J. Appl. Prob. **53,** 953–956 (2016)
*doi:*10.1017/jpr.2016.54

AN EXTENSION OF A CONVERGENCE THEOREM FOR MARKOV CHAINS ARISING IN POPULATION GENETICS

MARTIN MÖHLE,* *Eberhard Karls Universität Tübingen*
MORIHIRO NOTOHARA,** *Nagoya City University*

Abstract

An extension of a convergence theorem for sequences of Markov chains is derived. For every positive integer N let $(X_N(r))_r$ be a Markov chain with the same finite state space S and transition matrix $\Pi_N = I + d_N Q + c_N B_N$, where I is the unit matrix, Q a generator matrix, $(B_N)_N$ a sequence of matrices, $\lim_{N\to\infty} c_N = \lim_{N\to\infty} d_N = 0$ and $\lim_{N\to\infty} c_N/d_N = 0$. Suppose that the limits $P := \lim_{m\to\infty}(I + d_N Q)^m$ and $G := \lim_{N\to\infty} P B_N P$ exist. If the sequence of initial distributions $P_{X_N(0)}$ converges weakly to some probability measure μ, then the finite-dimensional distributions of $(X_N([t/c_N]))_{t\geq 0}$ converge to those of the Markov process $(X_t)_{t\geq 0}$ with initial distribution μ, transition matrix $P\mathrm{e}^{tG}$ and $\lim_{N\to\infty}(I + d_N Q + c_N B_N)^{[t/c_N]} = P - I + \mathrm{e}^{tG} = P\mathrm{e}^{tG} = \mathrm{e}^{tG}P$ for all $t > 0$.

Keywords: Convergence; Markov chain; population genetics; separation of time-scales

2010 Mathematics Subject Classification: Primary 60F05; 92D10
Secondary 60J27; 92D25

1. Introduction and main result

A convergence theorem proved in [4] has been used in population genetics for various models including partial selfing, two-sex populations, strong migration, age-structure, and so on, and it provides many useful results; see [2]–[8]. This theorem is a generalization of the well known matrix equation $\lim_{N\to\infty}(I + B/N)^N = \mathrm{e}^B$. Let $d \in \mathbb{N} := \{1, 2, \ldots\}$, $A = (a_{ij})$ be a $d \times d$ matrix satisfying $\|A\| := \max_i \sum_j |a_{ij}| = 1$, and suppose that $P := \lim_{m\to\infty} A^m$ exists. Let $t, K \geq 0$, and $(c_N)_{N\in\mathbb{N}}$ be a sequence of positive real numbers satisfying $\lim_{N\to\infty} c_N = 0$. Then (see [4]) $\lim_{N\to\infty} \sup_{\|B\|\leq K} \|(A + c_N B)^{[t/c_N]} - (P + c_N B)^{[t/c_N]}\| = 0$. If $(B_N)_{N\in\mathbb{N}}$ is a sequence of $d \times d$ matrices such that $G = \lim_{N\to\infty} P B_N B$ exists, then $\lim_{N\to\infty}(A + c_N B_N)^{[t/c_N]} = P - I + \mathrm{e}^{tG} = P\mathrm{e}^{tG} = \mathrm{e}^{tG}P$ for all $t > 0$.

Set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and for every $N \in \mathbb{N}$ let $(X_N(r))_{r\in\mathbb{N}_0}$ be a time-homogeneous Markov chain with the same finite state space S and transition matrix $\Pi_N := A + c_N B_N$. If the sequence of initial probability measures $P_{X_N(0)}$ converge weakly to some probability measure μ, then the finite-dimensional distributions of the process $(X_N([t/c_N]))_{t\geq 0}$ converge to those of a continuous-time Markov process $(X_t)_{t\geq 0}$ with initial distribution $P_{X_0} = \mu$ and transition matrix $\Pi(t) = P - I + \mathrm{e}^{tG} = P\mathrm{e}^{tG} = \mathrm{e}^{tG}P$, $t > 0$. The limiting process jumps instantaneously from a state $i \in S$ at time $t = 0$ to a state $j \in S$ at time $t = 0+$ with probability p_{ij}, where p_{ij}

