PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 13, Fasc. 1 (1992), pp. 39–58

WEAK CONVERGENCE UNDER MAPPING

BY

WŁADYSŁAW SZCZOTKA (WROCŁAW)

Abstract. For a given random element X of a metric space S and a measurable mapping h of S into a metric space S_1 such that $P\{X \in D_h\} > 0$ we give the conditions for a sequence of random elements X_n, $n \geqslant 1$, of the space S under which the convergence $X_n \xrightarrow{D} X$ implies $h(X_n) \xrightarrow{D} h(X)$ (Lemma 1) and stronger conditions for $\{X_n\}$ under which the convergence $X_n \xrightarrow{D} X$ implies $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$ (Theorem 3). Here D_h is the set of discontinuities of h. The case $S = D[0, \infty)$, $h(x) = \sup_{0 \leqslant t < \infty} x(t)$ is considered in detail.

1. Introduction. One of the theorems most frequently used in applications of the weak convergence of probability measures is the continuous mapping theorem (CMT) (see [1], Theorem 5.1). It says that if μ and μ_n, $n \geqslant 1$, are probability measures on a metric space S, h is a measurable mapping of S into a metric space S_1, and D_h is the set of discontinuities of h, then the weak convergence $\mu_n \Rightarrow \mu$ and $\mu(D_h) = 0$ imply the weak convergence $\mu_n h^{-1} \Rightarrow \mu h^{-1}$.

In the queueing theory and the reliability theory many characteristics have the following form:

$$h(X) = \sup_{0 \leqslant t < \infty} X(t),$$

where X is a process with sample paths belonging to the space $C[0, \infty)$ (the space of all continuous real-valued functions on $[0, \infty)$) or to the space $D[0, \infty)$ (the space of right-continuous real-valued functions on $[0, \infty)$ with limit from the left). Considering the space $C[0, \infty)$ with the topology generated by the uniform convergence on compact sets we see that the mapping h is not continuous at each $x \in C[0, \infty)$ such that $h(x) < \infty$. To see this let us take any continuous function x such that $b \overset{\text{df}}{=} h(x) < \infty$ and define the functions x_n, $n \geqslant 1$, as $x_n(t) = x(t)$ for $0 \leqslant t \leqslant n$, $x_n(t) = x(n) + (t-n)(b+\varepsilon - x(n))$ for $n \leqslant t \leqslant n+1$ and $x_n(t) = b+\varepsilon$ for $t \geqslant n+1$, where $\varepsilon > 0$. It is obvious that x_n, $n \geqslant 1$, are continuous and $\sup_{0 \leqslant t \leqslant c} |x_n(t) - x(t)| \to 0$ as $n \to \infty$ for any $c > 0$, but $h(x_n) = b + \varepsilon \not\to b = h(x)$.

A similar example can be given in the space $D[0, \infty)$ considered with the metric defined by Lindvall in [3]. $D[0, \infty)$ with Lindvall's metric is a Polish metric space, and Lindvall's metric generates the Stone topology in $D[0, \infty)$ (see [3]). The above example shows that we cannot use CMT to the investigation of the convergence

$$\sup_{0 \leqslant t < \infty} X_n(t) \xrightarrow{D} \sup_{0 \leqslant t < \infty} X(t)$$

under $X_n \xrightarrow{D} X$ in $D[0, \infty)$ with the Stone topology. Thus the following problem arises: for which sequences $\{X_n\} \overset{\text{df}}{=} \{X_n, n \geqslant 1\}$ does the convergence $X_n \xrightarrow{D} X$ imply the convergence $\sup_{0 \leqslant t < \infty} X_n(t) \xrightarrow{D} \sup_{0 \leqslant t < \infty} X(t)$? We generally state this problem as follows: Given a measurable mapping h of a metric space S into a metric space S_1 and given a random element X of S we ask: for which sequences $\{X_n\}$ such that $X_n \xrightarrow{D} X$ does the convergence $h(X_n) \xrightarrow{D} h(X)$ hold? Obviously, if $P\{X \in D_h\} = 0$, then by CMT we know that $X_n \xrightarrow{D} X$ implies $h(X_n) \xrightarrow{D} h(X)$.

Our approach to the investigation of the stated problem is based on an approximation of h by a sequence of measurable mappings h_k, $k \geqslant 1$, of S into S_1 such that

(a) $$h_k(X_n) \xrightarrow{D} h_k(X) \quad \text{as } n \to \infty, \text{ for each } k \geqslant 1,$$

and

(b) $$h_k(X) \xrightarrow{D} h(X) \quad \text{as } k \to \infty.$$

Then, as Proposition 1 shows, under (a) and (b) the condition

(c) $$\lim_k \overline{\lim_n}\, \varrho_0\big(\mathscr{L}(h_k(X_n)), \mathscr{L}(h(X_n))\big) = 0$$

is equivalent to the condition

(d) $$h(X_n) \xrightarrow{D} h(X) \quad \text{as } n \to \infty,$$

where ϱ_0 is any metric in the space of probability measures which metrizes the weak topology.

Theorem 1, given in Section 3, shows that if S_1 is the Euclidean space, then in the situation $P\{X \in D_h\} = 0$ there exists a sequence of continuous mappings h_k, $k \geqslant 1$, of S into S_1 such that the conditions (a), (b) and (c) hold under $X_n \xrightarrow{D} X$. From this we infer (Theorem 2) that our approach contains the situation of CMT, i.e. Theorem 5.1 from [1].

One of the properties desirable from a practical point of view is the implication

$$[X_n \xrightarrow{D} X \text{ and } h(X_n) \xrightarrow{D} h(X)] \Rightarrow [(X_n, h(X_n)) \xrightarrow{D} (X, h(X))].$$

This implication is true when $P\{X \in D_h\} = 0$. Unfortunately, it is generally false (see [6], Example 1). The reason for which this implication is desirable is the following: Let X and X_n, $n \geqslant 1$, be processes which generate the 0-th and the n-th queueing systems, respectively, and let $h(X)$ and $h(X_n)$ be some characteristics of those systems (for example the process of waiting time). Then the validity of the above implication allows us to investigate the joint convergence $(h(X_n), f(X_n, h(X_n))) \xrightarrow{D} (h(X), f(X, h(X)))$ as $n \to \infty$, where $f(X, h(X))$ is another characteristic of the queueing system.

An attempt of a characterization of the above implication is given in [6]. Here we give more concrete conditions under which the mentioned implication holds (Theorems 2 and 3). Thus if there exists a sequence of measurable mappings h_k, $k \geq 1$, of S into S_1 such that $P\{X \in D_{h_k}\} = 0$, $k \geq 1$, and

(b′) $$h_k(X) \xrightarrow{P} h(X) \quad \text{as } k \to \infty,$$

then additionally under $X_n \xrightarrow{D} X$ the condition

(c′) $$\lim_k \overline{\lim_n}\, \alpha(h_k(X_n), h(X_n)) = 0$$

is equivalent to the condition

(d′) $$(X_n, h(X_n)) \xrightarrow{D} (X, h(X)) \quad \text{as } n \to \infty.$$

Here α is the metric which metrizes the convergence in probability. Theorem 1a, given in Section 3, shows that if S is separable and S_1 is the Euclidean space, then in the situation $P\{X \in D_h\} = 0$ there exists a sequence of continuous mappings h_k, $k \geq 1$, of S into S_1 such that conditions (b′) and (c′) hold under $X_n \xrightarrow{D} X$. This approach allows us to reduce the investigation of the weak convergence of the joint distribution of several characteristics to the weak convergence of each characteristic separately (see Corollary 3).

In Section 4 we consider a special case of S and h, i.e. $S = D[0, \infty)$ and $h(x) = \sup_{0 \leq t < \infty} x(t)$. Furthermore we give an application of the results obtained to investigating the asymptotic stationarity of queueing systems.

