

A note on the passage time of finite state Markov chains¹

Wenming Hong² Ke Zhou³

Abstract

Consider a Markov chain with finite state $\{0, 1, \dots, d\}$. We give the generation functions (or Laplace transforms) of absorbing (passage) time in the following two situations : (1) the absorbing time of state d when the chain starts from any state i and absorbing at state d ; (2) the passage time of any state i when the chain starts from the stationary distribution supposed the chain is time reversible and ergodic. Example shows that it is more convenient compared with the existing methods, especially we can calculate the expectation of the absorbing time directly.

Keywords: Markov chain, absorbing time, passage time, generation functions, Laplace transforms, eigenvalues, stationary distribution.

Mathematics Subject Classification (2010): 60E10, 60J10, 60J27, 60J35.

1 Discrete time

1.1 Absorbing time when the process starting from any state i

Consider the discrete time Markov chain $\{X_n\}_{n \geq 0}$ with finite states $\{0, 1, \dots, d\}$ and absorbing at state d , the transition probability matrix P is given by

$$P = \begin{pmatrix} r_0 & p_{0,1} & \cdots & p_{0,d-1} & p_{0,d} \\ q_{1,0} & r_1 & \cdots & p_{1,d-1} & p_{1,d} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ q_{d-1,0} & q_{d-1,1} & \cdots & r_{d-1} & p_{d-1,d} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}_{(d+1) \times (d+1)}.$$

For $0 \leq i \leq d$, let $\tau_{i,d}$ be the absorbing time of state d starting from i , i.e.,

$$\tau_{i,d} := \inf\{n \geq 1, X_n = d \text{ when } X_0 = i\},$$

and $f_i(s)$ be the generation function of $\tau_{i,d}$,

$$f_i(s) = \mathbb{E}s^{\tau_{i,d}} \quad \text{for } 0 \leq i \leq d, \quad (1.1)$$

we have

¹ The project is partially supported by NSFC Grant No.11131003.

²School of Mathematical Sciences & Laboratory of Mathematics and Complex Systems, Beijing Normal University, Beijing 100875, P.R. China. Email: wmhong@bnu.edu.cn

³ School of Statistics, University of International Business and Economics, Beijing 100029, P.R. China. Email: zhouke@uibe.edu.cn

Theorem 1.1. For $1 \leq j \leq d+1$, denote $A_j(s)$ as the $d \times d$ sub-matrix by deleting the $(d+1)^{\text{th}}$ row and the j^{th} column of the matrix $I_{d+1} - sP$. Then, for $0 \leq i \leq d$, we have

$$f_i(s) = (-1)^{d-i} \frac{\det A_{i+1}(s)}{\det A_{d+1}(s)}. \quad (1.2)$$

□

Remark 1.1. As a consequence (see Corollary 1.1 below), for the birth-death and more general the skip-free (upward jumps is only of unit size, and there is no restriction on downward jumps) Markov chain with finite state $\{0, 1, \dots, d\}$ and absorbing at state d , the absorbing time is distributed as a summation of d independent geometric (or exponential) random variables.

There are many authors give out different proofs to the results. For the birth and death chain, the well-known results can be traced back to Karlin and McGregor ([10], 1959) Keilson ([11], 1971; [12]). Kent and Longford ([13], 1983) proved the result for the discrete time version (nearest random walk) although they have not specified the result as usual form (section 2, [13]). Fill ([5], 2009) gave the first stochastic proof to both nearest random walk and birth and death chain cases via duality which was established in [3]. Diaconis and Miclo ([4], 2009) presented another probabilistic proof for birth and death chain. Gong et al ([8], 2012) gave a similar result in the case that the state space is \mathbb{Z}^+ . For the skip-free chain, Brown and Shao ([2], 1987) first proved the result in continuous time situation; Fill ([6], 2009) gave a stochastic proof to both discrete and continuous time cases also by using the duality, and considered the general finite-state Markov chain situation when the chain starts from state 0.