Received 2 September 2015; revision received 7 September 2015.
* Postal address: Mathematisches Institut, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany. Email address: martin.moehle@uni-tuebingen.de
** Postal address: Graduate School of Natural Sciences, Nagoya City University, Mizuho, Nagoya, 467-8501, Japan. Email address: noto@nsc.nagoya-cu.ac.jp

is the (i, j)-entry of the matrix P describing the instantaneous jumps. After that, it is described by a Markov process with infinitesimal generator G. The limiting process reflects two time-scale phenomena, where one occurs on a fast time scale and the other on a slow time scale. This theorem can be extended to more general time-scaling phenomena by a slight modification. The purpose of this paper is to prove Theorem 1 below. We call a matrix $Q = (q_{ij})$ a generator matrix if $q_{ij} \geq 0$ for all $i \neq j$ and $\sum_j q_{ij} = 0$ for all i. In Theorem 1 below, the matrix A in [4] is replaced by a matrix A_N depending on N of the form $A_N := I + d_N Q$, where $(d_N)_{N\in\mathbb{N}}$ is a sequence of positive real numbers tending to 0 slower than $(c_N)_{N\in\mathbb{N}}$ and Q is a generator matrix.

Theorem 1. *Let $(c_N)_{N\in\mathbb{N}}$ and $(d_N)_{N\in\mathbb{N}}$ be two sequences of positive real numbers such that $\lim_{N\to\infty} c_N = \lim_{N\to\infty} d_N = 0$ and $\lim_{N\to\infty} c_N/d_N = 0$. Furthermore, let $Q = (q_{ij})$ be a $d \times d$ generator matrix and suppose that $P := \lim_{m\to\infty}(I + d_N Q)^m$ exists for every $N \in \mathbb{N}$. If $(B_N)_{N\in\mathbb{N}}$ is a $d \times d$ matrix sequence such that $G := \lim_{N\to\infty} PB_N P$ exists, then $\lim_{N\to\infty}(I + d_N Q + c_N B_N)^{[t/c_N]} = P - I + \mathrm{e}^{tG} = P\mathrm{e}^{tG} = \mathrm{e}^{tG}P$ for all $t > 0$.*

For every $N \in \mathbb{N}$, let $(X_N(r))_{r\in\mathbb{N}_0}$ be a Markov chain with the same finite state space S and transition matrix $\Pi_N := I + d_N Q + c_N B_N$. If the sequence of initial distributions $P_{X_N(0)}$ converges weakly to some probability measure μ, then the finite-dimensional distributions of the process $(X_N([t/c_N]))_{t\geq 0}$ converge to those of a continuous-time Markov process $(X_t)_{t\geq 0}$ with initial distribution $P_{X_0} = \mu$ and transition matrix $\Pi(t) = P - I + \mathrm{e}^{tG} = P\mathrm{e}^{tG} = \mathrm{e}^{tG}P$, $t > 0$.

Note that the limit $P := \lim_{m\to\infty}(I + d_N Q)^m$ does not depend on N (see Lemma 2 below). A special case of Theorem 1 has already been used by Nordborg and Krone [6], but its proof needs some attention.

2. Proof

The proof of Theorem 1 is based on the following three lemmas.

