A unification of Prohorov's and Skorohod's ideas: convergence in distribution in nonmetric spaces

Adam Jakubowski*
Uniwersytet Mikołaja Kopernika, Toruń, Poland

August 15, 1994

Abstract

In the paper a new topology is defined of the space $\mathcal{P}(\mathcal{X})$ of tight probability distributions on a topological space $(\mathcal{X}, \tau)$. The only topological assumption imposed on $(\mathcal{X}, \tau)$ is that some countable family of continuous functions separates points of $\mathcal{X}$. This new sequential topology, defined by means of a variant of the a.s. Skorohod representation, is quite operational and from the point of view of nonmetric spaces proves to be more satisfactory than the weak topology. In particular, in this topology both the direct and the converse Prohorov's theorems are quite natural and hold in many spaces. The topology coincides with the usual topology of weak convergence in case when $(\mathcal{X}, \tau)$ is a metric space.

1 Convergence in distribution of random elements

It is a traditional point of view that the kind of convergence of probabilities encountered in weak limit theorems of probability theory is exactly the "weak convergence" of distributions of random elements, i.e. convergence $X_n \longrightarrow_{\mathcal{D}} X_0$ is defined as

$$Ef(X_n) \longrightarrow Ef(X_0), \quad \text{as } n \to +\infty, \tag{1}$$

for each bounded and continuous function f defined on the space $\mathcal{X}$, in which $X_0, X_1, \ldots$ take values ($f \in CB(\mathcal{X})$). Since the distributions $P_{X_n} = P \circ X_n^{-1}$

*Research supported by Komitet Badań Naukowych under the Grant No 2 1108 91 01, and completed while the author was visiting Université de Rennes I and Universität Bielefeld.

are measures on some σ-algebra of subsets of $\mathcal{X}$ (usually on the Borel or Baire σ-algebras), there is a tendency to avoid probabilistic formulation and consider an abstract convergence $\mu_n \Longrightarrow \mu_0$ rather than (1), where $\mu_n \Longrightarrow \mu_0$ means

$$\int_{\mathcal{X}} f(x)\,\mu_n(dx) \longrightarrow \int_{\mathcal{X}} f(x)\,\mu_0(dx), \quad f \in CB(\mathcal{X}). \tag{2}$$

The most successful step towards the abstract setting was done by Prohorov in his fundamental paper [16], and the complete theory when $\mathcal{X}$ is a Polish space has been given in excellent books by Parthasarathy [14] and Billingsley [2]. Within this theory, the crucial method for proving weak convergence is the following "three-stage procedure":

1. Check *relative compactness* of $\{\mu_n\}$, i.e. whether every subsequence $\{\mu_{n_k}\}$ contains a further subsequence $\{\mu_{n_{k_l}}\}$ weakly convergent to *some* limit.

2. By some other tools (characteristic functionals, finite dimensional convergence, martingale problem, etc.) *identify* all limiting points of weakly convergent subsequences $\{\mu_{n_k}\}$ with some distribution μ_0.

Then conclude $\mu_n \Longrightarrow \mu_0$.

It is worth to emphasize that this reasoning is based on the following property of the weak convergence (obvious, when definition (2) is in force):

> If every subsequence $\{\mu_{n_k}\}$ contains a further subsequence $\{\mu_{n_{k_l}}\}$ weakly convergent to μ_0, then the whole sequence $\{\mu_n\}$ converges weakly to μ_0. (3)

The main Prohorov's contribution was providing a very efficient criterion of relative compactness. Due to the *direct Prohorov's theorem*, a family $\{\mu_i\}_{i\in\mathbb{I}}$ of probability laws on a *metric* space $(\mathcal{S}, \mathcal{B}_{\mathcal{S}})$ is relatively compact, if it is *uniformly tight*, i.e. for every $\varepsilon > 0$ there is a *compact* set $K_\varepsilon \subset \mathcal{S}$ such that

$$\mu_i(K_\varepsilon) > 1 - \varepsilon, \quad i \in \mathbb{I}. \tag{4}$$

The *converse Prohorov's theorem* states that in *Polish* spaces relative compactness implies uniform tightness.

There exist, however, separable metric spaces for which the converse Prohorov's theorem is not valid [4], with rational numbers $\mathbb{Q}$ being the most striking example [15]. Let us notice that every probability measure on $(\mathbb{Q}, \mathcal{B}_{\mathbb{Q}})$ must be tight, and so, by LeCam's theorem ([13], [2]) weak convergence of probability measures on $\mathbb{Q}$ implies uniform tightness. LeCam's theorem holds also in arbitrary metric spaces, provided we restrict weak convergence to the space $\mathcal{P}(\mathcal{X})$ of *tight* probability measures on $\mathcal{X}$. We may summarize the theory for metric spaces by saying that in $\mathcal{P}(\mathcal{X})$ relative compactness is equivalent to *relative* *uniform tightness*, with the latter meaning that in every subsequence there is a further subsequence which is uniformly tight.

After leaving the (relatively) safe area of metric spaces, the abstract setting brings many disturbing problems, even if we remain in the world of random elements with tight distributions. Let us consider, for example, the infinite dimensional separable Hilbert space $(H, <, >)$ equipped with the weak topology $\tau_w = \sigma(H, H)$. It is a completely regular space (for it is a linear topological space), and since H with the norm topology is Polish, (H, τ_w) is also Lusin in the sense of Fernique ("espace separeé" in [6]). But Fernique [6] gives an example of an H-valued sequence $\{X_n\}$ satisfying

$$Ef(X_n) \longrightarrow f(0), \quad \text{as } n \to +\infty, \tag{5}$$

for each bounded and *weakly* continuous function $f : H \to \mathbb{R}^1$, and such that for each $K > 0$

$$\liminf_{n\to+\infty} P(\|X_n\| > K) = 1. \tag{6}$$

This means that on the space (H, τ_w) there are weakly convergent sequences (to $\mu_0 = \delta_0$ in (5)) with no subsequence being uniformly tight. It follows that the approach based on the direct Prohorov's theorem is no longer a universal tool for the weak convergence on neither completely regular nor Lusin spaces.

