PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 28, Fasc. 1 (2008), pp. 75–90

TWO TYPES OF MARKOV PROPERTY

BY

ARKADIUSZ KASPRZYK (WROCŁAW) AND WŁADYSŁAW SZCZOTKA (WROCŁAW)

Abstract. The paper gives some insight into the relations between two types of Markov processes – in the strict sense and in the wide sense – as well as into two aspects of periodicity. It concerns Markov processes with finite state space, the elements of which are complex numbers. Firstly it is shown that under some assumptions this space can be transformed in such a way that the resulting Markov process is also Markov in the wide sense. Next, sufficient conditions are given under which periodic homogeneous Markov chain is a periodically correlated process.

2000 AMS Mathematics Subject Classification: Primary: 60J10, 60J27; Secondary: 60G12.

Key words and phrases: Markov chains in the strict sense and in the wide sense, linear prediction, covariance function, cyclostationary distribution, periodically correlated process.

1. INTRODUCTION

The paper deals with a relation between the notion of a *Markov process in the strict sense* (MP) and a *Markov process in the wide sense* (WM), also called *wide-sense Markov*. These processes are considered both in a continuous time and a discrete time; in the latter case they are called *Markov chains* (MC). For this, let $\mathcal{T}$ be the set of time indices, i.e. $\mathcal{T} = [0, \infty)$ for a continuous time and $\mathcal{T} = \{0, 1, \dots\} \equiv \mathbb{N}$ for a discrete time. We say that the process $\{X_t,\ t \in \mathcal{T}\} \equiv \{X_t\}$ with values in $\mathbb{C}$ and mean zero is WM if for each positive integer m and $t > t_1 > t_2 > \dots > t_m$ the following holds:

$$\widehat{\mathbf{E}}[X_t | X_{t_1}, \dots, X_{t_m}] = \widehat{\mathbf{E}}[X_t | X_{t_1}],$$

where for the random vector $\mathbf{Y} = (Y_1, \dots, Y_m)^{\mathsf{T}}$ the notation $\widehat{\mathbf{E}}[X|Y_1, \dots, Y_m] \equiv \widehat{\mathbf{E}}[X|\mathbf{Y}]$ is used, which is the minimum mean-square error projection of random variable X on the linear space generated by coordinates of $\mathbf{Y}$, i.e. $\mathbf{E}\big|X - \widehat{\mathbf{E}}[X|\mathbf{Y}]\,\big|^2 = \min_{\mathbf{a} \in \mathbb{C}^m} \mathbf{E}|X - \mathbf{a}^{\mathsf{T}}\mathbf{Y}|^2$. Here and below we use the notation $\mathbf{x}^{\mathsf{T}}$ for transposition of a vector or matrix $\mathbf{x}$. If the discrete time WM process is stationary, then it is an autoregressive process of order one, denoted by AR(1). The notion of WM processes was introduced by J. L. Doob in [2], where he proved the following characterization of such processes. Process $\{X_t\}$ with $\mathbf{E}[X_t] = 0$ and $\mathbf{E}|X_t|^2 < \infty$ is WM iff its normalized autocovariance function $R(t,s) := \mathbf{E}[X_t\overline{X}_s](\mathbf{E}|X_s|^2)^{-1}$ satisfies the following functional equation, called the *triangular relation*:

$$R(t_3, t_1) = R(t_3, t_2)R(t_2, t_1) \quad \text{for each } t_3 \geqslant t_2 \geqslant t_1. \tag{1.1}$$

This characterization indicates the second order nature of WM property. Hence investigation of the autocovariance function of the given process must be performed in order to determine if the process is or not WM.

The main problem of the paper deals with comparing two types of Markov property (Section 3) and two types of periodicity (Section 5). As for the first problem we compare properties of Markov process in the strict sense and in the wide sense. Markov process in the strict sense has not to have numerical values, but if we assume so, it has not to have the mean zero or finite second moment. Therefore the strict MP has not to be Markov in the wide sense. In Section 4 we show that if the infinitesimal operator of MP is diagonalizable and at least one of its right eigenvectors is real and has distinct coordinates, then MP can be transformed to a real-valued WM process (Theorem 4.1). To get a complex-valued process it is enough to assume that there exists a right eigenvector with distinct coordinates, which can be also complex. A similar result is formulated for homogeneous Markov processes in a discrete time, i.e. for homogeneous Markov chains (HMC) (Theorem 4.2). In Section 5 we consider two types of periodicity of MP, namely, periodicity in the usual sense of HMC (see [1], p. 72) and periodically correlated Markov chains. Theorem 5.1 gives conditions under which a periodic HMC is periodically correlated. Furthermore, we give an asymptotic of autocovariance function in a general case of HMC.

2. PRELIMINARIES

In this section we introduce some notation and assumptions for homogeneous Markov processes (HMP). Next, in terms of this notation we recall some well-known results which will be used in next sections.

Let $\{X_t\}$ be a continuous time HMP with finite state space $\mathfrak{X} = \{x_0, \ldots, x_N\}$, $x_i \in \mathbb{C}$, an irreducible transition probability matrix $\mathbb{P}(t) = \big(p_{ij}(t)\colon 0 \leqslant i,j \leqslant N\big)$, where $p_{ij}(t) = \mathsf{P}\{X_t = x_j | X_0 = x_i\}$, and an initial distribution $\mathbf{p} = (p_0, \ldots, p_N)$, where $p_i = \mathsf{P}\{X_0 = x_i\}$, $i = 0, 1, \ldots, N$. Assume that the transition semigroup $\{\mathbb{P}(t)\colon t \geqslant 0\}$ is continuous at $t = 0$, i.e. $\lim_{t\to 0} \mathbb{P}(t) = \mathbb{P}(0) = \mathbb{I}$, where $\mathbb{I}$ is the identity matrix. Then (see [1], p. 339) it has an infinitesimal operator $\mathbb{Q} = (q_{ij}\colon 0 \leqslant i,j \leqslant N)$ which is stable, i.e. $0 \leqslant -q_{ii} < \infty$, and conservative, i.e. $\mathbb{Q}\mathbf{1} = \mathbf{0}$, where $\mathbf{1} := (1, 1, \ldots, 1)^{\mathsf{T}} \in \mathbb{R}^{N+1}$. Moreover,

$$\mathbb{P}(t) = \exp(t\mathbb{Q}) = \sum_{k=0}^{\infty} \frac{t^k \mathbb{Q}^k}{k!}, \quad t \geqslant 0. \tag{2.1}$$

The distribution of X_t at any time $t \geqslant 0$ is defined as

$$\mathbf{p}(t) := \big(p_0(t), p_1(t), \ldots, p_N(t)\big), \quad \text{where} \quad p_i(t) := \mathsf{P}\{X_t = x_i\}.$$

Its stationary distribution (if it exists) is denoted by $\pi = (\pi_0, \pi_1, \ldots, \pi_N)$. Hence we have $\mathbf{p} = \mathbf{p}(0)$, $\mathbf{p}(t) = \mathbf{p}\mathbb{P}(t)$, $\mathbf{p}(t+s) = \mathbf{p}(t)\mathbb{P}(s)$ and $\pi = \pi\mathbb{P}(t)$ for all $t \geqslant 0$.

