

## STATIONARY FLOWS AND UNIQUENESS OF INVARIANT MEASURES

BY

FRANÇOIS BACCELLI (*Département d'Informatique, École Normale Supérieure, 45 rue d'Ulm,  
F-75230 Paris Cedex 05, France*)

AND

TAKIS KONSTANTOPOULOS (*School of Mathematical Sciences, Heriot-Watt University,  
Edinburgh EH14 4AS, UK*)

**Abstract.** We consider a quadruple  $(\Omega, \mathcal{A}, \vartheta, \mu)$ , where  $\mathcal{A}$  is a  $\sigma$ -algebra of subsets of  $\Omega$ , and  $\vartheta$  is a measurable bijection from  $\Omega$  into itself that preserves a finite measure  $\mu$ . For each  $B \in \mathcal{A}$ , we define and study the measure  $\mu_B$  obtained by integrating on  $B$  the number of visits to a set of the trajectory of a point of  $\Omega$  before returning to  $B$ . In particular, we obtain a generalization of Kac's formula and discuss its relation to discrete-time Palm theory. Although classical in appearance, its use in obtaining uniqueness of invariant measures of various stochastic models seems to be new. We apply the concept to countable Markov chains and Harris processes in general state space.

**1. Introduction.** This paper started with the following question. It is classical that, for a Markov chain  $(X_n, n \geq 0)$  with a countable state space  $S$  possessing a positive recurrent state  $b \in S$ , there is at least one invariant probability measure  $\pi^{(b)}$  on  $S$  which is defined by the usual “cycle formula”: Let  $t_b$  be the first return time to  $b$  and set

$$\pi^{(b)}(A) = \frac{1}{E_b t_b} E_b \sum_{n=0}^{t_b-1} \mathbf{1}(X_n \in A),$$

where  $\mathbf{1}(\cdot)$  denotes the indicator function. To show that  $\pi^{(b)}$  is the only invariant probability measure, we need, in addition, to ensure that the only positive recurrent states are those that communicate with  $b$  (this holds, for instance, if the chain is irreducible). There are several proofs of uniqueness, ranging from analytic [by means of the Perron-Frobenius theorem which itself can be proved in a number of ways—see, e.g., Lind and Marcus [9] for a geometric proof] to probabilistic [by means of applying the Doeblin coupling construction; this requires, in addition, aperiodicity—see, e.g., Norris [12, Theorem

---

Received February 15, 2008.

2000 *Mathematics Subject Classification.* Primary 37A05, 60J10; Secondary 37A50, 60G10.

Research supported by an EPSRC grant.

*E-mail address:* Francois.Bacelli@ens.fr

*E-mail address:* takis@ma.hw.ac.uk

©2010 Brown University  
Reverts to public domain 28 years from publication

1.8.3]. The question we posed is whether there is a way to prove uniqueness directly from the way that  $\pi^{(b)}$  is constructed by the cycle formula. If so, can we do this for Markov chains in a general state space?

In answering the question, we abstracted the problem and lifted it to a general measurable space  $(\Omega, \mathcal{A})$  endowed with a measurable bijective transformation  $\vartheta$  that preserves some finite measure  $\mu$ . The point of view appears to be new, although the tools used below are quite natural in ergodic theory and in the construction of Palm probabilities. The origin of these tools can be traced, as far as we can tell, to a paper by Kac [7] where the so-called Kac formula (see Corollary 2.10) is proved. In Section 2 we define, for each  $B \in \mathcal{A}$ , the forward and backward hitting times of  $B$  by the iterates of  $\vartheta$  (called  $T_B^+$  and  $T_B^-$ , respectively) and the measures

$$\mu_B^\pm(A) = \int_B d\mu \sum_{n=0}^{T_B^\pm - 1} \mathbf{1}_{\vartheta^{-n}A}.$$

Theorem 2.4 states that the two measures are identical, called  $\mu_B$ , for simplicity, and that

$$\mu_B(A) = \mu(A, T_B^+ < \infty) = \mu(A, T_B^- < \infty).$$

Thus, for all  $B$  such that  $\mu(B) > 0$ , Poincaré's recurrence lemma (Lemma 2.1 in this paper) implies that  $\mu_B = \mu$ . In addition, Section 2 contains several facts about the measures  $\mu_B$  such as their relation to discrete-time Palm theory and their behaviour as functions of  $B$ . In Section 3, we consider a Markov chain on a countable set  $S$ . Assuming irreducibility and positive recurrence, the previous observation immediately yields a unique probability measure  $\pi$  on  $S$  such that  $\pi P = \pi$ , which answers the original question. Finally, in Section 4, we consider a Harris chain and show uniqueness of the invariant probability measure constructed by means of cycles away from a recurrent regeneration set  $R$ .

A few words about the notation used in the paper: we find it less cumbersome to sometimes replace the symbol for intersection between two sets by a simple comma (this is customary in probability theory). So, for instance,  $\mu(A, T_B^+ < \infty)$  should be read as  $\mu(A \cap \{T_B^+ < \infty\})$ .

**2. The master formula.** Let  $(\Omega, \mathcal{A})$  be a measurable space endowed with an automorphism  $\vartheta$ , that is, a measurable function  $\vartheta : \Omega \rightarrow \Omega$  such that the inverse function  $\vartheta^{-1} : \Omega \rightarrow \Omega$  exists and is also a measurable. As in ergodic theory [see, e.g., Petersen [16]], we shall consider the group  $(\vartheta^n, n \in \mathbb{Z})$  acting on  $\Omega$ .

For  $A, B \in \mathcal{A}$  define the following functions:

$$T_B^+ \equiv T_B(\omega; \vartheta) := \inf\{n \geq 1 : \vartheta^n \omega \in B\}, \quad T_B^- \equiv T_B(\omega; \vartheta^{-1}), \quad (2.1a)$$

$$M_B^+(A) \equiv M_B(A, \omega; \vartheta) := \sum_{0 \leq n < T_B(\omega; \vartheta)} \mathbf{1}(\vartheta^n \omega \in A), \quad M_B^-(A) \equiv M_B(A, \omega; \vartheta^{-1}). \quad (2.1b)$$

They are both measurable and may assume the value  $+\infty$ .

Consider the forward trajectory, or orbit,  $(\omega, \vartheta\omega, \vartheta^2\omega, \dots)$  of the point  $\omega \in \Omega$ . Then  $M_B^+(A)(\omega)$  gives the number of times the forward orbit visits the set  $A$  up to (and not

including) the time  $T_B^+(\omega)$  it visits the set  $B$  for the first time. A similar interpretation can be given for  $M_B^-(A)(\omega)$  with respect to the backwards trajectory  $(\omega, \vartheta^{-1}\omega, \vartheta^{-2}\omega, \dots)$ .

