Applying Brownian motion to the study of birth-death chains.

Greg Markowsky
gmarkowsky@gmail.com
(054)279-5828
Pohang Mathematics Institute
POSTECH
Pohang, 790-784
Republic of Korea

Abstract
Basic properties of Brownian motion are used to derive two results concerning birth-death chains. First, the probability of extinction is calculated. Second, sufficient conditions on the transition probabilities of a birth-death chain are given to ensure that the expected value of the chain converges to a limit. The theory of Brownian motion local time figures prominently in the proof of the second result.

Keywords: Birth-death chain, Markov chain, Brownian motion, local time.

1. Introduction
Let $X_m$ be a Markov chain taking values on the nonnegative integers with the following transition probabilities for $n \neq 0$

$$p_{nj} = \begin{cases} r_n & \text{if } j = n + 1 \\ l_n & \text{if } j = n - 1 \\ 0 & \text{if } |n - j| \neq 1. \end{cases}$$

Implicit here is the fact that $r_n + l_n = 1$. We suppose further for simplicity that $X_0 = k$ almost surely, for some $k \in \mathbb{N}$. $X_m$ is essentially a random walk on the nonnegative integers, moving to the right from state $n$ with probability $r_n$ and to the left with probability $l_n$. We refer to such a Markov chain as a birth-death chain. This name comes from considering $X_m$ as the number of
members in a population, where at each step either a new member is born or an old member dies, causing the process to increase or decrease by 1. We can assume \( p_{00} = 1 \) and \( p_{0j} = 0 \) for any \( j \neq 0 \), as when the population reaches 0 it is considered to have gone extinct with no possibility of regeneration. The purpose of this paper is to introduce a method of using properties of Brownian motion to deduce two fundamental theorems concerning birth-death chains. The first theorem, presented in the next section, gives the probability that a birth-death chain goes extinct at some finite time. The second theorem, presented in Section 3, gives sufficient conditions for \( E[X_m] \) to converge as \( m \to \infty \). The properties of Brownian motion which will be utilized are standard and can be found in many references on Brownian motion, such as [2] or [4].

We will now introduce the basic setup. Let \( t_0 := 1 \) and

\[
(2) \quad t_n := \frac{l_1 l_2 \ldots l_n}{r_1 r_2 \ldots r_n}
\]

for \( n > 0 \). Define a sequence \( \{x_n\}_{n=0}^{\infty} \) recursively by setting \( x_0 = 0 \), and having defined \( x_n \) let \( x_{n+1} = x_n + t_n \). Since the sequence \( \{x_n\} \) is increasing it converges to a limit \( x_\infty \), possibly infinite, as \( n \to \infty \). Let \( B_t \) be a Brownian motion starting at \( x_k \) and stopped at the first time \( T_\Delta \) it hits 0 or \( x_\infty \). The recurrence properties of Brownian motion imply that \( T_\Delta < \infty \) almost surely. We define a sequence of stopping times \( T_m \) which are, roughly speaking, the successive hitting times of \( A := \{x_n\}_{n=0}^{\infty} \). More rigorously, \( T_m \) is defined recursively by setting \( T_0 = 0 \), and having defined \( T_m \) we let \( T_{m+1} = \inf_{t>T_m} \{B_t \in A, B_t \neq B_{T_m}\} \). We see that the variables \( B_{T_0}, B_{T_1}, B_{T_2}, \ldots \) form a random process taking values in \( A \). The strong Markov property of Brownian motion, together with the standard exit distribution of Brownian motion from an interval, imply that

\[
P(B_{T_{m+1}} = x_{n+1} | B_{T_m} = x_n) = \frac{x_n - x_{n-1}}{x_{n+1} - x_{n-1}} = \frac{t_{n-1}}{t_{n-1} + t_n} = \frac{1}{1 + t_n/r_n} = r_n
\]

and, likewise,

\[
P(B_{T_{m+1}} = x_{n-1} | B_{T_m} = x_n) = l_n
\]
If we define $\phi$ on $\{x_n\}_{n=0}^{\infty}$ by $\phi(x_n) = n$, we see that $\phi(B_{T_0}), \phi(B_{T_1}), \phi(B_{T_2}), \ldots$ is a realization of our original birth-death chain. The picture below gives an example, where we have oriented the time axis vertically and the space axis horizontally.

