Stochastic Analysis and Applications, 24: 685–702, 2006

ISSN 0736-2994 print/1532-9356 online
DOI: 10.1080/07362990600632045

Lumpability and Commutativity of Markov Processes

Jianjun Paul Tian
Mathematical Biosciences Institute, The Ohio State University,
Columbus, Ohio, USA

D. Kannan
Department of Mathematics, University of Georgia, Athens, Georgia, USA

Abstract: We introduce the concepts of lumpability and commutativity of a continuous time discrete state space Markov process, and provide a necessary and sufficient condition for a lumpable Markov process to be commutative. Under suitable conditions we recover some of the basic quantities of the original Markov process from the jump chain of the lumped Markov process.

Keywords: Commutativity; Lumpability; Markov process.

Mathematics Subject Classification: 60J10; 60J27; 60J75.

1. INTRODUCTION

The theory of Markov processes in the discrete state space has been among the important tools in the study of natural phenomena. This is particularly so in the research of modern biology, as can be seen in the utility of Markov processes in coalescent theory [3]. There also are numerous biological models which exhibit the power of Markov chains (see, for example, [1] and [7]). However, when we consider more

Received July 1, 2005; Accepted November 28, 2005
Address correspondence to Paul Tian, Mathematical Biosciences Institute, The Ohio State University, Columbus, Ohio 43210, USA; E-mail: tianjj@mbi.ohio-state.edu

realistic and/or complicated biological phenomena, it becomes necessary to develop some new techniques for the underlying Markov process itself to study these phenomena. The lumping of Markov processes is one such very useful technique. The stochastic processes literature is abundant with several papers, and a book by Kemeny and Snell [2], which exploit the lumpability of the discrete-time Markov processes or chains. As to the continuous-time Markov processes, there is almost no work involving its lumpability. When one of the authors (Tian) developed the colored coalescent theory (see Lin-Tian [4, 7]), the authors there found the lumping technique to be very crucial for their work. Therefore, we focus, in the present paper, on developing some of the basics of the lumpability of continuous-time Markov chains.

One standard approach in the study of a continuous-time Markov process with discrete state space is to analyze its jump chains. When a Markov process is lumped into a Markov process with a comparatively smaller state space, we end up with two different jump chains, one corresponding to the original process and the other to the lumped process. It is simpler to use the smaller jump chain to capture some of the fundamental qualities of the original Markov process. Toward this goal, certain conditions need to be imposed on the Markov processes. One such condition turns out to be the commutativity of Markov processes, by which we mean the commutativity of the diagram composed of four Markov processes (the underlying Markov process and its jump chain, and the lumped Markov process and its jump chain). In the present paper, we provide certain condition(s) for the commutativity of a lumpable Markov process. We also find hypotheses to recover some of the basic quantities of the underlying Markov process.

2. LUMPABILITY OF MARKOV PROCESSES

The text of Kemeny and Snell [2] defines the lumped chain of a given discrete-time finite state space Markov chain and discusses some preliminary results. We extend this notion of lumpability to continuous-time finite state space Markov chains and characterize lumpability in terms of the infinitesimal generator. For the sake of completeness and ready reference, we begin by recalling from [2] the definition and basic results that are useful for the present work.

Definition 2.1. Let $\{X_n\}$ be a Markov chain with state space $S = \{e_1, e_2, \ldots, e_r\}$ and initial vector π. Given a partition $\overline{S} = \{E_1, E_2, \ldots, E_v\}$ of the state space S, a new chain $\overline{X}_n$ can be defined as follows: At the jth step, the state of the new chain is the set E_k when E_k contains the state of the jth step of the original chain. Assign now the transition probabilities for $\overline{X}_n$ at each step as follows: The initial distribution is

$$\mathbb{P}\{\overline{X}_0 = E_i\} = \mathbb{P}_\pi\{X_0 \in E_i\}.$$

Given the initial state, the transition probability for step one is

$$\mathbb{P}\{\overline{X}_1 = E_j \mid \overline{X}_0 = E_i\} = \mathbb{P}_\pi\{X_1 \in E_j \mid X_0 \in E_i\}.$$

In general, for the nth step,

$$\begin{aligned}&\mathbb{P}\{\overline{X}_n = E_t \mid \overline{X}_{n-1} = E_s, \dots, \overline{X}_1 = E_j, \overline{X}_0 = E_i\}\\ &\quad = \mathbb{P}_\pi\{X_n \in E_t \mid X_{n-1} \in E_s, \dots, X_1 \in E_j, X_0 \in E_i\}.\end{aligned}$$

We call this new chain $\overline{X}_n$, a *lumped chain* (of the Markov chain X_n).

It is obvious that a lumped chain of a given Markov chain need not be a Markov chain in general. We therefore have the following definition.

Definition 2.2. A Markov chain $X = \{X_n\}$ with state space S is said to be *lumpable* with respect to a partition $\overline{S}$ of S if, for every starting vector π, the lumped chain $\overline{X} = \{\overline{X_n}\}$ defined previously is a Markov chain with state space $\overline{S}$ and that the associated transition probabilities do not depend on the choice of π. The new Markov chain $\overline{X}$ is called the *lumped Markov chain* of X.

Kemeny and Snell gave a necessary and sufficient condition for a Markov chain to be lumpable, and we quote that condition in the following theorm.

