International Journal of Algebra and Computation
Vol. 29, No. 8 (2019) 1431–1449

DOI: 10.1142/S0218196719500577

Normal distributions of finite Markov chains

John Rhodes*,‡ and Anne Schilling†,§

*Department of Mathematics, University of California Berkeley
970 Evans Hall, Berkeley, CA 94720-3840, USA

†Department of Mathematics, University of California Davis
One Shields Avenue, Davis, CA 95616-8633, USA
‡rhodes@math.berkeley.edu, blvdbastille@gmail.com
§anne@math.ucdavis.edu

Received 1 March 2019
Accepted 6 July 2019
Published 31 July 2019

Communicated by B. Steinberg

We show that the stationary distribution of a finite Markov chain can be expressed as the sum of certain normal distributions. These normal distributions are associated to planar graphs consisting of a straight line with attached loops. The loops touch only at one vertex either of the straight line or of another attached loop. Our analysis is based on our previous work, which derives the stationary distribution of a finite Markov chain using semaphore codes on the Karnofsky–Rhodes and McCammond expansion of the right Cayley graph of the finite semigroup underlying the Markov chain.

Keywords: Markov chains; stationary distributions; semaphore codes; Kleene expressions; Karnofsky–Rhodes expansion; McCammond expansion; normal distributions.

Mathematics Subject Classification 2010: Primary 20M30, 60J10; Secondary 20M05, 60B15, 60C05

1. Introduction

In our previous paper [11], we developed a general theory to compute the stationary distribution of a finite Markov chain. Every finite state Markov chain $\mathcal{M}$ has a random letter representation, that is, a representation of a semigroup S acting on the left on the state space Ω [8]. Combining the Karnofsky–Rhodes and the McCammond expansion of the right Cayley graph of S, we were able to provide a construction of the stationary distribution using finite semigroup theory without the use of linear algebra. The construction relies on the concept of lumping; the distributions for the expanded graphs can be computed thanks to normal forms of

§Corresponding author.

the elements. The stationary distribution of the original Markov chain $\mathcal{M}$ is then obtained by lumping.

In this paper, we show that the stationary distribution of any finite Markov chain can be obtained from certain normal (or Gaußian) distributions. The normal distributions are derived from planar graphs by adding directed loops (or circles) to the straight line, which only touch the graph at one point. Let us outline the construction of these normal forms in the remainder of the introduction.

1.1. *Straight line*

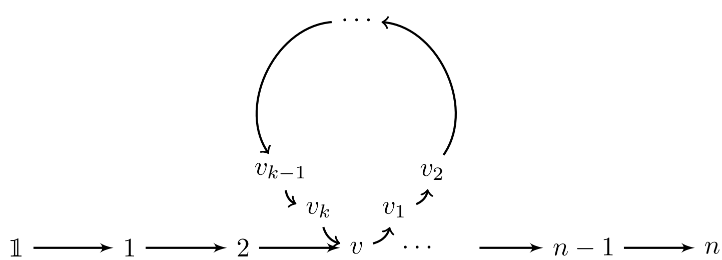
We start with a straight line starting at $\mathbb{1}$ with n further vertices:

$$\mathbb{1} \longrightarrow 1 \longrightarrow 2 \longrightarrow \cdots \longrightarrow n-1 \longrightarrow n$$

1.2. *Adding loops*

A loop is a sequence of vertices connected by edges $v_0 \longrightarrow v_1 \longrightarrow \ldots \longrightarrow v_k$ such that $v_0 = v_k$, but all other vertices v_i with $0 \leq i < k$ are distinct.

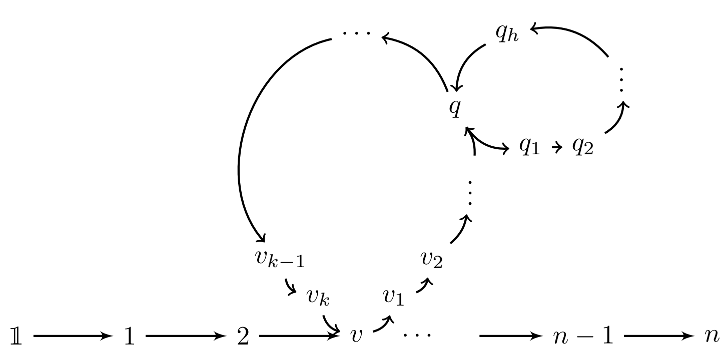
Add a loop ℓ to any vertex of the straight line constructed in Sec. 1.1 (except $\mathbb{1}$) with $k \geq 0$ new vertices, which only touches one existing vertex v.



The cut of ℓ is

$$v \longrightarrow v_1 \longrightarrow v_2 \longrightarrow \cdots \longrightarrow v_{k-1} \longrightarrow v_k$$

Continue to add loops at any vertex (except $\mathbb{1}$), including the new vertices. Multiple loops at a given vertex are allowed.

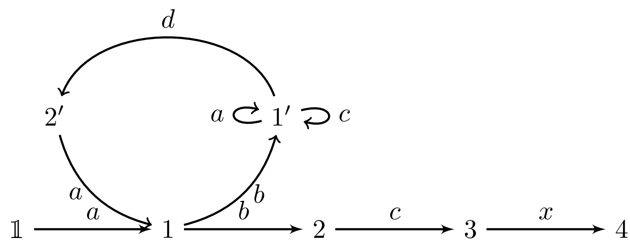


Let $\overline{G}$ be the directed graph obtained by this procedure. Notice that each such $\overline{G}$ can be drawn in the plane.

1.3. *Kleene expressions*

Given a finite alphabet A, assign a letter $a \in A$ to each arrow in the graph $\overline{G}$. The result is called a loop graph, denoted G.

Example 1.1. For the alphabet $A = \{a, b, c, d, x\}$, we might obtain



In general, this procedure gives a non-deterministic automata since different edges emitting from a vertex can be labeled by the same letter. In the above example, vertex 1 has two arrows labeled b coming out of it.

Denote the set of all paths in a loop graph G starting at $\mathbb{1}$ and ending at n (the last vertex on the initial straight line underlying G) by $\mathcal{P}_G$. Here a path is given by

$$\mathbb{1} \xrightarrow{a_1} v_1 \xrightarrow{a_2} \dots \xrightarrow{a_k} v_k = n,$$

where v_i are vertices in G and $a_i \in A$ are the labels on the edges.

There is a simple inductive way to describe $\mathcal{P}_G$ using Kleene expressions. Given a set L, define $L^0 = \{\varepsilon\}$ given by the empty string, $L^1 = L$, and recursively $L^{i+1} = \{wa \mid w \in L^i, a \in L\}$ for each integer $i > 0$. Then the Kleene star is

$$L^\star = \bigcup_{i \geq 0} L^i.$$

A Kleene expression only involves letters in A, concatenation, unions, and $\star$. To obtain a Kleene expression for $\mathcal{P}_G$, perform the following doubly recursive procedure:

Algorithm 1.

Induction basis: Start at vertex $\mathbb{1}$ and with the empty expression L.

