Karlsruher Institut für Technologie

Maße auf polnischen Räumen und Wassersteinmetriken

Bachelorarbeit von

Daniel Herbst

an der Fakultät für Mathematik

Erstgutachter:	Prof. Dr. Roland Schnaubelt
Zweitgutachter:	Prof. Dr. Dorothee Frey
Betreuender Mitarbeiter:	Richard Nutt M.Sc.

Abgabetermin: 30. September 2021

Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde, sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, den 30. September 2021

. . . .
(Daniel Herbst)

Inhaltsverzeichnis

1 Einleitung

Fehlt noch.

2 Grundlagen

In diesem Abschnitt möchten wir einige Begriffe und Schreibweisen festlegen sowie ein paar wenige Grundlagen vorstellen, die wir im weiteren Verlauf der Arbeit benötigen werden. In Teilabschnitt 2.1 führen wir zunächst zwei grundlegende Konzepte aus der Topologie ein, während wir uns in Teilabschnitt 2.2 neben der Definition einiger Begriffe mit der *Regularität* bzw. *schwachen Regularität* eine grundlegende Approximationseigenschaft von Maßen ansehen werden.

2.1 Topologische Grundlagen

Netze

Im Allgemeinen beschreiben Folgen und deren Konvergenz die Eigenschaften eines topologischen Raums nur unzureichend: Beispielsweise ist Folgenstetigkeit nicht notwendigerweise äquivalent zu Stetigkeit und zwei unterschiedliche Topologien auf einer gegebenen Menge können generell dieselben konvergenten Folgen haben (siehe etwa [Sim15, Beispiel 2.6.1]). Um solche Probleme im Folgenden zu umgehen, führen wir das Konzept des Netzes aus der mengentheoretischen Topologie ein. Netze verallgemeinern Folgen und können in vielerlei Hinsicht die Eigenschaften eines topologischen Raums besser erfassen. In diesem Teilabschnitt möchten wir auf einige Grundlagen zu Netzen eingehen, wobei wir uns überwiegend an [Sim15, Kapitel 2.6] orientieren.

Definition 2.1 (Gerichtete Menge)**.** Sei I eine Menge. Eine Relation $\preceq$ auf I mit

(a) $\forall \iota \in I: \ \iota \preceq \iota$
(b) $\forall \iota, \kappa, \lambda \in I: \ \iota \preceq \kappa$ und $\kappa \preceq \lambda$ impliziert $\iota \preceq \lambda$
(c) $\forall \iota, \kappa \in I: \ \exists \lambda \in I: \ \iota \preceq \lambda$ und $\kappa \preceq \lambda$

heißt *gerichtet*. In diesem Fall nennen wir das Paar $(I, \preceq)$ auch eine *gerichtete Menge*.

Definition 2.2 (Netz)**.** Sei $\mathcal{X}$ eine Menge und $(I, \preceq)$ eine gerichtete Menge. Ein *Netz* in $\mathcal{X}$ ist dann eine Abbildung $x\colon I \to \mathcal{X}$. In diesem Fall schreiben wir auch $x = (x_\iota)_{\iota \in I}$.

Definition 2.3 (Konvergenz von Netzen)**.** Sei $\mathcal{X}$ ein topologischer Raum und sei $x = (x_\iota)_{\iota \in I}$ ein Netz in $\mathcal{X}$. Wir sagen, dass x gegen $z \in \mathcal{X}$ konvergiert, falls es für jede Umgebung U von z ein derartiges $\iota_0 \in I$ gibt, dass $x_\iota \in U$ für alle $\iota_0 \preceq \iota \in I$ erfüllt ist. In diesem Fall schreiben wir auch $\lim_{\iota \to \infty} x_\iota = z$ oder $x_\iota \to z, \ \iota \to \infty$.

Bemerkung. Offenbar sind Folgen einfach Netze, bei denen $(I, \preceq) := (\mathbb{N}, \leq)$ gewählt wird. Der Konvergenzbegriff von Netzen stimmt dann auch mit dem von uns bekannten Konvergenzbegriff von Folgen in topologischen Räumen überein.

Mit dem folgenden Satz konkretisieren wir unsere zu Beginn dieses Abschnitts getroffene Aussage, dass Netze die Eigenschaften topologischer Räume besser als Folgen erfassen: Tatsächlich lässt sich jede Topologie einzig und allein durch ihre konvergenten Netze charakterisieren.

Satz 2.4. *Seien $\mathcal{X}$ und $\mathcal{Y}$ topologische Räume. Dann gelten die folgenden Aussagen:*

(a) Eine Teilmenge $C \subseteq \mathcal{X}$ ist genau dann abgeschlossen, wenn der Grenzwert $z \in \mathcal{X}$ eines jeden konvergenten Netzes $x = (x_\iota)_{\iota \in I}$ in $\mathcal{X}$ bereits in C liegt.

(b) Eine Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ ist genau dann stetig, wenn für jedes konvergente Netz $x = (x_\iota)_{\iota\in I}$ in $\mathcal{X}$

$$\lim_{\iota\to\infty} f(x_\iota) \;=\; f(\lim_{\iota\to\infty} x_\iota)$$

erfüllt ist.

(c) Zwei Topologien auf $\mathcal{X}$ sind genau dann identisch, wenn alle Netze in $\mathcal{X}$ bezüglich beider Topologien dasselbe Konvergenzverhalten aufweisen.

Beweis. Siehe [Sim15, Satz 2.6.3]. □

Bemerkung. Auch wenn man in spezielleren Fällen (wie etwa bei metrisierbarem $\mathcal{X}$) problemlos mit Folgen arbeiten kann, sei an dieser Stelle noch einmal angemerkt, dass im Allgemeinen keine der drei Aussagen gelten, wenn man „Netz“ durch „Folge“ ersetzt (vgl. [Sim15, Beispiel 2.6.1]).

Für reelle Netze können wir in Analogie zu Folgen auch den Limes Superior bzw. Inferior von Netzen definieren. Hierfür orientieren wir uns an [Meg98, Kapitel 2.1 und Aufgabe 2.55] und [AB06, Kapitel 2.4].

Definition 2.5 (Limes Superior und Inferior von Netzen)**.** Sei $x = (x_\iota)_{\iota\in I}$ ein Netz in $\mathbb{R}$. Dann definieren wir

$$\limsup_{\iota\to\infty} x_\iota \;:=\; \inf_{\iota\in I}\sup_{\iota\preceq\kappa} x_\kappa \quad\text{bzw.}\quad \liminf_{\iota\to\infty} x_\iota \;:=\; \sup_{\iota\in I}\inf_{\iota\preceq\kappa} x_\kappa$$

und nennen diese den *Limes Superior* bzw. *Limes Inferior* von x.

Bemerkung. Offenbar ist auch diese Definition mit den bekannten Begriffen für reelle Folgen kompatibel. Die meisten grundlegenden Eigenschaften der Konvergenz und des Limes Superior bzw. Inferior von reellen Folgen lassen sich auch auf reelle Netze übertragen, wie etwa die üblichen Rechenregeln für die Grenzwerte von Summen und Produkten reeller Folgen.

Außerdem gelten etwa für reelle Netze $x = (x_\iota)_{\iota\in I}$ und $y = (y_\iota)_{\iota\in I}$ die Ungleichungen

$$\begin{aligned}\limsup_{\iota\to\infty}\,(x_\iota + y_\iota) \;&\leq\; \limsup_{\iota\to\infty} x_\iota + \limsup_{\iota\to\infty} y_\iota \quad\text{und}\\ \liminf_{\iota\to\infty}\,(x_\iota + y_\iota) \;&\geq\; \liminf_{\iota\to\infty} x_\iota + \liminf_{\iota\to\infty} y_\iota,\end{aligned}$$

sofern die Summen jeweils definiert sind und mit Gleichheit, falls eines der beiden Netze konvergiert (vgl. [Meg98, Aufgabe 2.55 (e)]). Nach [Meg98, Aufgabe 2.55 (i)] ist für ein konvergentes Netz $x = (x_\iota)_{\iota\in I}$ in $\mathbb{R}$ ebenfalls

$$\lim_{\iota\to\infty} x_\iota \;=\; \limsup_{\iota\to\infty} x_\iota \;=\; \liminf_{\iota\to\infty} x_\iota.$$

Initialtopologie

Wir möchten nun noch ein weiteres Konzept aus der Topologie, die sogenannte *Initialtopologie* einführen. Auch wenn es sich hier um einen sehr elementaren Begriff handelt, wird dieser uns an späterer Stelle gerade in Kombination mit Satz 2.7 häufig ermöglichen, gewisse Topologien, mit denen wir arbeiten, und deren Konvergenzverhalten in knapper Form auszudrücken. Der folgende Abschnitt orientiert sich an [AB06, Kapitel 2.13].

Definition 2.6 (Initialtopologie)**.** Sei $\mathcal{X}$ eine Menge, A eine beliebige Indexmenge und für jedes $\alpha \in A$ sei $\mathcal{Y}_\alpha$ ein topologischer Raum sowie $f_\alpha\colon \mathcal{X} \to \mathcal{Y}_\alpha$ eine Abbildung. Die Initialtopologie auf $\mathcal{X}$ bezüglich der Abbildungen $f_\alpha,\ \alpha \in A$ ist dann die kleinste Topologie auf $\mathcal{X}$, bezüglich der alle Abbildungen $f_\alpha,\ \alpha \in A$ stetig sind.

Bemerkung (Produkttopologie). Sei A eine beliebige Indexmenge und sei für jedes $\alpha \in A$ ein topologischer Raum $\mathcal{X}_\alpha$ gegeben. Die Produkttopologie der $\mathcal{X}_\alpha,\ \alpha \in A$ auf dem kartesischen Produkt $\mathcal{X} := \prod_{\alpha\in A}\mathcal{X}_\alpha$ ist dann die Initialtopologie der kanonischen Projektionen $\pi_\beta\colon \mathcal{X} \to \mathcal{X}_\beta,\ x = (x_\alpha)_{\alpha\in A} \mapsto x_\beta,\ \beta \in A$.

Satz 2.7 (Konvergenz bezüglich der Initialtopologie). *In der Situation von Definition 2.6 konvergiert ein Netz $x = (x_\iota)_{\iota \in I}$ in $\mathcal{X}$ genau dann gegen ein $z \in \mathcal{X}$, wenn für alle $\alpha \in A$ die Konvergenz $f_\alpha(x_\iota) \to f_\alpha(z),\ \iota \to \infty$ gilt.*

Beweis. Wegen Satz 2.4 (b) folgt die Hinrichtung direkt aus der Stetigkeit der Abbildungen $f_\alpha,\ \alpha \in A$.

Für die Rückrichtung möchten wir zeigen, dass für alle Umgebungen V von z ein derartiges $\iota_0 \in I$ existiert, dass $x_\iota \in V$ für $\iota_0 \preceq \iota$ gilt. Man bemerke zunächst, dass die Menge alle Urbilder von offenen Mengen in $\mathcal{Y}_\alpha$ unter f_α für alle $\alpha \in A$ eine Subbasis der Topologie von $\mathcal{X}$ ist. Daher genügt es, Mengen der Form

$$\tilde{V} \ := \ \bigcap_{i=1}^{n} f_{\alpha_i}^{-1}(U_i) \ \subseteq \ \mathcal{X}$$

zu betrachten, wobei $\alpha_1, \ldots, \alpha_n \in A$ sind und U_i offene Umgebungen von $f_{\alpha_i}(z)$, $i \in \{1, \ldots, n\}$. Sei nun ein solches $\tilde{V}$ gegeben. Dann können wir für $i \in \{1, \ldots, n\}$ jeweils ein $\iota_0^{(i)} \in I$ finden, sodass $x_\iota \in f_{\alpha_i}^{-1}(U_i)$ für $\iota_0^{(i)} \preceq \iota$ gilt. Sukzessive Anwendung von Eigenschaft (c) in Definition 2.1 liefert ein $\iota_0 \in I$ mit $x_\iota \in \tilde{V}$ für $\iota_0 \preceq \iota$, was schließlich die Konvergenz $x_\iota \to z,\ \iota \to \infty$ bedeutet. □

Bemerkung. Man kann sich die Frage stellen, warum wir bei der Definition eines Netzes $(x_\iota)_{\iota \in I}$ fordern, dass die Indexmenge I *gerichtet* ist und etwa nicht lediglich partiell geordnet. Am Ende des vorigen Beweises lässt sich allerdings der Mehrwert, den wir durch Eigenschaft (c) in der Definition einer gerichteten Menge erhalten, sehr gut erkennen: Wir können ein ‚Maximum' von endlich vielen $\iota \in I$ finden und damit ähnlich argumentieren, als würden wir gerade mit Folgen arbeiten.

2.2 Borel-Maße auf topologischen Räumen

Nun wenden wir uns (endlichen) Borel-Maßen auf topologischen bzw. metrischen Räumen zu. Da diese den primären Betrachtungsgegenstand der Arbeit darstellen, möchten wir zunächst einige Begriffe definieren, Schreibweisen festlegen, und – zumindest im Falle von metrischen Räumen – eine im Folgenden sehr hilfreiche Eigenschaft dieser Maße, *Schwache Regularität* bzw. *Regularität*, vorstellen. Der Teilabschnitt folgt [Sim15, Kapitel 4.14].

Definition 2.8 (Borel-σ-Algebra). Sei $\mathcal{X}$ ein topologischer Raum mit Topologie $\mathcal{O}$. Dann definieren wir die *Borel-σ-Algebra*

$$\mathcal{B}(\mathcal{X}) = \mathcal{B} := \sigma(\mathcal{O})$$

als die kleinste σ-Algebra auf $\mathcal{X}$, die $\mathcal{O}$ umfasst.

Definition 2.9 (Borel-Maße). In der Situation von Definition 2.8 nennen wir Maße auf $\mathcal{B}(\mathcal{X})$ *Borel-Maße*. Wir führen außerdem die Schreibweisen

$$\begin{aligned} \mathcal{M}_+(\mathcal{X}) \ &:= \ \{\, \mu \colon \mathcal{B}(\mathcal{X}) \to [0, \infty) \mid \mu \text{ ist endliches Maß} \,\} \\ \mathcal{P}(\mathcal{X}) \ &:= \ \{\, \mu \colon \mathcal{B}(\mathcal{X}) \to [0, 1] \mid \mu \text{ ist Wahrscheinlichkeitsmaß} \,\} \end{aligned}$$

für die Menge aller *endlichen Borel-Maße* beziehungsweise *Borel-Wahrscheinlichkeitsmaße* auf $\mathcal{X}$ ein.

Definition 2.10 (Stetige beschränkte Funktionen). Sei $\mathcal{X}$ ein topologischer Raum. Dann schreiben wir

$$C_b(\mathcal{X}) \ := \ \{\, f \colon \mathcal{X} \to \mathbb{R} \mid f \text{ ist stetig und beschränkt} \,\}$$

für die Menge aller stetigen und beschränkten Funktionen auf $\mathcal{X}$. Ist $\mathcal{X}$ ein metrischer Raum, so schreiben wir außerdem

$$\begin{aligned} U_b(\mathcal{X}) &:= \{\, f\colon \mathcal{X} \to \mathbb{R} \mid f \text{ ist gleichmäßig stetig und beschränkt} \,\} \quad \text{und} \\ \operatorname{Lip}_b(\mathcal{X}) &:= \{\, f\colon \mathcal{X} \to \mathbb{R} \mid f \text{ ist lipschitzstetig und beschränkt} \,\}. \end{aligned}$$

Bemerkung. Sei $\mathcal{X}$ ein metrischer Raum. Sicherlich ist dann $\operatorname{Lip}_b(\mathcal{X}) \subseteq U_b(\mathcal{X}) \subseteq C_b(\mathcal{X})$, allerdings hängen $\operatorname{Lip}_b(\mathcal{X})$ und $U_b(\mathcal{X})$ im Gegensatz zu $C_b(\mathcal{X})$ von der Metrik selbst und nicht ausschließlich von der induzierten Topologie ab.

Wie angekündigt werden wir uns im Folgenden mit der *(schwachen) Regularität* von endlichen Borel-Maßen auseinandersetzen. Hierbei handelt es sich um eine Approximationseigenschaft, die es uns etwa erlaubt, gewisse Aussagen zunächst für abgeschlossene bzw. kompakte und für offene Mengen (die gewissermaßen ‚strukturierter' als lediglich messbare Mengen sind) zu zeigen, um diese anschließend auf ganz $\mathcal{B}(\mathcal{X})$ auszuweiten.

Definition 2.11 (Schwache Regularität von Maßen)**.** Sei $\mathcal{X}$ ein topologischer Raum und μ ein endliches Borel-Maß auf $\mathcal{X}$. Dann nennen wir $B \in \mathcal{B}(\mathcal{X})$ *schwach von innen bzw. von außen regulär*, falls

$$\mu(B) = \sup_{\substack{C \subseteq B \\ C \text{ abgeschlossen}}} \mu(C) \quad \text{bzw.} \quad \mu(B) = \inf_{\substack{U \supseteq B \\ U \text{ offen}}} \mu(U)$$

gelten. Wir nennen $B \in \mathcal{B}(\mathcal{X})$ *schwach regulär*, wenn B schwach von innen und von außen regulär ist. Sind alle $B \in \mathcal{B}(\mathcal{X})$ schwach regulär, so nennen wir μ ein *schwach reguläres Maß.*

Bemerkung. Ersetzen wir in der obigen Definition „abgeschlossen" durch „kompakt", so erhalten wir analog den Begriff der *Regularität.* Außerdem ist eine abgeschlossene (bzw. offene) Menge trivialerweise schwach regulär von innen (bzw. außen).

Im Falle von endlichen Maßen auf metrischen Räumen kann schwache Regularität recht leicht gezeigt werden, wie wir im Folgenden sehen werden.

Satz 2.12. *Ist $(\mathcal{X}, d)$ ein metrischer Raum, so ist jedes endliche Borel-Maß μ auf $\mathcal{X}$ schwach regulär.*

Für den Beweis des Satzes benötigen wir noch zwei Hilfssätze. Der erste Hilfssatz behandelt die recht elementare Tatsache, dass sich in metrischen Räumen abgeschlossene Mengen von außen gewissermaßen ‚beliebig gut' durch offene Mengen approximieren lassen. Diese Erkenntnis werden wir im weiteren Verlauf der Arbeit noch häufiger benötigen. Der zweite Hilfssatz, den wir vorstellen, wird uns ermöglichen, die schwache Regularität nur auf einem Erzeuger von $\mathcal{B}(\mathcal{X})$ zu zeigen. Im endgültigen Beweis des Satzes werden wir dann die abgeschlossenen Mengen als Erzeuger wählen.

Hilfssatz 2.13. *Sei $(\mathcal{X}, d)$ ein metrischer Raum und $C \subseteq \mathcal{X}$ eine abgeschlossene Teilmenge. Ferner definieren wir für $n \in \mathbb{N}$*

$$A_n := \left\{ y \in \mathcal{X} \;\middle|\; d(y, C) < \frac{1}{n} \right\} \quad \textit{und} \quad f_n\colon \mathcal{X} \to \mathbb{R},\ x \mapsto \max(0, 1 - nd(x, C)),$$

wobei wir $d(y, C) := \inf_{x \in C} d(y, x)$ setzen. Dann gilt:

(a) A_n ist offen für alle $n \in \mathbb{N}$.
(b) $C = \bigcap_{n \in \mathbb{N}} A_n$, insbesondere ist C also eine G_δ-Menge.
(c) Für alle n ist $f_n|_{A_n^c} = 0$ und f_n ist lipschitzstetig.
(d) $f_n \downarrow \mathbb{1}_C$.

Beweis. Aussage (a) folgt aus der Stetigkeit von $y \mapsto d(y, C)$.

Weiter ist $C \subseteq A_n$ für alle $n \in \mathbb{N}$ und damit $C \subseteq \bigcap_{n\in\mathbb{N}} A_n$. Umgekehrt gibt es für ein beliebiges $y \in \bigcap_{n\in\mathbb{N}} A_n$ eine Folge $(x_n)_n \in C^{\mathbb{N}}$ mit $x_n \to y$. Wegen der Abgeschlossenheit von C liegt y damit in C, sodass (b) gezeigt ist.

Aussage (c) ist klar (f_n ist als Komposition lipschitzstetiger Funktionen selbst lipschitzstetig).

