What is decidable about the Stochastic Reachability Problem?

George Kenison ✉
School for Logic and Computation, Technische Universität Wien

Abstract

We consider the open *Stochastic Reachability Problem*: given a stochastic matrix $K \in \mathbb{Q}^{d\times d}$, probability distributions $x, y \in \mathbb{Q}^d$, and a non-negative constant $r \in \mathbb{Q}$, determine whether there exists an $n \in \mathbb{N}$ such that $x^\top K^n y = r$. The restriction to random walks, the *Markov Reachability Problem*, is the task of determining whether there exists an $n \in \mathbb{N}$ such that the probability of travelling from a source state to a target state in n steps is equal to r. We establish decidability results for: walks on undirected graphs, hyperplane walks, Bernoulli walks on cycles, and circulant walks on certain finite groups. We illustrate this note with examples from self-organising lists and card shuffling.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Mathematics of computing → Graph theory; Computing methodologies → Symbolic and algebraic algorithms

Keywords and phrases Reachability Problems, Decision Problems, Markov Chains, Random Walks, The Skolem Problem, The Positivity Problem, Linear Recurrences

Digital Object Identifier 10.4230/LIPIcs...

Motivation

A dealer opens an unshuffled deck of cards and performs successive riffle shuffles on the deck. This shuffling process will continue until you call out 'Stop!' You want the dealer to stop at a given shuffle in order to be dealt a good hand. Can you decide when to make the call?

1 Introduction

Consider the following decision problem. Fix a stochastic matrix $K \in \mathbb{Q}^{d\times d}$, a pair of probability distributions $x, y \in \mathbb{Q}^d$, and a non-negative rational number r. Determine whether there exists an $n \in \mathbb{N}$ such that $x^\top K^n y$ is equal to r. The decidability of this *Stochastic Reachability Problem* is currently open.

The Stochastic Reachability Problem is a stochastic initialisation of the *Scalar Reachability Problem*: given a square matrix $M \in \mathbb{Q}^{d\times d}$, d-dimensional vectors $x, y \in \mathbb{Q}^d$, and a scalar value $\rho \in \mathbb{Q}$, determine whether there is an $n \in \mathbb{N}$ such that $x^\top M^n y = \rho$. The study of such matrix orbit and subspace hitting problems for linear dynamical systems has a long history. Seminal work by Harrison [21] introduced the *Orbit Problem*; given $M \in \mathbb{Q}^{d\times d}$ and vectors $x, y \in \mathbb{Q}^d$, determine whether there exists an $n \in \mathbb{N}$ such that $M^n x = y$. In two papers, Kannan and Lipton [25] showed that the Orbit Problem is decidable in polynomial time and then introduced the following generalisation [26]. Given a matrix $M \in \mathbb{Q}^{d\times d}$, a vector $x \in \mathbb{Q}^d$, and a subspace $V \subseteq \mathbb{Q}^d$, the *Generalised Orbit Problem* asks whether there exists an $n \in \mathbb{N}$ such that $M^n x \in V$. As noted in [25] the Generalised Orbit Problem is closely related to a fundamental and longstanding open problem in number theory, the *Skolem Problem* [16, 20]. Given a square matrix $M \in \mathbb{Q}^{d\times d}$ and d-dimensional vectors $x, y \in \mathbb{Q}^d$, the Skolem Problem asks whether there exists an $n \in \mathbb{N}$ such that $M^n x \in y^\perp$ where $y^\perp = \{v \in \mathbb{R}^d : v \cdot y = 0\}$.

The *Markov Reachability Problem* [3, 1], is a variant of the Stochastic Reachability Problem (K, x, y, r) where one restricts the probability distributions $x, y \in \mathbb{Q}^d$ to point

masses. Thus the problem instance (K, x, y, r) is a question about a random walk on the Markov chain associated with K. A *random walk*[1], is a sequence of states generated by the following process: start at an initial state; and then at each time instance move to a randomly selected neighbour of the current state. Briefly, the Markov Reachability Problem (K, x, y, r) asks one to determine whether there exists an $n \in \mathbb{N}$ such that the probability of travelling from a source state to a target state in n steps is equal to r. The decidability of the Markov Reachability Problem is likewise open.

For the avoidance of doubt, the Markov Reachability Problem does not ask whether the probability of transitioning from source to target in any number of steps is equal to r. This latter problem is trivially decidable: such probabilities are computable in polynomial time. The authors of [1] showed that the Skolem Problem reduces to the Markov Reachability Problem—hinting at the hardness of the latter.

Let us make clear the connection between number theory and the Skolem Problem. Observe that $M^n x \in y^\perp$ if and only if $y^\top M^n x = 0$ where $\langle y^\top M^n x \rangle_n$ is a linear recurrence sequence. That is to say, an equivalent formulation of the Skolem Problem asks whether the set $\{n \in \mathbb{N} : y^\top M^n x = 0\}$ is non-empty. A sequence $\langle u_n \rangle_{n=0}^\infty$ of real algebraic numbers is a *linear recurrence sequence* (sometimes a *C-finite* sequence) if its terms satisfy a recurrence relation

$$u_n = a_1 u_{n-1} + a_2 u_{n-2} + \cdots + a_\ell u_{n-\ell}$$

with fixed real algebraic constants $a_1, \ldots, a_\ell$ such that $a_\ell \neq 0$. Such a recurrence is said to have order ℓ and a sequence $\langle u_n \rangle_n$ satisfying the recurrence is wholly determined by its initial values $u_0, \ldots, u_{\ell-1}$.

▶ Remark 1. Herein our restriction to stochastic initialisations is not a narrowing specialisation in the following sense. Given a linear recurrence sequence $\langle u_n \rangle_n$, there is a stochastic matrix $\widetilde{K}$ and vectors $\tilde{x}, \tilde{y}$ such that for each $n \in \mathbb{N}$ one has $u_n = 0$ if and only if $\tilde{x}^\top \widetilde{K}^n \tilde{y} = 0$. The details of this argument are presented in [36, §3.3]. Thus the Skolem Problem is equivalent to the specialisation where one considers only linear recurrence sequences driven by the family of stochastic matrices.

It seems appropriate to make a small digression in order to discuss the structure of the set $\{n \in \mathbb{N} : u_n = 0\}$ when $\langle u_n \rangle_n$ is a linear recurrence sequence. Work by Skolem, Mahler, and Lech established that the set $\{n \in \mathbb{N} : u_n = 0\}$ is the union of a finite (possibly empty) set together with a finite (possibly zero) number of infinite arithmetic progressions. Briefly, the history of this remarkable result is a sequence of generalisations: proved by Skolem [46] for the field of rational numbers, the result was subsequently extended to the field of algebraic numbers by Mahler [31, 32], and then further extended to any field of characteristic 0 by Lech [30]. All known proofs of the Skolem–Mahler–Lech Theorem employ techniques from p-adic analysis. On the one hand, Berstel and Mignotte gave an effective method to obtain all of the arithmetic progressions in the statement of the theorem [5]. On the other hand, there is no known constructive method to produce the finite sporadic set of zeroes and so the decidability of the Skolem Problem is open.

There has been limited progress on the decidability of the Skolem Problem when one considers linear recurrence sequences of low order. Research by Mignotte, Shorey, and Tijdeman [35], and, independently, Vereshchagin [49], proved the following:

[1] Traditionally, random walks are defined on *reversible* Markov chains (see the Preliminaries), but we shall apply the term widely.

▶ **Proposition 2.** *Let $\langle u_n \rangle_n$ be a non-degenerate linear recurrence sequence. The Skolem Problem for $\langle u_n \rangle_n$ is decidable if $\langle u_n \rangle_n$ has at most three simple characteristic roots that are maximal in modulus.*

As a consequence, the Skolem Problem is known to be decidable for linear recurrences of order at most four. The aforementioned papers employ techniques from p-adic analysis and algebraic number theory and, in addition, Baker's theorem for linear forms in logarithms of algebraic numbers. Unfortunately the route taken via Baker's Theorem does not appear to extend easily to recurrences of higher order.

The next statement follows easily from the decidability results we establish herein.

▶ **Theorem 3.** *The Markov Reachability Problem is decidable for random walks on: undirected graphs, hyperplane arrangements, Bernoulli walks on cycles, and circulant walks on a class of finite groups.*

We make the following standing assumptions throughout (we refer the reader to the Preliminaries for the technical definitions). We shall assume that a graph is non-empty, finite, connected, and simple. The weights of edges in a weighted graph are non-negative and rational. Herein a Markov chain is irreducible.

