Chapter 1

Probability Measures on Metric Spaces

1.1 Tight measures

Let (E, d) denote a metric space, $\mathcal{O}(E)$, $\mathcal{A}(E)$, $\mathcal{K}(E)$ the systems of open, closed and compact subsets of E respectively. On (E, d) we have the notions of the *Borel σ-algebra*

$$\mathfrak{B}(E) := \sigma(\mathcal{O}(E)) = \sigma(\mathcal{A}(E))$$

of E and of a (Borel) measure on E, i.e. a non-negative extended real-valued σ-additive set function μ on $\mathfrak{B}(E)$ with the properties that $\mu(\emptyset) = 0$ and $\mu(K) < \infty$ for all $K \in \mathcal{K}(E)$.

Definition 1.1.1 A finite measure μ on E is called

(a) **regular** if for every $B \in \mathfrak{B}(E)$ and for every $\varepsilon > 0$ there exist $A \in \mathcal{A}(E)$ and $O \in \mathcal{O}(E)$ such that $A \subset B \subset O$ and $\mu(O) - \mu(A) < \varepsilon$,

and

(b) **tight** if

$$\mu(E) = \sup\{\mu(K) : K \in \mathcal{K}(E)\}\,.$$

Theorem 1.1.2 *Let μ be a finite measure on E. Then*

(i) *μ is regular.*
(ii) *If μ is tight then it must be* ***inner-regular*** *in the sense that for each $B \in \mathfrak{B}(E)$*

$$\mu(B) = \sup\{\mu(K) : K \in \mathcal{K}(E),\ \ K \subset B\}\,.$$

In particular, for finite measures the notions of tightness and inner-regularity coincide.

STRUCTURAL ASPECTS IN THE THEORY OF PROBABILITY - (Second Enlarged Edition)

Proof. (i). Let $\mathfrak{D} := \mathfrak{D}_\mu$ be the system of all $B \in \mathfrak{B}(E)$ with respect to which μ is regular. Then $\mathfrak{D}$ is a Dynkin system in the sense that $E \in \mathfrak{D}$, $B \in \mathfrak{D}$ implies that $\complement B \in \mathfrak{D}$, and whenever $(B_n)_{n\geqslant 1}$ is a disjoint sequence in $\mathfrak{D}$ then $B := \bigcup_{n\geqslant 1} B_n \in \mathfrak{D}$.

The proof of the first property is clear, and for the second one we observe that if $A \in \mathcal{A}(E)$ and $O \in \mathcal{O}(E)$ are chosen as in Definition 1.1.1 (a) then $\complement O \subset \complement B \subset \complement A$ and, noting that $\complement O \in \mathcal{A}(E)$ and $\complement A \in \mathcal{O}(E)$, we have

$$\mu(\complement A) - \mu(\complement O) = \mu(O) - \mu(A) < \varepsilon\,.$$

As for the third property, given $\varepsilon > 0$ we can find $A_n \in \mathcal{A}(E)$ and $O_n \in \mathcal{O}(E)$ with $A_n \subset B_n \subset O_n$ and

$$\mu(O_n) - \mu(A_n) < \frac{1}{2^{n+2}}\,\varepsilon$$

for all $n \in \mathbb{N}$. Let $O := \bigcup_{n\geqslant 1} O_n$, choose n_0 with $\mu\left(\bigcup_{n>n_0} A_n\right) < \varepsilon/4$ and put $A := \bigcup_{n=1}^{n_0} A_n$. Then $A \in \mathcal{A}(E)$, $O \in \mathcal{O}(E)$, $A \subset B \subset O$ and

$$\begin{aligned}\mu(O \setminus A) &\leqslant \sum_{n\geqslant 1} \mu(O_n \setminus A) \leqslant \sum_{n=1}^{n_0} \mu(O_n \setminus A) + \sum_{n>n_0} \mu(O_n)\\ &\leqslant \frac{\varepsilon}{4} + \sum_{n>n_0} \left(\mu(A_n) + \frac{1}{2^{n+2}}\,\varepsilon\right) \leqslant \frac{3}{4}\,\varepsilon \,<\, \varepsilon\,.\end{aligned}$$

Furthermore $\mathcal{A}(E) \subset \mathfrak{D}$. Indeed, given $A \in \mathcal{A}(E)$ for each $n \in \mathbb{N}$ we observe that

$$A^{\frac{1}{n}} := \left\{x \in E : d(x, A) < \frac{1}{n}\right\}$$

is open, and from $A^{\frac{1}{n}} \downarrow A$ (which holds as E is metric) it follows that $\mu(A^{\frac{1}{n}}) \downarrow \mu(A)$.

Now $\mathcal{A}(E)$ is $\cap$-stable, and therefore

$$\mathfrak{B}(E) = \sigma(\mathcal{A}(E)) = \mathfrak{D}(\mathcal{A}(E)) \subset \mathfrak{D} \subset \mathfrak{B}(E)\,,$$

whence $\mathfrak{D} = \mathfrak{B}(E)$. Here $\mathfrak{D}(\mathcal{A}(E))$ denotes the Dynkin hull of $\mathcal{A}(E)$.

(ii). Let $B \in \mathfrak{B}(E)$ and $\varepsilon > 0$. Using (i) there exists $A \in \mathcal{A}(E)$ with $A \subset B$ such that $\mu(B) - \mu(A) < \varepsilon/2$, and also $K \in \mathcal{K}(E)$ with $\mu(E) - \mu(K) < \varepsilon/2$. Then $A \cap K$ is a compact subset of B, and

$$\mu(B) - \mu(A \cap K) \leqslant \mu(B \setminus A) + \mu(\complement K) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

□

Corollary 1.1.3 *If* μ *is tight then for every downward filtered family* $(A_\iota)_{\iota\in I}$ *in* $\mathcal{A}(E)$

$$\mu\left(\bigcap_{\iota\in I} A_\iota\right) = \inf_{\iota\in I}\mu(A_\iota)\,.$$

Proof. From $A := \bigcap_{\iota\in I} A_\iota \subset A_\kappa$ we have $\mu(A) \leqslant \mu(A_\kappa)$ for all $\kappa \in I$, and hence $\mu(A) \leqslant \inf_{\iota\in I}\mu(A_\iota)$.

In the reverse direction, appealing to Theorem 1.1.2 (ii), to each $\varepsilon > 0$ there exists $K \in \mathcal{K}(E)$ with $K \subset \complement A$ such that

$$\mu(\complement K) - \mu(A) = \mu(\complement A) - \mu(K) < \varepsilon\,.$$

Now $K \subset \bigcup_{\iota\in I}\complement A_\iota$, and hence by compactness there exist $\iota_1, \iota_2, \ldots, \iota_n \in I$ with $K \subset \bigcup_{i=1}^{n}\complement A_{\iota_i}$. Also $(\complement A_\iota)_{\iota\in I}$ is an upward filtered family, and hence there exists $\iota_0 \in I$ such that $K \subset \complement A_{\iota_0}$. From $\mu(\complement K) - \mu(A) < \varepsilon$ it follows that

$$\mu(A_{\iota_0}) \leqslant \mu(\complement K) < \mu(A) + \varepsilon$$

and so ε being arbitrary we obtain $\inf_{\iota\in I}\mu(A_\iota) \leqslant \mu(A)$. □

Theorem 1.1.4 *Let* μ *be a tight measure on* E.

(i) *There exists a smallest closed subset* A_0 *of* E *with* $\mu(A_0) = \mu(E)$.
(ii) A_0 *is separable.*
(iii) $A_0 = \{x \in E : \mu(U) > 0$ *for all open neighbourhoods* U *of* $x\}$.

Proof. (i). The family

$$\{A \in \mathcal{A}(E) : \mu(A) = \mu(E)\}$$

is downward filtered, even $\cap$-stable. The result now follows from Corollary 1.1.3.

(ii). By Theorem 1.1.2 (ii) there exists a sequence $(K_n)_{n\geqslant 1}$ of compact and hence separable subsets of A_0 with $\mu(A_0) = \sup_{n\geqslant 1}\mu(K_n)$. Thus $A := (\bigcup_{n\geqslant 1} K_n)^-$ is separable and closed with $A \subset A_0$, from which it follows that

$$\mu(A) = \mu(A_0) = \mu(E)$$

and by (i), $A = A_0$ so that A_0 must be separable.

(iii). Write

$$B_0 := \{x \in E : \mu(U) > 0 \text{ for all open neighbourhoods } U \text{ of } x\}\,.$$

Given $x \in \complement A_0$ then $\complement A_0$ is an open neighbourhood of x with $\mu(\complement A_0) = 0$, and hence $x \in \complement B_0$. In the reverse direction given $x \in \complement B_0$ there exists an open neighbourhood U of x with $\mu(U) = 0$. Hence $\mu(\complement U) = \mu(E)$. Thus $A_0 \subset \complement U$ and hence $x \in \complement A_0$. □

Definition 1.1.5 The set A_0 in Theorem 1.1.4 is called the **support** of μ and will be denoted by $\text{supp}(\mu)$.

