Probability Theory — *A coupling approach to Doob's theorem*, by Alexei Kulik and Michael Scheutzow, communicated on 14 November 2014.

Abstract. — We provide a coupling proof of Doob's theorem which says that the transition probabilities of a regular Markov process which has an invariant probability measure μ converge to μ in the total variation distance. In addition we show that non-singularity (rather than equivalence) of the transition probabilities suffices to ensure convergence of the transition probabilities for μ-almost all initial conditions.

Key words: Markov process, invariant measure, coupling, convergence of transition probabilities, total variation distance.

Mathematics Subject Classification: 60J05, 60J25, 37L40.

1. Introduction

Doob's theorem, as formulated in [1], p. 43, states that if the stochastically continuous Markov semigroup P_t, $t \geq 0$ with Polish state space (E, d) has an invariant probability measure (ipm) μ and is t_0-*regular* for some $t_0 > 0$, then μ is unique and all transition probabilities converge to μ in the total variation distance. Here, t_0-regular means that all transition probabilities $P_{t_0}(x, .)$ are mutually equivalent. One common way to check t_0-regularity is to show that the Markov semigroup is irreducible and strong Feller, see [1], Proposition 4.1.1 (known as Khas'minskii's theorem).

In fact, Da Prato and Zabczyk formulate and prove Doob's theorem with respect to strong convergence (which is weaker than total variation convergence) and refer the reader to [4] and [3] for different proofs of total variation convergence. Neither of the proofs is short and elementary. In particular, neither of the proofs uses *coupling* which has been a powerful tool to prove convergence of transition probabilities in the past decades. The aim of this article is to provide a coupling proof of Doob's theorem. At the same time we generalize the result in various directions: instead of a Polish state space, we just require a mild condition on the measurable space $(E, \mathscr{E})$. Further, we allow the infimum of the times $t > 0$ for which $P_t(x, .)$ and $P_t(y, .)$ are equivalent to depend on the pair (x, y) without being uniformly bounded from above. We also show that we can replace equivalance of the transition probabilities by the much weaker property of non-singularity but in this case convergence of the transition probabilities only holds for μ-almost all initial conditions x. In the next section, we formulate the main results in the discrete time setting (in Remark 2 we say why the continuous time

claim follows) and provide an example showing that the weaker assumption in Corollary 1 does not guarantee the conclusion of Theorem 1.

2. Main results

Let X_n, $n \in \mathbb{Z}_+$ be a Markov chain with the state space $(E, \mathcal{E})$. The measurable space $(E, \mathcal{E})$ is assumed to be countably generated. We also assume that the diagonal $\Delta = \{(x, x), x \in E\}$ belongs to $\mathcal{E} \otimes \mathcal{E}$. A typical example of such a space is a Borel measurable space, e.g. a Polish space E endowed with the Borel σ-algebra $\mathcal{E}$.

Transition probabilities and n-step transition probabilities for X are denoted respectively by $P(x, dy)$ and $P_n(x, dy)$. The law of the sequence $\{X_n\}$ in $(E^\infty, \mathcal{E}^{\otimes \infty})$ with initial distribution $\mathrm{Law}(X_0) = \mu$ is denoted by $\mathbb{P}_\mu$, the respective expectation is denoted by $\mathbb{E}_\mu$; in case $\mu = \delta_x$ we write simply $\mathbb{P}_x$, $\mathbb{E}_x$.

Recall that an invariant probability measure for X is a probability measure μ on $(E, \mathcal{E})$ such that

$$(1) \qquad \mu(dy) = \int_E P(x, dy)\mu(dx).$$

Equivalently, a probability measure μ is invariant if the sequence $\{X_n, n \in \mathbb{Z}_+\}$ is strictly stationary under $\mathbb{P}_\mu$.