Nevertheless, since compacts in (H, τ_w) are metrisable, the direct Prohorov's theorem remains valid in (H, τ_w) (see [18]). But again the picture is not clear, since uniform tightness on (H, τ_w), i.e.

$$\lim_{K\to+\infty} \sup_n P(\|X_n\| > K) = 0, \tag{7}$$

implies relative compactness in topology strictly finer than the topology of weak convergence of measures on (H, τ_w), namely the topology of weak convergence of measures on H equipped with the sequential topology $(\tau_w)_s$ of weak convergence of elements of H. The direct proof of this fact is not difficult, but it seems to be more instructive to apply Theorem 1 of [8], which asserts that every sequence satisfying (7) contains a subsequence $\{X_{n_k}\}$ which admits the a.s. Skorohod representation: one can define on the Lebesgue interval $([0, 1], \mathcal{B}_{[0,1]}, \ell)$ H-valued random elements $Y_0, Y_1, \ldots$ such that

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \ldots \tag{8}$$

and for each $y \in H$ and each $\omega \in [0, 1]$

$$< y, Y_k(\omega) > \longrightarrow < y, Y_0(\omega) >, \quad \text{as } k \to \infty. \tag{9}$$

By the last line, for every *sequentially weakly* continuous function $f : H \to \mathbb{R}^1$ we have $f(Y_k(\omega)) \to f(Y_0(\omega))$, $\omega \in [0, 1]$, and if f is bounded,

$$Ef(X_{n_k}) = Ef(Y_k) \longrightarrow Ef(Y_0), \quad \text{as } k \to \infty. \tag{10}$$

One may rise a question whether there is a general notion of convergence in distribution which on a broad class of topological spaces shares the advantageous properties of the weak convergence of probability measures on metric spaces with respect to Prohorov's theorems.

In this paper we suggest a new definition of the convergence in distribution of random elements with *tight* laws, $\stackrel{*}{\Longrightarrow}$ say, which is defined by means of a variant of the a.s. Skorohod representation:

$\mu_n \stackrel{*}{\Longrightarrow} \mu_0$ iff every subsequence $\{n_k\}$ contains a further subsequence $\{n_{k_l}\}$ such that μ_0 and $\{\mu_{n_{k_l}} : l = 1, 2, \ldots\}$ admit a Skorohod representation defined on the Lebesgue interval and almost surely convergent "in compacts". (11)

(For precise definitions we refer to Section 3). Somewhat unexpectedly, this apparently very strong definition may be applied in most cases of interest, is quite operational and proves to be more satisfactory from the point of view of nonmetric spaces. In particular, $\mathcal{P}(\mathcal{X})$ equipped with the sequential topology determined by $\stackrel{*}{\Longrightarrow}$ has the following remarkable properties:

- "relatively compact" set of tight probability measures means exactly "relatively uniformly tight" (Theorem 3.5, Section 3);
- the converse Prohorov's theorem is quite natural and holds in many spaces (Theorems 4.1 – 4.5 and 4.7, Section 4);
- no assumptions like the T_3 (regularity) property are required for the space $\mathcal{X}$, what is very important in applications to sequential spaces (Section 2);
- on metric spaces the theory of the usual weak convergence of tight probability distributions remains unchanged (Theorem 3.8, Section 3).

2 Topological preliminaries

Let $(\mathcal{X}, \tau)$ be a topological space. Denote the convergence of sequences in τ-topology by " $\longrightarrow_\tau$" and by "τ_s" the sequential topology generated by τ-convergence. Recall that

$F \subset \mathcal{X}$ is τ_s-closed if F contains all limits of τ-convergent sequences of elements of F. (12)

Our basic assumption is:

There exists a countable family $\{f_i : \mathcal{X} \to [-1, 1]\}_{i \in \mathbb{I}}$ of τ-continuous functions, which separate points of $\mathcal{X}$. (13)

This condition is not restrictive and possesses several important implications which allow to built an interesting theory. As the most immediate consequence we obtain a convenient criterion for τ-convergence:

If $\{x_n\} \subset \mathcal{X}$ is relatively compact, and for each $i \in I\!I$ $f_i(x_n)$ converges to some number α_i, then x_n τ-converges to some x_0 and $f_i(x_0) = \alpha_i,\ i \in I\!I$. (14)

Assumption (13) defines a continuous mapping $\tilde{f} : \mathcal{X} \to [-1,1]^{I\!I}$ given by formula

$$\tilde{f}(x) = (f_i(x))_{i \in I\!I}. \tag{15}$$

By the separation property of the family $\{f_i\}_{i \in I\!I}$

$\mathcal{X}$ is a Hausdorff space (but need not be regular). (16)

There is an example of Hausdorff non-regular space, which will be referred to as "standard" and which is also suitable for our needs: take $\mathcal{X} = [0,1]$ and let the family of closed sets be generated by all sets closed in the usual topology and one *extra* set $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$. Then $\mathcal{X}$ is not a regular space [12], but still satisfies (13).

