LOOP-ERASED WALKS AND TOTAL POSITIVITY

SERGEY FOMIN

ABSTRACT. We consider matrices whose elements enumerate weights of walks in planar directed weighted graphs (not necessarily acyclic). These matrices are totally nonnegative; more precisely, all their minors are formal power series in edge weights with nonnegative coefficients. A combinatorial explanation of this phenomenon involves loop-erased walks. Applications include total positivity of hitting matrices of Brownian motion in planar domains.

1. INTRODUCTION

Then [Pooh] dropped two in at once, and leant over the bridge to see which of them would come out first; and one of them did; but as they were both the same size, he didn't know if it was the one which he wanted to win, or the other one.

A. A. Milne [17, Chapter VI]

Consider a stationary Markov process whose state space is a connected planar domain Ω, or a discrete subset thereof. In the continuous version, assume that all trajectories of this process are continuous; in the discrete version, the possible transitions from one state to another are described by a planar directed graph.

Let B be an absorbing subset of the boundary of Ω (i.e., once the process reaches a state in B, it stays there), and suppose that the points $a_1, a_2 \notin B$ and $b_1, b_2 \in B$ are such that any possible trajectory that goes from a_1 to b_2 must intersect any trajectory going from a_2 to b_1. See Figure 1.
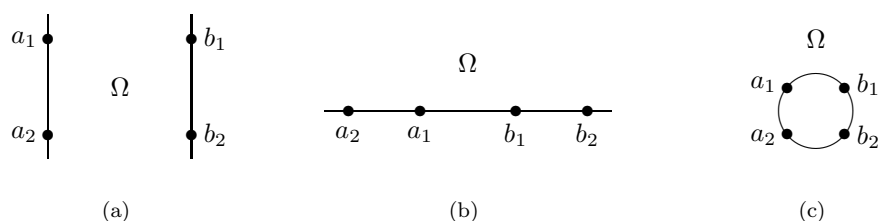

FIGURE 1. Some possible locations of a_1, a_2, b_1, b_2

Let us now consider two independent realizations of our Markov process, starting at points a_1 and a_2, respectively. Suppose that we know that the other endpoints of these trajectories are b_1 and b_2, although we do not know which of the two ended

up where; neither do we know how much time it took for each one to reach its destination. We then ask the usual Bayesian question: which of the two possibilities (i.e., $a_1 \rightsquigarrow b_1$, $a_2 \rightsquigarrow b_2$ or $a_1 \rightsquigarrow b_2$, $a_2 \rightsquigarrow b_1$) is more likely, and by how much? In particular, is it necessarily true that the first match-up is more likely than the second one?

This paper answers the last question—and various related ones—in the affirmative. To make it more precise, let x_{ij} denote the hitting probability (or the corresponding density) that the trajectory originating at a_i will hit the target B at location b_j. Then our claim amounts to the inequality $x_{11}x_{22} \geq x_{12}x_{21}$. In other words, the determinant of the submatrix $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ of the hitting matrix X of our Markov process is nonnegative. In fact, the determinant of *any* square submatrix of X can be shown to be nonnegative, so the hitting matrix is *totally nonnegative*. To explain this phenomenon, we provide a combinatorial interpretation of such determinants in the spirit of the celebrated Karlin-McGregor nonintersecting path approach; note that this approach cannot be applied directly to the case under consideration, since the trajectories are allowed to self-intersect. The crucial ingredient of our combinatorial construction is Lawler's concept of *loop-erased walk*.

To give a flavor of the main result of this paper, here is its simplest version, stated loosely: we prove that the 2×2 determinant $x_{11}x_{22} - x_{12}x_{21}$ is equal to the probability (or density) that the trajectories that started out at a_1 and a_2 will end up at b_1 and b_2, respectively, and furthermore the first trajectory will never intersect the loop-erased part of the second one.

We give conditions under which the minors of hitting matrices are *positive* (as opposed to nonnegative), so the matrices themselves are *totally positive*. This latter property immediately implies a number of others, for example, simplicity and positivity of the spectrum, and the variation-diminishing property.

The paper is organized as follows. Sections 2–5 are devoted to preliminaries of various kinds. Section 2 introduces walk matrices and hitting matrices of directed networks. Section 3 reviews classical results by Karlin-McGregor and Lindström on total positivity of walk matrices of acyclic directed networks or associated Markov chains. Section 4 gives an account of some of the results obtained in [6, thesis] on resistor networks and their Dirichlet-to-Neumann maps; although these results are not used in the rest of the paper, they provided the primary motivation for this investigation (cf. Acknowledgments below). Section 5 introduces loop-erased walks.

The main results are presented in Sections 6–7. We give combinatorial formulas for the minors of walk matrices and hitting matrices of directed networks (see Theorems 6.1, 6.4, 7.1, and 7.2) and Markov processes (see Theorems 7.4 and 7.5), and apply them to various examples of discrete and continuous Markov processes, such as one-dimensional Bernoulli process and two-dimensional Brownian motion.

2. Directed networks and associated matrices

Most of the material in this section is basic graph theory (see, e.g., [21, Sec. 4.7]), although some of the terminology is not standard.

A *directed network* $\Gamma = (V, E, w)$ is a directed graph (loops and multiple edges are allowed) with vertex set V and edge set E, together with a family of formal variables $\{w(e)\}_{e \in E}$, the *weights* of the edges.

The notation $a \xrightarrow{e} b$ will mean "edge e goes from vertex a to vertex b." More generally,

$$(2.1) \qquad a = a_0 \xrightarrow{e_1} a_1 \xrightarrow{e_2} a_2 \xrightarrow{e_3} \cdots \xrightarrow{e_m} a_m = b$$

will denote that the edges $e_1, \dots, e_m$ form a walk of length m from a to b that successively passes through $a_1, a_2, \dots$. We will sometimes shorten (2.1) to $a \xrightarrow{\pi} b$, where π denotes the walk. The weight of π is by definition given by $w(\pi) = \prod w(e_i)$. The degenerate walk $a \longrightarrow a$ of length 0 has weight 1.

No finiteness conditions are imposed on Γ; we only ask that the number of walks of any fixed length between any two vertices $a, b \in V$ is at most countable, so that the following formal power series is well defined:

$$(2.2) \qquad W(a, b) = \sum_m \sum w(e_1) \cdots w(e_m),$$

where the second summation goes over all walks of length m from a to b, as given in (2.1). In other words, $W(a, b)$ is simply the generating function for the weights of all walks from a to b. The matrix $W = (W(a, b))_{a,b \in V}$ will be called the *walk matrix* of the network Γ. (The term "Green function" would perhaps fit better, but it is already overused.)

One easily sees that $W = (I - Q)^{-1}$, where

$$Q = (Q(a, b))_{a,b \in V}, \quad Q(a, b) = \sum_{a \xrightarrow{e} b} w(e),$$

is the "weighted adjacency matrix" of the network, and I is the identity map/matrix. The identity $W = QW + I$ shows that for any $b \in V$, the function $W_b : a \mapsto W(a, b)$ is "Q-harmonic" on $V \setminus \{b\}$, i.e., satisfies $W_b(a) = \sum_c Q(a, c) W_b(c)$, for $a \neq b$.

Example 2.1. See Figure 2.

$$Q = \begin{pmatrix} 0 & q_1 & 0 \\ 0 & 0 & q_2 \\ q_3 & 0 & 0 \end{pmatrix} \qquad W = K \begin{pmatrix} 1 & q_1 & q_1 q_2 \\ q_2 q_3 & 1 & q_2 \\ q_3 & q_1 q_3 & 1 \end{pmatrix}$$
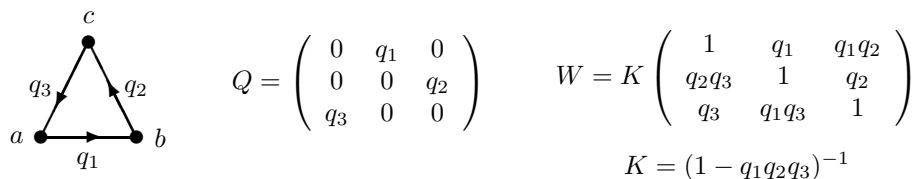
$$K = (1 - q_1 q_2 q_3)^{-1}$$

Figure 2. Directed network and its walk matrix

SERGEY FOMIN

In most applications, the weight variables $w(e)$ are specialized to numerical (usually nonnegative real) values. One should then be careful while introducing the walk matrix W, since the power series (2.2) may easily diverge. The following example is typical in this regard.