The paper is structured as follows. In Section 2 we recall preliminary material. The decidability of the Stochastic Reachability Problem is discussed: in Section 3 for undirected graphs; in Section 4 for hyperplane arrangements; in Section 5 for Bernoulli cycle graphs; and in Section 6 for circulant chains associated with both Abelian and ambivalent groups. In the self-contained Appendix A we give a brief discussion of the Stochastic Reachability Problem under standard assumptions (irreducibility and aperiodicity) from graph theory and probability theory.

2 Preliminaries

2.1 Linear algebra

A non-negative matrix $K \in \mathbb{Q}^{d\times d}$ is *stochastic* if $\sum_v K(u,v) = 1$ for each $u \in \{1,2,\ldots,d\}$. The spectrum of K lies in the unit disk and unity is always an element of the spectrum. Many of the following concepts and definitions are familiar in the study of non-negative matrices, but for our purpose we limit our discussion to the setting of stochastic matrices.

A stochastic matrix $K \in \mathbb{Q}^{d\times d}$ is *irreducible* if for each pair (u,v) there exists $m \in \mathbb{N}$ such that $K^m(u,v) > 0$. The *period* of a state $v \in V$ is the greatest common divisor of the elements in the set $\{m \in \mathbb{N} : K^m(v,v) > 0\}$. In fact, this set is non-empty and all the periods of an irreducible matrix are equal, so it is natural to speak of the (finite) *period* of the matrix. An irreducible matrix is *aperiodic* if it has period 1. Equivalently, K is irreducible and aperiodic if there exists an $n > 0$ such that $K^n(u,v) > 0$ for all (u,v).

▶ **Theorem 4** (Perron–Frobenius Theorem, [19])**.** *Let $K \in \mathbb{Q}^{d\times d}$ be an irreducible stochastic matrix with period h. First, each of the roots of unity $\mathrm{e}^{2\pi \mathrm{i}k/h}$ with $k \in \{0,1,\ldots,h-1\}$ is a simple eigenvalue of K and the remaining eigenvalues of K are strictly smaller in modulus. Second, let x and y be the respective left- and right-eigenvectors of K associated with the eigenvalue 1. Then the entries of x and y are all either positive or negative (up to scaling) and, in addition, these are the only eigenvectors of K with this property. Finally, K is similar to $\mathrm{e}^{2\pi \mathrm{i}/h}K$ and so the spectrum of K is invariant under the action of $\mathrm{e}^{2\pi \mathrm{i}/h}$.*

If, in addition, the stochastic matrix K is irreducible and aperiodic then one has the following.

▶ **Corollary 5.** *Suppose that $K \in \mathbb{Q}^{d\times d}$ is an irreducible and aperiodic stochastic matrix. First, unity is a simple eigenvalue of K that is strictly greater in modulus than all the other eigenvalues of K. Second, there is a unique positive left-eigenvector π (up to scaling) such that $\pi K = \pi$.*

In Appendix A we give a brief discussion of the Stochastic Reachability Problem when one assumes that K is irreducible and aperiodic.

2.2 Markov chains

Markov chains are fundamental models of randomness in probabilistic model checking [2]. A (discrete-time) Markov chain $\mathscr{C}$ on a finite state space S is a sequence of random variables $\langle X_n\rangle_{n=0}^{\infty}$ with each $X_n \in S$ that obeys the *Markov property* as follows. For each $n \in \mathbb{N}$ and sequence of states $\langle s_m\rangle_m$ in S,

$$\mathbb{P}(X_{n+1} = s_{n+1} \mid X_n = s_n, X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) = \mathbb{P}(X_{n+1} = s_{n+1} \mid X_n = s_n).$$

Here $\mathbb{P}$ is a probability measure. Since the next transition depends solely on the current state, such models are suitably said to be memoryless. We shall assume that the Markov chain is *homogeneous*; that is, the probability of transitioning from state s_n to state s_{n+1} is independent of the time instance n. Thus the transition probabilities are encoded using a (stochastic) *transition* matrix K whose rows and columns are indexed by the elements in S with entries $K(s,t) = \mathbb{P}(X_{n+1} = t \mid X_n = s)$ for each pair $(s,t) \in S^2$.

In the sequel we shall always assume that the transition matrix K is irreducible. This assumption is made without loss of generality as the questions of reachability can be decomposed into instances of time spent in transitory states and bottom strongly connected components (sometimes irreducible components or communication classes).

2.3 linear recurrence sequences

Whenever we refer to a sequence $\langle u_n\rangle_n$ as a linear recurrence sequence, we shall mean that $\langle u_n\rangle_n$ is a real linear recurrence sequence over $\mathbb{Q}$. For further information on the theory of recurrence sequences we refer the reader to [16].

Recall the Skolem–Mahler–Lech Theorem from earlier.

▶ **Theorem 6** (Skolem–Mahler–Lech)**.** *For a linear recurrence sequence $\langle u_n\rangle_n$, the set of terms where the sequence vanishes, $\{n \in \mathbb{N} : u_n = 0\}$, is given by the union of a finite set together with a finite number of infinite arithmetic progressions.*

It is useful to introduce the notion of degeneracy into our discussion of recurrence sequences. A recurrence sequence is *degenerate* if the ratio λ_i/λ_j of any two distinct characteristic roots of the sequence is a root of unity. Otherwise a linear recurrence sequence is *non-degenerate*. Any linear recurrence sequence can be effectively decomposed into an interleaving of finitely many non-degenerate sequences, some of which may be identically zero (see [5]). A non-degenerate non-zero linear recurrence sequence has only finitely many zeros; however, as mentioned earlier, there is no known method to compute this finite set.

The next result is considered folklore (see [20]).

▶ **Proposition 7.** *Let $\langle u_n\rangle_n$ be a non-degenerate linear recurrence sequence. The Skolem Problem for $\langle u_n\rangle_n$ is decidable if each of the characteristic roots of $\langle u_n\rangle_n$ is real.*

Let $m(X) = X^d - a_{d-1}X^{d-1} - \cdots - a_1X - a_0$ be the *characteristic polynomial* of the stochastic transition matrix $K \in \mathbb{Q}^{d\times d}$. Thus $\langle x^\top K^n y\rangle_n$ is a linear recurrence sequence that satisfies the recurrence relation

$$u_{n+d} - a_{d-1}u_{n+d-1} - \cdots - a_1u_{n+1} - a_0u_n = 0$$

of order d. We shall also call $m(X)$ the *characteristic polynomial* of the above recurrence relation. The characteristic polynomial of the recurrence relation of least order satisfied by a linear recurrence sequence is the *minimal polynomial* of said sequence. A constant sequence with $u_n = r \in \mathbb{Q}$ for each $n \in \mathbb{N}$ satisfies the relation $u_{n+1} = u_n$ with characteristic polynomial $X - 1$. It is easy to verify that the sequence $\langle x^\top K^n y - r\rangle_n$ satisfies a linear recurrence relation with characteristic polynomial given by $\operatorname{lcm}(m(X), X - 1) = m(X)$ [16, §1.1] where equality follows because K is a stochastic matrix and so $(X - 1) \mid m(X)$.

It is well known that the terms of a linear recurrence sequence can be written in a closed form as an exponential polynomial [16, §1.1]. Thus $x^\top K^n y - r = \sum_{k=1}^{d} A_k(n)\lambda_k^n$ where the polynomial coefficients $A_k(n)$ depend on the $x, y \in \mathbb{Q}^d$ and the *characteristic roots* λ_k are the roots of the characteristic polynomial m. In summary,

▶ **Lemma 8.** *Let (K, x, y, r) be a tuple as above. Then the sequence $\langle u_n\rangle_n$ with terms given by $u_n = y^\top K^n x - r$ is a linear recurrence sequence. Moreover, the characteristic polynomial of K and the characteristic polynomial of $\langle u_n\rangle_n$ are one and the same.*

2.4 reversible chains

A Markov chain $\mathscr{C}$ associated with the stochastic matrix K is *reversible* if there exists a probability distribution π that satisfies the *detailed balance equations* $\pi(u)K(u, v) = \pi(v)K(v, u)$. We note that the condition for reversibility is strictly stronger than the assumptions for the existence of a stationary distribution and it can be shown that the above probability distribution π is the unique stationary distribution for the chain $\mathscr{C}$ [40].

Let $\ell^2(\pi)$ be the Hilbert space of π-weighted square-summable real-valued functions with the usual inner product

$$\langle f, g\rangle_{\ell^2(\pi)} = \sum_{s\in S} f(s)\overline{g(s)}\pi(s)$$

and norm $\|f\|^2_{\ell^2(\pi)} = \sum_{s\in S} |f(s)|^2\pi(s)$. The matrix $K\colon \ell^2(\pi) \to \ell^2(\pi)$ is an operator so that $Kf(s) = \sum_{s'\in S} K(s, s')f(s')$ (in agreement with the standard matrix action). Then it is well-known (see, for example, [43, §1.3]) that the chain with transition matrix K and stationary distribution π is reversible if and only if K is *self-adjoint* on $\ell^2(\pi)$; that is, $\langle Kf, g\rangle_{\ell^2(\pi)} = \langle f, Kg\rangle_{\ell^2(\pi)}$ for each pair $f, g \in \ell^2(\pi)$. Thus we have the following.