2. Preliminaries. The paper uses the terminology of the weak convergence of probability measures, so most of the notation appearing here can be found in [1]. Here we introduce only specific notions, assumptions, and we formulate auxiliary facts. Throughout Sections 2 and 3 the letters S, $\tilde{S}$ and S_i, $1 \leq i \leq m$, denote metric spaces with metrics ϱ, $\tilde{\varrho}$ and ϱ_i, $1 \leq i \leq m$, respectively. The Cartesian product of metric spaces is considered with the product metric. For a Borel σ-field of subsets of a metric space we write $\mathscr{B}$ before the symbol denoting the space. For a mapping h the symbol D_h denotes the set of discontinuities of h. The set of all probability measures on $(S, \mathscr{B}(S))$ is denoted by $\mathscr{M}(S)$, and ϱ_P denotes the Prohorov metric on $\mathscr{M}(S)$. If S is separable, then $\nu(S)$ denotes the space of random elements of S defined on a fixed probability space. This space is considered with the metric α defined as (see [7])

$$\alpha(X, Y) = \inf\{\varepsilon\colon P\{\varrho(X, Y) \geq \varepsilon\} \leq \varepsilon\} \quad \text{for } X, Y \in \nu(S).$$

For the distribution of a random element we put $\mathscr{L}$ before a symbol denoting the random element. By $\Rightarrow$, $\xrightarrow{D}$, $\xrightarrow{P}$ we denote weak convergence of probability measures, convergence in distribution and convergence in probability of random elements. The Prohorov metric ϱ_P metrizes the weak topology in the space of probability measures on a fixed metric space and the metric α metrizes the topology of convergence in probability of random elements. Furthermore, the following relations hold (see [2]):

$$\alpha(X_1, Y_1) \leqslant \alpha(X_2, Y_2) \quad \text{when } \varrho(X_1, Y_1) \leqslant \varrho(X_2, Y_2) \text{ a.e.},$$

$$\varrho_P(\mathscr{L}(X), \mathscr{L}(Y)) \leqslant \alpha(X, Y) \quad \text{for } X, Y \in v(S).$$

In a few places we refer to the following fact:

PROPOSITION 1. *Let* $\{x_{k,n}, k, n \geqslant 1\}$, $\{x_k, k \geqslant 1\}$ *and* $\{y_n, n \geqslant 1\}$ *be an array and sequences of elements of the space* $\tilde{S}$, *respectively, such that: for each* $k \geqslant 1$, $x_{k,n} \to x_k$ *as* $n \to \infty$ *and* $x_k \to x \in \tilde{S}$ *as* $k \to \infty$. *Then the convergence*

$$\lim_k \overline{\lim_n} \tilde{\varrho}(x_{k,n}, y) = 0$$

is equivalent to the convergence $y_n \to x$ *as* $n \to \infty$.

Proof. The proof of the assertion is a consequence of the following inequalities which follow from the triangle inequality for a metric:

$$\tilde{\varrho}(y_n, x) \leqslant \tilde{\varrho}(y_n, x_{k,n}) + \tilde{\varrho}(x_{k,n}, x_k) + \tilde{\varrho}(x_k, x)$$

and

$$\tilde{\varrho}(x_{k,n}, y) \leqslant \tilde{\varrho}(x_{k,n}, x_k) + \tilde{\varrho}(x_k, x) + \tilde{\varrho}(x, y_n). \blacksquare$$

Before formulating other auxiliary facts let us introduce the notation of some conditions which are satisfied by an array $\{Z_{k,n}, k, n \geqslant 1\}$ and sequences $\{Z_k, k \geqslant 1\}$ and $\{Y_n, n \geqslant 1\}$ of random elements of S_1 and a random element Z of S_1.

A_1. for each $k \geqslant 1$, $Z_{k,n} \xrightarrow{D} Z_k$ as $n \to \infty$;

A_2. for each $i, k \geqslant 1$, $(Z_{i,n}, Z_{k,n}) \xrightarrow{D} (Z_i, Z_k)$ as $n \to \infty$;

B_1. $Z_k \xrightarrow{D} Z$ as $k \to \infty$;

B_1a. $Z_k \xrightarrow{P} Z$ as $k \to \infty$;

B_2. for each $i \geqslant 1$, $(Z_i, Z_k) \xrightarrow{D} (Z_i, Z)$ as $k \to \infty$;

C_1. $\lim_k \overline{\lim}_n \varrho_P(\mathscr{L}(Z_{k,n}), \mathscr{L}(Y_n)) = 0$;

C_1a. for each $\varepsilon > 0$, $\lim_k \overline{\lim}_n P\{\varrho_1(Z_{k,n}, Y_n) \geqslant \varepsilon\} = 0$;

C_2. for each $i \geqslant 1$, $\lim_k \overline{\lim}_n \varrho_P(\mathscr{L}(Z_{i,n}, Z_{k,n}), \mathscr{L}(Z_{i,n}, Y_n)) = 0$;

D_1. $Y_n \xrightarrow{D} Z$ as $n \to \infty$;

D_2. for each $i \geqslant 1$, $(Z_{i,n}, Y_n) \xrightarrow{D} (Z_i, Z)$ as $n \to \infty$.

Let us notice that the formulations of some of the above conditions need additional assumptions. Namely, in the conditions C_1a and C_2 the random elements Y_n and $Z_{k,n}$, $k \geqslant 1$, must be defined on a common probability space. Similarly, in the condition B_1a the random elements Z and Z_k, $k \geqslant 1$, must be defined on a common probability space. Moreover, the conditions A_2, B_1a, B_2, C_1a, C_2 and D_2 need the separability of S_1.

It is obvious that the following implications hold:

$$A_2 \Rightarrow A_1, \quad B_1a \Rightarrow B_2 \Rightarrow B_1, \quad C_1a \Rightarrow C_2 \Rightarrow C_1 \quad \text{and} \quad D_2 \Rightarrow D_1.$$

The implication $C_1a \Rightarrow C_2$ holds because the distance between $(Z_{i,n}, Z_{k,n})$ and $(Z_{i,n}, Y_n)$ in $S_1 \times S_1$ is equal to $\varrho_1(Z_{k,n}, Y_n)$, because of the inequality $\varrho_P(\mathscr{L}(X), \mathscr{L}(Y)) \leqslant \alpha(X, Y)$ when $X, Y \in v(S)$ holds and at last because of the following fact:

PROPOSITION 2. *The condition* C_1a *is equivalent to*

C_3. $\lim_k \overline{\lim}_n \alpha(Z_{k,n}, Y_n) = 0$.

The following fact gives more detailed relations between the above conditions:

PROPOSITION 3. (i) *If* A_1 *and* B_1 *hold, then* C_1 *and* D_1 *are equivalent.*

(ii) *If* S_1 *is separable and* A_2 *and* B_2 *hold, then* C_2 *and* D_2 *are equivalent.*

(iii) *If* S_1 *is separable and* A_2 *and* B_1a *hold, then* C_1a, C_2 *and* D_2 *are equivalent.*

Note 1. The assertion (i) is stronger than Theorem 4.2 from [1] where it has been shown that under A_1, B_1 and C_1a the condition D_1 holds.

Proof. The implications (i) and (ii) are immediate consequences of Proposition 1 with the following specifications: $\tilde{S} = \mathscr{M}(S)$ and $\tilde{\varrho} = \varrho_P$.

Now let us consider the implication (iii). Since B_1a implies B_2, so under A_2 and B_1a the equivalence of C_2 and D_2 follows from the implication (ii). Thus in view of the implication $C_1a \Rightarrow C_2$ it is enough to show the implication $C_2 \Rightarrow C_1a$ under A_2 and B_1a. But then D_2 holds. Hence, in view of the continuity of the metric, we get

$$\varrho_1(Z_{k,n}, Y_n) \xrightarrow{D} \varrho_1(Z_k, Z) \quad \text{as } n \to \infty.$$

Thus for almost all $\varepsilon > 0$ with respect to the Lebesgue measure we have

$$\lim_n P\{\varrho_1(Z_{k,n}, Y_n) > \varepsilon\} = P\{\varrho_1(Z_k, Z) > \varepsilon\}.$$

Let ε_1 be such that the above fails and let ε be such that $0 < \varepsilon < \varepsilon_1$ and the above holds. Then we have

$$\begin{aligned}\overline{\lim_n} P\{\varrho_1(Z_{k,n}, Y_n) > \varepsilon_1\} &\leqslant \lim_n P\{\varrho_1(Z_{k,n}, Y_n) > \varepsilon\} \\ &= P\{\varrho_1(Z_k, Z) > \varepsilon\}.\end{aligned}$$

Hence and from B_2a we get C_1a, which completes the proof. ∎

3. Weak convergence under mappings. Let h be a measurable mapping of S into S_1 and let X and X_n, $n \geq 1$, be random elements of S. Our main criterion for examination of the convergence $h(X_n) \xrightarrow{D} h(X)$ under $X_n \xrightarrow{D} X$ is the following

LEMMA 1. *Assume that there exists a sequence of measurable mappings h_k, $k \geq 1$, of S into S_1 such that*

(1) $$\text{for each } k \geq 1,\ h_k(X_n) \xrightarrow{D} h_k(X) \quad \text{as } n \to \infty,$$

(2) $$h_k(X) \xrightarrow{D} h(X) \quad \text{as } k \to \infty.$$

Then the following conditions are equivalent:

(3) $$\lim_k \overline{\lim_n}\, \varrho_P\big(\mathscr{L}(h_k(X_n)),\ \mathscr{L}(h(X_n))\big) = 0,$$

(4) $$h(X_n) \xrightarrow{D} h(X) \quad \text{as } n \to \infty.$$

Proof. The assertion is a consequence of the implication (i) of Proposition 3 with the following specification: $Z_{k,n} = h_k(X_n)$, $Z_k = h_k(X)$, $Z = h(X)$, $Y_n = h(X_n)$, $n, k \geq 1$.

