Random walk in a finite directed graph subject to a synchronizing road coloring

Kouji Yano and Kenji Yasutomi

Keywords and phrases: Markov chain, random walk, entropy, road coloring, coupling from the past.
AMS 2010 subject classifications: Primary 60J10; secondary 05C81; 37H10.

Abstract

A constructive proof is given to the fact that any ergodic Markov chain can be realized as a random walk subject to a synchronizing road coloring. Redundancy (ratio of extra entropy) in such a realization is also studied.

1 Introduction

A random walk in $\mathbb{R}$ is a process $(S_n)_{n \geq 0}$ which may be represented as

$$S_n = \xi_n + \xi_{n-1} + \cdots + \xi_1 + S_0, \quad n \geq 1$$

(1.1)

for some sequence $(\xi_n)_{n \geq 1}$ of IID (i.e., independent and identically distributed) random variables being independent of $S_0$. Note that equation (1.1) is equivalent to the recursion relation

$$S_n = \xi_n + S_{n-1}, \quad n \geq 1.$$  

(1.2)

We may introduce a natural analogue of random walk taking values in a finite set $V$, say, $\{1, \ldots, m\}$. Let $\Sigma$ denote the set of all mappings of $V$ into itself. A random walk in $V$ is a pair of processes $\{(X_n)_{n \geq 0}, (\phi_n)_{n \geq 1}\}$ such that $(\phi_n)_{n \geq 1}$ is a sequence of IID random variables taking values in $\Sigma$ and being independent of $X_0$ and such that

$$X_n = (\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1)(X_0), \quad n \geq 1.$$  

(1.3)

Note that equation (1.3) is equivalent to the recursion relation

$$X_n = \phi_n(X_{n-1}), \quad n \geq 1.$$  

(1.4)

It is obvious that, for each $n \geq 1$, the random variable $\phi_n$ is independent of $\sigma(X_j, \phi_j : j \leq n-1)$, since each $X_j$ is measurable with respect to $\sigma(X_0, \phi_1, \ldots, \phi_j)$.

It is now natural to extend the index set to $\mathbb{Z}$, the set of all integers, as follows.

---

1 Graduate School of Science, Kyoto University, Kyoto, JAPAN.
2 The research of this author was supported by KAKENHI (20740060).
3 Graduate School of Science and Engineering, Ritsumeikan University, Kusatsu, JAPAN.
Definition 1.1. A random walk in $V$ parametrized by $Z$ is a pair of processes $\{(X_n)_{n \in Z}, (\phi_n)_{n \in Z}\}$ which satisfies the following conditions:

(i) $(\phi_n)_{n \in Z}$ is a sequence of IID random variables taking values in $\Sigma$;
(ii) for each $n \in Z$, the random variable $\phi_n$ is independent of $\sigma(X_j, \phi_j : j \leq n - 1)$;
(iii) $X_n = \phi_n(X_{n-1})$ holds almost surely for all $n \in Z$.

If $\phi_n$'s have common law $\mu$ on $\Sigma$, such a random walk is called a $\mu$-random walk.

Our $\mu$-random walk may also be called a random walk in a finite directed graph subject to a road coloring. The reason will be explained in Section 5. Each element of $V$ will be called a site.

Proposition 1.2. For any Markov chain $(Y_n)_{n \in Z}$ in $V$, there exists a mapping law $\mu$.

A proof by means of rational approximation can be found in $[9]$. We shall give its constructive proof in the next section.

Note that, for a given Markov chain, there may exist several mapping laws. We may expect that we can take a nice mapping law in the following sense.

Definition 1.3. Let $\mu$ be a probability law on $\Sigma$ and denote by $\text{Supp}(\mu)$ the support of $\mu$. We say that $\mu$ is synchronizing (or simply sync) if there exists a finite sequence $\sigma_1, \ldots, \sigma_p$ of elements of $\text{Supp}(\mu)$ such that $\sigma_p \circ \sigma_{p-1} \circ \cdots \circ \sigma_1$ maps $V$ into a singleton.

Note that a $\mu$-random walk associated with a sync mapping law is utilized in Propp–Wilson’s sampling method of stationary law, which is called coupling from the past; we shall mention it briefly in Section 4.