▶ **Lemma 9.** *If $\mathscr{C}$ is reversible then all the eigenvalues of K are real and K is diagonalisable.*

Kolmogorov gave a necessary and sufficient condition for a Markov chain to be reversible [28, Theorem 1.7].

▶ **Theorem 10** (Kolmogorov's criterion)**.** *A Markov chain with stochastic matrix K is reversible if and only if its transition probabilities satisfy*

$$K(s_0, s_1)\cdots K(s_{n-2}, s_{n-1})K(s_{n-1}, s_0) = K(s_0, s_{n-1})K(s_{n-1}, s_{n-2})\cdots K(s_1, s_0)$$

for any finite sequence of states $s_0, s_1, \ldots, s_{n-1} \in S$.

Reversibility is a property of a number of commonly encountered Markov chains. For example, birth-and-death chains from probability theory are reversible. One important example from statistical physics, the Ehrenfest model, modelling idealised gas molecules moving between two connected chambers is a birth-and-death chain.

The *communication graphs* associated a Markov chain is the undirected graph with vertex set given by the states in the chain and edge set $E := \{e(u,v) : K(u,v) > 0 \text{ and } K(v,u) > 0\}$. The communication graph of any birth-and-death chain is a tree. In fact, this condition is sufficient for reversibility [45, Theorem 99].

▶ **Theorem 11.** *Suppose that the Markov chain $\mathscr{C}$ has a stationary distribution π and the communication graph of $\mathscr{C}$ is a tree. Then $\mathscr{C}$ is reversible with respect to π.*

Let us consider one final example whose underlying communication graph is not a tree.

▶ **Example 12.** Consider the random walk associated with the set of permutations of a finite list of items $I_1, I_2, \ldots, I_N$ placed in a linear array. Initialise the process with the items in some starting permutation. At each time instance n, an item I_j is selected at random and is swapped with the preceding item in the list. That is, unless I_j is the first item in the list, in which case nothing is changed.

The random walk in Example 12 is commonly called the *(adjacent) transposition scheme* [34, 23]. Suppose that each selection is independent and identically distributed, then the process is naturally modelled by a Markov chain. For clarity, we suppose that item I_j is selected with probability $w_j > 0$ at each time instance. It known that the Markov chain for this model is reversible. The proof of reversibility is a simple application of Kolmogorov's criterion; a formal argument is given in [40, Example 4.3.7].

2.5 circulant matrices

We begin by recalling standard material from the rich study of circulant matrices. For further exposition we direct the interested reader to the survey [29].

Fix $d \geq 2$ and let $\mathbb{C}^d$ denote the Euclidean d-dimensional complex vector space. In this section we will, at points, refer to elements of $\mathbb{C}^d$ as row vectors or column vectors depending on the suitability. Let $y = (y_0, y_1, \ldots, y_{d-1}) \in \mathbb{C}^d$. We define the shift operator $\sigma : \mathbb{C}^d \to \mathbb{C}^d$ by $\sigma(y_0, \ldots, y_{d-2}, y_{d-1}) = (y_{d-1}, y_0, \ldots, y_{d-2})$. Given the row vector $y \in \mathbb{C}^d$, let $\mathrm{circ}(y) \in \mathbb{C}^{d\times d}$ be the matrix with row $k \in \{1, \ldots, d\}$ given by $\sigma^{k-1}(y)$. The set of $d \times d$ circulant matrices $\mathrm{Circ}(d) \in \mathbb{C}^{d\times d}$ form a commutative ring under matrix multiplication and, in addition, this ring is isomorphic to the ring of $d \times d$ diagonal matrices.

Let $\omega = \mathrm{e}^{2\pi \mathrm{i}/d}$ be a primitive dth root of unity. For $\ell \in \{0, 1 \ldots, d-1\}$, let $x_\ell = d^{-1/2}(1, \omega^\ell, \ldots, \omega^{(d-1)\ell}) \in \mathbb{C}^d$. We form the Vandermonde matrix $Q \in \mathbb{C}^{d\times d}$ whose ℓth row (and column) is given by (x_ℓ) so that

$$Q = \frac{1}{\sqrt{d}} \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & \cdots & \omega^{d-2} & \omega^{d-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \omega^{d-2} & \cdots & \omega^{(d-2)^2} & \omega^{(d-2)(d-1)} \\ 1 & \omega^{d-1} & \cdots & \omega^{(d-1)(d-2)} & \omega^{(d-1)^2} \end{pmatrix}.$$

It can be shown that the set of x_ℓ form an orthogonal basis, from which it follows that Q is invertible. In addition, Q is both unitary and symmetric. As a consequence of the following result, all elements of $\mathrm{Circ}(d)$ have the same basis of orthonomal eigenvectors.

▶ **Proposition 13.** *The matrices in* $\mathrm{Circ}(d)$ *are simultaneously diagonalised by the unitary matrix* Q*. In more detail, if* $y = (y_0, y_1, \ldots, y_{d-1})$ *and* $V = \mathrm{circ}\,(y)$ *then* $Q^{-1}VQ = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{d-1})$ *where* $\lambda_\ell = \sum_{k=0}^{d-1} \omega^{k\ell} y_k$.

2.6 representation theory

Let G be a finite group. A *presentation* of G assigns to each element of the group a matrix such that the matrix assigned by the presentation to the product of two elements is given by the product of their matrices. More specifically, a group presentation $\rho\colon G \to \mathrm{GL}(V)$ is a homomorphism where V is a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$ of dimension d_ρ. Without loss of generality we can always assume that $\rho(s)$ is a *unitary* matrix so that $\rho(s)^\dagger$, the conjugate transpose of $\rho(s)$, is equal to $\rho(s^{-1})$.

A presentation is *irreducible* if there are no proper invariant subsets of the action of ρ; that is, ρ is an irreducible presentation if for each $s \in G$ there is a subspace W of V such that $\rho(s)W \subseteq W$ then W is necessarily either 0 or V.

Let $f\colon G \to \mathbb{R}$ then the *Fourier transform of* f *at* ρ is given by $\hat{f}(\rho) := \sum_{s\in G} f(s)\rho(s)$. We decompose the function f using the Fourier inversion theorem transform: we have

$$f(t) = \frac{1}{|G|} \sum_{\rho} d_\rho \,\mathrm{tr}(\rho(t^{-1})\hat{f}(\rho_j))$$

where the summation is taken over the set of irreducible representations for G.

▶ **Example 14.** The irreducible representations of $\mathbb{Z}_p$, the integers modulo p, are all one-dimensional and each map is of the form $\rho_j(s) = \mathrm{e}^{2\pi \mathrm{i} js/p}$ for each $j \in \{0, 1, \ldots, p-1\}$. The Fourier transformation is the well-known discrete Fourier transform. Thus for $f\colon \mathbb{Z}_p \to \mathbb{R}$ we have $f(t) = \frac{1}{p}\sum_{k=0}^{p-1} \mathrm{e}^{-2\pi \mathrm{i} k/p} \hat{f}(k)$ where $\hat{f}(k) = \sum_{j=0}^{p-1} f(j)\mathrm{e}^{2\pi \mathrm{i} jk/p}$.

3 Reachability for Undirected Graphs

A *directed graph* $\mathscr{G} = (V, E)$ is a non-empty finite tuple of *vertices* V and *directed edges* E such that each edge $e(u, v)$ is determined by an ordered pair with $u, v \in V$. We shall assume throughout that $\mathscr{G}$ is simple; that is to say, there is at most one directed edge $e(u, v)$ associated with the ordered pair $(u, v) \in V^2$. Let $\mathscr{G}$ be a *weighted undirected graph*. Each edge $e(u, v) \in E(\mathscr{G})$ is endowed with a positive weight $w(u, v) = w(v, u) \in \mathbb{Q}$. We define the *random walk* associated with $\mathscr{G}$ as the Markov chain with stochastic transition matrix K with entries $K(u, v) = w(u, v)/\overline{w}(u)$ where $\overline{w}(u) := \sum_{v\in N(u)} w(u, v)$ is a sum over the neighbours (or adjacent vertices) of vertex u. It is easily seen that this construction produces a reversible Markov chain with associated stationary distribution $\pi(u) = \overline{w}(u)/\overline{w}$ where $\overline{w} := \sum_v \overline{w}(v)$ is twice the total edge-weight sum of $\mathscr{G}$.