However, the existing proofs we mentioned above are heavily relied on the initial state being 0, no matter the “analysis” method by Brown and Shao ([2], 1987) and the “stochastic” method by Fill ([6], 2009), etc.. The first purpose of this paper (Theorem 1.1 and 2.1) is to improve the result to the general situation: the chain starts from any state i (not just from state 0 only ([6], 2009)). In particular, the results generalize the well-known theorems for the birth-death (Karlin and McGregor [10], 1959) and the skip-free ([2] and [6]) Markov chain.

Before proving the Theorem, let us at first to recover the results for the skip-free (and then the birth-death) discrete time Markov chain (Fill [6], 2009).

Corollary 1.1. Assume $p_{i,j} = 0$ for $j - i > 1$. We have

$$f_0(s) = \prod_{i=0}^{d-1} \left[\frac{(1 - \lambda_i)s}{1 - \lambda_i s} \right], \quad (1.3)$$

where $\lambda_0, \dots, \lambda_{d-1}$ are the d non-unit eigenvalues of P .

In particular, if all of the eigenvalues are real and nonnegative, then the hitting time is distributed as the sum of d independent geometric random variables with parameters $1 - \lambda_i$.

Proof. Note that, 1 is an eigenvalue of P evidently. So on the one hand $\det(I_{d+1} - sP) = (1 - s) \prod_{i=0}^{d-1} (1 - \lambda_i s)$ (where $\lambda_0, \dots, \lambda_{d-1}$ are the d non-unit eigenvalues of P); on the other hand we have $\det(I_{d+1} - sP) = (1 - s) \times \det A_{d+1}(s)$ from (1.2); it's trivial to show that

$$\det A_{d+1}(s) = \prod_{i=0}^{d-1} (1 - \lambda_i s). \quad (1.4)$$

From the definition of A_1 , it is easy to see

$$\det A_1(s) = (-1)^d p_{0,1} p_{1,2} \cdots p_{d-1,d} s^d. \quad (1.5)$$

By (1.2) and (1.4) we have

$$\det A_1(1) = (-1)^d f_0(1) \cdot \det A_{d+1}(1) = (-1)^d f_0(1) \cdot \prod_{i=0}^{d-1} (1 - \lambda_i).$$

And from (1.5), we get $\det A_1(1) = (-1)^d p_{0,1} p_{1,2} \cdots p_{d-1,d}$. Recall that $f_0(1) = 1$, by (1.1), we obtain

$$p_{0,1} p_{1,2} \cdots p_{d-1,d} = \prod_{i=0}^{d-1} (1 - \lambda_i). \quad (1.6)$$

Then by (1.5) and (1.6)

$$\det A_1(s) = (-1)^d \prod_{i=0}^{d-1} (1 - \lambda_i) s^d, \quad (1.7)$$

and (1.3) holds from (1.4) and (1.7) directly. \square

Remark 1.2. *The following example shows that Theorem 1.1 is more convenient compared with the existing methods in Corollary 1.1, especially we can calculate the expectation of the absorbing time directly from (1.2).*

Consider a Markov chain with $d + 1$ states $\{0, 1, 2, \dots, d\}$ whose transition matrix P can be given by

$$P = \begin{pmatrix} q & p & \cdots & 0 & 0 \\ q & 0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ q & 0 & \cdots & 0 & p \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}_{(d+1) \times (d+1)},$$

where $p + q = 1$.

Corollary 1.2. *For $0 \leq i \leq d$,*

$$f_i(s) = \frac{p^{d-i} s^{d-i} (1 - s) + p^d q s^{d+1}}{1 - s + p^d q s^{d+1}}, \quad (1.8)$$

and we have

$$\mathbb{E}\tau_{i,d} = \frac{1 - p^{d-i}}{p^d q}. \quad (1.9)$$

Proof. Take full advantage of $p + q = 1$, we can calculate that

$$\det A_{i+1}(s) = (-1)^{d-i} \frac{p^{d-i} s^{d-i} (1 - s) + p^d q s^{d+1}}{1 - ps},$$

$$\det A_{d+1}(s) = \frac{1 - s + p^d q s^{d+1}}{1 - ps}.$$

We obtain (1.8) by using Theorem 2.1. Recall that $\mathbb{E}\tau_{i,d} = f'_i(1)$, we can get (1.9) by some calculation easily. \square

