ulm university universität
uulm

Fakultät für Mathematik und Wirtschaftswissenschaften
Institut für Angewandte Analysis

Topologische und funktionalanalytische Aspekte der schwachen Konvergenz von Maßen und Anwendungen in der Theorie des optimalen Transports

Bachelorarbeit an der Universität Ulm

Vorgelegt von:
Stefan Christian Wagner
stefan.wagner@uni-ulm.de

Gutachter:
Prof. Dr. Rico Zacher

2018

Fassung 13. November 2018

Satz: PDF-LaTeX 2_ε

Inhaltsverzeichnis

Übersicht

In dieser Arbeit möchten wir einige topologische Strukturen auf Maßen untersuchen. Zum Einstieg betrachten wir in **Kapitel 1** signierte Maße auf allgemeinen messbaren Räumen. Im Gegensatz zu „normalen“ Maßen lässt sich für Maße mit Vorzeichen eine Vektorraumstruktur definieren, welche zusammen mit der sogenannten Totalvariationsnorm einen Banachraum bildet.

Allerdings stellt sich in der Anwendung heraus, dass diese Norm oft zu restriktiv ist, weshalb wir in **Kapitel 2** einen schwächeren Konvergenzbegriff einführen. Dieser basiert auf dem Testen mit stetigen beschränkten Funktionen. Damit wir von Stetigkeit sprechen können, betrachten wir ab diesem Kapitel Maße auf metrischen Räumen. Wir charakterisieren den Konvergenzbegriff, untersuchen auf Metrisierbarkeit und beweisen den Satz von Prochorov, der ein notwendiges und hinreichendes Kriterium für Kompaktheit liefert.

Etwas funktionalanalytischer wird es in **Kapitel 3**: Hier beantworten wir die Frage, ob die schwache Konvergenz von Maßen mit einer passenden schwach-* Konvergenz übereinstimmt. Mit diesem Ansatz gelingt es uns, den Satz von Prochorov alternativ mit dem Satz von Banach-Alaoglu zu beweisen.

Als nächstes möchten wir in **Kapitel 4** die schwache Konvergenz in der Theorie des optimalen Transports anwenden. Hierbei untersuchen wir die moderne Formulierung des Monge-Kantorovich-Problems auf Lösbarkeit. Dabei handelt es sich um ein Minimierungsproblem, welches mit Wahrscheinlichkeitsmaßen formuliert wird und damit sehr allgemein gehalten wird.

Mit Hilfe dieser Theorie erhalten wir in **Kapitel 5** einen weiteren Abstandsbegriff: die Wassersteindistanz. Nach einigen metrischen Eigenschaften ergeben sich Zusammenhänge mit der schwachen Konvergenz.

1 Signierte Maße

Maße haben einen entscheidenden Nachteil. Sie bilden lediglich einen positiven konvexen Kegel, das heißt, dass die Menge der Maße nur unter Bildung endlicher Summen und positiver Skalierung abgeschlossen ist. Unser Ziel in diesem Kapitel wird es sein, die sogenannten signierten Maße einzuführen, das heißt es wird auf die Positivität verzichtet, um somit eine Vektorraumstruktur zu erzeugen. Schließlich werden wir diesen Raum normieren und feststellen, dass es sich dabei sogar um einen Banachraum handelt. Sei im Verlauf des Kapitels $(\Omega, \mathfrak{A})$ ein messbarer Raum.

1.1 Einführung

In diesem Abschnitt werden wir zuerst die signierten Maße einführen und einige ihrer Eigenschaften darstellen. Wir orientieren uns dabei an [Els09, §1 in Kapitel VII].

Definition 1.1.1. *Signiertes Maß.* Eine Abbildung $\mu : \mathfrak{A} \to \mathbb{R}$ heißt *signiertes Maß* auf $(\Omega, \mathfrak{A})$, wenn folgende Eigenschaften erfüllt sind:

1. $\mu(\mathfrak{A}) \subset \mathbb{R} \cup \{-\infty\}$ oder $\mu(\mathfrak{A}) \subset \mathbb{R} \cup \{\infty\}$.
2. *σ-Additivität:* Für jede Folge $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ mit $A_i \cap A_j = \emptyset$ für alle $i \neq j$ konvergiert $\sum_{k=1}^{\infty} \mu(A_k)$ in $\mathbb{R} \cup \{-\infty, \infty\}$ und es gilt
$$\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k).$$
Hierbei setzen wir $\infty + a := \infty$ für alle $a \in \mathbb{R} \cup \{\infty\}$ und $-\infty + a = -\infty$ für alle $a \in \mathbb{R} \cup \{-\infty\}$. Den Fall $\infty - \infty$ schließen wir durch die erste Eigenschaft aus.
3. $\mu(\emptyset) = 0$.

Definition 1.1.2. Ein signiertes Maß μ heißt *endlich*, falls $\mu(\mathfrak{A}) \subset \mathbb{R}$. Die Menge der endlichen signierten Maße auf $(\Omega, \mathfrak{A})$ bezeichnen wir mit $\mathcal{M}(\mathfrak{A})$. Die Menge der endlichen Maße bezeichnen wir mit $\mathcal{M}^+(\mathfrak{A})$.

Bemerkung 1.1.3. Zu signierten Maßen.

1. Mit einem Korollar (siehe Satz A.1.1) aus dem Steinitzschen Umordnungssatz erhalten wir, dass für alle Folgen $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ die Reihe $\sum_{k=1}^{\infty} \mu(A_k)$ bereits absolut konvergiert.
2. Durch den Übergang von Maßen zu signierten Maßen verlieren wir Eigenschaften wie die Monotonie und die σ-Subadditivität. Ebenfalls müssen wir eine neuartige Definition der Nullmengen von μ vornehmen.
3. Da die mengenweise Addition[1] zweier signierten Maße im Allgemeinen nicht wohldefiniert ist, können wir auf der Menge der signierten Maße keine sinnvolle Vektorraumstruktur definieren. Falls jedoch mindestens eines der beiden Maße endlich ist, ist die Addition wohldefiniert. Insbesondere ist die Menge $\mathcal{M}(\mathfrak{A})$ ein Vektorraum über $\mathbb{R}$.

Definition 1.1.4. Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Eine Menge $A \in \mathfrak{A}$ heißt

1. *Nullmenge* unter μ, falls für jedes $B \subset A$ mit $B \in \mathfrak{A}$ bereits $\mu(B) = 0$ folgt.
2. *positiv* unter μ, falls für jedes $B \subset A$ mit $B \in \mathfrak{A}$ bereits $\mu(B) \geq 0$ folgt.
3. *negativ* unter μ, falls für jedes $B \subset A$ mit $B \in \mathfrak{A}$ bereits $\mu(B) \leq 0$ folgt.

Lemma 1.1.5. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Eine abzählbare Vereinigung von positiven (bzw. negativen) Mengen unter μ bleibt positiv (bzw. negativ).*

Beweis. Sei $(P_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ eine Folge von positiven Mengen. Setze

$$\tilde{P}_k := P_k \setminus \bigcup_{l=1}^{k-1} P_k$$

für jedes $k \in \mathbb{N}$. Dann ist $\tilde{P}_i \cap \tilde{P}_j = \emptyset$ für $i \neq j$ und $\bigcup_{k=1}^{\infty} P_k = \bigcup_{k=1}^{\infty} \tilde{P}_k$. Sei $A \subset \bigcup_{k=1}^{\infty} P_k$ $\mathfrak{A}$-messbar. Wegen $\tilde{P}_k \cap A \subset P_k$ und der Positivität von P_k gilt $\mu(\tilde{P}_k \cap A) \geq 0$ jedes $k \in \mathbb{N}$. Insgesamt erhalten wir

$$\mu(A) = \mu\left(\bigcup_{k=1}^{\infty} \tilde{P}_k \cap A\right) = \sum_{k=1}^{\infty} \mu(\tilde{P}_k \cap A) \geq 0.$$

[1] Wir setzen $(\mu_1 + \mu_2)(A) := \mu_1(A) + \mu_2(A)$ und $(\lambda\mu)(A) := \lambda\mu(A)$ für $A \in \mathfrak{A}$ und $\lambda \in \mathbb{R}$.

Falls $(P_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ eine Folge von negativen Mengen ist, gilt $\mu(\tilde{P}_k \cap A) \leq 0$ und damit auch $\mu(A) \leq 0$. $\square$

Eigenschaften 1.1.6. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$.*

1. *Sei $B \in \mathfrak{A}$ mit $\mu(B) \neq \pm\infty$. Dann folgt $\mu(A) \neq \pm\infty$ für jede $\mathfrak{A}$-messbare Menge $A \subset B$.*
2. *Sei $A, B \in \mathfrak{A}$ mit $A \subset B$ und $\mu(A) \neq \pm\infty$. Dann gilt $\mu(B \backslash A) = \mu(B) - \mu(A)$.*
3. *Falls $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$ steigend, das heißt es gilt $A_n \subset A_{n+1}$ für jedes $n \in \mathbb{N}$. Dann folgt $\mu(A_n) \to \mu(\bigcup_{k=1}^{\infty} A_k)$ für $n \to \infty$.*
4. *Falls $(A_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$ fallend, das heißt es gilt $A_n \supset A_{n+1}$ für jedes $n \in \mathbb{N}$ und falls es ein $n_0 \in \mathbb{N}$ mit $\mu(A_{n_0}) \neq \pm\infty$ gibt, dann folgt $\mu(A_n) \to \mu(\bigcap_{k=1}^{\infty} A_k)$ für $n \to \infty$.*

Die Eigenschaften 3. und 4. heißen σ-Stetigkeit von μ.

Beweis.

1. Angenommen $\mu(A) = \pm\infty$. Dann wäre mit der σ-Additivität
$$\mu(B) = \mu(A) + \mu(B \setminus A) = \pm\infty + \mu(B \setminus A) = \pm\infty,$$
was der Voraussetzung widerspricht.
2. Es gilt mit der σ-Additivität
$$\mu(B \setminus A) = \mu(B \setminus A) + \mu(A) - \mu(A) = \mu(B) - \mu(A).$$
3. Mit der σ-Additivität von μ gilt
$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} A_k\right) &= \mu\left(A_1 \cup \bigcup_{k=1}^{\infty} A_{k+1} \setminus A_k\right) = \mu(A_1) + \sum_{k=1}^{\infty} \mu(A_{k+1} \setminus A_k) \\
&= \lim_{n\to\infty} \mu(A_1) + \sum_{k=1}^{n-1} \mu(A_{k+1} \setminus A_k) \\
&= \lim_{n\to\infty} \mu\left(A_1 \cup \bigcup_{k=1}^{n-1} A_{k+1} \setminus A_k\right) \\
&= \lim_{n\to\infty} \mu\left(\bigcup_{k=1}^{n} A_k\right) = \lim_{n\to\infty} \mu(A_n).
\end{aligned}$$

4. O.B.d.A. ist $n_0 = 1$, sonst betrachte die Folge $(A_{n+n_0-1})_{n\in\mathbb{N}}$. Wir definieren die Folge $B_n := A_1 \setminus A_n$. Dann ist die Folge $(B_n)_{n\in\mathbb{N}}$ aufsteigend und wir erhalten mit der dritten Eigenschaft $\mu(B_n) \to \mu(\bigcup_{k=1}^\infty B_k)$ für $n \to \infty$. Mit der σ-Additivität gilt

$$\mu(A_1) = \mu(A_1 \setminus A_n) + \mu(A_n) = \mu(B_n) + \mu(A_n)$$

für jedes $n \in \mathbb{N}$. Mit der ersten Eigenschaft sind beide Summanden endlich. Wegen der Konvergenz von $\mu(B_n) \to \mu(\bigcup_{k=1}^\infty B_k)$ konvergiert also schon

$$\mu(A_n) \to \mu(A_1) - \mu\left(\bigcup_{k=1}^\infty B_k\right)$$

für $n \to \infty$. Wegen $\bigcup_{k=1}^\infty B_k \subset A_1$ können wir die zweite Eigenschaft anwenden und erhalten mit den De Morganschen Gesetzen, dass

$$\mu(A_1) - \mu\left(\bigcup_{k=1}^\infty B_k\right) = \mu(A_1) - \mu\left(A_1 \setminus \bigcap_{k=1}^\infty A_k\right) = \mu\left(\bigcap_{k=1}^\infty A_k\right)$$

gilt. Also folgt $\mu(A_n) \to \mu(\bigcap_{k=1}^\infty A_k)$ für $n \to \infty$.

□

1.2 Die Hahn-Jordan-Zerlegung

Falls wir eine reellwertige Funktion f betrachten, können wir f in einen positiven Anteil $f^+ := \max(f, 0)$ und einen negativen Anteil $f^- := -\min(f, 0)$ zerlegen, sodass $f = f^+ - f^-$ gilt. Unser Ziel in diesem Abschnitt wird es sein, ein signiertes Maß μ in Maße μ^+ und μ^- zu zerlegen, sodass analog $\mu = \mu^+ - \mu^-$ gilt. Wir orientieren uns beim folgenden Hilfslemma und beim eigentlichen Satz an [Els09, Lemma 1.7 und Satz 1.8 in Kapitel VII].

Lemma 1.2.1. Hilfslemma. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Sei $A \in \mathfrak{A}$ mit $\mu(A) > -\infty$ gegeben. Dann gibt es für jedes $l \in \mathbb{N}$ eine Menge $B \in \mathfrak{A}$, welche die Eigenschaften*

1. $B \subset A$
2. $\mu(A) \leq \mu(B)$

3. $\mu(C) \geq -\frac{1}{l}$ *für alle* $C \in \mathfrak{A}$ *mit* $C \subset B$

erfüllt.

Beweis. Angenommen die Behauptung ist falsch. Dann gibt es ein $l_0 \in \mathbb{N}$, sodass es für jede Menge $B \in \mathfrak{A}$ mit $B \subset A$ und $\mu(A) \leq \mu(B)$ eine $\mathfrak{A}$-messbare Menge $C \subset B$ mit $\mu(C) < -\frac{1}{l_0}$ gibt. Wir konstruieren nun eine Folge $(C_i)_{i\in\mathbb{N}} \in \mathfrak{A}$, die die Eigenschaften

1. $C_i \cap C_j = \emptyset$ für alle $i \neq j$
2. $\mu(C_i) < -\frac{1}{l_0}$
3. $C_i \subset A$ für alle $i \in \mathbb{N}$.

erfüllt. Die Konstruktion nehmen wir induktiv vor. Für $n = 1$ erhalten wir durch die Wahl von $B = A$ eine messbare Menge $C_1 \subset A$ mit $\mu(C_1) < -\frac{1}{l_0}$. Angenommen es sind bereits $C_1, \dots, C_n$ konstruiert. Durch die Wahl von $B = A \setminus \bigcup_{i=1}^n C_k \subset A$ gilt

$$\mu(A) < \mu(A) + \sum_{i=1}^n \frac{1}{l_0} < \mu(A) - \sum_{i=1}^n \mu(C_k) = \mu(A) - \mu\left(\bigcup_{i=1}^n C_k\right) = \mu\left(A \setminus \bigcup_{i=1}^n C_k\right)$$

und wir erhalten nach Annahme ein $C_{n+1} \subset B = A \setminus \bigcup_{i=1}^n C_k$, das disjunkt zu allen $C_1, \dots, C_n$ ist und $C_{n+1} < -\frac{1}{l_0}$ erfüllt. Wegen $C_i \subset A$ für jedes $i \in \mathbb{N}$ und $\mu(A) > -\infty$ erhalten wir mit

$$\begin{aligned}-\infty < \mu(A) &= \mu\left(\left(A \setminus \bigcup_{i=1}^\infty C_i\right) \cup \bigcup_{i=1}^\infty C_i\right) = \mu\left(A \setminus \bigcup_{i=1}^\infty C_i\right) + \sum_{i=1}^\infty \mu(C_i) \\ &< \mu\left(A \setminus \bigcup_{i=1}^\infty C_i\right) - \sum_{i=1}^\infty \frac{1}{l_0} = -\infty\end{aligned}$$

einen Widerspruch. □

Satz 1.2.2. Hahnscher Zerlegungssatz. *Sei* μ *ein signiertes Maß auf* $(\Omega, \mathfrak{A})$. *Dann gibt es disjunkte Mengen* $P \in \mathfrak{A}$ *positiv und* $N \in \mathfrak{A}$ *negativ, sodass* $\Omega = P \cup N$. *Die Aufteilung von* Ω *in* P *und* N *nennen wir* Hahn-Zerlegung.

Beweis. Setze

$$\alpha := \sup_{\substack{A\in\mathfrak{A} \\ \mu(A)\neq-\infty}} \mu(A).$$

Da $\emptyset \in \mathfrak{A}$ gilt sofort $\alpha \geq 0$. Wir nehmen $\alpha < \infty$ an, also dass $\mu(\mathfrak{A}) \subset \mathbb{R} \cup \{-\infty\}$ gilt. Ansonsten betrachte $-\mu \in \mathcal{M}^+(\mathfrak{A})$. Wir zeigen, dass wir eine Supremumsfolge aus positiven Mengen finden können. Wähle zunächst eine beliebige Supremumsfolge $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ mit $\mu(A_k) \neq -\infty$ sodass $\mu(A_k) \to \alpha$ für $k \to \infty$ konvergiert.

Sei $k \in \mathbb{N}$ fest. Wir konstruieren nun induktiv eine Folge $(D_l^k)_{l\in\mathbb{N}_0} \subset \mathfrak{A}$. Setze $D_0^k := A_k$. Seien $D_0^k, \dots, D_{l-1}^k \in \mathfrak{A}$ bereits konstruiert. Mit Lemma 1.2.1 erhalten wir ein $D_l^k \in \mathfrak{A}$ mit $D_l^k \subset D_{l-1}^k$ und $\mu(C) \geq -\frac{1}{l}$ für alle $C \in \mathfrak{A}$ mit $C \subset D_l^k$. Zusätzlich gilt

$$\mu(D_l^k) \geq \mu(D_{l-1}^k) \geq \ldots \geq \mu(D_0^k) = \mu(A_k). \tag{1.1}$$

Setze nun $P_k := \bigcap_{l=1}^{\infty} D_l^k \in \mathfrak{A}$. Wir zeigen, dass P_k eine positive Menge ist. Sei $C \subset P_k$ messbar. Nach Wahl der D_l^k gilt $\mu(C) \geq -\frac{1}{l}$ wegen $C \subset D_l^k$ für jedes $l \in \mathbb{N}$. Damit ist $\mu(C) \geq 0$ und P_k positiv. Betrachte mit der σ-Stetigkeit des signierten Maßes μ (Eigenschaften 1.1.6) und (1.1)

$$\alpha \geq \mu(P_k) = \lim_{l\to\infty} \mu(D_l^k) \geq \mu(A_k) \to \alpha$$

für $k \to \infty$. Damit folgt insbesondere, dass $\mu(P_k) \to \alpha$ für $k \to \infty$. Als abzählbare Vereinigung bleibt $P := \bigcup_{k=1}^{\infty} P_k$ mit Lemma 1.1.5 positiv. Setze $N := \Omega \setminus P \in \mathfrak{A}$. Wir zeigen, dass N negativ ist. Angenommen es gibt ein $A \subset N$ messbar mit $\mu(A) > 0$. Für jedes $n \in \mathbb{N}$ gilt

$$\alpha \geq \mu(P) = \mu((P \setminus P_n) \cup P_n) = \mu(P \setminus P_n) + \mu(P_n) \geq \mu(P_n) \to \alpha$$

für $n \to \infty$. Wir haben benutzt, dass $P \setminus P_n$ eine Teilmenge der positiven Menge P ist. Also $\mu(P) = \alpha$. Die Abschätzung

$$\alpha \geq \mu(P \cup A) = \mu(P) + \mu(A) = \alpha + \mu(A) > \alpha$$

liefert einen Widerspruch. $\square$

Bemerkung 1.2.3. Falls $\tilde{N}$ und $\tilde{P}$ eine weitere Hahn-Zerlegung von Ω unter einem signierten Maß μ ist, so folgt bereits, dass die symmetrischen Differenzen[2] $P \triangle \tilde{P}$ und $N \triangle \tilde{N}$ jeweils μ-Nullmengen sind.

Die nächsten Punkte basieren auf [Els09, Abschnitt 3 in Kapitel VII].

[2] $P \triangle \tilde{P} := P \setminus \tilde{P} \cup \tilde{P} \setminus P$

Definition 1.2.4. Seien μ, ν Maße auf $(\Omega, \mathfrak{A})$. μ und ν heißen zueinander *singulär*, falls es ein $A \in \mathfrak{A}$ gibt, sodass $\mu(A) = 0$ und $\nu(\Omega \setminus A) = 0$.

Definition 1.2.5. Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Sei P und N als Hahn-Zerlegung von Ω unter μ gegeben.

1. Das Maß
$$\mu^+ : \mathfrak{A} \to [0, \infty) \cup \{\infty\} \text{ mit } A \mapsto \mu(A \cap P)$$
heißt *positive Variation* von μ.
2. Das Maß
$$\mu^- : \mathfrak{A} \to [0, \infty) \cup \{\infty\} \text{ mit } A \mapsto -\mu(A \cap N)$$
heißt *negative Variation* von μ.
3. Das Maß $|\mu| := \mu^+ + \mu^-$ heißt *Variation von μ.*
4. Die Zerlegung $\mu = \mu^+ - \mu^-$ heißt *Jordan-Zerlegung* von μ.

Wegen Bemerkung 1.2.3 sind $\mu^+, \mu^-, |\mu|$ und damit auch die Jordan-Zerlegung unabhängig von der Wahl von N und P und damit wohldefiniert. μ^+ und μ^- sind zueinander singulär.

Lemma 1.2.6. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Weiterhin seien Maße μ_1, μ_2 auf $(\Omega, \mathfrak{A})$ gegeben, sodass $\mu = \mu_1 - \mu_2$ gilt und μ_1 und μ_2 zueinander singulär sind. Dann folgt bereits $\mu_1 = \mu^+$ und $\mu_2 = \mu^-$.*

Beweis. Da μ_1 und μ_2 zueinander singulär sind, gibt es nach Definition eine messbare Menge P, sodass $\mu_1(\Omega \setminus P) = 0$ und $\mu_2(P) = 0$ gilt. Wir setzen $N := \Omega \setminus P$ und zeigen, dass P und N die Hahn-Zerlegung von Ω ist. Sei B eine messbare Teilmenge von P. Dann ist B eine μ_2-Nullmenge und wir erhalten

$$\mu(B) = \mu_1(B) - \mu_2(B) = \mu_1(B) \geq 0.$$

Also ist P eine positive Menge und analog N eine negative Menge. Da $A \cap (\Omega \setminus P)$ eine μ_1-Nullmenge und $A \cap P$ eine μ_2-Nullmenge ist, folgt

$$\begin{aligned}\mu^+(A) = \mu(A \cap P) &= \mu_1(A \cap P) - \mu_2(A \cap P) = \mu_1(A \cap P) \\ &= \mu_1(A \cap P) + \mu_1(A \cap (\Omega \setminus P)) = \mu_1(A)\end{aligned}$$

für $A \in \mathfrak{A}$. Damit ist $\mu^+ = \mu_1$ und analog $\mu^- = \mu_2$. $\square$

Eigenschaft 1.2.7. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Dann gilt die Ungleichung $|\mu(A)| \leq |\mu|(A)$ für jedes $A \in \mathfrak{A}$.*

Beweis. Die Aussage folgt direkt mit der Definition der Variation von μ.

$$|\mu(A)| = |\mu^+(A) - \mu^-(A)| \leq |\mu^+(A)| + |\mu^-(A)| = \mu^+(A) + \mu^-(A) = |\mu|(A)$$

□

Um die Variation von einem signierten Maß μ an einem $A \in \mathfrak{A}$ auszuwerten, muss die Hahn-Zerlegung von Ω unter μ bekannt sein. Wir zeigen nun eine Gleichung, mit der die Variation - auch ohne den Zerlegungssatz zu benutzen - definiert werden könnte.

Lemma 1.2.8. *Sei μ ein signiertes Maß auf $(\Omega, \mathfrak{A})$. Dann gilt für jedes $A \in \mathfrak{A}$:*

$$|\mu|(A) = \max \left\{ \sum_{k=1}^{\infty} |\mu(A_k)| : (A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}, \bigcup_{k\in\mathbb{N}} A_k = A, A_j \cap A_i = \emptyset \textit{ für alle } i \neq j \right\}$$

Beweis. Sei $A \in \mathfrak{A}$ fest. Sei $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ mit $\bigcup_{k\in\mathbb{N}} A_k = A$ paarweise disjunkt. Dann gilt mit Eigenschaft 1.2.7 und der Sigma-Additivität von $|\mu|$

$$\sum_{k=1}^{\infty} |\mu(A_k)| \leq \sum_{k=1}^{\infty} |\mu|(A_k) = |\mu|(A).$$

Durch Supremumsbildung über alle Partitionen von A erhalten wir bereits die untere Abschätzung für $|\mu|(A)$. Sei andererseits $P, N \in \mathfrak{A}$ die Hahn-Zerlegung von Ω unter μ. Betrachte nun die Folge

$$A_1 := P \cap A,\ A_2 := N \cap A \text{ und } A_k := \emptyset \text{ für alle } k \geq 3.$$

Dann bildet $(A_k)_{k\in\mathbb{N}}$ eine Partition von A. Wegen der Positivität von P und der Negativität von N gilt $\mu(A_1) \geq 0$ und $\mu(A_2) \leq 0$ und wir erhalten

$$\begin{aligned} |\mu(A)| &= \mu^+(A) + \mu^-(A) = \mu(A_1) - \mu(A_2) \\ &= |\mu(A_1)| + |\mu(A_2)| = \sum_{k=1}^{\infty} |\mu(A_k)|. \end{aligned}$$

Damit wird das Supremum angenommen und es handelt sich tatsächlich um ein Maximum. □

Definition 1.2.9. Die Menge der signierten Maße auf $(\Omega, \mathfrak{A})$ statten wir mit der Relation $\leq$ aus. Dabei setzen wir $\mu_1 \leq \mu_2$ genau dann, wenn $\mu_1(A) \leq \mu_2(A)$ für alle $A \in \mathfrak{A}$ gilt. Diese Relation auf den signierten Maßen bildet eine partielle Ordnung gemäß Definition A.1.2.

1.3 Der Banachraum der endlichen signierten Maße

Wir werden nun den Vektorraum $\mathcal{M}(\mathfrak{A})$ mit einer passenden Norm ausstatten und die Vollständigkeit zeigen. Zunächst zeigen wir zusätzliche Eigenschaften der Variation der signierten Maße.

Definition 1.3.1. *Totalvariationsabstand.* Wir setzen

$$\|\cdot\|_{\mathrm{TV}} : \mathcal{M}(\mathfrak{A}) \to [0, \infty) \text{ mit } \mu \mapsto |\mu|(\Omega).$$

Diese Definition nehmen wir nur auf den endlichen signierten Maßen vor, da eine Norm nach Definition nicht den Wert $+\infty$ annehmen soll.

Eigenschaft 1.3.2. *Seien μ und ν signierte Maße auf $(\Omega, \mathfrak{A})$ und $\lambda \in \mathbb{R}$. Dann gilt $|\lambda\mu| = |\lambda||\mu|$ und $|\mu + \nu| \leq |\mu| + |\nu|$.*

Beweis. Folgt direkt mit Lemma 1.2.8, der Dreiecksungleichung der Betragsfunktion und der Subadditivität des Supremums. □

Wir zeigen nun, dass es sich beim Totalvariationsabstand tatsächlich um eine Norm auf $\mathcal{M}^+(\mathfrak{A})$ handelt:

Satz 1.3.3. $(\mathcal{M}(\mathfrak{A}), \|\cdot\|_{TV})$ *bildet einen normierten Raum.*

Beweis. Für die Definitheit sei $\mu \in \mathcal{M}(\mathfrak{A})$ mit $\|\mu\|_{\mathrm{TV}} = 0$. Also ist

$$0 = \|\mu\|_{\mathrm{TV}} = |\mu|(\Omega) = \mu^+(\Omega) + \mu^-(\Omega).$$

Da μ^+ und μ^- Maße sind, folgt $\mu^+(\Omega) = 0 = \mu^-(\Omega)$. Also gilt $\mu = \mu^+ - \mu^- = 0$. Homogenität und Dreiecksungleichung folgen direkt mit Eigenschaft 1.3.2 durch Einsetzen von Ω. □

Lemma 1.3.4. *Auf $\mathcal{M}(\mathfrak{A})$ sind die Normen $\|\cdot\|_{TV}$ und $\|\cdot\|_\infty$ äquivalent, wobei*

$$\|\cdot\|_\infty : \mathcal{M}(\mathfrak{A}) \to \mathbb{R} \textit{ mit } \mu \mapsto \sup_{A\in\mathfrak{A}} |\mu(A)|.$$

Das bedeutet, Konvergenz in der $\|\cdot\|_{TV}$-Norm lässt sich durch die gleichmäßige Konvergenz über alle messbaren Mengen charakterisieren.

Beweis. Mit Eigenschaft 1.2.7 und der Monotonie der Variation von μ erhalten wir

$$\|\mu\|_\infty = \sup_{A\in\mathfrak{A}} |\mu(A)| \le \sup_{A\in\mathfrak{A}} |\mu|(A) \le |\mu|(\Omega) = \|\mu\|_{\mathrm{TV}}.$$

Betrachte andererseits für μ die Hahn-Zerlegung P und N. Wegen $\mu(P) \ge 0$ und $\mu(N) \le 0$ gilt

$$\|\mu\|_{\mathrm{TV}} = \mu(P) - \mu(N) = |\mu(P)| + |\mu(N)| \le 2 \sup_{A\in\mathfrak{A}} |\mu(A)| = 2\|\mu\|_\infty.$$

Somit sind die Normen äquivalent. □

Satz 1.3.5. *$(\mathcal{M}(\mathfrak{A}), \|\cdot\|_{TV})$ bildet einen Banachraum.*

Beweis. Zu zeigen ist die Vollständigkeit von $\mathcal{M}(\mathfrak{A})$ unter der $\|\cdot\|_{\mathrm{TV}}$ Norm. Sei hierzu $(\mu_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathcal{M}(\mathfrak{A})$. Sei $\varepsilon > 0$. Es folgt die Existenz eines $n \in \mathbb{N}$, sodass für alle $k, l \ge n$ und alle $A \in \mathfrak{A}$

$$\begin{aligned} |\mu_k(A) - \mu_l(A)| &= |(\mu_k - \mu_l)(A)| \le |\mu_k - \mu_l|(A) \\ &\le |\mu_k - \mu_l|(\Omega) = \|\mu_k - \mu_l\|_{\mathrm{TV}} < \varepsilon \end{aligned} \tag{1.2}$$

gilt. Damit ist $(\mu_k(A))_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$ und wir erhalten wegen der Vollständigkeit von $\mathbb{R}$ die Wohldefiniertheit von

$$\mu : \mathfrak{A} \to \mathbb{R} \text{ mit } A \mapsto \lim_{k\to\infty} \mu_k(A).$$

Für $k \to \infty$ erhalten wir mit der Stetigkeit des Betrages $|\mu(A) - \mu_l(A)| \le \varepsilon$ für jedes $A \in \mathfrak{A}$ und $l \ge n$. Das heißt also, dass μ_k gleichmäßig über alle messbaren Mengen konvergiert. Falls wir also zeigen, dass es sich bei μ um ein endliches, signiertes Maß handelt, folgt mit Lemma 1.3.4 die Konvergenz von $\mu_k \to \mu$ in der $\|\cdot\|_{\mathrm{TV}}$-Norm für $k \to \infty$.