Schließlich fällt f_n und für $x \in C$ gilt $f_n(x) = 1$. Für $x \in C^{\mathsf{c}}$ ist $d(x, C) > 0$ und damit

$$f_n(x) = \max(0, 1 - nd(x, C)) \to 0, \quad n \to \infty,$$

womit auch Aussage (d) folgt. □

Hilfssatz 2.14. *In der Situation von Definition 2.11 ist*

$$\mathcal{S} := \{\, B \in \mathcal{B}(\mathcal{X}) \mid B \text{ schwach regulär bzgl. } \mu \,\}$$

eine σ-Algebra.

Beweis. Wir stellen eine Anpassung des Beweises von [Sim15, Hilfssatz 4.5.5] vor, wo die entsprechende Aussage mit „regulär“ anstelle „schwach regulär“ für kompakte $\mathcal{X}$ bewiesen wird.

Offensichtlich liegen $\emptyset$ und $\mathcal{X}$ in $\mathcal{S}$. Sei $B \in \mathcal{S}$ und damit schwach regulär von innen und von außen. Wegen $\mu(\mathcal{X}) < \infty$ gilt dann

$$\mu(B^{\mathsf{c}}) \;=\; \mu(\mathcal{X}) - \mu(B) \;=\; \mu(\mathcal{X}) - \sup_{\substack{C \subseteq B\\ C^{\mathsf{c}} \in \mathcal{O}}} \mu(C) \;=\; \inf_{\substack{C \subseteq B\\ C^{\mathsf{c}} \in \mathcal{O}}} (\mu(\mathcal{X}) - \mu(C)) \;=\; \inf_{\substack{U \supseteq B^{\mathsf{c}}\\ U \in \mathcal{O}}} \mu(U)$$

sowie analog

$$\mu(B^{\mathsf{c}}) \;=\; \mu(\mathcal{X}) - \mu(B) \;=\; \mu(\mathcal{X}) - \inf_{\substack{U \supseteq B\\ U \in \mathcal{O}}} \mu(C) \;=\; \sup_{\substack{U \supseteq B\\ U \in \mathcal{O}}} (\mu(\mathcal{X}) - \mu(U)) \;=\; \sup_{\substack{C \subseteq B^{\mathsf{c}}\\ C^{\mathsf{c}} \in \mathcal{O}}} \mu(U).$$

Also ist auch B^{c} schwach regulär.

Es bleibt nun zu zeigen, dass für $(B_n)_n \in \mathcal{S}^{\mathbb{N}}$ auch $B := \bigcup_{n\in\mathbb{N}} B_n$ schwach regulär ist. Hierfür beweisen wir zunächst die schwache Regularität von innen. Sei dazu $\varepsilon > 0$. Für $n \in \mathbb{N}$ gibt es jeweils abgeschlossene Mengen $C_n \subseteq B_n$ mit $\mu(B_n) - \mu(C_n) < \frac{\varepsilon}{3^n}$. Wir wählen nun N so groß, dass $\mu\left(B \setminus \bigcup_{n=1}^N B_n\right) < \frac{\varepsilon}{2}$ ist (was wegen $\mu(B) < \infty$ immer geht). Für die abgeschlossene Menge $C := \bigcup_{n=1}^N C_k$ gilt dann die Ungleichung

$$\begin{aligned}
\mu(B) - \mu(C) = \mu(B \setminus C) \;&=\; \mu\left(\left(B \setminus \bigcup_{n=1}^N B_n\right) \cup \left(\bigcup_{n=1}^N B_n \setminus C\right)\right)\\
&\le\; \mu\left(B \setminus \bigcup_{n=1}^N B_n\right) + \sum_{n=1}^N \mu(B_n \setminus C_n)\\
&<\; \frac{\varepsilon}{2} + \sum_{n=1}^\infty \frac{\varepsilon}{3^n} \;=\; \varepsilon,
\end{aligned}$$

also ist B schwach regulär von innen.

Ferner existieren für alle n offene Mengen $U_n \supseteq B_n$ mit $\mu(U_n) - \mu(B_n) < \frac{\varepsilon}{2^n}$. Wir setzen $U := \bigcup_{n\in\mathbb{N}} U_n$ und berechnen

$$\mu(U) - \mu(B) \;\le\; \sum_{n=1}^\infty \mu(U_n \setminus B) \;\le\; \sum_{n=1}^\infty \mu(U_n \setminus B_n) \;<\; \varepsilon.$$

Weil U offen ist, folgt insgesamt die schwache Regularität von μ. □

Ausgestattet mit den Hilfssätzen 2.13 und 2.14 können wir jetzt zum Beweis von Satz 2.12 übergehen.

Beweis von Satz 2.12. Es ist nun zu zeigen, dass für jedes endliche Borel-Maß μ auf $\mathcal{X}$ die Menge

$$\mathcal{S} := \{\, B \in \mathcal{B}(\mathcal{X}) \mid B \text{ schwach regulär bzgl. } \mu \,\}$$

bereits ganz $\mathcal{B}(\mathcal{X})$ ist. Da $\mathcal{S}$ nach Hilfssatz 2.14 eine σ-Algebra ist und $\mathcal{B}(\mathcal{X})$ von den abgeschlossenen Mengen erzeugt wird, genügt es zu zeigen, dass diese in $\mathcal{S}$ enthalten sind.

Sei $C \in \mathcal{B}(\mathcal{X})$. Dann ist C sicherlich schwach regulär von innen. Nun verwenden wir die offenen Mengen A_n, $n \in \mathbb{N}$ aus Hilfssatz 2.13. Wegen $A_n \downarrow C$ und $\mu(\mathcal{X}) < \infty$ folgt mit der Maßstetigkeit von oben die Konvergenz $\mu(A_n) \downarrow \mu(C)$, sodass C auch schwach regulär von außen ist. □

Bemerkung. Handelt es sich in der Situation von Satz 2.12 bei $\mathcal{X}$ um einen polnischen Raum, so werden wir zu einem späteren Zeitpunkt (in Folgerung 5.11) noch sehen, dass hier tatsächlich sogar Regularität gilt.

Schließlich möchten wir noch kurz eine direkte Folgerung aus Satz 2.12 vorstellen, die etwas schwächere hinreichende Bedingungen für die Gleichheit zweier endlicher Maße auf metrischen Räumen bereitstellt. Im weiteren Verlauf wird sich dies noch als nützlich erweisen.

Folgerung 2.15. *Sei $(\mathcal{X}, d)$ ein metrischer Raum und seien $\mu, \nu \in \mathcal{M}_+(\mathcal{X})$. Dann sind die folgenden Aussagen äquivalent:*

(i) $\mu = \nu$.
(ii) Für alle $f \in C_b(\mathcal{X})$ *ist* $\int f \,\mathrm{d}\mu = \int f \,\mathrm{d}\nu$.
(iii) Für alle $f \in U_b(\mathcal{X})$ *ist* $\int f \,\mathrm{d}\mu = \int f \,\mathrm{d}\nu$.
(iv) Für alle $f \in \mathrm{Lip}_b(\mathcal{X})$ *ist* $\int f \,\mathrm{d}\mu = \int f \,\mathrm{d}\nu$.
(v) Für alle abgeschlossenen Mengen $C \subseteq \mathcal{X}$ *ist* $\mu(C) = \nu(C)$.

Beweis. Die Implikationen (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) sind klar und (v) ⇒ (i) folgt aus Satz 2.12.

(iv) ⇒ (v): Gelte (iv) und sei $C \subseteq \mathcal{X}$ abgeschlossen. Dann gilt für die Funktionen $f_n \in U_b(\mathcal{X})$ aus Hilfssatz 2.13

$$\int f_n \,\mathrm{d}\mu = \int f_n \,\mathrm{d}\nu, \quad n \in \mathbb{N}.$$

Wegen $|f_n| \leq 1$ und $f_n \downarrow \mathbb{1}_C$ folgt mit dem Satz von Lebesgue

$$\int f_n \,\mathrm{d}\mu \;\to\; \mu(C) \quad \text{und} \quad \int f_n \,\mathrm{d}\nu \;\to\; \nu(C)$$

und damit gilt (v). □

3 Polnische Räume

Eine Klasse topologischer Räume wird in den späteren Abschnitten dieser Arbeit eine zentrale Rolle spielen: die der polnischen Räume.

Definition 3.1 (Polnischer Raum)**.** Ein *polnischer Raum* ist ein separabler und vollständig metrisierbarer topologischer Raum.

Bemerkung. Es sei hier anzumerken, dass Polnizität tatsächlich eine rein topologische Eigenschaft ist. Wir fordern für einen polnischen Raum $\mathcal{X}$ nämlich nur die Existenz einer Metrik, die die Topologie von $\mathcal{X}$ erzeugt und bezüglich der $\mathcal{X}$ vollständig ist, möchten uns aber die Flexibilität bewahren, diese Metrik nicht zu fixieren und bei Bedarf zwischen verschiedenen solchen Metriken zu wechseln. Es kann durchaus Metriken geben, die $\mathcal{X}$ zwar metrisieren, jedoch nicht vollständig. Beispielsweise ist $\mathbb{R}$ ausgestattet mit der Standardtopologie offensichtlich ein polnischer Raum, denn $\mathbb{R}$ ist separabel und die euklidische Metrik metrisiert $\mathbb{R}$ vollständig. Allerdings ist $(0,1)$ als Teilraum von $\mathbb{R}$ mit der euklidischen Metrik zwar separabel, aber nicht vollständig. Da $(0,1)$ homöomorph zu $\mathbb{R}$ ist, ist $(0,1)$ aber dennoch ein polnischer Raum.

Polnische Räume sind insofern interessant, dass es sich hierbei um eine verhältnismäßig allgemeine und elementare Klasse topologischer Räume handelt, die – wie wir in diesem Abschnitt sehen werden – etwa unter Bildung von G_δ-Teilmengen oder abzählbaren Produkten abgeschlossen ist und mit der sich daher sehr leicht umgehen lässt. Dennoch erlaubt es die Klasse der polnischen Räume in vielen Bereichen, nützliche und recht allgemeine Resultate zu erhalten sowie Pathologien zu vermeiden. Konkret werden wir insbesondere in Abschnitt 5 Beispiele hierfür sehen. Dieser Abschnitt orientiert sich überwiegend an den relevanten Teilen von [Sim15, Kapitel 4.14].

3.1 Der Satz von Alexandroff

Zunächst widmen wir uns einer grundlegenden Abgeschlossenheitseigenschaft von polnischen Räumen.

Satz 3.2 (Alexandroff)**.** *Sei $\mathcal{X}$ ein polnischer Raum. Dann ist $A \subseteq \mathcal{X}$ genau dann selbst ein polnischer Raum bezüglich der Teilraumtopologie, wenn $A \subseteq \mathcal{X}$ eine G_δ-Teilmenge ist.*

Beweis. Der Beweis der Hinrichtung folgt [Jor14, Satz 7], während der Beweis der Rückrichtung eine Anpassung des Beweises von Satz 4.14.6 aus [Sim15, Kapitel 4.14] ist.

Für die Hinrichtung sei $A \subseteq \mathcal{X}$ eine G_δ-Teilmenge und seien $U_n \subseteq \mathcal{X}$, $n \in \mathbb{N}$ offene Mengen mit $A = \bigcap_{n\in\mathbb{N}} U_n$, wobei wir außerdem $C_n := U_n^{\mathsf{c}}$ schreiben. Ferner sei d eine Metrik, die $\mathcal{X}$ vollständig metrisiert. Da A als Teilmenge eines separablen metrischen Raums selbst separabel ist, genügt es, die vollständige Metrisierbarkeit von A zu beweisen.

Für $x, y \in A$ definieren wir

$$\tilde{d}(x,y) \ := \ d(x,y) + \sum_{n=1}^{\infty} \min\left(\frac{1}{2^n}, \left|\frac{1}{d(x,C_n)} - \frac{1}{d(y,C_n)}\right|\right).$$

Offensichtlich ist $\tilde{d}$ eine Metrik auf A, die dieselbe Topologie wie d erzeugt.

Wir möchten nun zeigen, dass $(A, \tilde{d})$ vollständig ist. Sei dazu $(x_k)_k \in A^{\mathbb{N}}$ eine Cauchyfolge. Wegen $d \le \tilde{d}$ ist $(x_k)_k$ dann auch bezüglich d eine Cauchyfolge und die Vollständigkeit von $(\mathcal{X}, d)$ liefert die

Existenz eines Grenzwertes $x \in \mathcal{X}$. Angenommen $x \notin A$, so gibt es ein $m \in \mathbb{N}$ mit $x \in C_m$. Damit ist aber für alle $k \in \mathbb{N}$

$$\sup_{l \geq k} \tilde{d}(x_k, x_l) \;\geq\; \sup_{l \geq k} \sum_{n=1}^{\infty} \min\left(\frac{1}{2^n}, \left|\frac{1}{d(x_k, C_n)} - \frac{1}{d(x_l, C_n)}\right|\right) \;\geq\; \frac{1}{2^m} \;>\; 0,$$

was der Tatsache widerspricht, dass $(x_k)_k$ eine Cauchyfolge bezüglich $\tilde{d}$ ist. Also muss $x \in A$ gelten. Aus der Stetigkeit von $x \mapsto d(x, C_n)$ folgt mit dem Satz von Lebesgue unmittelbar die Konvergenz $\tilde{d}(x_k, x) \to 0$ für $k \to \infty$, also ist $(A, \tilde{d})$ vollständig und A damit selbst ein polnischer Raum.

Nun beweisen wir noch die Rückrichtung. Sei dazu $A \subseteq \mathcal{X}$ eine derartige Teilmenge, dass A bezüglich der Teilraumtopologie selbst polnisch ist. $\tilde{d}$ metrisiere im Folgenden A vollständig. Sicherlich ist $B_{1/k}(x_m) \subseteq A$ für alle $k, m \in \mathbb{N}$ relativ offen, also gibt es jeweils offene Mengen $V_{k,m} \subseteq \mathcal{X}$ mit

$$B_{1/k}(x_m) \;=\; V_{k,m} \cap A. \tag{3.1}$$

Wir möchten jetzt zeigen, dass

$$A \;=\; \overline{A} \cap \bigcap_{k \in \mathbb{N}} \left(\bigcup_{m \in \mathbb{N}} V_{k,m}\right) \tag{3.2}$$

gilt. Weil $\overline{A}$ nach Hilfssatz 2.13 als abgeschlossene Menge eine G_δ-Menge ist, folgt aus (3.2) direkt die Aussage (b).

Für jedes $m \in \mathbb{N}$ ist sicherlich $A \subseteq \bigcup_{m \in \mathbb{N}} V_{k,m}$, weshalb „$\subseteq$“ in (3.2) unmittelbar folgt. Wir nehmen nun an, dass z in der rechten Seite von (3.2) liegt. Dann gibt es für jedes $k \in \mathbb{N}$ ein $m_k \in \mathbb{N}$ mit $z \in V_{k,m_k}$. Ferner existiert wegen $z \in \overline{A}$ eine Folge $(y_k)_k \in A^{\mathbb{N}}$ mit $y_k \to z$, wobei wir zusätzlich $y_k \in \bigcap_{j=1}^{k} V_{j,m_j}$ fordern. Weil damit nach (3.1) auch für alle $k \in \mathbb{N}$

$$y_k \in \bigcap_{j=1}^{k} B_{1/k}(x_m)$$

gilt, ist $(y_k)_k$ eine Cauchyfolge bezüglich $\tilde{d}$. Aufgrund der Vollständigkeit von $(A, \tilde{d})$ gibt es also ein $y \in A$ mit $y_k \to y$, was $z = y \in A$ und damit „$\supseteq$“ in (3.2) impliziert. $\square$

3.2 Universalitätseigenschaft des Hilbertwürfels

Mit dem *Hilbertwürfel* möchten wir uns nun einen speziellen polnischen Raum ansehen, der sich gleich in gewisser Weise als universell für alle polnischen Räume erweisen wird.

Definition 3.3 (Hilbertwürfel)**.** Der *Hilbertwürfel* ist der topologische Raum $H = [0,1]^{\mathbb{N}}$ ausgestattet mit der Produkttopologie, also der Initialtopologie bezüglich aller Projektionen $\pi_n \colon H \to [0,1],\ x \mapsto x_n,\ n \in \mathbb{N}$.

Einige fundamentale Eigenschaften des Hilbertwürfels sind im folgenden Hilfssatz zusammengefasst.

Hilfssatz 3.4. *Für den Hilbertwürfel H gelten folgende Aussagen:*

(a) Eine Folge $(x^{(k)})_k \in H^{\mathbb{N}}$ konvergiert genau dann gegen ein $x \in H$, wenn alle Komponenten konvergieren, also, wenn $\lim_{n \to \infty} x_n^{(k)} = x_n$ für alle $n \in \mathbb{N}$ ist.

(b) Setzen wir für $x, y \in H$

$$\rho(x, y) := \max_{n \in \mathbb{N}} \frac{|x_n - y_n|}{2^n},$$

so definiert ρ eine Metrik, die H metrisiert.

(c) H ist kompakt.

Beweis. Aussage (a) folgt direkt aus Satz 2.7.

Für den Beweis von (b) bemerke man zunächst, dass Mengen der Form

$$U = \prod_{n=1}^{\infty} U_n, \quad U_n \subseteq [0,1] \text{ offen}, \qquad U_n = [0,1] \text{ für fast alle } n \tag{3.3}$$

eine Basis der Topologie von H bilden.

Offenbar wird durch ρ aus (b) eine Metrik auf H definiert. Es genügt also zu zeigen, dass ρ die Topologie von H induziert. Für $x \in H$ und $r > 0$ ist $B_r(x) = \prod_{n=1}^{\infty} B_{2^n r}(x_n) \subseteq H$ offen. Ist nun $U = \prod_{n=1}^{\infty} U_n$ wie in (3.3), so lässt sich leicht einsehen, dass es für jedes $x \in U$ ein $r > 0$ gibt mit $B_r(x) \subseteq U$, also ist U offen bezüglich der von ρ erzeugten Topologie und insgesamt wird H von ρ metrisiert.

Der aus der Topologie bekannte *Satz von Tychonoff* (vgl. [Sim15, Satz 2.7.1]) liefert unmittelbar die Kompaktheit von H. □

Bemerkung. Insbesondere folgt aus Hilfssatz 3.4, dass es sich bei H tatsächlich um einen polnischen Raum handelt, denn alle kompakten metrischen Räume sind insbesondere separabel und vollständig und damit auch polnisch.

H erfüllt nun die folgende Universalitätseigenschaft.

Satz 3.5 (Universalitätseigenschaft des Hilbertwürfels)**.** *Ein topologischer Raum $\mathcal{X}$ ist genau dann separabel und metrisierbar, wenn $\mathcal{X}$ homöomorph zu einer Teilmenge von H ist.*

Ferner ist $\mathcal{X}$ genau dann polnisch, wenn $\mathcal{X}$ homöomorph zu einer G_δ-Teilmenge von H ist.

Wir werden den Beweis von Satz 3.5 simultan für metrisierbare und separable bzw. für polnische Räume führen, benötigen hierfür aber zunächst noch einen Hilfssatz, der uns ermöglichen wird, einen beliebigen separablen und metrischen Raum in den Hilbertwürfel einzubetten.

Hilfssatz 3.6. *Sei $(\mathcal{X}, d)$ ein separabler metrischer Raum und sei $\mathcal{D} := \{\, x_n \mid n \in \mathbb{N} \,\} \subseteq \mathcal{X}$ eine abzählbare dichte Teilmenge. Außerdem setzen wir*

$$\varphi \colon \mathcal{X} \to H,\ x \mapsto (\min(1, d(x, x_n)))_n\,. \tag{3.4}$$

Dann ist φ ein Homöomorphismus zwischen $\mathcal{X}$ und $\varphi(\mathcal{X})$.

Beweis. Wir zeigen zunächst, dass φ injektiv ist. Seien dazu $y, z \in H$ mit $\varphi(y) = \varphi(z)$, also gilt

$$\min(1, d(y, x_n)) = \min(1, d(z, x_n)) \tag{3.5}$$

für alle $n \in \mathbb{N}$. Weil es eine Folge $(y_n)_n \in \mathcal{D}^{\mathbb{N}}$ mit $y_n \to y$ gibt und diese wegen (3.5) auch gegen z konvergiert, folgt die Gleichheit von y und z.