Theorem 1.1.6 *Let (E, d) be a separable complete metric space. Then every finite measure μ on E is tight.*

Proof. Let $\{x_k : k \in \mathbb{N}\}$ be a dense subset of E. Then for each $n \in \mathbb{N}$

$$\bigcup_{k \geqslant 1} B\left(x_k, \frac{1}{n}\right)^- = E\,,$$

where

$$B(x, \delta) := \{y \in E : d(x, y) < \delta\}$$

is the open ball of radius $\delta > 0$ with centre x. Choose $\varepsilon > 0$. Then to each $n \in \mathbb{N}$ there exists $k_n \in \mathbb{N}$ satisfying

$$\mu\left(E \setminus \bigcup_{k=1}^{k_n} B\left(x_k, \frac{1}{n}\right)^-\right) \leqslant \frac{\varepsilon}{2^n}\,.$$

The set

$$K := \bigcap_{n \geqslant 1} \bigcup_{k=1}^{k_n} B\left(x_k, \frac{1}{n}\right)^-$$

is closed and totally bounded. From the completeness of E it follows that K is compact. Finally

$$\mu(\complement K) = \mu\left(\bigcup_{n \geqslant 1}\left(E \setminus \bigcup_{k=1}^{k_n} B(x_k, \frac{1}{n})^-\right)\right) \leqslant \sum_{n \geqslant 1} \frac{\varepsilon}{2^n} = \varepsilon\,.$$

□

Theorem 1.1.7 *Let E, F be metric spaces, and $\varphi : E \to F$ a continuous mapping. If μ is a tight measure on E then the image measure $\varphi(\mu)$ of μ under φ is tight on F.*

Proof. Since φ is a continuous mapping it must be $\mathfrak{B}(E)$–$\mathfrak{B}(F)$-measurable, and hence $\varphi(\mu)$ is a finite measure on F. Given $\varepsilon > 0$ there exists a compact subset K of E with $\mu(\complement K) < \varepsilon$. Also $\varphi(K)$ is a compact subset of F, and

$$\begin{aligned}\varphi(\mu)(\complement\varphi(K)) &= \mu(\varphi^{-1}(\complement\varphi(K))) \\ &= \mu(\complement\varphi^{-1}(\varphi(K))) \\ &\leqslant \mu(\complement K) < \varepsilon\,.\end{aligned}$$

□

1.2 The topology of weak convergence

Although in the following discussion the set $M^{\mathrm{b}}(E)$ of all tight (finite Borel) measures on E and its subset $M^1(E) := \{\mu \in M^{\mathrm{b}}(E) : \mu(E) = 1\}$ of **probability measures** will remain the basic measure-theoretic objects, for some technical arguments we need a few facts on regular normed contents on E and related integrals. A **content** on E is a non-negative extended real-valued (finitely) additive set function μ on the algebra $\mathfrak{A}(\mathcal{O}(E))$ generated by $\mathcal{O}(E)$ satisfying $\mu(\emptyset) = 0$. Regular (finite) contents and probability contents on E are introduced in analogy to regular (finite) and probability measures on E.

Given a regular finite content μ on E, the μ-integral of a bounded real-valued function f on E is defined as follows. Let $\mathcal{P}$ be a partition of E consisting of finitely many pairwise disjoint sets $E_1, \ldots, E_n \in \mathfrak{A}(\mathcal{O}(E))$. We put

$$S_{\mathcal{P}} := \sum_{j=1}^{n} M_j \mu(E_j)$$

and

$$s_{\mathcal{P}} := \sum_{j=1}^{n} m_j \mu(E_j)\,,$$

where $M_j := \sup\{f(x) : x \in E_j\}$ and $m_j := \inf\{f(x) : x \in E_j\}$ for $j = 1, \ldots, n$. f is said to be **μ-integrable** if

$$\inf_{\mathcal{P}} S_{\mathcal{P}} = \sup_{\mathcal{P}} s_{\mathcal{P}}\,,$$

and in this case

$$\int f \,\mathrm{d}\mu := \inf_{\mathcal{P}} S_{\mathcal{P}}$$

is the **μ-integral** of f. Obviously every bounded continuous function f is μ-integrable, and

$$f \mapsto \int f \,\mathrm{d}\mu$$

defines a normed positive linear functional on the vector space $C^{\mathrm{b}}(E)$ of bounded continuous functions on E. Moreover, we have

Theorem 1.2.1 (F. Riesz)
There is a one-to-one correspondence

$$\mu \leftrightarrow L_\mu$$

between the set of regular finite (probability) contents μ *on* E *and the set of bounded (normed) positive linear functionals* L_μ *on* $C^{\mathrm{b}}(E)$ *given by*

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

for all $f \in C^{\mathrm{b}}(E)$.

Proof. The proof will be carried out only for the case in parentheses.
1. Let L be a normed positive linear functional on $C^{\mathrm{b}}(E)$, and let

$$\lambda_L(A) := \inf\{L(f) : f \in C^{\mathrm{b}}(E),\, f \geqslant \mathbb{1}_A\}$$

for every $A \in \mathcal{A}(E)$. $\lambda_L : \mathcal{A}(E) \to [0,1]$ is a smooth probability content in the sense of the following four properties

(a) $\lambda_L(\emptyset) = 0$, $\lambda_L(E) = 1$.
(b) $\lambda_L(A_1) \leqslant \lambda_L(A_2)$ for all $A_1, A_2 \in \mathcal{A}(E)$ with $A_1 \subset A_2$.
(c) $\lambda_L(A_1 \cup A_2) \leqslant \lambda_L(A_1) + \lambda_L(A_2)$ for all $A_1, A_2 \in \mathcal{A}(E)$, where equality holds whenever $A_1 \cap A_2 = \emptyset$.
(d) For all $A \in \mathcal{A}(E)$

$$\lambda_L(A) = \inf\{\lambda_L(O^-) : O \in \mathcal{O}(E),\, O \supset A\}\,.$$

2. Now, λ_L can be uniquely extended to a regular probability content $\mu_L : \mathfrak{A}(\mathcal{O}(E)) \to [0,1]$, and it turns out that

$$L(f) = \int f \,\mathrm{d}\mu_L$$

for all $f \in C^{\mathrm{b}}(E)$.

In order to verify this identity we pick $f \in C^{\mathrm{b}}(E)$ with $0 \leqslant f \leqslant 1$ and introduce the sets

$$G_i := \left\{x \in E : f(x) > \frac{i}{n}\right\} \in \mathcal{O}(E)$$

for all $i = 0, 1, \dots, n$, $n \geqslant 1$. Clearly, $G_0 \supset G_1 \supset \cdots \supset G_n = \emptyset$. Now we define functions $\alpha_i \in C\big([0,1]\big)$ by

$$\alpha_i := \begin{cases} \equiv 0 & \text{on } \left[0, \dfrac{i-1}{n}\right] \\ \text{linear} & \text{on } \left[\dfrac{i-1}{n}, \dfrac{i}{n}\right] \\ \equiv 1 & \text{on } \left[\dfrac{i}{n}, 1\right] \end{cases}$$

and functions f_i on E by

$$f_i := \alpha_i \circ f$$

for $i = 0, 1, \ldots, n,\ n \geqslant 1$. Then

$$\frac{1}{n}\sum_{i=1}^{n} \alpha_i(t) = t$$

for all $t \in [0,1]$, hence

$$\frac{1}{n}\sum_{i=1}^{n} f_i = f \qquad \text{and} \qquad \frac{1}{n}\sum_{i=1}^{n} L(f_i) = L(f)\,.$$

Since $f_i \geqslant \mathbb{1}_{G_i}$, and for any $A \in \mathcal{A}(E)$, $A \subset G_i$, $\mathbb{1}_{G_i} \geqslant \mathbb{1}_A$ we obtain that $f_i \geqslant \mathbb{1}_A$ and hence that

$$L(f_i) \geqslant \lambda_L(A) = \mu_L(A)\,.$$

From the regularity of μ_L we infer that

$$L(f_i) \geqslant \mu_L(G_i)$$

and thus

$$\begin{aligned} L(f) &\geqslant \frac{1}{n}\sum_{i=1}^{n} \mu_L(G_i) \;=\; \sum_{i=1}^{n}\left(\frac{i}{n} - \frac{i-1}{n}\right)\mu_L(G_i) \\ &= \sum_{i=1}^{n-1} \frac{i}{n}\left(\mu_L(G_i) - \mu_L(G_{i+1})\right) \\ &= \left(\sum_{i=1}^{n-1} \frac{i+1}{n}\mu_L(G_i \setminus G_{i+1})\right) - \frac{1}{n}\mu_L(G_1) \\ &\geqslant \left(\sum_{i=1}^{n-1} \int\limits_{G_i \setminus G_{i+1}} f \,\mathrm{d}\mu_L\right) - \frac{1}{n}\mu_L(G_1) \\ &= \int\limits_{G_1} f\,\mathrm{d}\mu_L - \frac{1}{n}\mu_L(G_1) \;\geqslant\; \int\limits_{E} f\,\mathrm{d}\mu_L - \frac{1}{n}\,. \end{aligned}$$

For $n \to \infty$ we obtain that

$$L(f) \geqslant \int f\,\mathrm{d}\mu_L$$

whenever $f \in C^{\mathrm{b}}(E)$ with $0 \leqslant f \leqslant 1$.

But since $f \in C_+(E)$ there exists a constant $c > 0$ such that $0 \leqslant cf \leqslant 1$, hence

$$L(f) = \frac{1}{c} L(cf) \geqslant \frac{1}{c} \int cf \, \mathrm{d}\mu_L = \int f \, \mathrm{d}\mu_L \,.$$

Moreover, if $f \in C^{\mathrm{b}}(E)$, there exists a constant $c_1 > 0$ satisfying $f + c_1 \geqslant 0$, hence

$$L(f) = L(f + c_1) - c_1 \geqslant \int (f + c_1) \, \mathrm{d}\mu_L - c_1 = \int f \, \mathrm{d}\mu_L \,.$$

Thus we have

$$L(f) \geqslant \int f \, \mathrm{d}\mu_L$$

for all $f \in C^{\mathrm{b}}(E)$. Replacing f by $-f$ yields the assertion.