Conversely, every irreducible and reversible Markov chain is associated with a random walk on a weighted undirected graph. Let K be the stochastic transition matrix associated with such a Markov chain. We define an undirected graph with vertex set equal to the set of states in the chain and $e(u, v)$ is an edge in the undirected graph if and only if $K(u, v) > 0$. In this setup each edge $e(u, v)$ is endowed with weight $w(u, v) := \pi(u)K(u, v)$.

▶ **Theorem 15.** *The Stochastic Reachability Problem with initialisation* (K, x, y, r) *is decidable if* $K \in \mathbb{Q}^{d\times d}$ *is the transition matrix of a Markov chain modelling a random walk on an undirected graph.*

From the above observations, Theorem 15 follows as a corollary to the next result for reversible Markov chains.

▶ **Proposition 16.** *The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable if K is associated with a reversible Markov chain.*

Proof. Since the Markov chain associated with K is reversible, the eigenvalues of K are all real and K is diagonalisable. Thus $K = P^{-1}DP$ where D is a diagonal matrix and P is a change of basis matrix. Then $x^\top K^n y = r$ if and only if $x^\top P^{-1} D^n P y - r = 0$. By Lemma 8, the question of whether $x^\top P^{-1} D^n P y - r$ vanishes for some $n \in \mathbb{N}$ is an instance of the Skolem Problem. Since K is reversible, the characteristic roots of the sequence $\langle x^\top P^{-1} D^n P y - r \rangle_n$ are all real (Lemmas 8 and 9). From Proposition 7, we deduce that this instance of the Skolem Problem is decidable. Hence the desired result. ◀

▶ Remark 17. Let us consider the threshold variant of the Stochastic Reachability Problem when K is diagonalisable (a considerably weaker assumption by comparison to reversibility). The problem of determining whether $x^\top K^n y - r \geq 0$ for all but finitely many $n \in \mathbb{N}$ is an instance of the *Ultimate Positivity Problem*. Ouaknine and Worrell [38] established decidability for simple linear recurrence sequences. A linear recurrence sequence is *simple* if each of the roots of the associated characteristic polynomial has algebraic multiplicity 1. Thus decidability is resolved in the case that K is diagonalisable since then roots of the minimal polynomial of $\langle x^\top K^n y - r \rangle_n$ are simple (Lemma 8).

4 Reachability for Hyperplane Arrangements

In the previous section, we employed the traditional interpretation of a random walk as an execution of a reversible Markov chain. In the sequel we shall abuse terminology and consider the execution of a finite Markov chain (not necessarily reversible) as a random walk on a weighted directed graph. In this section we consider a family of Markov chains that are, in general, not reversible: *walks on hyperplane arrangements*. The main result of this section is the following.

▶ **Theorem 18.** *The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable when $K \in \mathbb{Q}^{d \times d}$ is the transition matrix of a Markov chain modelling a random walk on a hyperplane arrangement.*

It is illuminating to consider prototypical examples from the literature on self-organising lists before we move on to hyperplane arrangements. The following setup is commonly painted in terms of *books on a shelf* or a *stack of cards*. Consider a finite list of items $I_1, I_2, \ldots, I_N$ placed in a linear array. At each time instance $t \in \mathbb{N}$, one of the items is called and an action is performed. If the called item is in position j, with $j \in \{1, \ldots, N\}$, then perform a permutation $\sigma_j \in S_N$ on the current array. If the specified permutations $\sigma_1, \sigma_2, \ldots, \sigma_N$ are such that from a given starting array one can reach any other by successive application of some or all of the σ_j (where repeat application is permissible) then the scheme $\langle \sigma_1, \ldots, \sigma_N \rangle$ is called *self-organising* [23].

Self-organising lists lend themselves nicely to probabilistic modelling when the selection process at each stage is random. In particular, when selection is a sequence of independent and identically distributed random variables the natural model is a Markov chain.

Self-organising lists appear in heuristic approaches to access time and information retrieval [34] and applications to such technologies as VLSI circuit simulations, data compression [41, 4, 42, 33], and communications networks. We refer the interested reader to the extensive

list of references given in [17]. Two prototypical schemes that reorganise a list so that, figuratively speaking, frequently accessed items are moved nearer to the start of the list include: Move-to-Front (Tsetlin's library) and transposition.

We described the (adjacent) transposition scheme in Example 12. The fact that the scheme is self-organising is easily apparent: any permutation is reachable from a given initial permutation. Since the scheme is associated with a reversible Markov chain, for such examples the Stochastic Reachability Problem is decidable (see Theorem 15).

Tsetlin introduced his model of a self-organising library, commonly *Tsetlin's library*, in his work on automata [48]. Tsetlin's library consists of a single shelf of books. At each time instance a book is selected and moved to the leftmost position on the shelf. If the selected book currently occupies the leftmost position on the shelf then nothing is changed. It is easily seen that there is a path in the Markov chain of length at most N connecting any pair of permutations of the shelved books. Thus Tsetlin's library is a self-organising scheme. In the computer science literature this sorting procedure is often referred to as the *Move-To-Front* scheme.

Hendricks modelled Tsetlin's library as a Markov chain so that the selection of book I_ℓ at time n occurs with probability $w_\ell > 0$ and $\sum_\ell w_\ell = 1$ [22, 23]. Hendricks went on to give an explicit formulation of the stationary distribution of this Markov chain.

Independent research by Donnelly [15], Kapoor and Reingold [27], and Phatarfod [39] established exact formulae for the eigenvalues of the chain. In fact, the eigenvalues are linear in the transition probabilities and their multiplicities satisfy elegant combinatorial formulae.

▶ **Theorem 19.** *The distinct eigenvalues of Tsetlin's library are indexed by subsets $D \subseteq \{1, 2, \ldots, N\}$ and $\lambda_D = \sum_{\ell \in D} w_\ell$. Further, the multiplicity of eigenvalue λ_D is given by the number of derangements on $N - |D|$ elements.*

In general, Tsetlin's library is not a reversible chain and so we cannot employ our previous decidability results for the Stochastic Reachability Problem in this instance. Nevertheless, since the eigenvalues of Tsetlin's library are all real-valued we obtain the following:

▶ **Proposition 20.** *The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable when $K \in \mathbb{Q}^{d \times d}$ is the transition matrix of a Markov chain modelling an instance of Tsetlin's library with rational weights.*

Proof. From Theorem 19, the eigenvalues of K are all real. Thus $x^\top K^n y = r$ for some $n \in \mathbb{N}$ if and only if $\{n \in \mathbb{N} : x^\top K^n y - r = 0\}$ is a non-empty set. As before, we see this an instance of the Skolem Problem. From Proposition 7, we deduce that this instance of the Skolem Problem is decidable. Hence the desired result. ◀

▶ Remark 21. Fill derives an exact and tractable formula for the probability of any permutation after any number of moves in [17, Theorem 2.1]. Thus Proposition 20 is a natural corollary of Fill's result.

A generalisation of Tsetlin's library is the *subset-move-to-front* self-organising scheme where, at each time instance, a subset of the N items is selected and the appropriate elements are moved to the front of the list (whilst preserving their relative order).

A further generalisation permits permutations indexed by block-ordered partitions of N elements. Label each item in the list with one of $m \in \mathbb{N}$ labels $T_1, \ldots, T_m$. The permutation denoted by $(T_1, \ldots, T_m)$ indicates the following action. First, move the items labelled T_1 to the front of the list whilst retaining their relative order. Second, move the items labelled T_2 behind the T_1 items, again, preserving their relative order. Continue this process until the

items labelled by T_m are left at the end of the list. Such operations are called *elementary pop shuffles*; more generally, a *pop shuffle* is given by any $\mathbb{C}$-linear combination of elementary pop shuffles (viewed as transformations on the appropriate vector space of permutations). Bidigare, Hanlon, and Rockmore [6] prove analogous spectral results to Theorem 19 in this wider setting. Further, those authors consider the pop shuffle model in the more abstract setting of random walks on hyperplane arrangements.

A *central real hyperplane arrangement* $\mathscr{A}$ (hereafter a hyperplane arrangement) is a collection of codimension-1 linear subspaces in Euclidean space $\mathbb{R}^d$. Let $\mathscr{L}(\mathscr{A}) := \{\cap\mathscr{E} : \mathscr{E} \subseteq \mathscr{A}\}$ be the *intersection poset* of $\mathscr{A}$; that is, the collection of subspaces of $\mathbb{R}^d$ given by intersections of some of the hyperplanes of $\mathscr{A}$ equipped with a partial ordering by reverse inclusion.

The *faces* of $\mathscr{A}$ are given by the intersections of open half-spaces and/or hyperplanes. The *chambers* of $\mathscr{A}$ are the connected components in the complement of the union of the hyperplanes. Hence each chamber of $\mathscr{A}$ is a d-dimensional face of $\mathscr{A}$.

