Séminaire Lotharingien de Combinatoire **85B** (2021)
Article #27, 12 pp.

Proceedings of the 33rd Conference on Formal Power Series and Algebraic Combinatorics (Ramat Gan)

Mixing time for Markov chain on linear extensions

John Rhodes[1] and Anne Schilling[*2]

[1]*Department of Mathematics, University of California, Berkeley, CA 94720, U.S.A.*
[2]*Department of Mathematics, UC Davis, One Shields Ave., Davis, CA 95616-8633, U.S.A.*

Abstract. We provide a general framework for computing mixing times of finite Markov chains when its minimal ideal is left zero. Our analysis is based on combining results by Brown and Diaconis with our previous work on stationary distributions of finite Markov chains. We introduce a new Markov chain on linear extensions of a poset with n vertices, which is a variant of the promotion Markov chain of Ayyer, Klee and the last author, and show that it has a mixing time $O(n \log n)$.

Keywords: Markov chains, mixing time, posets, linear extensions

1 Introduction

A *Markov chain* is a model that describes transitions between states in a state space according to certain probabilistic rules. The defining characteristic of a Markov chain is that the transition from one state to another only depends on the current state and the elapsed time, but not how the system arrived there. In other words, a Markov chain is "memoryless". Markov chains have an abundance of applications, from data analysis, population dynamics to traffic models.

For a Markov chain, the *stationary distribution* Ψ is the long-term limiting distribution. Mathematically speaking, it is the eigenvector of the transition matrix T of the Markov chain with eigenvalue one. That is $T\Psi = \Psi$. An important question is how quickly does the Markov chain converge to the stationary distribution. In Markov chain theory, distance is usually the total variation distance. If Ω is the state space, the total variation distance between two probability distributions ν and μ is defined as

$$\|\nu - \mu\| = \max_{A \subseteq \Omega} |\nu(A) - \mu(A)|.$$

For a given small $\epsilon > 0$, the *mixing time* t_{mix} is the smallest t such that

$$\|T^t \nu - \Psi\| \leqslant \epsilon,$$

independent of the initial distribution ν.

*anne@math.ucdavis.edu. Anne Schilling was partially supported by NSF grants DMS–1760329 and DMS–1764153.

In seminal work of Bidigare, Hanlon and Rockmore [5], which was continued by Diaconis, Brown, Athanasiadis, Björner, Chung and Graham, amongst others [11, 6, 10, 7, 8, 1, 14, 30], the special family of semigroups, now known as *left regular bands* first studied by Schützenberger [31] in the 1940s, was applied to random walks or Markov chains on hyperplane arrangements. In his 1998 ICM lecture [15], Diaconis discussed these developments. In Section 4.1, entitled *What is the ultimate generalization?*, he asks how far the semigroup techniques can be taken.

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 for example [22, Proposition 1.5] and [3, Theorem 2.3]. In this setting, there is a transition $s \stackrel{a}{\longrightarrow} s'$ with probability $0 \leqslant x_a \leqslant 1$, where $s, s' \in \Omega$, $a \in S$ and $s' = a.s$ is the action of a on the state s. It is enough to consider the semigroup S generated by the elements $a \in A$ with $x_a > 0$.

In the pursuit of finding Diaconis' ultimate generalization [15], the arguments in Brown and Diaconis [11] were generalized to Markov chains for $\mathscr{R}$-trivial semigroups [3]. In [25, 26], the current authors developed a general theory for computing the stationary distribution for any finite Markov chain. The theory uses semigroup methods such as the Karnofsky–Rhodes and McCammond expansion of a semigroup. These expansions give rise to loop graphs which immediately yield Kleene expressions for all paths from the root of the graph to elements in the minimal ideal of the semigroup. The Kleene expressions in turn give rational expressions for the stationary distribution.

In this paper we apply the findings of [25, 26] to study bounds on the mixing time of the Markov chain. In particular, Theorems 2.4 and 2.5 provide upper bounds for the mixing time directly from the rational expression of the stationary distribution in the case when the minimal ideal of the semigroup is left zero. In Section 3 we consider specific examples such as the Tsetlin library [13] and a new Markov chain on linear extensions of a finite poset with n vertices, which is a variant of the promotion Markov chain introduced in [2]. The new Markov chain on linear extensions has a mixing time of $O(n \log n)$ as compared to the mixing time of the model of Bubley and Dyer [12] with mixing time $O(n^3 \log n)$.