Let us observe that for any *compact* set $K \subset \mathcal{X}$ the image $\tilde{f}(K) \subset [-1,1]^{I\!I}$ is again compact and since $K = \tilde{f}^{-1}(\tilde{f}(K))$ we get

Every compact subset is $\sigma(f_i : i \in I\!I)$-measurable (hence is a Baire subset of $\mathcal{X}$) and is metrisable. (17)

In many cases $\sigma(f_i : i \in I\!I)$ is just the Borel σ-algebra. In any case every *tight* Borel probability measure on $(\mathcal{X}, \tau)$ is uniquely defined by its values on $\sigma(f_i : i \in I\!I)$. Moreover, every tight probability measure μ defined on $\sigma(f_i : i \in I\!I)$ can be uniquely extended to the whole σ-algebra of Borel sets. Hence if $X : (\Omega, \mathcal{F}, P) \to \mathcal{X}$ is $\sigma(f_i : i \in I\!I)$-measurable and the law of X (as the measure on $\sigma(f_i : i \in I\!I)$) is tight, then X is Borel-measurable if we replace $\mathcal{F}$ with its P-completion $\overline{\mathcal{F}}$. In particular, if $\{f'_i\}_{i \in I\!I'}$ is another family satisfying (13), then $X : (\Omega, \overline{\mathcal{F}}, P) \to \mathcal{X}$ is $\sigma(f'_i : i \in I\!I')$-measurable.

The above remarks show that our considerations do not depend essentially on the choice of the family $\{f_i\}_{i \in I\!I}$ satisfying (13). Therefore without loss of generality we may fix *some* family $\{f_i\}_{i \in I\!I}$ and shall restrict the attention to **random elements X such that $f_i(X),\ i \in I\!I$, are random variables and the law of X is tight, and to tight probability measures defined on $\sigma(f_i : i \in I\!I)$.** As in Section 1, the family of such measures will be denoted by $\mathcal{P}(\mathcal{X})$.

Every *tight* probability measure on $\mathcal{X}$ is the law of some $\mathcal{X}$-valued random element defined on the standard probability space $([0,1], \mathcal{B}_{[0,1]}, \ell)$. (18)

To see this, let us notice that $\tilde{f}$ is *one-to-one* and *continuous*, but (in general) is not a homeomorphism of $\mathcal{X}$ onto a subspace of $[0,1]^{\Pi}$. Nevertheless $\tilde{f}$ is a homeomorphic imbedding, if restricted to each *compact* subset $K \subset \mathcal{X}$, and so it is a *measurable isomorphism*, if restricted to each σ-compact subspace of $\mathcal{X}$. If μ is a tight probability measure, then it is concentrated on some σ-compact subspace $\mathcal{X}_1$ of $\mathcal{X}$, and $\mu \circ \tilde{f}^{-1}$ is a probability measure on $[0,1]^{\Pi}$, concentrated on the σ-compact subspace $\tilde{f}(\mathcal{X}_1)$. But it is well-known (see e.g. [3]) that then there exists a measurable mapping $Y : [0,1] \to [0,1]^{\Pi}$ such that

$$\mu \circ \tilde{f}^{-1} = \ell \circ Y^{-1}, \tag{19}$$

and, in particular, $Y \in \tilde{f}(\mathcal{X}_1)$ with probability one. It remains to take any $x_0 \in \mathcal{X}_1$ and define

$$X(\omega) = \begin{cases} \tilde{f}^{-1}(Y(\omega)), & \text{if } Y(\omega) \in \tilde{f}(\mathcal{X}_1); \\ x_0, & \text{otherwise.} \end{cases} \tag{20}$$

Using somewhat subtler reasoning than the one used in the proof of (17) we see that for *relatively compact* $K \subset \mathcal{X}$, the set $\tilde{f}^{-1}(\overline{\tilde{f}(K)})$ is both a τ-closed subset of $\mathcal{X}$ and the closure of K in the sequential topology τ_s. Hence we have

The closure of a relatively compact subset consists of limits of its convergent subsequences (but still need not be compact). (21)

Here again the standard example exhibits the pathology signalized in (21): the whole space $[0,1]$ is not compact, but it is a closure of a relatively compact set $[0,1] \setminus A$. Remark (21) affects the definition of uniform tightness where we cannot, in general, replace sequential compactness with measurability and relative compactness. In a similar way as (21) one can prove

$K \subset \mathcal{X}$ is compact iff it is sequentially compact. (22)

This in turn implies that

The sequential topology τ_s is the finest topology on $\mathcal{X}$ in which compact subsets are the same as in τ. (23)

To prove (23) let us observe first that $(\mathcal{X}, \tau_s)$ also satisfies (13), for τ-continuity implies *sequential* τ-continuity and so τ_s-continuity. By (22) compactness and sequential compactness are equivalent for both τ and τ_s. Since sequential compactness in τ and τ_s coincide, τ_s preserves the family of τ-compact subsets. It remains to prove that if $\tau' \supset \tau$, τ'-compacts coincide with τ-compacts and F is a τ'-closed subset, then F is τ_s-closed, i.e. satisfies (12). Suppose $\{x_n\} \subset F$ and $x_n \longrightarrow_\tau x_0$. Let $K = \{x_0, x_1, x_2, \ldots\}$. Then K is τ-compact, hence also τ'-compact. In particular, $F \cap K$ is τ'-compact, hence τ-compact, hence sequentially τ-compact, hence $x_0 \in K \cap F \subset F$ and $F \in \tau_s$.

The important corollary to (23) is

Any uniformly τ-tight sequence of random elements in $\mathcal{X}$ is uniformly τ_s-tight. (24)

Facts like (23) and (24) suggest that whenever we deal with uniform tightness (or Prohorov's theorem) sequential spaces satisfying (13) may be of special importance.

To define an "abstract" sequential topology on $\mathcal{X}$ one needs the notion of "convergence" of sequences.

