Chapter 2

Probability measures on topological spaces

As the duals of nuclear spaces are not metric spaces, to study Φ'-valued random variables or Φ'-valued stochastic processes we need to consider probability measures on general topological spaces. In Section 1 of this chapter, we first briefly recall some basic concepts about topological spaces. Then we establish some basic properties of Borel probability measures on general topological spaces. In Section 2 we study the weak convergence of Borel probability measures. In Section 3, we restrict ourselves to topological vector spaces and consider the Bochner functionals corresponding to cylinder measures. Finally in the last two sections we study two special topological spaces: $C([0,T],\Phi')$ and $D([0,T],\Phi')$ and probability measures. These two spaces will be our primary concern in the study of Φ'-valued stochastic processes with continuous sample paths and right-continuous sample paths respectively.

This chapter consists of basic material about probability measures on general topological vector spaces which we shall need in later chapters. For more detailed treatments we refer the reader to the books of Bilingsley [2], Ethier and Kurtz [9], Gel'fand and Vilenkin [12], Parthasarathy [43] and Xia [59]. Most of the material in Sections 4 and 5 is taken from Mitoma [41].

2.1 Probability measures on topological spaces

In this section we briefly present some basic concepts about topological spaces and consider probability measures on topological spaces. We shall study some special properties of the probability measures when the topology of the space has extra structure.

Now we define some special topological spaces.

Definition 2.1.1 *A topological space X is* **Hausdorff** *if $\forall x \neq y \in X$, there exist disjoint open sets G_1, G_2 such that $x \in G_1, y \in G_2$.*
X is **normal** *if*

i) $\forall x \in X$, $\{x\}$ is closed.

ii) For any disjoint closed sets F_1, F_2 there exist disjoint open sets G_1, G_2 such that $F_i \subset G_i, i = 1, 2$.

X is **completely regular** *if i) holds and*
iii) For any closed set F and $x_0 \in F^c$ there exists $f \in C_b(X)$ such that $0 \leq f(x) \leq 1\ \forall x \in X$, $f(x_0) = 0$ and $f|_F \equiv 1$, where $C_b(X)$ is the collection of all bounded continuous functions on X.

As we shall see later in this chapter that the topologies of $C([0,T], \Phi')$ and of $D([0,T], \Phi')$ are given by families of pseudometrics, the following theorem will be useful in the study of Φ'-valued processes.

Theorem 2.1.1 *Suppose that the topology of X is given by a family of pseudometrics $\{d_v : v \in \Gamma\}$, i.e., its neighborhoods are given by (1.1.3) with $p_{v_j}(x - x_0)$ replaced by $d_{v_j}(x, x_0)$, where d_v is a* **pseudometric** *if it satisfies the conditions of a metric (see Theorem 1.1.3 (c)) except that $d_v(x_1, x_2)$ can be 0 for $x_1 \neq x_2$. If the following separating condition holds:*

$$\forall\ x_1 \neq x_2\ \exists v \in \Gamma \text{ such that } d_v(x_1, x_2) > 0, \tag{2.1.1}$$

then X is a completely regular space.

Proof: Let $x_0 \in X$. For any $x_1 \neq x_0$, let $v \in \Gamma$ such that $\alpha \equiv d_v(x_1, x_0) > 0$. Then the neighborhood $\{x \in X : d_v(x, x_1) < \alpha/2\} \subset \{x_0\}^c$. This verifies the condition i) of Definition 2.1.1.

Let F be a closed set and $x_0 \in F^c$. As F^c is open, there exists a neighborhood

$$U = \{x \in X : d_{v_j}(x, x_0) < \epsilon_j, j = 1, \cdots, n\}$$

of x_0 such that $U \subset F^c$. Let $\epsilon = \min\{\epsilon_j : j = 1, \cdots, n\}$ and

$$d(x, y) = \max\{d_{v_j}(x, y) : j = 1, \cdots, n\},\ \forall x, y \in X.$$

Then

$$\{x \in X : d(x, x_0) < \epsilon\} \subset U \subset F^c.$$

Let

$$f(x) = \min\{1, d(x, x_0)/\epsilon\}.$$

It is easy to see that f satisfies the condition (iii) of Definition 2.1.1. ∎

Corollary 2.1.1 *i) Any metric space is completely regular.*
ii) Φ' is completely regular.

Now we study Borel measures on topological spaces. Let X be a topological space and $\mathcal{B}(X)$ (resp. $\mathcal{B}_0(X)$) be the σ-field (resp. field) generated by all open sets. A countably additive, positive, finite (resp. probability) measure μ on $\mathcal{B}(X)$ is called a **Borel measure** (resp. **Borel probability measure**). We denote the collection of all finite positive Borel measures (resp. Borel probability measures) on X by $\mathcal{M}(X)$ (resp. $\mathcal{P}(X)$).

Definition 2.1.2 $\mu \in \mathcal{M}(X)$ *is* **Radon** *if for any* $A \in \mathcal{B}(X)$

$$\mu(A) = \sup\{\mu(K) : K \text{ is compact and } F \subset A\}.$$

We present in next four theorems the relationship between Borel measures and bounded linear functionals on $C_b(X)$.

Theorem 2.1.2 *Let X be a Hausdorff topological space. Then $C_b(X)$ is a Banach space with norm*

$$\|f\| = \sup\{|f(x)| : x \in X\}.$$

Proof: It is easy to see that $C_b(X)$ is a TVS and $\|\cdot\|$ is a norm. Let $\{f_n\} \subset C_b(X)$ be a Cauchy sequence, i.e., for any $\epsilon > 0$ there exists N such that

$$|f_n(x) - f_m(x)| \leq \epsilon \quad \text{for any } n,\ m \geq N \text{ and } x \in X. \tag{2.1.2}$$

Then for any $x \in X$, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbf{R}$ and there exists $f(x) \in \mathbf{R}$ such that $f_n(x) \to f(x)$. By (2.1.2) we have

$$|f_n(x) - f(x)| \leq \epsilon \quad \text{for any } n \geq N \text{ and } x \in X. \tag{2.1.3}$$

As f_N is continuous, for any $x_0 \in X$ there exists a neighborhood U_ϵ of $x_0 \in X$ such that $x \in U_\epsilon$ implies

$$|f_N(x) - f_N(x_0)| \leq \epsilon. \tag{2.1.4}$$

Hence for any $x \in U_\epsilon$ we have

$$\begin{aligned} & |f(x) - f(x_0)| \\ \leq\ & |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ \leq\ & 3\epsilon. \end{aligned}$$

i.e. $f \in C(X)$. The boundedness of f and $f_n \to f$ follows from (2.1.3) directly. ∎

The proof of the following theorem is routine and we leave it to the reader.

Theorem 2.1.3 *Let X be a Hausdorff topological space. Then for any $\mu \in \mathcal{M}(X)$ there exists a unique $\ell \in C_b(X)'$ such that*
i) $\ell[f] \geq 0$ if $f \in C_b(X)$ and $f(x) \geq 0$ $\forall x \in X$.
ii) $\|\ell\|_{C_b(X)'} = \ell[1] = \mu(X)$.
iii)

$$\ell[f] = \int_X f(x)\mu(dx), \qquad \forall f \in C_b(X). \tag{2.1.5}$$

Remark 2.1.1 *Given $\ell \in C_b(X)'$, the relation (2.1.5) does not, in general, determine μ uniquely. The next theorem gives a sufficient condition for the uniqueness of μ given ℓ.*

Theorem 2.1.4 *Let X be a completely regular topological space. If μ and ν are two Radon probability measures such that*

$$\int_X f(x)\mu(dx) = \int_X f(x)\nu(dx), \qquad \forall f \in C_b(X), \tag{2.1.6}$$

then $\mu = \nu$.

Proof: For any compact set $K \subset X$ and $x \notin K$, let $f_x \in C_b(X)$ be given by Definition 2.1.1. Define a net

$$\Lambda \equiv \{\alpha = \{x_1, \cdots, x_n\} : n \in \mathbf{N}, x_j \notin K, 1 \leq j \leq n\}$$

whose order is given by set containing. For any $\alpha = \{x_1, \cdots, x_n\} \in \Lambda$, let

$$f_\alpha(x) = \max\{1 - f_{x_j}(x) : 1 \leq j \leq n\}.$$

Then $\{f_\alpha\} \subset C_b(X)$ is a nondecreasing net, $0 \leq f_\alpha \leq 1, f_\alpha|_K \equiv 0, f_\alpha \to 1_{K^c}$ and

$$\int_X f_\alpha(x)\mu(dx) \leq \mu(K^c). \tag{2.1.7}$$

On the other hand, for any compact $\tilde{K} \subset K^c$ and $\epsilon > 0$, we have

$$\tilde{K} \subset \cup_{\alpha\in\Lambda}\{x : f_\alpha(x) > 1 - \epsilon\},$$

and hence, there exist $n \in \mathbf{N}$ and $\alpha_j, j = 1, 2, \cdots, n$ such that

$$\tilde{K} \subset \cup_{j=1}^n\{x : f_{\alpha_j}(x) > 1 - \epsilon\}.$$

Let $\alpha^\epsilon \in \Lambda$ be such that $\alpha^\epsilon > \alpha_j$, $j = 1, 2, \cdots, n$. Then for any $\alpha \geq \alpha^\epsilon$, we have

$$\begin{aligned}\int_X f_\alpha(x)\mu(dx) &\geq (1-\epsilon)\mu(x : f_\alpha(x) > 1 - \epsilon) \\ &\geq (1-\epsilon)\mu\left(\cup_{j=1}^n\{x : f_{\alpha_j}(x) > 1 - \epsilon\}\right) \\ &\geq (1-\epsilon)\mu(\tilde{K}).\end{aligned} \tag{2.1.8}$$

As $\tilde{K}$ and ϵ are arbitrary, it follows from the Radonness of μ that (2.1.7) and (2.1.8) imply

$$\mu(K^c) = \lim_{\alpha} \int_X f_\alpha(x)\mu(dx). \tag{2.1.9}$$

It is obvious that (2.1.9) holds with μ replaced by ν. Hence $\mu(K^c) = \nu(K^c)$, i.e., $\mu(K) = \nu(K)$ for any compact subset K of X. By the Radonness again we have $\mu = \nu$. ∎

To prove the converse of Theorem 2.1.3 we need some extra structures on the topological space X and the following two lemmas.

Lemma 2.1.1 *(i) If X is a metric space, then X is normal.*
(ii) If X is a compact Hausdorff space, then X is normal.

Proof: (i) For any disjoint closed sets F_1, F_2, let

$$G_i = \{x : d(x, F_i) < d(x, F_{3-i})\} \qquad i = 1, 2.$$

Then G_1, G_2 satisfy the condition (ii) of Definition 2.1.1. The condition i) of Definition 2.1.1 follows from Theorem 2.1.1.
(ii) Let F be a closed set and $x \notin F$. For any $y \in F$ there exist two disjoint open set $G_{x,y}^{(i)}$, $i = 1, 2$ such that $x \in G_{x,y}^{(1)}$ and $y \in G_{x,y}^{(2)}$. As F is compact, there exist $y_1, \cdots, y_n$ such that $F \subset G_x^2$ where

$$G_x^2 = \cup_{j=1}^n G_{x,y_j}^{(2)}.$$

Let

$$G_x^1 = \cap_{j=1}^n G_{x,y_j}^{(1)}.$$

Then G_x^1 and G_x^2 are disjoint open sets and $x \in G_x^1$, $F \subset G_x^2$.

Let F_1, F_2 be disjoint closed sets. Let $G_x^i, i = 1, 2$ be given above with $x \in F_1$ and $F = F_2$. As F_1 is compact, there exist $x_1, \cdots, x_m$ such that $F_1 \subset G_1$ where

$$G_1 = \cup_{j=1}^m G_{x_j}^1.$$

Let

$$G_2 = \cap_{j=1}^m G_{x_j}^2.$$

Then G_1 and G_2 are disjoint open sets and $F_1 \subset G_1$, $F_2 \subset G_2$. ∎

Lemma 2.1.2 *If X is a normal topological space, then for any disjoint closed sets F_1, F_2, there exists $f \in C_b(X)$ such that $0 \le f(x) \le 1$ $\forall x \in X$ and $f|_{F_1} \equiv 0$, $f|_{F_2} \equiv 1$. In particular, X is completely regular.*

Proof: Let $G_{1/2}$ and $\tilde{G}_{1/2}$ be disjoint open sets containing F_1 and F_2 respectively. Then we have

$$F_1 \subset G_{1/2} \subset \bar{G}_{1/2} \subset \tilde{G}^c_{1/2}, F_2 \subset \tilde{G}_{1/2}$$

where $\bar{G}_{1/2}$ is the closure of $G_{1/2}$. Then F_1 and $G^c_{1/2}$ are disjoint closed sets, and F_2 and $\tilde{G}^c_{1/2}$ are disjoint closed sets. There exist open sets $G_{1/4}$ and $G_{3/4}$ such that

$$F_1 \subset G_{1/4} \subset \bar{G}_{1/4} \subset G_{1/2} \subset \bar{G}_{1/2} \subset G_{3/4} \subset \bar{G}_{3/4} \subset F_2^c.$$

By induction, there exists a family of open sets

$$\{G_r : r \in (0,1),\ r \text{ is dyadic rational}\}$$

such that (i) $r < s$ implies $\bar{G}_r \subset G_s$ and (ii) $F_1 \subset G_r$, $F_2 \cap \bar{G}_r = \phi$. Let $f(x) = \sup\{r : x \notin G_r\}$ with the convention that the supremum of the empty set is 0. We only need to verify the continuity of f.

Let $\alpha = f(x)$. If $\alpha \in (0,1)$, then $x \in G_{\alpha+\epsilon} \cap \bar{G}^c_{\alpha-\eta}$ for any small ϵ, η such that $\alpha + \epsilon$ and $\alpha - \eta$ are dyadic rationals. If $y \in G_{\alpha+\epsilon} \cap \bar{G}^c_{\alpha-\eta}$ then $|f(x) - f(y)| \leq \epsilon + \eta$.

If $\alpha = 0$, then $x \in \bar{G}^c_\eta$ for any small dyadic rational η. If $y \in \bar{G}^c_\eta$ then $|f(x) - f(y)| \leq \eta$. The case of $\alpha = 1$ can be verified similarly. ∎

It is easy to see that continuous function in a compact topological space is bounded. We shall denote $C_b(X)$ by $C(X)$ in the following theorem.

