Chin. J. Appl. Prob. Stat. 12:1 (1996), 55-59.

A COMMENT ON THE BOOK "CONTINUOUS-TIME MARKOV CHAINS" BY W.J. ANDERSON

Mu-Fa Chen

(Beijing Normal University)

Abstract. The book "Continuous-Time Markov Chains" by W. J. Anderson collects a large part of the development in the past thirty years. It is now a popular reference for the researchers on this subject or related fields. Unfortunately, due to a misunderstanding of the approximating methods, several results in the book are incorrectly stated or proved. Since the results are related to the present author's work, to whom it may be a duty to correct the mistakes in order to avoid further confusion. We emphasize the approximating methods because they are useful in many situations.

Throughout the note, take $E = \{0, 1, 2, \cdots\}$ and suppose that the Q-matrix $Q = (q_{ij} : i, j \in E)$ in the consideration is totally stable: $q_i := -q_{ii} < \infty$ for all $i \in E$. Refer to [1], [3] or [6] for further notation used below. Here, we mention only that the term "Q-process" is called "q-function" in [1].

1. Stochastic Comparability.

Given two Q-matrices $Q^{(1)} = \big(q_{ij}^{(1)}\big)$ and $Q^{(2)} = \big(q_{ij}^{(2)}\big)$, denote by $P^{\min(k)}(t) = \big(P_{ij}^{\min(k)}(t)\big)$, $k = 1, 2$ the corresponding minimal Q-processes. The Q-processes $P^{\min(1)}(t)$ and $P^{\min(2)}(t)$ are said to be **stochastically comparable** if

$$\sum_{j\geqslant k} P_{ij}^{\min(1)}(t) \leqslant \sum_{j\geqslant k} P_{mj}^{\min(2)}(t) \quad \text{for all } i \leqslant m \text{ and } k \geqslant 0. \tag{1.1}$$

From (1.1), we obtain

$$\sum_{j\geqslant k} q_{ij}^{(1)} \leqslant \sum_{j\geqslant k} q_{mj}^{(2)} \quad \text{for all } i \leqslant m \text{ and } k \in \{0, \cdots, i\} \cup \{m+1, m+2, \cdots\}. \tag{1.2}$$

1991 *Mathematics Subject Classification.* 60J27.

Key words and phrases. Q-process, approximation method, stochastically comparability..

Research supported in part by NSFC and the State Education Commission of China.

To see this, simply use the backward Kolmogorov equation to deduce that

$$\lim_{t\to 0}\frac{1}{t}\sum_{j\in A}P_{ij}^{\min(k)}(t)=\sum_{j\in A}q_{ij}^{(k)},\qquad i\notin A\subset E.$$

However, the proof of (1.1) $\Longrightarrow$ (1.2) given in [1; p.249] works only for the conservative case. The Kirstein's original theorem states that *if the Q-matrices $Q^{(1)}$ and $Q^{(2)}$ are both regular, then* (1.1) *and* (1.2) *are equivalent*. The theorem was extended in [1; Theorem 7.3.4], where it was claimed that if both $Q^{(1)}$ and $Q^{(2)}$ are conservative, then (1.2) implies (1.1) *without using the uniqueness assumption for the Q-processes* (By the way, we mentioned that in the last two sentences of part (1) (resp. part (2)) of Theorem 7.3.4 there, the items (a) and (b) have to be exchanged). Unfortunately, the extension is incorrect.

Counterexample. *Consider the conservative birth-death Q-matrices*

$$q_{i,i-1}^{(1)}=q_{i,i+1}^{(1)}=q_{i,i-1}^{(2)}=(i+1)^2,\quad q_{i,i+1}^{(2)}=\alpha(i+1)^2,\quad i\geqslant 0,\quad \alpha>1.$$

Then (1.2) *holds but* (1.1) *does not.*

Proof. a) Note that in the conservative case, condition (1.2) is reduced to that for all $i\leqslant m$,

$$\sum_{j\geqslant k}q_{ij}^{(1)}\leqslant\sum_{j\geqslant k}q_{mj}^{(2)}\quad\text{if } k\geqslant m+1\quad\text{and}\quad\sum_{j=0}^{k}q_{ij}^{(1)}\geqslant\sum_{j=0}^{k}q_{mj}^{(2)}\quad\text{if } k\leqslant i-1.\tag{1.3}$$