Given both a chamber C and a face F of $\mathscr{A}$, Tits [47] asserts that there is a unique chamber C' that has both F as a face and is closest to C in the metric that counts the number of hyperplanes in $\mathscr{A}$ separating C and C'. We call $C' := FC$ the *projection* of C onto F.

Given a hyperplane arrangement $\mathscr{A}$ as above, Bidigare, Hanlon, and Rockmore introduced a family of Markov chains called *hyperplane walks*; the state space of each member is precisely the set of chambers $\mathscr{C}(\mathscr{A})$ of $\mathscr{A}$. Let w be a probability measure on the set of faces $\mathscr{F}(\mathscr{A})$ of $\mathscr{A}$. At each time instance a face $F \in \mathscr{F}(\mathscr{A})$ is selected at random (F is chosen with probability $w(F)$). Then we update the state from the current chamber C to the new chamber FC. The stochastic transition matrix K for this hyperplane walk has entries given by

$$K(C, C') := \sum_{F \in \mathscr{F}(\mathscr{A}) \,:\, FC = C'} w(F)$$

where we have suppressed the dependence on w on the left-hand side.

The braid arrangements are a strict subclass of hyperplane arrangements. A *braid arrangement* is a set of hyperplanes $\{H_{i,j}\}_{i<j}$ such that each $H_{i,j} = \{(x_1, \ldots, x_N) : x_i = x_j\}$. The chambers of the braid arrangement are labelled by the $N!$ permutations of N elements. Further, the faces of such an arrangement are labelled by block-ordered partitions. Bidigare, Hanlon, and Rockmore showed that any Markov chain of a block-ordered permutation (and, in particular, Tsetlin's library) appears as a hyperplane walk on a braid arrangement.

The next theorem [6, Theorem 4.1] generalises the result in Theorem 19.

▶ **Theorem 22.** *Given a hyperplane arrangement $\mathscr{A}$ and a probability distribution w on $\mathscr{F}(\mathscr{A})$, the associated hyperplane walk has the following properties. The eigenvalues of the chain are indexed by the intersection poset $\mathscr{L}(\mathscr{A})$ such that the eigenvalue λ_W associated with $W \in \mathscr{L}(\mathscr{A})$ is given by*

$$\lambda_W = \sum_{F \in \mathscr{F}(\mathscr{A}) \,:\, F \subseteq W} w(F).$$

The multiplicity of eigenvalue λ_W is equal to the magnitude of the Möbius function on $\mathscr{L}(\mathscr{A})$ evaluated at W.

The proof of Theorem 18 is a straightforward corollary of Theorem 22.

Proof of Theorem 18. Consider that $x^\top K^n y = r$ for some $n \in \mathbb{N}$ if and only if $\{n \in \mathbb{N} : x^\top K^n y - r = 0\}$ is a non-empty set. As before, we see this an instance of the Skolem

Problem. By Theorem 22, the eigenvalues of K are all real and so we deduce this instance of the Skolem Problem is decidable (Proposition 7). Hence the desired result. ◀

▶ Remark 23. The insights in Bidigare, Hanlon and Rockmore's seminal paper [6] led to a veritable windfall of spectral results—analogous to Theorem 19—for random walks in abstract settings. These abstract settings include:

1. Brown and Diaconis' consideration of *non-centred* real hyperplane arrangements [10];
2. Brown's [9] results for *left regular bands*—semigroups whose elements are all *idempotent* (i.e., $x^2 = x$) and, in addition, satisfy the *cancellation property* $xyx = xy$ for all pairs x, y—and subsequent extension to all bands; and
3. Björner's eigenvalue formulae [8, 7] generalising Tsetlin's library from a single bookshelf to hierarchies of libraries.

Decidable instances of the Stochastic Reachability Problem are easily deduced from these spectral results.

▶ **Example 24.** A *riffle shuffle* cuts a deck of cards into two parts L and R and then interleaves the cards from said parts. The classical model for performing a single riffle shuffle is as follows [13]. First, a deck of n cards is partitioned into two by cutting the deck at the cth card with probability $\binom{n}{c}2^{-n}$. One of the parts is selected at random and the bottom card of this part is *dropped*. We repeat the process, dropping the bottom card from the selected (remaining) part on top of the previous drop, until all cards have been dropped. At a given interation, there are $|L|$ and $|R|$ cards in the (remaining) parts respectively. In this iterate part L is selected with probability $|L|/(|L| + |R|)$ and part R is selected with probability $|R|/(|L| + |R|)$.

Let $f(\sigma)$ be the probability obtaining permutation σ from a single riffle shuffle. Let K_f be the Markov chain associated with the random walk driven by f; that is, the random walk of successive riffle shuffles. The eigenvalues of this chain are $1, 2^{-1}, \ldots, 2^{-(n-1)}$ such that the multiplicity of 2^{-j} is equal to the number of permutations in S_n with $n - j$ cycles (an excellent exposition is given in [14]).

It is interesting that the random walk driven by the riffle shuffle (and likewise the generalisation that starts by partitioning a deck of cards into m parts) are related to the family of hyperplane walks. Indeed, each riffle shuffle walk appears as a reversed hyperplane walk on a braid arrangement [13].

5 Reachability for Bernoulli Cycles

In this section we consider a family of Markov chains where the Stochastic Reachability Problem is decidable. This family of chains does not fall into either of the previous families considered in this article.

Given a rational constant $0 \leq p \leq 1$, we define the entries of a circulant stochastic matrix $K \in \mathbb{Q}^{d\times d}$ with rows and columns indexed by $\{1, 2, \ldots, d\}$ as follows:

$$K(u,v) = \begin{cases} p & \text{if } v = u+1 \pmod d, \\ 1-p & \text{if } v = u-1 \pmod d, \\ 0 & \text{otherwise.} \end{cases}$$

Figuratively speaking, the associated random walk is determined by the sequence of flips of a weighted coin. We call the Markov chain associated to K a *Bernoulli cycle*. In this section we consider the Stochastic Reachability Problem (K, x, y, r) under the restriction that K is a Bernoulli cycle.

▶ **Theorem 25.** *The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable when $K \in \mathbb{Q}^{d\times d}$ is a Bernoulli cycle.*

Proof. There are two cases where we refer the reader to our previous discussions. First, in the event that $p \in \{0, 1\}$, K is an irreducible permutation matrix. In this instance the Stochastic Reachability Problem is trivially decidable. Second, if $p = 1/2$ then the Markov chain is reversible; this claim is a trivial consequence of Kolmogorov's criterion applied to a symmetric matrix. Thus the Stochastic Reachability Problem is decidable in this instance by Proposition 16.

It remains to consider $p \in (0, 1/2) \cup (1/2, 1)$. In fact, without loss of generality, we can assume that $p \in (1/2, 1)$ by using symmetry. So suppose that $p \in (0, 1/2)$. Let us take a short diversion and consider the function $f_p \colon [0, 2\pi) \to \mathbb{R}$ given by $f_p(\theta) := |\mathrm{e}^{2\mathrm{i}\theta}p + (1-p)|^2$. It is easily shown that

$$f_p(\theta) = 1 - 2p(1-p)(1-\cos(2\theta)).$$

Then the derivative of f_p with respect to θ is given by $f_p'(\theta) = -2p(1-p)\sin(2\theta)$. Further analysis shows that f_p achieves its global maxima at $\theta = 0, \pi$.

Let us return to the chain associated with K. By Proposition 13, the d eigenvalues of K are given by

$$\lambda_j = p\mathrm{e}^{2\pi\mathrm{i}j/d} + (1-p)\mathrm{e}^{-2\pi\mathrm{i}j/d} = \mathrm{e}^{-2\pi\mathrm{i}j/d}(p\mathrm{e}^{4\pi\mathrm{i}j/d} + (1-p))$$

for $j \in \{1, \ldots, d\}$. Observe that the modulus of eigenvalue λ_j of K is given by $f_p(\mathrm{e}^{2\pi\mathrm{i}j/d})$. In light of our analysis, there are two cases to consider. First, if d is even then K has two simple eigenvalues -1 and 1 that are strictly greater in modulus than all other elements of the spectrum. Second, if $d = 2k+1$ is odd then K has a simple maximal eigenvalue 1, $1 > |\lambda_k| = |\lambda_{k+1}|$, and all other eigenvalues are strictly smaller in modulus than λ_k and λ_{k+1}.