Theorem 2.1.5 *Let X be a compact Hausdorff topological space. If $\ell \in C(X)'$ such that $\ell[f] \geq 0$ for any non-negative continuous function f on X, then there exists a unique $\mu \in \mathcal{M}(X)$ such that $\mu(X) = \ell[1]$ and (2.1.5) holds.*

Proof: The uniqueness follows from Theorem 2.1.4. To prove the existence we define the following set function

$$\mu(F) = \inf\{\ell[f] : f \in C(X) \text{ and } f \geq 1_F\} \quad \text{for closed } F \subset X$$

and

$$\mu(A) = \sup\{\mu(F) : F \subset A \text{ closed}\} \quad \text{for any } A \subset X.$$

It is easy to see that μ is nondecreasing and $\mu(\emptyset) = 0$. Without loss of generality we assume that $||\ell||_{C(X)'} = 1$. Now we divide the proof into five steps:

1° For disjoint closed sets F_1 and F_2 we have

$$\mu(F_1 \cup F_2) = \mu(F_1) + \mu(F_2). \tag{2.1.10}$$

For any $\epsilon > 0$, there exist $f_i \in C(X)$ such that $f_i \geq 1_{F_i}$ and $\mu(F_i) \geq \ell[f_i] - \epsilon$, $i = 1, 2$. Then $f_1 + f_2 \geq 1_{F_1 \cup F_2}$ and hence

$$\mu(F_1 \cup F_2) \leq \ell[f_1 + f_2] = \ell[f_1] + \ell[f_2] \leq \mu(F_1) + \mu(F_2) + 2\epsilon. \qquad (2.1.11)$$

On the other hand, let $f \in C(X)$ such that $f \geq 1_{F_1 \cup F_2}$ and $\mu(F_1 \cup F_2) \geq \ell[f] - \epsilon$. It follows from Lemma 2.1.1 and Lemma 2.1.2 that there exists $f_0 \in C(X)$ such that $0 \leq f_0 \leq 1$, $f_0|_{F_1} \equiv 0$ and $f_0|_{F_2} \equiv 1$. Then $f_0 f \geq 1_{F_2}$, $(1 - f_0)f \geq 1_{F_1}$ and hence

$$\begin{aligned} \mu(F_1 \cup F_2) &\geq \ell[f] - \epsilon \\ &= \ell[f_0 f] + \ell[(1 - f_0)f] - \epsilon \\ &\geq \mu(F_1) + \mu(F_2) - \epsilon. \end{aligned} \qquad (2.1.12)$$

(2.1.10) then follows from (2.1.11) and (2.1.12).
2° Let

$$\mathcal{G} = \{B \subset X : \mu(A) = \mu(AB) + \mu(AB^c), \ \forall A \subset X\}. \qquad (2.1.13)$$

Then $\mathcal{G}$ is a field and $\mu|_{\mathcal{G}}$ is finitely additive.

It is obvious that $\mathcal{G}$ is closed under complementation. Let $B_1, B_2 \in \mathcal{G}$ and $A \subset X$. Then

$$\begin{aligned} & \mu(A(B_1 \cup B_2)) + \mu(A(B_1 \cup B_2)^c) \\ = \ & \mu(A(B_1 \cup B_2)B_1) + \mu(A(B_1 \cup B_2)B_1^c) + \mu(A(B_1 \cup B_2)^c) \\ = \ & \mu(AB_1) + \mu(AB_1^c B_2) + \mu(AB_1^c B_2^c) \\ = \ & \mu(AB_1) + \mu(AB_1^c) \\ = \ & \mu(A), \end{aligned}$$

i.e. $B_1 \cup B_2 \in \mathcal{G}$ and hence $\mathcal{G}$ is a field. Further, if B_1 and B_2 are disjoint, then

$$\mu(B_1 \cup B_2) = \mu((B_1 \cup B_2)B_1) + \mu((B_1 \cup B_2)B_1^c) = \mu(B_1) + \mu(B_2).$$

3° $\mathcal{B}_0(X) \subset \mathcal{G}$.

It follows from 2° that we only need to show that $\mathcal{G}$ contains all closed sets. Let F be closed and $A \subset X$. Then for any $\epsilon > 0$ there exist two disjoint closed sets F_1, F_2 such that $F_1 \subset AF, F_2 \subset AF^c$ and

$$\mu(AF) \leq \mu(F_1) + \epsilon \quad \text{and} \quad \mu(AF^c) \leq \mu(F_2) + \epsilon.$$

Therefore

$$\begin{aligned} \mu(AF) + \mu(AF^c) &\leq \mu(F_1) + \mu(F_2) + 2\epsilon \\ &= \mu(F_1 \cup F_2) + 2\epsilon \\ &\leq \mu(A) + 2\epsilon. \end{aligned} \qquad (2.1.14)$$

On the other hand, let $F_3 \subset A$ be a closed set such that $\mu(A) \leq \mu(F_3) + \epsilon$. Let $f_1 \in C(X)$ be such that

$$f_1 \geq 1_{FF_3} \quad \text{and} \quad \mu(FF_3) \geq \ell[f_1] - \epsilon.$$

Let $f_2 \in C(X)$ be such that

$$f_2 \geq 1_{\{f_1 \leq 1-\epsilon\} \cap F_3} \quad \text{and} \quad \mu(\{f_1 \leq 1-\epsilon\} \cap F_3) \geq \ell[f_2] - \epsilon.$$

Then for $x \in F_3$, we have either $x \in \{f_1 \leq 1-\epsilon\} \cap F_3$ or $x \in \{f_1 > 1-\epsilon\} \cap F_3$ and hence, either $f_2(x) \geq 1$ or $f_1(x) \geq 1-\epsilon$. Therefore

$$1_{F_3} \leq (1-\epsilon)^{-1} f_1 + f_2.$$

Hence

$$\begin{aligned}
& \mu(A) \leq \mu(F_3) + \epsilon \\
\leq\ & \ell[(1-\epsilon)^{-1} f_1 + f_2] + \epsilon \\
=\ & (1-\epsilon)^{-1} \ell[f_1] + \ell[f_2] + \epsilon \\
\leq\ & (1-\epsilon)^{-1}(\mu(FF_3) + \epsilon) + \mu(\{f_1 \leq 1-\epsilon\} \cap F_3) + 2\epsilon \\
\leq\ & (1-\epsilon)^{-1}(\mu(FA) + \epsilon) + \mu((FF_3)^c F_3) + 2\epsilon \\
\leq\ & (1-\epsilon)^{-1}(\mu(FA) + \epsilon) + \mu(F^c A) + 2\epsilon. \qquad (2.1.15)
\end{aligned}$$

Letting $\epsilon \to 0$, it follows from (2.1.14), (2.1.15) that

$$\mu(AF) + \mu(AF^c) = \mu(A).$$

4° μ is countably additive on $\mathcal{B}_0(X)$.

Let $\{B_j\} \subset \mathcal{B}_0(X)$ be a sequence of disjoint sets such that $\cup_j B_j \in \mathcal{B}_0(X)$. Then

$$\mu(\cup_j B_j) \geq \mu(\cup_{j=1}^n B_j) = \sum_{j=1}^n \mu(B_j)$$

and letting $n \to \infty$

$$\mu(\cup_j B_j) \geq \sum_j \mu(B_j). \qquad (2.1.16)$$

It follows from the definition of μ, the finite additivity of μ and $\mu(X) = 1$ that there exists a sequence of open sets G_j containing B_j such that $\mu(B_j) \geq \mu(G_j) - \epsilon 2^{-j}$, $j \in \mathbf{N}$, and a closed (and hence compact) set $F \subset \cup_j B_j$ such that $\mu(F) \geq \mu(\cup_j B_j) - \epsilon$. As $F \subset \cup_j G_j$, there exists n such that $F \subset \cup_{j=1}^n G_j$. Hence

$$\begin{aligned}
& \mu(\cup_j B_j) \leq \mu(F) + \epsilon \\
\leq\ & \mu(\cup_{j=1}^n G_j) + \epsilon = \mu\left(\cup_{j=1}^n \{G_j \setminus \cup_{i=1}^{j-1} G_i\}\right) + \epsilon
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} \mu\left(G_j \setminus \cup_{i=1}^{j-1} G_i\right) + \epsilon \leq \sum_{j=1}^{n} \mu(G_j) + \epsilon \\
&\leq \sum_{j=1}^{n} \left(\mu(B_j) + \epsilon 2^{-j}\right) + \epsilon \leq \sum_{j} \mu(B_j) + 2\epsilon. \qquad (2.1.17)
\end{aligned}
$$

The countable additivity of μ on $\mathcal{B}_0(X)$ follows from (2.1.16) and (2.1.17).
5° It follows from a theorem in standard measure theory that μ can be uniquely extended into a countably additive set function on $\mathcal{B}(X)$. We still denote the extension by μ. We only need show

$$
\ell[f] = \int_X f(x)\mu(dx) \qquad (2.1.18)
$$

for $f \in C(X)$ such that $0 \leq f \leq 1$.

For any $\epsilon > 0$, there exist $b_j \in (0,1), j = 1, \cdots, n$, and disjoint Borel sets $B_1, \cdots, B_n$ such that

$$
f \geq \sum_{j=1}^{n} b_j 1_{B_j} \quad \text{and} \quad \int_X f(x)\mu(dx) \leq \sum_{j=1}^{n} b_j \mu(B_j) + \epsilon.
$$

Let $F_j \subset B_j$ be closed such that $\mu(B_j) \leq \mu(F_j) + n^{-1}\epsilon, j = 1, \cdots, n$.

Next we prove by induction that there exist $f_j \in C(X)$ such that $0 \leq f_j \leq 1, f_j|_{F_j} \equiv 1, f_j|_{F_i} \equiv 0$ for any $1 \leq i \neq j \leq n$ and the sets $\{x : f_j(x) > 0\}, j = 1, 2, \cdots, n$, are disjoint.

The assertion is trivially true for $n = 1$. Suppose it is true for n. As the closed sets F_{n+1} and $\cup_{j=1}^{n} F_j$ are disjoint, there exist two disjoint open sets G and $\tilde{G}$ such that $F_{n+1} \subset G$ and $\cup_{j=1}^{n} F_j \subset \tilde{G}$. It follows from Lemma 2.1.2 that there exist $f_{n+1}, g \in C(X)$ such that a) $0 \leq f_{n+1}, g \leq 1$; b) $f_{n+1}|_{F_{n+1}} \equiv 1$, $f_{n+1}|_{G^c} \equiv 0$; c) $g|_{F_j} \equiv 1$ for all $1 \leq j \leq n$, $g|_{\tilde{G}^c} \equiv 0$. For $1 \leq j \leq n$, replacing f_j obtained from the induction assumption by $f_j g$ we see that our claim holds.

Therefore

$$
\begin{aligned}
&\int_X f(x)\mu(dx) \leq \sum_{j=1}^{n} b_j \mu(B_j) + \epsilon \\
&\leq \sum_{j=1}^{n} b_j \mu(F_j) + 2\epsilon \leq \sum_{j=1}^{n} b_j \ell[f_j \wedge (f/b_j)] + 2\epsilon \\
&= \sum_{j=1}^{n} \ell[(b_j f_j) \wedge f] + 2\epsilon = \ell\left[\sum_{j=1}^{n} \{(b_j f_j) \wedge f\}\right] + 2\epsilon \\
&= \ell\left[\left(\sum_{j=1}^{n} b_j f_j\right) \wedge f\right] + 2\epsilon \leq \ell[f] + 2\epsilon.
\end{aligned}
$$

i.e.

$$\int_X f(x)\mu(dx) \le \ell[f] \tag{2.1.19}$$

for $f \in C(X)$ such that $0 \le f \le 1$. Replacing f by $1-f$, we see that (2.1.19) becomes an equality, i.e. (2.1.18) holds. ∎

The following result is well known.

Theorem 2.1.6 *Let X be a Polish space and $\mu \in \mathcal{P}(X)$. Then μ is Radon.*

Proof: Let

$$\mathcal{G} = \left\{ B \in \mathcal{B}(X) : \mu(B) = \sup_{F \subset B, F \text{ closed}} \mu(F) = \inf_{B \subset G, G \text{ open}} \mu(F) \right\}.$$

If $B \in \mathcal{G}$, then clearly $B^c \in \mathcal{G}$. If $B_n \in \mathcal{G}$, there exist closed F_n and open G_n such that $F_n \subset B_n \subset G_n$ and $\mu(G_n \setminus F_n) \le \epsilon 2^{-n}$, $n = 1, 2, \cdots$. Let n_0 such that

$$\mu\left(\cup_{n=1}^\infty F_n \setminus \cup_{n=1}^{n_0} F_n\right) \le \epsilon/2.$$

Let $G = \cup_{n=1}^\infty G_n$ and $F = \cup_{n=1}^{n_0} F_n$. Then $F \subset \cup_{n=1}^\infty B_n \subset G$ and

$$\mu(G \setminus F) \le \sum_{n=1}^\infty \mu(G_n \setminus F_n) + \mu\left(\cup_{n=1}^\infty F_n \setminus F\right) \le \epsilon.$$

Hence $\mathcal{G}$ is a σ-field. For F closed, let $G_n = \{x : d(x, F) < n^{-1}\}$. Then G_n decreases to F and hence $F \in \mathcal{G}$. Therefore $\mathcal{G} = \mathcal{B}(X)$.

Since X is separable, there exists a countable set $\{x_n\}$ which is dense in X. Let

$$F_{nk} = \{x : d(x, x_n) \le k^{-1}\}, \qquad n, k \in \mathbf{N}.$$

Then $X = \cup_{n=1}^\infty F_{nk}$. As

$$1 = \lim_{N \to \infty} \mu\left(\cup_{n=1}^N F_{nk}\right),$$

there exists N_k such that

$$\mu\left(\cup_{n=1}^{N_k} F_{nk}\right) \ge 1 - \epsilon 2^{-k}.$$

Let

$$K = \cap_{k=1}^\infty \cup_{n=1}^{N_k} F_{nk}.$$

It is easy to see that K is compact. Further,

$$\mu(K^c) \le \sum_{k=1}^\infty \mu\left(\left(\cup_{n=1}^{N_k} F_{nk}\right)^c\right) \le \sum_{k=1}^\infty \epsilon 2^{-k} = \epsilon.$$

For any $B \in \mathcal{B}(X)$, there exists a closed set $F \subset B$ such that $\mu(B) \leq \mu(F)+\epsilon$. Hence $F \cap K \subset B$ is compact and

$$\mu(B) \leq \mu(F \cap K) + \mu(F \cap K^c) + \epsilon \leq \mu(F \cap K) + 2\epsilon.$$

This proves the Radonness of μ. ∎

2.2 Weak convergence of probability measures.

In this section, we introduce the weak convergence topology in the space of Borel probability measures on topological spaces. Then we give a sufficient condition for a sequence of probability measures to be weakly compact. At the end of this section, we state and prove a useful representation result due to Skorohod for weakly convergent sequence of probability measures on a Polish space.