Figure 1: The Brownian path pictured realizes the birth-death path $k, k+1, k, k+1, k, k-1, k, k-1, k-2, \ldots$

Given this framework, we are ready to prove several theorems. In the sequel, any reference to $X, B, x_n, T, \Delta, \phi$, etc. will refer to the definitions presented in this section.

2. The extinction probability of a birth-death chain

Perhaps the most fundamental question one can ask regarding a birth-death chain is whether the population must go extinct or not, that is, whether $P(X_m = 0 \text{ for some } m) = 1$ or $P(\lim_{m \to \infty} X_m = +\infty) > 0$. Let $P_k$ be the probability that the birth-death chain eventually hits 0 (recall $X_0 = k$ a.s.). We then have the following.

Theorem 1.

$$P_k = \frac{\sum_{j=k}^{\infty} t_j}{\sum_{j=0}^{\infty} t_j}$$
where this quotient is interpreted as being equal to 1 if the sums diverge.

This elegant theorem has a straightforward proof using recurrence relations; see [3] or [5]. A potentially pleasing aspect of the proof below, however, lies in giving a clear, visual intuition for the sums in (3).

**Proof of Theorem 1:** Recall that $x_\infty = \lim_{n \to \infty} x_n$ is given by

$$x_\infty = \sum_{j=0}^{\infty} t_j$$

(4)

If $x_\infty = \infty$, so that both sums in (3) diverge, then $B_{T_\Delta} = 0$ almost surely. This implies that the population dies out with probability 1. On the other hand, if $x_\infty < \infty$ then $P(B_{T_\Delta} = 0)$ is given by

$$\frac{x_\infty - x_k}{x_\infty - 0} = \sum_{j=k}^{\infty} \frac{t_j}{\sum_{j=0}^{\infty} t_j}$$

(5)

However, as in the first case, $P(B_{T_\Delta} = 0)$ is precisely $P_k$, the probability of extinction. This is because the Brownian motion hitting $x_\infty$ before 0 implies $B_{T_m} \to x_\infty$, hence $\phi(B_{T_m}) \to \infty$, whereas hitting 0 before $x_\infty$ implies $\phi(B_{T_\Delta}) = 0$ for some $m$. The two cases ($x_\infty = \infty$ and $x_\infty < \infty$) are illustrated in the following figure.
Figure 2: The left panel illustrates the situation in which $\sum_{j=0}^{\infty} t_j$ diverges. Eventually, $B_t$ hits 0 and the population goes extinct. The right panel illustrates the other scenario, in which $\sum_{j=0}^{\infty} t_j = x_\infty < \infty$. In this case, there is a positive probability that $B_t$ hits $x_\infty$ before 0, in which case the population never goes extinct.

This completes the proof of Theorem 1.

\[ \square \]

3. The long-term average of a birth-death chain

Recall that $B$ and $X$ are stopped upon reaching 0. It will therefore be convenient to let $X_m$ be defined to be 0 for all $m > m_0$, where $m_0$ is the smallest integer, if it exists, for which $X_{m_0} = 0$. Similarly, for convenience let $T_m = T_\Delta$ for all $m > m_0$, where $m_0$ is the smallest integer, if it exists, for which $B_{T_m} = 0$. In the case $r_i = l_i = \frac{1}{2}$ for all $i$, it is well known that $X_m$ is a martingale, and therefore $E[X_m] = E_0 = k$ for all $m$. This occurs despite the fact that $P(X_m = 0) \rightarrow 1$ as $m \rightarrow \infty$, as the average value of $X_m$ on $\{X_m \neq 0\}$ grows at exactly the right speed to balance the set of large probability upon which $X_m = 0$. Such behavior certainly does not hold.
for the general case, since we no longer have the martingale property, but we will see that the Brownian motion model presented above can shed light on the behavior of $E[X_m]$ as $m \to \infty$.

