Open mappings of probability measures and the Skorohod representation theorem

Vladimir I. Bogachev[1], **Alexander V. Kolesnikov**[2]

[1], [2]: Department of Mechanics and Mathematics,

Moscow State University, 119899 Moscow, Russia

Abstract

We prove that for a broad class of spaces X and Y (including all Souslin spaces), every open surjective mapping $f\colon X \to Y$ induces the open mapping $\mu \mapsto \mu \circ f^{-1}$ between the spaces of probability measures $\mathcal{P}(X)$ and $\mathcal{P}(Y)$. Connections with the Skorohod representation theorem and its generalizations are discussed.

AMS 1991 Subject classification:
Primary: 60B10, 54C10
Secondary: 28C15, 54C65
Keywords and phrases: weak convergence of probability measures, Skorohod representation, open mapping, continuous selection

Introduction

It was proved in [11] that every open surjection $f\colon X \to Y$ of compact spaces induces the open mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$, $\mu \mapsto \mu \circ f^{-1}$, between the spaces of probability measures equipped with the weak topology. According to [4], the same is true for Polish spaces. We extend the above result to a broad class of spaces that includes Souslin spaces as well as compact spaces. The second topic in this work is a discussion of connections between the aforementioned result and the Skorohod representation theorem for the weak convergence of measures together with its generalizations. Let us recall that A. V. Skorohod [32] proved that, for any sequence of probability measures μ_n on a complete separable metric space X weakly convergent to a measure μ_0, one can find Borel mappings ξ_n, $n = 0, 1, \dots$, from $[0, 1]$ to X such that $\xi_n \to \xi_0$ almost everywhere and the image of Lebesgue measure under the mapping ξ_n coincides with μ_n for all $n = 0, 1 \ldots$. D. Blackwell and L. Dubins [7] proved that the whole space $\mathcal{P}(X)$ of probability measures on X can be parameterized by mappings from $[0,1]$ to X with the preservation of the aforementioned correspondence, i.e., with every Borel probability measure μ on X, one can associate a Borel mapping $\xi_\mu\colon\ [0,1] \to X$ such that μ is the image of Lebesgue measure under ξ, and $\xi_{\mu_n} \to \xi_\mu$ almost everywhere if $\mu_n \to \mu$ weakly. Later the same was proved by X. Fernique [18]. In the case where $X = \mathbb{R}^n$, N. V. Krylov [22] has recently shown that the mappings ξ_μ can be taken almost everywhere differentiable. We give below a short derivation of the Blackwell–Dubins extension of the Skorohod theorem from Michael's selection theorem and the open mapping theorem for probability measures. On the other hand, we show how in the metric case, the Skorohod representation theorem can be used for proving that open mappings induce open mappings of probability measures. We also discuss the existence of continuous right inverse for the induced mappings between the spaces of probability measures.

1 Notation and terminology

All topological spaces considered below are assumed to be Hausdorff. The spaces that are homeomorphic to complete separable metric spaces are called Polish. Continuous images of Polish spaces are called Souslin spaces (or Souslin sets).

We denote by $\mathcal{B}(X)$ the Borel σ-field of a topological space X, i.e., the σ-field generated by all open sets. The Baire σ-field $\mathcal{B}_0(X)$ is generated by all sets of the form $\{f > 0\}$, where f is a continuous function on X. The set of all Borel (possibly, signed) measures on a topological space X is denoted by $\mathcal{M}(X)$; the symbol $\mathcal{M}_0(X)$ stands for the space of all Baire measures on X. We denote by $\mathcal{P}(X)$ the set of all probability Borel measures on X; $\mathcal{P}_0(X)$ is the set of all Baire probability measures on X. The subset of $\mathcal{P}(X)$ formed by all Radon (i.e., inner compactly regular) measures is denoted by $\mathcal{P}_r(X)$. The weak topology on $\mathcal{M}(X)$ and $\mathcal{M}_0(X)$ is the topology generated by the family of seminorms of the form $\mu \mapsto \left|\int \varphi\, d\mu\right|$, where φ is a bounded continuous function on X. Thus, a sequence of measures μ_n on X converges weakly to a measure μ if $\int \varphi\, d\mu_n \to \int \varphi\, d\mu$ for each bounded continuous function φ on X (about the weak topology, see [6] or [34]). The restrictions of the weak topology to $\mathcal{P}(X)$, $\mathcal{P}_0(X)$, and $\mathcal{P}_r(X)$ are called the weak topologies on the corresponding classes. The weak topology is Hausdorff on $\mathcal{M}_0(X)$. In the case where $\mathcal{B}(X) = \mathcal{B}\dashv(X)$ (which is always the case for completely regular Souslin spaces and for metric spaces), the weak topology on $\mathcal{P}(X)$ can be defined by means of the base generated by all sets of the form

$$U(\mu, V, \varepsilon) = \Big\{\nu \in \mathcal{P}(X)\colon\ \nu(V) > \mu(V) - \varepsilon\Big\}, \tag{1.1}$$

where $V \subset X$ is an open set and $\varepsilon > 0$. The base of the weak topology on $\mathcal{P}_0(X)$ is generated by the sets of the form (1.1) with co-zero sets V, i.e., $V = \{f > 0\}$, where f is a continuous function on X (see [8], [33]). For every completely regular space X, the weak topology on $\mathcal{P}_r(X)$ is Hausdorff, and the sets of the form (1.1) generate its base.

For any probability measure μ on X, the symbol $\mathcal{B}(X)_\mu$ stands for the Lebesgue completion of $\mathcal{B}(X)$ with respect to μ. The sets belonging to $\mathcal{B}(X)_\mu$ are called μ-measurable. A set A in a topological space X is called universally measurable if it is measurable with respect to every Borel measure μ on X, i.e., $A \in \mathcal{B}(X)_\mu$. It is known that every Souslin set is universally measurable (see [29]). If X and Y are topological spaces, μ is a Borel measure on X, and $f\colon X \to Y$ is a μ-measurable mapping (i.e., $f^{-1}(A) \subset \mathcal{B}(X)_\mu$ for all $A \in \mathcal{B}(Y)$), then the measure $\mu \circ f^{-1}$, defined on $\mathcal{B}(Y)$ by the formula

$$\mu \circ f^{-1}(A) = \mu\bigl(f^{-1}(A)\bigr),$$

is called the image of μ under the mapping f. For every continuous mapping $f\colon X \to Y$, the formula

$$\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y), \qquad \mu \mapsto \mu \circ f^{-1},$$

defines a mapping between the spaces of measures that is continuous in the weak topology.

We recall that a mapping f from a topological space X to a topological space Y is called open if it is continuous and takes all open sets to open ones. We say that a mapping $f\colon X \to Y$ has a continuous right inverse if there exists a continuous mapping $g\colon Y \to X$ such that $f\bigl(g(y)\bigr) = y$ for all $y \in Y$ (in other words, the multi-valued mapping f^{-1} has a continuous selection).

Lebesgue measure is denoted by λ. We denote by $\widehat{A}$ the closure of a set A and by ∂A the boundary of A.

We shall make use of the following classical result due to E. Michael [24].

Theorem 1.1. *Let X be a paracompact space and let M be a metrizable subset in a complete locally convex space E. Suppose that $\Phi\colon X \to 2^M$ is lower semi-continuous mapping with values in the set 2^M of all non-empty closed subsets of M such that $\Phi(x)$ is complete for each x. Then there exists a continuous mapping $f\colon X \to E$ such that $f(x)$ belongs to the closed convex hull of $\Phi(x)$ for all x. If M is complete, then the completeness of E may be omitted.*

We note that the last assertion follows from the first one, since E has a completion $\widetilde{E}$, and M is a complete metrizable subset of $\widetilde{E}$. A typical application of the above theorem is this: let $T\colon M \to X$ be a continuous affine mapping of a complete metrizable convex closed set M in a locally space E onto a metrizable set X in a locally convex space Y. Then $\Phi(x) = T^{-1}(x)$ satisfies the hypotheses of Theorem 1.1, hence T has a continuous right inverse.

2 Open mappings of probability measures

Lemma 2.1. *Let μ be a Radon probability measure on a completely regular space X and let $U \subset X$ be an open set. Then, for every $\varepsilon > 0$, there exists an open set V such that $\overline{V} \subset U$, $\mu(\partial V) = 0$, and $\mu(U \setminus V) < \varepsilon$. The same is true for every Borel probability measure if X is a perfectly normal space. Finally, the assertion is true for every Baire measure on an arbitrary space if U is a co-zero set.*

Proof. The exists a compact set $K \subset U$ with $\mu(U \backslash K) < \varepsilon$. By using that X is completely regular, one can find a continuous function on X such that $f|_K = 1$ and $f = 0$ outside U. Then there exists $c \in (0, 1)$ such that the set $\{f = c\}$ has μ-measure zero. Now we can take $V = \{f > c\}$, since $\partial V \subset \{f = c\}$. In the case where the set U has the form $\{f > 0\}$ for some continuous function f, a similar reasoning applies with c sufficiently small. $\square$

In particular, the assertion of Lemma 2.1 is true for every Borel measure on a universally measurable subspace of a completely regular Souslin space.

Lemma 2.2. *Let X and μ be the same as in Lemma 2.1 and let*

$$G_0 = \big\{\nu \in \mathcal{P}(X)\colon\ \nu(V_i) - \mu(V_i) > -\varepsilon_0,\ i = 1, \dots, n\big\},$$

where V_i are open sets in X and $\varepsilon_0 > 0$. Then, for every $\varepsilon > 0$, there exist pairwise disjoint open sets $W_1, W_2, \dots, W_m$ and a number $\delta > 0$ such that

(i) $\mu(V_i) - \sum_{\{j\colon\ W_j \subset V_i\}} \mu(W_j) < \varepsilon,\ i = 1, \dots, n,$

(ii) $G = \big\{\nu \in \mathcal{P}(X)\colon\ \nu(W_i) - \mu(W_i) > -\delta,\ i = 1, \dots, m\big\} \subset G_0.$

Proof. Let us prove the claim by induction in n. If $n = 1$, then the claim is trivial. Suppose now that for the sets $V_1, \cdots, V_{n-1}$ and the number $\varepsilon/2$, we have constructed pairwise disjoint open sets $\widetilde{W}_1, \dots, \widetilde{W}_{m_{n-1}}$ and a number δ_{n-1} such that (i) and (ii) are satisfied with δ_{n-1} in place of δ. Let $\varepsilon_1 < \dfrac{\min(\varepsilon, \varepsilon_0, \sigma_{n-1})}{2m_{n-1}}$ be positive. By Lemma 2.1, we can construct open sets W_j^1, $j = 1, \dots, m_{n-1}$, such that

$$\overline{W_j^1} \subset \widetilde{W}_j, \quad \mu(\widetilde{W}_j \setminus W_j^1) < \varepsilon_1, \quad \mu(\partial W_j^1) = 0.$$

Let us set $W = V_n \setminus \bigcup_{j=1}^{m_{n-1}} \overline{W_j^1}$. Now, for every j, we construct an open set W_j^2 such that

$$\overline{W_j^2} \subset W_j^1 \cap V_n, \quad \mu\big((W_j^1 \cap V_n) \setminus W_j^2\big) < \varepsilon_1, \quad \mu(\partial W_j^2) = 0.$$

Let us set $W_j^3 = W_j^1 \setminus \overline{W_j^2}$. Then the required sets are W, W_j^2, W_j^3, $j = 1, \dots, m_{n-1}$.