Let us notice that if S is a Polish metric space and S_1 is Euclidean, then there exists a sequence of continuous mappings h_k, $k \geq 1$, of S into S_1 such that $h_k(X) \to h(X)$ a.e. as $k \to \infty$. Unfortunately, to verify (3) we ought to know something more about h_k, $k \geq 1$. Thus we ought to give a construction of those mappings. The following theorem gives such a construction in the situation when S_1 is the Euclidean space and $P\{X \in D_h\} = 0$.

THEOREM 1. *If $X_n \xrightarrow{D} X$ and $P\{X \in D_h\} = 0$ where S_1 is the Euclidean space, then there exists a sequence of continuous mappings h_k, $k \geq 1$, of S into S_1 such that*

(5) $$h_k(X) \xrightarrow{D} h(X) \quad \text{as } k \to \infty,$$

(6) $$\lim_k \overline{\lim_n}\, \varrho_P\big(\mathscr{L}(h_k(X_n)),\ \mathscr{L}(h(X_n))\big) = 0.$$

Proof. Since S is a metric space, $\mathscr{L}(X)$ is regular. Hence for each $k \geq 1$ there exist closed sets F_k such that $F_k \subset D_h^c$ and $P\{X \in F_k\} \geq 1 - 1/k$, where D_k^c denotes the complement of D_h. Because of the continuity of h on D_h^c, h is continuous on each closed set F_k. Thus, by Tietze's theorem, for each $k \geq 1$ there exist continuous mappings h_k on S such that $h_k = h$ on F_k, $k \geq 1$. Moreover, in view of $P\{h_k(X) \neq h(X)\} \leq 1/k$ we have (5).

Now we show that the sequence of h_k, $k \geq 1$, satisfies (6). For clarity let us denote by μ and μ_n, $n \geq 1$, the distributions of X and X_n, $n \geq 1$, respectively. Then we have

$$\varrho_P(\mu_n h_k^{-1}, \mu_n h^{-1}) = \inf\{\varepsilon\colon \mu_n h^{-1}(F) \leq \mu_n h_k^{-1}(F^\varepsilon) + \varepsilon\}$$
$$\leq \inf\{\varepsilon\colon 0 \leq \mu_n h_k^{-1}(F^\varepsilon) + \varepsilon - \mu_n(\overline{h^{-1}F})\},$$

where $F^{\varepsilon} = \{x \in S_1: \varrho_1(x, F) < \varepsilon\}$, and F is closed. Hence

$$\overline{\lim_n}\, \varrho_P(\mu_n h_k^{-1}, \mu_n h^{-1})$$
$$\leqslant \overline{\lim_n} \inf\{\varepsilon: 0 \leqslant \mu_n h_k^{-1}(F^{\varepsilon}) + \varepsilon - \mu_n(\overline{h^{-1}F})\}$$
$$\leqslant \inf\{\varepsilon: 0 \leqslant \underline{\lim_n}\, \mu_n(h_k^{-1}F^{\varepsilon}) + \varepsilon - \overline{\lim_n}\, \mu_n(\overline{h^{-1}F})\}$$
$$\leqslant \inf\{\varepsilon: 0 \leqslant \mu h_k^{-1}(F^{\varepsilon}) + \varepsilon - \mu(\overline{h^{-1}F})\}$$
$$\leqslant \inf\{\varepsilon: 0 \leqslant \mu h_k^{-1}(F^{\varepsilon}) + \varepsilon - \mu(D_h \cup h^{-1}F)\}$$
$$\leqslant \inf\{\varepsilon: 0 \leqslant \mu h_k^{-1}(F^{\varepsilon}) + \varepsilon - \mu(h^{-1}F)\} = \varrho_P(\mu h_k^{-1}, \mu h^{-1}),$$

where F is closed. Hence, by (5), we obtain (6). ■

Assuming that S is separable we obtain the following strengthened version of Theorem 1:

THEOREM 1a. *If $X_n \xrightarrow{D} X$ and $P\{X \in D_h\} = 0$ where S_1 is the Euclidean space and S is a separable metric space, then there exists a sequence of continuous mappings h_k, $k \geqslant 1$, of S into S_1 such that*

(5a) $$h_k(X) \xrightarrow{P} h(X) \quad \textit{as } k \to \infty,$$

(6a) $$\textit{for each } \varepsilon > 0, \ \lim_k \overline{\lim_n}\, P\{\varrho_1(h_k(X_n), h(X_n)) \geqslant \varepsilon\} = 0.$$

Proof. Let us take the sequence of continuous mappings h_k, $k \geqslant 1$, chosen in the proof of Theorem 1. In view of $P\{h_k(X) \neq h(X)\} \leqslant 1/k$ the condition (5a) holds.

Now notice that the set of continuity points of the function $\varrho_1(h_k(x), h(x))$, $x \in S$, contains the set F_k and this function is equal to zero on F_k. Thus for fixed $\varepsilon > 0$, $k \geqslant 1$, and for each $x \in \partial F_k$ there exists an open sphere $N(x, \delta_x)$ with radius $\delta_x > 0$ such that $\varrho_1(h_k(y), h(y)) < \varepsilon$ for $y \in N(x, \delta_x)$. Let

$$G_k = \bigcup_{x \in \partial F_k} N(x, \delta_x) \cup F_k, \quad k \geqslant 1.$$

Then G_k, $k \geqslant 1$, are open and $G_k \supset F_k$, $k \geqslant 1$. Denoting by H_k the complement of G_k and using Theorem 2.1 from [1] we have

$$\overline{\lim_n}\, P\{\varrho_1(h_k(X_n), h(X_n)) \geqslant \varepsilon\} \leqslant \overline{\lim_n}\, P\{X_n \in H_k\} \leqslant P\{X \in H_k\} \leqslant 1/k.$$

Hence we get (6a), which completes the proof. ■

As a consequence of Theorem 1 and Lemma 1 we get the following well-known theorem:

THEOREM 2 (see [1], Theorem 5.1). *If $X_n \xrightarrow{D} X$ as $n \to \infty$ and $P\{X \in D_h\} = 0$, then $h(X_n) \xrightarrow{D} h(X)$ as $n \to \infty$.*

Proof. Let f be any bounded and continuous mapping of S_1 into $\mathbf{R}$. Then $f\circ h$ is a bounded mapping of S into $\mathbf{R}$ and in view of $D_{f\circ h} \subset D_h$ we get $P\{X \in D_{f\circ h}\} = 0$. Hence and from Theorem 1 we infer that there exists a sequence of real-valued mappings g_k, $k \geqslant 1$, on S which are bounded, continuous and satisfy the following conditions:

$$g_k(X) \xrightarrow{D} f\circ h(X) \quad \text{as } k \to \infty,$$

$$\lim_k \overline{\lim_n}\, \varrho_P\big(\mathscr{L}(g_k(X_n)),\ \mathscr{L}(f\circ h(X_n))\big) = 0.$$

Thus by Lemma 1 we obtain $f\circ h(X_n) \xrightarrow{D} f\circ h(X)$ as $n \to \infty$. Now, because $f\circ h$ is bounded, $\mathrm{E} f\circ h(X_n) \to \mathrm{E} f\circ h(X)$ as $n \to \infty$. This and the fact that f was an arbitrary, continuous and bounded mapping of S_1 into $\mathbf{R}$ give the convergence $h(X_n) \xrightarrow{D} h(X)$ as $n \to \infty$, which completes the proof. ∎