Theorem 2.3 (Kemeny–Snell [2]). *Let us denote the transition probability of the Markov chain X from state e_i to state e_j, $i, j = 1, \dots, r$, by p_{ij}. A necessary and sufficient condition for the Markov chain X to be lumpable with respect to the partition $\overline{S}$ is that for every pair of sets E_ξ and E_η, $\sum_{e_k \in E_\eta} p_{ik}$ has the same value for every e_i in E_ξ. These common values form the transition probabilities $p_{\xi\eta}$ for the lumped chain.*

A matrix formulation of this theorem will be more useful for our analysis. Toward this, we begin by introducing some notations. Associated with a partition $\overline{S} = \{E_1, E_2, \dots, E_v\}$ of the finite state space $S = \{e_1, e_2, \dots, e_r\}$, of the underlying Markov chain X_n, we introduce two useful matrices. Let U be the $v \times r$ matrix whose ξth row, $\xi = 1, 2, \dots, v$, is the probability vector having equal components for states in E_ξ and 0 elsewhere. Let V be the $r \times v$ matrix with the ηth column, $\eta = 1, 2, \dots, v$, is a vector with 1′s in the components corresponding to states in E_η, and 0 elsewhere.

Remark 2.4. It's easy to check that $UV = I_v$.

Now, Theorem 2.3 takes the following matrix form.

Theorem 2.5. *If P is the transition probability matrix of the Markov chain X_n, then X_n is lumpable with respect to the partition $\overline{S}$, if and only if*

$$VUPV = PV.$$

As stated in the introduction, our interest is in the continuous-time Markov chain/process with finite state space. We extend this notion of lumpability to the continuous-time case. Let us consider a Markov process $X(t)$ on S with transition probability $P(t) = (p_{ij}(t))$.

Definition 2.6. A continuous-time Markov chain $X(t)$ with a finite state space S is said to be *lumpable* with respect to the partition $\overline{S}$ if, for $e_i, e_j \in E_\xi$,

$$\sum_{e_k \in E_\eta} p_{ik}(t) = \sum_{e_k \in E_\eta} p_{jk}(t), \quad \text{for all } t \geq 0.$$

When $X(t)$ is lumpable, define its transition probabilities by

$$p_{\xi\eta}(t) = \sum_{e_k \in E_\eta} p_{ik}(t),$$

for $e_i \in E_\xi$. Denote this transition probability matrix by the $v \times v$-matrix $\overline{P}(t) = (p_{\xi\eta}(t))$.

Lemma 2.7. *$\overline{P}(t)$ defines a Markov process on $\overline{S}$.*

Proof. We need only check the Chapman-Kolmogorov semigroup property $\overline{P}(t)\overline{P}(s) = \overline{P}(t+s)$. Toward this, note that

$$\begin{aligned}\sum_\chi p_{\xi\chi}(t) p_{\chi\eta}(s) &= \sum_\chi \Big(\sum_{e_k \in E_\chi} p_{ik}(t) \Big) \cdot \Big(\sum_{e_j \in E_\eta} p_{kj}(s) \Big) \\ &= \sum_{e_j \in E_\eta} p_{ij}(t+s) \\ &= p_{\xi\eta}(t+s),\end{aligned}$$

giving us the needed semigroup property $\overline{P}(t)\,\overline{P}(s) = \overline{P}(t+s)$. □

We call the Markov process $\overline{X}(t)$ on $\overline{S}$ with the transition probability matrix $\overline{P}(t)$ a *lumping* of $X(t)$. In terms of matrices, we have the following lemma, which is a generalization of Theorem 2.3.

Lemma 2.8. *$X(t)$ is lumpable if and only if*

$$VUP(t)V = P(t)V.$$

Furthermore, when $X(t)$ is lumpable, the matrix

$$\overline{P}(t) = UP(t)V$$

is the transition probability matrix of the lumped process $\overline{X}(t)$.

The proof is a computation as in that of Theorem 6.3.5 in [2]. Therefore, we omit it.

Our motivation is to use the jump chains to study the underlying Markov processes. So, it is natural to consider characterizing lumpability by using the infinitesimal generator of the process. Let $Q = (q_{ik})$ be the infinitesimal generator of $X(t)$, i.e., $P(t) = e^{tQ}$. The following definition introduces the notion of lumpable infinitesimal generator.

Definition 2.9. We say that the (infinitesimal) generator Q is *lumpable* if

$$\sum_{e_k \in E_\eta} q_{ik} = \sum_{e_k \in E_\eta} q_{jk}$$

for $e_i, e_j \in E_\xi$.

The following result provides a simple characterization of the lumpability of the generator.

Theorem 2.10. 1. *The infinitesimal generator Q of $X(t)$ is lumpable if and only if $VUQV = QV$.*

2. *When Q is lumpable, the infinitesimal generator of its lumped chain is given by $\overline{Q} = UQV$.*

We now present a theorem which characterizes lumpability of a Markov process in terms of the associated infinitesimal generator. This result is taken from [5] and we reproduce the proof for the sake of completeness.

