft.

Definition 12.63. Nahe x^1 heißt das asymptotische Verhalten von u_2 *dominant*, das von u_1 *subdominant*. Eine analoge Klassifizierung gilt nahe x^0.

Beispiel 12.64. Wir betrachten die Gleichung

$$u'' = (x + \log x)\, u \tag{12.29}$$

für $x \to \infty$. Da das Integral $\int \frac{\log x}{x^{1/2}}\, dx$ für $x \to \infty$ divergiert, können wir *nicht* $f(x) = x$, $g(x) = \log x$ wählen.
Jedoch ist $f(x) = x + \log x$, $g(x) = 0$ eine mögliche Wahl. Man sieht dann, dass

$$f^{-1/4}\big(f^{-1/4}\big)'' = O\big(x^{-5/2}\big) \quad \text{für } x \to \infty,$$

also konvergiert $V(F)$ für $x \to \infty$. (12.29) hat folglich asymptotische Lösungen der Form

$$(x + \log x)^{-1/4} \exp\left(\pm \int (x + \log x)^{1/2}\, dx\right).$$

Dieses Ergebnis lässt sich verbessern. Dazu beobachten wir, dass

$$(x + \log x)^{1/2} = x^{1/2} + 1/2 x^{-1/2} \log x + O(x^{-1/2} (\log x)^2),$$

folglich

$$\int (x + \log x)^{1/2}\, dx = \frac{2}{3} x^{3/2} + x^{1/2} \log x - 2x^{1/2} + k + o(1)$$

für $x \to \infty$. Damit besitzt (12.29) eine eindeutige (subdominante) Lösung u_1 mit

$$u_1(x) \sim x^{-1/4 - \sqrt{x}} \exp\left(2x^{1/2} - \frac{2}{3} x^{3/2}\right)$$

für $x \to \infty$ und nicht eindeutige (dominante) Lösungen u_2 mit

$$u_2(x) \sim x^{-1/4 + \sqrt{x}} \exp\left(2x^{1/2} - \frac{2}{3} x^{3/2}\right)$$

für $x \to \infty$.

Der Fall $f < 0$

Wir schreiben jetzt f als $-f$ und nehmen an, dass das neue f überall auf I positiv ist.

Theorem 12.65. *Unter obigen Voraussetzungen und der zusätzlichen Bedingung $V_{x^*,x}(F) < \infty$ für ein fixiertes x^* mit $x^0 \leq x^* \leq x^1$ betrachten wir die Gleichung*

$$\frac{d^2u}{dx^2} = (-f(x) + g(x))\, u. \tag{12.30}$$

Diese besitzt zwei C^2-Lösungen u_0, u_1 mit

$$\begin{aligned} u_0(x) &= f^{-1/4}(x) \exp\left(i \int f^{1/2}(x)\, dx\right)(1 + \epsilon_0(x)), \\ u_1(x) &= f^{-1/4}(x) \exp\left(-i \int f^{1/2}(x)\, dx\right)(1 + \epsilon_1(x)), \end{aligned}$$

wobei

$$\begin{aligned} |\epsilon_j(x)| &\leq \exp V_{x^*,x}(F) - 1, \\ f^{-1/2}(x) \left|\epsilon_j'(x)\right| &\leq \exp V_{x^*,x}(F) - 1 \end{aligned}$$

für $j = 0, 1$ gilt. Ist g reellwertig, so sind die Lösungen u_0, u_1 komplex-konjugiert.

Beweis. Der Beweis ist analog dem vorigen Beweis. Die Integralgleichung für h lautet diesmal

$$h(y) = \frac{1}{2i} \int_{y^*}^{y} \left(1 - e^{2i(s-y)}\right) \psi(s)\,(1 + h(s))\; ds,$$

wobei $y = y^*$ für $x = x^*$ ist. Der fehlende Faktor von $\frac{1}{2}$ in den Fehlerabschätzungen erklärt sich aus der Abschätzung

$$\left|1 - e^{2i(s-y)}\right| \leq 2.$$

□

Die Wahl der Stelle x^* fixiert wegen

$$\epsilon_j(x) \to 0, \quad f^{-1/2}\epsilon_j'(x) \to 0 \text{ für } x \to x^*$$

und $j = 0, 1$ die Wahl der *Anfangsbedingungen*.

Bemerkung. Im Fall, dass g reellwertig ist, lassen sich reellwertige Lösungen u zu (12.30) in der Form

$$u(x) = A f^{-1/4}(x) \left(\sin\left(\int f^{1/2}(x)\, dx + \delta\right) + \epsilon(x)\right)$$

mit gewissen Konstanten A, δ und

$$|\epsilon(x)|, \;\; f^{-1/2}(x)\,|\epsilon'(x)| \leq \exp V_{x^*,x}(F) - 1$$

schreiben.