Recall that $\mathbb{Q}$ is *diagonalizable* (in [5] it is called *simple*) if there exist a diagonal matrix Λ and a nonsingular matrix $\mathbb{H}$ such that $\mathbb{Q} = \mathbb{H}\Lambda\mathbb{H}^{-1}$. In this case, $\mathbb{P}(t)$ is also diagonalizable, for each $t > 0$, by (2.1). Let $\lambda_0, \lambda_1, \ldots, \lambda_N$ be the eigenvalues of $\mathbb{Q}$ such that $\operatorname{Re}\lambda_0 \geqslant \operatorname{Re}\lambda_1 \geqslant \ldots \geqslant \operatorname{Re}\lambda_N$ and let $\mathbf{l}_0, \mathbf{l}_1, \ldots, \mathbf{l}_N$ and $\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N$ be the sequences of the corresponding left and right eigenvectors, respectively. Then the following holds:

PROPOSITION 2.1 (see [1], [5]). *If* $\mathbb{Q}$ *is diagonalizable*, *then*:

(i) $\Lambda = \operatorname{diag}(\lambda_0, \lambda_1, \ldots, \lambda_N)$;

(ii) *the systems of eigenvectors* $\{\mathbf{r}_j\}$ *and* $\{\mathbf{l}_j\}$ *can be chosen in such a way that* $\mathbb{H} = [\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N]$ *and* $\mathbb{H}^{-1} = [\mathbf{l}_0, \mathbf{l}_1, \ldots, \mathbf{l}_N]^{\mathsf{T}}$, *so they are quasi-biorthogonal*, *i.e.* $\mathbf{l}_j^{\mathsf{T}}\mathbf{r}_k = \delta_{jk}$ (see [5]); *if all eigenvalues are distinct*, *then* $\{\mathbf{r}_j\}$ *and* $\{\mathbf{l}_j\}$ *are uniquely determined*;

(iii) $\lambda_0 = 0 > \operatorname{Re}\lambda_1 \geqslant \ldots \geqslant \operatorname{Re}\lambda_N$ (the Perron–Frobenius theorem; see [1] or [5], [6], [9]);

(iv) $\mathbf{r}_0 = (1, 1, \ldots, 1) \in \mathbb{R}^{N+1}$ *and* $\mathbf{l}_0^{\mathsf{T}}$ *is the stationary distribution of* $\{X_t\}$, *i.e.* $\mathbf{l}_0^{\mathsf{T}} = \pi$.

If $\{\mathbf{r}_j\}$ and $\{\mathbf{l}_j\}$ are quasi-biorthogonal, then

$$\mathbb{Q} = \sum_{j=0}^{N} \lambda_j \mathbb{G}_j, \quad \text{where} \quad \mathbb{G}_j = \mathbf{r}_j \mathbf{l}_j^{\mathsf{T}}. \tag{2.2}$$

$\mathbb{G}_j$'s are called *constituent matrices* of $\mathbb{Q}$ (see [5]) and they have the following properties: $\mathbb{G}_j\mathbb{G}_k = \delta_{jk}\mathbb{I}$, $\sum_{j=0}^{N}\mathbb{G}_j = \mathbb{H}\mathbb{H}^{-1} = \mathbb{I}$, $\operatorname{rank}(\mathbb{G}_j) = 1$, and

$$\mathbb{G}_j \mathbf{r}_k = \delta_{jk}\mathbf{r}_k. \tag{2.3}$$

The relation (2.3) implies spectral representation for the transition probability matrix. Namely,

$$\mathbb{P}(t) = \sum_{r=0}^{N} \exp(t\lambda_r)\mathbb{G}_r = \mathbb{H}\mathbb{K}(t)\mathbb{H}^{-1}, \tag{2.4}$$

where $\mathbb{K}(t) = \operatorname{diag}\big(\exp(t\lambda_0), \exp(t\lambda_1), \ldots, \exp(t\lambda_N)\big) = \exp(t\Lambda)$. This implies that $\mathbb{Q} = \sum_{j=1}^{N}\lambda_j\mathbb{G}_j$ and the rank of $\mathbb{Q}$ is at most N. Furthermore, we have

$$\mathbb{P}(t) = \mathbb{G}_0 + \sum_{j=1}^{N} \exp(t\lambda_j)\mathbb{G}_j = \mathbf{1}\pi + \mathbb{Z}(t) \to \mathbf{1}\pi, \quad t \to \infty,$$

whenever the process $\{X_t\}$ is irreducible. This convergence implies that $\mathbb{Z}(t)$ converges to zero matrix as t tends to infinity. Moreover, $\pi\mathbb{Z}(t) = \mathbf{0}^\mathsf{T}$ and $\mathbb{Z}(t)\mathbf{1} = \mathbf{0}$. In order to formulate the next auxiliary result let us put $\mathbf{x} := (x_0, x_1, \ldots, x_N)^\mathsf{T}$ and $\mathbb{D} := \operatorname{diag}(\mathbf{x})$. Then, obviously, $\mathbb{D}\mathbf{1} = \mathbf{x}$.

LEMMA 2.1. *If $\{X_t,\, t \in [0,\infty)\}$ is HMP with initial distribution $\mathbf{p}$, then its moments $\mathbf{E_p}[X_t]$, $\mathbf{E_p}|X_t|^2$ and $\mathbf{E_p}[X_{t+h}\overline{X}_t]$ and its covariance function $\operatorname{Cov}(X_{t+h}, X_t)$ have the following form*:

$$\mu(t;\mathbf{p}) := \mathbf{E_p}[X_t] = \mathbf{p}(t)\mathbf{x}, \quad t \geqslant 0, \tag{2.5}$$

$$\mu_2(t;\mathbf{p}) := \mathbf{E_p}|X_t|^2 = \mathbf{p}(t)\overline{\mathbb{D}}\mathbf{x}, \quad t \geqslant 0, \tag{2.6}$$

$$\mu(t+h,t;\mathbf{p}) := \mathbf{E_p}[X_{t+h}\overline{X}_t] = \mathbf{p}(t)\overline{\mathbb{D}}\mathbb{P}(h)\mathbf{x}, \quad t \geqslant h \geqslant 0, \tag{2.7}$$

$$\Gamma(t+h,t;\mathbf{p}) := \operatorname{Cov}(X_{t+h}, X_t) = \mathbf{p}(t)\overline{\mathbb{D}}\big(\mathbb{I} - \mathbf{1}\mathbf{p}(t)\big)\mathbb{P}(h)\mathbf{x}. \tag{2.8}$$

Proof. Formula (2.5) comes directly from the definition of expectation of r.v. X_t. Formula (2.6) follows from $\overline{\mathbb{D}}\mathbf{x} = (|x_0|^2, |x_1|^2, \ldots, |x_N|^2)^\mathsf{T}$ and from the definition of expectation of r.v. $|X_t|^2$. To see (2.7) observe that

$$\begin{aligned}
\mathbf{E_p}[X_{t+h}\overline{X}_t] &= \sum_i \sum_j x_j \overline{x}_i \mathsf{P}(X_{t+h} = x_j, X_t = x_i) \\
&= \sum_i \sum_j x_j \overline{x}_i\, \mathsf{P}(X_h = x_j | X_0 = x_i) \mathsf{P}(X_t = x_i) \\
&= \sum_i p_i(t) \big(\sum_j p_{ij}(h) x_j\big) \overline{x}_i = \mathbf{p}(t)\overline{\mathbb{D}}\mathbb{P}(h)\mathbf{x}.
\end{aligned}$$