3. The injectivity of the correspondence $\mu \mapsto L_\mu$ can be seen as follows: Let μ, ν be regular probability contents of E satisfying

$$\int f \, \mathrm{d}\mu = \int f \, \mathrm{d}\nu$$

for all $f \in C^{\mathrm{b}}(E)$, and let $A \in \mathcal{A}(E)$. There exist decreasing sequences $(G_n)_{n \geqslant 1}$ and $(H_n)_{n \geqslant 1}$ in $\mathcal{O}(E)$ with $G_n \supset A$ and $H_n \supset A$ for all $n \geqslant 1$ such that

$$\lim_{n \to \infty} \mu(G_n) = \mu(A)$$

and

$$\lim_{n \to \infty} \nu(H_n) = \nu(A) \,.$$

But then $V_n := G_n \cap H_n \downarrow A$ and

$$\lim_{n \to \infty} \mu(V_n) = \mu(A)$$

as well as

$$\lim_{n \to \infty} \nu(V_n) = \nu(A) \,.$$

Choosing for every $n \geqslant 1$ a function $f_n \in C^{\mathrm{b}}(E)$ with the properties $0 \leqslant f_n \leqslant 1$, $f_n(A) = \{1\}$ and $f_n\big(\complement V_n\big) = \{0\}$ (the existence of which follows from $A \cap \complement V_n = \emptyset$ for all $n \geqslant 1$) we obtain

$$\int f_n \, \mathrm{d}\mu = \int_A f_n \, \mathrm{d}\mu + \int_{V_n \setminus A} f_n \, \mathrm{d}\mu = \mu(A) + \int_{V_n \setminus A} f_n \, \mathrm{d}\mu$$

and

$$\int_{V_n \setminus A} f_n \, d\mu \leqslant \mu(V_n \setminus A) = \mu(V_n) - \mu(A) \,,$$

hence

$$\lim_{n\to\infty} \int_{V_n \setminus A} f_n \, d\mu = 0 \,.$$

Therefore

$$\lim_{n\to\infty} \int f_n \, d\mu = \mu(A)$$

and

$$\lim_{n\to\infty} \int f_n \, d\nu = \nu(A) \,,$$

thus

$$\mu(A) = \nu(A) \quad \text{for all} \quad A \in \mathcal{A}(E)$$

and by the regularity of μ, ν also

$$\mu(B) = \nu(B) \quad \text{for all} \quad B \in \mathfrak{A}(\mathcal{A}(E)) = \mathfrak{A}(\mathcal{O}(E))$$

which implies that $\mu = \nu$. □

At a later stage we will apply the following consequences of the theorem.

Corollary 1.2.2 *If for measures* $\mu, \nu \in M^{\mathrm{b}}(E)$

$$\int f \, d\mu = \int f \, d\nu$$

holds whenever $f \in C^{\mathrm{b}}(E)$, *then* $\mu = \nu$ *(on* $\mathfrak{B}(E)$*).*

Corollary 1.2.3 *Let* (E, d) *be a compact metric space. There is a one-to-one correspondence*

$$\mu \leftrightarrow L_\mu$$

between the set $M^{\mathrm{b}}(E)$ *and the set* $L^1_+(C(E))$ *of positive normed linear functionals on* $C(E))$ *given by*

$$L_\mu(f) = \int f \, d\mu$$

for all $f \in C(E) = C^{\mathrm{b}}(E)$.

Proof. The proof follows directly from Theorem 1.2.1 by applying the fact that for compact E every regular finite content on E is in fact σ-additive and hence uniquely extendable to a measure in $M^{\mathrm{b}}(E)$. For the latter property see Theorem 1.3.1. □

We proceed to introducing a topology in $M^{\mathrm{b}}(E)$.

Definition 1.2.4 Given $\mu \in M^{\mathrm{b}}(E)$, $n \geqslant 1$, $f_1, f_2, \ldots, f_n \in C^{\mathrm{b}}(E)$ and $\varepsilon > 0$, define

$$V(\mu; f_1, f_2, \ldots, f_n; \varepsilon) := \left\{ \nu \in M^{\mathrm{b}}(E) : \left| \int f_i \, \mathrm{d}\mu - \int f_i \, \mathrm{d}\nu \right| < \varepsilon \text{ for all } i = 1, 2, \ldots, n \right\}.$$

The **weak topology** τ_{w} on $M^{\mathrm{b}}(E)$ is the uniquely determined topology for which

$$\left\{ V(\mu; f_1, f_2, \ldots, f_n; \varepsilon) : n \geqslant 1,\ f_1, f_2, \ldots, f_n \in C^{\mathrm{b}}(E),\ \varepsilon > 0 \right\}$$

is a neighbourhood system of μ for each $\mu \in M^{\mathrm{b}}(E)$.

Remark 1.2.5

(a) The weak topology on $M^{\mathrm{b}}(E)$ is Hausdorff due to Corollary 1.2.2.

(b) A net $(\mu_\iota)_{\iota \in I}$ in $M^{\mathrm{b}}(E)$ **converges weakly** (τ_{w}) to $\mu \in M^{\mathrm{b}}(E)$ whenever

$$\lim_{\iota \in I} \int f \, \mathrm{d}\mu_\iota = \int f \, \mathrm{d}\mu$$

for all $f \in C^{\mathrm{b}}(E)$; we write $\tau_{\mathrm{w}}\text{–}\lim_\iota \mu_\iota = \mu$.

(c) In the functional-analytic context of Appendix B.10 one introduces for the dual pair $(C^{\mathrm{b}}(E)', C^{\mathrm{b}}(E))$ of topological vector spaces the weak topology on $C^{\mathrm{b}}(E)'$. If E is compact, then Corollary 1.2.3 yields the homeomorphism

$$C^{\mathrm{b}}(E)'_+ \cong M^{\mathrm{b}}(E)$$

and consequently the coincidence of the weak topology restricted to $C^{\mathrm{b}}(E)'_+$ with the weak topology τ_{w} on $M^{\mathrm{b}}(E)$.

Definition 1.2.6 Let $\mu \in M^{\mathrm{b}}(E)$. A set $B \in \mathfrak{B}(E)$ is called a μ-**continuity** (μ-null boundary) set if $\mu(\partial B) = 0$ where $\partial B := B^- \setminus B^0$ $(\in \mathcal{A}(E))$.

Theorem 1.2.7 (Portemanteau theorem)
Let $(\mu_\iota)_{\iota \in I}$ be a net in $M^{\mathrm{b}}(E)$ and $\mu \in M^{\mathrm{b}}(E)$. The following statements are equivalent:

(i) $\tau_{\mathrm{w}}\text{-}\lim_\iota \mu_\iota = \mu$.
(ii) $\lim_{\iota \in I} \mu_\iota(E) = \mu(E)$ *and* $\limsup_{\iota \in I} \mu_\iota(A) \leqslant \mu(A)$ *for all* $A \in \mathcal{A}(E)$.
(iii) $\lim_{\iota \in I} \mu_\iota(E) = \mu(E)$ *and* $\liminf_{\iota \in I} \mu_\iota(O) \geqslant \mu(O)$ *for all* $O \in \mathcal{O}(E)$.
(iv) $\lim_{\iota \in I} \mu_\iota(B) = \mu(B)$ *for all μ-continuity sets B.*

Proof. (i) $\Rightarrow$ (ii). As $\mathbb{1}_E \in C^{\mathrm{b}}(E)$ we have $\lim_{\iota \in I} \mu_\iota(E) = \mu(E)$. Now consider $A \in \mathcal{A}(E)$. Then, as $A^{1/n} \downarrow A$ as $n \to \infty$, to each $\varepsilon > 0$ we can find $n \in \mathbb{N}$ with $\mu(A^{1/n}) - \mu(A) < \varepsilon$. Choose $f \in C^{\mathrm{b}}(E)$ with $0 \leqslant f \leqslant 1$, $f(A) = \{1\}$ and $f(\complement A^{1/n}) = \{0\}$. Then

$$\limsup_{\iota \in I} \mu_\iota(A) \leqslant \limsup_{\iota \in I} \int f \,\mathrm{d}\mu_\iota \leqslant \mu(A^{1/n}) < \mu(A) + \varepsilon$$

and hence

$$\limsup_{\iota \in I} \mu_\iota(A) \leqslant \mu(A)\,.$$

(ii) $\Leftrightarrow$ (iii). This follows by considering complements.

(ii), (iii) $\Rightarrow$ (iv). Let B be a μ-continuity set. Then

$$\begin{aligned}\limsup_{\iota \in I} \mu_\iota(B) \leqslant \limsup_{\iota \in I} \mu_\iota(B^-) \ &\leqslant\ \mu(B^-) \\ &= \mu(B^0)\ \leqslant\ \liminf_{\iota \in I} \mu_\iota(B^0)\ \leqslant\ \liminf_{\iota \in I} \mu_\iota(B)\end{aligned}$$

and this yields the result.

