

RANDOM WALK AND HYPERPLANE ARRANGEMENTS

KENNETH S. BROWN
Department of Mathematics
Cornell University
Ithaca, NY 14853

PERSI DIACONIS
Departments of Mathematics and ORIE
Cornell University
Ithaca, NY 14853

ABSTRACT. Let \mathcal{C} be the set of chambers of a real hyperplane arrangement. We study a random walk on \mathcal{C} introduced by Bidigare, Hanlon, and Rockmore. This includes various shuffling schemes used in computer science, biology, and card games. It also includes random walks on zonotopes and zonotopal tilings. We find the stationary distributions of these Markov chains, give good bounds on the rate of convergence to stationarity, and prove that the transition matrices are diagonalizable. The results are extended to oriented matroids.

1. INTRODUCTION

Let \mathcal{A} be a finite set of affine hyperplanes in $V = \mathbb{R}^n$. Then \mathcal{A} cuts V into regions called *chambers*. For example, there are 6 chambers (which are sectors) in Figure 1a, and there are 7 chambers in Figure 1b. The chambers are polyhedra (finite intersections of half-spaces) and hence have faces. For example, each chamber C in Figure 1a has 4 faces: C itself, 2 rays, and one point. We denote by \mathcal{F} the collection of all faces of the chambers. In Figure 1a, for example, \mathcal{F} has 13 elements: 6 chambers, 6 rays, and one point. The arrangement \mathcal{A} is called *central* if $\bigcap_{H \in \mathcal{A}} H \neq \emptyset$, as in Figure 1a.

\mathcal{F} admits a semigroup structure, whose definition will be recalled in Section 2 below. Of particular importance is the product FC for $F \in \mathcal{F}$ and $C \in \mathcal{C}$. This product is again a chamber, called the *projection* of C on F . It can be characterized as the nearest chamber to C having F as a face. Here “nearest” is defined in terms of the number of hyperplanes in \mathcal{A} separating C from FC . The projection operator $C \mapsto FC$ will be called the *action* of F on \mathcal{C} . See Figure 2 for a simple example (here F is a ray).

Figure 1a. A central arrangement of 3 hyperplanes in \mathbb{R}^2 .

Figure 1b. A non-central arrangement of 3 hyperplanes in \mathbb{R}^2 .

Figure 2. The projection of C on F .

Bidigare, Hanlon, and Rockmore [6], referred to hereafter as BHR, used the action of faces on chambers to define a random walk on \mathcal{C} .¹ Start with a probability measure w on \mathcal{F} . Then a step in the walk is given by:

(1.1) From $C \in \mathcal{C}$, choose F from the measure w and move to FC .

Thus the random walk started at a chamber C_0 is the process $(C_\ell)_{\ell \geq 0}$ with

$$(1.2) \quad C_\ell = F_\ell \cdots F_1 C_0,$$

where F_1, F_2, \dots are i.i.d. picks from w . This is simply random walk on the semi-group \mathcal{F} in the usual sense, with the starting state (and hence all future states) in the ideal $\mathcal{C} \subseteq \mathcal{F}$.

One can also describe the walk on \mathcal{C} by giving its transition matrix K :

$$(1.3) \quad K(C, C') = \sum_{FC = C'} w(F).$$

Remarkably, BHR found all the eigenvalues of K , which turn out to be real, non-negative, and linear in the entries of K . The multiplicities of the eigenvalues are given in terms of the Möbius function of the *intersection poset* \mathcal{S} (also called the *intersection lattice* in the central case). This is the set of all nonempty affine subspaces $W \subseteq V$ of the form $W = \bigcap_{H \in \mathcal{A}'} H$, where $\mathcal{A}' \subseteq \mathcal{A}$ is an arbitrary subset (possibly empty). We order \mathcal{S} by inclusion. [Warning: BHR and many other authors order \mathcal{S} by reverse inclusion.]

¹The reader should be warned that we will use this result only to prove Theorem 1.4.

One of our main results is a new proof of the BHR result, showing additionally that K is diagonalizable. Our proof is conceptual; it makes use of the algebraic topology of a regular cell complex Δ “dual” to \mathcal{A} . (In case \mathcal{A} is central, Δ is a convex polytope, called the *zonotope* associated to \mathcal{A} .) Here is the BHR result, combined with our improvement of it:

Theorem 1. *Let \mathcal{A} be a hyperplane arrangement in V , let \mathcal{F} be the set of faces, let \mathcal{S} be the intersection poset, and let w be a probability measure on \mathcal{F} . Then the matrix K defined at (1.3) is diagonalizable. For each $W \in \mathcal{S}$, there is an eigenvalue*

$$\lambda_W = \sum_{\substack{F \in \mathcal{F} \\ F \subseteq W}} w(F)$$

with multiplicity

$$m_W = |\mu(W, V)| = (-1)^{\text{codim}(W, V)} \mu(W, V),$$

where μ is the Möbius function of \mathcal{S} and $\text{codim}(W, V)$ is the codimension of W in V .

A second set of results proved here gives a description of the stationary distribution of the chain (1.1), together with a good estimate for the rate of convergence to stationarity. The estimate involves some of the eigenvalues (namely, the λ_H , $H \in \mathcal{A}$) and is surprisingly useful given that the chain is generally non-reversible.

We will say that the measure w *separates* the hyperplanes in \mathcal{A} , or simply that w is *separating*, if it is not concentrated on the faces in any one of them, i.e., for each $H \in \mathcal{A}$ there is a face F with $F \not\subseteq H$ and $w(F) > 0$.

Theorem 2. *Let \mathcal{A} be a hyperplane arrangement, let w be a probability measure on the set \mathcal{F} of faces, and let K be as in (1.3).*

- (a) *K has a unique stationary distribution π if and only if the measure w is separating.*
- (b) *Assume that w is separating. Sample without replacement from w , thereby getting an ordering F_1, \dots, F_m of $\{F \in \mathcal{F} : w(F) > 0\}$. Then the product $C = F_1 \cdots F_m$ in the semigroup \mathcal{F} is a chamber distributed from π .*
- (c) *Still assuming that w is separating, let K_C^ℓ be the distribution of the chain started from C after ℓ steps; then its total variation distance from π satisfies*

$$(1.4) \quad \|K_C^\ell - \pi\|_{\text{TV}} \leq \sum_{H \in \mathcal{A}} \lambda_H^\ell.$$

We remark that BHR give an estimate similar to (1.4), but involving all the eigenvalues λ_W and an alternating sum. We will discuss the connection between the two estimates in Section 4.

The remainder of the paper is organized as follows. In Section 2 we review definitions and facts about hyperplane arrangements. In Section 3 we discuss a number of examples. This section may be read first, for motivation. The examples include some previously studied card shuffling schemes, the classical Ehrenfests’ urn, random tiling, threshold graphs, and other examples. Section 4 contains the

proof of Theorem 2, in a more general setting; in particular, we consider walks driven by a stationary sequence on \mathcal{F} , not just a sequence of i.i.d. picks from \mathcal{F} . We prove Theorem 1 in Section 5. In Section 6 we explain how to extend all of the results to oriented matroids. Finally, there is an appendix which complements Section 2.

Some of the results of this paper extend to random walk on the chambers of a building. We will treat these in a separate paper.

Acknowledgements. We thank Phil Hanlon and Dan Rockmore for much interaction on this problem over the years. Jim Fill’s many results in the shuffling case have been an inspiration. Indeed, in preliminary work he has shown that many of his card shuffling arguments extend to the case of hyperplane walks; this gives a different conceptual argument for the result that the eigenvalues are positive reals and a somewhat complicated closed form formula for $K^\ell(C, C')$. Louis Billera, Vic Reiner, and Richard Stanley have been very helpful. We thank Susan Holmes for drawing the pictures.

2. REVIEW OF HYPERPLANE ARRANGEMENTS

The standard reference for this section is the book by Orlik and Terao [27]. Many of the results stated here can also be found in one or more of [9], [10], [11], and [38]. Throughout this section, \mathcal{A} denotes a finite set of affine hyperplanes in a finite-dimensional real vector space V . It will be convenient to write $\mathcal{A} = \{H_i\}_{i \in I}$ and to denote by H_i^+ and H_i^- the two open half-spaces determined by H_i . The choice of which one to call H_i^+ is arbitrary.

2A. Chambers and faces. “Face” in this paper will mean “relatively open face”. By definition, then, a face is a nonempty set $F \subseteq V$ of the form

$$F = \bigcap_{i \in I} H_i^{\sigma_i},$$

where $\sigma_i \in \{+, -, 0\}$ and $H_i^0 = H_i$. Equivalently, if we choose for each i an affine function $f_i : V \rightarrow \mathbb{R}$ such that H_i is defined by $f_i = 0$, then a face is a nonempty set defined by equalities and inequalities of the form $f_i > 0$, $f_i < 0$, or $f_i = 0$, one for each $i \in I$. The sequence $\sigma = (\sigma_i)_{i \in I}$ which encodes the definition of F is called the *sign sequence* of F and is denoted $\sigma(F)$.

The faces such that $\sigma_i \neq 0$ for all i are called *chambers*. They are convex open sets that partition the complement $V - \bigcup_{i \in I} H_i$. In general, a face F is open relative to its *support*, which is defined to be the affine subspace

$$\text{supp } F = \bigcap_{\sigma_i(F)=0} H_i.$$

In fact, the faces F with a given support W form the chambers of the hyperplane arrangement \mathcal{A}_W in W consisting of the intersections $H_i \cap W$ for those i such that $\sigma_i(F) \neq 0$. The arrangement \mathcal{A}_W is called the *restriction* of \mathcal{A} to W .

2B. Partial order. The *face poset* of \mathcal{A} is the set \mathcal{F} of faces, ordered as follows: Given $F, G \in \mathcal{F}$, we say that F is a *face* of G and write $F \leq G$ if for each $i \in I$ either $\sigma_i(F) = 0$ or $\sigma_i(F) = \sigma_i(G)$. In other words, the description of F by linear

equalities and inequalities is obtained from that of G by changing zero or more inequalities to equalities. See, for instance, [11] for other characterizations of the face relation. [Warning: Orlik and Terao use the reverse ordering on \mathcal{F} , i.e., they write $F \geq G$ if F is a face of G .]