Suppose that $\mu$ is sync and let $\{(X_n)_{n \in Z}, (\phi_n)_{n \in Z}\}$ be a $\mu$-random walk. We may assume without loss of generality that for any $x \in V$ there exists $\sigma \in \text{Supp}(\mu)$ such that $x \in \sigma(V)$; in fact, the Markov chain $(X_n)_{n \in Z}$ never visits such sites $x \in V$ that $x \notin \sigma(V)$ for any $\sigma \in \text{Supp}(\mu)$. Then we see that the Markov chain $(X_n)_{n \in Z}$ is ergodic, i.e., the following two conditions hold (see, e.g., $[5]$):
(i) the Markov chain is \emph{irreducible}, i.e., \(P(X_0 = x) > 0\) for all \(x \in V\) and for any \(x, y \in V\) there exists \(n \geq 1\) such that \(P(X_n = y|X_0 = x) > 0\);

(ii) the Markov chain is \emph{aperiodic}, i.e., for any \(x \in V\), the greatest common divisor of \(\{n \geq 1 : P(X_n = x|X_0 = x) > 0\}\) is one.

The condition (i) is obvious. The condition (ii) may be verified as follows. Let \(x \in V\). Take \(a \in V\) such that \(\sigma_p \circ \cdots \circ \sigma_1(V) = \{a\}\) and take \(q \geq 1\) such that \(P(X_q = x|X_0 = a) > 0\). Then the set \(\{n \geq 1 : P(X_n = x|X_0 = x) > 0\}\) contains all integers greater than \(p + q\), and hence its greatest common divisor is one.

The following theorem asserts that the converse is also true.

\textbf{Theorem 1.4} \cite{9}. \textit{Suppose that \((Y_n)_{n \in \mathbb{Z}}\) is an ergodic Markov chain. Then there exists a sync mapping law.}

To prove Theorem \ref{Theorem 1.4} the authors in \cite{9} utilized a profound graph-theoretic theorem, which was recently obtained by Trahtman \cite{7}, the complete solution to the \emph{road coloring problem}; we shall explain it briefly in Section \ref{Section 5}. In this chapter, we would like to give an elementary, self-contained and constructive proof of Theorem \ref{Theorem 1.4} without using Trahtman’s theorem.

The remainder of this chapter is as follows. In Section \ref{Section 2} we give constructive proofs to Proposition \ref{Proposition 1.2} and Theorem \ref{Theorem 1.4}. In Section \ref{Section 3} we study redundancy in random walk realization of a Markov chain. In Section \ref{Section 4} we mention the coupling from the past. In Section \ref{Section 5} we explain how our random walk is related to road coloring. In Section \ref{Section 6} we provide a summary and conclusion.

\section{A constructive proof of existence of sync mapping law}

A matrix \(Q = (q_{x,y})_{x,y \in V}\) with non-negative entries is called a \emph{transition matrix} if

\[
\sum_{y \in V} q_{x,y} = 1 \quad \text{for all } x \in V. \tag{2.7}
\]

We give a constructive proof of Proposition \ref{Proposition 1.2} for later use.

\textbf{A constructive proof of Proposition \ref{Proposition 1.2}.} It suffices to show that, for any transition matrix \(Q\), there exists a mapping law \(\mu\) for \(Q\), i.e.,

\[
q_{x,y} = \mu(\sigma \in \Sigma : \sigma(x) = y), \quad x, y \in V. \tag{2.8}
\]

We define

\[
E(Q) = \{(x, y) \in V \times V : q_{x,y} > 0\}. \tag{2.9}
\]
Let us prove the result by induction of $\sharp E(Q)$, where $\sharp A$ stands for the number of elements of $A$. It is obvious by (2.7) that $\sharp \{y \in V : q_{x,y} > 0 \} \geq 1$ for all $x \in V$, and hence that $\sharp E(Q) \geq \sharp V$.

Suppose that $\sharp E(Q) = \sharp V$. Then, by (2.7), it holds that $\sharp \{y \in V : q_{x,y} > 0 \} = 1$ for all $x \in V$. This shows that there exists $\sigma \in \Sigma$ such that

$$q_{x,y} = \begin{cases} 1 \quad \text{if } y = \sigma(x), \\ 0 \quad \text{otherwise.} \end{cases}$$

(2.10)

Thus the Dirac mass at $\sigma$ is as desired.