Lemma 1. *If A is a $d\times d$ matrix such that $P := \lim_{m\to\infty} A^m$ exists, then $\lim_{t\to\infty} \mathrm{e}^{t(A-I)} = P$.*

Proof. Fix $\varepsilon > 0$. Choose $m_0 = m_0(\varepsilon) \in \mathbb{N}$ such that $\|A^m - P\| < \varepsilon$ for all $m > m_0$. We have $\mathrm{e}^{t(A-I)} - P = \mathrm{e}^{-t}\mathrm{e}^{tA} - P = \mathrm{e}^{-t}\sum_{m=0}^{\infty}(t^m/m!)(A^m - P)$. Thus,

$$\begin{aligned}\|\mathrm{e}^{t(A-I)} - P\| &\leq \mathrm{e}^{-t}\sum_{m=0}^{m_0}\frac{t^m}{m!}\|A^m - P\| + \mathrm{e}^{-t}\sum_{m=m_0+1}^{\infty}\frac{t^m}{m!}\underbrace{\|A^m - P\|}_{\leq\varepsilon}\\ &\leq \Big(\sup_{0\leq m\leq m_0}\|A^m - P\|\Big)\mathrm{e}^{-t}\sum_{m=0}^{m_0}\frac{t^m}{m!} + \varepsilon\\ &\to \varepsilon \quad \text{as } t\to\infty.\end{aligned}$$

The assertion follows since $\varepsilon > 0$ can be chosen arbitrarily small. □

Lemma 2. (i) Let $0 < \delta < \infty$ and suppose that Q is a $d\times d$ matrix such that $P := \lim_{m\to\infty}(I + \delta Q)^m$ exists. Then $\lim_{t\to\infty} \mathrm{e}^{tQ} = P$.

(ii) If $0 < \delta_1, \delta_2 < \infty$, and if Q is a $d \times d$ matrix such that the two limits $P_1 := \lim_{m\to\infty}(I + \delta_1 Q)^m$ and $P_2 := \lim_{m\to\infty}(I + \delta_2 Q)^m$ exist, then $P_1 = P_2$.

https://doi.org/10.1017/jpr.2016.54 Published online by Cambridge University Press

Proof. Lemma 1, applied with $A := I + \delta Q$, yields $\lim_{t\to\infty} \mathrm{e}^{t\delta Q} = P$. Part (i) follows since $t \to \infty$ is equivalent to $t\delta \to \infty$. Part (ii) follows from (i) via $P_1 = \lim_{t\to\infty} \mathrm{e}^{tQ} = P_2$, completing the proof. □

Lemma 3. *Under the assumptions of Theorem 1, there exists a sequence of integers* $(M_N)_{N\in\mathbb{N}}$ *such that* $\lim_{N\to\infty} M_N c_N = 0$ *and* $\lim_{N\to\infty} \sup_{m\geq M_N} \|(I + d_N Q)^m - P\| = 0$.

Proof. Choose a sequence $(M_N)_{N\in\mathbb{N}}$ of positive integers such that $\lim_{N\to\infty} M_N d_N = \infty$, $\lim_{N\to\infty} M_N c_N = 0$, and $\lim_{N\to\infty} M_N d_N^2 = 0$; for example, one may choose $M_N := 1 + [\min((c_N d_N)^{-1/2}, (d_N^3)^{-1/2})]$ for all $N \in \mathbb{N}$. Set $A_N := I + d_N Q$. Since Q is a generator matrix and since $\lim_{N\to\infty} d_N = 0$, we conclude that $\|A_N\| = 1$ for all $N > N_0$ for some suitable $N_0 \in \mathbb{N}$. Without loss of generality assume that $N > N_0$ in the following. From

$$\|A_N^{m+1} - P\| = \|(A_N^m - P)A_N\| \ \leq\ \|A_N^m - P\|\,\|A_N\| \leq \|A_N^m - P\|,$$

it follows that the map $m \mapsto \|A_N^m - P\|$ is nonincreasing in m. Thus, it suffices to verify that $\lim_{N\to\infty} \|A_N^{M_N} - P\| = 0$. We have

$$\|A_N^{M_N} - P\| \leq \|A_N^{M_N} - \mathrm{e}^{M_N d_N Q}\| + \|\mathrm{e}^{M_N d_N Q} - P\|.$$