Induction step: Suppose one is at vertex $i \neq n$ (or $\mathbb{1}$) on the straight line path underlying G.

(1) Continue to the next vertex $i+1$ (or 1) on the straight line path underlying G and append the label a on the edge from $i \xrightarrow{a} i+1$ (or $\mathbb{1} \xrightarrow{a} 1$) to L.

(2) If there are loops $\ell_1, \ell_2, \ldots, \ell_k$ at vertex $i+1$ (or 1), append the formal expression

$$\{\ell_1, \ell_2, \ldots, \ell_k\}^\star$$

to L. The loops $\ell_1, \ell_2, \ldots, \ell_k$ are in one-to-one correspondence with the edges coming into vertex $i+1$.

(3) If $i+1 \neq n$, continue with the next induction step. Else stop and output L.

Algorithm 2. For each symbol ℓ_i in the expression for L, do the following:

(1) Consider the loop $\ell_i = \left(v_0 \xrightarrow{a_1} v_1 \xrightarrow{a_2} \cdots \xrightarrow{a_k} v_k = v_0\right)$ from vertex v_0 to v_0 in G. Consider the subgraph of G with straight line $v_1 \xrightarrow{a_2} \cdots \xrightarrow{a_k} v_k$ and all further loops that are attached to any of the vertices v_i in G. Attach $\mathbb{1}$ to v_1. The resulting graph $G^{(i)}$ is a new loop graph. Perform Algorithm 1 on $G^{(i)}$ to obtain a Kleene expression $L^{(i)}$. Replace the symbol ℓ_i in L by $L^{(i)}$.

(2) Continue this process until L does not contain any further expressions ℓ_i for some loop ℓ_i, that is, L only contains unions, $\star$ and elements in the alphabet A. Then the Kleene expression for $\mathcal{P}_G$ is L.

The resulting expressions can be made into unionless expressions by using Zimin words

$$\{a\}^\star = a^\star \quad \text{and} \quad \{a,b\}^\star = (a^\star b)^\star a^\star \quad \text{for } a,b \in A. \tag{1.1}$$

Expressions for larger unions can be obtained by induction using (1.1).

Example 1.2. Let G be as in Example 1.1. Then

$$L = a\ell_1^\star bcx,$$

where ℓ_1 is the loop attached to vertex 1. Cut this loop and continue the process to obtain

$$\ell_1 = b\{\ell_1', \ell_2'\}^\star da,$$

where ℓ_1' is the loop at vertex $1'$ labelled a and ℓ_2' is the loop at vertex $1'$ labelled c. We have $\ell_1' = a$ and $\ell_2' = c$, so that altogether we find

$$L = a(b\{a,c\}^\star da)^\star bcx = a(b(a^\star c)^\star a^\star da)^\star bcx,$$

where in the last step we used the Zimin words to get rid of the unions. This is a Kleene expression for $\mathcal{P}_G$.

See Example 3.8 for another example and also compare this construction to the definition of Pict in Definition 3.5.

Main results. We are now going to define normal distributions.

Definition 1.3 (Normal distribution). Let G be a loop graph with edges labeled by letters in the alphabet A. Associate the indeterminate x_a to $a \in A$. Then the normal distribution of G is defined as

$$\Psi_G = \sum_{p \in \mathcal{P}_G} \prod_{a \in p} x_a,$$

where the product is over all letters a in p.

We may use the Kleene expressions of the previous section for $\mathcal{P}_G$. The advantage in doing so is that one can immediately obtain rational expressions. Namely, using the geometric series, we find that

$$\sum_{s \in a^\star} \prod_{i \in s} x_i = \sum_{\ell=0}^{\infty} x_a^\ell = \frac{1}{1 - x_a}.$$

Similarly

$$\sum_{s \in \{a,b\}^\star} \prod_{i \in s} x_i = \sum_{s \in a^\star (b a^\star)^\star} \prod_{i \in s} x_i = \frac{1}{1 - x_a} \cdot \frac{1}{1 - \frac{x_b}{1 - x_a}} = \frac{1}{1 - x_a - x_b}.$$

In general, using the recursion (1.1) we derive by induction

$$\sum_{s \in \{a_1, a_2, \ldots, a_n\}^\star} \prod_{i \in s} x_i = \frac{1}{1 - x_{a_1} - x_{a_2} - \ldots - x_{a_n}}. \tag{1.2}$$

Our main theorem is the following.

Theorem 1.4. *The stationary distribution $\Psi^{\mathcal{M}}$ of a finite Markov chain $\mathcal{M}$ is the sum of normal distributions Ψ_G or certain limits of Ψ_G, where G is a loop graph.*

The proof of Theorem 1.4 is given in Sec. 3.3. A more precise version of Theorem 1.4 is stated in Theorem 3.9.

The paper is outlined as follows. In Sec. 2, we review the main results from [11], in particular the expressions for the stationary distribution of a finite Markov chain in terms of semaphore codes of the Karnofsky–Rhodes expansion of the right Cayley graph of the underlying semigroup. In Sec. 3, we review the McCammond expansion and its relation to semaphore codes and provide the definition of Pict. The map Pict is used to give a proof of Theorem 1.4. The original definition of Pict is due to McCammond, but the applications to random walks are due to the authors.

2. Stationary Distributions of Markov Chains

In this section, we provide definitions and review the necessary results we need from [11].

2.1. *Markov chains*

A Markov chain $\mathcal{M}$ consists of a finite or countable state space Ω together with transition probabilities $\mathcal{T}_{s',s}$ for the transition $s \longrightarrow s'$ for $s, s' \in \Omega$. The matrix $\mathcal{T} = (\mathcal{T}_{s',s})_{s,s'\in\Omega}$ is called the transition matrix, which is a column-stochastic matrix, meaning that the column sums of $\mathcal{T}$ are equal to one.

A Markov chain is irreducible if for any $s, s' \in \Omega$ there exists an integer m (possibly depending on s, s') such that $\mathcal{T}^m_{s',s} > 0$. In other words, one can get from any state s to any other state s' using only steps with positive probability. A state $s \in \Omega$ is called recurrent if the system returns to s in finitely many steps with probability one.

The stationary distribution of $\mathcal{M}$ is a vector $\Psi = (\Psi_s)_{s\in\Omega}$ such that $\mathcal{T}\Psi = \Psi$ and $\sum_{s\in\Omega} \Psi_s = 1$. In other words, Ψ is a right-eigenvector of $\mathcal{T}$ with eigenvalue one. If the Markov chain is irreducible, the stationary distribution is unique [8].