Let $k > \sharp V$ and suppose that all transition matrix $Q$ such that $\sharp E(Q) < k$ admits a mapping law. Let $Q$ be a transition matrix such that $\sharp E(Q) = k$. Write

$$\varepsilon = \min \{q_{x,y} : (x, y) \in E(Q)\}.$$

(2.11)

Since $\sharp E(Q) > \sharp V$, we see that $0 < \varepsilon < 1$. Take $(x, y) \in E(Q)$ such that $\varepsilon = q_{x,y}$, and take $\sigma \in \Sigma$ such that $(x, \sigma(x)) \in E(Q)$ for all $x \in V$. Define $\tilde{Q} = (\tilde{q}_{x,y})_{x,y \in V}$ by

$$\tilde{q}_{x,y} = \frac{1}{1 - \varepsilon} (q_{x,y} - \varepsilon 1_{\{\sigma(x) = y\}}).$$

(2.12)

Then we see that $\tilde{Q}$ is a transition matrix and that $\sharp E(\tilde{Q}) < k$. Now, by the assumption of the induction, we see that $\tilde{Q}$ admits a mapping law $\tilde{\mu}$. Therefore we conclude that $(1 - \varepsilon)\tilde{\mu} + \varepsilon \delta_\sigma$ is a mapping law for $Q$. The proof is now complete. \(\square\)

Utilizing Proposition 1.2 we give a constructive proof of Theorem 1.4.

A constructive proof of Theorem 1.4. Let $Q$ be the transition matrix of an ergodic Markov chain. Then we see that there exists $r \geq 1$ such that the $r$-th product $Q^r$ has positive entries.

Take $x_0 \in V$ arbitrarily and set $V_r = \{x_0\}$. If $V_k$ is defined for $k = r, r-1, \ldots, 1$, define $V_{k-1}$ recursively by

$$V_{k-1} = \{x \in V : (x, y) \in E(Q) \text{ for some } y \in V_k\},$$

(2.13)

where $E(Q)$ has been defined in (2.9). Note that

$$\sharp V_r \leq \sharp V_{r-1} \leq \cdots \leq \sharp V_0.$$  

(2.14)

Since $Q^r$ has positive entries, we see that $V_0 = V$.

For $k = r, r-1, \ldots, 1$, we pick $\sigma_k \in \Sigma$ so that $\sigma_k(x) \in V_k$ if $x \in V_{k-1}$ and $(x, \sigma_k(x)) \in E(Q)$ if $x \not\in V_{k-1}$. We then have

$$\sigma_r \circ \sigma_{r-1} \circ \cdots \circ \sigma_1 (V) = \{x_0\}.$$  

(2.15)
Let $\mu^{(1)}$ denote the uniform law on $\{\sigma_1, \ldots, \sigma_r\}$. Then we see that $\mu^{(1)}$ is a sync mapping law for a transition matrix $Q^{(1)} = (q^{(1)}_{x,y})_{x,y \in V}$ where

$$q^{(1)}_{x,y} = \mu^{(1)}(\sigma \in \Sigma : y = \sigma(x)) = \frac{1}{r} \sum_{k=1}^{r} 1_{\{y = \sigma_k(x)\}}. \quad (2.16)$$

Write

$$\varepsilon = \min \{ q_{x,y} : (x, y) \in E(Q) \} > 0 \quad (2.17)$$

and define $Q^{(2)} = (q^{(2)}_{x,y})_{x,y \in V}$ by

$$q^{(2)}_{x,y} = \frac{1}{1 - \varepsilon} (q_{x,y} - \varepsilon q^{(1)}_{x,y}). \quad (2.18)$$

Then $Q^{(2)}$ is a transition matrix, so that we may obtain a mapping law $\mu^{(2)}$ for $Q^{(2)}$ in the constructive way of the proof of Proposition 1.2 given above.

Now we define

$$\mu = \varepsilon \mu^{(1)} + (1 - \varepsilon) \mu^{(2)}, \quad (2.19)$$

which we have proved that is a sync mapping law for $Q$. The proof is therefore complete. \(\square\)

3 Redundancy in random walk realization

The uncertainty associated with information source may be measured by entropy (see, e.g., [2]). A Markov chain $Y = (Y_n)_{n \in \mathbb{Z}}$ with transition matrix $Q = (q_{x,y})_{x,y \in V}$ and with stationary law $\lambda$ has its entropy given by

$$h(Y) = - \sum_{x,y \in V} \lambda(x) q_{x,y} \log q_{x,y}, \quad (3.20)$$

where we adopt the binary logarithm $\log = \log_2$ for simplicity, and follow the usual convention: $0 \log 0 = 0$. For a probability law $\mu$ on $\Sigma$, an IID sequence $\phi = (\phi_n)_{n \in \mathbb{Z}}$ with common law $\mu$ has its entropy given by

$$h(\phi) = h(\mu) = - \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma). \quad (3.21)$$