Definition 2.2.1 *Let X be a topological space.*
i) A sequence $\{\mu_n\} \subset \mathcal{P}(X)$ **converges weakly** *to $\mu \in \mathcal{P}(X)$ if $\forall f \in C_b(X)$*

$$\lim_{n\to\infty} \int_X f(x)\mu_n(dx) = \int_X f(x)\mu(dx).$$

ii) $\{\mu_n\}$ is **tight** *if $\forall \epsilon > 0$ there exists a compact subset K_ϵ of X such that*

$$\mu_n(K_\epsilon) > 1 - \epsilon, \qquad \forall n \geq 1.$$

Lemma 2.2.1 (Banach-Alaoglu) *Let X be a Banach space with dual X'. The weak*-topology in X' is defined as the weakest topology such that for each $x \in X$, the map $f \in X' \to f[x] \in \mathbf{R}$ is continuous. Then the unit ball of X' is compact.*

Proof: It follows from Tychonoff's theorem that $K = \prod_{x\in X}[-\|x\|, \|x\|]$ is a compact subset of $\mathbf{R}^X \equiv \prod_{x\in X}\mathbf{R}$. Let $\pi : f \in X' \to \{f[x]\}_{x\in X} \in \mathbf{R}^X$ and $B = \pi S$, where S is the unit ball of X'. Then B is closed in $\mathbf{R}^X$. In fact, let $\{\{f_\alpha[x]\}_{x\in X}\}$ be a net in B such that $f_\alpha[x] \to f(x)$, $\forall x \in X$. Then $|f(x)| \leq \|x\|$ and, for any $x, y \in X$, $a, b \in \mathbf{R}$ we have

$$f(ax+by) = \lim_\alpha f_\alpha[ax+by] = \lim_\alpha (af_\alpha[x] + bf_\alpha[y]) = af(x) + bf(y),$$

and hence $\{f(x)\}_{x\in X} \in B$.

Further it is easy to see that π is an isomorphism between S and B. As $B \subset K$ is compact, we see that S is compact. ∎

Theorem 2.2.1 *Let X be a Hausdorff topological space. If $\{\mu_n\} \subset \mathcal{P}(X)$ is tight, then $\{\mu_n\}$ is relatively compact in the weak topology.*

Proof: Let $\{K_m\}$ be an increasing sequence of compact sets of X such that $\mu_n(K_m) > 1 - 2^{-m}$, $\forall n \geq 1$. For each m, let $\nu_{n,m}(B) = \mu_n(B)$ for any $B \in \mathcal{B}(K_m)$. Then $\{\nu_{n,m}\}_{n\geq 1}$ is a sequence of positive measures on K_m. By Theorem 2.1.3, $\{\nu_{n,m}\}_{n\geq 1}$ can be regarded as a sequence in the unit ball of $C(K_m)$. It follows from Lemma 2.2.1, the diagonal argument and Theorem 2.1.5 that there exists a sequence $\{n_k\}$ such that $\forall m \geq 1$, there is a positive Borel measure ρ_m on K_m satisfying: $\rho_m(K_m) \leq 1$, $\rho_{m'}|_{\mathcal{B}(K_m)} = \rho_m$ $\forall m' > m$, and $\forall f \in C(K_m)$,

$$\int_{K_m} f(x)\nu_{n_k,m}(dx) \to \int_{K_m} f(x)\rho_m(dx), \quad as\ k \to \infty. \tag{2.2.1}$$

Note that for any $B \in \mathcal{B}(X)$ we have

$$\rho_{m+1}(B \cap K_{m+1}) \geq \rho_{m+1}(B \cap K_m) = \rho_m(B \cap K_m),$$

and hence $\rho_m(B \cap K_m)$ increases, say to $\mu(B)$. It is easy to see that μ is nondecreasing and $\mu(\emptyset) = 0$. It follows from the monotone convergence theorem that for any disjoint $\{B_j\} \subset \mathcal{B}(X)$,

$$\begin{aligned} \mu(\cup_j B_j) &= \lim_{m\to\infty} \rho_m(\cup_j B_j \cap K_m) = \lim_{m\to\infty} \sum_j \rho_m(B_j \cap K_m) \\ &= \sum_j \lim_{m\to\infty} \rho_m(B_j \cap K_m) = \sum_j \mu(B_j). \end{aligned}$$

Further

$$\begin{aligned} 1 &\geq \mu(X) = \lim_{m\to\infty} \rho_m(K_m) = \lim_{m\to\infty} \lim_{k\to\infty} \nu_{n_k,m}(K_m) \\ &\geq \lim_{m\to\infty} \lim_{k\to\infty} (1 - 2^{-m}) = 1. \end{aligned}$$

Therefore $\mu \in \mathcal{P}(X)$. Finally for any $f \in C_b(X)$ we have $f|_{K_m} \in C(K_m)$ and hence

$$\begin{aligned} & \limsup_{k\to\infty} \left| \int_X f(x)\mu_{n_k}(dx) - \int_X f(x)\mu(dx) \right| \\ \leq & \limsup_{k\to\infty} \Big(\|f\|\mu_{n_k}(K_m^c) + \|f\|\mu(K_m^c) \\ & + \left| \int_{K_m} f(x)\mu_{n_k}(dx) - \int_{K_m} f(x)\mu(dx) \right| \Big) \\ \leq & \|f\| 2^{1-m}. \end{aligned}$$

Letting $m \to \infty$ we see that $\{\mu_{n_k}\}$ converges to μ weakly. ∎

Next we consider weak convergence in Polish spaces. The following theorem, due to Prohorov, is the converse of Theorem 2.2.1 for probability measures on Polish spaces. We need the following lemma.

Lemma 2.2.2 *The following four conditions are equivalent:*
i) $\mu_n \to \mu$ *weakly.*
ii) $\limsup_{n\to\infty} \mu_n(F) \le \mu(F)$ *for any closed set* F.
iii) $\liminf_{n\to\infty} \mu_n(G) \ge \mu(G)$ *for any open set* G.
iv) $\lim_{n\to\infty} \mu_n(B) = \mu(B)$ *for any* $B \in \mathcal{B}(X)$ *such that* $\mu(\partial B) = 0$, *where* ∂B *is the boundary of the set* B.

Proof: *i)* $\Rightarrow$ *ii)* Let $f_m(x) = \{md(x,F)\} \wedge 1$ for any $m \in \mathbf{N}$, $x \in X$. Then $f_m \in C_b(X)$ and f_m increases to 1_{F^c}. Hence

$$\limsup_{n\to\infty} \mu_n(F) \le \limsup_{n\to\infty} \int (1 - f_m(x))\mu_n(dx) = \int (1 - f_m(x))\mu(dx).$$

Therefore

$$\limsup_{n\to\infty} \mu_n(F) \le \lim_{m\to\infty} \int (1 - f_m(x))\mu(dx) = \mu(F).$$

It is easy to prove the equivalence of ii) and iii) by taking complements. That *ii)&iii)* implies *iv)* follows from the inequality

$$\begin{aligned}\liminf_{n\to\infty} \mu_n(B) &\ge \liminf_{n\to\infty} \mu_n(B^0) \ge \mu(B^0) = \mu(\bar{B}) \\ &\ge \limsup_{n\to\infty} \mu_n(\bar{B}) \ge \limsup_{n\to\infty} \mu_n(B)\end{aligned}$$

where B^0 and $\bar{B}$ are respectively the interior and closure of the set B.

Finally we show that *iv)* $\Rightarrow$ *i)*. Let $f \in C_b(X)$ be fixed. Note that for any $a < b$,

$$\partial\{x \in X : a \le f(x) < b\} \subset \{x \in X : f(x) = a \text{ or } f(x) = b\},$$

and the set

$$D = \{r \in \mathbf{R} : \mu\{x \in X : f(x) = r\} > 0\}$$

is countable. For any $\epsilon > 0$, let $r_1 \le -\|f\| < r_2 < \cdots < r_{m-1} < \|f\| \le r_m$ such that $r_j \notin D$, $\forall j = 1, 2, \cdots, m$, and $r_{j+1} - r_j < \epsilon$, $\forall j = 1, 2, \cdots, m-1$. Then

$$\int_X f(x)\mu(dx) \ge \sum_{j=1}^{m-1} r_j \mu(x \in X : r_j \le f(x) < r_{j+1})$$

$$= \lim_{n\to\infty} \sum_{j=1}^{m-1} r_j \mu_n(x \in X : r_j \le f(x) < r_{j+1})$$

$$
\begin{aligned}
&= \lim_{n\to\infty}\sum_{j=1}^{m-1}(r_j - r_{j+1})\mu_n(x\in X : r_j \le f(x) < r_{j+1}) \\
&\quad + \lim_{n\to\infty}\sum_{j=1}^{m-1} r_{j+1}\mu_n(x\in X : r_j \le f(x) < r_{j+1}) \\
&\ge -\epsilon + \limsup_{n\to\infty}\int_X f(x)\mu_n(dx).
\end{aligned}
$$

Therefore

$$\int_X f(x)\mu(dx) \ge \limsup_{n\to\infty}\int_X f(x)\mu_n(dx). \tag{2.2.2}$$

Replacing f by $\|f\| - f$ we have

$$\int_X f(x)\mu(dx) \le \liminf_{n\to\infty}\int_X f(x)\mu_n(dx).$$

(i) follows from the last two inequalities. ∎

Theorem 2.2.2 (Prohorov) *) Let X be a Polish space and let $\{\mu_n\}$ be a sequence of relatively compact Borel probability measures on X. Then $\{\mu_n\}$ is tight.*

Proof: Since X is separable, there exist open spheres $S_{1m}, S_{2m}, \cdots$ of radius $1/m$ such that $X = \cup_j S_{jm}$. First we show that for any $m \ge 1$ and $\eta > 0$, there exists $k(m)$ such that

$$\mu_n\left(\cup_{j=1}^{k(m)} S_{jm}\right) > 1 - \eta, \qquad \forall n \ge 1. \tag{2.2.3}$$

If (2.2.3) is not true, there exist $m_0 \ge 1$ and $\eta_0 > 0$, $\forall k \ge 1, \exists n_k \ge 1$ such that

$$\mu_{n_k}\left(\cup_{j=1}^{k} S_{jm_0}\right) \le 1 - \eta_0, \qquad \forall k \ge 1.$$

As $\{\mu_n\}$ is relatively compact, we assume that n_k increasing to infinity and $\mu_{n_k} \Rightarrow \mu$ in $\mathcal{P}(X)$. By Lemma 2.2.2, for any $J \ge 1$ we have

$$
\begin{aligned}
\mu\left(\cup_{j=1}^{J} S_{jm_0}\right) &\le \liminf_{k\to\infty}\mu_{n_k}\left(\cup_{j=1}^{J} S_{jm_0}\right) \\
&\le \liminf_{k\to\infty}\mu_{n_k}\left(\cup_{j=1}^{k} S_{jm_0}\right) \le 1 - \eta_0.
\end{aligned}
$$

Letting $J \to \infty$, then

$$1 = \mu(X) \le 1 - \eta_0,$$

and hence, (2.2.3) holds.

For any $\epsilon > 0$, taking $\eta = 2^{-m}\epsilon$ in (2.2.3), we define

$$K_\epsilon \equiv \cap_{m=1}^{\infty} \cup_{j=1}^{k(m)} \bar{S}_{jm}.$$

Then K_ϵ is compact and for any $n \geq 1$

$$\mu_n(K_\epsilon^c) \leq \sum_{m=1}^{\infty} \mu_n\left(\left\{\cup_{j=1}^{k(m)} \bar{S}_{jm}\right\}^c\right) \leq \sum_{m=1}^{\infty} 2^{-m}\epsilon = \epsilon$$

Therefore $\{\mu_n\}$ is tight. ∎

Next we present the relationship between converges in distribution and converges almost surely of random variables. It is easy to see that if $(\Omega, \mathcal{F}, P)$ is a probability space and ξ_n, ξ are measurable maps from $(\Omega, \mathcal{F})$ to $(X, \mathcal{B}(X))$ (i.e., X-valued random variables) such that $\xi_n \to \xi$ a.s., then $\mu_n \to \mu$ weakly, where $\mu_n = P(\xi_n)^{-1}$ and $\mu = P\xi^{-1}$ are probability measures on X. The following theorem is the converse of the above statement.

Theorem 2.2.3 (Skorohod) *Let X be a Polish space and $\{\mu_n\} \subset \mathcal{P}(X)$ converges to $\mu \in \mathcal{P}(X)$ weakly. Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and measurable maps ξ_n and ξ from $(\Omega, \mathcal{F})$ to $(X, \mathcal{B}(X))$ such that $\mu_n = P(\xi_n)^{-1}$, $\mu = P\xi^{-1}$ and $\xi_n \to \xi$ almost surely.*

Proof: Let $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}([0,1))$ and P be the Lebesgue measure. For any $x \in X$ and $r > 0$, let $B(x,r) = \{y \in X : d(y,x) < r\}$. We divide the proof into five steps.

1° Construct a family of partitions of X.

Let $\{x_k\}$ be a countable dense subset of X. For any $C \in \mathcal{B}(X)$ and $r > 0$, let

$$C_1^r = C \cap B(x_1, r) \quad \text{and} \quad C_{k+1}^r = (C \cap B(x_{k+1}, r)) \setminus \cup_{j=1}^k C_j^r, \ k \geq 0.$$

Then $\{C_j^r\}_{j\geq 1}$ is a partition of C while each of them has a diameter not larger than $2r$.

As for any $\nu \in \mathcal{P}(X)$, the set $\{r \in \mathbf{R} : \nu(\partial B(x_k, r)) > 0 \text{ for some } k \in \mathbf{N}\}$ is countable, we can choose a sequence $\{r_m\}$ decreasing to 0 such that

$$\begin{aligned} &\mu\left(\cup_{m=1}^\infty \cup_{k=1}^\infty \partial B(x_k, r_m)\right) \\ = \ &\mu_n\left(\cup_{m=1}^\infty \cup_{k=1}^\infty \partial B(x_k, r_m)\right) = 0, \quad \forall n \geq 1. \end{aligned} \tag{2.2.4}$$

For any $m \in \mathbf{N}$ and $(i_1, \cdots, i_m) \in \mathbf{N}^m$ we define $A_{i_1 \cdots i_m}$ inductively:

$$A_k = X_k^{r_1} \quad \text{and} \quad A_{i_1 \cdots i_m k} = (A_{i_1 \cdots i_m})_k^{r_{m+1}}, \quad k \geq 1.$$

It follows from Lemma 2.2.2 and (2.2.4) that

$$\mu_n(A_{i_1 \cdots i_m}) \to \mu(A_{i_1 \cdots i_m}) \tag{2.2.5}$$

for any $m \in \mathbf{N}$ and $(i_1, \cdots, i_m) \in \mathbf{N}^m$.
2° Construct a family of partitions of $[0,1)$ related to a probability measure ν ($\nu = \mu$ or μ_n) on X. Let

$$I_j^\nu = \left[\sum_{k=1}^{j-1} \nu(A_k), \sum_{k=1}^{j} \nu(A_k)\right), \qquad j \geq 1.$$

If $I^\nu_{i_1 \cdots i_m} = [\alpha, \beta)$ with $\beta - \alpha = \nu(A_{i_1 \cdots i_m})$, then

$$I^\nu_{i_1 \cdots i_m j} = \left[\alpha + \sum_{k=1}^{j-1} \nu(A_{i_1 \cdots i_m k}), \alpha + \sum_{k=1}^{j} \nu(A_{i_1 \cdots i_m k})\right), \qquad j \geq 1.$$

By (2.2.5) and induction, it is easy to see that, for any $m \in \mathbf{N}$ and $(i_1, \cdots, i_m) \in \mathbf{N}^m$, the left (resp. right) end point of $I^{\mu_n}_{i_1 \cdots i_m}$ tends to the left (resp. right) end point of $I^\mu_{i_1 \cdots i_m}$, as $n \to \infty$.
3° Construct random variables $\xi, \xi_1, \xi_2, \cdots$ from $[0,1)$ to X.