Two chambers are said to be *adjacent* if they have a common codimension 1 face. The *chamber graph* of \mathcal{A} has \mathcal{C} as vertex set, with edges defined by the adjacency relation. We write $d(C, C')$ for the distance between C and C' in this graph; it is the minimal length ℓ of a “gallery”

$$C = C_0, \dots, C_\ell = C',$$

where C_{i-1} and C_i are adjacent for $1 \leq i \leq \ell$. It is also equal to the number of hyperplanes in \mathcal{A} separating C from C' (cf. [11], §I.4E).

2C. Product. The set of faces also admits a semigroup structure: Given $F, G \in \mathcal{F}$, their *product* FG is the face with sign sequence

$$\sigma_i(FG) = \begin{cases} \sigma_i(F) & \text{if } \sigma_i(F) \neq 0 \\ \sigma_i(G) & \text{if } \sigma_i(F) = 0. \end{cases}$$

[Geometric interpretation: If we move on a straight line from a point of F toward a point of G , then FG is the face we are in after moving a small positive distance.] This product is the one referred to in §1 and used to define the action of faces on chambers. One can check that FC is a chamber if C is, and that it is the unique chamber having F as a face that is closest to C in the metric d defined above. [To see this, use the characterization of d in terms of separating hyperplanes.]

2D. Cell decomposition of the sphere. Assume throughout this subsection that \mathcal{A} is central, in which case each H_i can be taken to pass through the origin. We may further assume that \mathcal{A} is *essential*, i.e., that $\bigcap_{i \in I} H_i = \{0\}$. [Otherwise we can replace V by the quotient space $V/\bigcap_{i \in I} H_i$ without affecting any of the combinatorial objects of interest to us.] There is then a regular cell complex $\Sigma = \Sigma_{\mathcal{A}}$, homeomorphic to the sphere S^{n-1} ($n = \dim V$), whose cells correspond to the faces $F \neq \{0\}$ in \mathcal{F} .

Recall first that a (finite) *regular cell complex* is a compact Hausdorff space X , together with a finite collection $\{e_\alpha\}$ of subsets of X , such that:

- (i) Each e_α is homeomorphic to a closed ball.
- (ii) The relative interiors $\overset{\circ}{e}_\alpha$ partition X .
- (iii) For each α , the boundary $\dot{e}_\alpha = e_\alpha - \overset{\circ}{e}_\alpha$ is a union of cells (necessarily of lower dimension).

The e_α are called the *closed cells* of X , and the $\overset{\circ}{e}_\alpha$ are called the *open cells*. See [9], §4.7, for further information about regular cell complexes.

The simplest way to construct the complex Σ associated to \mathcal{A} is to put a metric on V and intersect the cells $F \neq \{0\}$ (which are cones) with the unit sphere in V . See Figure 3a. It is also possible to realize Σ as the boundary of a convex polytope $\widehat{\Sigma}$; see Figure 3b. We give a construction of $\widehat{\Sigma}$ in the appendix to this paper. Alternatively, one can first define the zonotope $Z = Z_{\mathcal{A}}$ (see below) and define $\widehat{\Sigma}$ to be the polar of Z . This is the approach taken in [9], Example 4.1.7, and [28], Corollary 7.18.

Figure 3a. Cell decomposition of the unit sphere.

Figure 3b. A polytope version of Figure 3a.

Note that the hyperplane chamber walk can be viewed as a walk on the chambers (maximal cells) of Σ . Each step consists of choosing a cell e (possibly empty²) from some distribution on the cells, and then moving from the current chamber c to the nearest chamber having e as a face. “Nearest” here refers to gallery distance, which can be defined for the chambers of Σ exactly as in §2B above.

2E. The zonotope dual to \mathcal{A} . A *zonotope* in a real vector space V is a Minkowski sum $Z = L_1 + \cdots + L_k$ of line segments, usually taken to be centered at the origin: $L_i = [-v_i, v_i]$. We may assume that the L_i are non-degenerate and that no two are parallel, i.e., that the v_i are nonzero and pairwise independent. The L_i are then uniquely determined by Z ; in fact, there is one for each parallelism class of edges of Z . The set of faces of the boundary of Z having an edge parallel to L_i is called the *i*th *zone* of Z .

Note that Z is the image of the cube $[-1, 1]^k$ under the linear map $\mathbb{R}^k \rightarrow V$ taking the k standard basis vectors e_1, \dots, e_k to v_1, \dots, v_k . Thus Z is the convex hull of the 2^k vectors $\sum_{1 \leq i \leq k} \pm v_i$, where the signs can be chosen arbitrarily. A simple example of a zonotope is a hexagon, obtained by projecting a cube in \mathbb{R}^3 onto a plane. See [9] or [38] for further information about zonotopes.

Returning to our central hyperplane arrangement \mathcal{A} in V , there is a zonotope $Z = Z_{\mathcal{A}}$ in the dual space V^* , with one zone for each hyperplane in \mathcal{A} , defined as follows: Choose $f_i \in V^*$ such that $H_i = \ker f_i$ and set

$$Z = \sum_{i \in I} [-f_i, f_i].$$

Equivalently, Z is the convex hull of the $2^{|I|}$ elements $\sum_{i \in I} \pm f_i$.

The poset of nonempty faces of Z is anti-isomorphic to the face poset \mathcal{F} of \mathcal{A} . This is proved in [9], Proposition 2.2.2, and [38], §7.3. It also follows by polarity theory ([38], §2.3) from the results about Σ stated above and proved in the appendix, since Z is in fact the polar of the polytope $\widehat{\Sigma}$ defined in the appendix. Thus Z has one vertex for each chamber C [that vertex being $\sum \sigma_i f_i$, where $\sigma = \sigma(C)$], one edge for each pair of adjacent chambers, etc. In particular, the 1-skeleton of Z is

²This is equivalent to the fact that $\mathcal{F} = \{ \emptyset \} \subseteq \mathcal{F}$.

the chamber graph of \mathcal{A} . Figure 4 shows a simple example, in which $V = \mathbb{R}^2$ and V^* is identified with V .

Figure 4. The zonotope Z

Note that the hyperplane chamber walk can be viewed as a walk on the vertices of Z . Each step consists of choosing a random e of Z from some measure on the faces, and then moving from a vertex v to the unique vertex of e closest to v (in the usual edge-path metric on the 1-skeleton of Z).

Remark. In some of the literature there is a slightly different definition of the zonotope associated to \mathcal{A} . Namely, one considers

$$Z' = \sum_{i \in I} [0, f_i],$$

or, equivalently, Z' is the convex hull of the $2^{|I|}$ elements $\sum_{i \in s} f_i$, where $s \subseteq I$ is an arbitrary subset. Note that Z' is obtained from Z by translating by $\sum_{i \in I} f_i$ and then multiplying by $1/2$. In particular, Z and Z' are combinatorially equivalent.

2F. The non-central case. For arbitrary \mathcal{A} , there is still a regular cell complex Δ dual to the arrangement. It is again a topological ball (though not necessarily a polytope), and its poset of cells is anti-isomorphic to the face poset \mathcal{F} of \mathcal{A} . See Figure 5 for a simple example.

Figure 5. The cell complex dual to an affine arrangement

We briefly recall the construction of Δ , which can be found in Ziegler [36], since it involves ideas that we will need later anyway. Let $\overline{V} = V \times \mathbb{R}$. For each $i \in I$ let \overline{H}_i be the linear hyperplane in \overline{V} spanned by $H_i \times 1$. Then the \overline{H}_i , together with the hyperplane $H_0 = V \times 0$, form a central arrangement $\overline{\mathcal{A}}$ in \overline{V} , said to be obtained from \mathcal{A} by *coning*. See Figure 6. Let $\overline{\mathcal{F}}$ be the face poset of $\overline{\mathcal{A}}$. Then the face poset \mathcal{F} of \mathcal{A} is isomorphic to the subset $\overline{\mathcal{F}}^+$ of $\overline{\mathcal{F}}$ consisting of the faces in the upper half space $V \times (0, \infty)$ of \overline{V} . In terms of sign sequences, we go from \mathcal{F} to $\overline{\mathcal{F}}^+$ by adjoining $+$ as the H_0 -component.

Figure 6. A central arrangement of 3 hyperplanes in \mathbb{R}^2 obtained by coning a non-central arrangement of 2 hyperplanes in \mathbb{R}^1 .

Let \overline{Z} be the zonotope associated to $\overline{\mathcal{A}}$, with cell poset $\overline{\mathcal{F}}^{\text{op}}$. Then the cells corresponding to $\overline{\mathcal{F}}^+$ give a subcomplex of \overline{Z} , and this is the desired Δ . We remark that Δ , though not a zonotope in the non-central case, can always be realized as the set of faces of a “zonotopal tiling”. (We will say more about zonotopal tilings in §3E.) For example, the complex Δ in Figure 5 above is combinatorially equivalent to the zonotopal tiling called Z_3 in §3E.

2G. Reflection arrangements. Finally, we briefly mention an important family of examples of central hyperplane arrangements. Assume that V is equipped with an inner product. Then every linear hyperplane H in V gives rise to a reflection s_H that fixes H pointwise and acts as -1 on the orthogonal complement. A *finite reflection group* in V is a finite group G of linear transformations such that G is generated by reflections s_H . The set of all H such that $s_H \in G$ is the *reflection arrangement* associated to G .

Reflection arrangements have a number of special properties; see, for instance, [11], Chapter I. For example, the chambers are always simplicial cones if the arrangement is essential, which we may assume without loss of generality. (A chamber of a reflection arrangement in \mathbb{R}^3 cannot, for instance, be the cone over a square.) It follows that the spherical cell-complex Σ of §2D is a simplicial complex. Moreover, the group G acts simply-transitively on \mathcal{C} , so that \mathcal{C} can be identified with G once a “fundamental chamber” is chosen. Thus the hyperplane chamber walk can be interpreted as a Markov chain on a group in the case of a reflection arrangement.