In the situation $X_n \xrightarrow{D} X$ as $n \to \infty$, Lemma 1 will be more suitable for an examination of $h(X_n) \xrightarrow{D} h(X)$ if we choose mappings h_k, $k \geqslant 1$, in a way such that $P\{X \in D_{h_k}\} = 0$, $k \geqslant 1$. Then in view of Theorem 2 we must only verify (2) and (3). For clarity let us introduce the following notions:

DEFINITION. A sequence of mappings h_k, $k \geqslant 1$, is said to *approximate h on X* either (i) *in probability* or (ii) *in distribution* if h_k, $k \geqslant 1$, are measurable mappings of S into S_1 such that

A. $P\{X \in D_{h_k}\} = 0$ for each $k \geqslant 1$,

and

Ba. $h_k(X) \xrightarrow{P} h(X)$ as $k \to \infty$ in the case (i),

B. $h_k(X) \xrightarrow{D} h(X)$ as $k \to \infty$ in the case (ii).

DEFINITION. A sequence of mappings h_k, $k \geqslant 1$, is said to *approximate h on the sequence $\{X_n\}$* either (i) *in probability* or (ii) *in distribution* if h_k are measurable mappings of S into S_1 such that

Ca. for each $\varepsilon > 0$, $\lim_k \overline{\lim}_n P\{\varrho_1(h_k(X_n), h(X_n)) > \varepsilon\} = 0$ in the case (i),

C. $\lim_k \overline{\lim}_n \varrho_P\big(\mathscr{L}(h_k(X_n)), \mathscr{L}(h(X_n))\big) = 0$ in the case (ii).

In this terminology Lemma 1 and Theorem 2 give the following

COROLLARY 1. *Let $X_n \xrightarrow{D} X$ as $n \to \infty$ and let $\{h_k\}$ approximate h on X in distribution. Then the following conditions are equivalent*:

C. *the sequence $\{h_k\}$ approximates h on $\{X_n\}$ in distribution,*

D. *$h(X_n) \xrightarrow{D} h(X)$ as $n \to \infty$.*

To obtain conditions under which the implication

$$[X_n \xrightarrow{D} X] \Rightarrow [(X_n, h(X_n)) \xrightarrow{D} (X, h(X))]$$

holds it is enough to use Corollary 1 by replacing h and h_k, $k \geqslant 1$, with mappings $\tilde{h}$ and $\tilde{h}_k$, $k \geqslant 1$, respectively, defined as $\tilde{h}(x) = (x, h(x))$ and $\tilde{h}_k(x) = (x, h_k(x))$. Then we have

COROLLARY 2. *Let $X_n \xrightarrow{D} X$ as $n \to \infty$ and let the sequence $\{\tilde{h}_k\}$ approximate $\tilde{h}$ on X in distribution. Then the condition that $\{\tilde{h}_k\}$ approximates $\tilde{h}$ on the sequence $\{X_n\}$ in distribution is equivalent to $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$ as $n \to \infty$.*

Let us notice that the condition that $\{h_k\}$ approximates h on X in distribution does not imply the condition that $\{\tilde{h}_k\}$ approximates $\tilde{h}$ on X in distribution. Similarly, the condition that $\{h_k\}$ approximates h on the sequence $\{X_n\}$ in distribution does not imply that $\{\tilde{h}_k\}$ approximates $\tilde{h}$ on $\{X_n\}$ in distribution. However, these implications hold if the approximation in distribution is replaced by the approximation in probability. But then the following problem of a relation between the conditions arises: $\{h_k\}$ approximates h on $\{X_n\}$ in probability and $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$. A solution of this problem is given by Theorem 3 below. Before stating it let us formulate the following lemma:

LEMMA 2. *Let $X_n \xrightarrow{D} X$ as $n \to \infty$ and let the metric spaces S and S_1 be separable. Furthermore, assume that there exists a sequence of mappings h_k, $k \geqslant 1$, approximating h on X in probability. Then the following conditions are equivalent:*

(7) *for each* $i \geqslant 1$, $\lim_k \overline{\lim}_n \varrho_P\big(\mathscr{L}(h_i(X_n), h_k(X_n)), \mathscr{L}(h_i(X_n), h(X_n))\big) = 0$,

(8) *for each* $\varepsilon > 0$, $\lim_k \overline{\lim}_n P\{\varrho_1(h_k(X_n), h(X_n)) \geqslant \varepsilon\} = 0$,

(9) *for each* $i \geqslant 1$, $(h_i(X_n), h(X_n)) \xrightarrow{D} (h_i(X), h(X))$ *as* $n \to \infty$.

Proof. Assume $Z = h(X)$, $Z_{k,n} = h_k(X_n)$, $Z_k = h_k(X)$ and $Y_n = h(X_n)$, $n, k \geqslant 1$. Then the convergences $X_n \xrightarrow{D} X$ as $n \to \infty$ and $h_k(X) \xrightarrow{P} h(X)$ as $k \to \infty$ and the condition A imply that the array $\{Z_{k,n}, k, n \geqslant 1\}$ and the sequences $\{Z_k, k \geqslant 1\}$ and $\{Y_n, n \geqslant 1\}$ satisfy the conditions A_2 and B_1a from Section 2. Hence using the implication (iii) of Proposition 3 we get the assertion. ∎

THEOREM 3. *Let $X_n \xrightarrow{D} X$ as $n \to \infty$ and let the metric spaces S and S_1 be separable. Furthermore, assume that $\{h_k\}$ approximates h on X in probability. Then the following conditions are equivalent:*

Ca. *the sequence $\{h_k\}$ approximates h on the sequence $\{X_n\}$ in probability,*

Da. $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$ *as* $n \to \infty$.

Proof. Let us define mappings $\tilde{h}$ and $\tilde{h}_k$, $k \geqslant 1$, of the space S into $S \times S_1$ as follows: $\tilde{h}(x) = (x, h(x))$ and $\tilde{h}_k(x) = (x, h_k(x))$ for $x \in S$ and $k \geqslant 1$. Obviously, these mappings are measurable and the set of discontinuities of $\tilde{h}_k$, i.e. $D_{\tilde{h}_k}$, is a subset of D_{h_k}, $k \geqslant 1$. Hence, by the condition A, we have $P\{X \in D_{\tilde{h}_k}\} = 0$ and, by the convergence $h_k(X) \xrightarrow{P} h(X)$ as $k \to \infty$, we get

$$\tilde{h}_k(X) = (X, h_k(X)) \xrightarrow{P} (X, h(X)) = \tilde{h}(X) \quad \text{as } k \to \infty.$$

The last facts mean that the sequence of mappings $\tilde{h}_k$, $k \geqslant 1$, approximates $\tilde{h}$ on X in probability. Thus using Lemma 2 we have the equivalence of the conditions

$\tilde{\mathrm{C}}$a. for each $\varepsilon > 0$, $\lim_k \overline{\lim}_n P\{\tilde{\varrho}(\tilde{h}_k(X_n), \tilde{h}(X_n)) \geqslant \varepsilon\} = 0$

and

$\tilde{\mathrm{D}}$. for each $i \geqslant 1$, $(\tilde{h}_i(X_n), \tilde{h}(X_n)) \xrightarrow{D} (\tilde{h}_i(X), \tilde{h}(X))$ as $n \to \infty$,

where $\tilde{\varrho}((x_1, y_1), (x_2, y_2)) = \varrho(x_1, x_2) + \varrho_1(y_1, y_2)$ for $x_i \in S$, $y_i \in S_1$, $i = 1, 2$. But the condition $\tilde{\mathrm{C}}$a is equivalent to the condition:

$$\text{for each } \varepsilon > 0, \ \lim_k \overline{\lim_n} P\{\varrho_1(h_k(X_n), h(X_n)) \geqslant \varepsilon\} = 0,$$

while the condition $\tilde{\mathrm{D}}$, in view of A, is equivalent to the convergence $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$ as $n \to \infty$. This completes the proof.

The following remark makes the conditions of Theorem 3 more clear to the investigation of the weak convergence of vector-valued mappings on the sequence $\{X_n\}$.