We choose $y_{i_1 \cdots i_m} \in A_{i_1 \cdots i_m}$ if it is non-empty and define

$$Z_m(\omega) = y_{i_1 \cdots i_m}, \qquad \forall m \geq 1, \ \omega \in I^\mu_{i_1 \cdots i_m},$$

and

$$Z_{m,n}(\omega) = y_{i_1 \cdots i_m}, \qquad \forall m \geq 1, \ \omega \in I^{\mu_n}_{i_1 \cdots i_m}.$$

For $\omega \in \Omega$, $m, k \geq 1$, we have $\omega \in I^\mu_{i_1 \cdots i_{m+k}} \subset I^\mu_{i_1 \cdots i_m}$ for some $(i_1, \cdots, i_{m+k}) \in \mathbf{N}^{m+k}$ such that $A_{i_1 \cdots i_{m+k}} \neq \emptyset$. Then, as $y_{i_1 \cdots i_{m+k}}, y_{i_1 \cdots i_m} \in A_{i_1 \cdots i_m}$, we have

$$d(Z_{m+k}(\omega), Z_m(\omega)) = d(y_{i_1 \cdots i_{m+k}}, y_{i_1 \cdots i_m}) \leq 2r_m \to 0, \qquad (2.2.6)$$

P-a.s. Therefore Z_m converges, say to ξ, a.s. as $m \to \infty$. Similarly we can prove that for each n, $Z_{n,m}$ converges to a random variable ξ_n a.s. as $m \to \infty$.
4° $\mu_n = P(\xi_n)^{-1}$, $\mu = P\xi^{-1}$.

Let $B \in \mathcal{B}(X)$ such that $\mu(\partial B) = P(\omega : \xi \in \partial B) = 0$. For each $m \in \mathbf{N}$, we denote by J_m the collection of all $(i_1, \cdots, i_m) \in \mathbf{N}^m$ such that $A_{i_1 \cdots i_m} \neq \emptyset$, $y_{i_1 \cdots i_m} \in B$. Then

$$\begin{aligned} P(\omega : \xi(\omega) \in B) &= \lim_{m \to \infty} P(\omega : Z_m(\omega) \in B) \\ &= \lim_{m \to \infty} P\left(\cup_{J_m} I^\mu_{i_1 \cdots i_m}\right) = \lim_{m \to \infty} \sum_{J_m} P\left(I^\mu_{i_1 \cdots i_m}\right) \\ &= \lim_{m \to \infty} \sum_{J_m} \mu(A_{i_1 \cdots i_m}) = \lim_{m \to \infty} \mu\left(\cup_{J_m} A_{i_1 \cdots i_m}\right) \\ &\leq \lim_{m \to \infty} \mu(x \in X : d(x, B) \leq 2r_m) \\ &= \mu(\bar{B}) = \mu(B). \end{aligned}$$

It follows from the same argument as in the proof of (2.2.2) that for any $f \in C_b(X)$

$$\int_X f(x)\mu(dx) \geq \int_X f(x)(P\xi^{-1})(dx), \tag{2.2.7}$$

and then by the same argument as in the proof of Lemma 2.2.2 we see that the above inequality becomes an equality. By Corollary 2.1.1 and Theorem 2.1.6, X is a completely regular space and μ, $P\xi^{-1}$ are two Radon measures on X. Hence, by Theorem 2.1.4, $\mu = P\xi^{-1}$. Similar arguments yield $\mu_n = P(\xi_n)^{-1}$.

5° $\xi_n \to \xi$ a.s. as $n \to \infty$.

Let Ω_0 be the collection of all end points of $I^{\mu}_{i_1\cdots i_m}$ for $m \in \mathbf{N}$ and $(i_1, \cdots, i_m) \in \mathbf{N}^m$. As Ω_0 is countable, $P(\Omega_0) = 0$. Let $\omega \notin \Omega_0$ be fixed. Then for any $m \in \mathbf{N}$ there exists $(i_1, \cdots, i_m) \in \mathbf{N}^m$ such that ω in the interior of $I^{\mu}_{i_1\cdots i_m}$. It follows from 2° that there exists n_m such that $n \geq n_m$ implies ω in the interior of $I^{\mu_n}_{i_1\cdots i_m}$ and hence $Z_{m,n}(\omega) = Z_m(\omega)$. Therefore for $n \geq n_m$

$$d(\xi_n(\omega), \xi(\omega)) \leq d(\xi_n(\omega), Z_{m,n}(\omega)) + d(Z_m(\omega), \xi(\omega)) \leq 4r_m$$

where the last inequality follows by taking $k \to \infty$ in (2.2.6). This shows that $\xi_n(\omega) \to \xi(\omega)$ a.s. ∎

2.3 Probability measures on linear topological vector spaces: The theorems of Minlos and Sazonov

In this section we study Borel probability measures on duals of linear topological vector spaces and their characteristic functions which will provide us with a powerful tool for dealing with some practical problems. Let X be a Hausdorff TVS with dual space X'.

Definition 2.3.1 *$A \subset X'$ is called a* **cylinder set associated with** *$(x_1, \cdots, x_n)$ if $x_1, \cdots, x_n \in X$ and there exists $B_n \in \mathcal{B}(\mathbf{R}^n)$ such that*

$$A = \{f \in X' : (f[x_1], \cdots, f[x_n]) \in B_n\}. \tag{2.3.1}$$

We denote it by $A \in \mathcal{C}_{x_1,\cdots,x_n}$. Let $\mathcal{C}$ be the collection of all cylinder sets.

It is easy to see that $\mathcal{C}_{x_1,\cdots,x_n} \subset \mathcal{B}(X')$ is a σ-field and $\mathcal{C} \subset \mathcal{B}(X')$ is a field. Then $\sigma(\mathcal{C}) \subset \mathcal{B}(X')$. On the other hand, if X is separable, then for any bounded subset B of X (see Definition 1.1.7) there exists a countable dense subset $B_0 \subset B$ and hence the seminorm q_B of X' can be written as

$$q_B(f) = \sup_{x \in B_0} |f[x]|$$

which is $\sigma(\mathcal{C})$-measurable. Therefore $\sigma(\mathcal{C}) = \mathcal{B}(X')$.

Definition 2.3.2 *A set function μ defined on $\mathcal{C}$ is called a* **cylinder probability measure on** X' *if for any $x_1, \cdots, x_n \in X$, $\mu|_{\mathcal{C}_{x_1,\cdots,x_n}}$ is a probability measure.*

From this definition, it is easy to see that any cylinder probability measure on X' is finitely additive on $\mathcal{C}$.

For any cylinder probability measure μ on X', we can define its **Bochner functional** as follows:

$$\hat{\mu}(x) = \int_{X'} e^{if[x]}\mu(df), \quad \forall x \in X.$$

If μ is countably additive, then $\hat{\mu}$ is called the **characteristic function of** μ.

From the finite dimensional results, it is easy to see that a cylinder probability measure μ is uniquely determined by its Bochner functional.

Theorem 2.3.1 *$F : X \to \mathbf{C}$ is the Bochner functional of a cylinder probability measure μ iff*
i) $F(0) = 1$.
ii) F is sequentially continuous at $0 \in X$.
iii) F is positive definite, i.e. $\forall n \in \mathbf{N}$, $x_j \in X$, α_j complex, $j = 1, 2, \cdots, n$ we have

$$\sum_{j,k=1}^{n} F(x_j - x_k)\alpha_j\bar{\alpha}_k \geq 0.$$

Proof: The necessity of the conditions follows from the same arguments as those for the characteristic functions of finite dimensional random variables. Now let $F : X \to \mathbf{C}$ satisfy the conditions i)-iii). For any $x_1, \cdots, x_n \in X$, let $F_{x_1,\cdots,x_n} : \mathbf{R}^n \to \mathbf{C}$ be given by

$$F_{x_1,\cdots,x_n}(u) = F\left(\sum_{j=1}^{n} u_j x_j\right), \quad \forall u \in \mathbf{R}^n.$$

Then $F_{x_1,\cdots,x_n}$ is a continuous positive definite function in $\mathbf{R}^n$ with $F_{x_1,\cdots,x_n}$ $(0) = 1$. Hence there exists a probability measure $\nu_{x_1,\cdots,x_n}$ on $\mathbf{R}^n$ such that

$$F_{x_1,\cdots,x_n}(u) = \int_{\mathbf{R}^n} \exp\left(i\sum_{j=1}^{n} u_j v_j\right) \nu_{x_1,\cdots,x_n}(dv), \quad \forall u \in \mathbf{R}^n.$$

If $A \in \mathcal{C}$ can be represented as

$$\begin{aligned} A &= \{f \in X' : (f[x_1], \cdots, f[x_n]) \in B_n\} \\ &= \{f \in X' : (f[y_1], \cdots, f[y_m]) \in B_m\}, \end{aligned}$$

we prove that

$$\nu_{x_1,\cdots,x_n}(B_n) = \nu_{y_1,\cdots,y_m}(B_m). \tag{2.3.2}$$

Without loss of generality we assume that $y_1, \cdots, y_m$ are linearly independent and there exist $a_{jk}, 1 \le j \le n, 1 \le k \le m$ such that

$$x_j = \sum_{k=1}^{m} a_{jk} y_k,\ j = 1, \cdots, n.$$

Let π be a linear map from $\mathbf{R}^m$ to $\mathbf{R}^n$ given by $u = \pi v$ such that

$$u_j = \sum_{k=1}^{m} a_{jk} v_k,\ j = 1, \cdots, n.$$

Let $\tilde{\nu}_{y_1,\cdots,y_m} = \nu_{y_1,\cdots,y_m}\pi^{-1} \in \mathcal{P}(\mathbf{R}^n)$. Then

$$\begin{aligned} &\int_{\mathbf{R}^n} \exp\left(i\sum_{j=1}^{n} t_j u_j\right) \tilde{\nu}_{y_1,\cdots,y_m}(du) \\ = &\int_{\mathbf{R}^m} \exp\left(i\sum_{j=1}^{n} t_j \sum_{k=1}^{m} a_{jk} v_k\right) \nu_{y_1,\cdots,y_m}(dv) \\ = &\ F\left(\sum_{k=1}^{m}\sum_{j=1}^{n} a_{jk} t_j y_k\right) = F\left(\sum_{j=1}^{n} t_j x_j\right) \\ = &\int_{\mathbf{R}^n} \exp\left(i\sum_{j=1}^{n} t_j u_j\right) \nu_{x_1,\cdots,x_n}(du) \end{aligned}$$

and (2.3.2) follows immediately. For each $A \in \mathcal{C}$ given by (2.3.1) we define

$$\mu(A) = \nu_{x_1,\cdots,x_n}(B_n).$$

Then μ is a well-defined cylinder probability measure on X'. It is easy to see that $F = \hat{\mu}$. ∎

Now we consider the countable additivity of cylinder probability measures on X' .

Lemma 2.3.1 *Let μ be a cylinder probability measure on X'. μ is countably additive on $\mathcal{C}$ iff for any sequence of cylinder sets $\{A_k\}$ with union X', we have*

$$\sum_k \mu(A_k) \ge 1. \tag{2.3.3}$$

Proof: The necessity of the condition is obvious. To prove the sufficiency, we note that if $C = \cup_k C_k$ is a decomposition of a cylinder set C into non-intersecting cylinder sets $C_1, C_2, \cdots$, then X' can be represented as the disjoint union of $X' - C, C_1, C_2, \cdots$ and hence

$$\mu(X' - C) + \sum_k \mu(C_k) \geq 1,$$

i.e.

$$\sum_k \mu(C_k) \geq \mu(C). \tag{2.3.4}$$

It follows from the finite additivity of μ we see that (2.3.4) becomes an equality. This proves the countable additivity of μ. ∎

Lemma 2.3.2 *Let μ be a cylinder probability measure on X'. If for any $\epsilon > 0$ there exists a compact subset K^ϵ of X' such that $\mu(C) \geq 1 - \epsilon$ for any open $C \in \mathcal{C}$ containing K^ϵ, then μ is countably additive on $\mathcal{C}$.*

Proof: Let $\{A_k\}$ be a sequence of cylinder sets with union X'. It follows from Theorem 2.1.6 that there exists a sequence of open cylinder sets $\{C_k\}$ such that $A_k \subset C_k$ and $\mu(C_k) \leq \mu(A_k) + \epsilon 2^{-k}$, $k \geq 1$. As $\{C_k\}$ is an open covering of K^ϵ, there exists k_0 such that $\{C_k\}_{1 \leq k \leq k_0}$ covers K^ϵ. Therefore

$$1 - \epsilon \leq \mu\left(\cup_{k=1}^{k_0} C_k\right) \leq \sum_{k=1}^{k_0} \mu(C_k) \leq \sum_{k=1}^{k_0} \mu(A_k) + \epsilon,$$

i.e. (2.3.3) holds and hence, μ is countably additive. ∎

Lemma 2.3.3 *Let μ be a cylinder probability measure on X' such that there exists an inner product $< \cdot, \cdot >_\mu$ on X, a constant ϵ such that*

$$|\hat{\mu}(x) - 1| \leq < x, x >_\mu + \epsilon, \qquad \forall x \in X.$$

Then for any $x_1, \cdots, x_n \in X$ and $M > 0$ we have

$$\mu\left(f \in X' : \sum_{j=1}^n f[x_j]^2 > M^2\right) \leq \frac{\sqrt{e}}{\sqrt{e} - 1}\left(\frac{1}{M^2}\sum_{j=1}^n < x_j, x_j >_\mu + \epsilon\right).$$

Proof: It follows from the Chebyshev inequality that

$$\begin{aligned}
&\mu\left(f\in X' : \sum_{j=1}^{n} f[x_j]^2 > M^2\right)\\
\le\ & \frac{\sqrt{e}}{\sqrt{e}-1}\int_{X'}\left\{1-\exp\left(-\sum_{j=1}^{n} f[x_j]^2/2M^2\right)\right\}\mu(df)\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\int_{X'}\int_{\mathbf{R}^n}\left\{1-\exp\left(i\sum_{j=1}^{n} f[x_j]u_j/M\right)\right\}\\
&\qquad (2\pi)^{-n/2}e^{-|u|^2/2}du\mu(df)\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\int_{\mathbf{R}^n}\left\{1-\hat{\mu}\left(\sum_{j=1}^{n} u_jx_j/M\right)\right\}\\
&\qquad (2\pi)^{-n/2}e^{-|u|^2/2}du\\
\le\ & \frac{\sqrt{e}}{\sqrt{e}-1}\int_{\mathbf{R}^n}\left(\left\langle\sum_{j=1}^{n} u_jx_j/M, \sum_{j=1}^{n} u_jx_j/M\right\rangle_{\mu}+\epsilon\right)\\
&\qquad (2\pi)^{-n/2}e^{-|u|^2/2}du\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\left(\frac{1}{M^2}\sum_{j=1}^{n} < x_j, x_j >_{\mu} +\epsilon\right).
\end{aligned}$$

∎

Of special importance for us is the case when X is a separable Hilbert space and the case when X is the countably Hilbertian nuclear space Φ. First we consider the case of $X = \Phi$.