Remark. The group G acts on the face poset \mathcal{F} . If the measure w on \mathcal{F} is separating and G -invariant, then the stationary distribution π is uniform. Moreover, the chamber walk in this case is a random walk on G in the usual sense, i.e., it consists of repeated multiplication by random elements of G , chosen from the measure

$$Q(g) = \sum_{F: \text{id} = g} w(F).$$

We thus have a reasonable collection of natural measures on groups, with explicitly analyzable random walks.

2H. Möbius function. Finally, we recall the definition of the function $\mu = \mu_{\mathcal{S}}$ that occurs in the statement of Theorem 1 (cf. [31], §3.7, or [27], §2.2). This is defined inductively by $\mu(V, V) = 1$ and, for $W \subsetneq V$,

$$(2.1) \quad \mu(W, V) = - \sum_{W \subseteq U \subsetneq V} \mu(W, U).$$

For example, if \mathcal{A} consists of 3 lines L_i through the origin in \mathbb{R}^2 , then the intersection lattice is

$$\mathbb{R}^2$$

$$L_1 \quad L_2 \quad L_3$$

$$\{0\}$$

In this case the Möbius numbers appearing in Theorem 1 are $\mu(\mathbb{R}^2, \mathbb{R}^2) = 1$, $\mu(L_i, \mathbb{R}^2) = -1$, and $\mu(\{0\}, \mathbb{R}^2) = 2$.

3. EXAMPLES

This section collects examples of hyperplane walks which have a natural alternative interpretation. We also make explicit the results of Theorems 1 and 2. In 3A we treat the dihedral arrangement, in 3B the Boolean arrangement, and in 3C the braid arrangement, with its many shuffling and computer science interpretations. In 3D we discuss an arrangement related to threshold graphs. In 3E we treat zonotopal tilings and introduce random walks on oriented matroids. Finally, we briefly mention some further examples in 3F.

Before beginning the examples, we make two remarks which sometimes simplify the computation of the stationary distribution π given in Theorem 2. First, instead of sampling without replacement, we could sample with replacement, stopping as soon as the product $F_1 \cdots F_m$ is a chamber. This gives the same distribution π , because we can strike out any factors in $F_1 \cdots F_m$ which have occurred earlier without affecting the value of the product. But sampling with replacement and then deleting repetitions is the same as sampling without replacement.

At the other extreme, we could remove from the pot even more than just the faces that are picked. Namely, if F_1, \dots, F_k have been picked and W is the support of the product $F_1 \cdots F_k$, then we can remove all faces F contained in W before picking the next face F_{k+1} . This follows from the same striking-out argument as above, the point being that a later factor contained in W will have no effect on the product.

3A. Dihedral arrangement. Let \mathcal{A} consist of m lines through the origin in \mathbb{R}^2 . [If the lines are equally spaced, this is an example of a reflection arrangement, the associated reflection group being dihedral of order $2m$.] There are $4m+1$ faces: $2m$ chambers, $2m$ rays, and the origin. Suppose, for this exposition, that the measure w is supported on the set of rays. One can then picture the walk as follows: There are $2m$ rooms in a circular house. A mouse lives in the walls R (the rays), occupying these with propensity $w(R)$. At each step of the walk, a cat is in one of the rooms and the mouse picks a wall; the cat then moves to the nearest room adjacent to that wall.

Note that the (1-dimensional) spherical complex Σ of §2D is a $2m$ -gon in this example. The chambers of the hyperplane arrangement correspond to the edges of Σ , and the rays correspond to the vertices. So we can visualize the walk as taking place on the edges of a $2m$ -gon, driven by a probability measure on the vertices. One can imagine here a queuing system with $2m$ service points arranged in a ring, corresponding to the vertices of a $2m$ -gon. A single server moves around the edges. Service requests come in with different propensities w_1, \dots, w_{2m} , and the server moves to the closest adjacent edge.

If the rays are chosen uniformly, $w(R) = 1/2m$ for all R , then the stationary distribution π is of course uniform. For general weights, w is separating unless it is supported on a pair $\pm R$ of opposite rays. If w is separating, Theorem 2 yields the following formula for π : Let C be a chamber bounded by rays R, R' , whose supports are the lines L, L' . Let \mathcal{Q} be the set of rays that are strictly on the same side of L as C , and define \mathcal{Q}' similarly in terms of L' . See Figure 7. Then we have

$$\pi(C) = w(R) \frac{w(\mathcal{Q})}{1 - w(R) - w(-R)} + w(R') \frac{w(\mathcal{Q}')}{1 - w(R') - w(-R')}.$$

[To see this, use the second variant of the sampling procedure described above. The sampling stops as soon as two rays have been picked, and we get C as the product $F_1 F_2$ if $F_1 = R$ and $F_2 \in \mathcal{Q}$ or if $F_1 = R'$ and $F_2 \in \mathcal{Q}'$.]

Figure 7

The eigenvalues given by Theorem 1 are as follows (see the example in §2H above): Each of the m lines L contributes an eigenvalue $\lambda_L = w(R) + w(-R)$ of multiplicity $1 = -\mu(L, V)$, where $\pm R$ are the rays in L . The whole plane $V = \mathbb{R}^2$ contributes the eigenvalue $\lambda_V = 1$ with multiplicity $\mu(V, V) = 1$. Finally, the trivial

subspace $\{0\}$ contributes the eigenvalue $\lambda_{\{0\}} = 0$ with multiplicity $\mu(\{0\}, V) = m - 1$.

Consider now the bound of Theorem 2 in three simple cases. Suppose first that $w(R) = 1/2m$ for each R , so that π is uniform, $\pi(C) = 1/2m$ for all C . Here $\lambda_L = 1/m$ so the bound becomes

$$\|K_C^\ell - \pi\|_{\text{TV}} \leq \sum_L \lambda_L^\ell = \frac{1}{m^{\ell-1}}.$$

It follows that for large m the distance to stationarity is small after two or three steps.

As a second example, suppose one weight is large and the others are small, e.g., $w(R_1) = 1/2$, $w(R) = 1/2(2m - 1)$ for $R \neq R_1$. Then the bound becomes

$$\|K_C^\ell - \pi\|_{\text{TV}} \leq \left(\frac{1}{2} + \frac{1}{2(2m-1)} \right)^\ell + (m-1) \left(\frac{1}{2m-1} \right)^\ell.$$

Again, a few steps suffice for convergence to stationarity, but the result is not quite as quick as in the uniform case.

As a third example, suppose the weights are proportional to $1, 1/2, \dots, 1/2m$, i.e.,

$$w(R_i) = \frac{1}{iH_{2m}}, \quad 1 \leq i \leq 2m, \quad H_{2m} = 1 + \frac{1}{2} + \dots + \frac{1}{2m} \sim \log m.$$

Assume here that the rays are enumerated so that $R_{i+m} = -R_i$ for $i = 1, \dots, m$. Then the bound becomes

$$\|K_C^\ell - \pi\|_{\text{TV}} \leq \sum_{i=1}^m \left(\frac{1}{H_{2m}} \left(\frac{1}{i} + \frac{1}{i+m} \right) \right)^\ell \leq \left(\frac{c}{\log m} \right)^\ell$$

for a universal constant c . Again, a few steps suffice to reach stationarity.

3B. Boolean arrangement. Let H_i be the coordinate hyperplane $x_i = 0$ in \mathbb{R}^n , $1 \leq i \leq n$. [The H_i again form a reflection arrangement, the group being $\{\pm 1\}^n$.] There are 3^n faces, one for each possible sign sequence, and 2^n of these are chambers; they are the orthants in \mathbb{R}^n and may be identified with the elements of $\{\pm 1\}^n$ (or with binary n -tuples). The polytope $\widehat{\Sigma}$ is a hyperoctahedron, and the dual zonotope Z is the cube $[-1, 1]^n$.

To picture the chamber walk, think of an element $x \in \{\pm 1\}^n$ as a landscape with n sites, each of which can be in one of two states. The action of a face F can be thought of as a ruler who conquers territory at sites in $s = \{i : \sigma_i(F) \neq 0\}$ and changes the territory in his own image. One may ask how the landscape evolves over time as territory is conquered by successive rulers F , chosen from some probability distribution on \mathcal{F} .

To make this more explicit, we pick the random face F as follows: First pick the subset $s \subseteq [n]$ from a probability distribution w_s . Then pick the nonzero components of $\sigma(F)$ from a probability distribution $P_s(\cdot)$ on $\{\pm 1\}^{|s|}$. We briefly discuss three examples.

Example 1. (Ehrenfests' urn) Suppose w is uniform on the singletons, $w_{\{i\}} = 1/n$, $w_{\{s\}} = 0$ otherwise. Suppose further that $P_s(\pm 1) = 1/2$. In other words, our measure

on \mathcal{F} is concentrated on the $2n$ coordinate rays and is uniform on these. The walk then evolves as, “Pick a coordinate of x at random, and half the time replace it by $+1$, half the time by -1 .” This is the same as the usual nearest neighbor random walk on the hypercube, with holding $1/2$. It has been extensively studied since its introduction by the Ehrenfests [22]. Here π is uniform, $\pi(x) = 1/2^n$, and the bound given by Theorem 2 is

$$\|K_x^\ell - \pi\| \leq n \left(1 - \frac{1}{n}\right)^\ell.$$

This shows that $n \log n + cn$ steps make variation distance smaller than e^{-c} . In fact, it is known that $(1/2)n \log n + cn$ steps are necessary and sufficient for convergence [18]. Thus the bound is good but not perfect.