In addition to the above, we assume that there is a finite measure  $\mu$  on  $(\Omega, \mathcal{A})$  which is preserved by  $\vartheta$ . First, recall the following standard result, a short proof of which (amongst the many possible ones, e.g., [16, 17, 18]) is given for completeness:

LEMMA 2.1 (Poincaré recurrence). Let  $\mu$  be a finite measure on  $(\Omega, \mathcal{A})$  which is preserved by  $\vartheta$ . Then, for all  $B \in \mathcal{A}$ ,

$$\mu(B) = \mu(B, T_B^- < \infty) = \mu(B, T_B^+ < \infty). \quad (2.2)$$

*Proof.* We have  $\{T_B^- = \infty\} = \bigcap_{n \geq 1} \vartheta^n B^c$ . Since  $\mu(B^c) < \infty$ , using the monotone convergence theorem, we obtain

$$\begin{aligned} \mu(B^c, T_B^- = \infty) &= \lim_{n \rightarrow \infty} \mu(B^c \cap \vartheta B^c \cap \dots \cap \vartheta^{n-1} B^c) \\ &= \lim_{n \rightarrow \infty} \mu(\vartheta B^c \cap \vartheta^2 B^c \cap \dots \cap \vartheta^n B^c) = \mu(T_B^- = \infty). \end{aligned}$$

Since both sides of this are finite, we have  $\mu(B, T_B^- = \infty) = \mu(T_B^- = \infty) - \mu(B^c, T_B^- = \infty) = 0$ . Therefore,

$$\mu(B) = \mu(B, T_B^- = \infty) + \mu(B, T_B^- < \infty) = \mu(B, T_B^- < \infty), \quad (2.3)$$

which concludes the proof of the first equality in (2.2). Working similarly, we obtain the equality for  $T_B^+$ .  $\square$

REMARK 2.2. In other words,  $T_B^+ < \infty$  and  $T_B^- < \infty$ ,  $\mu$ -a.e. on  $B$ :

$$\mu(B, T_B^- = \infty) = \mu(B, T_B^+ = \infty) = 0. \quad (2.4)$$

REMARK 2.3. Note also that

$$\mu(B) = 0 \iff \mu(T_B^+ < \infty) = 0 \iff \mu(T_B^- < \infty) = 0. \quad (2.5)$$

Indeed, if  $\mu(B) > 0$ , then, by (2.3),  $\mu(T_B^\pm < \infty) \geq \mu(B) > 0$ . On the other hand, if  $\mu(B) = 0$ , then  $\mu(T_B^+ < \infty) = \mu(\bigcup_{n \geq 1} \vartheta^{-n} B) \leq \sum_{n \geq 1} \mu(\vartheta^{-n} B) = \sum_{n \geq 1} \mu(B) = 0$ .

These are used in proving the main formulae of this paper:

THEOREM 2.4. If the finite measure  $\mu$  on  $(\Omega, \mathcal{A})$  is preserved by  $\vartheta$ , then, for all  $A, B \in \mathcal{A}$ ,

$$\mu_B^+(A) := \int_B M_B^+(A) d\mu = \int_A \mathbf{1}(T_B^- < \infty) d\mu, \quad (2.6a)$$

$$\mu_B^-(A) := \int_B M_B^-(A) d\mu = \int_A \mathbf{1}(T_B^+ < \infty) d\mu. \quad (2.6b)$$

Furthermore, the two measures  $\mu_B^+$  and  $\mu_B^-$  are identical.

*Proof.* We show the first relation; the second one can be proved in a similar way. Suppose first that  $\mu(B) > 0$ . Then using Fubini's theorem, and the assumption that  $\mu$

is preserved by  $\vartheta$  we have

$$\begin{aligned}
\int_B M_B^+(A) d\mu &= \mu(A \cap B) + \sum_{n \geq 1} \mu(B \cap \vartheta^{-1} B^c \cap \dots \cap \vartheta^{-n} B^c \cap \vartheta^{-n} A) \\
&= \mu(A \cap B) + \sum_{n \geq 1} \mu(\vartheta^n B \cap \vartheta^{n-1} B^c \cap \dots \cap B^c \cap A) \\
&= \mu(A \cap B) + \sum_{n \geq 1} \mu(T_B^- = n, A \cap B^c) \\
&= \mu(A \cap B) + \mu(A \cap B^c, T_B^- < \infty) \\
&= \mu(A \cap B, T_B^- = \infty) + \mu(A \cap B, T_B^- < \infty) + \mu(A \cap B^c, T_B^- < \infty) \\
&= \mu(A, T_B^- < \infty), \tag{2.7}
\end{aligned}$$

where we used (2.4) to obtain the last equality. If  $\mu(B) = 0$ , then the danger exists that the quantity  $M_B^+(A)$  may be infinite since  $T_B^+$  may assume the value  $+\infty$ . However, since  $T_B^+ < \infty$ ,  $\mu$ -a.e. on  $B$ , we have

$$\int_B M_B^+(A) d\mu \leq \int_B M_B^+(\Omega) d\mu = \int_B T_B^+ d\mu = 0.$$

On the other hand, by (2.5), we have  $\mu(T_B^- < \infty) = 0$  and so

$$\int_A \mathbf{1}(T_B^- < \infty) d\mu = 0.$$

Thus, if  $\mu(B) = 0$ , both sides of (2.6a) are equal to zero.

We next show that  $\mu_B^+ = \mu_B^-$  by showing that the right-hand sides of (2.6a) and (2.6b) are equal. Observe that, for all  $N \in \mathbb{Z}$ ,

$$\{T_B^+ = \infty\} = \{T_B^+ \circ \vartheta^N = \infty\}, \quad \mu - \text{a.e.},$$

in the sense that the  $\mu$ -measure of the symmetric difference of the two sets is zero. This follows from the fact that both sets have equal  $\mu$ -measure and that  $T_B^+ = \infty$  implies that  $T_B^+ \circ \vartheta^N = \infty$  if  $N > 0$  or is implied by it if  $N < 0$ . Therefore,<sup>1</sup>

$$\{T_B^+ = \infty\} = \bigcap_{N \in \mathbb{Z}} \{T_B^+ \circ \vartheta^N = \infty\}, \quad \mu - \text{a.e.}$$

But the last event can be written as  $\bigcap_{n \in \mathbb{Z}} \vartheta^n B^c$ . Since the sets  $\{T_B^+ = \infty\}$  and  $\bigcap_{n \in \mathbb{Z}} \vartheta^n B^c$  are  $\mu$ -a.e. equal, we have that, for all  $A \in \mathcal{A}$ ,

$$\mu(A, T_B^+ = \infty) = \mu(A \cap \bigcap_{n \in \mathbb{Z}} \vartheta^n B^c).$$

By the same argument, we have

$$\mu(A, T_B^- = \infty) = \mu(A \cap \bigcap_{n \in \mathbb{Z}} \vartheta^n B^c),$$

so the right-hand sides of (2.6a) and (2.6b) are identical. It follows that  $\mu_B^+ = \mu_B^-$ .  $\square$

REMARK 2.5. Since the two measures are the same, we shall simply use the symbol  $\mu_B$  for either of  $\mu_B^\pm$ .

<sup>1</sup>If  $B_0 = B_n$ ,  $\mu$ -a.e., for all  $n \geq 1$ , then  $B_0 = B$ ,  $\mu$ -a.e., where  $B = \bigcap_{n \geq 1} B_n$ . To see this notice that  $\mu(B_0^c \cap B) \leq \mu(B_0^c \cap B_1) = 0$  and  $\mu(B_0 \cap B^c) \leq \sum_{n \geq 1} \mu(B_0 \cap B_n^c) = 0$ .