Finally, equality (2.8) is an immediate consequence of (2.5) and (2.7). Namely,

$$\begin{aligned}
&\operatorname{Cov}(X_{t+h}, X_t) = \mu(t+h,t;\mathbf{p}) - \overline{\mu(t;\mathbf{p})}\mu(t+h;\mathbf{p}) \\
&\qquad = \mathbf{p}(t)\overline{\mathbb{D}}\mathbb{P}(h)\mathbf{x} - \mathbf{p}(t)\overline{\mathbb{D}}\mathbf{1}\mathbf{p}(t)\mathbb{P}(h)\mathbf{x} = \mathbf{p}(t)\overline{\mathbb{D}}\big(\mathbb{I} - \mathbf{1}\,\mathbf{p}(t)\big)\mathbb{P}(h)\mathbf{x},
\end{aligned}$$

which completes the proof of the lemma. ∎

All the theory above can be easily adopted to the discrete time case if the considered HMC is irreducible. The difference is that we consider the one-step transition probability matrix $\mathbb{P}$ instead of infinitesimal operator $\mathbb{Q}$. For this case we denote the eigenvalues of $\mathbb{P}$ by $\kappa_0, \kappa_1, \ldots, \kappa_N$, where $\kappa_0 = 1 > |\kappa_1| \geqslant \ldots \geqslant |\kappa_N|$, and the corresponding left and right eigenvectors by $\mathbf{l}_0, \mathbf{l}_1, \ldots, \mathbf{l}_N$ and $\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N$, respectively, as in the continuous case. The assertions (ii) and (iv) of Proposition 2.1 remain valid and the analogue of (2.4) can be expressed in the following proposition.

PROPOSITION 2.2 (see [1], p. 196). *If the transition probability matrix $\mathbb{P}$ of HMC is irreducible and diagonalizable*, *then its spectral representation has the* *following form*:

$$\mathbb{P}^n = \sum_{j=0}^{N} \kappa_j^n \mathbb{G}_j = \mathbb{H}\mathbb{K}^n\mathbb{H}^{-1},$$

where $\mathbb{K} = \operatorname{diag}(\kappa_0, \kappa_1, \ldots, \kappa_N)$ *and* $\mathbb{G}_j = \mathbf{r}_j \mathbf{l}_j^{\mathsf{T}}$.

LEMMA 2.2. *If* $\{X_t,\ t \in \mathbb{N}\}$ *is HMC with initial distribution* $\mathbf{p}$, *then the following holds*:

$$\mu(n; \mathbf{p}) = \mathbf{p}(n)\mathbf{x}, \quad n \geqslant 0, \tag{2.9}$$

$$\mu_2(n; \mathbf{p}) = \mathbf{p}(n)\overline{\mathbb{D}}\mathbf{x}, \quad n \geqslant 0, \tag{2.10}$$

$$\mu(n+h, n; \mathbf{p}) = \mathbf{p}(n)\overline{\mathbb{D}}\mathbb{P}(h)\mathbf{x}, \quad n \geqslant h \geqslant 0, \tag{2.11}$$

$$\Gamma(n+h, n; \mathbf{p}) = \mathbf{p}(n)\overline{\mathbb{D}}\big(\mathbb{I} - \mathbf{1}\mathbf{p}(n)\big)\mathbb{P}(h)\mathbf{x}, \quad n \geqslant h \geqslant 0. \tag{2.12}$$

The following theorem from [4] gives the form of the spectral density for stationary non-periodic HMC.

THEOREM 2.1 (Lai [4]). *Let the transition probability matrix* $\mathbb{P}$ *be irreducible*, *aperiodic and diagonalizable with stationary distribution* π. *Then the stationary HMC* $\{X_n\}$ *with the state space* $\mathfrak{X} = \{0, 1, \ldots, N\}$ *and the transition probability matrix* $\mathbb{P}$ *has the following spectral density function*:

$$f(\omega) = \frac{1}{2\pi} \sum_{j=1}^{N} \frac{\tau_j (1 - \kappa_j^2)}{|1 - \kappa_j e^{\mathbf{i}\omega}|^2}, \quad -\pi \leqslant \omega \leqslant \pi,$$

where $\tau_j = \pi\mathbb{D}\mathbb{G}_j\mathbf{x}$ *and* $\mathbf{i}^2 = -1$. *Moreover*, *this process is an autoregressive process of order one if the following conditions hold*:

(i) *all eigenvalues* κ_j *are distinct and real*,

(ii) *either* $\pi\mathbb{D}\mathbf{r}_j = 0$ *or* $\mathbf{l}_j^{\mathsf{T}}\mathbf{x} = 0$ *for all* $j = 2, 3, \ldots, N$ (*where* $\mathbf{r}_j$ *and* $\mathbf{l}_j$ *are the right and left eigenvectors, respectively, corresponding to* κ_j).

3. TWO TYPES OF MARKOV PROPERTY

This section deals with a relation between the Markov property and Markov property in the wide sense for process $\{X_t\}$ with complex state space and general time index set $\mathcal{T}$.

THEOREM 3.1. *Let* $\{X_t\}$ *be an HMP with complex state space*, $\mathbf{E}[X_t] = 0$ *and* $\mathbf{E}|X_t|^2 < \infty$ *for all* $t \in \mathcal{T}$. *If* $\mathbf{E}[X_t|X_s] = a(t,s)X_s$, *where* $a(t,s) \in \mathbb{C}$, *then* $\{X_t\}$ *is a Markov process in the wide sense.*

P r o o f. By the assumption that $\mathbf{E}[X_t|X_s] = a(t,s)X_s$ for some $a(t,s) \in \mathbb{C}$ and by the properties of conditional expectation and the linear prediction we have the following relations:

$$\mathbf{E}|X_t - a(t,s)X_s|^2 = \mathbf{E}\big|X_t - \mathbf{E}[X_t|X_s]\big|^2,$$
$$\mathbf{E}\big|X_t - \mathbf{E}[X_t|X_s]\big|^2 = \mathbf{E}\big|X_t - \mathbf{E}[X_t|X_r, r \leqslant s]\big|^2 \leqslant \mathbf{E}\big|X_t - \widehat{\mathbf{E}}[X_t|X_r, r \leqslant s]\big|^2$$
$$\leqslant \mathbf{E}\big|X_t - \widehat{\mathbf{E}}[X_t|X_s]\big|^2 \leqslant \mathbf{E}|X_t - a(t,s)X_s|^2 = \mathbf{E}\big|X_t - \mathbf{E}[X_t|X_s]\big|^2.$$

Hence

$$\widehat{\mathbf{E}}[X_t|X_r, r \leqslant s] \overset{\mathsf{P1}}{=} \widehat{\mathbf{E}}[X_t|X_s] \quad \text{and} \quad \widehat{\mathbf{E}}[X_t|X_s] = \mathbf{E}[X_t|X_s],$$

which gives the assertion of the theorem. ■

The following corollary is an immediate consequence of Theorem 3.1.

COROLLARY 3.1 (see [2]). *If* $\{X_t\}$ *is a Gaussian HMP with mean zero, then* $\{X_t\}$ *is a WM process.*

Reynolds gave an analytical condition for HMP with discrete state space for which $\mathbf{E}[X_t|X_s]$ is a linear function of X_s (see [7], formula (11)). Some queueing example of such a process is also given there.