The eigenvalues for this example were determined by Mark Kac [24] to be j/n with multiplicity $\binom{n}{j}$, $0 \leq j \leq n$. To see this from Theorem 1, observe that the intersection lattice \mathcal{S} here is isomorphic to the lattice \mathcal{L} of subsets $s \subseteq [n]$, with $W \in \mathcal{S}$ corresponding to $s = \{i : x_i \not\equiv 0 \text{ on } W\}$. One deduces

$$\lambda_W = \sum_{F \subseteq W} w(F) = 2|s| \cdot \frac{1}{2n} = \frac{|s|}{n}$$

and

$$m_W = |\mu_{\mathcal{S}}(W, V)| = |\mu_{\mathcal{L}}(s, [n])| = |(-1)^{n-|s|}| = 1.$$

This agrees with Kac’s result since there are $\binom{n}{j}$ subspaces W with $\lambda_W = j/n$.

Example 2. Consider the chain based on the same set of faces (the coordinate rays) but with general weights. Write $w_{\{i\}} = w_i$ and $P_i(1) = \theta_i$, $P_i(-1) = 1 - \theta_i$. The resulting measure on \mathcal{F} is separating if $w_i > 0$ for all i , and the stationary distribution is $\pi(x) = \prod P_i(x_i) = \prod \theta_i^{\varepsilon_i} (1 - \theta_i)^{1 - \varepsilon_i}$, where $\varepsilon_i = (x_i + 1)/2$. The bound for convergence is

$$\|K_x^\ell - \pi\| \leq \sum_{i=1}^n (1 - w_i)^\ell.$$

The rate of convergence depends on the shape of the weights. See [16] for many specific examples. Arguing as in Example 1, we obtain an eigenvalue

$$\lambda_s = \sum_{i \in s} w_i,$$

of multiplicity 1, for each subset $s \subseteq [n]$.

Example 3. (Changing landscape) Consider the $2n$ sets $\ell_i = \{1, 2, \dots, i\}$, $r_i = \{n, n-1, \dots, n-i+1\}$, $1 \leq i \leq n$. Let $w(\ell_i) = w(r_i) = 1/2n$. Suppose P_{ℓ_i} puts all $+1$ ’s in ℓ_i with probability 1 and P_{r_i} puts all -1 ’s in r_i with probability 1. This is a crude model of territory exchange: A force attacks from the left taking a uniform amount of territory and labeling it with $+1$. Attacks from the right label with -1 . The stationary distribution is supported on patterns of the form

$$\pi^{(j)} = \underbrace{\left(\underbrace{+1}_{\text{j}}, \underbrace{-1}_{\text{n-j}} \right)}_{\text{j}}, \quad 0 \leq j \leq n.$$

moreover, π is uniform on these, $\pi(x^{(j)}) = 1/(n+1)$.

[Sketch of proof: It suffices to show that $\pi\{N \leq j\} = (j+1)/(n+1)$, where $N(x)$ is the number of +1's in x . Now $\pi\{N \leq j\}$ is the probability p that, in sampling from the uniform distribution on $X = \{\ell_1, \dots, \ell_n, r_1, \dots, r_n\}$, an element of $R = \{r_{n-j}, \dots, r_n\}$ is chosen before an element of $L = \{\ell_{j+1}, \dots, \ell_n\}$. This probability is unchanged if we replace X by $R \cup L$, so it is simply the probability that a uniform pick from $R \cup L$ is in R , i.e., $p = |R|/(|R| + |L|) = (j+1)/(n+1)$.]

Theorem 2 gives

$$\|K_x^\ell - \pi\| \leq n \left(\frac{n-1}{2n} \right)^\ell \leq n \left(\frac{1}{2} \right)^\ell.$$

Thus $\log_2 n + c$ steps suffice to reach stationarity. In fact, it is easy to see that $K^2(x, x^{(j)}) \geq 1/4n$ for all x and j . A Doeblin argument then shows that $\|K_x^\ell - \pi\| \leq (3/4)^{\lfloor \ell/2 \rfloor}$. So the bound from Theorem 2 is slightly off in this example.

Finally, Theorem 1 gives an eigenvalue $(i+j)/2n$ for each subset $s \subseteq [n]$, where i is the size of the largest subset $\{1, \dots, i\}$ in s and j is the size of the largest subset $\{n-j+1, \dots, n\}$ of s ($0 \leq i, j \leq n$). Combining these according to the value of $k = i+j$, we obtain the following eigenvalues λ and multiplicities $m(\lambda)$:

λ	$m(\lambda)$
$\frac{1}{2n}$	1
$\frac{n-1}{2n}$	n
$\frac{k}{2n}$	$(k+1)2^{n-k-2}$ $0 \leq k \leq n-2$.

3C. Braid arrangement. One of the discoveries of BHR [6] (see also [5]) is that for the well-studied braid arrangement the action of faces on chambers captures a wide variety of shuffling schemes. The braid arrangement in \mathbb{R}^n consists of the $\binom{n}{2}$ hyperplanes H_{ij} given by $x_i - x_j = 0$ ($1 \leq i < j \leq n$). [This is again a reflection arrangement, the group being the symmetric group S_n on n letters.] The chambers can be identified with the $n!$ permutations, with $\tau \in S_n$ corresponding to the chamber

$$x_{\tau(1)} > x_{\tau(2)} > \cdots > x_{\tau(n)}.$$

The faces can be identified with ordered partitions $B = (B_1, \dots, B_k)$ of $[n]$. Here $\{B_1, \dots, B_k\}$ is a set partition in the usual sense, but the order matters. (We recover the chambers by taking $k = n$, so that each B_i is a singleton.) The spherical cell complex Σ is the barycentric subdivision of the boundary of an $(n-1)$ -simplex, and the zonotope Z is isomorphic to the permutohedron, which is the convex hull of the $n!$ permutations τ , viewed as vectors $(\tau(1), \dots, \tau(n))$.

The action of faces on chambers is most easily pictured by thinking of a permutation τ as the set of labels on a deck of n cards, with the card labeled $\tau(1)$ on top, and so on. The ordered partition B operates on τ by removing cards with labels in B_1 and placing them on top (keeping them in the same relative order), then removing cards with labels in B_2 and placing them next, and so on. Suppose, for example, that $n = 10$, $\tau = (1, 7, 3, 9, 10, 4, 5, 2, 6, 8)$, and $B = (\{2, 5\}, \{3, 4, 6, 10\}, \{7\}, \{1, 8, 9\})$; then B acting on τ gives $(5, 2, 3, 10, 4, 6, 7, 1, 9, 8)$.

We briefly describe two examples which have received much attention in other settings.

Example 1. (Random to top) Suppose the only ordered partitions that get positive mass are $(\{i\}, [n] - \{i\})$ with mass w_i , $1 \leq i \leq n$. The walk corresponds to repeatedly choosing i from w_i and then moving the card labeled i to the top. This is a well-studied model for dynamic rearrangement of files in computer science. Think of file folders being used with propensity w_i . One wants the frequently used files near the top. A simple self-organizing scheme for achieving this if the w_i are not known is to replace a folder on top after it is used. This scheme is called the Tsetlin library; see Dies [19] or Fill [23] for extensive reviews.

Assuming $w_i > 0$ for all i (or even all but one i), there is a unique stationary distribution π . It is given by Theorem 2 as sampling without replacement from the weights w_i :

$$\pi(\tau) = \frac{w_{\tau(1)} w_{\tau(2)} \cdots w_{\tau(n-1)}}{(1 - w_{\tau(1)})(1 - w_{\tau(1)} - w_{\tau(2)}) \cdots (1 - w_{\tau(1)} - \cdots - w_{\tau(n-2)})}.$$

This generally non-uniform distribution is easier to describe than to work with: Try to compute $\pi\{\tau(n) = 1\}$. Even the distribution π has its own literature; see [15], p. 174.

The bound from Theorem 2 is

$$\|K_\tau^\ell - \pi\| \leq \sum_{1 \leq i < j \leq n} (1 - w_i - w_j)^\ell.$$

A variety of special cases are analyzed in [16]. Suppose, for instance, that $w_i = 1/n$, $1 \leq i \leq n$. Then π is uniform, and the convergence rate is the same as for the top-to-random shuffle, in which the top card is repeatedly inserted at a random position.

The bound from Theorem 2 becomes

$$\|K_\tau^\ell - \pi\| \leq \binom{n}{2} \left(1 - \frac{2}{n}\right)^\ell,$$

so that $n(\log n + c)$ shuffles suffice to make the distance at most $e^{-c}/2$. More refined estimates are derived in [17], showing that this bound is sharp.

For this example with general weights w_i the eigenvalues were determined by Phatarfod [28]. Each subset $s \subseteq [n]$ contributes an eigenvalue

$$\lambda_s = \sum_{i \in s} w_i,$$

of multiplicity m_s equal to the number of permutations $\tau \in S_n$ with s as fixed-point set. In other words, m_s is the derangement number d_k , $k = n - |s|$, where d_k is the number of permutations in S_k with no fixed points. Note that $d_1 = 0$, so λ_s does not actually occur as an eigenvalue if $|s| = n - 1$.

It is shown in [6] and [5], by two different methods, how Theorem 1 gives Phatarfod's result. We briefly sketch a third method, since the ideas will be needed in a more difficult example below (§3D). This third method has the advantage that it can be used in cases where the Möbius numbers $\mu_{\mathcal{S}}(W, V)$ are not known.

A subspace $W \in \mathcal{S}$ is defined by zero or more equations of the form $x_i = x_j$. Let s_W be the set of $k \in [n]$ which do *not* occur in any of these equations. [If we identify the elements of \mathcal{S} with set partitions of $[n]$ in the usual way, then s_W is the union of the singleton blocks.] A straightforward application of Theorem 1 now gives eigenvalues λ_s ($s \subseteq [n]$) with multiplicity

$$n_s = \sum_{s_W = s} m_W,$$

where $m_W = |\mu_{\mathcal{S}}(W, V)|$. In particular, m_W depends only on the interval $[W, V]$ in the lattice \mathcal{S} . Now consider

$$(3.1) \quad N_s \stackrel{\text{def}}{=} \sum_{t \supseteq s} n_t = \sum_{s_W \supseteq s} m_W.$$

The W 's that occur here form a lattice isomorphic to the intersection lattice for the braid arrangement in $\mathbb{R}^{n-|s|}$. [Use the projection $\mathbb{R}^n \rightarrow \mathbb{R}^{n-|s|}$ which picks out the coordinates not in s .] It follows that the second sum in (3.1) is the sum of all multiplicities for the latter arrangement, hence it equals the total number of chambers.³ Thus N_s is the number of permutations of $[n] - s$ or, equivalently, the number of permutations of $[n]$ that fix s . If we now define m_s to be the number of permutations with s as fixed-point set, and if we set $M_s = \sum_{t \supseteq s} m_t$, we see that $N_s = M_s$ and hence $n_s = m_s$, as claimed.