Next we define lumping of Markov chains. Partition the state space Ω into $(\Omega_1, \ldots, \Omega_\ell)$ such that

$$\Omega_i \cap \Omega_j = \emptyset \quad \text{for } i \neq j \text{ and } \quad \Omega = \bigcup_{i=1}^{\ell} \Omega_i.$$

One may view such a partition as an equivalence relation $s \sim s'$ if $s, s' \in \Omega_i$ for some $1 \le i \le \ell$. We say that $\mathcal{M}$ can be lumped with respect to the partition $(\Omega_1, \ldots, \Omega_\ell)$ if the transition matrix $\mathcal{T}$ satisfies [8, Lemma 2.5] [5] for all $1 \le i, j \le \ell$

$$\sum_{t\in\Omega_j} \mathcal{T}_{t,s} = \sum_{t\in\Omega_j} \mathcal{T}_{t,s'} \qquad \text{for all } s, s' \in \Omega_i. \tag{2.1}$$

The lumped Markov chain is a random walk on the equivalence classes, whose stationary distribution labeled by w is $\sum_{s\sim w} \Psi_s$.

Every finite state Markov chain $\mathcal{M}$ has a random letter representation, that is, a representation of a semigroup S acting on the left on the state space Ω (see [8, Proposition 1.5] and [1, Theorem 2.3]). In this setting, we transition $s \xrightarrow{a} s'$ with probability $0 \le x_a \le 1$, where $s, s' \in \Omega$, $a \in S$ and $s' = a.s$ is the action of a on the state s. Let $A = \{a \in S \mid x_a > 0\}$. We assume that A generates S; if not, it suffices to consider the subsemigroup generated by A. Note that $\sum_{a\in A} x_a = 1$. The transition matrix $\mathcal{T}$ of $\mathcal{M}$ is the $|\Omega| \times |\Omega|$-matrix

$$\mathcal{T}_{s',s} = \sum_{\substack{a\in A \\ s\xrightarrow{a}s'}} x_a \quad \text{for } s, s' \in \Omega. \tag{2.2}$$

Note that we may assume that the action of S on Ω is faithful as this does not affect the random walk.

If S is a semigroup, then $S^{\mathbb{1}}$ denotes S with an adjoint identity $\mathbb{1}$ even if S already has an identity.

Definition 2.1 (Ideal). Let S be a semigroup. A two-sided ideal I (or ideal for short) is a subset $I \subseteq S$ such that $uIv \subseteq I$ for all $u, v \in S^{\mathbb{1}}$. Similarly, a left ideal I is a subset $I \subseteq S^{\mathbb{1}}$ such that $uI \subseteq I$ for all $u \in S^{\mathbb{1}}$.

If I, J are ideals of S, then $IJ \subseteq I \cap J$, so that $I \cap J \neq \emptyset$. Hence every finite semigroup has a unique minimal ideal denoted $K(S)$. As shown in [3, 6], the minimal ideal $K(S)$ of a finite semigroup S is the disjoint union of all the minimal left ideals of S and the Rees Theorem applies. By [1, Remark 2.8] the faithful left action of S on Ω is isomorphic to the left action of S on $K(S)$.

Let (S, A) be a semigroup S together with a choice of generators A for S. Define $\mathcal{M}(S, A)$ to be the Markov chain, where the transition $s \xrightarrow{a} s'$ for $s, s' \in S$ and $a \in A$ is given by $s' = as$ in the left Cayley graph with probability $0 < x_a \leq 1$. Note that we are assuming that all probabilities x_a for $a \in A$ are nonzero. Then it was shown in [4] (see also [1, Proposition 3.2]) that the recurrent states of $\mathcal{M}(S, A)$ are the elements in $K(S)$. Furthermore, the connected components of the recurrent states in the random walk are the minimal left ideals of S. The restriction of the random walk to any minimal left ideal is irreducible. Moreover, the chain so obtained is independent of the chosen minimal left ideal. This random walk and the random walk with states a left ideal L of $K(S)$ and S acting on the left made faithful, that is $x \xrightarrow{a} y$ for $x \in L$ and $y = ax$, are essentially the same. So we may not distinguish the two cases.

2.2. *Karnofsky–Rhodes expansion*

In this section, we define the right Cayley graph of a finite semigroup and its Karnofsky–Rhodes expansions.

Definition 2.2 (Right Cayley graph). Let (S, A) be a finite semigroup S together with a set of generators A. The right Cayley graph $\mathsf{RCay}(S, A)$ of S with respect to A is the rooted graph with vertex set $S^{\mathbb{1}}$, root $r = \mathbb{1} \in S^{\mathbb{1}}$, and edges $s \xrightarrow{a} s'$ for all $(s, a, s') \in S^{\mathbb{1}} \times A \times S^{\mathbb{1}}$, where $s' = sa$ in $S^{\mathbb{1}}$.

A path p in $\mathsf{RCay}(S, A)$ is a sequence

$$p = \left(v_1 \xrightarrow{a_1} \dots \xrightarrow{a_\ell} v_{\ell+1}\right),$$

where $v_i \in S^{\mathbb{1}}$ are vertices in $\mathsf{RCay}(S, A)$ and $v_i \xrightarrow{a_i} v_{i+1}$ are edges in $\mathsf{RCay}(S, A)$. The endpoint of p is $\tau(p) := v_{\ell+1}$. The length of the path p is $\ell(p) := \ell$, which equals the number of edges. A simple path is a path that does not visit any vertex twice. Empty paths are considered simple. A path which starts and ends at the same vertex is called a circuit. A circuit that is simple, when the last vertex is removed, is called a loop.

Definition 2.3 (Transition edges). An edge $s \xrightarrow{a} s'$ in the right Cayley graph $\mathsf{RCay}(S, A)$ is a transition edge if there is no directed path from s' to s in $\mathsf{RCay}(S, A)$. In other words, there does not exist any sequence $a_1, \ldots, a_k \in A$ with $k \geq 1$ such that $s'(a_1 \ldots a_k) = s$.

Let us now define the Karnofsky–Rhodes expansion of the right Cayley graph (see also [10, Definition 4.15] and [7, Sec. 3.4]). Let (A^+, A) be the free semigroup with generators A, where A^+ is the set of all words $a_1 \ldots a_\ell$ of length $\ell \geq 1$ over A with multiplication given by concatenation. When we write $[a_1 \ldots a_\ell]_S$, we mean the element in S when taking the product in the semigroup of the generators $a_i \in A$.

Definition 2.4 (Karnofksy–Rhodes expansion). The Karnofsky–Rhodes expansion $\mathsf{KR}(S, A)$ is obtained as follows. Start with the right Cayley graph $\mathsf{RCay}(A^+, A)$. Identify two paths in $\mathsf{RCay}(A^+, A)$

$$p := \left(\mathbb{1} \xrightarrow{a_1} v_1 \xrightarrow{a_2} \ldots \xrightarrow{a_\ell} v_\ell\right) \quad \text{and} \quad p' := \left(\mathbb{1} \xrightarrow{a'_1} v'_1 \xrightarrow{a'_2} \ldots \xrightarrow{a'_{\ell'}} v'_{\ell'}\right)$$