A long version of this paper has appeared in [27].

2 Mixing time

Let T be the *transition matrix* of a finite Markov chain. Assuming that the Markov chain is *ergodic* (meaning that it is irreducible and aperiodic), by the Perron–Frobenius Theorem there exists a unique *stationary distribution* Ψ and $T^t \nu$ converges to Ψ as $t \to \infty$ for any initial state ν. A Markov chain is irreducible if the graph of the Markov chain is strongly connected. It is aperiodic if the gcd of the cycle lengths in the graph of the Markov chain is one.

2.1 Upper bound

Brown and Diaconis [11] [10, Theorem 0] showed, for Markov chains associated to left regular bands, that the total variational distance from stationarity after t steps is bounded above by the probability $\Pr(\tau > t)$, where τ is the first time that the walk hits a certain ideal. The arguments in Brown and Diaconis [11] can be generalized to arbitrary finite Markov chains (not just those related to left regular bands). To state the details, we need some more notation. Let $\mathcal{M}(S, A)$ be a finite state Markov chain with state space Ω and transition matrix T associated to the semigroup S with generators A with probabilities $0 < x_a \leqslant 1$ for $a \in A$.

A two-sided *ideal* I (or ideal for short) is a subset $I \subseteq S$ such that $uIv \subseteq I$ for all $u, v \in S^{\mathbb{1}}$, where $S^{\mathbb{1}}$ is the semigroup S with identity $\mathbb{1}$ added (even if S already contains an identity). 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 nonempty minimal ideal denoted $K(S)$.

Assume that the minimal ideal $K(S)$ is *left zero*, that is, $xy = x$ for all $x, y \in K(S)$. This assumption implies that the Markov chain on the minimal ideal (given by the left action) is ergodic. Let τ be the random variable which is the time that the random walk is absorbed into the minimal ideal $K(S)$.

Theorem 2.1 ([3]). *Let S be a finite semigroup whose minimal ideal $K(S)$ is a left zero semigroup and let T be the transition matrix of the associated Markov chain. Then $\|T^t \nu - \Psi\| \leqslant \Pr(\tau > t)$.*

2.2 Ideals and semaphore codes

Let A be a finite alphabet, A^+ the set of all nonempty words in the alphabet A, and $A^\star$ the set of all words in the alphabet A. As shown in [28], ideals in A^+ are in bijection with semaphore codes [4]. A *prefix code* is a subset of A^+ such that all elements are incomparable in prefix order (meaning that no element is the prefix of any other element of the code). A *semaphore code* $\mathcal{S}$ is a prefix code such that $A\mathcal{S} \subseteq \mathcal{S}A^\star$. There is a natural left action on a semaphore code. If $u \in \mathcal{S} \subseteq A^+$ and $a \in A$, then au has a prefix in $\mathcal{S}$ (and hence a unique prefix of au). The left action $a.u$ is the prefix of au that is in $\mathcal{S}$. Assigning probability $0 \leqslant x_a \leqslant 1$ to $a \in A$, the left action on a semaphore code $\mathcal{S}$ defines a Markov chain with a countable state space $\mathcal{S}$.

The bijection between ideals $I \subseteq A^+$ and semaphore codes $\mathcal{S}$ over A is given as follows (see [28, Proposition 4.3]). If $u = a_1 a_2 \ldots a_j \in I \subseteq A^+$, find the (necessarily unique) index $1 \leqslant i \leqslant j$ such that $a_1 \ldots a_{i-1} \notin I$, but $a_1 \ldots a_i \in I$. Then $a_1 \ldots a_i$ is a code word and the set of all such words forms the semaphore code $\mathcal{S}$. Conversely, given a semaphore code $\mathcal{S}$, the corresponding ideal is $\mathcal{S}A^\star$.

In this setting, τ can be interpreted as the random variable given by the length of the semaphore code words. Let $\mathcal{S}$ be a semaphore code and I the ideal under the bijection described above. A semaphore code word $s = s_1 s_2 \dots s_\ell$ has the property that $s \in I$, but $s_1 \dots s_{\ell-1} \notin I$. Hence τ can be interpreted as the random variable given by the length ℓ.