Theorem 2.3.2 (Minlos's theorem) *A complex-valued function F on Φ is the characteristic function of a $\mu \in \mathcal{P}(\Phi')$ iff i) $F(0) = 1$, ii) F is positive definite, iii) F is continuous at 0 in Φ.*

Proof: As Φ is a metric space the sequential continuity is equivalent to the continuity. Hence the necessity of the conditions follows from Theorem 2.3.1. To prove the sufficiency, let μ be the cylinder probability measure on Φ' given by F. We only need to prove that μ is countably additive. By the continuity of F at 0 and a similar argument as in the proof of Lemma 1.3.1, for any $\epsilon > 0$, there exist $p > 0, \delta > 0$ such that

$$|F(\phi) - 1| \le \epsilon, \qquad \forall \phi \in \Phi \ s.t.\ \|\phi\|_p < \delta.$$

As in the finite dimensional case, it can be shown that $|F(\phi)| \le F(0)$ for any positive definite function F. Hence

$$|F(\phi) - 1| \le \epsilon + \frac{2 < \phi, \phi >_p}{\delta^2}, \qquad \forall\ \phi \in \Phi.$$

Let $q > p$ be such that the canonical injection ι from Φ_q to Φ_p is Hilbert-Schmidt. Let M be a constant such that

$$2\sqrt{e}\|\iota\|^2_{L_{(2)}(\Phi_q,\Phi_p)} = \delta^2 M^2 \epsilon(\sqrt{e}-1).$$

Define

$$K^\epsilon = \left\{ f \in \Phi' : \sup_{\phi\in\Phi, \|\phi\|_q \leq 1} |f[\phi]| \leq M \right\}.$$

Then K^ϵ is compact in Φ'. For any $A \in \mathcal{C}$ contained in $(K^\epsilon)^c$, let A be given by (2.3.1). By Schmidt orthonormalization of $\{x_j\}_{1\leq j\leq n}$, we assume that $\{x_j\}_{1\leq j\leq n}$ is an ONS in Φ_q. Then

$$B_n \subset \{u \in \mathbf{R}^n : |u| > M\}.$$

Otherwise, let $u \in B_n$ with $|u| \leq M$. Define $f \in \Phi'$ by

$$f[x] = \sum_{j=1}^n u_j < x, x_j >_q, \qquad \forall x \in \Phi.$$

Then $f \in A$ and

$$\sup_{\phi\in\Phi,\|\phi\|_q\leq 1} |f[\phi]| = \sup \left\{ \sum_{j=1}^n u_j < \phi, x_j >_q : \phi \in \Phi, \|\phi\|_q \leq 1 \right\} \leq M,$$

i.e., $f \in K^\epsilon$. This contradicts the fact that A is contained in $(K^\epsilon)^c$. Hence

$$A \subset \left\{ f \in \Phi' : \sum_{j=1}^n f[x_j]^2 > M^2 \right\}.$$

It follows from Lemma 2.3.3 that

$$\mu(A) \leq \frac{\sqrt{e}}{\sqrt{e}-1} \left(\frac{2}{\delta^2 M^2} \sum_{j=1}^n \|x_j\|_p^2 + \epsilon \right) \leq 4\epsilon.$$

Hence $\mu(C) \geq 1 - 4\epsilon$ for any $C \in \mathcal{C}$ containing K^ϵ and the countable additivity of μ follows from Lemma 2.3.2. ∎

Next we assume that X is a separable Hilbert space and identify X' with X by the Riesz representation theorem. In next theorem we consider the tightness for a sequence of probability measures on X.

Theorem 2.3.3 *Let $\{e_j\}$ be a CONS of X. $\{\mu_n\} \subset \mathcal{P}(X)$ is tight iff*
(a) For any $N \geq 1$,

$$\lim_{A\to\infty} \sup_n \mu_n \left\{ x \in X : \max_{1\leq i<N} | < x, e_i > | > A \right\} = 0$$

(b) For any $\delta > 0$

$$\lim_{N\to\infty} \sup_n \mu_n\{x \in X : r_N(x) \geq \delta\} = 0$$

where

$$r_N^2(x) = \sum_{j=N}^{\infty} < x, e_j >^2 .$$

Proof: "$\Rightarrow$" For any $\epsilon > 0$, let K_ϵ be a compact subset of X such that $\mu_n(K_\epsilon) > 1 - \epsilon$ for all $n \geq 1$. For any N there exists A such that

$$K_\epsilon \subset \left\{ x \in X : \max_{1\leq i<N} | < x, e_i > | \leq A \right\}$$

and hence (a) holds.

As $r_N(x)$ is uniformly continuous in $x \in K_\epsilon$ uniform for $N \geq 1$ and $r_N(x) \to 0$ as $N \to \infty$ for any $x \in X$. (b) follows easily.
"$\Leftarrow$" Let $N_0 = N_0(\epsilon, \delta)$ and $A_0 = A_0(\epsilon, \delta)$ be such that

$$\sup_n \mu_n\{x \in X : r_{N_0}(x) \geq \delta\} \leq \epsilon$$

and

$$\sup_n \mu_n \left\{ x \in X : \max_{1\leq i<N_0} | < x, e_i > | > A_0 \right\} \leq \epsilon.$$

Let $x^1, x^2, \cdots, x^{s(\delta)} \in X$ be such that

$$< x^j, e_i >= 0, \qquad \forall 1 \leq j \leq s(\delta) \text{ and } i \geq N_0$$

and for all $x \in X$ with $\max_{1\leq i<N_0} | < x, e_i > | \leq A_0$, we have

$$\min_{1\leq j<s(\delta)} \sum_{i=1}^{N_0-1} < x - x^j, e_i >^2 \leq \delta^2.$$

Therefore

$$\sup_n \mu_n \left\{ \cup_{j=1}^{s(\delta)} S(x^j, 2\delta) \right\} \geq 1 - 2\epsilon$$

where $S(x, \delta)$ is the sphere of radius δ around x. Replacing ϵ and δ by $2^{-m}\epsilon$ and m^{-1} respectively, we define

$$K_\epsilon = \cap_{m=1}^{\infty} \cup_{j=1}^{s(m^{-1})} S(x^j, 2m^{-1}).$$

Then K_ϵ is relatively compact in X and for any $n \geq 1$, we have

$$\mu_n(K_\epsilon^c) \leq \sum_{m=1}^{\infty} \left(1 - \mu_n \left\{ \cup_{j=1}^{s(m^{-1})} S(x^j, 2m^{-1}) \right\}\right) \leq \sum_{m=1}^{\infty} 2^{1-m}\epsilon = 2\epsilon.$$

i.e. $\{\mu_n\}$ is tight. ∎

The following corollary gives a convenient sufficient condition for the tightness of a sequence of probability measures on separable Hilbert space. The proof follows easily from Theorem 2.3.3.

Corollary 2.3.1 *Let $\{e_j\}$ be a CONS of X and $\{\mu_n\} \subset \mathcal{P}(X)$. If*

$$\sup_n \int r_1^2(x)\mu_n(dx) < \infty$$

and

$$\lim_{N\to\infty} \sup_n \int r_N^2(x)\mu_n(dx) = 0,$$

then $\{\mu_n\}$ is tight.

To study characteristic functions on X we introduce the S-topology.

Definition 2.3.3 *U is said to be an* **S-neighborhood of** $0 \in X$ *if there exists a positive definite self-adjoint nuclear operator S on X such that*

$$U = \{x \in X :< Sx, x >< 1\}.$$

Theorem 2.3.4 (Sazonov) *A complex-valued function F on X is the characteristic function of a $\mu \in \mathcal{P}(X)$ iff F satisfies the conditions (i), (iii) of Theorem 2.3.1 and F is continuous in S-topology.*

Proof: "$\Rightarrow$" Let K be a compact subset of X such that $\mu(K^c) < \epsilon$. Let S be the positive definite self-adjoint nuclear operator on X given by the following quadratic form

$$< Sx, x >= \epsilon^{-1} \int_K < x, y >^2 \mu(dy).$$

Then

$$|\hat{\mu}(y+h) - \hat{\mu}(y)| \le \int_X \left|e^{i<x,h>} - 1\right| \mu(dx)$$

$$= \int_X \sqrt{2(1 - cos < x, h >)}\mu(dx) \le \sqrt{\int_X 4\left|\sin\frac{< x, h >}{2}\right|^2 \mu(dx)}$$

$$\le \sqrt{\int_K < x, h >^2 \mu(dx) + 4\epsilon} = \sqrt{\epsilon < Sh, h > +4\epsilon} \le \sqrt{5\epsilon}$$

for h such that $< Sh, h >< 1$. Hence F is S-continuous.

"$\Leftarrow$" Let $\{e_j\}$ be a CONS of X. For any $n \ge 1$, let $J_n : \mathbf{R}^n \to X$ be given by $J_n u = \sum_{j=1}^n u_j e_j, \forall u \in \mathbf{R}^n$. Define $F_n(u) = F(J_n u), u \in \mathbf{R}^n$. It is easy to see that F_n is a characteristic function on $\mathbf{R}^n$. Let $\tilde{\mu}_n$ be the probability measure on $\mathbf{R}^n$ corresponding to F_n. Let $\mu_n = \tilde{\mu}_n \circ (J_n)^{-1}$. Then $\mu_n \in \mathcal{P}(X)$ and $\hat{\mu}_n = F_n$. Now we prove that $\{\mu_n\}$ is tight.

Note that, for any $N \geq 1$,

$$\begin{aligned} & \lim_{A\to\infty} \sup_n \mu_n \left\{ x \in X : \max_{1\leq i<N} | < x, e_i > | \leq A \right\} \\ = & \lim_{A\to\infty} \max_{1\leq n<N} \tilde{\mu}_n \left\{ u \in \mathbf{R}^n : \max_{1\leq i\leq n} |u_i| \leq A \right\} \\ = & \ 1. \end{aligned} \tag{2.3.5}$$

On the other hand, for any $\epsilon > 0$, let S_ϵ be a positive definite self-adjoint nuclear operator on X such that $< S_\epsilon y, y >< 1$ implies $|F(y) - 1| < \epsilon$. Therefore

$$|F(y) - 1| \leq 2 < S_\epsilon y, y > +\epsilon, \qquad \forall y \in X.$$

Hence

$$\begin{aligned} |\hat{\mu}_n(y) - 1| &= \left| F\left(\sum_{j=1}^n < y, e_j > e_j \right) - 1 \right| \\ &\leq 2 \left\| \sqrt{S_\epsilon} \sum_{j=1}^n < y, e_j > e_j \right\|^2 + \epsilon, \ \forall y \in X. \end{aligned} \tag{2.3.6}$$

It is clear that for $n < N$,

$$\mu_n\{x \in X : r_N(x) \geq \delta\} = 0. \tag{2.3.7}$$

We assume that $n \geq N$. It follows from (2.3.6) and Lemma 2.3.3 that

$$\begin{aligned} & \mu_n\{x \in X : r_N(x) \geq \delta\} = \mu_n \left\{ x \in X : \sum_{j=N}^n < x, e_j >^2 \geq \delta^2 \right\} \\ \leq & \ \frac{\sqrt{e}}{\sqrt{e}-1} \left(\frac{2}{\delta^2} \sum_{j=N}^n \left\| \sqrt{S_\epsilon} e_j \right\|^2 + \epsilon \right) \\ \leq & \ \frac{\sqrt{e}}{\sqrt{e}-1} \left(\frac{2}{\delta^2} \sum_{j=N}^\infty \left\| \sqrt{S_\epsilon} e_j \right\|^2 + \epsilon \right). \end{aligned} \tag{2.3.8}$$

By (2.3.7) and (2.3.8), we then have

$$\limsup_{N\to\infty} \sup_n \mu_n\{x \in X : r_N(x) \geq \delta\} \leq \frac{\sqrt{e}}{\sqrt{e}-1}\epsilon. \tag{2.3.9}$$

It follows from (2.3.5), (2.3.9), Theorem 2.3.3 and ϵ is arbitrary that $\{\mu_n\}$ is tight.

Let μ be a cluster point of $\{\mu_n\}$, i.e. there exists a subsequence $\{\mu_{n_k}\}$ converges to μ . It is easy to see that for any $x \in X$

$$\hat{\mu}(x) = \lim_{k\to\infty} \hat{\mu}_{n_k}(x) = \lim_{k\to\infty} F\left(\sum_{j=1}^{n_k} < x, e_j > e_j\right) = F(x).$$

This proves the assertion of Sazonov's theorem. ∎

Finally we give an example of a Bochner functional which is not a characteristic function.