Example 2.2. Consider a Markov chain with a countable set of states V. Whenever the transition probability $p(a, b)$ from state a to state b is nonzero, let $a \xrightarrow{e} b$ be an edge of E, and let its weight be $w(e) = t \cdot p(a, b)$, where t is a formal variable. Then $W = \sum_{m=0}^{\infty} t^m P^m = (I - tP)^{-1}$, where $P = (p(a, b))_{a,b \in V}$ is the transition matrix of the Markov chain. Note that we cannot simply put $t = 1$ and $w(e) = p(a, b)$ for $a \xrightarrow{e} b$, since this may lead to divergence of the power series for $W(a, b)$. If, however, this power series does converge, then

$$(2.3) \qquad W(a, b) = \left\langle \begin{array}{c} \text{expected number of times the process} \\ \text{visits } b \text{ provided it starts at } a \end{array} \right\rangle.$$

Example 2.3 (Bernoulli random walk)**.** Let $V = \mathbb{Z}$, with edges connecting n and $n + 1$ (in both directions) for all $n \in \mathbb{Z}$. Let the weights be given by

$$(2.4) \qquad w(e) = \begin{cases} p & \text{if } n \xrightarrow{e} n + 1; \\ q & \text{if } n + 1 \xrightarrow{e} n, \end{cases}$$

where $p + q = 1$ and $p > q > 0$. It is classically known (see, e.g., [20, Section 1, (11)]) that

$$(2.5) \qquad W(n, n + k) = \begin{cases} (p - q)^{-1} & \text{if } k \geq 0; \\ (p - q)^{-1} \left(\dfrac{p}{q}\right)^k & \text{if } k \leq 0. \end{cases}$$

Hitting matrices. In what follows, we will also need a variation of the notion of a walk matrix, which arises in situations where Γ comes equipped with a distinguished subset of vertices $\partial \Gamma \subset V$, called *boundary* of Γ. Accordingly, $\text{int}\Gamma = V - \partial \Gamma$ is the *interior* of Γ.

For $a \in V$ and $b \in \partial \Gamma$, let $X(a, b)$ be the weight generating function for the walks from a to b of nonzero length whose internal vertices all lie in the interior of Γ. Formally,

$$X(a, b) = \sum_{m > 0} \sum w(e_1) \cdots w(e_m),$$

where the second sum is over all walks (2.1) such that $a_i \notin \partial \Gamma$ for $0 < i < m$. The *hitting matrix* X is then defined by $X = (X(a, b))_{a \in V, b \in \partial \Gamma}$. Also, the term "hitting matrix" will sometimes be used for submatrices of X, with the notation $X_{A,B} = (X(a, b))_{a \in A, b \in B}$ for $A \subset V$, $B \subset \partial \Gamma$.

Example 2.4. See Figure 3.

Example 2.5. Let us continue with Example 2.2. Unlike in the case of walk matrices, it is now permissible to set $t=1$. Then $X(a, b)$ is nothing but the *hitting probability;* more precisely, it is the probability that the Markov process with initial state a visits the boundary after leaving a, and furthermore the first boundary state it hits is b.

$$A = \{a_1, a_2\}$$
$$B = \{b_1, b_2\}$$

$$X_{A,B} = \frac{1}{1 - q_1 q_2 q_3} \begin{pmatrix} q_4 q_5 & q_1 q_3 q_4 q_7 \\ q_1 q_2 q_5 q_6 & q_1 q_6 q_7 \end{pmatrix}$$
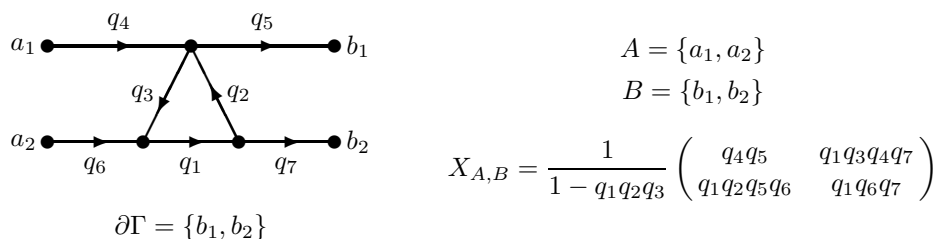
$$\partial\Gamma = \{b_1, b_2\}$$

FIGURE 3. Hitting matrix of a directed network

The matrix $X_{\partial\Gamma, \partial\Gamma}$ can be expressed in terms of the weighted adjacency matrix Q using the notion of *Schur complement.* Recall that for a matrix M of block structure $M = \begin{pmatrix} C & D \\ E & F \end{pmatrix}$, the Schur complement M/F is defined by $M/F = C - DF^{-1}E$. The following statement is a straightforward corollary of the definitions.

Proposition 2.6. *We have*

$$(2.6) \qquad I - X_{\partial\Gamma, \partial\Gamma} = (I - Q)/(I - Q_{int\Gamma, int\Gamma}),$$

where $Q_{int\Gamma, int\Gamma} = (Q(a,b))_{a,b \in int\Gamma}$. *(By a common abuse of notation, in formula (2.6), the letter I denotes identity matrices of three different sizes.)*

3. ACYCLIC GRAPHS AND TOTAL POSITIVITY

This section reviews some classical results concerning directed networks whose underlying graph is *acyclic.* Although these results will not be used in subsequent proofs, they will serve as the inspiration in the study of the general (non-acyclic) case.

For an acyclic network Γ, let the boundary $\partial\Gamma$ be the set of *sinks* in Γ:

$$\partial\Gamma = \{a \in V : \nexists a \xrightarrow{e} b\}.$$

Under these assumptions, the hitting matrix X is a submatrix of the walk matrix of Γ, and the entries $X(a,b)$ of X enumerate the paths from each vertex a to each sink b with respect to their weight.

For an arbitrary pair of totally ordered subsets $A = \{a_1, \dots, a_k\} \subset V$ and $B = \{b_1, \dots, b_k\} \subset \partial\Gamma$ of equal cardinality k, consider the corresponding $k \times k$ minor of the hitting matrix:

$$(3.1) \qquad \det(X_{A,B}) = \det\left(X(a_i, b_j)_{\substack{i=1,\dots,k \\ j=1,\dots,k}} \right).$$

The following fundamental observation goes back to Karlin and McGregor [14] (in the case of Markov chains) and Lindström [16].

Theorem 3.1. *In an acyclic network Γ, the minors of the hitting matrix are given by*

$$(3.2) \quad \det(X_{A,B}) = \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \sum_{\substack{\text{non-intersecting path families} \\ a_1 \xrightarrow{\pi_1} b_{\sigma(1)}, \ \dots, \ a_k \xrightarrow{\pi_k} b_{\sigma(k)}}} w(\pi_1) \cdots w(\pi_k),$$

where the first sum is over all permutations σ in the symmetric group S_k (interpreted as bijections $A \to B$), and the second sum runs over all families of vertex-disjoint paths which connect the vertices in A to the sinks in B, assigned to them

by σ. (Recall that the weight $w(\pi)$ of a path is by definition equal to the product of the weights of its edges.)

Theorem 3.1, which is proved by a path-switching argument, has many applications in the theory of stochastic processes, enumerative combinatorics, and beyond (see, e.g., [2] and references therein). Most of these applications are based on the following corollary, treating the case of *planar* acyclic networks.

Corollary 3.2. *Let Ω be a planar simply-connected domain whose boundary $\partial\Omega$ is a simple closed Jordan curve. Let Γ be an acyclic network, embedded into Ω in such a way that the edges of Γ do not intersect. Suppose that $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\} \subset \partial\Gamma$ lie on the boundary $\partial\Omega$, and furthermore these vertices appear in the order $a_1, \ldots, a_k, b_k, \ldots, b_1$ if traced counter-clockwise. Then*

$$(3.3) \qquad \det(X_{A,B}) = \sum_{\substack{\text{non-intersecting path families} \\ a_1 \xrightarrow{\pi_1} b_1, \ \ldots, \ a_k \xrightarrow{\pi_k} b_k}} w(\pi_1) \cdots w(\pi_k) \, .$$

In particular, if the weights of all edges are nonnegative, then $\det(X_{A,B}) \geq 0$, and furthermore $\det(X_{A',B'}) \geq 0$ for any subsets $A' \subset A$ and $B' \subset B$.

Thus any such matrix $X_{A,B}$ is *totally nonnegative* (cf., e.g., [10, 11]).

4. Resistor networks and Dirichlet-to-Neumann maps (after [6, 12])

This section reviews some well known and some fairly recent results from discrete potential theory—or, equivalently, the theory of resistor networks,—following D. V. Ingerman's thesis [12] and the paper [6] by E. Curtis, D. V. Ingerman, and J. Morrow.