(iv) $\Rightarrow$ (i). Let $f \in C^{\mathrm{b}}(E)$. Since $f(\mu)$ contains at most countably many atoms, to each $\varepsilon > 0$ there exists a strictly increasing sequence $(t_i)_{i=0,1,\ldots,k}$ in $\mathbb{R}$ with $f(\mu)(\{t_i\}) = 0$ for all $i = 0, 1, \ldots, k$, $t_i - t_{i-1} \leqslant \varepsilon$ for all $i = 1, 2, \ldots, k$, and $f(E) \subset [t_0, t_k[$. For each $i = 1, 2, \ldots, k$ put $B_i := f^{-1}([t_{i-1}, t_i[\,)$. Then $B_i \in \mathfrak{B}(E)$ and, since

$$\partial B_i \subset f^{-1}(\partial[t_{i-1}, t_i[\,) = f^{-1}(\{t_{i-1}, t_i\})\,,$$

we see that B_i is a μ-continuity set. We now define

$$g := \sum_{i=1}^{k} t_{i-1} \mathbb{1}_{B_i} \qquad \text{and} \qquad h := \sum_{i=1}^{k} t_i \mathbb{1}_{B_i}\,.$$

Then

$$g \leqslant f \leqslant g + \varepsilon \qquad \text{and} \qquad h - \varepsilon \leqslant f \leqslant h$$

and

$$\limsup_{\iota\in I}\int f\,\mathrm{d}\mu_\iota \leqslant \limsup_{\iota\in I}\int g\,\mathrm{d}\mu_\iota + \varepsilon\mu(E)$$
$$= \int g\,\mathrm{d}\mu + \varepsilon\mu(E) \leqslant \int f\,\mathrm{d}\mu + \varepsilon\mu(E)$$

and

$$\liminf_{\iota\in I}\int f\,\mathrm{d}\mu_\iota \geqslant \liminf_{\iota\in I}\int h\,\mathrm{d}\mu_\iota - \varepsilon\mu(E)$$
$$= \int h\,\mathrm{d}\mu - \varepsilon\mu(E) \geqslant \int f\,\mathrm{d}\mu - \varepsilon\mu(E)$$

as

$$\int g\,\mathrm{d}\nu = \sum_{i=1}^{k} t_{i-1}\nu(B_i) \qquad \text{and} \qquad \int h\,\mathrm{d}\nu = \sum_{i=1}^{k} t_i\nu(B_i)$$

for all $\nu \in M^{\mathrm{b}}(E)$, and E is a μ-continuity set as $\partial E = \emptyset$. It now follows that

$$\limsup_{\iota\in I}\int f\,\mathrm{d}\mu_\iota \leqslant \int f\,\mathrm{d}\mu \leqslant \liminf_{\iota\in I}\int f\,\mathrm{d}\mu_\iota$$

and this gives the desired equality. □

Corollary 1.2.8 *Let* $\mu \in M^{\mathrm{b}}(E)$. *Then each of the following sets is a* τ_{w}*-neighbourhood basis of* μ.

(a) $\{\nu \in M^{\mathrm{b}}(E) : |\nu(E) - \mu(E)| < \varepsilon$ *and* $\nu(A_i) < \mu(A_i) + \varepsilon$ *for all* $i = 1, 2, \ldots, n\}$, *where* $A_1, A_2, \ldots, A_n \in \mathcal{A}(E)$, $n \in \mathbb{N}$ *and* $\varepsilon > 0$.
(b) $\{\nu \in M^{\mathrm{b}}(E) : |\nu(E) - \mu(E)| < \varepsilon$ *and* $\nu(O_i) > \mu(O_i) - \varepsilon$ *for all* $i = 1, 2, \ldots, n\}$, *where* $O_1, O_2, \ldots, O_n \in \mathcal{O}(E)$, $n \in \mathbb{N}$ *and* $\varepsilon > 0$.
(c) $\{\nu \in M^{\mathrm{b}}(E) : |\nu(B_i) - \mu(B_i)| < \varepsilon$ *for all* $i = 1, 2, \ldots, n\}$, *where* $B_1, B_2, \ldots, B_n \in \mathfrak{B}(E)$ *are* μ*-continuity sets with* $n \in \mathbb{N}$ *and* $\varepsilon > 0$.

Theorem 1.2.9 *Let* $(\mu_\iota)_{\iota\in I}$ *be a net in* $M^{\mathrm{b}}(E)$ *with* $\tau_{\mathrm{w}}\text{-}\lim_\iota \mu_\iota = \mu \in M^{\mathrm{b}}(E)$. *Furthermore let* f *be a bounded Borel-measurable real-valued function on* E. *If the set* D_f *of discontinuity points of* f *is a* μ*-null set, then*

$$\lim_{\iota\in I}\int f\,\mathrm{d}\mu_\iota = \int f\,\mathrm{d}\mu\,.$$

Proof. Let $A \in \mathcal{A}(\mathbb{R})$. Since $f^{-1}(A)^- \subset D_f \cup f^{-1}(A)$ we can apply Theorem 1.2.7 to obtain

$$\begin{aligned}\limsup_{\iota\in I} f(\mu_\iota)(A) &= \limsup_{\iota\in I} \mu_\iota(f^{-1}(A)) \leqslant \limsup_{\iota\in I} \mu_\iota(f^{-1}(A)^-) \\ &\leqslant \mu(f^{-1}(A)^-) \leqslant \mu(D_f \cup f^{-1}(A)) = \mu(f^{-1}(A)) = f(\mu)(A)\,.\end{aligned}$$

In addition

$$\lim_{\iota\in I} f(\mu_\iota)(\mathbb{R}) = \lim_{\iota\in I} \mu_\iota(E) = \mu(E) = f(\mu)(\mathbb{R})\,.$$

A second application of Theorem 1.2.7 gives $\tau_{\mathrm{w}}\text{–}\lim_\iota f(\mu_\iota) = f(\mu)$. Now consider $\varphi \in C^{\mathrm{b}}(\mathbb{R})$ such that $\mathrm{Res}_B\,\varphi = \mathrm{id}_B$, where B is any bounded interval containing the bounded set $f(B)$. Since $\varphi \circ f = f$ we have

$$\lim_{\iota\in I} \int f\,\mathrm{d}\mu_\iota = \lim_{\iota\in I} \int \varphi\,\mathrm{d}f(\mu_\iota) = \int \varphi\,\mathrm{d}f(\mu) = \int f\,\mathrm{d}\mu\,.$$

□

Corollary 1.2.10 *Let $(\mu_\iota)_{\iota\in I}$ be a net in $M^{\mathrm{b}}(E)$ satisfying τ_{w}–$\lim_\iota \mu_\iota = \mu \in M^{\mathrm{b}}(E)$, and let $B \in \mathfrak{B}(E)$ be a μ-continuity set. Then for the corresponding measures induced on B we have $\tau_{\mathrm{w}}\text{–}\lim_\iota(\mu_\iota)_B = \mu_B$.*

Proof. Let $f \in C^{\mathrm{b}}(E)$. Then

$$\int f\,\mathrm{d}\nu_B = \int f 1\!\!1_B\,\mathrm{d}\nu$$

for all $\nu \in M^{\mathrm{b}}(E)$. From $D_{f1\!\!1_B} \subset \partial B$ we see that $D_{f1\!\!1_B}$ is a μ-null set. In addition $f1\!\!1_B$ is bounded and Borel measurable. Referring to Theorem 1.2.9 it follows that

$$\lim_{\iota\in I} \int f\,\mathrm{d}(\mu_\iota)_B = \lim_{\iota\in I} \int f1\!\!1_B\,\mathrm{d}\mu_\iota = \int f\,\mathrm{d}\mu_B\,.$$

□

Theorem 1.2.11 *The set $D(E) := \{\varepsilon_x : x \in E\}$ of Dirac measures on E is τ_{w}-closed in $M^{\mathrm{b}}(E)$, and $x \mapsto \varepsilon_x$ is a homeomorphism of E onto $D(E)$.*

Proof. Let $(x_\iota)_{\iota\in I}$ be a net in E such that $\lim_{\iota\in I} x_\iota = x \in E$. Then

$$\lim_{\iota\in I} \int f\,\mathrm{d}\varepsilon_{x_\iota} = \lim_{\iota\in I} f(x_\iota) = f(x) = \int f\,\mathrm{d}\varepsilon_x$$

for all $f \in C^{\mathrm{b}}(E)$, and thus $\tau_{\mathrm{w}}\text{–}\lim_\iota \varepsilon_{x_\iota} = \varepsilon_x$ which shows that $x \mapsto \varepsilon_x$ is a continuous mapping $E \to D(E)$. In addition $x \mapsto \varepsilon_x$ is injective, which is easily seen by simply choosing $B \in \mathfrak{B}(E)$ with $x \in B$ and $y \notin B$, and indeed $B = \{x\}$ will suffice.

Now suppose that $\tau_{\mathrm{w}}-\lim_\iota \varepsilon_{x_\iota} = \mu \in M^{\mathrm{b}}(E)$. From $\lim_{\iota\in I} \varepsilon_{x_\iota}(E) = \mu(E)$ we see that $\mu(E) = 1$, and hence $\operatorname{supp}(\mu) \neq \emptyset$. Choose $x \in \operatorname{supp}(\mu)$ and an open neighbourhood U of x. It follows from the properties of $\operatorname{supp}(\mu)$ and Theorem 1.2.7 that

$$\liminf_{\iota\in I} \varepsilon_{x_\iota}(U) \geqslant \mu(U) > 0$$

so that there exists $\iota_U \in I$ with $\varepsilon_{x_{\iota_U}}(U) > 0$, and hence $x_\iota \in U$ for all $\iota > \iota_U$. This shows that $\lim_{\iota\in I} x_\iota = x$, and in particular that $\varepsilon_x \mapsto x$ is continuous. It follows that $\mu = \varepsilon_x$, since τ_{w} is Hausdorff, and therefore $D(E)$ is τ_{w}-closed. □

Theorem 1.2.12 *Let E, F be metric spaces, and $\varphi : E \to F$ a continuous mapping. Then $\mu \mapsto \varphi(\mu)$ is a τ_{w}-continuous mapping from $M^{\mathrm{b}}(E)$ into $M^{\mathrm{b}}(F)$.*

Proof. According to Theorem 1.1.7 we see that $\varphi(\mu) \in M^{\mathrm{b}}(F)$ for all $\mu \in M^{\mathrm{b}}(E)$. Let $(\mu_\iota)_{\iota\in I}$ be a net in $M^{\mathrm{b}}(E)$ with $\tau_{\mathrm{w}}-\lim_\iota \mu_\iota = \mu \in M^{\mathrm{b}}(E)$. For each $f \in C^{\mathrm{b}}(F)$ we have $f \circ \varphi \in C^{\mathrm{b}}(E)$, and this implies that

$$\lim_{\iota\in I} \int f \,\mathrm{d}\varphi(\mu_\iota) = \lim_{\iota\in I} \int f \circ \varphi \,\mathrm{d}\mu_\iota = \int f \circ \varphi \,\mathrm{d}\mu = \int f \,\mathrm{d}\varphi(\mu)\,.$$ □

In the following we will show that we can restrict the study of weak convergence to bounded sequences. For this purpose we consider the metrizability of the space $M^{\mathrm{b}}(E)$ for an arbitrary metric space (E, d).