Als erstes erhalten wir $\mu(\emptyset) = \lim_{k\to\infty} \mu_k(\emptyset) = 0$. Für die σ-Additivität betrachte nun $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{A}$ paarweise disjunkt. Schreibe $A := \bigcup_{i=1}^{\infty} A_i$. Da es sich bei μ_n um ein signiertes Maß handelt, konvergiert die Reihe $\sum_{i=1}^{\infty} \mu_n(A_i)$. Also gibt es ein $\tilde{n} \in \mathbb{N}$, sodass

$$\begin{aligned}\left|\mu_n(A) - \sum_{i=1}^{k} \mu_n(A_i)\right| &= \left|\mu_n\left(\bigcup_{i=1}^{\infty} A_i\right) - \sum_{i=1}^{k} \mu_n(A_i)\right| \\ &= \left|\sum_{i=1}^{\infty} \mu_n(A_i) - \sum_{i=1}^{k} \mu_n(A_i)\right| < \varepsilon \qquad (1.3)\end{aligned}$$

für jedes $k \geq \tilde{n}$ gilt. Wir zeigen nun die Konvergenz der Reihe $\sum_{i=1}^{\infty} \mu(A_i)$ gegen $\mu(\bigcup_{i=1}^{\infty} A_i) = \mu(A)$. Mit der Dreiecksungleichung, (1.2) und (1.3) gilt

$$\begin{aligned}&\left|\sum_{i=1}^{k} \mu(A_i) - \mu(A)\right| \\ &\leq \left|\sum_{i=1}^{k} \mu(A_i) - \sum_{i=1}^{k} \mu_n(A_i)\right| + \left|\sum_{i=1}^{k} \mu_n(A_i) - \mu_n(A)\right| + |\mu_n(A) - \mu(A)| \\ &< \left|\sum_{i=1}^{k} \mu(A_i) - \sum_{i=1}^{k} \mu_n(A_i)\right| + 2\varepsilon = \lim_{j\to\infty} \left|\sum_{i=1}^{k} \mu_j(A_i) - \sum_{i=1}^{k} \mu_n(A_i)\right| + 2\varepsilon \\ &= \lim_{j\to\infty} \left|\mu_j\left(\bigcup_{i=1}^{k} A_i\right) - \mu_n\left(\bigcup_{i=1}^{k} A_i\right)\right| + 2\varepsilon \\ &= \left|\mu\left(\bigcup_{i=1}^{k} A_k\right) - \mu_n\left(\bigcup_{i=1}^{k} A_k\right)\right| + 2\varepsilon \leq 3\varepsilon\end{aligned}$$

für jedes $k \geq \max\{n, \tilde{n}\}$. Also folgt die $\sigma-$ Subadditivität von μ. Da $\mu_k(\mathfrak{A}) \subset \mathbb{R}$ für jedes $k \in \mathbb{N}$, gilt $\mu(\mathfrak{A}) \subset \mathbb{R}$. Somit ist $\mu \in \mathcal{M}(\mathfrak{A})$. Der Beweis ist angelehnt an [Els09, Satz 1.14 in Kapitel VII]. □

Bemerkung 1.3.6. Die Menge der signierten Maße $\mathcal{M}(\mathfrak{A})$ bildet einen Banach-Verband (siehe [Are14]). Weiterhin ist der konvexe Kegel

$$\{\mu \in \mathcal{M}(\mathfrak{A}) : \mu \geq 0\} = \mathcal{M}^+(\mathfrak{A}) \subset \mathcal{M}(\mathfrak{A})$$

generierend, das heißt es gilt

$$\mathcal{M}(\mathfrak{A}) = \mathcal{M}^+(\mathfrak{A}) - \mathcal{M}^+(\mathfrak{A}) := \{\mu_1 - \mu_2 : \mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{A})\}.$$

1.4 Die Totalvariationsnorm von signierten Maßen mit Dichten

Häufig werden signierte Maße betrachtet, die eine Dichte bezüglich einem gegebenen Maß besitzen. Im folgenden Abschnitt möchten wir Konvergenz dieser Maße mit der Konvergenz ihrer Dichten charakterisieren.

Lemma 1.4.1. *Sei μ ein Maß auf $(\Omega, \mathfrak{A})$ und $f : \Omega \to \mathbb{R} \cup \{\infty, -\infty\}$ quasi-integrierbar[3]. Dann ist die Abbildung $\nu : \mathfrak{A} \to \mathbb{R} \cup \{-\infty, \infty\}$ mit $A \mapsto \int_A f \, \mathrm{d}\mu$ ein signiertes Maß. Wir bezeichnen ν mit $f \odot \mu$.*

Beweis. Aus der Maßtheorie ist bekannt, dass die Abbildungen $\nu_1(A) := \int_A f^+ \, \mathrm{d}\mu$ und $\nu_2(A) := \int_A f^- \, \mathrm{d}\mu$ Maße auf $(\Omega, \mathfrak{A})$ sind, wobei nach Voraussetzung mindestens eines der beiden Maße endlich ist. Wegen Bemerkung 1.1.3 ist $\nu = \nu_1 - \nu_2$ ein Maß. $\square$

Eigenschaften 1.4.2. *Sei μ ein Maß auf $(\Omega, \mathfrak{A})$ und $f \in L^1(\Omega, \mathfrak{A}, \mu)$. Dann gelten die Gleichheiten*

1. $(f \odot \mu)^+ = f^+ \odot \mu$,
2. $(f \odot \mu)^- = f^- \odot \mu$,
3. $|f \odot \mu| = |f| \odot \mu$.

Damit ist insbesondere $\|f \odot \mu\|_{TV} = \|f\|_{L^1}$, das heißt eine gegebene Funktionenfolge $(f_k)_{k\in\mathbb{N}} \subset L^1(\Omega, \mathfrak{A}, \mu)$ konvergiert genau dann in L^1 gegen ein $f \in L^1(\Omega, \mathfrak{A}, \mu)$, wenn die Folge $(f_k \odot \mu)_{k\in\mathbb{N}}$ gegen $f \odot \mu$ in der $\| \cdot \|_{TV}$ Norm konvergiert.

Beweis. Wir setzen[4]

$$P := \{x \in \Omega : f(x) \geq 0\} \text{ und } N := \{x \in \Omega : f(x) < 0\}.$$

Für $B \subset P$ messbar gilt unmittelbar $(f \odot \mu)(B) = \int_B f \, \mathrm{d}\mu \geq 0$. Also ist P eine positive Menge bezüglich dem signierten Maß $f \odot \mu$. Analog ist N negativ. Also ist P und N bereits die Hahn-Zerlegung von $f \odot \mu$. Wir erhalten

$$(f \odot \mu)^+(A) = (f \odot \mu)(A \cap P) = \int_{A\cap P} f \, \mathrm{d}\mu = \int_A f^+ \, \mathrm{d}\mu = (f^+ \odot \mu)(A)$$

[3] Mindestens $\int f^+ \, \mathrm{d}\lambda$ oder $\int f^- \, \mathrm{d}\lambda$ ist endlich, wobei $f^+ := \max(f, 0)$ und $f^- := -\min(f, 0)$.

[4] Die Definition der beiden Mengen macht streng genommen keinen Sinn: Das Objekt f ist eine Äquivalenzklasse und Punktauswertungen lassen sich nicht sinnvoll definieren. Deshalb behandeln wir f als einen Vertreter dieser Äquivalenzklasse.

für $A \in \mathfrak{A}$. Analog folgt die Gleichung $(f \odot \mu)^- = f^- \odot \mu$ und damit auch

$$\begin{aligned} |f \odot \mu|(A) &= (f \odot \mu)^+(A) + (f \odot \mu)^-(A) = (f^+ \odot \mu)(A) + (f^- \odot \mu)(A) \\ &= \int_A f^+ \,\mathrm{d}\mu + \int_A f^- \,\mathrm{d}\mu = \int_A |f| \,\mathrm{d}\mu = (|f| \odot \mu)(A) \end{aligned}$$

für $A \in \mathfrak{A}$. □

Mit dem Umweg über die Theorie der signierten Maße haben wir nun mit Lemma 1.3.4 und mit Lemma 1.4.2 folgendes Resultat gezeigt:

Korollar 1.4.3. *Sei μ ein Maß auf $(\Omega, \mathfrak{A})$. Sei weiterhin $(f_k)_{k\in\mathbb{N}}$ eine Folge in $L^1(\Omega, \mathfrak{A}, \mu)$ und $f \in L^1(\Omega, \mathfrak{A}, \mu)$. Dann sind äquivalent:*

1. $f_k \to f$ *in* $L^1(\Omega, \mathfrak{A}, \mu)$ *für* $k \to \infty$.
2. $\int_A f \,\mathrm{d}\mu \to \int_A f_k \,\mathrm{d}\mu$ *für* $k \to \infty$ *gleichmäßig in* $A \in \mathfrak{A}$*, das heißt*

$$\sup_{A\in\mathfrak{A}} \left| \int_A f_k \,\mathrm{d}\mu - \int_A f \,\mathrm{d}\mu \right| \to 0$$

für $k \to \infty$.

Beweis. Betrachte $\nu_k := f_k \odot \mu$ und benutze die genannten Lemmata. □

1.5 Das Lebesgueintegral bezüglich signierten Maßen

Abschließend im Kapitel über signierte Maße werden wir noch das Integral bezüglich eines signierten Maßes anreißen. Diese Theorie benötigen wir für Kapitel 3, verzichten aber auf eine detaillierte Ausführung.

Definition 1.5.1. Sei $f : M \to \mathbb{R} \cup \{-\infty, \infty\}$ messbar. Falls das Integral über die positive Variation $\int f \,\mathrm{d}\mu^+$ und/oder über die negative Variation $\int f \,\mathrm{d}\mu^+$ endlich ist, setzen wir

$$\int f \,\mathrm{d}\mu := \int f \,\mathrm{d}\mu^+ - \int f \,\mathrm{d}\mu^-.$$

Falls μ bereits ein Maß ist, stimmt diese Definition offensichtlich mit der bisherigen Definition überein.

Bemerkung 1.5.2. Der Definition ist noch hinzuzufügen:

1. Das Integral ist sowohl in der Komponente der zu integrierenden Funktion, als auch in der Komponente des Maßes linear.
2. Das Integral ist im Allgemeinen in beiden Komponenten nicht monoton, es sei denn wir fordern für die jeweils andere Komponente Positivität.
3. Im Allgemeinen folgt aus $\int f \,\mathrm{d}\mu = 0$ nicht $f = 0$ für $|\mu|$-fast alle $x \in M$.
4. Für jedes $A \in \mathfrak{A}$ gilt

$$\begin{aligned}\mu(A) = \mu^+(A) - \mu^-(A) &= \int \mathbb{1}_A \,\mathrm{d}\mu^+ - \int \mathbb{1}_A \,\mathrm{d}\mu^- \\ &= \int \mathbb{1}_A \,\mathrm{d}(\mu^+ - \mu^-) = \int \mathbb{1}_A \,\mathrm{d}\mu.\end{aligned}$$

Eigenschaft 1.5.3. *Sei* $\mu \in \mathcal{M}(\mathfrak{A})$*,* $f \in \mathcal{M}(\Omega)$ *und* $A \in \mathfrak{A}$*. Dann gilt*

$$\left|\int_A f \,\mathrm{d}\mu\right| \leq |\mu|(A)\|f\|_\infty.$$

Beweis.

$$\begin{aligned}\left|\int_A f \,\mathrm{d}\mu\right| &= \left|\int_A f \,\mathrm{d}\mu^+ - \int_A f \,\mathrm{d}\mu^-\right| \leq \left|\int_A f \,\mathrm{d}\mu^+\right| + \left|\int_A f \,\mathrm{d}\mu^-\right| \\ &\leq \int_A \|f\|_\infty \,\mathrm{d}\mu^+ + \int_A \|f\|_\infty \,\mathrm{d}\mu^- \\ &= (\mu^+ + \mu^-)(A)\|f\|_\infty = |\mu|(A)\|f\|_\infty\end{aligned}$$

□

2 Schwache Konvergenz von Maßen

In Kapitel 1 haben wir bereits die Menge der endlichen signierten Maße $\mathcal{M}(\mathfrak{A})$ mit der durch die $\|\cdot\|_{\mathrm{TV}}$-Norm induzierten Topologie ausgestattet. Wir werden sehen, dass der daraus resultierende Konvergenzbegriff sehr restriktiv ist und bei vielen Anwendungen zu einer schwächeren Topologie übergegangen werden muss. Hierbei listen wir einige Charakterisierungen zur Konvergenz, die das Theorem von Portmanteau liefert, auf. Zusätzlich untersuchen wir die Topologie auf Metrisierbarkeit und führen den Satz von Prochorov aus, welcher die schwach kompakten Mengen charakterisiert.

Diese Theorie benötigt einen metrischen Raum und die zugehörige Borelsche σ-Algebra $\mathfrak{B}$ als messbaren Raum. Sei im Verlauf des Kapitels (M,d) ein metrischer Raum. Mit $\mathcal{M}^+(\mathfrak{B})$ bezeichnen wir die Menge der endlichen Maße auf $(\Omega,\mathfrak{B}) = (\Omega,\mathfrak{B}(\Omega))$. $C_b(M)$ sei die Menge der stetigen und beschränkten reellwertigen Funktionen auf M. Zusammen mit der Supremumsnorm $\|\cdot\|_\infty$ bildet $C_b(M)$ einen Banachraum.

2.1 Einführung der schwachen Topologie

Die nachfolgende Definition ist unter anderem in [Kle06, Definition 13.12] zu finden.

Definition 2.1.1. Sei die Abbildung $T_f : \mathcal{M}^+(\mathfrak{B}) \to \mathbb{R}$ mit $\nu \mapsto \int f \,\mathrm{d}\nu$ für $f \in C_b(M)$ gegeben. Wir bezeichnen die Initialtopologie[1] von $(T_f)_{f\in C_b(M)}$ als *schwache Topologie* auf $\mathcal{M}^+(\mathfrak{B})$. Damit konvergiert eine Folge $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ genau dann gegen ein $\mu \in \mathcal{M}^+(\mathfrak{B})$, wenn $\int f \,\mathrm{d}\mu_n \to \int f \,\mathrm{d}\mu$ für $n \to \infty$ und für alle $f \in C_b(M)$ gilt.

[1] Die gröbste Topologie, sodass T_f für jedes $f \in C_b(M)$ stetig ist.

Bemerkung 2.1.2. Im Folgenden möchten wir diese Topologie besser verstehen:

1. Falls $(\mu_k)_{k\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ schwach gegen $\mu \in \mathcal{M}^+(\mathfrak{B})$ konvergiert, folgt die Konvergenz der Gesamtmasse $\mu_k(M) \to \mu(M)$ für $k \to \infty$. Dies folgt direkt, wenn man mit $f = \mathbb{1}_M \in C_b(M)$ testet.

2. Falls wir von starker Konvergenz sprechen, meinen wie die Konvergenz in der $\|\cdot\|_{\mathrm{TV}}$-Norm. Aus der starken Konvergenz folgt die schwache Konvergenz, wie man leicht an folgender Abschätzung sehen kann:

$$\left|\int f \,\mathrm{d}\mu_n - \int f \,\mathrm{d}\mu\right| = \left|\int f \,\mathrm{d}(\mu_n - \mu)\right|$$
$$\leq \int \|f\|_\infty \,\mathrm{d}|\mu_n - \mu| = \|f\|_\infty \|\mu_n - \mu\|_{\mathrm{TV}}$$

   Wir haben die Linearität im Maß (Bemerkung 1.5.2) und die Abschätzung aus Eigenschaft 1.5.3 benutzt.

3. In Kapitel 3 wird sich herausstellen, dass es sich - falls (M, d) ein kompakter metrischer Raum ist - bei der schwachen Konvergenz auf $\mathcal{M}^+(\mathfrak{B})$ eigentlich um die schwach-* Konvergenz auf $C_b(M)'$ handelt. Da $C_b(M)$ im Allgemeinen nicht reflexiv ist, entspricht die schwache Konvergenz auf $\mathcal{M}^+(\mathfrak{B})$ also nicht dem Begriff der schwachen Konvergenz aus der Funktionalanalysis, bei dem mit Elementen aus dem Dualraum $\mathcal{M}^+(\mathfrak{B})'$ getestet wird.

4. Die aus der Wahrscheinlichkeitsrechnung bekannte Konvergenz von Zufallsvariablen in Verteilung entspricht der schwachen Konvergenz der zugehörigen Verteilungen.

Satz 2.1.3. *Die Menge $C_b(M)$ trennt den Raum der endlichen Maße $\mathcal{M}^+(\mathfrak{B})$. Das heißt, dass für alle $\mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{B})$ aus*

$$\int f \,\mathrm{d}\mu_1 = \int f \,\mathrm{d}\mu_2 \text{ für alle } f \in C_b(M)$$

bereits $\mu_1 = \mu_2$ folgt.

Beweis. Sei $\mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{B})$ und $A \subset M$ nicht leer und abgeschlossen. Mit Lemma $A.2.5$ erhalten wir eine gegen $\mathbb{1}_A$ punktweise konvergente Funktionenfolge $(f_k)_{k\in\mathbb{N}} \subset C_b(M)$ mit $f_k(M) \subset [0,1]$. Da μ_1 und μ_2 endliche Maße sind und die Funktionenfolge $(f_k)_{k\in\mathbb{N}}$ gleichmäßig durch $\mathbb{1}_M \in L^1(M, \mathfrak{B}, \mu_1) \cap L^1(M, \mathfrak{B}, \mu_2)$

beschränkt ist, können wir den Satz von Lebesgue anwenden und erhalten

$$\mu_1(A) = \int \mathbb{1}_A \,\mathrm{d}\mu_1 = \lim_{k\to\infty} \int f_k \,\mathrm{d}\mu_1 = \lim_{k\to\infty} \int f_k \,\mathrm{d}\mu_2 = \int \mathbb{1}_A \,\mathrm{d}\mu_2 = \mu_2(A).$$

Damit stimmt μ_1 und μ_2 auf allen abgeschlossenen Mengen überein. Die Aussage folgt dann mit dem Maßeindeutigkeitssatz, der Schnittstabilität abgeschlossener Mengen und der Endlichkeit von μ_1 und μ_2. Ein ähnlicher Beweis ist in [Els09, Satz 4.6 in Kapitel VIII] zu finden. □

Korollar 2.1.4. *Die schwache Topologie auf $\mathcal{M}^+(\mathfrak{B})$ bildet einen Hausdorffraum.*

Beweis. Seien $\mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{B})$ mit $\mu_1 \neq \mu_2$. Mit Satz 2.1.3 folgt die Existenz eines $f_0 \in C_b(M)$, sodass

$$\int f_0 \,\mathrm{d}\mu_1 \neq \int f_0 \,\mathrm{d}\mu_2$$

gilt. Setze $\varepsilon_0 := \frac{1}{2}|\int f_0 \,\mathrm{d}\mu_1 - \int f_0 \,\mathrm{d}\mu_2| > 0$ und definiere

$$\begin{aligned} U_1 &:= \left\{\mu \in \mathcal{M}^+(\mathfrak{B}) : \left|\int f_0 \,\mathrm{d}\mu_1 - \int f_0 \,\mathrm{d}\mu\right| < \varepsilon_0\right\} \text{ und} \\ U_2 &:= \left\{\mu \in \mathcal{M}^+(\mathfrak{B}) : \left|\int f_0 \,\mathrm{d}\mu_2 - \int f_0 \,\mathrm{d}\mu\right| < \varepsilon_0\right\}. \end{aligned}$$

Wir zeigen zunächst, dass $U_1 \cap U_2 = \emptyset$. Angenommen es gäbe ein $\mu \in \mathcal{M}^+(\mathfrak{B})$ mit $\mu \in U_1 \cap U_2$. Mit der Dreiecksungleichung ergibt sich der Widerspruch

$$\begin{aligned} \varepsilon_0 &= \frac{1}{2}\left|\int f_0 \,\mathrm{d}\mu_1 - \int f_0 \,\mathrm{d}\mu_2\right| \\ &\leq \frac{1}{2}\left|\int f_0 \,\mathrm{d}\mu_1 - \int f_0 \,\mathrm{d}\mu\right| + \frac{1}{2}\left|\int f_0 \,\mathrm{d}\mu_2 - \int f_0 \,\mathrm{d}\mu\right| < \frac{1}{2}\varepsilon_0 + \frac{1}{2}\varepsilon_0 = \varepsilon_0. \end{aligned}$$

Als Nächstes zeigen wir, dass U_1 und U_2 offene Mengen bezüglich der schwachen Topologie auf $\mathcal{M}^+(\mathfrak{B})$ sind. Nach Definition der schwachen Topologie ist die Abbildung $T_{f_0} : \mathcal{M}^+(\mathfrak{B}) \to \mathbb{R}$ gegeben mit $\mu \mapsto \int f_0 \,\mathrm{d}\mu$ stetig. Die Funktion

$$S : \mathcal{M}^+(\mathfrak{B}) \to \mathbb{R} \text{ mit } \mu \mapsto \left|\int f_0 \,\mathrm{d}\mu_1 - T_{f_0}(\mu)\right| = \left|\int f_0 \,\mathrm{d}\mu_1 - \int f_0 \,\mathrm{d}\mu\right|$$

ist als Verkettung stetiger Funktionen wiederum selbst stetig. Es ergibt sich für

U_1 die Darstellung

$$U_1 = \left\{ \mu \in \mathcal{M}^+(\mathfrak{B}) : \left| \int f_0 \, \mathrm{d}\mu_1 - \int f_0 \, \mathrm{d}\mu \right| < \varepsilon_0 \right\} = S^{-1}((-\infty, \varepsilon_0)),$$

wodurch U_1 (und genauso U_2) als Urbild einer offenen Menge unter einer stetigen Funktion selbst offen ist. Da offensichtlich $\mu_1 \in U_1$ und $\mu_2 \in U_2$, haben wir die Hausdorffeigenschaft gezeigt. Die Argumente aus diesem Beweis sind in [Els09, Definition 4.5 in Kapitel VIII] nachzulesen. □

Bemerkung 2.1.5. Wir werden später, unter der Annahme, dass M separabel ist, die schwache Topologie metrisieren. Dadurch würde automatisch die Hausdorffeigenschaft folgen.

2.2 Beispiele von schwach konvergenten Folgen

Beispiel 2.2.1. Sei $(a_k)_{k\in\mathbb{N}} \subset M$ eine gegen $a \in M$ konvergente Folge. Betrachte δ_{a_k}, wobei δ_x dem Dirac-Maß im Punkt x entspricht. Wir erwarten, dass δ_{a_k} auf natürliche Art und Weise gegen δ_a konvergiert. Dies ist aufgrund der Stetigkeit von $f \in C_b(M)$ der Fall:

$$\int_M f \, \mathrm{d}\delta_{a_k} = f(a_k) \to f(a) = \int_M f \, \mathrm{d}\delta_a \text{ für } k \to \infty$$

Beispiel 2.2.2. Wir betrachten als metrischen Raum die reellen Zahlen zusammen mit dem Absolutbetrag. Mit der Wahl von $a_k := \frac{1}{k}$ sieht man leicht, dass aus der schwachen Konvergenz im Allgemeinen nicht die mengenweise Konvergenz folgt. Es gilt zum Beispiel

$$\delta_{1/k}(\{0\}) = 0 \text{ für jedes } k \in \mathbb{N},$$

dennoch ist $\delta_0(\{0\}) = 1$. Insbesondere konvergiert die Folge $(\delta_{1/k})_{k\in\mathbb{N}}$ nicht in der $\|\cdot\|_{\mathrm{TV}}$-Norm, da einerseits aus dem starken Konvergenzbegriff die mengenweise Konvergenz folgen würde (siehe Lemma 1.3.4). Andererseits gilt

$$\|\delta_{1/k} - \delta_{1/l}\|_{\mathrm{TV}} = \delta_{1/k}(M) + \delta_{1/l}(M) = 2$$

für alle $k \neq l$.

Beispiel 2.2.3. Betrachte das kompakte Intervall $[0,1]$ zusammen mit der Betragsfunktion. Sei $\lambda : \mathfrak{B}([0,1]) \to [0,1]$ das eindimensionale Lebesguemaß. Wähle

$$\mu_n := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\frac{k}{n}}$$

für $n \in \mathbb{N}$. Das Lebesguemaß und die Maße $(\mu_n)_{n\in\mathbb{N}}$ haben eine völlig unterschiedliche Struktur. Dennoch konvergiert μ_n schwach gegen λ. Beachte, dass alle $f \in C_b([0,1])$ Riemann-Integrierbar sind. Wähle die ausgezeichnete[2] Partitionsfolge

$$\pi_n : 0 = \frac{0}{n} < \frac{1}{n} < \ldots < \frac{k}{n} < \ldots < \frac{n}{n} = 1$$

mit Zwischenstellen $\xi_k^n := \frac{k}{n}$ für $k \in \{0, \ldots, n-1\}$ und $n \in \mathbb{N}$. Da das Riemannintegral von stetigen Funktionen mit dem Lebesgueintegral auf Kompakta übereinstimmt, folgt die Behauptung

$$\int_{[0,1]} f \,\mathrm{d}\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) = \sum_{k=0}^{n-1} \left(\frac{k+1}{n} - \frac{k}{n}\right) f(\xi_k^n) \to \int_0^1 f \,\mathrm{d}x = \int_{[0,1]} f \,\mathrm{d}\lambda$$

für $n \to \infty$.

Beispiel 2.2.4. Im folgenden Beispiel möchten wir noch eine weitere schwach konvergente Funktionenfolge betrachten. Hierbei konstruieren wir eine zum Lebesguemaß absolutstetige Funktionenfolge, welche schwach, jedoch nicht stark konvergiert. Wir betrachten als metrischen Raum die reellen Zahlen zusammen mit dem euklidischen Abstand. Setze

$$A_k = \bigcup_{l=0}^{2^{k-1}-1} \left[\frac{1}{2^k} 2l, \frac{1}{2^k}(2l+1)\right] \text{ und } f_k := \mathbb{1}_{A_k} = \sum_{l=0}^{2^{k-1}-1} \mathbb{1}_{[\frac{1}{2^k}2l, \frac{1}{2^k}(2l+1)]}.$$

Wir betrachten die Maße $\mu_k := f_k \odot \lambda$. Es gilt also $\mu_k(A) = \lambda(A \cap A_k)$ für jedes messbare $A \in \mathfrak{B}$. Sei $g \in C_b(M)$ und $k \in \mathbb{N}$ fest. Dann gilt mit der Gleichheit von

[2] Das heißt: $|\pi_k| \to 0$ für $k \to \infty$.

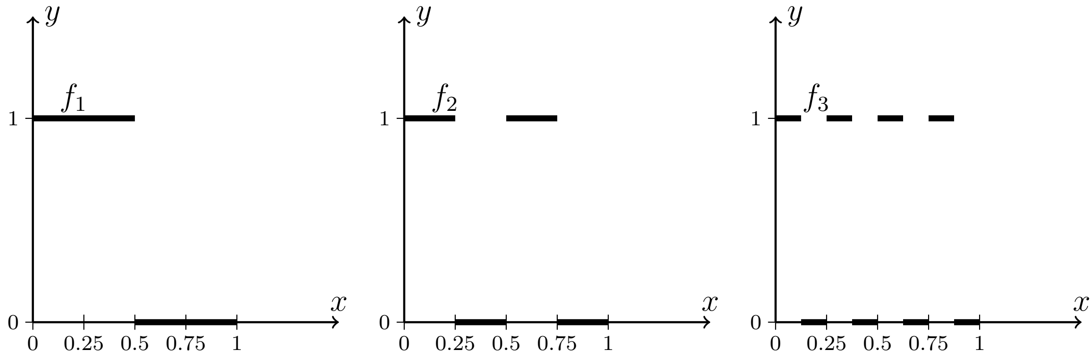

Abbildung 2.1: Die Funktionen f_1, f_2 und f_3.

Riemann und Lebesgueintegral auf kompakten Mengen von Funktionen in $C_b(M)$

$$\begin{aligned}\int_{[0,1]} g \,\mathrm{d}\mu_k &= \int_{[0,1]} g f_k \,\mathrm{d}\lambda = \sum_{l=0}^{2^{k-1}-1} \int_{[\frac{1}{2^k}2l, \frac{1}{2^k}(2l+1)]} g \,\mathrm{d}\lambda = \sum_{l=0}^{2^{k-1}-1} \int_{\frac{1}{2^k}2l}^{\frac{1}{2^k}(2l+1)} g \,\mathrm{d}x \\ &= \sum_{l=0}^{2^{k-1}-1} g(\xi_l^k) \left(\frac{1}{2^k}(2l+1) - \frac{1}{2^k}2l \right) \mathrm{d}x = \frac{1}{2} \sum_{l=0}^{2^{k-1}-1} g\left(\xi_l^k\right) \frac{1}{2^{k-1}} \,\mathrm{d}x.\end{aligned}$$

Die $\xi_l^k \in [\frac{1}{2^k}2l, \frac{1}{2^k}(2l+1)]$ existieren mit dem Mittelwertsatz der Integralrechnung. Beachte die ausgezeichnete Partitionenfolge mit Zwischenstellen ξ_l^k

$$\pi_k : 0 = \frac{0}{2^k} < \frac{2}{2^k} < \ldots < \frac{2l}{2^k} < \ldots < \frac{2^k}{2^k} = 1.$$

Die äquidistante Breite der Teilintervalle beträgt $\frac{1}{2^{k-1}}$. Damit konvergiert

$$\int_{[0,1]} g \,\mathrm{d}\mu_k \to \frac{1}{2} \int_0^1 g \,\mathrm{d}x = \int_{[0,1]} g \,\mathrm{d}\left(\frac{1}{2}\lambda\right)$$

für $k \to \infty$ wegen der Riemannintegrierbarkeit von $g \in C_b(M)$. Wir haben also gezeigt, dass $\mu_k \to \frac{1}{2}\lambda$ bezüglich der schwachen Topologie in $\mathcal{M}^+(\mathfrak{B})$ für $k \to \infty$.

Wir diskutieren noch das Konvergenzverhalten der Folge $(\mu_k)_{k\in\mathbb{N}}$ bezüglich der $\|\cdot\|_{\mathrm{TV}}$-Norm. Angenommen μ_k konvergiert gegen μ im starken Sinne. Dann wäre bereits $\mu = \frac{1}{2}\lambda = f \odot \lambda$ mit $f = \frac{1}{2}$. Mit Satz 1.4.2 folgt daraus bereits die L^1-Konvergenz von f_k gegen $\frac{1}{2}$. Mit der Umkehrung des Satzes von Lebesgue erhalten wir eine fast überall punktweise gegen $\frac{1}{2}$ konvergente Teilfolge μ_{n_k}. Allerdings nimmt f_k nur die Werte 0 und 1 an, was zu einem Widerspruch führt.

2.3 Das Theorem von Portmanteau

Im folgenden Abschnitt lernen wir ein Theorem kennen, welches die schwache Konvergenz von Folgen in $\mathcal{M}^+(\mathfrak{B})$ charakterisiert. Zuerst benötigen wir noch einige Hilfslemmata.

Lemma 2.3.1. *Sei $f : M \to \mathbb{R}$ $\mathfrak{B}$-messbar und $\mu \in \mathcal{M}^+(\mathfrak{B})$. Wir betrachten für alle $y \in \mathbb{R}$ die Niveaumengen*

$$B_y := \{x \in M : f(x) = y\}.$$

Dann ist $A := \{y \in \mathbb{R} : \mu(B_y) > 0\}$ abzählbar[3].