There exists $n \in \mathbb{N}$ such that $x^\top K^n y = r$ if and only if the linear recurrence sequence $\langle x^\top K^m y - r\rangle_m$ vanishes at index n. We can assume, without loss of generality, that $\langle x^\top K^m y - r\rangle_m$ is non-degenerate. Further, by Theorem 35 (Appendix A), all but the case $r = x^\top \Pi y$ are trivially decidable. (Here Π is the matrix whose rows are given by the stationary distribution of K, see Appendix A.) Thus we need only consider the case that $r = x^\top \Pi y$. From our spectral analysis of K and Lemma 8, the sequence $\langle x^\top K^m y - r\rangle_m$ has one simple characteristic root of maximal modulus when d is even and two simple characteristic roots of maximal modulus when d is odd. Thus we can apply Proposition 2, to deduce that these instances of the Stochastic Reachability Problem are decidable. We conclude that the Stochastic Reachability Problem is decidable for Bernoulli cycles. ◀

6 Reachability for G-circulant Chains

In this section we discuss decidable instances of the Stochastic Reachability Problem when the related Markov chain is driven by a *G-circulant* matrix where G is a finite group. This generalises our previous setting of walks on Bernoulli cycles ($\mathbb{Z}_d$-circulant chains). We begin by recalling terminology from the Fourier analysis of finite groups.

6.1 convolution

Let G be a finite group and $\{f(s)\}_{s\in G}$ a probability distribution on G so that $\sum_{s\in G} f(s) = 1$ and f is non-negative. Define the nth *convolution* of f recursively as follows:

$$f * f(t) = \sum_{s\in G} f(ts^{-1})f(s),$$
$$f^{*n}(t) = f^{*(n-1)} * f(t) \text{ for } n \in \{2, 3, \ldots\}.$$

For a random walk that starts at the identity and selects vertices according to the distribution f at each time instance, the quantity $f * f(t)$ represents the probability of the walk reaching vertex t after two steps. Similarly, $f^{*n}(t)$ is the probability of reaching vertex t after n steps.

Let K_f be the $|G| \times |G|$ stochastic matrix associated with the Markov chain model of the random walk driven by f. Then K_f is a *convolution matrix* with entries given by $K_f(s,t) := f(ts^{-1})$. Herein we shall assume that $\text{supp}(f)$ both generates G and is not contained in a coset of a proper normal subgroup of G. Under these mild assumptions the associated Markov chain is irreducible and aperiodic (see, for example, [44, Proposition 2.3] and the references listed in [11, pg. 23]). Without loss of generality we shall assume that every random walk starts at the identity e_G. Thus $K_f^n(e_G, t)$ is the probability of landing on vertex t after n steps.

▶ Remark 26. We call a probability distribution $f\colon G \to \mathbb{R}$ *symmetric* if $f(s) = f(s^{-1})$ for each $s \in G$. If f is symmetric then the associated stochastic transition matrix K_f is symmetric since $K_f(s,t) = f(ts^{-1}) = f(st^{-1}) = K_f(t,s)$ for each pair $s, t \in G$. We note the random walk driven by f is reversible in such cases. This claim follows as a trivial application of Kolmogorov's criterion. Thus by Proposition 16, the Stochastic Reachability Problem is decidable in such instances.

6.2 G-circulant matrices

Several accounts in the literature pay special attention to convolution schemes where f is a *class function*; that is to say, f is constant on conjugacy classes so that for each pair $s, t \in G$ one has $f(sts^{-1}) = f(t)$ [11, 12, 13]. A function that is both a class function and a probability distribution is sometimes called a *central probability* [44]. We note that the matrix K_f is *left invariant*; i.e., $K_f(hs, ht) = K_f(s,t)$ if and only if f is a class function. We follow Diaconis' convention [12, §3.E.4] and call such matrices *G-circulant*.

When f is a class function, the transform $\hat{f}(\rho)$ associated with an irreducible representation ρ is diagonal (see, for example, [12]). The next corollary follows from Schur's lemma, see [12, Corollary 1].

▶ **Corollary 27.** *Let G be a finite group and K_f a G-circulant matrix associated with a class function $f\colon G \to \mathbb{R}$. Then K_f is unitarily diagonalisable so that $K_f = \Phi^\dagger D \Phi$ where Φ is a unitary matrix and D is a diagonal matrix. The eigenvalues of K_f are indexed by the set of irreducible permutations of G so that*

$$\lambda_j = \frac{1}{d_{\rho_j}} \sum_{t\in G} f(t)\overline{\chi_j(t)}.$$

Here the eigenvalue λ_j has multiplicity $d_{\rho_j}^2$ and $\chi_j(t) := \text{tr}(\rho_j(t))$ is the character *associated with $\rho_j(t)$.*

Let G be a finite group. The family of such G-circulant matrices is closed under sum, product, transpose, conjugation, and inversion. In addition, all such matrices commute and are simultaneously diagonalisable.

▶ **Example 28** (Random walks on Abelian groups)**.** Consider $\mathbb{Z}_p$, the integers modulo p, as $p \in \mathbb{N}$ distinct states placed on the circumference of a circle. The symmetric random walk on $\mathbb{Z}_p$ permits a state transition that is either one step clockwise or one step anti-clockwise (both equally likely) at each time instance. This random walk is driven by the probability distribution f as follows:

$$f(\sigma) = \begin{cases} 1/2 & \text{if } \sigma \in \{1, p-1\}, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathbb{Z}_p$ is an Abelian group, each of its conjugacy classes consists of a single element. Thus f is a class function—in fact, by this same argument, any probability distribution on $\mathbb{Z}_p$ is a class function. Of course this argument holds for any Abelian group G: $sts^{-1} = t$ holds for all pairs $s, t \in G$. Thus any function from an Abelian group G into $\mathbb{R}$ is a class function.

One common application for random walks on groups is card shuffling: in such cases the walk takes place on S_N, the *symmetric group on N elements*, where N is the number of cards in the deck. At each time instance the current permutation $s \in S_N$ is recorded and the next state t is selected with probability $f(ts^{-1})$.

Recall that the conjugacy classes of S_N are precisely determined by *cycle types*. Since each element $\sigma \in S_N$ can be uniquely expressed as a product of disjoint cycles (up to ordering), two elements σ and τ in S_N have the same cycle type if they have the same number of cycles of equal length. It follows that the conjugacy classes are in a one-to-one correspondence with the partitions of N elements. This is easily seen when one considers that if, when written as a product of cycles, $\sigma = (a_{11}, \ldots, a_{1k_1})(a_{22}, \ldots, a_{2k_2}) \cdots (a_{\ell 1}, \ldots, a_{\ell k_\ell})$ then

$$\tau\sigma\tau^{-1} = (\tau(a_{11}), \ldots, \tau(a_{1k_1}))(\tau(a_{22}), \ldots, \tau(a_{2k_2})) \cdots (\tau(a_{\ell 1}), \ldots, \tau(a_{\ell k_\ell})).$$

▶ **Example 29** (Random transposition [11])**.** Consider the following shuffling scheme on N cards in a row. The cards are shuffled using random transpositions: a first card is randomly selected, then a second card is randomly selected, and then the two cards are transposed. Each selection is uniformly distributed. In the case that the first and second choice of cards coincide then no permutation occurs. The probability of performing permutation σ is given by

$$f(\sigma) = \begin{cases} 1/N & \text{if } \sigma = e_G, \\ 2/N^2 & \text{if } \sigma \text{ is a transposition,} \\ 0 & \text{otherwise.} \end{cases}$$

The identity and the set of transpositions on S_N are two conjugacy classes. Thus f is a class function. As a corollary of Theorem 31 below, the Markov Reachability Problem is decidable for the walk driven by random transposition.

▶ Remark 30. The conjugacy classes of S_N are particularly well-behaved. In fact, for each $N \in \mathbb{N}$, S_N is ambivalent. Here a group is *ambivalent* if for each $x \in G$, x and x^{-1} are conjugate elements.

▶ **Theorem 31.** *Let G be a finite ambivalent group. The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable when $K \in \mathbb{Q}^{|G| \times |G|}$ is G-circulant.*

We give two proofs of Theorem 31. The first proof, via character theory, follows from a technical lemma (see Problem 2.11 in [24]).

▶ **Lemma 32.** *Let G be a finite ambivalent group. Suppose that χ_j is the character associated with the irreducible representation ρ_j on G. Then χ_j is real-valued.*

Proof. Let $s \in G$. Suppose that ρ_j is irreducible, then $\rho_j(s)$ is similar to a diagonal representation $\text{diag}(\varepsilon_1, \ldots, \varepsilon_{d_j})$ [24, Lemma 2.15] where each ε_k is an nth root of unity for some $n \in \mathbb{N}$. Suppose that χ_j is the character associated with ρ_j. Then, by definition, $\chi_j(s) = \sum_{k=1}^{d_j} \varepsilon_k$ and immediately we have $\chi_j(s^{-1}) = \sum_{k=1}^{d_j} \varepsilon_k^{-1}$. Since $|\varepsilon_k| = 1$, we have that $\overline{\varepsilon_k} = \varepsilon_k^{-1}$ for each k. Thus $\chi_j(s^{-1}) = \overline{\chi_j(s)}$.

