Sequential Topologies in Probability Theory

Four papers on functional convergence

by

Adam Jakubowski

Preprint nr 1/95

Toruń, 1995

Contents

1 The a.s. Skorohod representation for subsequences in nonmetric spaces **1**
1.1 The a.s. Skorohod representation 1
1.2 Topological assumption and main theorem 3
1.3 Some examples 5

2 A unification of Prohorov's and Skorohod's ideas: convergence in distribution in nonmetric spaces **7**
2.1 Convergence in distribution of random elements 7
2.2 Topological preliminaries 10
2.2.1 Basic facts about L- and L^*- convergencies 12
2.3 The sequential topology of the convergence in distribution 13
2.4 Criteria of compactness and the converse Prohorov's theorem 16

3 A non-Skorohod topology on the Skorohod space **21**
3.1 Introduction 21
3.2 The topology S 24
3.3 Convergence in distribution on $(\mathbb{D}, S)$ 31
3.4 Uniform S-tightness and semimartingales 36

4 Convergence in various topologies for stochastic integrals driven by semimartingales **39**
4.1 Introduction 39
4.2 The results 41
4.3 Proofs 46
4.3.1 Basic lemma 46
4.3.2 Proof of Theorem 4.2.1 47
4.3.3 Proof of Theorems 4.2.4 –4.2.9 48
4.3.4 Proof of Theorem 4.2.10 49
4.3.5 Proof of Theorem 4.2.12 49

Bibliography **51**

Paper No 1.

The a.s. Skorohod representation for subsequences in nonmetric spaces

Abstract

It is shown that in a large class of topological spaces every uniformly tight sequence of random elements contains a subsequence which admits the usual a.s. Skorohod representation on the Lebesgue interval.

1.1 The a.s. Skorohod representation

Let $(\mathcal{X}, \rho)$ be a Polish space and let $X_1, X_2, \ldots$ be random elements taking values in $\mathcal{X}$ and converging in distribution to X_0:

$$X_n \underset{\mathcal{D}}{\longrightarrow} X_0. \tag{1.1}$$

In his famous paper [29], Skorohod proved that there exist $\mathcal{X}$-valued random elements $Y_0, Y_1, Y_2, \ldots$, defined on the unit interval $([0,1], \mathcal{B}_{[0,1]})$ equipped with the Lebesgue measure ℓ, such that

$$\text{the laws of } X_n \text{ and } Y_n \text{ coincide for } n = 0, 1, 2, \ldots, \tag{1.2}$$

$$\rho(Y_n(\omega), Y_0(\omega)) \longrightarrow 0, \quad \text{as } n \to +\infty, \text{ for each } \omega \in [0,1]. \tag{1.3}$$

Later, Dudley [7] extended the Skorohod representation to separable metric spaces, and Wichura [38] and Fernandez [10] proved its existence in nonseparable metric spaces, provided the limit has separable range (see also [8]). The price to be paid was larger space required by the definition of the representation.

It may be worth to emphasize that if we restrict our attention to convergence in distribution of random elements with *tight* (or *Radon*) distributions then even in arbitrary metric spaces the a.s. Skorohod representation exists in its original shape (on $[0,1]$). This is an easy consequence of the fact that each σ-compact metric space can be *homeomorphically* imbedded into a Polish space, and of Le Cam's theorem [20] asserting that in metric spaces any sequence $\{\mu_n\}$ of *tight* probability measures weakly convergent to a *tight* measure μ_0 is *uniformly tight*, i.e. for every $\varepsilon > 0$ there exists a compact subset K_ε such that

$$\mu_n(K_\varepsilon) > 1 - \varepsilon, \quad n = 1, 2, \ldots. \tag{1.4}$$

When we leave the safe area of metrisable spaces no result on the a.s. Skorohod representation seems to be known.

Let us consider, for example, the weak topology $\tau_w = \sigma(H, H)$ on the infinite dimensional separable Hilbert space $(H, <, >)$. Suppose that $X_n,\ n = 0, 1, \ldots$ take values in $(H, \mathcal{B}_{\tau_w})$ and that $X_n \quad \longrightarrow_{\mathcal{D}} X_0$ in this space, i.e.

$$Ef(X_n) \longrightarrow Ef(X_0), \quad \text{as } n \to +\infty, \tag{1.5}$$

for each bounded and *weakly* continuous function $f : H \to I\!R^1$.

We cannot find the a.s. Skorohod representation for X_n. Suppose, however, that while checking (1.5) we applied the classical procedure based on the direct Prohorov theorem. This means we were able to prove that for each $\varepsilon > 0$ there is a number $K_\varepsilon > 0$ such that

$$P(\|X_n\| > K_\varepsilon) < \varepsilon, \quad n = 1, 2, \ldots, \tag{1.6}$$

(uniform τ_w-tightness) and then we identified the limiting distribution, via e.g.

$$< y, X_n > \underset{\mathcal{D}}{\longrightarrow} < y, X_0 >, \quad , \quad \text{as } n \to +\infty,\ y \in H. \tag{1.7}$$

Consider the following theorem, which is a particular case of a much more general result proved in Section 1.2.

Theorem 1.1.1 *Let $X_1, X_2, \ldots$ be uniformly τ_w-tight, i.e. satisfy* (1.6).

Then there is a subsequence $\{n_k\}$ and H-valued random variables $Y_0, Y_1, \ldots$ defined on $([0,1], \mathcal{B}_{[0,1]}, \ell)$ such that

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \ldots, \tag{1.8}$$

$$< y, Y_k(\omega) > \longrightarrow < y, Y_0(\omega) >, \quad \text{as } k \to \infty,\ \omega \in [0,1],\ y \in H. \tag{1.9}$$

By the above theorem, if (1.5) and (1.6) hold, then *in every subsequence $\{X_{n_k}\}_{k \in I\!N}$ one can find a further subsequence $\{X_{n_{k_l}}\}_{l \in I\!N}$, for which the usual a.s. Skorohod representation on the Lebesgue interval exists.* Let us say that $\{X_n\}_{n \in I\!N}$ possesses the a.s. Skorohod representation *for subsequences.*

Notice that in practice the a.s. representation for subsequences is equally useful as the "full" representation. Typically one needs the Skorohod representation to prove convergence in distribution of some functionals of the underlying processes (see [4] for standard examples). In the simplest case the functional is a measurable mapping, g say, which is a.s. continuous with respect to the limiting law $\mathcal{L}(X_0)$. But it follows from the very definition of the weak convergence of probability laws that $g(X_n) \quad \longrightarrow_{\mathcal{D}} g(X_0)$ iff in every subsequence $\{g(X_{n_k})\}_{k \in I\!N}$ one can find a further subsequence $\{g(X_{n_{k_l}})\}_{l \in I\!N}$ converging in law to $g(X_0)$. Hence it is clear that the a.s. Skorohod representation for subsequences is just what we need.

On the other hand, Fernique [11, p.24-25] gives an example of an H-valued τ_w-weakly convergent (to 0) sequence of random elements with no subsequence being uniformly τ_w-tight. For such sequences our Theorem 1.1.1 cannot be applied. Nevertheless results like Theorem 1.1.1 work perfectly in cases when weak convergence does imply uniform tightness (e.g. in spaces of distributions - see Section 1.3) and even in the general case Theorem 1.1.1 may be applied every time we get weak convergence indirectly, i.e. first checking relative compactness (via uniform tightness and the direct Prohorov's theorem) and then identifying limits.

We aim at proving a general result on the existence of the a.s. Skorohod representation for subsequences, which covers most interesting cases, with emphasis on nonmetric spaces.

1.2 Topological assumption and main theorem

Let $(\mathcal{X}, \tau)$ be a topological space. Denote by " $\longrightarrow_\tau$" convergence of sequences in topology τ. The only assumption we impose on $(\mathcal{X}, \tau)$ is quite simple:

There exists a countable family $\{f_i : \mathcal{X} \to [-1,1]\}_{i\in I\!I}$ of τ-continuous functions, which separate points of $\mathcal{X}$. (1.10)

This condition is not very restrictive and possesses several nice consequences, which we list below together with some comments.

$\mathcal{X}$ is a Hausdorff space (but need not be regular). (1.11)

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 converges to some x_0 and $f_i(x_0) = \alpha_i,\ i \in I\!I$. (1.12)

$K \subset \mathcal{X}$ is compact iff it is sequentially compact (and then it is metrisable). (1.13)

The closure of a relatively compact subset consists of limits of its convergent subsequences (but still need not be compact). (1.14)

Therefore in the definition of uniform τ-tightness we cannot, in general, replace sequential compactness with measurability and relative compactness.

Finally, notice that in many cases $\sigma(f_i : i \in I\!I)$ is just the Borel σ-algebra. In any case compact sets are $\sigma(f_i : i \in I\!I)$-measurable and so 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 after P-completion of $\mathcal{F}$.

By the last property, we will restrict our attention to random elements X such that $f_i(X),\ i \in I\!I$, are random variables and the law of X is tight.

Theorem 1.2.1 *Let $(\mathcal{X}, \tau)$ be a sequential space satisfying (1.10) and let $X_1, X_2, \ldots$ be $\mathcal{X}$-valued random elements. Suppose for each $\varepsilon > 0$ there exists a* compact *subset $K_\varepsilon \subset \mathcal{X}$ such that*

$$P(X_n \in K_\varepsilon) > 1 - \varepsilon, \quad n = 1, 2, \ldots. \tag{1.15}$$

Then one can find a subsequence $\{X_{n_k}\}_{k\in I\!N}$ and $\mathcal{X}$-valued random elements $Y_0, Y_1, Y_2, \ldots$ defined on $([0,1], \mathcal{B}_{[0,1]}, \ell)$ such that

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \ldots, \tag{1.16}$$

$$Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty,\ \omega \in [0,1]. \tag{1.17}$$

Proof. Assumption (1.10) gives us the mapping

$$\mathcal{X} \ni x \mapsto \tilde{f}(x) = (f_i(x))_{i\in I\!I} \in [-1,1]^{I\!I}, \tag{1.18}$$

which is *one-to-one* and *continuous*, but (in general) is not a homeomorphism of $\mathcal{X}$ onto a subspace of $I\!R^{I\!I}$. 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 $\mathcal{X}$.

Let compact sets $K_m \subset \mathcal{X}$ be such that $K_m \subset K_{m+1}$, $m = 1, 2, \ldots$ and

$$P(X_n \in K_m) > 1 - 1/m, \quad n = 1, 2, \ldots. \tag{1.19}$$

Let $\tilde{\mu}_n = \mathcal{L}(\tilde{f}(X_n))$ and $\tilde{K}_m = \tilde{f}(K_m)$. Define on $I\!R^{I\!I}$ an integer-valued functional

$$\Phi(y) := \begin{cases} \min\{m : y \in \tilde{K}_m\} & \text{if } y \in \bigcup_{m=1}^{\infty} \tilde{K}_m \\ +\infty & \text{otherwise.} \end{cases} \tag{1.20}$$

Clearly, Φ is lower semicontinuous, i.e.

$$\liminf_{n\to\infty} \Phi(y_n) \geq \Phi(y_0), \tag{1.21}$$

whenever y_n converges in $I\!R^{I\!I}$ to y_0. Further, it follows from (1.19) that $\Phi < +\infty$ $\tilde{\mu}_n$-a.s., for each $n \in I\!N$, and that $\{\tilde{\mu}_n \circ \Phi^{-1}\}$ is a tight sequence of laws on $I\!N$. By the classical direct Prohorov's theorem we may extract a subsequence $\{n_k\}_{k\in I\!N}$ such that on the space $I\!R^{I\!I} \times I\!N$,

$$\tilde{\mu}_{n_k} \circ \Psi^{-1} \Longrightarrow \nu_0, \quad \text{as } k \to \infty,$$

where $\Psi(y) = (y, \Phi(y))$.

We need a slight refinement of the original Skorohod construction [29, Lemma 3.1.1].

Lemma 1.2.2 *Let $\mathcal{S}$ and $\mathcal{S}'$ be Polish spaces, and let $\Phi : \mathcal{S} \to \mathcal{S}'$ be measurable. Suppose*

$$(X_n, \Phi(X_n)) \underset{\mathcal{D}}{\longrightarrow} (X_0, Y_0) \quad \text{on } \mathcal{S} \times \mathcal{S}'. \tag{1.22}$$

Then there exist random elements $X_0', X_1', X_2', \ldots$ (in $\mathcal{S}$) and Y_0' (in $\mathcal{S}'$) defined on the standard probability space $([0,1], \mathcal{B}_{[0,1]}, \ell)$ and such that

$$\mathcal{L}(X_0', Y_0') = \mathcal{L}(X_0, Y_0); \tag{1.23}$$

$$\mathcal{L}(X_n') = \mathcal{L}(X_n), \quad n = 1, 2, \ldots; \tag{1.24}$$

$$(X_n'(\omega), \Phi(X_n'(\omega))) \longrightarrow (X_0'(\omega), Y_0'(\omega)) \quad \text{in } \mathcal{S} \times \mathcal{S}', \tag{1.25}$$

for ℓ-almost all $\omega \in [0,1]$.

Proof of the Lemma. This is an immediate consequence of the fact that for random elements with values in a separable metric space, the equality $(X, Y) \sim (X, \Phi(X))$ implies $Y = \Phi(X)$ a.s. (for details and the proof of a fact similar to Lemma 1.2.2 we refer to [34]). $\Box$

Proof of Theorem 1.2.1 (continued). By Lemma 1.2.2 we find an $I\!R^{I\!I}$-valued representation X_k' such that

$$\Big(X_k'(\omega), \Phi(X_k'(\omega))\Big) \longrightarrow \Big(X_0'(\omega), Y_0'(\omega)\Big) \ \ \ell\text{-a.s., as } k \to \infty, \tag{1.26}$$

and

$$\mathcal{L}(X_k') = \mathcal{L}(\tilde{f}(X_{n_k})), \quad k = 1, 2, \ldots. \tag{1.27}$$

Since $Y_0'(\omega) < +\infty$ ℓ-a.s., we have also

$$\sup_k \Phi(X_k'(\omega)) < +\infty, \quad \ell\text{-a.s.} \tag{1.28}$$

This implies that for ℓ-almost all ω, points $X'_k(\omega),\ k = 1, 2, \ldots$ remain inside a compact set $\tilde{K}_{m(\omega)} = \tilde{f}(K_{m(\omega)})$. Hence also $X'_0(\omega) \in \tilde{K}_{m(\omega)}$, and moreover,

$$\tilde{f}^{-1}(X'_k(\omega)) \underset{\tau}{\longrightarrow} \tilde{f}^{-1}(X'_0(\omega)).$$

Redefining (if necessary) X'_k on set $\bigcup_{k=0}^{\infty}(X'_k)^{-1}\Big(\bigcap_{m=1}^{\infty}\tilde{K}_m^c\Big)$ of ℓ-measure 0, we obtain the desired Skorohod representation for subsequence X_{n_k} in the form $Y_k = \tilde{f}^{-1}(X'_k),\ k = 0, 1, 2, \ldots$.

Notice that the distribution of Y_0 is tight: since Φ is lower semicontinuous we have

$$\Phi(X'_0(\omega)) \le Y'_0 \text{ a.s.},$$

and so

$$\begin{aligned} P(Y_0 \notin K_m) &= P(\tilde{f}^{-1}(X'_0) \notin K_m) \\ &= P(\Phi(X'_0) > m) \\ &\le P(Y'_0 > m) \to 0 \quad \text{as } m \to +\infty. \ \Box \end{aligned}$$

1.3 Some examples

Clearly, Theorem 1.1.1 is an example of application of Theorem 1.2.1. One can go further in this direction.

Theorem 1.3.1 *Let E be a separable Banach space and let E' be its topological dual. If $X_1, X_2, \ldots$ are E'-valued random elements such that the sequence $\{\|X_n\|\}$ of real random variables is bounded in probability, then along some subsequence $\{n_k\}$ there exists the Skorohod representation $Y_k,\ k = 0, 1, 2, \ldots$ such that*

$$< x, Y_k(\omega) > \longrightarrow < x, Y_0(\omega) >, \quad \text{as } k \to \infty,\ x \in E,\ \omega \in [0, 1]. \tag{1.29}$$

$\Box$

Somewhat different results arise when we consider $\mathcal{S}'$-valued (or $\mathcal{D}'$-valued) random elements or, more generally, random elements with values in the topological dual to a Frechét nuclear space (or to the strict inductive limit of a Frechét nuclear spaces).

For the sake of brevity we will formulate here results for the simpler case only. Let Φ be a Frechét nuclear space (see e.g. [28]). Let $\|\cdot\|_1 \le \|\cdot\|_2 \le \ldots$ be an increasing sequence of Hilbertian seminorms defining the topology on Φ. Denote by $(\Phi_p, \|\cdot\|_p)$ the Hilbert space arising by completion of the quotient space $\Phi/\|\cdot\|_p$ and by $(\Phi'_{-p}, \|\cdot\|_{-p})$ the topological dual of $(\Phi_p, \|\cdot\|_p)$. After obvious identification, Φ'_{-p} is a subset of Φ' and $\Phi' = \bigcup_{p=1}^{\infty}\Phi'_{-p}$. Φ' is equipped with the strong topology β, which on every Φ_{-p} is *strictly* weaker than the Hilbert topology of the norm $\|\cdot\|_{-p}$. The point is that the convergence of sequences in topology β may be defined in the following way:

$$x_n \underset{\beta}{\longrightarrow} x_0 \text{ iff } \|x_n - x_0\|_{-p} \to 0, \quad \text{as } n \to +\infty,\ \text{for some } p \in I\!N. \tag{1.30}$$

Theorem 1.3.2 *Let Φ' be the topological dual of a Frechét nuclear space Φ and let $X_1, X_2, \ldots$ be random elements with values in Φ'. Suppose that for every $\phi \in \Phi$ random variables $< \phi, X_n >$, $n = 1, 2, \ldots$ are uniformly tight. Then there exists a subsequence n_k and the Skorohod representation $Y_0, Y_1, Y_2, \ldots$ for this subsequence such that for each $\omega \in [0,1]$ one can find a number $p(\omega) \in \mathbb{N}$ with the property that*

$$\|Y_k(\omega) - Y_0(\omega)\|_{-p(\omega)} \longrightarrow 0, \quad \text{as } k \to \infty. \tag{1.31}$$

Proof. Standard arguments of the Minlos-Sazonov-type (see e.g. [22] or [11]) show that in Φ' "weak uniform tightness" implies usual uniform tightness: for each $\varepsilon > 0$ there are numbers $q_\varepsilon \in \mathbb{N}$ and $K_\varepsilon > 0$ such that

$$P(\|X_n\|_{-q_\varepsilon} \leq K_\varepsilon) > 1 - \varepsilon. \tag{1.32}$$

□

Corollary 1.3.3 *Any sequence convergent in distribution on $\mathcal{S}'$ or $\mathcal{D}'$ admits the a.s. Skorohod representation for subsequences.*

Proof. It is proved in [11] that on $\mathcal{S}'$ or $\mathcal{D}'$ relative compactness (in distribution) is equivalent to uniform tightness. □

Notice that our Theorem 1.2.1 may be viewed as the strong version of the direct Prohorov's theorem. Indeed, if $f : \mathcal{X} \to \mathbb{R}^1$ is bounded and continuous and $Y_0, Y_1, Y_2, \ldots$ form the Skorohod representation for $\{X_{n_k}\}$, then

$$Ef(X_{n_k}) = Ef(Y_k) \longrightarrow Ef(Y_0), \quad \text{as } n \to +\infty, \tag{1.33}$$

and so $\mathcal{L}(X_{n_k})$ weakly converges to $\mathcal{L}(Y_0)$ in the classical sense. But (1.33) holds also for all *sequentially* continuous and bounded f! It means that in the nonmetric case the direct Prohorov's theorem may give relative compactness in the stronger topology than the original one. Similar observation can be found in [11] where it was proved that convergence in distribution on $\mathcal{D}'$ equipped with the weak topology coincides with convergence in distribution with respect to the strong topology. This is not surprising in view of the fact that convergence *of sequences* in the weak topology on $\mathcal{D}'$ (and $\mathcal{S}'$) implies convergence in the strong topology.