Die Stetigkeit von φ folgt direkt, weil die Komponentenfunktionen $x \mapsto \min(1, d(x, x_n))$ jeweils stetig sind. Seien nun $(y_k)_k \in \mathcal{X}^{\mathbb{N}}$ und $y \in \mathcal{X}$ mit $\varphi(y_k) \to \varphi(y)$, also

$$\min(1, d(y_k, x_n)) \ \to\ \min(1, d(y, x_n)), \quad k \to \infty \tag{3.6}$$

für alle $n \in \mathbb{N}$. Wählt man nun $\varepsilon < 1$ beliebig, so existiert ein $n \in \mathbb{N}$ mit $d(y, x_n) \le \varepsilon$. Wegen (3.6) ist dann auch $d(y_k, x_n) \to d(y, x_n),\ k \to \infty$. Es gilt also die Ungleichung

$$\limsup_{k\to\infty} d(y_k, y) \ \le\ \limsup_{k\to\infty} d(y_k, x_n) + d(y, x_n) \ \le\ 2\varepsilon$$

und mit $\varepsilon \to 0$ folgt die Stetigkeit von φ^{-1}, womit wir nachgewiesen haben, dass φ ein Homöomorphismus auf sein Bild ist. □

Mit Hilfe von Hilfssatz 3.6 können nun die Aussagen von Satz 3.5 nachgewiesen werden.

Beweis von Satz 3.5. Da wir in Hilfssatz 3.6 bereits gesehen haben, dass jeder separable metrisierbare Raum homöomorph zu einer Teilmenge des Hilbertwürfels H ist und jede Teilmenge von H separabel und metrisierbar ist, folgt der erste Teil von Satz 3.5.

Sei $\mathcal{X}$ nun ein polnischer Raum und fixiere eine Metrik d, die $\mathcal{X}$ vollständig metrisiert. Nach Hilfssatz 3.6 ist $\mathcal{X}$ via $\varphi\colon \mathcal{X} \to H$ aus (3.6) homöomorph zu einer Teilmenge von H und da H selbst polnisch ist, muss $\varphi(\mathcal{X}) \subseteq H$ nach Satz 3.2 eine G_δ-Teilmenge sein. Ferner folgt aus demselben Satz, dass jede G_δ-Teilmenge von H polnisch ist. □

Insbesondere lässt sich unter Verwendung von Satz 3.5 recht mühelos eine weitere Abgeschlossenheitseigenschaft polnischer Räume beweisen.

Folgerung 3.7. *Sei $\mathcal{X}_n$ für jedes $n \in \mathbb{N}$ ein topologischer Raum. Dann gelten folgende Implikationen:*

(a) Ist $\mathcal{X}_n$ separabel und metrisierbar für jedes $n \in \mathbb{N}$, so ist auch $\prod_{n=1}^\infty \mathcal{X}_n$ mit der Produkttopologie separabel und metrisierbar.

(b) Ist $\mathcal{X}_n$ polnisch für jedes $n \in \mathbb{N}$, so ist auch $\prod_{n=1}^\infty \mathcal{X}_n$ mit der Produkttopologie polnisch.

Beweis. Wir zeigen zunächst Aussage (a). Seien alle $\mathcal{X}_n$ separabel und metrisierbar. Nach Satz 3.5 bzw. Hilfssatz 3.6 existieren dann Einbettungen $\varphi_n\colon \mathcal{X}_n \to H$, $n \in \mathbb{N}$. Nun ist aber

$$\prod_{n=1}^\infty \mathcal{X}_n \;\cong\; \prod_{n=1}^\infty \varphi_n(\mathcal{X}_n) \;\subseteq\; \prod_{n=1}^\infty H \;\cong\; H$$

und Satz 3.5 liefert, dass $\prod_{n=1}^\infty \mathcal{X}_n$ separabel und metrisierbar ist.

Für Teilaufgabe (b) setzen wir zusätzlich die Polnizität aller $\mathcal{X}_n$ voraus. Dann ist $\varphi_n(\mathcal{X}_n)$ für jedes $n \in \mathbb{N}$ eine G_δ-Teilmenge von H, also existieren für alle $k, n \in \mathbb{N}$ derartige offene Mengen $U_n^{(k)} \subseteq H$, dass $\varphi_n(\mathcal{X}_n) = \bigcap_{k\in\mathbb{N}} U_n^{(k)}$ ist. Da nun insbesondere auch

$$V_n^{(k)} \;:=\; \left(\prod_{m=1}^{n-1} H \times U_n^{(k)} \times \prod_{m=n+1}^{\infty} H\right) \;\subseteq\; \prod_{n=1}^\infty H$$

für alle $k, n \in \mathbb{N}$ offen ist, folgt

$$\prod_{n=1}^\infty \mathcal{X}_n \;\cong\; \prod_{n=1}^\infty \bigcap_{k\in\mathbb{N}} U_n^{(k)} \;=\; \bigcap_{n\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} V_n^{(k)} \;\subseteq\; \prod_{n=1}^\infty H \;\cong\; H,$$

also ist $\prod_{n=1}^\infty \mathcal{X}_n$ homöomorph zu einer G_δ-Teilmenge von H und damit nach Satz 3.5 polnisch. □

Bemerkung. Die Aussagen aus Folgerung 3.7 gelten auch für endliche Produkte, denn für alle übrigen $n \in \mathbb{N}$ können wir $\mathcal{X}_n$ in diesem Fall einfach als einen einzigen Punkt wählen.

4 Schwache Topologie und schwache Konvergenz

In diesem Kapitel werden wir endliche Borel-Maße auf topologischen Räumen selbst mit einer Topologie bzw. einem Konvergenzbegriff versehen, dem der sogenannten *schwachen Konvergenz*. Diese verallgemeinert die aus der Stochastik bekannte *Verteilungskonvergenz* auf allgemeinere Räume als $\mathbb{R}$.

Nachdem in Teilabschnitt 4.1 die schwache Topologie und schwache Konvergenz zunächst definiert werden, präsentieren wir in Teilabschnitt 4.2 eine alternative funktionalanalytische Sichtweise auf diese und charakterisieren die schwache Konvergenz in Satz 4.3 mit dem *Portmanteau-Theorem*, sofern der Grundraum metrisierbar ist.

Die grundlegenden Definitionen und Bemerkungen in den Abschnitten 4.1 und 4.2 folgen [Vee58, Abschnitte 1 bis 3], während wir uns beim Beweis von Satz 4.4 in Abschnitt 4.3 an [Sim15, Satz 4.14.4] orientieren.

4.1 Definitionen

Definition 4.1 (Schwache Topologie)**.** Sei $\mathcal{X}$ ein topologischer Raum. Dann ist die *schwache Topologie* auf $\mathcal{M}_+(\mathcal{X})$ definiert als die Initialtopologie bezüglich der Abbildungen

$$I_f \colon \mathcal{M}_+(\mathcal{X}) \to \mathbb{R},\ \mu \mapsto \int f \,\mathrm{d}\mu, \quad f \in C_b(\mathcal{X}).$$

Falls nicht explizit angegeben, sei $\mathcal{M}_+(\mathcal{X})$ im Folgenden immer mit der schwachen Topologie versehen.

Definition 4.2 (Schwache Konvergenz von Maßen)**.** In der Situation von Definition 4.1 heißt ein Netz $(\mu_\iota)_{\iota \in I}$ von Maßen in $\mathcal{M}_+(\mathcal{X})$ genau dann *schwach konvergent* gegen ein Maß $\mu \in \mathcal{M}_+(\mathcal{X})$, wenn $\lim_{\iota\to\infty} \mu_\iota = \mu$ bezüglich der schwachen Topologie gilt. In diesem Fall schreiben wir auch

$$\mu_\iota \xrightarrow{w} \mu.$$

Bemerkung. Aus Satz 2.7 folgt direkt, dass ein Netz $(\mu_\iota)_{\iota \in I}$ in $\mathcal{M}_+(\mathcal{X})$ genau dann schwach gegen $\mu \in \mathcal{M}_+(\mathcal{X})$ konvergiert, wenn

$$\lim_{\iota\to\infty} \int f \,\mathrm{d}\mu_\iota = \int f \,\mathrm{d}\mu$$

für alle $f \in C_b(\mathcal{X})$ gilt.

Häufig betrachten wir nur Teilräume von $\mathcal{M}_+(\mathcal{X})$ wie etwa $\mathcal{P}(\mathcal{X})$. Mit der schwachen Topologie auf solchen Teilräumen bezeichnen wir dann einfach die entsprechende Teilraumtopologie.

4.2 Beziehung zur Schwach-*-Konvergenz

Angesichts des Namens der *schwachen Konvergenz* bzw. *schwachen Topologie* stellt sich unmittelbar die Frage, ob diese in irgendeiner Beziehung zur schwachen Konvergenz bzw. schwachen Topologie in Banachräumen aus der Funktionalanalysis steht. Tatsächlich handelt es sich eher um eine Art von Schwach-$*$-Konvergenz, wie wir im Folgenden sehen werden.

Sei dazu zunächst $\mathcal{X}$ ein topologischer Raum. Dann ist $(C_b(\mathcal{X}), \|\cdot\|_\infty)$ ein normierter Vektorraum und wir betrachten nun die Abbildung

$$\Phi\colon \mathcal{M}_+(\mathcal{X}) \to C_b(\mathcal{X})^*,\ \mu \mapsto \left[l_\mu\colon C_b(\mathcal{X}) \to \mathbb{R},\ f \mapsto \int f \,\mathrm{d}\mu\right], \tag{4.1}$$

die wegen $\left|\int f \,\mathrm{d}\mu\right| \leq \mu(\mathcal{X}) \|f\|_\infty$ und $\mu(\mathcal{X}) < \infty$ wohldefiniert ist und einem endlichen Maß $\mu \in \mathcal{M}_+(\mathcal{X})$ ein zugehöriges lineares Funktional aus dem Dualraum $C_b(\mathcal{X})^*$ von $C_b(\mathcal{X})$ zuordnet. Ist $C_b(\mathcal{X})^*$ mit der Schwach-$*$-Topologie ausgestattet, so entspricht schwache Konvergenz in $\mathcal{M}_+(\mathcal{X})$ trivialerweise der Schwach-$*$-Konvergenz der jeweiligen Bilder unter Φ.

Ist $\mathcal{X}$ etwa metrisierbar, so können wir sogar noch eine stärkere Aussage formulieren:

Hilfssatz 4.3. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Dann ist die Abbildung Φ aus* (4.1) *ein Homöomorphismus auf sein Bild.*

Beweis. Φ ist injektiv, denn für $\mu, \nu \in \mathcal{M}_+(\mathcal{X})$ folgt aus $l_\mu = l_\nu$ wegen Folgerung 2.15 die Gleichheit $\mu = \nu$. Da schwache Konvergenz in $\mathcal{M}_+(\mathcal{X})$ exakt der Schwach-$*$-Konvergenz der jeweiligen Bilder unter Φ entspricht, liefert Satz 2.4 die Behauptung. ☐

Unter obigen Voraussetzungen lässt sich $\mathcal{M}_+(\mathcal{X})$ also in $C_b(\mathcal{X})^*$ einbetten. Mittels dieser Identifikation können nun Aussagen aus der Funktionalanalysis auch für $\mathcal{M}_+(\mathcal{X})$ fruchtbar gemacht werden, indem von topologischen Eigenschaften von $C_b(\mathcal{X})^*$ mit der Schwach-$*$-Konvergenz auf die entsprechenden Eigenschaften von $\mathcal{M}_+(\mathcal{X})$ geschlossen wird. Konkrete Anwendungen hiervon werden wir bereits in Abschnitt 5 sehen.

4.3 Charakterisierungen schwacher Konvergenz

Im Falle metrisierbarer Grundräume liefert der folgende Satz eine umfassende Charakterisierung der schwachen Konvergenz.

Satz 4.4 (Portmanteau)**.** *Sei $(\mathcal{X}, d)$ ein metrischer Raum, $(\mu_\iota)_{\iota\in I}$ ein Netz in $\mathcal{M}_+(\mathcal{X})$ und $\mu \in \mathcal{M}_+(\mathcal{X})$. Dann sind die folgenden Aussagen äquivalent:*

(i) $\mu_\iota \xrightarrow{w} \mu$.

(ii) Für alle $f \in U_b(\mathcal{X})$ gilt $\lim_{\iota\to\infty} \int f \,\mathrm{d}\mu_\iota = \int f \,\mathrm{d}\mu$.

(iii) Für alle $f \in \mathrm{Lip}_b(\mathcal{X})$ gilt $\lim_{\iota\to\infty} \int f \,\mathrm{d}\mu_\iota = \int f \,\mathrm{d}\mu$.

(iv) Es ist $\lim_{\iota\to\infty} \mu_\iota(\mathcal{X}) = \mu(\mathcal{X})$ und für alle abgeschlossenen Mengen $C \subseteq \mathcal{X}$ gilt

$$\limsup_{\iota\to\infty} \mu_\iota(C) \ \leq\ \mu(C).$$

(v) Es ist $\lim_{\iota\to\infty} \mu_\iota(\mathcal{X}) = \mu(\mathcal{X})$ und für alle offenen Mengen $U \subseteq \mathcal{X}$ gilt

$$\liminf_{\iota\to\infty} \mu_\iota(U) \ \geq\ \mu(U).$$

(vi) Für alle $B \in \mathcal{B}(\mathcal{X})$ mit $\mu(\partial B) = 0$ ist

$$\lim_{\iota\to\infty} \mu_\iota(A) \ =\ \mu(A).$$

Beweis. Wir präsentieren hier eine Verallgemeinerung des Beweises von Satz 4.14.4 in [Sim15, Kapitel 4.14] mit endlichen Maßen anstelle von Wahrscheinlichkeitsmaßen und Netzen anstelle von Folgen. Die Implikationen „(i) ⇒ (ii) ⇒ (iii)“ sind klar.

(iii) ⇒ (iv): Sei $C \subseteq \mathcal{X}$ abgeschlossen und seien $f_m,\ m \in \mathbb{N}$ die Funktionen aus Hilfssatz 2.13. Diese sind lipschitzstetig und beschränkt. Dann gilt für alle $m \in \mathbb{N}$

$$\mu_\iota(C) \ =\ \int \mathbb{1}_C \,\mathrm{d}\mu_\iota \ \leq\ \int f_m \,\mathrm{d}\mu_\iota \ \to\ \int f_m \,\mathrm{d}\mu, \quad \iota \to \infty,$$

also

$$\limsup_{\iota\to\infty} \mu_\iota(C) \;\leq\; \int f_m \,\mathrm{d}\mu.$$

Wegen $f_m \downarrow \mathbb{1}_C$ und $|f_m| \leq 1$ liefert der Satz von Lebesgue die Konvergenz

$$\int f_m \,\mathrm{d}\mu \;\to\; \int \mathbb{1}_C \,\mathrm{d}\mu \;=\; \mu(C), \quad m \to \infty,$$

woraus

$$\limsup_{\iota\to\infty} \mu_\iota(C) \;\leq\; \mu(C)$$

folgt. Außerdem ist

$$\mu_\iota(\mathcal{X}) \;=\; \int \mathbb{1}_{\mathcal{X}} \,\mathrm{d}\mu_\iota \;\to\; \int \mathbb{1}_{\mathcal{X}} \,\mathrm{d}\mu \;=\; \mu(\mathcal{X}), \quad \iota \to \infty.$$

(iv) $\Leftrightarrow$ (v): Es gelte (iv). Sei U offen, also $C := U^{\mathrm{c}}$ abgeschlossen. Dann erhalten wir

$$\begin{aligned}\liminf_{\iota\to\infty} \mu_\iota(U) \;&=\; \liminf_{\iota\to\infty} (\mu_\iota(\mathcal{X}) - \mu_\iota(C)) \;=\; \mu(\mathcal{X}) - \limsup_{\iota\to\infty} \mu_\iota(C)\\ &\geq\; \mu(\mathcal{X}) - \mu(C) \;=\; \mu(U)\end{aligned}$$

und damit (v). Die andere Richtung zeigt man analog.

(iv), (v) $\Rightarrow$ (vi): Sei $A \in \mathcal{B}(\mathcal{X})$ mit $\mu(\partial A) = 0$. Wegen $A^\circ \subseteq A \subseteq \overline{A}$ und $\partial A = \overline{A} \setminus A^\circ$ gilt $\mu(A^\circ) = \mu(A) = \mu(\overline{A})$. Ferner liefern die Annahmen

$$\limsup_{\iota\to\infty} \mu_\iota(\overline{A}) \;\leq\; \mu(\overline{A}) \quad \text{und} \quad \liminf_{\iota\to\infty} \mu_\iota(A^\circ) \;\geq\; \mu(A^\circ).$$

Daraus folgt

$$\limsup_{\iota\to\infty} \mu_\iota(A) \;\leq\; \mu(A) \;\leq\; \liminf_{\iota\to\infty} \mu_\iota(A),$$

also insgesamt

$$\lim_{\iota\to\infty} \mu_\iota(A) \;=\; \mu(A).$$

(vi) $\Rightarrow$ (i): Sei $f \in C_b(\mathcal{X})$ und $a < b \in \mathbb{R}$ mit $a < f < b$. Die Menge

$$S := \{\, c \in (a,b) \mid \mu(\{f = c\}) > 0 \,\},$$

ist abzählbar, da $S_n := \left\{\, c \in (a,b) \;\middle|\; \mu(\{f = c\}) > \frac{1}{n} \,\right\}$ für jedes $n \in \mathbb{N}$ endlich ist und $S = \bigcup_{n\in\mathbb{N}} S_n$ gilt. Damit können wir für jedes $m \in \mathbb{N}$ Zahlen $c_j^{(m)} \notin S$ mit

$$a = c_0^{(m)} < \cdots < c_{2m}^{(m)} = b, \qquad c_{j+1}^{(m)} - c_j^{(m)} \leq \frac{b-a}{m} \tag{4.2}$$

finden. Wir setzen

$$A_j^{(m)} := \left\{ c_j^{(m)} < f \leq c_{j+1}^{(m)} \right\}, \qquad j \in \{0, \ldots, 2m-1\}.$$

Die Stetigkeit von f impliziert $\partial A_j^{(m)} \subseteq \left\{ f = c_j^{(m)} \right\} \cup \left\{ f = c_{j+1}^{(m)} \right\}$, woraus sich

$$\mu(\partial A_j^{(m)}) = 0, \qquad j \in \{0, \ldots, 2m-1\}$$

ergibt. Für $m \in \mathbb{N}$ schreiben wir

$$u_m := \sum_{j=0}^{2m-1} c_j^{(m)} \mathbb{1}_{A_j^{(m)}}$$

und Aussage (v) führt dann auf

$$\int u_m \,\mathrm{d}\mu_\iota \;=\; \sum_{j=0}^{2m-1} c_j^{(m)} \mu_\iota(A_j^{(m)}) \;\to\; \sum_{j=0}^{2m-1} c_j^{(m)} \mu(A_j^{(m)}) \;=\; \int u_m \,\mathrm{d}\mu, \quad \iota \to \infty. \tag{4.3}$$

Außerdem folgen aus (4.2) die Ungleichungen

$$\left| \int f \,\mathrm{d}\mu_\iota - \int u_m \,\mathrm{d}\mu_\iota \right| \;\leq\; \frac{b-a}{m}, \qquad \left| \int f \,\mathrm{d}\mu - \int u_m \,\mathrm{d}\mu \right| \;\leq\; \frac{b-a}{m} \tag{4.4}$$

für $\iota \in I$. Mit (4.4) gilt nun für alle $m \in \mathbb{N}$, $\iota \in I$

$$\begin{aligned} \left| \int f \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu_\iota \right| \;&\leq\; \left| \int f \,\mathrm{d}\mu - \int u_m \,\mathrm{d}\mu \right| + \left| \int u_m \,\mathrm{d}\mu - \int u_m \,\mathrm{d}\mu_\iota \right| \\ &\quad + \left| \int u_m \,\mathrm{d}\mu_\iota - \int f \,\mathrm{d}\mu_\iota \right| \\ &\leq\; 2 \cdot \frac{b-a}{m} + \left| \int u_m \,\mathrm{d}\mu_\iota - \int u_m \,\mathrm{d}\mu \right|. \end{aligned} \tag{4.5}$$

(4.3) und (4.5) liefern also letztendlich für alle m

$$\begin{aligned} \limsup_{\iota \to \infty} \left| \int f \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu_\iota \right| \;&\leq\; 2 \cdot \frac{b-a}{m} + \limsup_{\iota \to \infty} \left| \int u_m \,\mathrm{d}\mu_\iota - \int u_m \,\mathrm{d}\mu \right| \\ &=\; 2 \cdot \frac{b-a}{m} \;\to\; 0, \quad m \to \infty, \end{aligned}$$

was die schwache Konvergenz $\mu_\iota \xrightarrow{w} \mu$ impliziert. □

Bemerkung (Verteilungskonvergenz). Betrachte den Fall $\mathcal{X} := \mathbb{R}$. Sei nun $(\mathbb{P}_n)_n \in \mathcal{P}(\mathbb{R})^{\mathbb{N}}$, $\mathbb{P} \in \mathcal{P}(\mathbb{R})$ und seien $F_n \colon \mathbb{R} \to [0,1]$, $x \mapsto \mathbb{P}_n((-\infty, x])$, $F \colon \mathbb{R} \to [0,1]$, $x \mapsto \mathbb{P}((-\infty, x])$ die zugehörigen Verteilungsfunktionen. In der Stochastik spricht man davon, dass $(\mathbb{P}_n)_n$ in Verteilung gegen $\mathbb{P}$ konvergiert, falls

$$\lim_{n \to \infty} F_n(x) \;=\; F(x)$$

für alle $x \in \mathbb{R}$ gilt, in denen F stetig ist [Hen16, Definition 6.1].