Example 2.3.1 *Let H be an infinite dimensional Hilbert space with norm $\|\cdot\|$. For any $x \in H$, let $F(x) = exp(-\frac{1}{2}\|x\|^2)$.*

It is clear that $F(0) = 1$ and F is sequentially continuous at $0 \in X$. Let $\{\xi_j\}$ be a sequence of i.i.d. random variables with common standard normal distribution and let $\{e_j\}$ be a CONS of H. For each $x \in H$, let

$$\eta(x) = \sum_{j=1}^{\infty} < x, e_j > \xi_j.$$

Then, $\forall n \in \mathbf{N}$, $x_j \in X$, α_j complex, $j = 1, 2, \cdots, n$ we have

$$\begin{aligned}\sum_{j,k=1}^{n} F(x_j - x_k)\alpha_j\bar{\alpha}_k &= \sum_{j,k=1}^{n} E\exp\left(i\eta(x_j - x_k)\right)\alpha_j\bar{\alpha}_k \\ = \quad E\left|\sum_{j=1}^{n}\exp\left(i\eta(x_j)\right)\alpha_j\right|^2 &\geq 0.\end{aligned}$$

It follows from Theorem 2.3.1 that F is a Bochner functional on H. Now we show that F is not continuous in S-topology. Otherwise $x \in H \to \|x\| \in \mathbf{R}$ is continuous in S-topology and hence, for any $\epsilon > 0$, there exists positive definite self-adjoint Hilbert-Schmidt operator Q_ϵ on H such that $\|Q_\epsilon x\| \leq 1$ implies $\|x\| \leq \epsilon$. Therefore

$$\|x\| \leq \epsilon\|Q_\epsilon x\| \qquad \forall x \in H.$$

Hence $(Q_\epsilon)^{-1}$ is a bounded linear functional on H and, by Theorem 1.2.3, the identity map $id_H = (Q_\epsilon)^{-1}Q_\epsilon$ on H is Hilbert-Schmidt. This contradicts the fact that H is of infinite dimension. Hence, by Sazonov's theorem F is not a characteristic function.

2.4 $C([0,T],\Phi')$ and $D([0,T],\Phi')$

Let X be a topological vector space whose topology is given by a family of seminorms $\{\|\cdot\|_v : v \in \Gamma\}$. In this section we fix $T > 0$ and let $C([0,T],X)$ (resp. $D([0,T],X)$ be the collection of all continuous (resp. right continuous with left limit) maps from $[0,T]$ to X.

There is a large class of stochastic processes arising from practical problems whose sample paths are either continuous or right continuous with left limits. To study the convergence property for these processes, we need to investigate the structures of the spaces $C([0,T],X)$ and $D([0,T],X)$ in which the sample paths of these processes can be regarded as points. In most applications, X can be chosen as $\mathbf{R}^n$, a Hilbert space, a Banach space or the dual of a countable Hilbertian nuclear space. We shall state the results for both spaces but leave the proof for the continuous space case to the reader.

Let Λ be the set of strictly increasing continuous maps λ from [0,T] onto itself such that

$$\gamma(\lambda) = \sup_{0\le s<t\le T} \left|\log \frac{\lambda(s)-\lambda(t)}{s-t}\right| < \infty.$$

For any $v \in \Gamma$, let

$$d_v(f,g) = \inf_{\lambda\in\Lambda} \sup_{0\le t\le T} \{\|f(t) - g(\lambda(t))\|_v + \gamma(\lambda)\}, \quad \forall f,\ g \in D([0,T],X).$$

It is easy to see that for any $v \in \Gamma$, d_v is a pseudometric on $D([0,T],X)$. We define the topology of $D([0,T],X)$ by this family $\{d_v : v \in \Gamma\}$ of pseudometrics.

Theorem 2.4.1 *Let $\{f_n\} \subset D([0,T],X)$ and $f \in D([0,T],X)$. Then for any $v \in \Gamma$, the following statements are equivalent:*
a) $d_v(f_n, f) \to 0$, as $n \to \infty$.
b) There exists $\{\lambda_n\} \subset \Lambda$ such that

$$\sup_{0\le t\le T} \|f_n(t) - f(\lambda_n(t))\|_v + \gamma(\lambda_n) \to 0.$$

c) There exists $\{\lambda_n\} \subset \Lambda$ such that

$$\sup_{0\le t\le T} \|f_n(t) - f(\lambda_n(t))\|_v + \sup_{0\le t\le T} |\lambda_n(t) - t| \to 0.$$

Proof: By the definition of $d_v(f_n, f)$, there exists $\{\lambda_n\} \subset \Lambda$ such that

$$d_v(f_n, f) \le \sup_{0\le t\le T} \|f_n(t)) - f(\lambda_n(t))\|_v + \gamma(\lambda_n) \le d_v(f_n, f) + n^{-1}.$$

Therefore a) and b) are equivalent. "*b*) ⇒ *c*)" follows from the inequality

$$\sup_{0\le t\le T} |\lambda_n(t) - t| \le T \max\left\{e^{\gamma(\lambda_n)} - 1, 1 - e^{-\gamma(\lambda_n)}\right\}. \tag{2.4.1}$$

Finally, we prove "*c*) ⇒ *b*)". Let $\{\lambda_n\}$ be given by c). Then for $m \ge 1$, there exists N_m such that

$$\sup_{0\le t\le T} \|f_n(t) - f(\lambda_n(t))\|_v + \sup_{0\le t\le T} |\lambda_n(t) - t| < m^{-1}, \quad \forall n \ge N_m.$$

Let $\tau_0 = 0$ and for $k \ge 1$

$$\tau_k = \inf\left\{t \in (\tau_{k-1}, T] : \|f(t) - f(\tau_{k-1})\|_v > m^{-1}\right\} \tag{2.4.2}$$

with the convention that the infimum over the empty set is equal to T. As $f \in D([0,T], X)$ it is easy to see that there exists k_0 such that $\tau_{k_0} = T$. Denote $(\lambda_n)^{-1}\tau_k$ by $\eta_{k,n}$ and define $\mu_{n,m} \in \Lambda$ to be piecewise linear on [0,T] such that $\mu_{n,m}(\eta_{k,n}) = \tau_k$ for all $k \le k_0$. Note that for $k \le k_0$,

$$\lim_{n\to\infty} |\eta_{k,n} - \tau_k| = \lim_{n\to\infty} |\eta_{k,n} - \lambda_n(\eta_{k,n})| = 0,$$

and hence

$$\gamma(\mu_{n,m}) = \max\left\{\left|\log \frac{\eta_{k+1,n} - \eta_{k,n}}{\tau_{k+1} - \tau_k}\right| : 0 \le k < k_0\right\} \to 0$$

as $n \to \infty$. Therefore there exists $\tilde{N}_m$ such that

$$\gamma(\mu_{n,m}) < m^{-1}, \quad \forall n \ge \tilde{N}_m. \tag{2.4.3}$$

Further for $n \ge N_m$

$$\begin{aligned}
&\sup_{0\le t\le T} \|f_n(t) - f(\mu_{n,m}(t))\|_v \\
\le\ & \sup_{0\le t\le T} \|f_n(t) - f(\lambda_n(t))\|_v + \sup_{0\le t\le T} \|f(\lambda_n(t)) - f(\mu_{n,m}(t))\|_v \\
\le\ & m^{-1} + \max_{0\le k<k_0} \sup_{\eta_{k,n}\le t<\eta_{k+1,n}} \|f(\lambda_n(t)) - f(\mu_{n,m}(t))\|_v \\
\le\ & m^{-1} + \max_{0\le k<k_0} \sup_{\eta_{k,n}\le t<\eta_{k+1,n}} (\|f(\lambda_n(t)) - f(\tau_k)\|_v \\
& \qquad + \|f(\mu_{n,m}(t)) - f(\tau_k)\|_v) \\
\le\ & m^{-1} + 2\max_{0\le k<k_0} \sup_{\tau_k\le t<\tau_{k+1}} \|f(t) - f(\tau_k)\|_v \le 3m^{-1}.
\end{aligned} \tag{2.4.4}$$

Let $n_m = \max(N_m, \tilde{N}_m)$. Without loss of generality, we assume that the sequence $\{n_m\}$ strictly increases. Let $\tilde{\lambda}_n = \mu_{n,m}$ for m such that $n_m \le n < n_{m+1}$. Then $\tilde{\lambda}_n \in \Lambda$ and, by (2.4.3) and (2.4.4) we have

$$\sup_{0\le t\le T} \|f_n(t) - f(\tilde{\lambda}_n(t))\|_v + \gamma(\tilde{\lambda}_n) \le 4m^{-1} \to 0$$

where $m \to \infty$ such that $n_m \leq n < n_{m+1}$. ∎

Corollary 2.4.1 *If $f_n \to f$ in $D([0,T],X)$, then*
i) for any continuity point $t \in [0,T]$ of f, $f_n(t) \to f(t)$ in X;
ii) for any $v \in \Gamma$

$$\sup_{0\leq t\leq T} \|f_n(t)\|_v \to \sup_{0\leq t\leq T} \|f(t))\|_v \qquad in\ \mathbf{R}.$$

Proof: i) follows from c) of Theorem 2.4.1 directly. For ii), let $\{\lambda_n\}$ be given by c) of Theorem 2.4.1. Note that

$$\begin{aligned} & \left| \sup_{0\leq t\leq T} \|f_n(t)\|_v - \sup_{0\leq t\leq T} \|f(t)\|_v \right| \\ = & \left| \sup_{0\leq t\leq T} \|f_n(t)\|_v - \sup_{0\leq t\leq T} \|f(\lambda_n(t))\|_v \right| \\ \leq & \sup_{0\leq t\leq T} \|f_n(t) - f(\lambda_n(t))\|_v \to 0. \end{aligned}$$ ∎

For any $v \in \Gamma$, let

$$\|f\|_{(v)} = \sup_{0\leq t\leq T} \|f(t)\|_v, \qquad \forall f \in C([0,T],X).$$

It is easy to see that for any $v \in \Gamma$, $\|\cdot\|_{(v)}$ is a seminorm on $C([0,T],X)$. We define the topology of $C([0,T],X)$ by this family of seminorms.

Theorem 2.4.2 *i) $C([0,T],X)$ is a sequentially complete topological vector space. Further, if X is a (separable) Banach space, then so is $C([0,T],X)$.*
ii) $D([0,T],X)$ is sequentially complete. Further, if X is a (separable) Banach space, then $D([0,T],X)$ is a complete (separable) metric space.

Proof: We only prove the results for $D([0,T],X)$. Let $\{f_n\} \subset D([0,T],X)$ be a Cauchy sequence. Then $\forall v \in \Gamma$, there exists a strictly increasing sequence $\{n_k\}$ such that

$$d_v(f_n, f_m) \leq 2^{-k-1}, \qquad \forall n, m \geq n_k. \tag{2.4.5}$$

Let $g_k = f_{n_k}$. Then there exists $\{\lambda_k\} \subset \Lambda$ such that

$$\sup_{0\leq t\leq T} \|g_{k+1}(t) - g_k(\lambda_k(t))\|_v + \gamma(\lambda_k) \leq 2^{-k}. \tag{2.4.6}$$

By (2.4.1) and (2.4.6) we can easily prove that ($\circ$ denoting composition)

$$\mu_k(t) \equiv \lim_{n\to\infty} (\lambda_k \circ \lambda_{k+1} \circ \cdots \circ \lambda_{k+n})(t)$$

exists uniformly for $t \in [0,T]$ and

$$\gamma(\mu_k) \leq \sum_{j=k}^{\infty} \gamma(\lambda_j) \leq 2^{1-k}. \tag{2.4.7}$$

By (2.4.6) again, we have

$$\begin{aligned} &\sup_{0\leq t\leq T} \|g_{k+1}(\mu_{k+1}(t)) - g_k(\mu_k(t))\|_v \\ = &\sup_{0\leq t\leq T} \|g_{k+1}(\mu_{k+1}(t)) - g_k(\lambda_k(\mu_{k+1}(t)))\|_v \leq 2^{-k}. \end{aligned} \tag{2.4.8}$$

Therefore the limit, denoted by $g(t)$, of $g_k(\mu_k(t))$ exists uniformly for $t \in [0,T]$ as $k \to \infty$ and hence g is right continuous with left limit. It follows from (2.4.7) and (2.4.8) that $g_k \to g$ in $D([0,T],X)$. Hence by (2.4.5), $f_n \to g$ in $D([0,T],X)$, i.e., $D([0,T],X)$ is sequentially complete.

If X is a Banach space, then it follows from Theorem 1.1.3 that $D([0,T], X)$ is a complete metric space. Now suppose that X is separable and let $\mathbf{T}_0$ and X_0 be countable dense subset of $[0,T]$ and X respectively. Let A be the collection of all functions of the form

$$f(t) = x_k, \qquad t \in [t_{k-1}, t_k),\ k = 1, 2, \cdots, n$$

where $t_0 = 0, t_n = T, \{t_k\}_{1\leq k<n} \subset \mathbf{T}_0$ is an increasing sequence and $\{x_k\}_{1\leq k\leq n} \subset X_0$. Then A is countable. For any $f \in D([0,T],X)$, $m \geq 1$, we define $0 = \tau_0 < \tau_1 < \cdots < \tau_{k_0} = T$ by (2.4.2). Taking $t_0 = 0$, $t_{k_0} = T$, $t_j \in \mathbf{T}_0$, $1 \leq j < k_0$, such that $0 \leq \tau_j - t_j \leq m^{-1}\min(\tau_i - \tau_{i-1} : 1 \leq i \leq k_0)$ and $x_j \in X_0, 1 \leq j \leq k_0$, such that $\|f(\tau_{j-1}) - x_j\| < m^{-1}$. Let

$$f_m(t) = x_k, \qquad t \in [t_{k-1}, t_k),\ k = 1, 2, \cdots, k_0$$

and define $\lambda_m \in \Lambda$ to be piecewise linear such that $\lambda_m(t_j) = \tau_j$, $0 \leq j \leq k_0$. It is easy to see that $f_m \in A$,

$$\gamma(\lambda_m) \leq \log\frac{m}{m-1} \quad \text{and} \quad \sup_{0\leq t\leq T} \|f_m(t) - f(\lambda_m(t))\| \leq 2m^{-1},$$

i.e. A is dense in $D([0,T],X)$. ∎

The next two theorems give criteria for the compactness of the subsets of $D([0,T],X)$ and $C([0,T],X)$. First we suppose that X is a Banach space. To characterize the compact sets of $C([0,T],X)$, we define the following moduli: $\forall\delta > 0$ and $f \in C([0,T],X)$

$$w_f(\delta; X) = \sup\{\|f(t) - f(s)\| : s, t \in [0,T] \text{ and } |t-s| < \delta\}$$

To characterize the compact sets of $D([0,T],X)$, we define the following moduli: $\forall\delta > 0$ and $f \in D([0,T],X)$

$$w'_f(\delta;X) = \inf_{\{t_i\}} \max_{1\le i\le n} \sup\{\|f(t)-f(s)\| : s,t \in [t_{i-1},t_i)\}$$

where the infimum is taken over the finite partitions $0 = t_0 < t_1 < \cdots < t_n = T$, $t_i - t_{i-1} > \delta, i = 1,2,\cdots,n$.