The above remarks may also suggest that identifying convergence in distribution with weak convergence of laws is not completely justified for some quite good spaces. We refer to Paper II for further discussion on this topic.

Finally let us mention that one of the main motivations to prove Theorem 1.2.1 was to deal with "really" sequential topology on the Skorohod space $\mathbb{D}$. The reader may find information on this non-Skorohod and nonmetric topology in Paper III.

Acknowledgment. The author would like to thank T. Bojdecki, B. Gołdys and S. Kwapień for valuable discussions, which influenced the paper in various ways.

Paper No 2.

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

Abstract

A new topology has been defined on 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.

2.1 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

$$Ef(X_n) \longrightarrow Ef(X_0), \quad \text{as } n \to +\infty, \tag{2.1}$$

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 $P_{X_n} = P \circ X_n^{-1}$ 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 (2.1), where $\mu_n \Longrightarrow \mu_0$ means

$$\int_{\mathcal{X}} f(x)\, \mu_n(dx) \longrightarrow \int_{\mathcal{X}} f(x)\, \mu_0(dx), \quad f \in CB(\mathcal{X}). \tag{2.2}$$

The most successful step towards the abstract setting was done by Prohorov in his fundamental paper [26], and the complete theory when $\mathcal{X}$ is a Polish space has been given in excellent books by Parthasarathy [24] and Billingsley [3]. 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 (2.2) is in force):

$$\text{If every subsequence } \{\mu_{n_k}\} \text{ contains a further subsequence } \{\mu_{n_{k_l}}\} \text{ weakly convergent to } \mu_0, \text{ then the whole sequence } \{\mu_n\} \text{ converges weakly to } \mu_0. \tag{2.3}$$

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

$$\mu_i(K_\varepsilon) > 1 - \varepsilon, \quad i \in \mathbb{I}. \tag{2.4}$$

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 [5], with rational numbers $\mathbb{Q}$ being the most striking example [25]. Let us notice that every probability measure on $(\mathbb{Q}, \mathcal{B}_{\mathbb{Q}})$ must be tight, and so, by LeCam's theorem ([20], [3]) 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 séparé" in [11]). But Fernique [11] gives an example of an H-valued sequence $\{X_n\}$ satisfying

$$Ef(X_n) \longrightarrow f(0), \quad \text{as } n \to +\infty, \tag{2.5}$$

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

$$\liminf_{n\to+\infty} P(\|X_n\| > K) = 1. \tag{2.6}$$

This means that on the space (H, τ_w) there are weakly convergent sequences (to $\mu_0 = \delta_0$ in (2.5)) 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 [32]). But again the picture is not clear, since uniform tightness on (H, τ_w), i.e.

$$\lim_{K\to+\infty} \sup_n P(\|X_n\| > K) = 0, \tag{2.7}$$

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 1.1.1 from Paper I, which asserts that every sequence satisfying (2.7) 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 $Y_0, Y_1, \dots$ such that

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \dots \tag{2.8}$$

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

$$< y, Y_k(\omega) > \longrightarrow < y, Y_0(\omega) >, \quad \text{as } k \to \infty. \tag{2.9}$$

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

$$Ef(X_{n_k}) = Ef(Y_k) \longrightarrow Ef(Y_0), \quad \text{as } k \to \infty. \tag{2.10}$$

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 $\{\mu_{n_{k_l}} : l = 1, 2, \dots\}$ admit a Skorohod representation defined on the Lebesgue interval and almost surely convergent "in compacts". (2.11)

(For precise definitions we refer to Section 2.3). 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 2.3.5, Section 2.3);
- the converse Prohorov's theorem is quite natural and holds in many spaces (Theorems 2.4.1 – 2.4.5 and 2.4.7, Section 2.4);
- 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 2.2);
- on metric spaces the theory of the usual weak convergence of tight probability distributions remains unchanged (Corollary 2.3.8, Section 2.3).

2.2 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. (2.12)

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 $\mathcal{X}$. (2.13)

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 $i \in \mathbb{I}$ $f_i(x_n)$ converges to some number α_i, then x_n τ-converges to some x_0 and $f_i(x_0) = \alpha_i,\ i \in \mathbb{I}$. (2.14)

Assumption (2.13) defines a continuous mapping $\tilde{f} : \mathcal{X} \to [-1,1]^{\mathbb{I}}$ given by formula

$$\tilde{f}(x) = (f_i(x))_{i\in\mathbb{I}}. \tag{2.15}$$

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

$\mathcal{X}$ is a Hausdorff space (but need not be regular). (2.16)

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 [18], but still satisfies (2.13).

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

Every compact subset is $\sigma(f_i : i \in \mathbb{I})$-measurable (hence is a Baire subset of $\mathcal{X}$) and is metrisable. (2.17)

In many cases $\sigma(f_i : i \in \mathbb{I})$ 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 $\sigma(f_i : i \in \mathbb{I})$. Moreover, every tight probability measure μ defined on $\sigma(f_i : i \in \mathbb{I})$ can be uniquely extended to the whole σ-algebra of Borel sets. Hence if $X : (\Omega, \mathcal{F}, P) \to \mathcal{X}$ is $\sigma(f_i : i \in \mathbb{I})$-measurable and the law of X (as the measure on $\sigma(f_i : i \in \mathbb{I})$) is tight, then X is Borel-measurable if we replace $\mathcal{F}$ with its P-completion $\overline{\mathcal{F}}$. In particular, if $\{f'_i\}_{i\in\mathbb{I}'}$ is another family satisfying (2.13), then $X : (\Omega, \overline{\mathcal{F}}, P) \to \mathcal{X}$ is $\sigma(f'_i : i \in \mathbb{I}')$-measurable.

The above remarks show that our considerations do not depend essentially on the choice of the family $\{f_i\}_{i\in\mathbb{I}}$ satisfying (2.13). Therefore without loss of generality we may fix *some* family $\{f_i\}_{i\in\mathbb{I}}$ and shall restrict the attention to **random elements X such that $f_i(X),\ i \in \mathbb{I}$, are random variables and the law of X is tight, and to tight probability measures defined on $\sigma(f_i : i \in \mathbb{I})$.** As in Section 2.1, 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 $([0,1], \mathcal{B}_{[0,1]}, \ell)$. (2.18)

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 $[0,1]^{I\!I}$. 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 $[0,1]^{I\!I}$, concentrated on the σ-compact subspace $\tilde{f}(\mathcal{X}_1)$. But it is well-known (see e.g. [4]) that then there exists a measurable mapping $Y : [0,1] \to [0,1]^{I\!I}$ such that

$$\mu \circ \tilde{f}^{-1} = \ell \circ Y^{-1}, \tag{2.19}$$

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

$$X(\omega) = \begin{cases} \tilde{f}^{-1}(Y(\omega)), & \text{if } Y(\omega) \in \tilde{f}(\mathcal{X}_1); \\ x_0, & \text{otherwise.} \end{cases} \tag{2.20}$$

Using somewhat subtler reasoning than the one used in the proof of (2.17) 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). (2.21)

Here again the standard example exhibits the pathology signalized in (2.21): the whole space $[0,1]$ is not compact, but it is a closure of a relatively compact set $[0,1] \setminus A$. Remark (2.21) affects the definition of uniform tightness where we cannot, in general, replace sequential compactness with measurability and relative compactness. In a similar way as (2.21) one can prove

$K \subset \mathcal{X}$ is compact iff it is sequentially compact. (2.22)

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. (2.23)

To prove (2.23) let us observe first that $(\mathcal{X}, \tau_s)$ also satisfies (2.13), for τ-continuity implies *sequential* τ-continuity and so τ_s-continuity. By (2.22) 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 (2.12). 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 (2.23) is

Any uniformly τ-tight sequence of random elements in $\mathcal{X}$ is uniformly τ_s-tight. (2.24)

Facts like (2.23) and (2.24) suggest that whenever we deal with uniform tightness (or Prohorov's theorem) sequential spaces satisfying (2.13) may be of special importance.

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

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

We say that $\mathcal{X}$ is a space of type $\mathcal{L}$ (Fréchet, [13]), if among all sequences of elements of $\mathcal{X}$ a class $\mathcal{C}(\to)$ of "convergent" sequences is distinguished, and to each convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ 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. (2.25)

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 $k \to \infty$. (2.26)

It is easy to see that in the space $\mathcal{X}$ of type $\mathcal{L}$ the statement paralleling (2.12):

$F \subset \mathcal{X}$ is *closed* if F contains all limits of "$\longrightarrow$"-convergent sequences of elements of F. (2.27)

defines a topology, $\mathcal{O}(\to)$ say. This topology defines in turn a new (in general) class of convergent sequences, which can be called convergent "a posteriori" (Urysohn, [37]), 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 $G \in \mathcal{O}(\to)$ eventually all elements of the sequence $\{x_n\}$ belong to G. Kantorowich *et al* [16, Theorem 2.42, p. 51] and Kisyński [17] 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*. (2.28)

We see that convergence *a posteriori* shares property (2.3) 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. (2.29)

If the $\mathcal{L}$-convergence "$\longrightarrow$" satisfies also (2.29), then we say that $\mathcal{X}$ is of type $\mathcal{L}^*$ and will denote such convergence by "$\xrightarrow{*}$". 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 (2.25) and (2.26), we can *weaken* this convergence to convergence "$\xrightarrow{*}$" satisfying additionally (2.28), and the latter convergence is already the usual convergence of sequences in the topological space $(\mathcal{X}, \mathcal{O}(\to)) \equiv (\mathcal{X}, \mathcal{O}(\xrightarrow{*}))$. At least two examples of such a procedure are well-known:

Example 2.2.1 If "$\longrightarrow$" denotes the convergence "almost surely" of real random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, then "$\xrightarrow{*}$" is the convergence "in probability".

Example 2.2.2 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. (2.30)

A function f on $\mathcal{X}$ is $\mathcal{O}(\stackrel{*}{\to})$-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 $x_n \longrightarrow x_0$). (2.31)

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 [9] or [2].

2.3 The sequential topology of the convergence in distribution

The reason we are interested in topological spaces satisfying (2.13) is Theorem 1.2.1 from Paper I (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 [29].

Theorem 2.3.1 *Let $(\mathcal{X}, \tau)$ be a topological space satisfying (2.13) and let $\{\mu_n\}_{n\in I\!N}$ 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*

$$X_{n_k} \sim Y_k, \quad k = 1, 2, \dots, \tag{2.32}$$

$$Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty,\ \omega \in [0,1]. \tag{2.33}$$

Let us notice that contrary to the metric case under (2.13) 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

$$C(K) = \{\omega : Y_k(\omega) \underset{\tau}{\longrightarrow} Y_0(\omega), \quad \text{as } k \to \infty\} \cap \bigcap_{k=1}^{\infty} \{\omega : Y_k(\omega) \in K\}, \tag{2.34}$$

where $K \subset \mathcal{X}$ is compact. This becomes obvious when we observe that by property (2.14) 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

$$P(C(K_\varepsilon)) > 1 - \varepsilon. \tag{2.35}$$

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 2.3.2 *Convergence almost surely "in compacts" implies uniform tightness.*

The a.s. convergence (2.33) has been established exactly the way described above. If the representation $Y_0, Y_1, Y_2, \ldots$ satisfies (2.32) and the convergence (2.33) 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 2.3.1 in the following form:

Theorem 2.3.3 *Let $(\mathcal{X}, \tau)$ be a topological space satisfying (2.13) and let $\{\mu_n\}_{n\in I\!N}$ be a uniformly tight sequence of laws on $\mathcal{X}$. Then there exists a subsequence $\mu_{n_1}, \mu_{n_2}, \ldots$ 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 2.1 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 $\mu_0, \mu_{n_1}, \mu_{n_2}, \ldots$ admit the strong a.s. Skorohod representation defined on the Lebesgue interval. (2.36)

As an immediate corollary to Theorem 2.3.3 we obtain the direct Prohorov's theorem for "$\stackrel{*}{\Longrightarrow}$".

Theorem 2.3.4 *If $(\mathcal{X}, \tau)$ satisfies (2.13), 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 (2.25), (2.26) and (2.29). Notice that (2.25) holds by (2.18), and that (2.29) is exactly condition (2.3) 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 (2.10), we see that for any *sequentially continuous* and bounded function $f : (\mathcal{X}, \tau_s) \to I\!R^1$, the mapping

$$\mathcal{P}(\mathcal{X}) \ni \mu \mapsto \int_{\mathcal{X}} f(x)\, \mu(dx) \in I\!R^1, \tag{2.37}$$

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 2.1.

Theorem 2.3.5 *Suppose $(\mathcal{X}, \tau)$ satisfies (2.13). 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. (2.38)

The class of relatively $\mathcal{O}$-compact sets coincides with the class of relatively uniformly τ-tight sets. (2.39)

Proof. Relation (2.39) gives us the family of relatively compact subsets and (2.38) helps us to identify limiting points. This information fully determines an $\mathcal{L}^*$-convergence. □

Remark 2.3.6 Analysing Fernique's example quoted in Introduction shows that (2.39) 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 2.4.

Remark 2.3.7 In many respects the topological space $(\mathcal{P}(\mathcal{X}), \mathcal{O}(\stackrel{*}{\Longrightarrow}))$ is as good as $(\mathcal{X}, \tau)$ is: the property (2.13) is hereditary. To see this, take as the separating functions

$$h_{(i_1,i_2,\ldots,i_m)}(\mu) = \int_{\mathcal{X}} f_{i_1}(x) f_{i_2}(x) \ldots f_{i_m}(x)\, \mu(dx), \tag{2.40}$$

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 2.3.5 does not contain the case of an arbitrary metric space, since in nonseparable spaces condition (2.13) may fail. However we have

Corollary 2.3.8 *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 [20], [3]. 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 [3] that the weak topology on $\mathcal{P}(\mathcal{X})$ is metrisable and so is sequential. □

Remark 2.3.9 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 (2.28).

Remark 2.3.10 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 [36]), coincides with weak convergence on metric spaces and by means of the Prohorov's theorem is operational and easy in handling.

2.4 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 2.4.1 *Suppose that $(\mathcal{X}, \tau)$ satisfies (2.13) and possesses a fundamental system $\{K_m\}$ of compact subsets. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following statements are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{2.41}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{2.42}$$

Proof. In view of Theorem 2.3.4 we have to prove that (2.41) implies (2.42). Suppose (2.42) does not hold. Then there is $\varepsilon > 0$ such that for each m one can find $\mu_m \in \mathcal{K}$ satisfying

$$\mu_m(K_m^c) > \varepsilon. \tag{2.43}$$

By $\stackrel{*}{\Longrightarrow}$-relative compactness there exists a subsequence μ_{m_k} admitting a strong a.s. Skorohod representation. By Corollary 2.3.2 $\{\mu_{m_k}\}_{k\in I\!N}$ is uniformly tight. This contradicts (2.43). □

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

$$\sup_{x\in K} h_k(x) < +\infty, \ k \in I\!K, \tag{2.44}$$

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

$$h_k(x_0) \leq \liminf_{n\to\infty} h_k(x_n) < +\infty, \ k \in I\!K. \tag{2.45}$$

Notice that under (2.45) any relatively compact set K satisfies (2.44) and is contained in some set of the form

$$K = \bigcap_{k\in I\!K} \{x : h_k(x) \leq C_k\}. \tag{2.46}$$

Moreover, under both (2.44) and (2.45) every set of the form (2.46) is *sequentially compact.*

Theorem 2.4.2 *Let $(\mathcal{X}, \tau)$ satisfies (2.13). Suppose compactness in $(\mathcal{X}, \tau)$ is given by boundedness with respect to a countable family $\{h_k\}_{k\in I\!K}$ of lower semicontinuous functionals. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{2.47}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{2.48}$$

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.

$$\lim_{C\to\infty} \sup_{\mu\in\mathcal{K}} \mu(\{x : h_k(x) > C\}) = 0. \tag{2.49}$$

PROOF. Conditions (2.48) and (2.49) are obviously equivalent and implication (2.48) $\Rightarrow$ (2.47) is proved in Theorem 2.3.4. In order to prove that (2.47) implies (2.49) 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

$$\mu_N(\{x : h_k(x) > N\}) \geq \varepsilon, \quad N \in I\!\!N. \tag{2.50}$$

If *some* subsequence of μ_N admits a strong a.s. Skorohod representation, it must be uniformly tight and (2.50) cannot hold along this subsequence. This shows that $\mathcal{K}$ is not $\stackrel{*}{\Longrightarrow}$-relatively compact. □

It is worth to emphasize that Theorem 2.4.2 *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 $k \in I\!\!N$

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

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 (2.45) 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 2.4.2. For example on the Skorohod space $I\!\!D([0,1] : I\!\!R^1)$ there exists (see Paper III) a minimal functional topology which is non-metrisable but satisfies (2.44) and (2.45), hence by our Theorem 2.4.2 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:

$$g_{k,j+1} \leq g_{k,j}, \quad k \in I\!\!K, j \in I\!\!N. \tag{2.51}$$

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

$$\lim_{j\to\infty} \sup_n g_{k,j}(x_n) = 0. \tag{2.52}$$

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

$$\lim_{j\to\infty} \sup_{x \in K} g_{k,j}(x) = 0, \tag{2.53}$$

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 (2.45) implies (2.52) and (2.44) and lower semicontinuity of h_k give conditional compactness in (2.53). Recall that in general in spaces satisfying (2.13) relative compactness does not imply conditional compactness. In metric spaces, however, it does and so e.g. Skorohod topology J_2 [29] (and not only J_1) satisfies the converse Prohorov theorem, as we can see from the following result.