We note that χ_j is a class function and, since G is an ambivalent group, we have that $\chi_j(s) = \chi_j(s^{-1}) = \overline{\chi_j(s)}$ for each $s \in G$, from which the desired result follows. ◀

Proof of Theorem 31. Suppose that G is a finite ambivalent group and $f \colon G \to \mathbb{R}$ is a probability distribution on G. Recall the formula for the eigenvalues of K given in Corollary 27: $\lambda_j = \frac{1}{d_{\rho_j}} \sum_{t \in G} f(t)\overline{\chi_j(t)}$. Here the eigenvalue λ_j has multiplicity $d_{\rho_j}^2$. By Lemma 32, χ_j is real-valued for each j. It follows that each of the eigenvalues of K_f is real.

The Stochastic Reachability Problem with initialisation (K_f, x, y, r) is equivalent to determining whether the linear recurrence sequence $\langle x^\top K_f^n y - r \rangle_n$ vanishes at some index. We note this is a decidable instance of the Skolem Problem as each of the characteristic roots of this sequence are real (see Proposition 7 and Lemma 8). ◀

▶ Remark 33. An alternative proof of Theorem 31 uses the theory of reversible Markov chains. We note that a central probability f on an ambivalent group G is symmetric. That is to say, $f(ts^{-1}) = f(st^{-1})$ since every element and its inverse are conjugate. Thus the associated stochastic transition matrix K_f is symmetric. We then apply the observation in Remark 26: by Kolmogorov's criterion, K_f determines a reversible Markov chain. The decidability of Theorem 31 follows from Proposition 16.

Since the number of irreducible representations is equal to the number of conjugacy classes of G, the irreducible representations of G are all one-dimensional maps if and only if G is an Abelian group [11, Theorem 8 in §2.D]. When G is an Abelian group the family of unitary maps on G commute and so the class of unitary maps on G are simultaneously diagonalisable (generalising the property of $\mathbb{Z}_d$-circulant matrices).

It is interesting to consider the threshold variant of the Stochastic Reachability Problem for G-circulant matrices when G is Abelian. Decidability of $x^\top K^n y \geq r$ for all but finitely many $n \in \mathbb{N}$ is equivalent to determining whether the terms in the linear recurrence sequence $\langle x^\top K^n y - r \rangle_n$ are non-negative for all but finitely many $n \in \mathbb{N}$. This is an instance of the *Ultimate Positivity Problem* that we first met in Remark 17.

▶ **Theorem 34.** *Suppose that G is a finite Abelian group. Fix a tuple (K, x, y, r) where K is a G-circulant matrix. It is decidable whether $x^\top K^n y \geq r$ for all but finitely many $n \in \mathbb{N}$.*

Proof. Let G be a finite Abelian group. Then each of the irreducible presentations of G is a one-dimensional map and so for each j we have $d_j = 1$. (Note that a G-circulant matrix is diagonalisable if and only if G is Abelian.) By Corollary 27, K is diagonalisable. We deduce that each of the roots of the minimal polynomial associated with $\langle x^\top K^n y - r \rangle_n$ is simple (Lemma 8). Decidability of the Ultimate Positivity Problem for such linear recurrence sequences was established by Ouaknine and Worrell in [38] (Remark 17). Thus we have the desired result. ◀

References

1 S. Akshay, T. Antonopoulos, J. Ouaknine, and J. Worrell. Reachability problems for Markov chains. *Inf. Process. Lett.*, 115(2):155–158, February 2015. doi:10.1016/j.ipl.2014.08.013.

2 Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, Cambridge, MA, 2008. With a foreword by Kim Guldstrand Larsen.

3 D. Beauquier, A. Rabinovich, and A. Slissenko. A logic of probability with decidable model checking. *J. Log. Comput.*, 16(4):461–487, 2006. doi:10.1093/logcom/exl004.

4 Jon Louis Bentley, Daniel D. Sleator, Robert E. Tarjan, and Victor K. Wei. A locally adaptive data compression scheme. *Comm. ACM*, 29(4):320–330, 1986. doi:10.1145/5684.5688.

5 J. Berstel and M. Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bull. Soc. Math. France*, 104(2):175–184, 1976. doi:10.24033/bsmf.1823.

6 Pat Bidigare, Phil Hanlon, and Dan Rockmore. A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Mathematical Journal*, 99(1), jul 1999. doi:10.1215/s0012-7094-99-09906-4.

7 Anders Björner. Note: Random-to-front shuffles on trees. *Electron. Commun. Probab.*, 14(4):36–41, 2009. doi:10.1214/ECP.v14-1445.

8 Anders Björner. Random walks, arrangements, cell complexes, greedoids, and self-organizing libraries. In Martin Grötschel, Gyula O. H. Katona, and Gábor Sági, editors, *Building Bridges: Between Mathematics and Computer Science*, pages 165–203. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-85221-6_5.

9 Kenneth S. Brown. Semigroups, Rings, and Markov Chains. *Journal of Theoretical Probability*, 13(3):871–938, 2000. doi:10.1023/A:1007822931408.

10 Kenneth S. Brown and Persi Diaconis. Random walks and hyperplane arrangements. *Ann. Probab.*, 26(4):1813–1854, 1998. doi:10.1214/aop/1022855884.

11 Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988.

12 Persi Diaconis. Patterned matrices. In *Matrix theory and applications (Phoenix, AZ, 1989)*, volume 40 of *Proc. Sympos. Appl. Math.*, pages 37–58. Amer. Math. Soc., Providence, RI, 1990. doi:10.1090/psapm/040/1059483.

13 Persi Diaconis. Mathematical developments from the analysis of riffle shuffling. In *Groups, combinatorics & geometry (Durham, 2001)*, pages 73–97. World Sci. Publ., River Edge, NJ, 2003. doi:10.1142/9789812564481_0005.

14 Persi Diaconis and Jason Fulman. Carries, shuffling, and an amazing matrix. *Amer. Math. Monthly*, 116(9):788–803, 2009. doi:10.4169/000298909X474864.

15 Peter Donnelly. The heaps process, libraries, and size-biased permutations. *J. Appl. Probab.*, 28(2):321–335, 1991. doi:10.2307/3214869.

16 G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward. *Recurrence sequences*, volume 104 of *Math. Surveys Monogr.* Amer. Math. Soc., Providence, RI, 2003. doi:10.1090/surv/104.

17 James Allen Fill. An exact formula for the move-to-front rule for self-organizing lists. *J. Theoret. Probab.*, 9(1):113–160, 1996. doi:10.1007/BF02213737.

18 P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

19 F. Gantmacher. *The Theory of Matrices Vol. II*. AMS Chelsea Publishing Series. Amer. Math. Soc., 2000.

20 V. Halava, T. Harju, M. Hirvensalo, and J. Karhumäki. Skolem's problem–on the border between decidability and undecidability. Technical report, Turku Centre for Computer Science, 2005.

21 Michael A. Harrison. *Lectures on linear sequential machines*. Academic Press, New York-London, 1969.

22 W. J. Hendricks. The stationary distribution of an interesting Markov chain. *Journal of Applied Probability*, 9(1):231–233, 1972.

23 W. J. Hendricks. An account of self-organizing systems. *SIAM J. Comput.*, 5(4):715–723, 1976. doi:10.1137/0205050.

24 Irving Martin Isaacs. *Character theory of finite groups*. AMS, Providence, RI, 1976.

25 Ravindran Kannan and Richard J. Lipton. The orbit problem is decidable. In Raymond E. Miller, Seymour Ginsburg, Walter A. Burkhard, and Richard J. Lipton, editors, *Proceedings of the 12th Annual ACM Symposium on Theory of Computing, April 28-30, 1980, Los Angeles, California, USA*, pages 252–261. ACM, 1980. doi:10.1145/800141.804673.

26 Ravindran Kannan and Richard J. Lipton. Polynomial-time algorithm for the orbit problem. *J. ACM*, 33(4):808–821, 1986. doi:10.1145/6490.6496.

27 Sanjiv Kapoor and Edward M. Reingold. Stochastic rearrangement rules for self-organizing data structures. *Algorithmica*, 6(2):278–291, 1991. doi:10.1007/BF01759046.

28 F. Kelly. *Reversibility and stochastic networks*. John Wiley & Sons, Ltd., Chichester, 1979. Wiley Series in Probability and Mathematical Statistics.