The last norm converges to 0 as $N \to \infty$ by Lemma 2, since $\lim_{N\to\infty} M_N d_N = \infty$. Moreover, since $\|A_N\| = 1$ and $\|\mathrm{e}^{d_N Q}\| = 1$, we have

$$\begin{aligned}
\|A_N^{M_N} - (\mathrm{e}^{d_N Q})^{M_N}\| &\leq M_N \|A_N - \mathrm{e}^{d_N Q}\| \\
&= M_N \left\| \sum_{m=2}^{\infty} \frac{(d_N Q)^m}{m!} \right\| \\
&= M_N O(d_N^2) \\
&\to 0 \quad \text{as } N \to \infty.
\end{aligned}$$

□

Proof of Theorem 1. The proof is given by a one point improvement of the proof provided in [4]. In [4, Appendix] it is shown that for fixed $t \geq 0$ and $\varepsilon > 0$, if we choose $M \in \mathbb{N}$ such that $\|A^m - P\| < \varepsilon$ for all $m \geq M$, and set $n := [t/c_N]$, then

$$\|(A + c_N B)^n - (P + c_N B)^n\| \leq \|A^n - P\| + S_1 + S_2,$$

where $S_1 \sim \varepsilon \mathrm{e}^t (t + 1)$ and $S_2 \sim 2M c_N \mathrm{e}^t (t + 2)$ as $N \to \infty$. Since $M c_N$ tends to 0 as $N \to \infty$, we conclude that $\lim_{N\to\infty} \sup_{\|B\|\leq K} \|(A + c_N B)^n - (P + c_N B)^n\| = 0$.

In our situation, the matrix A in [4] is replaced by $A_N := I + d_N Q$. We should say that, by Lemma 3, there exists $N_\varepsilon \in \mathbb{N}$ and a sequence of positive integers $(M_N)_{N\in\mathbb{N}}$ such that $\lim_{N\to\infty} M_N c_N = 0$ and $\|A_N^m - P\| < \varepsilon$ for all $N > N_\varepsilon$ and all $m \geq M_N$. It follows that $\lim_{N\to\infty} \sup_{\|B\|\leq K} \|(A_N + c_N B)^{[t/c_N]} - (P + c_N B)^{[t/c_N]}\| = 0$. The rest of the proof is the same as in [4]. □

Example 1. (i) We have $\Pi_N = I + Q/N^\alpha + B_N/N$; that is, $c_N = 1/N$ and $d_N = 1/N^\alpha$ with $0 < \alpha < 1$. Then we can choose $M_N := 1 + [N^\alpha \log N]$ for all $N \in \mathbb{N}$.

(ii) $\Pi_N = I + Q/\log N + B_N/N$; that is, $c_N = 1/N$ and $d_N = 1/\log N$. Then we can choose $M_1 := 1$ and $M_N := 1 + [(\log N)(\log\log N)]$ for all $N \in \mathbb{N}$ with $N \geq 2$.

Example 1(i) was used by Nordborg and Krone [6], where the matrix Q comprised fast migration occurring at a rate proportional to $1/N^\alpha$ and B_N corresponded to coalescent events

https://doi.org/10.1017/jpr.2016.54 Published online by Cambridge University Press

which occur at a rate proportional to $1/N$. Example 1(ii) is more artificial, where a fast phenomenon (for example, migration) occurs at a rate proportional to $1/\log N$.