Lemma 2.4.1 *Let K be a compact subset of X and $\delta > 0$. Define $A(K,\delta)$ to be the collection of $f \in D([0,T],X)$ which is of the form $f(t) = x_j$ for $t \in [t_{j-1},t_j)$, $j = 1,\cdots,m$, where $t_j - t_{j-1} \ge \delta, x_j \in K, t_0 = 0$ and $t_m = T$. Then $A(K,\delta)$ is relatively compact.*

Proof: Let $\{f_n\}$ be a sequence in $A(K,\delta)$. As $m(f_n) \le T/\delta$, $\forall n \ge 1$, taking a subsequence if necessary, we assume that $m(f_n) = m$ for all $n \ge 1$. By a diagonal argument there exists a subsequence $\{f_{n_k}\}$ such that $\forall j = 1,2,\cdots,m, t_j(f_{n_k}) \to t_j$ and $x_j(f_{n_k}) \to x_j \in K$. Since $t_j(f_{n_k}) - t_{j-1}(f_{n_k}) \ge \delta$, we have $t_j - t_{j-1} \ge \delta$, $\forall j = 1,2,\cdots,m$. Define f as in the statement of the lemma and define $\lambda_{n_k} \in \Lambda$ to be piecewise linear such that $\lambda_{n_k}(t_j(f_{n_k})) = t_j, j = 0,1,\cdots,m$. Then

$$\begin{aligned}
&\gamma(\lambda_{n_k}) + \sup_{0\le t\le T} \|f_{n_k}(t) - f(\lambda_{n_k}(t))\| \\
\le\; & \max_{1\le j\le m} \left|\log \frac{t_j - t_{j-1}}{t_j(f_{n_k}) - t_{j-1}(f_{n_k})}\right| \\
&+ \max_{1\le j\le m} \sup\{\|f_{n_k}(t) - f(\lambda_{n_k}(t))\| : t_{j-1}(f_{n_k}) \le t < t_j(f_{n_k})\} \\
\to\; & 0, \qquad \text{as } k \to \infty.
\end{aligned}$$

i.e. $A(K,\delta)$ is relatively compact. ∎

Theorem 2.4.3 *Let X be a Banach space. Then*
i) $A \subset C([0,T],X)$ is relatively compact iff
a) There exists a relatively compact subset B of X such that $f(t) \in B$ for any $t \in [0,T]$, $f \in A$,
b) $\sup\{w_f(\delta;X) : f \in A\}$ as $\delta \to 0+$.
ii) $A \subset D([0,T],X)$ is relatively compact iff
a) There exists a relatively compact subset B of X such that $f(t) \in B$ for any $t \in [0,T]$, $f \in A$,
b) $\sup\{w'_f(\delta;X) : f \in A\}$ as $\delta \to 0+$.

Proof: ii) "$\Leftarrow$" For any $m \in \mathbf{N}$, let $\delta_m > 0$ be such that

$$\sup\{w'_f(\delta_m;X) : f \in A\} \le m^{-1}.$$

Let $K = \bar{B}$ and $A_m = A(K, \delta_m)$. Then by Lemma 2.4.1, A_m is relatively compact in $D([0,T],X)$. For any $f \in A$, there exists a partition $0 = t_0 < t_1 < \cdots < t_n = T$ such that $t_i - t_{i-1} > \delta_m$ and $\|f(t) - f(s)\| \le 2m^{-1}$, $\forall s, t \in [t_{i-1}, t_i), i = 1, 2, \cdots, n$. Define $\tilde{f}(t) = f(t_{i-1})$ for $t \in [t_{i-1}, t_i), i = 1, 2, \cdots, n$. Then for any $f \in A$ we have $\tilde{f} \in A_m$ and

$$d(\tilde{f}, f) \le \sup_{0\le t\le T} \|f(t) - \tilde{f}(t)\| \le 2m^{-1}.$$

Therefore $A \subset \cap_{m=1}^{\infty}\{f \in D([0,T],X) : d(f, A_m) \le 2m^{-1}\}$ which is compact and hence, A is relatively compact.
"$\Rightarrow$" a) Let

$$B = \{x \in X : x = f(t) \text{ or } x = f(t-) \text{ for some } t \in [0,T] \text{ and } f \in A\}.$$

For any sequence $\{x_n\}$ in B, we have $x_n = f_n(t_n)$ or $x_n = f_n(t_n-)$ for some $t_n \in [0,T]$ and $f_n \in A$. We may assume $x_n = f_n(t_n)$. Otherwise we only need to replace x_n by $y_n = f_n(s_n)$ for some $s_n < t_n$ to be sufficiently close to t_n. By the compactness of $[0,T]$ and A, without loss of generality we may assume that there exist $t \in [0,T]$, $f \in D([0,T],X)$ such that $t_n \to t$ and $f_n \to f$. By Theorem 2.4.1, there exists $\{\lambda_n\} \subset \Lambda$ such that

$$\sup_{0\le s\le T} \|f_n(s) - f(\lambda_n(s))\| + \sup_{0\le s\le T} |\lambda_n(s) - s| \to 0.$$

Note that since

$$\begin{aligned} & \|x_n - f(t)\| \wedge \|x_n - f(t-)\| \\ \le & \sup_{0\le s\le T} \|f_n(s) - f(\lambda_n(s))\| \\ & + \|f(\lambda_n(t_n)) - f(t)\| \wedge \|f(\lambda_n(t_n)) - f(t-)\| \to 0, \qquad (2.4.9) \end{aligned}$$

$\{x_n\}$ has either $f(t)$ or $f(t-)$ as a cluster point. Therefore B is relatively compact in X.
b) If b) is not satisfied, there exist $\epsilon_0 > 0$, $f_n \in A$, $\delta_n \to 0$ such that $w'_{f_n}(\delta_n; X) \ge \epsilon_0$. Without loss of generality, we assume that f_n converges to some f in $D([0,T],X)$. Let τ_k be defined by (2.4.2) with m^{-1} replaced by $\epsilon_0/4$. Then $0 = \tau_0 < \tau_1 < \cdots < \tau_{k_0} = T$ is a partition of $[0,T]$. For any n sufficiently large such that $\delta_n < \min\{\tau_k - \tau_{k-1} : k = 1, 2, \cdots, k_0\}$ we have

$$\epsilon_0 \le w'_{f_n}(\delta_n; X) \le \max_{1\le k\le k_0} \sup\{\|f_n(t) - f_n(s)\| : s, t \in [\tau_{k-1}, \tau_k)\}.$$

Therefore there exist an integer k, a subsequence $\{f_{n_j}\}$ of $\{f_n\}$ and two convergent sequences $\{t_{n_j}\}, \{s_{n_j}\} \subset [\tau_{k-1}, \tau_k)$ such that

$$\|f_{n_j}(t_{n_j}) - f_{n_j}(s_{n_j})\| \ge 3\epsilon_0/4.$$

It follows from (2.4.9) that there exist $s,t \in [\tau_{k-1},\tau_k], x = f(s-)$ or $f(s)$, $y = f(t-)$ or $f(t)$ such that $x, y \neq f(\tau_k)$ and $\|f_{n_j}(t_{n_j}) - f_{n_j}(s_{n_j})\| \to \|x-y\|$. Hence $\|x - y\| \geq 3\epsilon_0/4$. On the other hand, by the definition of τ_k, we have $\|x - y\| \leq \epsilon_0/2$. This contradiction verifies that b) holds. ∎

Corollary 2.4.2 *Let X, Y be two Banach spaces and let $g : X \to Y$ be a continuous map vanishing in a neighborhood of $0 \in X$. Then the map G from $D([0,T],X)$ to $D([0,T],Y)$ given by*

$$(Gf)(t) = \sum_{s\leq t} g(\Delta f(s)) \tag{2.4.10}$$

is continuous, where $\Delta f(s) = f(s) - f(s-)$.

Proof: Let f_n converge to f in $D([0,T],X)$. As the set $\{r > 0 : \exists t \in [0,T]$ s.t. $\|\Delta f(t)\| = r\}$ is countable and g vanishes in a neighborhood of 0, g vanishes in $\{x \in X : \|x\| \leq r\}$ for some $r > 0$ and the set $\{t \in [0,T] : \|\Delta f(t)\| = r\}$ is empty. For any $h \in D([0,T],X)$, it is easy to see that the set $\{t \in [0,T] : \|\Delta h(t)\| > r\}$ contains only finitely many elements and we denote it by

$$J_h = \left\{t^1(h) < t^2(h) < \cdots < t^{m(h)}(h)\right\}.$$

As a consequence, Gf in (2.4.10) is a well-defined element in $D([0,T],Y)$.

Let $\{\lambda_n\} \subset \Lambda$ be such that

$$\sup_{0\leq t\leq T} \|f_n(t) - f(\lambda_n(t))\| + \sup_{0\leq t\leq T} |\lambda_n(t) - t| \to 0. \tag{2.4.11}$$

First we prove that there exists $\delta > 0$ such that

$$t^j(f_n) - t^{j-1}(f_n) \geq \delta, \qquad \forall n \geq 1 \text{ and } 1 < j \leq m(f_n). \tag{2.4.12}$$

If (2.4.12) does not hold, then for some n_k, $2 \leq j_k \leq m(f_{n_k})$ such that $t^{j_k}(f_{n_k}) - t^{j_k-1}(f_{n_k}) < k^{-1}$ for $k \geq 1$. Then for any partition in the definition of $w'_{f_{n_k}}(k^{-1}, X)$, there exists j (equals to j_k or $j_k - 1$) such that $t^j(f_{n_k})$ is in the interior of one of the partition subintervals and hence

$$w'_{f_{n_k}}(k^{-1}, X) \geq \|\Delta f_{n_k}(t^j(f_{n_k}))\| \geq r.$$

By Theorem 2.4.3, the above inequality contradicts the fact that $\{f_n\}_{n\geq1}$ is relatively compact. Therefore (2.4.12) holds.

For $1 \leq j \leq m(f)$, as

$$\|\Delta f_n(\lambda_n^{-1}(t^j(f))) - \Delta f(t^j(f))\| \leq 2 \sup_{0\leq t\leq T} \|f_n(\lambda_n^{-1}(t)) - f(t)\| \to 0 \tag{2.4.13}$$

we see that there exists n_j such that $\lambda_n^{-1}(t^j(f)) \in J_{f_n}$, $\forall n \geq n_j$. As a consequence of (2.4.12) and (2.4.13), there exists N such that $m(f_n) \geq m(f)$, $\forall n \geq N$.

Proceeding similarly we have that for any $1 \leq j \leq m(f)$

$$\liminf_{n\to\infty} \|\Delta f(\lambda_n(t^j(f_n)))\| = \liminf_{n\to\infty} \|\Delta f_n(t^j(f_n))\| \geq r. \tag{2.4.14}$$

If we have a subsequence n_k such that $\lambda_{n_k}(t^j(f_{n_k}))$ decreases (or increases) to t, it is easy to see that $\Delta f(\lambda_{n_k}(t^j(f_{n_k}))) \to 0$ which contradicts (2.4.14). Therefore there exist n_j and t such that $\lambda_n(t^j(f_n)) = t$ for all $n \geq n_j$. By (2.4.14) again we see that $t \in J_f$. Therefore the collection of all cluster points of the set $\{m^j(f_n) : 1 \leq j \leq m(f), n \geq 1\}$ is contained in J_f.

Based on the facts obtained above, it is easy to see that for sufficiently large n, $m(f_n) = m(f)$ and

$$t^j(f_n) \to t^j(f) \text{ and } \Delta f_n(t^j(f_n)) \to \Delta f(t^j(f)), \quad 1 \leq j \leq m(f). \tag{2.4.15}$$

The conclusion of the corollary follows immediately from (2.4.15). ∎

Next, we consider the case of $X = \Phi'$, the dual of a CHNS. To characterize the compact sets of $C([0,T],\Phi')$, we define the following moduli:
(a) For any $f \in C([0,T],\Phi')$, $\phi \in \Phi$, $\delta > 0$, let $w_f(\delta,\phi) = w_{f(\cdot)[\phi]}(\delta;,\mathbf{R})$.
(b) For any $f \in C([0,T],\Phi_{-p})$, $\delta > 0$, let $w_f(\delta,p) = w_f(\delta;\Phi_{-p})$.
(c) For any $f \in C([0,T],\mathbf{R})$, $\delta > 0$, let $w_f(\delta) = w_f(\delta;\mathbf{R})$.

To characterize the compact sets of $D([0,T],\Phi')$, we define the following moduli:
(a) For any $f \in D([0,T],\Phi')$, $\phi \in \Phi$, $\delta > 0$, let $w'_f(\delta,\phi) = w'_{f(\cdot)[\phi]}(\delta;\mathbf{R})$.
(b) For any $f \in D([0,T],\Phi_{-p})$, $\delta > 0$, let $w'_f(\delta,p) = w'_f(\delta;\Phi_{-p})$.
(c) For any $f \in D([0,T],\mathbf{R})$, $\delta > 0$, let $w'_f(\delta) = w'_f(\delta;\mathbf{R})$.

The next two results, due to Mitoma [41], will be used extensively in the rest of this book.

Theorem 2.4.4 (Mitoma) *The following statements are equivalent:*
a) A is relatively compact in $D([0,T],\Phi')$ (resp. $C([0,T],\Phi')$).
b) For any $\phi \in \Phi$, $\{f(\cdot)[\phi] : f \in A\}$ is relatively compact in $D([0,T],\mathbf{R})$ (resp. $C([0,T],\mathbf{R})$).
c) There exists $p \in \mathbf{N}$ such that A is relatively compact in $D([0,T],\Phi_{-p})$ (resp. $C([0,T],\Phi_{-p})$).

Proof: For $\phi \in \Phi$, it is easy to see that the map $\pi_\phi : D([0,T],\Phi') \to D([0,T],\mathbf{R})$ is continuous. Also the canonical injection from $D([0,T],\Phi_{-p})$ to $D([0,T],\Phi')$ is continuous. Therefore $(c) \Rightarrow (a) \Rightarrow (b)$ follows immediately. Now we show that $(b) \Rightarrow (c)$.

Applying Theorem 2.4.3 to the relatively compact set $\{f(\cdot)[\phi] : f \in A\}$ with $X = \mathbf{R}$, we have

$$V(\phi) \equiv \sup_{f\in A}\sup_{0\le t\le T} |f(t)[\phi]| < \infty, \ \forall \phi \in \Phi.$$

It is easy to verify that V satisfies the conditions of Lemma 1.3.1 and hence, there exist $\theta > 0$ and $r \geq 0$ such that

$$V(\phi) \le \theta\|\phi\|_r, \qquad \forall\phi\in\Phi.$$

By the nuclearity of Φ, there exist $q \geq p \geq r$ such that the canonical injections $\Phi_q \to \Phi_p \to \Phi_r$ are Hilbert-Schmidt. Let $\{\phi_j^p\}$, $\{\phi_j^q\} \subset \Phi$ be CONS' of Φ_p and Φ_q respectively. Define

$$M^2 = \theta^2 \sum_j \|\phi_j^p\|_r^2 \quad \text{and} \quad B = \left\{ x \in \Phi' : \sum_j x[\phi_j^p]^2 \le M^2 \right\}.$$

Then $B \subset \Phi_{-p}$ and B is compact in Φ_{-q}. Note that

$$\begin{aligned}\sup_{f\in A}\sup_{0\le t\le T}\sum_j |f(t)[\phi_j^p]|^2 &\le \sum_j \sup_{f\in A}\sup_{0\le t\le T} |f(t)[\phi_j^p]|^2 \\ &\le \theta^2\sum_j \|\phi_j^p\|_r^2 = M^2,\end{aligned}$$

i.e., $f(t) \in B$, $\forall t\in[0,T]$, $f\in A$, and hence A satisfies the first condition of Theorem 2.4.3 ii).