2.1 Basic facts about $\mathcal{L}$- and $\mathcal{L}^*$- convergencies

We say that $\mathcal{X}$ is a space of type $\mathcal{L}$ (Fréchet, [7]), if among all sequences of elements of $\mathcal{X}$ a class $\mathcal{C}(\rightarrow)$ of "convergent" sequences is distinguished, and to each convergent sequence $\{x_n\}_{n\in I\!N}$ exactly one point x_0 (called "limit": $x_n \longrightarrow x_0$) is attached in such a way that

For every $x \in \mathcal{X}$, the constant sequence $(x, x, \ldots)$ is convergent to x. (25)

If $x_n \longrightarrow x_0$ and $1 \leq n_1 < n_2 < \ldots$, then the subsequence $\{x_{n_k}\}$ converges, and to the same limit: $x_{n_k} \longrightarrow x_0$, as $k \to \infty$. (26)

It is easy to see that in the space $\mathcal{X}$ of type $\mathcal{L}$ the statement paralleling (12):

$F \subset \mathcal{X}$ is *closed* if F contains all limits of "$\longrightarrow$"-convergent sequences of elements of F. (27)

defines a topology, $\mathcal{O}(\rightarrow)$ say. This topology defines in turn a new (in general) class of convergent sequences, which can be called convergent "a posteriori" (Urysohn, [20]), in order to distinguish from the original convergence (= convergence "a priori"). So $\{x_n\}$ converges *a posteriori* to x_0, if for every open set $G \in \mathcal{O}(\rightarrow)$ eventually all elements of the sequence $\{x_n\}$ belong to G. Kantorowich *et al* [10, Theorem 2.42, p. 51] and Kisyński [11] proved that this is equivalent to the following condition:

Every subsequence $x_{n_1}, x_{n_2}, \ldots$ of $\{x_n\}$ contains a further subsequence $x_{n_{k_1}}, x_{n_{k_2}}, \ldots$ convergent to x_0 *a priori*. (28)

We see that convergence *a posteriori* shares property (3) with the weak convergence of measures, i.e. satisfies condition

If every subsequence $x_{n_1}, x_{n_2}, \ldots$ of $\{x_n\}$ contains a further subsequence $x_{n_{k_1}}, x_{n_{k_2}}, \ldots$ convergent to x_0, then the whole sequence $\{x_n\}$ is convergent to x_0. (29)

If the $\mathcal{L}$-convergence "$\longrightarrow$" satisfies also (29), then we say that $\mathcal{X}$ is of type $\mathcal{L}^*$ and will denote such convergence by "$\stackrel{*}{\longrightarrow}$". Within this terminology, another immediate consequence of Kantorovich-Kisyński's theorem is that in spaces of type $\mathcal{L}^*$ convergence *a posteriori* coincides with convergence *a priori*.

It follows that given convergence "$\longrightarrow$" satisfying (25) and (26), we can *weaken* this convergence to convergence "$\stackrel{*}{\longrightarrow}$" satisfying additionally (28), and the latter convergence is already the usual convergence of sequences in the topological space $(\mathcal{X}, \mathcal{O}(\rightarrow)) \equiv (\mathcal{X}, \mathcal{O}(\stackrel{*}{\rightarrow}))$. At least two examples of such a procedure are well-known:

Example 2.1 If "$\longrightarrow$" denotes the convergence "almost surely" of real random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, then "$\stackrel{*}{\longrightarrow}$" is the convergence "in probability".

Example 2.2 Let $\mathcal{X} = I\!\!R^1$ and take a sequence $\varepsilon_n \searrow 0$. Say that $x_n \longrightarrow x_0$, if for each $n \in I\!\!N$, $|x_n - x_0| < \varepsilon_n$, i.e. x_n converges to x_0 at given rate $\{\varepsilon_n\}$. Then "$\stackrel{*}{\longrightarrow}$" means usual convergence of real numbers.

The following obvious properties of sequential spaces will be used throughout the paper without annotation:

A set $K \subset \mathcal{X}$ is "$\longrightarrow$"-relatively compact iff it is "$\stackrel{*}{\longrightarrow}$"-relatively compact. (30)

A function f on $\mathcal{X}$ is $\mathcal{O}(\stackrel{*}{\rightarrow})$-continuous iff it is "$\stackrel{*}{\longrightarrow}$"-sequentially continuous (equivalently: "$\longrightarrow$"-sequentially continuous), i.e. $f(x_n)$ converges to $f(x_0)$ whenever $x_n \stackrel{*}{\longrightarrow} x_0$ (or $x_n \longrightarrow x_0$). (31)

Finally, let us notice that if $(\mathcal{X}, \tau)$ is a Hausdorff topological space, then $\tau \subset \tau_s \equiv \mathcal{O}(\rightarrow_\tau)$, and in general this inclusion may be strict. In particular, the space of sequentially continuous functions may be larger than the space of τ-continuous functions.

For more information on sequential spaces we refer to [5] or [1].

3 The sequential topology of the convergence in distribution

The reason we are interested in topological spaces satisfying (13) is Theorem 3 from [8] (restated below) which may be considered both as a strong version of the direct Prohorov's theorem and a generalization of the original Skorohod construction [17].

Theorem 3.1 *Let $(\mathcal{X}, \tau)$ be a topological space satisfying (13) and let $\{\mu_n\}_{n\in\mathbb{N}}$ be a uniformly tight sequence of laws on $\mathcal{X}$. Then there exists a subsequence $n_1 < n_2 < \dots$ and $\mathcal{X}$-valued random elements $Y_0, Y_1, Y_2, \dots$ defined on $([0,1], \mathcal{B}_{[0,1]}, \ell)$ such that*

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \dots, \tag{32}$$

$$Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty, \ \omega \in [0,1]. \tag{33}$$

Let us notice that contrary to the metric case under (13) alone we do not know whether the set of convergence

$$\{\omega : Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty\}$$

is measurable. What we know is measurability of sets of the form

$$C(K) = \{\omega : Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty\} \cap \bigcap_{k=1}^{\infty} \{\omega : Y_k(\omega) \in K\}, \tag{34}$$

where $K \subset \mathcal{X}$ is compact. This becomes obvious when we observe that by property (14) we have

$$C(K) = \{\omega : \tilde{f}(Y_k(\omega)) \to \tilde{f}(Y_0(\omega)), \quad \text{as } k \to \infty\} \cap \bigcap_{k=1}^{\infty} \{\omega : Y_k(\omega) \in K\}.$$