Now, it is easy to check that our birth-death Q-matrices satisfy (1.3).

b) Next, for a given birth-death Q-matrix: $q_{i,i-1}=(i+1)^2$ and $q_{i,i+1}=\beta(i+1)^2$, $\beta>0$, it is well known that the Q-process is unique iff $\beta\leqslant 1$. Applying this to our example, we see that the $Q^{(1)}$-process is unique but not the $Q^{(2)}$-process. Thus, $\sum_{j\geqslant 0}P_{ij}^{\min(1)}(t)=1>\sum_{j\geqslant 0}P_{mj}^{\min(2)}(t)$ and so (1.1) fails. □

To explain what was wrong in the proof of [1; Theorem 7.3.4], we should introduce some approximating methods.

2. Approximating Methods.

In the study of Markov chains, there are several different approximating methods. Among them, the simplest one is as follows. Take $E_n\subset E$, $E_n\uparrow E$ and let $\{Q_n\}$ be the truncated Q-matrices of Q. That is,

$$q_{ij}^{(n)}=\begin{cases}q_{ij}, & \text{if } i,j\in E_n\\ 0, & \text{otherwise}\end{cases}\tag{2.1}$$

Then we have

$$P_{ij}^{\min(n)}(t)\uparrow P_{ij}^{\min}(t)\quad\text{as } n\uparrow\infty\ \text{ for all } i,j\in E_n \text{ and } t\geqslant 0.\tag{2.2}$$

Usually, one chooses $\{E_n\}$ so that each Q_n is a bounded Q-matrix. Note that the Q-matrices Q_n often become non-conservative even though the original Q-matrix is usually assumed to be conservative in practice.

It is known that for bounded Q-matrices, (1.1) and (1.2) are equivalent. Hence, the main step of the proof of [1; Theorem 7.3.4] is again using a sequence of Q-processes with bounded Q-matrices to approximate the minimal Q-process. However, in order to keep (1.2), the truncated Q-matrices given by (2.1) is not suitable. One adopts a different choice:

$$q_{ij}^{(n)} = \begin{cases} q_{ij}, & \text{if } i, j \leqslant n-1 \\ \sum_{k \geqslant n} q_{ik}, & \text{if } i \leqslant n-1 \text{ and } j = n \\ 0, & \text{otherwise.} \end{cases} \tag{2.3}$$

Two different approximation methods were introduced in the study of reaction-diffusion processes [2, 3, 4]. The first one is a truncating from E to $E_n := \{i \in E : q_i \leqslant n\}$,

$$q_{ij}^{(n)} = I_{E_n}(i) q_{ij}, \qquad i,\ j \in E, \tag{2.4}$$

where I_A is the indicator of set A. The second one is stopped at the n-th row,

$$q_{ij}^{(n)} = \begin{cases} q_{ij}, & \text{if} \quad i \leqslant n \\ q_{nj}, & \text{if} \quad i > n. \end{cases} \tag{2.5}$$

In both cases, the resulting Q-matrices $Q_n = \big(q_{ij}^{(n)}\big)$ are conservative if so is the original $Q = (q_{ij})$. The latter one was used to keep the Lipschitz property of the corresponding semi-group [2]. The former one is especially powerful to deal with the uniqueness of Q-processes. It was originally used in [3] to prove [1; Corollary 2.2.15, Corollary 2.2.16 and Theorem 7.5.5]. It should be pointed out that the results of [1; §2.2] starting from Proposition 2.2.12 are mainly taken (but without mentioned) from [3] restricted to the particular case of Markov chains.

From my knowledge, a new point was made by the author in [1; Proposition 2.2.14], where (2.1), (2.3) and (2.4) are unified into a general form. Take $E_n \uparrow E$ and let $Q_n = \big(q_{ij}^{(n)}\big)$ be a Q-matrix such that

$$q_{ij}^{(n)} = \begin{cases} q_{ij}, & \text{if} \quad i,\ j \in E_n \\ 0, & \text{if} \quad i \notin E_n. \end{cases} \tag{2.6}$$