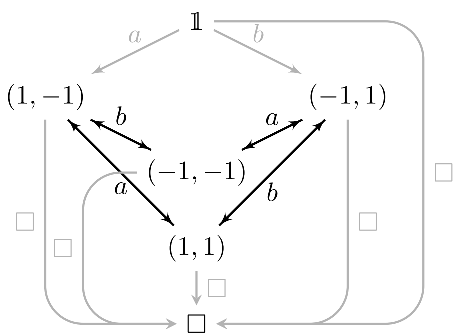
in $\mathsf{KR}(S, A)$ if and only if the corresponding paths in $\mathsf{RCay}(S, A)$

$$[p]_S := \left(\mathbb{1} \xrightarrow{a_1} [v_1]_S \xrightarrow{a_2} \ldots \xrightarrow{a_\ell} [v_\ell]_S\right) \quad \text{and}$$

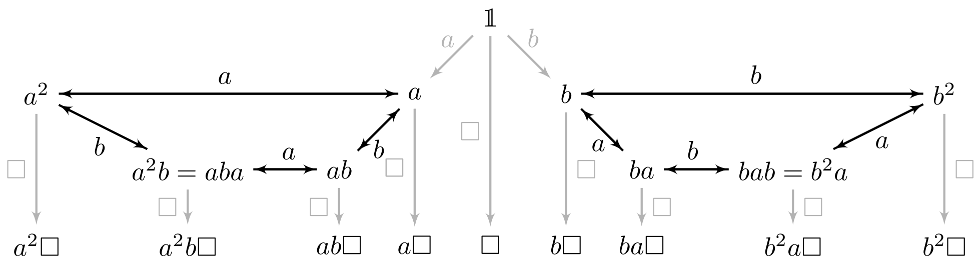
$$[p']_S := \left(\mathbb{1} \xrightarrow{a'_1} [v'_1]_S \xrightarrow{a'_2} \ldots \xrightarrow{a'_{\ell'}} [v'_{\ell'}]_S\right),$$

where $v_i = a_1 a_2 \ldots a_i$ and $v'_i = a'_1 a'_2 \ldots a'_i$, end at the same vertex $[v_\ell]_S = [v'_{\ell'}]_S$ and in addition the set of transition edges of $[p]_S$ and $[p']_S$ in $\mathsf{RCay}(S, A)$ is equal.

Example 2.5. Consider the right Cayley graph of the Klein 4-group $Z_2 \times Z_2$ with zero with generators $\{a, b, \square\}$, where $a = (1, -1)$, $b = (-1, 1)$, and $\square$ is the zero. The right Cayley graph $\mathsf{RCay}(Z_2 \times Z_2 \cup \{\square\}, \{a, b, \square\})$ is



where all three arrows $a, b, \square$ fix the vertex $\square$ at the bottom. Transition edges are indicated in grey. Double edges mean that right multiplication by the label for either vertex yields the other vertex. The Karnofsky–Rhodes expansion of this right Cayley graph is given by



where arrows $a, b, \square$ fix all the vertices at the bottom.

Proposition 2.6 ([11, Proposition 2.15]). $\mathsf{KR}(S, A)$ *is the right Cayley graph of a semigroup, also denoted by* $\mathsf{KR}(S, A)$.

2.3. *Stationary distribution*

We now review the main results of [11], which give the stationary distribution for any Markov chain $\mathcal{M}(S, A)$ for a finite semigroup with chosen generators (S, A). Recall that $\mathcal{M}(S, A)$ is the random walk on the unique minimal ideal $K(S)$ of S. More precisely, the random walk is given by the left action of S on $K(S)$.

To state our results for the stationary distribution, we first need to review the semaphore codes associated to (S, A) [2]. The semaphore code $\mathcal{S}(S, A)$ is the set of all words $a_1a_2 \ldots a_\ell \in A^+$ such that $[a_1a_2 \ldots a_\ell]_S \in K(S)$, but $[a_1a_2 \ldots a_{\ell-1}]_S \notin K(S)$. Semaphore codes are closely related to normal forms of the McCammond expansion, see Sec. 3.1.

The main results are the following.

Theorem 2.7 ([11, Corollary 2.28]). *The Markov chain* $\mathcal{M}(S, A)$ *is the lumping of* $\mathcal{M}(\mathsf{KR}(S, A))$ *with stationary distribution*

$$\Psi_w^{\mathcal{M}(S,A)} = \sum_{\substack{v \in \mathsf{KR}(S,A) \\ [v]_S = w}} \Psi_v^{\mathcal{M}(\mathsf{KR}(S,A))} \quad \text{for all } w \in (S, A).$$

The next result is non-trivial. It requires the assumption that the minimal ideal $K(S)$ is left zero, that is, $xy = x$ for all $x, y \in K(S)$.

Theorem 2.8 ([11, Theorem 2.12]). *If* $K(S)$ *is left zero, the stationary distribution of the Markov chain* $\mathcal{M}(\mathsf{KR}(S, A))$ *is given by*

$$\Psi_w^{\mathcal{M}(\mathsf{KR}(S,A))} = \sum_{\substack{s \in \mathcal{S}(S,A) \\ [s]_{\mathsf{KR}(S,A)} = w}} \prod_{a \in s} x_a \quad \text{for all } w \in K(\mathsf{KR}(S, A)).$$

As outlined in [11, Sec. 2.9], the case when $K(S)$ is not left zero can be constructed from the case when $K(S)$ is left zero using the flat operation. That is, one adds an additional generator $\Box$ to the alphabet A, which acts as zero. The associated probability is $x_\Box$. The elements in the minimal ideal $K(\mathsf{KR}(S\cup\{\Box\}, A\cup\{\Box\}))$ are of the form $w\Box$, where $w \in \mathsf{KR}(S, A)$. Since $\Box v = \Box$ for all $v \in \mathsf{KR}(S, A)$, we indeed have that $K(\mathsf{KR}(S \cup \{\Box\}, A \cup \{\Box\}))$ is left zero and hence Theorem 2.8 applies. Then [11, Corollary 2.33]

$$\Psi_w^{\mathcal{M}(\mathsf{KR}(S,A))} = \lim_{x_\Box \to 0} \Psi_w^{\mathcal{M}(\mathsf{KR}(S\cup\{\Box\},A\cup\{\Box\}))}. \tag{2.3}$$

3. Normal Distributions for Random Walks

In this section, we prove Theorem 1.4. By Theorems 2.7 and 2.8 and Eq. (2.3), the stationary distribution $\Psi_w^{\mathcal{M}(S,A)}$ is the sum of terms of the form $\prod_{a\in s} x_a$, where $s \in \mathcal{S}(S, A)$ (or limits of such expressions). In Sec. 3.1, we will explain how the semaphore code $\mathcal{S}(S, A)$ is related to the McCammond expansion $\mathsf{Mc}\circ\mathsf{KR}(S, A)$. In Sec. 3.2, we will then define the map Pict on $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ to deduce that $\Psi_w^{\mathcal{M}(S,A)}$ is a sum of normal forms. A proof of Theorem 1.4 is given in Sec. 3.3. Theorem 3.9 is a more precise version of Theorem 1.4.

3.1. *The McCammond expansion and semaphore codes*

Let us now turn to the McCammond expansion [9, 10] of the Karnofsky–Rhodes expansion of the right Cayley graph of (S, A). Recall that a simple path in $\mathsf{KR}(S, A)$ is a path that does not visit any vertex twice. Empty paths are considered simple.