Next, besides the measures  $\mu$  and  $\mu_B$ , we are going to consider the restriction  $\mu|_B$  of  $\mu$  on the set  $B$ , defined by

$$\mu|_B(A) := \mu(B \cap A). \quad (2.8)$$

We let

$$\Omega_B^+ := \{T_B^+ < \infty\},$$

and define the measurable function  $\vartheta^{T_B^+} : \Omega \rightarrow \Omega$  by

$$(\vartheta^{T_B^+})(\omega) := \begin{cases} \vartheta^{T_B^+(\omega)}(\omega), & \omega \in \Omega_B^+ \\ \omega, & \text{otherwise.} \end{cases}$$

We define  $\vartheta^{-T_B^+}$ ,  $\vartheta^{T_B^-}$ , and  $\vartheta^{-T_B^-}$  in a similar way.

We shall see that  $\mu|_B$  is preserved by  $\vartheta^{T_B^+}$  and  $\vartheta^{T_B^-}$ .

**PROPOSITION 2.6.** If the finite measure  $\mu$  on  $(\Omega, \mathcal{A})$  is preserved by  $\vartheta$ , then its restriction  $\mu|_B$  on some  $B \in \mathcal{A}$  is preserved by  $\vartheta^{T_B^+}$  and by  $\vartheta^{T_B^-}$ ; i.e., for all  $A, B \in \mathcal{A}$ ,

$$\mu|_B((\vartheta^{T_B^+})^{-1}A) = \mu|_B((\vartheta^{T_B^-})^{-1}A) = \mu|_B(A).$$

*Proof.* By the definition of  $\vartheta^{T_B^+}$  we have

$$(\vartheta^{T_B^+})^{-1}A = \{T_B^+ = \infty, A\} \cup \bigcup_{n=1}^{\infty} \{T_B^+ = n, \vartheta^{-n}A\}.$$

Hence

$$\mu|_B((\vartheta^{T_B^+})^{-1}A) = \mu(B \cap A, T_B^+ = \infty) + \sum_{n=1}^{\infty} \mu(B \cap \vartheta^{-n}A, T_B^+ = n).$$

Due to (2.4) the first term is zero and so

$$\begin{aligned} \mu|_B((\vartheta^{T_B^+})^{-1}A) &= \sum_{n=1}^{\infty} \mu(B \cap \vartheta^{-n}A, T_B^+ = n) \\ &= \sum_{n=1}^{\infty} \mu(B \cap \vartheta^{-n}A \cap \vartheta^{-1}B^c \cap \dots \cap \vartheta^{-(n-1)}B^c \cap \vartheta^{-n}B) \\ &= \sum_{n=1}^{\infty} \mu(B \cap \vartheta^{-1}B^c \cap \dots \cap \vartheta^{-(n-1)}B^c \cap \vartheta^{-n}(A \cap B)) \\ &= \sum_{n=1}^{\infty} \mu(\vartheta^n B \cap \vartheta^{n-1}B^c \cap \dots \cap \vartheta B^c \cap A \cap B) \\ &= \sum_{n=1}^{\infty} \mu(T_B^- = n, A \cap B) = \mu(T_B^- < \infty, A \cap B) = \mu(A \cap B), \end{aligned}$$

where the latter equality follows again from (2.4). The assertion about  $T_B^-$  is proved in the same manner.  $\square$

**PROPOSITION 2.7.** If the finite measure  $\mu$  on  $(\Omega, \mathcal{A})$  is preserved by  $\vartheta$ , then, for all  $B \in \mathcal{A}$ , the measure  $\mu_B$ , defined by either (2.6a) or (2.6b), is also preserved by  $\vartheta$ .

*Proof.* Note that

$$M_B^+(\vartheta^{-1}A) - M_B^+(A) = \mathbf{1}_{(\vartheta^{T_B^+})^{-1}A} - \mathbf{1}_A.$$

Integrating both sides against  $\mu$  on the set  $B$  we obtain

$$\int_B M_B^+(\vartheta^{-1}A)d\mu - \int_B M_B^+(A)d\mu = \mu(B \cap (\vartheta^{-T_B^+})^{-1}A) - \mu(B \cap A).$$

From the definition of the measure  $\mu_B^+$ , we have that the left-hand side equals  $\mu_B^+(\vartheta^{-1}A) - \mu_B^+(A)$ , and from the definition of the measure  $\mu_{|B}$  we have that the right-hand side equals  $\mu_{|B}((\vartheta^{T_B^+})^{-1}A) - \mu_{|B}(A)$ , so:

$$\mu_B^+(\vartheta^{-1}A) - \mu_B^+(A) = \mu_{|B}((\vartheta^{T_B^+})^{-1}A) - \mu_{|B}(A).$$

But the right-hand side is zero, thanks to Proposition 2.6. Hence  $\mu_B^+$  is preserved by  $\vartheta$ .  $\square$

REMARK 2.8. Some results do not require the invertibility of  $\vartheta$ . For instance, recall formula (2.2):  $\mu(B) = \mu(B, T_B^+ < \infty) = \mu(B, \bigcup_{n=1}^{\infty} \vartheta^{-n}B)$ . This holds for any  $\mu$ -preserving measurable map  $\vartheta$ . However, the main formulae (2.6a)–(2.6b) that exhibit the “duality” between forward and backward iterates of  $\vartheta$ , do require invertibility. On the other hand, even without using Theorem 2.4 and Propositions 2.6–2.7, we can show that the measure  $\nu_B(A) := \mu(A, T_B^- < \infty)$  satisfies  $\nu_B(\vartheta A) = \nu_B(A)$  directly. To do this, note that  $\{T_B^- < \infty\} = \bigcup_{n \geq 1} \vartheta^n B$ , so  $\vartheta^{-1}\{T_B^- < \infty\} = \bigcup_{n \geq 0} \vartheta^n B = B \cup \{T_B^- < \infty\}$  and write

$$\begin{aligned} \mu(\vartheta A, T_B^- < \infty) &= \mu(A, B \cup \{T_B^- < \infty\}) \\ &= \mu((A \setminus B) \cup (A \cap B), B \cup \{T_B^- < \infty\}) \\ &= \mu(A \setminus B, T_B^- < \infty) + \mu(A \cap B, T_B^- < \infty) \\ &= \mu(A, T_B^- < \infty). \end{aligned}$$

This, incidentally, gives a second proof of Proposition 2.7.

2.1. *The case of a probability measure.* Assume now that  $\mu$  is a probability measure, i.e.,  $\mu(\Omega) = 1$ . Write  $P$  instead of  $\mu$  and let  $E$  denote integration with respect to it. Thus Theorem 2.4 reads

$$P_B(A) = E\mathbf{1}_B M_B^+(A) = P(A, T_B^+ < \infty).$$

Note that  $P_B$  is not necessarily a probability measure. We then have a number of important corollaries, the first of which concerns the total mass of  $P_B$ .