Remark 1. (*Extension to the countable infinite case.*) Let ℓ^∞ be the Banach space of all $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{R}^\mathbb{N}$ with $\|x\| := \sup_{i\in\mathbb{N}} |x_i| < \infty$. Let $A = (a_{ij})_{i,j\in\mathbb{N}}$ be the linear operator from ℓ^∞ to ℓ^∞ defined by $(Ax)_i := \sum_{j\in\mathbb{N}} a_{ij}x_j$ for $x = (x_i)_{i\in\mathbb{N}} \in \ell^\infty$, where we assume that $\|A\| := \sup_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} |a_{ij}| < \infty$. The set L of all such linear operators A is a complete metric space with $\|AB\| \le \|A\|\,\|B\|$ for all $A, B \in L$. For $A \in L$, the exponential can be defined by $\mathrm{e}^A := \sum_{n=0}^\infty A^n/n!$ as in the finite-dimensional case. All matrices in [6] can be reconsidered as linear operators from ℓ^∞ to ℓ^∞. If their norms are finite then all lemmas and theorems in this paper are still valid.

Remark 2. (*Convergence of the semigroups.*) For $t \ge 0$ and $N \in \mathbb{N}$, define the linear operators S_N, $T_t^{(N)}$, and T_t via $S_N f(i) := \mathbb{E}(f(X_N(r+1)) \mid X_N(r) = i) = \sum_{j\in S}(\Pi_N)_{ij} f(j)$, $T_t^{(N)} := S_N^{[t/c_N]}$, and $T_t f(i) := \sum_{j\in S}(P\mathrm{e}^{tG})_{ij} f(j)$. Note that $(T_t^{(N)})_{t\ge0}$ and $(T_t)_{t\ge0}$ are the semigroups of the processes $(X_N([t/c_N]))_{t\ge0}$ and $(X_t)_{t\ge0}$, respectively. Under the conditions of Theorem 1, it follows that, for all $t > 0$, $N \in \mathbb{N}$, and $f : S \to \mathbb{R}$,

$$\begin{aligned} \|T_t^{(N)} f - T_t f\| &= \sup_{i\in S} |S_N^{[t/c_N]} f(i) - T_t f(i)| \\ &\le \sup_{i\in S} \sum_{j\in S} |f(j)|\,|(\Pi_N^{[t/c_N]})_{ij} - (P\mathrm{e}^{tG})_{ij}| \\ &\le \|f\|\,\|\Pi_N^{[t/c_N]} - P\mathrm{e}^{tG}\| \\ &\to 0 \quad \text{as } N \to \infty, \end{aligned}$$

since $\|f\| < \infty$. Thus, $\lim_{N\to\infty} \|T_t^{(N)} f - T_t f\| = 0$ for all $f : S \to \mathbb{R}$ and all $t > 0$. Note that, however, the latter convergence does not hold for $t = 0$. It is, hence, not permitted to apply convergence results such as [1, p. 168, Theorem 2.6] in order to verify convergence in the Skorokhod topology. The paths of the limiting process $(X_t)_{t\ge0}$ are in general not right-continuous at $t = 0$.

Acknowledgement

M. Notohara wishes to thank Dr Shimizu for helpful discussion and advice.

References

[1] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* John Wiley, New York.
[2] Hössjer, O. (2011). Coalescence theory for a general class of structured populations with fast migration. *Adv. Appl. Prob.* **43,** 1027–1047.
[3] Kaj, I., Krone, S. M. and Lascoux, M. (2001). Coalescent theory for seed bank models. *J. Appl. Prob.* **38,** 285–300.
[4] Möhle, M. (1998). A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. *Adv. Appl. Prob.* **30,** 493–512.
[5] Möhle, M. (1998). Coalescent results for two-sex population models. *Adv. Appl. Prob.* **30,** 513–520.
[6] Nordborg, M. and Krone, S. M. (2002). Separation of time scales and convergence to the coalescent in structured populations. In *Modern Developments in Theoretical Population Genetics,* eds M. Slatkin and M. Veuille, Oxford University Press, pp. 194–232.
[7] Pollak, E. (2011). Coalescent theory for age-structured random mating populations with two sexes. *Math. Biosci.* **233,** 126–134.
[8] Sampson, K. Y. (2006). Structured coalescent with nonconservative migration. *J. Appl. Prob.* **43,** 351–362.