Recall that $\mathcal{A} = \{x_n\}_{n=0}^{\infty}$. Let $\phi : \mathcal{A} \to \mathbb{R}^+$ be extended to a continuous function from $\mathbb{R}^+$ to $\mathbb{R}^+$ by defining $\phi$ to be linear on each interval $(x_{n-1}, x_n)$. Alternatively, we may think of $x_n = x(n)$ as a function from $\mathbb{N}$ to $\mathbb{R}$ which can be extended by linear interpolation to an increasing function from $\mathbb{R}^+$ to $\mathbb{R}^+$. In this case, $\phi$ is simply $x^{-1}$. $\phi$ is therefore a piecewise linear function, and $\phi'$ exists on $\mathbb{R}^+ - \mathcal{A}$. Let $\phi'_{n}$ be the value of $\phi'$ on $(x_{n-1}, x_n)$. We will prove the following theorem.

**Theorem 2.** If $\phi'_{\infty} = \lim_{n \to \infty} \phi'_{n}$ exists, then

$$\lim_{m \to \infty} E[X_m] = x_k \phi'_{\infty}$$

Note that we are allowing $\phi'_{\infty} = +\infty$ or 0. Writing $\phi'_{\infty}$ and $x_k$ in terms of the $l'_n$'s and $r_n$'s shows that the following statement is equivalent to Theorem 2.

If $t_{\infty} := \lim_{n \to \infty} \frac{l_1 \ldots l_k}{r_1 \ldots r_n}$ exists then $\lim_{m \to \infty} E[X_m]$ exists, and

$$\lim_{m \to \infty} E[X_m] = \frac{1 + \frac{l_1}{r_1} + \ldots + \frac{l_{k-1}}{r_1 \ldots r_{k-1}}}{t_{\infty}}$$

The bulk of the rest of this section is devoted to the proof of this theorem. We will simplify initially by assuming $\sum_{n=1}^{\infty} |\phi'_{n+1} - \phi'_{n}| < \infty$; this condition will be removed at the end of the proof. For the case in which there is a positive probability that the population never goes extinct, it is easy to see that $E[X_m] \to \infty$ as $m \to \infty$, and that $\phi'_{\infty}$ exists and is equal to $+\infty$, so that (6) is valid. We will therefore assume that $P(X_m = 0$ for some $m) = 1$.

Note that $\phi'_{n+1} = \frac{1}{t_n}$, and $x_{n+1} - x_n = t_n$, so that $\phi'_{n+1}(x_{n+1} - x_n) = 1$. Note also that $x_1 \phi'_1 = 1$. This allows us to perform the following manipulations to obtain an expression which will be more convenient for the purposes of the proof.

$$x_k \phi'_{\infty} = x_k (\phi'_{\infty} - \phi'_k) + x_k \phi'_k + k - x_1 \phi'_1 - \sum_{n=1}^{k-1} \phi'_{n+1}(x_{n+1} - x_n)$$
\[
= k + x_k (\phi'_{\infty} - \phi'_k) + \sum_{n=1}^{k-1} (\phi'_{n+1} - \phi'_n)x_n
\]

The last equality uses summation by parts; see [1]. We see that the conclusion of the theorem is equivalent to showing

\[
\lim_{m \to \infty} E[X_m] = k + x_k (\phi'_{\infty} - \phi'_k) + \sum_{n=1}^{k-1} (\phi'_{n+1} - \phi'_n)x_n
\]

This is what we will prove. We will proceed through several lemmas, and will need properties of Brownian motion local time, which is the density of the occupation measure of Brownian motion with respect to Lebesgue measure. That is, the local time \( L^x_t \) satisfies

\[
L^x_t dx = \int_0^t 1_{B_s \in dx} ds
\]

It is well known that \( L^x_t \) exists and that

\[
L^x_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1_{|B_s - x| < \varepsilon} ds
\]

almost surely. See [2] for a comprehensive treatment of local time, or the more general reference [4]. The following is an extension of Tanaka’s formula, Theorem VI.1.2 in [4].

**Lemma 1.** Almost surely, for any stopping time \( T \),

\[
\phi(B_T) = k + \int_0^T \phi'(B_s) dB_s + \sum_{n=1}^{\infty} \frac{(\phi'_{n+1} - \phi'_n)}{2} L^n_T
\]

where \( L^n_T \) denotes the local time of \( B_t \) at \( x_n \) at time \( T \).