Here $q_{ij}^{(n)}$ $(i \in E_n,\ j \notin E_n)$ are left to be freedom[1]. Moreover, the author proved that

$$\begin{aligned} &P_{ij}^{\min (n)}(t) \leqslant P_{ij}^{\min (n+1)}(t) \qquad \text{for all } i,\ j \in E_n \text{ and } t \geqslant 0 \\ &\text{and } \lim_{n\to\infty} P_{ij}^{\min (n)}(t) = P_{ij}^{\min}(t) \qquad \text{for all } i, j \text{ and } t \geqslant 0. \end{aligned} \tag{2.7}$$

[1]Addition in Proof (This and the next footnotes are made on December 24, 2014). Certainly, the additional conditions that $q_{ij}^{(n)} \geqslant 0$ for all $i \in E_n, j \neq E_n$ and $\sum_j q_{ij}^{(n)} \leqslant 0$ for every $i \in E_n$ are required here.

Since the monotonicity of $P_{ij}^{\min (n)}(t)$ in n depends on i and j, the conclusion (2.7) is weaker than the following statement:

$$P_{ij}^{\min (n)}(t) \uparrow P_{ij}^{\min}(t) \quad \text{as } \ n \to \infty, \quad \text{for all } i, j, t \geqslant 0.$$

Unfortunately, the author neglected the difference of these two statements and incorrectly wrote the latter one as the conclusion of [1; Proposition 2.2.14].[2] This

[2] Here are some counterexamples to the proposition. Let $Q = (q_{ij} : i, j \in E)$ be a conservative matrix on $E = \{0, 1, 2, \ldots\}$. For each $n \geqslant 1$, set $E_n = \{0, 1, \ldots, n\}$ and define

$$q_{ij}^{(n)} = I_{E_n}(i) q_{ij}, \qquad i, j \in E$$

which is (2.4) here and is a special case of (2.6) in [1]. This is a bounded and conservative Q-matrix on E and hence the corresponding process is unique. For each n, denote by $P^{(n)}(t)$ the corresponding process. Of course,

$$\sum_{j\in E} P_{ij}^{(n)}(t) = 1, \qquad i \in E,\ t \geqslant 0.$$

Recall that [1; Proposition 2.2.14] says that

$$P_{ij}^{(n)}(t) \uparrow P_{ij}^{\min}(t), \qquad i, j \in E,\ t \geqslant 0$$

where $P^{\min}(t)$ is the minimal process determined by the original Q. If the proposition were true, then the monotone convergence theorem would imply that

$$1 = \sum_{j\in E} P_{ij}^{(n)}(t) \uparrow \sum_{j\in E} P_{ij}^{\min)}(t) \qquad \text{as } n \to \infty.$$

This leads to a contradiction since we do not assume here the Q-process to be unique.

As far as I know, among the different approximating methods listed at the beginning of this section, the conclusion of [1; Proposition 2.2.14] holds only for the one defined by (2.1).

Clearly, the proof of [1; Corollary 2.2.15] is incomplete since at the last step, one uses the incorrect [1; Proposition 2.2.14].

However, [1; Corollary 2.2.16] is correct. Honestly, I would to say that this is one of my favourite contribution to the theory of Markov Chains. It appeared first in my Chinese book [4] in 1986 and then in the paper [3; Theorem 16)] in the same year. This paper is cited in [1], but the originality of the result is not mentioned. As we know, the earlier known criterion for the uniqueness is asking to solve a equation with infinite variables. This is usually not practicable. Actually, it costs me more than 5 years to find out such a powerful sufficient condition. A short story about this and others is included in Chapter 9 of my book "*Eigenvalues, Inequalities, and Ergodic Theory*". Springer 2005. In which, a number of papers are included for the applications of this result. A new application of the result to genetic study is given in "Transformation Markov jump processes" by Z.M. Ma et al. There the authors apply the result to the continuous state space. Actually, my original result is stated for general state space, not necessarily discrete one. Except the 1986's publications mentioned above, the same result was also published several times later: in the paper "*On three classical problems for Markov chains with continuous time parameters*, J. Appl. Prob. 28(1991)2, 305-320; in the enlarged version [6; pages 81 and 84] (2nd ed. 2004) of my Chinese book [4]; in a textbook "*Introduction to Stochastic Processes* (in Chinese, 2006. Higher Education Press, Beijing)" by me and Yong-Hua Mao. Actually, this result was taught in a course on Stochastic Processes for undergraduate/graduate students at my university every year since 1989.

leads to the incorrect proof of [1; Theorem 7.3.4]. More precisely, in the latter case, we have

$$\lim_{n\to\infty}\sum_{j\geqslant k}P_{ij}^{\min\,(n)}(t)=\sum_{j\geqslant k}P_{ij}^{\min}(t).$$

But under (2.7), we may only have

$$\varliminf_{n\to\infty}\sum_{j\geqslant k}P_{ij}^{\min\,(n)}(t)\geqslant\sum_{j\geqslant k}P_{ij}^{\min}(t).$$

The inequality can be appeared as shown by the above counterexample.