Example 2. (Riffle shuffle) Consider next the 2-block ordered partitions $(s, [n] - s)$, $\emptyset \subsetneq s \subsetneq [n]$, together with the one-block partition $([n])$. We assign weight $1/2^n$ to each of the $2^n - 2$ two-block ordered partitions, and we assign weight $2/2^n =$

³One could also note that this is a Zeckendorf sum, [25], so formula (5.6) is valid.

$1/2^{n-1}$ to the one-block partition. The corresponding shuffling mechanism consists of inverse riffle shuffles. In an ordinary riffle shuffle a deck of cards is divided into two piles which are riffled together. The inverse chooses a set s of cards which are removed (“unriffled”) and placed on top. Here s can be \emptyset or $[n]$, in which case the deck is unchanged; these cases both correspond to the action of the one-block partition. Thus the effect of our choice of weights is that the 2^n subsets $s \subseteq [n]$ are all equally likely to be unriffled. This corresponds to the Gilbert–Shannon–Reeds measure, in which the subset to be riffled is chosen uniformly (see [2], where reference to earlier work is given).

The stationary distribution π is uniform. The convergence rate is the same for ordinary and inverse shuffles. The bound from Theorem 2 gives

$$\|K_\tau^\ell - \pi\| \leq \binom{n}{2} \left(\frac{1}{2}\right)^\ell.$$

Thus the distance to uniformity is less than 2^{-c} after $2 \log_2 n + c - 1$ steps. A more exact analysis is available [2] showing that the variation distance rapidly cuts down from 1 to 0 at about $\ell = (3/2) \log_2 n$. Thus the general bound from Theorem 2 again gives a quite good result in this case, though not the best possible. When $n = 52$, for example, $2 \log_2 n \doteq 11.4$, but in fact about seven shuffles suffice to mix up the deck.

See [6] and [5] for many other examples, including the a -shuffles of [2], together with a detailed description of the eigenvalues. The a -shuffles were shown in [2] to give the celebrated descent algebra of L. Solomon [30]. Connections between the chamber walk and the descent algebra for general reflection groups are developed in Bidigare [5] and Bergeron–Bergeron–Howlett–Taylor [4].

3D. Threshold arrangement. Our next example is a random walk on an interesting family of graphs,⁴ called *threshold graphs*. These can be characterized in many different ways (see [26]), of which we mention three. Let G be a graph with vertex set V .

- (1) G is a threshold graph if and only if there exist real numbers $w_v \geq 0$ ($v \in V$) and $t \geq 0$ such that the edges of G are the pairs uv with $w_u + w_v > t$.
- (2) Let $d = (d_v)_{v \in V}$ be the degree sequence of G , i.e., d_v is the number of edges having v as a vertex. Then G is a threshold graph if and only if it is the unique graph on V with degree sequence d .
- (3) G is a threshold graph if and only if it can be constructed from the empty graph by repeatedly adjoining either an isolated vertex or a dominating vertex. [Recall that a vertex of a graph is called *dominating* if it is connected by an edge to all other vertices.]

For example, the graph shown in Figure 8 is a threshold graph, being the unique graph on $\{1, 2, 3, 4, 5\}$ with degree sequence $d = (1, 4, 2, 2, 3)$. The condition in (1) holds with weights $w = (1, 5, 2, 2, 4)$ and $t = 5$. And G can be constructed from the empty graph by adjoining 3 and 4 as isolated vertices, then adjoining 5 as a dominating vertex, then 1 as an isolated vertex, and finally 2 as a dominating vertex.

⁴Graphs have also been called *completely symmetric* [1] and *uniquely partitionable* [26].

Figure 8. A threshold graph

Notice that, by (3), a threshold graph can be represented (usually in more than one way) by a signed permutation $\tau(1)^{\varepsilon(1)}\tau(2)^{\varepsilon(2)}\dots\tau(n)^{\varepsilon(n)}$. Here τ is a permutation and $\varepsilon(i) = \pm 1$. The corresponding graph is obtained by adjoining the vertices in the order $\tau(n), \tau(n-1), \dots, \tau(1)$, and making $\tau(i)$ dominating if $\varepsilon(i) = +1$ and isolated if $\varepsilon(i) = -1$. For example, the graph in Figure 8 is represented by each of the eight signed permutations $2^+1^-5^+3^-4^\pm$, $2^+1^-5^+4^-3^\pm$. And the complete graph on $[n]$ can be represented by $n!$ signed permutations, as can its complement, the discrete graph. [By the *complement* of a graph G we mean the graph G' with the same vertex set but the complementary edge set.] These examples suggest that the number t_n of threshold graphs on $[n]$ is much less than $2^n n!$. In fact, it is known [3] that

$$\frac{t_n}{n!} \sim \left(\frac{1}{\log 2} - 1 \right) \left(\frac{1}{\log 2} \right)^n,$$

with an absolute error that tends to 0 exponentially fast as $n \rightarrow \infty$.

Consider now the hyperplane arrangement in \mathbb{R}^n consisting of the $\binom{n}{2}$ hyperplanes $x_i + x_j = 0$, one for each two-element subset ij of $[n]$. The zonotope Z dual to this arrangement has been studied extensively; see [26], [32], and further references cited there. It is convenient here to take Z to be the zonotope called Z' in §2E. Identifying \mathbb{R}^n with its dual, we see that Z is the convex hull of vectors d_E , where E is an arbitrary collection of 2-element subsets ij of $[n]$ and

$$d_E = \sum_{ij \in E} (e_i + e_j).$$

[Here e_1, \dots, e_n denotes the standard basis of \mathbb{R}^n .] Now E may be viewed as the set of edges of a graph G on $[n]$, and $d_E = (d_1, \dots, d_n)$ is the degree sequence of G . Thus Z is the convex hull of the set of degree sequences of graphs on $[n]$.

The first major result about Z is that its vertices are the degree sequences of the threshold graphs on $[n]$. Hence the chambers of our hyperplane arrangement can be identified with threshold graphs. Explicitly, a chamber with sign sequence (σ_{ij}) corresponds to the graph with edge set $\{ij : \sigma_{ij} = +\}$. The face poset \mathcal{F} seems quite complicated, but we can single out a subset of \mathcal{F} whose action generates an interesting random walk:

Assume $n \geq 4$. Then each of the $2n$ coordinate rays R_i^\pm is a face. For example

R_1^+ is defined by

$$\begin{aligned} x_1 + x_j &> 0 \quad \text{for } 2 \leq j \leq n \\ x_i + x_j &= 0 \quad \text{for } 2 \leq i, j \leq n, i \neq j. \end{aligned}$$

The action of R_i^ε on chambers is easily pictured: Given a threshold graph, R_i^+ adds edges to make the vertex i dominating, and R_i^- deletes edges to make i isolated.

As in Example 2 of §3B, we can describe a probability measure on these $2n$ coordinate rays by giving weights $w_i \geq 0$ ($1 \leq i \leq n$, $\sum w_i = 1$) and “coin-tossing parameters” θ_i ($0 \leq \theta_i \leq 1$). The resulting walk then evolves as follows: At each stage there is a threshold graph. Pick a vertex i according to the weights w_i . Then flip a coin with probability θ_i of heads. If heads comes up, add edges to make i dominating; otherwise delete all edges involving i .

If $w_i > 0$ for all i (or even all but one i), then our measure on \mathcal{F} is separating and there is a unique stationary distribution π . It can be described as follows: Sample without replacement from the weights w_i to get an ordering i_1, \dots, i_n of $[n]$. Pick signs $\varepsilon(i) = \pm 1$, where $\varepsilon(i) = +1$ with probability θ_i . Then $i_1^{\varepsilon(i_1)} i_2^{\varepsilon(i_2)} \dots i_n^{\varepsilon(i_n)}$ is a random signed permutation whose associated graph G is distributed from π .

If $w_i = 1/n$ and $\theta_i = 1/2$ for all i , then π is simply the measure on threshold graphs induced by the uniform distribution on signed permutations. In other words,

$$\pi(G) = \frac{s(G)}{2^n n!},$$

where $s(G)$ is the number of signed permutations representing G . For example, $\pi(G) = 1/2^n$ if G is the complete graph or the discrete graph. Note, in particular, that π is not uniform.

The convergence bound given by Theorem 2 is exactly the same as in Example 1 of §3C (weighted random-to-top shuffle):

$$\|K_G^\ell - \pi\| \leq \sum_{ij} (1 - w_i - w_j)^\ell,$$

where the sum is taken over all 2-element subsets ij of $[n]$. In particular, $n \log n + cn$ steps suffice to reach stationarity if $w_i = 1/n$ for all i .

Finally, one can work out the eigenvalues and multiplicities by a slight variation on the method used in the random-to-top example. We omit the details and simply state the result: For each subset $s \subseteq [n]$ there is an eigenvalue

$$\lambda_s = \sum_{i \in s} w_i,$$

with multiplicity m_s equal to the number of threshold graphs on $[n]$ with s as the set of isolated vertices. Equivalently, if $n - |s| = k$, then $m_s = \tau_k$, where τ_k is the number of threshold graphs on $[k]$ with no isolated vertices. Thus τ_k is an analogue of the derangement number d_k .

We have $\tau_0 = 1$, $\tau_1 = 0$ [so that $m_s = 0$ if $|s| = n - 1$, and λ_s does not actually occur], and $\tau_k = t_k/2$ for $k \geq 2$ where, as above, t_k is the number of threshold graphs on $[k]$. This follows from the fact that, by the characterization

(3) of threshold graphs, G has no isolated vertices if and only if its complement G' has at least one isolated vertex.