Theorem 2.11. *A necessary and sufficient condition for the Markov process $X(t)$ to be lumpable is that its infinitesimal generator Q is lumpable. When Q is lumpable, we have $\overline{P}(t) = e^{t\overline{Q}}$.*

Proof. When $\xi \neq \eta$ and $e_i, e_j \in E_\xi$, we have

$$\lim_{t\to 0+} \frac{p_{\xi\eta}(t)}{t} = \lim_{t\to 0+} \sum_{e_k \in E_\eta} \frac{p_{ik}(t)}{t}$$
$$= \sum_{e_k \in E_\eta} \lim_{t\to 0+} \frac{p_{jk}(t)}{t}.$$

Since $X(t)$ and $\overline{X}(t)$ are finite state Markov processes, the limits exist and hence

$$\sum_{e_k \in E_\eta} q_{ik} = \sum_{e_k \in E_\eta} q_{jk}.$$

Now when $\xi = \eta$ and $e_i \in E_\xi$, we have

$$\begin{aligned}\lim_{t\to 0+} \frac{p_{\xi\xi}(t)-1}{t} &= \lim_{t\to 0+} \frac{\sum_{e_k\in E_\xi} p_{ik}(t)-1}{t}\\ &= \lim_{t\to 0+}\left(\frac{p_{ii}(t)-1}{t} + \sum_{e_k\in E_\xi, i\neq k} \frac{p_{ik}(t)}{t}\right)\\ &= \sum_{e_k\in E\xi} q_{ik}.\end{aligned}$$

The value of the limit does not depend on the choice of e_i in E_ξ. Thus, we have shown that, if the Markov process $X(t)$ is lumpable, then its infinitesimal generator Q is lumpable.

To prove the converse, assume now that Q is lumpable. We want to show that $VUP(t)V = P(t)V$. To see this, we use the power series expansion of $P(t) = e^{tQ}$:

$$P(t) = e^{tQ} = I + tQ + \frac{t^2}{2!}Q^2 + \frac{t^3}{3!}Q^3 + \cdots.$$

First, we have $VUV = VI_v = I_rV$. Then, right-multiplying both sides of the equality $VUQV = QV$ by UQV, we get $VUQVUQV = QVUQV$. Then, using the equality $VUQV = QV$, we get $VUQ^2V = Q^2V$. Thus, we have inductively,

$$VUQ^nV = Q^nV, \quad \text{for all } n = 0, 1, 2, \ldots.$$

Hence, $VUP(t)V = P(t)V$ and $P(t)$ is therefore lumpable. Finally, since

$$(UQV)^n = UQVUQV(UQV)^{n-2} = UQ^2V(UQV)^{n-2} = UQ^nV, \quad n \geq 2,$$

we get $\overline{P}(t) = e^{t\overline{Q}}$. □

The following theorem shows that when a Markov process is lumpable and has a stationary distribution, then the lumped process also has a stationary distribution, and provides the relation between these stationary distributions.

Theorem 2.12. *Let $X(t)$ be an irreducible continuous time Markov chain with stationary distribution π. If it is lumpable with respect to a partition of the state space, then the lumped chain also has a stationary distribution $\bar{\pi}$ whose components can be obtained from π by adding corresponding components in the same cell of the partition.*

Proof. Suppose $P(t)$ is the transition matrix of the underlying Markov process $X(t)$. The partition of the state space is represented by the matrices U and V introduced earlier in this section. Then the lumped process $\overline{X}(t)$ has the transition probability matrix given by

$$\overline{P}(t) = UP(t)V.$$

Since $X(t)$ is irreducible, we have

$$P(t) \to W = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}, \quad \text{as } t \to \infty,$$

where π is the row probability vector, which is the stationary distribution of $X(t)$. Thus,

$$\overline{P}(t) \to UWV, \quad \text{as } t \to \infty,$$

which gives us the conclusion. □

Remark 2.13. 1. This result can be reworded in terms of the capacity of a subset of S (or states in $\overline{S}$). Let $\pi = (\pi_1, \ldots, \pi_r)'$ be the stationary distribution of the Markov process $X(t)$ and A be any subset of the state space S. The *capacity* C_A of $A \subset S$ is defined by

$$C_A := \sum_{e_i \in A} \pi_i.$$

Thus, $C_{e_k} = \pi_k, \ 1 \le k \le r$.

2. Now we can state this theorem in more precise terminology. If the underlying lumpable Markov process $X(t)$ has a stationary distribution $\pi = (\pi_1, \ldots, \pi_r)'$, then the lumped process $\overline{X}(t)$ also has a stationary distribution given by $\bar{\pi} = (C_{E_1}, \ldots, C_{E_v})'$.

The following result and example that conclude this section shows that the reversibility of the basic Markov process generally does not imply the reversibility of the lumped process. We begin by recalling the definition of the reversibility.

Definition 2.14. A Markov process $X(t)$ on a (discrete) state space $\mathscr{S}$ is said to be *reversible* if there is a positive measure $\pi(x)$, $x \in \mathscr{S}$, such that the following *detailed balance equations* are satisfied

$$\pi(x)Q_{xy} = \pi(y)Q_{yx}, \quad \text{for all } x, y \in \mathscr{S},$$

or equivalently,

$$\pi(x)P_{xy}(t) = \pi(y)P_{yx}(t), \quad \text{for all } t > 0, \text{ and } x, y \in \mathscr{S}.$$

Remark 2.15. We shall consider the reversibility with respect to the invariant/stationary measure $\pi(x)$. Since our state space is finite, we assume the process to be ergodic. The reversibility is equivalent to $P(t) = P^*(t)$, where the adjoint $P^*(t)$ is defined by $P^*_{xy}(t) = \frac{\pi(y)P_{yx}(t)}{\pi(x)}$.