Now suppose for each $\varepsilon > 0$ there is a compact set K_ε such that

$$P(C(K_\varepsilon)) > 1 - \varepsilon. \tag{35}$$

Then the set of convergence contains a measurable set of full probability and one can say that Y_k **converges to Y_0 almost surely "in compacts"**. In particular we have

Corollary 3.2 *Convergence almost surely "in compacts" implies uniform tightness.*

The a.s. convergence (33) has been established exactly the way described above. If the representation $Y_0, Y_1, Y_2, \dots$ satisfies (32) and the convergence (33) is strengthened to the almost sure convergence "in compacts", then we will call it **"the strong a.s. Skorohod representation".** Using this terminology we may rewrite Theorem 3.1 in the following form:

Theorem 3.3 *Let $(\mathcal{X}, \tau)$ be a topological space satisfying (13) and let $\{\mu_n\}_{n\in\mathbb{N}}$ be a uniformly tight sequence of laws on $\mathcal{X}$. Then there exists a subsequence $\mu_{n_1}, \mu_{n_2}, \dots$ which admits the strong a.s. Skorohod representation defined on $([0,1], \mathcal{B}_{[0,1]}, \ell)$.*

We are also ready to give a formal definition of the convergence "$\stackrel{*}{\Longrightarrow}$" introduced in Section 1 for elements of $\mathcal{P}(\mathcal{X})$:

$\mu_n \stackrel{*}{\Longrightarrow} \mu_0$ if every subsequence $\{n_k\}$ contains a further subsequence $\{n_{k_l}\}$ such that $\mu_0, \mu_{n_1}, \mu_{n_2}, \dots$ admit the strong a.s. Skorohod representation defined on the Lebesgue interval. (36)

As an immediate corollary to Theorem 3.3 we obtain the direct Prohorov's theorem for "$\stackrel{*}{\Longrightarrow}$".

Theorem 3.4 *If $(\mathcal{X}, \tau)$ satisfies (13), then in $\mathcal{P}(\mathcal{X})$ relative uniform tightness implies relative compactness with respect to "$\stackrel{*}{\Longrightarrow}$".*

The space $\mathcal{P}(\mathcal{X})$ with the induced convergence "$\stackrel{*}{\Longrightarrow}$" is of $\mathcal{L}^*$ type, i.e. "$\stackrel{*}{\Longrightarrow}$" satisfies (25), (26) and (29). Notice that (25) holds by (18), and that (29) is exactly condition (3) which allows to apply the standard "three-stage procedure" of verifying convergence.

Let us say that the topology $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ is "induced by the strong a.s. Skorohod representation".

By the reasoning similar to the one given before (10), we see that for any *sequentially continuous* and bounded function $f : (\mathcal{X}, \tau_s) \to I\!R^1$, the mapping

$$\mathcal{P}(\mathcal{X}) \ni \mu \mapsto \int_{\mathcal{X}} f(x)\,\mu(dx) \in I\!R^1, \tag{37}$$

is *sequentially* continuous (hence: continuous) with respect to $\mathcal{O}(\stackrel{*}{\Longrightarrow})$. In particular, $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ is finer than the sequential topology given by the usual weak convergence of elements of $\mathcal{P}(\mathcal{X}, \tau_s)$. The standard example shows that in general these two topologies do not coincide. But even if they do, the definition using the strong a.s. Skorohod representation is more operational. Moreover, we have a nice characterization of relative $\stackrel{*}{\Longrightarrow}$-compactness, as announced in Section 1.

Theorem 3.5 *Suppose $(\mathcal{X}, \tau)$ satisfies (13). Then the topology $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ induced by the strong a.s. Skorohod representation is the only sequential topology $\mathcal{O}$ on $\mathcal{P}(\mathcal{X})$ satisfying:*

$\mathcal{O}$ is finer than the topology of weak convergence of measures. (38)

The class of relatively $\mathcal{O}$-compact sets coincides with the class of relatively uniformly τ-tight sets. (39)

Proof. Relation (39) gives us the family of relatively compact subsets and (38) helps us to identify limiting points. This information fully determines an $\mathcal{L}^*$-convergence. □

Remark 3.6 Analysing Fernique's example quoted in Introduction shows that (39) is not valid in the space $\mathcal{P}((H, \tau_w))$ equipped with the topology of weak convergence. It follows the topology $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ may be *strictly* finer than the topology of weak convergence (or weak topology) on $\mathcal{P}(\mathcal{X})$ and the converse Prohorov's theorem holds in many spaces — see section 4.

Remark 3.7 In many respects the topological space $(\mathcal{P}(\mathcal{X}), \mathcal{O}(\stackrel{*}{\Longrightarrow}))$ is as good as $(\mathcal{X}, \tau)$ is: the property (13) is hereditary. To see this, take as the separating functions

$$h_{(i_1,i_2,\ldots,i_m)}(\mu) = \int_{\mathcal{X}} f_{i_1}(x) f_{i_2}(x) \ldots f_{i_m}(x)\, \mu(dx), \tag{40}$$

for all finite sequences $(i_1, i_2, \ldots, i_m)$ of elements of $I\!I$. Hence we may consider within our framework "random distributions" as well.