Theorem 2.4.3 *Let $(\mathcal{X}, \tau)$ satisfies (2.13). Suppose conditions (2.51) – (2.53) determine conditional compactness in $(\mathcal{X}, \tau)$. Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$$\mathcal{K} \text{ is } \stackrel{*}{\Longrightarrow}\text{-relatively compact.} \tag{2.54}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{2.55}$$

For each $k \in I\!\!K$

$$\lim_{j\to\infty} \sup_{\mu\in\mathcal{K}} \mu(\{x : g_{k,j}(x) > \varepsilon\}) = 0, \quad \varepsilon > 0. \tag{2.56}$$

Proof. Similarly as before, it is enough to show that if (2.56) 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 (2.51) and (2.52), for each $k \in I\!\!K$ and almost surely,

$$\limsup_{l\to\infty} g_{k,j_l}(X_l) \le \lim_{j\to\infty} \limsup_{l\to\infty} g_{k,j}(X_l) = 0. \tag{2.57}$$

If (2.56) 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

$$\mu_j(\{x : g_{k,j}(x) > \varepsilon\}) \ge \varepsilon. \tag{2.58}$$

If X_l is the a.s. Skorohod representation for some subsequence μ_{j_l} then by (2.57)

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

hence (2.58) 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 nonnegative measurable functionals such that

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

$$\sup_{x\in K} q_{m_0}(x) \le C_{m_0} < +\infty. \tag{2.59}$$

Notice this implies

$$\sup_{m\ge m_0} \sup_{x\in K} q_m(x) \le C_{m_0},$$

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

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

Theorem 2.4.4 *Let $(\mathcal{X}, \tau)$ satisfies (2.13) and (2.59). Then for $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$ the following conditions are equivalent:*

$$\mathcal{K} \text{ is } \overset{*}{\Longrightarrow}\text{-relatively compact.} \tag{2.60}$$

$$\mathcal{K} \text{ is uniformly } \tau\text{-tight.} \tag{2.61}$$

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

$$\sup_{\mu \in \mathcal{K}} \mu(\{x : q_{m_0}(x) > C\}) < \varepsilon. \tag{2.62}$$

Proof. We apply the standard strategy. If (2.62) is not satisfied, then there is $\varepsilon > 0$ such that for every M and for some $\mu_M \in \mathcal{K}$

$$\mu_M(\{x : q_M(x) > M\}) \geq \varepsilon. \tag{2.63}$$

If $\{X_k\}$ is the strong a.s. Skorohod representation for some subsequence μ_{M_k}, then it is tight (by Corollary 2.3.2). and so for some m_0 and C

$$P(q_{m_0}(X_k) \leq C) = \mu_{M_k}(\{x : q_{m_0}(x) \leq C\}) > 1 - \varepsilon, \quad k = 1, 2, \ldots. \tag{2.64}$$

Hence for k satisfying $M_k > C$ and $M_k > m_0$ we get from (2.63) and (2.64)

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

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 [25]). 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 2.4.5 *Let $(\mathcal{X}, \tau)$ be an S-P space satisfying (2.13). 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 (2.21) 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, [25], pp. 109-110. □

Corollary 2.4.6 *Any S-P space satisfying (2.13) has the property that the closure of a relatively compact set is compact and consists of the set itself and its limiting points.* □

Theorem 2.4.7 *Let* $(\mathcal{X}_n, \tau_n),\ n = 1, 2, \ldots$ *be S-P spaces satisfying (2.13). 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 [16] for Kisyński's theorem.

Paper No 3.

A non-Skorohod topology on the Skorohod space

Abstract

A new topology (called S) is defined of the space $I\!D$ of functions $x : [0,1] \to I\!R^1$ which are right-continuous and admit limits from the left at each $t > 0$. This topology converts $I\!D$ into a linear topological space but cannot be metricized. Nevertheless, S is quite natural and shares many useful properties with the traditional Skorohod's topology J_1. In particular, on the space $\mathcal{P}(I\!D)$ of laws of stochastic processes with trajectories in $I\!D$ the topology S induces a sequential topology for which both the direct and the converse Prohorov's theorems are valid, the a.s. Skorohod representation for subsequences exists and finite dimensional convergence outside a countable set holds.

3.1 Introduction

Let $I\!D = I\!D([0,1] : I\!R^1)$ be the space of functions $x : [0,1] \to I\!R^1$ which are right-continuous and admit limits from the left at each $t > 0$. We are going to study a new sequential topology on $I\!D$ generated by naturally arising criteria of relative compactness. The novelty is that this topology cannot be metricized. Nevertheless we shall show how to build a complete and satisfactory theory of the convergence in distribution with respect to this topology.

Despite metric topologies are sequential, the process of defining topology through description of the family of convergent sequences is not the common approach, especially in probability theory. We refer to Paper II for rather extensive discussion of sequential methods in weak limit theorems of probability theory. Here we shall point out only that

> every time one proves a limit theorem via the direct Prohorov's theorem, one obtains a result for the *sequential* topology generated by the original one.

Since Prohorov's theorems are accepted tools of probability theory the same should happen to sequential methods, as they fit the original Prohorov's and Skorohod's ideas much better than the theory based on weak-$*$ convergence of distributions.

In order to show our motivations we begin with a simple, well-known example. Let $\mathbb{V}^+ \subset \mathbb{D}$ consists of *nonnegative* and *nondecreasing* functions $v : [0,1] \to I\!R^+$. Suppose that for some subset $K \subset \mathbb{V}^+$ we have:

$$\sup_{v \in K} v(1) < +\infty. \tag{3.1}$$

Let $\mathbb{Q} \subset [\nvdash, \nVdash]$ be countable dense and let $1 \in \mathbb{Q}$. By (3.1) we may find a sequence $\{v_n\}_{n \in I\!N} \subset K$ such that for each $q \in \mathbb{Q}$

$$v_n(q) \to \tilde{v}(q),$$

where $\tilde{v}(q)$ are numbers satisfying $\tilde{v}(q') \leq \tilde{v}(q'')$, $q' < q''$, $q', q'' \in \mathbb{Q}$. Hence the function

$$v_0(t) = \begin{cases} \inf\limits_{\substack{q>t \\ q\in\mathbb{Q}}} \tilde{v}(q) & \text{if} \quad t < 1; \\ \tilde{v}(1) & \text{if} \quad t = 1, \end{cases} \tag{3.2}$$

belongs to $\mathbb{V}^+$. Moreover, we have as $n \to \infty$

$$v_n(t) \to v_0(t), \tag{3.3}$$

provided $t = 1$ or t is a point of continuity of v_0: $v_0(t) = v_0(t-)$. Define finite measures on $([0,1], \mathcal{B}_{[0,1]})$ by formula

$$\mu_n([0,t]) = v_n(t), \ t \in [0,1], \ n = 0, 1, 2, \ldots,$$

and observe that (3.3) is equivalent to the weak convergence of μ_n's, i.e. convergence of μ_n's considered as continuous linear functionals on the space $C([0,1] : \mathbb{R}^1)$ of continuous functions on $[0,1]$ equipped with the weak-$*$ topology:

$$\mu_n \Rightarrow \mu_0 \quad \text{iff} \quad \int f(t)\, d\mu_n(t) \to \int f(t)\, d\mu_0(t), \ f \in C([0,1] : \mathbb{R}^1).$$

It follows that condition (3.1) when restricted to $\mathbb{V}^+$ is a criterion of relative compactness for some, quite natural topology.

A very similar procedure may be performed for the space $\mathbb{D}$. Suppose that

$$\sup_{x\in K} \sup_{t\in[0,1]} |x(t)| \leq C_K < +\infty, \tag{3.4}$$

and that for all $a < b, \ a, b \in \mathbb{R}$

$$\sup_{x\in K} N^{a,b}(x) \leq C_K^{a,b} < +\infty, \tag{3.5}$$

where $N^{a,b}$ is the usual number of up-crossing given levels $a < b$. (Recall that $N^{a,b}(x) \geq k$ if one can find numbers $0 \leq t_1 < t_2 < \ldots < t_{2k-1} < t_{2k} \leq 1$ such that $x(t_{2i-1}) < a$ and $x(t_{2i}) > b$, $i = 1, 2, \ldots, k$). Let, as previously, $\mathbb{Q} \subset [\nvdash, \nVdash]$, $\mathbb{Q} \ni \nVdash$, be countable dense. By (3.4) we can extract a sequence $\{x_n\}_{n\in\mathbb{N}} \subset K$ such that, as $n \to \infty$

$$x_n(q) \to \tilde{x}(q), \ q \in \mathbb{Q}. \tag{3.6}$$

Now the construction of the limiting function x_0 is not as easy as in the case of $\mathbb{V}^+$ and one has to use (3.5) in essential way to see that

$$x_0(t) = \begin{cases} \lim\limits_{\substack{q\to t+ \\ q\in\mathbb{Q}}} \tilde{x}(q) & \text{if} \quad t < 1; \\ \tilde{x}(1) & \text{if} \quad t = 1, \end{cases} \tag{3.7}$$

is well-defined and belongs to $\mathbb{D}$. And whether x_n converges to x_0 in some topology on $\mathbb{D}$ is not clear at all.

Meyer and Zheng [23] considered on $I\!\!D$ so-called pseudo-path topology and proved that (3.4) and (3.5) imply conditional compactness of K in this topology. However, the pseudo-path topology was shown to be equivalent on $I\!\!D$ to the convergence in (Lebesgue) measure. And so neither (3.4) nor (3.5) form the proper description of the relative compactness in the pseudo-path topology, for it is easy to find a sequence $\{x_n\}$ of elements of $I\!\!D$ which is convergent in measure to an element $x_0 \in I\!\!D$ and is such that for $K = \{x_n : n = 0, 1, 2, \ldots\}$ both (3.4) and (3.5) are not satisfied.

In the present paper we construct a topology on $I\!\!D$, say S, which possesses the following properties.

- $K \subset I\!\!D$ is relatively S-compact iff both (3.4) and (3.5) hold.
- S is sequential and cannot be metricized. $(I\!\!D, S)$ is a linear topological space.
- The σ-field $\mathcal{B}_S$ of Borel subsets for S coincides with the usual σ-field generated by projections (or evaluations) on $I\!\!D$: $\mathcal{B}_S = \sigma(\pi_t : t \in [0, 1])$.
- The set $\mathcal{P}(I\!\!D, S)$ of *S-tight* probability measures is exactly the set of distributions of stochastic processes with trajectories in $I\!\!D$: $\mathcal{P}(I\!\!D, S) = \mathcal{P}(I\!\!D)$.
- S is weaker than Skorohod's J_1-topology. Since the latter is Polish, S is Lusin in the sense of Fernique. Being a linear topological space, $(I\!\!D, S)$ is also a completely regular topological space.

Hence it is possible to use the full power of the existing theory for regular Lusin spaces ([11], [12]). The point is that in some Lusin spaces the direct Prohorov's theorem is not the proper tool for proving limit theorems based on the weak-$*$ convergence of probability measures (and, consequently, there is no the a.s. Skorohod representation). And in such a case the classical weak topology on $\mathcal{P}(I\!\!D, S)$ becomes difficult to handle. Therefore we suggest an alternative approach, based on the topology induced by the a.s. Skorohod representation.

- On $\mathcal{P}(I\!\!D)$ there exists a unique sequential topology $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ (where $\stackrel{*}{\Longrightarrow}$ denotes the convergence determining the topology) which is finer than the weak topology and for which $\stackrel{*}{\Longrightarrow}$-relative compactness coincides with uniform S-tightness. In particular, for $\mathcal{O}(\stackrel{*}{\Longrightarrow})$ both the direct and the converse Prohorov's theorems are valid.
- Let $X_n \ \stackrel{*}{\longrightarrow}_{\mathcal{D}}\ X_0$ means that the laws of processes X_n converge in our sense: $\mathcal{L}(X_n) \stackrel{*}{\Longrightarrow} \mathcal{L}(X_0)$. Suppose $X_n \ \stackrel{*}{\longrightarrow}_{\mathcal{D}}\ X_0$. Then in each subsequence $\{X_{n_k}\}_{k \in I\!\!N}$ one can find a further subsequence $\{X_{n_{k_l}}\}_{l \in I\!\!N}$ such that:
  - $\{X_0\} \cup \{X_{n_{k_l}} : l = 1, 2, \ldots\}$ admit the usual a.s. Skorohod representation on $([0, 1], \mathcal{B}_{[0,1]})$;
  - outside some *countable* set $\mathbb{Q}_{\nvdash} \subset [\nvdash, \nVdash)$ all finite dimensional distributions of $\{X_{n_{k_l}}\}$ converge to those of X_0.

The last statement improves a corresponding result due to Meyer and Zheng [23], where finite dimensional convergence outside a set of Lebesgue measure null was shown via results on the pseudo-path topology.

Results on finite dimensional convergence are of quite different flavour than in the case of the Skorohod J_1-topology: for every $t \neq 1$ the projection $\pi_t : I\!\!D \to I\!\!R^1$, $\pi_t(x) = x(t)$, is *nowhere* S-continuous and so the standard continuity arguments cannot be applied. There exists, however, a simple procedure, completely paralleling the one described for $\mathbb{D}$ in (3.4) – (3.7) and allowing to determine the finite dimensional distributions of the limiting process (hence identifying the limits). Suppose that $\{X_n\}$ is a uniformly S-tight sequence. Choose a dense countable subset $\mathbb{Q} \subset [0,1]$, $\mathbb{Q} \ni 1$, and extract a subsequence $\{X_{n_k}\}$ such that for each finite sequence $q_1 < q_2 < \ldots < q_m$ of elements of $\mathbb{Q}$ we have

$$(X_{n_k}(q_1), X_{n_k}(q_2), \ldots, X_{n_k}(q_m)) \underset{\mathcal{D}}{\longrightarrow} \tilde{\nu}_{(q_1,q_2,\ldots,q_m)}, \tag{3.8}$$

where $\tilde{\nu}_{(q_1,q_2,\ldots,q_m)}$ is a probability distribution on $\mathbb{R}^m$. Then one can prove that for each finite sequence $t_1 < t_2 < \ldots < t_m$ and each approximating sequence $q_{1,l} < q_{2,l} < \ldots < q_{m,l}$, $q_{j,l} \searrow t_j$, as $l \to \infty$, ($q_{m,l} = 1$ if $t_m = 1$), probability distributions $\tilde{\nu}_{(q_{1,l},q_{2,l},\ldots,q_{m,l})}$ weakly converge to some limit $\nu_{(t_1,t_2,\ldots,t_m)}$. Moreover, there is a stochastic process X_0 with trajectories in $\mathbb{D}$ such that for every finite sequence $t_1 < t_2 < \ldots < t_m$

$$(X_0(t_1), X_0(t_2), \ldots, X_0(t_m)) \sim \nu_{(t_1,t_2,\ldots,t_m)}, \tag{3.9}$$

and $X_{n_k} \xrightarrow{*}_{\mathcal{D}} X_0$. This simple procedure was known to hold in the Skorohod J_1-topology ([36]); it is interesting to see that it preserves validity at much lower level.

The paper is organized as follows.

In Section 3.2 we define the topology S and give its basic properties.

In Section 3.3 we apply to the space $\mathcal{P}(\mathbb{D}, \mathcal{S})$ the machinery developed in Paper II which is suitable for spaces with countable continuous separation property (as it is the case). In particular we prove all the results announced above.

In Section 3.4 we show how the theory can be applied to sets of semimartingales satisfying so-called Condition UT which implies uniform S-tightness and which is known to be important in limit theorems for stochastic integrals and stochastic differential equations ([14], [30],[19], [31], Paper IV).

3.2 The topology S

We shall define the topology S in several steps, following the strategy described in Section 2.2 of Paper II (the reader is referred to that paper for definitions and notation). Here is the description of subsequent steps.

1. Give equivalent reformulations for the (potential) criteria of compactness (3.4) and (3.5)—Lemma 3.2.1.
2. Find an $\mathcal{L}$-convergence for which (3.4) and (3.5) are criteria of relative compactness—Lemma 3.2.7.
3. Define S-topology and construct an $\mathcal{L}^*$-convergence using the Kantorovich-Kisyński recipe.
4. Show there exists a countable family of S-continuous functions separating points of $\mathbb{D}$ and conclude the topology S is suitable for needs of probability theory.
5. Investigate in some detail other properties of the topology S (Propositions 3.2.14 and 3.2.15).

Let us denote by $\|x\|_\infty$ the supremum norm on $I\!D$:

$$\|x\|_\infty = \sup_{t\in[0,1]} |x(t)|,$$

and by $\|v\|$ the total variation of v:

$$\|v\|(t) = \sup\{|v(0)| + \sum_{i=1}^{m} |v(t_i) - v(t_{i-1})| : 0 = t_0 < t_1 < \ldots < t_m = t,\ m \in I\!N\}.$$

For $a < b$ let the number of up-crossings $N^{a,b}$ be defined as above (see the lines after formula (3.5)) and for $\eta > 0$ let the number N_η of η-oscillations be defined by the relation: $N_\eta(x) \geq k$ iff one can find numbers $0 \leq t_1 \leq t_2 \leq \ldots \leq t_{2k-1} \leq t_{2k} \leq 1$ such that $|x(t_{2i-1}) - x(t_{2i})| > \eta$, $i = 1, 2, \ldots, k$. Finally, let $\mathbb{V} = \mathbb{V}^+ - \mathbb{V}^+$.