Theorem 2.16. *Let $X(t)$ be a continuous-time Markov process on the state space $\mathcal{S} := \{e_1, \ldots, e_r\}$ with a symmetric infinitesimal generator and the invariant measure $\pi(x)$, $x \in \mathcal{S}$. Let $\overline{X}(t)$ be the lumped Markov process with respect to a lumped state space $\overline{\mathcal{S}} := \{E_1, \ldots, E_v\}$ and the corresponding stationary measure $\bar{\pi}$ as obtained in Theorem 2.12. Then, $\overline{X}$ is reversible with respect to $\bar{\pi}$.*

Proof. Since the infinitesimal generator is symmetric, the invariant measure $\pi(x) = \frac{1}{r}$. Furthermore, the transition probability matrix is also symmetric. The stationary measure of the lumped process, $\bar{\pi}_\eta$, is $\frac{|E_\eta|}{r}$. We therefore could carry out a straight forward calculation to prove the theorem:

$$\begin{aligned}
\bar{p}^*_{\xi\eta}(t) &= \frac{\bar{\pi}_\eta \bar{p}_{\eta\xi}(t)}{\bar{\pi}_\xi} \\
&= \frac{|E_\eta|}{|E_\xi|}\Big(\sum_{e_i \in E_\xi} p_{ki}(t)\Big) \\
&= \frac{1}{|E_\xi|}\sum_{e_k \in E_\eta}\sum_{e_i \in E_\xi} p_{ki}(t) \\
&= \frac{1}{|E_\xi|}\sum_{e_i \in E_\xi}\sum_{e_k \in E_\eta} p_{ki}(t) \\
&= \frac{1}{|E_\xi|}\sum_{e_i \in E_\xi}\sum_{e_k \in E_\eta} p_{ik}(t), \quad \text{since } P(t) \text{ is symmetric,} \\
&= \frac{1}{|E_\xi|}(|E_\xi|)\Big(\sum_{e_k \in E_\eta} p_{ik}(t)\Big) \\
&= \bar{p}_{\xi\eta}(t).
\end{aligned}$$

□

Example 2.17. Generally speaking, when a lumpable Markov process is reversible, the lumped process is not necessarily reversible. We look at a simple example now. Let's take a Markov process, $X(t)$, which has three states e_1, e_2, and e_3, and it is lumpable with respect to state space partition $E_1 = e_1$ and $E_2 = e_2, e_3$. If π_i, $i = 1, 2, 3$ denotes the stationary measure of $X(t)$, then the stationary measure of lumped process is π_1, $\pi_2 + \pi_3$. If the lumped process is reversible, it should satisfy the following conditions:

$$p_{12}(t) + p_{13}(t) = \frac{(\pi_2 + \pi_3)p_{21}(t)}{\pi_1},$$

and

$$p_{21}(t) = \frac{\pi_1(p_{12}(t) + p_{13}(t))}{\pi_2}.$$

This implies clearly that the reversibility of $X(t)$ is not enough for the reversibility of the lumped process.

3. COMMUTATIVITY OF MARKOV PROCESSES

When a Markov process $X(t)$ is lumpable, we associate with it (i) its jump chain (that is, the embedded chain) J, (ii) its lumped Markov process $\overline{X}(t)$, and (iii) the jump chain $\overline{J}$ of the lumped process $\overline{X}(t)$. Our interest is to compute quantities concerning the original processes from those of the associated jump chains. It thus becomes important to study Figure 1.

Definition 3.1. If the jump chain $\overline{J}$ of the lumped process $\overline{X}(t)$ is the same as the lumped chain of the jump chain J of $X(t)$, that is, if Figure 1 is commutative, we say the Markov process $X(t)$ is *commutative* with respect to lumping (or a given partition of the state space).

We now state and prove the main result of the present paper, which characterizes the commutativity of a lumpable Markov process.

Theorem 3.2. *Let $X(t)$ be a Markov process on the state space $S = \{e_1, e_2, \dots, e_r\}$ and with infinitesimal generator $Q = (q_{ij})$. Suppose $X(t)$ is lumpable with respect to the partition $\overline{S} = \{E_1, E_2, \dots, E_v\}$ of S. Then $X(t)$ is commutative if and only if*

$$q_{\xi\xi} = \sum_{e_u \in E_\eta} q_{iu} = q_{ii}, \quad \textit{for } e_i \in E_\eta, \tag{1}$$

in other words, $q_{ij} = 0$, if $e_i, e_j \in E_\eta$ and $i \neq j$.

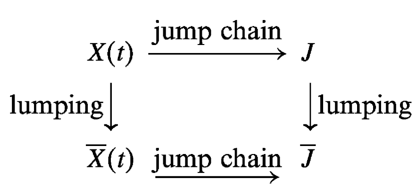

Figure 1. Commutativity diagram.

Proof. We first prove the sufficiency of the condition. Since $X(t)$ is lumpable, the transition rate from E_ζ to E_η is $q_{\zeta\eta} \equiv q_{iE_\eta} = \sum_{e_u \in E_\eta} q_{iu}$ for all $e_i \in E_\zeta$. Hence, the matrix $(q_{\zeta\eta})$ is the infinitesimal generator for the lumped process $\overline{X}(t)$. The transition probability matrix for the jump chain of $\overline{X}(t)$ is given by $(p_{\zeta\eta})$, where

$$\begin{aligned} p_{\zeta\eta} &= \frac{\sum_{e_j \in E_\eta} q_{ij}}{\sum_{e_k \in E_\zeta} q_{ik}}, \quad \text{for any } e_i \in E_\zeta, \\ &= \frac{\sum_{e_j \in E_\eta} q_{ij}}{q_{ii}}, \quad \text{if } \zeta \neq \eta, \\ p_{\zeta\eta} &= 0, \quad \text{if } \zeta = \eta, \end{aligned}$$

since $q_{ij} = 0$, for $e_i, e_j \in E_\zeta$ and $i \neq j$.