Tatsächlich ist die Verteilungskonvergenz von $(\mathbb{P}_n)_n$ gegen $\mathbb{P}$ gerade äquivalent zur schwachen Konvergenz $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$. Dies lässt sich analog zur Äquivalenz (i) $\Leftrightarrow$ (vi) in Satz 4.4 beweisen, wobei für den Beweis der Rückrichtung passende Intervalle als $A_j^{(m)}$ verwendet werden müssen.

5 Eigenschaften der schwachen Topologie

Nachdem wir im vorigen Abschnitt die schwache Topologie definiert haben, möchten wir uns nun einige Eigenschaften dieser ansehen, wenn man weitere Bedingungen an den Grundraum $\mathcal{X}$ stellt. In Teilabschnitt 5.1 wird eine Möglichkeit vorgestellt, den Grundraum $\mathcal{X}$ in $\mathcal{M}_+(\mathcal{X})$ einzubetten und in Teilabschnitt 5.2 sehen wir, dass sich bestimmte topologische Eigenschaften des Grundraums $\mathcal{X}$ wie Metrisierbarkeit, Separabilität, Polnizität und Kompaktheit unter gewissen Voraussetzungen auf $\mathcal{M}_+(\mathcal{X})$ bzw. $\mathcal{P}(\mathcal{X})$ übertragen. In Teilabschnitt 5.3 findet sich mit dem *Satz von Prokhorov* eine sehr hilfreiche Charakterisierung der Kompaktheit von Teilmengen von $\mathcal{P}(\mathcal{X})$.

5.1 Einbettung des Grundraums

Sofern der Grundraum $\mathcal{X}$ metrisierbar ist, liefert der folgende Satz eine Möglichkeit, $\mathcal{X}$ in $\mathcal{M}_+(\mathcal{X})$ einzubetten. Wir orientieren uns dabei an [Vee58, Hilfssätze 3.2 und 3.3].

Satz 5.1. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Für $x \in \mathcal{X}$ schreiben wir $\delta_x \in \mathcal{P}(\mathcal{X})$ für das Dirac-Maß im Punkt x und definieren*

$$\delta \colon \mathcal{X} \to \mathcal{M}_+(\mathcal{X}),\ x \mapsto \delta_x.$$

Außerdem setzen wir $\Delta := \delta(\mathcal{X}) = \{\, \delta_x \mid x \in \mathcal{X} \,\}$. Dann gelten die folgenden Aussagen:

(a) δ ist ein Homöomorphismus zwischen $\mathcal{X}$ und Δ.
(b) $\Delta \subseteq \mathcal{M}_+(\mathcal{X})$ ist sequentiell abgeschlossen.

Beweis. Wir beweisen zunächst Aussage (a). Weil $\mathcal{X}$ metrisierbar ist und die einelementigen Mengen $\{x\}$, $x \in \mathcal{X}$ daher in $\mathcal{B}(\mathcal{X})$ enthalten sind, ist δ sicherlich injektiv. Wegen Satz 2.4 (b) genügt es also nachzuweisen, dass für jedes Netz $(x_\iota)_{\iota \in I}$ in $\mathcal{X}$ sowie $z \in \mathcal{X}$ die Äquivalenz

$$x_\iota \to z \qquad \Longleftrightarrow \qquad \delta_{x_\iota} \xrightarrow{w} \delta_z, \quad \iota \to \infty$$

gilt.

Sei $(x_\iota)_{\iota \in I}$ ein Netz in $\mathcal{X}$, das gegen $z \in \mathcal{X}$ konvergiert und sei $f \in C_b(\mathcal{X})$. Dann ist

$$\int f \, \mathrm{d}\delta_{x_\iota} \;=\; f(x_\iota) \;\to\; f(z) = \int f \, \mathrm{d}\delta_z, \quad \iota \to \infty,$$

also konvergiert δ_{x_ι} schwach gegen δ_z. Gelte umgekehrt $\delta_{x_\iota} \xrightarrow{w} \delta_z$. Wähle eine Metrik d, die $\mathcal{X}$ metrisiert und setze $f := \min(1, d(\cdot, z)) \in C_b(\mathcal{X})$. Dann ist

$$\min(1, d(x_\iota, z)) \;=\; \int f \, \mathrm{d}\delta_{x_\iota} \;\to\; \int f \, \mathrm{d}\delta_z = 0,$$

was $x_\iota \to z$ nach sich zieht.

Nun möchten wir Aussage (b) zeigen. Sei $(x_n)_n \in \mathcal{X}^{\mathbb{N}}$ und sei $\mu \in \mathcal{M}_+(\mathcal{X})$ mit $\delta_{x_n} \xrightarrow{w} \mu$, $n \to \infty$. Zunächst gilt sicherlich $\mu \in \mathcal{P}(\mathcal{X})$. Außerdem muss $(x_n)_n$ eine konvergente Teilfolge besitzen. Ansonsten ist nämlich $S := \{\, x_n \mid n \in \mathbb{N} \,\} \subseteq \mathcal{X}$ eine unendliche und abgeschlossene Teilmenge. Das Portmanteau-Theorem (Satz 4.4) liefert also für jede unendliche Teilmenge $\tilde{S} \subseteq S$ die Abschätzung

$$1 \;=\; \limsup_{n \to \infty} \delta_{x_n}(\tilde{S}) \;\leq\; \mu(\tilde{S}) \;\leq\; 1,$$

was einen Widerspruch darstellt. Also existiert eine Teilfolge $(x_{n_k})_k$ und $x \in \mathcal{X}$ mit $x_{n_k} \to x$, $k \to \infty$. Damit ist aber auch $\delta_{x_{n_k}} \xrightarrow{w} \delta_x$, weshalb mit Folgerung 2.15 die Gleichheit $\mu = \delta_x \in \Delta$ folgt. □

5.2 Beziehung zu Eigenschaften des Grundraums

Als Grundbedingung möchten wir in diesem Teilabschnitt immer die Metrisierbarkeit von $\mathcal{X}$ fordern. Unter dieser Voraussetzung werden die folgenden Äquivalenzen unsere zentralen Erkenntnisse sein:

- $\mathcal{M}_+(\mathcal{X})$ ist genau dann metrisierbar und separabel, wenn $\mathcal{X}$ separabel ist.
- $\mathcal{M}_+(\mathcal{X})$ ist genau dann polnisch, wenn $\mathcal{X}$ polnisch ist.
- $\mathcal{P}(\mathcal{X})$ ist genau dann metrisierbar und kompakt, wenn $\mathcal{X}$ kompakt ist.

Abgesehen von Satz 5.8, bei dem wir im Wesentlichen [van03] folgen, orientieren wir uns in diesem Abschnitt an [Vee58].

Wir wenden uns nun der ersten unserer drei Hauptaussagen zu.

Satz 5.2. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Dann ist $\mathcal{M}_+(\mathcal{X})$ genau dann metrisierbar und separabel, wenn $\mathcal{X}$ separabel ist.*

Für den Beweis benötigen wir zunächst noch ein paar Hilfssätze, die wir im Folgenden vorstellen möchten.

Hilfssatz 5.3. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Dann ist $(C_b(\mathcal{X}), \|\cdot\|_\infty)$ ein Banachraum. Ist $\mathcal{X}$ kompakt, so ist $C_b(\mathcal{X})$ separabel.*

Beweis. Sicherlich ist $(C_b(\mathcal{X}), \|\cdot\|_\infty)$ ein normierter Vektorraum. Siehe [Sim15, Satz 2.1.6] für die Vollständigkeit und [Sim15, Satz 2.3.7] für die Separabilität bei kompaktem $\mathcal{X}$. ☐

Hilfssatz 5.4. *Sei $(\mathcal{X}, d)$ ein totalbeschränkter metrischer Raum. Dann ist $U_b(\mathcal{X})$ ein separabler Banachraum.*

Beweis. Sei $\tilde{\mathcal{X}}$ die Vervollständigung von $\mathcal{X}$. Diese ist kompakt und damit ist $C_b(\tilde{\mathcal{X}})$ nach Hilfssatz 5.3 ein separabler Banachraum. Jedes $f \in U_b(\mathcal{X})$ besitzt eine eindeutige Fortsetzung $\tilde{f} \in C_b(\tilde{\mathcal{X}})$ mit $\|f\|_\infty = \|\tilde{f}\|_\infty$. Damit wird über die Zuordnung

$$E\colon U_b(\mathcal{X}) \to C_b(\tilde{\mathcal{X}}),\ f \mapsto \tilde{f}$$

eine isometrische Einbettung von $U_b(\mathcal{X})$ nach $C_b(\tilde{\mathcal{X}})$ definiert. Wegen $C_b(\tilde{\mathcal{X}}) = U_b(\tilde{\mathcal{X}})$ (vgl. [Sim15, Satz 2.4.1]) ist $g|_{\mathcal{X}} \in U_b(\mathcal{X})$ für alle $g \in C_b(\tilde{\mathcal{X}})$ und damit auch $g = E(g|_{\mathcal{X}})$. Also ist E surjektiv und damit eine Isometrie zwischen $U_b(\mathcal{X})$ und $C_b(\tilde{\mathcal{X}})$. Hilfssatz 5.3 liefert nun die Behauptung. ☐

Hilfssatz 5.5. *Sei $R := \mathbb{R}^{\mathbb{N}}$ ausgestattet mit der Produkttopologie von $\mathbb{R}$, also mit der Initialtopologie der Projektionen $\pi_k\colon R \to \mathbb{R},\ x = (x_n)_n \mapsto x_k,\ k \in \mathbb{N}$. Dann ist R polnisch.*

Beweis.

$$R \;\cong\; (0,1)^{\mathbb{N}} \;=\; \bigcap_{n\in\mathbb{N}} \left(\prod_{k=1}^{n} (0,1) \times \prod_{k=n+1}^{\infty} [0,1] \right)$$

ist homöomorph zu einer G_δ-Teilmenge des Hilbertwürfels H, worauf mit Satz 3.5 unmittelbar die Behauptung folgt. ☐

Ausgestattet mit den obigen Hilfssätzen können wir nun zum Beweis von Satz 5.2 übergehen. Im Wesentlichen werden wir eine Einbettung von $\mathcal{M}_+(\mathcal{X})$ nach R konstruieren, die uns direkt die Behauptung liefert.

Beweis von Satz 5.2. Für die Hinrichtung fixieren wir auf $\mathcal{X}$ eine Metrik d, die $\mathcal{X}$ metrisiert. Nach Hilfsssatz 3.6 existiert dann ein Homöomorphismus $\varphi\colon \mathcal{X} \to H$ von $\mathcal{X}$ auf eine Teilmenge $\varphi(\mathcal{X})$ des Hilbertwürfels H. Wir wählen nun die Metrik ρ aus Hilfssatz 3.4 auf H. Aufgrund der Kompaktheit von H ist (H, ρ) totalbeschränkt und damit auch jede Teilmenge von H. Also können wir über ρ eine Metrik $\tilde{d}$ auf $\mathcal{X}$ einführen, sodass $\mathcal{X}$ von $\tilde{d}$ metrisiert wird und $(\mathcal{X}, \tilde{d})$ totalbeschränkt ist. Im Folgenden sei auf $\mathcal{X}$ diese Metrik fixiert. Hilfssatz 5.4 liefert uns, dass $(U_b(\mathcal{X}), \|\cdot\|_\infty)$ ein separabler

Banachraum ist. Sei nun $\mathcal{D} := \{ f_n \mid n \in \mathbb{N} \} \subseteq U_b(\mathcal{X})$ eine abzählbare dichte Teilmenge, die wir so wählen, dass $f_1 = \mathbb{1}_{\mathcal{X}}$ ist.

Wir setzen

$$T\colon \mathcal{M}_+(\mathcal{X}) \to R,\ \mu \mapsto \left(\int f_n \,\mathrm{d}\mu \right)_n$$

und möchten nachweisen, dass T ein Homöomorphismus auf sein Bild ist. Da R nach Hilfssatz 5.5 separabel und metrisierbar ist, übertragen sich diese Eigenschaften dann auch auf $\mathcal{M}_+(\mathcal{X})$.

Wir zeigen zunächst die Injektivität. Seien dazu $\mu, \nu \in \mathcal{M}_+(\mathcal{X})$ mit $T(\mu) = T(\nu)$. Sei außerdem $g \in U_b(\mathcal{X})$. Dann gibt es eine Folge $(g_n)_n \in \mathcal{D}^{\mathbb{N}}$ mit $\|g - g_n\|_\infty \to 0,\ n \to \infty$, was auch

$$\int g_n \,\mathrm{d}\mu \to \int g \,\mathrm{d}\mu \quad \text{und} \quad \int g_n \,\mathrm{d}\nu \to \int g \,\mathrm{d}\nu, \quad n \to \infty$$

impliziert. Da aber $\int g_n \,\mathrm{d}\mu = \int g_n \,\mathrm{d}\nu$ für alle $n \in \mathbb{N}$ gilt, folgt auch $\int g \,\mathrm{d}\mu = \int g \,\mathrm{d}\nu$. Folgerung 2.15 liefert nun die Gleichheit $\mu = \nu$.

Um den Beweis der Hinrichtung abzuschließen, genügt es wegen Satz 2.4 zu beweisen, dass für jedes Netz $(\mu_\iota)_{\iota \in I}$ in $\mathcal{M}_+(\mathcal{X})$ und $\mu \in \mathcal{M}_+(\mathcal{X})$ die Äquivalenz

$$\mu_\iota \xrightarrow{w} \mu \quad \iff \quad T(\mu_\iota) \to T(\mu), \quad \iota \to \infty$$

gilt.

Gelte $\mu_\iota \xrightarrow{w} \mu$, was nach Satz 4.4 äquivalent dazu ist, dass $\lim_{\iota \to \infty} \int f \,\mathrm{d}\mu_\iota = \int f \,\mathrm{d}\mu$ für alle $f \in U_b(\mathcal{X})$ ist. Offenbar impliziert dies die Konvergenz $T(\mu_\iota) \to T(\mu),\ \iota \to \infty$.

Sei nun umgekehrt $T(\mu_\iota) \to T(\mu),\ \iota \to \infty$, was gleichbedeutend damit ist, dass

$$\lim_{\iota \to \infty} \int f_n \,\mathrm{d}\mu_\iota = \int f_n \,\mathrm{d}\mu \tag{5.1}$$

für alle $n \in \mathbb{N}$ gilt. Wegen $f_1 = \mathbb{1}_{\mathcal{X}}$ gilt auch

$$\mu_\iota(\mathcal{X}) \ = \ \int \mathbb{1}_{\mathcal{X}} \,\mathrm{d}\mu_\iota \ \to \ \int \mathbb{1}_{\mathcal{X}} \,\mathrm{d}\mu \ = \ \mu(\mathcal{X}), \quad \iota \to \infty$$

und damit können wir eine Konstante $C < \infty$ und ein $\iota_0 \in I$ finden, sodass $\mu_\iota(\mathcal{X}) \leq C$ für alle $\iota_0 \preceq \iota$ ist. Sei $g \in U_b(\mathcal{X})$ fest aber beliebig. Dann gibt es eine Folge $(g_n)_n \in \mathcal{D}^{\mathbb{N}}$ mit $\|g - g_n\|_\infty \to 0,\ n \to \infty$ und für jedes $n \in \mathbb{N},\ \iota_0 \preceq \iota$ folgt

$$\begin{aligned}
\left| \int g \,\mathrm{d}\mu_\iota - \int g \,\mathrm{d}\mu \right| \ &\leq \ \left| \int g \,\mathrm{d}\mu_\iota - \int g_n \,\mathrm{d}\mu_\iota \right| + \left| \int g_n \,\mathrm{d}\mu_\iota - \int g_n \,\mathrm{d}\mu \right| \\
&\quad + \left| \int g_n \,\mathrm{d}\mu - \int g \,\mathrm{d}\mu \right| \\
&\leq \ 2C \,\|g - g_n\|_\infty + \left| \int g_n \,\mathrm{d}\mu_\iota - \int g_n \,\mathrm{d}\mu \right|,
\end{aligned}$$

was schließlich mit (5.1) auf

$$\limsup_{\iota \to \infty} \left| \int g \,\mathrm{d}\mu_\iota - \int g \,\mathrm{d}\mu \right| \ \leq \ 2C \,\|g - g_n\|_\infty \ \to \ 0, \quad n \to \infty$$

führt. Damit gilt

$$\int g \,\mathrm{d}\mu_\iota \ \to \ \int g \,\mathrm{d}\mu, \quad \iota \to \infty$$

für alle $g \in U_b(\mathcal{X})$, was nach Satz 4.4 die schwache Konvergenz $\mu_\iota \xrightarrow{w} \mu,\ \iota \to \infty$ impliziert.

Die Rückrichtung folgt direkt aus Satz 5.1, da die Metrisierbarkeit und Separabilität von $\mathcal{M}_+(\mathcal{X})$ direkt auch die Separabilität von $\mathcal{X} \cong \Delta \subseteq \mathcal{M}_+(\mathcal{X})$ nach sich zieht. □

Wir werden uns nun der zweiten Aussage vom Beginn dieses Teilabschnitts widmen.

Satz 5.6. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Dann ist $\mathcal{M}_+(\mathcal{X})$ genau dann polnisch, wenn $\mathcal{X}$ polnisch ist.*

Für die Hinrichtung zeigen wir in einem Hilfssatz zunächst den wichtigen Spezialfall, dass der Grundraum $\mathcal{X}$ kompakt ist und weiten dies anschließend auf die Klasse aller polnischen Räume aus.

Hilfssatz 5.7. *Sei $\mathcal{X}$ ein kompakter metrisierbarer topologischer Raum. Dann ist $\mathcal{M}_+(\mathcal{X})$ polnisch.*

Beweis. Wir wählen zunächst eine Metrik d, die $\mathcal{X}$ vollständig metrisiert. Da $\mathcal{X}$ insbesondere separabel ist, können wir zunächst genau wie im Beweis von Satz 5.2 verfahren (abgesehen davon, dass wir anstelle von $\tilde{d}$ direkt mit d arbeiten). Es sei außerdem angemerkt, dass nach [Sim15, Satz 2.4.1] aus der Kompaktheit von $\mathcal{X}$ die Gleichheit $C_b(\mathcal{X}) = U_b(\mathcal{X})$ folgt. Da $C_b(\mathcal{X})$ nach Hilfssatz 5.3 separabel ist, existieren abzählbare dichte Teilmengen $\mathcal{D}_1 \subseteq C_b(\mathcal{X})$ und $\mathcal{D}_2 \subseteq \{ f \in C_b(\mathcal{X}) \mid f \geq 0 \}$. Setzen wir

$$\mathcal{D} \;:=\; \{ f_n \mid n \in \mathbb{N} \} \;:=\; \{\mathbb{1}_{\mathcal{X}}\} \cup \mathcal{D}_1 \cup \mathcal{D}_2,$$

so erhalten wir also wieder die Abbildung

$$T\colon \mathcal{M}_+(\mathcal{X}) \to R,\; \mu \mapsto \left(\int f_n \,\mathrm{d}\mu \right)_n,$$

die ein Homöomorphismus auf ihr Bild ist.