We use the usual relations for probability measures μ, ν on $(E, \mathcal{E})$: μ and ν are equivalent (notation $\mu \sim \nu$) if each of them is absolutely continuous w.r.t. the other; μ and ν are singular (notation $\mu \perp \nu$) if there exists $A \in \mathcal{E}$ such that $\mu(A) = 1$, $\nu(A) = 0$; otherwise μ and ν are non-singular (notation $\mu \not\perp \nu$). The total variation distance between probability measures μ, ν on $(E, \mathcal{E})$ is the total variation of the signed measure $\mu - \nu$ (notation $\|\mu - \nu\|$).

THEOREM 1. *Assume that for each x, $y \in E$ there exists $n = n_{x,y}$ such that*

$$(2) \qquad P_n(x, \cdot) \sim P_n(y, \cdot).$$

Then there exists at most one ipm for the chain X. If an ipm μ exists, then for every $x \in E$

$$(3) \qquad \|P_n(x, \cdot) - \mu\| \to 0, \quad n \to \infty.$$

The proof of Theorem 1 is given in Section 3 below. The following theorem shows to what extent the basic assumption (2) can be relaxed.

THEOREM 2. *Let X be a Markov chain which has an ipm μ. Assume further that for $\mu \otimes \mu$-almost all $(x, y) \in E \times E$ there exists $n = n_{x,y}$ such that*

$$(4) \qquad P_n(x, \cdot) \not\perp P_n(y, \cdot).$$

Then (3) holds true for μ-almost all $x \in E$.

The following corollary is simple and straightforward.

COROLLARY 1. *Assume that for each $x, y \in E$ there exists $n = n_{x,y}$ such that (4) holds true.*

Then there exists at most one ipm for the chain X. If an ipm μ exists, then (3) holds true for μ-a.a. $x \in E$.

To prove uniqueness of an ipm (which is the only addition to Theorem 2), let us assume that there exist two different ipm's μ_1, μ_2, and consider the averaged ipm $\mu = (1/2)(\mu_1 + \mu_2)$. Applying Theorem 2 first to μ_1 and then to μ, we get a contradiction: because μ_1 is absolutely continuous w.r.t. μ, we get that for μ_1-a.a. $x \in E$ the transition probabilities $P_n(x, \cdot)$ converge both to μ_1 and to μ, but $\mu_1 \neq \mu$. □

REMARK 1. Note that the condition of Theorem 2 alone does not yield uniqueness of the ipm for X: a simple counter-example is given by a chain with a finite state space with at least two mutually disconnected classes of states.

REMARK 2. Note that the results of Theorem 1, Theorem 2, and Corollary 1 are also true in the continuous time case when P_t, $t \geq 0$ is a Markov semigroup (no regularity in t is required). To see this, note that uniqueness of an ipm μ for the discretized chain P_n, $n \in \mathbb{N}_0$ implies uniqueness of an ipm for P_t, $t \geq 0$ and that for any (discrete or continuous time) Markov semigroup (P_r) with ipm μ, the function $r \mapsto \|P_r(x, .) - \mu\|$ is non-increasing.

EXAMPLE 1. The following example shows that the assumptions of Corollary 1 do not imply the conclusion of Theorem 1. Equip $E = \mathbb{Z}_+$ with the discrete σ-algebra $\mathscr{E}$ and define the transition probabilities by $p_{0,0} = 1$, $p_{i,i-1} = 1/3$ and $p_{i,i+1} = 2/3$ for $i \geq 1$. Then $\mu = \delta_0$ is the unique invariant probability measure, the assumptions of Corollary 1 hold but $P_n(i, .)$ does not converge to μ for any $i \neq 0$.

REMARK 3. Note that in the discrete case (i.e. E is finite or countably infinite) the assumptions in Theorem 2 and in Corollary 1 are also necessary for the respective conclusion to hold. This is not true for Theorem 1 however as the example $E = \mathbb{Z}_+$ with $p_{0,0} = p_{0,1} = 1/2$, $p_{i,i-1} = 2/3$ and $p_{i,i+1} = 1/3$ for $i \geq 1$ shows.