Lemma 3.2.1 *Let $K \subset I\!D$ and suppose that*

$$\sup_{x\in K} \|x\|_\infty < +\infty. \tag{3.10}$$

Then the statements (i) *and* (ii) *below are equivalent:*

(i) *For each $a < b$*

$$\sup_{x\in K} N^{a,b}(x) < +\infty. \tag{3.11}$$

(ii) *For each $\eta > 0$*

$$\sup_{x\in K} N_\eta(x) < +\infty. \tag{3.12}$$

Moreover, either set of conditions (3.10)+(3.11) and (3.10)+(3.12) is equivalent to

(iii) *For each $\varepsilon > 0$ and for each $x \in K$ there exists $v_{x,\varepsilon} \in \mathbb{V}$ such that*

$$\sup_{x\in K} \|x - v_{x,\varepsilon}\|_\infty \leq \varepsilon, \tag{3.13}$$

and

$$\sup_{x\in K} \|v_{x,\varepsilon}\|(1) < +\infty. \tag{3.14}$$

Proof. Let us observe first that **(iii)** implies (3.10):

$$\|x\|_\infty \leq \varepsilon + \|v_{x,\varepsilon}(x)\|(1). \tag{3.15}$$

Then the chain of implications **(iii)**$\Rightarrow$**(ii)**$\Rightarrow$ **(i)** follows by inequalities

$$N_\eta(x) \leq \frac{v_{x,\varepsilon}}{\eta - 2\varepsilon}, \quad \eta > 2\varepsilon > 0,\ x \in I\!D, \tag{3.16}$$

and

$$N^{a,b}(x) \leq N_{b-a}(x), \quad b > a,\ x \in I\!D. \tag{3.17}$$

Now assume (3.10). First we shall prove **(i)**$\Rightarrow$**(ii)**. Let $C_\infty = \sup_{x\in K} \|x\|_\infty$ and suppose that for some $\eta > 0$ and every $n \in I\!N$ there is $x_n \in K$ such that $N_\eta(x_n) \geq n$. In particular, for some $0 \le t_{n,1} < t_{n,2} < \ldots < t_{n,2n-1} < t_{n,2n} \le 1$ we have $|x_n(t_{n,2i}) - x_n(t_{n,2i-1})| > \eta$, $i = 1, 2, \ldots, n$. Let $a_1 < a_2 < \ldots < a_R$ be an $\eta/2$-net for the interval $I = [-C_\infty, C_\infty]$, i.e. for each $x \in I$ there is a_j such that $|x - a_j| < \eta/2$. In every interval $I_{n,i}$ with ends in points $x_n(t_{n,2i-1})$ and $x_n(t_{n,2i})$ there are at least two points a_{j-1} and a_j which belong to $I_{n,i}$ and are distinct from the interval's ends. Let $M_j(n)$ be the number of i's such that both a_{j-1} and a_j belong to $I_{n,i}$. It follows that for some j_0

$$\sup_n M_{j_0}(x_n) = +\infty. \tag{3.18}$$

But $N^{a_{j_0-1},a_{j_0}}(x_n)$ is greater or equal to the integer part of $(M_{j_0}(x_n) - 1)/2$ and so by (3.18)

$$\sup_{x \in K} N^{a_{j_0-1},a_{j_0}}(x) \ge \sup_n N^{a_{j_0-1},a_{j_0}}(x_n) = +\infty.$$

It remains to prove **(ii)⇒(iii)**. The construction of $v_{x,\varepsilon}$ is, in some sense, standard. For $\varepsilon > 0$ let us define

$$\begin{aligned} \tau_0^\varepsilon(x) &= 0 && (3.19)\\ \tau_k^\varepsilon(x) &= \inf\{t > \tau_{k-1}^\varepsilon(x) : |x(t) - x(\tau_{k-1}^\varepsilon(x))| > \varepsilon\},\ k = 1, 2, \ldots. && (3.20)\end{aligned}$$

(where by convention $\inf \emptyset = +\infty$) and let

$$v_\varepsilon(x)(t) = x(\tau_k^\varepsilon(x)) \ \text{ if } \tau_k^\varepsilon(x) \le t < \tau_{k+1}^\varepsilon(x),\ t \in [0,1], k = 0, 1, 2, \ldots. \tag{3.21}$$

Then by the very definition

$$\begin{aligned} \|x - v_\varepsilon(x)\|_\infty &\le \varepsilon, && (3.22)\\ \|v_\varepsilon(x)\|(1) &\le \|x\|_\infty (2N_{\varepsilon/2}(x) + 1), && (3.23)\end{aligned}$$

and the lemma follows. □

Corollary 3.2.2 *For each $t \in [0,1]$ the mapping $I\!\!D \ni x \mapsto (v_\varepsilon(x))(t)$ defined by (3.21) is $\mathcal{F}_{t+}$-measurable, where $\{\mathcal{F}_{t+}\}_{t \in [0,1]}$ is the natural right-continuous filtration on the canonical space $I\!\!D$: $\mathcal{F}_{t+} = \cap_{u>t} \sigma(\pi_s : 0 \le s \le u)$. Hence $v_\varepsilon(X)$ is an adapted stochastic process provided X is adapted to a right-continuous filtration.*

Before introducing a convergence in $I\!\!D$ which generates the S-topology, let us recall some facts on the weak-$*$ topology on $\mathbb{V}$. Any element $v \in \mathbb{V}$ determines a signed measure ν on $([0,1], \mathcal{B})$ given by the formula

$$\nu([0,t]) = v(t), \quad t \in [0,1].$$

Since the set of signed measures can be identified with the dual of the Banach space $C([0,1] : I\!\!R^1)$, $\mathbb{V}$ can be equipped with the weak-$*$ topology. Convergence of elements of $\mathbb{V}$ in this topology will be denoted by $\longrightarrow_w$. In particular, $v_n \longrightarrow_w v_0$ means that for every continuous function $f : [0,1] \to I\!\!R^1$

$$\int_{[0,1]} f(t)\, dv_n(t) \underset{w}{\longrightarrow} \int_{[0,1]} f(t)\, dv_0(t). \tag{3.24}$$

Definition 3.2.3 We shall write $x_n \longrightarrow_S x_0$ if for every $\varepsilon > 0$ one can find elements $v_{n,\varepsilon}$, $n = 0, 1, 2, \ldots$ which are ε-uniformly close to x_n's and weakly-$*$ convergent:

$$\|x_n - v_{n,\varepsilon}\|_\infty \le \varepsilon, \qquad n = 0, 1, 2, \ldots, \tag{3.25}$$

$$v_{n,\varepsilon} \underset{w}{\longrightarrow} v_{0,\varepsilon}, \qquad \text{as } n \to +\infty. \tag{3.26}$$

Remark 3.2.4 (3.26) implies that

$$v_{n,\varepsilon}(t) \to v_{0,\varepsilon}(t), \tag{3.27}$$

for each t outside a countable set $D_\varepsilon \subset [0, 1)$. Taking $\varepsilon = 1, 1/2, 1/3, \ldots$ we obtain

$$x_n(t) \to x_0(t), \tag{3.28}$$

for each $t \in [0, 1] \setminus \bigcup_{m=1}^\infty D_{1/m}$. Hence the limit for $\longrightarrow_S$ is determined uniquely. Further, step functions are dense in $\mathbb{D}$ for the uniform topology and so for the constant sequence $x_n \equiv x_0$, $n = 1, 2, \ldots$, we have $x_n \longrightarrow_S x_0$. Since also a subsequence of a sequence convergent in the sense of $\longrightarrow_S$ is (obviously) $\longrightarrow_S$ -convergent, we conclude that $\longrightarrow_S$ is an $\mathcal{L}$-convergence. It follows we have enough information to define a topology.

Definition 3.2.5 A set $F \subset \mathbb{D}$ is closed in S-topology, if it contains all limits of its $\longrightarrow_S$ -convergent subsequences, i.e. if $x_n \in F$, $n = 1, 2, \ldots$, and $x_n \longrightarrow_S x_0$ then $x_0 \in F$. The convergence of sequences in S-topology will be denoted by $\overset{*}{\longrightarrow}_S$.

Remark 3.2.6 Similarly as in many other cases, $\overset{*}{\longrightarrow}_S$, being an $\mathcal{L}^*$-convergence, may be weaker than the original $\mathcal{L}$-convergence $\longrightarrow_S$. This is not a real problem in view of the Kantorovich-Kisyński recipe [16],[17]:

$$x_n \overset{*}{\longrightarrow}_S x_0 \text{ if, and only if, in every subsequence } \{n_k\} \text{ one can find a further subsequence } \{n_{k_l}\} \text{ such that } x_{n_{k_l}} \longrightarrow_S x_0. \tag{3.29}$$

In particular, relative $\overset{*}{\longrightarrow}_S$ -compactness of $K \subset \mathbb{D}$ and relative $\longrightarrow_S$ -compactness of K coincide, as well as sequential $\overset{*}{\longrightarrow}_S$ -continuity ($\equiv$ S-continuity) of $g : \mathbb{D} \to \mathbb{R}^1$ means the same as sequential $\longrightarrow_S$ -continuity of g.

The reason for our interest in the S-topology is that (3.4) and (3.5) provide criteria of relative compactness for topology S.

Lemma 3.2.7 *If (3.4) and (3.5) hold for $K \subset \mathbb{D}$ then there exists a sequence $\{x_n\} \subset K$ and $x_0 \in \mathbb{D}$ such that $x_n \longrightarrow_S x_0$.*

Conversely, if in every sequence $\{x_n\}$ one can find a subsequence $\{x_{n_k}\}$ and $x_0 \in \mathbb{D}$ such that $x_{n_k} \longrightarrow_S x_0$, then K satisfies both conditions (3.4) and (3.5).

Proof. Suppose (3.4) and (3.5) are satisfied for $K \subset \mathbb{D}$. Fix for the time being $\varepsilon > 0$ and consider the map $v_\varepsilon(x)$ defined by (3.21). By Lemma 3.2.1

$$\sup\{\|v_\varepsilon(x)\|(1) : x \in K\} < +\infty, \tag{3.30}$$

hence the set $\{v_\varepsilon(x) : x \in K\}$ is a relatively $\longrightarrow_w$ -compact subset of $\mathbb{V}$ and we can extract a sequence $\{x_{\varepsilon,n}\} \subset K$ such that $v_\varepsilon(x_{\varepsilon,n}) \longrightarrow_w v_\varepsilon$, for some $v_\varepsilon \in \mathbb{V}$.

Now let us set $\varepsilon = 1, 1/2, 1/3, \ldots$ and apply the diagonal procedure in order to find a sequence $\{x_n\} \subset K$ such that for every $m \in I\!N$

$$v_{1/m}(x_n) \underset{w}{\longrightarrow} v_{1/m}. \tag{3.31}$$

Let $\mathbb{Q} \subset [\nvdash, \nVdash]$ consists of those t for which

$$v_{1/m}(x_n)(t) \to v_{1/m}(t), \quad \text{as } n \to +\infty, m = 1, 2, \ldots. \tag{3.32}$$

We have $1 \in \mathbb{Q}$ and $[0,1] \setminus \mathbb{Q}$ is at most countable, hence $\mathbb{Q}$ is dense. In particular, for any $x \in I\!D$,

$$\|x\|_\infty = \sup_{t \in \mathbb{Q}} |x(t)|. \tag{3.33}$$

By (3.32)

$$\begin{aligned}
&|v_{1/m}(t) - v_{1/k}(t)| \\
&\quad = \lim_{n\to\infty} |v_{1/m}(x_n)(t) - v_{1/k}(x_n)(t)| \\
&\quad \le \limsup_{n\to\infty} |v_{1/m}(x_n)(t) - x_n(t)| + |x_n(t) - v_{1/k}(x_n)(t)| \\
&\quad \le \frac{1}{m} + \frac{1}{k},
\end{aligned}$$

and by (3.33)

$$\|v_{1/m} - v_{1/k}\|_\infty \le \frac{1}{m} + \frac{1}{k}.$$

It follows that $v_{1/m}$ *uniformly* converges to some $x_0 \in I\!D$ and $\|x_0 - v_{1/m}\| \le 1/m$. Hence $x_n \longrightarrow_S x_0$.

The converse part follows immediately from Lemma 3.2.8 below. □

Lemma 3.2.8 *Suppose* $x_n \overset{*}{\longrightarrow}_S x_0$. *Then*

$$\begin{aligned}
\sup_{n \in I\!N} \|x_n\|_\infty &\le C_\infty < +\infty, && (3.34)\\
\sup_{n \in I\!N} N^{a,b}(x_n) &\le C^{a,b} < +\infty,\ a < b, a, b \in I\!R^1, && (3.35)\\
\sup_{n \in I\!N} N_\eta(x_n) &\le C_\eta < +\infty,\ \eta > 0, && (3.36)\\
\sup_{n \in I\!N} \|v_\varepsilon(x_n)\|(1) &\le C_\varepsilon < +\infty,\ \varepsilon > 0. && (3.37)
\end{aligned}$$

Proof. By (3.29) we may assume $x_n \longrightarrow_S x_0$. Because of inequalities (3.15), (3.16) and (3.17), it is enough to prove (3.37) only. But by the special way of construction of $v_\varepsilon(x)$, for any $y \in I\!D$ satisfying $\|y - x\|_\infty \le \varepsilon/3$, one has

$$\|y\|(1) \ge \frac{1}{3}\|v_\varepsilon(x)\|(1). \tag{3.38}$$

If $v_{n,\varepsilon/3}$ is such that $\|x_n - v_{n,\varepsilon/3}\|_\infty \le \varepsilon/3$, $n = 0, 1, 2, \ldots$, and $v_{n,\varepsilon/3} \longrightarrow_w v_{0,\varepsilon/3}$, then

$$+\infty > 3 \sup_n \|v_{n,\varepsilon/3}\|(1) \ge \sup_n \|v_\varepsilon(x_n)\|(1). \quad \square$$

In Remark 3.2.4 we showed that $x_n \longrightarrow_S x_0$ implies pointwise convergence outside a countable set $D \subset [0,1)$. By (3.29) we have

Corollary 3.2.9 *If $x_n \stackrel{*}{\longrightarrow}_S x_0$, then in each subsequence $\{x_{n_k}\}$ one can find a further subsequence $\{x_{n_{k_l}}\}$ and a countable set $D \subset [0,1)$ such that*

$$x_{n_{k_l}}(t) \longrightarrow x_0(t), \quad t \in [0,1] \setminus D. \tag{3.39}$$

Given Corollary 3.2.9 we have lower semicontinuity of many useful functionals on $I\!D$.

Corollary 3.2.10 *If $x_n \stackrel{*}{\longrightarrow}_S x_0$ then:*

$$\begin{aligned}
\liminf_{n\in I\!N} \|x_n\|_\infty &\geq \|x_0\|_\infty, && (3.40)\\
\liminf_{n\in I\!N} N^{a,b}(x_n) &\geq N^{a,b}(x_0), \quad a<b,\ a,b\in I\!R^1, && (3.41)\\
\liminf_{n\in I\!N} N_\eta(x_n) &\geq N_\eta(x_0), \quad \eta>0, && (3.42)\\
\liminf_{n\in I\!N} \|v_\varepsilon(x_n)\|(1) &\geq \|v_\varepsilon(x_0)\|(1), \quad \varepsilon>0. && (3.43)
\end{aligned}$$

In view of (3.34) and (3.39) we have continuity for integral functionals.

Corollary 3.2.11 *Let $\Phi : [0,1] \times I\!R^1 \to I\!R^1$ be measurable and such that for each $t \in [0,1]$ $\Phi(t,\cdot)$ is continuous and for each $C > 0$*

$$\sup_{t\in[0,1]} \sup_{|x|\leq C} |\Phi(t,x)| < +\infty. \tag{3.44}$$

Let μ be an atomless *finite measure on $[0,1]$. Then the mapping*

$$I\!D \ni x \mapsto \int_{[0,1]} \Phi(t, x(t))\, d\mu(t) \in I\!R^1, \tag{3.45}$$

is S-continuous.

An important particular case is the Lebesgue measure on $[0,1]$ and

$$\Phi(t,x) = \frac{1}{\delta} 1\!\!1_{[u,u+\delta]}(t) \cdot x,$$

which gives the S-continuity of mappings

$$I\!D \ni x \mapsto x_u^\delta = \frac{1}{\delta} \int_{[u,u+\delta]} x(t)dt. \tag{3.46}$$

Since for $u < 1$

$$\lim_{\delta\searrow 0} x_u^\delta = x(u) = \pi_u(x)$$

and π_1 is S-continuous, we conclude that Borel subsets of $(I\!D, S)$ coincide with $\sigma(\pi_u : u \in [0,1])$. In addition, running u and δ over rational numbers in $[0,1]$ we get a *countable family of S-continuous functions which separate points in $I\!D$*. In particular, any *S-compact* subset of $I\!D$ is metrisable. Another useful statement implied by continuity of (3.46) is that $(I\!D, S)$ is a Hausdorff space.

It should be, however, emphasized that contrary to Skorohod's topologies, the evaluations π_u are for $u < 1$ *nowhere* S-continuous. To see this, take $x = x_0 \in I\!\!D$ and $0 \leq u < 1$ and define

$$x_n(t) = x_0(t) + 1\!\!\mathrm{I}_{[u,u+(1/n))}(t).$$

Clearly, $x_n \longrightarrow_S x_0$, but $x_n(u) = x_0(u) + 1 \not\to x_0(u)$.

The fact that $(I\!\!D, S)$ is a *linear topological space* in a trivial way follows from the very definition of convergence $\longrightarrow_S$. Since any linear topological space (with the T_1-property) is *completely regular*, so does $(I\!\!D, S)$. But S is *not metrisable*. In order to prove the latter statement, let us consider an example.

Example 3.2.12 Let for $m, n \in I\!\!N$ $x_{m,n}(t) = m1\!\!\mathrm{I}_{[1/2,1/2+1/n)}(t)$, $t \in [0,1]$. Then for each fixed $m \in I\!\!N$ we have $x_{m,n} \longrightarrow_S 0$, as $n \to +\infty$. Let $n_m \to \infty$, as $m \to \infty$. The sequence $\{x_{m,n_m}\}_{m \in I\!\!N}$ does not contain any $\stackrel{*}{\longrightarrow}_S$ -convergent subsequence, for $\liminf_{m\to\infty} \|x_{m,n_m}\|_\infty = +\infty$. And for any convergence generated by a metric there should exist a sequence $\{n_m\}$ such that $x_{m,n_m} \stackrel{*}{\longrightarrow}_S 0$.

We summarize all obtained results in

Theorem 3.2.13 *The Skorohod space $I\!\!D$ equipped with the sequential topology S is a linear Hausdorff topological space which cannot be metricized. Moreover:*

(i) *There exists a countable family of S-continuous functions which separate points in $I\!\!D$.*

(ii) *Compact subsets $K \subset I\!\!D$ are metrisable.*

(iii) *A subset $K \subset I\!\!D$ is relatively S-compact if either of equivalent sets of conditions (3.10)+(3.11), (3.10)+(3.12) and (3.13)+(3.14) is satisfied.*

(iv) *S-Borel subsets $\mathcal{B}_S$ coincide with the standard σ-algebra generated by evaluations (projections).*

We conclude this section with two additional properties of the S-topology.