Beweis. Setze $A_k := \{y \in \mathbb{R} : \mu(B_y) > \frac{1}{k}\}$ für $k \in \mathbb{N}$. Dann gilt mit dem archimedischen Axiom $A = \bigcup_{k=1}^{\infty} A_k$. Da die abzählbare Vereinigung von abzählbaren Mengen abzählbar bleibt, genügt es, die Abzählbarkeit von A_k für jedes $k \in \mathbb{N}$ zu zeigen. Sei $k \in \mathbb{N}$ fest. Wir zeigen, dass A_k sogar nur endlich viele Elemente enthält. Angenommen es gäbe $(y_i)_{i\in\mathbb{N}} \subset A_k$ paarweise verschieden. Da f eine Abbildung ist, sind $(B_{y_i})_{i\in\mathbb{N}}$ paarweise disjunkt. Wir erhalten, da μ ein endliches Maß ist, den Widerspruch

$$\infty > \mu(M) \geq \mu\left(\bigcup_{i=1}^{\infty} B_{y_i}\right) = \sum_{i=1}^{\infty} \mu(B_{y_i}) > \sum_{i=1}^{\infty} \frac{1}{k} = \infty.$$

□

Lemma 2.3.2. *Sei $f : M \to \mathbb{R}$ stetig und $y \in \mathbb{R}$. Dann gilt*

$$\partial f^{-1}([y, \infty)) \subset f^{-1}(\{y\}).$$

Beweis. Sei $x \in \partial f^{-1}([y, \infty)]$. Dann gibt es Folgen

$$(x_k)_{k\in\mathbb{N}} \subset f^{-1}([y, \infty)) \text{ und } (\tilde{x_k})_{k\in\mathbb{N}} \subset M \setminus f^{-1}([y, \infty))$$

mit $x_k, \tilde{x_k} \to x$ für $k \to \infty$. Das heißt es gilt $f(x_k) \geq y$ und $f(\tilde{x_k}) < y$ für jedes $k \in \mathbb{N}$. Mit der Stetigkeit von f folgt für $k \to \infty$, dass $y \leq f(x) \leq y$ und damit $y \in f^{-1}(\{y\})$ gilt. □

[3] Wir verstehen unter Abzählbarkeit, dass A entweder endlich oder gleichmächtig zu den natürlichen Zahlen $\mathbb{N}$ ist.

Definition 2.3.3. Für $A \subset M$ nicht leer und $\varepsilon > 0$ setzen wir

$$A_\varepsilon := \{x \in M : \operatorname{dist}(x, A) < \varepsilon\}.$$

A_ε bezeichnen wir als *offene ε-Umgebung* von A.

Eigenschaft 2.3.4. Approximationseigenschaft. *Sei $(a_k)_{k\in\mathbb{N}} \subset [0, \infty)$ eine Nullfolge. Dann gilt $A = \bigcap_{k\in\mathbb{N}} A_{a_k}$ für jedes nicht leere und abgeschlossene $A \subset M$, wobei wir $A_0 := A$ setzen. Außerdem gilt $\mu(A_{a_k}) \to \mu(A)$ für $k \to \infty$ für jedes $\mu \in \mathcal{M}^+(\mathfrak{B})$.*

Beweis. Wegen $A \subset A_{a_k}$ folgt direkt $A \subset \bigcap_{k\in\mathbb{N}} A_{a_k}$. Sei andererseits $x \in A_{a_k}$ für alle $k \in \mathbb{N}$. Dann ist nach Definition $0 \leq \operatorname{dist}(x, A) \leq a_k \to 0$ für $k \to \infty$ und damit $\operatorname{dist}(x, A) = 0$. Da A abgeschlossen ist, folgt $x \in A$. Mit der σ-Stetigkeit von μ folgt die Approximationseigenschaft des Maßes. □

Satz 2.3.5. Portmanteau. *Seien $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ und $\mu \in \mathcal{M}^+(\mathfrak{B})$. Dann sind folgende Aussagen äquivalent:*

1. *$\mu_n \to \mu$ für $n \to \infty$ in der schwachen Topologie.*
2. *$\int f \,\mathrm{d}\mu_n \to \int f \,\mathrm{d}\mu$ für $n \to \infty$ für alle $f \in C_b(M)$ mit $f(M) \subset [0, 1]$.*
3. *$\lim\limits_{n\to\infty} \mu_n(M) = \mu(M)$ und $\int f \,\mathrm{d}\mu \leq \liminf\limits_{n\to\infty} \int f \,\mathrm{d}\mu_n$ für alle unterhalbstetigen $f : M \to [0, \infty) \cup \{\infty\}$.*
4. *$\lim\limits_{n\to\infty} \mu_n(M) = \mu(M)$ und $\limsup\limits_{n\to\infty} \mu_n(C) \leq \mu(C)$ für alle abgeschlossenen Mengen $C \subset M$.*
5. *$\lim\limits_{n\to\infty} \mu_n(M) = \mu(M)$ und $\liminf\limits_{n\to\infty} \mu_n(U) \geq \mu(U)$ für alle offenen Mengen $U \subset M$.*
6. *$\lim\limits_{n\to\infty} \mu_n(A) = \mu(A)$ für alle $A \in \mathfrak{B}$ mit $\mu(\partial A) = 0$.*

Beweis.

„1. $\Rightarrow$ 4.“

Sei zunächst $(\mu_n)_{n\in\mathbb{N}}$ schwach konvergent gegen μ. Sei $C \subset M$ abgeschlossen. Die Konvergenz von $\mu_n(M)$ gegen $\mu(M)$ folgt durch Testen mit $\mathbb{1}_M$. Wegen des Lemmas von Urysohn für Umgebungen (Lemma A.2.5) gibt es eine Funktionenfolge $(f_k)_{k\in\mathbb{N}} \subset C_b(M)$ mit $\mathbb{1}_C \leq f_k \leq \mathbb{1}_{C_{\frac{1}{k}}}$ für jedes $k \in \mathbb{N}$.

Beachte die Monotonie des „lim sup“ und „lim inf“ und die Approximations-eigenschaft 2.3.4:

$$\begin{aligned}\limsup_{n\to\infty}\mu_n(C) &= \limsup_{n\to\infty}\int \mathbb{1}_C \,\mathrm{d}\mu_n \leq \liminf_{k\to\infty}\limsup_{n\to\infty}\int f_k \,\mathrm{d}\mu_n \\ &\leq \liminf_{k\to\infty}\int f_k \,\mathrm{d}\mu \leq \liminf_{k\to\infty}\int \mathbb{1}_{C_{\frac{1}{k}}} \,\mathrm{d}\mu \\ &= \liminf_{k\to\infty}\mu(C_{\frac{1}{k}}) = \mu(C)\end{aligned}$$

„4. $\Rightarrow$ 5.“

Sei U offen. Da $M \setminus U$ abgeschlossen ist, können wir die Voraussetzung anwenden und erhalten

$$\begin{aligned}\liminf_{n\to\infty}\mu_n(U) &= \liminf_{n\to\infty}(\mu_n(M) - \mu_n(M\setminus U)) = \mu(M) - \limsup_{n\to\infty}\mu_n(M\setminus U) \\ &\geq \mu(M) - \mu(M\setminus U) = \mu(U).\end{aligned}$$

„5. $\Rightarrow$ 4.“ folgt analog zu „4. $\Rightarrow$ 5.“.

„4. und 5. $\Rightarrow$ 6.“

Sei $A \in \mathfrak{B}$ mit $\mu(\partial A) = 0$. Dann gilt mit $\overline{A} \setminus \mathring{A} = \partial A$:

$$\begin{aligned}\limsup_{n\to\infty}\mu_n(A) &\leq \limsup_{n\to\infty}\mu_n(\overline{A}) \leq \mu(\overline{A}) = \mu(\mathring{A}\cup\partial A) \\ &= \mu(\mathring{A}) + \mu(\partial A) = \mu(\mathring{A}) \leq \liminf_{n\to\infty}\mu_n(\mathring{A}) \\ &\leq \liminf_{n\to\infty}\mu_n(A) \leq \limsup_{n\to\infty}\mu_n(A)\end{aligned}$$

Also gilt überall Gleichheit und der Grenzwert

$$\lim_{n\to\infty}\mu_n(A) = \mu(\mathring{A}) = \mu(\overline{A})$$

existiert. Mit $\mu(\mathring{A}) \leq \mu(A) \leq \mu(\overline{A}) = \mu(\mathring{A})$ folgt die Behauptung.

„6. $\Rightarrow$ 2.“

Sei $f \in C_b(M)$ mit $f(M) \subset [0,1]$. Wir setzen

$$\begin{aligned}\varphi_k &: [0,1] \to [0,\infty) \text{ mit } y \mapsto \mu_k(\{x \in M : f(x) \geq y\}) \text{ und} \\ \varphi &: [0,1] \to [0,\infty) \text{ mit } y \mapsto \mu(\{x \in M : f(x) \geq y\})\end{aligned}$$

für jedes $k \in \mathbb{N}$. Die Abbildungen φ_k und φ sind monoton fallend und damit messbar. Mit Lemma 2.3.1 erhalten wir die Abzählbarkeit von

$$A := \{y \in \mathbb{R} : \mu(f^{-1}(\{y\})) > 0\}.$$

Damit ist A eine Lebesgue Nullmenge. Somit gilt für alle $y \in [0,1] \setminus A$ mit Lemma 2.3.2 und der Monotonie des Maßes μ

$$\mu(\partial\{x \in M : f(x) \geq y\}) \leq \mu(f^{-1}(\{y\})) = 0.$$

Nach Voraussetzung erhalten wir, dass φ_k λ-fast-überall gegen φ konvergiert für $k \to \infty$. Wegen $\mu(\partial M) = \mu(\emptyset) = 0$ erhalten wir, dass die Folge $(\mu_k(M))_{k\in\mathbb{N}}$ konvergiert und damit durch eine Konstante $C \in \mathbb{R}$ beschränkt ist. Insbesondere ist auch $\mu(M) \leq C$. Mit der Monotonie des Maßes sind $(\varphi_k)_{k\in\mathbb{N}}$ und φ gleichmäßig durch $C\mathbb{1}_{[0,1]} \in L^1(\mathbb{R}, \mathfrak{B}, \lambda)$ beschränkt. Mit dem Prinzip von Cavalieri (Satz A.3.5) und dem Satz von Lebesgue erhalten wir

$$\begin{aligned}\int f \,\mathrm{d}\mu_k = \int_{[0,\infty]} \mu(\{x \in \Omega : f(x) \geq y\}) \,\mathrm{d}\lambda(y) = \int_{[0,\infty]} \varphi_k \,\mathrm{d}\lambda \\ \to \int_{[0,\infty]} \varphi \,\mathrm{d}\lambda = \int f \,\mathrm{d}\mu\end{aligned}$$

für $k \to \infty$.

„2. $\Rightarrow$ 1.“

Sei $f \in C_b(M)$. Setze

$$g := \frac{1}{2\|f\|_\infty} f + \frac{1}{2} \in C_b(M),$$

also $g(M) \subset [0,1]$. Insbesondere ist

$$f = 2\|f\|_\infty g - \|f\|_\infty.$$

Dann folgt, da $\mathbb{1}_M \in C_b(M)$ und $\mathbb{1}_M \subset [0,1]$, mit Voraussetzung

$$\begin{aligned}\int f \,\mathrm{d}\mu_n = 2\|f\|_\infty \int g \,\mathrm{d}\mu_n - \|f\|_\infty \int 1 \,\mathrm{d}\mu_n \\ \to 2\|f\|_\infty \int g \,\mathrm{d}\mu - \|f\|_\infty \int 1 \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu \text{ für } n \to \infty.\end{aligned}$$

Wir haben nun die Äquivalenz von 1, 2, 4, 5 und 6 gezeigt, das heißt es bleibt noch die Äquivalenz zu 3 zu zeigen:

„3. $\Rightarrow$ 5.“

Sei $U \subset M$ offen. Dann ist $\mathbb{1}_U$ unterhalbstetig und es gilt

$$\mu(U) = \int \mathbb{1}_U \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int \mathbb{1}_U \,\mathrm{d}\mu_n = \liminf_{n\to\infty} \mu_n(U).$$

„5. $\Rightarrow$ 3.“

Sei $f : M \to [0,\infty) \cup \{\infty\}$ unterhalbstetig. Dann ist $\{x \in M : f(x) \leq t\}$ abgeschlossen und damit $\{x \in M : f(x) > t\}$ offen für jedes $t \geq 0$. Mit Cavalieri A.3.5, der Monotonie des Integrals und dem Lemma von Fatou erhalten wir

$$\begin{aligned}\int f \,\mathrm{d}\mu &= \int_{[0,\infty)} \mu(\{x \in M : f(x) > t\}) \,\mathrm{d}\lambda(t) \\ &\leq \int_{[0,\infty)} \liminf_{n\to\infty} \mu_n(\{x \in M : f(x) > t\}) \,\mathrm{d}\lambda(t) \\ &\leq \liminf_{n\to\infty} \int_{[0,\infty)} \mu_n(\{x \in M : f(x) > t\}) \,\mathrm{d}\lambda(t) = \liminf_{n\to\infty} \int f \,\mathrm{d}\mu_n.\end{aligned}$$

Damit sind alle Äquivalenzen gezeigt. Der Beweis ist angelehnt an [Kel17, Theorem 2.2]. $\square$

Bemerkung 2.3.6. Falls f nicht negativ und unterhalbstetig ist, zeigt Portmanteau, dass das Funktional $\mu \mapsto \int f \,\mathrm{d}\mu$ schwach unterhalbstetig ist. Dies wird im vierten Kapitel von Bedeutung sein. Zudem lässt sich die schwache Konvergenz noch weiter beschreiben: Falls man $M = \mathbb{R}$ betrachtet, kann man mit dem Satz von Helly-Bray die schwache Konvergenz mit der Konvergenz der Verteilungsfunktionen charakterisieren. Falls man $M = \mathbb{R}^n$ betrachtet, liefert der Stetigkeitssatz von Lévy einen Zusammenhang der schwachen Konvergenz mit der Konvergenz der Fouriertransformierten, bzw. der charakteristischen Funktionen der Maße.

Korollar 2.3.7. *Eine Folge $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ konvergiert genau dann schwach gegen das Null-Maß, wenn $\mu_n(A) \to 0$ für alle $A \in \mathfrak{B}$ für $n \to \infty$ gilt.*

Beweis. Folgt direkt aus Portmanteau (Satz 2.3.5). $\square$

Beispiel 2.3.8. In Beispiel 2.2.3 haben wir gesehen, dass

$$\frac{1}{n}\sum_{k=0}^{n-1}\delta_{\frac{k}{n}} \to \lambda$$

schwach für $n \to \infty$ konvergiert. Mit Portmanteau (Satz 2.3.5) erhalten wir, dass sich das Lebesguemaß eines $A \in \mathfrak{B}([0,1])$, dessen Rand eine λ-Nullmenge ist, beliebig genau mit der Folge[4]

$$a_n := \frac{1}{n}\#\left\{k \in \{0,\dots,n-1\} : \frac{k}{n} \in A\right\}$$

approximieren lässt, das heißt $a_n \to \lambda(A)$ für $n \to \infty$. Auf die Bedingung an den Rand von A können wir nicht verzichten. Falls wir zum Beispiel $A = \mathbb{Q} \cap [0,1]$ mit $\lambda(\partial(\mathbb{Q} \cap [0,1])) = \lambda([0,1]) = 1$ betrachten, so wäre $a_n = 1$ für jedes $n \in \mathbb{N}$. Allerdings ist $\lambda(\mathbb{Q}) = 0$, das heißt auf die Voraussetzung an den topologischen Rand lässt sich nicht verzichten.

Aus diesem Beispiel lässt sich noch eine weitere interessante Aussage ableiten: Falls $A, B \in \mathfrak{B}([0,1])$, $\lambda(\partial A) = \lambda(\partial B) = 0$ und $A \cap \mathbb{Q} = B \cap \mathbb{Q}$ gilt, folgt bereits $\lambda(A) = \lambda(B)$.

2.4 Die Prochorov-Metrik

Für den Beweis des Theorems von Prochorov führen wir auf $\mathcal{M}^+(\mathfrak{B})$ eine Metrik ein, die unter bestimmten Voraussetzungen die schwache Topologie metrisiert. Dabei orientieren wir uns im gesamten Abschnitt an [Kel17, Abschnitt 2.2]; dort wird die Prochorov-Metrik für Wahrscheinlichkeitsmaße eingeführt.

Definition 2.4.1. *Die Prochorov-Metrik.*

1. Definiere die Abbildung $d_P^* : \mathcal{M}^+(\mathfrak{B}) \times \mathcal{M}^+(\mathfrak{B}) \to [0,\infty)$ mit

$$(\mu_1,\mu_2) \mapsto \inf\{\varepsilon > 0 : \mu_1(A) \le \mu_2(A_\varepsilon) + \varepsilon \text{ für alle } A \in \mathfrak{B}\}.$$

2. Definiere anschließend die Abbildung $d_P : \mathcal{M}^+(\mathfrak{B}) \times \mathcal{M}^+(\mathfrak{B}) \to [0,\infty)$ mit $(\mu_1,\mu_2) \mapsto \max\{d_P^*(\mu_1,\mu_2), d_P^*(\mu_2,\mu_1)\}$.

[4] Das Symbol # beschreibt das Zählmaß.

Wegen der Endlichkeit der Maße ist die Menge, worüber das Infimum gebildet wird, nicht leer. Dadurch ist die Abbildung d_P^* wohldefiniert. Wir bezeichnen d_P als Prochorov-Metrik auf $\mathcal{M}^+(\mathfrak{B})$.

Um Eigenschaften herzuleiten sammeln wir zunächst einige Eigenschaften für die ε-Umgebung einer Menge:

Eigenschaften 2.4.2. *Sei $A \subset M$ nicht leer und $\varepsilon, \alpha > 0$. Dann gilt*

1. $(A_\varepsilon)_\alpha \subset A_{\varepsilon+\alpha}$,
2. $(M \setminus A_\varepsilon)_\varepsilon \subset M \setminus A$,
3. $(\overline{A})_\varepsilon = A_\varepsilon$.

Beweis.

1. Sei $x \in (A_\varepsilon)_\alpha$, das heißt $\operatorname{dist}(x, A_\varepsilon) < \alpha$. Damit gibt es ein $y \in A_\varepsilon$ mit $d(x,y) < \alpha$. Weiterhin gibt es ein $z \in A$ mit $d(y,z) < \varepsilon$. Mit der Dreiecksungleichung folgt
$$\operatorname{dist}(x, A) \leq d(x,z) \leq d(x,y) + d(y,z) < \alpha + \varepsilon.$$
Also ist $x \in A_{\varepsilon+\alpha}$.
2. Sei $x \in (M \setminus A_\varepsilon)_\varepsilon$. Das heißt es gilt $\operatorname{dist}(x, M \setminus A_\varepsilon) < \varepsilon$. Also existiert ein $y \in M \setminus A_\varepsilon$, sodass $\operatorname{dist}(x,y) < \varepsilon$ ist. Wegen $y \in M \setminus A_\varepsilon$ folgt $\operatorname{dist}(y, A) \geq \varepsilon$. Wir nehmen an, dass $x \in A$ ist. Damit folgt der Widerspruch
$$\varepsilon > \operatorname{dist}(y,x) \geq \operatorname{dist}(y, A) \geq \varepsilon.$$
3. Sei $x \in (\overline{A})_\varepsilon$. Dann gibt es ein $y \in \overline{A}$, sodass $d(x,y) < \varepsilon$ gilt. Wähle ein $z \in A$ mit $d(y,z) < \varepsilon - d(x,y)$. Dann folgt
$$\operatorname{dist}(x, A) \leq d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + \varepsilon - d(x,y) = \varepsilon$$
und damit die Behauptung $x \in A_\varepsilon$. Die andere Inklusion folgt direkt mit $A \subset \overline{A}$.

□

Lemma 2.4.3. *Sei* $\mu_1(M) = \mu_2(M)$. *Dann gilt* $d_P^*(\mu_1, \mu_2) = d_P^*(\mu_2, \mu_1)$ *für alle* $\mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{B})$. *Damit ist insbesondere* $d_P^* = d_P$.

Beweis. Sei $\alpha := d_P^*(\mu_1, \mu_2) + \frac{1}{k} > 0$ für festes $k \in \mathbb{N}$. Damit ist

$$\inf\{\varepsilon > 0 : \mu_1(B) \leq \mu_2(B_\varepsilon) + \varepsilon \text{ für alle } B \in \mathfrak{B}\} < \alpha.$$

Sei $A \in \mathfrak{B}$ beliebig. Dann gilt insbesondere $\mu_1(M \setminus A_\alpha) \leq \mu_2((M \setminus A_\alpha)_\alpha) + \alpha$. Betrachte die Voraussetzung und die zweite Eigenschaft aus 2.4.2 zusammen mit der Monotonie des Maßes

$$\begin{aligned}\mu_2(A) &= \mu_2(M) - \mu_2(M \setminus A) = \mu_1(M) - \mu_2(M \setminus A)\\ &\leq \mu_1(M) - \mu_2((M \setminus A_\alpha)_\alpha) \leq \mu_1(M) - \mu_1(M \setminus A_\alpha) + \alpha = \mu_1(A_\alpha) + \alpha.\end{aligned}$$

Also ist $d_P^*(\mu_2, \mu_1) < \alpha = d_P^*(\mu_1, \mu_2) + \frac{1}{k}$. Für $k \to \infty$ folgt

$$d_P^*(\mu_2, \mu_1) \leq d_P^*(\mu_1, \mu_2).$$

Analog folgt die andere Ungleichung und damit die Gleichheit. □

Lemma 2.4.4. *Bei* d_P^* *genügt es abgeschlossene Mengen zu betrachten, das heißt es gilt*

$$d_P^*(\mu_1, \mu_2) = \inf\{\varepsilon > 0 : \mu_1(C) \leq \mu_2(C_\varepsilon) + \varepsilon \textit{ für alle } C \textit{ abgeschlossen}\}.$$

Beweis. Wir zeigen die Gleichheit der beiden Mengen

$$\begin{aligned}&\{\varepsilon > 0 : \mu_1(C) \leq \mu_2(C_\varepsilon) + \varepsilon \text{ für alle } C \text{ abgeschlossen}\} \text{ und}\\ &\{\varepsilon > 0 : \mu_1(A) \leq \mu_2(A_\varepsilon) + \varepsilon \text{ für alle } A \in \mathfrak{B}\},\end{aligned}$$

über welche das Infimum gebildet wird. Wir zeigen, dass die erste Menge in der zweiten enthalten ist, die andere Inklusion ist klar. Sei $\varepsilon > 0$, sodass die Ungleichung $\mu_1(C) \leq \mu_2(C_\varepsilon) + \varepsilon$ für alle abgeschlossenen $C \in \mathfrak{B}$ gilt. Sei $A \in \mathfrak{B}$. Dann gilt mit der dritten Eigenschaft aus 2.4.2 und der Monotonie von μ_2

$$\mu_1(A) \leq \mu_1(\overline{A}) \leq \mu_2((\overline{A})_\varepsilon) + \varepsilon = \mu_2(A_\varepsilon) + \varepsilon$$

und es folgt die Behauptung. □

Lemma 2.4.5. *Die Abbildung d_P genügt den Eigenschaften einer Metrik.*

Beweis.

1. *Symmetrie:* Folgt direkt aus der Definition.
2. *Definitheit:* Sei $d_P(\mu_1, \mu_2) = 0$ für $\mu_1, \mu_2 \in \mathcal{M}^+(\mathfrak{B})$. Damit ist insbesondere $d_P^*(\mu_1, \mu_2) = 0$. Für alle abgeschlossenen $A \in \mathfrak{B}$ und alle $k \in \mathbb{N}$ gilt

$$\mu_1(A) \leq \mu_2(A_{\frac{1}{k}}) + \frac{1}{k}.$$

   Mit der Approximationseigenschaft 2.3.4 erhalten wir

$$\mu_1(A) \leq \liminf_{k\to\infty} \left(\mu_2(A_{\frac{1}{k}}) + \frac{1}{k} \right) = \mu_2(A).$$

   Die Abschätzung $\mu_2(A) \leq \mu_1(A)$ folgt analog. Mit dem Maßeindeutigkeitssatz und der Endlichkeit beider Maße folgt die Definitheit.
3. *Dreiecksungleichung:* Sei $\mu_1, \mu_2, \mu_3 \in \mathcal{M}^+(\mathfrak{B})$. Setze

$$d_P^*(\mu_1, \mu_2) \leq d_P(\mu_1, \mu_2) < \varepsilon := d_P(\mu_1, \mu_2) + \frac{1}{k} \text{ und}$$
$$d_P^*(\mu_2, \mu_3) \leq d_P(\mu_2, \mu_3) < \alpha := d_P(\mu_2, \mu_3) + \frac{1}{k}$$

   für $k \in \mathbb{N}$. Sei $A \in \mathfrak{B}$. Mit der ersten Eigenschaft aus 2.4.2 und der Monotonie des Maßes folgt

$$\mu_1(A) \leq \mu_2(A_\varepsilon) + \varepsilon \leq \mu_3((A_\varepsilon)_\alpha) + \varepsilon + \alpha \leq \mu_3(A_{\varepsilon+\alpha}) + \varepsilon + \alpha.$$

   Für $k \to \infty$ erhalten wir

$$d_P^*(\mu_1, \mu_3) \leq \varepsilon + \alpha = d_P(\mu_1, \mu_2) + d_P(\mu_2, \mu_3) + \frac{2}{k} \to d_P(\mu_1, \mu_2) + d_P(\mu_2, \mu_3).$$

   Analog gilt $d_P^*(\mu_3, \mu_1) \leq d_P(\mu_1, \mu_2) + d_P(\mu_2, \mu_3)$ und somit die zu zeigende Ungleichung

$$d_P(\mu_1, \mu_3) = \max\{d_P^*(\mu_1, \mu_3), d_P^*(\mu_3, \mu_1)\} \leq d_P(\mu_1, \mu_2) + d_P(\mu_2, \mu_3).$$

□

Bemerkung 2.4.6. Für $a, b \in M$ ist $d_P(\delta_a, \delta_b) = \min(d(a,b), 1)$. Dies wird in [Els09, Beispiel 4.32 in Abschnitt VIII] gezeigt.

Als nächstes zeigen wir Zusammenhänge der Prochorov-Metrik mit der schwachen Konvergenz:

Satz 2.4.7. *Sei $(\mu_k)_{k\in\mathbb{N}}$ eine in $(\mathcal{M}^+(\mathfrak{B}), d_P)$ gegen $\mu \in \mathcal{M}^+(\mathfrak{B})$ konvergente Folge. Dann konvergiert μ_k schwach gegen μ.*

Beweis. Wähle $\varepsilon_k := d_P(\mu_k, \mu) + \frac{1}{k} > d_P(\mu_k, \mu) \geq d_P^*(\mu_k, \mu)$. Sei $A \subset M$ abgeschlossen. Mit der Approximationseigenschaft 2.3.4 und $\varepsilon_k \to 0$ für $k \to \infty$ folgt

$$\limsup_{k\to\infty} \mu_k(A) \leq \limsup_{k\to\infty} \left(\mu(A_{\varepsilon_k}) + \varepsilon_k\right) = \mu(A).$$

Da ebenfalls $d_P^*(\mu, \mu_k) < \varepsilon_k$ gilt, erhalten wir für den ganzen Raum $M = M_{\varepsilon_k}$ die Abschätzung

$$\mu(M) \leq \liminf_{k\to\infty}(\mu_k(M) + \varepsilon_k) = \liminf_{k\to\infty} \mu_k(M) \leq \limsup_{k\to\infty} \mu_k(M) \leq \mu(M).$$

Damit folgt die Konvergenz von $\mu_k(M)$ gegen $\mu(M)$. Mit dem Satz von Portmanteau (Satz 2.3.5) folgt die schwache Konvergenz. Der Beweis ist in [Kel17, Satz 2.8.] nachzulesen. □

Satz 2.4.8. *Sei (M,d) separabel. Sei $(\mu_k)_{k\in\mathbb{N}}$ eine schwach gegen $\mu \in \mathcal{M}^+(\mathfrak{B})$ konvergente Folge in $\mathcal{M}^+(\mathfrak{B})$. Dann konvergiert $(\mu_k)_{k\in\mathbb{N}}$ bereits in $(\mathcal{M}^+(\mathfrak{B}), d_P)$ gegen μ.*

Beweis. Sei $\varepsilon > 0$. Da (M,d) separabel ist, können wir eine dichte Folge $(a_k)_{k\in\mathbb{N}}$ in M wählen. Definiere rekursiv die Partition

$$A_1 := B_{\frac{\varepsilon}{2}}(a_1) \text{ und } A_k := B_{\frac{\varepsilon}{2}}(a_k) \setminus \bigcup_{l=1}^{k-1} A_l \text{ für } k \geq 2.$$

Da $\sum_{k=1}^{\infty} \mu(A_k) = \mu(\bigcup_{k\in\mathbb{N}} A_k) = \mu(M) < \infty$ konvergiert die Reihe absolut und wir können ein $N \in \mathbb{N}$ wählen, sodass

$$\sum_{k=N+1}^{\infty} \mu(A_k) = \mu\left(\bigcup_{k=N+1}^{\infty} A_k\right) < \varepsilon. \tag{2.1}$$

Für jedes $I \subset \{1, \ldots, N\}$ definieren wir die offene Menge $B_I := \left(\bigcup_{i\in I} A_i\right)_\varepsilon$. Mit Portmanteau (Satz 2.3.5) gilt $\liminf_{n\to\infty} \mu_k(B_I) \geq \mu(B_I)$. Damit gibt es ein n_I, sodass

$$\mu_n(B_I) > \mu(B_I) - \varepsilon \tag{2.2}$$

für alle $n \geq n_I$. Wähle zudem ein $n_M \in \mathbb{N}$, sodass

$$|\mu(M) - \mu_n(M)| < \varepsilon \tag{2.3}$$

für alle $n \geq n_M$. Setze

$$n_0 := \max(\{n_I : I \subset \{1, \ldots N\}\} \cup \{n_M\}).$$

Sei nun $C \in \mathfrak{B}$ fest. Setze $I_0 := \{k \in \{1, \ldots, N\} : A_k \cap C \neq \emptyset\}$. Wir zeigen, dass $B_{I_0} \subset C_{2\varepsilon}$. Sei hierfür $x \in B_{I_0}$. Das heißt es gibt ein $i_0 \in I_0$, sodass $\operatorname{dist}(x, A_{i_0}) < \varepsilon$. Weiter gibt es ein $\hat{x} \in A_{i_0}$, sodass $d(x, \hat{x}) < \varepsilon$ gilt. Nach der Definition von I_0 gibt es ein $y \in C \cap A_{i_0}$. Mit der Dreiecksungleichung folgt

$$\operatorname{dist}(x, C) \leq d(x, y) \leq d(x, \hat{x}) + d(\hat{x}, a_{i_0}) + d(a_{i_0}, y) < \varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = 2\varepsilon.$$

Damit gilt also $B_{I_0} \subset C_{2\varepsilon}$. Ebenfalls gilt $C \subset B_{I_0} \cup \bigcup_{k=N+1}^\infty A_k$. Sei $n \geq n_0 \geq n_{I_0}$. Dann gilt mit (2.1), (2.2) und der Subadditivität des Maßes

$$\begin{aligned} \mu(C) &\leq \mu\left(B_{I_0} \cup \bigcup_{k=N+1}^\infty A_k\right) \leq \mu(B_{I_0}) + \mu\left(\bigcup_{k=N+1}^\infty A_k\right) \leq \mu(B_{I_0}) + \varepsilon \\ &\leq \mu_n(B_{I_0}) + 2\varepsilon \leq \mu_n(C_{2\varepsilon}) + 2\varepsilon. \end{aligned} \tag{2.4}$$

Also folgt $d_P^*(\mu, \mu_n) < 2\varepsilon$. Insbesondere folgt mit (2.4) auch

$$\mu(M \setminus C_{2\varepsilon}) \leq \mu_n((M \setminus C_{2\varepsilon})_{2\varepsilon}) + 2\varepsilon. \tag{2.5}$$

Betrachte andererseits mit der zweiten Eigenschaft aus 2.4.2, (2.3), (2.5) und der Monotonie des Maßes

$$\begin{aligned} \mu_n(C) &= \mu_n(M) - \mu_n(M \setminus C) \leq |\mu_n(M) - \mu(M)| + \mu(M) - \mu_n(M \setminus C) \\ &\leq \varepsilon + \mu(M) - \mu_n(M \setminus C) \leq \varepsilon + \mu(M) - \mu_n((M \setminus C_{2\varepsilon})_{2\varepsilon}) \\ &\leq \varepsilon + \mu(M) - \mu(M \setminus C_{2\varepsilon}) + 2\varepsilon = \mu(C_{2\varepsilon}) + 3\varepsilon \leq \mu(C_{3\varepsilon}) + 3\varepsilon. \end{aligned}$$

Damit gilt also $d_P^*(\mu_n, \mu) < 3\varepsilon$. Wir haben nun also $d_P(\mu_n, \mu) < 3\varepsilon$ gezeigt und erhalten die Konvergenz. Der Beweis orientiert sich an [Kel17, Satz 2.8.] und [Els09, Satz 4.35 in Abschnitt VIII]. □

Satz 2.4.7 und Satz 2.4.8 liefern also - unter der Voraussetzung, dass (M, d) separabel ist - die Metrisierbarkeit der schwache Topologie auf $\mathcal{M}^+(\mathfrak{B})$ durch die Prochorov-Metrik.