COROLLARY 2.9. If  $P$  is a probability measure on  $(\Omega, \mathcal{A})$ , then

$$P_B(\Omega) = E[T_B^+ \mathbf{1}_B] = P(T_B^- < \infty) = E[T_B^- \mathbf{1}_B] = P(T_B^+ < \infty) \leq 1. \quad (2.9)$$

*Proof.* Let  $A = \Omega$  in (2.6a):

$$\int_B M_B^+(\Omega)dP = \int_{\Omega} \mathbf{1}(T_B^- < \infty)dP.$$

But  $M_B^+(\Omega) = T_B^+$  and so the left-hand side equals  $E[T_B^+ \mathbf{1}_B]$ , proving the first equality in (2.9). The other equalities follow from the last part of Theorem 2.4.  $\square$

For  $A, B \in \mathcal{A}$ , we let, as usual,  $P(A|B) := P(A \cap B)/P(B)$ , provided that  $P(B) > 0$ . For a random variable  $X : \Omega \rightarrow \mathbb{R}$  let  $E(X|B) := E(X\mathbf{1}_B)/P(B)$  if  $P(B) > 0$  and  $E(|X|\mathbf{1}_B) < \infty$ .

COROLLARY 2.10 (Kac formula). If  $P$  is a probability measure on  $(\Omega, \mathcal{A})$  and  $B \in \mathcal{A}$  with  $P(B) > 0$ , then

$$E(T_B^+|B) = \frac{1}{P(B|T_B^- < \infty)} = E(T_B^-|B) = \frac{1}{P(B|T_B^+ < \infty)}.$$

*Proof.* From Corollary 2.9 we have  $E[T_B^+\mathbf{1}_B] \leq 1$ , and, by assumption,  $P(B) > 0$ ; therefore  $E(T_B^+|B) = E[T_B^+\mathbf{1}_B]/P(B)$  can be defined. From Corollary 2.9 again, we have  $E[T_B^+\mathbf{1}_B] = P(T_B^- < \infty)$ . Note also that we can define  $P(B|T_B^- < \infty)$  since, by Poincaré’s recurrence,  $P(T_B^- < \infty) \geq P(T_B^- < \infty, B) = P(B) > 0$ . Hence

$$E(T_B^+|B) = \frac{E[T_B^+\mathbf{1}_B]}{P(B)} = \frac{P(T_B^- < \infty)}{P(B, T_B^- < \infty)} = \frac{1}{P(B|T_B^- < \infty)}.$$

The second formula follows in the same way. □

Corollary 2.10 is slightly more general than Kac’s formula [7], many proofs of which are known: see, e.g., Petersen [16] or Pollicott and Yuri [17].

LEMMA 2.11. The distribution of  $T_B^+$  is the same as the distribution of  $T_B^-$ .

*Proof.* For each  $n \geq 1$ , we have

$$\begin{aligned} P(T_B^+ > n) &= P(\vartheta^{-n}B^c \cap \dots \cap \vartheta^{-1}B^c) \\ &= P(\vartheta B^c \cap \dots \cap \vartheta^n B^c) = P(T_B^- > n). \end{aligned} \quad \square$$

2.2. *Connections with discrete-time Palm theory.* We fix a set  $B \in \mathcal{A}$  and define the point process  $N(\omega)$  as the random set

$$N(\omega) := \{n \in \mathbb{Z} : \omega \in \vartheta^n B\},$$

and refer to the set  $\mathbb{Z}$  as the set of times. The rate of the point process is

$$\lambda_B := P(B)$$

which will be assumed to be positive. We define the Palm probability<sup>2</sup>  $P_B^0$  with respect to  $N$  as a probability measure on  $(\Omega, \mathcal{A})$  that puts a point of  $N$  at the origin of time, i.e.,

$$P_B^0(A) := P(A|0 \in N) = P(A|B). \tag{2.10}$$

A basic property of the Palm probability measure is that it is invariant under a group of transformations that places the origin of time at any of the points of the point process  $N$ . To make this precise, fix  $B \in \mathcal{A}$  and let  $(S_m(\omega), m \in \mathbb{Z})$  be an enumeration of the elements of  $N(\omega)$  in a way that  $\dots S_{-1}(\omega) < S_0(\omega) \leq 0 < S_1(\omega) < S_2(\omega) < \dots$ . Thus,  $S_1 = T_B^+$ . Let

$$\varphi_m := \vartheta^{S_m}, \quad m \in \mathbb{Z}.$$

---

<sup>2</sup>See [3] for an exposition of the theory in discrete and continuous time. The point here is that, in discrete time, Palm theory is simple and a simple consequence of the extension of Kac’s formula.

It is easy to see that  $\varphi_m \circ \varphi_k = \varphi_{m+k}$  for all  $m, k \in \mathbb{Z}$ . If  $\omega \in B$ , then  $S_0(\omega) = 0$  and so  $\varphi_0(\omega) = \omega$ . In other words,  $\varphi_0$  is the identity on  $B$ . Therefore  $\varphi_m \circ \varphi_{-m}$  is the identity on  $B$  and so the family  $(\varphi_m, m \in \mathbb{Z})$  is a group of transformations on  $B$ . Proposition 2.6 states that  $P(\varphi_1^{-1}A \cap B) = P(A \cap B)$  or  $P_B^0(\varphi_{-1}A) = P_B(A)$ ; this implies that  $P_B^0(\varphi_m A) = P_B(A)$  for all  $m \in \mathbb{Z}$ .

In the language of Palm probabilities, we can also see that Theorem 2.4 gives, under an additional assumption, the so-called Palm inversion formula:

LEMMA 2.12. If  $B \in \mathcal{A}$  is such that  $P(T_B^+ < \infty) = 1$ , then  $P_B = P$ , and if  $P_B^0$  is defined as in (2.10), then

$$P(A) = \lambda_B E_B^0 \left[ \sum_{n=0}^{T_B^+ - 1} \mathbf{1}_A \circ \vartheta^n \right], \quad (2.11)$$

and  $\lambda_B = P(B)$  is positive.

*Proof.* From the assumption  $P(T_B^+ < \infty) = 1$  and (2.6a) we have

$$P_B(A) = E[\mathbf{1}_B M_B^+(A)] = P(A, T_B^+ < \infty) = P(A).$$

Therefore,

$$E \left[ \mathbf{1}_B \sum_{n=0}^{T_B^+ - 1} \mathbf{1}_A \circ \vartheta^n \right] = P(A).$$

Since  $P(T_B^+ < \infty) > 0$  it follows from (2.5) that  $\lambda_B > 0$  and so the left-hand side of the above equals  $P(B) E[\sum_{n=0}^{T_B^+ - 1} \mathbf{1}_A \circ \vartheta^n | B] = \lambda_B E_B^0[\sum_{n=0}^{T_B^+ - 1} \mathbf{1}_A \circ \vartheta^n]$ .  $\square$

REMARK 2.13. We stress two important aspects of this last lemma.

- (i) First, it says that if  $P(T_B^+ < \infty) = 1$ , then the sub-probability measure  $P_B$  is identical to the probability measure  $P$ .
- (ii) Second, it says that  $P$  is fully determined from  $P_B^0$  through the inversion formula (2.11). As such, it is the discrete-time analogue of Slivnyak's inverse construction [20] showing that the law of a stationary point process is fully determined by its Palm measure.