We have observed that the evaluations at $t < 1$ are nowhere continuous. It is therefore interesting that the evaluations still can be used for identification of the limit.

Proposition 3.2.14 *Let $\mathbb{Q} \subset [\nvdash,\nVdash]$, $1 \in \mathbb{Q}$, be dense. Suppose $\{x_n\}$ is relatively S-compact and*

$$x_n(q) \to x_0(q), \quad \text{as } n \to +\infty,\ q \in \mathbb{Q}. \tag{3.47}$$

Then $x_n \stackrel{*}{\longrightarrow}_S x_0$.

Proof. Suppose along a subsequence $\{n_k\}$ we have $x_{n_k} \longrightarrow_S y_0$. Then for q' in some dense subset $\mathbb{Q}' \subset [\nvdash,\nVdash)$, $x_n(q') \to y_0(q')$. If $y_0 \neq x_0$, then for some $\eta > 0$ and $(u,v) \subset [0,1)$ one has $|y_0(t) - x_0(t)| > \eta$ for $t \in (u,v)$. Let $u < t_1 < t_2 < \ldots < v$ be such that:

1. $t_1, t_3, \ldots, t_{2m-1}, \ldots \in \mathbb{Q}$,

2. $t_2, t_4, \ldots, t_{2m}, \ldots \in \mathbb{Q}'$.

Then for each $m \in I\!\!N$ $N_{\eta/2}(x_{n_k}) \geq 2m$ for k large enough, hence $N_{\eta/2}(x_{n_k}) \to +\infty$ and $\{x_{n_k}\}$ cannot be S-convergent. Hence $y_0 = x_0$. □

Clearly, not all topologies on $I\!\!D$ possess the property investigated in Proposition 3.2.14. For example, it is easy to find a sequence $\{x^n\} \subset I\!\!D$ which converges in measure to $x^0 \equiv 0$ and is such that $x^n(q) \to 1$ for each rational $q \in [0,1]$. Hence Proposition 3.2.14 is not valid for the "pseudo-path" topology.

Another interesting feature of the S-topology is the continuity of the smoothing operation $I\!\!D \ni x \mapsto s_\mu(x) \in C([0,1] : I\!\!R^1)$, where

$$s_\mu(x)(t) = \int_0^t x(s)\, d\mu(s), \tag{3.48}$$

and μ is an atomless finite measure on $[0,1]$ (e.g. given by a density $p_\mu(s)$).

Proposition 3.2.15 *The operation $s_\mu : (I\!\!D, S) \to (C, \|\cdot\|_\infty)$ is continuous.*

Proof. Suppose $x_n \longrightarrow_S x_0$. Corollary 3.2.11 shows that for every $t \in [0,1]$ $s_\mu(x_n)(t) \to s_\mu(x_0)(t)$. Hence only relative compactness of $\{s_\mu(x_n)\} \subset C([0,1] : I\!\!R^1)$ has to be verified. But this is straightforward:

$$\sup_n \sup_{t<u<t+\delta} |s_\mu(x_n)(u) - s_\mu(x_n)(t)| \leq \sup_n \|x_n\|_\infty \times \sup_{t\in[0,1]} \mu([t, t+\delta]) \to 0,$$

when $\delta \to 0$. □

3.3 Convergence in distribution on $(I\!\!D, S)$

In the previous section we checked the equality $\mathcal{B}_S = \sigma\{\pi_t : t \in [0,1]\}$. It follows that every probability measure on $(I\!\!D, \mathcal{B}_S)$ is *tight*. Let us denote the set of such measures by $\mathcal{P}(I\!\!D)$. Further, the notions "random element in $(I\!\!D, \mathcal{B}_S)$" and "stochastic process with trajectories in $I\!\!D$" are synonymous and we see that the theory developed in Section 3.2 applies to usual objects.

It is a nice feature of Lemma 3.2.7 that we have

Proposition 3.3.1 *A family of stochastic processes $\{X_\alpha\}$ with trajectories in $I\!\!D$ is uniformly S-tight if, and only if, the family of random variables $\{\|X_\alpha\|_\infty\}$ is uniformly tight and for each pair $a < b$, $a, b \in I\!\!R^1$, the family of random variables $\{N^{a,b}(X_\alpha)\}$ is uniformly tight.* □

Since S-compact subsets of $I\!\!D$ are metrisable, uniform S-tightness of $\mathcal{K} \subset \mathcal{P}(I\!\!D)$ implies relative compactness of $\mathcal{K}$ with respect to the weak-$*$ topology on $\mathcal{P}(I\!\!D)$ (the direct Prohorov's theorem). The converse, however, is not clear at all. And we are not going to investigate this question. Instead we shall equip $\mathcal{P}(I\!\!D)$ with a sequential topology induced by S, which is finer than the weak-$*$ topology, for which both the direct and the converse Prohorov's theorem hold and which is very close to the a.s. Skorohod representation. The advantage of such an approach is evident: we have effective tools (Prohorov's theorems, the a.s. Skorohod representation), more (in general) continuous functionals and if the converse Prohorov's theorem is valid for the weak-$*$ topology, then the convergence of sequences in both topologies coincides (see Paper II).

The advertized "new" convergence in distribution will be denoted by $\xrightarrow{*}_{\mathcal{D}}$ and $X_n \xrightarrow{*}_{\mathcal{D}} X_0$ will mean the following:

Definition 3.3.2 In every subsequence $\{X_{n_k}\}$ one can find a further subsequence $\{X_{n_{k_l}}\}$ and stochastic processes $\{Y_l\}$ defined on $([0,1], \mathcal{B}_{[0,1]}, \ell)$ such that

$$Y_l \sim X_{n_{k_l}}, \quad l = 1, 2, \ldots, \tag{3.49}$$

for each $\omega \in [0,1]$

$$Y_l(\omega) \ \xrightarrow[S]{*} \ Y_0(\omega), \quad \text{as } l \to \infty, \tag{3.50}$$

and for each $\varepsilon >$ there exists an S-compact subset $K_\varepsilon \subset I\!D$ such that

$$P(\{\omega \in [0,1] : Y_l(\omega) \in K_\varepsilon, l = 1, 2, \ldots\}) > 1 - \varepsilon. \tag{3.51}$$

One can say that (3.50) and (3.51) describe "the almost sure convergence in compacts" and that (3.49), (3.50) and (3.51) define the *strong* a.s. Skorohod representation *for subsequences* ("strong" because of condition (3.51)).

We shall write $\mu_n \overset{*}{\Longrightarrow} \mu_0$ whenever $\mu_n = \mathcal{L}(X_n)$ and $X_n \ \xrightarrow{*}_{\mathcal{D}} \ X_0$.

It follows from Lemma 3.2.7 and Corollary 3.2.10 that on $(I\!D, S)$ there exists a countable family of lower continuous functionals $\{\xi_i\}_{i \in I\!I}$ such that $K \subset I\!D$ is conditionally compact iff each ξ_i is bounded on K, $i \in I\!I$. Therefore we can apply Theorem 2.4.2 of Paper II to obtain the direct and converse Prohorov's theorems for $\mathcal{O}_S(\overset{*}{\Longrightarrow})$.

Theorem 3.3.3 *Let $\{X_\alpha\}$ be a family of stochastic processes with trajectories in $I\!D$. The following statements are equivalent.*

- (i) *$\{X_\alpha\}$ is relatively compact with respect to " $\xrightarrow{*}_{\mathcal{D}}$ " on $(I\!D, S)$.*
- (ii) *$\{X_\alpha\}$ is uniformly S-tight.*
- (iii) *$\{\|X_\alpha\|_\infty\}$ is a uniformly tight family as well as for each $a < b$ $\{N^{a,b}(X_\alpha)\}$ is uniformly tight.*
- (iv) *$\{\|X_\alpha\|_\infty\}$ is uniformly tight and for each $\eta > 0$ $\{N_\eta(X_\alpha)\}$ is uniformly tight.*
- (v) *For each $\varepsilon > 0$ the family $\{\|v_\varepsilon(X_\alpha)\|(1)\}$ is uniformly tight (where v_ε is defined by (3.21)).*

We have shown the topology $\mathcal{O}_S(\overset{*}{\Longrightarrow})$ induced by S is as good as the weak-$*$-star topology on $\mathcal{P}(I\!D, J_1)$, at least from the point of view of proving limit theorems. Theorem 3.3.3 does not end the list of (formal) similarities between the mentioned topologies. We shall point three other properties of the topology $\mathcal{O}_S(\overset{*}{\Longrightarrow})$ which belong to standard tools of limit theory. It will be proved first that convergence of finite dimensional distributions can be used for identification of limits in $\overset{*}{\Longrightarrow}$-convergence (although the projections are nowhere continuous).

Theorem 3.3.4 *Let $\mathbb{Q} \subset [\nvdash, \nVdash]$ be dense, $1 \in \mathbb{Q}$. Suppose that for each finite subset $\mathbb{Q}_\nvdash = \{\shortparallel_\nvdash < \shortparallel_\nvDash < \ldots < \shortparallel_\gtrdot\} \subset \mathbb{Q}$ we have as $n \to \infty$*

$$(X_n(q_1), X_n(q_2), \ldots, X_n(q_m)) \ \xrightarrow[\mathcal{D}]{} \ (X_0(q_1), X_0(q_2), \ldots, X_0(q_m)), \tag{3.52}$$

where X_0 is a stochastic process with trajectories in $I\!D$.

If $\{X_n\}$ is relatively compact with respect to $\xrightarrow{*}_{\mathcal{D}}$, then $X_n \ \xrightarrow{*}_{\mathcal{D}} \ X_0$.

PROOF. Let $\psi_0(x) = \|x\|_\infty$ and let $\psi_i(x) = N_{1/i}(x)$, $i = 1, 2, \ldots$. Define also $\Psi(x) = (\psi_i(x))_{i=0,1,2,\ldots} \in I\!\!R^\infty$ and $\Phi(x) = (x(q))_{q\in\mathbb{Q}} \in I\!\!R^{\mathbb{Q}}$.

By Theorem 3.3.3 $\{\Psi(X_n)\}$ is uniformly tight in $I\!\!R^\infty$. Hence at least along some subsequence we have in $I\!\!R^\infty \times I\!\!R^{\mathbb{Q}}$

$$(\Phi(X_n), \Psi(X_n)) \underset{\mathcal{D}}{\longrightarrow} (\Phi(X_0), Z). \tag{3.53}$$

By Skorohod's theorem on the a.s. representation there exist random elements (U_n, V_n), $n = 0, 1, 2, \ldots$, defined on $([0, 1], \mathcal{B}_{[0,1]}, \ell)$ and such that

$$(U_n, V_n) \sim (\Phi(X_n), \Psi(X_n)), \quad n = 1, 2, \ldots, \tag{3.54}$$

$$(U_0, V_0) \sim (\Phi(X_0), Z), \tag{3.55}$$

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

$$U_n(\omega) \to U_0(\omega) \text{ in } I\!\!R^\infty,\ V_n(\omega) \to V_0(\omega) \text{ in } I\!\!R^{\mathbb{Q}}. \tag{3.56}$$

We claim that

$$\begin{array}{l}\text{there exists a measurable mapping } \Theta : I\!\!R^\infty \to I\!\!D \text{ such that } \Theta(U_n) \sim X_n,\\ n = 0, 1, 2, \ldots, \text{ and } V_n = \Psi \circ \Theta(U_n),\ \ell\text{-a.s., } n = 1, 2, \ldots.\end{array} \tag{3.57}$$

Φ is a measurable and one-to-one mapping from $I\!\!D$ into $I\!\!R^\infty$. We know S-compact subsets of $I\!\!D$ are metrisable, hence Φ maps each S-compact subset K onto a *Borel* subset $\Phi(K)$ of $I\!\!R^\infty$. The same holds also for any σ-compact subset of $I\!\!D$. Since laws of all $X_0, X_1, X_2, \ldots$ are S-tight, we can find a common σ-compact support K_0, i.e. $P(X_n \in K_0) = 1$, $n = 0, 1, 2, \ldots$. If we set $\Theta(y) = \Phi^{-1}(y)$ for $y \in \Phi(K_0)$ and $\Theta(y) = 0$ outside $\Phi(K_0)$, then $X_n = \Theta(\Phi(X_n))$ almost surely on the probability space where X_n is defined and $U_n = \Phi(\Theta(U_n))$ ℓ-a.s. We have also $\Psi(X_n) = \Psi(\Theta(\Phi(X_n)))$ almost surely and so $V_n = \Psi(\Theta(U_n))$ ℓ-a.s. by (3.54).

Choose $\omega \in [0, 1]$ in a "good" subset of full measure and consider $\{x_n = \Theta(U_n(\omega))\}_{n\in I\!\!N} \subset I\!\!D$. We have $x_n(q) \to x_0(q)$, $q \in \mathbb{Q}$, and for each $i = 0, 1, 2, \ldots$ $\psi_i(x_n) \to (V_0(\omega))_i$, hence $\sup_n \psi_i(x_n) < +\infty$. By Proposition 3.2.14 $x_n \xrightarrow{*}_S x_0$. Hence $\Theta(U_n)$ is the a.s. Skorohod representation for X_n (in fact, for a subsequence of X_n) and so $X_n \xrightarrow{*}_{\mathcal{D}} X_0$. □

Only a slightly modified proof give us

Theorem 3.3.5 *Let $\mathbb{Q} \subset [\nvdash, \nVdash]$ be dense, $1 \in \mathbb{Q}$. Suppose $\{X_n\}$ is a uniformly S-tight sequence and that for each finite sequence $q_1 < q_2 < \ldots < q_m$ of elements of $\mathbb{Q}$ we have*

$$(X_n(q_1), X_n(q_2), \ldots, X_n(q_m)) \underset{\mathcal{D}}{\longrightarrow} \tilde{\nu}_{(q_1,q_2,\ldots,q_m)}, \tag{3.58}$$

where $\tilde{\nu}_{(q_1,q_2,\ldots,q_m)}$ is a probability distribution on $I\!\!R^m$. Then for each finite sequence $t_1 < t_2 < \ldots < t_m$ and each approximating sequence $q_{1,l} < q_{2,l} < \ldots < q_{m,l}$, $q_{j,l} \searrow t_j$, as $l \to \infty$, ($q_{m,l} = 1$ if $t_m = 1$), probability distributions $\tilde{\nu}_{(q_{1,l},q_{2,l},\ldots,q_{m,l})}$ weakly converge to some limit $\nu_{(t_1,t_2,\ldots,t_m)}$. Moreover, there is a stochastic process X_0 with trajectories in $I\!\!D$ such that for every finite sequence $t_1 < t_2 < \ldots < t_m$

$$(X_0(t_1), X_0(t_2), \ldots, X_0(t_m)) \sim \nu_{(t_1,t_2,\ldots,t_m)}, \tag{3.59}$$

and $X_{n_k} \xrightarrow{*}_{\mathcal{D}} X_0$.

Remark 3.3.6 The Skorohod J_1-topology also possesses the properties described in Theorems 3.3.4 and 3.3.5 (see [36]).

An especially convenient tool for investigations of the topology $\mathcal{O}_S(\overset{*}{\Longrightarrow})$ is provided by the following decomposition.

Theorem 3.3.7 *A family $\{X_\alpha\}$ of stochastic processes with trajectories in $I\!D$ is uniformly S-tight if, and only if, for each $\varepsilon > 0$ we can decompose processes X_α in the following way:*

$$X_\alpha(t) = R_{\alpha,\varepsilon}(t) + U_{\alpha,\varepsilon}(t) - V_{\alpha,\varepsilon}(t), \quad t \in [0,1], \tag{3.60}$$

where all processes $R_{\alpha,\varepsilon}$, $U_{\alpha,\varepsilon}$ and $V_{\alpha,\varepsilon}$ are adapted to the natural right-continuous filtration generated by X_α, trajectories of $R_{\alpha,\varepsilon}$ are uniformly small:

$$\|R_{\alpha,\varepsilon}\|_\infty \le \varepsilon, \tag{3.61}$$

$U_{\alpha,\varepsilon}$ and $V_{\alpha,\varepsilon}$ are nonnegative and nondecreasing (i.e. have paths in $\mathbb{V}^+$) and both $\{U_{\alpha,\varepsilon}(1)\}$ and $\{V_{\alpha,\varepsilon}(1)\}$ are uniformly tight families of random variables.

Proof. By Theorem 3.3.5, the family $\{X_\alpha\}$ is uniformly S-tight iff for each $\varepsilon > 0$ the family of random variables $\{\|v_\varepsilon(X_\alpha)\|(1)\}$ is uniformly tight, where processes $v_\varepsilon(X_\alpha)(t)$ are defined by (3.21) and are adapted by Corollary 3.2.2. Hence it is enough to set $R_\alpha(t) = X_\alpha(t) - v_\varepsilon(X_\alpha)(t)$ and decompose $v_\varepsilon(X_\alpha)(t)$ into a difference of two increasing processes $U_{\alpha,\varepsilon}(t)$ and $V_{\alpha,\varepsilon}(t)$ in such a way that

$$\|v_\varepsilon(X_\alpha)\|(1) = U_{\alpha,\varepsilon}(1) + V_{\alpha,\varepsilon}(1). \ \square$$

Remark 3.3.8 It follows from (3.60) that any fact on increasing processes may contribute to the knowledge of S-topology. The very close relation between the weak topology on $\mathbb{V}^+$ and the S-topology on $I\!D$ becomes clear, if we realize that the topology S induces on $\mathbb{V}^+$ exactly the topology of weak convergence (notice, however, that on $\mathbb{V} = \mathbb{V}^+ - \mathbb{V}^+$ the topology S is weaker!). So we can consider the S-topology on $I\!D$ as a natural extension of the notion of weak convergence of elements of $\mathbb{V}^+$.