2.5 Der Satz von Prochorov

Nun gehen wir zum Theorem von Prochorov über. Zum Einstieg zeigt ein kleines Lemma, wie sich die Kompaktheit im metrischen Raums M auf die schwache Kompaktheit in $\mathcal{M}^+(\mathfrak{B})$ überträgt. Hieraus motivierten wir das Prinzip der gleichmäßigen Straffheit, welches schließlich unter bestimmten Voraussetzungen die Kompaktheit charakterisiert.

Lemma 2.5.1. *Sei (M, d) ein separabler metrischer Raum und $K \subset M$. Dann ist $\{\delta_x : x \in K\} \subset \mathcal{M}^+(\mathfrak{B})$ genau dann (relativ) schwach kompakt, wenn K (relativ) kompakt ist.*

Beweis. Sei $\Delta := \{\delta_x : x \in M\} \subset \mathcal{M}^+(\mathfrak{B})$ die Menge aller Dirac-Maße auf M. Wir bezeichnen mit $i : M \to \Delta$ die kanonische Inklusionsabbildung $x \mapsto \delta_x$. Wir zeigen, dass i ein Homöomorphismus ist, wobei wir Δ mit der Spurtopologie der schwachen Topologie ausstatten. Mit Satz 2.4.7 und 2.4.8 ist gezeigt, dass es sich bei $\mathcal{M}^+(\mathfrak{B})$ zusammen mit der schwachen Topologie um einen metrischen Raum handelt. Es genügt also Folgenstetigkeit für i und i^{-1} zu zeigen und wir brauchen nicht zu Netzen übergehen. Mit Beispiel 2.2.2 ist direkt die Stetigkeit von i klar. Sei $(a_k)_{k \in \mathbb{N}} \subset M$ und $a \in M$, sodass $\delta_{a_k} \to \delta_a$ schwach konvergiert. Mit Bemerkung 2.4.6 folgt $\min(d(a_k, a), 1) \to 0$ und damit $a_k \to a$ für $k \to \infty$. Damit ist auch i^{-1} stetig. Wir haben gezeigt, dass Δ - zusammen mit der Spurtopologie der schwachen Topologie - topologisch äquivalent zu M - zusammen mit der von d erzeugten Topologie - ist. Damit folgt die zu zeigende Aussage. □

Definition 2.5.2. Ein Maß $\mu \in \mathcal{M}^+(\mathfrak{B})$ heißt *straff*, falls es für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset M$ gibt, sodass $\mu(M \setminus K) < \varepsilon$ gilt. Eine Familie $(\mu_i)_{i \in I}$ in $\mathcal{M}^+(\mathfrak{B})$ über eine beliebige Indexmenge I heißt *gleichmäßig straff*, falls es für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset M$ gibt, sodass $\mu_i(M \setminus K) < \varepsilon$ für jedes $i \in I$ gilt.

Ergänzend zeigen wir nun eine zusätzliche Äquivalenz zu unserem Einstiegslemma. Dies zeigt, dass sich die Definition der gleichmäßigen Straffheit schon einmal gut mit der Kompaktheit von Mengen bestehend aus Diracmaßen verhält.

Lemma 2.5.3. *Sei (M, d) ein separabler metrischer Raum und $C \subset M$. Dann ist $\delta_C := \{\delta_x : x \in C\}$ genau dann gleichmäßig straff, wenn δ_C relativ schwach kompakt ist.*

Beweis. Mit dem Einstiegslemma 2.5.1 genügt es, die relative Kompaktheit von C in M zu betrachten. Sei zunächst δ_C gleichmäßig straff. Wähle $\varepsilon_0 := \frac{1}{2}$. Nach Voraussetzung gibt es ein kompaktes K, sodass $\delta_a(M \setminus K) < \frac{1}{2}$ für jedes $a \in C$ gilt. Da $\delta_a(M \setminus K) \in \{0, 1\}$, folgt bereits $\delta_a(M \setminus K) = 0$. Also ist $C \subset K$ und damit C relativ kompakt. Sei andererseits nun C relativ kompakt, also ist $\overline{C}$ kompakt. Für jedes $\varepsilon > 0$ und für jedes $a \in C$ gilt $\delta_a(M \setminus \overline{C}) = 0 < \varepsilon$. □

In vollständigen metrischen Räumen haben wir den Vorteil, dass wir Kompaktheit mit dem Begriff der Totalbeschränktheit charakterisieren können (siehe Satz A.2.1). Dies nutzen wir aus, da wir $\mathcal{M}^+(\mathfrak{B})$ metrisiert haben. Hierzu noch wenige Hilfslemmata.

Lemma 2.5.4. *Sei $A \subset M$ nicht leer und totalbeschränkt. Dann gibt es eine Folge $(a_l)_{l \in \mathbb{N}} \subset A$, sodass es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass*

$$A \subset \bigcup_{l=1}^{N} B_\varepsilon(a_k)$$

gilt. Insbesondere gilt $A \subset \overline{\{a_l : l \in \mathbb{N}\}}$.

Beweis. Wähle $x_1^k, \ldots, x_{n_k}^k \in A$ für jedes $k \in \mathbb{N}$, sodass

$$A \subset \bigcup_{l=1}^{n_k} B_{\frac{1}{k}}(x_l^k)$$

gilt. Setze

$$(a_k)_{k \in \mathbb{N}} := (x_1^1, \ldots, x_{n_1}^1, x_1^2, \ldots, x_{n_2}^2, x_1^3, \ldots, x_{n_3}^3, \ldots).$$

Mit dem archimedischen Axiom gibt es ein $k_0 \in \mathbb{N}$ mit $\frac{1}{k_0} < \varepsilon$. Setze $N := \sum_{l=1}^{k_0} n_l$.

Dann folgt

$$A \subset \bigcup_{l=1}^{n_{k_0}} B_{\frac{1}{k_0}}(x_l^{k_0}) \subset \bigcup_{l=1}^{N} B_{\frac{1}{k_0}}(a_l) \subset \bigcup_{l=1}^{N} B_\varepsilon(a_l).$$

Weiterhin gibt es für jedes $a \in A$ ein $l^* \in \{1, \dots, N\}$, sodass $a \in B_\varepsilon(a_{l^*})$. Damit folgt $A \subset \overline{\{a_l : l \in \mathbb{N}\}}$. □

Ein weiteres kleines Hilfslemma:

Lemma 2.5.5. *Seien $x_1, \dots, x_n \in M$ und $\alpha > \varepsilon > 0$. Dann gilt*

$$\left(M \setminus \bigcup_{i=1}^{n} \overline{B_\alpha(x_i)}\right)_\varepsilon \subset \left(M \setminus \bigcup_{i=1}^{n} B_\alpha(x_i)\right)_\varepsilon \subset M \setminus \bigcup_{i=1}^{n} \overline{B_{\alpha-\varepsilon}(x_i)} \subset M \setminus \bigcup_{i=1}^{n} B_{\alpha-\varepsilon}(x_i).$$

Beweis. Zu zeigen ist nur die zweite Inklusion: Sei ein $y \in M$ gegeben mit $\operatorname{dist}(y, M \setminus \bigcup_{i=1}^n B_\alpha(x_i)) < \varepsilon$. Somit gibt es ein $x \in M$ mit $d(x, x_i) \geq \alpha$ für alle $i \in \{1, \dots, n\}$ und $d(x, y) < \varepsilon$. Mit der Dreiecksungleichung folgt

$$d(x_i, y) \geq d(x_i, x) - d(y, x) > \alpha - \varepsilon$$

für alle $i \in \{1, \dots, n\}$. Damit ist $y \in M \setminus \bigcup_{i=1}^n \overline{B_{\alpha-\varepsilon}(x_i)}$. □

Wir kommen nun zur eigentlichen Aussage. Wir zeigen die beiden Implikationen separat, damit deutlich wird, welche Voraussetzungen wir jeweils tatsächlich benötigen.

Satz 2.5.6. Prochorov. *Sei (M, d) vollständig und separabel und $C \subset \mathcal{M}^+(\mathfrak{B})$ totalbeschränkt. Dann ist C gleichmäßig straff.*

Beweis. Um die offenen Bälle in (M, d) und in $(\mathcal{M}^+(\mathfrak{B}), d_P)$ besser unterscheiden zu können, bezeichnen wir die offene Kugel um das Maß μ mit Radius ε (bezüglich der Prochorov-Metrik) mit $B_\varepsilon^{d_P}(\mu)$. Sei $\varepsilon \in (0, 1)$. Da C totalbeschränkt ist, erhalten wir mit Lemma 2.5.4 eine Folge $(\mu_k)_{k \in \mathbb{N}}$ in $(\mathcal{M}^+(\mathfrak{B}), d_P)$, sodass es für alle $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$ mit

$$C \subset \bigcup_{l=1}^{n_k} B^{d_P}_{\varepsilon 2^{-(k+1)}}(\mu_l)$$

gibt. Durch die Wahl von

$$A_l^k := B^{d_P}_{\varepsilon 2^{-(k+1)}}(\mu_l) \cap C$$

für $l \in \{1, \ldots, n_k\}$ gilt insbesondere

$$C = \bigcup_{l=1}^{n_k} A_l^k \tag{2.6}$$

für jedes $k \in \mathbb{N}$.. Da (M, d) separabel ist, gibt es eine dichte Folge $(x_i)_{i\in\mathbb{N}}$ in M.Insbesondere gilt $M = \bigcup_{i=1}^{\infty} B_{2^{-(k+1)}}(x_i)$ für $k \in \mathbb{N}$. Beachte, dass

$$M \setminus \bigcup_{i=1}^{j} B_{2^{-(k+1)}}(x_i) \searrow \emptyset \text{ für } j \to \infty \text{ gilt,}$$

woraus zusammen mit der σ-Stetigkeit und der Monotonie des Maßes μ_l für festes $l \in \{1, \ldots, n_k\}$ die Existenz eines $M_k^l \in \mathbb{N}$ folgt, sodass

$$\mu_l\left(M \setminus \bigcup_{i=1}^{M_k^l} B_{2^{-(k+1)}}(x_i)\right) < \varepsilon 2^{-(k+1)}.$$

Setze $M_k := \max\{M_k^1, \ldots, M_k^{n_k}\}$, womit wir für alle $l \in \{1, \ldots, n_k\}$ die Ungleichung

$$\mu_l\left(M \setminus \bigcup_{i=1}^{M_k} B_{2^{-(k+1)}}(x_i)\right) \leq \mu_l\left(M \setminus \bigcup_{i=1}^{M_k^l} B_{2^{-(k+1)}}(x_i)\right) < \varepsilon 2^{-(k+1)} \tag{2.7}$$

erhalten. Setze

$$W_k := \bigcup_{i=1}^{M_k} \overline{B_{2^{-k}}(x_i)} \text{ und } W := \bigcap_{k=1}^{\infty} W_k.$$

Wir zeigen nun, dass W die gesuchte kompakte Menge ist, mit der die gleichmäßige Straffheit folgt. Sei $k \in \mathbb{N}$ und $\mu \in C$. Damit gibt es mit (2.6) ein $l \in \{1, \ldots, n_k\}$ mit $\mu \in A_l^k$. Mit der Definition der A_l^k ist $d_P(\mu, \mu_l) < \varepsilon 2^{-(k+1)}$. Mit Lemma 2.5.5, der Monotonie des Maßes μ_l, $\varepsilon < 1$ und Gleichung (2.7) folgt

$$\begin{aligned}
\mu(M \setminus W_k) &= \mu\left(M \setminus \bigcup_{i=1}^{M_k} \overline{B_{2^{-k}}(x_i)}\right) \leq \mu_l\left(M \setminus \left(\bigcup_{i=1}^{M_k} \overline{B_{2^{-k}}(x_i)}\right)_{\varepsilon 2^{-(k+1)}}\right) + \varepsilon 2^{-(k+1)} \\
&\leq \mu_l\left(M \setminus \left(\bigcup_{i=1}^{M_k} B_{2^{-k}}(x_i)\right)_{\varepsilon 2^{-(k+1)}}\right) + \varepsilon 2^{-(k+1)} \\
&\leq \mu_l\left(M \setminus \bigcup_{i=1}^{M_k} B_{2^{-k}-\varepsilon 2^{-(k+1)}}(x_i)\right) + \varepsilon 2^{-(k+1)} \\
&\leq \mu_l\left(M \setminus \bigcup_{i=1}^{M_k} B_{2^{-k}-2^{-(k+1)}}(x_i)\right) + \varepsilon 2^{-(k+1)} \\
&= \mu_l\left(M \setminus \bigcup_{i=1}^{M_k} B_{2^{-(k+1)}}(x_i)\right) + \varepsilon 2^{-(k+1)} \\
&< \varepsilon 2^{-(k+1)} + \varepsilon 2^{-(k+1)} = \varepsilon 2^{-k}.
\end{aligned} \tag{2.8}$$

Schließlich gilt mit der Subadditivität von μ und (2.8)

$$\begin{aligned}
\mu(M \setminus W) = \mu\left(M \setminus \bigcap_{k=1}^{\infty} W_k\right) &= \mu\left(\bigcup_{k=1}^{\infty} M \setminus W_k\right) \\
&\leq \sum_{k=1}^{\infty} \mu(M \setminus W_k) < \sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon.
\end{aligned}$$

Zu zeigen bleibt nur noch die Kompaktheit von W. Die Menge W_k ist abgeschlossen für alle $k \in \mathbb{N}$. Damit ist auch W als Schnitt abgeschlossener Mengen abgeschlossen. Da (M, d) ein vollständiger metrischer Raum ist, ist auch $(W, d_{|W \times W})$ vollständig. Mit Satz A.2.1 genügt es also, die Totalbeschränktheit von W zu zeigen: Sei $\tilde{\varepsilon} > 0$. Wähle $k_0 \in \mathbb{N}$, sodass $2^{-k_0} < \frac{\tilde{\varepsilon}}{2}$. Dann gilt

$$W = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{M_k} \overline{B_{2^{-k}}(x_i)} \subset \bigcup_{i=1}^{M_{k_0}} \overline{B_{2^{-k_0}}(x_i)} \subset \bigcup_{i=1}^{M_{k_0}} B_{\tilde{\varepsilon}}(x_i)$$

Also wird W von endlich vielen Bällen mit Radius $\tilde{\varepsilon}$ überdeckt. Der Beweis ist angelehnt an [Kel17, Satz 2.11.]. $\square$

Satz 2.5.7. Prochorov. *Sei $(\mu_k)_{k\in\mathbb{N}}$ eine gleichmäßig straffe, beschränkte*[5] *Folge in $\mathcal{M}^+(\mathfrak{B})$. Dann gibt es ein von außen reguläres Maß (gemäß Definition A.3.4) $\mu \in \mathcal{M}^+(\mathfrak{B})$ und eine schwach gegen μ konvergente Teilfolge $(\mu_{n_k})_{k\in\mathbb{N}}$.*

Beweis. Um den Beweis möglichst übersichtlich zu gestalten, werden wir die Argumente in einzelne Schritte gliedern.

Schritt 1. *Definition geeigneter Mengensysteme.*

Mit der gleichmäßigen Straffheit von $(\mu_k)_{k\in\mathbb{N}}$ gibt es für jedes $m \in \mathbb{N}$ eine kompakte Menge $\tilde{K}_m \subset M$, sodass $\mu_n(M \setminus \tilde{K}_m) < \frac{1}{m}$ für jedes $n \in \mathbb{N}$. Setze

$$K_m := \bigcup_{i=1}^{m} \tilde{K}_i$$

für jedes $m \in \mathbb{N}$. Dann ist K_m kompakt und $K_m \subset K_{m+1}$ und es gilt

$$\begin{aligned} \mu_n(M \setminus K_m) &= \mu_n\left(M \setminus \bigcup_{i=1}^{m} \tilde{K}_i\right) = \mu_n\left(\bigcap_{i=1}^{m} M \setminus \tilde{K}_i\right) \\ &\le \mu_n\left(M \setminus \tilde{K}_m\right) < \frac{1}{m}. \end{aligned} \tag{2.9}$$

Als kompakte Teilräume von M sind alle K_m insbesondere separabel und wir können abzählbare $L_m \subset M$ wählen, sodass $\overline{L_m} = K_m$. Setze $L := \bigcup_{m=1}^{\infty} L_m$. Dann ist L als abzählbare Vereinigung abzählbarer Mengen selbst abzählbar und $K := \bigcup_{m=1}^{\infty} K_m = \bigcup_{m=1}^{\infty} \overline{L_m} \subset \overline{L}$. Betrachte die Mengensysteme

$$\begin{aligned} \mathfrak{F} &:= \{B_\varepsilon(x) : \varepsilon \in \mathbb{Q}, \varepsilon > 0, x \in L\} \text{ und} \\ \mathfrak{D} &:= \left\{\bigcup_{i=1}^{n} (\overline{B_i} \cap K_{m_i}) : n \in \mathbb{N}, B_1, \dots, B_n \in \mathfrak{F}, m_1, \dots, m_n \in \mathbb{N}\right\} \cup \{\emptyset\}. \end{aligned}$$

$\mathfrak{F}$ enthält offene Mengen, $\mathfrak{D}$ kompakte Mengen. Weiter enthalten beide Systeme nur abzählbar viele Mengen und $\mathfrak{D}$ ist nach Definition $\cup$-stabil[6]. Wir nehmen an[7], dass es ein $n^* \in \mathbb{N}$ gibt, sodass $\mu_{n^*}(M) > 0$. Damit gilt für $m^* \in \mathbb{N}$ mit

[5] Beschränktheit bezüglich der TV-Norm, das heißt, dass $(\mu_k(M))_{k\in\mathbb{N}}$ beschränkt ist.
[6] Das Mengensystem $\mathfrak{D}$ heißt vereinigungsstabil (schreibe $\cup$-stabil), falls aus $A, B \in \mathfrak{D}$ bereits $A \cup B \in \mathfrak{D}$ folgt.
[7] Sonst wäre $(\mu_k)_{k\in\mathbb{N}}$ die konstante Folge des Null-Maßes und die Folgerung ist trivial.

$\frac{1}{m^*} < \mu_{n^*}(M)$ mit (2.9), dass

$$\mu_{n^*}(K_{m^*}) = \mu_{n^*}(M) - \mu_{n^*}(M \setminus K_{m^*}) > \frac{1}{m^*} - \mu_{n^*}(M \setminus K_{m^*}) > 0.$$

Also ist K_{m^*} nicht leer. Damit ist auch $L \neq \emptyset$ und $\mathfrak{F} \neq \emptyset$. Wir erhalten

$$\bigcup_{B \in \mathfrak{F}} B = M.$$

Für festes $m \in \mathbb{N}$ ist also insbesondere $\mathfrak{F}$ eine offene Überdeckung der kompakten Menge $K_m \subset M$. Also können wir ein $r \in \mathbb{N}$ und $B_1, \ldots, B_r \in \mathfrak{F}$ wählen, sodass $K_m \subset \bigcup_{i=1}^r B_i$ gilt. Insbesondere gilt

$$K_m = \bigcup_{i=1}^r \overline{B_i} \cap K_m \in \mathfrak{D}. \tag{2.10}$$

Schritt 2. *Finden einer Teilfolge.*

Beachte, dass $(\mu_n(A))_{n \in \mathbb{N}}$ für jedes $A \in \mathfrak{D}$ eine beschränkte Folge in $\mathbb{R}$ darstellt. Mit dem Cantor'schen Diagonalargument (Lemma A.2.6) finden wir eine Indexfolge $(n_k)_{k \in \mathbb{N}} \subset \mathbb{N}$, sodass $(\mu_{n_k}(A))_{k \in \mathbb{N}}$ für jedes $A \in \mathfrak{D}$ konvergiert. Setze nun

$$\alpha : \mathfrak{D} \to [0, \infty) \text{ mit } D \mapsto \lim_{k \to \infty} \mu_{n_k}(D).$$

Für $D_1, D_2 \in \mathfrak{D}$ mit $D_1 \subset D_2$ folgt direkt mit der Monotonie der Maße μ_{n_k}, dass $\alpha(D_1) \leq \alpha(D_2)$ gilt.

Schritt 3. *Fortsetzung[8] von α auf ganz $\mathfrak{P}(M)$.*

Im nächsten Schritt werden wir α zu einem äußeren Maß φ fortsetzen und später zeigen, dass $\varphi_{|\mathfrak{B}}$ ein Maß ist, wogegen $(\mu_{n_k})_{k \in \mathbb{N}}$ schwach konvergiert. Hierfür setzen wir

$$\beta : \mathfrak{U} := \{U \subset M : U \text{ offen}\} \to [0, \infty] \text{ mit } U \mapsto \sup_{\substack{D \in \mathfrak{D} \\ D \subset U}} \alpha(D) \text{ und}$$
$$\varphi : \mathfrak{P}(M) \to [0, \infty] \text{ mit } A \mapsto \inf_{\substack{U \text{ offen} \\ U \supset A}} \sup_{\substack{D \in \mathfrak{D} \\ D \subset U}} \alpha(D) = \inf_{\substack{U \text{ offen} \\ U \supset A}} \beta(U).$$

Das Supremum in der Definition von β ist endlich, da mit der Beschränkheit von

[8] Wir bezeichnen die Potenzmenge von M mit $\mathfrak{P}(M)$.

$(\mu_k)_{k\in\mathbb{N}}$

$$\alpha(D) = \lim_{k\to\infty} \mu_{n_k}(D) \le \sup_{k\in\mathbb{N}} \mu_{n_k}(M) < \infty$$

unabhängig von D gilt.

Schritt 4. *φ ist ein äußeres Maß.*

Es gilt $\emptyset \in \mathfrak{D}$. Damit ist $\alpha(\emptyset) = \lim_{k\to\infty} \mu_{n_k}(\emptyset) = 0$. Weiter ist direkt $\beta(\emptyset) = \varphi(\emptyset) = 0$. Die Monotonie von φ folgt direkt aus

$$\{U \subset M : U \supset A, U \text{ offen}\} \supset \{U \subset M : U \supset B, U \text{ offen}\}$$

für $A \subset B \subset M$. Analog ist β monoton, woraus wir folgern[9], dass $\varphi_{|\mathfrak{U}} = \beta$. Die Hauptarbeit liegt nun darin, die σ-Subadditivität zu zeigen.

Schritt 4a. *Subadditivität von β für endlich viele offene Mengen.*

Seien $U, V \subset M$ offen. Falls $U = M$ gilt, ist

$$\beta(U \cup V) = \beta(M) \le \beta(M) + \beta(V) = \beta(U) + \beta(V).$$

Analog folgt der Fall $V = M$. Seien nun $U, V \ne M$. Sei $D \subset U \cup V$ mit $D \in \mathfrak{D}$ beliebig. Setze

$$\begin{aligned} A_U &:= \{x \in M : \operatorname{dist}(x, M \setminus U) \ge \operatorname{dist}(x, M \setminus V)\} \cap D \text{ und} \\ A_V &:= \{x \in M : \operatorname{dist}(x, M \setminus U) \le \operatorname{dist}(x, M \setminus V)\} \cap D. \end{aligned} \tag{2.11}$$

Man sieht direkt, dass $D = A_U \cup A_V$. Weiter sind die Mengen A_U und A_V abgeschlossen, da die Distanzabbildung stetig ist und D ebenfalls abgeschlossen ist. Wir zeigen nun, dass $A_U \subset U$. Sei $x \in A_U$. Angenommen $x \in M \setminus U$. Dann ist mit (2.11)

$$0 = \operatorname{dist}(x, M \setminus U) \ge \operatorname{dist}(x, M \setminus V) \ge 0.$$

Da $M \setminus V$ abgeschlossen ist, gilt bereits $x \in M \setminus V$. Allerdings ist

$$x \in D \setminus U \subset (U \cup V) \setminus U \subset V,$$

was zu einem Widerspruch führt. Analog folgt $A_V \subset V$.

[9] Es ist a priori nicht klar, ob $\varphi_{|\mathfrak{D}} = \alpha$ gilt.

Wir zeigen nun, dass es ein $E \in \mathfrak{D}$ gibt, welches $A_U \subset E \subset U$ erfüllt. Beachte mit der Definition von $\mathfrak{D}$ die Darstellung

$$D = \bigcup_{i=1}^{n} (\overline{B_i} \cap K_{m_i})$$

mit $n \in \mathbb{N}, B_1, \ldots, B_n \in \mathfrak{F}$ und $m_1, \ldots, m_n \in \mathbb{N}$. Setze $m := \max\{m_1, \ldots, m_n\}$. Dann gilt $D \subset K_m$, da die Folge von Kompakta aufsteigend gewählt wurde. Mit $K = \bigcup_{m=1}^{\infty} K_m$ gilt insbesondere

$$A_U \subset D \cap U \subset K_m \cap U \subset K \cap U.$$

Für jedes $x \in A_U$ wähle $\varepsilon_x > 0$, sodass $B_{\varepsilon_x}(x) \subset U$. Da L dicht in K liegt, können wir ein $a_x \in L$ wählen, sodass $d(a_x, x) < \frac{\varepsilon_x}{2}$. Wegen $\overline{\mathbb{Q}} = \mathbb{R}$ können wir ein $r_x > 0, r_x \in \mathbb{Q}$ wählen, sodass $d(a_x, x) < r_x < \frac{\varepsilon_x}{2}$. Dann gilt $B_{r_x}(a_x) \in \mathfrak{F}$ und insbesondere $x \in B_{r_x}(a_x)$. Für jedes $y \in B_{r_x}(a_x)$ gilt mit der Dreiecksungleichung

$$d(y, x) \leq d(y, a_x) + d(a_x, x) < r_x + \frac{\varepsilon_x}{2} < \varepsilon_x.$$

Also ist $x \in B_{r_x}(a_x) \subset B_{\varepsilon_x}(x) \subset U$ und damit $(B_{r_x}(a_x))_{x \in A_U}$ eine offene Überdeckung von A_U. Wegen der Kompaktheit von A_U gibt es eine endliche Teilüberdeckung, das heißt es gibt endlich viele $x_1, \ldots, x_n \in A_U$, sodass

$$A_U \subset \bigcup_{k=1}^{n} B_{r_{x_k}}(a_{x_k}) \subset U.$$

Setze nun $E := \bigcup_{k=1}^{n} B_{r_{x_k}}(a_{x_k}) \cap K_m \in \mathfrak{D}$. Dann ist

$$A_U = A_U \cap D \subset A_U \cap K_m \subset E \subset U,$$

das heißt E erfüllt die gewünschte Eigenschaft. Analog gibt es ein $F \in \mathfrak{D}$ mit $A_V \subset F \subset V$. Mit der bereits gezeigten Monotonie von α, mit $E, F, E \cup F \in \mathfrak{D}$ und der Subbadditivität folgt

$$\begin{aligned}\alpha(D) = \alpha(A_U \cup A_V) \leq \alpha(E \cup F) &= \lim_{k \to \infty} \mu_{n_k}(E \cup F) \\ &\leq \lim_{k \to \infty} \mu_{n_k}(E) + \lim_{k \to \infty} \mu_{n_k}(F) = \alpha(E) + \alpha(F) \leq \beta(U) + \beta(V).\end{aligned}$$

Durch Supremumsbildung über sämtliche D folgt

$$\beta(U \cup V) = \sup_{\substack{D \in \mathfrak{D} \\ D \subset U \cup V}} \alpha(D) \leq \beta(U) + \beta(V).$$

Induktiv ergibt sich die Subadditivität von β für endlich viele offene Mengen.

Schritt 4b. *Subadditivität von β für offene Mengenfolgen.*

Sei nun $(U_k)_{k\in\mathbb{N}}$ eine Folge von offenen Mengen in M. Sei $D \subset \bigcup_{k=1}^{\infty} U_k$ mit $D \in \mathfrak{D}$. Dann gibt es wegen der Kompaktheit von D bereits endlich viele $U_{k_1}, \ldots, U_{k_n}$ ($k_1, \ldots, k_n \in \mathbb{N}$ paarweise verschieden), die D überdecken. Also gilt mit der Definition von β und Positivität von φ

$$\alpha(D) \leq \beta\left(\bigcup_{i=1}^{n} U_{k_i}\right) \leq \sum_{i=1}^{n} \beta(U_{k_i}) \leq \sum_{i=1}^{\infty} \beta(U_i).$$

Supremumsbildung auf der linken Seiten über alle $D \in \mathfrak{P}(M)$ mit $D \subset \bigcup_{k=1}^{\infty} U_k$ liefert die Subadditivität für offene Mengenfolgen.

Schritt 4c. *Subadditivität von φ.*

Sei $(A_k)_{k\in\mathbb{N}} \subset \mathfrak{B}$ und $\varepsilon > 0$. Nach Definition von φ gibt es $U_k \subset M$ offen mit $U_k \supset A_k$ und

$$\beta(U_k) \leq \varphi(A_k) + \varepsilon\frac{1}{2^k} \text{ für jedes } k \in \mathbb{N}.$$

Also gilt

$$\varphi\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \beta\left(\bigcup_{k=1}^{\infty} U_k\right) \leq \sum_{k=1}^{\infty} \beta(U_k) \leq \sum_{k=1}^{\infty}\left(\varphi(A_k) + \varepsilon\frac{1}{2^k}\right) = \sum_{k=1}^{\infty} \varphi(A_k) + \varepsilon.$$

Für $\varepsilon \to 0$ folgt die σ-Subadditivität[10]. Wir haben also gezeigt, dass φ ein äußeres Maß auf M ist.