A $\mu$-random walk $(X, \phi) = \{(X_n)_{n \in \mathbb{Z}}, (\phi_n)_{n \in \mathbb{Z}}\}$ with stationary law $\lambda$ is a Markov chain whose transition matrix $\overline{Q} = (\overline{q}(x,\nu),(y,\sigma))_{(x,\nu),(y,\sigma) \in V \times \Sigma}$ and stationary law $\overline{\lambda}$ given by

$$\overline{q}(x,\nu),(y,\sigma) = \mu(\sigma)1_{\{y = \sigma(x)\}}, \quad \overline{\lambda}((x, \nu)) = \mu(\nu)\lambda(w \in V : x = \nu(w)). \quad (3.22)$$
Now its entropy \( h(X, \phi) \) is computed as
\[
    h(X, \phi) = - \sum_{(x, \nu), (y, \sigma) \in V \times \Sigma} \lambda((x, \nu)) \overline{q}_{(x, \nu), (y, \sigma)} \log \overline{q}_{(x, \nu), (y, \sigma)} \tag{3.23}
\]
\[
    = - \sum_{x, y \in V, \sigma \in \Sigma} \left\{ \sum_{\nu \in \Sigma} \mu(\nu) \lambda(x \in V : x = \nu(w)) \right\} 1_{\{y=\sigma(x)\}} \mu(\sigma) \log \mu(\sigma) \tag{3.24}
\]
\[
    = - \sum_{x \in V, \sigma \in \Sigma} \lambda(x) \left\{ \sum_{y \in V} 1_{\{y=\sigma(x)\}} \right\} \mu(\sigma) \log \mu(\sigma) \tag{3.25}
\]
\[
    = - \sum_{x \in V} \lambda(x) \left\{ \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma) \right\} \tag{3.26}
\]
\[
    = - \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma). \tag{3.27}
\]

Thus we obtain
\[
    h(X, \phi) = h(\phi) = h(\mu). \tag{3.28}
\]

If the Markov chain \( Y \) is identical in law to \( X \) for some \( \mu \)-random walk \((X, \phi)\), we have
\[
    h(\mu) \geq h(Y). \tag{3.29}
\]

In fact, by (2.8), we have \( \mu(\sigma) \leq q_{x,y} \) if \( y = \sigma(x) \), and hence we see that
\[
    h(\mu) = - \sum_{x \in V} \lambda(x) \sum_{\sigma \in \Sigma} \mu(\sigma) \log \mu(\sigma) \tag{3.29}
\]
\[
    = - \sum_{x, y \in V} \lambda(x) \sum_{\sigma \in \Sigma : y=\sigma(x)} \mu(\sigma) \log \mu(\sigma) \tag{3.30}
\]
\[
    \geq - \sum_{x, y \in V} \lambda(x) \sum_{\sigma \in \Sigma : y=\sigma(x)} \mu(\sigma) \log q_{x,y} \tag{3.31}
\]
\[
    = - \sum_{x, y \in V} \lambda(x) q_{x,y} \log q_{x,y} = h(Y). \tag{3.32}
\]

The inequality (3.28) shows that any \( \mu \)-random walk realization of \( Y \) requires some extra entropy, the extent of which may be measured by
\[
    r(\mu; Y) := \frac{h(\mu) - h(Y)}{h(\mu)}. \tag{3.33}
\]

This quantity \( r(\mu; Y) \) is called the (relative) redundancy in the \( \mu \)-random walk realization of the Markov chain \( Y \). We denote the totality of all possible redundancies by
\[
    \rho(Y) = \{ r(\mu; Y) : \mu \text{ is a mapping law for } Y \}. \tag{3.34}
\]

**Theorem 3.1.** For a Markov chain \( Y \), the following assertions hold:

(i) the set \( \rho(Y) \) has finite minimum \( r(Y) \geq 0 \) and maximum \( R(Y) \leq 1 \);
(ii) for any \( r(Y) \leq r \leq R(Y) \), there exists a mapping law \( \mu \) for \( Y \) such that \( r(\mu; Y) = r \).