Clearly, they are pairwise disjoint. For $i = 1, 2, \dots, n-1$, we have

$$\mu(V_i) - \sum_{j\colon \widetilde{W}_j \subset V_i} \bigl(\mu(W_j^2) + \mu(W_j^3)\bigr) = \mu(V_i) - \sum_{j\colon \widetilde{W}_j \subset V_i} \mu(W_j^1)$$
$$< \mu(V_i) - \sum_{j\colon \widetilde{W}_j \subset V_i} \mu(\widetilde{W}_j) + m_{n-1}\varepsilon_1 < \frac{\varepsilon}{2} + m_{n-1}\varepsilon_1 < \varepsilon.$$

We also have

$$\mu(V_n) - \mu(W) - \sum_j \mu(W_j^2) < m_{n-1}\varepsilon_1 < \varepsilon.$$

Therefore, condition (i) is fulfilled. Let $\delta < \dfrac{\min(\varepsilon_0, \delta_{n-1})}{4(m_{n-1}+1)}$ be positive. In order to show (ii) it remains to prove that

$$\nu(\widetilde{W}_j) - \mu(\widetilde{W}_j) > -\delta_{n-1}, \quad \forall\, j = 1, \dots, m_{n-1},$$

and that $\nu(V_n) - \mu(V_n) > -\varepsilon_0$ if $\nu \in G$. Indeed, one has

$$\nu(\widetilde{W}_j) - \mu(\widetilde{W}_j) > \nu(W_j^1) - \mu(W_j^1) - \varepsilon_1$$
$$\geq \nu(W_j^2) + \nu(W_j^3) - \mu(W_j^2) - \mu(W_j^3) - \varepsilon_1 > -2\delta - \varepsilon_1 > -\delta_{n-1}$$

and

$$\nu(V_n) - \mu(V_n) \geq \nu(W) + \sum_j \nu(W_j^1 \cap V_n) - \mu(W) - \sum_j \mu(W_j^1 \cap V_n)$$
$$> \nu(W) + \sum_j \nu(W_j^2) - \mu(W) - \sum_j \mu(W_j^2) - m_{n-1}\varepsilon_1$$
$$> -(m_{n-1}+1)\delta - m_{n-1}\varepsilon_1 > -\varepsilon_0.$$

The lemma is proven. $\square$

Theorem 2.3. *Let X and Y be completely regular spaces and let $f\colon X \to Y$ be a surjective open mapping that satisfies the following local conservativity condition: for every open set $V \subset X$ and every Radon probability measure ν on $f(V)$, there exists a Radon probability measure μ on V with $\mu \circ f^{-1} = \nu$. Then the mapping $\widehat{f}\colon \mathcal{P}_r(X) \to \mathcal{P}_r(Y)$ is an open surjection. The same is true for the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ if X and Y are perfectly normal and f satisfies the above condition with "Radon measure" replaced by "Borel measure". Finally, the same is true for arbitrary spaces and Baire measures if f is a continuous surjection taking co-zero sets to co-zero sets and the above local conservativity condition is fulfilled for Baire measures.*

Proof. The surjectivity of $\widehat{f}$ follows by definition. Let $\mu_0 \in \mathcal{P}_r(X)$ and let

$$U = \bigl\{\mu : \mu(V_i) - \mu_0(V_i) > -\varepsilon,\ i = 1, \dots, n\bigr\}, \quad \varepsilon > 0,$$

be its weak neighborhood, where the sets $V_i \subset X$ are open and $\varepsilon > 0$. According to Lemma 2.2, we may assume that the sets V_i are pairwise disjoint. Let $\eta_0 = \mu_0 \circ f^{-1}$. The sets $f(V_i)$ are open. By Lemma 2.2, there exist pairwise disjoint open sets $W_1, \dots, W_m \subset Y$ such that

$$\eta_0\bigl(f(V_i)\bigr) - \sum_{W_j \subset f(V_i)} \eta_0(W_j) < \frac{\varepsilon}{2}.$$

Let $\varepsilon_1 < \dfrac{\varepsilon}{2m}$ be positive and let us fix a neighborhood of η_0 of the form

$$O = \bigl\{\eta \in \mathcal{P}_r(Y)\colon\ \eta(W_j) - \eta_0(W_j) > -\varepsilon_1,\ j = 1, \dots, m\bigr\}.$$

We shall prove that every measure $\eta \in O$ has a preimage in U. We set $I_j = \{i : W_j \subset f(V_i)\}$. For every $i \in I_j$, we find nonnegative numbers α_{ij} such that $\sum_{i \in I_j} \alpha_{ij} = 1$ and $\alpha_{ij}\eta_0(W_j) \geq \mu_0\bigl(f^{-1}(W_j) \cap V_i\bigr)$. By our hypothesis, for every $i \in I_j$, there exist preimages μ_{ij} of the measure $\eta|_{W_j}$ with respect to the mapping $f|_{V_i}$, i.e., nonnegative Radon measures μ_{ij} on V_i such that $\mu_{ij} \circ f^{-1} = \eta|_{W_j}$. Let $\widetilde{\mu}$ be a preimage of $\eta|_{Y \setminus \cup_j W_j}$ with respect to f. Let

$$\mu = \widetilde{\mu} + \sum_j \sum_{i \in I_j} \alpha_{ij}\mu_{ij}.$$

Clearly, $\nu = \mu \circ f^{-1}$. Let us show that $\mu \in U$. Indeed, one has

$$\begin{aligned}
\mu(V_i) &\geq \sum_{j\colon\, i \in I_j} \alpha_{ij}\mu_{ij}(V_i) = \sum_{j\colon\, i \in I_j} \alpha_{ij}\eta(W_j) \\
&> \sum_{j\colon\, i \in I_j} \alpha_{ij}\Bigl(\eta_0(W_j) - \varepsilon_1\Bigr) \geq \sum_{j\colon\, i \in I_j} \alpha_{ij}\eta_0(W_j) - m\varepsilon_1 \\
&\geq \sum_{j\colon\, i \in I_j} \mu_0\Bigl(f^{-1}(W_j) \cap V_i\Bigr) - m\varepsilon_1 \\
&= \mu_0(V_i) - \mu_0\Bigl(V_i \setminus \bigcup_{j\colon\, i \in I_j} f^{-1}(W_j)\Bigr) - m\varepsilon_1 \\
&\geq \mu_0(V_i) - \eta_0\Bigl(f(V_i) \setminus \bigcup_{j\colon\, W_j \in f(V_i)} W_j\Bigr) - m\varepsilon_1 \\
&\geq \mu_0(V_i) - \frac{\varepsilon}{2} - m\varepsilon_1 > \mu_0(V_i) - \varepsilon.
\end{aligned}$$

The first assertion is proven. Two other statements are proved by a similar reasoning. $\square$

The local conservativity condition in Theorem 2.3 is satisfied in many cases, for example, for all continuous mappings between Souslin spaces, see, e.g., [8, Corollary 6.1.9], [9, Ch. IX, §2, Proposition 9]. Moreover, the same is true for Borel mappings, which follows by the measurable selection theorem (see, e.g., [12, Appendix 3] or [29, Theorem 13, Ch. II, §3, p. 127]). More generally, it suffices that X and Y be Borel isomorphic to Polish spaces or that the closed sets in X be $\mathcal{K}$-analytic (see [15], [23], [30]).

In the case where X is a Polish space, the assertion of Theorem 2.3 was proved in [5]. A result of this type was obtained in [4, Proposition 2.10] for certain special mappings f between metric spaces; some ideas of that proof have been used above. In [4], the question about the validity of this assertion in the case of Borel subsets of separable metric spaces was raised.

Corollary 2.4. *Let X and Y be completely regular Souslin spaces and let $f\colon X \to Y$ be an open surjective mapping. Then the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ is an open surjection for the weak topology.*

Proof. It suffices to recall that $\mathcal{P}(X) = \mathcal{P}_r(X)$ and $\mathcal{P}(Y) = \mathcal{P}_r(Y)$ for completely regular Souslin spaces and that every continuous mapping between Souslin spaces is locally conservative in the sense of Theorem 2.3. □

Corollary 2.5. *Let $f\colon X \to Y$ be an open surjective mapping of complete metric spaces. Then the mapping $\mathcal{P}_r(X) \to \mathcal{P}_r(Y)$ is an open surjection.*

Proof. Let $V \subset X$ be open, and let ν be a Radon probability measure on $f(V)$. We have to show that there exists a Radon probability measure μ on V such that $\mu\circ f^{-1} = \nu$. Let $\varepsilon > 0$ and let $E \subset f(V)$ be compact with $\nu(E) > 1-\varepsilon$. Recall that V is homeomorhic to a complete metric space (see [14, Theorem 4.3.23]). According to [14, 5.5.11], there exits a compact set $K \subset V$ such that $f(K) = E$. Now the assertion follows by [8, Theorem 6.1.8]. □

Remark 2.6. Theorem 2.3 implies the result of [11] where the spaces X and Y were compact. Indeed, let $f\colon X \to Y$ be a an open mapping, let $U \subset X$ be open, and let ν be a Radon probability measure on $V = f(U)$. In order to show that there is a Radon probability measure μ on U such that $\mu \circ f^{-1} = \nu$, it suffices to show that for every $\varepsilon > 0$, there is a compact set $S \subset U$ with $\nu\bigl(f(S)\bigr) > 1-\varepsilon$ (see, e.g., [8, Theorem 6.1.8]). Let us take a compact set $K \subset V$ with $\nu(K) > 1-\varepsilon$. For every point $k \in K$, we pick some $u_k \in U \cap f^{-1}(k)$ and find a neighborhood U_k of u_k such that $\overline{U_k} \subset U$. The open sets $f(U_k)$ cover K, hence one can take a finite subcover $f(U_{k_1}), \dots, f(U_{k_n})$. Then the set $S = \cup_{i=1}^n \overline{U_{k_i}}$ is compact and $K \subset f(S)$, whence $\nu\bigl(f(S)\bigr) > 1-\varepsilon$.