Proof of Theorem 1.1 By decomposing the first step, for $0 \leq i \leq d$, the generation function of $\tau_{i,d}$ satisfies,

$$f_i(s) = r_i s f_i(s) + p_{i,i+1} s f_{i+1}(s) + p_{i,i+2} s f_{i+2}(s) + \cdots + p_{i,d-1} s f_{d-1}(s) + p_{i,d} s q_{i,i-1} s f_{i-1}(s) + q_{i,i-2} s f_{i-2}(s) + \cdots + q_{i,0} s f_0(s).$$

These system of equations are linear with respect to $f_0(s), f_1(s), \dots, f_{d-1}(s)$. Using Cramer's Rule, we can get (1.2) by solving from these equations. \square

1.2 Passage time when starting from the stationary distribution

Consider a discrete time Markov chain $\{X_n\}_{n \geq 1}$ with finite states $\{0, 1, \dots, d\}$ starting from the stationary distribution $\pi := (\pi_0, \pi_1, \dots, \pi_d)$, the transition probability matrix \hat{P} is given by

$$\hat{P} = \begin{pmatrix} r_0 & p_{0,1} & \cdots & p_{0,d-1} & p_{0,d} \\ q_{1,0} & r_1 & \cdots & p_{1,d-1} & p_{1,d} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ q_{d,0} & q_{d,1} & \cdots & q_{d,d-1} & r_d \end{pmatrix}_{(d+1) \times (d+1)}.$$

In addition, write

$$D = \begin{pmatrix} 1 & -\pi_1 & \cdots & -\pi_{d-1} & -\pi_d \\ 0 & 1 - r_1 s & \cdots & -p_{1,d-1} s & -p_{1,d} s \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & -q_{d,1} s & \cdots & -q_{d,d-1} s & 1 - r_d s \end{pmatrix}, \quad (1.10)$$

and

$$D_0 = \begin{pmatrix} \pi_0 & -\pi_1 & \cdots & -\pi_{d-1} & -\pi_d \\ q_{1,0} s & 1 - r_1 s & \cdots & -p_{1,d-1} s & -p_{1,d} s \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ q_{d,0} s & -q_{d,1} s & \cdots & -q_{d,d-1} s & 1 - r_d s \end{pmatrix}. \quad (1.11)$$

Theorem 1.2. *When the chain starts from X_0 with the stationary distribution π , and*

$$\tau := \inf\{n \geq 0, X_n = 0\}, \quad (1.12)$$

be the passage time of state 0. Denote $g_\pi(s)$ as the generation function of τ , i.e., $g_\pi(s) = \mathbb{E}_\pi s^\tau$; we have

$$g_\pi(s) = \frac{\det D_0}{\det D},$$

where D and D_0 are given in (1.10) and (1.11) respectively.

Specifically, if the chain is time reversible and ergodic, Brown([1], 1999) point out the elegant connection between the passage time starting from the stationary distribution and the interlacing eigenvalues theorem of linear algebra. Recently, this result also proved by Fill and Lyzinski([7], 2013) with a stochastic method. In what follows, we will reprove it directly as a corollary of Theorem 1.2.

Corollary 1.3. *Let $\lambda_1, \dots, \lambda_d$ be the d non-unit eigenvalues of \widehat{P} (we assume $\lambda_0 = 1$), and $\gamma_1, \dots, \gamma_d$ be the d eigenvalues of \widehat{P}_0 , which is the sub-matrix obtained by deleting the first row and the first column of \widehat{P} . Then we have*

$$g_\pi(s) = \left(\prod_{i=1}^d \frac{1 - \gamma_i}{1 - \lambda_i} \right) \prod_{i=1}^d \left(\frac{1 - \lambda_i s}{1 - \gamma_i s} \right).$$

