Combinatorics, Probability and Computing (2005) 14, 269–273. © 2005 Cambridge University Press
doi:10.1017/S0963548304006352 Printed in the United Kingdom

Lifting Markov Chains to Random Walks on Groups

ONN CHAN[†] and T. K. LAM[‡]

Department of Mathematics, National University of Singapore,
Kent Ridge, S119260, Republic of Singapore
(e-mail:onn.chan@ubs.com,taokai@yahoo.com)

Received 17 November 2000; revised 8 December 2003

1. Introduction and motivation

In a 1947 paper [6], M. Kac derived the eigenvalues and eigenvectors of the probability transition matrix associated with the Ehrenfest urn model with n balls, in which a ball is selected at random and moved to the other urn. The connection (see, for instance, [1, p. 19]) between the Markov chain defined by the Ehrenfest model and the nearest-neighbour uniform random walk on the abelian group $\mathbb{Z}_2^n$ prompted M. Kac to ask when a Markov chain may be lifted to a random walk on a group. More specifically, given a Markov chain $\{X_m\}_{m\geqslant 0}$ with state space $S = \{1, 2, \ldots, n\}$ and probability transition matrix $P = [p_{ij}]$ (here $p_{ij} := P(X_{r+1} = j | X_r = i)$ for all $r \geqslant 0$), we say that the Markov chain $\{X_m\}_{m\geqslant 0}$ or equivalently P *lifts* to a random walk on a finite group G if there exists a probability measure μ on G and a surjective map $L : G \to S$ such that, for all $i, j \in S$, and for each $g \in L^{-1}(i)$,

$$p_{ij} = \sum_{h \in L^{-1}(j)} \mu(g^{-1}h).$$

The question is: Can we find a random walk on some finite group G with probability measure μ and a surjective map $L : G \to S$ so that the image under L of the random walk on G is the given Markov chain? For the random walk on G, elements $g_1, g_2, \cdots \in G$ are selected independently with respect to the probability measure μ on G and the position of the walk, starting at $g_0 \in G$, after r steps in given by $g_0 g_1 g_2 \ldots g_r$ for all $r \geqslant 1$.

[†] Current address: Block 2, Everton Park, #02-51, Singapore 081002, Republic of Singapore.
[‡] Current address: 86, Marlborough St, #9, Boston, MA 02116, USA (e-mail: tklam@alum.mit.edu).

https://doi.org/10.1017/S0963548304006352 Published online by Cambridge University Press

If the Markov chain $\{X_m\}_{m\geqslant 0}$ lifts to a random walk on some finite group G, one may regard the Markov chain as being obtained by an amalgamation (given by $L : G \to S$) of some state spaces (group elements) in G. Several works [2, 3, 5] have investigated such functionals that may be derived from Markov chains. In the question at hand, the extra ingredient is the presence of the group structure.

An important issue in the study of Markov chains is the rate of convergence to stationarity. Much progress has been made in this area for random walks on groups through the use of Fourier analytic methods from group representations. A good reference for this is the book [1] by P. Diaconis. Knowing that a Markov chain lifts to a random walk on a group thus provides an additional method of analysing the Markov chain. The analysis is facilitated whenever the probability measure on the group is a class function. A special case is that of an abelian group where any probability measure is a class function. In this note, we shall show that a probability transition matrix P lifts to a random walk on a group if and only if P admits a stationary probability distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ in which each stationary probability π_i is a positive rational number.

2. Main results

When the transition matrix P is an $n \times n$ doubly stochastic matrix it may be lifted to a random walk on S_n, the symmetric group (or permutation group) on n symbols. This is an exercise in [1, p. 20]. This gives a sufficient condition for when P may be lifted. For completeness, we provide a proof of this result.

Lemma 2.1. *Let P be an $n \times n$ doubly stochastic matrix. Then P lifts to a random walk on S_n.*

Proof. By Birkhoff's theorem (see, for instance, [4, p. 527]), we may write $P = \sum_{\sigma \in S_n} \alpha_\sigma P_\sigma$ as a convex nonnegative linear combination of permutation matrices P_σ. We take the measure μ on S_n to be

$$\mu(\sigma) = \alpha_\sigma \quad \text{for all } \sigma \in S_n$$

and define $L : S_n \to \{1, 2, \ldots, n\}$ by

$$L(\sigma) = \sigma(1).$$

It is easy to check that this is indeed a lifting of P to the symmetric group S_n. □

The next lemma gives a necessary condition for the lifting of P.