In the next section, we shall reinforce Theorem 2.3 in the metric case.

Lemma 2.7. *Let X and Y be Polish spaces and let $f\colon X \to Y$ be an open surjective mapping. Then the corresponding mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ has a continuous right inverse in the weak topology, i.e., there exists a continuous in the weak topology mapping $\Psi\colon \mathcal{P}(Y) \to \mathcal{P}(X)$ such that $\widehat{f}\big(\Psi(\nu)\big) = \nu$ for all $\nu \in \mathcal{P}(Y)$. The same is true for arbitrary complete metric spaces X and Y if we consider $\mathcal{P}_r(X)$ and $\mathcal{P}_r(Y)$ in place of $\mathcal{P}(X)$ and $\mathcal{P}(Y)$, respectively.*

Proof. We know that the mapping $\widehat{f}$ is surjective and open. Since this mapping is linear and the space $\mathcal{P}_r(X)$ is convex, metrizable, and complete in the locally convex space $\mathcal{M}_r(X)$ with the weak topology, then the existence of a continuous right inverse follows by Michael's Theorem 1.1. In our situation, we apply a special case of this theorem that yields the existence of a continuous single-valued selection of a lower semicontinuous multi-valued mapping from a complete metric space S_1 to closed convex subsets of a metrizable closed convex set S_2 in a locally convex space. Then it follows that, for any open surjective linear mapping $\varphi\colon A \to B$, where A and B are convex, metrizable, complete sets in locally convex spaces, there exists a continuous mapping $\psi\colon B \to A$ such that $\varphi \circ \psi$ is the identity mapping on B. $\square$

Corollary 2.8. *For every universally measurable set Y in a Polish space Z, there exist a subset X of the space $\mathcal{R}$ of the irrational numbers in $[0,1]$ and a continuous surjective mapping $f\colon X \mapsto Y$ such that the mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ has a continuous in the weak topology right inverse $g\colon \mathcal{P}(Y) \to \mathcal{P}(X)$.*

Proof. There exists an open surjection $F\colon \mathcal{R} \to Z$ (see [14, Ch. 4, 4.2.D]; recall that $\mathcal{R}$ is homeomorphic to $\mathbb{N}^{\mathbb{N}}$). Every measure μ on Y extends to Z by setting $\mu(Z\backslash Y) = 0$, which is possible by the universal measurability of Y. Let $\Psi\colon \mathcal{P}(Z) \to \mathcal{P}(\mathcal{R})$ be the mapping from Lemma 2.7. Let us put $X = F^{-1}(Y)$. If measures μ_n on Y converge weakly to μ, then the restrictions of the measures $\Psi(\mu_n)$ to X converge in the weak topology of the space $\mathcal{P}(X)$, since $\Psi(\mu)(X) = 1$. Hence one can take for f the restriction of F to X and for $g(\mu)$ the restriction of $\Psi(\mu)$ to X. $\square$

Remark 2.9. Many problems in probability theory and the theory of random processes (e.g., the problem of solving stochastic differential equations) fit the following general framework (see, e.g., [15], [16]). Given a measurable mapping $f\colon X \to Y$ between measurable spaces X and Y and a probability measure ν on Y, find a measure μ on X such that $\nu = \mu \circ f^{-1}$. In the case where X and Y are topological spaces and the spaces of probability measures $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ are equipped with the weak topologies, the existence of a continuous right inverse for the induced mapping $\widehat{f}\colon X \to Y$ ensures a continuous selection of a solution μ^ν of the above problem even if a solution is not unique. It would be interesting to extend Lemma 2.7 to larger classes of spaces.

The following example shows that for the existence of a continuous right inverse of $\widehat{f}$, it is not necessary that f be open. We denote by C the standard Cantor set, which consists of all numbers in the interval $[0,1]$, whose triadic decompositions do not contain 1 (see [14, Example 3.1.28]).

Example 2.10. *There exists a continuous mapping $F\colon C \to T$ of the Cantor set C onto the infinite-dimensional torus T (or onto the Hilbert cube) such that the corresponding mapping $\widehat{F}\colon \mathcal{P}(C) \to \mathcal{P}(T)$ has a continuous right inverse in the weak topology.*

Proof. Let us define a mapping $\tau\colon C \to T$ by using the partition of the space T into "cylinders" $T^n_{i_1,i_2,\dots,i_n}$ constructed as follows. Let $T = \prod\limits_{k=1}^{\infty} I_k$, where I_k is the k-th copy of $I = \{t\colon t = e^{2\pi i\psi},\ \psi \in [0,1]\}$. For every n and every $k \le n$, we shall consider the partition of I_k into 2^{n-k+1} equal parts $I^1_k = \{e^{2\pi i\psi},\ \psi \in [\frac{i-1}{2^{n-k+1}}, \frac{i}{2^{n-k+1}})\}$, $i = 1,\dots,2^{n-k+1}$. The space T is partitioned into all possible subsets of the form $T_n^{i_1,i_2,\dots,i_n} = \prod\limits_{k=1}^{n} I_k^{i_k} \prod\limits_{k=n+1}^{\infty} I_k$. One gets a sequence of decreasing partitions Γ_n of the space T into disjoint sets, the maximal of the diameters of which tends to zero as n is growing.

The mapping τ is constructed as the limit of the mappings τ_n. For constructing τ_n, we divide consecutively the Cantor set into disjoint subsets as indicated below. The mapping τ_0 takes the whole C to some fixed point of the space T. Suppose that the construction is carried out till the n-th step inclusive. At the n-th step T is partitioned into disjoint subsets of the n-th rank $T^n_{i_1,i_2,\dots,i_n}$, $1 \le i_k \le 2^{n-k+1}$, the Cantor set is partitioned into $2^{\sum_{1\le k\le n} n-k+1}$ equal disjoint parts $C^n_{i_1,i_2,\dots,i_n}$, $1 \le i_k \le 2^{n-k+1}$. The mapping τ_n takes all of $C^n_{i_1,i_2,\dots,i_n}$ to an arbitrary point $x^n_{i_1,i_2,\dots,i_n} \in T^n_{i_1,i_2,\dots,i_n}$. Then each of $T^n_{i_1,i_2,\dots,i_n}$ is subdivided into 2^{n+1} disjoint "cylinders" of the $(n+1)$-th rank. Let $a = \inf\{t \in C^n_{i_1,i_2,\dots,i_n}\}$, $b = \sup\{t \in C^n_{i_1,i_2,\dots,i_n}\}$. We partition $C^n_{i_1,i_2,\dots,i_n}$ into disjoint sets $C^n_{i_1,i_2,\dots,i_n} \cap [+(b-a)\frac{j-1}{2^{n+1}}, a+(b-a)\frac{j}{2^{n+1}}]$, $1 \le j \le 2^{n+1}$. The collections of 2^{n+1} sets obtained in this way are put in a one-to-one correspondence, the index of every "cylinder" of subdivision is given to the corresponding part of the Cantor set. This operation is carried out for all indices $\{n, i_1, i_2, \dots, i_n\}$. The mapping τ_{n+1} is determined by the conditions: τ_{n+1} is constant on $C^{n+1}_{i_1,i_2,\dots,i_{n+1}}$ and takes its value in $T^{n+1}_{i_1,i_2,\dots,i_{n+1}}$. All mappings τ_n are continuous, and the sequence τ_n is fundamental in the uniform metric. Thus, we get in the limit a continuous surjective mapping τ. It is readily seen that the limit does not depend on a choice of the points $x^n_{i_1,i_2,\dots,i_n}$.

Let us fix a probability measure μ on T. It has the natural preimage ν with respect to the mapping τ, which we call the preimage of μ with respect to τ and which is defined as follows. For an arbitrary n and an arbitrary set of indices

$i_1, i_2, \ldots, i_n$, let us put

$$\nu^{(n)} = \sum_{i_1, i_2, \ldots, i_n} \mu(T^n_{i_1,i_2,\ldots,i_n}) \delta_{t^n_{i_1,i_2,\ldots,i_n}}, \tag{2.2}$$

where δ_t is Dirac's measure at the point t and $t^n_{i_1,i_2,\ldots,i_n} = \inf\{t : t \in C^n_{i_1,i_2,\ldots,i_n}\}$. We observe that for the weak convergence of a sequence of probability measures on the Cantor set, it is necessary and sufficient that there exist the limits of the values of these measures at every set $C^n_{i_1,i_2,\ldots,i_n}$. The necessity follows from the fact that the boundaries of those sets are empty. The sufficiency follows from the easily verified fact that the sequence of the integrals of every continuous function against these measures is fundamental, which implies the existence of the weak limit. This observation yields the existence of the weak limit ν of the sequence $\nu^{(n)}$, since whenever $m \geq n$ we have $\nu^{(m)}(C^n_{i_1,i_2,\ldots,i_n}) = \nu^{(n)}(C^n_{i_1,i_2,\ldots,i_n})$. For establishing the equality $\mu = \nu \circ \tau^{-1}$, it suffices to prove that the sequence of measures $\nu_n \circ \tau^{-1}$ converges weakly to μ. This follows by the uniform continuity of every continuous function f on T and the equalities

$$\int_T f\, d\mu - \int_T f\, d(\nu_n \circ \tau^{-1}) = \sum_{i_1,i_2,\ldots,i_n} \int_{T^n_{i_1,i_2,\ldots,i_n}} \Big(f - f\big(\tau(t^n_{i_1,i_2,\ldots,i_n})\big)\Big)\, d\mu.$$

If measures μ_n on T converge weakly to μ and $\mu(\Sigma) = 0$, then for their preimages ν_n with respect to τ and any fixed m, we have

$$\lim_{n\to\infty} \nu_n(C^m_{i_1,i_2,\ldots,i_m}) = \lim_{n\to\infty} \mu_n(T^m_{i_1,i_2,\ldots,i_m}) = \mu(T^m_{i_1,i_2,\ldots,i_m}) = \nu(C^m_{i_1,i_2,\ldots,i_m}).$$

Thus, in this case the measures ν_n converge weakly to ν.