Schritt 5. *Die Einschränkung von φ auf $\mathfrak{B}$ ist ein Maß.*

Mit dem Satz von Carathéodory (Satz A.3.3) ist φ eingeschränkt auf die σ-Algebra

$$\Sigma := \{A \subset M : \varphi(Q) = \varphi(Q \cap A) + \varphi(Q \cap (M \setminus A)) \text{ für alle } Q \subset M\}$$

[10] Das Argument mit $\varepsilon \to 0$ funktioniert natürlich nur, wenn $\sum_{k=1}^{\infty} \varphi(A_k) < \infty$ ist. Ansonsten folgt die Aussage direkt.

ein Maß. Da wir ein Maß auf $\mathfrak{B}$ konstruieren möchten, zeigen wir, dass $\mathfrak{B} \subset \Sigma$ ist. Da Σ eine σ-Algebra ist und $\mathfrak{B}$ von den offenen Mengen erzeugt wird, genügt es zu zeigen, dass alle abgeschlossenen Mengen in Σ enthalten sind. Sei also $A \subset M$ abgeschlossen. Es genügt für jedes $Q \subset M$ die Ungleichung

$$\varphi(Q \cap A) + \varphi(Q \cap (M \setminus A)) \leq \varphi(Q)$$

zu zeigen, denn die andere Abschätzung folgt automatisch mit der bereits gezeigten σ-Subadditivität.

Schritt 5a. *Abschätzung für offene Q.*

Sei zunächst $Q \subset M$ offen. Dann ist $Q \cap (M \setminus A)$ offen und wir können nach Definition von β ein $D_1 \in \mathfrak{D}$ mit $D_1 \subset Q \cap (M \setminus A)$ wählen, sodass

$$\beta(Q \cap (M \setminus A)) \leq \alpha(D_1) + \varepsilon \tag{2.12}$$

gilt. Da D_1 kompakt und damit $Q \cap (M \setminus D_1)$ offen ist, können wir weiter $D_2 \in \mathfrak{D}$ mit $D_2 \subset Q \cap (M \setminus D_1)$ wählen, sodass

$$\beta(Q \cap (M \setminus D_1)) \leq \alpha(D_2) + \varepsilon \tag{2.13}$$

gilt. Für $x \in D_2$ folgt $x \notin D_1$. Damit sind also D_1 und D_2 disjunkt. Da $A \subset M \setminus D_1$, $D_1 \cup D_2 \subset Q$ und $D_1 \cup D_2 \in \mathfrak{D}$ erhalten wir mit (2.12) und (2.13)

$$\begin{aligned}
\varphi(Q \cap A) + \varphi(Q \cap (M \setminus A)) &\leq \beta(Q \cap (M \setminus D_1)) + \beta(Q \cap (M \setminus A)) \\
&\leq \varepsilon + \alpha(D_2) + \varepsilon + \alpha(D_1) = 2\varepsilon + \lim_{k \to \infty} \mu_{n_k}(D_2) + \lim_{k \to \infty} \mu_{n_k}(D_1) \\
&= 2\varepsilon + \lim_{k \to \infty} \mu_{n_k}(D_1 \cup D_2) = 2\varepsilon + \alpha(D_1 \cup D_2) \\
&\leq 2\varepsilon + \sup_{\substack{D \in \mathfrak{D} \\ D \subset U}} \alpha(D) \leq 2\varepsilon + \beta(Q).
\end{aligned} \tag{2.14}$$

Für $\varepsilon \to 0$ folgt $\varphi(Q \cap A) + \varphi(Q \cap (M \setminus A)) \leq \beta(Q)$.

Schritt 5b. *Abschätzung für beliebige Q.*

Sei nun $Q \subset M$ beliebig. Wähle für $\varepsilon > 0$ ein offenes $U \supset Q$ mit $\beta(U) \leq \varphi(Q) + \varepsilon$. Beachte mit der Monotonie von φ und (2.14)

$$\varphi(Q \cap A) + \varphi(Q \cap (M \setminus A)) \leq \varphi(U \cap A) + \varphi(U \cap (M \setminus A)) \leq \beta(U) \leq \varphi(Q) + \varepsilon.$$

Für $\varepsilon \to 0$ folgt die Abschätzung. Also enthält Σ einen Erzeuger von $\mathfrak{B}$ und damit auch $\mathfrak{B}$ selbst. Mit dem Satz von Carathéodory ist $\varphi_{|\Sigma}$ ein Maß und damit auch insbesondere $\mu := \varphi_{|\mathfrak{B}}$ ein Maß. Nach Definition von φ ist μ von außen regulär.

Schritt 6. *Schwache Konvergenz von $(\mu_{n_k})_{k\in\mathbb{N}}$ gegen μ.*

Nun zeigen wir schließlich noch die schwache Konvergenz von $(\mu_{n_k})_{k\in\mathbb{N}}$ gegen μ. Sei $U \subset M$ offen. Dann gilt wegen der Monotonie des Maßes μ_{n_k}

$$\mu(U) = \beta(U) = \sup_{\substack{D\in\mathfrak{D}\\ D\subset U}} \alpha(D) = \sup_{\substack{D\in\mathfrak{D}\\ D\subset U}} \lim_{k\to\infty} \mu_{n_k}(D) \leq \liminf_{k\to\infty} \mu_{n_k}(U).$$

Insbesondere gilt

$$\mu(M) \leq \liminf_{k\to\infty} \mu_{n_k}(M) = \liminf_{k\to\infty} \|\mu_{n_k}\|_{\mathrm{TV}} \leq \sup_{k\in\mathbb{N}} \|\mu_k\|_{\mathrm{TV}} < \infty,$$

da die Folge $(\|\mu_k\|_{\mathrm{TV}})_{k\in\mathbb{N}}$ beschränkt ist und M offen ist. Also ist $\mu \in \mathcal{M}^+(\mathfrak{B})$. Es gilt mit (2.10) die Inklusion $\{K_i : i \in \mathbb{N}\} \subset \mathfrak{D}$, weshalb mit der gleichmäßigen Straffheit

$$\begin{aligned}
\mu(M) = \sup_{D\in\mathfrak{D}} \alpha(D) \geq \sup_{m\in\mathbb{N}} \alpha(K_n) &= \sup_{m\in\mathbb{N}} \left(\lim_{k\to\infty} \mu_{n_k}(K_m) + \frac{1}{m} - \frac{1}{m} \right) \\
&\geq \sup_{m\in\mathbb{N}} \left(\limsup_{k\to\infty} \mu_{n_k}(K_m) + \mu_{n_k}(M \setminus K_m) - \frac{1}{m} \right) \\
&= \sup_{m\in\mathbb{N}} \left(\limsup_{k\to\infty} \mu_{n_k}(M) - \frac{1}{m} \right) = \limsup_{k\to\infty} \mu_{n_k}(M).
\end{aligned}$$

gilt. Also folgt mit dem Satz von Portmanteau (Satz 2.3.5) die schwache Konvergenz von $(\mu_{n_k})_{k\in\mathbb{N}}$ gegen μ für $k \to \infty$. Der Beweis ist an [Els09, Satz 4.22 in Abschnitt VIII] angelehnt. □

Korollar 2.5.8. Prochorov. *Sei (M, d) separabel und vollständig. Sei weiter eine Menge $C \subset \mathcal{M}^+(\mathfrak{B})$ gegeben. Dann sind diese Aussagen äquivalent:*

1. *C ist relativ schwach kompakt.*
2. *C ist relativ schwach folgenkompakt.*
3. *C ist beschränkt und gleichmäßig straff.*

Beweis. Die Abbildung d_P metrisiert $\mathcal{M}^+(\mathfrak{B})$ zusammen mit der schwache Topologie (Satz 2.4.7 und Satz 2.4.8). Das heißt die kompakten Mengen in der schwachen Topologie auf $\mathcal{M}^+(\mathfrak{B})$ werden mit Satz A.2.1 bereits durch die Eigenschaft der Teilfolgen und die Totalbeschränktheit charakterisiert und wir können auf den

Übergang zu Netzen und Teilnetzen verzichten. Mit Satz 2.5.6 und Satz 2.5.7 sind die Aussagen äquivalent. □

Bemerkung 2.5.9. Der Satz von Prochorov lässt sich auch allgemeiner auf sogenannten polnischen Räumen formulieren. Dabei heißt ein topologischer Raum (X, τ) polnisch, wenn er sich metrisieren lässt, sodass der zugrundeliegende metrische Raum vollständig und separabel ist. Falls wir $M = \mathbb{R}$ betrachten und wir uns auf die Menge der Wahrscheinlichkeitsmaße beschränken, lässt sich ein vereinfachter Beweis von Satz 2.5.7 mithilfe des Satzes von Hally-Bray führen.

Korollar 2.5.10. *Der Raum* $(\mathcal{M}^+(\mathfrak{B}), d_p)$ *ist vollständig.*

Beweis. Sei $(\mu_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $(\mathcal{M}^+(\mathfrak{B}), d_p)$. Dann ist $\{\mu_k : k \in \mathbb{N}\}$ totalbeschränkt und insbesondere beschränkt. Mit Satz 2.5.6 ist $(\mu_k)_{k\in\mathbb{N}}$ gleichmäßig straff und wir erhalten mit Satz 2.5.7 eine konvergente Teilfolge. Damit konvergiert bereits $(\mu_k)_{k\in\mathbb{N}}$. □

Korollar 2.5.11. *Sei* (M, d) *vollständig und separabel. Sei* $\mu \in \mathcal{M}^+(\mathfrak{B})$*. Dann ist* μ *straff und von außen regulär.*

Beweis. Da $\{\mu\}$ schwach kompakt ist, folgt direkt mit Satz 2.5.6, dass μ straff ist. Damit ist mit Satz 2.5.7 das Maß μ von außen regulär. □

Wir werden später den Satz von Ulam zitieren (3.1.2), welcher auch die Regularität von innen liefert.

Korollar 2.5.12. *Sei* (M, d) *vollständig und separabel. Dann ist der metrische Raum* (M, d) *genau dann kompakt, wenn* $(\mathcal{M}^+(\mathfrak{B}), d_p)$ *schwach kompakt ist.*

Beweis. Falls (M, d) vollständig ist, so ist jede Folge in $\mathcal{M}^+(\mathfrak{B})$ gleichmäßig straff und wir erhalten mit Satz 2.5.7 die Kompaktheit von $(\mathcal{M}^+(\mathfrak{B}), d_p)$. Für die Rückrichtung betrachte man entsprechende Diracmaße. □

Im vierten Kapitel werden wir die schwache Konvergenz und den Satz von Prochorov anwenden, um das Monge-Kantorovich-Minimierungsproblem zu lösen.

3 Zusammenhang mit der schwach-* Konvergenz

In diesem Kapitel werden wir die schwache Konvergenz in Zusammenhang mit der aus der Funktionalanalysis bekannten schwach-* Konvergenz bringen. Dabei benutzen wir den Satz von Ulam und Rieszscher Darstellungssatz, dessen Beweise aber den Rahmen dieser Arbeit sprengen würden. Weiter stellt sich heraus, dass der für Kompaktheit hinreichende Abschnitt des Satzes von Prochorov (Satz 2.5.7) alternativ mit dem Satz von Banach-Alaoglu bewiesen werden kann. Sei im Verlauf dieses Abschnittes (M, d) ein metrischer Raum.

3.1 Der Dualraum der stetigen Funktionen

Lemma 3.1.1. *Für alle $\mu \in \mathcal{M}(\mathfrak{B})$ setzen wir*

$$T_\mu : C_b(M) \to \mathbb{R} \text{ mit } f \mapsto \int f \,\mathrm{d}\mu.$$

Dann ist $T_\mu \in C_b(M)'$ und es gilt die Abschätzung

$$\|T_\mu\|_{C_b(M)'} \leq \|\mu\|_{TV}.$$

Beweis. Die Linearität folgt direkt mit der Linearität des Integrals. Zu zeigen bleibt die Beschränktheit von T_μ. Sei $f \in C_b(M)$. Dann gilt mit Eigenschaft 1.5.3

$$|T_\mu(f)| = \left| \int f \,\mathrm{d}\mu \right| \leq |\mu|(\Omega)\|f\|_\infty = \|\mu\|_{\mathrm{TV}}\|f\|_\infty.$$

Also folgt $\|T_\mu\|_{C_b(M)'} \leq \|\mu\|_{\mathrm{TV}}$. □

Das heißt also, dass die Abbildung $T : \mathcal{M}(\mathfrak{B}) \to C_b(M)'$ eine stetige Einbettung $\mathcal{M}(\mathfrak{B}) \subset C_b(M)'$ impliziert. $T_{|\mathcal{M}^+(\mathfrak{B})}$ ist mit der Trennungseigenschaft aus Lemma 2.1.3 injektiv.

Wir sehen also sofort, dass für ein Maß $\mu \in \mathcal{M}^+(\mathfrak{B})$ und eine Folge von Maßen $(\mu_k)_{k\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ die Äquivalenz der Aussagen

1. $\mu_k \to \mu$ in der schwachen Topologie für $k \to \infty$
2. $T_{\mu_k} \to T_\mu$ in der schwach-* Topologie von $C_b(M)$ für $k \to \infty$

gilt. Wir zeigen nun, dass die endlichen signierte Maße - falls (M, d) vollständig und separabel ist - tatsächlich auch isometrisch in $C_b(M)'$ eingebettet sind. Hierzu benötigen wir zunächst eine Regularitätsaussage:

Satz 3.1.2. Satz von Ulam. *Sei (M, d) separabel und vollständig. Dann ist jedes $\mu \in \mathcal{M}^+(\mathfrak{B})$ regulär.*

Beweis. Nachzulesen in [Els09, Satz 1.16 in Kapitel VIII.]. □

Lemma 3.1.3. *Sei (M, d) separabel und vollständig und $\mu \in \mathcal{M}(\mathfrak{B})$. Dann gilt $\|T_\mu\|_{C_b(M)'} = \|\mu\|_{TV}$.*

Beweis. Mit Lemma 3.1.1 erhalten wir $\|T_\mu\|_{C_b(M)'} \le \|\mu\|_{\mathrm{TV}}$. Wir wählen die Hahn-Zerlegung $P, N \in \mathfrak{B}$ für μ (Satz 1.2.2). Sei nun $\varepsilon > 0$. Mit Satz 3.1.2 ist die Variation $|\mu|$ regulär und wir erhalten eine kompakte Menge $K \subset P$ und eine offene Menge $O \supset P$ mit

$$|\mu|(O \setminus P) \le \frac{\varepsilon}{4} \text{ und } |\mu|(P \setminus K) \le \frac{\varepsilon}{4}.$$

Durch Skalierung erhalten wir mit Urysohn (Lemma A.2.4) eine stetige Abbildung $f : M \to [-1, 1]$ mit den Eigenschaften[1]

$$f(x) = 1 \text{ für alle } x \in K \text{ und } f(x) = -1 \text{ für alle } x \in M \setminus O.$$

Betrachte, dass $\|f\|_\infty = 1$ gilt. Nun erhalten wir mit Eigenschaft 1.5.3, der Positi-

[1]Falls $K = \emptyset$ wähle $f = -1$; falls $M \setminus O = \emptyset$ wähle $f = 1$.

vität von K und der Negativität von $M \setminus O$:

$$\begin{aligned}
\|\mu\|_{\mathrm{TV}} - T_\mu(f) &= \mu(P) - \mu(N) - \int f \,\mathrm{d}\mu \\
&= \mu(P) - \mu(N) - \int_K f \,\mathrm{d}\mu - \int_{M\setminus O} f \,\mathrm{d}\mu - \int_{O\setminus K} f \,\mathrm{d}\mu \\
&= \mu(P) - \mu(N) - \mu(K) + \mu(M \setminus O) - \int_{O\setminus K} f \,\mathrm{d}\mu \\
&\leq \mu(P) - \mu(N) - \mu(K) + \mu(M \setminus O) + |\mu|(O \setminus K) \\
&= |\mu|(P) - |\mu|(K) + |\mu|(N) - |\mu|(M \setminus O) + |\mu|(P \setminus K) + |\mu|(O \setminus P) \\
&= 2|\mu|(P \setminus K) + 2|\mu|(O \setminus P) \leq \varepsilon
\end{aligned}$$

Also ist $\|T_\mu\|_{C_b(M)'} = \|\mu\|_{\mathrm{TV}}$. □

Korollar 3.1.4. *Falls (M, d) vollständig und separabel ist, so ist T eine isometrische Einbettung.*

Wir fragen uns nun nach der Umkehrung.

Satz 3.1.5. Rieszscher Darstellungssatz. *Sei (M, d) kompakt. Dann gibt es für jedes $\varphi \in C_b(M)'$ genau ein $\mu \in \mathcal{M}(\mathfrak{B})$, sodass*

$$\varphi(f) = \int f \,\mathrm{d}\mu$$

für jedes $f \in C_b(M)$ gilt. Weiterhin gilt

$$\|\mu\|_{TV} = \|\varphi\|_{C_b(M)'}.$$

Wir können also $\mathcal{M}(\mathfrak{B})$ (mit der schwachen Topologie) mit $C_b(M)'$ (mit der schwach- Topologie) identifizieren.*

Beweis. Nachzulesen in [Sch12]. □

Bemerkung 3.1.6. Auf die Kompaktheit des metrischen Raums lässt sich im Allgemeinen nicht verzichten.

3.2 Der Satz von Prochorov - alternativer Beweis

Um einen alternativen Beweis des Satzes von Prochorov zu geben, erinnern wir uns zuerst an den Satz von Banach-Alaoglu:

Satz 3.2.1. Banach-Alaoglu. *Sei X ein Banachraum. Dann ist*

$$B_{X'} = \{x' \in X' : \|x'\|_{X'} \leq 1\}$$

kompakt bezüglich der schwach- Topologie.*

Beweis. Nachzulesen in [Are18, Theorem 26.3]. □

Bemerkung 3.2.2. Der Satz handelt zunächst nur von topologischer Kompaktheit. Das heißt wir erhalten von Folgen in $B_{X'}$ nur schwach-* konvergente Teilnetze (welche im Allgemeinen keine Teilfolgen sind). Allerdings lassen sich, falls X separabel ist, sogar schwach-* konvergente Teilfolgen bilden (siehe [Are18, Proposition 19.4]). Tatsächlich lässt sich in diesem Fall die schwach-* Topologie z.B. durch die Metrik

$$d(x_1', x_2') := \sum_{k=1}^{\infty} 2^{-k}((x_1'(x_k)) - x_2'(x_k))$$

metrisieren, wobei $(x_k)_{k\in\mathbb{N}}$ eine dichte Folge in X ist.

Die nachfolgenden Punkte orientieren sich an [vG03, Theorem 5.2 und Lemma 5.4].

Lemma 3.2.3. *Sei (M, d) ein kompakter metrischer Raum. Dann ist der Banachraum $C_b(M)$ separabel.*

Beweis. Der Beweis benutzt das Theorem von Stone-Weierstraß. Nachzulesen in [MF11, Lemma 3.102]. □

Im Folgenden bezeichnen wir mit $\mathcal{M}_1(\mathfrak{B})$ die Menge der Wahrscheinlichkeitsmaße auf $\mathfrak{B}$. Wir werden den Satz von Prochorov zuerst für Wahrscheinlichkeitsmaße auf kompakten metrischen Räumen zeigen, dann zu nicht notwendigerweise kompakten Räumen übergehen und schließlich das Resultat für endliche Maße verallgemeinern.

Satz 3.2.4. Satz von Prochorov für kompakte metrische Räume. *Sei* (M, d) *ein kompakter metrischer Raum. Dann ist* $\mathcal{M}_1(\mathfrak{B})$ *schwach folgenkompakt.*

Beweis. Sei $(\mu_k)_{k\in\mathbb{N}}$ eine Folge in $\mathcal{M}_1(\mathfrak{B})$. Mit dem Rieszschen Darstellungssatz (Satz 3.1.5) lässt sich die beschränkte Folge $(\mu_k)_{k\in\mathbb{N}}$ als beschränkte Folge in $C(M)'$ einbetten. Da M kompakt ist, ist $C(M)$ separabel (Lemma 3.2.3) und wir erhalten eine schwach konvergente Teilfolge von $(\mu_k)_{k\in\mathbb{N}}$ (Satz von Banach-Alaoglu 3.2.1 und Bemerkung 3.2.2). Damit ist $\mathcal{M}_1(\mathfrak{B})$ schwach folgenkompakt. □

Wir möchten nun das Resultat auf allgemeine metrische Räume übertragen. Hierzu können wir den gegebenen metrischen Raum kompaktifizieren, um das eben gezeigte Resultat zu verwenden.

Lemma 3.2.5. *Die Menge* $Y := [0,1]^{\mathbb{N}} := \{(a_l)_{l\in\mathbb{N}} \subset [0,1]\}$ *ist zusammen mit der Abbildung*

$$\delta : Y \times Y \to [0,1] \text{ mit } (a,b) \mapsto \sum_{l=1}^{\infty} 2^{-l} |a_l - b_l|$$

ein kompakter metrischer Raum.

Beweis. Wir zeigen nur die Kompaktheit. Dies folgt direkt mit dem Satz von Tychonoff; wir werden aber dennoch ein direktes Argument geben. Sei dazu die Folge $(a^k)_{k\in\mathbb{N}} \subset Y$ gegeben. Die Folgen $(a_l^k)_{k\in\mathbb{N}}$ sind für jedes $l \in \mathbb{N}$ beschränkt und wir erhalten mit Cantor's Diagonalargument (Satz A.2.6) eine Indexfolge $(n_k)_{k\in\mathbb{N}} \subset \mathbb{N}$ und ein $a \in Y$, sodass

$$a_l^{n_k} \to a_l \text{ für alle } l \in \mathbb{N} \text{ und } k \to \infty. \tag{3.1}$$

Zu zeigen bleibt die Konvergenz von a^{n_k} gegen a in der Metrik δ. Sei dazu $\varepsilon > 0$. Wegen $\sum_{l=1}^{\infty} 2^{-l} = 1$ können wir ein $n \in \mathbb{N}$ wählen, sodass

$$\sum_{l=n+1}^{\infty} 2^{-l} < \frac{\varepsilon}{2} \tag{3.2}$$

ist. Wegen (3.1) erhalten wir für jedes $l \in \{1, \ldots, n\}$ ein $m_l \in \mathbb{N}$, sodass für alle $k \geq m_l$

$$|a_l^k - a_l| < \frac{\varepsilon}{2n} \tag{3.3}$$

gilt. Sei nun $k \geq m^* := \max\{m_1, \dots, m_n\}$. Dann gilt mit (3.2) und (3.3)

$$\delta(a^k, a) = \sum_{l=1}^{\infty} 2^{-l} |a_l^k - a_l| \leq \sum_{l=1}^{n} |a_l^k - a_l| + \sum_{l=n+1}^{\infty} 2^{-l} < n\frac{\varepsilon}{2n} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Bemerkung 3.2.6. Wir haben im Beweis von Lemma 3.2.5 gesehen, dass im metrischen Raum (Y, δ) aus der punktweisen Konvergenz bereits die Konvergenz in der Metrik δ resultiert. Die Umkehrung folgt einfach, woraus also die von der Metrik δ erzeugten Topologie der Topologie der punktweisen Konvergenz entspricht.

Satz 3.2.7. Kompaktifizierung. *Sei (M, d) ein separabler metrischer Raum. Dann gibt es einen kompakten metrischen Raum (X, δ) und eine Abbildung $T : M \to X$, sodass T ein Homöomorphismus zwischen M und $T(M)$ ist*[2].

Beweis. Sei $(a_k)_{k\in\mathbb{N}} \subset M$ eine dichte Folge. Wähle (Y, δ) wie in Lemma 3.2.5 und setze

$$T : M \to Y \text{ mit } x \mapsto (\min\{d(x, a_l), 1\})_{l\in\mathbb{N}} \in Y.$$

1. *Hilfsaussage: Für $C \subset M$ abgeschlossen und $x \in M \setminus C$ gibt es ein $\varepsilon > 0$ und ein $l \in \mathbb{N}$, sodass*

   $$T(x)_l \leq \frac{\varepsilon}{3} \text{ und } T(y)_l \geq \frac{2\varepsilon}{3} \text{ für jedes } y \in C \text{ gilt.}$$

   Sei dazu C abgeschlossen und $x \in M \setminus C$. Wähle

   $$\varepsilon := \min\{\operatorname{dist}(x, C), 1\}.$$

   Da C abgeschlossen ist, folgt $\varepsilon > 0$. Wähle mit der Dichtheit der Folge $(a_k)_{k\in\mathbb{N}}$ ein $l \in \mathbb{N}$, sodass

   $$T(x)_l = \min\{d(a_l, x), 1) = d(a_l, x) \leq \frac{\varepsilon}{3}$$

[2]Die Abbildung T ist injektiv und stetig. Zudem ist $T^{-1} : T(M) \to M$ eine stetige Abbildung, wobei $T(M)$ mit der Metrik $\delta_{|T(M)\times T(M)}$ ausgestattet wird.

ist. Sei $y \in C$. Wir erhalten mit der Dreiecksungleichung

$$\begin{aligned} T(y)_l &= \min\{d(y, a_l), 1\} \geq \min\{d(x, y) - d(x, a_l), 1\} \\ &\geq \min\{\operatorname{dist}(x, C) - d(x, a_l), 1\} \geq \min\{\operatorname{dist}(x, C) - \frac{\varepsilon}{3}, 1\} \\ &\geq \min\{\varepsilon - \frac{\varepsilon}{3}, 1\} = \frac{2\varepsilon}{3}. \end{aligned}$$

2. *Injektivität:* Seien $x, y \in M$ verschieden. Wähle $C := \{x\}$ abgeschlossen. Mit der Hilfsaussage sind $T(x)_l$ und $T(y)_l$ für ein $l \in \mathbb{N}$ verschieden, woraus die Injektivität folgt.
3. *Stetigkeit von T:* Sei $(x_k)_{k\in\mathbb{N}} \subset M$ konvergent gegen ein $x \in M$. Dann konvergiert wegen der Stetigkeit des Minimums und der Metrik die Folge $T(x_k)_l$ gegen $T(x)_l$ für jedes $l \in \mathbb{N}$ und $k \to \infty$. Mit Bemerkung 3.2.6 folgt die Konvergenz von $T(x_k)$ gegen $T(x)$.
4. *Stetigkeit von $T^{-1} : T(M) \to M$.* Sei $(T(x_k))_{k\in\mathbb{N}} \subset T(M)$ konvergent gegen ein $T(x) \in T(M)$. Wir nehmen an, dass x_k nicht gegen x konvergiert. Die Negation der Definition von Konvergenz liefert ein $\varepsilon_0 > 0$ und eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, sodass $d(x_{n_k}, x) > \varepsilon$ für alle $k \in \mathbb{N}$ gilt. Damit gilt insbesondere:

$$x \in M \setminus \overline{\bigcup_{k=1}^{\infty} \{x_{n_k}\}}$$

Mit der Hilfsaussage erhalten wir ein $\varepsilon > 0$ und ein $l \in \mathbb{N}$, sodass $T(x_l) \leq \frac{\varepsilon}{3}$ und $T(x_{n_k})_l \geq \frac{2\varepsilon}{3}$ für alle $k \in \mathbb{N}$. Damit konvergiert nicht $T(x_{n_k})_l$, nicht $T(x_{n_k})$ und damit auch nicht $T(x_k)$, was der Annahme widerspricht. Damit ist $T^{-1} : T(M) \to M$ stetig.

$\square$

Wir schließen uns bereits jetzt der Konvention an, dass im Kontext des optimalen Transports das Bildmaß eines Maßes μ unter einer Abbildung f als *„Pushforward Measure"* bezeichnet wird. Wir notieren dieses als $f_{\#}\mu$.

Satz 3.2.8. *Sei (M, d) ein vollständiger, separabler metrischer Raum und eine Folge $(\mu_k)_{k\in\mathbb{N}} \subset \mathcal{M}_1(\mathfrak{B})$ gleichmäßig straff. Dann gibt es eine gegen ein Maß $\mu \in \mathcal{M}_1(\mathfrak{B})$ schwach konvergente Teilfolge $(\mu_{n_k})_{k\in\mathbb{N}}$.*

Beweis. Wir wählen mit Lemma 3.2.7 eine Kompaktifizierung von (M, d), das heißt wir erhalten einen metrischen Raum (Y, δ) und eine Abbildung $T : M \to Y$, die einen Homöomorphismus zwischen M und $T(X)$ bildet. Mit Hilfe der Kompaktifizierung von (M, d) können wir nun folgende Strategie anwenden: Wir „liften“ die Folge $(\mu_k)_{k\in\mathbb{N}}$ in den metrischen Raum (Y, δ), wählen dort die Teilfolge und gehen anschließend wieder zurück zu unserem ursprünglichen metrischen Raum.

Sei dazu $\nu_k := T_{\#}\mu_k$. Die Maße ν_k bilden Wahrscheinlichkeitsmaße auf der Borelschen σ-Algebra von (Y, δ). Mit Satz 3.2.4 können wir eine gegen ein Wahrscheinlichkeitsmaß ν schwach konvergente Teilfolge $(\nu_{n_k})_{k\in\mathbb{N}}$ wählen. Wähle mit der gleichmäßigen Straffheit für jedes $n \in \mathbb{N}$ eine kompakte Menge K_n, sodass $\mu_k(K_n) > 1 - \frac{1}{n}$ für jedes $k \in \mathbb{N}$ gilt. Dann gilt mit Portmanteau (Satz 2.3.5) und der Kompaktheit von $T(K_n)$

$$\nu(T(K_n)) \geq \limsup_{k\to\infty} \nu_{n_k}(T(K_n)) = \limsup_{k\to\infty} \mu_{n_k}(K_n) \geq 1 - \frac{1}{n}.$$

Wir wählen nun $E := \bigcup_{n=1}^{\infty} T(K_n) \in \mathfrak{B}(Y)$ mit $E \subset T(M)$. Die Messbarkeit von E folgt aus der Kompaktheit von $T(K_n)$. Dann ist

$$\nu(E) \geq \sup_{n\in\mathbb{N}} \nu(T(K_n)) = 1.$$

Wir haben nun eine messbare Menge in $T(M)$ gefunden, dessen Komplement eine ν-Nullmenge ist. Diese Konstruktion war nötig, da $T(M)$ im Allgemeinen nicht $\mathfrak{B}(Y)$-messbar ist. Wir definieren nun das Wahrscheinlichkeitsmaß

$$\tilde{\nu}(A) := \nu(A \cap E) \text{ für alle } A \in \mathfrak{B}(T(M)).$$

Die Abbildung $\tilde{\nu}$ ist wohldefiniert, da aus $A \in \mathfrak{B}(T(M))$ die Existenz einer Menge $B \in \mathfrak{B}(Y)$ folgt[3], sodass $A = B \cap T(M)$. Damit ist $A \cap E = B \cap E$ $\mathfrak{B}(Y)$-messbar.