**Proof:** Note that \( \phi''(x) = \sum_{n=1}^{\infty} (\phi'_{n+1} - \phi'_n) \delta_{x_n}(x) \) in the sense of distributions, where \( \delta_{x_n}(x) = \delta_0(x - x_n) \) denotes the Dirac delta function at point \( x_n \). Lemma [3] is therefore seen to be a special case of the Itô-Tanaka formula, Theorem VI.1.5 of [4], provided that \( \phi \) can be realized as a difference of two convex functions. However, any piecewise-linear function can be realized as the difference of two convex functions, provided that the points of nondifferentiability do not accumulate. We may argue as follows. \( \phi' \) is a piecewise
constant function, which is therefore of bounded variation on bounded intervals, and as such we may write $\phi' = f - g$ where $f$ and $g$ are nondecreasing. Let $F$ and $G$ be antiderivatives of $f$ and $g$ chosen so that $\phi = F - G$. Then $F$ and $G$ are convex, and the result follows.

Applying this lemma to the stopping time $T_k$, we immediately obtain

$$E[X_m] = k + \sum_{n=1}^{\infty} \frac{(\phi_{n+1}' - \phi_n')}{2} E[L_{T_m}^n],$$

Note that the convergence of the sum at this point is not an issue, since $L_{T_m}^n = 0$ for $n > m + k$. Using the identity $\lfloor 13 \rfloor$ does not seem to be an effective way to calculate $E[X_m]$, due to the difficulty of obtaining information about $T_m$. Nonetheless, we do know that $T_m \nearrow T_\Delta$ as $m \to \infty$, and this implies

$$\lim_{m \to \infty} E[X_m] = k + \sum_{n=1}^{\infty} \frac{(\phi_{n+1}' - \phi_n')}{2} E[L_{\infty}^n],$$

provided that $E[L_{\infty}^n]$ can be bounded uniformly, which we will show soon to be the case. We should mention that it was in obtaining $\lfloor 14 \rfloor$ that we used the assumption that $\sum_{n=1}^{\infty} |\phi_{n+1}' - \phi_n'| < \infty$. This is because a priori the quantities $E[L_{T_m}^n]$ may be growing in some strange way that causes problems if $\sum_{n=1}^{\infty} |\phi_{n+1}' - \phi_n'| = \infty$. We will return to this point at the end of the proof.

In light of $\lfloor 14 \rfloor$, we must compute $E[L_{\infty}^{x_n}]$.

**Lemma 2.**

$$E[L_{\infty}^{x_n}] = 2 \min(x_k, x_n)$$

**Proof:** One may derive this through standard calculations involving the probability density function of $B_t$, but the following is a quicker and easier proof. Let us suppose first that $n = k$. From Tanaka’s formula, $E[L_{\infty}^{x_k}] = \lim_{t \to \infty} E[|B_t - x_k|]$. Furthermore, $B_t$ is a martingale, so $E[(B_t - x_k)] = 0$ for all $t$. It follows from this that

$$E[L_{\infty}^{x_k}] = 2 \lim_{t \to \infty} E[\max(-(B_t - x_k), 0)]$$

$$= 2 \lim_{t \to \infty} \left( x_k P(B_t = 0) + \int_0^{x_k} (x_k - x) P(B_t \in dx) \right)$$
As \( \lim_{t \to \infty} P(B_t = 0) = 1 \), we can conclude that \( E[L_{\infty}^x] = 2x \). Now suppose \( n \neq k \). Note that, if we let \( T_{x_n} = \inf \{ t : B_t = x_n \} \), then

\[
L_{\infty}^x = L_{x_n} + L_{\infty}(B \circ \theta_{T_{x_n}}) \mathbb{1}_{T_{x_n} < T_\Delta} = L_{\infty}(B \circ \theta_{T_{x_n}}) 1_{T_{x_n} < T_\Delta}
\]

where \( \theta \) denotes the standard shift operator and \( L_t(B \circ \theta_{T_{x_n}}) \) is the local time of the shifted process \( B \circ \theta_{T_{x_n}} \). Let \( E_{x_j} \) denote expectation with respect to a Brownian motion \( W \) which starts at \( x_j \) and is stopped upon hitting 0.