It is clear that Markov process with mean zero has not to be Markov process in the wide sense. To see this it is enough to consider HMP with autocovariance function not satisfying the triangular condition (1.1). The autocovariance function for HMP with $\mathbb{Q}$ diagonalizable with state space $\mathfrak{X} = \{0, 1, \ldots, N\}$ is given in [8] in the following form:

$$\gamma(h) = \Gamma(t+h, t; \pi) = \sum_{j=1}^{N} \tau_j \exp(\lambda_j h),$$

where $\tau_j = \pi \mathbb{D} \mathbb{G}_j \mathbf{x}$. It is obvious that this function satisfies the triangular condition (1.1) iff the given sum reduces to the single term, e.g. $\tau_j = 0$ for all $j > 1$.

The following corollary is an immediate consequence of Theorems 2.1 and 3.1.

COROLLARY 3.2. *Let* $\{X_n\}$ *be a stationary HMC with mean zero and with transition probability matrix* $\mathbb{P}$ *satisfying the conditions of Theorem* 2.1 *jointly with conditions* (i) *and* (ii). *Then* $\{X_n\}$ *is a WM process in a discrete time.*

4. TRANSFORMATION OF MARKOV PROCESS TO A MARKOV PROCESS IN THE WIDE SENSE

4.1. Continuous time case. Let $\{X_t\}$ be a continuous time HMP with infinitesimal operator $\mathbb{Q}$ and let A and B mean the following conditions:

A. The infinitesimal operator $\mathbb{Q}$ is diagonalizable and at least one of its right eigenvectors, say $\mathbf{r} = (r_0, r_1, \ldots, r_N)^{\mathsf{T}}$, is such that all its coordinates are distinct, i.e. $r_i \neq r_j$ for $i \neq j$.

B. The eigenvector satisfying the condition A is real, i.e. $\mathbf{r} \in \mathbb{R}^{N+1}$.

Condition B implies that the eigenvalue λ corresponding to $\mathbf{r}$ is also real, i.e. $\lambda \in \mathbb{R}$. For $\mathbf{r}$ satisfying the condition A let us define transformation $g\colon \mathfrak{X} \to \widetilde{\mathfrak{X}} = \{r_0, r_1, \ldots, r_N\}$ as follows: $g(x_i) = r_i$, where r_i is the i-th coordinate of the eigenvector $\mathbf{r}$. Thus we can identify the state space $\widetilde{\mathfrak{X}}$ with coordinates of $\mathbf{r}$. Let us associate with an HMP $\{X_t\}$ with the infinitesimal operator $\mathbb{Q}$ an HMP $\{\widetilde{X}_t\}$ with state space $\widetilde{\mathfrak{X}}$ and the same infinitesimal operator $\mathbb{Q}$.

THEOREM 4.1. *If the infinitesimal operator $\mathbb{Q}$ satisfies the condition* A, *then the HMP $\{\widetilde{X}_t\}$ is a WM process for any initial distribution $\mathbf{p}$ such that $\mathbf{pr} = 0$. If the initial distribution $\mathbf{p}$ is the stationary distribution π for $\{X_t\}$, then $\{\widetilde{X}_t\}$ is WM with the covariance function of the following form*:

$$\gamma(h) := \Gamma(t+h, t; \pi) = e^{\lambda h}\mu_2, \tag{4.1}$$

where $\mu_2 := \mathbf{E}_\pi|X_0|^2 \equiv \sum_{i=0}^{N} \pi_i |x_i|^2$, $\pi = (\pi_0, \pi_1, \ldots, \pi_N)$ and λ is the eigenvalue corresponding to the eigenvector $\mathbf{r}$.

If additionally the condition B *is satisfied, then $\{\widetilde{X}_t\}$ is a real-valued WM process.*

Proof. We shall show that for the process $\{\widetilde{X}_t\}$ we have $\mathbf{E}_\mathbf{p}[\widetilde{X}_t] = 0$ and the triangular property (1.1) holds for any initial distribution $\mathbf{p}$. First observe that from Lemma 2.1 we get the formulas for moments of the process $\{\widetilde{X}_t\}$. Namely,

$$\begin{aligned} \mu(t;\mathbf{p}) &:= \mathbf{E}_\mathbf{p}[\widetilde{X}_t] = \mathbf{p}(t)\mathbf{r}, \quad t \geqslant 0, \\ \mu_2(t;\mathbf{p}) &:= \mathbf{E}_\mathbf{p}|\widetilde{X}_t|^2 = \mathbf{p}(t)\overline{\mathbb{D}}\mathbf{r}, \quad t \geqslant 0, \\ \mu(t+h,t;\mathbf{p}) &:= \mathbf{E}_\mathbf{p}[\widetilde{X}_{t+h}\overline{\widetilde{X}_t}] = \mathbf{p}(t)\overline{\mathbb{D}}\mathbb{P}(h)\mathbf{r}, \quad t \geqslant h \geqslant 0, \end{aligned}$$

where $\mathbb{D} = \operatorname{diag}(\mathbf{r})$ and $\overline{\mathbb{D}}$ is the conjugate of $\mathbb{D}$. Notice that $\mathbf{r} \neq \mathbf{r}_0 = \mathbf{1}$ as we demand that $\mathbf{r}$ has all coordinates distinct. Thus, if λ is the eigenvalue corresponding to $\mathbf{r}$, then $\operatorname{Re}\lambda < 0$. Moreover, by (2.2), properties of $\mathbb{G}_j$'s and the relation (2.3) we get $\mathbb{P}(h)\mathbf{r} = e^{\lambda h}\mathbf{r}$, where λ is the eigenvalue corresponding to $\mathbf{r}$. This in turn gives

$$\mu(t+h,t;\mathbf{p}) = e^{\lambda h}\mu_2(t;\mathbf{p}) \tag{4.2}$$

and

$$\mu(t;\mathbf{p}) = e^{\lambda t}\mu(0;\mathbf{p}) = e^{\lambda t}\mathbf{pr} = 0 \tag{4.3}$$

by the assumption that $\mathbf{pr} = 0$. Hence the normalized autocovariance is of the form $R(u,t) = \mu(u,t;\mathbf{p})\mu_2(t;\mathbf{p})^{-1} = e^{\lambda(u-t)}$, $u \geqslant t$, which implies

$$R(u,s) = e^{\lambda(u-s)} = e^{\lambda(u-t)}e^{\lambda(t-s)} = R(u,t)R(t,s), \quad u \geqslant t \geqslant s.$$

Thus the triangular property holds.

Now observe that $\pi = \mathbf{l}_0^{\mathsf{T}}$ and $\mathbf{r} \neq \mathbf{r}_0$, so $\pi\mathbf{r} = 0$, which gives $\mu(t;\pi) = e^{\lambda t}\mu(0;\pi) = 0$ for all $t \geqslant 0$. Similarly, $\mu(t+h,t;\pi) = e^{\lambda h}\mu_2(t;\pi)$. If $\mathbf{p} = \pi$, then $\{\widetilde{X}_t\}$ is stationary and $\mu_2(t;\pi) = \mu_2(0;\pi) = \mu_2$, which implies

$$\gamma(h) = \mu(t+h,t;\pi) = e^{\lambda h}\mu_2.$$

Thus the first part of the theorem is proved.

If condition B is satisfied, then the state space $\widetilde{\mathfrak{X}}$ consists of real numbers and the eigenvalue λ corresponding to $\mathbf{r}$ must be also real. Consequently, $\Gamma(u,t;\mathbf{p})$ is also real. This completes the proof of the theorem. ■

4.2. Discrete time case. It is clear that Theorem 4.1 holds for a discrete time HMC under the following reformulations of conditions A and B:

A$'$. The one-step transition probability matrix $\mathbb{P}$ of HMC is diagonalizable and at least one of its right eigenvectors, say $\mathbf{r} = (r_0, \ldots, r_N)^{\mathsf{T}}$, is such that $r_i \neq r_j$ for $i \neq j$.