2.3. *Additional properties of  $P_B(A)$  as a function of  $B$ .* As a function of  $A$ , the quantity  $P_B(A) = E\mathbf{1}_B M_B(A)$  is a measure (i.e., a countably additive set function with total mass  $P_B(\Omega) \leq 1$ ). On the other hand, the function  $B \mapsto P_B(A)$  can be thought of as a *pre-capacity*. By this we mean that it satisfies the following inequality:<sup>3</sup>

PROPOSITION 2.14.

$$P_{B_1 \cup B_2}(A) + P_{B_1 \cap B_2}(A) \leq P_{B_1}(A) + P_{B_2}(A).$$

The point here is that this is not obvious from the definition of  $P_B(A)$  as  $E\mathbf{1}_B M_B^+(A)$ . But, owing to Theorem 2.4,  $P_B(A) = P(A, T_B^+ < \infty)$  and this is what facilitates the proof of the above inequality. In fact, we shall prove the stronger statement:

<sup>3</sup>The notion of capacity [cf. Schneider and Weil [19], Molchanov [11]] requires additional continuity properties which are possible only under additional topological structure on  $\Omega$  and we shall not be concerned with this here.



PROPOSITION 2.15.

$$P(T_{B_1 \cap B_1}^+ = \infty, T_{B_1}^+ < \infty, T_{B_2}^+ < \infty, A) = P_{B_1}(A) + P_{B_2}(A) - P_{B_1 \cap B_2}(A) - P_{B_1 \cup B_2}(A). \quad (2.12)$$

*Proof.* Indeed, consider the ‘forward trajectory’ of a fixed element  $\omega$  of  $\Omega$ , i.e., the set:

$$\Psi(\omega) := \{\vartheta^n \omega : n \geq 1\}.$$

Then, for any  $B \in \mathcal{A}$ ,

$$\{\omega \in \Omega : T_B^+(\omega) < \infty\} = \{\omega \in \Omega : \Psi(\omega) \cap B \neq \emptyset\},$$

which is abbreviated as  $\{\Psi \cap B \neq \emptyset\}$ , as usual. So the left-hand side of (2.12) equals

$$L = P(A, \Psi \cap B_1 \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset, \Psi \cap B_2 \neq \emptyset).$$

It is easy to see that

$$\begin{aligned} & \{\Psi \cap B_1 \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset\} \\ &= \{\Psi \cap B_1 \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset, \Psi \cap B_2 \neq \emptyset\} \cup \{\Psi \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset\}, \end{aligned}$$

and the union is disjoint. (The trajectory of a point avoids  $B_1 \cap B_2$  but hits  $B_1$  if it either does so by also hitting  $B_2$  or by not hitting it.) Therefore the left-hand side of (2.12) equals

$$L = P(A, \Psi \cap B_1 \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset) - P(A, \Psi \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset).$$

Consider these last two terms separately. First, consider

$$\begin{aligned} P(A, \Psi \cap B_1 \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset) &= P(A, \Psi \cap B_1 \cap B_2 = \emptyset) - P(A, \Psi \cap B_1 = \emptyset) \\ &= P(A, \Psi \cap B_1 \neq \emptyset) - P(A, \Psi \cap B_1 \cap B_2 \neq \emptyset) \\ &= P(A, T_{B_1}^+ < \infty) - P(A, T_{B_1 \cap B_2}^+ < \infty) \\ &= P_{B_1}(A) - P_{B_1 \cap B_2}(A). \end{aligned}$$

Then, consider

$$\begin{aligned} P(A, \Psi \cap B_2 = \emptyset, \Psi \cap B_1 \neq \emptyset) &= P(A, \Psi \cap B_2 = \emptyset) - P(A, \Psi \cap (B_1 \cup B_2) = \emptyset) \\ &= P_{B_1 \cup B_2}(A) - P_{B_2}(A). \end{aligned}$$

Combining the two terms above, we obtain (2.12).  $\square$

**3. Uniqueness of the invariant distribution in Markov chains.** Suppose that

$$P = [p_{i,j}]$$

is a stochastic matrix on a countable state space  $S = \{a, b, c, \dots, i, j, \dots\}$ , i.e.,

$$p_{i,j} \geq 0, \quad \sum_{k \in S} p_{i,k} = 1, \quad i, j \in S.$$

Assume that it is

- (i) irreducible (each  $i$  communicates with each  $j$  in  $S$ ),
- (ii) positive recurrent (starting from some  $i$  the expected return time to  $i$  has finite expectation).

These properties depend entirely on the matrix  $P$ . Let  $(X_0, X_1, \dots)$  be a Markov chain with values in  $S$  such that  $P(X_{n+1} = j | X_n = i) = p_{i,j}$  for all  $i, j \in S$  and all  $n \geq 0$ . A standard reference for these notions is Brémaud [4].

**THEOREM 3.1.** If (i) and (ii) above hold, then there is a unique probability  $\pi$  on  $S$  such that  $\pi P = \pi$ . This  $\pi$  is necessarily given by

$$\pi(a) = \frac{E_b \sum_{n=0}^{t_b-1} \mathbf{1}(X_n = a)}{E_b t_b}, \quad a \in S, \quad (3.1)$$

where  $E_b$  denotes expectation conditional on  $X_0 = b$ , and

$$t_b := \inf\{n \geq 1 : X_n = b\},$$

for an arbitrary  $b \in S$ .

*Proof.* We wish to show this by using the idea developed in the previous section. First, let  $\pi^{(b)}$  denote the specific probability measure defined by (3.1), i.e.,

$$\pi^{(b)}(a) := \frac{E_b \sum_{n=0}^{t_b-1} \mathbf{1}(X_n = a)}{E_b t_b}. \quad (3.2)$$

It is well known [see, e.g., Brémaud [4]] that  $\pi^{(b)}$  is invariant for  $P$ , i.e., that it satisfies  $\pi^{(b)} P = \pi^{(b)}$ . It is important to note that  $\pi^{(b)}$  depends entirely on the stochastic matrix  $P$  and the choice of the state  $b$  (hence, the superscript).

To show uniqueness, we work at the level of sequences, i.e., with the space  $\Omega = S^{\mathbb{Z}}$ , whose elements are denoted by  $\omega = (\omega_n, n \in \mathbb{Z})$ , equipped with the cylinder  $\sigma$ -algebra  $\mathcal{A}$ . We consider the natural shift

$$\vartheta : (n \mapsto \omega_n) \mapsto (n \mapsto \omega_{n+1}),$$

which is  $\mathcal{A}$ -measurable and invertible. We are thus in the setup of the earlier section. Consider a probability  $\pi$  on  $S$  which is invariant for  $P$ , i.e., it satisfies

$$\pi P = \pi,$$

and let  $P$  be the probability measure on  $(S^{\mathbb{Z}}, \mathcal{A})$  defined by

$$P(\{\omega \in \Omega : \omega_m = i_m, \dots, \omega_n = i_n\}) = \pi(i_m) p_{i_m, i_{m+1}} \cdots p_{i_{n-1}, i_n}, \\ i_m, \dots, i_n \in S, \quad m, n \in \mathbb{Z}, \quad m \leq n. \quad (3.3)$$