On the other hand, the transition matrix (p_{ij}) of the jump chain of $X(t)$ is

$$\begin{aligned} p_{ij} &= \frac{q_{ij}}{q_{ii}}, \quad \text{if } i \neq j, \\ p_{ij} &= 0, \quad \text{if } i = j. \end{aligned}$$

We now lump the transition probability matrix (p_{ij}) with respect to the same partition and denote the transition probability matrix of the resulting Markov chain by $(\bar{p}_{\xi\eta})$. Then

$$\begin{aligned} \bar{p}_{\zeta\eta} &= \sum_{e_j \in E_\eta} p_{ij} = \sum_{e_j \in E_\eta} \frac{q_{ij}}{q_{ii}}, \quad \text{for any } e_i \in E_\zeta, \\ &= \frac{\sum_{e_j \in E_\eta} q_{ij}}{q_{ii}}, \quad \text{if } \zeta \neq \eta, \\ \bar{p}_{\zeta\eta} &= 0, \quad \text{if } \zeta = \eta. \end{aligned}$$

This proves the sufficiency of the condition.

It remains to prove that the condition is necessary. Since $X(t)$ is lumpable, it follows from Theorem 2.10 that there are matrices U and V, such that $\overline{P}(t) = U\overline{P}(t)V$ and $\overline{Q} = UQV$. The transition probability matrix for the jump chain J of $X(t)$ is $D^{-1}Q + I$, where D is the diagonal matrix, which has the same diagonal as that of Q except for the sign of each entry, that is, $D = -diag(Q)$. Denote the diagonal of $\overline{Q}$ by $\overline{D}$. Then by the commutativty of the process, we have

$$U(D^{-1}Q + I)V = \overline{D}^{-1}(UQV) + I.$$

Since $UV = I$, this reduces to the following equation

$$UD^{-1}QV = \overline{D}^{-1}UQV.$$

We need to deal now with this equation. For convenience, we assume the partition of the state space has the order-preserved form

$$
\begin{aligned}
E_1 &= \{e_1, e_2, \dots, e_{k_1}\}, \\
E_2 &= \{e_{k_1+1}, e_{k_1+2}, \dots, e_{k_2}\}, \\
&\cdots\cdots\cdots\cdots\cdots \\
E_v &= \{e_{k_{v-1}+1}, e_{k_{v-1}+2}, \dots, e_{k_v}\},
\end{aligned}
$$

where $r = k_v$. If the partition of the state space is different, we just need permutations of the state space so as to get the order-preserved partition. For the order-preserved partition, U and V have simple forms and UQV is

$$
UQV = \begin{pmatrix}
\sum_{j=1}^{k_1} q_{1j} & \sum_{j=k_1+1}^{k_2} q_{1j} & \cdots & \sum_{j=k_{v-1}+1}^{k_v} q_{1,j} \\
\sum_{j=1}^{k_1} q_{k_1+1,j} & \sum_{j=k_1+1}^{k_2} q_{k_1+1,j} & \cdots & \sum_{j=k_{v-1}+1}^{k_v} q_{k_1+1,j} \\
\cdots & \cdots & \cdots & \cdots \\
\sum_{j=1}^{k_1} q_{k_{v-1}+1,j} & \sum_{j=k_1+1}^{k_2} q_{k_{v-1}+1,j} & \cdots & \sum_{j=k_{v-1}+1}^{k_v} q_{k_{v-1}+1,j}
\end{pmatrix}.
$$

It is easy to see

$$
D^{-1}QV = \begin{pmatrix}
q_1^{-1}\sum_{j=1}^{k_1} q_{1j} & q_1^{-1}\sum_{j=k_1+1}^{k_2} q_{1j} & \cdots & q_1^{-1}\sum_{j=k_{v-1}+1}^{k_v} q_{1,j} \\
q_2^{-1}\sum_{j=1}^{k_1} q_{2j} & q_2^{-1}\sum_{j=k_1+1}^{k_2} q_{2j} & \cdots & q_2^{-1}\sum_{j=k_{v-1}+1}^{k_v} q_{2j} \\
\cdots & \cdots & \cdots & \cdots \\
q_r^{-1}\sum_{j=1}^{k_1} q_{rj} & q_r^{-1}\sum_{j=k_1+1}^{k_2} q_{rj} & \cdots & q_r^{-1}\sum_{j=k_{v-1}+1}^{k_v} q_{rj}
\end{pmatrix}
$$

From this, we have

$$
\begin{pmatrix}
k_1^{-1}\sum_{i=1}^{k_1} q_i^{-1}\sum_{j=1}^{k_1} q_{ij} = -1 \\
(k_2-k_1)^{-1}\sum_{i=k_1+1}^{k_2} q_i^{-1}\sum_{j=k_1+1}^{k_2} q_{ij} = -1 \\
\cdots\cdots\cdots\cdots \\
(k_v-k_{v-1})^{-1}\sum_{i=k_{v-1}+1}^{k_v} q_i^{-1}\sum_{j=k_{v-1}+1}^{k_v} q_{ij} = -1
\end{pmatrix}.
$$

Observing

$$
k_1^{-1}\sum_{i=1}^{k_1}\sum_{j=1, j\neq i}^{k_1} q_i^{-1}q_{ij} = 0,
$$

we note that

$$
\sum_{i=1}^{k_1}\sum_{j=1, j\neq i}^{k_1} q_i^{-1}q_{ij} = 0.
$$

Since $q_i \neq 0$ and $q_i^{-1}q_{ij} \geq 0$, for $i \neq j$, we get $q_{ij} = 0$, $e_i, e_j \in E_1$.