Nun möchten wir nachweisen, dass $T(\mathcal{M}_+(\mathcal{X})) \subseteq R$ abgeschlossen ist, was uns sofort die Behauptung liefert, denn R ist nach Hilfssatz 5.5 selbst polnisch.

Wir verwenden die Charakterisierung von Abgeschlossenheit durch Netze aus Satz 2.4 (a). Ist $(\mu_\iota)_{\iota \in I}$ ein Netz in $\mathcal{M}_+(\mathcal{X})$ mit $T(\mu_\iota) \to r,\; \iota \to \infty$ für ein $r = (r_n)_n \in R$, so müssen wir zeigen, dass r in $T(\mathcal{M}_+(\mathcal{X}))$ enthalten ist. Sicherlich gilt $r_n = \lim_{\iota \to \infty} \int f_n \,\mathrm{d}\mu_\iota$ für alle $n \in \mathbb{N}$ und analog zum Beweis von Satz 5.2 können wir wieder eine Konstante $C < \infty$ und ein $\iota_0 \in I$ finden, sodass $\mu_\iota(\mathcal{X}) \leq C$ für alle $\iota_0 \preceq \iota$ ist.

Wir setzen nun $A := \operatorname{span}_{\mathbb{R}} \mathcal{D}$, was ein dichter Untervektorraum von $C_b(\mathcal{X})$ ist. Auf A definieren wir

$$l\colon A \to \mathbb{R},\; f \mapsto \lim_{\iota \to \infty} \int f \,\mathrm{d}\mu_\iota.$$

Offenbar ist l wohndefiniert und linear. Da $l(f) \leq C \left\| f \right\|_\infty$ für alle $f \in C_b(\mathcal{X})$ gilt, folgt $l \in A^*$. Nach dem Satz von Hahn-Banach (vgl. [Sim15, Folgerung 5.5.2]) existiert nun eine (eindeutige) Fortsetzung von l auf ganz $C_b(\mathcal{X})$, die wir ab jetzt l nennen. Es ist $l \in C_b(\mathcal{X})^*$ mit $\|l\|_{C_b(\mathcal{X})^*} = \|l\|_{A^*}$. Offenbar gilt $l(f) \geq 0$ für alle $f \in C_b(\mathcal{X})$ mit $f \geq 0$ und damit gibt es nach dem Darstellungssatz von Riesz-Markov (vgl. [Sim15, Satz 4.8.8]) ein Maß $\mu \in \mathcal{M}_+(\mathcal{X})$ mit

$$l(f) \;=\; \int f \,\mathrm{d}\mu, \quad f \in U_b(\mathcal{X}).$$

Also folgt

$$\lim_{\iota \to \infty} \int f_n \,\mathrm{d}\mu_\iota \;=\; r_n \;=\; l(f_n) \;=\; \int f_n \,\mathrm{d}\mu$$

für alle $n \in \mathbb{N}$. In völliger Analogie zum Beweis von Satz 5.2 folgt nun die schwache Konvergenz $\mu_\iota \xrightarrow{w} \mu,\; \iota \to \infty$, die wiederum $T(\mu_\iota) \to T(\mu)$ impliziert. Demnach ist $r = T(\mu) \in T(\mathcal{M}_+(\mathcal{X}))$ und insgesamt haben wir gezeigt, dass $T(\mathcal{M}_+(\mathcal{X})) \subseteq R$ abgeschlossen ist. $\square$

Mit Hilfe von Hilfssatz 5.7 kann jetzt der Beweis von Satz 5.6 geführt werden.

Beweis von Satz 5.6. Wir widmen uns zunächst der Hinrichtung des Beweises. Analog zum Beweis von Satz 5.2 fixieren wir auf $\mathcal{X}$ eine Metrik $\tilde{d}$, bezüglich der $\mathcal{X}$ totalbeschränkt ist und bezeichnen die Vervollständigung von $\mathcal{X}$ bezüglich dieser Metrik mit $\tilde{\mathcal{X}}$. $\tilde{\mathcal{X}}$ ist als kompakter metrischer Raum polnisch, weshalb $\mathcal{X} \subseteq \tilde{\mathcal{X}}$ nach Satz 3.2 eine G_δ-Teilmenge ist. Indem wir ein Maß $\mu \in \mathcal{M}_+(\mathcal{X})$ mittels $\mu(B) := \mu(B \cap \mathcal{X}),\; B \in \mathcal{B}(\tilde{\mathcal{X}})$ auf $\tilde{\mathcal{X}}$ fortsetzen, können wir $\mathcal{M}_+(\mathcal{X})$ als Teilmenge von $\mathcal{M}_+(\tilde{\mathcal{X}})$

auffassen (in diesem Fall entspricht die Teilraumtopologie von $\mathcal{M}_+(\mathcal{X}) \subseteq \mathcal{M}_+(\tilde{\mathcal{X}})$ tatsächlich der schwachen Topologie von $\mathcal{M}_+(\mathcal{X})$). Weil $\mathcal{M}_+(\tilde{\mathcal{X}})$ nach Hilfssatz 5.7 polnisch ist, genügt es also nach Satz 3.2 zu zeigen, dass $\mathcal{M}_+(\mathcal{X}) \subseteq \mathcal{M}_+(\tilde{\mathcal{X}})$ eine G_δ-Teilmenge ist.

Da $\mathcal{X} \subseteq \tilde{\mathcal{X}}$ eine G_δ-Teilmenge ist, existieren offene Mengen $U_k \subseteq \tilde{\mathcal{X}},\ k \in \mathbb{N}$ so, dass $\mathcal{X} = \bigcap_{k\in\mathbb{N}} U_k$ gilt. Insbesondere ist damit auch $\mathcal{M}_+(\mathcal{X}) = \bigcap_{k\in\mathbb{N}} \mathcal{M}_+(U_k)$. Für $k, r \in \mathbb{N}$ setzen wir nun

$$V_{k,r} := \left\{ \mu \in \mathcal{M}_+(\tilde{\mathcal{X}}) \,\middle|\, \mu(U_k^{\mathsf{c}}) < \frac{1}{r} \right\}.$$

Mit diesen Mengen folgt dann $\mathcal{M}_+(U_k) = \bigcap_{r\in\mathbb{N}} V_{k,r}$ für alle $k \in \mathbb{N}$ und insbesondere auch

$$\mathcal{M}_+(\mathcal{X}) = \bigcap_{k\in\mathbb{N}} \bigcap_{r\in\mathbb{N}} V_{k,r}.$$

Es genügt also zu beweisen, dass $V_{k,r}$ für alle $k, r \in \mathbb{N}$ offen ist, was äquivalent zur Abgeschlossenheit von

$$C_{k,r} := V_{k,r}^{\mathsf{c}} = \left\{ \mu \in \mathcal{M}_+(\tilde{\mathcal{X}}) \,\middle|\, \mu(U_k^{\mathsf{c}}) \geq \frac{1}{r} \right\}$$

für alle $k, r \in \mathbb{N}$ ist. Wir verwenden hierfür die Charakterisierung abgeschlossener Mengen durch Netze aus Satz 2.4 (a). Fixiere $k, r \in \mathbb{N}$. Sei nun $(\mu_\iota)_{\iota\in I}$ ein Netz in $C_{k,r}$, das schwach gegen ein $\mu \in \mathcal{M}_+(\tilde{\mathcal{X}})$ konvergiert. Dann ist $\mu_\iota(U_k^{\mathsf{c}}) \geq \frac{1}{r}$ für alle $\iota \in I$ und damit nach dem Portmanteau-Theorem (Satz 4.4 (iv))

$$\mu(U_k^{\mathsf{c}}) \geq \limsup_{\iota\to\infty} \mu_\iota(U_k^{\mathsf{c}}) \geq \frac{1}{r},$$

was $\mu \in C_{k,r}$ impliziert. Daher ist $C_{k,r}$ für alle $k, r \in \mathbb{N}$ abgeschlossen und der Beweis vollständig.

Die Rückrichtung folgt wieder unmittelbar aus Satz 5.1: Sei $\mathcal{M}_+(\mathcal{X})$ polnisch, so können wir $\mathcal{X} \cong \Delta \subseteq \mathcal{M}_+(\mathcal{X})$ betrachten. Weil Δ wegen der Metrisierbarkeit von $\mathcal{M}_+(\mathcal{X})$ nun auch abgeschlossen und damit nach Satz 2.13 eine G_δ-Teilmenge von $\mathcal{M}_+(\mathcal{X})$ ist, liefert Satz 3.2 die Behauptung. ☐

Bemerkung. Satz 5.2 und Satz 5.6 gelten auch, wenn man $\mathcal{M}_+(\mathcal{X})$ durch $\mathcal{P}(\mathcal{X})$ ersetzt (für Satz 5.6 folgt die Hinrichtung dann etwa aus der Abgeschlossenheit von $\mathcal{P}(\mathcal{X}) \subseteq \mathcal{M}_+(\mathcal{X})$).

Abschließend soll noch die dritte der zu Beginn formulierten Äquivalenzen gezeigt werden.

Satz 5.8. *Sei $\mathcal{X}$ ein metrisierbarer topologischer Raum. Dann ist $\mathcal{P}(\mathcal{X})$ genau dann metrisierbar und kompakt, wenn $\mathcal{X}$ kompakt ist.*

Beweis. Im Wesentlichen stellen wir den Beweis von Proposition 5.3 in [van03] vor.

Für die Hinrichtung sei $\mathcal{X}$ zunächst kompakt. Dann ist $\mathcal{X}$ insbesondere separabel und damit ist $\mathcal{P}(\mathcal{X})$ nach Satz 5.2 metrisierbar. Außerdem setzen wir

$$A := \left\{ l \in C_b(\mathcal{X})^* \,\middle|\, \|l\|_{C_b(\mathcal{X})^*} \leq 1,\ l(\mathbb{1}_{\mathcal{X}}) = 1,\ \forall f \in C_b(\mathcal{X}) \text{ mit } f \geq 0 : l(f) \geq 0 \right\}$$

und analog zu (4.1)

$$\Phi\colon \mathcal{P}(\mathcal{X}) \to A,\ \mu \mapsto \left[l_\mu\colon C_b(\mathcal{X}) \to \mathbb{R},\ f \mapsto \int f \,\mathrm{d}\mu \right].$$

Der Darstellungssatz von Riesz-Markov (vgl. [Sim15, Satz 4.8.8]) liefert die Bijektivität von Φ, also ist Φ ein Homöomorphismus zwischen $\mathcal{P}(\mathcal{X})$ und $A \subseteq C_b(\mathcal{X})^*$ mit der Schwach-$*$-Topologie. Mit dem Satz von Banach-Alaoglu ([Sim15, Satz 5.8.1]) folgt nun, dass A als abgeschlossene Teilmenge des Kompaktums $\left\{ l \in C_b(\mathcal{X})^* \,\middle|\, \|l\|_{C_b(\mathcal{X})^*} \leq 1 \right\}$ selbst kompakt bezüglich der Schwach-$*$-Topologie ist, was schließlich die Kompaktheit von $\mathcal{P}(\mathcal{X})$ nach sich zieht.

Die Rückrichtung ist wieder eine direkte Konsequenz aus Satz 5.1: Ist $\mathcal{P}(\mathcal{X})$ metrisierbar und kompakt, so muss $\mathcal{X} \cong \Delta$ als abgeschlossene Teilmenge von $\mathcal{P}(\mathcal{X})$ selbst kompakt sein. ☐

Insbesondere haben wir in diesem Abschnitt gesehen, dass $\mathcal{M}_+(\mathcal{X})$ für metrisierbares und separables $\mathcal{X}$ selbst metrisierbar ist. Dies bedeutet auch, dass die schwache Topologie in diesem Fall vollständig durch konvergente Folgen charakterisiert wird. Wir müssen im weiteren Verlauf (bei metrisierbarem und separablem $\mathcal{X}$) also nicht mehr mit dem Konzept des Netzes arbeiten, das wir zu Beginn eingeführt haben. Stattdessen können wir einfach immer Folgen verwenden, um Stetigkeit, Abgeschlossenheit, etc. nachzuweisen.

5.3 Straffheit und der Satz von Prokhorov

Das zentrale Resultat dieses Abschnitts stellt der *Satz von Prokhorov* (Satz 5.12) dar, der für einen polnischen Raum $\mathcal{X}$ eine einfache Charakterisierung der Kompaktheit von Teilmengen von $\mathcal{P}(\mathcal{X})$ bereitstellt. Hierfür werden wir zunächst das Konzept der *Straffheit* einführen, das auch in der Formulierung des Satzes von Prokhorov enthalten ist. Wir orientieren uns im folgenden Abschnitt an [Sim15, Kapitel 4.14].

Definition 5.9. Sei $\mathcal{X}$ ein polnischer Raum und $\mu \in \mathcal{M}_+(\mathcal{X})$. Dann nennen wir eine Teilmenge $A \subseteq \mathcal{M}_+(\mathcal{X})$ *straff*, falls es für jedes $\varepsilon > 0$ eine kompakte Teilmenge $K_\varepsilon \subseteq \mathcal{X}$ gibt mit

$$\mu(K_\varepsilon) \geq \mu(\mathcal{X}) - \varepsilon$$

für alle $\mu \in A$. Ferner heißt ein Maß $\mu \in \mathcal{M}_+(\mathcal{X})$ *straff*, falls $\{\mu\}$ straff ist.

Im folgenden Satz werden wir sehen, dass wir es tatsächlich sehr häufig mit straffen Maßen zu tun haben.

Satz 5.10. *Sei $\mathcal{X}$ ein polnischer Raum und $\mu \in \mathcal{M}_+(\mathcal{X})$. Dann ist μ straff.*

Beweis. Sei d eine Metrik, die $\mathcal{X}$ vollständig metrisiert und sei $\mathcal{D} := \{\, x_m \mid m \in \mathbb{N} \,\}$ eine abzählbare dichte Teilmenge von $\mathcal{X}$. Für $k \in \mathbb{N}$ gilt also $\bigcup_{m\in\mathbb{N}} B_{1/k}(x_m) = \mathcal{X}$ und Maßstetigkeit von unten impliziert $\lim_{M\to\infty} \mu(\bigcup_{m=1}^{M} B_{1/k}(x_m)) = \mu(\mathcal{X})$.

Sei nun $\varepsilon > 0$. Dann finden wir natürliche Zahlen $M_1 \leq M_2 \leq \dots$ mit

$$\mu\left(\bigcup_{m=1}^{M_k} B_{1/k}(x_m)\right) \;\geq\; \mu(\mathcal{X}) - \frac{\varepsilon}{2^k}$$

für alle $k \in \mathbb{N}$. Wir setzen nun

$$S_\varepsilon \;:=\; \bigcap_{k\in\mathbb{N}} \left(\bigcup_{m=1}^{M_k} B_{1/k}(x_m)\right) \quad \text{und} \quad K_\varepsilon := \overline{S_\varepsilon}.$$

Offenbar ist S_ε totalbeschränkt und damit ist $K_\varepsilon = \overline{S_\varepsilon}$ kompakt (dies wird etwa in [Sim15, Satz 2.3.8] bewiesen, hier geht die Vollständigkeit von $(\mathcal{X}, d)$ ein). Außerdem berechnen wir

$$\begin{aligned} \mu(K_\varepsilon) \;\geq\; \mu(S_\varepsilon) \;\geq\; \mu(\mathcal{X}) - \sum_{k=1}^{\infty} \left(\mu(\mathcal{X}) - \mu\left(\bigcup_{m=1}^{M_k} B_{1/k}(x_m)\right)\right) \\ \geq\; \mu(\mathcal{X}) - \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} \;=\; \mu(\mathcal{X}) - \varepsilon. \end{aligned}$$

Weil $\varepsilon > 0$ beliebig gewählt war, ist μ also straff. □

Bemerkung. Da endliche Vereinigungen kompakter Mengen wieder kompakt sind, folgt aus dem vorigen Hilfssatz auch, dass endliche Teilmengen von $\mathcal{M}_+(\mathcal{X})$ straff sind.

Ohne große Anstrengungen liefert der vorige Satz die Regularität von endlichen Borel-Maßen auf polnischen Räumen.

Folgerung 5.11. *Ist $\mathcal{X}$ ein polnischer Raum, so ist jedes $\mu \in \mathcal{M}_+(\mathcal{X})$ regulär.*

Beweis. Sei $\mu \in \mathcal{M}_+(\mathcal{X})$ und $B \in \mathcal{B}(\mathcal{X})$. Wegen Satz 2.12 wissen wir bereits, dass

$$\mu(B) = \sup_{\substack{C \subseteq B \\ C \text{ abgeschlossen}}} \mu(C) \quad \text{bzw.} \quad \mu(B) = \inf_{\substack{U \supseteq B \\ U \text{ offen}}} \mu(U)$$

gilt. Sei nun $\varepsilon > 0$. Dann gibt es eine abgeschlossene Teilmenge $C \subseteq B \subseteq \mathcal{X}$ mit $\mu(B \setminus C) < \varepsilon$. Außerdem gibt es nach Satz 5.10 eine kompakte Menge $\tilde{K} \subseteq \mathcal{X}$ mit $\mu(\tilde{K}) \geq 1 - \varepsilon$. Setzen wir nun $K := \tilde{K} \cap C$, so ist K ebenfalls kompakt und $K \subseteq B$ mit

$$\mu(B \setminus K) \ = \ \mu((B \setminus C) \cap (B \setminus \tilde{K})) \ \leq \ \mu(B \setminus C) + \mu(K^{*\mathsf{c}}) \ \leq \ 2\varepsilon.$$

Daher folgt

$$\mu(B) = \sup_{\substack{K \subseteq B \\ K \text{ kompakt}}} \mu(K)$$

und insgesamt ist μ regulär. ☐

Nun können wir den zentralen Satz dieses Teilabschnitts formulieren und beweisen.

Satz 5.12 (Prokhorov). *Sei $\mathcal{X}$ ein polnischer Raum und $A \subseteq \mathcal{P}(\mathcal{X})$. Dann sind folgende Aussagen äquivalent:*

(i) A ist straff.
(ii) $\overline{A} \subseteq \mathcal{P}(\mathcal{X})$ ist kompakt.

Beweis. Sei zunächst $A \subseteq \mathcal{P}(\mathcal{X})$ straff. Wir möchten nun beweisen, dass $\overline{A}$ kompakt ist, was wegen der Metrisierbarkeit von $\mathcal{P}(\mathcal{X})$ (siehe Satz 5.2 bzw. Satz 5.6) äquivalent dazu ist, dass jede Folge $(\mu_n)_n \in A^{\mathbb{N}}$ eine schwach konvergente Teilfolge besitzt.

Aufgrund der Straffheit von A gibt es für jedes $m \in \mathbb{N}$ ein Kompaktum $K^{(m)} \subseteq \mathcal{X}$, das für alle $n \in \mathbb{N}$

$$\mu_n(K^{(m)}) \geq 1 - \frac{1}{m+1} \tag{5.2}$$

erfüllt. Ohne Einschränkung dürfen wir außerdem annehmen, dass

$$K^{(m)} \subseteq K^{(m+1)} \tag{5.3}$$

für alle $m \in \mathbb{N}$ gilt.