Consider also the random variables

$$X_n(\omega) := \omega_n, \quad \omega \in \Omega, \quad n \in \mathbb{Z}.$$

Under  $P$ , the sequence  $(X_n)$  is a Markov chain with transition probability matrix  $P$ . The assumption  $\pi = \pi P$  implies that the measure  $P$  defined by (3.3) is preserved by  $\vartheta$ . By Proposition 2.7, for any  $B \in \mathcal{A}$ , the measures

$$P_B(\cdot) = E \mathbf{1}_B M_B^+(\cdot) = \int_B M_B^+(\cdot) dP,$$

where  $T_B^+, M_B^+$  are given by (2.1a)–(2.1b), are also preserved by  $\vartheta$ . Fix some  $b \in S$ , and consider the set

$$B = \{\omega \in \Omega : \omega_0 = b\}. \quad (3.4)$$

Observe that

$$t_b(\omega) := \inf\{n \geq 1 : \omega_n = b\} = T_B^+(\omega).$$

By Theorem 2.4,

$$P_B(A) = E_{1_B}M_B^+(A) = P(A, T_B^+ < \infty) = P(A, t_b < \infty), \quad A \in \mathcal{A}. \quad (3.5)$$

By (i) and (ii) we have

$$P(t_b < \infty) = P(T_B^+ < \infty) = 1,$$

and so (3.5) yields

$$P_B(A) = P(A), \quad A \in \mathcal{A},$$

while Corollary 2.9 gives

$$E_{1_B}T_B^+ = 1.$$

Therefore,

$$P(A) = P_B(A) = E_{1_B}M_B^+(A) = \frac{E_{1_B}M_B^+(A)}{E_{1_B}T_B^+} = \frac{E_b \sum_{n=0}^{t_b-1} \mathbf{1}(\vartheta^n \omega \in A)}{E_b t_b}. \quad (3.6)$$

Apply this identity to the set

$$A := \{\omega \in \Omega : \omega_0 = a\},$$

for an arbitrary  $a \in S$ . Then the left-hand side of (3.6) equals  $\pi(a)$  while the right-hand side equals  $\pi^{(b)}(a)$  by the definition (3.2) of  $\pi^{(b)}$ . Therefore,  $\pi(a) = \pi^{(b)}(a)$ , for all  $a \in S$ . In other words, an *arbitrary* invariant probability measure  $\pi$  must necessarily be equal to the *specific* measure  $\pi^{(b)}$ ; hence the uniqueness.  $\square$

REMARK 3.2. The last argument directly proves that, for the case of Markov chains,

$$\frac{E_b \sum_{n=0}^{t_b-1} \mathbf{1}(X_n = a)}{E_b t_b} = \frac{E_c \sum_{n=0}^{t_c-1} \mathbf{1}(X_n = a)}{E_c t_c}.$$

This is the so-called exchange formula of (discrete-index) Palm theory; see also [8].

REMARK 3.3. In essence, uniqueness follows from the following two facts: First, thanks to the Markov setting considered here, the law of a cycle starting from a given state until the chain returns to this state is uniquely determined by the transition matrix. Second,  $P$  can be uniquely determined by its Palm measure: see Lemma 2.12 and Remark 2.13.

REMARK 3.4. A direct consequence of these ideas is the following well-known fact: If  $b$  is not a positive recurrent state of a (any) Markov chain, then  $\pi(b) = 0$  for any invariant probability measure  $\pi$ . To see this, let  $\pi$  be an invariant probability measure. Construct  $(\Omega, \mathcal{A}, \vartheta, P)$  as above, and let  $B$  be as in (3.4). Then

$$E[\mathbf{1}(X_0 = b)t_b] = E[\mathbf{1}_B T_B^+] \leq 1,$$

where the last inequality is due to Corollary 2.9. If  $\pi(b)$  were positive, we would obtain  $E_b t_b \leq 1/\pi(b) < \infty$ , in contradiction to the positive recurrence of  $b$ .

**4. Uniqueness of invariant distribution in Harris chains.** A Markov chain in a general state space  $S$ , assumed to be Polish (i.e., a separable metrizable space which is complete under an appropriate metric), with Borel sets  $\mathcal{S}$ , is defined via its transition kernel  $K(x, B)$  which is measurable in  $x \in S$  and is a probability measure in  $B \in \mathcal{S}$ . We let  $K^n(x, B)$  be the  $n$ -th fold convolution, i.e.,  $K^1(x, B) = K(x, B)$ ,  $K^{n+1}(x, B) = \int_S K(x, dy)K^n(y, B)$ ,  $n \geq 1$ . By Kolmogorov's theorem, we can uniquely define probability measures  $P_x$ ,  $x \in S$ , on the space of sequences  $(x_n, n \in \mathbb{Z}_+) \in S^{\mathbb{Z}_+}$  such that  $P_x(x_0 = x) = 1$  and  $P_x(x_{n+1} \in \cdot | x_n, x_{n-1}, \dots) = K(x_n, \cdot)$ . A measure  $\mu$  on  $(S, \mathcal{S})$  is called invariant for  $K$  if  $\mu(B) = \int_S \mu(dx)K(x, B)$ ,  $B \in \mathcal{S}$ . We are interested in showing uniqueness of an invariant probability measure for  $K$ , under the condition that the chain is positive Harris recurrent, which we shall define below. First, we present some historical overview, along with definitions. Doeblin [5] first introduced the condition that there exists a reference probability measure on  $(S, \mathcal{S})$  and a positive integer  $\ell$ , such that the probability  $K^\ell(x, B)$  is bounded from below by a positive number  $\varepsilon$  for all  $x \in S$  and all Borel sets  $B$  which are assigned a reference probability larger than some nonnegative number  $\delta < 1$ . Under this condition, he proved an "ergodic theorem", in the sense that  $K^n(x, \cdot)$  converges, as  $n \rightarrow \infty$ . Harris [6] introduced the condition that there exists a sigma-finite measure  $\varphi$  on  $(S, \mathcal{S})$  such that every set  $B \in \mathcal{S}$  is visited infinitely often by the Markov chain, starting from any initial state. This condition (known as recurrence in the sense of Harris) turns out to be weaker than Doeblin's, but equivalent to the one introduced by Orey [14] [see also Orey [15]] which is the one we adopt here [see also Asmussen [1]]:

There exists a set  $R \in \mathcal{S}$ , a probability measure  $\lambda$  on  $(S, \mathcal{S})$ , a positive integer  $\ell$ , and  $\varepsilon > 0$ , such that

- (i)  $P_x(\mathbf{t}_R < \infty) = 1$  for all  $x \in S$ , where  $\mathbf{t}_R := \inf\{n \in \mathbb{N} : x_n \in R\}$ .
- (ii)  $K^\ell(x, \cdot) \geq \varepsilon\lambda(\cdot)$ , for all  $x \in R$ .

The chain is called *positive Harris recurrent* if, in addition to (i) and (ii) we also have

- (iii)  $E_\lambda \mathbf{t}_R < \infty$ , where  $E_\lambda$  denotes expectation with respect to  $P_\lambda(\cdot) := \int_S \lambda(dx)P_x(\cdot)$ .