3. PROOFS OF THEOREM 1 AND THEOREM 2

An auxiliary construction. Denote for $N \in \mathbb{N}$, $p \in (0, 1)$

$$C_{N,p} = \{(x, y) : \|P_N(x, \cdot) - P_N(y, \cdot)\| \leq 2(1 - p)\}.$$

Note that $C_{N,p} \in \mathscr{E} \otimes \mathscr{E}$ by Lemma 1(i) below. Since

$$\nu \not\perp \bar{\nu} \Leftrightarrow \|\nu - \bar{\nu}\| < 2,$$

the assumption of Theorem 2 (and therefore the stronger assumption of Theorem 1, as well) yield that there exist $N \in \mathbb{N}$, $p \in (0,1)$ such that

$$(\mu \otimes \mu)(C_{N,p}) > 0.$$

In the sequel we fix these values N, p and write simply C instead of $C_{N,p}$. In addition we assume that $N = 1$. Remark 2 shows that this is no loss of generality.

An outline of the method. Our aim, in fact, is to prove the convergence

$$(5) \qquad \|P_n(x_1, \cdot) - P_n(x_2, \cdot)\| \to 0, \quad n \to \infty$$

either for all $(x_1, x_2) \in E \times E$ in the case considered in Theorem 1, or for $\mu \otimes \mu$-a.a. $(x_1, x_2) \in E \times E$ in the case considered in Theorem 2. Once (5) is proved, the required convergence (3) follows using the representation (1) and the triangle inequality. The following fact is well-known ([5], p. 14): for any two random elements ξ_1, ξ_2, defined on a same probability space $(\Omega, \mathscr{F}, \mathbb{P})$, valued in $(E, \mathscr{E})$, and such that $\mathrm{Law}(\xi_i) = \nu_i$, $i = 1, 2$, one has

$$(6) \qquad \|\nu_1 - \nu_2\| \leq 2\mathbb{P}(\xi_1 \neq \xi_2).$$

Hence for any sequence $\{Z_n = (Z_n^1, Z_n^2), n \in \mathbb{Z}_+\}$ such that the laws of $\{Z_n^i, n \in \mathbb{Z}_+\}$, $i = 1, 2$ equal respectively $\mathbb{P}_{x_i}$, $i = 1, 2$, one has a bound

$$\|P_n(x_1, \cdot) - P_n(x_2, \cdot)\| \leq 2\mathbb{P}(Z_n^1 \neq Z_n^2).$$

This is the essence of the famous *coupling* approach which dates back to W. Döblin [2]: to prove (5) one should construct a sequence Z which verifies the above assumption (such a sequence is usually called a *coupling* for X) in such a way that

$$(7) \qquad \mathbb{P}(Z_n^1 \neq Z_n^2) \to 0, \quad n \to \infty.$$

Construction of the coupling. The sequence Z will be taken as a Markov chain on $E \times E$ with suitably constructed transition probability. The first part of this construction is based on the fact that a proper choice of the pair (ξ_1, ξ_2) may turn inequality (6) into an identity. This fact, sometimes called the Coupling Lemma, is well known; the law of any such pair (ξ_1, ξ_2) is called a *maximal coupling*. We refer to [5] Section 1.4, where the construction of a maximal coupling based on the *splitting representation* of a random variable is given. In our framework we use essentially the same construction, but with modifications which are caused by the necessity (a) to deal with transition probabilities instead of measures, and therefore to take care of measurability issues; (b) to manage properly the law of the pair "outside of the diagonal". Namely, we have the following statement (the proof is given in the Appendix).