Let Φ^x be the homeomorphisms of T defined as follows. Let φ^x be the rotation of the circle by an angle of $2\pi x$. We define Φ^x on the space T by

$$\Phi^x(t_1, t_2, \ldots, t_n, \ldots) = \big(\varphi^x(t_1), \varphi^x(t_2), \ldots, \varphi^x(t_n), \ldots\big).$$

We assume that the domain of the parameter x is the set C. For a fixed measure μ, we denote by ν^x the natural preimage of the measure $\mu \circ (\Phi^x)^{-1}$ with respect to τ. Let us consider the mapping

$$\Psi(x,t) \colon, C^2 \mapsto T, \;\; \Psi(x,t) = (\Phi^x)^{-1} \circ \tau(t).$$

We shall show that the measure γ_μ on C^2 defined by the equality

$$\gamma_\mu(B) = \int_K \nu^x(B_x)\, \lambda_C(dx),$$

where λ_C is some probability Borel measure on C and B_x is the section of the set B at the point x, is well-defined. To this end, it suffices to show that the function $\eta_B\colon\ x \mapsto \nu^x(B_x)$ is Borel for every Borel set B. The class $\mathcal{E}$ of all Borel sets with such a property is σ-additive, i.e., it contains the empty set, is closed with respect to taking unions of increasing sequences, $A \cup B \in \mathcal{E}$ if $A,\, B \in \mathcal{E}$ and $A \cap B = \emptyset$, and $B \backslash A \in \mathcal{E}$ if $A \subset B$. Since the smallest σ-additive class containing all open sets coincides with the Borel σ-field (see [25, I.4.5]), it suffices to prove our claim for open B. In the case of the space C^2 one can restrict oneself by open-closed B. Let us consider the function $\eta_B^{(n)}\colon\ x \mapsto \nu^{x(n)}(B_x)$, where $\nu^{x(n)}$ is the n-th pre-limit measure of the measure ν^x constructed by formula (2.2). It is seen from this formula that the function $\eta_B^{(n)}$ is Borel for Borel B. Indeed, the function $\mu \circ \big(\Phi(x)\big)^{-1}(T^n_{i_1,i_2,\ldots,i_n})$ is continuous everywhere except for at most countably many points. For all x, the measures $\nu^{x(n)}$ converge weakly to ν^x. If B is open-closed, then its boundary is empty. This means that $\eta_B = \lim_{n\to\infty} \eta_B^{(n)}$. Therefore, the measure γ_μ is well-defined.

Now for a measure λ_C we take any probability measure without points of positive measure. The mapping $\Psi(x,t)$ transforms γ_μ to μ. This follows from the relations

$$\begin{aligned}
\gamma_\mu \circ \Psi^{-1}(A) &= \int\limits_C \nu^x\big(\Psi^{-1}(A)\big)_x\, \lambda_C(dx) = \int\limits_C \nu^x\Big(\tau^{-1}\big(\Phi^x(A)\big)\Big)\, \lambda_C(dx) \\
&= \int\limits_C (\nu^x \circ \tau^{-1})\big(\Phi^x(A)\big)\, \lambda_C(dx) = \int\limits_C \mu \circ (\Phi^x)^{-1}\big(\Phi^x(A)\big)\, \lambda_C(dx) \\
&= \int\limits_C \mu(A)\, \lambda_C(dx) = \mu(A).
\end{aligned}$$

Suppose now that the measures μ_n converge weakly to μ and that f is a continuous function on C^2. Then

$$\int\limits_C f(x,t)\, \gamma_{\mu_n}(dx dt) = \int\limits_C \int\limits_C f(x,t)\, \nu_n^x(dt)\, \lambda_C(dx).$$

Except for at most countable set of values x we have

$$\int\limits_C f(x,t)\, \nu_n^x(dt) \to \int\limits_C f(x,t)\, \nu^x(dt),$$

whence we obtain by Lebesgue's dominated convergence theorem that

$$\int\limits_C f(x,t)\, \gamma_{\mu_n}(dx dt) \to \int\limits_C f(x,t) \gamma_\mu(dx dt).$$

Thus, the measures γ_{μ_n} converge weakly to γ_μ. The mapping Ψ is continuous. It is known that the space C^2 is homeomorphic to C (see [14, 3.1.28]). Let $\varphi\colon\ C^2 \to C$ be the corresponding homeomorphism. Then we take for F the mapping $\Psi \circ \varphi^{-1}$, and for a desirable right inverse we take the mapping $G\colon\ \mu \to \gamma_\mu \circ \varphi^{-1}$. An analogous construction is possible for the Hilbert cube $[0,1]^\infty$ in place of T. $\square$

This example is not a corollary of Lemma 2.7, since T cannot be the image of C under an open mapping.

Corollary 2.11. *For every universally measurable set Y in a Polish space, there exist a subset X of the Cantor set C and a continuous surjective mapping $f\colon\ X \mapsto Y$ such that the mapping $\widehat{f}\colon\ \mathcal{P}(X) \to \mathcal{P}(Y)$ has a continuous in the weak topology right inverse $g\colon\ \mathcal{P}(Y) \to \mathcal{P}(X)$. In the case of compact Y the set X can be chosen compact.*

Proof. Let us embed Y homeomorphically into the infinite-dimensional torus T, which is possible by the Uryson theorem (see [14, Theorem 4.2.10]). Every measure μ on Y extends to T by setting $\mu(T\backslash Y) = 0$, which is possible by the universal measurability of Y. Let $F\colon\ C \to T$ and $G\colon\ \mathcal{P}(T) \to \mathcal{P}(C)$ be the mappings constructed in Example 2.10. Let us put $X = F^{-1}(Y)$. If measures μ_n on Y converge weakly to μ, then the restrictions of the measures $G(\mu_n)$ to X converge in the weak topology of the space $\mathcal{P}(X)$, since $G(\mu)(X) = 1$. Hence one can take for f the restriction of F to X and for $g(\mu)$ the restriction of $G(\mu)$ to X. Clearly, X is compact if Y is compact. $\square$

3 Skorohod's parameterization of the weak convergence

In this section, we prove the following theorem and consider its applications. Althought, this theorem can be obtained directly from the result in [7] (or in [18]), where it is formulated for complete separable metric spaces, we give a different and a shorter proof, which employs, however, the above results.

Theorem 3.1. *Let X be a universally measurable set in a Polish space. Then to every probability measure μ on X, one can associate a Borel mapping $\xi_\mu\colon\ [0,1] \to X$ such that μ is the image of Lebesgue measure under the mapping ξ_μ, and $\lim\limits_{n\to\infty} \xi_{\mu_n}(t) = \xi_\mu(t)$ for almost all t, provided the measures μ_n converge weakly to the measure μ.*

We shall say that a topological space X has the strong Skorohod representation property if the assertion of Theorem 3.1 is valid for this space.

Lemma 3.2. *Let X be a space with the strong Skorohod representation property. Then every universally measurable subset Y of X has this property as well.*

Proof. Making use of the universal measurability of Y, every measure μ on Y can be extended to X by setting $\mu(X\backslash Y) = 0$. Let $\xi_\mu\colon [0,1] \to X$ be a Borel mapping corresponding to μ in the definition of the strong Skorohod representation property. There exists a Borel set $B \subset Y$ of full μ-measure. Let us put $\eta_\mu(t) = \xi_\mu(t)$ if $t \in \xi_\mu^{-1}(B)$ and $\eta_\mu(t) = z$ if $t \notin \xi_\mu^{-1}(B)$, where z is an arbitrary point in Y. Then $\lambda\bigl(\xi_\mu^{-1}(B)\bigr) = 1$, hence $\eta_\mu(t) = \xi_\mu(t)$ for almost all t from $[0,1]$, whence $\lambda \circ \eta_\mu^{-1} = \lambda \circ \xi_\mu^{-1}$. If probability measures μ_n on Y converge weakly to a measure μ, then their extensions to X converge weakly to the extension of μ, whence $\lim\limits_{n\to\infty} \xi_{\mu_n}(t) = \xi_\mu(t)$ almost everywhere. Therefore, $\lim\limits_{n\to\infty} \eta_{\mu_n}(t) = \eta_\mu(t)$ almost everywhere. □

Lemma 3.3. *Let f be a continuous mapping from a topological space X with the strong Skorohod representation property onto a topological space Y such that the corresponding mapping $\widehat{f}\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ is surjective and has a continuous right inverse in the weak topology. Then Y has the strong Skorohod representation property.*

Proof. Let $\Psi\colon \mathcal{P}(Y) \to \mathcal{P}(X)$ be a mapping continuous in the weak topology such that $\widehat{f} \circ \Psi$ is the identity mapping on $\mathcal{P}(Y)$. For each $\nu \in \mathcal{P}(Y)$ we set $\eta_\nu(t) = f\bigl(\xi_{\Psi(\nu)}(t)\bigr)$, where ξ is a parameterization of measures from $\mathcal{P}(X)$ by Borel mappings of the interval $[0,1]$ to X. Then

$$\lambda \circ \eta_\nu^{-1} = \bigl(\lambda \circ \xi_{\Psi(\nu)}^{-1}\bigr) \circ f^{-1} = \Psi(\nu) \circ f^{-1} = \nu,$$

since $f \circ \Psi(y) = y$ for all $y \in Y$. If measures ν_n converge weakly to a measure ν on Y, then the measures $\Psi(\nu_n)$ converge weakly to the measure $\Psi(\nu)$ on X, hence $\xi_{\Psi(\nu_n)}(t) \to \xi_{\Psi(\nu)}(t)$ for almost all t in $[0,1]$, whence $\eta_{\nu_n}(t) \to \eta_\nu(t)$ for all such t by the continuity of f. □

Proof of Theorem 3.1. By Lemma 3.2, Corollary 2.8, and Lemma 3.3, it suffices to show that the interval $[0,1]$ has the strong Skorohod representation property. In this case, one can take for ξ_μ the mapping

$$\xi_\mu(t) = \sup_{x\in[0,1]} \bigl\{x : \mu[0,x) \le t\bigr\}, \tag{3.3}$$

for which our claim is verified directly (see [31, Ch. III, §8]).

Remark 3.4. It would be also enough to prove that the Hilbert cube $Q = [0,1]^\infty$ has the strong Skorohod representation property, since by the Uryson theorem (see [14, Theorem 4.2.10]), every Polish space is homeomorphic to a subset of Q. Hence it suffices to have at least one compact metric space Z with the strong Skorohod representation property that can be mapped onto Q by the aid of an open mapping. Since the space $\mathbb{R}^n$ has the strong Skorohod representation property by Krylov's theorem [22], it remains to find a compact set in $\mathbb{R}^n$ that can be openly mapped onto Q. Such compacta indeed exist [2]; moreover, one can find a compact set $Z \subset \mathbb{R}^3$ with the desired property (see [1, Ch. 4, §4] or [2]).