B$'$. The eigenvector satisfying the condition A$'$ is real, i.e. $\mathbf{r} \in \mathbb{R}^{N+1}$.

Analogously to the continuous case, condition B$'$ implies that the eigenvalue κ of $\mathbb{P}$ corresponding to $\mathbf{r}$ is real, i.e. $\kappa \in \mathbb{R}$. Let us associate with an HMP $\{X_n\}$ with state space $\mathfrak{X} = \{x_0, x_1, \ldots, x_N\}$ and transition probability matrix $\mathbb{P}$ an HMP $\{\widetilde{X}_n\}$ with state space $\widetilde{\mathfrak{X}} = \{r_0, r_1, \ldots, r_N\}$ and the same transition probability matrix $\mathbb{P}$. Now, we have the following analogue of Theorem 4.1 for the discrete time case.

THEOREM 4.2. *If the transition probability matrix* $\mathbb{P}$ *satisfies the condition* A$'$, *then the HMC* $\{\widetilde{X}_n\}$ *is a WM process for any initial distribution* $\mathbf{p}$ *such that* $\mathbf{pr} = 0$. *If the initial distribution* $\mathbf{p}$ *is the stationary distribution* π, *then the covariance function of* $\{\widetilde{X}_n\}$ *is of the following form*:

$$\gamma(h) = \kappa^h \mu_2, \qquad h \in \mathbb{N}, \tag{4.1$'$}$$

where $\mu_2 := \mathbf{E}_\pi|X_0|^2 \equiv \sum_{i=0}^N \pi_i |x_i|^2$, $\pi = (\pi_0, \pi_1, \ldots, \pi_N)$ *and* κ *is the eigenvalue corresponding to the eigenvector* $\mathbf{r}$.

If additionally the condition B$'$ *is satisfied*, *then* $\{\widetilde{X}_n\}$ *is a real-valued WM process.*

Proof. The proof carries over from the proof of Theorem 4.1 with the only difference that e^λ should be replaced with eigenvalue κ of $\mathbb{P}$. Then, for example, formulas (4.2) and (4.3) take the following form:

$$\mu(n+h,n;\mathbf{p}) = \kappa^h \mu_2(n;\mathbf{p}) \tag{4.2$'$}$$

and

$$\mu(n;\mathbf{p}) = \kappa^n \mu(0;\mathbf{p}) = \kappa^n \mathbf{pr} = 0. \tag{4.3$'$}$$

Then the normalized autocovariance function is of the form $R(n,m) = \kappa^{n-m}$, $n \geqslant m$. Thus the triangular relation (1.1) obviously holds.

In the stationary case we have $\mu_2(n;\pi) = \mu_2(0,\pi) \equiv \mu_2$. Moreover, $\pi\mathbf{r} = 0$, and hence $\mu(n;\pi) = 0,\ n \geqslant 0$, which gives the formula (4.1′). This completes the proof of the theorem. ■

5. TWO TYPES OF PERIODICITY

In this section we consider a relation between two types of periodicity of HMC. The first one is the periodicity of HMC in the usual sense as defined, for example, in [1], p. 72. The second type is defined by periodicity of the covariance function. We say that the process $\{X_t\}$ is *periodically correlated* with period $T > 0$ if

$$\mathbf{E}[X_t] = \mathbf{E}[X_{t+T}], \qquad t \in \mathcal{T},$$

and

$$\mathbf{E}[X_t\overline{X}_s] = \mathbf{E}[X_{t+T}\overline{X}_{s+T}], \qquad t, s \in \mathcal{T},$$

where T is the smallest number with this property.

Let $\{X_n\}$ be irreducible HMC with finite state space $\mathfrak{X} = \{x_0, x_1, \dots, x_N\}$. Thus it is positive recurrent and its stationary distribution is unique. We shall show that if $\{X_n\}$ is periodic with period T, then it is periodically correlated with some period T' if its initial distribution is of special form, i.e. it is cyclostationary (see Definition 5.1). We can reduce the period of the autocovariance function by choosing a special cyclostationary distribution. Furthermore, we investigate an asymptotic of the autocovariance function of $\{X_n\}$ depending on its initial distribution.

We assume that the HMC $\{X_n\}$ with state space $\mathfrak{X} = \{x_0, x_1, \dots, x_N\}$, where $N < \infty,\ x_r \in \mathbb{C}$, is periodic with period $T > 1$. Then $\mathfrak{X}$ consists of T disjoint subspaces $\mathfrak{X}_0, \mathfrak{X}_1, \dots, \mathfrak{X}_{T-1}$, called *cyclic classes* (see [1]), such that $\mathfrak{X} = \bigcup_{i=0}^{T-1} \mathfrak{X}_i$ and $\sum_{x\in\mathfrak{X}_i} \mathsf{P}(X_n = x \mid X_{n-1} = y) = 1$ for $y \in \mathfrak{X}_{i-1}$, $i = 1, 2, \dots, T$ with $\mathfrak{X}_T = \mathfrak{X}_0$. Let D_i denote the set of indices of elements from the i-th cyclic class and assume, for simplicity, that for each $i = 0, 1, \dots, T-2$ the elements of the cyclic class $\mathfrak{X}_i$ have smaller indices than elements of the class $\mathfrak{X}_{i+1}$. Then the transition probability matrix $\mathbb{P}$ has the following superdiagonal block form:

$$\mathbb{P} = \begin{bmatrix} & \mathbb{S}_{0,1} & & \cdots & \\ & & \mathbb{S}_{1,2} & \cdots & \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ & & \cdots & & \mathbb{S}_{T-2,T-1} \\ \mathbb{S}_{T-1,0} & & \cdots & & \end{bmatrix}, \tag{5.1}$$

where the empty entries are zeros. The matrix $\mathbb{S}_{i,i+1}$ represents the probability transitions in one step from the cyclic class $\mathfrak{X}_i$ to $\mathfrak{X}_{i+1}$, i.e. $\sum_{s\in D_{i+1}} p_{r,s} = 1$ for

$r \in D_i$, where $D_T := D_0$. If $|\mathfrak{X}_i| = N_i$, then the block $\mathbb{S}_{i,i+1}$ has size $N_i \times N_{i+1}$ for $i = 0, 1, \ldots, T-2$ and $\mathbb{S}_{T-1,0}$ has size $N_{T-1} \times N_0$, i.e. the matrix $\mathbb{P}$ is $(N_0, N_1, \ldots, N_{T-1})$-superdiagonal (in the terminology of [6]). Then

$$\mathbb{P}^T = \operatorname{diag}(\mathbb{P}^{(0)}, \mathbb{P}^{(1)}, \ldots, \mathbb{P}^{(T-1)}),$$

where $\mathbb{P}^n = \mathbb{P}^{n-1}\mathbb{P}$, $n \geqslant 0$, and

$$\mathbb{P}^{(i)} := \mathbb{S}_{i,i+1} \ldots \mathbb{S}_{T-1,0}\mathbb{S}_{0,1} \ldots \mathbb{S}_{i-1,i}.$$

For each $i = 0, 1, \ldots, T-1$ the block $\mathbb{P}^{(i)}$ is of size $N_i \times N_i$ and is a transition probability matrix of the irreducible and non-periodic HMC $\{X_n^{(i)} = X_{Tn}\}$ with the state space $\mathfrak{X}_i$. Let $\mathbf{1}_i := (1, 1, \ldots, 1)^{\mathsf{T}} \in \mathbb{R}^{N_i}$, $i = 0, 1, \ldots, T-1$. The following lemma will be used in the proof of the next results. Its assertions can be found in [3], however not stated explicitly, so we show our simple proof of it.