Theorem 3.3.9 *Let $\{V_\alpha\}$ be a family of stochastic processes with trajectories in $\mathbb{V}^+$. Suppose that $\{V_\alpha(1)\}$ is uniformly tight. Then there exists a sequence $\{V_n\} \subset \{V_\alpha\}$, an increasing process V_0 and a countable subset $D \subset [0,1)$ such that for all finite sets $\mathbb{Q}_m = \{q_1 < q_2 < \ldots < q_m\} \subset [0,1] \setminus \mathbb{D}$*

$$(V_n(q_1), V_n(q_2), \ldots, V_n(q_m)) \underset{\mathcal{D}}{\longrightarrow} (V_0(q_1), V_0(q_2), \ldots, V_0(q_m)). \tag{3.62}$$

Proof. Let μ_α be a random measure on $([0,1], \mathcal{B}_{[0,1]})$ given by formula

$$\mu_\alpha([0,t],\omega) = \frac{V_\alpha(t,\omega)}{1 + V_\alpha(1,\omega)}, \quad t \in [0,1]. \tag{3.63}$$

Since μ_α takes values in the space $\mathcal{M}_{\le 1} = \mathcal{M}_{\le 1}([0,1])$ of measures on compact space $[0,1]$ with total mass smaller than 1, we can extract a sequence μ_n such that on the space $\mathcal{M}_{\le 1} \times I\!R^+$

$$(\mu_n, V_n(1)) \underset{\mathcal{D}}{\longrightarrow} (\mu_0, Z_0). \tag{3.64}$$

Since $\mathcal{M}_{\leq 1} \times \mathbb{R}^+$ is a Polish space we may assume without loss of generality that almost surely

$$\mu_n(\cdot, \omega) \Longrightarrow \mu_0(\cdot, \omega), \quad V_n(1, \omega) \to Z_0(\omega). \tag{3.65}$$

Let us consider the mapping $[0,1] \ni t \mapsto E\mu_0([0,t], \omega) = (E\mu_0)([0,t])$. Outside a countable set $D \subset [0,1)$ we have $(E\mu_0)(\{t\}) = 0$, hence $\mu_0(\{t\}, \omega) = 0$ for almost all ω, and so by (3.65) and for $t \notin D$

$$\mu_n([0,t], \omega) \to \mu_0([0,t], \omega), \tag{3.66}$$

or

$$V_n(t, \omega) \to (1 + Z_0(\omega))\mu_0([0,t], \omega) =: V_0(t, \omega). \tag{3.67}$$

The theorem has been proved. □

Now we are ready to prove a corresponding result for $(\mathbb{D}, S)$ which (in some sense) improves Theorem 5 of [23], where finite dimensional convergence outside a set of Lebesgue measure 0 was obtained (but only convergence in distribution with respect to the "pseudo-path topology" was assumed).

Theorem 3.3.10 *Let $\{X_\alpha\}$ be a uniformly S-tight family of stochastic processes with trajectories in $\mathbb{D}$. Then there exists a sequence $\{X_n\} \subset \{X_\alpha\}$, a process X_0 with trajectories in $\mathbb{D}$ and a countable subset $D \subset [0,1)$ such that for all finite sets $\mathbb{Q}_0 = \{q_1 < q_2 < \ldots < q_m\} \subset [0,1] \setminus \mathbb{D}$*

$$(X_n(q_1), X_n(q_2), \ldots, X_n(q_m)) \underset{\mathcal{D}}{\longrightarrow} (X_0(q_1), X_0(q_2), \ldots, X_0(q_m)). \tag{3.68}$$

In particular, $X_n \xrightarrow{*}_{\mathcal{D}} X_0$.

Proof. According to (3.60), for each $m \in \mathbb{N}$ we can decompose

$$X_\alpha = R_{\alpha,1/m} + U_{\alpha,1/m} - V_{\alpha,1/m}$$

with $U_{\alpha,1/m}$ and $V_{\alpha,1/m}$ uniformly tight on $(\mathbb{V}^+, \mathbb{S})$. By tightness assumptions one can find a subsequence such that on the space $(\mathbb{D}, S) \times (\mathbb{V}^+, \mathbb{S})^{\mathbb{N}}$

$$\begin{aligned} &(X_n, U_{n,1}, V_{n,1}, \ldots, U_{n,1/m}, V_{n,1/m}, \ldots) \\ &\qquad \underset{\mathcal{D}}{\xrightarrow{*}} (X_0, U_1, V_1, \ldots, U_m, V_m, \ldots). \end{aligned}$$

It means that passing again to a subsequence we have the a.s. Skorohod representation, i.e. without loss of generality we may assume that

$$\begin{aligned} X_n(\omega) &\underset{S}{\xrightarrow{*}} X_0(\omega) \quad \text{in} \quad \mathbb{D}, \\ U_{n,1/m}(\omega) &\underset{S}{\xrightarrow{*}} U_m(\omega) \quad \text{in} \quad \mathbb{V}^+, \quad m \in \mathbb{N}, \\ V_{n,1/m}(\omega) &\underset{S}{\xrightarrow{*}} V_m(\omega) \quad \text{in} \quad \mathbb{V}^+, \quad m \in \mathbb{N}. \end{aligned}$$

Since $X_n(\omega) \xrightarrow{*}_S X_0(\omega)$ and because of (3.60) we have

$$\|U_m(\omega) - V_m(\omega) - X_0(\omega)\|_\infty \leq 2/m$$

and so for every finite subset $\{t_1 < t_2 < \ldots < t_r\} \subset [0,1]$ we have as $m \to \infty$

$$\begin{aligned}(U_m(t_1) - V_m(t_1), U_m(t_2) - V_m(t_2), \ldots, U_m(t_r) - V_m(t_r)) \\ \longrightarrow_{\mathcal{D}} \quad (X_0(t_1), X_0(t_2), \ldots, X_0(t_r)).\end{aligned}$$

On the other hand, applying Theorem 3.3.9 we see that as $n \to \infty$

$$\begin{aligned}(U_{n,1/m}(q_1), V_{n,1/m}(q_1), U_{n,1/m}(q_2), V_{n,1/m}(q_2), \ldots, U_{n,1/m}(q_r), V_{n,1/m}(q_r)) \\ \longrightarrow_{\mathcal{D}} \quad (U_m(q_1), V_m(q_1), U_m(q_2), V_m(q_2), \ldots, U_m(q_r), V_m(q_r))\end{aligned}$$

for every finite subset $\mathbb{Q}_m = \{q_1 < q_2 < \ldots < q_r\} \subset [0,1]$ outside a countable subset $D_m \subset [0,1)$. Hence outside $D = D_1 \cup D_2 \cup \ldots$ we have (3.68). $\square$

3.4 Uniform S-tightness and semimartingales

Theorem 3.3.3 provides several sets of conditions equivalent to uniform S-tightness. Inequalities (3.16) and (3.17) suggest the easiest way of proving uniform S-tightness: one has to check whether families $\{\|X_\alpha\|_\infty\}$ and $\{N^{a,b}(X_\alpha)\}$, for each $a < b$, are bounded in probability. For example, if X is a supermartingale, then one can use the classical Doob's inequalities (see [6], Ch. VI):

$$P(\sup_{t\in[0,1]} |X(t)| \geq \lambda) \leq 3\lambda^{-1} \sup_{t\in[0,1]} E|X(t)|, \tag{3.69}$$

$$EN^{a,b}(X) \leq \frac{1}{b-a}(|a| + \sup_{t\in[0,1]} E|X(t)|). \tag{3.70}$$

It follows immediately that any sequence $\{X_n\}$ of supermartingales satisfying

$$\sup_{t\in[0,1]} E|X_n(t)| < +\infty, \tag{3.71}$$

is uniformly S-tight. The most general result of such kind concerns semimartingales and belongs to Stricker [33]. We shall restate it using terminology of the paper [14] and the setting of S-topology.

Let $\{X_\alpha\}$ be a family of stochastic processes with trajectories in $\mathbb{D}$, with X_α defined on the stochastic basis $(\Omega^\alpha, \mathcal{F}^\alpha, \{\mathcal{F}^\alpha_t\}_{t\in[0,1]}, P^\alpha)$ and adapted to filtration $\{\mathcal{F}^\alpha_t\}_{t\in[0,1]}$. We say that *Condition UT* holds for $\{X_\alpha\}$, if the family of elementary stochastic integrals $\{\int H^\alpha_- \, dX_\alpha(1)\}$ with integrands bounded by 1 is uniformly tight. (By an elementary stochastic integral with integrands bounded by 1 we mean random variable of the form

$$\sum_{i=1}^m H^\alpha(t_{i-1}) \left(X_\alpha(t_i) - X_\alpha(t_{i-1})\right),$$

where $m \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_m = T$, $H^\alpha(t_i) \leq 1$ and $H^\alpha(t_i)$ is $\mathcal{F}^\alpha(t_i)$-measurable for $i = 0, 1, \ldots, m$. Condition UT was considered for the first time in [33]. The reader may find conditions which follow the line of (3.71) and are sufficient for Condition UT in [14]. For equivalent reformulations in terms of predictable characteristics we refer to [21]. Here we stress the fact that the family consisting of a single process X satisfies Condition UT if, and only if, X is a semimartingale (see e.g. [27]). Therefore in what follows we shall deal with semimartingales only.

The essential step in the proof of Theorem 2 of [33] gives us

Theorem 3.4.1 *Condition UT implies uniform S-tightness.*

By Theorem 3.3.10 any set $\{X_\alpha\}$ of semimartingales satisfying Condition UT contains a sequence $X_n \xrightarrow{*}_{\mathcal{D}} X_0$ for which finite dimensional convergence holds outside a countable set $D \subset [0,1)$: for every finite subset $\{0 \leq q_1 < q_2 < \ldots < q_m \leq 1\} \subset \mathbb{Q} = [\nvdash, \nVdash] \setminus \mathbb{D}$

$$(X^n(q_1), X^n(q_2), \ldots, X^n(q_m)) \underset{\mathcal{D}}{\longrightarrow} (X^0(q_1), X^0(q_2), \ldots, X^0(q_m)) \tag{3.72}$$

on the space $I\!R^m$. If (3.72) is satisfied for every finite subset $\{0 \leq q_1 < q_2 < \ldots < q_m \leq 1\}$ of some set $\mathbb{Q}$, we shall write $X_n \longrightarrow_{\mathcal{D}_f(\mathbb{Q})} X_0$.

Since $\mathbb{Q} = [\nvdash, \nVdash] \setminus \mathbb{D}$ is dense in $[0,1]$ we can apply Lemma 1.3 of [14] and conclude that X_0 is a semimartingale. Let us observe that the pseudo-path topology gives the finite dimensional convergence over a set of full Lebesgue measure only and so our methods give an improvement of Theorem 2 in [33].

Theorem 3.4.2 *Suppose Condition UT holds for a family of semimartingales $\{X_\alpha\}$. Then there exists a sequence $\{X_n\} \subset \{X_\alpha\}$ and a semimartingale X_0 such that $X_n \xrightarrow{*}_{\mathcal{D}} X_0$ and $X_n \longrightarrow_{\mathcal{D}_f(\mathbb{Q})} X_0$, where $1 \in \mathbb{Q}$ and the complement of $\mathbb{Q}$ in $[0,1]$ is at most countable.*

The two subsequent results are based on Theorems 3.3.4 and 3.3.5 and *are not valid* for the pseudo-path topology.

Theorem 3.4.3 *Suppose $\mathbb{Q} \subset [\nvdash, \nVdash]$ is dense, $1 \in \mathbb{Q}$ and $X_n \longrightarrow_{\mathcal{D}_f(\mathbb{Q})} X_0$, where X_0 has trajectories in $I\!D$.*

If Condition UT is satisfied for $\{X_n\}$, then X_0 is a semimartingale and $X_n \xrightarrow{*}_{\mathcal{D}} X_0$.

Theorem 3.4.4 *Let $\mathbb{Q}$ be as in Theorem 3.4.3. Suppose for each finite sequence $q_1 < q_2 < \ldots < q_m$ of elements of $\mathbb{Q}$ we have*

$$(X_n(q_1), X_n(q_2), \ldots, X_n(q_m)) \underset{\mathcal{D}}{\longrightarrow} \tilde{\nu}_{(q_1, q_2, \ldots, q_m)}, \tag{3.73}$$

where $\tilde{\nu}_{(q_1, q_2, \ldots, q_m)}$ is a probability distribution on $I\!R^m$.

If Condition UT holds for $\{X_n\}$, then there exists a semimartingale X_0 such that $X_n \xrightarrow{*}_{\mathcal{D}} X_0$. Moreover, for each finite sequence $t_1 < t_2 < \ldots < t_m$ and each approximating sequence $q_{1,l} < q_{2,l} < \ldots < q_{m,l}$, $q_{j,l} \searrow t_j$, as $l \to \infty$, ($q_{m,l} = 1$ if $t_m = 1$), probability distributions $\tilde{\nu}_{(q_{1,l}, q_{2,l}, \ldots, q_{m,l})}$ weakly converge to the joint distribution of $(X_0(q_1), X_0(q_2), \ldots, X_0(q_m))$.

Eventually, let us notice that the topology S arises in a quite natural manner in limit theorems for the Ito stochastic integrals (see Paper IV).

Paper No 4.

Convergence in various topologies for stochastic integrals driven by semimartingales

Abstract

We generalize existing limit theory for stochastic integrals driven by semimartingales and with left-continuous integrands. Joint Skorohod convergence is replaced with joint finite dimensional convergence *plus* assumption excluding the case when oscillations of the integrand appear immediately before oscillations of the integrator. Integrands may converge in a very weak topology. It is also proved that convergence of integrators implies convergence of stochastic integrals with respect to the same topology.

4.1 Introduction

Let us begin with a simple example demonstrating one of central difficulties in limit theory for integrals with discontinuous integrators.

Example 4.1.0 Let $k^0(t) = \mathbb{I}_{[1/2,1]}(t)$, $k^n(t) = \mathbb{I}_{[1/2-1/n,1]}(t)$, $n = 1, 2, \ldots$ and let $x^n(t) = \mathbb{I}_{[1/2,1]}(t)$, $n = 0, 1, 2, \ldots$. Then $k^n \to k^0$ and $x^n \to x^0$ in the Skorohod space $\mathbb{D} = \mathbb{D}([0,1] : \mathbb{R}^1)$, but

$$\int k^n_- \, dx^n \equiv x^0 \not\to \int k^0_- \, dx^0 \equiv 0.$$

(Here—and in the sequel—we consider integrals over the interval excluding 0, i.e.

$$(\int k_- \, dx)(t) = \int_{]0,t]} k(s-) \, dx(s),$$

with $k(0-) = 0$).

One can eliminate such pathological situation by assuming *joint* convergence of (k^n, x^n), i.e. convergence in $\mathbb{D}([0,1] : \mathbb{R}^2)$. A very general result in this direction was proved in [14, Theorem 2.6].

Theorem 4.1.1 *For each $n \in \mathbb{N}$, let X^n be a semimartingale with respect to the stochastic basis $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}^n_t\}_{t\in[0,1]}, P^n)$ and let K^n be adapted to $\{\mathcal{F}^n_t\}_{t\in[0,1]}$ and with trajectories in $\mathbb{D}$. Assume that*

$$(K^n, X^n) \underset{\mathcal{D}}{\longrightarrow} (K^0, X^0) \tag{4.1}$$

on the space $\mathbb{D}([0,1]:\mathbb{R}^2)$*. Then* X^0 *is a semimartingale with respect to the natural filtration generated by* (K^0, X^0) *and*

$$\int K^n_- \, dX^n \underset{\mathcal{D}}{\longrightarrow} \int K^0_- \, dX^0 \text{ on } \mathbb{D}([0,1]:\mathbb{R}^1), \tag{4.2}$$

as well as

$$(K^n, X^n, \int K^n_- \, dX^n) \underset{\mathcal{D}}{\longrightarrow} (K^0, X^0, \int K^0_- \, dX^0) \text{ on } \mathbb{D}([0,1]:\mathbb{R}^3), \tag{4.3}$$

provided *so called* Condition UT *holds, i.e. the family of elementary stochastic integrals* $\{\int H^n_- \, dX^n(1)\}$ *with integrands bounded by* 1 *is uniformly tight.*

To be explicit, Condition UT means that the family of all random variables of the form

$$\sum_{i=1}^{m} H^n_{t_{i-1}} \left(X^n_{t_i} - X^n_{t_{i-1}} \right)$$

is uniformly tight, where $m \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_m = T$, $H^n_{t_i} \le 1$ and $H^n_{t_i}$ is $\mathcal{F}^n_{t_i}$-measurable for $i = 0, 1, \ldots, m$. Condition UT was considered for the first time in [33]. The reader may find sufficient conditions for Condition UT in [14] and equivalent reformulations in [21]. Here we shall mention only that Condition UT plays also a crucial role in approximation of solutions of stochastic differential equations. For the corresponding results in this area as well as for interesting examples we refer to [30], [31] and [19].

Theorem 4.1.1 suffices for most applications related to stability problems of stochastic differential equations. On the other hand, within limit theory for stochastic integrals there exist phenomena which are not covered by this theorem.

Example 4.1.2 Normalized sums of moving averages with summable *positive* coefficients of i.i.d. random variables with laws belonging to domain of attraction of an α-stable law ($\alpha < 2$) in general *do not converge* in functional manner when $\mathbb{D}$ is equipped with the usual Skorohod's J_1 topology. But they *do converge* to an α-stable Lévy's motion if we consider another, weaker topology on $\mathbb{D}$, known as M_1 (see [1], also [29] for definitions of Skorohod's topologies).

There exists a satisfactory theory of stochastic integration with respect to α-stable processes (see e.g. [15]). It follows that for some naturally arising integrators the requirement of convergence in the usual Skorohod topology may be too strong.

Example 4.1.3 Let, as in Example 4.1.0, $x^n(t) = \mathbb{1}_{[1/2,1]}(t)$, $n = 0, 1, 2, \ldots$ and $k^0(t) = \mathbb{1}_{[1/2,1]}(t)$. The difference will be in the choice of k^n:

$$k^n(t) = \mathbb{1}_{[1/2+1/n,1]}(t), \quad n = 1, 2, \ldots.$$

As before we have $k^n \to k^0$ and $x^n \to x^0$ in $\mathbb{D}$ with the standard topology and $(k^n, x^n) \not\to (k^0, x^0)$ in $\mathbb{D}([0,1]:\mathbb{R}^2)$. But this time

$$\int k^n_- \, dx^n \equiv 0 \to \int k^0_- \, dx^0 \equiv 0.$$

Example 4.1.4 The preceding example may seem to be artificial and related to the extremely simple structure of involved processes. To convince the reader the example illustrates a general rule, let us consider a much less obvious fact.