A resistor network is essentially an undirected graph (V, E) (i.e., a graph in which each edge $a \xrightarrow{e} b$ is paired with an edge $b \xrightarrow{e'} a$) together with a *conductivity* function γ satisfying $\gamma(e) = \gamma(e')$. The conductivities $\gamma(e)$ can be viewed as either positive reals, or as independent variables taking positive real values.

All our networks will be presumed connected.

For a "potential" function $u : V \to \mathbb{R}$, the corresponding "current" function I_u (more precisely, the function that gives currents out of each vertex, for a given collection of potentials) is given by

$$I_u(a) = \sum_{a \xrightarrow{e} b} \gamma(e)(u(a) - u(b)) \, .$$

The *Kirchhoff matrix* $K = (K(a,b))_{a,b \in V}$ of a resistor network represents the linear map $u \mapsto I_u$ from potentials on V to the corresponding currents. Thus

$$K(a,b) = \begin{cases} -\displaystyle\sum_{a \xrightarrow{e} b} \gamma(e) & \text{if } a \neq b, \\[2mm] \displaystyle\sum_{a \xrightarrow{e} c, c \neq a} \gamma(e) & \text{if } a = b. \end{cases}$$

Suppose that the set of vertices V is partitioned into two disjoint subsets $\partial\Gamma$ and $int\Gamma$, as above. The potential functions $u : V \to \mathbb{R}$ satisfying *Kirchhoff's Law*

$$I_u(a) = \sum_b K(a,b)u(b) = 0 , \quad \text{for } a \in int\Gamma ,$$

are called γ-*harmonic*. In other words, the value of u at an interior node a should equal the weighted average of its values at the neighbors of a.

The values of a γ-harmonic function u at the interior nodes of Γ are uniquely determined by the values of u at the boundary nodes. This allows us to define the *Dirichlet-to-Neumann map* Λ of the resistor network as the map that sends a function $f : \partial\Gamma \to \mathbb{R}$ to the current out of the boundary nodes of the unique γ-harmonic continuation of f. The Dirichlet-to-Neumann map is represented by the *response matrix* $\Lambda = (\Lambda(a,b))_{a,b\in\partial\Gamma}$ of the network. A straightforward calculation yields the following formula.

Proposition 4.1 ([6, Theorem 3.2]). *The response matrix Λ is the Schur complement in the Kirchhoff matrix K:*

$$\Lambda = K/K_{int\Gamma,int\Gamma} ,$$

where $K_{int\Gamma,int\Gamma}$ is the submatrix $(K(a,b))_{a,b\in int\Gamma}$.

For further discussion of Dirichlet-to-Neumann maps, see [6] and references therein.

The comparison of Propositions 4.1 and 2.6 shows that the relationship between the Kirchhoff matrix and the response matrix is very similar to the relationship between the weighted adjacency matrix and the hitting matrix $X_{\partial\Gamma,\partial\Gamma}$. This analogy is not accidental: we will see very soon that the response matrix of a resistor network is closely related to the hitting matrix of the Markov chain associated to the network in the standard way [3, 8]. Recall that this Markov chain has the vertices in V as its states, and the numbers

$$p(a,b) = \frac{\sum_{a\xrightarrow{e}b} \gamma(e)}{\sum_{a\xrightarrow{e}c} \gamma(e)}$$

as transition probabilities. The following statement is a straightforward consequence of the definitions.

Proposition 4.2 ([12, (6.6)]). *The hitting matrix $X_{\partial\Gamma,\partial\Gamma}$ of the Markov chain associated to a resistor network is related to network's response matrix Λ by*

(4.1) $$X_{\partial\Gamma,\partial\Gamma} = I - K_0^{-1}\Lambda,$$

where $K_0 = (K_0(a,b))_{a,b\in\partial\Gamma}$ is the diagonal part of a principal submatrix of the Kirchhoff matrix: $K_0(a,b) = K(a,b)\delta_{ab}$.

Formula (4.1) shows that the hitting matrix of a resistor network is, up to renormalization, its response matrix, that is, the matrix of the Dirichlet-to-Neumann map.

Next, let us look at the minors of the response (or hitting) matrix of a resistor network. Since the underlying directed graph of the network is almost never acyclic (indeed, any two adjacent vertices a and b give rise to a cycle $a \xrightarrow{e} b \xrightarrow{e'} a$), the nonintersecting-path formulas from Section 3 do not apply. Instead, we will refer to a remarkable determinantal formula discovered by D. Ingerman (see [6, Lemma 4.1]), reproduced below without proof.

Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$ be two disjoint ordered subsets of $\partial\Gamma$ of the same cardinality k. Let

$$\Lambda_{A,B} = (\Lambda(a_i, b_j))_{\substack{i=1,\ldots,k \\ j=1,\ldots,k}}$$

denote the corresponding submatrix of the response matrix Λ.

Theorem 4.3 (D. Ingerman). *The minor $\det(\Lambda_{A,B})$ of Λ is given by*

(4.2)

$$\det(\Lambda_{A,B}) = (-1)^k \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \sum_{\pi = (\pi_1, \ldots, \pi_k) \in \mathcal{C}_\sigma(A,B)} w(\pi_1) \cdots w(\pi_k) \frac{\det(K_\pi)}{\det(K_{int\Gamma, int\Gamma})} \,,$$

where, for a permutation $\sigma \in S_k$, we denote

(4.3)

$$\mathcal{C}_\sigma(A,B) = \left\{ \pi = (\pi_1, \ldots, \pi_k) : \begin{array}{l} \textit{paths } a_1 \xrightarrow{\pi_1} b_{\sigma(1)}, \ \ldots, \ a_k \xrightarrow{\pi_k} b_{\sigma(k)} \ \textit{are} \\ \textit{vertex-disjoint and stay in the interior of } \Gamma \end{array} \right\} \,,$$

and K_π denotes the submatrix of K whose rows and columns are labelled by the interior vertices that do not lie on any of the paths in π.

For a connected resistor network with positive conductivities, the submatrix $K_{int\Gamma, int\Gamma}$ is always invertible (cf. Proposition 4.1), so the denominator in (4.2) does not vanish.

It follows from Proposition 4.2 that the corresponding minors of the hitting matrix of the network are given by

$$\det(X_{A,B}) \quad = \quad \frac{(-1)^k}{\prod_{a \in A} K(a,a)} \det(\Lambda_{A,B})$$

(4.4)

$$= \quad \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \sum_{\pi \in \mathcal{C}_\sigma(A,B)} \frac{w(\pi_1) \cdots w(\pi_k) \det(K_\pi)}{\det(K_{int\Gamma, int\Gamma}) \prod_{a \in A} K(a,a)} \,,$$

under the assumptions of Theorem 4.3.

Just as in Section 3—and for exactly the same reasons,—formulas (4.2) and (4.4) simplify considerably in the case of a *planar* resistor network, yielding the following results.

Corollary 4.4. *Assume that a connected resistor network Γ with positive conductivities is embedded into a simply-connected planar domain Ω (whose boundary $\partial\Omega$ is a simple Jordan curve) so that the edges of Γ do not intersect, and $\partial\Gamma \subset \partial\Omega$. Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$ be two equinumerous subsets of boundary vertices which lie on non-overlapping segments of $\partial\Omega$; we presume that the vertices in A are ordered clockwise, while those in B are ordered counter-clockwise. (See Figure 4.) Then*

$$(4.5) \qquad \det(\Lambda_{A,B}) = (-1)^k \sum_{\pi \in \mathcal{C}_{id}(A,B)} w(\pi_1) \cdots w(\pi_k) \cdot \frac{\det(K_\pi)}{\det(K_{int\Gamma, int\Gamma})} \,,$$

where we use the notation of Theorem 4.3, and id ∈ S_k is the identity permutation. Consequently,

$$(4.6) \qquad \det(X_{A,B}) \;=\; \frac{\displaystyle\sum_{\pi \in \mathcal{C}_{id}(A,B)} w(\pi_1) \cdots w(\pi_k) \det(K_\pi)}{\det(K_{int\Gamma, int\Gamma}) \displaystyle\prod_{a \in A} K(a,a)} \; .$$

It is not hard to show [6, Lemma 3.1] that for any connected (not necessarily planar) resistor network and any subset of vertices $V' \subset V$, the determinant of the corresponding principal submatrix $K_{V'} = (K(a,b))_{a,b \in V'}$ of the Kirchhoff matrix K is *positive*. Thus all determinants appearing in the right-hand sides of (4.2)–(4.6) are positive, and we obtain the following corollary.

Corollary 4.5. *Under the assumptions of Corollary 4.4, $(-1)^k \det(\Lambda_{A,B}) \geq 0$ and $\det(X_{A,B}) \geq 0$.*

The inequality $(-1)^k \det(\Lambda_{A,B}) \geq 0$ in Corollary 4.5 was first obtained by E. Curtis, E. Mooers, and J. Morrow [7] (for "well-connected circular planar networks") and by Y. Colin de Verdière [5] (general case). A short proof based on Theorem 4.3 was given in [6].