Lemma 1.2.13 *The mapping $\rho : M^{\mathrm{b}}(E) \times M^{\mathrm{b}}(E) \to \mathbb{R}_+$ given by*

$$\varrho(\mu,\nu) = \inf\bigl\{\varepsilon > 0 : \mu(B) \leqslant \nu(B^\varepsilon) + \varepsilon \text{ and } \nu(B) \leqslant \mu(B^\varepsilon) + \varepsilon \text{ for all } B \in \mathfrak{B}(E)\bigr\}$$

*for all $\mu, \nu \in M^{\mathrm{b}}(E)$ is a metric on $M^{\mathrm{b}}(E)$, called the **Prokhorov metric**. (It should be noted that the existence of numbers $\varepsilon > 0$ satisfying the above is guaranteed by the boundedness of μ, ν.)*

Proof. It is clear that $\rho(\mu,\nu) \geqslant 0$, $\rho(\mu,\nu) = \rho(\nu,\mu)$ and $\rho(\mu,\mu) = 0$ for all $\mu, \nu \in M^{\mathrm{b}}(E)$. Now suppose $\rho(\mu,\nu) = 0$. For each $A \in \mathcal{A}(E)$ we know that $A^{1/n} \downarrow A$. Then $\mu(A) = \nu(A)$ and hence by Theorem 1.1.2 (i), $\mu = \nu$.

Finally, to prove that ρ satisfies the triangle inequality, consider $\lambda, \mu, \nu \in M^{\mathrm{b}}(E)$ and $\alpha, \beta > 0$ with $\rho(\lambda,\mu) < \alpha$ and $\rho(\mu,\nu) < \beta$. By definition of $\rho(\mu,\nu)$ we have

$$\mu(B) \leqslant \nu(B^\beta) + \beta$$

for all $B \in \mathfrak{B}(E)$. From

$$(B^\alpha)^\beta \cup (B^\beta)^\alpha \subset B^{\alpha+\beta}$$

it follows that

$$\lambda(B) \leqslant \mu(B^\alpha) + \alpha \leqslant \nu((B^\alpha)^\beta) + \beta + \alpha \leqslant \nu(B^{\alpha+\beta}) + (\alpha+\beta)$$

and correspondingly

$$\nu(B) \leqslant \lambda(B^{\alpha+\beta}) + (\alpha+\beta)\,.$$

Thus $\rho(\lambda,\nu) \leqslant \alpha+\beta$. Now since $\alpha > \rho(\lambda,\mu)$ and $\beta > \rho(\mu,\nu)$ were chosen arbitrarily it follows that

$$\rho(\lambda,\nu) \leqslant \rho(\lambda,\mu) + \rho(\mu,\nu)\,.$$

□

Theorem 1.2.14 *The Prokhorov metric induces the weak topology* τ_w *on* $M^\mathrm{b}(E)$.

Proof. In each τ_w-neighbourhood U of $\mu \in M^\mathrm{b}(E)$ there is an open ρ-ball centred in μ. Without loss of generality we may assume U to be chosen as in Corollary 1.2.8 (a). Now since μ is continuous from above and $A_j^\delta \downarrow A_j$ as $\delta \downarrow 0$ there exists $\delta \in]0, \varepsilon/2[$ such that

$$\mu(A_j^\delta) < \mu(A_j) + \frac{\varepsilon}{2} \quad \text{for all} \quad j = 1,2,\ldots,n\,.$$

Choose $\nu \in M^\mathrm{b}(E)$ satisfying $\rho(\mu,\nu) < \delta$. Since $E^\delta = E$ it follows that

$$|\nu(E) - \mu(E)| \leqslant \delta < \varepsilon\,.$$

Furthermore

$$\nu(A_j) \leqslant \mu(A_j^\delta) + \delta < \mu(A_j) + \frac{\varepsilon}{2} + \delta < \mu(A_j) + \varepsilon$$

for all $j = 1,2,\ldots,n$, and this implies that $\nu \in U$.

For the reverse direction it is to be shown that each open ρ-ball centred on μ with radius $\varepsilon > 0$ contains a τ_w-neighbourhood U. Let $\delta \in]0, \varepsilon/4[$. Since μ is tight, by Theorem 1.1.4 there exists a σ-compact set $G \subset E$ with $\mu(G) = \mu(E)$. To each $x \in G$ there corresponds $\delta(x) \in]0, \delta/2[$ with $\mu(\partial B(x,\delta(x))) = 0$. Note that

$$\partial B(x,\eta) \subset \{y \in E : d(x,y) = \eta\}\,.$$

For each $n \in \mathbb{N}$ there exist finitely many $\eta > 0$ such that $\mu(\partial B(x,\eta)) \geqslant 1/n$. Thus there exist only countably many η with $\mu(\partial B(x,\eta)) \neq 0$.

Since $G \subset \bigcup_{x\in G} B(x, \delta(x))$, the σ-compactness of G yields the existence of a sequence $(x_n)_{n\geqslant 1}$ with

$$G \subset \bigcup_{n\geqslant 1} B(x_n, \delta(x_n))\,.$$

For each $n \in \mathbb{N}$ put $G_n := B(x_n, \delta(x_n))$. Then $(G_n)_{n\geqslant 1}$ is a sequence of μ-continuity sets with diameter less than δ, and

$$G \subset \bigcup_{n\geqslant 1} G_n\,.$$

Hence there exists $k \in \mathbb{N}$ such that

$$\mu\left(\bigcup_{n=1}^{k} G_n\right) > \mu(E) - \delta\,.$$

Put

$$\mathfrak{C} := \left\{\bigcup_{j\in J} G_j : J \subset \{1, 2, \ldots, k\}\right\}.$$

Now $\mathfrak{C}$ is a finite system of μ-continuity sets since

$$\partial\left(\bigcup_{j\in J} G_j\right) \subset \bigcup_{j\in J} G_j^- \setminus \bigcup_{j\in J} G_j \subset \bigcup_{j\in J} \partial G_j\,.$$

It follows from Corollary 1.2.8 that

$$U := \{\nu \in M^{\mathrm{b}}(E) : |\nu(E) - \mu(E)| < \delta,\, |\nu(C) - \mu(C)| < \delta \text{ for all } C \in \mathfrak{C}\}$$

is a τ_{w}-neighbourhood of μ.

We wish to show that for each $\nu \in U$ we have $\rho(\mu, \nu) < \varepsilon$. Put $C_0 := \bigcup_{n=1}^{k} G_n$. Since $C_0 \in \mathfrak{C}$ we have

$$\nu(C_0) > \mu(C_0) - \delta > \mu(E) - 2\delta > \nu(E) - 3\delta\,.$$

Now choose $B \in \mathfrak{B}(E)$. Then

$$C := \bigcup_{\{j\leqslant k\,:\, G_j\cap B\neq\emptyset\}} G_j \in \mathfrak{C}\,,$$

$C \subset B^\delta$ (as diam $G_j < \delta$) and $B \subset C \cup \complement C_0$ (where $C \cap \complement C_0 = \emptyset$). Then

$$\mu(B) \leqslant \mu(C) + \mu(\complement C_0) < \nu(C) + \delta + \delta \leqslant \nu(B^\delta) + 2\delta < \nu(B^{4\delta}) + 4\delta$$

and analogously

$$\nu(B) \leqslant \nu(C) + \nu(\complement C_0) < \mu(C) + \delta + 3\delta \leqslant \mu(B^{4\delta}) + 4\delta\,.$$

Thus

$$\rho(\mu, \nu) \leqslant 4\delta < \varepsilon\,.$$

□

Application 1.2.15 *Let* $(X_n)_{n\geqslant 0}$ *be a sequence of* E*-valued random variables on a probability space* $(\Omega, \mathfrak{A}, \mathsf{P})$ *with distributions* $\mathsf{P}_{X_n} \in M^1(E)$ *for* $n \geqslant 0$. *Then*

$$X_n \to X_0 \quad \mathsf{P}\text{-}stochastically$$

implies

$$\mathsf{P}_{X_n} \to \mathsf{P}_{X_0} \quad weakly\ as \ \ n \to \infty.$$

In fact, given $\varepsilon > 0$ there exists $n_\varepsilon \geqslant 1$ such that

$$\mathsf{P}([d(X_n, X_0) \geqslant \varepsilon]) < \varepsilon \quad \text{for all} \quad n \geqslant n_\varepsilon\,.$$

Let $B \in \mathfrak{B}(E)$. Then

$$\begin{aligned}[X_n \in B] &\subset [d(X_n, X_0) \geqslant \varepsilon] \cup \big([d(X_n, X_0) < \varepsilon] \cap [X_0 \in B^\varepsilon]\big)\\ &\subset [d(X_n, X_0) \geqslant \varepsilon] \cup [X_0 \in B^\varepsilon]\end{aligned}$$

and analogously,

$$[X_0 \in B] \subset [d(X_n, X_0) \geqslant \varepsilon] \cup [X_n \in B^\varepsilon]\,.$$

For $n \geqslant n_\varepsilon$ follows

$$\mathsf{P}_{X_n}(B) \leqslant \varepsilon + \mathsf{P}_{X_0}(B^\varepsilon)$$

as well as

$$\mathsf{P}_{X_0}(B) \leqslant \varepsilon + \mathsf{P}_{X_n}(B^\varepsilon)$$

whenever $B \in \mathfrak{B}(E)$. But this implies

$$\rho(\mathsf{P}_{X_n}, \mathsf{P}_{X_0}) < \varepsilon\,,$$

thus Theorem 1.2.14 yields the assertion.