On the other hand, for $j \geq 1$, note that

$$\lim_{\delta\to 0+}\sup_{f\in A} w'_f(\delta,\phi_j^q) = 0$$

and

$$\sup_{f\in A} w'_f(\delta,\phi_j^q)^2 \le \sup_{f\in A}\sup_{0\le t\le T} 4|f(t)[\phi_j^q]|^2 \le 4M^2\|\phi_j^q\|_p^2$$

is summable. Hence, by the dominated convergence theorem,

$$\begin{aligned}\lim_{\delta\to 0+}\sup_{f\in A} w'_f(\delta,q)^2 &\le \lim_{\delta\to 0+}\sum_j \sup_{f\in A} w'_f(\delta,\phi_j^q)^2 \\ &= \sum_j \lim_{\delta\to 0+}\sup_{f\in A} w'_f(\delta,\phi_j^q) = 0. \qquad (2.4.16)\end{aligned}$$

Therefore $A \subset D([0,T],\Phi_{-q})$ and by Theorem 2.4.3, A is relatively compact in $D([0,T],\Phi_{-q})$. ∎

2.5 Probability measures on $D([0,T],\Phi')$

In this section, we study the weak compactness for sequences of Borel probability measures on $C([0,T],\Phi')$ and $D([0,T],\Phi')$. We need the following lemma.

Lemma 2.5.1 *Let $\{\mu_n\} \subset \mathcal{P}(D([0,T],\Phi'))$ be such that for any $\phi \in \Phi$, $\{\mu_n \pi_\phi^{-1}\}$ is tight in $\mathcal{P}(C([0,T],\mathbf{R}))$. Then for any $\epsilon > 0$ there exist $p \in \mathbf{N}$, $M > 0$ such that*

$$\mu_n \left\{ f \in D([0,T],\Phi') : \sup_{t\in[0,T]} \|f(t)\|_{-p} \leq M \right\} \geq 1-\epsilon, \ \forall n \geq 1. \qquad (2.5.1)$$

Proof: Let $\mathcal{X} = D([0,T],\Phi')$ and

$$V(\phi) = \sup_{n\geq 1} \int_{\mathcal{X}} \frac{\sup_t |f(t)[\phi]|}{1+\sup_t |f(t)[\phi]|} \mu_n(df), \quad \phi \in \Phi.$$

It is easy to verify that V satisfies the conditions (1), (2) and the first half of (3) in Lemma 1.3.1. To apply that lemma we only need to show that $V(m^{-1}\phi) \to 0$ as $m \to \infty$ for any $\phi \in \Phi$. It follows from the tightness of $\{\mu_n \pi_\phi^{-1}\}$ and Theorem 2.4.3 that $\forall \eta > 0$, there exists $m(\eta)$ such that

$$\mu_n \left\{ f \in D([0,T],\Phi') : \sup_{t\in[0,T]} |f(t)[\phi]| \leq \sqrt{m(\eta)} \right\} \geq 1-\eta, \ \forall n \geq 1.$$

Hence for any $m \geq m(\eta)$ we have

$$\begin{aligned} V(m^{-1}\phi) &= \sup_{n\geq 1} \left(\int_{\mathcal{X}} \frac{\sup_t |f(t)[\phi/m]|}{1+\sup_t |f(t)[\phi/m]|} 1_{\sup_t |f(t)[\phi]| \leq \sqrt{m(\eta)}} \mu_n(df) \right. \\ &\quad \left. + \int_{\mathcal{X}} \frac{\sup_t |f(t)[\phi/m]|}{1+\sup_t |f(t)[\phi/m]|} 1_{\sup_t |f(t)[\phi]| > \sqrt{m(\eta)}} \mu_n(df) \right) \\ &\leq \frac{1}{1+\sqrt{m}} + \sup_{n\geq 1} \mu_n \left\{ \sup_t |f(t)[\phi]| > \sqrt{m(\eta)} \right\} \\ &\leq \frac{1}{1+\sqrt{m}} + \eta. \end{aligned}$$

Therefore $V(m^{-1}\phi) \to 0$ as $m \to \infty$. It follows from Lemma 1.3.1 that V is continuous in Φ. Hence for any $\eta > 0$ there exist $r \in \mathbf{N}$, $\delta > 0$ such that

$$V(\phi) < \eta, \qquad \forall \phi \in \Phi \text{ such that } \|\phi\|_r < \delta.$$

Then for $\|\phi\|_r < \delta$ we have

$$\sup_{n\geq 1} \int_{\mathcal{X}} \sup_t \left| 1 - e^{if(t)[\phi]} \right| \mu_n(df) \quad \leq \quad \sup_{n\geq 1} \sqrt{\eta} \mu_n \left\{ f : \sup_t |f(t)[\phi]| \leq \sqrt{\eta} \right\}$$

$$
\begin{aligned}
&+\sup_{n\geq 1} 2\mu_n\left\{f : \sup_t |f(t)[\phi]| > \sqrt{\eta}\right\}\\
\leq\ & \sqrt{\eta} + \frac{1+\sqrt{\eta}}{\sqrt{\eta}}V(\phi) \leq 3\sqrt{\eta}.
\end{aligned}
$$

Therefore for any $\phi \in \Phi$, we have

$$\sup_{n\geq 1}\int_{\mathcal{X}} \sup_t \left|1 - e^{if(t)[\phi]}\right| \mu_n(df) \leq 3\sqrt{\eta} + \frac{2\|\phi\|_r^2}{\delta^2}.$$

Let $p > r$ be such that the canonical injection from Φ_p to Φ_r is Hilbert-Schmidt. It follows from similar arguments as in the proof of Lemma 2.3.3 that $\forall n \geq 1$

$$
\begin{aligned}
& \frac{\sqrt{e}}{\sqrt{e}-1}\left(\frac{2}{M^2\delta^2}\sum_{j=1}^{\infty}\|\phi_j^p\|_r^2 + 3\sqrt{\eta}\right)\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\lim_{d\to\infty}\left(\frac{2}{M^2\delta^2}\sum_{j=1}^{d}\|\phi_j^p\|_r^2 + 3\sqrt{\eta}\right)\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\lim_{d\to\infty}\int_{\mathbf{R}^d}\left(2\left\langle\sum_{j=1}^{d}u_j\phi_j^p/M\delta, \sum_{j=1}^{d}u_j\phi_j^p/M\delta\right\rangle_r + 3\sqrt{\eta}\right)\\
& \qquad (2\pi)^{-d/2}e^{-|u|^2/2}du\\
\geq\ & \frac{\sqrt{e}}{\sqrt{e}-1}\lim_{d\to\infty}\int_{\mathbf{R}^d}\int_{\mathcal{X}}\sup_t\left|1-\exp\left(if(t)\left[\sum_{j=1}^{d}u_j\phi_j^p/M\right]\right)\right|\mu_n(df)\\
& \qquad (2\pi)^{-d/2}e^{-|u|^2/2}du\\
\geq\ & \frac{\sqrt{e}}{\sqrt{e}-1}\lim_{d\to\infty}\int_{\mathcal{X}}\sup_t\left|\int_{\mathbf{R}^d}\left(1-\exp\left(if(t)\left[\sum_{j=1}^{d}u_j\phi_j^p/M\right]\right)\right)\right.\\
& \qquad \left.(2\pi)^{-d/2}e^{-|u|^2/2}du\right|\mu_n(df)\\
=\ & \frac{\sqrt{e}}{\sqrt{e}-1}\lim_{d\to\infty}\int_{\mathcal{X}}\sup_t\left|1-\exp\left(-\sum_{j=1}^{d}f(t)[\phi_j^p]^2/2M^2\right)\right|\mu_n(df)\\
\geq\ & \frac{\sqrt{e}}{\sqrt{e}-1}\int_{\mathcal{X}}\left(1-\exp\left(-\sup_t\|f(t)\|_{-p}^2/2M^2\right)\right)\mu_n(df)\\
\geq\ & \mu_n\left\{f\in D([0,T],\Phi') : \sup_{t\in[0,T]}\|f(t)\|_{-p} > M\right\}.
\end{aligned}
$$

For any $\epsilon > 0$, taking M and η such that

$$\epsilon = \frac{\sqrt{e}}{\sqrt{e}-1}\left(\frac{2}{M^2\delta^2}\sum_{j=1}^{\infty}\|\phi_j^p\|_r^2 + 3\sqrt{\eta}\right)$$

we see that (2.5.1) holds. ∎

Theorem 2.5.1 (Mitoma) *Let* $\{\mu_n\}$ *be a sequence in* $\mathcal{P}(D([0,T],\Phi'))$ *(resp.* $\mathcal{P}(C([0,T],\Phi'))$*) such that, for any* $\phi \in \Phi$, $\{\mu_n \pi_\phi^{-1}\}$, *as a sequence of Borel probability measures on* $D([0,T],\mathbf{R})$ *(resp.* $C([0,T],\mathbf{R})$*), is tight. Then* $\{\mu_n\}$ *is tight in* $D([0,T],\Phi')$ *(resp.* $C([0,T],\Phi')$*).*

Proof: Let ϵ, p and M be the same as those in Lemma 2.5.1. Let $q > p$ be such that the canonical injection from Φ_q to Φ_p is Hilbert-Schmidt. Let $\{\phi_j^q\}$ be a CONS of Φ_q. Define

$$C^\epsilon = \left\{ f \in D([0,T],\Phi') : \sup_{0\le t\le T} \|f(t)\|_{-p} \le M \right\}$$

and

$$B^\epsilon = \{x \in \Phi_{-q} : \|x\|_{-p} \le M\}.$$

It is easy to see that $C^\epsilon \in \mathcal{B}(D([0,T],\Phi'))$ and B^ϵ is a compact subset of Φ_{-q}. Further, by the same arguments as in the proof of Theorem 2.4.4 we have $C^\epsilon \subset D([0,T],\Phi_{-q})$.

For $j \ge 1$, it follows from the tightness of $\{\mu_n \pi_{\phi_j^q}^{-1}\}$ that for any $\epsilon > 0$ there exists a compact subset K_j^ϵ of $D([0,T],\mathbf{R})$ such that

$$\left(\mu_n \pi_{\phi_j^q}^{-1}\right)(K_j^\epsilon) > 1 - \frac{\epsilon}{2^j}.$$

Letting

$$K^\epsilon = C^\epsilon \cap \left(\cap_{j=1}^\infty \pi_{\phi_j^q}^{-1} K_j^\epsilon\right) \subset D([0,T],\Phi_{-q})$$

we have

$$\lim_{\delta\to 0+} \sup_{f\in K^\epsilon} w_f'(\delta,\phi_j^q) \le \lim_{\delta\to 0+} \sup_{x\in K_j^\epsilon} w_x'(\delta) = 0$$

and

$$\{f(t) : f \in K^\epsilon, t \in [0,T]\} \subset B^\epsilon.$$

It follows from similar argument as in (2.4.16) that

$$\lim_{\delta\to 0} \sup_{f\in K^\epsilon} w_f'(\delta,q) = 0$$

and hence K^ϵ is relatively compact in $D([0,T],\Phi_{-q})$. Further,

$$\begin{aligned}\mu_n(K^\epsilon) &= 1 - \mu_n\left(\cup_{j=1}^\infty (\pi_{\phi_j^q}^{-1} K_j^\epsilon)^c \cup (C^\epsilon)^c\right) \\ &\ge 1 - \sum_{j=1}^\infty \frac{\epsilon}{2^j} - \epsilon = 1 - 2\epsilon.\end{aligned}$$

As the canonical map from $D([0,T],\Phi_{-q})$ to $D([0,T],\Phi')$ is continuous, we see that K^ϵ is relatively compact in $D([0,T],\Phi')$ and hence $\{\mu_n\}$ is tight. ∎

Theorem 2.5.2 *Let $\{\mu_n\}$ be a sequence of Borel probability measures on $D([0,T],\Phi')$ (resp. $C([0,T],\Phi')$). Let $q \geq p$ be such that the canonical injection from Φ_{-p} to Φ_{-q} is Hilbert-Schmidt. Suppose that*
a) $\forall \phi \in \Phi$, $\{\mu_n \pi_\phi^{-1}\}$, as a sequence of Borel probability measures on $D([0,T],\mathbf{R})$ (resp. on $D([0,T],\mathbf{R})$), is tight.
b) For any $\epsilon > 0$ there exists a constant M such that $\forall n \geq 1$

$$\mu_n \left\{ f \in D([0,T],\Phi') : \sup_{0\leq t\leq T} \|f(t)\|_{-p} \leq M \right\} \geq 1-\epsilon.$$

Then $\{\mu_n\}$, regarded as a sequence of Borel probability measures on $D([0,T],\Phi_{-q})$ (resp. $C([0,T],\Phi_{-q})$), is tight.

Proof: Let K^ϵ be given as in the proof of the last theorem. Then K^ϵ is relatively compact in $D([0,T],\Phi_{-q})$ while q does not depend on ϵ under our present assumption b). We only need to show that $\{\mu_n\}$ can be regarded as a sequence of Borel probability measures on $D([0,T],\Phi_{-q})$. It follows from the same argument as in the proof of (2.4.16) that the identity map from K^ϵ (with the restricted topology of $D([0,T],\Phi')$) to $D([0,T],\Phi_{-q})$ is continuous. For each $B \in \mathcal{B}(D([0,T],\Phi_{-q}))$ we have $B \cap K^\epsilon \in \mathcal{B}(D([0,T],\Phi'))$. Define $\tilde{\mu}_n(B)$ as the limit of $\mu_n(B \cap K^\epsilon)$ as $\epsilon \to 0$. Then $\tilde{\mu}_n \in \mathcal{P}(D([0,T],\Phi_{-q}))$ and $\tilde{\mu}_n(B) = \mu_n(B)$ for any $B \in \mathcal{B}(D([0,T],\Phi')) \cap \mathcal{B}(B([0,T],\Phi_{-q}))$. Therefore $\{\mu_n\}$ can be regarded as a sequence of Borel probability measures on $D([0,T],\Phi_{-q})$. ∎