Let X be a semimartingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,1]}, P)$ and let K be adapted to $\{\mathcal{F}_t\}_{t\in[0,1]}$ and with trajectories in $I\!\!D$. Choose a partition $\tau = \{0 = t_0 < t_1 < t_2 < \ldots < t_m = 1\}$ of $[0,1]$ and define τ-discretizations of the integrand K:

$$K^\tau(t) = K(t_k), \text{ if } t_k \le t < t_{k+1},\ k = 0, 1, \ldots, m-1,\ K^\tau(1) = K(1). \tag{4.4}$$

Then it follows from the "dominated convergence" theorem (see [6, VIII.14]) that the elementary stochastic integrals $\int K^\tau_-\, dX$ converge *uniformly in probability* to $\int K_-\, dX$ when τ condenses in a suitable manner. Here again the pair (K^τ, X) does not converge jointly to (K, X) unless the relations between K and X are very special.

The purpose of the present note is to provide a criterion for finite dimensional convergence of stochastic integrals and to demonstrate how this criterion can be used in particular situations to obtain results on functional convergence with respect to various topologies on the space $I\!\!D$.

4.2 The results

Let us denote by $\longrightarrow_{\mathcal{D}_f(\mathbb{Q})}$ the finite dimensional convergence over set $\mathbb{Q}$. For example $(K^n, X^n) \longrightarrow_{\mathcal{D}_f(\mathbb{Q})} (K^0, X^0)$ means that for every finite subset $\{0 \le t_1 < t_2 < \ldots < t_m \le 1\} \subset \mathbb{Q}$

$$\begin{aligned}(K^n(t_1), X^n(t_1), K^n(t_2), X^n(t_2) \ldots, K^n(t_m), X^n(t_m))\\ \underset{\mathcal{D}}{\longrightarrow} \quad (K^0(t_1), X^0(t_1), K^0(t_2), X^2(t_2) \ldots, K^0(t_m), X^0(t_m))\end{aligned}$$

on the space $I\!\!R^{2m}$.

It is finite dimensional convergence of (K^n, X^n) on a dense set $\mathbb{Q} \subset [\nvdash, \nVdash]$ and Condition UT for $\{X^n\}$ what implies that X^0 is a semimartingale with respect to the natural filtration generated by (K^0, X^0) (see [14, Lemma 1.3]). In what follows we will use this fact without further reference.

For $k \in I\!\!D$, let $N_\eta(k)$ be the number of η-oscillations of k in the interval $[0,1]$. More precisely: $N_\eta(k) \ge m$ if there are points $0 \le t_1 \le t_2 \le \ldots t_{2m-1} \le t_{2m} \le 1$ such that $|k(t_{2j}) - k(t_{2j-1})| > \eta$, $j = 1, 2, \ldots, m$. By reasons to be clear later we shall say that a sequence $\{K^n\}$ of processes with trajectories in $I\!\!D$ is *uniformly S-tight* if both $\{\|K^n\|_\infty = \sup_{t\in[0,1]} K^n(t)\}$ and $\{N_\eta(K^n)\}$, for each $\eta > 0$, are uniformly tight sequences of random variables.

We will also need a variant of the well-known modulus of continuity ω''. For $k, x \in I\!\!D$ let

$$\omega''_\delta(k, x) = \sup\{\min(|k(s) - k(t)|, |x(t) - x(u)|) : 0 \le s < t < u \le (s + \delta) \wedge 1\}.$$

Theorem 4.2.1 *For each $n \in I\!\!N$, let X^n be a semimartingale with respect to the stochastic basis $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}^n_t\}_{t\in[0,1]}, P^n)$, and let K^n be adapted to $\{\mathcal{F}^n_t\}_{t\in[0,1]}$ and with trajectories in $I\!\!D$. Let $\mathbb{Q} \subset [\nvdash, \nVdash]$, $0, 1 \in \mathbb{Q}$, be dense.*

Suppose Condition UT holds for $\{X^n\}$, $\{K^n\}$ is uniformly S-tight and we have joint finite dimensional convergence over $\mathbb{Q}$:

$$(K^n, X^n) \underset{\mathcal{D}_f(\mathbb{Q})}{\longrightarrow} (K^0, X^0), \tag{4.5}$$

where both K^0 and X^0 have trajectories in $I\!\!D$. Further, suppose there are no oscillations of K^n's preceding oscillations of X^n's:

$$\lim_{\delta\searrow 0} \limsup_{n\to\infty} P^n(\omega''_\delta(K^n, X^n) > \varepsilon) = 0, \quad \varepsilon > 0. \tag{4.6}$$

Then we have

$$\int K^n_- \, dX^n \underset{\mathcal{D}_f(\mathbb{Q})}{\longrightarrow} \int K^0_- \, dX^0. \tag{4.7}$$

The proof (as well as proofs of other results below) is deferred to the next section.

Remark 4.2.2 Stricker [33] proved that Condition UT implies uniform S-tightness. Hence all methods of verifying Condition UT apply to uniform S-tightness as well.

It was announced in the Introduction that Theorem 4.2.1 may serve as a tool for identification of the limit in various kinds of functional convergence of stochastic integrals. By "functional" we mean convergence in distribution with respect to any topology τ on $I\!\!D$ such that relative compactness (in law) and finite dimensional convergence over a dense subset $\mathbb{Q} \subset [\nvdash, \nVdash]$, $0, 1 \in \mathbb{Q}$, imply convergence in law. By the result due to Topsøe [36] Skorohod's J_1 topology *is* functional in our sense. On the other hand, so called "pseudo-path" topology considered by Meyer and Zheng [23] is not "functional", for it is known that convergence of sequences in this topology is just the convergence in (Lebesgue) measure. And it is easy to find a sequence $\{x^n\} \subset I\!\!D$ which converges in measure to $x^0 \equiv 0$ and is such that $x^n(q) \to 1$ for each rational $q \in [0, 1]$.

Corollary 4.2.3 *Suppose all assumptions of Theorem 4.2.1 are in force and we know that the laws of stochastic integrals $\int K^n_- dX^n$ are relatively compact when $I\!\!D$ is equipped with some topology τ generating "functional" convergence.*

Then the sequence $\{\int K^n_- \, dX^n\}_{n\in I\!\!N}$ converges in law with respect to τ (and the limiting process is $\int K^0_- \, dX^0$).

Given uniform τ-tightness of X^n's the task of verifying uniform tightness of $\{\int K^n_- dX^n\}_{n\in I\!\!N}$ can be quite easy. This is so, for instance, in Theorem 4.2.10 below. For metric topologies on $I\!\!D$ we have, however, results more direct than Corollary 4.2.3. We begin with the classical Skorohod's J_1 topology, to emphasize the generality of the present approach.

Theorem 4.2.4 *Let K^n's, X^n's and $\mathbb{Q}$ be as in Theorem 4.2.1. Suppose $\{X^n\}$ fulfills Condition UT, $\{K^n\}$ is uniformly S-tight and finite dimensional convergence (4.5) holds.*

If (4.6) is satisfied and $X^n \longrightarrow_{\mathcal{D}} X^0$ on the space $(I\!\!D, J_1)$ then on the space $(I\!\!D([0,1] : I\!\!R^2), J_1)$

$$(X^n, \int K^n_- \, dX^n) \underset{\mathcal{D}}{\longrightarrow} (X^0, \int K^0_- \, dX^0). \tag{4.8}$$

Remark 4.2.5 Comparing to Theorem 4.1.1 we require very weak convergence of K^n's (in S-topology—see below) and much less information on the joint convergence - finite dimensional convergence and (4.6).

Remark 4.2.6 The crucial property (4.6) (or (4.19) below) holds if either X^n's or K^n's are C-tight, i.e. are uniformly J_1-tight with all limiting laws concentrated on $C([0,1] : I\!R^1)$. Hence our Theorem 4.2.4 generalizes Theorem 4.7 of [19] and is a step in a similar direction as Theorem 4.8 *ibid.*, with dramatically simpler formulation.

Without any change in the proof one can obtain limit results for a variety of topologies on $I\!D$. Let ρ be a metric on $I\!D$ such that the topology $\mathcal{O}_\rho$ generated by ρ is coarser than Skorohod's J_1 topology, but rich enough to preserve the same family of Borel subsets. This guarantees that all probability measures on $(I\!D, \mathcal{O}_\rho)$ are tight and that $X^n \longrightarrow_{\mathcal{D}} X^0$ on $(I\!D, \mathcal{O}_\rho)$ if, and only if, $X^n, n = 0, 1, 2, \ldots$ admit the almost surely convergent Skorohod representation on the Lebesgue interval $([0,1], \mathcal{B}_{[0,1]}, \ell)$. In addition we assume that ρ satisfies

$$\rho(x, y) \leq C\|x - y\|_\infty, \ x, y \in I\!D. \tag{4.9}$$

for some $C > 0$

From the point of view of limit theorems it is natural to consider only metrics which are consistent with convergence of elementary integrals. To explain this notion, take a partition $\tau = \{0 = t_0 < t_1 < t_2 < \ldots < t_m = 1\}$, a sequence $a^\tau = (a_0, a_1, \ldots, a_m) \in I\!R^m$ and $x \in I\!D$ and define

$$(\int a^\tau_- \, dx)(t) = \sum_{k=1}^{m} a_{k-1}(x(t_k \wedge t) - x(t_{k-1} \wedge t)).$$

The consistency means that for every τ and every sequence $(a^n)^\tau \to (a^0)^\tau$ (in $I\!R^m$), $\rho(x^n, x^0) \to 0$ implies

$$\rho(\int (a^n)^\tau_- \, dx^n, \int (a^0)^\tau_- \, dx^0) \to 0. \tag{4.10}$$

(At least metrics generating Skorohod's topologies J_1 and M_1 are consistent with convergence of elementary integrals).

Let us say that ρ is *compatible with integration* if all above requirements are satisfied.

Theorem 4.2.7 *Let K^n's, X^n's and $\mathbb{Q}$ be as in Theorem 4.2.1 and let metric ρ on $I\!D$ be compatible with integration. Suppose $\{X^n\}$ fulfills Condition UT, $\{K^n\}$ is uniformly S-tight and finite dimensional convergence (4.5) holds.*

If (4.6) is satisfied and $X^n \longrightarrow_{\mathcal{D}} X^0$ on the space $(I\!D, \mathcal{O}_\rho)$ then on the same space

$$\int K^n_- \, dX^n \underset{\mathcal{D}}{\longrightarrow} \int K^0_- \, dX^0. \tag{4.11}$$

Remark 4.2.8 Suppose all processes are defined on the same probability space $(\Omega, \mathcal{F}, P)$ and in assumptions of Theorem 4.2.7 we replace relation (4.5) with

$$K^n(t) \underset{\mathcal{P}}{\longrightarrow} K^0(t), \ X^n(t) \underset{\mathcal{P}}{\longrightarrow} X^0(t), \quad t \in \mathbb{Q}, \tag{4.12}$$

and $X^n \longrightarrow_{\mathcal{D}} X^0$ with

$$\lim_{n\to\infty} P(\rho(X^n, X^0) > \varepsilon) = 0, \quad \varepsilon > 0. \tag{4.13}$$

Then it follows from the proof of Theorem 4.2.7 that (4.11) changes to

$$\lim_{n\to\infty} P(\rho(\int K^n_- \, dX^n, \int K^0_- \, dX^0) > \varepsilon) = 0, \quad \varepsilon > 0. \tag{4.14}$$

In a similar way one can transform Theorem 4.2.1.

Dealing with metrics ρ generating topologies strictly finer than the Skorohod's J_1 topology is difficult since such topologies (if any interesting) may be nonseparable. Despite this we have a result for the convergence uniformly in probability, generalizing Theorem 1.9 of [21].

Theorem 4.2.9 *Let K^n's, X^n's and $\mathbb{Q}$ be as in Theorem 4.2.1, with the additional assumption that all processes are defined on the same probability space: $(\Omega^n, \mathcal{F}^n, P^n) = (\Omega, \mathcal{F}, P)$. Suppose Condition UT holds for X^n's, the sequence $\{K^n\}$ is uniformly S-tight, (4.6) is satisfied and*

$$K^n(t) \underset{\mathcal{P}}{\longrightarrow} K^0(t), \quad t \in \mathbb{Q}, \quad \|\mathbb{X}^{\ltimes} - \mathbb{X}^{\nvdash}\|_\infty \underset{\mathcal{P}}{\longrightarrow} \nvdash. \tag{4.15}$$

Then also

$$\| \int K^n_- \, dX^n - \int K^0_- \, dX^0 \|_\infty \underset{\mathcal{P}}{\longrightarrow} 0. \tag{4.16}$$

In fact in Theorem 4.2.1 we have more than finite dimensional convergence only: stochastic integrals already "functionally" converge with respect to an ultraweak topology on $I\!D$, introduced in Paper III and called there "S-topology". This non-Skorohod sequential topology is not metrisable, but it is still good enough to build a satisfactory theory of the convergence in distribution. In particular:

- $K \subset I\!D$ is S-relatively compact iff

$$\sup_{k\in K} \sup_{t\in[0,1]} |k(t)| \le C_K < +\infty, \quad \sup_{k\in K} N_\eta(k) \le C_\eta < +\infty, \ \eta > 0. \tag{4.17}$$

- The σ-field $\mathcal{B}_S$ of Borel subsets for S coincides with the usual σ-field generated by projections (or evaluations) on $I\!D$: $\mathcal{B}_S = \sigma(\pi_t : t \in [0,1])$.

- The set $\mathcal{P}(I\!D, S)$ of S-*tight* probability measures is exactly the set of distributions of stochastic processes with trajectories in $I\!D$: $\mathcal{P}(I\!D, S) = \mathcal{P}(I\!D)$. And for a family $\{K^n\}$ of stochastic processes uniform tightness with respect to S coincides with the uniform S-tightness introduced at the beginning of this section.

- S is weaker than Skorohod's J_1-topology. Since the latter is Polish, S is Lusin in the sense of Fernique. Even more is true: $(I\!D, S)$ is a linear topological space and so is completely regular.

- On $\mathcal{P}(I\!D)$ there exists a unique sequential topology $\mathcal{O}(\overset{*}{\Longrightarrow})$ (where $\overset{*}{\Longrightarrow}$ denotes the convergence determining the topology) which is finer than the S-weak topology and for which $\overset{*}{\Longrightarrow}$-relative compactness coincides with uniform S-tightness. In particular, for $\mathcal{O}(\overset{*}{\Longrightarrow})$ both the direct and the converse Prohorov's theorems are valid.

- Let $X_n \overset{*}{\longrightarrow}_{\mathcal{D}} X_0$ means that the laws of processes X_n converge in our sense: $\mathcal{L}(X_n) \overset{*}{\Longrightarrow} \mathcal{L}(X_0)$. Suppose $X_n \overset{*}{\longrightarrow}_{\mathcal{D}} X_0$. Then in each subsequence $\{X_{n_k}\}_{k\in I\!N}$ one can find a further subsequence $\{X_{n_{k_l}}\}_{l\in I\!N}$ such that:

- $\{X_0\} \cup \{X_{n_{k_l}} : l = 1, 2, \ldots\}$ admit the usual a.s. Skorohod representation on $([0,1], \mathcal{B}_{[0,1]})$;
- outside some *countable* subset of $[0,1)$ all finite dimensional distributions of $\{X_{n_{k_l}}\}$ converge to those of X_0.

• There are many S-continuous functionals. As examples may serve mappings of the form $I\!D \ni x \mapsto \int_0^1 \Phi(x(s))\, d\mu(s) \in I\!R^1$, where μ is a finite atomless measure on $[0,1]$ and Φ is continuous.

Theorem 4.2.10 *In assumptions of Theorem 4.2.1 we have*

$$\int K_-^n\, dX^n \underset{\mathcal{D}}{\longrightarrow} \int K_-^0\, dX^0.$$

on the space $(I\!D, S)$.

We have noticed that uniform S-tightness implies convergence of finite dimensional distributions outside a countable subset of $[0,1)$ and for some subsequence. By Remark 4.2.2 Condition UT also possesses this property. Hence under tightness assumptions only we can always extract a subsequence (K^{n_k}, X^{n_k}) and a dense set $\mathbb{Q}' \subset \mathbb{R}^{\nVdash}$, $1 \in \mathbb{Q}'$, for which joint finite dimensional convergence over $\mathbb{Q}'$ holds. It is however possible that $0 \notin \mathbb{Q}'$ and this fact may influence the convergence of stochastic integrals.

Example 4.2.11 Let $k^n(t) = 1 \neq 0$ and let $x^n(t) = \mathbb{I}_{[1/n,1]}(t)$, $n = 1, 2, \ldots$. Then all assumptions of Theorem 4.2.1 are satisfied, except that (4.5) holds for $\mathbb{Q}' = (\nvdash, \nVdash]$. But

$$\int k_-^n\, dx^n = x^n \not\to 0 = \int k_-^0\, dx^0,$$

in any topology which generates "functional" convergence.

There is an easy way to overcome this difficulty. Let us consider an embedding $I\!D\,([0,1] : I\!R^1) \ni x \mapsto \tilde{x} \in I\!D\,([-1,1] : I\!R^1)$ given by

$$\tilde{x}(t) = \begin{cases} x(t) & \text{if } t \in [0,1]; \\ 0 & \text{if } t \in [-1,0). \end{cases} \tag{4.18}$$

and let

$$\widetilde{\omega}''_\delta(k, x) = \omega''_\delta(\tilde{k}, \tilde{x}),$$

with ω''_δ redefined on $I\!D\,([-1,1] : I\!R^1)$ in a natural manner.

Theorem 4.2.12 *Let K^n and X^n be as in Theorem 4.2.1. Suppose Condition UT holds for $\{X^n\}$, K^n's are uniformly S-tight and*

$$\lim_{\delta \searrow 0} \limsup_{n\to\infty} P^n(\widetilde{\omega}''_\delta(K^n, X^n) > \varepsilon) = 0, \quad \varepsilon > 0. \tag{4.19}$$

Then along some subsequence $\{n_k\}$

$$\int K_-^{n_k}\, dX^{n_k} \underset{\mathcal{D}}{\longrightarrow} \int K_-^0\, dX^0, \tag{4.20}$$

on the space $(I\!D, S)$, where K^0 has trajectories in $I\!D$ and X^0 is a semimartingale.

If, in addition, $\{X^n\}$ is uniformly $\mathcal{O}_\rho$-tight for metric ρ compatible with integration, then (4.20) may be strengthen to convergence on $(I\!D, \mathcal{O}_\rho)$

Remark 4.2.13 The above theorem may be viewed as a specific criterion of compactness for sets of stochastic integrals: the closure (in a suitable topology) still contains stochastic integrals only.