2.3 Rational expressions for stationary distributions

Let $\mathcal{M}(S, A)$ be the Markov chain associated to the finite semigroup S with generators in A. Assume that its minimal ideal $K(S)$ is left zero, so that $K(S)$ can be taken as the state space Ω of the Markov chain. Denote by $\mathcal{S}(S, A)$ the semaphore code associated to $K(S)$ (see Section 2.2). For a word $s \in A^+$, we denote by $[s]_S$ the image of the word in the alphabet A in S. The following theorem is stated in [26, Corollaries 2.23 & 2.28].

Theorem 2.2 ([26]). *If $K(S)$ is left zero, the stationary distribution of the Markov chain $\mathcal{M}(S, A)$ labeled by $w \in K(S)$ is given by*

$$\Psi_w(x_1, \dots, x_n) = \sum_{\substack{s \in \mathcal{S}(S,A) \\ [s]_S = w}} \prod_{a \in s} x_a. \tag{2.1}$$

In [25, 26], we developed a strategy using loop graphs to compute the expressions in Theorem 2.2 as rational functions in the probabilities x_a for $a \in A$. This is done in several steps:

1. We used the McCammond and Karnofsky–Rhodes expansion $\mathsf{Mc} \circ \mathsf{KR}(S, A)$ of the right Cayley graph $\mathsf{RCay}(S, A)$ of the semigroup S with generators A. In this paper we do not require the details of these definitions, except that the right Cayley graph as well as its expansions are rooted graphs with root $\mathbb{1}$. The Karnofsky–Rhodes expansion is another right Cayley graph, whereas the McCammond expansion is only an automata. For the precise definition of the Karnofsky–Rhodes expansion, we refer the reader to [24, Definition 4.15], [23, Section 3.4], [25, Section 2.4], and [29, Section 2]. For the definition of the McCammond expansion, we refer the reader to [24, Section 2.7] and [25, Section 2.5]. The Markov chain $\mathcal{M}(S, A)$ is a *lumping* [22] of the Markov chains associated to the expansions.

2. The stationary distributions of the Markov chains associated to the expansions can be expressed using *loop graphs* G, see [25]. A loop graph is a straight line path from $\mathbb{1}$ to an endpoint s with directed loops of any finite length attached recursively to any vertex (besides $\mathbb{1}$ and s). In this way [25, Theorem 1.4]

$$\Psi_w(x_1, \dots, x_n) = \sum_G \Psi_G(x_1, \dots, x_n), \tag{2.2}$$

where the sum is over certain loop graphs G with end point s such that $[s]_S = w$. Here [25, Definition 1.3]

$$\Psi_G(x_1, \dots, x_n) = \sum_p \prod_{a \in p} x_a, \tag{2.3}$$

where the sum is over all path p in G starting at $\mathbb{1}$ and ending in s.

3. There is a *Kleene expression* for the set of all paths from $\mathbb{1}$ to s in G. The Kleene expression immediately yields a rational expression for the stationary distribution $\Psi_G(x_1, \ldots, x_n)$ and hence $\Psi_w(x_1, \ldots, x_n)$ by (2.2).

Remark 2.3. An important property of the above construction is that in the series expansion of the rational expression for $\Psi_w(x_1, \ldots, x_n)$ (resp. $\Psi_G(x_1, \ldots, x_n)$) the total degree of each term corresponds to the length of the underlying semaphore code word in (2.1) (resp. the underlying path in G in (2.3)).

2.4 Mixing time via truncation of Kleene expressions

As stated in Theorem 2.1, $\mathsf{Pr}(\tau \geqslant t)$ provides an upper bound on the mixing time in the setting that $K(S)$ is left zero. As discussed in Section 2.2, τ can be interpreted as the random variable given by the length of the semaphore code words or paths in the loop graph. To compute $\mathsf{Pr}(\tau \geqslant t)$, one needs to compute the sum of probabilities of all paths of length weakly greater than t. By Remark 2.3, the length of the paths is given by the total degree in the probability variables $x_1, \ldots, x_n$ for the generators $a_1, \ldots, a_n$ of the semigroup S. Hence we obtain $\mathsf{Pr}(\tau \geqslant t)$ by truncating the rational function for the stationary distribution to total degree weakly bigger than t.