It is remarkable that the multiplicities m_s can be obtained, as in the random-to-top example, with virtually no knowledge about the intersection lattice \mathcal{S} . For completeness, however, we give a brief description of the elements of \mathcal{S} . A subspace $W \in \mathcal{S}$ of dimension r is determined by the following data:

- (i) A subset $s_0 \subseteq [n]$ which is either empty or else has at least 3 elements.
- (ii) A set partition $\{B_1, \dots, B_r\}$ of the complement $[n] - s_0$.
- (iii) For each block B_i with $|B_i| \geq 2$, a set partition of B_i into two parts.

Write $i \smile j$ if i and j are in the same part of some block B_k with $|B_k| \geq 2$ and $i \frown j$ if they are in different parts of some B_k . Then the subspace W corresponding to (i)–(iii) is given by the (redundant) system of equations

$$\begin{aligned} x_i &= 0 && \text{if } i \in s_0 \\ x_i &= x_j && \text{if } i \smile j \\ x_i &= -x_j && \text{if } i \frown j. \end{aligned}$$

To prove that elements of \mathcal{S} correspond to data as in (i)–(iii), let $W \in \mathcal{S}$ be defined by equations of the form $x_i + x_j = 0$. Let Γ be the graph on $[n]$ whose edges are the ij such that $x_i + x_j = 0$ on W . Then $x_i = x_j$ on W (resp., $x_i = -x_j$ on W) if i and j can be joined by a path in Γ of even (resp., odd) length. If Γ contains a cycle of odd length, it follows that $x_i = 0$ on W for all vertices i in the connected component Γ_0 containing that cycle. Moreover, Γ_0 is the complete graph on $s_0 = \{i : x_i = 0 \text{ on } W\}$, and all other connected components are bipartite. The remaining details are left to the reader.

Remark. Although the face poset \mathcal{F} is complicated, its atoms (i.e., the faces that are rays) are easy to describe: Each $W \in \mathcal{S}$ of dimension $r = 1$ contributes two rays. So we get one ray for each ordered pair (s, t) of disjoint nonempty subsets of $[n]$ which either cover $[n]$ or omit at least 3 elements. In addition, we get the $2n$ coordinate rays if $n \geq 4$; these correspond to the case $|s_0| = n - 1$. This description of the rays in \mathcal{F} or, equivalently, the vertices of the polytope $\widehat{\Sigma}$ (§2D), is equivalent to the description of the facets of the zonotope Z given in [26], Theorem 3.3.17.

3E. Zonotopes, tilings, and oriented matroids. As explained in Section 2E, the dual of a central hyperplane arrangement is a convex polyhedron called a zonotope. The hyperplane chamber walks become walks on the vertices of the zonotope. In this section we show how to analyze random walk on the vertices of a tiling of a zonotope. Given a zonotope $Z = \sum_{i=1}^n [-v_i, v_i]$, a *zonotopal tiling* of Z is a polyhedral subdivision of Z in which all of the faces are translates of zonotopes of the form $\sum_{i \in s} [-v_i, v_i]$, where $s \subseteq \{1, \dots, n\}$. Figure 9 shows two examples. See [38], §7.5, or [9], §2.2, for further information about zonotopal tilings.

A walk on the vertices of a zonotopal tiling is driven by a probability distribution $w(\cdot)$ on the faces making up the tiling. From a given vertex x , choose a face F from $w(\cdot)$ and move to the vertex of F closest to x . If the zonotopal tiling is a projection of a higher dimensional zonotope (Figure 9a), the walk is just projected and no new theory is needed. As Figure 9b suggests, not all zonotopal tilings are projections

[Pictures to be pasted in]

Figure 9. Two zonotopal tilings of a 10-gon

However, there is a very similar combinatorial structure called an oriented matroid which covers general zonotopal tilings. As we show in Section 6, Theorems 1 and 2 carry over to oriented matroids. The present example may thus motivate the extra work.

Recall that a face of a hyperplane arrangement $\{H_i\}_{i=1}^n$ can be coded as a sequence of $\{0, \pm\}$ symbols of length n . These are the signs σ in $F = \bigcap H_i^{\sigma_i}$. An *oriented matroid* is a set $\mathcal{X} \subseteq \{0, \pm\}^n$ satisfying

- (1) $0 \in \mathcal{X}$.
- (2) If $x \in \mathcal{X}$ then $-x \in \mathcal{X}$.
- (3) If $x, y \in \mathcal{X}$ then $x \cdot y \in \mathcal{X}$, where $x \cdot y = \begin{cases} x_i & x_i \neq 0 \\ y_i & x_i = 0. \end{cases}$
- (4) Given $x, y \in \mathcal{X}$, let $S(x, y) = \{i : x_i = -y_i \neq 0\}$. For every $j \in S(x, y)$ there is a $z \in \mathcal{X}$ with $z_j = 0$ and $z_i = (x \cdot y)_i = (y \cdot x)_i$ for $i \notin S(x, y)$.

It is not hard to check that the sign sequences of a central hyperplane arrangement form an oriented matroid. There is also a concept of *affine oriented matroid*, generalizing affine hyperplane arrangements, cf. [9], §4.5.

A *chamber* of an oriented matroid \mathcal{X} is an element $y \in \mathcal{X}$ with no zero coordinates. (We may assume without loss of generality that such elements exist.) Note that if y is a chamber and $x \in \mathcal{X}$ then $x \cdot y$ is a chamber. Thus if $w(\cdot)$ is a probability distribution on \mathcal{X} we may define a Markov chain $K(x, y)$ on the chambers of \mathcal{X} via

$$(3.2) \quad K(x, y) = \sum_{z \cdot x = y} w(z).$$

Section 6 shows that Theorems 1 and 2 hold for this chain.

Return now to zonotopal tilings. According to the Bohne–Dress theorem (see [38], 7.32), a zonotopal tiling of a fixed zonotope corresponds to a certain affine oriented matroid. The correspondence is such that the walk on the matroid yields the walk described above on the vertices of the tiling. We are thus in a position to find the eigenvalues and rates of convergence.

We conclude this section with an example of a family of tilings of a $2n$ -gon where all the details can be carried through. This example was suggested by Louis Billera. The family is built up by repeatedly adding a new layer of tiles as in the following

pictures.

Z_2 Z_3 Z_4 Z_5

[pictures to be pasted in]

n	2	3	4	5
t_n	1	3	6	10
v_n	4	7	11	16
e_n	4	9	16	25

At stage n , a new family of $n - 1$ rhombus tiles is added on the top and left of the previous Z_{n-1} . This results in a rhombic tiling Z_n of a $2n$ -gon with $t_n = \binom{n}{2}$ tiles, $v_n = \binom{n+1}{2} + 1$ vertices, and $e_n = n^2$ edges. These formulae are easily proved by induction.

These tilings are all projections of a three-dimensional zonotope. Nevertheless, we will present the associated affine oriented matroid, which is here realized by an affine hyperplane arrangement, and use it to analyze the walk. The hyperplane arrangement \mathcal{A}_n associated to the tiling Z_n may be represented as n lines in \mathbb{R}^2 in general position. Thus the examples above correspond to

n 2 3 4 5

Here the chambers of \mathcal{A}_n correspond to vertices of the tiling Z_n ; the segments of lines in \mathcal{A}_n correspond to edges of Z_n , with the half infinite line segments corresponding to the $2n$ bounding edges of Z_n ; and the vertices of \mathcal{A}_n correspond to the two-dimensional tiles of Z_n .

The correspondence may be seen here directly (without the Bohne–Dress theorem) by drawing the lines on top of the picture of Z_n as follows:

The dotted segments (pseudo-lines) can be straightened out to be the lines shown. See [38], §7.5, for further details.

We will analyze the walk on the vertices of Z_n driven by $\binom{n}{2}$ weights on the 2-dimensional tiles. This corresponds to a walk on the chambers of \mathcal{A}_n driven by weights on the $\binom{n}{2}$ vertices. Each vertex corresponds to the intersection of two lines. Labeling the lines L_1, \dots, L_n , let w_{ij} be the associated weights. Suppose for simplicity that all $w_{ij} > 0$. Then the weights are separating and so the walk has a unique stationary distribution described in Theorem 2. Here, it can be shown that the stationary distribution is supported on chambers of \mathcal{A}_n that meet the convex hull of the vertices of \mathcal{A}_n .

Theorem 1 shows that the eigenvalues and their multiplicities are

λ	$m(\lambda)$
1	1
λ_{L_i}	1 $1 \leq i \leq n$
w_{ij}	1 $i \neq j$

with $\lambda_{L_i} = \sum_{j \neq i} w_{ij}$. This gives $1 + n + \binom{n}{2} = \binom{n+1}{2} + 1$ eigenvalues.

Theorem 2 shows

$$\|K_x^\ell - \pi\| \leq \sum_{i=1}^n \lambda_{L_i}^\ell.$$

For example, if $w_{ij} \equiv \frac{1}{\binom{n}{2}}$, then $\lambda_{L_i} = 2/n$ and the bound becomes

$$\|K_x^\ell - \pi\| \leq n \left(\frac{2}{n}\right)^\ell.$$

Thus, for large n , two steps suffice for convergence.

As a second example, let $w_{i,i+1} = 1/(n-1)$ for $1 \leq i \leq n-1$, and $w_{ij} = 0$ otherwise. These weights are separating for $n \geq 4$ and the bound in Theorem 2 becomes

$$\|K_x^\ell - \pi\| \leq 2 \left(\frac{1}{n-1}\right)^\ell + (n-2) \left(\frac{2}{n-1}\right)^\ell.$$

Here again, the walk converges after $\ell = 2$ steps.

3F. Further examples. There are many further examples of hyperplane arrangements where the chambers can be indexed by a natural class of combinatorial objects, cf. Orlik–Terao [27], Stanley [33], and Ziegler [38]. We briefly mention two: (a) The set of all regular cubical tilings of a fixed zonotope and (b) the Shi arrangement.