1.3 The Prokhorov theorem

As before (E, d) is a metric space, and the space $M^{\mathrm{b}}(E)$ is given the topology τ_{w}, or equivalently the Prokhorov metric ρ. Hence $(M^{\mathrm{b}}(E), \rho)$ is a metric space.

A finite content μ on E is called **inner-regular** if

$$\mu(B) = \sup\{\mu(K) : K \in \mathcal{K}(E),\ \ K \subset B\}$$

for all $B \in \mathfrak{B}(E)$.

Theorem 1.3.1 *Every inner-regular content μ on E is σ-additive, so that $\mu \in M^{\mathrm{b}}(E)$.*

Proof. All that needs to be shown is that μ is $\emptyset$-continuous. Let $(B_n)_{n\geqslant 1}$ be a sequence in $\mathfrak{B}(E)$ with $B_n \downarrow \emptyset$, and ε be given. For each $n \in \mathbb{N}$ there exists a compact set $K_n \subset B_n$ such that

$$\mu(B_n) - \mu(K_n) < \frac{1}{2^n}\,\varepsilon\,.$$

Put $L_n := \bigcap_{i=1}^n K_i$. Then

$$\mu(B_n) - \mu(L_n) \leqslant \sum_{i=1}^{n}(\mu(B_i) - \mu(K_i)) < \varepsilon$$

for all $n \in \mathbb{N}$, and $L_n \downarrow \emptyset$. From the finite intersection property there exists $n_0 \in \mathbb{N}$ such that $L_n = \emptyset$ for all $n \geqslant n_0$, and hence $\mu(L_n) = 0$ for all $n \geqslant n_0$. Then $\mu(B_n) < \varepsilon$ for all $n \geqslant n_0$, and we have shown that $\lim_{n\to\infty} \mu(B_n) = 0$. □

Corollary 1.3.2 *Let $(\mu_n)_{n\geqslant 1}$ be an increasing sequence in $M^{\mathrm{b}}(E)$ satisfying*

$$\sup_{n\geqslant 1} \mu_n(E) < \infty\,.$$

Then $\sup_{n\geqslant 1} \mu_n \in M^{\mathrm{b}}(E)$.

Proof. Write $\mu(B) := \sup_{n\geqslant 1} \mu_n(B)$ for all $B \in \mathfrak{B}(E)$. Clearly μ is a finite content on $\mathfrak{B}(E)$. From Theorem 1.1.2 we have that μ_n is inner-regular, and so μ is an inner-regular content, and the result follows from Theorem 1.3.1. □

Theorem 1.3.3 *Let (E, d) be a compact metric space. Then for each $a > 0$*

$$M^{(a)}(E) := \{\mu \in M^{\mathrm{b}}(E) : \mu(E) \leqslant a\}$$

is τ_{w}-compact.

Proof. According to Appendix B.11 (Alaoglu, Bourbaki), the set

$$V^{(a)} := \{\varphi \in C^{\mathrm{b}}(E)' : \|\varphi\| \leqslant a\}$$

is weakly compact, i.e. compact with respect to the topology $\sigma(C^{\mathrm{b}}(E)', C^{\mathrm{b}}(E))$. Hence

$$V_+^{(a)} := \{\varphi \in V^{(a)} : \varphi \geqslant 0\} = \bigcap_{f \in C_+^b(E)} \{\varphi \in V^{(a)} : \varphi(f) \geqslant 0\}$$

is weakly compact. Furthermore by Corollary 1.2.3 the mapping

$$\mu \mapsto \left(f \mapsto \int f \,\mathrm{d}\mu\right)$$

is a bijection from $M^{(a)}(E)$ onto $V_+^{(a)}$. From the definition of τ_{w} it follows that it is a homeomorphism, and hence that $M^{(a)}(E)$ is τ_{w}-compact. □

Definition 1.3.4 A set $H \subset M^{\mathrm{b}}(E)$ is called **uniformly tight** if

(a) $\sup\{\mu(E) : \mu \in H\} < \infty$;
(b) to each $\varepsilon > 0$ there exists a compact set $K \subset E$ such that

$$\mu(\complement K) < \varepsilon \quad \text{for all} \quad \mu \in H\,.$$

Theorem 1.3.5 *Suppose* (E, d) *is separable complete. For each* $H \subset M^{\mathrm{b}}(E)$ *the following statements are equivalent:*

(i) H *is uniformly tight.*
(ii) (a) $\sup\{\mu(E) : \mu \in H\} < \infty$*;*
(b) *For all* $\varepsilon > 0$ *and* $n \in \mathbb{N}$ *there exist* $x_1, x_2, \ldots, x_k \in E$ *such that*

$$\mu(\complement B_n) < \varepsilon \quad \textit{for all} \quad \mu \in H$$

where $B_n := \bigcup_{j=1}^k B\left(x_j, \frac{1}{n}\right)$.

Proof. (i) ⇒ (ii). To each $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $\mu(\complement K) < \varepsilon$ for all $\mu \in H$. Also to each $n \in \mathbb{N}$ there exist $x_1, x_2, \ldots, x_k \in E$ such that $K \subset B_n$, so that

$$\mu(\complement B_n) < \varepsilon$$

for all $\mu \in H$.

(ii) ⇒ (i). Let $\varepsilon > 0$. From the assumption for each $n \in \mathbb{N}$ there exists $B_n \subset E$ such that B_n is the finite union of $\frac{1}{n}$-balls and

$$\mu(\complement B_n) < \frac{1}{2^n}\,\varepsilon$$

for all $\mu \in H$. Put $L := \bigcap_{n\geqslant 1} B_n$. Then $\mu(\complement L) < \varepsilon$ for all $\mu \in H$. Now L is totally bounded, and hence so is L^-. As E is complete it follows that L^- is compact. The result now follows from the fact that

$$\mu(\complement L^-) \leqslant \mu(\complement L) < \varepsilon \quad \text{for all} \quad \mu \in H\,.$$

□

Lemma 1.3.6 *Let $A \in \mathcal{A}(E)$, $\mu_\iota \in M^{\mathrm{b}}(E)$ for all $\iota \in I$, $\mu \in M^{\mathrm{b}}(E)$, $\mu_A \in M^{\mathrm{b}}(A)$. If $\tau_{\mathrm{w}}-\lim_\iota \mu_\iota = \mu$ and $\tau_{\mathrm{w}}-\lim_\iota \mathrm{Res}_A\, \mu_\iota = \mu_A$ then $\mu_A \leqslant \mathrm{Res}_A\, \mu$.*

Proof. We can use Theorem 1.1.2 (ii) to deduce that $\mathrm{Res}_A\, \mu \in M^{\mathrm{b}}(A)$. Consider a continuous function $g : A \to [0,1]$. By Tietze's extension theorem there exists a continuous function $f : E \to [0,1]$ with $\mathrm{Res}_A\, f = g$. Then

$$\int g \,\mathrm{d}\mu_A = \lim_{\iota\in I} \int g \,\mathrm{d}(\mathrm{Res}_A\, \mu_\iota) = \lim_{\iota\in I} \int_A f \,\mathrm{d}\mu_\iota$$

and

$$\int g \,\mathrm{d}(\mathrm{Res}_A\, \mu) = \int_A f \,\mathrm{d}\mu\,.$$

But in slight modification of Corollary 1.2.2 one obtains that for measures $\mu, \nu \in M^{\mathrm{b}}(E)$ satisfying

$$\int f \,\mathrm{d}\mu \leqslant \int f \,\mathrm{d}\nu \quad \text{for all} \quad f \in C^{\mathrm{b}}(E) \quad \text{with} \quad 0 \leqslant f \leqslant 1$$

$\mu \leqslant \nu$ holds. Therefore we need only prove that

$$\lim_{\iota\in I} \int_A f \,\mathrm{d}\mu_\iota \leqslant \int_A f \,\mathrm{d}\mu\,.$$

However this inequality follows immediately from the Portemanteau theorem 1.2.7, as clearly $\tau_{\mathrm{w}}-\lim_\iota f\cdot\mu_\iota = f\cdot\mu$. □

Theorem 1.3.7 (Prokhorov)
Consider $H \subset M^{\mathrm{b}}(E)$.

(i) *If H is uniformly tight then H is τ_{w}-relatively compact.*
(ii) *If E is separable complete then any τ_{w}-relatively compact set H is uniformly tight.*

Proof. (i). Let $(K_n)_{n\geqslant 1}$ be an increasing sequence of compact subsets of E satisfying $\mu(\complement K_n) < 1/n$ for all $\mu \in H$. Let $(\mu_k)_{k\geqslant 1}$ be a sequence in H. We have to show that $(\mu_k)_{k\geqslant 1}$ possesses a τ_{w}-convergent subsequence. From

$$a := \sup\{\mu(K_n) : \mu \in H,\, n \in \mathbb{N}\} \leqslant \sup\{\mu(E) : \mu \in H\} < \infty$$

appealing to Theorem 1.3.3 we have that $\{\mathrm{Res}_{K_n} \mu : \mu \in H\}$ $(\subset M^{(a)}(K_n))$ is τ_{w}-relatively compact in $M^{\mathrm{b}}(K_n)$. By the metrizability of $M^{\mathrm{b}}(E)$, which is the content of Theorem 1.2.14, a diagonal argument provides for each $n \in \mathbb{N}$ a measure $\nu'_n \in M^{\mathrm{b}}(K_n)$ with

$$\tau_{\mathrm{w}}\text{–}\lim_k \mathrm{Res}_{K_n} \mu_k = \nu'_n\,.$$

Put

$$\nu_n(B) := \nu'_n(B \cap K_n) \quad \text{for all} \quad B \in \mathfrak{B}(E)\,.$$