Similarly, we can get $q_{ij} = 0,\ i \neq j\ \ e_i,\ e_j \in E_\eta$. This establishes the necessity of the condition and hence completes the proof of the theorem. □

We shall now proceed to find corresponding relations between the underlying Markov process and the lumped Markov process (and their respective jump chains). Continuing the topic of stationary distributions, the following lemma gives the relation between the stationary distribution of the Markov process $X(t)$ and the stationary distribution of its jump chain.

Lemma 3.3. *Let $X(\cdot)$ be an irreducible Markov process with infinitesimal generator Q and stationary distribution α. If the stationary distribution of its jump chain J is denoted by π, then*

$$\alpha = \left(\sum_{k=1}^{r} \frac{\pi_k}{q_k}\right)^{-1} \pi D^{-1},$$

where D is the negative of the diagonal of Q.

Proof. Writing $Q = -D + A$, the transition matrix for the jump chain is given by $D^{-1}A$. Then

$$\frac{I + D^{-1}A + (D^{-1}A)^2 + \cdots + (D^{-1}A)^n}{n+1} \rightarrow \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}, \quad \text{as } n \rightarrow \infty,$$

and

$$\pi(D^{-1}A) = \pi.$$

Now set $\pi D^{-1} = \beta$, then $\beta D = \pi$, and $\beta D = \beta A$, so $\beta Q = 0$. Also,

$$\begin{aligned} \beta P(t) &= \beta e^{Qt} \\ &= \beta\left(I + Qt + \frac{1}{2!}Q^2t^2 + \frac{1}{3!}Q^3t^3 + \cdots\right) \\ &= \beta + \beta Qt + \frac{1}{2!}\beta Q^2t^2 + \frac{1}{3!}\beta Q^3t^3 + \cdots \\ &= \beta. \end{aligned}$$

So β is an invariant measure. Since $X(t)$ is irreducible, β can be normalized, too, as a stationary probability vector α.

$$\beta = (\pi_1\pi_2\cdots\pi_n)\begin{pmatrix} q_1^{-1} & & \\ & \ddots & \\ & & q_n^{-1} \end{pmatrix} = \left(\frac{\pi_1}{q_1}\cdots\frac{\pi_n}{qn}\right),$$

and hence

$$\alpha = \left(\sum_{k=1}^{n} \frac{\pi_k}{q_k}\right)^{-1} \left(\frac{\pi_1}{q_1} \cdots \frac{\pi_n}{q_n}\right)$$

$$= \left(\sum_{k=1}^{n} \frac{\pi_k}{q_k}\right)^{-1} \pi D^{-1}. \qquad \square$$

When a Markov process is commutative with respect to a lumping, we can, as the following theorem shows, use the stationary distribution of the jump chain of the lumped process to recover some information about the stationary distribution of the underlying Markov process.

Theorem 3.4. *Let $X(t)$ be an irreducible Markov process which is commutative with respect to some lumping and α be an invariant measure of $X(t)$. Define the matrix Z by $Z := \begin{pmatrix} \alpha \\ \vdots \\ \alpha \end{pmatrix}$. Let $D := -\,diag(Q)$, where Q is the infinitesimal generator of $X(t)$, and $\bar{\pi}$ be the stationary distribution of the jump chain $\overline{J}$ of the lumped Markov process $\overline{X}(t)$. Define $\overline{W} = \begin{pmatrix} \bar{\pi} \\ \vdots \\ \bar{\pi} \end{pmatrix}$. Then,*

$$UZDV = \overline{W},$$

where recall that U and V are the lumping matrices.

Proof. Theorem 3.3 gives us an invariant measure $\bar{\alpha}$ for the lumped Markov process $\overline{X}(t)$. Let π denote the stationary distribution of jump chain J. Then,

$$\overline{W} = UWV, \quad W = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}. \quad \text{set } \overline{Z} = \begin{pmatrix} \bar{\alpha} \\ \vdots \\ \bar{\alpha} \end{pmatrix}.$$

Note that $\overline{Z} = UZV$. As before, decompose the infinitesimal generator Q as $Q = -D + A$ and let $\overline{Q}$ be the infinitesimal generator of $\overline{X}(t)$. Set $\overline{Q} = -\overline{D} + \overline{A}$. Then, we have $\overline{Q} = UQV$. Since the Markov process $X(t)$ is commutative, we have

$$\overline{Q} = U(-D + A)V = -UDA + UAV,$$

i.e., D and A are also lumpable matrices with the given partition

$$\overline{Q} = -\overline{D} + \overline{A}, \quad \overline{D} = UDV, \quad \overline{A} = UAV, \quad \text{and} \quad VUDV = DV.$$

Now $\overline{Z} = \overline{WD}^{-1}$, then $\overline{ZD} = \overline{W}$. Therefore,

$$UZV \cdot UDV = \overline{W}, \quad \text{that is, } UZDV = \overline{W}. \qquad \square$$

For an irreducible Markov process, the mean first return time is an interesting quatity. We shall give several statements about the mean first return time of lumped Markov processes and related quatities. We use notations that are self-explanatory. For example, $h(E_\zeta, E_\zeta)$ is the mean first return time for a lumped Markov process, while $h(e_i, e_i)$ is the mean first return time for a given Markov process. We use the upper case letters for that of Markov chains or jump chains. For example, $H(E_\zeta, E_\zeta)$ is the mean first return time for a lumped jump chain, while $H(e_i, e_i)$ is the mean first return time for a given Markov chain or jump chain.