Das Maß $\mu := T_{\#}^{-1}\tilde{\nu}$ ist ein wohldefiniertes Wahrscheinlichkeitsmaß auf M, da T^{-1} stetig ist. Zu zeigen bleibt die Konvergenz von μ_{n_k} gegen μ. Sei dazu $C \subset M$ abgeschlossen. Wegen der Stetigkeit von T^{-1} ist $T(C)$ abgeschlossen in $T(M)$ und wir erhalten ein abgeschlossenes $D \subset Y$, sodass $T(C) = D \cap T(M)$. Weiter gilt mit $T^{-1}(D) = C$ und $\nu(Y \setminus E) = 0$:

$$\begin{aligned}\limsup_{k\to\infty} \mu_{n_k}(C) &= \limsup_{k\to\infty} \mu_{n_k}(T^{-1}(D)) = \limsup_{k\to\infty} \nu_{n_k}(D) \leq \nu(D) \\ &= \nu(D \cap E) = \nu(T(C) \cap E) = \tilde{\nu}(T(C)) = \tilde{\nu}((T^{-1})^{-1}(C)) = \mu(C)\end{aligned}$$

Also folgt mit Portmanteau (Satz 2.3.5) die Konvergenz. □

[3] Eigenschaft der Spur-σ-Algebra.

Wir verallgemeinern nun das Resultat auf endliche Maße:

Satz 3.2.9. *Sei* (M, d) *ein vollständiger, separabler metrischer Raum. Sei weiter* $(\mu_k)_{k\in\mathbb{N}} \subset \mathcal{M}^+(\mathfrak{B})$ *gleichmäßig straff und beschränkt. Dann gibt es eine gegen ein Maß* $\mu \in \mathcal{M}^+(\mathfrak{B})$ *schwach konvergente Teilfolge* $(\mu_{n_k})_{k\in\mathbb{N}}$.

Beweis. Falls es eine Teilfolge gibt, die konstant das Null-Maß ist, so folgt die zu zeigende Aussage. Wir können also nach Auswahl einer Teilfolge annehmen, dass $\mu_k(M) > 0$ für jedes $k \in \mathbb{N}$. Mit Satz 3.2.8 gibt es ein Maß $\tilde{\mu} \in \mathcal{M}_1(\mathfrak{B})$ und eine Indexfolge $(n_k)_{k\in\mathbb{N}}$, sodass $\frac{1}{\mu_{n_k}(M)}\mu_{n_k}$ schwach gegen $\tilde{\mu}$ konvergiert. Nach Auswahl einer weiteren Teilfolge konvergiert $\mu_{n_k}(M)$ gegen ein $C \geq 0$. Mit der Definition der schwachen Konvergenz folgt direkt die schwache Konvergenz von μ_{n_k} gegen $C\tilde{\mu}$. $\square$

4 Optimaler Transport

Im Folgenden möchten wir das sogenannte Monge-Kantorovich-Problem einführen. Dabei handelt es sich um das relaxierte Mongeproblem, auf welches wir aber nicht näher eingehen werden. Die Relaxierung wird durchgeführt, um schönere Existenzaussagen zu erhalten. Diese folgen direkt aus dem Satz von Prochorov, den wir im zweiten Kapitel kennen gelernt haben. Bei beiden Fragestellungen handelt es sich um optimale Transportprobleme. Seien im Verlauf des Kapitels (M_1, d_1) und (M_2, d_2) jeweils metrische Räume. Mit $\mathfrak{B}_1$ und $\mathfrak{B}_2$ bezeichnen wir die jeweiligen Borelschen σ-Algebren. Weiterhin seien $\mathcal{M}_1(\mathfrak{B}_1)$ und $\mathcal{M}_1(\mathfrak{B}_2)$ die Menge der Wahrscheinlichkeitsmaße auf $(M_1, \mathfrak{B}_1)$ und $(M_2, \mathfrak{B}_2)$.

4.1 Kopplungen zwischen Maßen

Das Monge-Kantorovich-Minimierungsproblem beruht auf Transportplänen, die wir über die Theorie der Kopplungen einführen. Wir orientieren uns dabei an [Vil09, Definition 1.1].

Lemma 4.1.1. *Seien $\mu_1 \in \mathcal{M}_1(\mathfrak{B}_1)$, $\mu_2 \in \mathcal{M}_1(\mathfrak{B}_2)$ und $\pi \in \mathcal{M}_1(\mathfrak{B}_1 \otimes \mathfrak{B}_2)$. Dann sind folgende Aussagen äquivalent:*

1. *Für jedes $A_1 \in \mathfrak{B}_1$ und für jedes $A_2 \in \mathfrak{B}_2$ gilt*
$$\mu_1(A_1) = \pi(A_1 \times M_2) \text{ und } \mu_2(A_2) = \pi(M_1 \times A_2).$$
2. *Für die Projektionen $p_i : M_1 \times M_2 \to M_i$ mit $(x_1, x_2) \mapsto x_i$ für $i \in \{1, 2\}$ gilt $p_{1\#}\pi = \mu_1$ und $p_{2\#}\pi = \mu_2$.*
3. *Für alle $f_1 \in L^1(M_1, \mathfrak{B}_1, \mu_1)$ und $f_2 \in L^1(M_2, \mathfrak{B}_2, \mu_2)$ gilt*
$$\int_{M_1 \times M_2} (f_1(x_1) + f_2(x_2))\, \mathrm{d}\pi(x_1, x_2) = \int_{M_1} f_1(x_1)\, \mathrm{d}\mu_1(x_1) + \int_{M_2} \mu_2(x_2)\, \mathrm{d}\mu_2(x_2).$$

Beweis.

„1. $\Leftrightarrow$ 2.“

Sei $A_1 \in \mathfrak{B}_1$ und $A_2 \in \mathfrak{B}_2$. Dann gilt

$$\begin{aligned} p_{1\#}\pi(A_1) &= \pi(p_1^{-1}(A_1)) = \pi(A_1 \times M_2) \text{ und} \\ p_{2\#}\pi(A_2) &= \pi(p_2^{-1}(A_2)) = \pi(M_1 \times A_2) \end{aligned}$$

mit der Definition des Bildmaßes, womit direkt die Äquivalenz folgt.

„1. und 2. $\Rightarrow$ 3.“

Mit der Transformationsformel gilt

$$\begin{aligned} &\int_{M_1 \times M_2} (f_1(x_1) + f_2(x_2)) \,\mathrm{d}\pi(x_1, x_2) \\ &\quad = \int_{M_1 \times M_2} f_1 \circ p_1(x_1, x_2) \,\mathrm{d}\pi(x_1, x_2) + \int_{M_1 \times M_2} f_2 \circ p_2(x_1, x_2) \,\mathrm{d}\pi(x_1, x_2) \\ &\quad = \int_{M_1} f_1(x_1) \,\mathrm{d}p_{1\#}\pi(x_1) + \int_{M_2} f_2(x_2) \,\mathrm{d}p_{2\#}\pi(x_2) \\ &\quad = \int_{M_1} f_1(x_1) \,\mathrm{d}\mu_1(x_1) + \int_{M_2} f_2(x_2) \,\mathrm{d}\mu_2(x_2). \end{aligned}$$

„3. $\Rightarrow$ 1.“

Sei $A_1 \in \mathfrak{B}_1$. Dann gilt

$$\begin{aligned} \mu_1(A_1) &= \int_{M_1} \mathbb{1}_{A_1}(x_1) \,\mathrm{d}\mu_1(x_1) \\ &= \int_{M_1 \times M_2} \mathbb{1}_{A_1 \times M_2}(x_1, x_2) \,\mathrm{d}\pi(x_1, x_2) = \pi(A_1 \times M_2). \end{aligned}$$

Analog folgt $\mu_2(A_2) = \pi(M_2, A_2)$ für alle $A_2 \in \mathfrak{B}_2$.

$\square$

Definition 4.1.2. Seien $\mu_1 \in \mathcal{M}_1(\mathfrak{B}_1)$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}_2)$. Ein Wahrscheinlichkeitsmaß π auf $\mathfrak{B}_1 \otimes \mathfrak{B}_2$, das die äquivalenten Aussagen von Lemma 4.1.1 erfüllt, heißt *Kopplung* zwischen μ_1 und μ_2. Wir setzen

$$\Pi(\mu_1, \mu_2) := \{\pi \in \mathcal{M}_1(\mathfrak{B}_1 \otimes \mathfrak{B}_2) : \pi \text{ Kopplung zwischen } \mu_1 \text{ und } \mu_2\}.$$

Für Teilmengen $A_1 \subset \mathcal{M}_1(\mathfrak{B}_1)$ und $A_2 \subset \mathcal{M}_1(\mathfrak{B}_2)$ bezeichnen wir mit

$$\Pi(A_1, A_2) := \{\pi \in \Pi(\mu_1, \mu_2) : \mu_1 \in A_1, \mu_2 \in A_2\}$$

die Menge aller Kopplungen zwischen A_1 und A_2.

Bemerkung 4.1.3. Für jedes $\mu_1 \in \mathcal{M}_1(\mathfrak{B}_1)$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}_2)$ gilt stets

$$\mu_1 \otimes \mu_2 \in \Pi(\mu_1, \mu_2).$$

Somit ist $\Pi(\mu_1, \mu_2)$ nie leer.

Wir kommen nun zu einer Kompaktheitsaussage für Kopplungen, mit welcher wir später das Monge-Kantorovich-Problem lösen können:

Eigenschaften 4.1.4. *Seien (M_1, d_1) und (M_2, d_2) vollständig und separabel. Dann gilt*

1. *Falls $A_1 \subset \mathcal{M}_1(\mathfrak{B}_1)$ und $A_2 \subset \mathcal{M}_1(\mathfrak{B}_2)$ schwach abgeschlossen sind, so ist auch $\Pi(A_1, A_2)$ schwach abgeschlossen.*
2. *Falls $A_1 \subset \mathcal{M}_1(\mathfrak{B}_1)$ und $A_2 \subset \mathcal{M}_1(\mathfrak{B}_2)$ relativ schwach kompakt sind, so ist auch $\Pi(A_1, A_2)$ relativ schwach kompakt.*
3. *Falls $A_1 \subset \mathcal{M}_1(\mathfrak{B}_1)$ und $A_2 \subset \mathcal{M}_1(\mathfrak{B}_2)$ schwach kompakt sind, so ist auch $\Pi(A_1, A_2)$ schwach kompakt.*

Beweis.

1. Sei $(\pi_k)_{k\in\mathbb{N}} \subset \Pi(A_1, A_2)$ schwach konvergent gegen ein $\pi \in \mathcal{M}_1(\mathfrak{B}_1 \otimes \mathfrak{B}_2)$. Sei $f \in C_b(M_1)$. Mit der Transformationsformel und $f \circ p_1 \in C_b(M_1 \times M_2)$ folgt
$$\int_{M_1} f \,\mathrm{d}p_{1\#}\pi_k = \int_{M_1\times M_2} f \circ p_1 \,\mathrm{d}\pi_k \to \int_{M_1\times M_2} f \circ p_1 \,\mathrm{d}\pi = \int_{M_1} f \,\mathrm{d}p_{1\#}\pi$$
für $k \to \infty$. Das heißt, dass $p_{1\#}\pi_k$ schwach gegen $p_{1\#}\pi$ konvergiert. Da $p_{1\#}\pi_k \in A$ und A schwach abgeschlossen, folgt $p_{1\#}\pi \in A_1$ und analog auch $p_{2\#}\pi \in A_2$. Also ist $\pi \in \Pi(p_{1\#}\pi, p_{2\#}\pi) \subset \Pi(A_1, A_2)$, womit die Abgeschlossenheit gezeigt ist.
2. Sei $\varepsilon > 0$. Mit dem Satz von Prochorov (Satz 2.5.6) sind die Mengen A_1 und A_2 jeweils gleichmäßig straff, das heißt es existieren kompakte Mengen

$K_1 \subset M_1$, $K_2 \subset M_2$, sodass für alle $\mu_1 \in A_1$ und $\mu_2 \in A_2$

$$\mu_1(M_1 \setminus K_1) < \frac{\varepsilon}{2} \text{ und } \mu_2(M_2 \setminus K_2) < \frac{\varepsilon}{2}$$

gilt. Beachte, dass $K_1 \times K_2$ kompakt in $M_1 \times M_2$ ist. Weiterhin gilt

$$(M_1 \times M_2) \setminus (K_1 \times K_2) = (M_1 \setminus K_1) \times M_2 \cup M_1 \times (M_2 \setminus K_2).$$

Sei $\pi \in \Pi(A_1, A_2)$, also gibt es $\mu_1 \in A_1$ und $\mu_2 \in A_2$ mit $\pi \in \Pi(\mu_1, \mu_2)$. Mit der Subadditivität von π folgt

$$\begin{aligned} \pi((M_1 \times M_2) \setminus (K_1 \times K_2)) &\leq \pi((M_1 \setminus K_1) \times M_2) + \pi(M_1 \times (M_2 \setminus K_2) \\ &= \mu_1(M_1 \setminus K_1) + \mu_2(M_2 \setminus K_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Somit ist $\Pi(A_1, A_2)$ gleichmäßig straff. Mit dem Satz von Prochorov (Satz 2.5.7) folgt, dass $\Pi(A_1, A_2)$ relativ kompakt ist.

3. Folgt direkt mit der zweiten und dritten Eigenschaft.

Der Beweis ist an [Vil09, Lemma 4.4] angeleht. □

Bemerkung 4.1.5. Da $\{\mu_1\}$ und $\{\mu_2\}$ jeweils kompakt sind, ist auch automatisch $\Pi(\mu_1, \mu_2)$ kompakt.

Wir haben nun die nötigen Werkzeuge erarbeitet, um zur eigentlichen Problemstellung überzugehen:

4.2 Das Monge-Kantorovich-Problem

Problemstellung 4.2.1. *Das Monge–Kantorovich-Minimierungsproblem.* Sei

$$c : M_1 \times M_2 \to (-\infty, \infty) \cup \{-\infty, \infty\}$$

und $\mu_1 \in \mathcal{M}_1(\mathfrak{B}_1)$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}_2)$. Wir untersuchen

$$\inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{M_1 \times M_2} c(x_1, x_2) \, \mathrm{d}\pi(x_1, x_2).$$

Damit das Problem überhaupt wohl gestellt ist, sollte man die $\mathfrak{B}_1 \otimes \mathfrak{B}_2$-Messbarkeit von c sicherstellen. Weiterhin ist die Quasiintegrierbarkeit von c unter allen Kopplungen $\Pi(\mu_1, \mu_2)$ zu fordern.

Wir interpretieren μ_1 und μ_2 als Masseverteilungen auf M_1 und M_2. Dabei soll μ_1 in μ_2 überführt werden. Der Ausdruck $c(x, y)$ gibt an, wie viel der Transport von einem Punkt $x \in M_1$ zu einem Punkt $y \in M_2$ kostet. Daher bezeichnen wir die Funktion c als Kostenfunktion. Die Kopplung π ist die gesuchte Transportabbildung. Hier gibt $\pi(A \times B)$ an, wie viel Masse von einem Bereich A in M_1 zu einem Bereich B in M_2 transportiert wird.

Wir kommen nun schon zu der ersten Existenzaussage. Diese benutzt zum einen die Kompaktheit von $\Pi(\mu_1, \mu_2)$ (Eigenschaften 4.1.5) und zum anderen die Unterhalbstetigkeit des Funktionals (Portmanteau, Satz 2.3.5)

$$\mathcal{E} : \Pi(\mu_1, \mu_2) \to [0, \infty) \cup \{\infty\} \text{ mit } \pi \mapsto \int c \, \mathrm{d}\pi.$$

Satz 4.2.2. *Seien (M_1, d_1) und (M_2, d_2) vollständig und separabel. Dann ist das Monge-Kantorovich-Minimierungsproblem für eine unterhalbstetige, nichtnegative Kostenfunktion c lösbar. Das heißt es gibt ein $\pi^* \in \Pi(\mu_1, \mu_2)$, sodass*

$$\int c \, \mathrm{d}\pi^* = \min_{\pi \in \Pi(\mu_1, \mu_2)} \int c \, \mathrm{d}\pi.$$

Beweis. Mit Bemerkung 4.1.3 ist $\Pi(\mu_1, \mu_2) \neq \emptyset$. Das heißt wir können eine Infimumsfolge $(\pi_k)_{k \in \mathbb{N}} \subset \Pi(\mu_1, \mu_2)$ mit

$$\int c \, \mathrm{d}\pi_k \to \alpha := \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int c \, \mathrm{d}\pi$$

für $k \to \infty$ wählen. Mit Bemerkung 4.1.5 ist $\Pi(\mu_1, \mu_2)$ kompakt, das heißt wir finden eine Teilfolge $(\pi_{n_k})_{k \in \mathbb{N}}$, die gegen ein $\pi^* \in \Pi(\mu_1, \mu_2)$ schwach konvergiert. Mit Portmanteau (Satz 2.3.5) folgt

$$\alpha \leq \int c \, \mathrm{d}\pi^* \leq \liminf_{k \to \infty} \int c \, \mathrm{d}\pi_k = \alpha$$

für die unterhalbstetige Funktion c. Damit folgt die Behauptung. □

Bemerkung 4.2.3. Die Lösbarkeit sagt noch nichts über die Endlichkeit des Infimums aus. Um dieses Problem zu vermeiden, könnte man zum Beispiel zusätzlich $c \in L^1(M_1 \times M_2, \mathfrak{B}_1 \otimes \mathfrak{B}_2, \mu_1 \otimes \mu_2)$ fordern.

Im nächsten Schritt überlegen wir uns, unter welchen Voraussetzungen das Problem auch für Kostenfunktionen c, die auch negative Werte annehmen können, lösbar ist.

Lemma 4.2.4. *Seien (M_1, d_1) und (M_2, d_2) vollständig und separabel. Sei weiter $c : M_1 \times M_2 \to [0, \infty]$ unterhalbstetig, $a_1 \in L^1(M_1, \mathfrak{B}_1, \mu_1)$ und $a_2 \in L^1(M_2, \mathfrak{B}_2, \mu_2)$ (nicht notwendigerweise positiv). Dann gibt es einen gemeinsamer Minimierer des Monge-Kantorovich-Minimierungsproblems zu der Kostenfunktion c und der Kostenfunktion*

$$\tilde{c} : M_1 \times M_2 \to \mathbb{R} \cup \{\infty\} \textit{ mit } (x_1, x_2) \mapsto c(x_1, x_2) + a_1(x_1) + a_2(x_2).$$

Beweis. Mit Satz 4.2.2 erhalten wir einen Minimierer $\pi^* \in \Pi(\mu_1, \mu_2)$ des Monge-Kantorovich-Minimierungsproblems mit der Kostenfunktion c. Wir zeigen nun, dass π^* ebenfalls ein Minimierer des Problems mit der Kostenfunktion $\tilde{c}$ ist. Mit der Definition der Kopplung 4.1.1 erhalten wir

$$\begin{aligned}
\int \tilde{c} \, \mathrm{d}\pi^* &= \int c \, \mathrm{d}\pi^* + \int (a_1(x_1) + a_2(x_2)) \, \mathrm{d}\pi^*(x_1, x_2) \\
&= \int c \, \mathrm{d}\pi^* + \int a_1 \, \mathrm{d}\mu_1 + \int a_2 \, \mathrm{d}\mu_2 \\
&= \min_{\pi \in \Pi(\mu_1, \mu_2)} \left(\int c \, \mathrm{d}\pi + \int a_1 \, \mathrm{d}\mu_1 + \int a_2 \, \mathrm{d}\mu_2 \right) = \min_{\pi \in \Pi(\mu_1, \mu_2)} \int \tilde{c} \, \mathrm{d}\pi,
\end{aligned}$$

womit die Aussage gezeigt ist. □

Satz 4.2.5. *Seien (M_1, d_1) und (M_2, d_2) vollständig und separabel. Sei weiterhin $c : M_1 \times M_2 \to \mathbb{R} \cup \{\infty\}$ unterhalbstetig. Weiterhin gebe es ein $a_1 \in L^1(M_1, \mathfrak{B}_1, \mu_1)$ und ein $a_2 \in L^1(M_2, \mathfrak{B}_2, \mu_2)$ jeweils oberhalbstetig, sodass*

$$a_1(x_1) + a_2(x_2) \leq c(x_1, x_2) \textit{ für alle } (x_1, x_2) \in M_1 \times M_2$$

gilt. Dann ist das Monge-Kantorovich-Minimierungsproblem lösbar.

Beweis. Beachte

$$f : M_1 \times M_2 \to \mathbb{R} \cup \{\infty\} \text{ mit } (x_1, x_2) \mapsto c(x_1, x_2) - a_1(x_1) - a_2(x_2).$$

Mit der Unterhalbstetigkeit von $-a_1$ und $-a_2$ ist auch f als Addition unterhalbstetiger Funktionen selbst unterhalbstetig und weiter nicht negativ. Mit Satz 4.2.4

gibt es einen gemeinsamen Minimierer von f und der Kostenfunktion c, die durch

$$c(x_1, x_2) = f(x_1, x_2) + a_1(x_1) + a_2(x_2).$$

gegeben ist. Der Beweis orientiert sich an [Vil09, Theorem 4.1]. □

Bemerkung 4.2.6. Die beiden bewiesenen Existenzsätze sagen nichts über Eindeutigkeit aus. Diese ist unter den geforderten Voraussetzugen auch nicht zu erwarten, wie man zum Beispiel durch die Wahl von c als konstante Funktion sieht.

4.3 Beispiele

Zum besseren Verständnis diskutieren wir nun einige Beispiele.

Beispiel 4.3.1. Wir betrachten $M_1 = [0,1]$ zusammen mit der euklidischen Metrik und $M_2 = \{0,1\}$ zusammen mit der diskreten Metrik. Wir betrachten einmal das Lebesguemaß λ auf $[0,1]$ und für $\alpha \in (0,1)$ die Konvexkombination

$$\alpha\delta_0 + (1-\alpha)\delta_1 \in \mathcal{M}_1(\mathfrak{B}(M_2)).$$

Sei $A \in \mathfrak{B}([0,1])$ und $\pi \in \Pi(\lambda, \alpha\delta_0 + (1-\alpha)\delta_1)$. Dann gilt

$$\pi(A \times \{0\}) + \pi(A \times \{1\}) = \pi(A \times \{0,1\}) = \lambda(A).$$

Das heißt also, dass die Maße $\pi(\cdot \times \{0\})$ und $\pi(\cdot \times \{1\})$ absolut stetig bezüglich λ sind. Mit dem Satz von Radon-Nikodým gibt es somit $f_0,\ f_1 : [0,1] \to [0,\infty)$ messbar und nicht negativ, sodass

$$\pi(A \times \{0\}) = \int_A f_0 \,\mathrm{d}\lambda \text{ und } \pi(A \times \{1\}) = \int_A f_1 \,\mathrm{d}\lambda$$

für jedes $A \in \mathfrak{B}([0,1])$ gilt. Da die Radon-Nikodým Dichtefunktion λ-fast überall eindeutig ist und

$$\int_A 1 \,\mathrm{d}\lambda = \int_A f_0 \,\mathrm{d}\lambda + \int_A f_1 \,\mathrm{d}\lambda = \int_A (f_0 + f_1) \,\mathrm{d}\lambda$$

gilt, folgt bereits $f_0 + f_1 = 1$ λ-fast überall[1]. Weiterhin gilt

$$\|f_0\|_{L^1} = \int_{[0,1]} f_0 \,\mathrm{d}\lambda = \pi([0,1], \{0\}) = \alpha.$$

Andererseits sieht man schnell, dass für alle $g_0,\ g_1 : [0,1] \to [0,1]$ messbar mit $g_0 + g_1 = 1$ und $\|g_0\|_{L^1} = \alpha$ durch

$$\pi(C) := \int_{p_0^{-1}(C)} g_0 \,\mathrm{d}\lambda + \int_{p_1^{-1}(C)} g_1 \,\mathrm{d}\lambda$$

eine Kopplung zwischen λ und $\alpha\delta_0 + (1-\alpha)\delta_1$ gegeben ist. Dabei ist $p_i : x \mapsto (x, i)$ für $i = 0, 1$. Das heißt also, dass die Kopplungen genau durch zwei Dichten mit entsprechenden Eigenschaften gegeben sind. Sei nun $h : [0,1] \to \mathbb{R}$ messbar. Wir setzen

$$c : (x, y) \mapsto \mathbb{1}_{\{0\}}(y) h(x)$$

und erhalten

$$\int c \,\mathrm{d}\pi = \int c(x,0) f_0(x) \,\mathrm{d}\lambda(x) + \int c(x,0) f_1(x) \,\mathrm{d}\lambda(x) = \int h(x) f_0(x) \,\mathrm{d}\lambda(x)$$

für alle Kopplungen π mit konstruierten Dichten f_0 und f_1. Unter dieser Wahl ergibt sich nun das spezielle Minimierungsproblem

$$\begin{aligned} &\text{Minimiere } \int hf \,\mathrm{d}\lambda, \\ &\quad \text{sodass } f : [0,1] \to [0,1] \text{ messbar mit } \|f\|_{L^1} = \alpha. \end{aligned}$$

Falls h unterhalbstetig ist, ist durch Wahl der Metrik auf M_2 auch c unterhalbstetig. Mit Satz 4.2.2 erhalten wir in diesem Fall einen Minimierer.

Beispiel 4.3.2. Wir versehen $M_1 := \{1, \ldots, n\}$ und $M_2 := \{1, \ldots, m\}$ mit der diskreten Metrik. Da jede Teilmenge von M_1 und M_2 offen ist, entsprechen die Borelschen σ-Algebren den Potenzmengen. Weiter besitzen Maße in $\mathcal{M}_1(\mathfrak{P}(M_1))$

[1]Die Größe $f_0(x)$ (bzw. $f_1(x)$) gibt den Anteil der Masse im Punkt x an, welche zum Punkt 0 (bzw. 1) transportiert wird.

und $\mathcal{M}_1(\mathfrak{P}(M_2))$ Zähldichten und wir können wie folgt identifizieren:

$$\mathcal{M}_1(\mathfrak{P}(M_1)) = \left\{ x \in \mathbb{R}^n : \sum_{k=1}^{n} x_k = 1 \text{ und } x_k \geq 0 \right\}$$
$$\mathcal{M}_1(\mathfrak{P}(M_2)) = \left\{ x \in \mathbb{R}^m : \sum_{k=1}^{m} x_k = 1 \text{ und } x_k \geq 0 \right\}$$
$$\mathcal{M}_1(\mathfrak{P}(M_1 \times M_2)) = \left\{ A \in \mathbb{R}^{n \times m} : \sum_{k=1}^{n} \sum_{l=1}^{m} A_{kl} = 1 \text{ und } A_{kl} \geq 0 \right\}$$

Die Kopplungen zwischen Zähldichten $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$ entsprechen der Menge

$$\left\{ A \in \mathbb{R}^{n \times m} : \sum_{k=1}^{n} A_{kl} = y_l,\ \sum_{l=1}^{m} A_{kl} = x_k \text{ und } A_{kl} \geq 0 \right\}.$$

In diesem Fall lässt sich also das Monge-Kantorovich-Minimierungsproblem unter einer Kostenfunktion $c : M_1 \times M_2 \to (-\infty, \infty) \cup \{\infty\}$ auch als lineares Programm schreiben. Bei der Menge der Punkte, die die Nebenbedingungen erfüllen, handelt es sich um einen sogenannten Transportpolyeder.

$$\begin{aligned} \text{Minimiere } & \sum_{k=1}^{n} \sum_{l=1}^{m} c(k,l) A_{kl}, \\ \text{sodass } & \sum_{k=1}^{n} A_{kl} = y_l \text{ für } l = 1, \ldots, m \\ & \sum_{l=1}^{m} A_{kl} = x_k \text{ für } k = 1, \ldots, n \\ & A_{kl} \geq 0,\ A \in \mathbb{R}^{n \times m}. \end{aligned}$$

Um Existenz eines Minimierers zu zeigen, benötigen wir die Unterhalbstetigkeit der Kostenfunktion c, welche aber in unserem Setting immer gegeben ist.

Wir möchten nun untersuchen, was passiert, wenn wir die diskrete Metrik als Kostenfunktion wählen. Dies lässt sich dadurch interpretieren, dass es irrelevant ist, „wie weit“ Masse transportiert wird. Nur die Information, ob Maße transportiert wird, geht in die Zählung der Gesamtkosten ein.

Lemma 4.3.3. *Sei (M, d) ein metrischer Raum mit Borelscher σ-Algebra $\mathfrak{B}$. Dann gilt*

$$\inf_{\pi \in \Pi(\mu_1,\mu_2)} \int c \,\mathrm{d}\pi = \frac{1}{2}\|\mu_1 - \mu_2\|_{TV} \text{ mit } c : (x_1, x_2) \mapsto \begin{cases} 0 & \text{falls } x_1 = x_2 \\ 1 & \text{falls } x_1 \neq x_2. \end{cases}$$

für alle $\mu_1, \mu_2 \in \mathcal{M}_1(\mathfrak{B})$.

Beweis. Da c unterhalbstetig ist, kennen wir bereits mit Satz 4.2.2 - falls (M, d) vollständig und separabel ist - die Existenz eines Minimierers. Um die Gleichheit zu zeigen, geben wir dennoch einen Minimierer an. Wir setzen $\nu := \mu_1 - \mu_2 \in \mathcal{M}(\mathfrak{B})$. Wir betrachten die Hahn-Zerlegung P und N von ν (Satz 1.2.2). Damit erhalten wir

$$\begin{aligned}
\nu^+(M) &= (\mu_1 - \mu_2)^+(M) = (\mu_1 - \mu_2)(P) = \mu_1(P) - \mu_2(P) \\
&= (1 - \mu_2(P)) - (1 - \mu_1(P)) = \mu_2(N) - \mu_1(N) \\
&= -(\mu_1 - \mu_2)(N) = (\mu_1 - \mu_2)^-(M) = \nu^-(M),
\end{aligned}$$

da μ_1 und μ_2 Wahrscheinlichkeitsmaße sind. Somit gilt

$$\frac{1}{2}\|\mu_1 - \mu_2\|_{\mathrm{TV}} = \frac{1}{2}\|\nu\|_{\mathrm{TV}} = \frac{1}{2}(\nu^+(M) + \nu^-(M)) = \nu^+(M) = \nu^-(M).$$

Sei $\pi \in \Pi(\mu_1, \mu_2)$ eine Kopplung. Mit der Disjunktheit von P und N erhalten wir

$$P \times N \subset \{(x_1, x_2) \in M^2 : x_1 \neq x_2\}.$$

Weiter gilt mit der Monotonie des Maßes π

$$\begin{aligned}
\int c \,\mathrm{d}\pi &= \pi(\{(x_1, x_2) \in M^2 : x_1 \neq x_2\}) \\
&\geq \pi(P \times N) \geq \pi(P \times N) - \pi(N \times P) \\
&= \pi(P \times N) + \pi(P \times P) - (\pi(P \times P) + \pi(N \times P) \\
&= \pi(P \times M) - \pi(M \times P) = \mu_1(P) - \mu_2(P) \\
&= \frac{1}{2}\|\mu_1 - \mu_2\|_{\mathrm{TV}}.
\end{aligned}$$

Wir sehen, dass Gleichheit genau dann gilt, wenn

$$\pi(\{(x_1, x_2) \in M^2 : x_1 \neq x_2\}) = \pi(P \times N) \text{ und } \pi(N \times P) = 0$$

gilt[2]. Dies macht auch Sinn, da genau nur der Transport von P nach N stattfinden soll, damit es sich um eine optimale Kopplung handelt. Wir werden nun einen Minimierer angeben. Sei $d : M \to M \times M$ gegeben mit $x \mapsto (x, x)$. Wir setzen[3]

$$\pi^* := d_\# \mu_2(\cdot \cap P) + d_\# \mu_1(\cdot \cap N) + \frac{1}{\nu^+(M)}(\nu^+ \otimes \nu^-).$$

Wir zeigen zuerst $\pi^* \in \Pi(\mu_1, \mu_2)$. Für jedes $A \in \mathfrak{B}$ gilt wegen $\nu^+(M) = \nu^-(M)$, dass

$$\begin{aligned}
\pi^*(A \times M) &= \mu_2(d^{-1}(A \times M) \cap P) + \mu_1(d^{-1}(A \times M) \cap N) + \frac{1}{\nu^+(M)}\nu^+(A)\nu^-(M) \\
&= \mu_2(A \cap P) + \mu_1(A \cap N) + \nu^+(A) \\
&= \mu_2(A \cap P) + \mu_1(A \cap N) + (\mu_1 - \mu_2)(P \cap A) \\
&= \mu_1(A \cap N) + \mu_1(A \cap P) = \mu_1(A).
\end{aligned}$$

Analog erhalten wir $\pi^*(M \times A) = \mu_2(A)$. Wegen $\pi^*(M \times M) = \mu_1(M) = 1$ ist π^* ein Wahrscheinlichkeitsmaß und damit eine Kopplung. Wir zeigen nun, dass π^* tatsächlich ein Minimierer ist. Wegen $d^{-1}(\{(x_1, x_2) \in M^2 : x_1 \neq x_2\}) = \emptyset$ und $\nu^+(N) = 0$ gilt

$$\begin{aligned}
\int c \, \mathrm{d}\pi &= \pi^*(\{(x_1, x_2) \in M^2 : x_1 \neq x_2\}) \\
&= \frac{1}{\nu^+(M)}(\nu^+ \otimes \nu^-)(\{(x_1, x_2) \in M^2 : x_1 \neq x_2\}) \\
&= \frac{1}{\nu^+(M)} \int_M -\nu(N \setminus \{x\}) \, \mathrm{d}\nu^+(x) \\
&= \frac{1}{\nu^+(M)} \int_P -\nu(N \setminus \{x\}) \, \mathrm{d}\nu^+(x) = \frac{1}{\nu^+(M)} \int_P -\nu(N) \, \mathrm{d}\nu^+(x) \\
&= \frac{1}{\nu^+(M)} \int_P \nu^-(M) \, \mathrm{d}\nu^+(x) = \nu^-(M) = \frac{1}{2}\|\mu_1 - \mu_2\|_{\mathrm{TV}}.
\end{aligned}$$

Damit folgt die zu zeigende Behauptung. □

[2] Die zweite Eigenschaft folgt bereits aus der ersten.