Sei nun also $(\mu_n)_n \in A^{\mathbb{N}}$. In Analogie zu Satz 5.8 lässt sich einsehen, dass

$$\mathcal{M}_{+,\leq 1}(K^{(m)}) \ := \ \left\{ \mu\colon \mathcal{B}(K^{(m)}) \to [0,1] \;\middle|\; \mu \text{ ist endliches Maß und } \mu(K^{(m)}) \leq 1 \right\}$$

für alle $m \in \mathbb{N}$ kompakt bezüglich der schwachen Topologie ist. Damit existiert ein Maß $\nu^{(1)} \in \mathcal{M}_{+,\leq 1}(K^{(1)})$ und eine Teilfolge $(\mu^{(1)}_{n_j})_j$ von $(\mu_n)_n$ mit

$$\mu^{(1)}_{n_j}\Big|_{K^{(1)}} \ \xrightarrow{w} \ \nu^{(1)}.$$

Auf dieselbe Art und Weise lassen für alle $m \geq 2$ sukzessive Maße $\nu^{(m)} \in \mathcal{M}_{+,\leq 1}(K^{(m)})$ und Teilfolgen $(\mu^{(m)}_{n_j})_j$ von $(\mu^{(m-1)}_{n_j})_j$ mit

$$\mu^{(m)}_{n_j}\Big|_{K^{(m)}} \ \xrightarrow{w} \ \nu^{(m)}, \quad j \to \infty$$

finden. Nun diagonalisieren wir und erhalten mit $\mu_{n_j} := \mu^{(j)}_{n_j}$ eine Teilfolge $(\mu_{n_j})_j$ von $(\mu_n)_n$, die

$$\mu_{n_j}\big|_{K^{(m)}} \ \xrightarrow{w} \ \nu^{(m)}, \quad j \to \infty$$

für alle $m \in \mathbb{N}$ erfüllt. Außerdem können die Maße $\nu^{(m)},\ m \in \mathbb{N}$ via $\nu^{(m)}(B) := \nu^{(m)}(B \cap K^{(m)}),\ B \in \mathcal{B}(\mathcal{X})$ einfach auf $\mathcal{X}$ fortgesetzt werden. Mit dieser Konvention können wir nun beweisen, dass die Maße $\nu^{(m)},\ m \in \mathbb{N}$ in dem Sinne monoton wachsen, dass

$$\nu^{(m)}(B) \le \nu^{(m+1)}(B) \tag{5.4}$$

für alle $B \in \mathcal{B}(\mathcal{X})$ und $m \in \mathbb{N}$ gilt.

Wegen schwacher Regularität (vgl. Satz 2.12) genügt es, die Ungleichung (5.4) für abgeschlossene Mengen zu zeigen. Sei also $m \in \mathbb{N}$ fixiert, $C \subseteq \mathcal{X}$ abgeschlossen und seien $f_k \colon \mathcal{X} \to \mathbb{R},\ k \in \mathbb{N}$ die gleichmäßig stetigen, beschränkten Funktionen aus Hilfssatz 2.13 (wobei irgendeine Metrik auf $\mathcal{X}$, die $\mathcal{X}$ metrisiert, fixiert sei). Unter Ausnutzung von (5.3) folgt nun

$$\int_{K^{(m)}} f_k \,\mathrm{d}\mu_{n_j} \ \le\ \int_{K^{(m+1)}} f_k \,\mathrm{d}\mu_{n_j}$$

für alle $k, j \in \mathbb{N}$. Wegen

$$\int_{K^{(m)}} f_k \,\mathrm{d}\mu_{n_j} \to \int f_k \,\mathrm{d}\nu^{(m)} \quad \text{und} \quad \int_{K^{(m+1)}} f_k \,\mathrm{d}\mu_{n_j} \to \int f_k \,\mathrm{d}\nu^{(m+1)}, \quad j \to \infty$$

ist damit auch

$$\int f_k \,\mathrm{d}\nu^{(m)} \ \le\ \int f_k \,\mathrm{d}\nu^{(m+1)}$$

für alle $k \in \mathbb{N}$, was wiederum aufgrund der Konvergenz $f_k \downarrow \mathbb{1}_C$ und dem Satz von Lebesgue Ungleichung (5.4) impliziert.

Schlussendlich möchten wir nun nachweisen, dass durch

$$\nu \colon \mathcal{B}(\mathcal{X}) \to [0,1],\ B \mapsto \sup_{m \in \mathbb{N}} \nu^{(m)}(B)$$

ein Wahrscheinlichkeitsmaß auf $\mathcal{X}$ definiert wird, gegen das $(\mu_{n_j})_j$ schwach konvergiert.

Zunächst verifizieren wir, dass ν ein Wahrscheinlichkeitsmaß ist. $\nu(\emptyset) = 0$ ist klar. Weil $\nu^{(m)}$ in $\mathcal{M}_{+,\le 1}(K^{(m)})$ liegt und wir (5.2) angenommen haben, gilt

$$\nu^{(m)}(\mathcal{X}) = \nu^{(m)}(K^{(m)}) \in [1 - \frac{1}{m+1}, 1]$$

für alle $m \in \mathbb{N}$ und damit $\nu(\mathcal{X}) = 1$. Für die σ-Additivität seien $B_k \in \mathcal{B}(\mathcal{X}),\ k \in \mathbb{N}$ paarweise disjunkt. Dann gilt

$$\begin{aligned} \nu\left(\dot{\bigcup_{k\in\mathbb{N}}} B_k\right) &= \sup_{m\in\mathbb{N}} \nu^{(m)}\left(\dot{\bigcup_{k\in\mathbb{N}}} B_k\right) = \sup_{m\in\mathbb{N}} \sum_{k=1}^{\infty} \nu^{(m)}(B_k) \\ &\overset{(5.4)}{=} \sum_{k=1}^{\infty} \sup_{m\in\mathbb{N}} \nu^{(m)}(B_k) = \sum_{k=1}^{\infty} \nu(B_k), \end{aligned}$$

wobei in der vorletzten Gleichung der Satz von Beppo Levi verwendet wurde. Insgesamt haben wir also $\nu \in \mathcal{P}(\mathcal{X})$ nachgewiesen.

Jetzt bleibt lediglich zu zeigen, dass $(\mu_{n_j})_j$ tatsächlich schwach gegen ν konvergiert. Hierfür möchten wir die Charakterisierung (iv) des Portmanteau-Theorems (Satz 4.4) verwenden. Sei dazu $C \subseteq \mathcal{X}$ abgeschlossen. Zunächst liefert ebendieser Satz für alle $m \in \mathbb{N}$ die Abschätzung $\limsup_{j\to\infty} \mu_{n_j}(C \cap K^{(m)}) \le \nu^{(m)}(C)$ und wegen (5.2) ist zudem $\mu_{n_j}(C \cap K^{(m)\mathsf{c}}) \le \frac{1}{m+1}$ für alle $j, m \in \mathbb{N}$. Insgesamt folgt daraus

$$\begin{aligned} \limsup_{j\to\infty} \mu_{n_j}(C) &= \lim_{m\to\infty} \limsup_{j\to\infty} \left(\mu_{n_j}(C \cap K^{(m)}) + \mu_{n_j}(C \cap K^{(m)\mathsf{c}})\right) \\ &\le \lim_{m\to\infty} \left(\nu^{(m)}(C) + \frac{1}{m+1}\right) = \nu(C), \end{aligned}$$

was den Beweis der Hinrichtung von Satz 5.12 abschließt.

Für die Rückrichtung sei nun $\overline{A}$ kompakt. Außerdem sei d eine Metrik, die $\mathcal{X}$ vollständig metrisiert und $\mathcal{D} := \{ x_n \mid n \in \mathbb{N} \} \subseteq \mathcal{X}$ eine abzählbare dichte Teilmenge.

Wir behaupten, dass für alle $\delta > 0$ ein solches $M_\delta \in \mathbb{N}$ existiert, dass

$$\mu\Big(\bigcup_{m=1}^{M_\delta} B_\delta(x_m)\Big) \;>\; 1-\delta \tag{5.5}$$

für alle $\mu \in A$ gilt. Denn falls es kein derartiges M_δ gibt, so lässt sich ein $\delta > 0$ finden, für das für alle $M \in \mathbb{N}$ ein $\mu_M \in A$ existiert mit

$$\mu_M\Big(\bigcup_{m=1}^{M} B_\delta(x_m)\Big) \;\leq\; 1-\delta.$$

Insbesondere bedeutet das natürlich auch, dass wir für alle $M \in \mathbb{N}$ und $N \geq M$

$$\mu_N\Big(\bigcup_{m=1}^{M} B_\delta(x_m)\Big) \;\leq\; 1-\delta$$

abschätzen dürfen. Aufgrund der Kompaktheit von $\overline{A}$ gibt es eine Teilfolge $(\mu_{N_j})_j$ von $(\mu_N)_N$, die einen schwachen Grenzwert $\mu \in \mathcal{P}(\mathcal{X})$ besitzt. Fixiere nun ein $M \in \mathbb{N}$. Dann ist $\bigcup_{m=1}^{M} B_\delta(x_m)$ offen und daher liefert Charakterisierung (v) des Portmanteau-Theorems (Satz 4.4)

$$\mu\Big(\bigcup_{m=1}^{M} B_\delta(x_m)\Big) \;\leq\; \liminf_{j\to\infty} \mu_{N_j}\Big(\bigcup_{m=1}^{M} B_\delta(x_m)\Big) \;\leq\; 1-\delta. \tag{5.6}$$

Wegen $\bigcup_{m=1}^{\infty} B_\delta(x_m) = \mathcal{X}$ und Maßstetigkeit von unten folgt aus (5.6) unmittelbar

$$\mu(\mathcal{X}) \;\leq\; 1-\delta \;<\; 1,$$

was einen Widerspruch dazu darstellt, dass es sich bei μ um ein Wahrscheinlichkeitsmaß auf $\mathcal{X}$ handelt. Also muss unsere Behauptung in (5.5) tatsächlich gelten. Sei nun ein $\varepsilon > 0$ gegeben. Wir setzen

$$S_\varepsilon \;:=\; \bigcap_{k\in\mathbb{N}} \bigcup_{m=1}^{M_{\varepsilon/2^k}} B_{\varepsilon/2^k}(x_m)$$

und verfahren nun ähnlich wie im Beweis von Satz 5.10: Offensichtlich ist S_ε total beschränkt und damit auch $K_\varepsilon := \overline{S_\varepsilon}$ kompakt (vgl. [Sim15, Satz 2.3.8]). Außerdem gilt für alle $\mu \in A$

$$\mu(K_\varepsilon) \;\geq\; \mu(S_\varepsilon) \;=\; \mu\left(\bigcap_{k\in\mathbb{N}} \bigcup_{m=1}^{M_{\varepsilon/2^k}} B_{\varepsilon/2^k}(x_m)\right) \;\geq\; 1-\sum_{k=1}^{\infty}\frac{\varepsilon}{2^k} \;=\; 1-\varepsilon$$

und damit ist A straff. $\square$

6 Wassersteinmetriken und optimaler Transport

Eine unserer zentralen vorigen Erkenntnisse war, dass für einen polnischen Raum $\mathcal{X}$ auch $\mathcal{M}_+(\mathcal{X})$ bzw. $\mathcal{P}(\mathcal{X})$ polnisch ist (siehe Satz 5.6). Im zugehörigen Beweis haben wir allerdings explizit gar keine Metrik angegeben, die $\mathcal{M}_+(\mathcal{X})$ vollständig metrisiert. Analysiert man den Beweis schrittweise, so lässt sich die Metrik D, die uns in Satz 5.6 zu unserer Erkenntnis verhalf, allerdings rekonstruieren: Es gilt nämlich

$$D(\mu,\nu) \;:=\; \max_{n\in\mathbb{N}} \frac{\left|h\left(\int f_n \,\mathrm{d}\mu\right) - h\left(\int f_n \,\mathrm{d}\nu\right)\right|}{2^n}, \quad \mu,\nu \in \mathcal{M}_+(\mathcal{X})$$

für einen Homöomorphismus $h\colon \mathbb{R} \to (0,1)$ und eine dichte Teilmenge $\{\, f_n \mid n \in \mathbb{N} \,\} \subseteq U_b(\mathcal{X})$, die $\mathbb{1}_{\mathcal{X}}$ enthält. Dabei muss eine Metrik d auf $\mathcal{X}$ gewählt sein, bezüglich der $(\mathcal{X}, d)$ totalbeschränkt ist. Offenbar ist diese Metrik kompliziert und unhandlich, denn $\mathcal{X}$ muss etwa bereits selbst mit einer recht speziellen Metrik versehen sein. Zudem ist die Metrik in dem Sinne etwas unintuitiv, dass es völlig unklar bleibt, ob und gegebenenfalls wie sich die Geometrie von $(\mathcal{X}, d)$ auf $(\mathcal{M}_+(\mathcal{X}), D)$ überträgt.

In diesem Abschnitt betrachten wir ausschließlich Wahrscheinlichkeitsmaße und werden uns von rein topologischen zu etwas geometrischeren Überlegungen hinwenden. Damit geht einher, dass in diesem Abschnitt Polnizität häufig als Eigenschaft metrischer Räume und nicht topologischer Räume angesehen wird. Das bedeutet, dass auf einem polnischen Raum $\mathcal{X}$ immer auch eine Metrik d fixiert ist, der (etwa in der Anwendung) eine besondere Bedeutung zukommt, und die $\mathcal{X}$ vollständig metrisiert. Wir sprechen in diesem Abschnitt dann davon, dass $(\mathcal{X}, d)$ ein *polnischer metrischer Raum* ist. Im folgenden Abschnitt werden wir nun für einen polnischen metrischen Raum $(\mathcal{X}, d)$ eine Klasse geometrisch besonders anschaulicher Metriken auf $\mathcal{P}(X)$ – die sogenannten *Wassersteinmetriken* – vorstellen, die die schwache Topologie metrisieren. Hierbei werden wir an vielen Stellen unsere bisherigen Erkenntnisse – in erster Linie den Satz von Prokhorov – beim Einsatz in konkreten Anwendungssituationen sehen können.

Zunächst werden wir in den Teilabschnitten 6.1 und 6.2 einige grundlegende Denkweisen, Begriffe und Ergebnisse aus der Theorie des optimalen Transports vorstellen, die wir anschließend in Teilabschnitt 6.3 verwenden, um die Wassersteinmetriken zu beschreiben. Dieser Abschnitt orientiert sich überwiegend an [Vil09, Kapitel 4 bis 6].

6.1 Kopplungen

Beim optimalen Transport orientieren wir uns an der Formulierung von Monge, der mit seiner Arbeit [Mon81] das Gebiet begründete. Die Ausgangssituation in Monges Formulierung können wir uns wie folgt vorstellen (vgl. [Vil09, Seiten 41 und 42]): Wir nehmen an, dass wir eine gewisse Menge an Erde aus dem Boden fördern müssen, um diese anschließend an Stellen zu bringen, an denen sie weiterverarbeitet werden kann. Die räumlichen Verteilungen vor und nach dem Transport sind hierbei festgelegt und werden jeweils durch Wahrscheinlichkeitsmaße μ bzw. ν auf Räumen $\mathcal{X}$ bzw. $\mathcal{Y}$ modelliert. Allgemeiner sollen Wahrscheinlichkeitsmaße in diesem Abschnitt also nicht den Zufall, sondern Verteilungen von Gütern (oder allgemeiner Masse) in einem Raum darstellen.

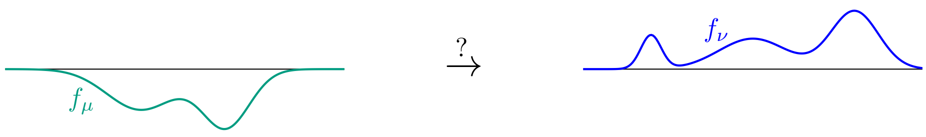


Abbildung 6.1: Mögliche Dichten f_μ, f_ν von μ und ν, die die räumliche Verteilung der Erde bzw. Masse (hier in $\mathbb{R}$) darstellen.

Wir möchten nun eine Möglichkeit vorstellen, mittels *Kopplungen* einen Transportplan zwischen zwei Masseverteilungen μ und ν zu modellieren, der spezifiziert, woher wie viel Masse wohin transportiert werden soll.

Definition 6.1 (Kopplung)**.** Seien $\mathcal{X}$, $\mathcal{Y}$ topologische Räume und $A \subseteq \mathcal{P}(\mathcal{X})$, $B \subseteq \mathcal{P}(\mathcal{Y})$. Wir definieren dann

$$\Pi(A, B) \ := \ \{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \mid \pi(\cdot \times \mathcal{Y}) \in A \quad \text{und} \quad \pi(\mathcal{X} \times \cdot) \in B\}.$$

Für $\mu \in \mathcal{P}(\mathcal{X})$ und $\nu \in \mathcal{P}(\mathcal{Y})$ definieren wir außerdem $\Pi(\mu, \nu) \ := \ \Pi(\{\mu\}, \{\nu\})$ und nennen die Elemente von $\Pi(\mu, \nu)$ *Kopplungen* von μ und ν oder auch *Transportpläne* zwischen μ und ν.

Eine Kopplung $\pi \in \Pi(\mu, \nu)$ spezifiziert nun in dem Sinne einen Transportplan zwischen μ und ν, dass wir für $A \in \mathcal{B}(\mathcal{X})$ und $B \in \mathcal{B}(\mathcal{Y})$ die Zahl $\pi(A \times B) \in [0, 1]$ als die Masse interpretieren, die von A nach B transportiert wird. Die Bedingungen an die Marginalverteilungen $\pi(\cdot \times \mathcal{Y}) = \mu$ und $\pi(\mathcal{X} \times \cdot) = \nu$ sichern dann, dass die gesamte Masse von μ abtransportiert wird und nach dem Transport die spezifizierte Endverteilung der Massen vorliegt.

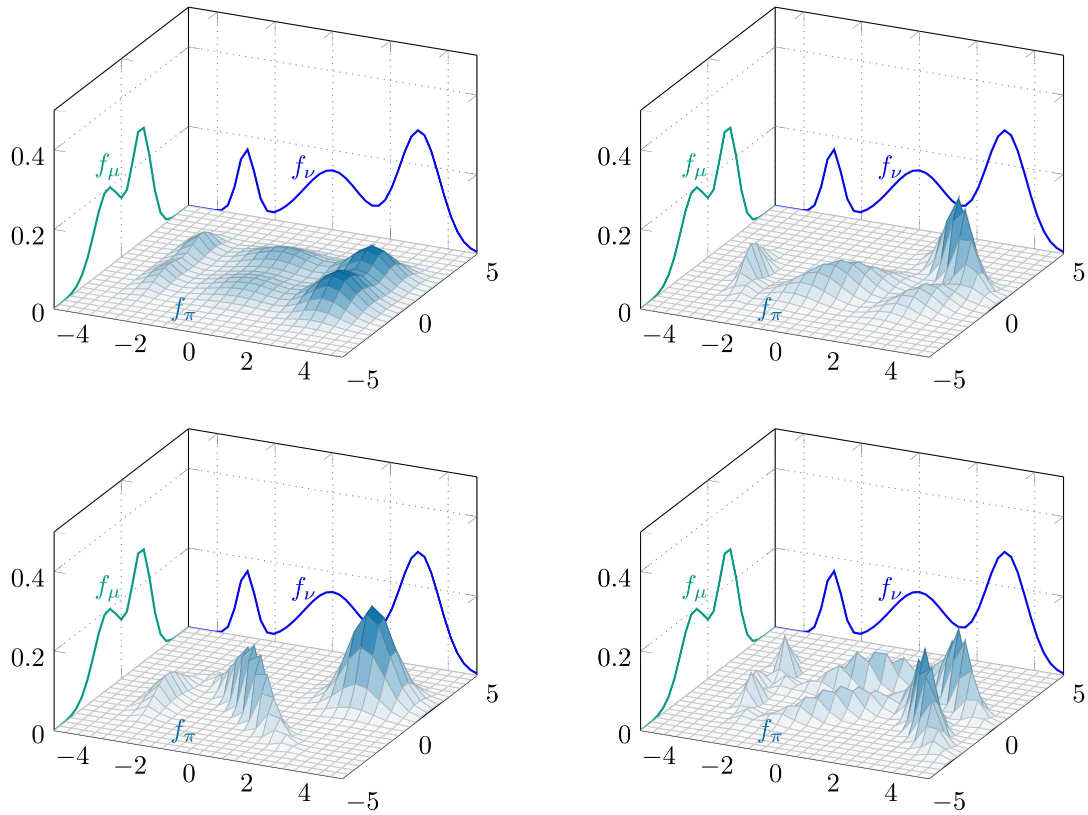


Abbildung 6.2: Mögliche Dichten f_π von Kopplungen zweier reeller Wahrscheinlichkeitsmaße μ und ν mit Dichten f_μ, f_ν.

In den beiden nächsten Hilfssätzen halten wir einige grundlegende Eigenschaften von Kopplungen fest, die sich im weiteren Verlauf als nützlich erweisen werden.

Hilfssatz 6.2. *Seien $\mathcal{X}$, $\mathcal{Y}$ polnische Räume und $A \subseteq \mathcal{P}(\mathcal{X})$ sowie $B \subseteq \mathcal{P}(\mathcal{Y})$. Dann gelten die folgenden Implikationen:*

(a) Sind A und B straff, so ist auch $\Pi(A, B) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ straff.
(b) Sind A und B abgeschlossen, so ist auch $\Pi(A, B) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ abgeschlossen.
(c) Sind A und B kompakt, so ist auch $\Pi(A, B) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ kompakt.