Note that (2.5) is not contained in (2.6). In general if

$$\lim_{n\to\infty}q_{ij}^{(n)}=q_{ij},\qquad i,j\in E, \tag{2.8}$$

then by the backward Kolmogorov equation and the Fatou's lemma, we have

$$\varliminf_{n\to\infty}P_{ij}^{\min\,(n)}(t)\geqslant P_{ij}^{\min}(t),\qquad i,j\in E,\ \ t\geqslant 0 \tag{2.9}$$

(cf. [6; Lemma 5.14] for example). For such a general setup, the conclusion (2.9) is certainly far away from (2.7), we even do not know that the limit $\lim_{n\to\infty}P_{ij}^{\min(n)}(t)$ exists or not and can not say that $\varliminf_{n\to\infty}P_{ij}^{\min\,(n)}(t)$ provides us a Q-process. However, whenever $Q=(q_{ij})$ being regular, we do have

$$\lim_{n\to\infty}P_{ij}^{\min\,(n)}(t)=P_{ij}^{\min}(t),\qquad i,j\in E,\ \ t\geqslant 0.$$

3. Extended Kirstein's Theorem.

The correct extension of the Kirstein's Theorem may be stated as follows.

Theorem. *The condition* (1.1) *always implies* (1.2). *Conversely,* (1.2) *implies* (1.1) *provided either* $Q^{(1)}$ *and* $Q^{(2)}$ *are bounded or* $Q^{(2)}$ *being regular.*

Because of (2.7) or (2.9), the Q-matrix $Q^{(1)}$ can be arbitrary. The proof of the conclusion is the same as the proof of [5; Theorem 7] or [6; Theorem 5.31] and hence is omitted here. By the way, we mention that for [5; Theorem 7] or [6; Theorem 5.31], the regularity assumption for the first q-pair can be removed.

The next result shows that in general it is impossible to remove the uniqueness assumption. Recall that a Q-process $P(t)=(P_{ij}(t))$ is said to be **monotone** if (1.1) holds for the same process $P(t)$.

Proposition. *Given a single birth Q-matrix (i.e., $q_{i,i+1}>0$ and $q_{ij}=0$ for all $j>i+1$, $i\geqslant 0$), the corresponding minimal Q-process is monotone iff the Q-process is unique and* (1.2) *holds.*

Proof. The sufficiency as well as the necessity of (1.2) follow from the above theorem. To prove the necessity of the uniqueness, suppose that (1.1) holds. Then,

$\sum_j P_{ij}^{\min}(t)$ is increasing in i and so is its Laplace transform $\sum_j P_{ij}^{\min}(\lambda)\,(\lambda > 0)$. Hence

$$z_i(\lambda) := 1 - \lambda \textstyle\sum_j P_{ij}^{\min}(\lambda), \qquad \lambda > 0$$

is decreasing in i. On the other hand, it is known that $z_i(\lambda)$ is increasing in i and it is indeed strictly increasing except $z_i(\lambda) \equiv 0$ [6; Proof of Theorem 3.16]. Therefore, $z_i(\lambda) \equiv z_0(\lambda) = 0$ and hence the process is unique. $\square$

The above proposition shows that the example given in [1; p.251] and furthermore [1; Corollary 7.4.3] are also incorrect.

4. Application to the Uniqueness of Q-processes.

To illustrate an application of the comparison technique, we present the following result.