4.3 Proofs

4.3.1 Basic lemma

Lemma 4.3.1 *Suppose $\{K^n\}$ is uniformly S-tight, $\{X^n\}$ satisfies Condition UT and (4.6) holds. Then for any sequence $\tau_m = \{0 = t_{m,0} < t_{m,1} < \ldots < t_{m,k_m} = 1\}$ of partitions of $[0,1]$ such that*

$$|\tau_m| = \max\{t_{m,k} - t_{m,k-1} : k = 1, 2, \ldots, k_m\} \to 0, \tag{4.21}$$

we have

$$\lim_{m\to\infty} \sup_n P(\sup_{t\in[0,1]} |\int (K^n)^{\tau_m}_- \, dX^n(t) - \int K^n_- \, dX^n(t)| > \varepsilon) = 0, \ \varepsilon > 0. \tag{4.22}$$

Proof. Recall that $(K^n)^{\tau_m}$ is the discretization of K^n given by formula (4.4). If $n \in \mathbb{N}$ is fixed and $m \to \infty$ then by the "dominated convergence" theorem

$$\lim_{m\to\infty} P(\sup_{t\in[0,1]} |\int (K^n)^{\tau_m}_- \, dX^n(t) - \int K^n_- \, dX^n(t)| > \varepsilon) = 0, \ \varepsilon > 0. \tag{4.23}$$

It follows that we may replace $\int K^n_- \, dX^n$ with $\int (K^n)^{\tau_{m_n}}_- \, dX^n$, if m_n is large enough. Summarizing, it is enough to prove that for each $\varepsilon > 0$ and m_n such that $|\tau_{m_n}| < |\tau_m|$

$$\lim_{m\to\infty} \sup_{n\in\mathbb{N}} P(\max_{t\in[0,1]} |\int (K^n)^{\tau_m}_- \, dX^n(t) - \int (K^n)^{\tau_{m_n}}_- \, dX^n(t)| > \varepsilon) = 0. \tag{4.24}$$

Let us fix $\eta > 0$, $n \in \mathbb{N}$, τ_m and τ_{m_n}, $|\tau_{m_n}| < |\tau_m|$. To make the formulas more readable, let us change slightly the notation and set

$$\begin{aligned} \tau_m &= \{0 = t_0 < t_1 < \ldots < t_{k_m} = 1\}; \\ \tau_{m_n} &= \{0 = s_0 < s_1 < \ldots < s_{k_{m_n}} = 1\}. \end{aligned}$$

For $k = 1, \ldots, k_m$ and $j = 0, 1, 2, \ldots$ define

$$\sigma_{k,0} = t_{k-1} \tag{4.25}$$

$$\sigma_{k,j+1} = \min\{s_{l-1} \geq \sigma_{k,j} : |K^n(s_{l-1}) - K^n(\sigma_{k,j})| > \eta\} \wedge t_k. \tag{4.26}$$

(We use the convention that $\min \emptyset = 1$). Then for $t \in [0,1]$ we can decompose

$$\begin{aligned} &\int (K^n)^{\tau_m}_- \, dX^n(t) - \int (K^n)^{\tau_{m_n}}_- \, dX^n(t) = \\ &= \sum_{i=1}^{k} \sum_{j=1}^{\infty} \sum_{\sigma_{i,j-1} \leq s_{l-1} < \sigma_{i,j}} (K^n(\sigma_{i,0}) - K^n(s_{l-1}))(X^n(s_l \wedge t) - X^n(s_{l-1} \wedge t)). \end{aligned}$$

Let us observe that

$$\begin{aligned}
&\sum_{\sigma_{i,j-1}\le s_{l-1}<\sigma_{i,j}} (K^n(\sigma_{i,0}) - K^n(s_{l-1}))(X^n(s_l\wedge t) - X^n(s_{l-1}\wedge t)) \\
&= \sum_{\sigma_{i,j-1}\le s_{l-1}<\sigma_{i,j}} (K^n(\sigma_{i,0}) - K^n(\sigma_{i,j-1}))(X^n(s_l\wedge t) - X^n(s_{l-1}\wedge t)) \\
&\quad + \sum_{\sigma_{i,j-1}\le s_{l-1}<\sigma_{i,j}} (K^n(\sigma_{i,j-1}) - K^n(s_{l-1}))(X^n(s_l\wedge t) - X^n(s_{l-1}\wedge t)) \\
&= (K^n(\sigma_{i,0}) - K^n(\sigma_{i,j-1}))(X^n(\sigma_{i,j}\wedge t) - X^n(\sigma_{i,j-1}\wedge t)) \\
&\quad + \sum_{\sigma_{i,j-1}\le s_{l-1}<\sigma_{i,j}} (K^n(\sigma_{i,j-1}) - K^n(s_{l-1}))(X^n(s_l\wedge t) - X^n(s_{l-1}\wedge t))
\end{aligned}$$

Finally, we have

$$\begin{aligned}
&\int (K^n)_-^{\tau_m}\, dX^n(t) - \int (K^n)_-^{\tau_{m_n}}\, dX^n(t) = \\
&= \sum_{i=1}^{k}\sum_{j=2}^{\infty} (K^n(\sigma_{i,0}) - K^n(\sigma_{i,j-1}))(X^n(\sigma_{i,j}\wedge t) - X^n(\sigma_{i,j-1}\wedge t)) \\
&\quad + \sum_{i=1}^{k}\sum_{j=1}^{\infty}\sum_{\sigma_{i,j-1}\le s_{l-1}<\sigma_{i,j}} (K^n(\sigma_{i,j-1}) - K^n(s_{l-1}))(X^n(s_l\wedge t) - X^n(s_{l-1}\wedge t)) \\
&= I_\eta^n(t) + J_\eta^n(t)
\end{aligned}$$

Using definitions of $N_\eta(\cdot)$ and $\omega''_\delta(\cdot,\cdot)$ given at the beginning of Section 3.2 and taking into account that for fixed ω in the sum $I_\eta^n(t,\omega)$ there is no more than $N_\eta(K^n(\omega))$ nonzero summands we can estimate

$$\begin{aligned}
&\sup_{t\in[0,1]} |I_\eta^n(t)| \\
&\le N_\eta(K^n)\cdot \\
&\quad \cdot \sup_{\substack{1\le i\le k_m \\ j\in I\!N,\ t\in[0,1]}} \{\max\{|K^n(\sigma_{i,0}) - K^n(\sigma_{i,j-1})|, |X^n(\sigma_{i,j}\wedge t) - X^n(\sigma_{i,j-1}\wedge t)|\}\cdot \\
&\qquad \cdot \min\{|K^n(\sigma_{i,0}) - K^n(\sigma_{i,j-1})|, |X^n(\sigma_{i,j}\wedge t) - X^n(\sigma_{i,j-1}\wedge t)|\}\} \\
&\le N_\eta(K^n)\cdot 2(\|K^n\|_\infty + \|X^n\|_\infty)\cdot \omega''_{|\tau_m|}(K^n, X^n).
\end{aligned}$$

Since Condition UT implies tightness of $\|X^n\|_\infty$ (see e.g. [14], Lemma 1.2) we see that when η is fixed and $|\tau_m| \to 0$ then $\sup_{t\in[0,1]} |I_\eta^n(t)|$ converges in probability to 0 uniformly in n.

It remains to prove that by the choice of η random variables $\sup_{t\in[0,1]} |J_\eta^n(t)|$ can be made as small as desired (in probability, uniformly in n). But the processes $\eta^{-1}\cdot J_\eta^n$ are elementary stochastic integrals appearing in the definition of Condition UT. By Lemma 1.1 [14] the family of random variables $\{\sup_{t\in[0,1]} \eta^{-1}\cdot |J_\eta^n(t)| : n \in I\!N, \eta > 0\}$ is uniformly tight and we obtain the required property. □

4.3.2 Proof of Theorem 4.2.1

If $\mathbb{Q}$ is not countable, replace it with its proper countable dense subset containing 0 and 1. Let us choose a sequence $\{\tau_m\}$ of partitions of $[0,1]$ such that $|\tau_m| \to 0$, $\tau_m \subset \tau_{m+1} \subset \mathbb{Q}$, $m = 1, 2, \ldots$ and $\cup_{m=1}^{\infty} \tau_m = \mathbb{Q}$.

Let $\mathbb{Q}_{\nvdash} = \{\shortmid\shortmid_{\nVdash} < \shortmid\shortmid_{\nvDash} \ldots < \shortmid\shortmid_{\searrow}\} \subset \mathbb{Q}$ and let m be so large that $\mathbb{Q}_{\nvdash} \subset \tau_{\gtrdot} = \{\nvdash = \approx_{\nvdash} < \approx_{\nVdash} < \ldots < \approx_{\neg_{\gtrdot}}\}$ as well as

$$\sup_n P(\sup_{q \in \mathbb{Q}_{\nvdash}} |\int (K^n)_-^{\tau_m}\, dX^n(q) - \int K_-^n\, dX^n(q)| > \varepsilon) < \varepsilon, \tag{4.27}$$

for $n = 0, 1, 2, \ldots$ (the latter by Lemma 4.3.1). For each $t \in \tau_m$ the integral $\int (K^n)_-^{\tau_m}\, dX^n(t)$ is a continuous function of the vector

$$(K^n(0), X^n(0), K^n(t_1), X^n(t_1), \ldots, K^n(t_{k_m}), X^n(t_{k_m})).$$

Hence (4.5) implies

$$\int (K^n)_-^{\tau_m}\, dX^n \underset{\mathcal{D}_f(\tau_m)}{\longrightarrow} \int (K^0)_-^{\tau_m}\, dX^0.$$

and by $\mathbb{Q}_{\nvdash} \subset \tau_{\gtrdot}$ also

$$\int (K^n)_-^{\tau_m}\, dX^n \underset{\mathcal{D}_f(\mathbb{Q}_{\nvdash})}{\longrightarrow} \int (K^0)_-^{\tau_m}\, dX^0. \tag{4.28}$$

The theorem follows now by (4.27) and (4.28). □

4.3.3 Proof of Theorems 4.2.4–4.2.9

We shall prove Theorem 4.2.7 first.

A variant of the Skorohod representation theorem is necessary.

Lemma 4.3.2 *For each subsequence n_k there exists a further subsequence n_{k_l} and random elements $L^0, L^1, \ldots$ with values in $I\!R^{\mathbb{Q}}$ and $Y^0, Y^1, \ldots$ with values in $(I\!D, \mathcal{O}_\rho)$ and defined on the Lebesgue interval such that*

$$((L^0(q), Y^0(q))_{q \in \mathbb{Q}}, Y^0) \sim ((K^0(q), X^0(q))_{q \in \mathbb{Q}}, X^0),$$

for each $l = 1, 2, \ldots$

$$((L^l(q), Y^l(q))_{q \in \mathbb{Q}}, Y^l) \sim ((K^{n_{k_l}}(q), X^{n_{k_l}}(q))_{q \in \mathbb{Q}}, X^{n_{k_l}}),$$

and for almost every $\omega \in [0, 1]$

$$L^l(q, \omega) \to L^0(q, \omega),\ \ Y^l(q, \omega) \to Y^0(q, \omega), q \in \mathbb{Q},$$

and

$$\rho(Y^l(\omega), Y^0(\omega)) \to 0.$$

Proof. We have separate information on joint finite dimensional convergence and convergence on $(I\!D, \mathcal{O}_\rho)$. By tightness of both components, in each subsequence we may extract a further subsequence such that the joint convergence holds. Let (U^l, V^l, Z^l) be the Skorohod representation for such subsequence. In particular, we have

$$(V^l, Z^l) \sim (h(X^{n_{k_l}}), X^{n_{k_l}}), \tag{4.29}$$

where $h : I\!D \to I\!R^{\mathbb{Q}}$ is a measurable mapping given by $h(x) = (x(q))_{q \in \mathbb{Q}}$. Hence (4.29) implies that $V^l = h(Z^l)$ ℓ-a.s. and so the lemma follows. □

Because of Lemma 4.3.2 we may and do assume that $\rho(X^n(\omega), X^0(\omega)) \to 0$ and that $K^n(q,\omega) \to K^0(q,\omega)$, $X^n(q,\omega) \to X^0(q,\omega)$, $q \in \mathbb{Q}$. Using the consistency of ρ with respect to elementary integrals we get for each fixed τ_m

$$\rho(\int (K^n)_-^{\tau_m}\, dX^n, \int (K^0)_-^{\tau_m}\, dX^0) \to 0 \text{ a.s.}$$

Further, by (4.9) and (4.22) we have for each $\varepsilon > 0$ and as $m \to \infty$

$$\begin{aligned} &\sup_{n\in\mathbb{N}} P(\rho(\int (K^n)_-^{\tau_m}\, dX^n, \int K_-^n\, dX^n) > \varepsilon) \\ &\qquad \leq \sup_{n\in\mathbb{N}} P(\| \int (K^n)_-^{\tau_m}\, dX^n - \int K_-^n\, dX^n \|_\infty > \frac{\varepsilon}{C}) \to 0. \end{aligned}$$

Similarly, by (4.23)

$$\begin{aligned} &P(\rho(\int (K^0)_-^{\tau_m}\, dX^0, \int K_-^0\, dX^0) > \varepsilon) \\ &\qquad \leq P(\| \int (K^0)_-^{\tau_m}\, dX^0 - \int K_-^0\, dX^0 \|_\infty > \frac{\varepsilon}{C}) \to 0, \quad \text{as } m \to \infty, \varepsilon > 0. \end{aligned}$$

Hence $\int K_-^{n_{k_l}}\, dX^{n_{k_l}} \longrightarrow_{\mathcal{D}} \int K_-^0\, dX^0$ on $(\mathbb{D}, \mathcal{O}_\rho)$. This concludes the proof, for n_{k_l} was a subsequence of an *arbitrary* subsequence n_k. □

The proof of Theorem 4.2.9 is essentially the same (except it does not require reduction via the a.s. Skorohod representation). □

4.3.4 Proof of Theorem 4.2.10

It has been proved in Paper III that the convergence $\overset{*}{\Longrightarrow}$ in $\mathcal{P}(\mathbb{D})$ (induced by S-topology) is "functional" in our sense, i.e. finite dimensional convergence and relative $\overset{*}{\Longrightarrow}$-compactness imply $\overset{*}{\Longrightarrow}$-convergence. Hence in view of (4.7) it suffices to prove S-uniform tightness of $\{\int K_-^n\, dX^n\}$. Using [21, Lemma 1.6], we check that processes $\{\int K_-^n\, dX^n\}$ satisfy Condition UT. Now the mentioned already result due to Stricker [33] (see Remark 4.2.2) gives us S-uniform tightness of $\{\int K_-^n\, dX^n\}$ and finishes the proof of Theorem 4.2.10. □

4.3.5 Proof of Theorem 4.2.12

Let us embed our processes into the space $\mathbb{D}\,[-1,1] : \mathbb{R}^1)$ via mapping (4.18). Then we have finite dimensional convergence on $\mathbb{Q}_{\nVdash} = \mathbb{Q} \cup [-\nVdash, \nvdash)$. Since the left end of the interval of integration belons to $\mathbb{Q}_{\nVdash}$, it is possible to apply previous results. It remains to observe that

$$\int_{]-1,t]} K^n(s-)\, dX^n(s) = \int_{]0,t]} K^n(s-)\, dX^n(s),\ n = 1, 2, \ldots.$$

□

Bibliography

[1] Avram, F., Taqqu, M. Weak convergence of sums of moving averages in the α-stable domain of attraction, *Ann. Probab.* **20** (1992) 483–503.

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

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

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

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

[6] Dellacherie, C., Meyer, P.A. **Probabilités et potentiel**, t. 2, Hermann, Paris 1980.

[7] Dudley, R.M., Distances of probability measures and random variables, Ann. Math. Stat., **39** (1968) 1563–1572.

[8] Dudley, R.M., An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions, in: A. Beck at al. Eds., Probability in Banach Spaces V, Lecture Notes in Math. **1153** (1985) 141–178.

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

[10] Fernandez, P.J., Almost surely convergent versions of sequences which converge weakly, Bol. Soc. Brasil. Math. **5** (1974) 51–61.

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

[12] Fernique, X., Convergence en loi de variables alétoires et de fonctions alétoires, propriétes de compacité des lois, II, in: J. Azéma, P.A. Meyer, M. Yor, (Eds.) *Séminaire de Probabilités XXVII, Lecture Notes in Math.* v. **1557**, pp. 216–232, Springer, Berlin 1993.

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

[14] Jakubowski, A., Mémin, J., Pages, G. Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbb{D}^1$ de Skorokhod, *Probab. Th. Rel. Fields* **81** (1989) 111–137.

[15] Janicki, A., Weron, A. **Simulation and Chaotic Behaviour of α-Stable Stochastic Processes**, Marcel Dekker, New York 1994.

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

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

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

[19] Kurtz, T., Protter, P. Weak limit theorems for stochastic integrals and stochastic differential equations, *Ann. Probab.* **19** (1991) 1035–1070.

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

[21] Mémin, J., Słomiński, L. Condition UT et stabilité en loi des solutions d'equations différentieles stochastiques, *Sém. de Probab. XXV, Lecture Notes in Math.* v. **1485**, pp. 162–177, Springer, Berlin 1991.

[22] Meyer, P.A., Le théorème de continuite de P. Lévy sur les espaces nucléaires (d'aprés X. Fernique). Séminaire Bourbaki, 18^e année, no 311, (1965-66), 1–14.

[23] Meyer, P.A., Zheng, W.A. Tightness criteria for laws of semimartingales, *Ann. Inst. Henri Poincaré* B **20** (1984) 353–372.

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

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

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

[27] Protter, Ph., **Stochastic Integration and Differential Equations. A New Approach.** 2nd ed., Springer 1992.

[28] Schaefer, H.H., **Topological Vector Spaces**, Springer, Berlin 1970.

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

[30] Słomiński, L. Stability of strong solutions of stochastic differential equations, *Stoch. Proc. Appl.* **31** (1989) 173–202.

[31] Słomiński, L. Stability of stochastic differential equations driven by general semimartingales, preprint (1994), to appear in *Dissert. Math.*.

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

[33] Stricker, C. Lois de semimartingales et critères de compacité, *Séminares de probabilités XIX. Lect. Notes in Math.* **1123**, Springer, Berlin 1985.

[34] Szczotka, W., A note on Skorokhod representation, Bull. Pol. Acad. Sci. Math. **38** (1990) 35–39.

[35] Topsøe, F. A criterion for weak convergence of measures with an application to convergence of measures on $\mathbb{D}[0,1]$, *Math. Scand.* **25** (1969) 97–104.

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

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

[38] Wichura, M.J., On the construction of almost uniformly convergent random variables with given weakly convergent image laws, Ann. Math. Statist. **41** (1970) 284–291.

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