Lemma 2.2. *Let P be an $n \times n$ probability transition matrix associated with a Markov chain on the state space $S = \{1, 2, \ldots, n\}$. Suppose P lifts to a random walk on a finite group G with probability measure μ through the map $L : G \to \{1, 2, \ldots, n\}$. Then P has a stationary distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, where each π_i is a positive rational number. Moreover, the ratios satisfy*

$$\pi_1 : \pi_2 : \ldots : \pi_n = |L^{-1}(1)| : |L^{-1}(2)| : \ldots : |L^{-1}(n)|.$$

https://doi.org/10.1017/S0963548304006352 Published online by Cambridge University Press

Proof. We first note that if P lifts to a random walk on G, then the image under L of a stationary distribution, say π^*, of a random walk on G with respect to μ is also a stationary distribution of P. Here the image $\pi^{L(*)}$ is defined, for each $i \in S$, by

$$\pi^{L(*)}(i) = \sum_{g \in L^{-1}(i)} \pi^*(g).$$

We observe that $\pi^{L(*)}$ is a probability distribution on S. Also

$$\begin{aligned}\sum_{i \in S} \pi^{L(*)}(i) p_{ij} &= \sum_{i \in S} \left(\sum_{g \in L^{-1}(i)} \pi^*(g) \right) \left(\sum_{h \in L^{-1}(j)} \mu(g^{-1}h) \right) \\ &= \sum_{h \in L^{-1}(j)} \left(\sum_{i \in S} \sum_{g \in L^{-1}(i)} \pi^*(g) \mu(g^{-1}h) \right) \\ &= \sum_{h \in L^{-1}(j)} \pi^*(h) \\ &= \pi^{L(*)}(j),\end{aligned}$$

showing that $\pi^{L(*)}$ is a stationary distribution of P.

Now the uniform distribution π^U on G is a stationary distribution of the random walk on G with respect to μ. It follows that its image $\pi^{L(U)}$ is a stationary distribution of P. Moreover $\pi^{L(U)}(i) = \sum_{g \in L^{-1}(i)} \pi^U(g) = |L^{-1}(i)|/|G|$ for each $i \in S$. Hence the result. □

It turns out that the condition given in the previous lemma is not only necessary but also sufficient. To show this, we lift P to a Markov chain with a doubly stochastic probability transition matrix P'. So suppose we are given a vector $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ of positive rational numbers such that

$$\pi P = \pi.$$

For each $i = 1, 2, \ldots, n$, let

$$m_i = M\pi_i$$

where M is a suitably large integer so that every m_i is a positive integer. We construct an $N \times N$ doubly stochastic probability transition matrix $P' = [p'_{ij}]$ where $N = \sum m_i$. This $N \times N$ matrix is divided into n^2 blocks by partitioning the index set $\{1, 2, \ldots, N\}$ into n sections of lengths $m_1, m_2, \ldots, m_n$. The entries in the (k, l)th block are all equal to $\frac{p_{kl}}{m_l}$. Clearly, for each i with $m_1 + \cdots + m_{k-1} < i \leqslant m_1 + \cdots + m_{k-1} + m_k$ ($m_0 := 0$),

$$\sum_{j=1}^{N} p'_{ij} = \sum_{l=1}^{n} m_l \frac{p_{kl}}{m_l} = 1.$$

For each j with $m_1 + \cdots + m_{l-1} < j \leqslant m_1 + \cdots + m_{l-1} + m_l$,

$$\sum_{i=1}^{N} p'_{ij} = \sum_{k=1}^{n} m_k \frac{p_{kl}}{m_l} = \frac{1}{m_l} \sum_{k=1}^{n} m_k p_{kl} = 1,$$

https://doi.org/10.1017/S0963548304006352 Published online by Cambridge University Press

since $(m_1, m_2, \ldots, m_n)P = (m_1, m_2, \ldots, m_n)$. This shows that P' is a doubly stochastic probability transition matrix of a Markov chain with state space, say, $\{1, 2, \ldots, N\}$. We have constructed a Markov chain with a uniform stationary distribution on a big state space $\{1, 2, \ldots, N\}$ which can be thought of as the disjoint union of n sets $T_1, T_2, \ldots, T_n$ where each T_i has m_i elements. From an element in T_i we go to an element in T_j with probability p_{ij}/m_j.