Remark 3.5. Let X be a complete separable metric space. Then the mapping $h\colon\ \mu \mapsto \xi_\mu$ is a homeomorphism from the space $\mathcal{P}(X)$ with the weak topology to the metric space $L^0(\lambda, X)$ of measurable mappings $\psi\colon\ [0,1] \to X$ with the metric

$$d(\psi,\varphi) = \int_0^1 \frac{d_X r\big(\psi(t),\varphi(t)\big)}{d_X\big(\psi(t),\varphi(t)\big)+1}\,dt,$$

where d_X is a metric in X. The metric d corresponds to the topology of convergence in measure, i.e., a sequence ψ_n converges in this topology to ψ if for every $\delta > 0$, we have $\lim\limits_{n\to\infty} \lambda\Big(t\colon\ d_X\big(\psi_n(t),\psi(t)\big) > \delta\Big) = 0$. Clearly, $\lambda \circ h(\mu)^{-1} = \mu$, i.e., h is a continuous right inverse for the mapping $\psi \mapsto \lambda \circ \psi^{-1}$ from $L^0(\lambda, X)$ onto $\mathcal{P}(X)$. We recall that in nontrivial cases the almost everywhere convergence cannot be topologized, i.e., one cannot define a topology, in which the convergent sequences are precisely the sequences convergent almost everywhere. For the absence of a topologization, it suffices to have a sequence that converges in measure but converges at no point. Indeed, by the Riesz theorem, every subsequence in this sequence has a subsequence convergent almost everywhere to one and the same limit, which, in the presence of a topology, would imply the convergence of the whole sequence.

Proposition 3.6. *Let X and Y be universally measurable sets in Polish spaces and let $f\colon\ X \to Y$ be an open surjective mapping. Then the mappings*

$$\widehat{f}\colon\ \mathcal{P}(X) \to \mathcal{P}(Y) \quad \textit{and} \quad f_*\colon\ L^0(\lambda, X) \to L^0(\lambda, Y),\ \varphi \mapsto f\circ\varphi$$

are open and surjective.

Proof. We shall first consider the mapping f_*. Let $V \subset L^0(\lambda, X)$ be an open set. In order to prove that $f_*(V)$ is open, it suffices to show that every neighborhood of $f\circ\varphi$, where $\varphi \in V$, contains a neighborhood $W \subset L^0(\lambda, Y)$, belonging to $f_*(V)$. Let us denote by d_Y the metric in Y and look for W of the form $U_Y(f\circ\varphi,\varepsilon)$, where

$$U_Y(\zeta,\delta) = \Big\{u \in L^0(\lambda, Y)\colon\ \lambda\Big(t \in [0,1]\colon\ d_Y\big(u(t),\zeta(t)\big) > \delta\Big) < \delta\Big\}.$$

Analogously we define the neighborhoods $U_X(\varphi,\delta)$ in $L^0(\lambda, X)$. Let us choose a neighborhood $U_X(\varphi,\delta) \subset V$. By Lusin's theorem, there exists a compact set $K \subset [0,1]$ with $\lambda(K) > 1 - \delta/4$, on which the mapping φ is continuous. Note that the validity of Lusin's theorem for measurable mappings with values in X and Y follows by the universal measurability of these spaces and the validity of Lusin's theorem for mappings with values in compact metric spaces, since for each $c > 0$, there exists a compact set $B \subset X$ such that $\lambda \circ \varphi^{-1}(X \backslash B) < c$. Let us choose in K disjoint compact parts $K_1, \ldots, K_n$ such that $\lambda\Bigl(K \backslash \bigcup_{j=1}^n K_j\Bigr) < \delta/4$ and the oscillation of φ on every part K_i does not exceed $\delta/4$, i.e., $d_X\bigl(\varphi(t), \varphi(s)\bigr) \leq \delta/4$ for all t, $s \in Z_i$. The compact sets $\varphi(K_i)$ have $\delta/4$-neighborhoods V_i. Since f is open, the sets $f(V_i)$ are open. There exists a positive number $\varepsilon < \delta/(4n)$ such that for every $i = 1, \ldots, n$, the open ε-neighborhood of the compact set $f\bigl(\varphi(K_i)\bigr)$ is contained in $f(V_i)$. Let $u \in U_Y\bigl(f \circ \varphi, \varepsilon\bigr)$. Making use of Lusin's theorem and the definition of $U_Y(f \circ \varphi, \varepsilon)$, we choose a compact part Z_i in every compact set K_i such that $\lambda(K_i \backslash Z_i) \leq \varepsilon$, $d_Y\bigl(u(t), f \circ \varphi(t)\bigr) \leq \varepsilon$ on $\bigcup_{i=1}^n Z_i$, and u is continuous on Z_i. Denote by ν the measure $\lambda \circ u^{-1}$. By the measurable choice theorem, there exist Borel mappings $g_i\colon\ f(V_i) \to V_i$ such that $f\bigl(g_i(y)\bigr) = y$ for ν-almost all $y \in f(V_i)$ (see, e.g., [12, Appendix 3] or [29, Theorem 13, Ch. II, §3, p. 127]). This means that $f\Bigl(g_i\bigl(u(t)\bigr)\Bigr) = u(t)$ for almost all $t \in Z_i$, since $u(t) \in f(V_i)$ if $t \in Z_i$ by virtue of our choice of ε and the estimate $d_Y\bigl(u(t), f \circ \varphi(t)\bigr) < \varepsilon$ for $t \in Z_i$. Let us define a mapping $\psi\colon\ [0,1] \to Y$ as follows: for $t \in Z_i$ we put $\psi(t) = g_i\bigl(u(t)\bigr)$, and outside $\bigcup_{i=1}^n Z_i$ we put $\psi(t) = g\bigl(u(t)\bigr)$, where $g\colon\ Y \to X$ is a Borel mapping such that $f\bigl(g(y)\bigr) = y$ for ν-almost all y (this mapping exists by the measurable selection theorem). Then, by the definition of the measure ν, we have $f\bigl(\psi(t)\bigr) = u(t)$ for almost all t in $[0,1]$. In addition, $d_X\bigl(\psi(t), \varphi(t)\bigr) \leq \delta/2$ for $t \in Z_i$ due to our choice of Z_i and the inclusion of $\psi(t)$ into the $\delta/4$-neighborhood of the set $\varphi(Z_i)$. Since $\lambda\Bigl(\bigcup_{j=1}^n (K_j \backslash Z_j)\Bigr) \leq \delta/4$, one has $\lambda\Bigl(\bigcup_{j=1}^n Z_j\Bigr) > 1 - \delta$. Hence $\psi \in U_X(\varphi, \delta)$. The surjectivity of f_* is obvious from the proof, since for every $u \in L^0(\lambda, Y)$ with $\lambda \circ u^{-1} = \nu$, we have $u = f \circ g$ almost everywhere, where $g\colon\ Y \to X$ is the aforementioned Borel mapping with $f\bigl(g(y)\bigr) = y$ ν-a.e.

Let us prove that $\widehat{f}$ is surjective and open. Let $h_1\colon\ \mathcal{P}(X) \to L^0(\lambda, X)$ and $h_2\colon\ \mathcal{P}(Y) \to L^0(\lambda, Y)$ be the homeomorphic embeddings mentioned in Remark 3.5. Then the mapping $\widehat{f} = h_2^{-1} \circ f_* \circ h_1$ is surjective. It follows from the reasoning above that it is open. Indeed, for the measure μ of the form $\mu = \lambda \circ \varphi^{-1}$, the sets $W(\mu, \delta) = \bigl\{\lambda \circ \zeta^{-1}\colon\ \zeta \in U_X(\varphi, \delta)\bigr\}$ form a basis of neighborhoods of μ in the weak topology, and the set $\widehat{f}\bigl(W(\mu, \delta)\bigr)$ contains a neighborhood $\bigl\{\lambda \circ v^{-1}\colon\ v \in U_Y(f \circ \varphi, \varepsilon)\bigr\}$, which belongs, as shown above, to $\widehat{f}\bigl(W(\mu, \delta)\bigr)$. $\square$

Remark 3.7. If X and Y are Polish spaces, then the mapping $f_*\colon\ L^0(\lambda, X) \to L^0(\lambda, Y)$ in the above proposition has a continuous right inverse g_*. It suffices to put $g_* = \xi \circ g \circ i_Y$, where $\xi\colon\ \mathcal{P}(X) \to L^0(\lambda, X)$ from Theorem 3.1, g is a continuous right inverse for $\widehat{f}$, and $i_Y\colon\ L^0(\lambda, Y) \to \mathcal{P}(Y)$, $i_Y(\psi) = \lambda \circ \psi^{-1}$. This can be also deduced from Michael's selection theorem.

Lemma 3.8. *Let S be a discrete space, i.e., every point in S is open. Then S has the strong Skorohod representation property.*

Proof. We may assume that S is completely ordered. Each Radon measure μ on S is concentrated on some countable set $\{\omega_i\}$. We take the unit interval $I = [0,1]$ with Lebesgue measure λ for a probability space. Let us represent I as a union of disjoint semi-open intervals $[a_i, b_i)$ and the point 1 such that $b_i - a_i = \mu(\omega_i)$ and if $\omega_{i_1} < \omega_{i_2}$, then $b_{i_1} \leq a_{i_2}$. Let us set $\xi_\mu(x) = \omega_i$ if $x \in [a_i, b_i)$. Since every sequence of Radon measures on S is concentrated on at most a countable set and the weak convergence $\mu_n \to \mu$ implies that $\mu_n(A) \to \mu(A)$ for every $A \subset S$, it is readily seen that ξ_μ is the desired parameterization.