A walk on tilings. Let $Z = [-v_1, v_1] + \cdots + [-v_n, v_n]$ be a d -dimensional zonotope in \mathbb{R}^d with n zones. A zonotopal tiling of Z is called *cubical* if every face of the tiling is a translate of a parallelepiped of the form $\sum_{i \in s} [-v_i, v_i]$, where s is a subset of $[n] = \{1, \dots, n\}$ such that $\{v_i\}_{i \in s}$ is linearly independent. A zonotopal tiling is *regular* if it arises as an appropriate projection of a $(d+1)$ -dimensional zonotope. (See [38] or [9] for the precise definition.) For example, the left tiling in Figure 9 is regular, but the right one is not. Billera and Sturmfels [7] have proved that the set of all regular cubical tilings of a fixed zonotope Z is itself the set of vertices of a second zonotope \widehat{Z} . Thus the walk on the vertices of the zonotope \widehat{Z} becomes a walk on the set of regular cubical tilings of Z .

To define \widehat{Z} , consider the subsets $s \subseteq [n]$ such that $\{v_i : i \in s\}$ is a minimal linearly dependent set. Each such s gives rise to a linear relation $\sum_{i=1}^n \alpha_i v_i = 0$, unique up to scalar multiplication. Let $\alpha_s = (\alpha_1, \dots, \alpha_n)$. Then \widehat{Z} is the zonotope in \mathbb{R}^n defined by

$$\widehat{Z} = \sum_s [-\alpha_s, \alpha_s].$$

It has one zone for each minimal dependent subset of $\{v_i\}$. Note that \widehat{Z} is $(n-d)$ -dimensional, because the linear span of the α_s is the kernel of the linear surjection $\mathbb{R}^n \rightarrow \mathbb{R}^d$ given by $e_i \mapsto v_i$.

One can give a more explicit formula for \widehat{Z} by noting that each minimal dependent subset of $\{v_i\}$ can be extended to a set of $d+1$ vectors that span \mathbb{R}^d . Writing these as the columns of a matrix, we can find the essentially unique linear relation among them by forming the cross product of the rows of the matrix (which form a set of d independent vectors in \mathbb{R}^{d+1}). This leads to the equivalent definition

$$\widehat{Z} = \sum_s [-\beta_s, \beta_s],$$

where now s ranges over the $(d+1)$ -subsets $s_1 < \cdots < s_{d+1}$ of $[n]$ and

$$\beta_s = \sum_{i=1}^{d+1} (-1)^i \det(v_{s_1}, \dots, v_{s_{i-1}}, v_{s_{i+1}}, \dots, v_{s_{d+1}}) e_{s_i}.$$

Note that some of the β_s may be 0 and some of the nonzero ones may be scalar multiples of others. So the number of zones of \widehat{Z} may be less than $\binom{n}{d+1}$.

As an example, there are 8 regular cubical tilings of an octagon (Figure 10). The zonotope \widehat{Z} is then itself an octagon, as shown. More generally, consider all regular cubical tilings of a regular $2n$ -gon. The bounding zonotope Z is generated by n vectors in \mathbb{R}^2 . The zonotope \widehat{Z} is $(n-2)$ -dimensional, generated by $\binom{n}{3}$ vectors α_s in \mathbb{R}^n . To go further, one would have to understand the geometry of \widehat{Z} , identifying natural families of faces and thus walks on the regular tilings.

The papers [37] and [21] study *all* cubical tilings of a $2n$ -gon and connect these to the higher Bruhat orders of Manin and Schechtman. For $n \leq 5$ all such tilings are regular, so that Fig. 3 of [37] gives a picture of the regular cubical tilings of a 10-gon.

Shi arrangement. The Shi arrangement in \mathbb{R}^n consists of the $2 \cdot \binom{n}{2}$ hyperplanes given by $\alpha_i - \alpha_j = 0$ and $\alpha_i - \alpha_j = 1$, $i < j$. This is one of several “deformations”

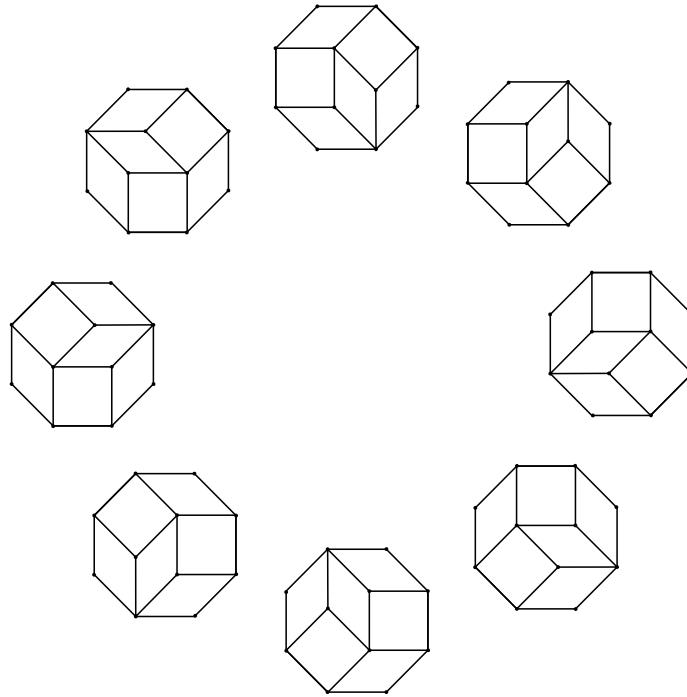


Figure 10. An octagon of tilings of an octagon.

of the braid arrangement discussed in [33]. It is known that there are $(n + 1)^{n-1}$ chambers, which may be put into bijective correspondence with the set of labeled trees on $n + 1$ vertices as well as with the set of “parking functions” (see [33] and further references cited there). There are many such bijections. Unfortunately, despite extensive effort, we have not found one such that the natural walk on the chambers seems natural when viewed as a walk on labeled trees.

4. STATIONARY DISTRIBUTION AND BOUNDS

In this section we prove Theorem 2 and some extensions. The argument handles input from a stationary process on \mathcal{F} , not just i.i.d. input. We begin with some general observations that use no special properties of the hyperplane situation.

4A. Iteration of random mappings. Let \mathcal{F} be an arbitrary finite semigroup and let \mathcal{C} be a finite set on which \mathcal{F} operates. Thus we have a function $\mathcal{F} \times \mathcal{C} \rightarrow \mathcal{C}$, denoted $(f, c) \mapsto fc$, such that $f(gc) = (fg)c$ for $f, g \in \mathcal{F}$, $c \in \mathcal{C}$. Note that this setup encompasses an arbitrary family of mappings $\mathcal{C} \rightarrow \mathcal{C}$, since we could simply take \mathcal{F} to be the semigroup that they generate under composition.

Let $\dots, F_{-2}, F_{-1}, F_0, F_1, F_2, \dots$ be a stationary \mathcal{F} -valued process, e.g., an i.i.d. sequence. (Recall that “stationary” means that the distribution of the sequence is shift-invariant.) We can use this to define a process $(C_\ell)_{\ell \geq 0}$ (not necessarily Markov) on \mathcal{C} : Fix a starting point C_0 and set

$$(4.1) \quad C_\ell = F_{-\ell} \cdots F_{-2} F_{-1} C_0$$

Our description of the limiting distribution for C_ℓ will make use of the infinite product

$$\prod_{i=1}^{\infty} F_i = \lim_{\ell \rightarrow \infty} F_1 \cdots F_\ell$$

where, in this discrete setting, a sequence is said to converge if and only if it is eventually constant. In order for this to make sense, we assume (F_i) satisfies

$$(IP) \quad \text{The infinite product } \prod_{i=1}^{\infty} F_i \text{ exists almost surely.}$$

Example. If \mathcal{F} is the face semigroup of a hyperplane arrangement, then the sequence of partial products $F_1 \cdots F_\ell$ is increasing with respect to the face relation, so (IP) holds.

Let $F_1^\infty = \prod_{i=1}^{\infty} F_i$ and $F_1^\ell = \prod_{i=1}^{\ell} F_i$. Our proof of Theorem 2 will be based on the following simple observation:

Theorem 3. *Assume the stationary sequence $(F_i)_{i \in \mathbb{Z}}$ satisfies (IP). Fix $C_0 \in \mathcal{C}$, let C_ℓ be defined by (4.1) for $\ell \geq 1$, and let π_ℓ be the distribution of C_ℓ , i.e., $\pi_\ell(c) = P\{C_\ell = c\}$. Let π be the distribution of $F_1^\infty C_0$. Then $\pi_\ell \rightarrow \pi$ as $\ell \rightarrow \infty$. More precisely,*

$$(4.2) \quad \|\pi_\ell - \pi\|_{\text{TV}} \leq P\{F_1^\ell \neq F_1^\infty\}.$$

Proof. By stationarity, C_ℓ has the same distribution as $F_1^\ell C_0$. Since $F_1^\ell C_0 \rightarrow F_1^\infty C_0$ a.s., it follows that $\pi_\ell \rightarrow \pi$. To prove (4.2), recall that, by definition of total variation distance,

$$\|\pi_\ell - \pi\| = \max_{\mathcal{D} \subseteq \mathcal{C}} |\pi_\ell(\mathcal{D}) - \pi(\mathcal{D})|.$$

We have $\pi_\ell(\mathcal{D}) = P\{F_1^\ell C_0 \in \mathcal{D}\}$ and $\pi(\mathcal{D}) = P\{F_1^\infty C_0 \in \mathcal{D}\}$. Break up both events according to whether or not $F_1^\ell = F_1^\infty$:

$$\begin{aligned} \pi_\ell(\mathcal{D}) &= P\{F_1^\ell = F_1^\infty, F_1^\ell C_0 \in \mathcal{D}\} + P\{F_1^\ell \neq F_1^\infty, F_1^\ell C_0 \in \mathcal{D}\} \\ \pi(\mathcal{D}) &= P\{F_1^\ell = F_1^\infty, F_1^\infty C_0 \in \mathcal{D}\} + P\{F_1^\ell \neq F_1^\infty, F_1^\infty C_0 \in \mathcal{D}\}. \end{aligned}$$

The two first terms are equal and the two second terms are at most $P\{F_1^\ell \neq F_1^\infty\}$. \square

Remark. The infinite product F_1^∞ acts on \mathcal{C} as an infinite composite $f_1 \circ f_2 \circ \cdots$ of random maps. The idea of using such infinite composites to construct limiting distributions has occurred in a variety of contexts. See, for instance, Letac [25] and Chamayou–Letac [12]. In the setting of the hyperplane chamber walk, one can even verify that the Propp–Wilson [34] monotonicity condition holds with respect to the weak Bruhat order [9] on \mathcal{C} . This means that monotone coupling from the past can be used to draw exact samples from the stationary distribution π .