[3] Die ersten beiden Summanden beschreiben diejenige Masse von μ_1, die nicht transportiert werden muss, da sie auch bei μ_2 an selber Stelle „benötigt“ wird. Der dritte Summand beschreibt, dass die zu transportierende Masse gleichmäßig verteilt wird.

5 Der Wassersteinabstand

5.1 Einführung und erste Eigenschaften

Unter einer gegebenen Kostenfunktion c lässt sich der Ausdruck

$$\inf_{\pi \in \Pi(\mu_1, \mu_2)} \int c \, \mathrm{d}\pi$$

als minimaler Aufwand interpretieren, um das Maß μ_1 in das Maß μ_2 zu überführen. Dies entspricht einem Abstandsbegriff. Sind die beiden Maße „ähnlich“, muss wenig Masse transportiert werden - sind sie sehr „verschieden“, muss mehr Masse transportiert werden. Wir haben im Beispiel 4.3.3 gesehen, dass durch Wahl der diskreten Metrik als Kostenfunktion der Totalvariationsabstand entsteht. Die diskrete Metrik ist sehr stark und auch die resultierende $\|\cdot\|_{\mathrm{TV}}$-Norm ist in der Anwendung oft zu restriktiv. Es ist also naheliegend, statt des diskreten Abstandes die gegebene Metrik als Kostenfunktion zu setzen. Wir erhoffen uns dadurch einen besseren, schwächeren Abstandsbegriff. Durch eine zusätzliche Potenz p ergibt sich dadurch der Wassersteinabstand W_p. Sei im Folgenden (M, d) ein vollständiger und separabler metrischer Raum und $\mathfrak{B}$ die zugehörige Borelsche σ-Algebra.

Definition 5.1.1. *Wassersteinabstand.* Sei $p \in [1, \infty)$. Wir setzen

$$W_p(\mu_1, \mu_2) := \left(\inf_{\pi \in \Pi(\mu_1, \mu_2)} \int d(x_1, x_2)^p \, \mathrm{d}\pi(x_1, x_2) \right)^{\frac{1}{p}} \in [0, \infty) \cup \{\infty\}$$

für $\mu_1, \mu_2 \in \mathcal{M}_1(\mathfrak{B})$, wobei wir $\infty^{\frac{1}{p}} := \infty$ festlegen. W_1 wird auch *Kantorovich-Rubinstein-Distanz* genannt.

Bemerkung 5.1.2. Da die Distanzfunktion und die p-te Potenz stetig ist, erhalten wir für $\mu_1, \mu_2 \in \mathcal{M}_1(\mathfrak{B})$ mit Satz 4.2.2 eine optimale Kopplung $\pi^* \in \Pi(\mu_1, \mu_2)$. Das heißt es gilt $W_p(\mu_1, \mu_2)^p = \int d(x_1, x_2)^p \, \mathrm{d}\pi^*(x_1, x_2)$.

Die Definition des Wassersteinabstandes ist zum Beispiel in [Kel17, Definition 5.2] zu finden. Dort befindet sich auch das nachfolgende Lemma:

Lemma 5.1.3. *Sei* $\infty > p > q \geq 1$ *und* $\mu_1, \mu_2 \in \mathcal{M}_1(\mathfrak{B})$. *Dann gilt*

$$W_q(\mu_1, \mu_2) \leq W_p(\mu_1, \mu_2).$$

Beweis. Falls $W_p(\mu_1, \mu_2) = \infty$ gilt, ist die Folgerung trivial. Sei $W_p(\mu_1, \mu_2) < \infty$ und π^* die optimale Kopplung zwischen μ_1 und μ_2 bezüglich $d(\cdot, \cdot)^p$. Dann gilt

$$\begin{aligned} W_q(\mu_1, \mu_2)^q &= \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int d(x_1, x_2)^q \, \mathrm{d}\pi^*(x_1, x_2) \leq \int d(x_1, x_2)^q \, \mathrm{d}\pi^*(x_1, x_2) \\ &\leq \left(\int d(x_1, x_2)^p \, \mathrm{d}\pi^*(x_1, x_2) \right)^{\frac{q}{p}} \left(\int 1 \, \mathrm{d}\pi^*(x_1, x_2) \right)^{\frac{p-q}{p}} \\ &= \left(\int d(x_1, x_2)^p \, \mathrm{d}\pi^*(x_1, x_2) \right)^{\frac{q}{p}} = W_p(\mu_1, \mu_2)^q \end{aligned}$$

mit der Hölder-Ungleichung für $\frac{1}{\frac{p}{q}} + \frac{1}{\frac{p}{p-q}} = 1$. Da die q-te Wurzel monoton ist, folgt die Behauptung. □

Beispiel 5.1.4. Im Allgemeinen ist $W_p(\mu_1, \mu_2)$ nicht endlich. Betrachte hierfür den einelementigen metrischen Raum, der die 0 enthält und die natürlichen Zahlen zusammen mit dem euklidischen Abstand. Als Borelsche σ-Algebren ergeben sich die jeweiligen Potenzmengen. Sei $\mu_1 = \delta_0$ und

$$\mu_2 : \mathfrak{P} \to [0, 1] \text{ mit } A \mapsto C \sum_{k=1}^{\infty} \frac{1}{k^2} \delta_k(A).$$

Dabei sei $C > 0$ so gewählt[1], dass $\mu_2 \in \mathcal{M}_1(\mathfrak{P}(\mathbb{N}))$. Dann ist die einzige Kopplung[2] zwischen μ_1 und μ_2 durch $\mu_1 \otimes \mu_2$ gegeben. Wir betrachten den Fall $p = 1$. Mit Tonelli folgt

$$\begin{aligned} W_1(\mu_1, \mu_2) &= \int d(j, k) \, \mathrm{d}(\mu_1 \otimes \mu_2)(j, k) = \int \int d(j, k) \, \mathrm{d}\mu_1(j) \, \mathrm{d}\mu_2(k) \\ &= \int k \, \mathrm{d}\mu_2(k) = C \sum_{k=1}^{\infty} \frac{1}{k^2} k = \infty. \end{aligned}$$

Der Wassersteinabstand ist also daher nicht endlich, da μ_2 kein endliches erstes Moment besitzt.

[1]Da $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, ist $C = \frac{6}{\pi^2}$.

[2]Einen eleganten Beweis, dass es hier nur die triviale Kopplung gibt, liefern wir mit Korollar 5.2.9.

Es ist also sinnvoll, sich auf Wahrscheinlichkeitsmaße zu beschränken, die endliches p-tes Moment besitzen. Zuerst benötigen wir ein Lemma, um diesen Begriff für Maße auf allgemeinen metrischen Räumen zu definieren.

Lemma 5.1.5. *Sei $\mu \in \mathcal{M}_1(\mathfrak{B})$ und $p \in [1, \infty)$. Falls*

$$\int d(x, x_0)^p \, \mathrm{d}\mu(x) < \infty$$

für ein $x_0 \in M$ gilt, so folgt die Endlichkeit bereits für alle $x_0 \in M$.

Beweis. Sei $x_1 \in M$. Dann gilt mit der Dreiecksungleichung der Metrik und der Minkowski-Ungleichung

$$\begin{aligned}\left(\int d(x, x_1)^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} &\leq \left(\int (d(x, x_0) + d(x_0, x_1))^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} \\ &\leq \left(\int d(x, x_0)^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} + \left(\int d(x_0, x_1)^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} \\ &= \left(\int d(x, x_0)^p \, \mathrm{d}\mu(x)\right)^{\frac{1}{p}} + d(x_0, x_1) < \infty.\end{aligned}$$

□

Definition 5.1.6. Sei $p \in [1, \infty)$. Wir sagen, $\mu \in \mathcal{M}_1(\mathfrak{B})$ besitzt *endliches p-tes Moment*, falls die Eigenschaft aus Lemma 5.1.5 für ein und damit alle $x_0 \in M$ erfüllt ist. Die Menge

$$P_p(M) := \{\mu \in \mathcal{M}_1(\mathfrak{B}) : \mu \text{ besitzt endliches } p\text{-tes Moment}\}$$

bezeichnen wir als *Wassersteinraum der Ordnung p*.

Bemerkung 5.1.7. P_p enthält alle Diracmaße und ist damit nicht leer.

Lemma 5.1.8. *Für alle $\mu_1, \mu_2 \in P_p(M)$ und alle $\pi \in \Pi(\mu_1, \mu_2)$ gilt*

$$\int d(x_1, x_2)^p \, \mathrm{d}\pi(x_1, x_2) < \infty.$$

Insbesondere ist also $W_p(\mu_1, \mu_2) < \infty$.

Beweis. Sei $x_0 \in M$. Mit der Minkowski-Ungleichung und Lemma 4.1.1 folgt

$$\left(\int d(x_1,x_2)^p \,\mathrm{d}\pi(x_1,x_2)\right)^{\frac{1}{p}} \leq \left(\int (d(x_1,x_0)+d(x_0,x_2))^p \,\mathrm{d}\pi(x_1,x_2)\right)^{\frac{1}{p}}$$
$$\leq \left(\int d(x_1,x_0)^p \,\mathrm{d}\pi(x_1,x_2)\right)^{\frac{1}{p}} + \left(\int d(x_0,x_2)^p \,\mathrm{d}\pi(x_1,x_2)\right)^{\frac{1}{p}}$$
$$\leq \left(\int d(x_1,x_0)^p \,\mathrm{d}\mu_1(x_1)\right)^{\frac{1}{p}} + \left(\int d(x_0,x_2)^p \,\mathrm{d}\mu_2(x_2)\right)^{\frac{1}{p}} < \infty.$$

□

Lemma 5.1.9. *Für $\infty > p > q \geq 1$ gilt $P_p(M) \subset P_q(M)$.*

Beweis. Sei $\mu \in P_p$ und $x_0 \in M$. Ähnlich zum Beweis von Lemma 5.1.3 benutzen wir die Hölder-Ungleichung für $\frac{1}{\frac{p}{q}} + \frac{1}{\frac{p}{p-q}} = 1$.

$$\int d(x,x_0)^q \,\mathrm{d}\mu(x) \leq \left(\int d(x,x_0)^p \,\mathrm{d}\mu(x)\right)^{\frac{q}{p}} \left(\int 1 \,\mathrm{d}\mu(x)\right)^{\frac{p-q}{p}} < \infty$$

Damit ist $\mu \in P_q(M)$. □

5.2 Das Desintegrationstheorem

Die Vermutung ist naheliegend, dass $(P_p(M), W_p)$ für alle $p \in [1,\infty)$ einen metrischen Raum bildet. Dies werden wir in Abschnitt 5.3 positiv beantworten. Für die Dreiecksungleichung benötigen wir das sogenannte Desintegrationstheorem, auf welches wir in diesem Abschnitt eingehen werden. Hierzu führen wir den Begriff des stochastischen Kerns ein. Einige Begriffe lassen sich in [Pau17, Abschnitt 6] nachlesen.

Definition 5.2.1. Seien (X, d_X) und (Y, d_Y) metrische Räume. Ein *stochastischer Kern* zwischen $(X, \mathfrak{B}(X))$ und $(Y, \mathfrak{B}(Y))$ ist eine Abbildung

$$\mu : X \times \mathfrak{B}(Y) \to [0,1]$$

mit folgenden Eigenschaften.

1. $\mu(x,\cdot)$ ist für jedes $x \in X$ ein Wahrscheinlichkeitsmaß.

2. $\mu(\cdot, A)$ ist für jedes $A \in \mathfrak{B}(Y)$ $\mathfrak{B}(X)$-messbar.

Wir schreiben $\mu_x(A) := \mu(x, A)$ für jedes $(x, A) \in X \times \mathfrak{B}(Y)$.

Lemma 5.2.2. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Sei $\mu^1 \in \mathcal{M}_1(\mathfrak{B}(X))$ und μ^2 ein stochastischer Kern zwischen $(X, \mathfrak{B}(X))$ und $(Y, \mathfrak{B}(Y))$. Dann ist die Abbildung ν definiert durch*

$$\nu(A) := \int \mu_x^2(A) \,\mathrm{d}\mu^1(x)$$

für $A \in \mathfrak{B}(Y)$ ein Wahrscheinlichkeitsmaß auf Y. Weiter gilt

$$\int f \,\mathrm{d}\nu = \int \int f(y) \,\mathrm{d}\mu_x^2(y) \,\mathrm{d}\mu^1(x)$$

für alle $\mathfrak{B}(Y)$ messbaren Abbildungen $f : Y \to [0, \infty)$.

Beweis. Die Gleichheit folgt direkt mit maßtheoretischer Induktion über den Aufbau von Treppenfunktionen. □

Weiter ist es möglich, mehrere Übergangskerne zusammenfügen.

Lemma 5.2.3. *Sei (X, d) ein metrischer Raum. Sei weiter $n \in \mathbb{N}$ und (X_i, d_i) für $i = 1, \ldots, n$ metrische Räume. Wir betrachten stochastische Kerne μ^i zwischen $(X, \mathfrak{B}(X))$ und $(X_i, \mathfrak{B}(X_i))$. Dann ist die Abbildung*

$$\mu : X \times \bigotimes_{k=1}^{n} \mathfrak{B}(X_k) \to [0, 1] \text{ mit } (x, A) \mapsto \left(\bigotimes_{k=1}^{n} \mu_x^k \right)(A)$$

wiederum selbst ein stochastischer Kern.

Beweis. Nach Definition ist μ_x für jedes $x \in X$ ein Wahrscheinlichkeitsmaß. Zu zeigen bleibt also die Messbarkeit. Wir setzen

$$\mathfrak{D} := \left\{ A \in \bigotimes_{k=1}^{n} \mathfrak{B}(X_i) : x \mapsto \mu_x(A) \text{ ist messbar} \right\}.$$

Da die Nullabbildung messbar ist, ist $\emptyset \in \mathfrak{D}$. Sei weiter $A \in \mathfrak{D}$. Dann ist die Abbildung $x \mapsto \mu_x(X \setminus A) = 1 - \mu_x(A)$ messbar und damit $X \setminus A \in \mathfrak{D}$. Falls

$(A_k)_{k\in\mathbb{N}} \subset \mathfrak{D}$ paarweise disjunkt sind, so ist mit der σ-Additivität von μ_x

$$x \mapsto \mu_x\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu_x(A_k)$$

messbar als punktweiser Grenzwert messbarer Funktionen. Somit ist $\mathfrak{D}$ ein Dynkin-System. Sei $A_1 \times \ldots \times A_n \in \mathcal{E} := \{B_1 \times \ldots \times B_n : B_i \in \mathfrak{B}(X_i) \text{ für } i = 1, \ldots n\}$. Dann ist

$$\mu_x(A_1 \times \ldots \times A_n) = \prod_{k=1}^{n} \mu_x^k(A_k)$$

als Produkt messbarer Funktionen messbar. Damit ist also $\mathcal{E} \subset \mathfrak{D}$. Das erzeugte Dynkin-System $\mathcal{D}(\mathcal{E})$ ist mit der $\cap$-Stabilität von $\mathcal{E}$ gleich der erzeugten σ-Algebra $\sigma(\mathcal{E})$ und wir erhalten

$$\bigotimes_{k=1}^{n} \mathfrak{B}(X_k) = \sigma(\mathcal{E}) = \mathcal{D}(\mathcal{E}) \subset \mathfrak{D}.$$

Damit ist $x \mapsto \mu_x(A)$ für alle $A \in \bigotimes_{k=1}^{n} \mathfrak{B}(X_k)$ messbar. □

Satz 5.2.4. Desintegrationstheorem nach Rokhlins. *Seien (X, d_X) und (Y, d_Y) vollständige und separable metrische Räume. Sei $T : X \to Y$ messbar und weiter $\mu \in \mathcal{M}_1(\mathfrak{B}(X))$. Dann gibt es einen stochastischen Kern ν zwischen $(Y, \mathfrak{B}(Y))$ und $(X, \mathfrak{B}(X))$, sodass*

$$\mu(A) = \int \nu_y(A)\,\mathrm{d}T_\#\mu(y)$$

für jedes $A \in \mathfrak{B}(X)$ gilt. Weiter gilt

$$\nu_y(X \setminus T^{-1}(\{y\})) = 0 \text{ für } T_\#\mu\text{-fast alle } y \in M_2.$$

Beweis. Nachzulesen in [Kel17, Theorem 2.23]. □

Lemma 5.2.5. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Weiter sei $x \in X$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}(Y))$. Dann gilt*

$$(\delta_x \otimes \mu_2)(A) = \mu_2(\{y \in Y : (x, y) \in A\})$$

für alle $A \in \mathfrak{B}(X) \otimes \mathfrak{B}(Y)$.

Beweis. Sei zunächst $A = A_1 \times A_2 \in \mathfrak{B}(X) \otimes \mathfrak{B}(Y)$. Dann gilt

$$\begin{aligned}(\delta_x \otimes \mu_2)(A_1 \times A_2) = \delta_x(A_1)\mu_2(A_2) &= \begin{cases} \mu_2(A_2) & x \in A_1 \\ \mu_2(\emptyset) & \text{sonst} \end{cases} \\ &= \mu_2(\{y \in \Omega_2 : (x, y) \in A_1 \times A_2\}).\end{aligned}$$

Mit der Eindeutigkeit des Produktmaßes folgt für alle $A \in \mathfrak{B}(X) \otimes \mathfrak{B}(Y)$ die Gleichheit. □

Wir möchten nun die Theorie der stochastischen Kerne und das Desintegrationstheorem verwenden, um Kopplungen auf eine neue Art und Weise darzustellen.

Korollar 5.2.6. *Seien (M_1, d_1) und (M_2, d_2) vollständige und separable metrische Räume mit $\mu_1 \in \mathcal{M}_1(\mathfrak{B}(M_1))$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}(M_2))$. Sei weiter π eine Kopplung zwischen μ_1 und μ_2. Dann gibt es einen stochastischen Kern ν zwischen M_1 und M_2, sodass*

$$\pi(A) = \int (\delta_x \otimes \nu_x)(A) \, \mathrm{d}\mu_1(x)$$

für alle $A \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ gilt.

Beweis. Mit dem Desintegrationstheorem (Satz 5.2.4) gibt es für die Projektion $p_1 : (x, y) \mapsto x$ einen stochastischen Kern $\tilde{\nu}$ zwischen M_1 und $M_1 \times M_2$, sodass

$$\pi(A) = \int \tilde{\nu}(x, A) \, \mathrm{d}p_{1\#}\pi(x) = \int \tilde{\nu}(x, A) \, \mathrm{d}\mu_1(x)$$

für $A \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ gilt. Zusätzlich gilt

$$\begin{aligned}0 &= \tilde{\nu}(x, (M_1 \times M_2) \setminus p_1^{-1}(\{x\})) = \tilde{\nu}(x, (M_1 \times M_2) \setminus (\{x\} \times M_2)) \\ &= \tilde{\nu}(x, (M_1 \setminus \{x\}) \times M_2)\end{aligned}$$

für μ_1-fast alle $x \in M_2$. Wir setzen

$$\nu : M_1 \times \mathfrak{B}(M_2) \to [0, 1] \text{ mit } (x, B) \mapsto \tilde{\nu}(x, \{x\} \times B) = \tilde{\nu}(x, M_1 \times B).$$

Wegen $\tilde{\nu}(x, \{x\} \times M_2) = 1$ ist ν_x für jedes $x \in M_1$ ein Wahrscheinlichkeitsmaß. Die Messbarkeit von $\tilde{\nu}$ in x überträgt sich auf ν, womit ν ein stochastischer Kern ist. Wir zeigen nun die geforderte Eigenschaft. Sei dazu $A \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$. Mit der Monotonie des Maßes ist $A \cap (M_1 \setminus \{x\}) \times M_2$ eine $\tilde{\nu}_x$-Nullmenge. Mit der Darstellung aus Lemma 5.2.5 gilt

$$\begin{aligned}\tilde{\nu}_x(A) &= \tilde{\nu}_x(A \cap (\{x\} \times M_2)) + \tilde{\nu}_x(A \cap (M_1 \setminus \{x\} \times M_2)) \\ &= \tilde{\nu}_x(A \cap (\{x\} \times M_2)) = \tilde{\nu}_x(\{x\} \times \{y \in M_2 : (x,y) \in A\}) \\ &= \nu_x(\{y \in M_2 : (x,y) \in A\}) = (\delta_x \otimes \nu_x)(A)\end{aligned}$$

für μ_1-fast alle $y \in M_2$. Damit sind insbesondere die Integrale $\int \tilde{\nu}(x, A) \, \mathrm{d}\mu_1(x)$ und $\int (\delta_x \otimes \nu_x)(A) \, \mathrm{d}\mu_1(x)$ gleich, womit die Behauptung folgt. Der Beweis orientiert sich an [Kel17, Korollar 3.6.]. □

Bemerkung 5.2.7. Der Ausdruck $\delta_x \otimes \nu_x$ beschreibt den relativen Transport ausgehend vom Punkt x. Dabei gibt $\nu_x(B)$ für $B \in \mathfrak{B}(M_2)$ den Anteil der Masse in x an, der in den Bereich B transportiert wird. Durch Aufintegrieren mit dem Maß μ_1 ergibt sich der Gesamttransport.

Bemerkung 5.2.8. Analog folgt für jedes $\pi \in \Pi(\mu_1, \mu_2)$ die Existenz eines stochastischen Kerns ν zwischen $(M_2, \mathfrak{B}(M_2))$ und $(M_1, \mathfrak{B}(M_1))$, sodass

$$\pi(A) = \int (\nu_y \otimes \delta_y)(A) \, \mathrm{d}\mu_2(y)$$

für alle $A \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ gilt.

Korollar 5.2.9. *Sei $x_0 \in M_1$ und $\mu_2 \in \mathcal{M}_1(\mathfrak{B}_2)$. Dann gibt es nur die triviale Kopplung $\delta_{x_0} \otimes \mu_2$ zwischen δ_{x_0} und μ_2. Insbesondere gilt*

$$W_p(\delta_{x_0}, \mu_2) = \left(\int d(x_0, x)^p \, \mathrm{d}\mu_2 \right)^{\frac{1}{p}}.$$

Beweis. Sei $\pi \in \Pi(\mu_1, \delta_{x_0})$. Mit Bemerkung 5.2.6 erhalten wir einen stochastischen Kern, sodass

$$\pi(A) = \int (\delta_x \otimes \nu_x)(A) \, \mathrm{d}\delta_{x_0}(x) = (\delta_{x_0} \otimes \nu_{x_0})(A)$$

für alle $A \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ gilt. Wegen $\mu_2 = p_{2\#}\pi = \nu_{x_0}$ folgt also $\pi = \delta_{x_0} \otimes \mu_2$. Die zweite Aussage folgt direkt mit dem Satz von Fubini. □

5.3 Metrische Eigenschaften

Wir können nun Kopplungen mithilfe von stochastischen Kernen darstellen. Dies ermöglicht uns später, wenn wir die Dreiecksungleichung beweisen wollen, zwei Kopplungen „aneinander zu kleben[3]“. Zunächst benötigen wir aber noch ein einfaches Lemma, um die Symmetrie zu zeigen. Sei im Folgenden (M, d) ein vollständiger und separabler metrischer Raum.

Lemma 5.3.1. *Sei $f : M \times M \to M \times M$ mit $(x_1, x_2) \mapsto (x_2, x_1)$. Dann ist π genau dann eine Kopplung zwischen μ_1 und μ_2, wenn $f_\#\pi$ eine Kopplung zwischen μ_2 und μ_1 ist. Das heißt es gilt $\Pi(\mu_1, \mu_2) = \{f_\#\pi : \pi \in \Pi(\mu_2, \mu_1)\}$.*

Beweis. Sei $\pi \in \Pi(\mu_1, \mu_2)$ und $A \in \mathfrak{B}$. Dann gilt

$$f_\#\pi(A \times M) = \pi(f^{-1}(A \times M)) = \pi(M \times A) = \mu_2(A).$$

Analog folgt $f_\#\pi(M \times A) = \mu_1(A)$. Damit ist $f_\#\pi \in \Pi(\mu_2, \mu_1)$. Die Rückrichtung funktioniert genauso. $\square$

Satz 5.3.2. *$(P_p(M), W_p)$ bildet für jedes $p \in [1, \infty)$ einen metrischen Raum.*

Beweis. Wir zeigen die drei Axiome der Metrik.

1. *Positive Definitheit:* $W_p(\mu_1, \mu_2) \geq 0$ folgt direkt mit der Nichtnegativität der Metrik. Sei $W_p(\mu_1, \mu_2) = 0$. Mit Bemerkung 5.1.2 erhalten wir eine Kopplung $\pi^* \in \Pi(\mu_1, \mu_2)$ mit

$$0 = \int d(x_1, x_2)^p \, \mathrm{d}\pi^*(x_1, x_2).$$

Mit der Definitheit von d gilt $x_1 = x_2$ für π^*-fast alle $(x_1, x_2) \in M \times M$. Weiter gilt mit Lemma 4.3.3

$$\begin{aligned} 0 &= \pi^*(\{(x_1, x_2) \in M \times M : x_1 \neq x_2\}) \\ &\geq \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int \mathbb{1}_{\{(x_1,x_2) \in M \times M : x_1 \neq x_2\}} \, \mathrm{d}\pi = \frac{1}{2}\|\mu_1 - \mu_2\| \geq 0. \end{aligned}$$

Mit der Definitheit von $\|\cdot\|_{\mathrm{TV}}$ folgt die Gleichheit von μ_1 und μ_2.

[3] Es sei auf das „Gluing lemma“ in [Vil09] verwiesen.

2. *Symmetrie:* Mit Lemma 5.3.1 mit der dort definierten Abbildung f und der Transformationsformel und der Symmetrie der Metrik folgt

$$\begin{aligned} W_p(\mu_1,\mu_2)^p &= \inf_{\pi\in\Pi(\mu_1,\mu_2)} \int d(x_1,x_2)^p \,\mathrm{d}\pi(x_1,x_2) \\ &= \inf_{\pi\in\Pi(\mu_1,\mu_2)} \int d(f(x_2,x_1))^p \,\mathrm{d}\pi(x_1,x_2) \\ &= \inf_{\pi\in\Pi(\mu_1,\mu_2)} \int d(x_2,x_1)^p \,\mathrm{d}f_\#\pi(x_1,x_2) \\ &= \inf_{\tilde\pi\in\{f_\#\pi:\pi\in\Pi(\mu_1,\mu_2)\}} \int d(x_2,x_1)^p \,\mathrm{d}\tilde\pi(x_1,x_2) \\ &= \inf_{\tilde\pi\in\Pi(\mu_2,\mu_1)} \int d(x_2,x_1)^p \,\mathrm{d}\tilde\pi(x_1,x_2) = W_p(\mu_2,\mu_1)^p. \end{aligned}$$

3. *Dreiecksungleichung:* Seien $\mu_1,\mu_2,\mu_3 \in P_p(M)$. Seien mit Bemerkung 5.1.2 $\pi_1 \in \Pi(\mu_1,\mu_2)$ und $\pi_2 \in \Pi(\mu_2,\mu_3)$ jeweils optimale Kopplungen. Wir erhalten mit dem Lemma 5.2.6 und Bemerkung 5.2.8 stochastische Kerne ν^1 und ν^2 auf $(M,\mathfrak{B})$, sodass

$$\pi_1(A) = \int (\nu_y^1 \otimes \delta_y)(A)\,\mathrm{d}\mu_2(y) \text{ und } \pi_2(A) = \int (\delta_y \otimes \nu_y^2)(A)\,\mathrm{d}\mu_2(y)$$

für alle $A \in \mathfrak{B}(M)\otimes\mathfrak{B}(M)$ gilt. Nach Lemma 5.2.3 ist auch $\nu_x := \nu_x^1 \otimes \nu_x^2$ ein stochastischer Kern. Wir definieren die Abbildung

$$\hat\pi : \mathfrak{B}(M)\otimes\mathfrak{B}(M) \mapsto [0,1] \text{ mit } C \mapsto \int \nu_x(C)\,\mathrm{d}\mu_2(x).$$

Mit Lemma 5.2.2 ist $\hat\pi$ ein Wahrscheinlichkeitsmaß. Wir zeigen nun, dass $\hat\pi$ eine Kopplung zwischen μ_1 und μ_3 ist. Sei $A \in \mathfrak{B}$. Dann gilt

$$\begin{aligned} \hat\pi(A\times M) &= \int \nu_x(A\times M)\,\mathrm{d}\mu_2(x) = \int \nu_x^1(A)\nu_x^2(M)\,\mathrm{d}\mu_2(x) \\ &= \int \nu_x^1(A)\,\mathrm{d}\mu_2(x) = \int \nu_x^1(A)\delta_x(M)\,\mathrm{d}\mu_2(x) \\ &= \int (\nu_x^1\otimes\delta_x)(A\times M)\,\mathrm{d}\mu_2(x) = \pi_1(A\times M) = \mu_1(A). \end{aligned}$$

Analog folgt $\hat\pi(M\times A) = \pi_2(M\times A) = \mu_3(A)$, womit also $\hat\pi \in \Pi(\mu_1,\mu_3)$.