LEMMA 5.1. *Let $\{X_n\}$ be an irreducible, periodic HMC with period T and a finite state space $\mathfrak{X} = \bigcup_{i=0}^{T-1} \mathfrak{X}_i$, where $\mathfrak{X}_i$'s are its cyclic classes. Let the row vectors $\pi^{(i)} \in \mathbb{R}^{N_i}$ be some probability distributions on $\mathfrak{X}_i$, $i = 0, 1, \ldots, T-1$. Then the following statements are equivalent*:

(i) $\pi = T^{-1}(\pi^{(0)}, \pi^{(1)}, \ldots, \pi^{(T-1)})$ *is a stationary distribution of the process* $\{X_n\}$;

(ii) $\pi^{(i)}\mathbb{S}_{i,i+1} = \pi^{(i+1)}$ *for* $i = 0, 1, \ldots, T-2$ *and* $\pi^{(T-1)}\mathbb{S}_{T-1,0} = \pi^{(0)}$;

(iii) $\pi^{(i)}$ *is a stationary distribution of* $\{X_n^{(i)}\}$ *for* $i = 0, 1, \ldots, T-1$.

Proof. The equivalence of (i) and (ii) is obvious. If (i) holds, then $\pi\mathbb{P} = \pi$ and $\pi\mathbb{P}^T = \pi$. Hence for each $i = 0, 1, \ldots, T-1$ the equality $\pi^{(i)}\mathbb{P}^{(i)} = \pi^{(i)}$ holds, which gives (iii). Suppose now that (iii) holds and $\mathbf{p}$ is the stationary distribution of $\{X_n\}$. If we divide the vector $\mathbf{p}$ into T subvectors of length N_i, i.e. $\mathbf{p} = (\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(T-1)})$, such that $\mathbf{p}^{(i)} \in \mathbb{R}^{N_i}$, then also $\mathbf{p}^{(i)}\mathbb{P}^{(i)} = \mathbf{p}^{(i)}$ for each i. However, by uniqueness of the stationary distribution of the irreducible and positive recurrent HMC, applied to the process $\{X_n^{(i)}\}$, we get $\mathbf{p}^{(i)} = q_i\pi^{(i)}$ for some constant q_i. Moreover, $\mathbf{p}^{(i)}\mathbb{S}_{i,i+1} = \mathbf{p}^{(i+1)}$ and $q_i\pi^{(i)}\mathbb{S}_{i,i+1} = q_{i+1}\pi^{(i+1)}$. Therefore, postmultiplying each side of this equation by $\mathbf{1}_{i+1}$ we get $q_i = q_{i+1}$, $i = 0, 1, \ldots, T-2$, which implies $q_i = T^{-1}$. This shows that $\mathbf{p} = \pi$, which completes the proof of the lemma. ∎

DEFINITION 5.1. Let $\{X_n\}$ be an irreducible periodic HMC with period $T > 1$, state space $\mathfrak{X} = \bigcup_{i=0}^{T-1} \mathfrak{X}_i$, where $\mathfrak{X}_i$'s are its cyclic classes, and with the transition probability matrix $\mathbb{P}$ of the form (5.1). We say that a probability distribution $\mathbf{p}$ on the state space $\mathfrak{X}$ is a *cyclostationary distribution* for $\{X_n\}$ if

$$\mathbf{p}\mathbb{P}^T = \mathbf{p}.$$

Let $\pi^{(i)}$ be the stationary distribution of $\{X_n^{(i)}\}$, $i = 0, 1, \dots, T-1$, and define the mapping $\pi\colon \mathbb{C}^T \to \mathbb{C}^{N+1}$ in the following way:

If $\mathbf{a} = (a_0, a_1, \dots, a_{T-1}) \in \mathbb{C}^T$, then

$$\pi(\mathbf{a}) \equiv \pi(a_0, a_1, \dots, a_{T-1}) := (a_0\pi^{(0)}, a_1\pi^{(1)}, \dots, a_{T-1}\pi^{(T-1)}).$$

In the next two lemmas we give a characterization of cyclostationary distributions together with some important properties. Before this, assume the following notation. Let $\langle m\rangle := m \bmod T$ for the integer m. For vector $\mathbf{a} = (a_0, \dots, a_{T-1})$ define $\mathbf{a}_n := (a_{T-n}, \dots, a_{T-1}, a_0, \dots, a_{T-n-1})$ for $0 \leqslant n < T$, and then $\mathbf{a}_n := \mathbf{a}_{\langle n\rangle}$ for all $n \in \mathbb{Z}$.

LEMMA 5.2. *Let $\{X_n\}$ be an HMC which is irreducible, periodic with period $T > 1$, with finite state space $\mathfrak{X} = \bigcup_{i=0}^{T-1} \mathfrak{X}_i$, where $\mathfrak{X}_i$'s are its cyclic classes, and with the transition probability matrix $\mathbb{P}$ of the form* (5.1). *Then, for a probability distribution $\mathbf{p}$ on $\mathfrak{X}$, the following statements are equivalent*:

(i) *$\mathbf{p}$ is a cyclostationary distribution of $\{X_n\}$*;

(ii) *$\mathbf{p} = \pi(q_0, \dots, q_{T-1})$ for some probability vector $\mathbf{q} = (q_0, \dots, q_{T-1})$*;

(iii) *$\mathbf{p}\mathbb{P}^n$ is cyclostationary for all $n \geqslant 0$.*

If we additionally assume that $\mathbb{P}$ is of full rank, then (i), (ii) *and* (iii) *are equivalent to the following condition*:

(iv) *$\mathbf{p}\mathbb{P}^n$ is cyclostationary for some $n \in \mathbb{N}$.*

Proof. First we show the equivalence (i) $\Leftrightarrow$ (ii). Suppose (i) holds and divide the vector $\mathbf{p}$ into T subvectors of length N_i, i.e. $\mathbf{p} = (\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \dots, \mathbf{p}^{(T-1)})$. Then directly from Definition 5.1 we see that $\mathbf{p}^{(i)}\mathbb{P}^{(i)} = \mathbf{p}^{(i)}$ for each i. This implies $\mathbf{p}^{(i)} = q_i\pi^{(i)}$ for some vector $(q_0, q_1, \dots, q_{T-1})$, which must be the probability distribution. The converse implication is obvious.

The implication (ii) $\Rightarrow$ (iii) is a consequence of Lemma 5.1 (ii), which yields

$$\pi(\mathbf{q})\mathbb{P}^n = \pi(\mathbf{q}_n). \tag{5.2}$$

The converse implication is trivial.

Now we shall prove the equivalence (iv) $\Leftrightarrow$ (ii) under the assumption that $\mathbb{P}$ is full. Suppose that (iv) holds, i.e. $\mathbf{p}\mathbb{P}^n = \pi(\mathbf{q}_n)$ for some probability vector $\mathbf{q}$. This in view of (5.2) implies that there exists a row vector $\mathbf{e} = (e_1, e_2, \dots, e_N)$, $\sum e_j = 0$, for which $\mathbf{p} = \pi(\mathbf{q}) + \mathbf{e}$, and then

$$\mathbf{p}\mathbb{P}^n = \pi(\mathbf{q}_n) + \mathbf{e}\mathbb{P}^n.$$

Thus $\mathbf{e}\mathbb{P}^n = \mathbf{0}$, which is possible only for $\mathbf{e} = \mathbf{0}$ if $\mathbb{P}$ is full. Hence (iv) implies (ii).