Proposition. *Let $Q^{(1)}$ and $Q^{(2)}$ be conservative Q-matrices satisfying* (1.2). *If $Q^{(2)}$ is regular, then so is $Q^{(1)}$. In particular, if there exists a non-negative function φ such that $\varphi_i \uparrow \infty$ as $i \uparrow \infty$ and moreover*

$$\sum_j q_{ij}^{(2)}(\varphi_j - \varphi_i) \leqslant c(1 + \varphi_i), \qquad i \in E \tag{4.1}$$

for some constant $c \geqslant 0$. Then both $Q^{(1)}$ and $Q^{(2)}$ are regular.

Proof. a) By using (2.3) and restricting to the finite space $E_n = \{0, 1, \cdots, n\}$, we obtain

$$P_{in}^{(n,1)}(t) \leqslant P_{in}^{(n,2)}(t), \qquad i \leqslant n-1, \quad t \geqslant 0,$$

where $\big(P_{ij}^{(n,k)}(t)\big)$ is the Q-process corresponding the truncating Q-matrix $Q_n^{(k)}$, $k = 1,\ 2$. This means that

$$\sum_{j \geqslant n} P_{ij}^{(n,1)}(t) \leqslant \sum_{j \geqslant n} P_{ij}^{(n,2)}(t), \qquad i \leqslant n-1, \quad t \geqslant 0. \tag{4.2}$$

Next, denote by $\big(\bar{P}_{ij}^{(n,k)}(t)\big)$ $(k = 1,\ 2)$ the Q-process obtained by using (2.4). Then, by the localization theorem [6; Theorem 2.13], we get

$$\sum_{j \leqslant n-1} P_{ij}^{(n,k)}(t) = \sum_{j \leqslant n-1} \bar{P}_{ij}^{(n,k)}(t), \qquad i \leqslant n-1, \quad k = 1,\ 2.$$

Equivalently,

$$\sum_{j \geqslant n} P_{ij}^{(n,k)}(t) = \sum_{j \geqslant n} \bar{P}_{ij}^{(n,k)}(t), \qquad i \leqslant n-1, \quad k = 1,\ 2.$$

Combining this with (4.2), we obtain

$$\sum_{j \geqslant n} \bar{P}_{ij}^{(n,1)}(t) \leqslant \sum_{j \geqslant n} \bar{P}_{ij}^{(n,2)}(t), \qquad i \leqslant n-1, \quad t \geqslant 0.$$

The first assertion now follows from [1; Corollary 2.2.15].

b) Let (4.1) holds. Applying [1; Corollary 2.2.16] to $E_n = \{i : \varphi_i \leqslant n\}$ and $x_i = 1 + \varphi_i$, it follows that the $Q^{(2)}$-process is unique. Hence the second assertion follows from the first one. $\square$

We remark that the condition (4.1) is necessary for the uniqueness of single birth processes [3; Remark 23]. Refer to [4; Remark 3.5.1] or [6; Remark 3.20] for a proof.

Finally, the proof of [1; Theorem 7.5.5] is also incorrect since the resulting bounded Q-matrices do not possess the coupling relation. The original proof uses (2.4). Refer to [3] or [4; Chapter 5] for details.

Acknowledgement. The author is greatly indebted to Prof. Yong-Long Dai, Prof. Feng-Yu Wang and Mr. Yu-Hui Zhang for their comments on the first version of the paper. The proof a) of the last proposition is due to Y. H. Zhang, which improves the original conclusion.

5. References.

[1] Anderson, W. J. (1991), *Continuous-Time Markov Chains*, Springer Series in Statistics.

[2] Chen, M. F. (1985), *Infinite-dimensional reaction-diffusion processes*, Acta Math. Sin. New Ser.**1**:3, 261–273.

[3] Chen, M. F.(1986), *Couplings of jump processes*, Acta Math. Sin. New Ser.**2**:2, 123–136.

[4] Chen, M. F.(1986), *Jump Processes and Particle Systems* (In Chinese), Beijing Normal Univ. Press, Beijing.

[5] Chen, M. F.(1991), *On coupling of jump processes*, Chin. Ann. Math. **12(B)**:4, 385–399.

[6] Chen, M. F. (1992), *From Markov Chains to Non-Equilibrium Particle Systems*, World Scientific.

[7] Kirstein, B. M.(1976), *Monotonicity and comparability of time-homogeneous Markov processes with discrete state space*, Math. Operationsforsch. Statist. **7**, 151–168.

Department of Mathematics, Beijing Normal University, Beijing 100875, The People's Republic of China.