Lemma 3.5 (First Return Time). *Let $X(t)$ be an irreducible Markov process and $u = (u_1 \cdots u_r)$ be stationary distribution of $X(t)$, E_ζ is a cell of lumping partition, then the mean first return time is given by*

$$h(E_\zeta, E_\zeta) = \frac{1}{\sum_{e_k \in E_\zeta} u_k \cdot \sum_{e_j \in E_\zeta} q_{ij}}, \quad e_i \in E_\zeta.$$

Proof. By Theorem 2.2, we know that the component of stationary distribution of lumped chain $\overline{X}(t)$ corresponding to cell E_ζ to be $\sum_{e_k \in E_\zeta} u_k$, and the waiting rate of E_ζ is $\sum_{e_j \in E_\zeta} q_{ij}$, for any $e_i \in E_\zeta$. Then we have the formula as in the theorem for the mean first return time. $\square$

Theorem 3.6. *If an irreducible and lumpable Markov chain satisfies the commutativity condition, then for every lumping partition cell E_ζ*

$$h(E_\zeta, E_\zeta) \leq h(e_k, e_k), \quad \textit{for any } e_k \in E_\zeta,$$

and

$$h(E_\zeta, E_\zeta) \leq \frac{1}{|E_\zeta|^2} \sum_{e_k \in E_\zeta} h(e_k, e_k).$$

Proof. Let $u = (u_1 u_2 \ldots u_r)$ be the stationary distribution for $X(t)$, $Q = (q_{ij})$ be the infinitesmal generator for $X(t)$, and we set $q_{ii} = -q_i$, then the mean first return time is

$$h(e_k, e_k) = \frac{1}{q_k} \frac{1}{u_k}$$

$$h(E_\zeta, E_\zeta) = \frac{1}{\sum_{e_k \in E_\zeta} u_k \cdot \sum_{e_j \in E_\zeta} q_{ij}} = \frac{1}{q_i} \frac{1}{\sum_{e_k \in E_\zeta} u_k}, \quad \text{for any } e_i \in E_\zeta.$$

Since every $u_k > 0$, then

$$h(E_\zeta, E_\zeta) \le h(e_k, e_k), \quad \text{for any } e_k \in E_\zeta.$$

Now, consider

$$\sum_{e_k \in E_\zeta} h(e_k, e_k) = \sum_{e_k \in E_\zeta} \frac{1}{q_k u_k} = \frac{1}{q_i} \sum_{e_k \in E_\zeta} \frac{1}{u_k}.$$

Using the standard inequality $\left(a_1 + \cdots + a_n\right)\left(\frac{1}{a_1} + \cdots + \frac{1}{a_n}\right) \ge n^2$, $a_i > 0$, we have

$$\sum_{e_k \in E_\zeta} \frac{1}{u_k} \cdot \sum_{e_k \in E_\zeta} u_k \ge |E_\zeta|^2,$$

$$\frac{1}{\sum_{e_k \in E_\zeta} u_k} \le \frac{1}{|E_\zeta|^2} \sum_{e_k \in E_\zeta} \frac{1}{u_k},$$

$$\frac{1}{q_i} \frac{1}{\sum_{e_k \in E_\zeta} u_k} \le \frac{1}{|E_\zeta|^2} \frac{1}{q_i} \sum_{e_k \in E_\zeta} \frac{1}{u_k}, \quad \text{for any } e_i \in E_\zeta.$$

Therefore,

$$h(E_\zeta, E_\zeta) \le \frac{1}{|E_\zeta|^2} \sum_{e_k \in E_\zeta} h(e_k, e_k).$$ □

Let us now recall the definition of the frequency for the Markov process $X(t)$ on a weighted graph [6]

$$f := \lim_{t \to \infty} \frac{N(t)}{t}$$

where the quantity $N(t)$ for $t > 0$ is

$$N(t) = E \text{ (the number of jumps of } X(t) \text{ up to time } t).$$

Actually, this definition works for any Markov process.

Theorem 3.7. *If an irreducible lumpable Markov chain X satisfies the commutativity condition, then the frequency for $X(t)$ and $\overline{X}(t)$ is the same.*

Proof. By the definition, the frequency for $X(t)$ is given by

$$\begin{aligned} f_X &= \lim_{t \to \infty} \frac{E \text{ (the number of jumps of } X(t) \text{ up to time } t)}{t} \\ &= \sum_i u_i q_i, \quad \text{(by Theorem 2.1 in [6])}, \end{aligned}$$

where $u = (u_1 \cdots u_r)$ is the stationary distribution for $X(t)$, and

$$\begin{aligned} f_{\overline{X}} &= \sum_{\zeta}\left(\left(\sum_{e_k\in E_\zeta} u_k\right)\cdot\left(\sum_{e_j\in E_\zeta} q_{ij}\right)\right) \quad (\text{fix } e_i \in E_\zeta) \\ &= \sum_{\zeta}\left(\left(\sum_{e_k\in E_\zeta} u_k\right)q_i\right) \\ &= \sum_{\zeta}\left(\left(\sum_{e_k\in E_\zeta} u_k q_\zeta\right)\right. \quad (\text{write } q_i \text{ as } q_\zeta,\, e_i \in E_\zeta) \\ &= f_X. \end{aligned}$$

□

Corollary 3.8. *We have*

$$H(E_\zeta, E_\zeta) \le H(e_k, e_k), \quad \textit{for any } e_k \in E_\zeta$$

and

$$H(E_\zeta, E_\zeta) \le \frac{1}{|E_\zeta|^2}\sum_{e_k\in E_\zeta} H(e_k, e_k),$$

where $H(\cdot,\cdot)$ is the mean first return times (steps) for the jump chain.