Then $\nu_n \in M^{\mathrm{b}}(E)$ for all $n \in \mathbb{N}$. Also from Lemma 1.3.6, $\nu'_n \leqslant \mathrm{Res}_{K_n} \nu'_{n+1}$ and (recall that $K_n \subset K_{n+1}$)

$$\nu_n(B) \leqslant \nu'_{n+1}(B \cap K_n) \leqslant \nu_{n+1}(B)$$

for all $B \in \mathfrak{B}(E)$ which says that $(\nu_n)_{n\geqslant 1}$ is an increasing sequence. Moreover,

$$\nu_n(E) = \nu'_n(K_n) = \lim_{k\to\infty} \mu_k(K_n) \leqslant \liminf_{k\to\infty} \mu_k(E) < \infty$$

for all $n \in \mathbb{N}$. Now appeal to Corollary 1.3.2 to obtain

$$\nu := \sup_{n\geqslant 1} \nu_n \in M^{\mathrm{b}}(E)$$

and moreover

$$\nu(E) = \sup_{n\geqslant 1} \nu_n(E) \leqslant \liminf_{k\to\infty} \mu_k(E)\,.$$

Let $A \in \mathcal{A}(E)$. Applying the Portemanteau theorem 1.2.7 we get

$$\begin{aligned}\nu(A) \geqslant \nu_n(A) = \nu'_n(A \cap K_n) &\geqslant \limsup_{k\geqslant 1} \mu_k(A \cap K_n) \\ &\geqslant \limsup_{k>1} \mu_k(A) - \frac{1}{n}\end{aligned}$$

for all $n \in \mathbb{N}$, where for the second inequality we have used the fact that $A \cap K_n$ is closed, and for the third that

$$\mu(A \cap K_n) \geqslant \mu(A) - \frac{1}{n}$$

for all $\mu \in H$. Hence

$$\nu(A) \geqslant \limsup_{k\geqslant 1} \mu_k(A)$$

and

$$\nu(E) = \lim_{k\to\infty} \mu_k(E)\,.$$

A further application of the Portemanteau theorem 1.2.7 gives τ_{w}–$\lim_k \mu_k = \nu$. Thus every sequence in H has a τ_{w}-convergent subsequence, and this just says that H is τ_{w}-relatively compact.

(ii). Assume that H is τ_{w}-relatively compact and at the same time fails to be uniformly tight. The mapping

$$\mu \mapsto \mu(E) = \int \mathbb{1}_E \,\mathrm{d}\mu$$

is continuous, and this implies that

$$\sup\{\mu(E) : \mu \in H\} < \infty\,.$$

Let $\mathcal{F} = \{F \subset E : |F| < \infty\}$. Theorem 1.3.5 implies that there exist $\varepsilon > 0$ and $n \in \mathbb{N}$ such that to each $F \in \mathcal{F}$ there is $\mu_F \in H$ with

$$\mu_F\left(E \setminus \left(\bigcup_{x\in F} B\left(x, \frac{1}{n}\right)\right)\right) \geqslant \varepsilon\,.$$

The net $(\mu_F)_{F\in\mathcal{F}}$ contains a τ_{w}-convergent subnet $(\mu_{F_\iota})_{\iota\in I}$ with limit μ say. For each $\iota \in I$ define

$$B_\iota := \bigcup_{x\in F_\iota} B\left(x, \frac{1}{n}\right)\,.$$

Now $(\complement B_\iota)_{\iota\in I}$ is a downward filtered family in $\mathcal{A}(E)$ with $\bigcap_{\iota\in I} \complement B_\iota = \emptyset$. It follows from Corollary 1.1.3 that

$$\inf_{\iota\in I} \mu(\complement B_\iota) = 0\,.$$

On the other hand applying the Portemanteau theorem we have

$$\mu(\complement B_\kappa) \geqslant \limsup_{\iota\in I} \mu_{F_\iota}(\complement B_\kappa) \geqslant \limsup_{\iota\in I} \mu_{F_\iota}(\complement B_\iota) \geqslant \varepsilon$$

for all $\kappa \in I$, and this is a contradiction. □

1.4 Convolution of measures

In this section (E, d) will denote a separable complete metric Abelian group which means that E is an Abelian group (with binary operation denoted by addition and 0 as neutral element), (E, d) is a separable and complete metric space with distance function d, and the mapping $(x, y) \mapsto x - y$ from $E \times E$ into E is continuous. Along with (E, d), $(E \times E, d \times d)$ is also a separable complete metric Abelian group. Here the metric $d \times d$ is defined by $d \times d((x, y), (u, v)) := \max\{d(x, u), d(y, v)\}$ for all $(x, y), (u, v) \in E$. A prominent example of a separable complete metric Abelian group is a separable Banach space $(E, \|\cdot\|)$ over $\mathbb{R}$, where the distance function is given by $d(x, y) := \|x - y\|$ for all $x, y \in E$.

For separable complete metric groups E we have that each finite measure on E is tight (Theorem 1.1.6) and that the notions "uniform tightness" and "τ_{w}-relative compactness" are equivalent (Prokhorov's theorem 1.3.7).

Theorem 1.4.1 *Let (E, d) be a separable complete metric Abelian group.*

(i) $\mathfrak{B}(E \times E) = \mathfrak{B}(E) \otimes \mathfrak{B}(E)$.
(ii) *The mapping $(x, y) \mapsto m(x, y) := x + y$ from $E \times E$ into E is $(\mathfrak{B}(E) \otimes \mathfrak{B}(E), \mathfrak{B}(E))$-measurable.*

Proof. (i). $\mathcal{O}(E)$ is a generator of $\mathfrak{B}(E)$ with the exhaustion property which says that there exists a sequence $(O_n)_{n\geqslant 1}$ in $\mathcal{O}(E)$ such that $O_n \uparrow E$. Thus $\mathcal{O}(E) \times \mathcal{O}(E)$ is a generator of $\mathfrak{B}(E) \otimes \mathfrak{B}(E)$. Furthermore $\mathcal{O}(E) \times \mathcal{O}(E) \subset \mathcal{O}(E \times E)$ from which follows that $\mathfrak{B}(E) \otimes \mathfrak{B}(E) \subset \mathfrak{B}(E \times E)$. Now choose a countable dense subset D of E, so that $D \times D$ is a countable dense subset of $E \times E$. The open balls in E and $E \times E$ centred at points in D and $D \times D$ respectively and with rational radii make up countable bases $\mathcal{B}$ and $\mathcal{C}$ in $\mathcal{O}(E)$ and $\mathcal{O}(E \times E)$ respectively. Now

$$B((x, y), r) = B(x, r) \times B(y, r)$$

for the corresponding balls, and hence $\mathcal{B} \times \mathcal{B} \supset \mathcal{C}$. Now $\mathcal{B}$ and $\mathcal{C}$ are generators of $\mathfrak{B}(E)$ and $\mathfrak{B}(E \times E)$ respectively with the exhaustion property, which implies that $\mathfrak{B}(E) \otimes \mathfrak{B}(E) \supset \mathfrak{B}(E \times E)$.

(ii). The mapping m from $E \times E$ into E is continuous, and hence $(\mathfrak{B}(E \times E), \mathfrak{B}(E))$-measurable. Now apply (i). $\square$

Application 1.4.2 *Let X, Y be E-valued random variables on a probability space $(\Omega, \mathfrak{A}, \mathsf{P})$. Then by Theorem 1.4.1 (ii) the mapping*

$\omega \mapsto (X+Y)(\omega) = X(\omega) + Y(\omega)$ from Ω into E is also an E-valued random variable on $(\Omega, \mathfrak{A}, \mathsf{P})$, since $X + Y = m \circ (X, Y)$.

Definition 1.4.3 For $\mu, \nu \in M^{\mathrm{b}}(E)$ we refer to the measure $\mu * \nu := m(\mu \otimes \nu)$ on E as the **convolution** of μ and ν.

We have the following properties of the convolution mapping.

Properties 1.4.4

1.4.4.1 *For all $f \in C^{\mathrm{b}}(E)$*

$$\int f \, \mathrm{d}(\mu * \nu) = \int \left(\int f(x+y)\, \mu(\mathrm{d}x) \right) \nu(\mathrm{d}y) = \int \left(\int f(x+y)\, \nu(\mathrm{d}y) \right) \mu(\mathrm{d}x)$$

which follows from Fubini's theorem.

1.4.4.2 *In particular, for all $B \in \mathfrak{B}(E)$*

$$(\mu * \nu)(B) = \int \mu(B - y)\, \nu(\mathrm{d}y) = \int \nu(B - x)\, \mu(\mathrm{d}x)\,.$$

1.4.4.3 *For all $B, C \in \mathfrak{B}(E)$ we have $B + C \in \mathfrak{B}(E)$ and*

$$(\mu * \nu)(B + C) \geqslant \mu(B)\nu(C)$$

The latter can easily be seen from the inequalities

$$(\mu * \nu)(B + C) = (\mu \otimes \nu)(m^{-1}(B + C)) \geqslant (\mu \otimes \nu)(B \times C) = \mu(B)\nu(C)\,.$$

1.4.4.4 *The convolution is commutative and associative, that is*

$$\mu * \nu = \nu * \mu$$

and

$$(\lambda * \mu) * \nu = \lambda * (\mu * \nu)$$

for all $\lambda, \mu, \nu \in M^{\mathrm{b}}(E)$.

The proof uses 1.4.4.1 in conjunction with the injectivity of the mapping

$$\mu \mapsto \left(f \mapsto \int f \, \mathrm{d}\mu \right)$$

from $M^{\mathrm{b}}(E)$ into $C^{\mathrm{b}}(E)'_+$.