Theorem 3.5 does not contain the case of an arbitrary metric space, since in nonseparable spaces condition (13) may fail. However we have

Corollary 3.8 *If $\mathcal{X}$ is a metric space, then in $\mathcal{P}(\mathcal{X})$ the weak topology and $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ coincide.*

Proof. Let us observe that in $\mathcal{P}(\mathcal{X})$ the a.s. Skorohod representation for *full* sequences does exist. This is an easy consequence of the fact that each σ-compact metric space can be *homeomorphically* imbedded into a Polish space, and of LeCam's theorem [13], [2]. Following the proof of LeCam's theorem one can also prove that in metric spaces almost sure convergence of random elements with tight laws implies almost sure convergence "in compacts". Hence in $\mathcal{P}(\mathcal{X})$ the sequential topology of weak convergence and $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ coincide. But it is well known [2] that the weak topology on $\mathcal{P}(\mathcal{X})$ is metrisable and so is sequential. □

Remark 3.9 One may prefer the stronger convergence defined by means of the Skorohod representation for the full sequence: $\mu_n \Longrightarrow_{Sk} \mu_0$ if on $([0,1], \mathcal{B}_{[0,1]}, \ell)$ there exists the strong a.s. Skorohod representation $Y_0, Y_1, \ldots$ for $\mu_0, \mu_1, \ldots$. However, by the very definition "$\Longrightarrow_{Sk}$" is only $\mathcal{L}$-convergence and so is not a topological notion, while "$\stackrel{*}{\Longrightarrow}$" is the $\mathcal{L}^*$-convergence obtained from "$\Longrightarrow_{Sk}$" by Kantorovich-Kisyński's recipe (28).

Remark 3.10 The definition of the topology induced by the strong a.s. Skorohod representation may seem to be not the most natural one. But $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ fulfills all possible "portmanteau" theorems (see [19]), coincides with weak convergence on metric spaces and by means of the Prohorov's theorem is operational and easy in handling.

4 Criteria of compactness and the converse Prohorov's theorem

To make the direct Prohorov's theorem working, one needs efficient criteria of checking sequential compactness. It will be seen that given such criteria relative uniform tightness is equivalent to uniform tightness and the converse Prohorov's theorem easily follows.

We begin with spaces $(\mathcal{X}, \tau)$ possessing a fundamental system of compact subsets, i.e. an increasing sequence $\{K_m\}_{m\in I\!N}$ of compact subsets of $\mathcal{X}$ such that every convergent sequence $x_n \longrightarrow_\tau x_0$ is contained in some K_{m_0} (equivalently: every compact subset is contained in some K_{m_0}). Locally compact spaces with countable basis serve here as the most important, but not the only example. For instance, balls $K_m = \{x : \|x\| \leq m\}$ form the fundamental system of compact subsets in a Hilbert space H with either the weak topology τ_w or the sequential topology $(\tau_w)_s$ generated by the weak convergence in H. The same is true in a topological dual E' of a separable Banach space E.

Theorem 4.1 *Suppose that $(\mathcal{X}, \tau)$ satisfies (13) and possesses a fundamental system $\{K_m\}$ of compact subsets. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following statements are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{41}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{42}$$

Proof. In view of Theorem 3.4 we have to prove that (41) implies (42). Suppose (42) does not hold. Then there is $\varepsilon > 0$ such that for each m one can find $\mu_m \in \mathcal{K}$ satisfying

$$\mu_m(K_m^c) > \varepsilon. \tag{43}$$

By $\stackrel{*}{\Longrightarrow}$-relative compactness there exists a subsequence μ_{m_k} admitting a strong a.s. Skorohod representation. By Corollary 3.2 $\{\mu_{m_k}\}_{k\in I\!N}$ is uniformly tight. This contradicts (43). □

As the next step we will consider a more general scheme in which compactness means boundedness with respect to some countable family of lower semicontinuous functionals. More precisely, we suppose that there exists a countable family of measurable nonnegative functionals $\{h_k\}_{k\in I\!K}$ such that

$$\sup_{x\in K} h_k(x) < +\infty,\ k \in I\!K, \tag{44}$$

implies relative compactness of K, and if $x_n \longrightarrow_\tau x_0$ then

$$h_k(x_0) \leq \liminf_{n\to\infty} h_k(x_n) < +\infty,\ k \in I\!K. \tag{45}$$

Notice that under (45) any relatively compact set K satisfies (44) and is contained in some set of the form

$$K = \bigcap_{k\in I\!\!K} \{x : h_k(x) \leq C_k\}. \tag{46}$$

Moreover, under both (44) and (45) every set of the form (46) is *sequentially compact.*

Theorem 4.2 *Let $(\mathcal{X}, \tau)$ satisfies (13). Suppose compactness in $(\mathcal{X}, \tau)$ is given by boundedness with respect to a countable family $\{h_k\}_{k\in I\!\!K}$ of lower semi-continuous functionals. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{47}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{48}$$

For each $k \in I\!\!K$ the set $\{\mu \circ h_k^{-1} : \mu \in \mathcal{K}\} \subset \mathcal{P}(I\!\!R^+)$ is uniformly tight, i.e.

$$\lim_{C\to\infty} \sup_{\mu\in\mathcal{K}} \mu(\{x : h_k(x) > C\}) = 0. \tag{49}$$

Proof. Conditions (48) and (49) are obviously equivalent and implication (48) $\Rightarrow$ (47) is proved in Theorem 3.4. In order to prove that (47) implies (49) suppose that for some $k \in I\!\!K$ there is $\varepsilon > 0$ such that for each N one can find $\mu_N \in \mathcal{K}$ with the property

$$\mu_N(\{x : h_k(x) > N\}) \geq \varepsilon, \quad N \in I\!\!N. \tag{50}$$

If *some* subsequence of μ_N admits a strong a.s. Skorohod representation, it must be uniformly tight and (50) cannot hold along this subsequence. This shows that $\mathcal{K}$ is not $\stackrel{*}{\Longrightarrow}$-relatively compact. □

It is worth to emphasize that Theorem 4.2 *completely* generalizes the ordinary converse Prohorov's theorem. To see this, take Polish space $(\mathcal{X}, \rho)$ and choose in it a countable dense subset $D = \{x_1, x_2, \ldots\}$. Set for $k \in I\!\!N$

$$h_k(x) = \inf\{N : x \in \overline{\bigcup_{i=1}^N K_\rho(x_i, 1/k)}.$$

Then every functional h_k is bounded on $K \subset \mathcal{X}$ if, and only if, K is totally ρ-bounded, hence conditionally compact by completeness of $(\mathcal{X}, \rho)$. The property (45) follows by the very definition of h_k.