Let $\Psi_w^{\geqslant t}(x_1, \ldots, x_n)$ be the truncation of the formal power series associated to the rational function $\Psi_w(x_1, \ldots, x_n)$ to terms of degree weakly bigger than t In addition, let $\Psi_w^{<t}(x_1, \ldots, x_n)$ be the truncation of the formal power series associated to the rational function $\Psi_w(x_1, \ldots, x_n)$ to terms of degree strictly smaller than t. Note that

$$\Psi_w(x_1, \ldots, x_n) = \Psi_w^{<t}(x_1, \ldots, x_n) + \Psi_w^{\geqslant t}(x_1, \ldots, x_n).$$

Theorem 2.4 ([27])**.** *Suppose the Markov chain satisfies the conditions of Theorem 2.1. If $\Psi_w(x_1, \ldots, x_n)$ is represented by a rational function such that each term of degree ℓ in its formal power sum expansion corresponds to a semaphore code word s of length ℓ with $[s]_S = w$, we have*

$$\mathsf{Pr}_w(\tau \geqslant t) = \frac{\Psi_w^{\geqslant t}(x_1, \ldots, x_n)}{\Psi_w(x_1, \ldots, x_n)} = 1 - \frac{\Psi_w^{<t}(x_1, \ldots, x_n)}{\Psi_w(x_1, \ldots, x_n)}.$$

For each $w \in K(S)$, we can also give an explicit formula for the expected number of steps $E_w[\tau]$ it takes to reach the endpoint of w using the Cauchy–Euler operator. By Markov's inequality [22] $\mathsf{Pr}(\tau > t) \leqslant \frac{E[\tau]}{t+1}$, this also yields a bound on the mixing time.

Theorem 2.5 ([27])**.** *With the same assumptions as in Theorem 2.4, we have*

$$E_w[\tau] = \left(\sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i} \right) \ln \Psi_w(x_1, \ldots, x_n).$$

Remark 2.6. Note that the formal expression for $\Psi_w(x_1, \ldots, x_n)$ cannot be manipulated using that $x_1 + \cdots + x_n = 1$ when using Theorems 2.4 and 2.5.

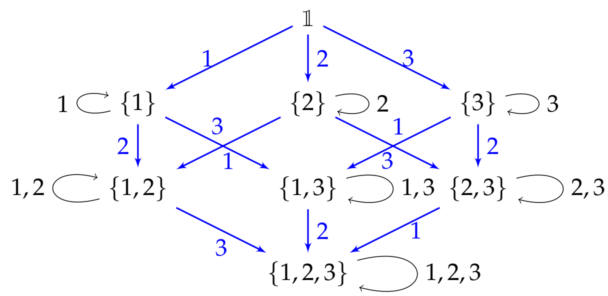


Figure 1: The right Cayley graph $\mathsf{RCay}(S, A)$ with $S = P(3)$ and $A = \{1,2,3\}$.

3 Markov chain on linear extensions

3.1 The Tsetlin library

The Tsetlin library [13] is a Markov chain whose states are all permutations S_n of n books (on a shelf). Given $\pi \in S_n$, construct $\pi' \in S_n$ from π by removing book a from the shelf and inserting it to the front. In this case write $\pi \xrightarrow{a} \pi'$. Let $0 < x_a \leqslant 1$ be probabilities for each $1 \leqslant a \leqslant n$ such that $\sum_{a=1}^{n} x_a = 1$. In the Tsetlin library Markov chain, we transition $\pi \xrightarrow{a} \pi'$ with probability x_a. The stationary distribution for the Tsetlin library was derived by Hendricks [18, 17]

$$\Psi_\pi = \prod_{i=1}^{n} \frac{x_{\pi_i}}{1 - \sum_{j=1}^{i-1} x_{\pi_j}} \qquad \text{for all } \pi \in S_n. \tag{3.1}$$

In [26, Section 3.1], the stationary distribution was derived using right Cayley graphs and their Karnofsky–Rhodes and McCammond expansions.

Consider the semigroup $P(n)$, which consists of the set of all non-empty subsets of $[n] := \{1, 2, \ldots, n\}$. Multiplication in $P(n)$ is union of sets. We pick as generators $[n]$. Then the right Cayley graph $\mathsf{RCay}(P(n), [n])$ is the Boolean poset with $\mathbb{1}$ as root. The right Cayley graph for $P(3)$ is depicted in Figure 1. Except for the loops at a given vertex, all edges are transitional. Hence $\mathsf{Mc} \circ \mathsf{KR}(P(n), [n]) = \mathsf{KR}(P(n), [n])$ is a tree with leaves given by the permutations S_n of $[n]$. The case $n = 3$ is depicted in Figure 2.