Definition 3.1 (McCammond expansion). The McCammond expansion $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ of $\mathsf{KR}(S, A)$ is the graph with vertex set V, which is the set of simple paths in $\mathsf{KR}(S, A)$. The edges are given by

$$E := \{(p, a, q) \in V \times A \times V \mid \tau(q) = \tau(p)a,\ \ell(q) \leq \ell(p) + 1,$$
$$q \text{ is an initial segment of } p \text{ if } \ell(q) \leq \ell(p)\}.$$

In other words, if the path pa in $\mathsf{KR}(S, A)$ is simple, then $q = pa$. Otherwise $\tau(pa) = v$ is a vertex of p and then q is the initial segment of p up to and including v.

Remark 3.2. Note that $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ has a spanning tree T with the same vertex set as $\mathsf{Mc}\circ\mathsf{KR}(S, A)$, but only those edges $(p, a, q) \in E$ such that $\ell(q) = \ell(p) + 1$.

Example 3.3. The McCammond expansion of $\mathsf{KR}(S, A)$ of Example 2.5 is given in Fig. 1.

By Remark 3.2, the McCammond expansion $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ has a spanning tree T. In this tree, the vertices are naturally labeled by the sequence of edge labels in the path from $\mathbb{1}$ to the vertex. More concretely, if

$$p = \left(\mathbb{1} \xrightarrow{a_1} v_1 \xrightarrow{a_2} \cdots \xrightarrow{a_\ell} v_\ell\right)$$

is a path in T, then the vertex v_ℓ is naturally labeled by $a_1 \ldots a_\ell$. Hence the corresponding vertex v_ℓ in $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ has a normal form given by $a_1 \ldots a_\ell$.

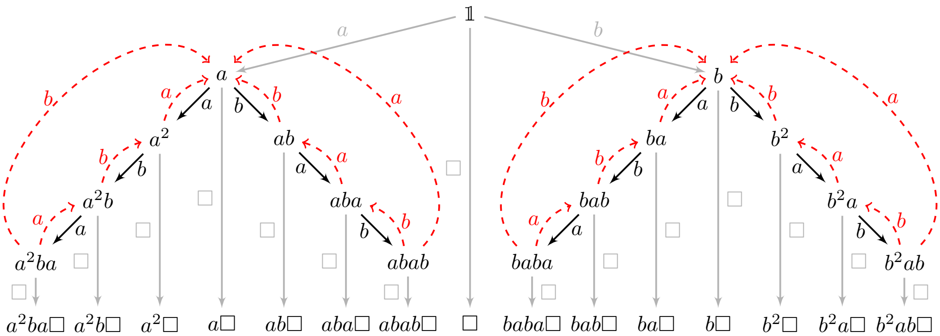


Fig. 1. (Color online) The McCammond expansion of $\mathsf{KR}(S, A)$ of Example 2.5. Transition edges are grey. The edges $(p, a, q) \in E$ with $\ell(q) = \ell(p) + 1$ are solid, whereas the edges with $\ell(q) \le \ell(p)$ are dashed (and red). The spanning tree T is obtained by removing all the dashed (red) arrows.

Remark 3.2. Also ensures that $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ has the unique simple path property, defined as follows.

Definition 3.4 (Unique simple path property). A rooted graph $(\Gamma, \mathbb{1})$ with root $\mathbb{1}$ has the unique simple path property if for each vertex v in Γ there is a unique simple path from the root $\mathbb{1}$ to v.

Elements in the semaphore code $\mathcal{S}(S, A)$ are paths in $\mathsf{Mc}\circ\mathsf{KR}(S, A)$ (rather than in T) starting at $\mathbb{1}$ and ending in $K(S)$. They are also in natural correspondence with words $a_1 \dots a_\ell \in A^+$ such that $[a_1 \dots a_\ell]_S \in K(S)$ and $[a_1 \dots a_{\ell-1}]_S \notin K(S)$. From the semaphore code, one can obtain the normal form by stripping away all loops in the path.

3.2. *Definition of* Pict

We are now going to define the map Pict from the set of tuples (Γ, p), where Γ is a graph with the unique simple path property and p is a simple path in Γ starting at $\mathbb{1}$, to the set of loop graphs. The straight line, that the loop graph is based on, will correspond to p. The map Pict was first defined by McCammond (we give a simplified definition here).

Definition 3.5 (McCammond). Let Γ be a graph with the unique simple path property and p a simple path in Γ starting at $\mathbb{1}$. Then $\mathsf{Pict}(\Gamma, p)$ is defined by the principle of induction.
Induction basis: Set $P = p$ and start at vertex $v_0 = \mathbb{1}$.
Induction step: Suppose one is at vertex $v_0 \neq \tau(p)$ on path p. Take the edge e from v_0 to v_1 in p.

(1) If there is no edge in Γ coming into v_1 besides e, continue with the unique next vertex in p, now denoted v_1 (with the current vertex v_1 relabeled v_0), unless $v_1 = \tau(p)$. If $v_1 = \tau(p)$, then output $\mathsf{Pict}(\Gamma, p) = P$.

(2) Otherwise there is at least one edge $e' \neq e$ in Γ going into v_1, given by $e' = \left(v' \xrightarrow{a} v_1\right)$ for some $a \in A$. Since Γ has the unique simple path property by assumption, there must be a unique simple path starting at $\mathbb{1}$ going to v_0 along the path p followed by the path p' starting at v_0, going along e to v_1, and ending at v'.
   (a) Run the induction on p' in a subgraph Γ' of Γ, consisting of all edges and vertices on circuits containing a vertex of p'. Note that p' is simple in Γ'. The output is $P' = \mathsf{Pict}(\Gamma', p')$.
   (b) Modify P by attaching P' disjointly except at v_1 and adding edge e' from v' in P' back to v_1.

(3) Repeat step (2) for each edge $e' \neq e$ at vertex v_1.

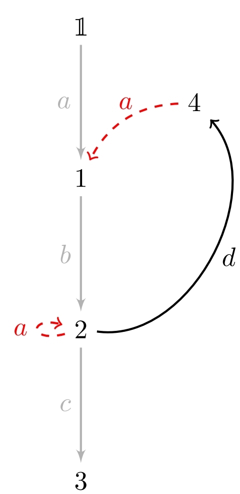
(4) Continue with the induction step unless $v_1 = \tau(p)$. If $v_1 = \tau(p)$, then output $\mathsf{Pict}(\Gamma, p) = P$.

Remark 3.6. If Γ is a rooted graph with the unique simple path property, then Γ with some edges removed (and any vertices that are no longer connected to the root $\mathbb{1}$) still has the unique simple path property. This is the case since either the unique simple path from $\mathbb{1}$ to v is still there or the vertex v is now disconnected from $\mathbb{1}$ and has hence been removed.