Topologically complete spaces and non-metrisable σ-compact spaces like (H, τ_w) does not end the list of cases covered by Theorem 4.2. For example on the Skorohod space $I\!\!D([0,1] : I\!\!R^1)$ there exists (see [9]) a minimal functional topology which is non-metrisable but satisfies (44) and (45), hence by our Theorem 4.2 is as good as Polish space (at least from the probabilistic point of view). In fact, the present paper may be considered as an attempt to find a general framework in which that topology can be placed naturally.

"Countable boundedness" is not a universal criterion for compactness. In general we do not know any criterion which could pretend to universality. Therefore any particular case must be carefully analysed. We will show three examples of such an analysis.

The first type of results has been suggested by topologies on function spaces in which conditional compactness can be described in terms of "moduli of continuity". A rough generalization is that on a topological space $(\mathcal{X}, \tau)$ a double array $\{g_{k,j}\}_{k\in I\!K, j\in I\!N}$ (where $I\!K$ is countable) of nonnegative measurable functionals is given and that the functionals possess the following properties:

$$g_{k,j+1} \leq g_{k,j}, \quad k \in I\!K, j \in I\!N. \tag{51}$$

If $x_n \longrightarrow_\tau x_0$ then for each $k \in I\!K$

$$\lim_{j\to\infty} \sup_n g_{k,j}(x_n) = 0. \tag{52}$$

If for each $k \in I\!K$

$$\lim_{j\to\infty} \sup_{x\in K} g_{k,j}(x) = 0, \tag{53}$$

then $K \subset \mathcal{X}$ is *conditionally* compact.

Clearly, the new scheme contains the previous one. If we set

$$g_{k,j}(x) = \frac{1}{j} h_k(x), \quad k \in I\!K, j \in I\!N,$$

then (45) implies (52) and (44) and lower semicontinuity of h_k give conditional compactness in (53). Recall that in general in spaces satisfying (13) relative compactness does not imply conditional compactness. In metric spaces, however, it does and so e.g. Skorohod topology J_2 [17] (and not only J_1) satisfies the converse Prohorov theorem, as we can see from the following result.

Theorem 4.3 *Let $(\mathcal{X}, \tau)$ satisfies (13). Suppose conditions (51) – (53) determine conditional compactness in $(\mathcal{X}, \tau)$. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$\mathcal{K}$ is $\stackrel{*}{\Longrightarrow}$-relatively compact. (54)

$\mathcal{K}$ is uniformly τ-tight. (55)

For each $k \in I\!K$

$$\lim_{j\to\infty} \sup_{\mu\in\mathcal{K}} \mu(\{x : g_{k,j}(x) > \varepsilon\}) = 0, \quad \varepsilon > 0. \tag{56}$$

PROOF. Similarly as before, it is enough to show that if (56) is not satisfied then one can find in $\mathcal{K}$ a sequence with no subsequence admitting a strong a.s. Skorohod representation. Let us observe first that if $X_l \longrightarrow_\tau X_0$ a.s. and $j_l \to \infty$ then by (51) and (52), for each $k \in I\!\!K$ and almost surely,

$$\limsup_{l\to\infty} g_{k,j_l}(X_l) \leq \lim_{j\to\infty} \limsup_{l\to\infty} g_{k,j}(X_l) = 0. \tag{57}$$

If (56) is not satisfied, then there are $k \in I\!\!K$ and $\varepsilon > 0$ such that for each $j \in I\!\!N$ one can find $\mu_j \in \mathcal{K}$ satisfying

$$\mu_j(\{x : g_{k,j}(x) > \varepsilon\}) \geq \varepsilon. \tag{58}$$

If X_l is the a.s. Skorohod representation for some subsequence μ_{j_l} then by (57)

$$\mu_{j_l}(\{x : g_{k,j_l}(x) > \varepsilon\}) \to 0,$$

hence (58) cannot hold. □

The second type of results is motivated by the structure of compact subsets in the space of distributions $\mathcal{S}'$ or, more generally, the topological dual of a Fréchet nuclear space.

Suppose that on $(\mathcal{X}, \tau)$ there exists a decreasing sequence $\{q_m\}_{m\in I\!\!N}$ of non-negative measurable functionals such that

$K \subset \mathcal{X}$ is *conditionally compact* if for some $m_0 \in I\!\!N$

$$\sup_{x\in K} q_{m_0}(x) \leq C_{m_0} < +\infty. \tag{59}$$

Notice this implies

$$\sup_{m\geq m_0} \sup_{x\in K} q_m(x) \leq C_{m_0},$$

but it may happen that for some $m < m_0$

$$\sup_{x\in K} q_m(x) = +\infty.$$

Theorem 4.4 *Let $(\mathcal{X}, \tau)$ satisfies (13) and (59). Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{60}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{61}$$

For each $\varepsilon > 0$ one can find $m_0 \in I\!\!N$ and $C > 0$ such that

$$\sup_{\mu\in\mathcal{K}} \mu(\{x : q_{m_0}(x) > C\}) < \varepsilon. \tag{62}$$

Proof. We apply the standard strategy. If (62) is not satisfied, then there is $\varepsilon > 0$ such that for every M and for some $\mu_M \in \mathcal{K}$

$$\mu_M(\{x : q_M(x) > M\}) \geq \varepsilon. \tag{63}$$

If $\{X_k\}$ is the strong a.s. Skorohod representation for some subsequence μ_{M_k}, then it is tight (by Corollary 3.2). and so for some m_0 and C

$$P(q_{m_0}(X_k) \leq C) = \mu_{M_k}(\{x : q_{m_0}(x) \leq C\}) > 1 - \varepsilon, \quad k = 1, 2, \ldots. \tag{64}$$