Without using the ergodic theorem, and based on Section 2, we give a proof of the following:

**THEOREM 4.1.** A positive Harris recurrent chain possesses a unique invariant probability measure.

Note that this theorem, without the positivity condition, is proved by Harris [6], Orey [14] and in the paper of Athreya and Ney [2]; the latter authors introduced the idea of splitting to provide a construction of the Harris chain on an enlarged probability space. This was also done, independently, by Nummelin [13]. [See also Meyn and Tweedie [10].] Our contribution here is that the uniqueness proof is a direct corollary of the results of Section 2, in the case of positive Harris recurrence, with an appeal to the ergodic theorem.

*Proof of Theorem 4.1.* Suppose first that  $\ell = 1$ . We first describe the construction of a Markov chain  $(x_n, n \geq 0)$  with kernel  $K$  on an enlarged probability space as in

Athreya and Ney [2] and Nummelin [13]. Fix a probability measure  $\lambda$  and a number  $\varepsilon > 0$  for which condition (ii) above holds. Let  $p$  be the Bernoulli probability measure on the two-element set  $\{0, 1\}$  such that

$$p(1) = \varepsilon, \quad p(0) = 1 - \varepsilon.$$

We consider a Markov chain  $((x_n, \zeta_n), n \geq 0)$  with values in  $S_* := S \times \{0, 1\}$  defined by the following kernel (here  $\alpha, \beta$  range in  $\{0, 1\}$  and  $x, y$  in  $S$ ):

$$\begin{aligned} \text{If } x \notin R, \text{ then } K_*((x, \alpha), (dy, \beta)) &= K(x, dy) p(\beta). \\ \text{If } x \in R, \text{ then } \begin{cases} K_*((x, 1), (dy, \beta)) &= \lambda(dy) p(\beta), \\ K_*((x, 0), (dy, \beta)) &= Q(x, dy) p(\beta), \end{cases} \end{aligned}$$

where  $Q$  is the kernel on  $S$  defined by

$$Q(x, \cdot) := \frac{K(x, \cdot) - \varepsilon\lambda(\cdot)}{1 - \varepsilon}.$$

Condition (ii) is responsible for the nonnegativity of  $Q$ . We let  $\mathbb{P}_{(x, \alpha)}$ , where  $(x, \alpha) \in S_*$ , be the probability measure on the canonical space  $S_*^{\mathbb{Z}^+}$  (defined by means of Kolmogorov's theorem), making the coordinate process  $((x_n, \zeta_n), n \geq 0)$  a Markov chain with kernel  $K_*$  and initial state  $(x, \alpha)$ . We let

$$\mathbb{P}_x := \varepsilon\mathbb{P}_{(x, 1)} + (1 - \varepsilon)\mathbb{P}_{(x, 0)}.$$

It is easy to see that, under  $\mathbb{P}_x$ , the sequence  $(x_n, n \geq 0)$  is a Markov chain in  $S$  with kernel  $K$  and that the sequence  $(\zeta_n, n \geq 0)$  is a sequence of i.i.d. Bernoulli random variables with common law  $p(\cdot)$ ; the two sequences are not independent under  $\mathbb{P}_x$ . The initial idea for this construction comes from the paper of Athreya and Ney [2]. Here we follow the formulation and construction of Thórisson [21, Chapter 10, pp. 365-369] who shows that the times at which the pair process  $(x_n, \zeta_n)$  visits the set

$$R \times \{1\}$$

are  $\mathbb{P}_x$ -a.s. finite for all  $x \in S$ . Owing to the definition of the kernel  $K_*$ , it follows that at the  $\mathbb{P}_x$ -a.s. finite random time

$$t := \inf\{n \geq 1 : (x_{n-1}, \zeta_{n-1}) \in R \times \{1\}\} \tag{4.1}$$

the pair process regenerates with  $(x_t, \zeta_t)$  having law  $\lambda \times p$ . This enables us to define the probability measure

$$\pi_*(C) := \frac{\mathbb{E}_{\lambda \times p} \sum_{n=0}^{t-1} \mathbf{1}(x_n \in C)}{\mathbb{E}_{\lambda \times p} t}, \quad C \in \mathcal{S}, \tag{4.2}$$

which can be seen to be invariant for  $K$  [see Meyn and Tweedie [10] and Asmussen [1]].

We now wish to show that  $\pi_*$  is the unique invariant distribution for  $K$ . Let  $\pi$  be a probability measure on  $(S, \mathcal{S})$  which is invariant for the kernel  $K$ . On the enlarged probability space

$$\Omega := S_*^{\mathbb{Z}}$$

of two-sided sequences

$$\omega = (\omega_n = (x_n, \zeta_n), n \in \mathbb{Z}),$$

consider the probability measure  $\mathbb{P}$  under which  $(x_n, \zeta_n), n \in \mathbb{Z}$ , is a stationary Markov chain with kernel  $K_*$  and initial distribution  $\pi \times p$ . Define the natural shift  $\vartheta$  on  $\Omega$ :

$$\vartheta : (n \mapsto (x_n, \zeta_n)) \mapsto (n \mapsto (x_{n+1}, \zeta_{n+1})).$$

The invariance of  $\pi$  for  $K$  implies that  $\mathbb{P}$  is preserved by  $\vartheta$ . We are thus in the setup of Section 2. As in (2.1a), we shall consider the random time  $T_B^+$  and, as in (2.6a), we will work with the probability measure

$$\mathbb{P}_B(\cdot) = \mathbb{E} \mathbf{1}_B M_B^+(\cdot) = \int_B M_B^+(\cdot) d\mathbb{P}.$$

Consider next the set  $B \in \mathcal{A}$  defined by

$$B := \{\omega \in \Omega : \omega_{-1} \in R \times \{1\}\}.$$

Note that, as in (2.1a),

$$\begin{aligned} T_B^+(\omega) &= \inf\{n \geq 1 : \vartheta^n \omega \in B\} \\ &= \inf\{n \geq 1 : (\vartheta^n \omega)_{-1} \in R \times \{1\}\} \\ &= \inf\{n \geq 1 : \omega_{n-1} \in R \times \{1\}\} = \mathfrak{t}(\omega), \end{aligned}$$

as seen by (4.1). The rest proceeds as in the proof of Theorem 3.1. Since  $\mathbb{P}_x(\mathfrak{t} < \infty) = 1$  for all  $x$ , we have  $\mathbb{P}(\mathfrak{t} < \infty) = 1$ , i.e.  $\mathbb{P}(T_B^+ < \infty) = 1$ , and so, by Theorem 2.4,

$$\mathbb{P}_B(A) = \mathbb{P}(A), \quad A \in \mathcal{A}.$$

By Corollary 2.9,

$$\mathbb{E} \mathbf{1}_B T_B^+ = 1.$$

Therefore,

$$\mathbb{P}(A) = \mathbb{P}_B(A) = \mathbb{E} \mathbf{1}_B M_B^+(A) = \frac{\mathbb{E} \mathbf{1}_B M_B^+(A)}{\mathbb{E} \mathbf{1}_B T_B^+}.$$