Remark 1. Let h^i, $1 \leqslant i \leqslant m$, be measurable mappings of a separable metric space S into separable metric spaces S_i, $1 \leqslant i \leqslant m$, respectively, and let h be the mapping on S defined as

$$h(x) = (h^1(x), h^2(x), \ldots, h^m(x)) \quad \text{for } x \in S.$$

If for each h^i, $1 \leqslant i \leqslant m$, there exists a sequence of measurable mappings h_k^i, $k \geqslant 1$, of S into S_i such that this sequence approximates h^i on X in probability, then the sequence of mappings h_k, $k \geqslant 1$, defined as

$$h_k(x) = (h_k^1(x), h_k^2(x), \ldots, h_k^m(x)) \quad \text{for } x \in S$$

approximates the mapping h on X in probability. Similarly, if for each h^i, $1 \leqslant i \leqslant m$, the sequence of mappings h_k^i, $k \geqslant 1$, approximates h^i on the sequence $\{X_n\}$ in probability, then the sequence of mappings h_k, $k \geqslant 1$, approximates h on the sequence $\{X_n\}$ in probability.

Note 2. The identity mapping is approximated in probability by the sequence of identity mappings on each X and each $\{X_n\}$.

As an immediate consequence of Remark 1 and Theorem 3 we get the following corollary:

COROLLARY 3. *Let $X_n \xrightarrow{D} X$ as $n \to \infty$ where S is separable and let h^i, $1 \leqslant i \leqslant m$, be measurable mappings of S into separable metric spaces S_i, respectively. If furthermore for each h^i, $1 \leqslant i \leqslant m$, there exists a sequence of* *mappings* h_k^i, $k \geqslant 1$, *which approximates* h^i *on* X *in probability, then the convergence*

$$(10) \quad (X_n, h^1(X_n), h^2(X_n), \ldots, h^m(X_n)) \xrightarrow{D} (X, h^1(X), \ldots, h^m(X)) \quad \text{as } n \to \infty$$

holds iff the sequence $\{h_k^i, k \geqslant 1\}$ *approximates* h^i *on the sequence* $\{X_n\}$ *in probability for each* i, $1 \leqslant i \leqslant m$.

In queueing theory one considers queueing systems which are periodic in time. Here in place of the convergence $\mathscr{L}(X_n) \Rightarrow \mathscr{L}(X)$ we put the convergence

$$n^{-1} \sum_{i=1}^{n} \mathscr{L}(X_i) \Rightarrow \mathscr{L}(X) \quad \text{as } n \to \infty.$$

In this situation the problem formulated in Section 1 takes the following form: for which sequences $\{X_n\}$ does the convergence

$$(11) \qquad n^{-1} \sum_{i=1}^{n} \mathscr{L}(X_i) \Rightarrow \mathscr{L}(X) \quad \text{as } n \to \infty$$

give the convergence

$$(12) \qquad n^{-1} \sum_{i=1}^{n} \mathscr{L}(X_i, h(X_i)) \Rightarrow \mathscr{L}(X, h(X)) \quad \text{as } n \to \infty?$$

An answer to this problem follows from Theorem 3 and takes the following form:

COROLLARY 4. *Assume that* (11) *holds and that there exists a sequence of mappings* h_k, $k \geqslant 1$, *approximating* h *on* X *in probability. Then* (12) *holds iff for each* $\varepsilon > 0$

$$(13) \qquad \lim_k \overline{\lim_n}\, n^{-1} \sum_{i=1}^{n} P\{\varrho_1(h_k(X_i), h(X_i)) \geqslant \varepsilon\} = 0.$$

4. The case $S = D(T)$. We restrict ourselves to the case where S is the space of right-continuous functions on $T \subset R$, i.e. $S = D(T)$. In this situation for any measurable mapping h of $D(T)$ into a separable metric space S_1 we indicate an example of a sequence of mappings h_k, $k \geqslant 1$, which approximates h on a random element X of the space $D(T)$ in probability. Next we restrict ourselves to the case where the mapping h at $x \in D(T)$ is equal to $\sup_{t \geqslant 0} x(t)$ if it is finite and zero otherwise. In this situation we give conditions on X and on a sequence $\{X_n\}$ of random elements of $D(T)$ under which the sequence of mappings h_k,

$$h_k(x) = \sup_{0 \leqslant t \leqslant c_k} x(t), \quad k \geqslant 1, \text{ where } c_k \uparrow \infty,$$

approximates h on X and on $\{X_n\}$ in probability.

Let T be a subinterval of the real line. T can be finite or infinite and, if finite, open or closed. Let $D(T)$ be the set of all right-continuous real-valued functions on T with limits from the left. Let $D(T)$ have Skorohod's J_1-topology or its natural extension to non-compact intervals: a sequence $\{x_n, n \geqslant 1\}$ converges to x in $D(T)$ if the restrictions of x_n converge to the restrictions of x in $D[a, b]$ for each compact interval $[a, b] \subset T$ such that a and b are continuity points of x or endpoints of T. This mode of convergence agrees with the previous extension of the J_1-topology given by Stone and Lindvall (see [8]). In the case $T = [0, \infty)$ we can consider $D(T)$ with Lindvall's metric defined in [3] or with Whitt's metric defined in [8] while in the case $T = (-\infty, \infty)$ we consider $D(T)$ with Whitt's metric. $D(T)$ with the mentioned metrics is a Polish metric space.

Let $\{c_k\} = \{c_k, k \geqslant 1\}$ be an increasing sequence of positive numbers tending to infinity and let $\{r_k\} = \{r_k, k \geqslant 1\}$ be the sequence of mappings of $D(T)$ into $D(T)$ defined as

$$r_k(x)(t) = \begin{cases} x(t) & \text{for } 0 \leqslant t < c_k, \\ x(c_k) & \text{for } t \geqslant c_k \end{cases}$$

when $T = [0, \infty)$, and as

$$r_k(x)(t) = \begin{cases} x(-c_k) & \text{for } t < -c_k, \\ x(t) & \text{for } -c_k \leqslant t < c_k, \\ x(c_k) & \text{for } t \geqslant c_k \end{cases}$$

when $T = (-\infty, \infty)$.

Let h be a mapping of $D(T)$ into a separable metric space S_1 and let h_k, $k \geqslant 1$, be mappings of $D(T)$ into S_1 defined as $h_k(x) = h(r_k(x))$ for $x \in D(T)$. Obviously, the mappings h_k, $k \geqslant 1$, are measurable and the following fact holds:

Remark 2. Let X be a random element of $D(T)$. The sequence of mappings h_k, $k \geqslant 1$, defined above approximates h on X in probability if

$$h(r_k(X)) \xrightarrow{p} h(X) \quad \text{as } k \to \infty$$

and

$$P\{X(c_k) = X(c_k-)\} = 1 \quad \text{for each } k \geqslant 1$$

when $T = [0, \infty)$, while

$$P\{X(c_k) = X(c_k-)\} = P\{X(-c_k) = X(-c_k-)\} = 1 \quad \text{for each } k \geqslant 1$$

when $T = (-\infty, \infty)$.

Henceforth, let h be defined on $D(T)$ as follows: h at $x \in D(T)$ is equal to $\sup_{t \geqslant 0} x(t)$ if it is finite and h is equal to zero otherwise. Obviously, this mapping is measurable. As in Remark 2, let us define mappings h_k, $k \geqslant 1$, on

$D(T)$ as

$$h_k(x) = \sup_{0 \leqslant t \leqslant c_k} x(t),$$

where $\{c_k\}$ is an increasing sequence of positive numbers tending to infinity. The case where h is defined at $x \in D(-\infty, \infty)$ as

$$h(x) = \sup_{t \leqslant 0} x(t)$$

if it is finite and zero otherwise reduces itself to the earlier case, i.e.

$$h(x) = \sup_{t \geqslant 0} x(t).$$

Henceforth, let X and X_n, $n \geqslant 1$, be random elements of $D(T)$ satisfying the condition

a_1. $P\{\sup_{t \geqslant 0} X(t) < \infty\} = 1$ and $P\{\sup_{t \geqslant 0} X_n(t) < \infty\} = 1$ for $n \geqslant 1$.

Immediately from Remark 2 we get the following fact:

Remark 3. If X satisfies the condition a_1 and the condition

a_2. $P\{X(c_k) = X(c_k -)\} = 1$, $k \geqslant 1$,

then the sequence of mappings h_k, $k \geqslant 1$, approximates the mapping h on X in probability.