Moreover, if \( Y \) is ergodic, then the following assertion also holds:

(iii) for any \( r(Y) < r < R(Y) \), there exists a sync mapping law \( \mu \) for \( Y \) such that \( r(\mu; Y) = r \).

Proof. Let us remark on several basic facts about the entropy. Since \( \Sigma \) is a finite set, the totality of probability measures on \( \Sigma \), which is denoted by \( \mathcal{P}(\Sigma) \), is equipped with the total variation topology. It is well-known that \( \mathcal{P}(\Sigma) \) is compact and that \( \mu_n \rightarrow \mu \) if and only if \( \mu_n(\sigma) \rightarrow \mu(\sigma) \) for all \( \sigma \in \Sigma \). By definition (3.24), the function \( \mathcal{P}(\Sigma) \ni \mu \mapsto h(\mu) \) is continuous.

(i) Let \( \mathcal{P}(Y) \) denote the set of all mapping laws for \( Y \). It is obvious that \( \mathcal{P}(Y) \) is a compact convex subset of \( \mathcal{P}(\Sigma) \). Since \( h(\mu) \geq h(Y) > 0 \) for all \( \mu \in \mathcal{P}(Y) \), and since \( t \mapsto (t-h(Y))/t \) is continuous in \( t \geq h(Y) \), we see that \( \mathcal{P}(Y) \ni \mu \mapsto r(\mu; Y) \) is continuous. Hence we see that the set \( \rho(Y) \) has finite minimum \( r(Y) \) and maximum \( R(Y) \).

(ii) Take \( \mu^{(1)}, \mu^{(2)} \in \mathcal{P}(Y) \) such that \( r(Y) = r(\mu^{(1)}; Y) \) and \( R(Y) = r(\mu^{(2)}; Y) \). Let \( 0 \leq p \leq 1 \). Then \( \mu_p := p\mu^{(1)} + (1-p)\mu^{(2)} \) also belongs to \( \mathcal{P}(Y) \). Since \([0,1] \ni p \mapsto r(\mu_p; Y)\) is continuous, we see that \( \rho(Y) \) contains all \( r \) such that \( r(Y) < r < R(Y) \). Thus we obtain (ii).

(iii) Suppose that \( Y \) is ergodic. Theorem 1.4 implies that there exists a sync mapping law \( \mu^{(0)} \) for \( Y \). Let \( r(Y) < r < R(Y) \) and take \( r^{(1)}, r^{(2)} \) such that \( r(Y) < r^{(1)} < r \) and \( r < r^{(2)} < R(Y) \). By (ii), we may take mapping laws \( \mu^{(1)} \) and \( \mu^{(2)} \) for \( Y \) such that \( r(\mu^{(1)}; Y) = r^{(1)} \) and \( r(\mu^{(2)}; Y) = r^{(2)} \). Now we may take \( \varepsilon > 0 \) small enough such that

\[
\begin{align*}
  r((1-\varepsilon)\mu^{(1)} + \varepsilon\mu^{(0)}; Y) &< r((1-\varepsilon)\mu^{(2)} + \varepsilon\mu^{(0)}; Y).
\end{align*}
\]

Hence we may take \( 0 < p < 1 \) such that the mapping law \( \mu \) defined by

\[
\mu = (1-\varepsilon)(p\mu^{(1)} + (1-p)\mu^{(2)}) + \varepsilon\mu^{(0)}
\]

satisfies \( r(\mu; Y) = r \). This shows that \( \mu \) is a sync mapping law for \( Y \). Therefore the proof is complete.

Example 3.2. Let \( V = \{1, 2, 3\} \) and let

\[
Q = \begin{bmatrix}
  q_{1,1} & q_{1,2} & q_{1,3} \\
  q_{2,1} & q_{2,2} & q_{2,3} \\
  q_{3,1} & q_{3,2} & q_{3,3}
\end{bmatrix} = \begin{bmatrix}
  0 & 0 & 1 \\
  1/2 & 0 & 1/2 \\
  1/2 & 1/2 & 0
\end{bmatrix}.
\]

(3.37)

The Markov chain \( Y \) with transition matrix \( Q \) has a unique stationary law

\[
\lambda = [\lambda(1), \lambda(2), \lambda(3)] = \frac{1}{9} [3, 2, 4].
\]

(3.38)
A simple computation leads to \( h(Y) = 2/3 \). For a mapping law \( \mu \) for \( Y \), elements which may possibly be contained in \( \text{Supp}(\mu) \) are the following four:

\[
\begin{align*}
\sigma^{(1)} &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \\
\sigma^{(2)} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \\
\sigma^{(3)} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \\
\sigma^{(4)} &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.
\end{align*}
\]  