The prior calculation together with (17) and the strong Markov property of Brownian motion imply that

\[
E[L_{\infty}^x] = P(T_{x_n} < T_\Delta) E_{x_n}[L_{\infty}^x]
\]

The general result follows from noting that \( P(T_{x_n} < T_\Delta) \) is 1 if \( x_n < x_k \) and \( \frac{x_n}{x_k} \) if \( x_n > x_k \).

Combining (14) and Lemma 2 gives

\[
\lim_{m \to \infty} E[X_m] = k + \sum_{n=1}^{k-1} (\phi'_n - \phi'_n) x_n + x_k \sum_{n=k}^{\infty} (\phi'_n - \phi'_n)
\]

Since \( \sum_{n=k}^{\infty} (\phi'_n - \phi'_n) = (\phi'_\infty - \phi'_k) \), we are done in this case. It remains only to remove the restriction that \( \sum_{n=1}^{\infty} |\phi'_n - \phi'_n| < \infty \). The following lemma is key.

**Lemma 3.** For any \( m \), and any \( n \geq k \), \( E[L_{T_{x_n}}^x] \geq E[L_{T_{x_n}}^x] \).

**Proof:** In fact, we may prove somewhat more, namely that if \( T \) is any stopping time with \( B_T \in A \) almost surely, then \( E[L_{T_{x_n}}^x] \geq E[L_{T_{x_n}}^x] \). Let \( E_{x_j} \) and \( W \) be as in the proof of Lemma 2. Using Lemma 2 and the strong Markov property of Brownian motion, we obtain

\[
E[L_{T_{x_n}}^x] = E[L_{T_{x_n}}^x] - E \left[ \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{T_{x_n}}^{\infty} 1_{(-\epsilon, \epsilon)}(B_s - x_n)ds \right]
\]

\[
= 2x_k - \sum_{j=1}^{\infty} P(B_T = x_j) E_{x_j} \left[ \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{0}^{\infty} 1_{(-\epsilon, \epsilon)}(W_s - x_n)ds \right]
\]

\[
= 2x_k - \sum_{j=1}^{n-1} P(B_T = x_j) x_j - \sum_{j=n}^{\infty} P(B_T = x_j) x_n
\]
Similarly,

\begin{equation}
E[L_{T_m}^{x_{n+1}}] = 2x_k - \sum_{j=1}^{n} P(B_T = x_j)x_j - \sum_{j=n+1}^{\infty} P(B_T = x_j)x_{n+1}
\end{equation}

The conclusion of the lemma now follows from the fact that \(x_{n+1} > x_n\). \(\Box\)

We may now complete the proof of the theorem. Recall (13), and observe that \(E[L_{T_m}^{x_n}] = 0\) for \(n > k + m\), since \(X_m \leq m + k\). This means that (13) is in fact a finite sum.

\begin{equation}
E[X_m] = k + \sum_{n=1}^{m+k} \frac{(\phi'_{n+1} - \phi'_n)}{2} E[L_{T_m}^{x_n}]
\end{equation}

\begin{equation}
= k + \sum_{n=1}^{k-1} \frac{(\phi'_{n+1} - \phi'_n)}{2} E[L_{T_m}^{x_n}] + \sum_{n=k}^{k+m} \frac{(\phi'_{n+1} - \phi'_n)}{2} E[L_{T_m}^{x_n}]
\end{equation}

The indices of the first sum in the final expression of (23) are independent of \(m\). This implies that the sum converges as \(m \to \infty\), since \(E[L_{T_m}^{x_n}] \to 2x_n\) as \(m \to \infty\) for \(n \leq k\). We must show that the second sum converges as \(m \to \infty\). We use summation by parts again, which gives

\begin{equation}
\sum_{n=k}^{k+m} \frac{(\phi'_{n+1} - \phi'_n)}{2} E[L_{T_m}^{x_n}] = \frac{1}{2} \left( \phi'_{k+m+1} E[L_{T_m}^{x_{k+m+1}}] - \phi'_k E[L_{T_m}^{x_k}] \right)
\end{equation}

\begin{equation}
- \sum_{n=k}^{k+m} \phi'_{n+1} (E[L_{T_m}^{x_{n+1}}] - E[L_{T_m}^{x_n}])
\end{equation}