29 I. Kra and S. Simanca. On circulant matrices. *Notices Amer. Math. Soc.*, 59(3):368–377, 2012. doi:10.1090/noti804.

30 C. Lech. A note on recurring series. *Ark. Mat.*, 2:417–421, 1953.

31 K. Mahler. Eine arithmetische Eigenschaft der Taylor-koeffizienten rationaler Funktionen. *Proc. Akad. Wet. Amst.*, 38:50–69, 1935.

32 K. Mahler and J. Cassels. On the Taylor coefficients of rational functions. *Math. Proc. Cambridge Philos. Soc*, 52(1):39–48, 1956. doi:10.1017/S0305004100030966.

33 Giovanni Manzini. An analysis of the Burrows-Wheeler transform. *J. ACM*, 48(3):407–430, 2001. doi:10.1145/382780.382782.

34 John McCabe. On serial files with relocatable records. *Operations Research*, 13(4):609–618, 1965.

35 M. Mignotte, T. Shorey, and R. Tijdeman. The distance between terms of an algebraic recurrence sequence. *J. Reine Angew. Math.*, pages 63–76, 1984.

36 J. Ouaknine and J. Worrell. Decision problems for linear recurrence sequences. In *Reachability problems*, volume 7550 of *Lecture Notes in Comput. Sci.*, pages 21–28. Springer, Heidelberg, 2012. doi:10.1007/978-3-642-33512-9_3.

37 J. Ouaknine and J. Worrell. Positivity problems for low-order linear recurrence sequences. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 366–379. ACM, New York, 2014. doi:10.1137/1.9781611973402.27.

38 J. Ouaknine and J. Worrell. Ultimate positivity is decidable for simple linear recurrence sequences. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, volume 8573 of *Lecture Notes in Comput. Sci.*, pages 330–341. Springer, 2014. doi:10.1007/978-3-662-43951-7_28.

39 R. M. Phatarfod. On the matrix occurring in a linear search problem. *Journal of Applied Probability*, 28(2):336–346, jun 1991. doi:10.2307/3214870.

40 Sheldon M. Ross. *Introduction to probability models*. Elsevier/Academic Press, Amsterdam, eleventh edition, 2014.

41 B. Ya. Ryabko. Data compression by means of a "book stack". *Problemy Peredachi Informatsii*, 16(4):16–21, 1980.

42 B. Ya. Ryabko, R. Nigel Horspool, and Gordon V. Cormack. Comments to: "A locally adaptive data compression scheme" [Comm. ACM **29** (1986), no. 4, 320–330] by J. L. Bentley, D. D. Sleator, R. E. Tarjan and V. K. Wei. *Comm. ACM*, 30(9):792–794, 1987. doi:10.1145/30401.315747.

43 L. Saloff-Coste. Lectures on finite Markov chains. In P. Bernard, editor, *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996*, pages 301–413. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. doi:10.1007/BFb0092621.

44 L. Saloff-Coste. Random walks on finite groups. In H. Kesten, editor, *Probability on Discrete Structures*, pages 263–346. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. doi:10.1007/978-3-662-09444-0_5.

45 R. Serfozo. *Basics of applied stochastic processes*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2009. doi:10.1007/978-3-540-89332-5.

46 T. Skolem. Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen und diophantischer Gleichungen. *8de Skand. Mat. Kongress, Stockholm (1934)*, pages 163–188, 1934.

47 Jacques Tits. *Buildings of spherical type and finite BN-pairs*. Lecture Notes in Mathematics, Vol. 386. Springer-Verlag, Berlin-New York, 1974.

48 M.L. Tsetlin. Finite automata and modeling the simplest forms of behavior. In M.L. Tsetlin, editor, *Automaton Theory and Modeling of Biological Systems*, volume 102 of *Mathematics in Science and Engineering*, pages 3–83. Elsevier, 1973. doi:10.1016/S0076-5392(08)60818-8.

49 N. Vereshchagin. Occurrence of zero in a linear recursive sequence. *Math. Notes Acad. Sci. USSR*, 38(2):609–615, Aug 1985. doi:10.1007/BF01156238.

A Reachability for Irreducible and Aperiodic Markov Chains

In this self-contained section we draw parallels between the Stochastic Reachability Problem and the study of convergence properties of Markov chains in probability theory. In particular, we narrow our focus to irreducible and aperiodic Markov chains.

A probability vector π such that $\pi K = \pi$ is called a *stationary distribution* of the matrix K. If K is both irreducible and aperiodic then there exists a unique stationary distribution as stated in Corollary 5. Here irreducibility ensures the existence and uniqueness of the stationary distribution. Aperiodicity of K is a sufficient condition for the stationary distribution to be a *steady state*; that is to say, $\lim_{n\to\infty} K^n(u,v) = \pi(v)$ for each u [43].

▶ **Theorem 35.** *The Stochastic Reachability Problem with initialisation (K, x, y, r) is decidable if $K \in \mathbb{Q}^{d\times d}$ is an irreducible and aperiodic matrix and $r \neq x^\top \Pi y$ where Π is the matrix whose rows are given by the stationary distribution of K.*

Proof. As K is an aperiodic stochastic matrix it has a unique stationary distribution $\pi K = \pi$ with $\sum \pi(j) = 1$. Further, the remaining eigenvalues of K have modulus at most $0 \leq \varepsilon < 1$. Then $x^\top K^n y = x^\top \Pi y + \mathrm{O}(\varepsilon^n)$. If $r \neq x^\top \Pi y$ then one can determine in finite time whether there exists $n \in \mathbb{N}$ such that $x^\top K^n y = r$. ◀

Let us discuss the general case for strongly connected Markov chains—the chains whose stochastic matrices are *irreducible*, but not necessarily aperiodic. For an irreducible stochastic matrix K with period $\rho \geq 2$ we note that, subject to the permutation of its rows and columns, K can be written in block form as

$$K = \begin{pmatrix} 0 & & & & & K_{1,d} \\ K_{2,1} & 0 & & & & \\ & K_{3,2} & & & & \\ & & & & & \\ & & & \ddots & 0 & \\ & & & & K_{d,d-1} & 0 \end{pmatrix}$$

in order that K^ρ admits a diagonal square block decomposition where each diagonal block is an aperiodic component. Then studying the matrix $K^{m\rho}$ reduces to a block-by-block consideration of powers of an aperiodic subsystem. Thus, for sufficiently large $m \in \mathbb{N}$, it is clear that there is a single $\ell \in \{0, 1, \ldots, \rho - 1\}$ such that $K^{m\rho+\ell}(1,2)$ is non-zero. Thus the Stochastic Reachability Problem reduces to asking whether there exists an $m \in \mathbb{N}$ such that $x^\top K^{m\rho+\ell} y = x^\top K^{m\rho+\ell} y = r$. That is to say, the Stochastic Reachability Problem for (K, w, y, r) with irreducible transition matrix K reduces to the Scalar Reachability Problem with (M, w, y, r) where $w^\top = x^\top K^\ell$ and $M = K^\rho$ is block diagonal. The above decomposition and reduction process is described in detail in [18, pp.341–342].

▶ Remark 36. **1.** Under the assumptions of irreducibility and aperiodicity, for a given initialisation of the Stochastic Reachability Problem there is a single value, $r = x^\top \Pi y$, for which decidability is at least as hard as the Skolem Problem. We should also note that, under these assumptions, there are two initialisations where decidability is trivial. For example, take $x = \pi$ then for each $n \in \mathbb{N}$ we have $x^\top K^n = \pi$ and so decidability of the Stochastic Reachability Problem is trivial for the case (K, π, y, r). In a similar vein, if we take $y = \frac{1}{d}\mathbb{1}$ to be the uniform distribution on the vertices of the chain then $K^n y = \frac{1}{d}\mathbb{1}$ for each $n \in \mathbb{N}$ and so the decidability of the Stochastic Reachability Problem is trivial for the case $(K, x, \frac{1}{d}\mathbb{1}, r)$.

2. It is motivating to consider the Stochastic Reachability Problem when $y = \pi$, the stationary distribution of an irreducible chain. In the case that the chain is also aperiodic, a fundamental result in probability theory is the (uniform) convergence from an initial distribution to the stationary distribution [43]. In some sense the Stochastic Reachability Problem is the analogous decision problem: given an initial distribution x, determine whether a given correlation is achieved between the nth-step distribution $x^\top K^n$ and π for some $n \in \mathbb{N}$.
3. One might naturally consider a *threshold* variant of this decision problem. For example, given an initialisation (K, x, y, r), determine whether the inequality $x^\top K^n y \geq r$ holds for each $n \in \mathbb{N}$ (or for all but finitely many $n \in \mathbb{N}$). These decision problems appear frequently in the literature of linear recurrence sequences as the Positivity Problem and the Ultimate Positivity Problem, respectively [38, 37]. In light of the previous observations for irreducible and aperiodic chains, the hard cases to determine such thresholds occur when $r = x^\top \Pi y$.