Lemma 3.9. *The Baire space $B(\tau) = \prod_{i=1}^\infty X_i$, where X_i is a copy of a discrete space S of cardinality τ, has the strong Skorohod representation property.*

Proof. Let μ be a Radon probability measure on $B(\tau)$. We denote by $P_n \colon B(\tau) \to X_n$ the projection to the n-th factor. By Lemma 3.8, we obtain a parameterization $\xi^\mu = \xi_{\mu\circ P_1^{-1}} \colon [0,1] \to X_1$ for the projection to the first factor. Since $\mu \circ P_1^{-1}$ is Radon, the relation $\mu \circ P_1^{-1}(\omega_{i_1}) \neq 0$ holds for at most a countable collection of points $\omega_{i_1} \in X_1$, $i_1 = 1, 2, \ldots$. For every ω_{i_1}, we set

$$B_{i_1} = \{(x_1, x_2, \ldots, x_m, \ldots) \in B(\tau) \colon x_1 = \omega_{i_1}\}.$$

Let μ_{i_1} denote the conditional measure $\mu|_{B_{i_1}}/\mu \circ P_1^{-1}(\omega_{i_1})$. At all other points, we choose for conditional measures μ_ω arbitrary fixed probability measures. For every Borel set $A \subset B(\tau)$, we have

$$\mu(A) = \int \mu_{i_1}(A \cap B_{i_1})\, \mu \circ P_1^{-1}(dx_1).$$

If a sequence of Radon probability measures μ^n converges weakly to μ, then $\mu^n \circ P_1^{-1}(\omega_{i_1}) \neq 0$ for all n sufficiently large. Since B_{i_1} is open and closed, we conclude that the measures $\mu^n_{i_1}$ converge weakly to μ_{i_1}.

We observe that every space B_{i_1} is homeomorphic to $B(\tau)$. Let us apply the same procedure to each B_{i_1}, projecting it to X_2. For every measure μ_{i_1}, we consider its projection $\mu_{i_1} \circ P_2^{-1}$ to X_2. Let us denote by $\xi^\mu_{i_1}$ the Skorohod representation for μ_{i_1} in accordance with Lemma 3.8. For the points ω_{i_2} with $\mu_{i_1} \circ P_2^{-1}(\omega_{i_2}) \neq 0$, we construct sets

$$B_{i_1,i_2} = \Big\{(x_1, x_2, \ldots, x_m, \ldots) \colon \ x_1 = \omega_1, x_2 = \omega_2\Big\}$$

that are open and closed and obtain the conditional measures

$$\mu_{i_1,i_2} = \mu_{i_1}|_{B_{i_1,i_2}}/\mu_{i_1} \circ P_2^{-1}(\omega_{i_2}).$$

The weak convergence of $\{\mu^n\}$ to μ yields the almost everywhere convergence of $\xi^{\mu^n}_{i_1}$ to $\xi^\mu_{i_1}$ and the weak convergence of the measures $\mu^n_{i_1,i_2}$ to μ_{i_1,i_2}. Repeating this process countably many times, we obtain a sequence of indices $i_1, i_2, \dots, i_k, \dots$, each of which assumes at most countably many values, the measures $\mu_{i_1,i_2,\dots,i_k}$ on the sets $B_{i_1,i_2,\dots,i_k}$, and the Borel mappings $\xi^\mu_{i_1,i_2,\dots,i_k}\colon [0,1] \to X_{k+1}$ such that $\lambda \circ (\xi^\mu_{i_1,i_2,\dots,i_k})^{-1} = \mu_{i_1,i_2,\dots,i_k} \circ P^{-1}_{k+1}$. The weak convergence of a sequence $\{\mu^n\}$ to μ yields the almost everywhere convergence of $\xi^{\mu^n}_{i_1,i_2,\dots,i_k}$ to $\xi^\mu_{i_1,i_2,\dots,i_k}$ and the weak convergence of $\mu^n_{i_1,i_2,\dots,i_k}$ to $\mu_{i_1,i_2,\dots,i_k}$. In addition, for every Borel set B, we have the relation

$$\mu(B) = \int_{X_1} \mu \circ P_1^{-1}(dx_1) \int_{X_2} \mu_{i_1} \circ P_2^{-1}(dx_2) \cdots \int_{X_n} \mu_{i_1,\dots,i_n}(B \cap B_{i_1 \cdots i_n})\, \mu_{i_1,\dots,i_{n-1}} \circ P_n^{-1}(dx_n) \tag{3.4}$$

Now we define $\eta^\mu\colon [0,1]^\infty \to B(\tau)$ by setting

$$\eta^\mu(t_1, t_2, \dots, t_n, \dots) = \Big(\xi^\mu(t_1), \xi^\mu_{i_1}(t_2), \cdots, \xi^\mu_{i_1,i_2,\dots,i_{n-1}}(t_n), \dots\Big),$$

where every index i_n depends on $t_1, \dots, t_{n-1}$ as follows:

$$\xi^\mu(t_1) = \omega_{i_1}, \cdots, \xi^\mu_{i_1,i_2,\dots,i_{n-1}}(t_n) = \omega_{i_1,i_2,\dots,i_n}, \dots .$$

It is readily seen that η^μ is a Borel mapping. Let us verify that the countable product of Lebesgue measures $P = \bigotimes_{n=1}^\infty \lambda_n$ is transformed by η^μ to μ. It suffices to consider the sets of the form $B = B_1 \times B_2 \times \cdots B_n \times X_{n+1} \times X_{n+2} \cdots$, for which relation (3.4) takes the form

$$\mu(B) = \int_{B_1} \mu \circ P_1^{-1}(dx_1) \int_{B_2} \mu_{i_1} \circ P_2^{-1}(dx_2) \cdots \int_{B_n} \mu_{i_1,\dots,i_{n-1}} \circ P_n^{-1}(dx_n). \tag{3.5}$$

Indeed,

$$P\big((\eta^\mu)^{-1}(B)\big) = \int_{(\xi^\mu)^{-1}(B_1)} \lambda(dt_1) \int_{(\xi^\mu_{i_1})^{-1}(B_2)} \lambda(dt_2) \cdots \int_{(\xi^\mu_{i_1,\cdots,i_{n-1}})^{-1}(B_n)} \lambda(dt_n). \tag{3.6}$$

Comparing (3.5) and (3.6), we arrive at the desired equality $P\big((\eta^\mu)^{-1}(B)\big) = \mu(B)$. The family of mappings η^μ is the strong Skorohod representation. This is readily seen from the above mentioned properties of the mappings $\xi^\mu_{i_1,\dots,i_n}$. Clearly, we could take the unit interval with Lebesgue measure for a probability space. $\square$

Theorem 3.10. *Every complete metric space Y has the strong Skorohod representation property.*

Proof. According to [3, Ch. 6, Problem 150], for some cardinality τ, there exists an open surjection $f\colon B(\tau) \to Y$. The mapping $\widehat{f}\colon \mathcal{P}_r\big(B(\tau)\big) \to \mathcal{P}_r(Y)$ is open by Corollary 2.5. Now the claim follows by Lemma 3.9 and Lemma 3.3, since $\widehat{f}$ has a continuous right inverse by the above cited selection theorem of Michael, which is applicable, because the space $\mathcal{P}_r(M)$ with the weak topology is complete metrizable for every complete metric space M. $\square$

4 Additional Remarks

Recall that a topological space X is called a Prohorov space if every compact in the weak topology set M of probability measures on X is uniformly tight, i.e., for each $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset X$ such that $\mu(X\backslash K_\varepsilon) < \varepsilon$ for all $\mu \in M$. If every weakly convergent sequence of probability measures on X is uniformly tight, then X is called sequentially Prohorov (see [8] concerning the various versions of the Prohorov property). According to Prohorov's theorem [27], every complete separable metric space has the Prohorov property. Note that the Skorohod theorem implies that such a space X is sequentially Prohorov. Indeed, if mappings $\xi_n\colon [0,1] \to X$ converge almost everywhere to a mapping ξ, then by the theorems of Egorov and Lusin, for every $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset [0,1]$ such that all mappings ξ_n are continuous on K_ε and converge uniformly. The uniform tightness of the measures $\lambda \circ \xi_n^{-1}$ is readily seen from this. On the other hand, not every Borel subset of the interval has the Prohorov property. For example, as shown by D. Preiss [26], the space of rational numbers in the interval with its natural topology is not Prohorov. Thus, for universally measurable subsets of complete separable metric spaces, the strong Skorohod representation property is weaker than the Prohorov property, although the latter can be deduced from it for complete spaces. If X is a nonmetrizable Lusin space (i.e., a continuous one-to-one image of a Polish space), then it may fail to have the strong Skorohod representation property. As an example we mention the space l^2 with the weak topology. In this example, however, the space is not sequentially Prohorov (see the work [17], whose reasoning shows also the absence of the strong Skorohod representation property, and [20]); a Prohorov space with the same property is constructed in Example 4.1 below.

The definition of the strong Skorohod representation property can be extended by replacing the interval with Lebesgue measure by the class of all probability spaces (then such a property can be called the strong Skorohod representation property in the wide sense). This extension is reasonable for those topological spaces, on which not every probability measure is an image of Lebesgue measure. By analogy with the sequential Prohorov property and in the spirit of the original Skorohod theorem one can consider the sequential Skorohod representation property, i.e., a possibility to parameterize by mappings weakly convergent sequences of probability measures. For example, it was shown in [13] that every non complete separable metric space X possesses the sequential Skorohod representation property in the wide sense, (see [20] for additional related references). An intermediate property: a parameterization of weakly compact families of measures. Finally, the last two properties can be considered only for uniformly tight families of measures. It would be interesting to investigate nontrivial connections between the aforementioned properties. Some results in this direction have been recently obtained in [20]. In addition, it is of interest to consider non metrizable spaces with the strong Skorohod representation property and operations on them (for example, products, projective and inductive limits, etc.). It has been proved in [20] that if a Hausdorff space X admits a continuous injection to $\mathbb{R}^\infty$, then, for every uniformly tight sequence of Radon probability measures μ_n on X, one can find Borel mappings $\xi_n\colon [0,1] \to X$ with the distributions μ_n such that there is a subsequence $\xi_{n_i}\colon [0,1] \to X$ that converge almost everywhere (i.e., one has the Skorohod representation for subsequences). We shall give an example showing that even for Prohorov spaces, this may fail for the whole sequence $\{\mu_n\}$. In addition, we obtain below a similar parameterization of all measures on X simultaneously.

We recall that the locally convex space X is a strict inductive limit of an increasing sequence of locally convex spaces X_n that are linear subspaces in X if the convex neighborhoods of the origin in X are exactly those convex sets whose intersections with every X_n are open in X_n (see [28, Ch. V]).