To obtain a bound on the mixing time, we compute $E[\tau]$ from the Karnofsky–Rhodes expansion of the right Cayley graph. The ideal consists of the leaves of $\mathsf{KR}(P(n), [n])$, which are labeled by permutations in S_n. The Kleene expression for all paths from $\mathbb{1}$ to $12 \ldots n$ is given by $11^\star 2\{1,2\}^\star 3\{1,2,3\}^\star \ldots \{1, 2, \ldots, n-1\}^\star n$. Hence we obtain (compare with (3.1))

$$\Psi_{12\ldots n}(x_1, \ldots, x_n) = \frac{x_1 \cdots x_n}{(1 - x_1)(1 - x_1 - x_2) \cdots (1 - x_1 - \cdots - x_{n-1})}$$

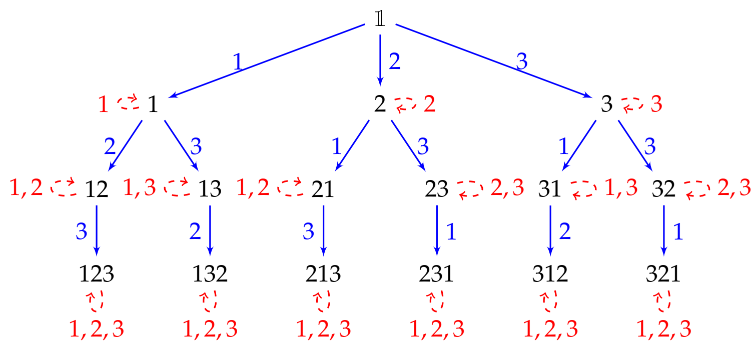


Figure 2: The Karnofsky–Rhodes expansion of the right Cayley graph of Figure 1.

and by Theorem 2.5

$$E_{12\ldots n}[\tau] = n + \frac{x_1}{1-x_1} + \frac{x_1+x_2}{1-x_1-x_2} + \cdots + \frac{x_1+\cdots+x_{n-1}}{1-x_1-\cdots-x_{n-1}}. \tag{3.2}$$

If $x_i = \frac{1}{n}$ for all $1 \leqslant i \leqslant n$, we hence have

$$E_G[\tau] = n + \frac{1}{n-1} + \frac{2}{n-2} + \cdots + \frac{n-1}{1} = n\left(\sum_{i=1}^{n} \frac{1}{i}\right). \tag{3.3}$$

The last equality can be proved by induction on n. It is well-known that the sequence $t_n = \sum_{i=1}^{n} \frac{1}{i} - \ln(n)$ approaches the Euler–Mascheroni constant γ as $n \to \infty$. Therefore $E[\tau] = E_G[\tau] \leqslant n\ln(n) + n\gamma$ and by Markov's inequality $\|T^t\nu - \pi\| \leqslant \frac{n\ln(n)+n\gamma}{t+1}$.

3.2 Markov chain on linear extensions

Let P be a *partially ordered set*, also known as a *poset*, on n elements with partial order $\preccurlyeq$. A partial order must be reflexive ($a \preccurlyeq a$ for all $a \in P$), antisymmetric ($a \preccurlyeq b$ and $b \preccurlyeq a$ implies $a = b$ for $a, b \in P$), and transitive ($a \preccurlyeq b$ and $b \preccurlyeq c$ implies $a \preccurlyeq c$ for $a, b, c \in P$). We assume that the elements of P are labeled by integers in $[n] := \{1, 2, \ldots, n\}$ such that if $i, j \in P$ with $i \preccurlyeq j$ then $i \leqslant j$ as integers. Let $\mathcal{L} := \mathcal{L}(P)$ be the set of *linear extensions* of P defined as $\mathcal{L}(P) = \{\pi \in S_n \mid i \prec j \text{ in } P \implies \pi_i^{-1} < \pi_j^{-1} \text{ as integers}\}$.