Proof of Corollary 1.3 It is easy to see that (recall $\gamma_1, \dots, \gamma_d$ are the eigenvalues of \widehat{P}_0),

$$\det D = \prod_{i=1}^d (1 - \gamma_i s). \quad (1.13)$$

In what follows we will show

$$\det D_0 = \pi_0 \prod_{i=1}^d (1 - \lambda_i s). \quad (1.14)$$

Define $e_1 = (1, 0, \dots, 0)$, and recall that $\pi = (\pi_0, \pi_1, \dots, \pi_d)$. Then, we have

$$\det D_0 = \begin{vmatrix} 0 & \pi \\ e_1^T & I - s\widehat{P} \end{vmatrix}.$$

If we let $\Pi = \text{diag}(\pi_0, \pi_1, \dots, \pi_d)$, the reversibility of \widehat{P} implies that $\Pi^{1/2}\widehat{P}\Pi^{-1/2}$ is a real symmetric matrix. Thus there exist an orthogonal matrix U such that

$$U\Pi^{1/2}\widehat{P}\Pi^{-1/2}U^T = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_d). \quad (1.15)$$

We can calculate that

$$\begin{aligned} & \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \Pi^{1/2} \end{pmatrix} \begin{pmatrix} 0 & \pi \\ e_1^T & I - s\widehat{P} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \Pi^{-1/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U^T \end{pmatrix} \\ &= \begin{pmatrix} 0 & \pi\Pi^{-1/2}U^T \\ U\Pi^{1/2}e_1^T & U\Pi^{1/2}(I - s\widehat{P})\Pi^{-1/2}U^T \end{pmatrix} = \begin{pmatrix} 0 & \pi\Pi^{-1/2}U^T \\ U\Pi^{1/2}e_1^T & I - s\text{diag}(\lambda_0, \lambda_1, \dots, \lambda_d) \end{pmatrix} \end{aligned}$$

It is easy to prove that $\lambda_0 = 1$ is the unique eigenvalue of maximum modulus of \widehat{P} . So the geometric multiplicity of \widehat{P} corresponding to λ_0 is one ([10] P500 Perron's Theorem). On the one hand, $e_1 U \Pi^{1/2}$ is a left eigenvector corresponding to λ_0 . π is also the left eigenvector of λ_0 . Because $\|e_1 U \Pi^{1/2}\| = \|\pi\| = 1$, we have $e_1 U \Pi^{1/2} = \pi$. So

$$U\Pi^{1/2}e_1^T = (e_1 U \Pi^{1/2})^T = \pi^T. \quad (1.16)$$

On the other hand, $\Pi^{-1/2}U^T e_1^T$ is a right eigenvector corresponding to eigenvalue λ_0 , and $\mathbf{1} = \{1, 1, \dots, 1\}$ is also the right eigenvector of λ_0 . Because $\|\Pi^{-1/2}U^T e_1^T\| = \|\mathbf{1}\| = 1$, we have $\Pi^{-1/2}U^T e_1^T = \mathbf{1}$, and

$$\pi \Pi^{-1/2}U^T e_1^T = \pi \mathbf{1} = 1 \quad (1.17)$$

By (1.15) and $\pi = \pi \widehat{P}$,

$$\pi \Pi^{-1/2}U^T = \pi \widehat{P} \Pi^{-1/2}U^T = \pi \Pi^{-1/2}U^T \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_d).$$

Because $\lambda_i \neq 1$ for $i = 1, 2, \dots, d$, by (1.17), $\pi \Pi^{-1/2}U^T$ must be equal to e_1 . By (1.16), we obtain

$$\det D_0 = \begin{vmatrix} 0 & \pi \Pi^{-1/2}U^T \\ U \Pi^{1/2} e_1^T & I - s \widehat{P} \end{vmatrix} = \begin{vmatrix} 0 & e_1 \\ \pi^T & I - s \widehat{P} \end{vmatrix} = \pi_0 \prod_{i=1}^d (1 - \lambda_i s),$$

which we get (1.14). Combine (1.13) and (1.14), we obtain

$$g_\pi(s) = \frac{\pi_0 \prod_{i=1}^d (1 - \lambda_i s)}{\prod_{i=1}^d (1 - \gamma_i s)}.$$