LEMMA 1 (The Coupling Lemma for transition probabilities). *Let $(E, \mathscr{E})$ be a countably generated measurable space.*

Then for any Markov kernel $P(x, dy)$ on $(E, \mathcal{E})$ there exists a Markov kernel $Q((x_1, x_2), dy_1 dy_2)$ on $(E \times E, \mathcal{E} \otimes \mathcal{E})$ such that for every $(x_1, x_2) \in E \times E$:

(i) $Q((x_1, x_2), \Delta) = 1 - (1/2)\|P(x_1, \cdot) - P(x_2, \cdot)\|;$
(ii) *the measure $Q((x_1, x_2), dy_1 dy_2)$ restricted to $(E \times E) \backslash \Delta$ is absolutely continuous w.r.t. $P(x_1, dy_1) \otimes P(x_2, dy_2)$.*

REMARK 4. We refer a reader to [6] for another version of the Coupling Lemma which takes into account measurability issues; the measurable space $(E, \mathcal{E})$ therein is assumed to be a Borel one.

Denote

$$R((x_1, x_2), dy_1 dy_2) = P(x_1, dy_1) \otimes P(x_2, dy_2),$$

which is just the transition probability of the Markov chain in $E \times E$ whose components are independent and each of the components is a Markov chain with the transition probability $P(x, dy)$. Such a chain is usually called an *independent coupling*, and below we denote it by $W = \{W_n, n \in \mathbb{Z}_+\}$.

Finally, we define the transition probability for Z by

$$S((x_1, x_2), .) := \begin{cases} Q((x_1, x_2), .) & \text{if } (x_1, x_2) \in C \\ R((x_1, x_2), .) & \text{otherwise,} \end{cases}$$

with the set C defined at the beginning of the proof. By construction, Z is a coupling for X, and our aim is to prove (7) with $Z_0 = (x_1, x_2)$ either for all (x_1, x_2) in the case of Theorem 1, or for $\mu \otimes \mu$-a.a. (x_1, x_2) in the case of Theorem 2.

PROOF OF (7). Note that $\Delta \subset C$, and for any point $(x, x) \in \Delta$

$$Q((x, x), \Delta) = 1.$$

Hence, by the construction the following property holds: once Z hits Δ, all the subsequent values of Z a.s. stay in Δ ("once the components are coupled they stay coupled"). Therefore

$$\mathbb{P}(Z_n^1 \neq Z_n^2) = \mathbb{P}(T > n), \quad T = \inf\{m : Z_m^1 = Z_m^2\},$$

with the usual convention $\inf \emptyset = \infty$.

Consider the sequence of stopping times

$$\tau_0 = 0, \quad \tau_k = \inf\{n > \tau_{k-1} : Z_n \in C\}, \quad k \geq 1,$$

and *assume* for a moment that we know that

(8) $$\tau_k < \infty, \quad k \geq 1$$

with probability 1. Clearly, for any $k \geq 1$

$$\{T > \tau_k\} = \bigcup_{n=1}^{\infty} \{\tau_k = n, Z_1 \notin \Delta, \ldots, Z_n \notin \Delta\} \in \mathscr{F}_{\tau_k};$$

here $\{\mathscr{F}_n\}$ denotes the natural filtration of the sequence Z. By the construction of Z,

$$\mathbb{P}(Z\tau_{k+1} \in \Delta \,|\, \mathscr{F}_{\tau_k}) \quad \begin{cases} \geq p, & T > \tau_k, \\ = 1, & \text{otherwise.} \end{cases}$$

Hence for the sequence $p_k = \mathbb{P}(T \leq \tau_k)$, $k \geq 1$ one has

$$p_{k+1} \geq p(1 - p_k) + p_k, \quad k \geq 1,$$

and therefore $p_k \to 1$, $k \to \infty$, which together with (8) yields (7).

PROOF OF (8): THE RECURRENCE LEMMA. We have reached the last and the crucial step in the proof: we need to prove that the set C, which in a sense is "favorable for the subsequent coupling attempt", is a.s. visited by Z infinitely often. We use the following lemma, whose proof is given in the Appendix.

LEMMA 2 (The Recurrence Lemma). *Assume that the Markov chain X satisfies the condition of Theorem 2.*
Then for any $B \in \mathscr{E}$ with $\mu(B) > 0$, for μ-a.a. $x \in E$

(9) $$\mathbb{P}_x(X_n \in B \text{ infinitely often}) = 1.$$

If, in addition, the condition of Theorem 1 holds true, then (9) holds true for every $x \in E$.