The graph Γ' in the Induction step (2)(a) in the definition of Pict can be obtained in two steps. First remove all incoming and outgoing edges on the vertices along the path p from $\mathbb{1}$ to v_1, except the edges on the path p itself. Remove all vertices that have become disconnected in this process. By the remark above, the resulting graph still has the unique simple path property. In this graph, all simple paths go through the vertex v_1. Hence we may make v_0 the root (removing all vertices $\mathbb{1}$ up to v_0 along p). The result is Γ', which still has the unique simple path property.

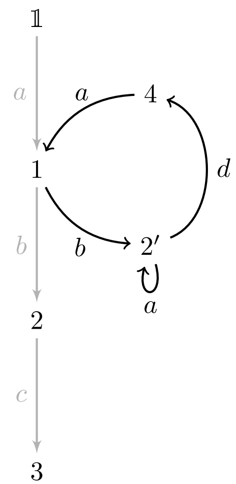
Example 3.7. Let $p = (\mathbb{1} \xrightarrow{a} 1 \xrightarrow{b} 2 \xrightarrow{c} 3)$ in

$\Gamma =$



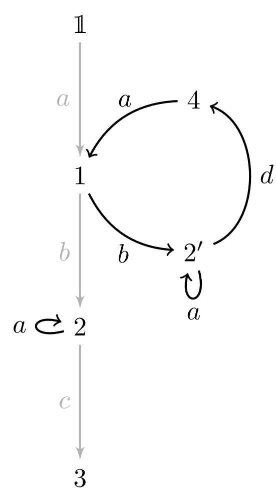
To compute $\mathsf{Pict}(\Gamma, p)$, we start with $P = p$, $v_0 = \mathbb{1}$ and $v_1 = 1$. We are in step (2) of the Induction step with $e = (\mathbb{1} \xrightarrow{a} 1)$ and $e' = (4 \xrightarrow{a} 1)$. Then $p' = (\mathbb{1} \xrightarrow{a} 1 \xrightarrow{b} 2 \xrightarrow{d} 4)$ and Γ' is Γ with the arrow labelled a from $v' = 4$ to $v_1 = 1$ removed. Also $P' = \mathsf{Pict}(\Gamma', p')$ is p' with a loop labelled a at vertex 2. Attaching P' at $v_1 = 1$ (with its vertex 2 relabelled to $2'$ to avoid repetition) and adding edge e' we obtain

$P =$



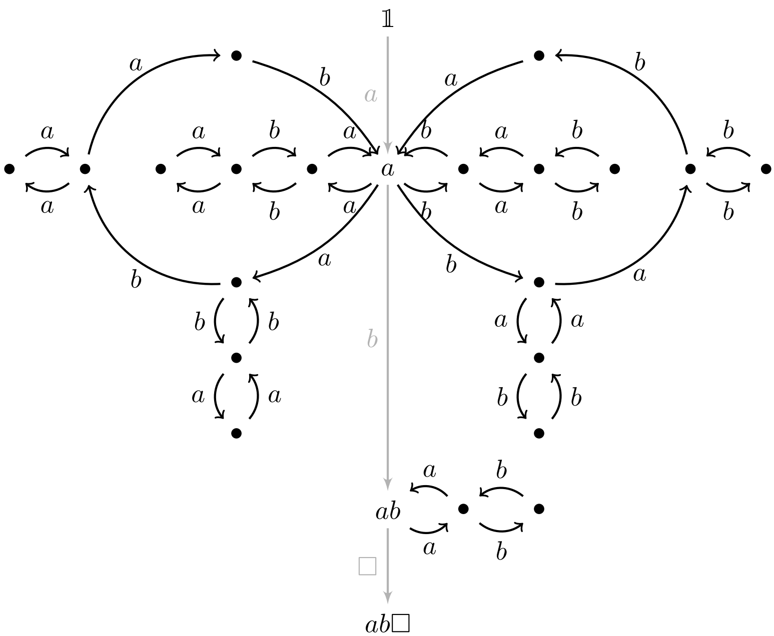
Since there are no further edges going into vertex $v_1 = 1$, we continue with the induction along p. This means that we set $v_0 = 1$, $v_1 = 2$ and $e = (1 \xrightarrow{b} 2)$. Besides e, there is only one other arrow going into $v_1 = 2$ in Γ, namely $e' = (2 \xrightarrow{a} 2)$. In this case $p' = 1 \xrightarrow{b} 2$ and Γ' is Γ with $\mathbb{1}$ and the arrows $\mathbb{1} \xrightarrow{a} 1$, $4 \xrightarrow{a} 1$ and $2 \xrightarrow{a} 2$ removed. Hence the new P with $P' = \mathsf{Pict}(\Gamma', p')$ added is

$\mathsf{Pict}(\Gamma, p) = P =$



The remaining induction steps do not change this P, which is hence also $\mathsf{Pict}(\Gamma, p)$.

Example 3.8. Consider the McCammond expansion $\Gamma = \mathsf{Mc} \circ \mathsf{KR}(S, A)$ of Example 3.3 (see also Fig. 1) and the path in the McCammond tree T given by $ab\square$. Then $\mathsf{Pict}(\Gamma, ab\square)$ is given by



Following the algorithm explained in Sec. 1.3, a Kleene expression for $\mathcal{P}_{\mathsf{Pict}(\Gamma, ab\square)}$ is given by

$$L = a\{\ell_1, \ell_2, \ell_3, \ell_4\}^\star b \ell_5^\star \square,$$

where

$$\ell_1 = a(b(aa)^\star b)^\star b(aa)^\star ab,$$

$$\ell_2 = a(b(aa)^\star b)^\star a,$$

$$\ell_3 = b(a(bb)^\star a)^\star a(bb)^\star ba,$$

$$\ell_4 = b(a(bb)^\star a)^\star b,$$

$$\ell_5 = a(bb)^\star a.$$

Hence

$$\Psi_{\mathsf{Pict}(\Gamma, ab\square)} = \frac{x_a x_b x_\square}{\left(1 - \frac{x_a^2 x_b^2}{\left(1 - \frac{x_b^2}{1-x_a^2}\right)(1-x_a^2)} - \frac{x_a^2}{1 - \frac{x_b^2}{1-x_a^2}} - \frac{x_a^2 x_b^2}{\left(1 - \frac{x_a^2}{1-x_b^2}\right)(1-x_b^2)} - \frac{x_b^2}{1 - \frac{x_a^2}{1-x_b^2}}\right) \times \left(1 - \frac{x_a^2}{1-x_b^2}\right)}$$

$$= \frac{x_a x_b x_\square (1 - x_b^2)}{\left(1 - \frac{2x_a^2 x_b^2}{1-x_a^2-x_b^2} - \frac{x_a^2(1-x_a^2)}{1-x_a^2-x_b^2} - \frac{x_b^2(1-x_b^2)}{1-x_a^2-x_b^2}\right)(1 - x_a^2 - x_b^2)},$$

$$= \frac{x_a x_b x_\square (1 - x_b^2)}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2}.$$

Using that $x_a + x_b + x_\square = 1$, we find that in the limit $x_\square \to 0$

$$\lim_{x_\square \to 0} \Psi_{\mathsf{Pict}(\Gamma, ab\square)} = \frac{1}{8}(1 - x_b^2).$$