Proof. By Theorem 3.7, the frequencies for $X(t)$ and $\overline{X}(t)$ are the same $f_X = f_{\overline{X}}$. We also know that

$$H(e_i, e_i) = f_X h(e_i, e_i),$$

and

$$H(E_\zeta, E_\zeta) = f_{\overline{X}} h(E_\zeta, E_\zeta).$$

This concludes the proof. □

Theorem 3.9. *Let $X(t)$ be an absorbing chain. If $X(t)$ is lumpable (lumping only states of the same kind), then for every cell E_ζ, the mean absorption time of $e_k \in E_\zeta$ and that of E_ζ are the same.*

Proof. Denote the transition function of $X(t)$ as $P(t) = (p_{ij}(t))$ and lumping partition matrices as usual by U and V, then $\overline{X}(t)$ has transition function $UP(t)V$. Also, the absorption time T_{e_k} has mean

$$E(T_{e_k}) = \sum_{e_j\in S_1}\int_0^\infty p_{kj}(t)dt,$$

where S_1 is the subset whose elements are all non-absorbing states. For the lumped process,

$$E(T_{E_\zeta}) = \sum_{E_\eta \in \overline{S}_1} \int_0^\infty p_{\zeta\eta}(t)dt,$$

where $\overline{S}_1$ is the subset of the state space partition whose elements are all non-absorbing lumped states. By lumpability

$$E(T_{E_\zeta}) = \sum_{E_\eta \in \overline{S}_1} \int_0^\infty p_{\zeta\eta}(t)dt = \sum_{E_\eta \in \overline{S}_1} \int_0^\infty \Big(\sum_{e_j \in E_\eta} p_{kj}(t) \Big) dt, \quad (e_k \in E_\eta)$$

$$= \sum_{E_\eta \in \overline{S}_1} \sum_{e_j \in E_\eta} \int_0^\infty p_{kj}(t)dt = \sum_{e_j \in \overline{S}_1} \int_0^\infty p_{kj}(t)dt = E(T_{e_k}). \qquad \square$$

Lemma 3.10. *Let $X(t)$ be a simple continuous time random walk on a connected graph and $X(t)$ be lumpable. Then the mean hitting time from cell E_ζ to cell E_η is given by*

$$h(E_\zeta, E_\eta) = \frac{1}{|E_\eta|} \Big(\sum_{e_j \in E_\eta} h(e_k, e_j) - \sum_{e_j \in E_\eta, e_j \neq e_i} h(e_i, e_j) \Big), \quad e_j \neq e_i$$

when $e_k \in E_\zeta$, $e_i \in E_\eta$.

Proof. Let A be the incidence matrix of graph G and D be the degree matrix of G. Then $Q = -D + A$. Also note that the ergodic distribution for $X(t)$ is given by $u = (\frac{1}{n} \cdots \frac{1}{n})$, where n is the order (number of vertice) of G, and $P(t) = e^{Qt} = (p_{ij}(t))_{n \times n}$. Then for the lumped cells E_ζ and E_η $(\zeta \neq \eta)$

$$\begin{aligned}
h(E_\zeta, E_\eta) &= \frac{\int_0^\infty (p_{\eta\eta}(t) - u_\eta)dt - \int_0^\infty (p_{\zeta\eta}(t) - u_\eta)dt}{u_\eta} \\
&= \frac{\int_0^\infty (p_{\eta\eta}(t) - p_{\zeta\eta}(t))dt}{u_\eta} \\
&= \frac{\int_0^\infty \big(\sum_{e_j \in E_\eta} p_{ij}(t) - \sum_{e_j \in E_\eta} p_{kj}(t) \big) dt}{\sum_{e_j \in E_\eta} \pi_j} \\
&= \frac{\int_0^\infty \big(\sum_{e_j \in E_\eta} p_{ij}(t) - \sum_{e_j \in E_\eta} p_{jj}(t) + \sum_{e_j \in E_\eta} p_{jj}(t) - \sum_{e_j \in E_\eta} p_{kj}(t) \big) dt}{|E_\eta| \frac{1}{n}} \\
&= \frac{1}{|E_\eta| \frac{1}{n}} \left(\sum_{e_j \in E_\eta} \int_0^\infty (p_{jj}(t) - p_{kj}(t))dt - \sum_{e_j \in E_\eta} \int_0^\infty (p_{jj}(t) - p_{ij}(t))dt \right) \\
&= \frac{1}{|E_\eta|} \Big(\sum_{e_j \in E_\eta} h(e_k, e_j) - \sum_{e_j \in E_\eta, e_j \neq e_i} h(e_i, e_j) \Big). \qquad \square
\end{aligned}$$

Actually, this formula holds for any symmetric random walk or symmetric Markov chain. (Symmetry means infinitesimal generator is symmetric matrix.)

ACKNOWLEDGMENT

The first author would like to thank Professor Barry Arnold for his many important suggestions.

REFERENCES

1. Kannan, D. 1983. A Markov chain analysis of predator strategy in a model-mimic system. *Bull. Math. Biol.* 45:347–400.
2. Kemeny, J.G., and J.L. Snell. 1976. *Finite Markov Chains*. New York: Springer-Verlag.
3. Kingman, J.F.C. 1982. The coalescent processes. *Stochastic Processes and Their Applications* 13:235–248.
4. Tian, J., and X.-S. Lin. 2005. Colored coalescent theory. *Discrete and Continuous Dynamical Systems* 833–845.
5. Tian, J., and X.-S. Lin. Colored genealogical trees and coalescent theory. In Review.
6. Tian, J., and X.-S. Lin. Continuous time Markov processes on graphs. In Review.
7. Ewens, W.J. 2004. *Mathematical Population Genetics*. Second edition. Springer.