Beweis. Der Beweis von Aussage (a) orientiert sich an [Vil09, Hilfssatz 4.4]. Fixiere ein beliebiges $\varepsilon > 0$ und wähle kompakte Teilmengen $K_{\varepsilon/2} \subseteq \mathcal{X}$ sowie $L_{\varepsilon/2} \subseteq \mathcal{Y}$ so, dass $\mu(K_{\varepsilon/2}) \geq 1 - \frac{\varepsilon}{2}$ für alle $\mu \in A$ und $\nu(L_{\varepsilon/2}) \geq 1 - \frac{\varepsilon}{2}$ für alle $\nu \in B$ gilt. Dann ist $K_{\varepsilon/2} \times L_{\varepsilon/2} \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ auch kompakt und für jede Kopplung π von μ und ν gilt

$$\begin{aligned} \pi(K_{\varepsilon/2} \times L_{\varepsilon/2}) \;&=\; 1 - \pi(K^{\mathsf{c}}_{\varepsilon/2} \times \mathcal{Y} \cup \mathcal{X} \times L^{\mathsf{c}}_{\varepsilon/2}) \\ &\geq\; 1 - \mu(K^{\mathsf{c}}_{\varepsilon/2}) - \nu(L^{\mathsf{c}}_{\varepsilon/2}) \;\geq\; 1 - \varepsilon, \end{aligned}$$

also ist auch $\Pi(A, B) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ straff.

Da mit $\mathcal{X}$ und $\mathcal{Y}$ auch $\mathcal{X} \times \mathcal{Y}$ polnisch ist (siehe Folgerung 3.7), können wir für Aussage (b) mit der Charakterisierung von Abgeschlossenheit durch Folgen arbeiten. Seien also A und B abgeschlossen und sei $(\pi_k)_k \in \Pi(A, B)^{\mathbb{N}}$ eine schwach konvergente Folge mit Grenzwert $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$. Dann ist für alle abgeschlossenen Teilmengen $C \subseteq \mathcal{X}$ auch $C \times \mathcal{Y} \subseteq \mathcal{X} \times \mathcal{Y}$ abgeschlossen und das Portmanteau-Theorem (Satz 4.4) liefert

$$\limsup_{k \to \infty} \pi_k(C \times \mathcal{Y}) \;\leq\; \pi(C \times \mathcal{Y}).$$

Damit konvergiert $(\pi_k(\cdot \times \mathcal{Y}))_k \in A^{\mathbb{N}}$ wiederum nach dem Portmanteau-Theorem schwach gegen $\pi(\cdot \times \mathcal{Y})$ und die Abgeschlossenheit von $A \subseteq \mathcal{P}(\mathcal{X})$ liefert, dass $\pi(\cdot \times \mathcal{Y})$ in A enthalten ist. Analog folgt $\pi(\mathcal{X} \times \cdot) \in B$, also ist insgesamt $\pi \in \Pi(A, B)$. Daher ist $\Pi(A, B)$ abgeschlossen.

Durch eine Anwendung des Satzes von Prokhorov (Satz 5.12) können wir aus (a) und (b) unmittelbar auch Aussage (c) folgern, denn dieser liefert, dass sowohl in $\mathcal{P}(\mathcal{X})$ und $\mathcal{P}(\mathcal{Y})$ als auch in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ kompakte Teilmengen gerade den Teilmengen entsprechen, die straff und abgeschlossen sind. □

Hilfssatz 6.3. *Seien $\mathcal{X}$, $\mathcal{Y}$ metrisierbare topologische Räume und seien $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$, $\mu \in \mathcal{P}(\mathcal{X})$, $(\nu_k)_k \in \mathcal{P}(\mathcal{Y})^{\mathbb{N}}$ und $\nu \in \mathcal{P}(\mathcal{Y})$ mit*

$$\mu_k \xrightarrow{w} \mu \quad \textit{und} \quad \nu_k \xrightarrow{w} \nu, \quad k \to \infty.$$

Sei außerdem $\pi_k \in \Pi(\mu_k, \nu_k)$ für alle $k \in \mathbb{N}$ und $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ so, dass

$$\pi_k \xrightarrow{w} \pi, \quad k \to \infty.$$

Dann ist $\pi \in \Pi(\mu, \nu)$.

Beweis. Wie im Beweis von Hilfssatz 6.2 gilt hier auch

$$\pi_k(\cdot \times \mathcal{Y}) \;\xrightarrow{w}\; \pi(\cdot \times \mathcal{Y}) \quad \text{und} \quad \pi_k(\mathcal{X} \times \cdot) \;\xrightarrow{w}\; \pi(\mathcal{X} \times \cdot), \quad k \to \infty.$$

Wegen $\pi_k(\cdot \times \mathcal{Y}) = \mu_k$ und $\pi_k(\mathcal{X} \times \cdot) = \nu_k$ für alle $k \in \mathbb{N}$ folgt damit aber insbesondere auch $\pi(\cdot \times \mathcal{Y}) = \mu$ und $\pi(\mathcal{X} \times \cdot) = \nu$, also ist π in $\Pi(\mu, \nu)$ enthalten. □

6.2 Optimaler Transport

Im Allgemeinen wird es sehr viele unterschiedliche Transportpläne zwischen zwei Maßen μ und ν geben. Offenbar können wir aber intuitiv bereits sagen, dass manche Transportpläne gewissermaßen

‚besser' als andere sind: Ist etwa $\mathcal{X}$ ein Grundraum und wir betrachten Transportpläne zwischen $\mu \in \mathcal{P}(X)$ und sich selbst, so können wir die triviale Kopplung $\pi_1(A \times B) := \mu(A \cap B)$ wählen, die jede Masse einfach genau an ihrem Platz lässt, oder wir können die unabhängige Kopplung $\pi_2(A \times B) := \mu(A) \times \mu(B)$ nehmen, bei der die Masse in jedem Punkt ‚gleichmäßig' verteilt wird. Wenn μ nicht in einem Punkt konzentiert ist, handelt es sich hierbei um zwei verschiedene Kopplungen, aber natürlich wäre die triviale Kopplung π_1 zu präferieren, wenn wir den ‚Arbeitsaufwand' minimieren möchten.

In der nächsten Definition führen wir eine Möglichkeit ein, den Aufwand eines Transportplans zu quantifizieren.

Definition 6.4 (Kosten eines Transportplans)**.** Seien $\mathcal{X}$, $\mathcal{Y}$ topologische Räume, $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$ und $c\colon \mathcal{X} \times \mathcal{Y} \to [0, \infty]$ eine $\mathcal{B}(\mathcal{X} \times \mathcal{Y})$-messbare Kostenfunktion. Die Kosten von $\pi \in \Pi(\mu, \nu)$ möchten wir dann mit

$$C(\pi) \;:=\; \int c \,\mathrm{d}\pi \;=\; \int_{\mathcal{X} \times \mathcal{Y}} c(x, y)\, \pi(\mathrm{d}(x, y)) \in\; [0, \infty]$$

bezeichnen.

$c(x, y)$ soll in diesem Fall für $x \in \mathcal{X}$ und $y \in \mathcal{Y}$ den Aufwand angeben, der für den Transport einer Einheitsmasse von x nach y nötig ist. Der Modellierung liegt die Annahme zugrunde, dass die Kosten, um eine Masse m von x nach y zu transportieren, gerade $m \cdot c(x, y)$ ist.

Im Falle stetiger Kostenfunktionen liefert uns der folgende Satz die Existenz eines optimalen Transportplans, der die Kosten über alle möglichen Transportpläne minimiert.

Satz 6.5 (Existenz eines optimalen Transportplans)**.** *Seien $\mathcal{X}$, $\mathcal{Y}$ polnische Räume, $\mu \in \mathcal{P}(\mathcal{X})$, $\nu \in \mathcal{P}(\mathcal{Y})$ und $c\colon \mathcal{X} \times \mathcal{Y} \to [0, \infty]$ eine stetige Kostenfunktion. Dann existiert eine Kopplung $\hat{\pi}$ von μ und ν, die*

$$C(\hat{\pi}) \;=\; \min_{\pi \in \Pi(\mu, \nu)} C(\pi)$$

erfüllt. Wir nennen $\hat{\pi}$ eine optimale Kopplung *von μ und ν oder auch einen* optimalen Transportplan *zwischen μ und ν.*

Bemerkung. Der Beweis folgt [Vil09, Satz 4.1 und Hilfssatz 4.3], wo eine allgemeinere Version des Satzes vorgestellt wird, die lediglich die Unterhalbstetigkeit von c fordert sowie unter gewissen Bedingungen die Negativität von c zulässt. Für unsere Zwecke reicht obige Formulierung allerdings aus.

Für den Beweis von Satz 6.5 wird noch ein Hilfssatz benötigt.

Hilfssatz 6.6. *In der Situation von Satz 6.5 gilt für jede schwach konvergente Folge $(\pi_k)_k \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})^{\mathbb{N}}$ mit Grenzwert $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ die Ungleichung*

$$C(\pi) \;\leq\; \liminf_{k \to \infty} C(\pi_k).$$

Beweis. Setzen wir $c_l := \min(c, l) \in C_b(\mathcal{X})$ für alle $l \in \mathbb{N}$, so gilt $c_l \uparrow c$ punktweise. Sei nun $(\pi_k)_k \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})^{\mathbb{N}}$ mit $\pi_k \xrightarrow{w} \pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ eine schwach konvergente Folge. Dann gilt $\int c_l \,\mathrm{d}\pi_k \;\leq\; \int c \,\mathrm{d}\pi_k$ für alle $k, l \in \mathbb{N}$ und damit auch

$$\int c_l \,\mathrm{d}\pi \;=\; \lim_{k \to \infty} \int c_l \,\mathrm{d}\pi_k \;\leq\; \liminf_{k \to \infty} \int c \,\mathrm{d}\pi_k$$

für alle $l \in \mathbb{N}$. Der Satz von Beppo Levi liefert also

$$\int c \,\mathrm{d}\pi \;=\; \lim_{l \to \infty} \int c_l \,\mathrm{d}\pi \;\leq\; \liminf_{k \to \infty} \int c \,\mathrm{d}\pi_k$$

und es folgt die Behauptung. □

Beweis von Satz 6.5. Zunächst sind $\{\mu\} \subseteq \mathcal{P}(\mathcal{X})$ und $\{\nu\} \subseteq \mathcal{P}(\mathcal{Y})$ kompakt, also ist $\Pi(\mu,\nu) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ wegen Hilfssatz 6.2 (c) ebenfalls kompakt. Sei nun $(\pi_k)_k \in \Pi(\mu,\nu)^{\mathbb{N}}$ mit

$$C(\pi_k) \;\to\; \inf_{\pi \in \Pi(\mu,\nu)} C(\pi), \quad k \to \infty.$$

Dann gibt es eine Teilfolge $(\pi_{k_l})_l$ und ein $\hat{\pi} \in \Pi(\mu,\nu)$ mit $\pi_{k_l} \xrightarrow{w} \hat{\pi}$. Nach Hilfssatz 6.6 gilt

$$C(\hat{\pi}) \;\leq\; \liminf_{k\to\infty} C(\pi_k) \;=\; \inf_{\pi \in \Pi(\mu,\nu)} C(\pi),$$

was bedeutet, dass $\hat{\pi}$ ein optimaler Transportplan zwischen μ und ν ist. □

Im nächsten Teilabschnitt werden wir noch das folgende Resultat aus der Theorie des optimalen Transports benötigen, das uns etwa unter gewissen Voraussetzungen liefert, dass schwache Grenzwerte von Folgen optimaler Kopplungen selbst wieder optimal sind. Da der Beweis des Satzes sehr technisch ist und den Umfang dieser Arbeit sprengen würde, zitieren wir ihn lediglich.

Satz 6.7. *Seien $\mathcal{X}$, $\mathcal{Y}$ polnische Räume, sei $c\colon \mathcal{X} \times \mathcal{Y} \to [0,\infty)$ eine stetige Kostenfunktion und seien $A \subseteq \mathcal{P}(\mathcal{X})$ und $B \subseteq \mathcal{P}(\mathcal{Y})$ kompakte Teilmengen. Dann ist*

$$\hat{\Pi}(A,B) \;:=\; \{\,\pi \in \Pi(A,B) \mid \pi \textit{ ist eine optimale Kopplung}\,\} \;\subseteq\; \mathcal{P}(\mathcal{X} \times \mathcal{Y})$$

kompakt.

Beweis. Siehe [Vil09, Folgerung 5.21]. □

6.3 Wassersteinmetriken

Mit den Ergebnissen aus den Teilabschnitten 6.1 und 6.2 sind wir nun bereit, die Wassersteinmetriken W_p für $p \in [1,\infty)$ zu definieren. Der wesentliche Gedanke hierbei ist, für einen polnischen metrischen Raum $(\mathcal{X}, d)$ die Distanz zwischen zwei Wahrscheinlichkeitsmaßen μ und $\nu \in \mathcal{P}(\mathcal{X})$ über die Kosten eines optimalen Transportplans zwischen μ und ν zu definieren, wobei die Kostenfunktion eine Potenz von d ist.

In diesem Teilabschnitt werden wir mit Satz 6.11 das zentrale Resultat dieses Abschnitts formulieren und beweisen: Unter gewissen Voraussetzungen metrisiert W_p die schwache Topologie und $(\mathcal{P}(X), W_p)$ ist selbst ein polnischer metrischer Raum.

Definition 6.8. Sei $(\mathcal{X}, d)$ ein polnischer metrischer Raum, sei $p \in [1,\infty)$ und seien $\mu, \nu \in \mathcal{P}(\mathcal{X})$. Dann nennen wir

$$W_p(\mu,\nu) \;:=\; \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y))\right)^{1/p} \;\in\; [0,\infty]$$

die *p-te Wassersteindistanz* zwischen μ und ν.

Zunächst muss nachgewiesen werden, dass die Wassersteinmetrik tatsächlich ihren Namen verdient.

Satz 6.9. *In der Situation von Definition 6.8 erfüllt W_p die Eigenschaften einer Metrik, abgesehen davon, dass $W_p(\mu,\nu) = \infty$ möglich ist.*

Beweis. Offensichtlich ist W_p symmetrisch.

Für die positive Definitheit seien $\mu, \nu \in \mathcal{P}(\mathcal{X})$ mit $W_p(\mu,\nu) = 0$. Dann ist

$$W_p(\mu,\nu)^p \;=\; \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y)) \;=\; 0$$

und da die Zuordnung $(x,y) \mapsto d(x,y)$ stetig ist, existiert nach Satz 6.5 eine Kopplung $\hat{\pi} \in \Pi(\mu,\nu)$ mit
$$\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \hat{\pi}(\mathrm{d}(x,y)) \;=\; 0.$$
Damit folgt $\hat{\pi}(\{\,(x,x) \mid x \in \mathcal{X}\,\}) = \hat{\pi}(\{d(x,y)^p = 0\}) = 1$ und für jedes $B \in \mathcal{B}(\mathcal{X})$ gilt $\mu(B) = \hat{\pi}(B \times \mathcal{X}) = \hat{\pi}(\mathcal{X} \times B) = \nu(B)$, also $\mu = \nu$.

Es bleibt die Dreiecksungleichung zu zeigen. Seien dazu $\mu_1, \mu_2, \mu_3 \in \mathcal{P}(\mathcal{X})$ und seien $\hat{\pi}_{1,2}$ und $\hat{\pi}_{2,3}$ optimale Kopplungen von μ_1 und μ_2 bzw. μ_2 und μ_3. Nach [Vil09, Seiten 23 und 24] gibt es nun ein Wahrscheinlichkeitsmaß $\pi \in \mathcal{P}(\mathcal{X}^3)$ mit $\pi(\cdot \times \mathcal{X}) = \hat{\pi}_{1,2}$ und $\pi(\mathcal{X} \times \cdot) = \hat{\pi}_{2,3}$. Damit gilt
$$\begin{aligned} W_p(\mu_1,\mu_3) \;&\leq\; \left(\int_{\mathcal{X}\times\mathcal{X}} d(x,z)^p \, \pi(\mathrm{d}(x,y,z))\right)^{1/p} \\ &\leq\; \left(\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y,z))\right)^{1/p} + \left(\int_{\mathcal{X}\times\mathcal{X}} d(y,z)^p \, \pi(\mathrm{d}(x,y,z))\right)^{1/p} \\ &=\; W_p(\mu_1,\mu_2) + W_p(\mu_2,\mu_3), \end{aligned}$$
wobei wir bei der zweiten Abschätzung die Minkowski-Ungleichung anwenden. □

Bemerkung. Seien $x_0, y_0 \in \mathcal{X}$. Wegen $\Pi(\delta_{x_0}, \delta_{y_0}) = \left\{\delta_{(x_0,y_0)}\right\} \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{X})$ folgt
$$W_p(\delta_{x_0}, \delta_{y_0}) \;=\; \left(\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \delta_{(x_0,y_0)}(\mathrm{d}(x,y))\right)^{1/p} \;=\; d(x_0,y_0).$$
Die Abbildung
$$(\mathcal{X}, d) \to (\mathcal{P}(\mathcal{X}), W_p), \quad x \mapsto \delta_x$$
ist also eine isometrische Einbettung. Insbesondere ist dies auch vor dem Kontext von Satz 5.1 zu betrachten, wo wir nachgewiesen haben, dass obige Abbildung ein Homöomorphismus auf ihr Bild ist, sofern $\mathcal{P}(X)$ mit der schwachen Topologie versehen ist. Wird die schwache Topologie also tatsächlich von W_p metrisiert, so bietet W_p eine Möglichkeit, die Geometrie von $(\mathcal{X}, d)$ gewissermaßen auch auf $\mathcal{P}(X)$ zu übertragen.

Bemerkung. Seien $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, $1 \leq p \leq q < \infty$ und $r \in [1,\infty]$ so, dass $\frac{1}{q/p} + \frac{1}{r} = 1$. Eine Anwendung der Hölder-Ungleichung liefert dann
$$\begin{aligned} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y)) \;&\leq\; \left(\int_{\mathcal{X}\times\mathcal{X}} (d(x,y)^p)^{q/p} \, \pi(\mathrm{d}(x,y))\right)^{p/q} \left(\int_{\mathcal{X}\times\mathcal{X}} \pi(\mathrm{d}(x,y))\right)^{1/r} \\ &=\; \left(\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^q \, \pi(\mathrm{d}(x,y))\right)^{p/q}, \end{aligned}$$
was
$$\left(\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y))\right)^{1/p} \;\leq\; \left(\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^q \, \pi(\mathrm{d}(x,y))\right)^{1/q}$$
und damit
$$W_p(\mu,\nu) \;\leq\; W_q(\mu,\nu), \quad \mu,\nu \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$$
impliziert.

Beispiel (W_2-Distanz von Normalverteilungen)**.** Sei $\mathcal{X} := \mathbb{R}$ versehen mit der von $|\cdot|$ induzierten Metrik und seien
$$\mu_1 \;:=\; \mathcal{N}(\lambda_1, \sigma_1^2), \quad \mu_2 \;:=\; \mathcal{N}(\lambda_2, \sigma_2^2) \;\in\; \mathcal{P}(\mathbb{R})$$
Normalverteilungen mit Erwartungswerten $\lambda_1, \lambda_2 \in \mathbb{R}$ und Varianzen $\sigma_1^2, \sigma_2^2 \in (0,\infty)$. Wir möchten $W_2(\mu_1,\mu_2)$ berechnen.