Applying the inequalities in Corollary 4.5 to arbitrary subsets $A' \subset A$ and $B' \subset B$ of the same cardinality, we arrive at the following result.

Corollary 4.6. *Under the assumptions of Corollary 4.4, the submatrices $-\Lambda_{A,B}$ and $X_{A,B}$ of the hitting matrix X and the negated response matrix $-\Lambda$ are totally nonnegative. Furthermore, all these submatrices are totally positive (i.e., all their minors are positive) provided that the set $\mathcal{C}_{id}(A,B)$ is not empty for any allowable choice of A and B, i.e., provided there exists at least one family of vertex-disjoint paths connecting A and B through the interior of Ω.*

The original motivation for this paper was to provide a combinatorial explanation for this total positivity phenomenon, in the spirit of Corollary 3.2. Such an explanation is given in Section 7, where total nonnegativity of hitting matrices of directed planar networks is established by purely combinatorial means.

5. Loop-erased walks

The concept of loop-erased walk (on an undirected graph) has been extensively used in the study of random walks, following the work of G. Lawler (see [15, Section 7] and references therein). Here is the definition, adapted, in the most straightforward way, to the case of oriented graphs.

Definition 5.1. Let π be a walk given by

$$(5.1) \qquad a_0 \xrightarrow{e_1} a_1 \xrightarrow{e_2} a_2 \xrightarrow{e_3} \cdots \xrightarrow{e_m} a_m \, .$$

The *loop-erased part* of π, denoted $\mathrm{LE}(\pi)$, is defined recursively as follows. If π does not have self-intersections (i.e., all vertices a_i are distinct), then $\mathrm{LE}(\pi) = \pi$. Otherwise, set $\mathrm{LE}(\pi) = \mathrm{LE}(\pi')$, where π' is obtained from π by removing the first loop it makes; more precisely, find $a_i = a_j$ with the smallest value of j, and remove the segment $a_i \xrightarrow{e_{i+1}} a_{i+1} \rightarrow \cdots \xrightarrow{e_j} a_j$ from π to obtain π'.

The loop erasure operator LE maps arbitrary walks to self-avoiding walks, i.e., walks without self-intersections. In the case of a directed network $\Gamma = (V, E, w)$, this operator projects the weight function on the set of walks, defined in Section 2 by $w(\pi) = \prod w(e_i)$, onto the corresponding function, denoted by $\tilde{w}$, on self-avoiding walks $\varkappa$:

$$\tilde{w}(\varkappa) = \sum_{\mathrm{LE}(\pi) = \varkappa} w(\pi) \,.$$

Although it will not be needed in the sequel, the author cannot resist a temptation to mention here a remarkable theorem, first explicitly stated in complete generality by R. Pemantle [18]. This theorem asserts, speaking loosely, that in the special case of a *simple random walk* on an undirected graph, the probability measure induced by loop erasure on self-avoiding walks from a vertex a to a vertex b coincides with the measure obtained by choosing, uniformly at random, a spanning tree of the underlying graph, and then selecting the unique path in the tree that connects a and b. See [18] for further details.

6. Minors of walk matrices

In this section, we prove the master theorem on minors of walk matrices (Theorem 6.1), state its simplified version applicable to walk matrices of planar directed networks (Theorem 6.4), and give a number of applications.

The notation and terminology introduced in Section 2 is kept throughout. Thus $\Gamma = (V, E, w)$ is a directed network with the walk matrix $W = (W(a, b))_{a,b \in V}$.

Similarly to Sections 3 and 4, let us choose a pair of totally ordered subsets $A = \{a_1, \ldots, a_k\} \subset V$ and $B = \{b_1, \ldots, b_k\} \subset V$, not necessarily disjoint, of the same cardinality k. Let us denote by $W_{A,B} = (W(a_i, b_j))_{\substack{i=1,\ldots,k \\ j=1,\ldots,k}}$ the corresponding $k \times k$ submatrix of the walk matrix W.

From the definition of the determinant, we have

$$(6.1) \qquad \det(W_{A,B}) = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \sum_{a_1 \xrightarrow{\pi_1} b_{\sigma(1)}, \, \ldots, \, a_k \xrightarrow{\pi_k} b_{\sigma(k)}} w(\pi_1) \cdots w(\pi_k) \,,$$

where the first sum is over all permutations $\sigma \in S_k$, and the second sum runs over all families of walks $\pi_1, \ldots, \pi_k$, which connect elements of A to the elements of B assigned to them by permutation σ.

Theorem 6.1. *The minors of the walk matrix W are given by the formula*

$$(6.2) \qquad \det(W_{A,B}) = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \sum_{i < j \implies \pi_j \cap \mathrm{LE}(\pi_i) = \emptyset} w(\pi_1) \cdots w(\pi_k) \,,$$

obtained by restricting the second summation in (6.1) *to the families of walks $\pi_1, \ldots, \pi_k$ satisfying the following condition: for any $1 \le i < j \le k$, the walk π_j has no common vertices with the loop-erased part of π_i.*

Before embarking on the proof of this theorem, let us pause for a couple of comments.

First, note that the left-hand side of (6.2) is invariant, up to a sign, under changes of total orderings of the sets A and B. It follows that the right-hand side possesses

the same kind of invariance, which is not at all obvious, since the condition imposed there on families of walks involves these orderings in a nontrivial way.[1]

Another symmetry of the left-hand side of (6.2) that does not manifest itself on the right-hand side is the invariance with respect to "time reversal." In other words, we can redirect all the edges backwards, while keeping their weights. The walk matrix is then transposed, so its minors remain the same, with the roles of A and B interchanged. This transformation of the network will replace ordinary loop erasure by "backwards loop erasure" (tracing a walk backwards, and erasing loops as they appear), and it is not immediately clear that this modification will leave the expression on the right-hand side of (6.2) invariant (but it will).

Proof. We will prove (6.2) by constructing a sign-reversing involution on the set of summands appearing on the right-hand side of (6.1), for which the condition $i < j \Longrightarrow \pi_j \cap \mathrm{LE}(\pi_i) = \emptyset$ is violated. The argument will be first presented for the special case $k = 2$, and then extended to a general k.

The following general construction will be useful in the proof. Let $a \xrightarrow{\pi} b$ be an arbitrary walk, and let v be a vertex lying on its loop-erased part $\mathrm{LE}(\pi)$. The edge sequence of $\mathrm{LE}(\pi)$ is canonically a subsequence of the edge sequence of π; in the latter sequence, let us identify the unique entry of the form $v' \xrightarrow{e} v$ which is contained in the loop-erased subsequence. (The case $v = a$ is an exception, to be kept in mind.) The walk π is now partitioned at the end of the entry e into two walks $a \xrightarrow{\pi'(v)} v \xrightarrow{\pi''(v)} b$ such that

(a) the last entry in the edge sequence of $\pi'(v)$ contributes to $\mathrm{LE}(\pi'(v))$;

(b) $\pi''(v)$ does not visit any vertices which lie on $\mathrm{LE}(\pi'(v))$, except for v.

The conditions (a)–(b) uniquely determine the partition $(\pi'(v), \pi''(v))$ of the walk π; in the special case $v = a$, we set $\pi'(v)$ to be the trivial path $v \to v$, and let $\pi''(v) = \pi$.

Let us get back to the proof. Assume $k = 2$, and let the walks $a_1 \xrightarrow{\pi_1} b_{\sigma(1)}$ and $a_2 \xrightarrow{\pi_2} b_{\sigma(2)}$ be such that π_2 and $\mathrm{LE}(\pi_1)$ pass through at least one common vertex. Among all such vertices, choose the one (call it v) which is closest to a_1 along the self-avoiding walk $\mathrm{LE}(\pi_1)$. We then partition π_1 into $a_1 \xrightarrow{\pi_1'(v)} v \xrightarrow{\pi_1''(v)} b_{\sigma(1)}$, following the rules above (see (a)–(b)). Let us denote $L = \mathrm{LE}(\pi_1'(v))$. With this notation, condition (b) can be restated as follows:

(b1) $\pi_1''(v)$ does not visit any vertices which lie on L, except for v.

Let us now split π_2 at the point of its first visit to v. More formally, we define the partition $a_2 \xrightarrow{\pi_2'} v \xrightarrow{\pi_2''} b_{\sigma(2)}$ of π_2 by requiring that π_2' does visit v before arriving at its endpoint. By the choice of v,

(b2) π_2'' does not visit any vertices which lie on L, except for v;

(c) π_2' does not visit any vertices which lie on L, except for ending at v.