The condition (ii) implies (iii) and the implication (iii) $\Rightarrow$ (iv) is trivial. This completes the proof of the lemma. ∎

LEMMA 5.3. *Let the conditions of Lemma* 5.2 *hold. Then for all initial distributions* $\mathbf{p}$ *of* $\{X_n\}$ *and for all integers* $0 \leqslant h < T$ *the following convergence holds*:

$$\lim_{n\to\infty} \mathbf{p}\mathbb{P}^{nT+h} = \pi(q_{T-h}\dots q_{T-1}, q_0, \dots, q_{T-h-1}) \equiv \pi(q_h), \tag{5.3}$$

where $q_i = \mathsf{P}(X_0 \in \mathfrak{X}_i)$. *If additionally* $\mathsf{P}(X_0 \in \mathfrak{X}_i) = T^{-1}$ *holds true for all* $i = 0, 1, \dots, T-1$, *then*

$$\lim_{n\to\infty} \mathbf{p}\mathbb{P}^{n} = \pi(T^{-1}, T^{-1}, \dots, T^{-1}) = \pi.$$

P r o o f. It is enough to show (5.3) for $h = 0$; then the general form for $h \geqslant 1$ comes from (5.2). Observe that $q_i = \mathsf{P}(X_0 \in \mathfrak{X}_i) = \sum_{j\in D_i} p_j$. Then $\mathbf{p} = (q_0\widetilde{\mathbf{p}}^{(0)}, q_1\widetilde{\mathbf{p}}^{(1)}, \dots, q_{T-1}\widetilde{\mathbf{p}}^{(T-1)})$ where for each $i = 0, 1, \dots, T-1$ the vector $\widetilde{\mathbf{p}}^{(i)}$ is a probability distribution on the cyclic class $\mathfrak{X}_i$. Now, observe that

$$\begin{aligned}\mathbf{p}\mathbb{P}^{nT} &= \big(q_0\widetilde{\mathbf{p}}^{(0)}(\mathbb{P}^{(0)})^n, q_1\widetilde{\mathbf{p}}^{(1)}(\mathbb{P}^{(1)})^n, \dots, q_{T-1}\widetilde{\mathbf{p}}^{(T-1)}(\mathbb{P}^{(T-1)})^n\big)\\ &\to (q_0\pi^{(0)}, q_1\pi^{(1)}, \dots, q_{T-1}\pi^{(T-1)}), \quad n\to\infty,\end{aligned}$$

which completes the proof of the lemma. ■

THEOREM 5.1. *Let* $\{X_n\}$ *be an irreducible and periodic HMC with period* $T > 1$ *and with a transition probability matrix* $\mathbb{P}$ *of the form* (5.1). *If an initial distribution* $\mathbf{p}$ *is cyclostationary for* $\{X_n\}$, *then* $\{X_n\}$ *is periodically correlated with period* T' *which is a divisor of* T. *If* $\mathbf{p} = \pi(q_0, q_1, \dots, q_{T-1})$, *then* T' *is a period of a sequence* $\{q_0, q_1, \dots\}$, *where we define* $q_m := q_{\langle m\rangle}$ *for* $m \geqslant T$.

P r o o f. Let $B(n, h; \mathbf{q}) := \mathrm{Cov}\big(X_{n+h}, X_n; \pi(\mathbf{q})\big)$, where $n, h \geqslant 0$. Using the formulas for moments of HMC given in Lemma 2.2 we get the following

$$\begin{aligned}B(n, h; \mathbf{q}) &= \mu\big(n+h, n; \pi(\mathbf{q})\big) - \mu\big(n+h; \pi(\mathbf{q})\big)\overline{\mu\big(n; \pi(\mathbf{q})\big)}\\ &= \pi(\mathbf{q})\mathbb{P}^n\overline{\mathbb{D}}\big(\mathbb{I} - \mathbf{1}\pi(\mathbf{q})\mathbb{P}^n\big)\mathbb{P}^h\mathbf{x}.\end{aligned}$$

Thus in view of (5.2) we get

$$B(n, h; \mathbf{q}) = \pi(\mathbf{q}_n)\overline{\mathbb{D}}\big(\mathbb{I} - \mathbf{1}\pi(\mathbf{q}_n)\big)\mathbb{P}^h\mathbf{x},$$

where $\mathbf{q}_n := (q_{T-n}\dots q_{T-1}, q_0, \dots, q_{T-n-1})$ for $0 \leqslant n < T$ and $\mathbf{q}_n := \mathbf{q}_{\langle n\rangle}$ for $n > T$. Thus the function $B(n, h; \mathbf{q})$ is periodic in n with the period T' of the sequence $\{q_0, q_1, \dots\}$. Obviously, T' is a divisor of T. ■

Let $\{X_n^{\circ}\}$ be an irreducible and stationary HMC with transition probability matrix $\mathbb{P}$, and $\big\{X_n^{\circ(k)}\big\}$ be a stationary HMC with transition probability matrix $\mathbb{P}^{(k)}$. It means that their initial distributions are stationary distributions π and $\pi^{(k)}$, respectively.

LEMMA 5.4. *Let $\{X_n\}$ be an irreducible and periodic HMC with period $T > 1$. Then for any $x \in \mathfrak{X}_k$ and $h \geqslant 0$ the following convergence holds*:

$$\begin{aligned}&\mathrm{Cov}(X_{nT+h}, X_{nT}|X_0 = x)\\ &\qquad \to \sum_{i\in D_k} (\mathbf{E}[X_h\overline{X}_0|X_0 = x_i] - \overline{\mu^{(k)}}\mathbf{E}[X_h|X_0 = x_i])\pi_i^{(k)}, \quad n \to \infty,\end{aligned}$$

where $\mu^{(k)} := \mathbf{E}[X_0^{\circ(k)}] = \sum_{i\in D_k} x_i\pi_i^{(k)}$.

P r o o f. Notice that for any $n \geqslant 1$ and $h \geqslant 0$ the following relations hold:

$$\mathbf{E}[X_{nT+h}\overline{X}_{nT}|X_0 = x] = \sum_{i\in D_k} \mathbf{E}[X_h\overline{X}_0|X_0 = x_i]\mathsf{P}(X_{nT} = x_i|X_0 = x),$$

$$\mathbf{E}[X_{nT+h}|X_0 = x] = \sum_{i\in D_k} \mathbf{E}[X_h|X_0 = x_i]\mathsf{P}(X_{nT} = x_i|X_0 = x), \quad h \geqslant 0.$$

But

$$\mathsf{P}(X_{nT} = x_i|X_0 = x) \to \pi_i^{(k)} \quad \text{for } i \in D_k \text{ and } x \in \mathfrak{X}_k.$$

Hence, by the definition of the covariance

$$\begin{aligned}&\mathrm{Cov}(X_{nT+h}, X_{nT}|X_0 = x)\\ &\qquad = \mathbf{E}[X_{nT+h}\overline{X}_{nT}|X_0 = x] - \mathbf{E}[X_{nT+h}|X_0 = x]\mathbf{E}[\overline{X}_{nT}|X_0 = x]\end{aligned}$$

and by the above equalities, we get the assertion of the lemma. ∎

LEMMA 5.5. *Let $\{X_n\}$ be an irreducible and periodic HMC with period $T > 1$ and with transition probability matrix $\mathbb{P}$ given by* (5.1). *Then for any non-negative integer h and $n \to \infty$ the following convergences hold*:

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{E}[X_{k+h}\overline{X}_k] \to \mathbf{E}[X_0^\circ\overline{X}_h^\circ], \tag{5.4}$$

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{E}[X_k] \to \mathbf{E}[X_0^\circ], \tag{5.5}$$

$$\frac{1}{T}\sum_{i=0}^{T-1} \mathbf{E}[X_{n+i}] \to \mathbf{E}[X_0^\circ]. \tag{5.6}$$

If the initial distribution $\mathbf{p} = (p_0, p_1, \dots, p_N)$ is such that $q_k := \sum_{i\in D_k} p_i = T^{-1}$ for all $k = 0, 1, \dots, T-1$, then

$$\mathbf{E}[X_n] \to \mathbf{E}[X_0^\circ]. \tag{5.7}$$

P r o o f. Notice that

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}[X_k\overline{X}_{k+h}] &= \sum_i \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}[X_k\overline{X}_{k+h}|X_k = x_i]\mathsf{P}(X_k = x_i)\\ &= \sum_i \mathbf{E}[X_0\overline{X}_h|X_0 = x_i]\frac{1}{n}\sum_{k=1}^{n}\mathsf{P}(X_k = x_i)\\ &\to \sum_i \mathbf{E}[X_0\overline{X}_h|X_0 = x_i]\pi_i = \mathbf{E}[X_0^\circ\overline{X}_h^\circ].\end{aligned}$$

This proves the convergence (5.4). The proof of the second convergence runs over in a similar way to the above.

To prove the convergence (5.6) notice first that

$$\frac{1}{T}\sum_{i=0}^{T-1}\mathbf{E}[X_{n+i}] = \Big(\frac{1}{T}\sum_{i=0}^{T-1}\mathbf{p}(n+i)\Big)\mathbf{x} = \Big(\frac{1}{T}\sum_{i=0}^{T-1}\mathbf{p}\mathbb{P}^{n+i}\Big)\mathbf{x},$$

where $\mathbf{p}$ is the initial distribution of $\{X_n\}$. Notice that

$$\sum_{i=0}^{T-1}\mathbf{p}(n+i) = \sum_{i=0}^{T-1}\mathbf{p}(m_nT+i)$$

for some positive integer m_n. By (5.3) we have the following convergence:

$$\lim_{n\to\infty}\frac{1}{T}\sum_{i=0}^{T-1}\mathbf{p}\mathbb{P}^{m_nT+i} = \frac{1}{T}\sum_{i=0}^{T-1}\pi(\mathbf{q}_{\langle i\rangle}),$$

where $\mathbf{q} = (q_0, q_1, \ldots, q_{T-1})$, and $q_i = \mathsf{P}(X_0 \in \mathfrak{X}_i)$. Moreover,

$$\frac{1}{T}\sum_{i=0}^{T-1}\pi(\mathbf{q}_{\langle i\rangle}) = \frac{1}{T}\pi\big(\sum_{i=0}^{T-1}\mathbf{q}_{\langle i\rangle}\big) = \frac{1}{T}\pi\big(1, 1, \ldots, 1\big) = \pi.$$

This implies

$$\Big(\frac{1}{T}\sum_{i=0}^{T-1}\mathbf{p}(n+i)\Big)\mathbf{x} \to \pi\mathbf{x} = \mathbf{E}[X_0^\circ],$$

which proves the convergence (5.6).

The convergence (5.7) comes from (5.3), where for each $h = \langle h\rangle$ we have

$$\mathbf{E}[X_{mT+h}] = \mathbf{p}\mathbb{P}^{mT+h}\mathbf{x} \to \frac{1}{T}\pi(1, 1, \ldots, 1)\mathbf{x} = \pi\mathbf{x} = \mathbf{E}[X_0^\circ],$$

by the assumption that $q_0 = q_1 = \ldots = q_{T-1} = 1/T$. This completes the proof of the theorem. ∎

THEOREM 5.2. *Let $\{X_n\}$ be an HMC, irreducible, recurrent, periodic with period $T > 1$ and transition probability matrix $\mathbb{P}$ given by* (5.1). *Then for any nonnegative integer h the following convergence holds:*

$$\frac{1}{n}\sum_{k=0}^{n}\operatorname{Cov}\Big(X_{k+h},\frac{1}{T}\sum_{i=0}^{T-1}X_{k-T+i}\Big)\to\frac{1}{T}\sum_{i=0}^{T-1}\gamma(h+T-i), \tag{5.8}$$

where

$$\gamma(h)=\operatorname{Cov}(X_0^\circ,X_h^\circ)=\mathbf{E}[X_0^\circ\overline{X}_h^\circ]-|\mathbf{E}[X_0^\circ]|^2 \tag{5.9}$$

and $\{X_n^\circ\}$ is the stationary process corresponding to $\{X_n\}$.

If the initial distribution $\mathbf{p}=(p_0,p_1,\ldots,p_N)$ is such that $q_k:=\sum_{i\in D_k}p_i=T^{-1}$ for all $k=0,1,\ldots,T-1$, then

$$\frac{1}{n}\sum_{k=0}^{n}\operatorname{Cov}(X_k,X_{k+h})\to\gamma(h), \tag{5.10}$$

where $\gamma(h)$ is given by (5.9).

P r o o f. It is well known that the convergences

$$x_n\to x\quad\text{and}\quad\frac{1}{n}\sum_{k=1}^{n}y_k\to y$$

imply the convergence

$$\frac{1}{n}\sum_{k=1}^{n}x_ky_k\to xy.$$

Hence, using the form $\operatorname{Cov}(X_k,X_l)=\mathbf{E}[X_k\overline{X}_l]-\mathbf{E}[X_k]\mathbf{E}[\overline{X}_l]$ and convergences (5.4)–(5.6) in Lemma 5.5, we get (5.8). The convergence (5.10) follows from (5.7) in Lemma 5.5. This completes the proof of the theorem. ■

REFERENCES

[1] P. Brémaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulations and Queues*, Springer, New York 1999.
[2] J. L. Doob, *Stochastic Processes*, Wiley, New York 1953.
[3] M. Iosifescu, *Finite Markov Processes and Their Applications*, Wiley, New York 1980.
[4] C. D. Lai, *First order autoregressive Markov processes*, Stochastic Process. Appl. 7 (1978), pp. 65–72.
[5] P. Lancaster, *Theory of Matrices*, Academic Press, New York 1969.
[6] H. Minc, *Nonnegative Matrices*, Wiley, New York 1988.
[7] J. F. Reynolds, *On the autocorrelation and spectral functions of queues*, J. Appl. Probab. 5 (1968), pp. 467–475.

[8] J. F. Reynolds, *Some theorems on the transient covariance of Markov chains*, J. Appl. Probab. 9 (1972), pp. 214–218.

[9] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer, New York 1981.

Institute of Mathematics
Wrocław University
pl. Grunwaldzki 2/4
50-384 Wrocław, Poland
E-mail: kasprzyk@math.uni.wroc.pl

Institute of Mathematics
Wrocław University
pl. Grunwaldzki 2/4
50-384 Wrocław, Poland
E-mail: szczotka@math.uni.wroc.pl

Received on 13.12.2006;
revised version on 22.4.2007