Example 4.1. *Let $X = \mathbb{R}^\infty_0 = \{(x_1, \dots, x_n, 0, 0, \dots)\}$ be the strict inductive limit of the spaces $\mathbb{R}^n$. Then X is a Prohorov Souslin space, but there exists a weakly convergent sequence of probability measures on X that has no Skorohod representation.*

Proof. Let J be a closed interval with the end points a and b that belong to $\mathbb{R}^n$ for some n, but do not belong to $\mathbb{R}^{n-1}$. Given $\varepsilon > 0$, let us define the mapping $F(J,\varepsilon)$ from J to R^{n+1} as follows: the point $(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \cdots, \frac{a_n+b_n}{2})$ is sent to $(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \cdots, \frac{a_n+b_n}{2}, \varepsilon)$, $F(J,\varepsilon)(a) = a$, $F(J,\varepsilon)(b) = b$, and $F(J,\varepsilon)$ is linear on the intervals $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$. We equip $I = [0,1]$ with Lebesgue measure μ. We shall construct a sequence of measures μ_m that are the images of μ under certain homeomorphisms from I to some polygonal lines I_m. The process of construction consists of a sequence of steps. At the n-th step, we add to the previously constructed collection of measures some finite number of elements. At the first step, we set:

$$I_1 = F([0,1], 2^{-1})I, \quad \mu_1 = \mu \circ F^{-1}([0,1], 2^{-1}).$$

Let us describe the n-th step. In consists, in turn, of the n sub-steps. At the k-th sub-step, we obtain a collection of the polygonal lines $I_{n;i_1,\dots,i_k}$, where $1 \le i_j \le 2^{j-1}$. Every polygonal line $I_{n;i_1,\dots,i_k}$ will support some measure $\mu_{n;i_1,\dots,i_k}$. Each $I_{n;i_1,\dots,i_k}$ consists of the 2^k intervals $I^l_{n;i_1,\dots,i_k} = [a^l_{n;i_1,\dots,i_k}, b^l_{n;i_1,\dots,i_k}]$, $1 \le l \le 2^k$, and $\mu_{n;i_1,\dots,i_k}(I^l_{i_1,\dots,i_k}) = 2^{-k}$. At the first sub-step, we set

$$I_{n;1} = F([0,1], 2^{-n})I, \quad \mu_{n;1} = \mu \circ F^{-1}([0,1], 2^{-n}),$$

$$I^1_{n;1} = F([0,1], 2^{-n})[0, 2^{-1}], \quad I^2_{n;1} = F([0,1], 2^{-n})[2^{-1}, 1].$$

Suppose that the $(k-1)$'s sub-step is completed. Let us set

$$I^{2i_k-1}_{n;i_1,\dots,i_k} = F([I^{i_k}_{n;i_1,\dots,i_{k-1}}], 2^{-n})\Big[a^{i_k}_{n;i_1,\dots,i_{k-1}}, \frac{a^{i_k}_{n;i_1,\dots,i_{k-1}} + b^{i_k}_{n;i_1,\dots,i_{k-1}}}{2}\Big],$$

$$I^{2i_k}_{n;i_1,\dots,i_k} = F([I^{i_k}_{n;i_1,\dots,i_{k-1}}], 2^{-n})\Big[\frac{a^{i_k}_{n;i_1,\dots,i_{k-1}} + b^{i_k}_{n;i_1,\dots,i_{k-1}}}{2}, b^{i_k}_{n;i_1,\dots,i_{k-1}}\Big].$$

If $l \ne 2i_k - 1$ and $l \ne 2i_k$, then we put

$$I^l_{n;i_1,\dots,i_k} = \Big[a^r_{n;i_1,\dots,i_{k-1}}, \frac{a^r_{n;i_1,\dots,i_{k-1}} + b^r_{n;i_1,\dots,i_{k-1}}}{2}\Big], \quad l = 2(r-1)+1,$$

$$I^l_{n;i_1,\dots,i_k} = \Big[\frac{a^r_{n;i_1,\dots,i_{k-1}} + b^r_{n;i_1,\dots,i_{k-1}}}{2}, b^r_{n;i_1,\dots,i_{k-1}}\Big], \quad l = 2r.$$

The polygonal line $I_{n;i_1,\cdots,i_k}$ is the union of the intervals $I^l_{n;i_1,\cdots,i_k}$. The measure $\mu_{n;i_1,\dots,i_k}$ is the image of the measure $\mu_{n;i_1,\dots,i_{k-1}}$ under the mapping

$$F_{n;i_k}\colon\ I_{n;i_1,\dots,i_{k-1}} \to I_{n;i_1,\dots,i_k}$$

that coincides with $F[I^{i_k}_{n;i_1,\dots,i_{k-1}}, 2^{-n}]$ on the interval $I^{i_k}_{n;i_1,\dots,i_{k-1}}$, and is identity on its complement. The n-th sub-step is completed, and we add the measures $\mu_{n;i_1,\dots,i_n}$ to our sequence. Then we proceed to the $(n+1)$-th step. Finally, we numerate the measures we have constructed in the natural order and obtain the sequence $\{\mu_m\}$. The sequence $\{\mu_m\}$ converges weakly to μ. Indeed, it suffices to show that for every open set U, one has $\underline{\lim}_{m\to\infty}\mu_m(U) \ge \mu(U)$. Let P^n be the natural projection to $\mathbb{R}^n$. It is clear that the measures $\mu_m \circ (P^n)^{-1}$ converge weakly to μ as $m \to \infty$. Hence $\underline{\lim}_{m\to\infty}\mu_m \circ (P^n)^{-1}(U) \ge \mu(U)$. But $|\mu_m(U) - \mu_m \circ (P^n)^{-1}(U)| \le 2^{-n}$. This proves our claim.

Suppose now that there exists the corresponding Skorohod parameterization $\{\xi_{\mu_m}\}$ for $\{\mu_m\}$. For every n, let us consider the set $\{A_n\colon\ \xi_{\mu_m} \in \mathbb{R}^n, \forall\, m \in \mathbb{N}\}$. Since the sequence ξ_{μ_m} converges almost surely, one has $P(\cup_n A_n) = 1$. Let us pick N such that $P(A_N) > 0$. Taking the restrictions of the random variables ξ_{μ_m} to A_N, we obtain a sequence of measures $\tau_m = P|_{A_N} \circ {\xi_{\mu_m}}^{-1}$ with the following properties:

the measure $\mu_m - \tau_m$ is nonnegative, the measures τ_m are concentrated on $\mathbb{R}^N$ and converge weakly on $\mathbb{R}^N$. However, as one can easily see, the limit of such measures can be only zero, which contradicts the fact that $P(A_N) > 0$. Indeed, the space $\mathbb{R}^N$ can be partitioned into sets $A_1, \ldots, A_k, A$ such that $A_1, \ldots, A_k$ are open and disjoint, A is the complement of their union, $\mu(A) = 0$, and for every A_i, there exists infinitely many values m_l such that $\tau_{m_l}(A_i) = 0$. Now it remains to note that τ_{m_i} cannot be bigger than μ. $\square$

We shall say that a topological space X has the weak Skorohod representation property if, for every Radon probability measure μ on X, there exists a μ-measurable mapping $\xi_\mu\colon\ \Omega \to X$ on a probability space (Ω, P) such that the law of ξ_μ is μ, and for every uniformly tight sequence of Radon probability measures μ_n on X, there is a subsequence in $\{\xi_{\mu_n}\}$ that converges almost surely.

The idea of the next lemma goes back to [20].

Lemma 4.2. (i) *Let X be a topological space that has a continuous injection to $\mathbb{R}^\infty$ and let $\Phi\colon\ X \to Y$ be a Borel mapping with values in a metrizable Souslin space Y such that the sets $\Phi(A)$ and $\Phi^{-1}(B)$ have compact closures for all compact sets A and B. Then X has the weak Skorohod representation property.*

(ii) *Let f be a continuous surjection of a space X with the weak Skorohod representation property onto a space Y such that the preimages of compact sets are compact. Then Y has the weak Skorohod representation property.*

Proof. (i) Let $F\colon\ X \to \mathbb{R}^\infty$ be an injective continuous mapping and let

$$\Gamma\colon\ X \to \mathbb{R}^\infty \times Y, \quad \Gamma(x) = \big(F(x), \Phi(x)\big).$$

Let us consider the Skorohod representation (ξ_μ, η_μ) for the measures of the form $\mu\circ\Gamma^{-1}$ on the space $\mathbb{R}^\infty \times Y$, where μ is a Radon probability measure on X. Clearly, in this case, one can take the unit interval with Lebesgue measure for a probability space (Ω, P). We shall show that the mapping $\mu \mapsto \zeta_\mu := F^{-1} \circ \xi_\mu$ is the desired representation on X. Since μ is concentrated on a countable union of compact sets in X, the set $F(X)$ is $\mu \circ F^{-1}$-measurable, and the mapping F^{-1} on $F(X)$ is μ-measurable (in fact, the compact sets in X are metrizable due to our hypothesis, hence Radon measures on X are concentrated on Souslin sets). The image of P under the mapping ξ_μ coincides with $\mu \circ F^{-1}$, hence $P \circ \zeta_\mu^{-1} = \mu$. We observe that

$$\eta_\mu = \Phi \circ F^{-1}(\xi_\mu) \quad \text{a.e.} \tag{4.7}$$

Indeed, let us apply the mapping

$$T\colon\ (x,y) \mapsto \big(\Phi \circ F^{-1}(x), y\big), \quad \big(F(X), Y\big) \to (Y, Y),$$

to the random variable (ξ_μ, η_μ). The image of Lebesgue measure under the mapping $T \circ (\xi_\mu, \eta_\mu)$ coincides with the measure $\mu \circ \Gamma^{-1} \circ T^{-1}$, which is the image of μ under the mapping $G\colon x \mapsto \bigl(\Phi(x), \Phi(x)\bigr)$, $X \to Y \times Y$. Clearly, $\mu \circ G^{-1}$ is concentrated on the diagonal $\Delta = \{(y,y)\colon y \in Y\}$. Hence $T\bigl(\xi_\mu(t), \eta_\mu(t)\bigr) \in \Delta$ for a.e. $t \in [0,1]$, which yields (4.7). By the properties of Φ, we obtain that if $\{\mu_n\}$ is a uniformly tight sequence of Radon probability measures on X, then the sequence $\{\mu_n \circ \Gamma^{-1}\}$ is uniformly tight as well. Hence there exists a a subsequence $(\xi_{n_i}, \Phi \circ F^{-1}(\eta_{n_i}))$ that converges almost surely in $\mathbb{R}^\infty \times Y$. Then, for almost all ω, the sequence $\Phi \circ F^{-1}(\xi_{n_i})(\omega)$ is convergent, hence, by the properties of Φ, the sequence $F^{-1}(\xi_{n_i})$ belongs to some compact set $K(\omega)$ in the space X. By its convergence in $\mathbb{R}^\infty$, it converges also in X, since the topologies of X and $\mathbb{R}^\infty$ coincide on $K(\omega)$.