Hence for k satisfying $M_k > C$ and $M_k > m_0$ we get from (63) and (64)

$$\begin{aligned} 1 - \varepsilon &\geq \mu_{M_k}(\{x : q_{M_k}(x) \leq M_k\}) \\ &\geq \mu_{M_k}(\{x : q_{M_k}(x) \leq C\}) \\ &\geq \mu_{M_k}(\{x : q_{m_0}(x) \leq C\}) > 1 - \varepsilon, \end{aligned}$$

what is a contradiction. □

Usually results valid for $\mathcal{S}'$ hold also for space $\mathcal{D}'$, despite its more complicated structure. The reason is that $\mathcal{D}'$ can be identified with a closed subset of a countable product of duals to Fréchet nuclear spaces and that the properties under consideration are preserved when passing to closed subspaces and countable products. This is exactly the case with our "Prohorov spaces". Recall that $(\mathcal{X}, \tau)$ is "Prohorov space" if every conditionally compact subset $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ (with $\mathcal{P}(\mathcal{X})$ equipped with the weak topology) is uniformly τ-tight (see [15]). Since we know that $\mathcal{O}(\overset{*}{\Rightarrow})$ may be strictly finer than the weak topology, the corresponding notion for $(\mathcal{P}(\mathcal{X}), \mathcal{O}(\overset{*}{\Rightarrow}))$ may be different. Therefore we say that $(\mathcal{X}, \tau)$ is **an S-P space**, if every $\overset{*}{\Rightarrow}$-relatively compact subset of $\mathcal{P}(\mathcal{X})$ is uniformly τ-tight.

The present section contains several standard examples of S-P spaces. We conclude the paper with formal statement of some properties of S-P spaces.

Theorem 4.5 *Let $(\mathcal{X}, \tau)$ be an S-P space satisfying (13). If $C \subset \mathcal{X}$ is either closed or G_δ, then $(C, \tau|_C)$ is again S-P space.*

Proof. The only nontrivial part is proving that if G is open and $\mathcal{K} \subset \mathcal{P}(G)$ is $\overset{*}{\Rightarrow}$-relatively compact (in $\mathcal{P}(G)$!), then $\mathcal{K}$ is uniformly $\tau|_G$-tight. Since relative compactness in $\mathcal{P}(G)$ means also relative compactness in $\mathcal{P}(\mathcal{X})$, by the S-P property we get uniform τ-tightness of $\mathcal{K}$. By (21) the closure $\overline{\mathcal{K}}$ in $\mathcal{P}(\mathcal{X})$ (which consists of limiting points of $\mathcal{K}$) is uniformly τ-tight and so sequentially compact, both in $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(G)$ (the latter by relative compactness in $\mathcal{P}(G)$). Since in our case sequential compactness is equivalent to compactness, it is now possible to repeat step by step the reasoning given in the proof of Theorem 1, [15], pp. 109-110. □

Corollary 4.6 *Any S-P space satisfying (13) has the property that the closure of a relatively compact set is compact and consists of the set itself and its limiting points.* □

Theorem 4.7 *Let* $(\mathcal{X}_n, \tau_n),\ n = 1, 2, \ldots$ *be S-P spaces satisfying (13). Then the product space* $\prod_{n=1}^{\infty}(X_n, \tau_n)$ *is an S-P space.* □

Acknowledgement. The author would like to thank Professor Kisyński for information on the independent source [10] for Kisyński's theorem.

References

[1] Arkhangel'skiĭ, A.V. and Pontryagin, L.S., (Eds.), **General Topology I**, *Encyklopaedia of Mathematical Sciences*, **17** Springer, Berlin 1990.

[2] Billingsley, P., **Convergence of Probability Measures**, Wiley, New York 1968.

[3] Billingsley, P., **Weak Convergence of Measures: Applications in Probability**, SIAM, Philadelphia 1971.

[4] Davies, R.O., A non-Prohorov space, *Bull. London Math. Soc.* **3** (1971) 341–342.

[5] Engelking, R., **General Topology**, Helderman, Berlin 1989.

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

[7] Fréchet, M., Sur la notion de voisinage dans les ensembles abstraits, *Bull. Sci. Math.* **42** (1918) 138-156.

[8] Jakubowski, A., The a.s. Skorohod representation for subsequences in non-metric spaces, (1994), submitted.

[9] Jakubowski, A., A non-Skorohod topology on the Skorohod space, (1994), in preparation.

[10] Kantorowich, L.V., Vulih, B.Z. & Pinsker, A.G., **Functional Analysis on Semiordered Spaces**, GIT-TL, Moscow 1950 (in Russian).

[11] Kisyński, J., Convergence du type L, *Colloq. Math.* **7** (1960) 205–211.

[12] Kuratowski, K., **Topology, Vol. I**, Academic Press, New York 1966.

[13] Le Cam, L., Convergence in distribution of stochastic processes, *Univ. California Publ. Statist.* **2** no. 11 (1957) 207–236.

[14] Parthasarathy, K.R., **Probability Measures on Metric Spaces**, Academic Press, New York 1967.

[15] Preiss, D., Metric spaces in which Prohorov's theorem is not valid, *Z. Wahrscheinlichkeitstheorie* **27** (1973) 109–116.

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

[17] Skorohod, A.V., Limit theorems for stochastic processes, *Theor. Probability Appl.* **1** (1956) 261–290.

[18] Smolyanov, O. & Fomin, S.V., Measures on linear topological spaces, *Russ. Math. Surveys* **31** (1976) 1–53.

[19] Topsøe, F., **Topology and Measure**, *Lecture Notes in Math.* **133**, Springer, Berlin 1970.

[20] Urysohn, P., Sur les classes L de M. Fréchet, *Enseign. Math.* **25** (1926) 77–83.

Author's address:

Adam Jakubowski
Nicholas Copernicus University
Faculty of Mathematics and Informatics
ul. Chopina 12/18
87-100 Toruń, Poland

E-mail: adjakubo@mat.uni.torun.pl