Now let us notice that the condition

a_3. for each $\varepsilon > 0$, $\lim_k \overline{\lim_n} P\{\sup_{t \geqslant 0} X_n(t) - \sup_{0 \leqslant t \leqslant c_k} X_n(t) > \varepsilon\} = 0$,

means that the sequence of mappings h_k, $k \geqslant 1$, approximates the mapping h on $\{X_n\}$ in probability.

Obviously, the condition a_3 is weaker than the condition

$\tilde{a}_3$. $\lim_k \overline{\lim_n} P\{h_k(X_n) \neq h(X_n)\} = 0$.

However, as we see soon, $\tilde{a}_3$ is equivalent to a_3 in some class of stochastic processes X_n, $n \geqslant 1$ (see Lemma 3 and Note 3).

LEMMA 3. *Assume that the following conditions hold:*

a_4. *for each* $k \geqslant 1$, $X_n(c_k) \xrightarrow{D} X(c_k)$ *as* $n \to \infty$,

a_5. $X(c_k) \xrightarrow{P} -\infty$ *as* $k \to \infty$,

a_6. *for any* $\varepsilon > 0$ *there exists* $b > 0$ *such that for all* $n, k \geqslant 1$

$$P\{\sup_{t \geqslant 0} X_n(t) < -b\} < \varepsilon \quad \textit{and} \quad P\{\sup_{0 < t < \infty} (X_n(t + c_k) - X_n(c_k)) > b\} < \varepsilon.$$

Then the condition $\tilde{a}_3$ *holds.*

Proof. We put

$$\theta_{k,n} = \sup_{0\leqslant t\leqslant c_k} X_n(t), \quad n, k \geqslant 1,$$

$$\eta_{k,n} = \sup_{0\leqslant t<\infty} (X_n(t+c_k)-X_n(c_k)), \quad n, k \geqslant 1.$$

Now let us notice that for any $x \in D(T)$ and $c > 0$ we have

$$\begin{aligned}\sup_{t\geqslant 0} x(t) - \sup_{0\leqslant t\leqslant c} x(t) &= \max\{\sup_{0\leqslant t\leqslant c} x(t), \sup_{t>c} x(t)\} - \sup_{0\leqslant t\leqslant c} x(t)\\ &= \max\{0, \sup_{t>c} x(t) - \sup_{0\leqslant t\leqslant c} x(t)\}\\ &= \max\{0, \sup_{0<t<\infty} (x(t+c)-x(c))+x(c) - \sup_{0\leqslant t\leqslant c} x(t)\}.\end{aligned}$$

Hence

$$P\{\sup_{t\geqslant 0} X_n(t) - \sup_{0\leqslant t\leqslant c_k} X_n(t) > 0\} = P\{\eta_{k,n} - \theta_{k,n} + X_n(c_k) > 0\}.$$

But by a_6 the above does not exceed

$$2\varepsilon + P\{\eta_{k,n} - \theta_{k,n} + X_n(c_k) > 0,\ \eta_{k,n} \leqslant b,\ \theta_{k,n} \geqslant -b\} \leqslant 2\varepsilon + P\{X_n(c_k) > -2b\}.$$

Hence and from a_4 we have

$$\begin{aligned}\overline{\lim_n} P\{\sup_{t\geqslant 0} X_n(t) - \sup_{0\leqslant t\leqslant c_k} X_n(t) > 0\} &\leqslant 2\varepsilon + \overline{\lim_n} P\{X_n(c_k) > -2b\}\\ &\leqslant 2\varepsilon + \overline{\lim_n} P\{X_n(c_k) \geqslant -2b\} \leqslant 2\varepsilon + P\{X(c_k) \geqslant -2b\}.\end{aligned}$$

Now by a_5 we get

$$\lim_k \overline{\lim_n} P\{\sup_{t\geqslant 0} X_n(t) - \sup_{0\leqslant t\leqslant c_k} X_n(t) > 0\} \leqslant 2\varepsilon.$$

Since ε was arbitrary, we get the assertion. ∎

Note 3. From the proof of Lemma 3 it follows that if

$$\lim_k \overline{\lim_n} P\{X_n(c_k) > -b\} = 0 \quad \text{for each } b > 0$$

and a_6 hold, then $\tilde{a}_3$ holds.

The following remark gives the sufficient conditions for a_6.

Remark 4. The condition a_6 holds whenever the sequence $\{\sup_{t\geqslant 0} X_n(t), n \geqslant 1\}$ is tight and one of the following conditions (a), (b) or (c) holds:

(a) the sequence $\{\sup_{0<t<\infty}(X_n(t+c_k)-X_n(c_k)),\ n, k \geqslant 1\}$ is tight;

(b) X_n, $n \geqslant 1$, have stationary increments;

(c) for each $n, k \geqslant 1$ and $x > 0$ the following inequality holds:

$$P\{\sup_{0<t<\infty} (X_n(t+c_k)-X_n(c_k)) > x\} \leqslant P\{\sup_{t\geqslant 0} X_n(t) > x\}.$$

Now, using Remarks 2–4, Lemma 3 and Theorem 3 we get

COROLLARY 5. *Let X and X_n, $n \geqslant 1$, be random elements of $D(T)$ satisfying* a_1 *and* a_2. *Then the following implications hold:*

(i_1) *The condition* $\tilde{a}_3$ *and the convergence* $X_n \xrightarrow{D} X$ *imply the convergence* a_7. $(X_n, h(X_n)) \xrightarrow{D} (X, h(X))$ *as* $n \to \infty$.

(i_2) *The conditions* a_7 *and* a_5 *and one of the conditions* (a), (b) *or* (c) *given in Remark* 4 *imply* $\tilde{a}_3$.

To consider the asymptotic stationarity of queueing systems it is useful to have an analogue of Corollary 5 in the case where $D(-\infty, \infty)$ is replaced by the space $R^{\infty}_{-\infty}$, i.e. the space of sequences $x = \{x_k, -\infty < k < \infty\}$, where $x_k \in R$, while h at $x \in R^{\infty}_{-\infty}$ is equal to $\sup_{j\leqslant 0} x_j$ if it is finite and zero otherwise. Obviously, each element x of $R^{\infty}_{-\infty}$ can be meant as an element x of $D(-\infty, \infty)$ if we write $x(t) = x_{[t]}$, $t \in R$. Now, since $\sup_{j\leqslant 0} x_j = \sup_{j\geqslant 0} x_{-j} = \sup_{t\geqslant 0} x(-t)$, the case considered now is a special case of the case considered in Corollary 5. In spite of this we rewrite Corollary 5 in a suitable form for our later applications.

Let S_2 be a separable metric space and let (ξ, Y) and (ξ_n, Y_n), $n \geqslant 1$, be random elements of the metric space $R^{\infty}_{-\infty} \times S_2$ such that ξ and ξ_n, $n \geqslant 1$, are random elements of $R^{\infty}_{-\infty}$ while Y and Y_n, $n \geqslant 1$, are random elements of S_2. The random elements ξ and ξ_n, $n \geqslant 1$, are written as $\xi = \{\xi_k, -\infty < k < \infty\}$ and $\xi_n = \{\xi_{n,k}, -\infty < k < \infty\}$, $n \geqslant 1$, where ξ_k and $\xi_{n,k}$ are random variables. Henceforth we assume that ξ and ξ_n, $n \geqslant 1$, satisfy the following condition:

$$(14) \qquad P\{\sup_{j\leqslant 0} \xi_j < \infty\} = 1 \quad \text{and} \quad P\{\sup_{j\leqslant 0} \xi_{n,j} < \infty\} = 1, \quad n \geqslant 1.$$

Now, compiling Corollary 5, Theorem 3 and Corollary 3 we get

COROLLARY 6. *Under the assumed conditions the following implications hold:*

(i_1) *If* $(\xi_n, Y_n) \xrightarrow{D} (\xi, Y)$ *as* $n \to \infty$ *and*

$$(15) \qquad \lim_k \overline{\lim_n} P\{\sup_{j\leqslant 0} \xi_{n,j} - \sup_{-k\leqslant j\leqslant 0} \xi_{n,j} > 0\} = 0,$$

then

$$(16) \qquad (\xi_n, \sup_{j\leqslant 0} \xi_{n,j}, Y_n) \xrightarrow{D} (\xi, \sup_{j\leqslant 0} \xi_j, Y) \quad \text{as } n \to \infty.$$