Now we can finish the whole proof; consider first the case of Theorem 1. The independent coupling W verifies the assumptions of Lemma 2 with $E \times E$ instead of E and $\mu \otimes \mu$ instead of μ. Because $(\mu \otimes \mu)(C) > 0$, this yields

(10) $$\mathbb{P}(W_n \in C \text{ infinitely often}) = 1$$

for all initial values $W_0 = (x_1, x_2) \in E \times E$.

Observe that up to τ_1 the law of Z coincides with the law of the independent coupling W up to its first visit to C, hence by (10)

$$\mathbb{P}(\tau_1 < \infty) = 1$$

for all initial values $Z_0 = (x_1, x_2) \in E \times E$. Hence Z a.s. performs at least one "coupling attempt". If this attempt is successful, i.e. $T \leq \tau_1 + 1$, then a.s. $\tau_2 = \tau_1 + 1, \tau_3 = \tau_2 + 1, \ldots$ because $\Delta \subset C$ and, once the components of Z are coupled, they stay coupled. In that case (8) holds true. If "the first coupling

attempt is not successful", the chain Z afterwards again performs as the independent coupling W up to the time moment τ_2. Applying Lemma 2 once again and the strong Markov property of Z, we get

$$\mathbb{P}(\tau_2 < \infty) = 1.$$

Iterating this argument, we get (8) for all initial values $Z_0 = (x_1, x_2) \in E \times E$, which completes the proof of Theorem 1.

Let us proceed with Theorem 2; in that case it is convenient to prove (8) for a version of Z with $\mathrm{Law}(Z_0) = \mu \otimes \mu$. The argument, completely the same as above, proves that $\tau_1 < \infty$ a.s., and if "the first coupling attempt is successful" then (8) holds true. Consider the law of Z_{τ_1+1} conditioned by the event that "the first coupling attempt is not successful"; that is, $\{Z_{\tau_1+1} \notin \Delta\}$. By the choice of the law of Z_0 and the construction of the kernel Q, this law is absolutely continuous w.r.t. $\mu \otimes \mu$. Using this and the strong Markov property of Z at the stopping time $\tau_1 + 1$, we apply Lemma 2 once again and get

$$\mathbb{P}(\tau_2 < \infty) = 1.$$

Iterating this argument, we get (8) for Z with $\mathrm{Law}(Z_0) = \mu \otimes \mu$, which completes the proof of Theorem 2. $\square$

A. Proofs of the auxiliary lemmas

A.1. Proof of Lemma 1

Denote

$$\Lambda(x_1, x_2; dy) = (1/2)(P(x_1, dy) + P(x_2, dy)).$$

Then for every x_1, x_2 $P_i(x_i, dy) \ll \Lambda(x_1, x_2; dy)$, $i = 1, 2$. Let us show that respective Radon-Nikodym derivatives can be chosen in a jointly measurable way; that is, there exist measurable functions $f_i : E \times E \times E \to \mathbb{R}^+$, $i = 1, 2$ such that

$$P(x_i, A) = \int_A f_i(x_1, x_2, y)\Lambda(x_1, x_2; dy), \quad i = 1, 2, \ x_1, x_2 \in E, \ A \in \mathscr{E}.$$

Let $\mathscr{E}_0$ be a countable algebra which generates $\mathscr{E}$, then the countable class $\mathscr{H}$, which consists of all functions representable in the form of a finite sum $\sum_k c_k 1_{A_k}$, $\{c_k\} \subset \mathbb{Q}$, $\{A_k\} \subset \mathscr{E}_0$, is dense in $L_1(E, \lambda)$ for any probability measure λ on $(E, \mathscr{E})$.