Because $g_\pi(s)$ is a generation function, $g_\pi(1) = 1$. So

$$\pi_0 = \frac{\prod_{i=1}^d (1 - \gamma_i)}{\prod_{i=1}^d (1 - \lambda_i)},$$

which complete the proof. \square

Proof of Theorem 1.2 Denote $g_i(s)$ as the generation function of passage time of state 0 when the chain is starting from i . By the Markov property, we have

$$g_\pi(s) = \pi(0)g_0(s) + \pi(1)g_1(s) + \dots + \pi(d)g_d(s). \quad (1.18)$$

Obviously, $g_0(s) = 1$. By decomposing the first step, for $1 \leq i \leq d$, $g_i(s)$ satisfies

$$\begin{aligned} g_i(s) &= r_i s g_i(s) + p_{i,i+1} s g_{i+1}(s) + \dots + p_{i,d-1} s g_{d-1}(s) + p_{i,d} s g_d(s) \\ &\quad + q_{i,i-1} s g_{i-1}(s) + q_{i,i-2} s g_{i-2}(s) + \dots + q_{i,0} s. \end{aligned}$$

These system of equations together with (1.18) are linear with respect to $g_\pi(s), g_1(s), g_2(s), \dots, g_d(s)$. Use Cramer's Rule, we can get $g_\pi(s)$ by solving from these equations as

$$g_\pi(s) = \frac{\det D_0}{\det D}.$$

\square

Remark 1.3. Actually, if we define for $i = 1, 2, \dots, d$

$$\tau_i := \inf\{n \geq 0, X_n = i\}, \quad (1.19)$$

be the passage time of state i . Denote $g_\pi^i(s)$ as the generation function of τ_i , i.e., $g_\pi^i(s) = \mathbb{E}_\pi s^{\tau_i}$, we can obtain the formula for $g_\pi^i(s)$ with the corresponding modification for the D and D_0 , the proof is almost line by line with regard of $g_i(s) = 1$ this time.

2 Continuous time

We can write the counterpart results for the continuous time Markov chain with finite states $\{0, 1, \dots, d\}$ easily. The proof is similar as in section 1 and so we omit the details.

2.1 Starting from any fixed state

Define $\{X_t\}_{t \geq 0}$ being the (continuous time) Markov chain with finite states $\{0, 1, \dots, d\}$ and absorbing at state d , the generator Q is given by

$$Q = \begin{pmatrix} -\gamma_0 & \alpha_{0,1} & \cdots & \alpha_{0,d-1} & \alpha_{0,d} \\ \beta_{1,0} & -\gamma_1 & \cdots & \alpha_{1,d-1} & \alpha_{1,d} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \beta_{d-1,0} & \beta_{d-1,1} & \cdots & -\gamma_{d-1} & \alpha_{d-1,d} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{(d+1) \times (d+1)}.$$

Let $\tau_{i,d}$ be the absorbing time of state d starting from i and $\tilde{f}_i(s)$ be the Laplace transform of $\tau_{i,d}$. i.e.

$$\tilde{f}_i(s) = \mathbb{E}e^{-s\tau_{i,d}}.$$

Theorem 2.1. For $1 \leq j \leq d+1$, we denote $\tilde{A}_j(s)$ as the $d \times d$ sub-matrix by deleting the $(d+1)^{th}$ row and the j^{th} column of the matrix $sI_{d+1} - Q$. Then, for $0 \leq i \leq d$ we have

$$\tilde{f}_i(s) = (-1)^{d-i} \frac{\det \tilde{A}_{i+1}}{\det \tilde{A}_{d+1}}. \quad (2.1)$$

□

Immediately, we recover the results for the skip-free continuous time Markov chain (Brown and Shao [2], 1987).