(3.39)

Set \( p = \mu(\sigma^{(1)}) \). A simple computation leads to

\[
\mu(\sigma^{(2)}) = p, \quad \mu(\sigma^{(3)}) = \mu(\sigma^{(4)}) = 1/2 - p.
\]  

(3.40)

Thus we obtain

\[
h(\mu) = 2f(p) + 2f(1/2 - p),
\]  

(3.41)

where \( f(t) = -t \log t \). Since the variable \( p \) may vary in \([0, 1/2]\), we see that \( h(\mu) \) ranges \([1, 2]\), where the minimum \( h(\mu) = 1 \) is attained at \( p = 0 \) and \( 1/2 \) and the maximum \( h(\mu) = 2 \) at \( p = 1/4 \). Hence we obtain

\[
r(Y) = 1/3, \quad R(Y) = 2/3.
\]  

(3.42)

In this case, for all \( p \in [0, 1/2] \), the mapping law \( \mu \) is sync; in fact, \( \sigma^{(1)} \circ \sigma^{(2)}(V) = \{1\} \) and \( \sigma^{(3)} \circ \sigma^{(4)} \circ \sigma^{(4)} \circ \sigma^{(3)}(V) = \{3\} \).

Before closing this section, we mention the following theorem, which provides a necessary and sufficient condition for zero minimum redundancy.

**Theorem 3.3** ([9]). Suppose that \( Y \) is ergodic. Then \( r(Y) = 0 \) if and only if \( Y \) is \( p \)-uniform, i.e., there exists a probability law \( \nu \) on \( V \) and a family \( \{\tau_x : x \in V\} \) of permutations of \( V \) such that

\[
q_{x,y} = \nu(\tau_x(y)), \quad x, y \in V.
\]  

(3.43)

For the proof of Theorem 3.3 see [9].

### 4 Coupling from the past

In some practical problems, we sometimes need to simulate the stationary law of an ergodic Markov chain. As a powerful method for the simulation, Propp–Wilson’s coupling from the past is widely known; see [6] and also [3] and [4]. The fundamental idea is to utilize a random walk realization associated with a sync mapping law. Let us explain it briefly.

Let an ergodic Markov chain be given and suppose that we find a sync mapping law \( \mu \) for the Markov chain. Then a \( \mu \)-random walk \( \{(X_n)_{n \in \mathbb{Z}}, (\phi_n)_{n \in \mathbb{Z}}\} \) is a realization of the Markov chain. Let \( (\sigma_1, \ldots, \sigma_p) \) be a finite sequence of elements of \( \text{Supp}(\mu) \) such that
\( \sigma_p \circ \cdots \circ \sigma_1(V) \) is a singleton. The latest time when the exact sequence \((\sigma_p, \sigma_{p-1}, \ldots, \sigma_1)\) can be found in \((\phi_0, \phi_{-1}, \ldots)\) will be denoted by

\[
T = \sup \{ k \in \mathbb{Z} : 0 \geq k + p - 1, \ \phi_{k+p-1} = \sigma_p, \ldots, \phi_k = \sigma_1 \}. \tag{4.44}
\]

Here we understand that \( \sup \emptyset = -\infty \). Note that \( T \) is finite almost surely. This random time \( T \) plays a role of stopping time in the sense that

\[
\{ T = k \} \in \sigma (\phi_0, \phi_{-1}, \ldots, \phi_k) \quad \text{for} \quad 0 \geq k + p - 1. \tag{4.45}
\]

Since \( \sigma_p \circ \cdots \circ \sigma_1(V) \) is a singleton, we see that \( \phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_T \) maps \( V \) into a singleton. Thus it holds that

\[
X_0 = \phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_T(x) \quad \text{a.s.} \tag{4.46}
\]

for all \( x \in V \). This shows the following: We pick a sequence \( f_0, f_{-1}, \ldots \) from the law \( \mu \) up to the latest time \( T \) when \( (f_{T+p-1}, \ldots, f_T) = (\sigma_p, \ldots, \sigma_1) \). Then the resulting site \( f_0 \circ f_{-1} \circ \cdots \circ f_T(x) \), which does not depend on the choice of \( x \in V \), is a sample point from the stationary law, which is as desired.