Consider a Radon-Nikodym derivative

$$\rho_{x_1, x_2}^1(y) = \frac{P(x_1, dy)}{\Lambda(x_1, x_2; dy)},$$

then for every $\varepsilon > 0$ there exists $h \in \mathscr{H}$ such that

$$\sup_{A \in \mathscr{E}_0} \int_A (\rho^1_{x_1, x_2}(y) - h(y)) \Lambda(x_1, x_2; dy) < \frac{\varepsilon}{2}.$$

Observe that this relation is equivalent to

(11) $$\sup_{A \in \mathscr{E}_0} \left(P(x_1, A) - \int_A h(y) \Lambda(x_1, x_2; dy) \right) < \frac{\varepsilon}{2},$$

and yields

$$\int_E |\rho_{x_1, x_2}(y) - h(y)| \Lambda(x_1, x_2; dy) < \varepsilon.$$

Fix some enumeration of the class $\mathscr{H} = \{h_m, m \in \mathbb{N}\}$, and denote for $n \geq 1$ by $m(x_1, x_2; n)$ the minimal $m \geq 1$ such that (11) holds true for $h = h_m$ with $\varepsilon = 2^{-n-1}$. Then $m(\cdot; n) : E \times E \to \mathbb{N}$ is measurable, and therefore

$$f_1^n(x_1, x_2, y) = h_{m(x_1, x_2; n)}(y)$$

is measurable as a function $E \times E \times E \to \mathbb{R}$. When x_1, x_2 are fixed, the sequence $\{f_1^n(x_1, x_2, y), n \geq 1\}$ converges to $\rho^1_{x_1, x_2}(y)$ for $\Lambda(x_1, x_2; \cdot)$-a.a. y: this follows from the Borel-Cantelli lemma because, by the construction,

$$\|f_1^n(x_1, x_2, \cdot) - \rho^1_{x_1, x_2}\|_{L_1(E, \Lambda(x_1, x_2; \cdot))} \leq 2^{-n}.$$

Therefore the function

$$f_1(x_1, x_2, y) = \begin{cases} \lim_{n \to \infty} f_1^n(x_1, x_2, y), & \text{if the limit exists} \\ 0, & \text{otherwise} \end{cases}$$

gives the required measurable version of the Radon-Nikodym derivative for $P(x_1, dy)$ (the construction for $P(x_2, dy)$ is the same).

We finish the proof by repeating essentially the construction from [5] Section 1.4, based on the splitting representation for probability laws. Write

$$g(x_1, x_2, y) = \min(f_1(x_1, x_2, y), f_2(x_1, x_2, y)),$$

$$p(x_1, x_2) = \int_E g(x_1, x_2, y) \Lambda(x_1, x_2; dy),$$

$$\Theta(x_1, x_2; dy) = \frac{g(x_1, x_2, y)}{p(x_1, x_2)} \Lambda(x_1, x_2; dy)$$

with the convention that (anything)$/0 = 1$. Then we have representations

$$P(x_i, dy) = p(x_1, x_2)\Theta(x_1, x_2; dy) + (1 - p(x_1, x_2))\Sigma_i(x_1, x_2; dy), \quad i = 1, 2$$

with probability kernels Θ, Σ_1, Σ_2 and measurable $p : E \times E \to [0, 1]$.

Note that the mapping $E \ni x \mapsto (x, x) \in E \times E$ is $\mathscr{E} - \mathscr{E} \otimes \mathscr{E}$ measurable, and denote by $Q_1((x_1, x_2), dy_1 dy_2)$ the image of $\Theta(x_1, x_2; dy)$ under this mapping. Denote also by $Q_2((x_1, x_2), dy_1 dy_2)$ the product of the measures $\Sigma_i(x_1, x_2; dy_i)$, $i = 1, 2$. Then

$$Q((x_1, x_2), \cdot) = p(x_1, x_2) Q_1((x_1, x_2), \cdot) + (1 - p(x_1, x_2)) Q_2((x_1, x_2), \cdot)$$

is the required kernel; observe that the assertion (ii) now holds true because $Q_2((x_1, x_2), \cdot)$ is chosen to be a product measure with the components $\Sigma_i(x_1, x_2; \cdot) \ll P(x_i, \cdot)$, $i = 1, 2$. $\qquad\square$