It is easy to see that P can be lifted to P' by defining $L' : \{1, 2, \ldots, N\} \to \{1, 2, \ldots, n\}$ to be

$$L'(i) = k \quad \text{where } m_1 + \cdots + m_{k-1} < i \leqslant m_1 + \cdots + m_{k-1} + m_k.$$

Furthermore, by Lemma 2.1, P' itself can be lifted to a random walk on S_N by the map $L'' : S_N \to \{1, 2, \ldots, N\}$ where $L''(\sigma) = \sigma(1)$ for $\sigma \in S_N$. Now the composition $L' \circ L'' : S_N \to \{1, 2, \ldots, n\}$ is a lifting of P to a random walk on S_N. Thus we have shown the following result.

Theorem 2.3. *Let P be an $n \times n$ probability transition matrix. The Markov chain associated with P can be lifted to a random walk on a finite group G if and only if P has a stationary probability distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ where each π_i is a positive rational number.*

Remarks. **(1)** Lemma 2.2 gives explicit information on $|G|$, the size of the group into which the Markov chain is lifted. Given a stationary probability distribution π, let $(m_1, m_2, \ldots, m_n)$ be a sequence of positive integers with $\gcd(m_1, m_2, \ldots, m_n) = 1$ (where gcd denotes the greatest common divisor) such that

$$m_1 : m_2 : \ldots : m_n = \pi_1 : \pi_2 : \ldots : \pi_n.$$

Then $|G|$ is a positive integral multiple of $\sum_i m_i$.

(2) Note that in the construction of the doubly stochastic matrix P', the size of the group where the lifting takes place is $|S_N| = N!$, which can be rather large. In general, it may be possible to lift the Markov chain to random walk on a smaller group, as in the next example.

Example. Consider the Markov chain with state space $\{1, 2\}$ and probability transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

The stationary probability distribution is $\pi = (\frac{1}{3}, \frac{2}{3})$.

Following the construction described to obtain P', we have

$$P' = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \end{pmatrix}.$$

https://doi.org/10.1017/S0963548304006352 Published online by Cambridge University Press

Observe that from Birkhoff's theorem P' has the decomposition

$$P' = \frac{1}{8}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + \frac{1}{4}\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix} + \frac{1}{4}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix} + \frac{3}{8}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}.$$

This shows that P lifts to a random walk on S_3 with probability measure μ given by $\mu((1)) = \frac{1}{8}$, $\mu((12)) = \mu((13)) = \frac{1}{4}$, $\mu((23)) = \frac{3}{8}$ and $\mu((123)) = \mu((132)) = 0$. The lifting from P to the random walk on S_3 is effected by $L : S_3 \to \{1,2\}$, where $L^{-1}(1) = \{(1),(23)\}$ and $L^{-1}(2) = \{(12),(13),(123),(132)\}$.

Note that in this lifting, μ is not a class function and the support of μ does not generate an abelian subgroup in S_3.

However, we can do better. In fact, P admits a lifting to $\mathbb{Z}_3 = \{0,1,2\}$ which is an abelian group by $\hat{L} : \mathbb{Z}_3 \to \{1,2\}$ where $\hat{L}^{-1}(1) = \{0\}$ and $\hat{L}^{-1}(2) = \{1,2\}$. The probability measure $\hat{\mu}$ on $\mathbb{Z}_3$ is given by $\hat{\mu}(0) = \frac{1}{2}$ and $\hat{\mu}(1) = \hat{\mu}(2) = \frac{1}{4}$.

To conclude this note, we observe that it is of interest to determine when a Markov chain can be lifted to a random walk on an abelian group or a group whose probability measure is a class function. In these cases, the Fourier analysis using group representations simplifies to computations involving the irreducible characters of the group.

Acknowledgements

The authors would like to thank the referee for helpful suggestions that result in a more succinct proof of Lemma 2.2 and improved exposition of the article overall.

References

[1] Diaconis, P. (1988) *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics, Hayward, CA.
[2] Diaconis, P. and Shahshahani, M. (1987) Time to reach stationarity in the Bernoulli–Laplacian diffusion model. *SIAM J. Math. Anal.* **18** 208–218.
[3] Heller, A. (1965) On stochastic processes derived from Markov chains. *Ann. Math. Statist.* **36** 1286–1291.
[4] Horn, R. A. and Johnson, C. R. (1992) *Matrix Analysis*, Cambridge University Press.
[5] Howe, E. C. and Johnson, C. R. (1989) Aggregation of Markov processes: Axiomatization. *J. Theoret. Probab.* **2** 201–208.
[6] Kac, M. (1947) Random walk and the theory of Brownian motion. *Amer. Math. Monthly* **54** 369–391.

https://doi.org/10.1017/S0963548304006352 Published online by Cambridge University Press