This method can be applied to simulate a Gibbs distribution. In this case, a sync mapping law can be constructed with the help of monotonicity structure of the state space \( V \).

**Remark 4.1.** The identity (4.46) implies that, for each \( n \in \mathbb{Z} \), the random variable \( X_n \) is measurable with respect to \( \sigma (\phi_j : j \leq n) \). One can ask what happens when \( \mu \) is not sync. The following theorem answers this question.

**Theorem 4.2** (Yano [8]). Suppose that the Markov chain \( (X_n)_{n \in \mathbb{Z}} \) is ergodic and that \( \mu \) is not sync. Then, for each \( n \in \mathbb{Z} \), the random variable \( X_n \) is not measurable with respect to \( \sigma (\phi_j : j \leq n) \).

For the proof of Theorem 4.2 see [8].

## 5 Random walk and road coloring

Let us explain how our \( \mu \)-random walk is related to road coloring.

First, let us introduce some notations in graph theory. A finite directed graph is the pair \((V, A)\) of finite sets \( V \) and \( A \) associated with mappings \( i : A \to V \) and \( t : A \to V \). Each element of \( V \) will be called a site (or a node) and each element \( a \) of \( A \) will be called a (oneway) road (or an arrow) which runs from \( i(a) \) to \( t(a) \). For \( a \in A \), the site \( i(a) \) (resp. \( t(a) \)) will be called the initial (resp. terminal) site of \( a \). For \( x \in V \), the number of roads running from \( x \), namely,

\[
O(x) = \sharp \{ a \in A : i(a) = x \}, \tag{5.47}
\]
will be called the outdegree at the site $x$. If $O(x)$ does not depend on $x \in V$, the directed graph $(V, A)$ is called of constant outdegree. A path from $x \in V$ to $y \in V$ is a word $w = (a_1, \ldots, a_n)$ of roads such that $a_1$ runs from $x$ to $i(a_2)$, $a_2$ to $i(a_3)$, $\ldots$, $a_{n-1}$ to $i(a_n)$, and $a_n$ to $y$. The number $L(w) = n$ is called the length of the path $w = (a_1, \ldots, a_n)$. The directed graph $(V, A)$ is called strongly connected if, for any $x, y \in V$, there exists a path from $x$ to $y$. The directed graph $(V, A)$ is called aperiodic if, for any $x \in V$, the greatest common divisor of the set of $L(w) \geq 1$ among all paths $w$ from $x$ to itself is one.

Second, we introduce some notations in road coloring. Suppose that $(V, A)$ is of constant outdegree and denote the common outdegree by $d$. A road coloring of $(V, A)$ is a partition of $A$ into $d$ disjoint subsets $C = \{c_1, \ldots, c_d\}$ such that, for each $x \in V$, each color $c_k$ contains one and only one road whose initial site is $x$. For a finite sequence $s = (c_1, \ldots, c_p)$ of elements of $C$, a path $w = (a_1, \ldots, a_p)$ is said to be along $s$ if $a_k \in c_k$ for all $k = 0, 1, \ldots, p$. The following notion originates Adler, Goodwyn and Weiss [1].

**Definition 5.1.** A road coloring $C$ of $(V, A)$ is called sync if there exists a finite sequence $c_1, \ldots, c_p$ of elements of $C$ such that all paths along $(c_1, \ldots, c_p)$ have common terminal site.

Let us give an example.

**Example 5.2.** Let $V = \{1, 2, 3\}$ and $A = \{a^{(x,k)} : x \in V, k = 1, 2\}$ and define the initial and terminal sites of each road as follows:

|   | $a^{(1,1)}$ | $a^{(2,1)}$ | $a^{(3,1)}$ | $a^{(1,2)}$ | $a^{(2,2)}$ | $a^{(3,2)}$ |
|---|-------------|-------------|-------------|-------------|-------------|-------------|
| $i(a)$ | 1           | 2           | 3           | 1           | 2           | 3           |
| $t(a)$ | 3           | 3           | 1           | 3           | 1           | 2           |

Take the road coloring $C = \{c^{(1)}, c^{(2)}\}$ defined by

$c^{(1)} = \{a^{(x,1)} : x \in V\}$, $c^{(2)} = \{a^{(x,2)} : x \in V\}$. (5.49)

Now it is obvious that the road coloring $C$ is sync; in fact, all paths along $(c^{(1)}, c^{(2)}, c^{(2)}, c^{(1)})$ have common terminal site $3$.