Sei dazu $\pi \in \Pi(\mu_1, \mu_2)$ beliebig und sei $(Z_1, Z_2)^\top$ ein Zufallsvektor, der π als Verteilung besitzt. Insbesondere existieren auch derartige Zufallsvariablen X_1, X_2 mit Verteilung $\mathcal{N}(0,1)$, dass

$$Z_1 = \lambda_1 + \sigma_1 X_1 \quad \text{und} \quad Z_2 = \lambda_2 + \sigma_2 X_2$$

gilt. Nun ist

$$\begin{aligned} \int_{\mathbb{R}^2} |x-y|^2 \, \pi(\mathrm{d}(x,y)) &= \mathbb{E}\left[(Z_1 - Z_2)^2\right] = \mathbb{E}\left[(\lambda_1 + \sigma_1 X_1 - \lambda_2 - \sigma_2 X_2)^2\right] \\ &= \lambda_1^2 + \sigma_1^2 + \lambda_2^2 + \sigma_2^2 - 2\lambda_1\lambda_2 - 2\sigma_1\sigma_2\mathbb{E}\left[X_1 X_2\right] \\ &\geq (\lambda_1 - \lambda_2)^2 + (\sigma_1 - \sigma_2)^2 \end{aligned}$$

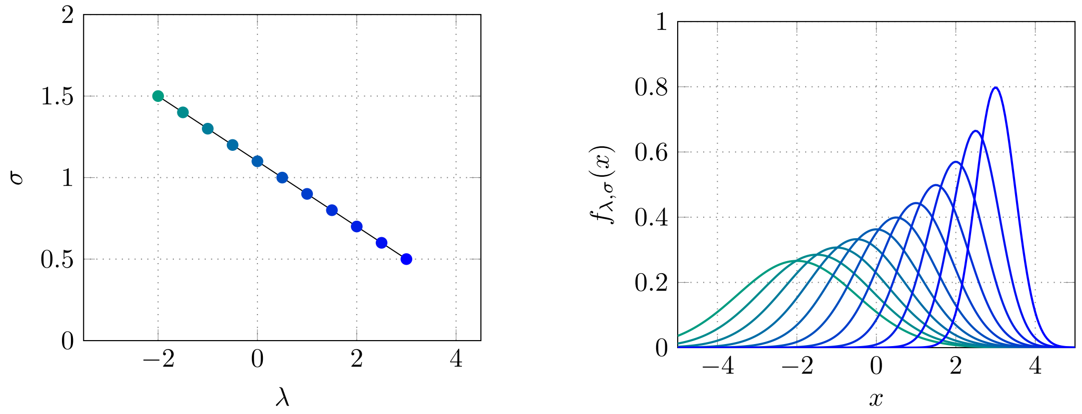
mit Gleichheit in der letzten Abschätzung, falls $\mathbb{E}\left[X_1 X_2\right] = 1$, also falls fast sicher $X_1 = X_2$ gilt. Damit ist

$$W_2(\mu_1, \mu_2) = \left(\inf_{\pi \in \Pi(\mu_1,\mu_2)} \int_{\mathbb{R}^2} |x-y|^2 \, \pi(\mathrm{d}(x,y)) \right)^{1/2} = \left\| \begin{pmatrix} \lambda_1 \\ \sigma_1 \end{pmatrix} - \begin{pmatrix} \lambda_2 \\ \sigma_2 \end{pmatrix} \right\|_2$$

und die Abbildung

$$\mathcal{N} \colon (\mathbb{R} \times (0,\infty), \|\cdot\|_2) \to (\mathcal{P}(\mathbb{R}), W_2), \ (\lambda, \sigma) \mapsto \mathcal{N}(\lambda, \sigma)$$

ist eine isometrische Einbettung. Dies bedeutet insbesondere, dass die Wahrscheinlichkeitsmaße zu aufeinanderfolgenden Lebesgue-Dichten $f_{\lambda,\sigma}$ in Abbildung 6.3b in W_2 äquidistant liegen.



(a) Parametervektoren (λ, σ).

(b) Dichten $f_{\lambda,\sigma}$ von $\mathcal{N}(\lambda, \sigma)$.

Abbildung 6.3: Visualisierung von $\mathcal{N}$.

Auch abgesehen davon, dass es sich bei W_p streng genommen um keine Metrik handelt, metrisiert W_p die schwache Konvergenz für jeden polnischen metrischen Raum $(\mathcal{X}, d)$ noch nicht. In [Vil09] wird dies gelöst, indem $\mathcal{P}(\mathcal{X})$ auf eine Teilmenge eingeschränkt wird, auf der W_p lediglich endliche Werte annimmt, und ein leicht abgeändertes Konzept von schwacher Konvergenz in dieser Teilmenge einführt (siehe [Vil09, Definitionen 6.4 und 6.8]). Wir umgehen diesen zusätzlichen technischen Aufwand, indem wir ab sofort die Beschränktheit von d annehmen. In der Praxis lässt sich für einen metrischen Raum $(\mathcal{X}, d)$ immer eine äquivalente, beschränkte Metrik $\tilde{d}$ finden, die für kleine Werte d selbst entspricht. Beispielsweise lässt sich

$$\tilde{d} := \min(d, d_{\max})$$

für eine gewisse maximale Distanz $d_{\max} \in (0, \infty)$ wählen.

Bei der folgenden Dualitätsformel, die wir für den Beweis unseres zentralen Satzes an einer Stelle benötigen werden, handelt es sich wieder um ein Resultat, das wir angesichts des umfangreichen und technischen Beweises lediglich zitieren werden.

Satz 6.10 (Dualitätsformel von Kantorovich-Rubinstein)**.** *Sei $(\mathcal{X}, d)$ ein polnischer metrischer Raum, wobei d eine beschränkte Metrik sei. Wir definieren*

$$\mathrm{Lip}_1(\mathcal{X}) \;:=\; \{\, f\colon \mathcal{X} \to \mathbb{R} \mid f \text{ ist } 1\text{-lipschitzstetig} \,\}\,.$$

Dann gilt für jedes Paar $\mu, \nu \in \mathcal{P}(\mathcal{X})$ von Wahrscheinlichkeitsmaßen

$$W_1(\mu, \nu) \;=\; \sup_{f \in \mathrm{Lip}_1(\mathcal{X})} \left(\int f \,\mathrm{d}\mu - \int f \,\mathrm{d}\nu \right).$$

Beweis. Siehe [Vil09, Bemerkung 6.5]. □

Wir sind nun bereit, uns dem zentralen Satz dieses Abschnitts zu widmen.

Satz 6.11. *Sei $(\mathcal{X}, d)$ ein polnischer metrischer Raum, wobei d eine beschränkte Metrik sei, und sei $p \in [1, \infty)$. Dann metrisiert W_p die schwache Topologie und $(\mathcal{P}(\mathcal{X}), W_p)$ ist auch ein polnischer metrischer Raum.*

Der Beweis orientiert sich an [Vil09, Satz 6.9 und Folgerung 6.13], wobei sich der Aufwand durch die Annahme der Beschränktheit von d stellenweise verringert. Außerdem ist noch ein letzter Hilfssatz vonnöten, in dessen Beweis die obige Dualitätsformel eingeht.

Hilfssatz 6.12. *In der Situation von Satz 6.11 sei $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$ eine Cauchyfolge bezüglich W_p. Dann ist $\{\, \mu_k \mid k \in \mathbb{N} \,\}$ straff.*

Beweis. Sei $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$ eine Cauchyfolge bezüglich W_p, also

$$\lim_{k \to \infty} \sup_{l \geq k} W_p(\mu_k, \mu_l) \;=\; 0.$$

Wegen $W_1 \leq W_p$ ist $(\mu_k)_k$ dann auch insbesondere bezüglich W_1 eine Cauchyfolge.

Fixiere nun ein beliebiges $\varepsilon > 0$. Dann gibt es ein derartiges $n \in \mathbb{N}$, dass

$$W_1(\mu_n, \mu_k) \;<\; \varepsilon^2$$

für alle $k \geq n$ gilt. Offenbar können wir damit insbesondere auch für jedes $k \in \mathbb{N}$ ein $j \in \{1, \ldots, n\}$ mit $W_1(\mu_j, \mu_k) < \varepsilon^2$ finden. Wegen Satz 5.10 ist $\{\mu_1, \ldots, \mu_n\}$ straff, also existiert ein Kompaktum $K \subseteq \mathcal{X}$ so, dass

$$\mu_j(K) \;\geq\; 1 - \varepsilon, \quad j \in \{1, \ldots, n\}$$

gilt. Da K kompakt ist, existieren endlich viele $x_1, \ldots, x_m \in \mathcal{X}$ mit $K \subseteq U := \bigcup_{i=1}^m B_\varepsilon(x_i)$. Außerdem schreiben wir im Folgenden

$$U_\varepsilon \;:=\; \bigcup_{i=1}^m B_{2\varepsilon}(x_i) \;\supseteq\; \{\, x \in \mathcal{X} \mid d(x, U) < \varepsilon \,\}$$

und setzen

$$\varphi\colon \mathcal{X} \to \mathbb{R},\; x \mapsto \max\left(1 - \frac{d(x, U)}{\varepsilon}, 0\right).$$

Offensichtlich ist $\mathbb{1}_U \leq \varphi \leq \mathbb{1}_{U_\varepsilon}$ und φ ist $1/\varepsilon$-lipschitzstetig.

Für $k \in \mathbb{N}$ sei nun $j \in \{1, \ldots, n\}$ mit $W_1(\mu_j, \mu_k) < \varepsilon^2$. Da $\varepsilon\varphi$ 1-lipschitzstetig ist, liefert die Dualitätsformel von Kantorovich-Rubinstein (Satz 6.10)

$$W_1(\mu_j, \mu_k) \;\geq\; \varepsilon \left(\int \varphi \,\mathrm{d}\mu_k - \int \varphi \,\mathrm{d}\mu_j \right).$$

Also können wir abschätzen:

$$\begin{aligned}
\mu_k(U_\varepsilon) \;=\; \int \mathbb{1}_{U_\varepsilon}\,\mathrm{d}\mu_k \;&\geq\; \int \varphi\,\mathrm{d}\mu_k \;=\; \int \varphi\,\mathrm{d}\mu_j + \left(\int \varphi\,\mathrm{d}\mu_k - \int \varphi\,\mathrm{d}\mu_j\right)\\
&\geq\; \int \varphi\,\mathrm{d}\mu_j - \frac{W_1(\mu_j,\mu_k)}{\varepsilon} \;\geq\; \mu_j(U) - \frac{W_1(\mu_j,\mu_k)}{\varepsilon}\\
&\geq\; \mu_j(K) - \frac{W_1(\mu_j,\mu_k)}{\varepsilon} \;\geq\; 1-\varepsilon - \frac{W_1(\mu_j,\mu_k)}{\varepsilon}\\
&\geq\; 1-\varepsilon-\frac{\varepsilon^2}{\varepsilon} \;=\; 1-2\varepsilon.
\end{aligned}$$

Da wir zu Beginn $\varepsilon > 0$ beliebig gewählt haben, können wir die gesamte Argumentation analog für $\varepsilon/2^{(l+1)}$, $l \in \mathbb{N}$ führen und erhalten damit für jedes $l \in \mathbb{N}$ eine endliche Menge $\left\{x_1^{(l)}, \dots, x_{m_l}^{(l)}\right\} \subseteq \mathcal{X}$ mit

$$\mu_k\left(\bigcup_{i=1}^{m_l} B_{\varepsilon/2^l}(x_i^{(l)})\right) \;\geq\; 1-\frac{\varepsilon}{2^l}, \quad k \in \mathbb{N}.$$

Nun können wir analog zu den Beweisen von Satz 5.10 und Satz 5.12 vorgehen: Wir setzen

$$S_\varepsilon \;:=\; \bigcap_{l\in\mathbb{N}} \bigcup_{i=1}^{m_l} B_{\varepsilon/2^l}(x_i^{(l)}) \quad \text{und} \quad K_\varepsilon := \overline{S_\varepsilon}.$$

Da S_ε totalbeschränkt ist, ist K_ε kompakt (hier geht wieder die Vollständigkeit von $(\mathcal{X}, d)$ ein) und für alle $k \in \mathbb{N}$ gilt

$$\mu_k(K_\varepsilon) \;\geq\; \mu_k(K_\varepsilon) \;\geq\; 1-\varepsilon,$$

also ist $\{\, \mu_k \mid k \in \mathbb{N}\,\}$ straff. □

Beweis von Satz 6.11. Wegen Satz 5.6 ist lediglich zu beweisen, dass $\mathcal{P}(X)$ vollständig durch W_p metrisiert wird.

Zuerst zeigen wir, dass $\mathcal{P}(X)$ von W_p metrisiert wird. Es genügt, die Äquivalenz

$$\mu_k \xrightarrow{w} \mu \quad\Longleftrightarrow\quad W_p(\mu_k,\mu) \to 0, \quad k \to \infty$$

für jede Folge $(\mu_k)_k \in \mathcal{P}(X)^{\mathbb{N}}$ und $\mu \in \mathcal{P}(X)$ nachzuweisen. Für beide Richtungen verwenden wir die elementare Tatsache, dass für $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$ und $\mu \in \mathcal{P}(\mathcal{X})$ die Äquivalenz

$$\mu_k \xrightarrow{w} \mu \quad\Longleftrightarrow\quad \begin{array}{l}\text{jede Teilfolge } (\mu_{k'})_{k'} \text{ von } (\mu_k)_k \text{ hat}\\ \text{eine Teilfolge } (\mu_{k''})_{k''} \text{ mit } \mu_{k''} \xrightarrow{w} \mu\end{array} \tag{6.1}$$

gilt.

Wir möchten zunächst die Hinrichtung zeigen. Sei dazu $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$ und $\mu \in \mathcal{P}(\mathcal{X})$ mit $\mu_k \xrightarrow{w} \mu$. Sei außerdem $\pi_k \in \Pi(\mu_k, \mu)$ für jedes $k \in \mathbb{N}$ eine optimale Kopplung von μ_k mit sich selbst. Insbesondere sind dann $A := \{\, \mu_k \mid k \in \mathbb{N}\,\} \cup \{\mu\}$ und $B := \{\mu\}$ kompakt und Satz 6.7 liefert, dass

$$\hat{\Pi}(A,B) \;:=\; \{\, \pi \in \Pi(A,B) \mid \pi \text{ ist eine optimale Kopplung}\,\} \;\subseteq\; \mathcal{P}(\mathcal{X}\times\mathcal{X})$$

kompakt ist. Sei $(\pi_{k'})_{k'}$ nun eine Teilfolge von $(\pi_k)_k$. Da π_k für jedes $k \in \mathbb{N}$ in $\hat{\Pi}(A,B)$ liegt, hat $(\pi_{k'})_{k'}$ eine Teilfolge $(\pi_{k''})_{k''}$, die schwach gegen ein $\pi' \in \hat{\Pi}(A,B)$ konvergiert. Wegen Hilfssatz 6.3 muss π' eine optimale Kopplung von μ und μ sein, also

$$\int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p\,\pi'(\mathrm{d}(x,y)) \;=\; W_p(\mu,\mu)^p \;=\; 0.$$

Daher gilt $\pi'(\{\,(x,x) \mid x \in \mathcal{X}\,\}) = 1$ und für alle $A, B \in \mathcal{B}(\mathcal{X})$ folgt $\pi'(A \times B) = \mu(A \cap B)$, weshalb π' einfach die triviale Kopplung π von μ und μ ist. Weil $\{\, \pi_k \mid k \in \mathbb{N}\,\} \subseteq \hat{\Pi}(A,B)$ gilt, ist $\overline{\{\, \pi_k \mid k \in \mathbb{N}\,\}}$

kompakt und $\{ \pi_k \mid k \in \mathbb{N} \}$ damit nach dem Satz von Prokhorov (Satz 5.12) straff. Insgesamt impliziert (6.1) nun also die schwache Konvergenz

$$\pi_k \xrightarrow{w} \pi, \quad k \to \infty.$$

Da die Zuordnung $(x, y) \mapsto d(x, y)^p$ wegen der Beschränktheit von d aber in $C_b(\mathcal{X} \times \mathcal{X})$ enthalten ist, folgt

$$W_p(\mu_k, \mu)^p \;=\; \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi_k(\mathrm{d}(x,y)) \;\to\; \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y)) \;=\; 0, \quad k \to \infty,$$

also auch $W_p(\mu_k, \mu) \to 0$.

Für die Rückrichtung sei nun $(\mu_k)_k \in \mathcal{P}(\mathcal{X})^{\mathbb{N}}$ und $\mu \in \mathcal{P}(\mathcal{X})$ so, dass $W_p(\mu_k, \mu) \to 0$. Nach Hilfssatz 6.12 ist $\{ \mu_k \mid k \in \mathbb{N} \}$ dann straff, nach dem Satz von Prokhorov ist $\overline{\{ \mu_k \mid k \in \mathbb{N} \}}$ also kompakt. Sei $(\mu_{k'})_{k'}$ nun eine Teilfolge von $(\mu_k)_k$. Dann existiert eine derartige Teilfolge $(\mu_{k''})_{k''}$ von $(\mu_{k'})_{k'}$ und ein Wahrscheinlichkeitsmaß $\mu'' \in \mathcal{P}(\mathcal{X})$, dass

$$\mu_{k''} \xrightarrow{w} \mu'', \quad k'' \to \infty. \tag{6.2}$$

Für $k'' \in \mathbb{N}$ sei $\pi_{k''}$ eine optimale Kopplung von $\mu_{k''}$ und μ. Da $\{ \pi_{k''} \mid k'' \in \mathbb{N} \}$ nach Hilfssatz 6.2 straff ist, liefert der Satz von Prokhorov die Existenz einer Teilfolge $(\pi_{k'''})_{k'''}$ von $(\pi_{k''})_{k''}$, die schwach gegen ein $\pi''' \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ konvergiert. Nach Hilfssatz 6.3 ist π''' eine Kopplung von μ'' und μ. Durch eine Anwendung von Hilfssatz 6.6 können wir wie folgt abschätzen:

$$\begin{aligned} W_p(\mu'', \mu)^p \;&=\; \inf_{\pi \in \Pi(\mu'', \mu)} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi(\mathrm{d}(x,y)) \\ &\le\; \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi'''(\mathrm{d}(x,y)) \\ &\le\; \liminf_{k''' \to \infty} \int_{\mathcal{X}\times\mathcal{X}} d(x,y)^p \, \pi_{k'''}(\mathrm{d}(x,y)) \;=\; \liminf_{k''' \to \infty} W_p(\mu_{k'''}, \mu)^p \;=\; 0. \end{aligned}$$

Da dies $\mu'' = \mu$ impliziert, gilt wegen (6.1) und (6.2) also $\mu_k \xrightarrow{w} \mu,\ k \to \infty$.

Es bleibt zu zeigen, dass $(\mathcal{P}(\mathcal{X}), W_p)$ vollständig ist. Sei dazu $(\mu_k)_k$ eine Cauchyfolge in $\mathcal{P}(\mathcal{X})$ bezüglich W_p. Nach Hilfssatz 6.12 ist $\{ \mu_k \mid k \in \mathbb{N} \}$ straff. Der Satz von Prokhorov liefert nun ein $\mu \in \mathbb{N}$ und eine Teilfolge $(\mu_{k'})_{k'}$ von $(\mu_k)_k$ mit $\mu_{k'} \xrightarrow{w} \mu$ und damit auch $W_p(\mu_{k'}, \mu) \to 0,\ k' \to \infty$. Also ist auch $W_p(\mu_k, \mu) \to 0,\ k \to \infty$ und $(\mathcal{P}(\mathcal{X}), W_p)$ vollständig. $\square$

Literatur

[AB06] Charalambos D. Aliprantis und Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Third Edition. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2006. ISBN: 978-3-540-29587-7.

[Hen16] Norbert Henze. *Wahrscheinlichkeitstheorie*. 2016. URL: `https://www.math.kit.edu/stoch/~henze/media/wt-ss2016-handout-final.pdf`.

[Jor14] Jordan Bell. *Polish spaces and Baire spaces*. 2014. URL: `http://individual.utoronto.ca/jordanbell/notes/polish.pdf`.

[Meg98] Robert E. Megginson. *An Introduction to Banach Space Theory*. Bd. 183. Graduate Texts in Mathematics. New York, NY: Springer, 1998. ISBN: 0-387-98431-3.

[Mon81] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais*. Imprimerie Royale, 1781.

[Sim15] Barry Simon. *Real Analysis*. Bd. Part 1. A Comprehensive Course in Analysis. Providence, Rhode Island: AMS American Mathematical Society, 2015. ISBN: 978-1-4704-1099-5.

[van03] Onno van Gaans. *Probability measures on metric spaces*. 2002/03. URL: `https://www.math.leidenuniv.nl/~vangaans/jancol1.pdf`.

[Vee58] Veeravalli S. Varadarajan. „Weak Convergence of Measures on Separable Metric Spaces“. In: *Sankhyā: The Indian Journal of Statistics* Vol. 19, No. 1/2 (1958), S. 15–22. URL: `https://www.jstor.org/stable/25048364`.

[Vil09] Cédric Villani. *Optimal Transport: Old and New*. Bd. 338. Grundlehren der mathematischen Wissenschaften. Berlin: Springer, 2009. ISBN: 978-3-540-71049-3.