Let us assume that \(\phi'_\infty < \infty\), and let \(\epsilon > 0\) be given. We may choose \(N > k\) such that \(\phi'_n \in (\phi_\infty - \epsilon, \phi_\infty + \epsilon)\) for all \(n \geq N\). Having chosen this, we may choose \(M > N - k\) such that \(2x_k \geq E[L_{T_m}^{x_n}] \geq 2x_k - \epsilon\) for all \(n \in [k, N], m \geq M\). Using the fact that \(E[L_{T_m}^{x_{k+m+1}}] = 0\), and setting \(\overline{\phi}' = \sup_{j>0} \phi'_j\), we see that for \(m > M\)

\begin{equation}
\sum_{n=k+1}^{k+m} \frac{(\phi'_{n+1} - \phi'_n)}{2} E[L_{T_m}^{x_n}]
\end{equation}
\[
\leq \frac{1}{2} \left( -\phi_k'(2x_k - \varepsilon) + \sum_{n=k}^{N} \phi_{n+1}'(E[L_{T_m}^{x_n}] - E[L_{T_m}^{x_n+1}]) \\
+ \sum_{n=N+1}^{k+m} \phi_{n+1}'(E[L_{T_m}^{x_n}] - E[L_{T_m}^{x_n+1}]) \right) \\
\leq \frac{1}{2} \left( -\phi_k'(2x_k - \varepsilon) + \overline{\phi}(E[L_{T_m}^{x_k}] - E[L_{T_m}^{x_{N+1}}]) \\
+ (\phi'_{\infty} + \varepsilon)(E[L_{T_m}^{x_{N+1}}] - E[L_{T_m}^{x_{k+m+1}}]) \right) \\
\leq \frac{1}{2} \left( -\phi_k'(2x_k - \varepsilon) + \overline{\phi}\varepsilon + (\phi'_{\infty} + \varepsilon)2x_k \right)
\]

This shows that

\[
\limsup_{m \to \infty} \sum_{n=k}^{k+m} \frac{\phi'_{n+1} - \phi'_n}{2} E[L_{T_m}^{x_n}] \leq x_k(\phi'_{\infty} - \phi'_k) \tag{25}
\]

Proceeding similarly, we can obtain

\[
\liminf_{m \to \infty} \sum_{n=k+1}^{k+m+1} \frac{\phi'_{n+1} - \phi'_n}{2} E[L_{T_m}^{x_n}] \geq x_k(\phi'_{\infty} - \phi'_k) \tag{26}
\]

Together, these prove the desired convergence. The case \(\phi'_{\infty} = +\infty\) is similar but easier and is omitted. This completes the proof of Theorem 2. \(\square\)

We conclude with a simple but counterintuitive example. Let \(l_n = \frac{n}{2n+1}, r_n = \frac{n+1}{2n+1}\) for \(n \geq 1\). Then \(t_n = \frac{1}{n+1}\), so that \(t_{\infty} = 0\). On the other hand, \(x_{\infty} = 1 + \sum_{n=1}^{\infty} t_n = \infty\). We see that the birth-death chain \(X_m\) built upon these transition probabilities has an extinction probability of 1, but \(E[X_m] \to \infty\) as \(m \to \infty\).

4. Acknowledgements

I am grateful for support from the Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant #2009-0094070). I would also like to thank the referee for comments which improved the exposition, as well as for asking an interesting question which led to the development of Section 3.
5. References

References

[1] Lang, S., 1997, *Undergraduate Analysis*, Springer-Verlag, New York.

[2] Marcus, M., Rosen, J., 2006, *Markov Processes, Gaussian Processes, and Local Times*, Cambridge University Press, New York.

[3] Norris, J. R., 1997, *Markov chains*, Cambridge University Press, New York.

[4] Revuz, D., Yor, M., 1968, *Continuous Martingales and Brownian Motion*, third edition, Springer, Berlin.

[5] Syski, R., 1992, *Passage Times for Markov Chains*, IOC Press, Amsterdam.