(ii) Let $\mu \mapsto \xi_\mu$ be a parameterization of Radon probability measures on X as required in the definition of the weak Skorohod representation property. It follows from our assumption on f (see, e.g., [8, Theorem 6.1.8]) that for every Radon probability measure ν on Y, there exists a Radon probability measure μ_ν on X such that $\mu_\nu \circ f^{-1} = \nu$. Set $\eta_\nu = f \circ \xi_{\mu_\nu}$. If a sequence of Radon probability measures ν_n on Y is uniformly tight, then the sequence μ_{ν_n} is uniformly tight as well. It follows that $\{\eta_{\nu_n}\}$ has a subsequence that converges almost surely. □

Let us consider examples of spaces with the weak Skorohod representation property. More examples can be produced by using the fact that every universally measurable subset of a space with weak Skorohod representation property has this property as well.

Example 4.3. *Let $X = \prod_{m=1}^\infty X^m$, where each space X_m admits a continuous injection in $\mathbb{R}^\infty$ and has a fundamental sequence $\{K_n^m\}$ of compacts sets (i.e., every compact subset of X_m is contained in one of the K_n^m's). Then X has the weak Skorohod representation property.*

Proof. Let K_n^m be such that $K_n^m \subset K_{n+1}^m$ and let

$$\Phi\colon\ X \to \mathbb{N}^\infty, \quad \bigl(\Phi(x)^m\bigr) = \{\min n : x^m \in K_n^m\}.$$

It remains to apply Lemma 4.2. □

Note that the space of tempered distributions $\mathcal{S}'$ has a foundamental sequence of compact sets; the same is true for all separable reflexive Banach spaces with the weak topology, and the duals to separable Banach spaces with the weak-$*$ topology. Example 4.3 covers also the space of distributions $\mathcal{D}'$ (which is homeomorphic to a closed subspace of a countable product of spaces having fundamental sequences of compact sets).

Example 4.4. *Let a locally convex space X be the strict inductive limit of an increasing sequence of its closed metrizable Souslin linear subspaces X_j. Then X has the weak Skorohod property. The same is true for any countable product of the above inductive limits.*

Proof. It is known (see, e.g., [8, Example 8.1.14]) that if a sequence of probability measures μ_n on X converges weakly to a measure μ, then, for every ε, there exists a number j such that $\mu_n(X_j) > 1-\varepsilon$ for all n. Let ϱ_α be a family of seminorms defining the topology in X. Since X_j is metrizable, there exists a countable family of seminorms ϱ^j_m, $m=1,2,\ldots$, whose restrictions on X_j generate the topology of X_j. Let us consider a new topology τ on X by means of the seminorms ϱ^j_m, $j=1,2,\ldots$, $m=1,2,\ldots$. This is a locally convex metrizable topology. Let $F\colon\ X\to(X,\tau)$ be the identity mapping and let $\Phi(x)=\{\min j\colon\ x\in X_j\}$. Applying the same reasoning as in Lemma 4.2 (which is not applicable directly) and taking into account that τ coincides with the initial topology on every X_j, we obtain the desired assertion. The second claim is analogous. $\square$

Note that Example 4.4 covers the space $\mathcal{D}(\mathbb{R}^d)$ of smooth compactly supported functions on $\mathbb{R}^d$ (with its natural topology of the inductive limit of the spaces $C_0^\infty(U_n)$, where U_n is the centered open ball of radius n).

Let us make several remarks about the properties of mappings ξ_μ, which can parameterize measures on a Polish space X. The mappings constructed by N. V. Krylov [22] in the case of $\mathbb{R}^n$ are almost everywhere differentiable. Clearly, these mappings cannot be always made continuous. As shown in [21], a probability measure μ on X is the image of Lebesgue measure on $[0,1]$ under a continuous mapping precisely when the topological support of μ (the smallest closed set of full measure) is connected and locally connected compact. The mappings defined by formula (3.3) are monotone. Hence the mappings ξ_μ constructed in our proof of Theorem 3.1 belong to the Skorohod space $D(X)$, which consists of the mappings from $[0,1]$ to X that are right continuous and have left-side limits at all points in $(0,1]$. If X is a complete separable metric space, then $D(X)$ is equipped with a certain natural complete separable metric d_0 (see [6], [19]). However, even in the case where $X=[0,1]$, there exists no parameterization of probability measures on $[0,1]$ by mappings of the class $D([0,1])$, for which the weak convergence of measures implies the convergence of the corresponding mappings in the Skorohod topology (see [6]). For example, let μ_n on $[0,1]$ be defined by $\mu_n(\{0\})=1/n$, $\mu_n(\{1\})=1-1/n$. This sequence converges weakly (even in the variation norm) to Dirac's measure at 1, but the distances with respect to d_0 between any mappings from $D([0,1])$ that induce these measures equal 1.

The authors are grateful to N. V. Krylov, V. I. Ponomarev, M. Röckner, and V. V. Fedorchuk for useful discussions. We thank the Department of Mathematics of the Bielefeld University and the SFB 343, where this paper has been completed, for support and hospitality. This work was supported in part by the Russian Foundation of Fundamental Research projects 96–15–96865 and 97–01–00932, the INTAS Grant 94–378, the INTAS-RFBR Grant 95-0099, the SFB 343, and DFG Grant 436 RUS 113/343/0(R).

References

[1] Aleksandrov P.S., Pasynkov B.A., Introduction to the dimension theory, Nauka, Moscow, 1973 (in Russian).

[2] Anderson R.D., Monotone interior dimension-raising mappings, Duke Math. J. 19 (1952), 359–366.

[3] Arkhangelskii A.V., Ponomarev V.I., Fundamentals of general topology. Problems and exercises, Reidel Publishing Co., Dordrecht – Boston, Mass., 1984.

[4] Banakh T.O., Topology of spaces of probability measures, Matematichni Studii, Pratsi Lviv, Mathem. Collective, 5 (1995), 65–87 (in Russian).

[5] Banakh T.O., Radul T.N., Geometry of mappings of spaces of probability measures, Matematichni Studii (1998) (in Russian).

[6] Billingsley P., Convergence of probability measures, Wiley, New York, 1968.

[7] Blackwell D., Dubins L.E., An extension of Skorohod's almost sure representation theorem, Proc. Amer. Math. Soc. 89 (1983), no. 4, 691–692.

[8] Bogachev V.I., Measures on topological spaces, J. Math. Sci. (New York) 91 (1998), no. 4, 3033–3156.

[9] Bourbaki N., Intégration, Hermann, Paris.

[10] Cuesta-Albertos J.A., Matrán-Bea C., Stochastic convergence through Skorohod representation theorems and Wasserstein distances, In: First International Conference on Stochastic Geometry, Convex Bodies and Empirical Measures (Palermo, 1993). Rend. Circ. Mat. Palermo (2) Suppl. No. 35 (1994), 89–113.

[11] Ditor S., Eifler L., Some open mapping theorems for measures, Trans. Amer. Math. Soc. 164 (1972), 287–293.

[12] Dynkin E.B., Yushkevich A.A., Controlled Markov processes, Nauka, Moscow, 1975 (in Russian); English transl.: Springer, Berlin, 1979.

[13] Dudley R.M., Distances of probability measures and random variables, Ann. Math. Stat. 39 (1967), 1563–1572.

[14] Engelking R., General topology, Warszawa, 1977.

[15] Ershov M. P., Extension of measures and stochastic equations, Teor. Veroyatnost. i ee Primen. 19 (1974), no. 3, 457–471 (in Russian); English transl.: Theor. Probab. Appl. 19 (1974), 431–444.

[16] Ershov M. P., The Choquet theorem and stochastic equations, Analysis Math. 1 (1975), no. 4, 259–271.

[17] Fernique X., Processus linéaires, processus généralisés, Ann. Inst. Fourier, Grenoble, 17 (1967), 1–92.

[18] Fernique X., Un modèle presque sûr pour la convergence en loi, C. R. Acad. Sci. Paris, Sér. 1, 306 (1988), 335–338.

[19] Jakubowski A., On the Skorokhod topology, Ann. Inst. H. Poincare Probab. Statist. 22 (1986), no. 3, 263–285

[20] Jakubowski A., The almost sure Skorokhod representation for subsequences in nonmetric spaces, Teor. Veroyatn. i Primenen. 42 (1997), no. 1, 209–217; English transl.: Theory Probab. Appl. 42 (1997), no. 1, 167–174.

[21] Kolesnikov A.V., On continuous images of Lebesgue measure, Matem. Zametki, 1999 (in Russian); English transl.: Math. Notes (1999).

[22] Krylov N.V., On SPDEs and superdiffusions, Ann. Probab. 25 (1997), 1789–1809.

[23] Lembcke J., On simultaneous preimage measures on Hausdorff spaces, Lecture Notes in Math. 945 (1982), 110–115.

[24] Michael E., A selection theorem, Proc. Amer. Math. Soc. 17 (1966), 1404–1406.

[25] Neveu J., Bases mathématiques du calcul des probabilités, Masson et Cie, Paris, 1964.

[26] Preiss D., Metric spaces in which Prohorov's theorem is not valid, Z. Wahrsch. theor. verw. Geb. 27 (1973), 109–116.

[27] Prohorov Yu.V., Convergence of random processes and limit theorems in probability theory, Theor. Probab. Appl. 1 (1956), 157–214.

[28] Robertson A.P., Robertson W., Topological vector spaces, Cambridge University Press, 1966.

[29] Schwartz L., Radon measures on arbitrary topological spaces and cylindrical measures, Oxford Univ. Press, London, 1973.

[30] Sion M., On uniformization of sets in topological spaces, Trans. Amer. Math. Soc. 96 (1960), 237–245.

[31] Shiryaev A.N., Probability, Nauka, Moscow, 1989 (in Russian); English transl.: Springer, 1990.

[32] Skorohod A.V., Limit theorems for stochastic processes, Theor. Probab. Appl. 1 (1956), no. 3, 261–290.

[33] Topsøe F., Topology and measure, Lecture Notes in Math. 133 (1970).

[34] Vakhania N.N., Tarieladze V.I., Chobanyan S.A., Probability distributions in Banach spaces, Nauka, Moscow, 1984 (in Russian); English transl.: Kluwer Academic Publishers, 1991.