(i_2) *On the contrary, if the convergence*

$$(\xi_n, \sup_{j\leqslant 0} \xi_{n,j}) \xrightarrow{D} (\xi, \sup_{j\leqslant 0} \xi_j) \quad \text{as } n \to \infty$$

holds and furthermore $\xi_{-k} \xrightarrow{P} -\infty$ *as* $k \to \infty$ *and the sequence* $\{\sup_{j\leqslant 0}(\xi_{n,j-k} - \xi_{n,-k}),\ n, k \geqslant 1\}$ *is tight, then* (15) *holds true.*

Now we give an analogue of Corollary 6 in the case where instead of the convergence $\mathscr{L}(\xi_n, Y_n) \Rightarrow \mathscr{L}(\xi, Y)$ we consider the convergence

$$(17)\qquad n^{-1} \sum_{i=1}^{n} \mathscr{L}(\xi_i, Y_i) \Rightarrow (\xi, Y) \quad \text{as } n \to \infty,$$

written as $\mathscr{L}(\xi, Y) \stackrel{m}{\Longrightarrow} \mathscr{L}(\xi, Y)$ and called the *weak convergence in mean.*

Corollary 6a. *If* ξ *and* ξ_n, $n \geqslant 1$, *satisfy* (14), *then the following implications hold*:

(i_1) *If* (17) *holds and*

$$(18)\qquad \lim_k \overline{\lim_n}\, n^{-1} \sum_{i=1}^{n} P\{\sup_{j \leqslant 0} \xi_{i,j} - \sup_{-k \leqslant j \leqslant 0} \xi_{i,j} > 0\} = 0,$$

then

$$(19)\qquad \mathscr{L}(\xi_n, \sup_{j \leqslant 0} \xi_{n,j}, Y_n) \stackrel{m}{\Longrightarrow} \mathscr{L}(\xi, \sup_{j \leqslant 0} \xi_j, Y) \quad \text{as } n \to \infty.$$

(i_2) *On the contrary, if the convergences*

$$\mathscr{L}(\xi_n, \sup_{j \leqslant 0} \xi_{n,j}) \stackrel{m}{\Longrightarrow} \mathscr{L}(\xi, \sup_{j \leqslant 0} \xi_j) \text{ as } n \to \infty \quad \text{and} \quad \xi_{-k} \stackrel{P}{\longrightarrow} -\infty \text{ as } k \to \infty$$

hold and the sequence of probability measures

$$n^{-1} \sum_{i=1}^{n} \mathscr{L}(\sup_{j \leqslant 0}(\xi_{i,j-k} - \xi_{i,-k})), \quad n, k \geqslant 1,$$

is tight, then (18) *holds.*

Example 1. Here we illustrate an application of Corollaries 6 and 6a to prove Theorem 1 from [4]. The proof of this theorem given here is easier and clearer than the proof given in [4]. Besides, the case (ii) of Theorem 4 formulated below is stronger than the case (ii) of Theorem 1 from [4]. Moreover, Condition AB formulated below and Condition AB in mean are free of an initial condition w_1, however Condition AB under the assumptions of Theorem 4 is equivalent to Condition AB from [4]. Other applications of Theorem 4 are given in [5].

Let $(\boldsymbol{v}, \boldsymbol{u}) = \{(v_k, u_k), k \geqslant 1\}$ be a generic sequence for a single server queue (see [4]), let w_k be the waiting time of the k-th unit, and let $(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u}) = \{(w_k, v_k, u_k), k \geqslant 1\}$. Furthermore, denote by $(\boldsymbol{v}^0, \boldsymbol{u}^0) = \{(v_k^0, u_k^0), k \geqslant 1\}$ a stationary representation of $(\boldsymbol{v}, \boldsymbol{u})$ in the sense of weak convergence or weak convergence in mean (see [4]), and by $(\boldsymbol{v}^*, \boldsymbol{u}^*) = \{(v_k^*, u_k^*), -\infty < k < \infty\}$ a two-sided stationary sequence such that $\{(v_k^*, u_k^*), k \geqslant 1\}$ and $(\boldsymbol{v}^0, \boldsymbol{u}^0)$ have the same distribution. Also let $X_k = v_k - u_k$, $S_0 = 0$, $S_k = X_1 + X_2 + \ldots + X_k$, $k \geqslant 1$, and $X_k^* = v_k^* - u_k^*$, $S_0^* = 0$, $S_k^* = \sum_{i=k+1}^{0} X_i^*$ for $k < 0$.

We say that (v, u) *satisfies Condition* AB or *Condition* AB *in mean* if

$$\lim_k \overline{\lim_n} P\{\max_{k\leqslant j\leqslant n} (S_n - S_{n-j}) > 0\} = 0$$

or

$$\lim_k \overline{\lim_n} n^{-1} \sum_{i=1}^{n} P\{\max_{k\leqslant j\leqslant i} (S_i - S_{i-j}) > 0\} = 0,$$

respectively.

THEOREM 4 (see [4], Theorem 1). *Let (v, u) be either* (i) *weakly asymptotically stationary or* (ii) *weakly asymptotically stationary in mean and assume that it satisfies Condition* AB *in the case* (i) *and Condition* AB *in mean in the case* (ii). *Furthermore, let the stationary representation (v^0, u^0) in both cases be such that $S^*_{-k} \to -\infty$ a.e. as $k \to \infty$. Then the sequence (w, v, u) is weakly asymptotically stationary in the case* (i) *and weakly asymptotically stationary in mean in the case* (ii). *Moreover, the stationary representation of (w, v, u) is given by* (4.14) *in* [4].

Proof. First we show that $S_n \xrightarrow{P} -\infty$ as $n \to \infty$ in the case (i) and $n^{-1} \sum_{j=1}^{n} P\{S_j > -a\} \to 0$ as $n \to \infty$, for any $a > 0$, in the case (ii). For that purpose let us notice that Condition AB and Condition AB in mean imply

$$P\{S_n > 0\} \to 0 \quad \text{and} \quad n^{-1} \sum_{j=1}^{n} P\{S_j > 0\} \to 0 \quad \text{as } n \to \infty,$$

respectively. Hence for any $\varepsilon > 0$ and $b > 0$ there exists n_1 such that for $n > n_1$ we have $P\{S_n > b\} < \varepsilon$ in the first case and $n^{-1} \sum_{j=1}^{n} P\{S_j > b\} < \varepsilon$ in the second case. But for any $a > 0$ we have

$$P\{S_n < -a\} \geqslant P\{S_n < -a, S_{n-k} < a/2\} \geqslant P\{S_n - S_{n-k} < -a/2, S_{n-k} < a/2\}$$
$$\geqslant P\{S_n - S_{n-k} < -a/2\} - P\{S_{n-k} > a/2\}$$

and

$$n^{-1} \sum_{j=1}^{n} P\{S_j < -a\} \geqslant n^{-1} \sum_{j=k}^{n} P\{S_j - S_{j-k} < -a/2\} - n^{-1} \sum_{j=k}^{n} P\{S_{j-k} > a/2\}.$$

By the weak asymptotic stationarity of $v - u$ in the first case and the weak asymptotic stationarity in mean in the second case, for any $a > 0$, $\varepsilon > 0$ and $k \geqslant 1$ there exists n_k such that for $n > n_k$ we have

$$P\{S_n - S_{n-k} < -a\} \geqslant P\{S^*_{-k} < -a\} - \varepsilon \quad \text{in the first case}$$

and

$$n^{-1} \sum_{j=k}^{n} P\{S_j - S_{j-k} < -a\} \geqslant P\{S^*_{-k} < -a\} - \varepsilon \quad \text{in the second case.}$$

Now, in view of $S^*_{-k} \to -\infty$ a.e. we see that for any $\varepsilon > 0$ there exists k_0 such that for $k > k_0$ we have $P\{S^*_{-k} < -a\} \geqslant 1 - \varepsilon$. Compiling the above facts we infer that for any $\varepsilon > 0$ and $a > 0$ there exists n_0 such that $P\{S_n < -a\} \geqslant 1-3\varepsilon$ in the first case and $n^{-1}\sum_{j=1}^{n} P\{S_j < -a\} \geqslant 1-3\varepsilon$ in the second case, which gives the required convergences.