Mit Lemma 5.2.2, Fubini und der Minkowskiungleichung[4] gilt

$$\begin{aligned}
W(\mu_1, \mu_3) &\leq \left(\int d(x_1, x_3)^p \,\mathrm{d}\hat{\pi}(x_1, x_3) \right)^{\frac{1}{p}} \\
&= \left(\int \int d(x_1, x_3)^p \,\mathrm{d}\nu_{x_2}(x_1, x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&= \left(\int \int \int d(x_1, x_3)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\nu^2_{x_2}(x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&\leq \left(\int \int \int d(x_1, x_2)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\nu^2_{x_2}(x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&+ \left(\int \int \int d(x_2, x_3)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\nu^2_{x_2}(x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}}.
\end{aligned}$$

Mit der Optimalität von μ_1 gilt

$$\begin{aligned}
&\left(\int \int \int d(x_1, x_2)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\nu^2_{x_2}(x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&= \left(\int \int d(x_1, x_2)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&= \left(\int \int \int d(x_1, x)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\delta_{x_2}(x) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&= \left(\int \int d(x_1, x)^p \,\mathrm{d}(\nu^1_{x_2} \otimes \delta_{x_2})(x_1, x) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} \\
&= \left(\int d(x_1, x)^p \,\mathrm{d}\pi_1(x_1, x) \right)^{\frac{1}{p}} = W(\mu_1, \mu_2)
\end{aligned}$$

für den ersten Summanden. Analog gilt

$$\left(\int \int \int d(x_2, x_3)^p \,\mathrm{d}\nu^1_{x_2}(x_1) \,\mathrm{d}\nu^2_{x_2}(x_3) \,\mathrm{d}\mu_2(x_2) \right)^{\frac{1}{p}} = W_p(\mu_2, \mu_3)$$

für den zweiten Term. Somit folgt die Dreiecksungleichung für alle μ_1, μ_2, μ_3.

[4] Streng genommen können wir hier die Minkowski-Ungleichung nicht verwenden, da wir drei ineinander geschachtelte Integrale haben, die sich im Allgemeinen nicht zu einem einzigen Integral auf $M_1 \times M_2 \times M_3$ zusammenfassen lassen. Allerdings kann man den Beweis der Hölder-Ungleichung und der resultierenden Minkowski-Ungleichung kopieren und in diesem Setting anwenden.

Der Beweis zur Dreiecksungleichung ist an [Kel17, Satz 5.5.] angelehnt. □

Bemerkung 5.3.3. Es gibt einen elementaren Beweis der Dreiecksungleichung des Wassersteinabstandes (siehe [CD08]), der nicht das Desintegrationstheorem benutzt.

Lemma 5.3.4. *Für alle $\mu_1, \mu_2 \in P_1(M)$ und $f \in Lip^1(M)$ gilt die Ungleichung*

$$\int f \,\mathrm{d}\mu_1 - \int f \,\mathrm{d}\mu_2 \leq W_1(\mu_1, \mu_2),$$

wobei $Lip^1(M) := \{f : M \to \mathbb{R} : f$ ist Lipschitzstetig mit Konstante maximal $1\}$.

Beweis. Sei $\pi^* \in \Pi(\mu_1, \mu_2)$ eine optimale Kopplung. Dann gilt mit Lemma 4.1.1

$$\begin{aligned}\int f \,\mathrm{d}\mu_1 - \int f \,\mathrm{d}\mu_2 &= \int f(x_1) - f(x_2) \,\mathrm{d}\pi^*(x_1, x_2) \leq \int |f(x_1) - f(x_2)| \,\mathrm{d}\pi^*(x_1, x_2)\\ &\leq \int d(x_1, x_2) \,\mathrm{d}\pi^*(x_1, x_2) = W_1(\mu_1, \mu_2)\end{aligned}$$

für alle $f \in \mathrm{Lip}^1(M)$. □

Bemerkung 5.3.5. Es gilt sogar

$$\sup_{f \in \mathrm{Lip}^1(M)} \int f \,\mathrm{d}\mu_1 - \int f \,\mathrm{d}\mu_2 = W_1(\mu_1, \mu_2)$$

für alle $\mu_1, \mu_2 \in P_1(M)$. Man spricht hier von der *Kantorovich-Rubinstein Dualität*. Siehe auch [Vil09, Particular Case 5.16.].

Im nächsten Schritt möchten wir zeigen, dass sich die Vollständigkeit von (M, d) auf $(P_p(M), W_p)$ überträgt. Weiter möchten wir die Konvergenz bezüglich des Wassersteinabstandes charakterisieren. Hierzu benötigen wir folgendes Lemma:

Lemma 5.3.6. *Sei $\mu, \hat{\mu} \in P_p(M)$ und $\mu_k \in P_p(M)$ schwach konvergent gegen μ. Dann gilt*

$$W_p(\mu, \hat{\mu}) \leq \limsup_{k \to \infty} W_p(\mu_k, \hat{\mu}).$$

Beweis. Seien π_k optimale Kopplungen zwischen μ_k und $\hat{\mu}$. Da $\{\mu_k : k \in \mathbb{N}\} \cup \{\mu\}$ und $\{\hat{\mu}\}$ jeweils schwach kompakt sind, ist auch die Menge

$$\Pi(\{\mu_k : k \in \mathbb{N}\} \cup \{\mu\}, \{\hat{\mu}\}),$$

in welcher alle π_k enthalten sind, ebenfalls schwach kompakt (Eigenschaften 4.1.4). Somit gibt es eine schwach gegen ein π konvergente Teilfolge π_{n_k}. Wir zeigen nun, dass π eine Kopplung zwischen μ und $\hat{\mu}$ ist. Da

$$(\pi_{n_k})_{k\geq m} \subset \Pi(\{\mu_k : k \geq m\} \cup \{\mu\}, \{\hat{\mu}\}) \text{ für alle } m \in \mathbb{N},$$

$(\pi_{n_k})_{k\geq m}$ nach wie vor schwach gegen π konvergiert und $\Pi(\{\mu_k : k \geq m\} \cup \{\mu\}, \{\hat{\mu}\})$ schwach abgeschlossen ist (Eigenschaften 4.1.4), folgt

$$\pi \in \bigcap_{m=1}^{\infty} \Pi(\{\mu_k : k \geq m\} \cup \{\mu\}, \{\hat{\mu}\}).$$

Insbesondere ist für $p_1 : (x_1, x_2) \mapsto x_1$

$$p_{1\#}\pi \in \bigcap_{m=1}^{\infty} \{\mu_k : k \geq m\} \cup \{\mu\} = \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} \{\mu_k\} \cup \{\mu\}. \tag{5.1}$$

Angenommen $p_{1\#}\pi \neq \mu$. Dann können wir mit (5.1) eine Teilfolge (μ_{m_k}) wählen, sodass $\mu_{m_k} = p_{1\#}\pi$ für alle $k \in \mathbb{N}$ gilt. Allerdings konvergiert nach Voraussetzung $(\mu_{m_k})_{k\in\mathbb{N}}$ gegen μ, was zu einem Widerspruch führt. Somit ist $\pi \in \Pi(\mu, \hat{\mu})$. Mit Portmanteau (Satz 2.3.5) gilt

$$\begin{aligned} W_p(\mu, \hat{\mu})^p &\leq \int d(x_1, x_2)^p \, \mathrm{d}\pi(x_1, x_2) \leq \liminf_{k\to\infty} \int d(x_1, x_2)^p \, \mathrm{d}\pi_{n_k}(x_1, x_2) \\ &= \liminf_{k\to\infty} W_p^p(\mu_{n_k}, \hat{\mu}) \leq \limsup_{k\to\infty} W_p^p(\mu_k, \hat{\mu}) = \left(\limsup_{k\to\infty} W_p(\mu_k, \hat{\mu})\right)^p. \end{aligned}$$

Durch Anwenden der p-ten Wurzel folgt die Behauptung. □

Bemerkung 5.3.7. Es gilt sogar

$$W_p(\mu, \hat{\mu}) \leq \liminf_{k\to\infty} W_p(\mu_k, \hat{\mu}_k)$$

für schwach gegen μ und $\hat{\mu}$ konvergente Folgen $(\mu_k)_{k\in\mathbb{N}}$ und $(\hat{\mu}_k)_{k\in\mathbb{N}}$ in $P_p(M)$ (siehe [Kel17, Satz 4.23]). Das heißt die Wassersteinmetrik ist in beiden Komponenten schwach unterhalb stetig. Für unseren Zweck genügt jedoch das hier bewiesene Lemma.

Wir kommen nun zum Hauptresultat in diesem Abschnitt. Das folgende Lemma charakterisiert die Konvergenz in $(P_p(M), W_p)$:

Satz 5.3.8. *Der metrische Raum $(P_p(M), W_p)$ ist vollständig. Weiterhin sind für $(\mu_k)_{k\in\mathbb{N}} \subset P_p(M)$ und $\mu \in P_p(M)$ folgende Aussagen äquivalent:*

1. *$\mu_k \to \mu$ in $(P_p(M), W_p)$ für $k \to \infty$.*
2. *$\mu_k \to \mu$ schwach für $k \to \infty$ und*

$$\int d(x, x_0)^p \,\mathrm{d}\mu_k \to \int d(x, x_0)^p \,\mathrm{d}\mu \text{ für } k \to \infty$$

   für ein $x_0 \in M$.

Beweis. Sei zunächst $(\mu_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $(P_p(M), W_p)$. Die Prochorov-Metrik metrisiert die schwache Topologie (Satz 2.4.7 und Satz 2.4.8). Wir zeigen zunächst, dass $(\mu_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $(\mathcal{M}(M), d_P)$ ist. Sei $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, sodass für alle $k, l \geq N$ die Abschätzung (mit Lemma 5.1.3)

$$W_1(\mu_k, \mu_l) \leq W_p(\mu_k, \mu_l) < \varepsilon^2$$

gilt. Sei $C \subset M$ abgeschlossen. Wir bekommen mit dem Lemma von Urysohn für Umgebungen (Lemma A.2.5) eine Abbildung $f : M \to [0, 1]$, die lipschitzstetig mit Konstante $\frac{1}{\varepsilon}$ ist und $\mathbb{1}_C(x) \leq f(x) \leq \mathbb{1}_{C_\varepsilon}(x)$ für jedes $x \in M$ erfüllt. Da εf Lipschitzkonstante 1 besitzt, ist εf eine legitime Funktion für die Kantorovich-Rubinstein Dualität (Lemma 5.3.4) und wir erhalten

$$W_1(\mu_m, \mu_n) \geq \varepsilon \left(\int f \,\mathrm{d}\mu_m - \int f \,\mathrm{d}\mu_n \right).$$

Schließlich gilt

$$\begin{aligned}
\mu_m(C) &= \int \mathbb{1}_C \,\mathrm{d}\mu_m + \frac{1}{\varepsilon} W_1(\mu_m, \mu_n) - \frac{1}{\varepsilon} W_1(\mu_m, \mu_n) \\
&\leq \int \mathbb{1}_C \,\mathrm{d}\mu_m + \varepsilon - \frac{1}{\varepsilon} W_1(\mu_m, \mu_n) \\
&\leq \int \mathbb{1}_C \,\mathrm{d}\mu_m + \varepsilon - \left(\int f \,\mathrm{d}\mu_m - \int f \,\mathrm{d}\mu_n \right) \\
&\leq \int f \,\mathrm{d}\mu_m + \varepsilon - \left(\int f \,\mathrm{d}\mu_m - \int f \,\mathrm{d}\mu_n \right) \\
&= \varepsilon + \int f \,\mathrm{d}\mu_n \leq \varepsilon + \int \mathbb{1}_{C_\varepsilon} \,\mathrm{d}\mu_n = \varepsilon + \mu_n(C).
\end{aligned}$$

Mit Satz 2.4.4 erhalten wir direkt $d_P^*(\mu_m, \mu_n) \leq \varepsilon$. Da wir nur Wahrscheinlichkeitsmaße betrachten, folgt mit Lemma 2.4.3 bereits, dass $d_P^*(\mu_m, \mu_n) = d_P^*(\mu_n, \mu_m)$. Somit ist $(\mu_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $(\mathcal{M}^+(\mathfrak{B}), d_P)$ und wegen der Vollständigkeit (Satz 2.5.10) gibt es ein $\mu \in \mathcal{M}^+(\mathfrak{B})$, sodass $\mu_n \to \mu$ schwach für $n \to \infty$ konvergiert. Insbesondere folgt $1 = \|\mu_n\|_{\mathrm{TV}} \to \|\mu\|_{\mathrm{TV}}$, wodurch $\mu \in \mathcal{M}_1(\mathfrak{B})$. Wir zeigen nun, dass μ endliches p-tes Moment besitzt. Mit Korollar 5.2.9 und der umgekehrten Dreiecksungleichung erhalten wir

$$\begin{aligned} \left| \int d(x_0, x)^p \,\mathrm{d}\mu_m - \int d(x_0, x)^p \,\mathrm{d}\mu_n \right| &= |W_p(\mu_m, \delta_{x_0}) - W_p(\mu_n, \delta_{x_0})| \\ &\leq |W_p(\mu_m, \mu_n)| < \varepsilon \end{aligned}$$

für $m, n \geq N$. Also ist $(\int d(x_0, x)^p \,\mathrm{d}\mu_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$. Es folgt Konvergenz und insbesondere Beschränktheit. Da $d(x_0, \cdot)^p$ insbesondere unterhalbstetig ist, erhalten wir mit Portmanteau (Satz 2.3.5)

$$\int d(x_0, x)^p \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int d(x_0, x)^p \,\mathrm{d}\mu_n < \infty.$$

Womit also $\mu \in P_p(M)$.

Wir zeigen nun, dass tatsächlich μ_n gegen den Kandidaten μ in $(P_p(M), W_p)$ für $n \to \infty$ konvergiert. Sei $k \geq N$. Dann gilt mit Lemma 5.3.6

$$W_p(\mu, \mu_k) \leq \limsup_{\substack{l\to\infty \\ l\geq k}} W_p(\mu_l, \mu_k) \leq \varepsilon.$$

Somit ist die Konvergenz und die Vollständigkeit von $W_p(M)$ gezeigt. Ebenfalls haben wir bereits gezeigt, dass falls $\mu_k \to \mu$ in $(P_p(M), W_p)$ konvergiert, dass dann $\mu_k \to \mu$ bereits schwach konvergiert. Bisher wissen wir nur, dass das p-te Moment konvergiert, allerdings noch nicht, dass der Grenzwert tatsächlich dem p-ten Moment von μ entspricht. Dies folgt aber analog mit

$$\begin{aligned} 0 &\leq \left| \int d(x_0, x)^p \,\mathrm{d}\mu - \int d(x_0, x)^p \,\mathrm{d}\mu_n \right| \\ &= |W_p(\mu, \delta_{x_0}) - W_p(\mu_n, \delta_{x_0}| \leq |W_p(\mu, \mu_n)| \to 0 \end{aligned}$$

für $n \to \infty$. Der Beweis orientiert sich an [Vil09, Lemma 6.14]. Für die Rückrichtung sei auf [Kel17, Satz 5.7.] verwiesen. Diese benötigt ein Stabilitätsresultat, dessen Ausführung den Rahmen sprengen würde. □

A Appendix

Im Folgenden möchten wir einige Hilfsresultate liefern, die sich jedoch thematisch vom Hauptteil abgrenzen.

A.1 Allgemeine Begriffe

Satz A.1.1. *Sei* $(V, \|\cdot\|)$ *ein endlichdimensionaler, normierter Vektorraum. Sei* $(x_k)_{k\in\mathbb{N}}$ *eine Folge in* V*. Dann sind folgende Aussagen äquivalent:*

1. *Die Reihe* $\sum_{k=1}^{\infty} a_k$ *konvergiert absolut.*
2. *Für jede Bijektion* $\pi : \mathbb{N} \to \mathbb{N}$ *konvergiert* $\sum_{k=1}^{\infty} a_{\pi(k)}$.

Beweis. Ein Korollar aus dem Steinitzschen Umordnungssatz. □

Definition A.1.2. Sei V ein Vektorraum. Eine Relation $\leq$ auf V heißt *partielle Ordnung*, falls die folgenden Eigenschaften gelten:

1. *Reflexivität.* Es gilt $x \leq x$ für alle $x \in V$.
2. *Antisymmetrie.* Aus $x \leq y$ und $y \leq x$ folgt bereits $x = y$ für alle $x, y \in V$.
3. *Transitivität.* Aus $x \leq y$ und $y \leq z$ folgt bereits $x \leq z$ für alle $x, y, z \in V$.

Das Tupel $(V, \leq)$ bezeichnen wir als partiell geordneten Vektorraum.

A.2 Topologische Begriffe

Wir erinnern uns an folgende Charakterisierung von Kompaktheit.

Satz A.2.1. *Sei* (M, d) *ein metrischer Raum und* $K \subset M$*. Dann sind folgende Aussagen äquivalent:*

1. *K ist kompakt, das heißt für jede offene Überdeckung von K gibt es eine endliche Teilüberdeckung.*
2. *Jede Folge in K besitzt eine konvergente Teilfolge, dessen Grenzwert in K enthalten ist.*
3. *Der Raum $(K, d_{|K\times K})$ ist vollständig und totalbeschränkt, das heißt für jedes $\varepsilon > 0$ gibt es endlich viele $x_1, \ldots, x_n \in K$, sodass $K \subset \bigcup_{k=1}^n B_\varepsilon(x_k)$ gilt.*

Weiter sind für ein $C \subset M$ diese Aussagen äquivalent:

1. *C ist relativ kompakt, das heißt $\overline{C}$ ist kompakt.*
2. *C ist totalbeschränkt und $(\overline{C}, d_{|\overline{C}\times\overline{C}})$ ist vollständig.*
3. *Jede Folge in C besitzt eine konvergente Teilfolge.*

Definition A.2.2. *Distanzfunktion.* Sei (M, d) ein metrischer Raum. Wir setzen

$$\mathrm{dist} : M \times (\mathfrak{P}(M) \setminus \{\emptyset\}) \to [0, \infty) \text{ mit } (x, A) \mapsto \inf_{y\in A} d(x, y).$$

Lemma A.2.3. *Sei (M, d) ein metrischer Raum. Die Abbildung dist$(A, \cdot)$ ist für jedes feste, nichtleere $A \subset M$ Lipschitzstetig mit Lipschitzkonstante* 1.

Beweis. Sei $A \subset M$ mit $A \neq \emptyset$. Sei $x_1, x_2 \in M$. Dann gilt

$$\begin{aligned}\mathrm{dist}(x_1, A) - \mathrm{dist}(x_2, A) &= \inf_{y\in A} d(x_1, y) - \inf_{y\in A} d(x_2, y)\\ &\leq \inf_{y\in A}(d(x_1, x_2) + d(x_2, y)) - \inf_{y\in A} d(x_2, y) = d(x_1, x_2)\end{aligned}$$

Analog erhalten wir $\mathrm{dist}(x_2, A) - \mathrm{dist}(x_1, A) \leq d(x_1, x_2)$ und damit

$$|\mathrm{dist}(x_1, A) - \mathrm{dist}(x_2, A)| \leq d(x_1, x_2),$$

womit die Lipschitzstetigkeit gezeigt ist. □

Lemma A.2.4. Urysohn für metrische Räume. *Sei (M, d) ein metrischer Raum und $A, B \subset M$ abgeschlossen, nicht leer und disjunkt. Dann gibt es eine stetige Funktion $f : M \to [0, 1]$ mit $f(x) = 0$ für alle $x \in A$ und $f(x) = 1$ für alle $y \in B$.*

Beweis. Wähle $f : M \to [0, 1]$ mit $x \mapsto \frac{\mathrm{dist}(x,A)}{\mathrm{dist}(x,A)+\mathrm{dist}(x,B)}$. Da A und B abgeschlossen sind, gilt $\mathrm{dist}(x, A) = 0$ (bzw. $\mathrm{dist}(x, B) = 0$) genau dann, wenn $x \in A$ (bzw. $x \in B$). Da mit der Disjunktheit von A und B entweder $\mathrm{dist}(x, A) \neq 0$ oder

$\mathrm{dist}(x, B) \neq 0$ ist, ist f wohldefiniert. Die Stetigkeit folgt mit der Verkettung stetiger Funktionen (die Distanzabbildung ist nach Lemma A.2.3 stetig) wiederum selbst stetig. Falls $x \in A$ ist $f(x) = \frac{0}{0+\mathrm{dist}(x,B)} = 0$ und falls $x \in B$ ist $f(x) = \frac{\mathrm{dist}(x,A)}{\mathrm{dist}(x,A)+0} = 1$. Damit erfüllt f die geforderten Eigenschaften. □

Nun zeigen wir einen Spezialfall des Lemmas von Urysohn, wenn $B = M \setminus A_\varepsilon$ für ein $\varepsilon > 0$ betrachtet wird. Wir erhalten dadurch zusätzlich eine Lipschitzbedingung für f.

Lemma A.2.5. Urysohn: Spezialfall. *Sei (M, d) ein metrischer Raum und $A \subset M$ nicht leer und abgeschlossen. Sei $\varepsilon > 0$. Dann gibt es eine mit Konstante $\frac{1}{\varepsilon}$ lipschitzstetige Abbildung $f_\varepsilon : M \to [0, 1]$ mit*

$$f_\varepsilon(x) = 1 \text{ für alle } x \in A \text{ und } f_\varepsilon(x) = 0 \text{ für alle } x \in M \setminus A_\varepsilon.$$

Es gilt also $\mathbb{1}_A(x) \leq f_\varepsilon(x) \leq \mathbb{1}_{A_\varepsilon}(x)$ für jedes $x \in M$. Insgesamt erhalten wir $f_\varepsilon \to \mathbb{1}_A$ für $\varepsilon \to 0$ punktweise.

Beweis. Wir setzen

$$f : M \to [0, 1] \text{ mit } x \mapsto \begin{cases} 1 - \frac{1}{\varepsilon}\mathrm{dist}(x, A) & \mathrm{dist}(x, A) \leq \varepsilon \\ 0 & \text{sonst.} \end{cases}$$

Falls $x \in A$, dann gilt direkt $\mathrm{dist}(x, A) = 0$ und damit $f(x) = 1$. Falls $x \in M \setminus A_\varepsilon$ gilt $\mathrm{dist}(x, A) > \varepsilon$ und damit nach Definition $f(x) = 0$. Zu zeigen bleibt also nur die Lipschitzstetigkeit.

1. Für $x, y \in M \setminus A_\varepsilon$ gilt sofort $|f(x) - f(y)| = |0 - 0| \leq d(x, y)$.
2. Für $y \in M \setminus A_\varepsilon$ und $x \in A_\varepsilon$ ist $d(y, z) \geq \varepsilon$ für alle $z \in A$. Damit folgt

$$\begin{aligned} |f(x) - f(y)| = |f(x)| = 1 - \frac{1}{\varepsilon}\mathrm{dist}(x, A) &= \frac{1}{\varepsilon}\left(\varepsilon - \inf_{z\in A} d(x, z)\right) \\ &= \frac{1}{\varepsilon}\sup_{z\in A}(\varepsilon - d(x, z)) \leq \frac{1}{\varepsilon}\sup_{z\in A}(d(y, z) - d(x, z)) \\ &\leq \frac{1}{\varepsilon}\left(\sup_{z\in A} d(y, x) + d(x, z) - d(x, z)\right) = \frac{1}{\varepsilon}d(x, y). \end{aligned}$$

3. Für $x, y \in A_\varepsilon$ gilt wegen der Lipschitzstetigkeit von der Distanzabbildung

(Lemma A.2.3) mit Kostante 1

$$\begin{aligned}|f(x) - f(y)| &= |(1 - \frac{1}{\varepsilon}\text{dist}(x, A)) - (1 - \frac{1}{\varepsilon}\text{dist}(y, A))| \\ &= \frac{1}{\varepsilon}|\text{dist}(y, A) - \text{dist}(x, A)| \leq \frac{1}{\varepsilon}d(y, x).\end{aligned}$$

Da A abgeschlossen ist, erhalten wir $\mathbb{1}_{A_\varepsilon} \to \mathbb{1}_A$ für $\varepsilon \to 0$ punktweise. Mit dem Einschließungskriterium ist die Aussage gezeigt.

□

Lemma A.2.6. Cantor's Diagonalargument. *Sei $(X_p, \tau_p)_{p \in I}$ eine abzählbare Familie von folgenkompakten topologischen Räumen. Sei weiter $(x_k^p)_{k \in \mathbb{N}} \subset X_p$ für jedes $p \in I$ eine Folge. Dann gibt es eine streng monotone Indexfolge $(n_k)_{k \in \mathbb{N}} \subset \mathbb{N}$ und für jedes $p \in I$ ein $x^p \in X_p$, sodass*

$$x_{n_k}^p \to x^p \textit{ für } k \to \infty \textit{ und alle } p \in I \textit{ gilt.}$$

gilt.

Beweis. Nachzulesen in [Are18, Theorem 5.1]. □

A.3 Maßtheoretische Begriffe

Definition A.3.1. *Äußeres Maß.* Sei M eine nicht leere Menge. Eine Abbildung

$$\mu : \mathfrak{P}(M) \to [0, \infty) \cup \{\infty\}$$

heißt *äußeres Maß*, falls die folgenden Eigenschaften erfüllt sind:

1. $\mu(\emptyset) = 0$
2. *Monotonie.* $\mu(A) \leq \mu(B)$ für alle $A, B \subset M$ mit $A \subset B$.
3. *σ-Subadditivität.* $\mu(\bigcup_{k=1}^\infty A_k) \leq \sum_{k=1}^\infty \mu(A_k)$ für alle $(A_k)_{k \in \mathbb{N}} \subset M$.

Im Kontext von äußeren Maßen definieren wir den Begriff der Messbarkeit.

Definition A.3.2. Sei M eine Menge und μ ein äußeres Maß auf M. Eine Menge $A \subset M$ heißt *Carathéodory-messbar*, wenn $\mu(Q) = \mu(Q \cap A) + \mu(Q \cap (M \setminus A))$ für alle $Q \subset M$ gilt.

Satz A.3.3. Satz von Carathéodory. *Sei μ ein äußeres Maß. Dann ist die Menge $\mathfrak{A} := \{A \subset M : A$ Carathéodory-messbar$\}$ eine σ-Algebra und $\mu_{|\mathfrak{A}}$ ist ein Maß.*

Beweis. Nachzulesen in [Are17, Satz 13.3]. □

Definition A.3.4. Sei (M, d) ein metrischer Raum mit der zugehörigen Borelschen σ-Algebra $\mathfrak{B}$. Ein Maß $\mu \in \mathcal{M}^+(\mathfrak{B})$ heißt

1. von *innen regulär*, falls

$$\mu(A) = \sup_{\substack{K \subset A \\ K \text{ kompakt}}} \mu(K).$$

   für jedes $A \in \mathfrak{B}$ gilt. Das heißt das Maß jeder Menge kann von innen mit dem Maß kompakter Mengen approximiert werden.

2. von *außen regulär*, falls

$$\mu(A) = \sup_{\substack{U \supset A \\ U \text{ offen}}} \mu(U)$$

   für jedes $A \in \mathfrak{B}$ gilt. Das heißt das Maß jeder Menge kann von außen mit dem Maß offener Mengen approximiert werden.

3. *regulär*, falls es von innen und von außen regulär ist.

Lemma A.3.5. Prinzip von Cavalieri. *Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. Sei weiter $f : \Omega \to [0, \infty]$ messbar. Dann gilt*[1]

$$\int_\Omega f \,\mathrm{d}\mu = \int_{[0,\infty]} \mu(f^{-1}([y, \infty)) \,\mathrm{d}\lambda(y) = \int_{[0,\infty]} \mu(f^{-1}((y, \infty)) \,\mathrm{d}\lambda(y).$$

[1] Die Abbildung $\mu(f^{-1}([\cdot, \infty))$ ist monoton fallend und damit messbar.

Beweis. Mit dem Satz von Tonelli gilt

$$\begin{aligned}
\int_\Omega f \,\mathrm{d}\mu &= \int_\Omega \int_{[0,\infty)} \mathbb{1}_{\{\tilde{y}\in\mathbb{R}:\tilde{y}\leq f(x)\}}(y)\,\mathrm{d}\lambda(y)\,\mathrm{d}\mu(x) \\
&= \int_\Omega \int_{[0,\infty)} \mathbb{1}_{\{(\tilde{x},\tilde{y})\in\Omega\times\mathbb{R}:\tilde{y}\leq f(\tilde{x})\}}(x,y)\,\mathrm{d}\lambda(y)\,\mathrm{d}\mu(x) \\
&= \int_{[0,\infty)} \int_\Omega \mathbb{1}_{\{(\tilde{x},\tilde{y})\in\Omega\times\mathbb{R}:\tilde{y}\leq f(\tilde{x})\}}(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \\
&= \int_{[0,\infty)} \int_\Omega \mathbb{1}_{\{\tilde{x}\in\Omega:y\leq f(\tilde{x})\}}(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \\
&= \int_{[0,\infty)} \int_\Omega \mathbb{1}_{\{\tilde{x}\in\Omega:y\leq f(\tilde{x})\}}(x)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \\
&= \int_{[0,\infty)} \mu(\{x\in\Omega : y \leq f(x)\})\,\mathrm{d}\lambda(y).
\end{aligned}$$

Die zweite Gleichheit folgt analog mit $\lambda([0, f(x))) = \lambda([0, f(x)])$. $\square$

Literaturverzeichnis

[Are17] W. Arendt. Skript zur Vorlesung „Maßtheorie,“, Wintersemester 16/17.

[Are18] W. Arendt. Skript zur Vorlesung „Funktionalanalysis,“, Wintersemester 17/18.

[Are14] W. Arendt. Einführung in Banachverbände und positive Operatoren. https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/gerlach/ws1314/fa/positivity.pdf, 2014. [Aufgerufen am 14. August 18].

[CD08] Philippe Clement and Wolfgang Desch. An elementary proof of the triangle inequality for the Wasserstein metric. *Proc. Amer. Math. Soc.*, 136(1):333–339, 2008.

[Els09] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, fourth edition, 2009. Grundwissen Mathematik. [Basic Knowledge in Mathematics].

[Kel17] M. Kell. Skript zur Vorlesung „Theorie des optimalen Transports“. https://www.math.uni-tuebingen.de/arbeitsbereiche/gadr/personen/dr-martin-kell/vorlesungen-lectures/optimal-transport/skript-optimal-transport, 2017. [Aufgerufen am 9. Mai 18].

[Kle06] A. Klenke. *Wahrscheinlichkeitstheorie*. Masterclass. Springer Berlin Heidelberg, 2006.

[MF11] Petr Hájek Vicente Montesinos Václav Zizler (auth.) Marián Fabian, Petr Habala. *Banach Space Theory: The Basis for Linear and Nonlinear Analysis*. CMS Books in Mathematics. Springer-Verlag New York, 1 edition, 2011.

[Pau17] M. Pauly. Skript zur Vorlesung „Stochastik I“, Sommersemester 17.

[Sch12] M. Schelling. Der Dualraum von C(K). https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/gerlach/ss12/

`masssem/michael-schelling-arbeit.pdf`, 2012. [Aufgerufen am 12. August 18].

[vG03] Onno van Gaans. Probability measures on metric spaces. `https://www.math.leidenuniv.nl/~vangaans/jancol1.pdf`, 2003. [Aufgerufen am 12. September 18].

[Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.