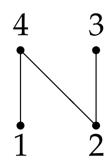
In computer science, linear extensions are also known as *topological sortings* [20, 21]. Computing the number of linear extensions is an important problem for real world applications [19]. For example, it relates to sorting algorithms. Suppose one wants to schedule a sequence of tasks based on their dependencies. Specifying that a certain task has to come before another task gives rise to a partial order. A linear extension gives a total order in which to perform the jobs. A recursive formula for the number of linear extensions for a given poset P was given in [16]. Brightwell and Winkler [9] showed that counting the number of linear extensions is #P-complete. Bubley and Dyer [12] provided an algorithm to (almost) uniformly sample the set of linear extensions of a finite poset of size n with mixing time $O(n^3 \log n)$. In [2], the promotion Markov chain was introduced, which is a random walk on the linear extensions of a finite poset P. Here we discuss a new Markov chain on linear extensions which has mixing time of order $O(n \log n)$.

Denote by $\mathcal{W}(P)$ the set of subwords of linear extensions in $\mathcal{L}(P)$ and set $A = [n]$. We define a semigroup on $\mathcal{W}(P)$ as follows. Let $w \in \mathcal{W}(P)$ and $a \in A$. Then define

$$wa = \begin{cases} w & \text{if } a \in w, \\ \mathsf{straight}(wa) & \text{if } a \notin w. \end{cases} \tag{3.4}$$

Here $\mathsf{straight}(wa)$ is defined as follows. If wa is a subword of a linear extension of P, then $\mathsf{straight}(wa) = wa$. If not, write $w = w_1 \dots w_k$ and find the largest $1 \leqslant j_1 \leqslant k$ such that $a \prec w_{j_1}$ in P. Interchange w_{j_1} and a. Repeat by finding the largest $1 \leqslant j_2 < j_1$ such that $a \prec w_{j_2}$. Interchange w_{j_2} and a. Repeat until no further element bigger than a exists to the left. The result is $\mathsf{straight}(wa)$.

Example 3.1. Let P be the poset on four vertices with cover relations $\{(1,4),(2,4),(2,3)\}$. Then its Hasse diagram is the following:



This poset has five linear extensions $\mathcal{L}(P) = \{1234, 1243, 2134, 2143, 2314\}$. Take $w = 234 \in \mathcal{W}(P)$ and $a = 1$. We have $1 \prec 4$, so $j_1 = 3$. Both 2 and 3 are incomparable to 1, so we find $\mathsf{straight}(wa) = 2314 \in \mathcal{L}(P)$.

Lemma 3.2 ([27])**.** *Let $a \in A$ and $w \in \mathcal{W}(P)$ such that $a \notin w$. Then $\mathsf{straight}(wa) \in \mathcal{W}(P)$.*

Proposition 3.3 ([27])**.** *The set $\mathcal{W}(P)$ together with the product as in* (3.4) *forms a semigroup.*

Define $(\mathcal{W}(P), A)$ to be the semigroup with product (3.4) and generators $A = [n]$.

Theorem 3.4 ([27])**.** *The semigroup $(\mathcal{W}(P), A)$ is $\mathcal{R}$-trivial.*

Example 3.5. The right Cayley graph of $(\mathcal{W}(P), A)$ for the poset of Example 3.1 is given in Figure 3.

The minimal ideal of $(\mathcal{W}(P), A)$ is the set of linear extensions $\mathcal{L}(P)$ of the poset P. Let $\mathcal{M}(\mathcal{W}(P), A)$ be the Markov chain on $\mathcal{L}(P)$ induced by the semigroup $(\mathcal{W}(P), A)$. More precisely, we transition from $\pi \in \mathcal{L}(P)$ to $a\pi \in \mathcal{L}(P)$ with probability x_a.

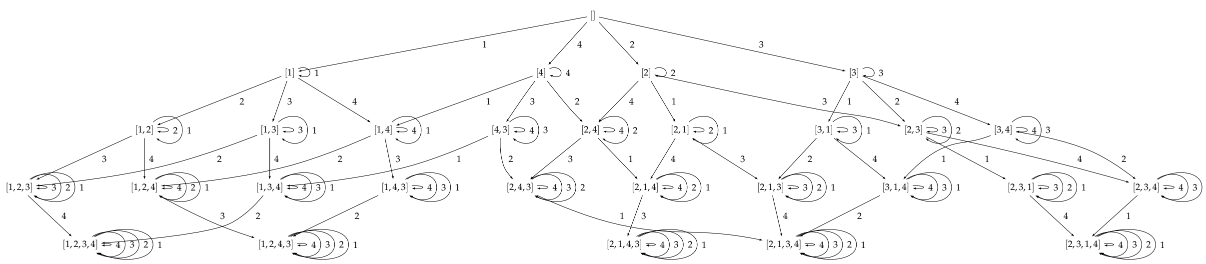


Figure 3: The right Cayley graph of $(\mathcal{W}(P), A)$ for the poset of Example 3.1.