Now define sequences $(\tilde{v}_n, \tilde{u}_n) = \{(\tilde{v}_{n,k}, \tilde{u}_{n,k}), -\infty < k < \infty\}$, $\xi_n = \{\xi_{n,k}, -\infty < k < \infty\}$, $n \geqslant 1$, and $\xi = \{\xi_k, -\infty < k < \infty\}$ in the following way:

$$\tilde{v}_{n,k} = v_{n+k},\ \tilde{u}_{n,k} = u_{n+k} \text{ for } k > -n+1 \quad \text{and} \quad \tilde{v}_{n,k} = \tilde{u}_{n,k} = 0 \text{ for } k \leqslant -n,$$

$$\xi_{n,k} = \sum_{i=k+1}^{0} (\tilde{v}_{n,i} - \tilde{u}_{n,i}) \text{ for } k < 0 \quad \text{and} \quad \xi_{n,k} = 0 \text{ for } k \geqslant 0,$$

$$\xi_k = \sum_{i=k+1}^{0} (v_i^* - u_i^*) \text{ for } k < 0 \quad \text{and} \quad \xi_k = 0 \text{ for } k \geqslant 0.$$

Then in the case (i) we have the convergences

$$(20) \qquad \mathscr{L}(\tilde{v}_n, \tilde{u}_n) \Rightarrow \mathscr{L}(v^*, u^*) \quad \text{and} \quad \mathscr{L}(\xi_n) \Rightarrow \mathscr{L}(\xi) \quad \text{as } n \to \infty,$$

while in the case (ii) we have (20) with $\overset{m}{\Longrightarrow}$ instead of $\Rightarrow$. Moreover, $\xi_{-k} = S^*_{-k} \to -\infty$ a.e. as $k \to \infty$. Since

$$w_{n+1} = \max\bigl(S_n + w_1, \max_{0 \leqslant j \leqslant n} (S_n - S_{n-j})\bigr), \quad n \geqslant 1,$$

we have

$$(21) \qquad w_{n+1} = \max(\xi_{n,-n} + w_1, \sup_{j \leqslant 0} \xi_{n,j}), \quad n \geqslant 1.$$

In view of Lemma 14 from [4], to show the convergence $\mathscr{L}(w_n, v_n, u_n) \Rightarrow \mathscr{L}(w^0, v^0, u^0)$ in the case (i) and the convergence $\overset{m}{\Longrightarrow}$ in the case (ii) it is enough to show the convergence $\mathscr{L}(w_{n+1}, v_n, u_n) \Rightarrow \mathscr{L}(w_1^0, v^0, u^0)$ in the case (i) and the convergence $\overset{m}{\Longrightarrow}$ in the case (ii). Here

$$(w_n, v_n, u_n) = \{(w_{n+k}, v_{n+k}, u_{n+k}), k \geqslant 1\}.$$

Define the mapping f: $R^{\infty}_{-\infty} \times R^{\infty}_{-\infty} \to R^{\infty}_{-\infty}$ as

$$f(x, y) = \{z_k, -\infty < k < \infty\},$$

where

$$x = \{x_k, -\infty < k < \infty\}, \quad y = \{y_k, -\infty < k < \infty\}$$

and

$$z_k = \sum_{i=k+1}^{0} (x_i - y_i) \text{ for } k < 0 \quad \text{and} \quad z_k = 0 \text{ for } k \geqslant 0.$$

Now let us notice that

$$\sup_{j \leqslant 0} \xi_{n,j} = h\bigl(f(\tilde{v}_n, \tilde{u}_n)\bigr),$$

where $h(x) = \sup_{j\leqslant 0} x_j$ if it is finite and zero otherwise. Hence

$$(w_{n+1}, \tilde{v}_n, \tilde{u}_n) = (\max(\xi_{n,-n}+w_1, \sup_{j\leqslant 0}\xi_{n,j}), \tilde{v}_n, \tilde{u}_n)$$
$$= \big(\max\big(\xi_{n,-n}+w_1, h(f(\tilde{v}_n, \tilde{u}_n))\big), \tilde{v}_n, \tilde{u}_n\big).$$

But in view of (20) and of the continuity of f we have the convergence

(22) $$\mathscr{L}(I_0\cdot(\xi_{n,-n}+w_1), f(\tilde{v}_n, \tilde{u}_n), \tilde{v}_n, \tilde{u}_n) \Rightarrow \mathscr{L}(0, f(v^*, u^*), v^*, u^*)$$

in the case (i) and the convergence $\overset{m}{\Longrightarrow}$ in the case (ii), where I_0 denotes the indicator of $\{0\}$. Now let us notice that

(23) $$P\{\sup_{j\leqslant 0}\xi_{n,j} - \sup_{-k\leqslant j\leqslant 0}\xi_{n,j} > 0\}$$
$$= P\{\sup_{j\leqslant -k}\xi_{n,j} - \sup_{-k\leqslant j\leqslant 0}\xi_{n,j} > 0\} \leqslant P\{\max_{k\leqslant j\leqslant n}(S_n - S_{n-j}) > 0\}.$$

Since (v, u) satisfies Condition AB with the initial condition w_1 in the case (i) and Condition AB in mean with the initial condition w_1 in the case (ii), by (23) the sequence $\{\xi_n\}$ satisfies condition (15) of Corollary 6 in the case (i) and condition (18) of Corollary 6a in the case (ii). Thus using Corollary 6 in the case (i) and Corollary 6a in the case (ii) with the following specification:

$$Y_n = (I_0\cdot(\xi_{n,-n}+w_1), \tilde{v}_n, \tilde{u}_n), \qquad Y = (0, v^*, u^*),$$
$$\xi_n = f(\tilde{v}_n, \tilde{u}_n), \qquad \xi = f(v^*, u^*), \qquad n \geqslant 1,$$

we get the convergence $\mathscr{L}(w_{n+1}, \tilde{v}_n, \tilde{u}_n) \Rightarrow \mathscr{L}(w_1^0, v^*, u^*)$ in the case (i) and the convergence $\overset{m}{\Longrightarrow}$ in the case (ii). This immediately implies the assertion. ∎

Remark 5. *Under the assumptions of Theorem 4 the Condition* AB *in the case* (i) *and the Condition* AB *in mean in the case* (ii) *are necessary for the weak convergence of* $\{\mathscr{L}(w_k)\}$ *as* $k\to\infty$ *in the case* (i) *and the weak convergence in mean of this sequence in the case* (ii).

Proof. If the sequence $\{\mathscr{L}(w_k)\}$ is either (i) weakly convergent or (ii) weakly convergent in mean, then $\{\mathscr{L}(w_k)\}$ is tight in the case (i) while the sequence $\{n^{-1}\sum_{i=1}^{n}\mathscr{L}(w_i), n\geqslant 1\}$ is tight in the case (ii). But

(24) $$\sup_{j<0}(\xi_{n,j-k} - \xi_{n,-k}) = \sup_{j<0}\sum_{i=j-k+1}^{-k}\xi_{n,i}$$
$$= \max_{0\leqslant j\leqslant n-k}(S_{n-k} - S_{n-k-j}) \leqslant w_{n-k+1}, \qquad 1\leqslant k\leqslant n.$$

Hence and from the second part of Corollary 6 we infer that the condition (15) is necessary in the case (i) of Remark 5, which implies that Condition AB is necessary in the case (i). Similarly, from the second part of Corollary 6a we see that the condition (18) is necessary in the case (ii) of Remark 5, which implies that Condition AB in mean is necessary in the case (ii). ∎

REFERENCES

[1] P. Billingsley, *Convergence of Probability Measures*, J. Wiley, New York 1968.
[2] R. M. Dudley, *Distances of probability measures and random variables*, Ann. Math. Statist. 39 (1968), pp. 1563–1572.
[3] T. Lindvall, *Weak convergence of probability measures and random functions in function space* $D[0, \infty)$, J. Appl. Probab. 10 (1973), pp. 109–121.
[4] W. Szczotka, *Stationary representation of queues. I*, Adv. in Probab. 18 (1986), pp. 815–848.
[5] — *Exponential approximation of waiting time and queue size for queues in heavy traffic* (in preparation).
[6] — *A note on Skorohod representation* (in preparation).
[7] W. Whitt, *Preservation of rates of convergence under mappings*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 29 (1980), pp. 34–44.
[8] — *Some useful functions for functional limit theorems*, Math. Oper. Res. 5, No 1 (1980), pp. 67–85.

Mathematical Institute
Wrocław University
pl. Grunwaldzki 2/4
50-384 Wrocław, Poland

Received on 10.12.1989