In a similar fashion, we find

$$\Psi_\square = x_\square \xrightarrow{x_\square \to 0} 0,$$

$$\Psi_{a\square} = \frac{x_a(1 - x_a^2 - x_b^2)x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a}{4},$$

$$\Psi_{aba\square} = \frac{x_a^2 x_b x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a}{8},$$

$$\Psi_{abab\square} = \frac{x_a^2 x_b^2 x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a x_b}{8},$$

$$\Psi_{a^2\square} = \frac{x_a^2(1 - x_a^2)x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a(1 + x_a)}{8},$$

$$\Psi_{a^2b\square} = \frac{x_a^2 x_b x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a}{8},$$

$$\Psi_{a^2ba\square} = \frac{x_a^3 x_b x_\square}{1 - 2x_a^2 - 2x_b^2 + (x_a^2 - x_b^2)^2} \xrightarrow{x_\square \to 0} \frac{x_a^2}{8}.$$

The stationary probabilities for the elements with a and b interchanged are obtained by symmetry. It is not hard to check that these probabilities sum to one as desired.

As noted in the introduction, $\mathsf{Pict}(\Gamma, p)$ is not necessarily deterministic. There can be several arrows leaving a vertex labeled by the same element $a \in A$. For example, vertex 1 in Example 3.7 has two arrows labeled b coming out.

One can make a non-deterministic automata $\mathcal{A}$ deterministic as follows. If $\mathcal{A}$ has states Q with start state $\mathbb{1}$ and final states F not containing $\mathbb{1}$, we make a deterministic automata $\mathsf{det}(\mathcal{A})$ accepting the same strings going from $\mathbb{1}$ to a member of F as follows. The states Q' of $\mathsf{det}(\mathcal{A})$ are the collection of subsets of Q determined a follows:

- $\{\mathbb{1}\}$ is in Q';
- if $Z \in Q'$, then $Z.a \in Q'$ for $a \in A$, where $Z.a = \{q \mid z \xrightarrow{a} q \in \mathcal{A} \text{ where } z \in Z\}$.

One continues by induction until the process adds no new subsets. For $\mathsf{det}(\mathcal{A})$, start in state $\{\mathbb{1}\}$. The final states are all the states of $\mathsf{det}(\mathcal{A})$ such that the intersection with F is non-empty.

With this definition, making $\mathsf{Pict}(\Gamma, p)$ deterministic gives the automata for (Γ, p) back.

3.3. *Proof of Theorem* 1.4

As explained in Sec. 2.1, any finite Markov chain $\mathcal{M}$ can be described as a Markov chain $\mathcal{M}(S, A)$ in terms of a finite semigroup S with generators A. Since by Theorem 2.7, $\Psi_w^{\mathcal{M}(S,A)}$ is the sum over $\Psi_v^{\mathcal{M}(\mathsf{KR}(S,A))}$, it suffices to prove the statement of Theorem 1.4 for $\Psi_v^{\mathcal{M}(\mathsf{KR}(S,A))}$. When $K(S)$ is not left zero, we may use the limiting construction of (2.3) to obtain $\Psi_v^{\mathcal{M}(\mathsf{KR}(S,A))}$ from the case in which the minimal ideal is left zero. Assuming that $K(S)$ is left zero, we have by Theorem 2.8

$$\Psi_w^{\mathcal{M}(\mathsf{KR}(S,A))} = \sum_{\substack{s \in \mathcal{S}(S,A) \\ [s]_{\mathsf{KR}(S,A)} = w}} \prod_{a \in s} x_a \quad \text{for all } w \in K(\mathsf{KR}(S, A)). \tag{3.1}$$

As explained in Sec. 3.1, there is a normal form associated to each semaphore code element $s \in \mathcal{S}(S, A)$. Namely, s is a path in $\mathsf{Mc} \circ \mathsf{KR}(S, A)$ starting at $\mathbb{1}$ and the normal form is the simple path with all loops stripped away from s; equivalently the normal form is the path in T starting at $\mathbb{1}$ and ending at $\tau(s)$, where T is the tree associated to the McCammond expansion $\mathsf{Mc} \circ \mathsf{KR}(S, A)$. In the tree T, a path p starting at $\mathbb{1}$ is also naturally in bijection with its endpoint $\tau(p)$. Hence we may identify vertex $t \in \mathsf{T}$ with the path from $\mathbb{1}$ to t in T or equivalently with the simple path from $\mathbb{1}$ to t in $\mathsf{Mc} \circ \mathsf{KR}(S, A)$. Therefore, we may rewrite the sum in (3.1) as

$$\Psi_w^{\mathcal{M}(\mathsf{KR}(S,A))} = \sum_{\substack{t \in \mathsf{T} \\ [t]_{\mathsf{KR}(S,A)} = w}} \left(\sum_{\substack{s \in \mathcal{S}(S,A) \\ \tau(s) = t}} \prod_{a \in s} x_a \right) \qquad \text{for all } w \in K(\mathsf{KR}(S, A)). \tag{3.2}$$

We claim that for a given $t \in \mathsf{T}$ with $[t]_{\mathsf{KR}(S,A)} \in K(\mathsf{KR}(S, A))$

$$\Psi_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)} = \sum_{\substack{s \in \mathcal{S}(S,A) \\ \tau(s) = t}} \prod_{a \in s} x_a. \tag{3.3}$$

Recall that by Definition 1.3

$$\Psi_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)} = \sum_{p \in \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)}} \prod_{a \in p} x_a.$$

Hence, (3.3) can be proved by establishing a bijection

$$\varphi \colon \{s \in \mathcal{S}(S, A) \mid \tau(s) = t\} \longrightarrow \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)}. \tag{3.4}$$

In fact, we are going to prove a slight generalization of (3.4). Namely, for any $t \in \mathsf{T}$ we will show that there is a bijection

$$\varphi \colon \{s \in \mathcal{P}_{\mathsf{Mc} \circ \mathsf{KR}(S,A)} \mid \tau(s) = t\} \longrightarrow \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)}, \tag{3.5}$$

where $\mathcal{P}_{\mathsf{Mc}\circ\mathsf{KR}(S,A)}$ is the set of paths in $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ starting at $\mathbb{1}$. Then (3.4) is the special case when $[t]_{\mathsf{KR}(S,A)} \in K(\mathsf{KR}(S,A))$.