Everything is now ready for the path-switching argument. Let us create new walks

$$\tilde{\pi}_1 \ : \ a_1 \xrightarrow{\pi_1'(v)} v \xrightarrow{\pi_2''} b_{\sigma(2)}$$

and

$$\tilde{\pi}_2 \ : \ a_2 \xrightarrow{\pi_2'} v \xrightarrow{\pi_1''} b_{\sigma(1)} \, .$$

---

[1] *Note added in revision.* This symmetry can be explained using D. B. Wilson's algorithm [22, 19] for constructing random spanning trees. This observation was related to me by several experts, including Y. Peres and O. Schramm, who based it on the setup of [1].

The map $(\pi_1, \pi_2) \mapsto (\tilde{\pi}_1, \tilde{\pi}_2)$ is the desired sign-reversing involution. The basic reason for this is the similarity of conditions (b1) and (b2), which allows us to interchange the portions $\pi_1''(v)$ and π_2''. It furthermore ensures that the new walk $\tilde{\pi}_1$ splits into $\tilde{\pi}_1'(v) = \pi_1'(v)$ and $\tilde{\pi}_1''(v) = \pi_2''$. In particular, the loop-erased part of the initial segment of π_1 remains invariant: $\mathrm{LE}(\tilde{\pi}_1'(v)) = L$. Now (b1) and (c) show that $\tilde{\pi}_2$ splits at v, as needed. As a result, applying the same procedure to $(\tilde{\pi}_1, \tilde{\pi}_2)$ recovers (π_1, π_2), and we are done.

The case of an arbitrary k is proved by the same argument combined with a careful choice of the pair of paths to which it is applied. Take a term on the right-hand side of (6.1) which corresponds to a k-tuple of walks $(\pi_1, \ldots, \pi_k)$. Assume that this term does not appear in (6.2). Thus the set of triples (i, j, v), $1 \leq i < j \leq k$, $v \in V$, such that the walks π_j and $\mathrm{LE}(\pi_i)$ pass through v is not empty. Among all such triples, let us choose lexicographically minimal, in the following order of priority:

- take the smallest possible value of i;
- for this i, let v to be as close as possible to a_i along $\mathrm{LE}(\pi_i)$, among all intersections with π_j, for all $j > i$;
- for these i and v, find the smallest $j > i$ such that π_j hits v.

We then proceed exactly as before, working with the pair of walks (π_i, π_j). Thus we partition the walk π_i into $a_i \xrightarrow{\pi_i'(v)} v \xrightarrow{\pi_i''(v)} b_{\sigma(i)}$, as prescribed by (a)–(b) above. Then locate

- the first visit of π_j to v, starting at a_j

and split π_j accordingly into $a_j \xrightarrow{\pi_j'} v \xrightarrow{\pi_j''} b_{\sigma(j)}$. Exchanging the portions $\pi_i''(v)$ and π_j'' of the walks π_i and π_j provides the desired sign-reversing involution. A straightforward verification is omitted. $\qquad\square$

Example 6.2. Consider the directed network in Example 2.1. Let $A = \{b, c\}$ and $B = \{a, b\}$. Then

$$(6.3) \qquad \det(W_{A,B}) = (1 - q_1 q_2 q_3)^{-2} \begin{Vmatrix} q_2 q_3 & 1 \\ q_3 & q_1 q_3 \end{Vmatrix} = -q_3(1 - q_1 q_2 q_3)^{-1}.$$

On the other hand, the only way a pair of walks π_1, π_2 connecting A and B can satisfy the conditions of Theorem 6.1 is the following: take any closed walk $b \xrightarrow{\pi_1} b$ (so the loop-erased part of π_1 will be the trivial walk $b \to b$ of length 0), and set π_2 to be the walk $c \to a$ of length 1. The corresponding permutation σ will be a transposition, so $\mathrm{sgn}(\sigma) = -1$. The weight of π_2 will be equal to q_3, and the generating function for the weight of π_1 will be $(1 - q_1 q_2 q_3)^{-1}$. So the right-hand side of (6.2) will be equal to $-q_3(1 - q_1 q_2 q_3)^{-1}$, in compliance with (6.3).

Example 6.3 (Bernoulli random walk). Let us now look at Example 2.3. The results we are going to obtain are not new, but the proofs illustrate quite well how Theorem 6.1 works in one-dimensional applications.

First, let $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$ be such that $a_1 < b_1 < a_2 < b_2$ and $a_2 - b_1 = k$, as shown below:
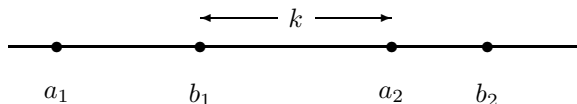
By formula (2.5), the left-hand side of (6.2) is

$$
\det(W_{A,B}) = \left\| \begin{array}{cc} (p-q)^{-1} & (p-q)^{-1} \\ (p-q)^{-1}q^k p^{-k} & (p-q)^{-1} \end{array} \right\| = (p-q)^{-2}(1 - q^k p^{-k}).
$$

Let us now look at the right-hand side. The only possible σ is the identity permutation. (It might seem that a mistake is being made here, since the scenario $a_1 \rightsquigarrow b_2$, $a_2 \rightsquigarrow b_1$ may occur without direct collision: the two "particles" can easily cross each other's ways without ever being at the same point at the same time. However, we only need to make sure that the *trajectories* intersect—in space, but not necessarily in time! This distinguishes the loop-erased switching from the conventional Karlin-McGregor argument.)

Any walk $a_1 \xrightarrow{\pi_1} b_1$ has the same loop-erased part, namely the shortest path from a_1 to b_1. To avoid intersecting that path, a walk $a_2 \xrightarrow{\pi_2} b_2$ must never hit b_1. The sum of weights of all such walks can be interpreted as

$$
E(a_2, b_2; b_1) = \left\langle \begin{array}{c} \text{expected number of visits that a random walk originating} \\ \text{at } a_2 \text{ makes to } b_2 \text{ before hitting } b_1 \text{ for the first time} \end{array} \right\rangle
$$

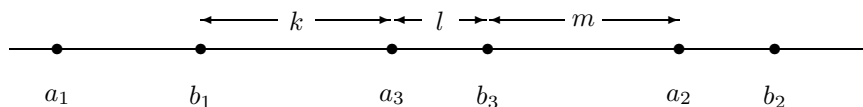(if a trajectory never hits b_1, then all visits to b_2 count toward $E(a_2, b_2; b_1)$). Thus

$$
\det(W_{A,B}) = (p-q)^{-2}(1 - q^k p^{-k}) = W(a_1, b_1)E(a_2, b_2; b_1) = (p-q)^{-1}E(a_2, b_2; b_1),
$$

and we conclude that $E(a_2, b_2; b_1) = (p-q)^{-1}(1 - q^k p^{-k})$ (assuming $b_1 < a_2 < b_2$ and $a_2 - b_1 = k$).

Continuing with this example, let us now swap a_2 and b_2, as shown below:
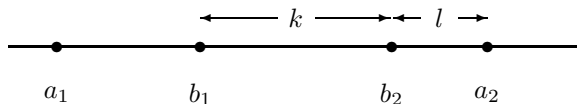
Analogous computations give $E(a_2, b_2; b_1) = (p-q)^{-1}q^l p^{-l}(1 - q^k p^{-k})$ for $b_1 < b_2 < a_2$ and $b_2 - b_1 = k$, $a_2 - b_2 = l$. Finally, let us add another pair of points, and change the labelling once again:

Then

$$
\det(W_{A,B}) = (p-q)^{-3} \left\| \begin{array}{ccc} 1 & 1 & 1 \\ (\frac{q}{p})^{k+l+m} & 1 & (\frac{q}{p})^m \\ (\frac{q}{p})^k & 1 & 1 \end{array} \right\| = (p-q)^{-3}(1-(\tfrac{q}{p})^k)(1-(\tfrac{q}{p})^m).
$$

By Theorem 6.1,

(6.4) $\det(W_{A,B}) = W(a_1,b_1)E(a_2,b_2;b_1)E(a_3,b_3;b_1,a_2),$

where $E(a_2,b_2;b_1)$ has the same meaning as before, and

$$
E(a_3,b_3;b_1,a_2) = \left\langle \begin{array}{l} \text{expected number of visits that a random walk which} \\ \text{originates at } a_3 \text{ makes to } b_3 \text{ before hitting either } b_1 \text{ or } a_2 \end{array} \right\rangle.
$$