A.2. Proof of Lemma 2

Denote

$$\psi(x) = \mathbb{P}_x(X_n \in B \text{ infinitely often}),$$

and consider a stationary version of X with $\mathrm{Law}(X_0) = \mu$. Then the sequence $\{\psi(X_n)\}$ is stationary. But, in addition, this sequence is a Lévy martingale: by the Markov property of X, we have with probability 1

$$\psi(X_n) = \mathbb{E}_{X_n} \mathbf{1}_{X_k \in B \text{ i.o.}} = \mathbb{E}[\mathbf{1}_{X_k \in B \text{ i.o.}, k \geq n} \mid \mathscr{F}_n] = \mathbb{E}[\mathbf{1}_{X_k \in B \text{ i.o.}, k \geq 0} \mid \mathscr{F}_n].$$

Then with probability 1

$$\psi(X_n) \to \mathbf{1}_{X_k \in B \text{ i.o.}, k \geq 0}, \quad n \to \infty,$$

and hence by stationarity of $\{\psi(X_n)\}$ we have $\psi(x) = 0$ or 1 for μ-a.a. $x \in E$. Denote

$$\Psi_0 = \{x : \psi(x) = 0\}, \quad \Psi_1 = \{x : \psi(x) = 1\},$$

and observe that because $\psi(x) \in [0, 1]$ in any case, by the martingale property of $\{\psi(X_n)\}$ one has for any $n \geq 1$

$$P_n(x, \Psi_0) = 1 \quad \text{for } \mu\text{-a.a. } x \in \Psi_0 \text{ and } P_n(x, \Psi_1) = 1 \text{ for } \mu\text{-a.a. } x \in \Psi_1.$$

Because of the assumption of the lemma (or Theorem 2), it is impossible that both Ψ_0 and Ψ_1 have positive measure μ. The identity $\mu(\Psi_1) = 0$ contradicts Birkhoff's ergodic theorem: with probability 1 we have

$$\frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_B(X_n) \to \eta, \quad N \to \infty$$

where $E\eta = \mu(B) > 0$. Therefore $\mu(\Psi_0) = 0$, which implies the required identity $\mu(\Psi_1) = 1$.

If, in addition, the condition of Theorem 1 holds true, then it follows from what we have just proved that for any $x \in E$ there exists $n = n_x$ such that $P_n(x, \Psi_1) = 1$. By the Markov property of X and the definition of ψ, this implies $\psi(x) = 1$.   □

REFERENCES

[1] G. DA PRATO - J. ZABCZYK, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press, Cambridge (1996).

[2] W. DOEBLIN, *Eléments d'une théorie générale des chaînes simples constantes de Markoff*, Ann. Sci. E.N.S., 57, 61–111, 1940.

[3] J. SEIDLER, *Ergodic behaviour of stochastic parabolic equations*, Czechoslovak Math. J., 47(122), 277–316, 1997.

[4] L. STETTNER, *Remarks on ergodic conditions for Markov processes on Polish spaces*, Bull. Polish Acad. Sci. Math., 42, 103–114, 1994.

[5] H. THORISSON, *Coupling, Stationarity, and Regeneration*, Springer, NY (2000).

[6] A. YU. VERETENNIKOV, *Coupling method for Markov chains under integral Doeblin type condition*, Theory Stoch. Processes, 8 (3–4), 383–391, 2002.

A. Kulik
Institute of Mathematics
NAS of Ukraine, 3
Tereshchenkivska str., 01601 Kyiv, Ukraine
kulik.alex.m@gmail.com

M. Scheutzow
Institut für Mathematik
MA 7-5, Technische Universität Berlin
Str. des 17. Juni 136, 10623 Berlin, Germany
ms@math.tu-berlin.de