Corollary 2.1. Assume $\alpha_{i,j} = 0$ for $j - i > 1$. We have

$$\varphi_d(s) = \prod_{i=0}^{d-1} \frac{\lambda_i}{\lambda_i + s},$$

where λ_i are the d non-zero eigenvalues of $-Q$.

In particular, if all of the eigenvalues are real and nonnegative, then the hitting time is distributed as the sum of d independent exponential random variables with parameters λ_i .

Proof. The proof is similar as Corollary 1.1, we can calculate that $\det \tilde{A}_{d+1} = \prod_{i=0}^{d-1} (\lambda_i + s)$, and $\det \tilde{A}_1 = (-1)^d \alpha_{0,1} \alpha_{1,2} \cdots \alpha_{d-1,d} = (-1)^d \prod_{i=0}^{d-1} \lambda_i$. □

2.2 Starting from the stationary distribution

If we consider a time reversible ergodic Markov chain with generator \widehat{Q} , let \widehat{Q}_0 be the sub-matrix which is obtained by deleting the first row and the first column of \widehat{Q} . We denote $\tilde{g}_\pi(s)$ as the Laplace transform of the hitting time of state 0 when the chain is starting from the stationary distribution π .

Theorem 2.2. *We have*

$$\tilde{g}_\pi(s) = \left(\prod_{i=1}^d \frac{\gamma_i}{\lambda_i} \right) \prod_{i=1}^d \left(\frac{\lambda_i + s}{\gamma_i + s} \right), \quad (2.2)$$

where $\lambda_1, \dots, \lambda_d$ are the d non-zero eigenvalues of $-\widehat{Q}$ (we assume $\lambda_0 = 0$), and $\gamma_1, \dots, \gamma_d$ are the d eigenvalues of $-\widehat{Q}_0$.

Acknowledgement

We appreciate Professor Daniel R. Jeske for his interesting example (see Corollary 1.2). We would also thank the referee's valuable suggestions.

References

- [1] Brown, M. Interlacing Eigenvalues in Time Reversible Markov Chains. *Math. Oper. Res.* **24**(4) (1999), 847–864.
- [2] Brown, M. and Shao, Y. S. Identifying coefficients in the spectral representation for first passage time distributions. *Probab. Eng. Inform. Sci.* **1** (1987), 69–74.
- [3] Diaconis, P. and Fill, J. A. Strong stationary times via a new form of duality. *Ann. Probab.* **18** (1990), 1483–1522.
- [4] Diaconis, P. and Miclo, L. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.* **22** (2009), 558–586.
- [5] Fill, J. A. The passage time distribution for a birth-and-death chain: Strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.* **22** (2009), 543–557.
- [6] Fill, J. A. On hitting times and fastest strong stationary times for skip-free and more general chains. *J. Theoret. Probab.* **22** (2009), 587–600.
- [7] Fill, J. A. and Lyzinski, V. Hitting Times and Interlacing Eigenvalues: A Stochastic Approach Using Intertwinings. to appear in *J. Theoret. Probab.*
- [8] Gong, Y. Mao, Y. H. and Zhang, C. Hitting time distributions for denumerable birth and death processes. *J. Theoret. Probab.* **25** (2012), 950–980.
- [9] Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, Cambridge, 1985.
- [10] Karlin, S. and McGregor, J. Coincidence properties of birth and death processes. *Pacific J. Math.* **9** (1959), 1109–1140.

- [11] Keilson, J. Log-concavity and log-convexity in passage time densities for of diffusion and birth-death processes. *J. Appl. Probab.* **8** (1971), 391–398.
- [12] Keilson, J. *Markov Chain Models—Rarity and Exponentiality*. Springer, New York, 1979.
- [13] Kent, J. T. and Longford, N. T. An eigenvalue decomposition for first hitting times in random walk. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **63** (1983), 71–84.
- [14] Zhou, K. Hitting Time Distribution for Skip-Free Markov Chains: A Simple Proof. *Statist. Probab. Lett.* **83(7)** (2013), 1782–1786.