Proposition 3.6 ([27])**.** *$\mathcal{M}(\mathcal{W}(P), A)$ is ergodic.*

The stationary distribution for $\mathcal{M}(\mathcal{W}(P), A)$ is given by

$$\Psi_\pi = \sum_{\substack{\sigma \in S_n \\ [\sigma]_{\mathcal{W}(P)} = \pi}} \left(\prod_{i=1}^{n} \frac{x_{\sigma_i}}{1 - \sum_{j=1}^{i-1} x_{\sigma_j}} \right) \qquad \text{for all } \pi \in \mathcal{L}(P).$$

Theorem 3.7 ([27])**.** *The expected value $E[\tau]$ for $\mathcal{M}(\mathcal{W}(P), A)$ is bounded above by $n \ln(n) + n\gamma$.*

Remark 3.8. Note that the Markov chain $\mathcal{M}(\mathcal{W}(P), A)$ is not identical to the promotion Markov chain in [2]. For example, left multiplication by 4 on 2143 in $(\mathcal{W}(P), \{1, 2, 3, 4\})$ for the poset in Example 3.1 yields 2143, whereas in the promotion Markov chain 2143 goes to 1243 under the promotion operator ∂_4 (see [2]). The full transition diagram for the new Markov chain is given in Figure 4.

Theorem 3.7 shows that the mixing time for $\mathcal{M}(\mathcal{W}(P), [n])$ is of order $O(n \log n)$. This does not take the computational complexity of computing the product (3.4) into account. For a word of length k, this involves up to k swaps.

Acknowledgements

We are grateful to Arvind Ayyer, Darij Grinberg, John Hunter, Stuart Margolis, Igor Pak, Dan Romik, Eric Severson, Benjamin Steinberg, and Andrew Waldron for discussions. The last author was partially supported by NSF grants DMS–1760329 and DMS–1764153. This material is based upon work supported by the Swedish Research Council under grant no. 2016-06596 while the author was in residence at Institut Mittag–Leffler in Djursholm, Sweden during Spring 2020.

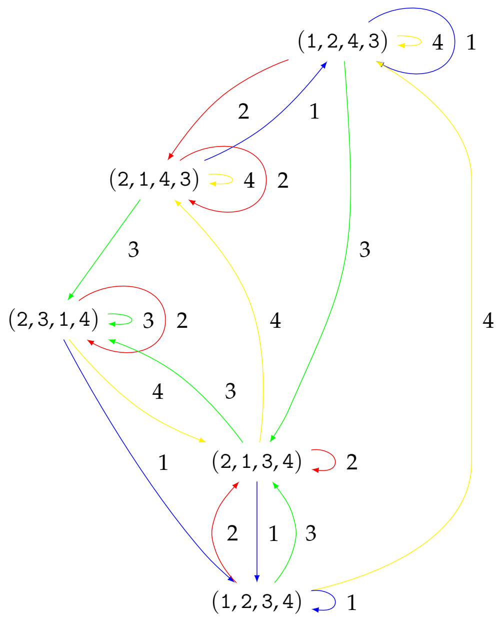


Figure 4: The Markov chain $\mathcal{M}(\mathcal{W}(P), [4])$ for the poset of Example 3.1

References

[1] C. A. Athanasiadis and P. Diaconis. "Functions of random walks on hyperplane arrangements". *Adv. in Appl. Math.* **45**.3 (2010), pp. 410–437. DOI.

[2] A. Ayyer, S. Klee, and A. Schilling. "Combinatorial Markov chains on linear extensions". *J. Algebraic Combin.* **39**.4 (2014), pp. 853–881. DOI.

[3] 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), pp. 169–231. DOI.

[4] J. Berstel, D. Perrin, and C. Reutenauer. *Codes and automata*. Vol. 129. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2010, pp. xiv+619.