Substituting $W(a_1,b_1) = (p-q)^{-1}$ and $E(a_2,b_2;b_1) = (p-q)^{-1}(1-(\tfrac{q}{p})^{k+l+m})$ into (6.4), we conclude that

$$
E(a_3,b_3;b_1,a_2) = \frac{(1-(\tfrac{q}{p})^k)(1-(\tfrac{q}{p})^m)}{(p-q)(1-(\tfrac{q}{p})^{k+l+m})}.
$$

Most interesting applications of Theorem 6.1 arise when the network Γ is embedded into a *planar* domain Ω, and the points $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ are chosen on the topological boundary of Ω in such a way that the only allowable permutation σ is the identity permutation; cf. Figures 1 and 4.
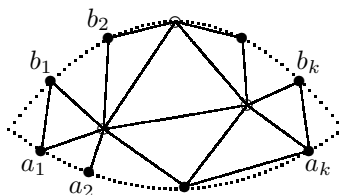

FIGURE 4. Planar network

The following corollary of Theorem 6.1 is useful in such situations,

Theorem 6.4. *Assume that the vertices $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ are chosen so that any walk from a_i to b_j intersects any walk from $a_{i'}$, $i' > i$, to $b_{j'}$, $j' < j$. Then the corresponding minor of the walk matrix is given by*

(6.5) $\det(W_{A,B}) = \sum w(\pi_1) \cdots w(\pi_k),$

where the sum runs over all families of walks $a_1 \xrightarrow{\pi_1} b_1, \ldots, a_k \xrightarrow{\pi_k} b_k$ such that for any $1 \le i < k$, the walk π_{i+1} has no common vertices with the loop-erased part of π_i. In particular, if all edge weights are nonnegative, then the matrix $W_{A,B}$ is totally nonnegative.

Proof. Consider the right-hand side of (6.2). First we note that the loop-erased walks $\text{LE}(\pi_i)$ are pairwise vertex-disjoint, and therefore σ must be the identity permutation, under the assumptions of Theorem 6.4.

It remains to show that the restrictions on the π_i may be relaxed as indicated. Assume that the walks $a_1 \xrightarrow{\pi_1} b_1, \ldots, a_k \xrightarrow{\pi_k} b_k$ satisfy the condition in Theorem 6.4. Suppose that the corresponding condition in Theorem 6.1 is violated, that

is, some walk π_j intersects $\mathrm{LE}(\pi_i)$, for some $i \leq j - 2$. We may assume that among all such violations, this one has the smallest value of $j - i$, which in particular means that π_{j-1} does not intersect $\mathrm{LE}(\pi_i)$. We can then combine segments of π_j and $\mathrm{LE}(\pi_i)$ to create a walk $a_i \xrightarrow{\pi_{ij}} b_j$. By assumptions of Theorem 6.4, π_{ij} must intersect the walk $a_{j-1} \xrightarrow{\mathrm{LE}(\pi_{j-1})} b_{j-1}$. On the other hand, neither $\mathrm{LE}(\pi_i)$ nor π_j intersects $\mathrm{LE}(\pi_{j-1})$, a contradiction. $\square$

Theorem 6.4 can be applied to the study of two-dimensional random walk. Let us briefly discuss two possible venues.

Example 6.5 (Random walk on a planar strip). Consider a Markov chain on the state space $V = \mathbb{Z} \times \{0, 1, \dots, N\}$ (see Figure 5a) satisfying the following conditions:

(6.6)
    (a) all allowable transitions are of the form $(x, y) \to (x \pm 1, y \pm 1)$;

    (b) the transition probabilities are translation-invariant along the x axis.

Let $\partial\Gamma_0 = \{(i, 0) : i \in \mathbb{Z}\}$ and $\partial\Gamma_1 = \{(j, N) : j \in \mathbb{Z}\}$ denote the lower and the upper boundaries of the strip. Choose two points $a = (i, 0) \in \partial\Gamma_0$ and $b = (i + k, N) \in \partial\Gamma_1$. Let us denote by $T(k) = W(a, b)$ the expected number of times the process passes through b provided it started at a. (To simplify matters, let us assume that $E(k) < +\infty$.) Then the infinite Toepliz matrix $T = (T(j - i))_{i,j \in \mathbb{Z}}$ is the same as the submatrix $(W(a, b))_{a \in \partial\Gamma_0, b \in \partial\Gamma_1}$ of the walk matrix W. Now Theorem 6.4 shows that T must be totally nonnegative. Furthermore, it provides combinatorial (or probabilistic) interpretations for the minors of T. For example, for any positive integers l and m, the 2×2 determinant $T(k)T(k + m - l) - T(k + m)T(k - l)$ is equal to the expected number of times a trajectory originating at $(0, l)$ hits the point $(k + m, N)$ before hitting (k, N) for the first time.

The same conclusions will of course hold if the transition probabilities/weights are not translation-invariant, except that the matrix in question will generally not be Toeplitz. It is natural to ask whether such matrices T can be obtained using some modification of the construction presented in Section 3. At present, we do not have a good answer to this question.
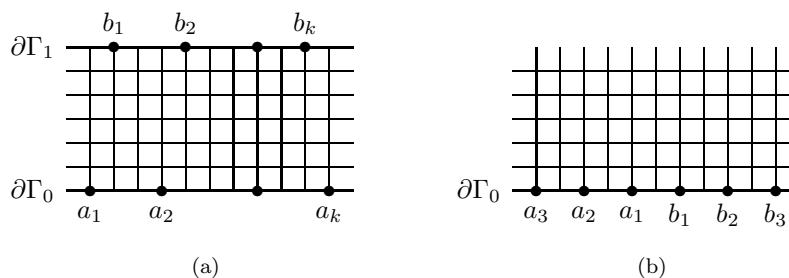

FIGURE 5. Two-dimensional random walks

Example 6.6 (Random walk on a half-plane). Let us consider random walk on the half-plane $\mathbb{Z} \times \{0, 1, \dots, N\}$ (see Figure 5b). As before, we assume conditions (6.6), and denote $\partial\Gamma_0 = \{(i, 0) : i \in \mathbb{Z}\}$.

Choose two points $a = (i, 0) \in \partial\Gamma_0$ and $b = (i + k, 0) \in \partial\Gamma_1$, and let $H(k) = W(a, b)$. Then the infinite Hankel matrix $T = (T(i + j))_{i,j \in \mathbb{Z}}$ will be totally nonnegative, for reasons similar to Example 6.5. Furthermore, it will be totally positive

provided all transition probabilities between neighbouring states are positive. Details are left to the reader.

7. MINORS OF HITTING MATRICES

The main goal of this section is to produce analogues of Theorems 6.1 and 6.4 for hitting matrices. As before in Section 2, we assume that the vertex set V is arbitrarily partitioned into two disjoint subsets: the boundary $\partial\Gamma$ and the interior $int\Gamma$. Choose a pair of totally ordered subsets $A = \{a_1, \dots, a_k\} \subset V$ and $B = \{b_1, \dots, b_k\} \subset \partial\Gamma$, and denote by $X_{A,B} = (X(a_i, b_j))_{\substack{i=1,\dots,k \\ j=1,\dots,k}}$ the corresponding hitting submatrix.

Recall that $\mathrm{LE}(\pi)$ denotes the loop-erased part of a walk π.

Theorem 7.1. *The minors of the hitting matrix are given by*

$$(7.1) \qquad \det(X_{A,B}) = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \sum_{a_1 \xrightarrow{\pi_1} b_{\sigma(1)}, \, \dots, \, a_k \xrightarrow{\pi_k} b_{\sigma(k)}} w(\pi_1) \cdots w(\pi_k) \,,$$

where the first sum is over all permutations $\sigma \in S_k$, and the second sum runs over all families of walks $\pi_1, \dots, \pi_k$ satisfying the following conditions:

- *π_i has nonzero length, begins at a_i, ends at $b_{\sigma(i)}$, and in the meantime does not pass through any boundary vertices;*
- *for any $1 \le i < j \le k$, the walks π_j and $\mathrm{LE}(\pi_i)$ have no common vertices in the interior of Γ.*

Note that Theorem 7.1 reduces to Theorem 3.1 if the network is acyclic.

Comparing to formula (4.4) (which has a narrower domain of applicability), Theorem 7.1 has the advantage of "polynomiality:" the right-hand side of (7.1), unlike that of (4.4), is manifestly a formal power series in the edge weights. This feature will be essential while extending the result to the continuous case.

Proof. This theorem can be proved by a direct argument similar to the one used in the proof of Theorem 6.1. To save an effort, we will instead use a simple observation that will reduce Theorem 7.1 to Theorem 6.1.