Third, we recall the road coloring problem. If a directed graph $(V, A)$ of constant outdegree admits a sync road coloring, then it is necessarily strongly connected and aperiodic. The converse was posed as a conjecture by Adler, Goodwyn and Weiss [1], which had been called the road coloring problem until it was completely solved by Trahtman [7].

**Theorem 5.3 (Trahtman [7]).** A directed graph which is of constant outdegree, strongly connected, and aperiodic, does admit a sync road coloring.

Fourth, let us explain how to understand our $\mu$-random walk by means of road coloring. Let $\mu$ be a probability law on $\Sigma$. Since $\Sigma$ is a finite set, the support of $\mu$ may be written as $\{\sigma^{(1)}, \ldots, \sigma^{(d)}\}$. We define the set $A$ of roads as the totality of $a^{(x,k)}$ for $x \in V$ and
$k = 1, \ldots, d$ where $a^{(x,k)}$ runs from $x$ to $\sigma^{(k)}(x)$. Thus the law $\mu$ induces naturally the road coloring $C = \{c^{(1)}, \ldots, c^{(d)}\}$ such that

$$c^{(k)} = \{a^{(x,k)} : x \in V\}. \quad (5.50)$$

It is now obvious that the probability law $\mu$ is sync in the sense of Definition 1.3 if and only if the road coloring $C$ is sync in the sense of Definition 5.1.

For a $\mu$-random walk $(X, \phi)$, the process $X$ moves from site to site in the directed graph $(V, A)$ via the equation $X_n = \phi_n(X_{n-1})$, being driven by the colors of roads indicated by $\phi$ which are randomly chosen from the road coloring $C$ induced by $\mu$. Thus we may call $(X, \phi)$ a $\mu$-random walk in the directed graph $(V, A)$ subject to the road coloring $C$.

Let $Y$ be a Markov chain and suppose that $Y$ is realized as $X$ of a $\mu$-random walk $(X, \phi)$ in the directed graph $(V, A)$ subject to the road coloring $C$ induced by $\mu$. Then, to each edge $(x, y) \in E(Y)$, there corresponds at least one road $a$ which runs from $x$ to $y$. For example, consider Example 3.2 with $p = 0$. In this case, we have $\text{Supp}(\mu) = \{\sigma^{(3)}, \sigma^{(4)}\}$, and

$$E(Y) = \{(1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}. \quad (5.51)$$

The set $A$ of roads induced by $\mu$ is given as $A = \{a^{(x,k)} : x \in V, k = 1, 2\}$, where the initial and terminal sites of each road are given as (5.48). Then we find that the road coloring induced by $\mu$ is nothing else but $C = \{c^{(1)}, c^{(2)}\}$ given as (5.49) in Example 5.2, where we note that $\sigma^{(3)}$ and $\sigma^{(4)}$ correspond to $c^{(1)}$ and $c^{(2)}$, respectively. Note that there exist two roads $a^{(1,1)}$ and $a^{(1,2)}$ which run from 1 to 3, which are colored differently from each other. See Figure 1 below for the illustration.

![Figure 1](image-url)

6 Conclusion

We have introduced a random walk in a finite set as a stochastic evolutionary process driven by a IID sequence of mappings. It can be understood as a random walk in a finite directed graph moving according to random road colors. Any ergodic Markov chain is proved to be realized, in a constructive way, by a random walk associated with a sync mapping law. The redundancy in random walk realization with a sync mapping law can be as close as desired to the minimum redundancy.
References

[1] R. L. Adler, L. W. Goodwyn, and B. Weiss. Equivalence of topological Markov shifts. *Israel J. Math.*, 27(1):48–63, 1977.

[2] P. Billingsley. *Ergodic theory and information*. Robert E. Krieger Publishing Co., Huntington, N.Y., 1978. Reprint of the 1965 original.

[3] O. Häggström. *Finite Markov chains and algorithmic applications*, volume 52 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2002.

[4] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[5] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.

[6] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

[7] A. N. Trahtman. The road coloring problem. *Israel J. Math.*, 172:51–60, 2009.

[8] K. Yano. Random walk in a finite directed graph subject to a road coloring. Preprint, arXiv:1005.0079, 2010.

[9] K. Yano and K. Yasutomi. Realization of ergodic markov chain as a random walk subject to a synchronizing road coloring. *J. Appl. Probab.*, to appear. Preprint, arXiv:1006.0534, 2010.