To define φ in (3.5), fix $t = a_1 \ldots a_k$, where $a_i \in A$ are the labels in the path in T. A path $s \in \mathcal{P}_{\mathsf{Mc}\circ\mathsf{KR}(S,A)}$ with $\tau(s) = t$, can be viewed as t with circuits $\ell_j^{(j)}$ interspersed. More precisely,

$$s = a_1 \left(\prod_{j\in J_1} \ell_1^{(j)} \right) a_2 \left(\prod_{j\in J_2} \ell_2^{(j)} \right) \cdots a_k \left(\prod_{j\in J_k} \ell_k^{(j)} \right),$$

where $\tau(a_1 \ldots a_i) = \tau(a_1 \ldots a_i \ell_i^{(j)})$ for all $1 \leq i \leq k$ and $j \in J_i$ and any initial subsequence of $\ell_i^{(j)}$ does not reach the vertex $a_1 \ldots a_i$. Here the sets J_i index the set of circuits $\{\ell_i^{(j)} \mid j \in J_i\}$ at vertex $a_1 \ldots a_i$ and either $J_i = \{1, 2, \ldots, n_i\}$ is a finite set or $J_i = \{1, 2, 3, \ldots\}$ is the set of positive integers. In other words, each $\ell_i^{(j)}$ is a circuit from vertex $a_1 \ldots a_i$ to itself, which does not pass through $a_1 \ldots a_i$ otherwise. The last step of $\ell_i^{(j)}$ is an edge in $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ that is not in T. Suppose by induction that

$$s' = a_1 \left(\prod_{j\in J_1} \ell_1^{(j)} \right) \cdots a_i \left(\prod_{j\in J_i'} \ell_i^{(j)} \right),$$

where $1 \leq i \leq k$ and $J_i' = \{1, 2, \ldots, n_i'\} \subseteq J_i$ or $J_i' = J_i$, is mapped to π in $\mathsf{Pict}(\mathsf{Mc}\circ\mathsf{KR}(S,A), a_1 \ldots a_i)$ under φ. We need to distinguish two cases.

Case $\boldsymbol{J_i' \subsetneq J_i}$. Let j be the smallest element in $J_i \setminus J_i'$. Recall that $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ has the unique simple path property. Hence the path p' in $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ from $v_0 = a_1 \ldots a_{i-1}$ through $v_1 = a_1 \ldots a_i$ to v', which is $a_1 \ldots a_i \ell_i^{(j)}$ with the last edge e' removed is a path in Γ' in the notation of Sec. 3.2. By induction this path is mapped to π' in $\mathcal{P}_{\mathsf{Pict}(\Gamma', p')}$. Hence

$$\varphi(s' \ell_i^{(j)}) = \pi\pi' \in \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc}\circ\mathsf{KR}(S,A), a_1 \ldots a_i)}$$

This corresponds to the induction step (2) in Definition 3.5.

Case $\boldsymbol{J_i' = J_i}$. If $i = k$, we are done. If $i < k$, we define

$$\varphi(s' a_{i+1}) = \pi a_{i+1} \in \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc}\circ\mathsf{KR}(S,A), a_1 \ldots a_{i+1})},$$

which is a well-defined path since the last step is along the straight line path and hence unique. This corresponds to the induction step (1) (if $J_i = \emptyset$) or step (4) (if $J_i \neq \emptyset$) in Definition 3.5.

This shows that φ is a well-defined map. It has an inverse φ^{-1} by mapping a path $\pi \in \mathcal{P}_{\mathsf{Pict}(\mathsf{Mc}\circ\mathsf{KR}(S,A), t)}$ to a path in $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ by just reading the labels of the edges. This indeed gives a path in $\mathsf{Mc}\circ\mathsf{KR}(S,A)$ by the construction of Pict.

Combining (3.2) and (3.3), we obtain

$$\Psi_w^{\mathcal{M}(\mathsf{KR}(S,A))} = \sum_{\substack{t \in \mathsf{T} \\ [t]_{\mathsf{KR}(S,A)} = w}} \Psi_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)},$$

which proves Theorem 1.4 since $\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)$ is a loop graph.

In summary, we proved the following theorem, which is a more detailed version of Theorem 1.4.

Theorem 3.9. *Let $\mathcal{M}(S,A)$ be a Markov chain associated to the finite semigroup with generators (S,A). If $K(S)$ is left zero, the stationary distribution is given by*

$$\Psi_w^{\mathcal{M}(S,A)} = \sum_{\substack{t \in \mathsf{T} \\ [t]_S = w}} \Psi_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S,A),t)} \quad \textit{for } w \in K(S),$$

where T is the spanning tree of $\mathsf{Mc} \circ \mathsf{KR}(S,A)$. Otherwise

$$\Psi_w^{\mathcal{M}(S,A)} = \sum_{\substack{t \in \mathsf{T} \\ [t]_S = w\square}} \lim_{x_\square \to 0} \Psi_{\mathsf{Pict}(\mathsf{Mc} \circ \mathsf{KR}(S \cup \{\square\}, A \cup \{\square\}),t)} \quad \textit{for } w \in K(S),$$

where T is the spanning tree of $\mathsf{Mc} \circ \mathsf{KR}(S \cup \{\square\}, A \cup \{\square\})$ and $\square$ acts as zero.

Acknowledgments

We are grateful to Jon McCammond and Ben Steinberg for discussions. The map Pict of Definition 3.5 is due to McCammond, told to the first author in 1994, written by the first author in 2008, and simplified here.

The first author thanks the Simons Foundation Collaboration Grants for Mathematicians for travel grant #313548. The second author was partially supported by NSF grants DMS–1760329 and DMS–1764153.

References

[1] A. Ayyer, A. Schilling, B. Steinberg and N. M. Thiéry, Markov chains, $\mathscr{R}$-trivial monoids and representation theory, *Internat. J. Algebra Comput.* **25**(1–2) (2015) 169–231.

[2] J. Berstel, D. Perrin and C. Reutenauer, *Codes and Automata*, Encyclopedia of Mathematics and its Applications, Vol. 129 (Cambridge University Press, Cambridge, 2010).

[3] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups, Vol. I*. Mathematical Surveys, No. 7 (American Mathematical Society, Providence, R.I., 1961).

[4] G. Högnäs and A. Mukherjea, *Probability Measures on Semigroups. Convolution Products, Random Walks, and Random Matrices*, Probability and its Applications. 2nd eds. edn. (Springer, New York, 2011).

[5] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Undergraduate Texts in Mathematics (Springer-Verlag, New York-Heidelberg, 1976).

[6] K. Krohn, J. Rhodes and B. Tilson, *Algebraic theory of machines, languages, and semigroups*, Caps. 2, eds. M. A. Arbib, K. Krohn, and J. L. Rhodes (Academic Press, New York, 1968) pp. 5–9.

[7] S. Margolis, F. Saliola, and B. Steinberg, Combinatorial topology and the global dimension of algebras arising in combinatorics, *J. Eur. Math. Soc.* (JEMS) **17**(12) (2015) 3037–3080.

[8] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov Chains and Mixing Times*, eds. J. G. Propp and D. B. Wilson (American Mathematical Society, Providence, RI, 2009).

[9] J. P. McCammond, Normal forms for free aperiodic semigroups. *Internat. J. Algebra Comput.* **11**(5) (2001) 581–625.

[10] J. McCammond, J. Rhodes and B. Steinberg, Geometric semigroup theory, preprint (2011), arxiv: 1104.2301.

[11] J. Rhodes and A. Schilling, Unified theory for finite Markov chains, *Adv. Math.* **347** (2019) 739–779.