Let us define a new network Γ' by splitting every boundary vertex $b \in \partial\Gamma$ into a source b' and a sink b'', converting all outgoing edges $b \to c$ into $b' \to c$, redirecting all incoming edges $a \to b$ into $a \to b''$, and keeping the edge weights intact. Let $\partial\Gamma'$ and $\partial\Gamma''$ be the sets of sources and sinks, respectively. Then the hitting matrix X of the original network Γ becomes a submatrix of the walk matrix W' of the transformed network Γ'; more precisely, $X(a, b) = W'(a', b'')$ for $a, b \in \partial\Gamma$, while $X(a, b) = W'(a, b'')$ for $a \in V \setminus \partial\Gamma$ and $b \in \partial\Gamma$. To complete the proof, it remains to carefully reformulate the statement of Theorem 6.1 for Γ' (with each $a_i \in \partial\Gamma$ replaced by a_i' and each b_j replaced by b_j'') in terms of Γ. $\square$

The following analogue of Theorem 6.4 is a corollary of Theorem 7.1.

Theorem 7.2. *Assume that the vertices $a_1, \dots, a_k \in V$ and $b_1, \dots, b_k \in \partial\Gamma$ have the property that any walk from a_i to b_j through the interior of Γ intersects any such walk from $a_{i'}$, $i' > i$, to $b_{j'}$, $j' < j$, at a point in the interior. Then*

$$(7.2) \qquad \det(X_{A,B}) = \sum w(\pi_1) \cdots w(\pi_k) \,,$$

where the sum runs over all families of walks $\pi_1, \dots, \pi_k$ satisfying the following conditions:

- π_i has nonzero length, begins at a_i, ends at b_i, and in the meantime does not pass through any boundary vertices;
- for any $1 \leq i < k$, the walk π_{i+1} has no common vertices with the loop-erased part of π_i in the interior of Γ.

In particular, if the edge weights are nonnegative, then the matrix $X_{A,B}$ is totally nonnegative.

The proof of Theorem 7.2 is similar to that of Theorem 6.4, and is omitted.

Example 7.3. In Example 2.4, let us verify Theorem 7.2. We have

$$\det(X_{A,B}) = (1 - q_1 q_2 q_3)^{-2} \left\| \begin{array}{cc} q_4 q_5 & q_1 q_3 q_4 q_7 \\ q_1 q_2 q_5 q_6 & q_1 q_6 q_7 \end{array} \right\| = (1 - q_1 q_2 q_3)^{-1} q_1 q_4 q_5 q_6 q_7 .$$

On the other hand, the right-hand side of (7.2) is the product of $X(a_1, b_1) = (1 - q_1 q_2 q_3)^{-1} q_4 q_5$ and the weight of the only walk $a_2 \xrightarrow{\pi_2} b_2$ that does not intersect the only self-avoiding walk $a_1 \xrightarrow{\pi_1} b_1$. This weight being equal to $q_1 q_6 q_7$, Theorem 7.2 checks.

Hitting matrices of Markov chains. If the weights of a directed network are transition probabilities (cf. Example 2.5), Theorem 7.2 has the following probabilistic interpretation.

As before, consider a Markov chain on the state space V with "boundary" $\partial\Gamma \subset V$. The entries of the hitting matrix X are the hitting probabilities: $X(a, b)$ is the probability that b is the first boundary state hit by the process that originates at a. (Here and below in Theorem 7.4, a "hit" must occur *after* the clock is started. In other words, if the process originates at a boundary state, it is not presumed to hit the boundary right away.)

Theorem 7.4. *Suppose that totally ordered subsets $A = \{a_1, \ldots, a_k\} \subset V$ and $B = \{b_1, \ldots, b_k\} \subset \partial\Gamma$ are such that any possible trajectory connecting a_i to b_j through $\mathrm{int}\Gamma$ intersects any trajectory connecting $a_{i'}$, $i' > i$, to $b_{j'}$, $j' < j$, at a point in $\mathrm{int}\Gamma$. Then the minor $\det(X_{A,B})$ of the hitting matrix X is equal to the probability that k independent trajectories $\pi_1, \ldots, \pi_k$ originating at $a_1, \ldots, a_k$, respectively, will hit the boundary $\partial\Gamma$ for the first time at the points $b_1, \ldots, b_k$, respectively, and furthermore, the trajectory π_{i+1} will have no common vertices with the loop-erased part of π_i in the interior of Γ, for $i = 1, \ldots, k - 1$.*

In particular, $\det(X_{A,B}) \geq 0$, and moreover the matrix $X_{A,B}$ is totally nonnegative.

To illustrate, the hitting matrices for two-dimensional random walks discussed in Examples 6.5 and 6.6 are totally nonnegative.

As an application of Theorem 7.4, we obtain another proof of total nonnegativity of response matrices of resistor networks (see Corollary 4.6).

Generalized hitting matrices. For $a \in V$ and $B \subset \partial\Gamma$, let $X(a, B)$ denote the corresponding hitting probability; that is, $X(a, B)$ is the probability that the process originating at a will first hit the boundary at some point $b \in B$. Using multilinearity of the determinant, we obtain the following modification of Theorem 7.4. (Other theorems above can also be given similar analogues.)

Theorem 7.5. *Assume that distinct states $a_1, \ldots, a_k \in V$ and disjoint subsets $B_1, \ldots, B_k \subset \partial\Gamma$ are such that any possible trajectory connecting a_i to B_j intersects*

any trajectory connecting $a_{i'}$, $i' > i$, to $B_{j'}$, $j' < j$, at a point in $\text{int}\Gamma$. Then the determinant $\det(X(a_i, B_j))$ is nonnegative, and moreover the matrix $(X(a_i, B_j))$ is totally nonnegative.

Moreover, $\det(X(a_i, B_j))$ is equal to the probability that k independent trajectories $\pi_1, \ldots, \pi_k$ originating at $a_1, \ldots, a_k$, respectively, will hit the boundary $\partial\Gamma$ for the first time at points which belong to $B_1, \ldots, B_k$, respectively, and furthermore, the trajectory π_{i+1} will have no common vertices with the loop-erased part of π_i in the interior of Γ, for $i = 1, \ldots, k-1$.

In order for the statement of Theorem 7.5 to make sense, the Markov process under consideration does not have to be discrete. Although the technical details remain to be worked out, it is clear that by passing to a limit in a discrete approximation, Theorem 7.5 can be generalized to (non-isotropic) Brownian motions on planar domains, or arbitrary simply connected Riemann manifolds with boundary, subject to technical assumptions. (See Figure 6.) The same type of argument can be used to justify well-definedness of the quantities involved; notice that in order to define a continuous analogue of the probability appearing in Theorem 7.5, we do not need the notion of loop-erased Brownian motion. Instead, we discretize the model, compute the probability, and then pass to a limit. One can further extend these results to densities of the corresponding hitting distributions.

The rest of this section is devoted to a couple of characteristic applications involving Brownian motion.
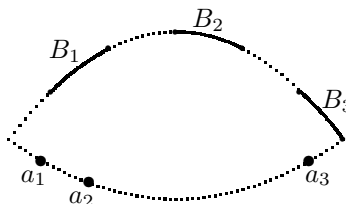

FIGURE 6. Theorem 7.5

Brownian motion in the quadrant with reflecting side. Let

$$\Omega = \{(x,y) \,:\, x \geq 0, y \geq 0\},$$

and consider the Brownian motion in Ω which reflects in (bounces off) [the positive part of] the x axis (see Figure 7a). The y half-axis is absorbing, i.e., the process stops once it reaches a state $(0, y)$, $y \geq 0$. This is equivalent to taking a Brownian motion in the half-plane $x \geq 0$ that stops upon hitting the y axis, and reflecting the portions of it that go into the lower quadrant $\{(x, y) \,:\, x \geq 0, y < 0\}$ over the x axis.