[5] P. Bidigare, P. Hanlon, and D. Rockmore. "A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements". *Duke Math. J.* **99**.1 (1999), pp. 135–174. DOI.

[6] L. J. Billera, K. S. Brown, and P. Diaconis. "Random walks and plane arrangements in three dimensions". *Amer. Math. Monthly* **106**.6 (1999), pp. 502–524. DOI.

[7] A. Björner. "Random walks, arrangements, cell complexes, greedoids, and self-organizing libraries". *Building bridges*. Vol. 19. Bolyai Soc. Math. Stud. Springer, Berlin, 2008, pp. 165–203. DOI.

[8] A. Björner. "Note: Random-to-front shuffles on trees". *Electron. Commun. Probab.* **14** (2009), pp. 36–41. DOI.

[9] G. Brightwell and P. Winkler. "Counting linear extensions". *Order* **8**.3 (1991), pp. 225–242. DOI.

[10] K. S. Brown. "Semigroups, rings, and Markov chains". *J. Theoret. Probab.* **13**.3 (2000), pp. 871–938. DOI.

[11] K. S. Brown and P. Diaconis. "Random walks and hyperplane arrangements". *Ann. Probab.* **26**.4 (1998), pp. 1813–1854. DOI.

[12] R. Bubley and M. Dyer. "Faster random generation of linear extensions". *Discrete Math.* **201**.1-3 (1999), pp. 81–88. DOI.

[13] M. L. Cetlin. "Finite automata and the simulation of the simplest forms of behavior". *Uspehi Mat. Nauk* **18**.4 (112) (1963), pp. 3–28.

[14] F. Chung and R. Graham. "Edge flipping in graphs". *Adv. in Appl. Math.* **48**.1 (2012), pp. 37–63. DOI.

[15] P. Diaconis. "From shuffling cards to walking around the building: an introduction to modern Markov chain theory". *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*. Extra Vol. I. 1998, pp. 187–204.

[16] P. Edelman, T. Hibi, and R. P. Stanley. "A recurrence for linear extensions". *Order* **6**.1 (1989), pp. 15–18. DOI.

[17] W. J. Hendricks. "The stationary distribution of an interesting Markov chain". *J. Appl. Probability* **9** (1972), pp. 231–233.

[18] W. J. Hendricks. "An extension of a theorem concerning an interesting Markov chain". *J. Appl. Probability* **10** (1973), pp. 886–890.

[19] A. Karzanov and L. Khachiyan. "On the conductance of order Markov chains". *Order* **8**.1 (1991), pp. 7–15. DOI.

[20] D. E. Knuth. *The art of computer programming. Vol. 1*. Fundamental algorithms, Third edition [of MR0286317]. Addison-Wesley, Reading, MA, 1997, pp. xx+650.

[21] D. E. Knuth. *The art of computer programming. Vol. 3*. Sorting and searching, Second edition [of MR0445948]. Addison-Wesley, Reading, MA, 1998, pp. xiv+780.

[22] D. A. Levin and Y. Peres. *Markov chains and mixing times*. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2017, pp. xvi+447.

[23] 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), pp. 3037–3080. DOI.

[24] J. McCammond, J. Rhodes, and B. Steinberg. "Geometric semigroup theory". 2011. arXiv: 1104.2301.

[25] J. Rhodes and A. Schilling. "Normal distributions of finite Markov chains". *Internat. J. Algebra Comput.* **29**.8 (2019), pp. 1431–1449. DOI.

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

[27] J. Rhodes and A. Schilling. "Bounds on mixing time of finite Markov chains". 2020. arXiv:2010.08879.

[28] J. Rhodes, A. Schilling, and P. V. Silva. "Random walks on semaphore codes and delay de Bruijn semigroups". *Internat. J. Algebra Comput.* **26**.4 (2016), pp. 635–673. DOI.

[29] J. Rhodes, A. Schilling, and P. V. Silva. "Holonomy theorem for finite semigroups". 2020. arXiv:2007.15552.

[30] F. Saliola. "Eigenvectors for a random walk on a left-regular band". *Adv. in Appl. Math.* **48**.2 (2012), pp. 306–311. DOI.

[31] M.-P. Schützenberger. "Sur certains treillis gauches". *C. R. Acad. Sci. Paris* **224** (1947), pp. 776–778.