1.4.4.5 *For each $x \in E$ and $B \in \mathfrak{B}(E)$*

$$(\mu * \varepsilon_x)(B) = \mu(B - x)\,.$$

1.4.4.6 *From* 1.4.4.5 *and* 1.4.4.2 *we have $\mu * \varepsilon_0 = \mu$ and $\varepsilon_x * \varepsilon_y = \varepsilon_{x+y}$ for all $x, y \in E$.*

Application 1.4.5 *Let* X, Y *be independent* E*-valued random variables on a probability space* $(\Omega, \mathfrak{A}, \mathsf{P})$. *Then* $\mathsf{P}_{X+Y} = \mathsf{P}_X * \mathsf{P}_Y$.

Proof. To show this we note that independence of X, Y gives $\mathsf{P}_{(X,Y)} = \mathsf{P}_X \otimes \mathsf{P}_Y$ which implies that

$$\mathsf{P}_{X+Y} = \mathsf{P}_{m\circ(X,Y)} = m(\mathsf{P}_{(X,Y)}) = m(\mathsf{P}_X \otimes \mathsf{P}_Y) = \mathsf{P}_X * \mathsf{P}_Y \ .$$

□

Theorem 1.4.6 (Support formula)
For $\mu, \nu \in M^{\mathrm{b}}(E)$

$$\mathrm{supp}(\mu * \nu) = (\mathrm{supp}(\mu) + \mathrm{supp}(\nu))^- \,.$$

Proof. First we note that

$$\begin{aligned}(\mu * \nu)((\mathrm{supp}(\mu) + \mathrm{supp}(\nu))^-) &= (\mu \otimes \nu)(m^{-1}((\mathrm{supp}(\mu) + \mathrm{supp}(\nu))^-)) \\ &\geqslant (\mu \otimes \nu)(\mathrm{supp}(\mu) \times \mathrm{supp}(\nu)) \;=\; \mu(E)\nu(E) \\ &= (\mu \otimes \nu)(E \times E) \;=\; (\mu * \nu)(E)\,.\end{aligned}$$

Thus by Theorem 1.1.4 (i)

$$\mathrm{supp}(\mu * \nu) \subset (\mathrm{supp}(\mu) + \mathrm{supp}(\nu))^- \,.$$

To prove the reverse inclusion, consider $x \in \mathrm{supp}(\mu)$, $U \in \mathfrak{V}(x)$, $y \in \mathrm{supp}(\nu)$ and $V \in \mathfrak{V}(y)$. Then $U + V \in \mathfrak{V}(x + y)$. Now Property 1.4.4.3 together with Theorem 1.1.4 (iii) gives

$$(\mu * \nu)(U + V) \geqslant \mu(U)\mu(V) > 0\,.$$

To each $W \in \mathfrak{V}(x+y)$ there exist $U \in \mathfrak{V}(x)$, $V \in \mathfrak{V}(y)$ with $U+V \subset W$. A further application of Theorem 1.1.4 (iii) gives $x+y \in \mathrm{supp}(\mu+\nu)$. Thus

$$\mathrm{supp}(\mu) + \mathrm{supp}(\nu) \subset \mathrm{supp}(\mu * \nu)$$

and this completes the proof. □

Corollary 1.4.7 *If* $\mu * \nu$ *is a Dirac measure then so is each of the measures* μ *and* ν.

Proof. We just observe that a measure is Dirac precisely when it has a single element support. □

Theorem 1.4.8 *Let* $(\mu_n)_{n\geqslant 1}$, $(\nu_n)_{n\geqslant 1}$ *be sequences with* $\tau_{\mathrm{w}}\!-\!\lim_n \mu_n = \mu$, $\tau_{\mathrm{w}}\!-\!\lim_n \nu_n = \nu$ *in* $M^{\mathrm{b}}(E)$. *Then* $\tau_{\mathrm{w}}\!-\!\lim_n \mu_n \otimes \nu_n = \mu \otimes \nu$.

Proof. Appealing to Prokhorov's theorem 1.3.7 to each $\varepsilon > 0$ there exist $K \in \mathcal{K}(E)$ with

$$\mu_n(\complement K) < \varepsilon \qquad \text{and} \qquad \nu_n(\complement K) < \varepsilon$$

and $\alpha > 0$ such that

$$\mu_n(E) \leqslant \alpha \qquad \text{and} \qquad \nu_n(E) \leqslant \alpha$$

for all $n \in \mathbb{N}$. Clearly $K \times K \in \mathcal{K}(E \times E)$, and

$$\begin{aligned}(\mu_n \otimes \nu_n)(K \times K) = \mu_n(K)\nu_n(K) &\geqslant (\mu_n(E) - \varepsilon)(\nu_n(E) - \varepsilon) \\ &> (\mu_n \otimes \nu_n)(E \times E) - \varepsilon(\mu_n(E) + \nu_n(E)).\end{aligned}$$

Therefore

$$(\mu_n \otimes \nu_n)(\complement(K \times K)) < 2\alpha\varepsilon$$

and

$$(\mu_n \otimes \nu_n)(E \times E) = \mu_n(E)\nu_n(E) \leqslant \alpha^2$$

for all $n \in \mathbb{N}$. Furthermore from Prokhorov's theorem 1.3.7 we see that $\{\mu_n \otimes \nu_n : n \in \mathbb{N}\}$ is τ_{w}-relatively compact.

Let $\lambda \in M^{\mathrm{b}}(E \times E)$ be a cluster point of the sequence $(\mu_n \otimes \nu_n)_{n \geqslant 1}$, that is, there exists a subsequence $(\mu_{n_k} \otimes \nu_{n_k})_{k \geqslant 1}$ with limit λ. Now

$$\begin{aligned}\mathcal{E} := \{B_1 \times B_2 \in \mathfrak{B}(E) \times \mathfrak{B}(E) :& \\ &B_1 \times B_2 \text{ is both } \mu \otimes \nu\text{- and a } \lambda\text{-continuity set}\}\end{aligned}$$

is an $\cap$-stable generator of $\mathfrak{B}(E) \otimes \mathfrak{B}(E) = \mathfrak{B}(E \times E)$. Let $B_1 \times B_2 \in \mathcal{E}$. Then

$$\partial(B_1 \times B_2) = (B_1^- \times \partial B_2) \cup (\partial B_1 \times B_2^-).$$

Therefore we have the following three cases to consider.

(1) B_1 is a μ-continuity set and B_2 is a ν-continuity set.
(2) $\mu(B_1^-) = 0$.
(3) $\nu(B_2^-) = 0$.

With the help of the Portemanteau theorem 1.2.7 we have

$$\begin{aligned}\lambda(B_1 \times B_2) = \lim_{k \to \infty} (\mu_{n_k} \otimes \nu_{n_k})(B_1 \times B_2) &= \lim_{k \to \infty} \mu_{n_k}(B_1)\nu_{n_k}(B_2) \\ = \mu(B_1)\nu(B_2) &= (\mu \otimes \nu)(B_1 \times B_2),\end{aligned}$$

$$\lambda(B_1 \times B_2) \leqslant \lim_{k \to \infty} \mu_{n_k}(B_1)\alpha = \mu(B_1)\alpha = 0 = (\mu \otimes \nu)(B_1 \times B_2).$$

and similarly in the remaining case.

Thus

$$\lambda(B_1 \times B_2) = (\mu \otimes \nu)(B_1 \times B_2) \quad \text{for all} \quad B_1 \times B_2 \in \mathcal{E}$$

and hence $\lambda = \mu \otimes \nu$. Hence the τ_{w}-relatively compact sequence $(\mu_n \otimes \nu_n)_{n \geqslant 1}$ has a unique cluster point $\mu \otimes \nu$, and it follows that $\tau_{\mathrm{w}}-\lim_n \mu_n \otimes \nu_n = \mu \otimes \nu$. □

Theorem 1.4.9 *Let (E, d) be a separable complete metric Abelian group. Then the space $(M^{\mathrm{b}}(E), \tau_{\mathrm{w}}, *)$ with the convolution $*$ defined above is a commutative metric semigroup, in the sense that*

(i) $(M^{\mathrm{b}}(E), \tau_{\mathrm{w}})$ *is a metric space (with the Prokhorov metric inducing the topology τ_{w}),*
(ii) $(M^{\mathrm{b}}(E), *)$ *is a commutative semigroup with neutral element ε_0, that is, $\varepsilon_0 * \mu = \mu$ for all $\mu \in M^{\mathrm{b}}(E)$, and*
(iii) *the mapping $(\mu, \nu) \mapsto \mu * \nu$ from $M^{\mathrm{b}}(E) \times M^{\mathrm{b}}(E)$ into $M^{\mathrm{b}}(E)$ is τ_{w}-continuous.*

Moreover,

(iv) $(M^1(E), \tau_{\mathrm{w}}, *)$ *is a subsemigroup of $(M^{\mathrm{b}}(E), \tau_{\mathrm{w}}, *)$.*

Proof. In view of the assumed metrizability it suffices to consider sequences. So given sequences $(\mu_n)_{n \geqslant 1}$, $(\nu_n)_{n \geqslant 1}$ with $\tau_{\mathrm{w}}-\lim_n \mu_n = \mu$, $\tau_{\mathrm{w}}-\lim_n \nu_n = \nu$ in $M^{\mathrm{b}}(E)$ we have by Theorem 1.4.8 that $\tau_{\mathrm{w}}-\lim_n \mu_n \otimes \nu_n = \mu \otimes \nu$. Then Theorem 1.2.12 gives

$$\tau_{\mathrm{w}}-\lim_n \mu_n * \nu_n = \tau_{\mathrm{w}}-\lim_n m(\mu_n \otimes \nu_n) = m(\mu \otimes \nu) = \mu * \nu\,.$$

This proves the continuity of $(\mu, \nu) \mapsto \mu * \nu$. The remaining assertions follow immediately from Properties 1.4.4.2 and 1.4.4.4 of the convolution.

□