But

$$\begin{aligned} \mathbb{E} \mathbf{1}_B T_B^+ &= \mathbb{E}[\mathfrak{t} \cdot \mathbf{1}(\omega_{-1} \in R \times \{1\})] = \mathbb{E}[\mathfrak{t} \mid \omega_{-1} \in R \times \{1\}] \cdot \mathbb{P}(\omega_{-1} \in R \times \{1\}) \\ &= \mathbb{E}_{\lambda \times p}[\mathfrak{t}] \mathbb{P}(\omega_{-1} \in R \times \{1\}). \end{aligned}$$

Similarly,

$$\mathbb{E} \mathbf{1}_B M_B^+(A) = \mathbb{P}(\omega_{-1} \in R \times \{1\}) \mathbb{E}_{\lambda \times p} \sum_{n=0}^{\mathfrak{t}-1} \mathbf{1}(\omega_n \in A).$$

Therefore,

$$\mathbb{P}(A) = \frac{\mathbb{E}_{\lambda \times p} \sum_{n=0}^{\mathfrak{t}-1} \mathbf{1}(\omega_n \in A)}{\mathbb{E}_{\lambda \times p}[\mathfrak{t}]},$$

for all  $A \in \mathcal{A}$ , the cylinder  $\sigma$ -field of  $S_*^{\mathbb{Z}}$ . Now let  $C \in \mathcal{S}$  and define  $A \in \mathcal{A}$  by

$$A := \{\omega \in \Omega : \omega_0 = (x_0, \zeta_0) \in C \times \{0, 1\}\}.$$

Then  $\mathbb{P}(A) = \pi(C)$  and

$$\frac{\mathbb{E}_{\lambda \times p} \sum_{n=0}^{\mathfrak{t}-1} \mathbf{1}(\omega_n \in A)}{\mathbb{E}_{\lambda \times p}[\mathfrak{t}]} = \frac{\mathbb{E}_{\lambda \times p} \sum_{n=0}^{\mathfrak{t}-1} \mathbf{1}(x_n \in C)}{\mathbb{E}_{\lambda \times p}[\mathfrak{t}]} = \pi_*(C),$$

by (4.2). Hence  $\pi(C) = \pi_*(C)$  for all  $C \in \mathcal{S}$ . Recall that  $\pi_*$  depends on  $\lambda$  and  $\varepsilon$ ; hence uniqueness follows.

The case  $\ell > 1$  is similar. All that is required is the construction of the extended process  $((x_n, \zeta_n), n \geq 1 - \ell)$  with values in  $S_* = S \times \{0, 1\}$ . This is again found in Thórisson [21, Chapter 10, pp. 365-369]. The difference with the  $\ell = 1$  case is that this extended process is an  $\ell$ th order Markov chain with transition kernel, say,  $K_*$ . The times at which the process visits the set  $R \times \{1\}$  are a.s. finite, regardless of the initial state. In fact, the process regenerates at the time

$$t = \inf\{n \geq 1 : (x_{n-\ell}, \zeta_{n-\ell}) \in R \times \{1\}\}$$

and  $(x_t, \zeta_t)$  has law  $\lambda \times p$ . As before, formula (4.2) defines an invariant probability measure  $\pi_*$  for  $K$ . The rest of the proof proceeds as before; we only sketch it here. On the enlarged probability space  $\Omega = S_*^{\mathbb{Z}^+}$  with the natural shift  $\vartheta$ , we let  $\mathbb{P}$  be the probability measure that makes the coordinate process  $((x_n, \zeta_n), n \in \mathbb{Z})$  an  $\ell$ -th order Markov chain with kernel  $K_*$  and marginal distribution  $\pi \times p$ . Also,  $\mathbb{P}$  is preserved by  $\vartheta$ . The set  $B$  is defined as  $\{\omega \in \Omega : \omega_{-\ell} \in R \times \{1\}\}$ , and, on choosing  $A$  as before, we can, exactly in the same manner, show that the arbitrary invariant probability measure  $\pi$  equals  $\pi_*$ .  $\square$

**Acknowledgments.** The authors would like to thank two anonymous referees whose comments substantially improved the appearance of this paper. Remark 3.4 was pointed out to them by one of the referees.

#### REFERENCES

- [1] ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd ed. Springer-Verlag. MR1978607 (2004f:60001)
- [2] ATHREYA, K.B. AND NEY, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245**, 493-501. MR511425 (80i:60092)
- [3] BACCELLI, F. AND BRÉMAUD, P. (2003). *Elements of Queueing Theory*. Springer-Verlag. MR1957884 (2003m:60001)
- [4] BRÉMAUD, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer-Verlag. MR1689633 (2000k:60137)
- [5] DOEBLIN, W. (1940). Éléments d'une théorie générale des chaînes simples constantes de Markoff. *Annales Scientifiques de l'École Normale Supérieure*, Sér. 3, **57**, 61-111. MR0004409 (3:3d)
- [6] HARRIS, T.E. (1956). The existence of stationary measures for certain Markov processes. *Proc. 3rd Berkeley Symposium Math. Stat. Prob.* **2**, 113-124, Univ. of California Press. MR0084889 (18:941d)
- [7] KAC, M. (1947). On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc.* **53**, 1002-1010. MR0022323 (9:194a)
- [8] KONSTANTOPOULOS, T. AND ZAZANIS, M. (1995). A discrete time proof of Neveu's exchange formula. *J. Appl. Probability* **32**, 917-921. MR1363333 (96h:60081)
- [9] LIND, D. AND MARCUS, B. (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press. MR1369092 (97a:58050)
- [10] MEYN, S.P. AND TWEEDIE, R.L. (1993). *Markov Chains and Stochastic Stability*. Springer-Verlag, London. MR1287609 (95j:60103)
- [11] MOLCHANOV, I. (2005). *Theory of Random Sets*. Springer-Verlag. MR2132405 (2006b:60004)
- [12] NORRIS, J.R. (1998). *Markov Chains*. Cambridge University Press. MR1600720 (99c:60144)
- [13] NUMMELIN, E. (1978). A splitting technique for Harris recurrent Markov chains. *Zeit. Wahr. v. Gebiete* **43**, 309-318. MR0501353 (58:18732)
- [14] OREY, S. (1959). Recurrent Markov chains. *Pacific J. Math.* **9**, 805-827. MR0125632 (23:A2931)

- [15] OREY, S. (1971). *Limit Theorems for Markov Chain Transition Probabilities*. Van Nostrand Reinhold, London. MR0324774 (48:3123)
- [16] PETERSEN, K. (1983). *Ergodic Theory*. Cambridge University Press. MR833286 (87i:28002)
- [17] POLLICOTT, M. AND YURI, M. (1998). *Dynamical Systems and Ergodic Theory*. Cambridge University Press. MR1627681 (99f:58130)
- [18] SINAI, YA.G. (1976). *Introduction to Ergodic Theory*. Princeton University Press. MR0584788 (58:28437)
- [19] SCHNEIDER, R. AND WEIL, W. (2000). *Stochastische Geometrie*. Teubner Stuttgart-Leipzig. MR1794753 (2001j:60026)
- [20] SLIVNYAK, I.M. (1962). Some properties of stationary flows of homogeneous random events. *Th. Prob. Appl.* **7**, 336-341.
- [21] THÓRISSON, H. (2000). *Coupling, Stationarity, and Regeneration*. Springer-Verlag. MR1741181 (2001b:60003)