Let the process start at time $t = 0$ at a point $(x, 0)$, $x > 0$ on the x axis. It is classically known (see, e.g., [9, Section 1.9] for two different proofs) that the Brownian motion in half-plane hits the y axis for the first time at a point $Y(x)$ that has the Cauchy distribution with appropriate parameter. Taking absolute values, we conclude that our process in Ω induces the hitting distribution on the half-line $\{(0, y) \,:\, y \geq 0\}$ which has the density

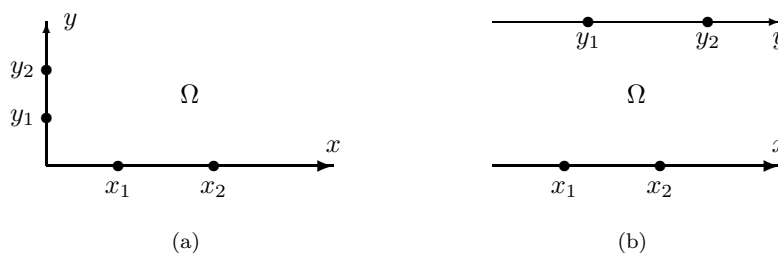$$K(x, y) = \frac{2x}{\pi(x^2 + y^2)}, \quad y \geq 0.$$

FIGURE 7. Brownian motion in domains with reflecting sides

It follows from the general results above that the kernel $K(x, y)$ is totally positive, i.e., any determinant of the form $\det(K(x_i, y_j))$ with $0 < x_1 < x_2 < \cdots$ and $0 < y_1 < y_2 < \cdots$ is positive. This result is not new. What seems to be new is the interpretation we give to such determinants in terms of loop-erased walks. To illustrate, let us fix $0 < x_1 < x_2$, and consider two independent trajectories π_1 and π_2 which start at points $(x_1, 0)$ and $(x_2, 0)$ and end up hitting the y axis at points $(0, Y(x_1))$ and $(0, Y(x_2))$, respectively. Assume that these two points are the points $(0, y_1)$ and $(0, y_2)$ with $0 < y_1 < y_2$, but we do not know which trajectory hit which of the two points. Then the conditional probability of the scenario $Y(x_1) = y_1$, $Y(x_2) = y_2$ is larger than that of $Y(x_1) = y_2$, $Y(x_2) = y_1$, and the difference of these conditional probabilities, which is given by

$$\frac{K(x_1, y_1)K(x_2, y_2) - K(x_1, y_2)K(x_2, y_1)}{K(x_1, y_1)K(x_2, y_2) + K(x_1, y_2)K(x_2, y_1)},$$

is equal to the conditional probability that π_2 *did not hit the loop-erased part of* π_1. (We repeat once again that the well-definedness of the loop-erased Brownian motion does not have to be justified in order for such probabilities to make perfect sense.) Computations show that

$$\left\| \begin{matrix} K(x_1, y_1) & K(x_1, y_2) \\ K(x_2, y_1) & K(x_2, y_2) \end{matrix} \right\| = \frac{4x_1 x_2 (x_2^2 - x_1^2)(y_2^2 - y_1^2)}{\pi^2 \prod_{i=1}^{2} \prod_{j=1}^{2} (x_i^2 + y_j^2)},$$

and therefore the conditional probability in question is given by the surprisingly simple formula

$$(7.3) \qquad P(\pi_1 \cap \mathrm{LE}(\pi_2) = \emptyset \,|\, \{Y(x_1), Y(x_2)\} = \{y_1, y_2\}) = \frac{(x_2^2 - x_1^2)(y_2^2 - y_1^2)}{(x_1^2 + y_2^2)(x_2^2 + y_1^2)}.$$

If we do not condition on the locations of the hitting points (still keeping the initial locations $x_1 < x_2$ fixed), then the probability that the trajectory π_2 originating at x_2 will not intersect the loop-erased part of the trajectory π_1 that starts at x_1 (or a similar quantity with π_1 and π_2 interchanged) is equal to

$$P(Y(x_1) < Y(x_2)) - P(Y(x_1) \geq Y(x_2)) = \int_0^\infty \int_{y_1}^\infty \left\| \begin{matrix} K(x_1, y_1) & K(x_1, y_2) \\ K(x_2, y_1) & K(x_2, y_2) \end{matrix} \right\| dy_2 dy_1.$$

Computing the integral yields

$$P(\pi_1 \cap \mathrm{LE}(\pi_2) = \emptyset) = -\frac{4}{\pi^2} \left(\mathrm{Li}_2(-\alpha) + \mathrm{Li}_2(1 - \alpha) + \ln(\alpha)\ln(1 + \alpha) + \frac{\pi^2}{12} \right),$$

where $\alpha = \dfrac{x_2}{x_1}$ and $\mathrm{Li}_2(t) = \displaystyle\sum_{n=1}^{\infty} \dfrac{t^n}{n^2}$ is the dilogarithm function. In particular, if $\alpha = \frac{\sqrt{5}+1}{2}$ (so the initial locations x_1 and x_2 are in golden ratio), then we obtain, using [4], the formula $P(\pi_1 \cap \mathrm{LE}(\pi_2) = \emptyset) = \frac{1}{3} - \frac{6}{\pi^2}\left(\ln(\frac{\sqrt{5}+1}{2})\right)^2$.

Brownian motion in a strip with reflecting side. In this example, we let

$$\Omega = \{(x,y) \, : \, x \in \mathbb{R}, 0 \le y \le 1\}$$

and consider the Brownian motion in Ω which is reflecting in the x axis (see Figure 7b). The process begins at a point $(x_0, 0)$ on the x axis and stops when it hits the line $y = 1$.

Standard computations (employing either conformal invariance of Brownian motion or the reflection principle) yield the well-known formula for the hitting density:

$$K(x_0, x_0 + x) = \frac{1}{e^{\frac{\pi x}{2}} + e^{-\frac{\pi x}{2}}} = \frac{1}{2\cosh(\frac{\pi x}{2})} \, .$$

Thus K is a totally positive translation-invariant kernel; equivalently, the function $K(x) = \dfrac{1}{2\cosh(\frac{\pi x}{2})}$ is a *Pólya frequency function.* This particular example of a PF function is of course well known; see, e.g., Karlin [13, § 7.1, example (b)]. The computation of probabilities related to loop-erased walks in this model is omitted.

References

1. I. Benjamini, R. Lyons, Y. Peres, and O. Schramm, Uniform spanning forests, *Ann. Probab.,* to appear.
2. W. Böhm and S. G. Mohanty, On the Karlin-McGregor theorem and applications, *Ann. Appl. Probab.* **7** (1997), 314–325. MR **98g:**60131
3. B. Bollobás, *Modern graph theory,* Springer-Verlag, New York, 1998. MR **99h:**05001
4. J. M. Borwein, D. M. Bradley, D. J. Broadhurst, and P. Lisoněk, Special values of multiple polylogarithms, *Trans. Amer. Math. Soc.,* to appear.
5. Y. Colin de Verdière, Réseaux électriques planaires, *Prépublications de l'Institut Fourier* **225** (1992), 1–20.
6. E. Curtis, D. V. Ingerman, and J. Morrow. Circular planar graphs and resistor networks, *Linear Algebra Appl.* **283** (1998), 115–150. MR **99k:**05096
7. E. Curtis, E. Mooers, and J. Morrow, Finding the conductors in circular networks from boundary measurements, *RAIRO Modél. Math. Anal. Numér.* **28** (1994), 781–814. MR **96i:**65110
8. P. G. Doyle and J. L. Snell, *Random walks and electric networks,* Math. Assoc. of America, 1984. MR **89a:**94023
9. R. Durrett, *Brownian motion and martingales in analysis,* Wadsworth, 1984. MR **87a:**60054
10. S. Fomin and A. Zelevinsky, Total positivity: tests and parametrizations, *Math. Intelligencer* **22** (2000), no. 1, 23–33. MR **2001b:**22014
11. F. R. Gantmacher and M. G. Krein, *Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme*, Akademie-Verlag, Berlin, 1960. MR **22:**5161 (Russian edition: Moscow-Leningrad, 1950. MR **14:**178)
12. D. V. Ingerman, Discrete and continuous inverse boundary problems on a disk, Ph.D. thesis, University of Washington, 1997.
13. S. Karlin, *Total positivity*, Stanford University Press, 1968. MR **37:**5667
14. S. Karlin and G. McGregor, Coincidence probabilities, *Pacific J. Math.* **9** (1959), 1141–1164. MR **22:**5072
15. G. Lawler, *Intersections of random walks,* Birkhäuser, 1991. MR **92f:**60122
16. B. Lindström, On the vector representations of induced matroids, *Bull. London Math. Soc.* **5** (1973), 85-90. MR **49:**95
17. A. A. Milne, *The house at Pooh corner,* E. P. Dutton & Co., 1928 ff.

18. R. Pemantle, Choosing a spanning tree for the integer lattice uniformly, *Ann. Probab.* **19** (1991), 1559–1574. MR **92g:**60014
19. J. G. Propp and D. B. Wilson, How to Get a Perfectly Random Sample from a Generic Markov Chain and Generate a Random Spanning Tree of a Directed Graph, *J. Algorithms* **27** (1998), 170–217. MR **99g:**60116
20. F. Spitzer, *Principles of random walk,* Van Nostrand, 1964. MR **30:**1521
21. R. P. Stanley, *Enumerative combinatorics, vol. 1,* 2d ed., Cambridge Univ. Press, 1997. MR **98a:**05001
22. D B. Wilson, Generating random spanning trees more quickly than the cover time, *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing,* 296–303, 1996. MR **97g:**68005

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109
E-mail address: fomin@math.lsa.umich.edu