4B. Proof of Theorem 2. We now specialize to the case where \mathcal{F} is the set of faces and \mathcal{C} the set of chambers of a hyperplane arrangement. We still allow, for the moment, the process on chambers to be driven by a stationary process on \mathcal{F} as above. If we assume that F_1^∞ is a.s. a chamber, then $F_1^\ell = F_1^\infty$ as soon as F_1^ℓ is a chamber. The bound (4.2) therefore yields

$$(4.3) \quad \|\pi_\ell - \pi\| \leq P\{F_1 \cdots F_\ell \notin \mathcal{C}\}.$$

Assume from now on that $(F_i)_{i \in \mathbb{Z}}$ consists of i.i.d. picks from a measure w on \mathcal{F} , so that (4.1) is the hyperplane chamber walk. If w is separating, then F_1^∞ is a.s. a chamber and (4.3) holds. Now $F_1 \cdots F_\ell \notin \mathcal{C}$ if and only if there is a hyperplane $H \in \mathcal{A}$ such that $F_i \subseteq H$ for $1 \leq i \leq \ell$. And $P\{F_i \subseteq H\} = \lambda_H$. Hence

$$(4.4) \quad P\{F_1 \cdots F_\ell \notin \mathcal{C}\} \leq \sum_{H \in \mathcal{A}} P\{F_i \subseteq H \text{ for } 1 \leq i \leq \ell\} = \sum_{H \in \mathcal{A}} \lambda_H^\ell.$$

Combining this with (4.3), we get part (c) of Theorem 2.

Remark. BHR [6] give a more careful analysis of $P\{F_1 \cdots F_\ell \notin \mathcal{C}\}$. They break up the event $\{F_1 \cdots F_\ell \notin \mathcal{C}\}$ according to the support W of $F_1 \cdots F_\ell$ and then use Möbius inversion to get

$$(4.5) \quad P\{F_1 \cdots F_\ell \notin \mathcal{C}\} = - \sum_{\substack{W \in \mathcal{S} \\ W \neq V}} \mu_{\mathcal{S}}(W, V) \lambda_W^\ell.$$

Combining (4.3), (4.4), and (4.5), we obtain

$$(4.6) \quad \|K_C^\ell - \pi\| \leq - \sum_{W \neq V} \mu(W, V) \lambda_W^\ell \leq \sum_{H \in \mathcal{A}} \lambda_H^\ell$$

so that (4.5) gives, in principle, a sharper bound than that of Theorem 2, using all the eigenvalues. On the other hand, the right side of (4.5) seems quite difficult to estimate (without using the second inequality in (4.6)), and we have not found any examples where we could use it to get a better bound than that of Theorem 2.

Returning to the proof of Theorem 2, still assuming w is separating, Theorem 3 gives the following description of the limiting distribution of the chain started at C_0 : Sample with replacement from w , stopping as soon as $F_1 \cdots F_m$ is a chamber. Then $F_1 \cdots F_m$ is distributed from π . As we explained at the beginning of §3, this is equivalent to the description of π in part (b) of Theorem 2. Since π is independent of the starting chamber C_0 , we have also proven half of (a): There is a unique stationary distribution if w is separating.

Finally, suppose w is not separating, and let H contain the support of w . Then Theorem 3 gives the limiting distribution π_{C_0} for the chain started at C_0 , and we see that π_{C_0} is concentrated on the chambers on the C_0 -side of H . Hence there are at least two stationary distributions, and Theorem 2 is proved.

Remarks. 1. One can give more precise results for general (non-separating) w : Let \mathcal{A}' be the set of $H \in \mathcal{A}$ which contain the support of w . Partition the chambers C according to which \mathcal{A}' chamber contains C . Then this gives the decomposition

of the hyperplane chamber walk into its components: The walk started in any \mathcal{A}' -chamber stays there and has a unique stationary distribution. In particular, the set of all stationary distributions is a simplex with one vertex for each \mathcal{A}' -chamber.

2. The weighted random-to-top shuffle with Markovian input is a standard item of study in the recent literature. Refer to Phatarfod–Dyte [29] and Dobrow–Fill [20] for results in this case. Note that for a stationary Markov chain as input, the process is driven by the time-reversed chain according to (4.1). Turning things around, suppose we are interested in the process $C_0, F_1 C_0, F_2 F_1 C_0, \dots$ with $(F_i)_{i=1}^\infty$ a stationary Markov chain. Let $\tilde{F}_1, \tilde{F}_2, \dots$ be the time-reversed process. The bound then becomes

$$\|\pi_\ell - \pi\| \leq P\{\tilde{F}_1 \dots \tilde{F}_\ell \notin \mathcal{C}\}.$$

As before, $\tilde{F}_1 \dots \tilde{F}_\ell \notin \mathcal{C}$ if and only if there is a hyperplane $H \in \mathcal{A}$ containing $\tilde{F}_1, \dots, \tilde{F}_\ell$; the probability of this event can be bounded by the “cover time” of the chain (\tilde{F}_i) . See [1] for a review of the literature on cover times.

5. DIAGONALIZATION

In this section we prove Theorem 1. After setting up some notation in 5A, we prove that K is diagonalizable in 5B, assuming, for simplicity, that \mathcal{A} is central. The proof gives the eigenvalues but not the multiplicities (which were calculated by BHR [6] in the central case). In 5C we show how the ideas in 5B lead naturally to a new proof of the BHR formula for the multiplicities. Finally, we treat the non-central case in 5D.

5A. Notation. For any finite set S , let $\mathbb{R}S$ denote the vector space of all real linear combinations $\sum_{s \in S} \alpha(s)s$ of elements of S . In particular, we have vector spaces $\mathbb{R}\mathcal{C}$ and $\mathbb{R}\mathcal{F}$ generated by the chambers and faces of a hyperplane arrangement. Note that $\mathbb{R}\mathcal{F}$ is an \mathbb{R} -algebra (the semigroup algebra of \mathcal{F}), and $\mathbb{R}\mathcal{C}$ is an $\mathbb{R}\mathcal{F}$ -module via the action of faces on chambers. Given a probability measure w on \mathcal{F} , we have an element

$$(5.1) \quad T = T_w = \sum_{F \in \mathcal{F}} w(F)F$$

of $\mathbb{R}\mathcal{F}$, which therefore acts as an operator on $\mathbb{R}\mathcal{C}$. Explicitly, given an element $\alpha = \sum_{C \in \mathcal{C}} \alpha(C)C \in \mathbb{R}\mathcal{C}$, we have

$$T(\alpha) = \sum_{\substack{F \in \mathcal{F} \\ C \in \mathcal{C}}} w(F)\alpha(C)FC = \sum_{C' \in \mathcal{C}} \beta(C')C',$$

where

$$\beta(C') = \sum_{F, C: FC = C'} w(F)\alpha(C) = \sum_{C \in \mathcal{C}} \alpha(C)K(C, C').$$

Here K is the transition matrix defined at (1.3). Thus if elements of $\mathbb{R}\mathcal{C}$ are viewed as row vectors indexed by \mathcal{C} , then T acts as right multiplication by the matrix K . In particular, the eigenvectors of T on $\mathbb{R}\mathcal{C}$ are the left eigenvectors of K .

5B. Diagonalizability. We already have an eigenvector $\Sigma\pi(C)C$ with eigenvalue 1 ($= \lambda_V$), where π is any stationary distribution for the chamber walk defined by K .⁵ Note that this eigenvector maps to 1 $\in \mathbb{R}$ under the linear map

$$\partial_0 : \mathbb{R}\mathcal{C} \rightarrow \mathbb{R}$$

given by $\partial_0(C) = 1$ for all $C \in \mathcal{C}$. Now ∂_0 is a homomorphism of $\mathbb{R}\mathcal{F}$ -modules, where each $F \in \mathcal{F}$ acts as the identity on \mathbb{R} , so $\ker \partial_0$ is an $\mathbb{R}\mathcal{F}$ -module. In particular, $\ker \partial_0$ is T -invariant, so T (and hence K) will be diagonalizable provided its restriction to $\ker \partial_0$ is diagonalizable.

Note that $\ker \partial_0$ is spanned by the differences $C - C'$ ($C, C' \in \mathcal{C}$). In fact, by connectivity of the chamber graph, $\ker \partial_0$ is spanned by differences $C - C'$ such that C and C' are adjacent. Let $\mathcal{F}_1 \subset \mathcal{F}$ be the set of codimension 1 faces, i.e., the set of faces whose support is a hyperplane. Our first task is to define a linear surjection $\partial_1 : \mathbb{R}\mathcal{F}_1 \rightarrow \ker \partial_0$, which sends $A \in \mathcal{F}_1$ to $\partial_1(A) = \pm(C - C')$, where C and C' are the two chambers having A as a face; here we must specify a rule for determining the ambiguous sign.⁶ Assume, for simplicity, that \mathcal{A} is central, so that \mathcal{A} can be taken to consist of linear hyperplanes. We will return to the general case in §5D.

Choose arbitrarily an orientation for the ambient vector space V . This means that we have a rule which associates a sign $\varepsilon = \pm 1$ to each ordered basis e_1, \dots, e_n of V , in such a way that two ordered bases have the same (resp., opposite) sign if the matrix relating them has positive (resp., negative) determinant. Similarly, each hyperplane $H \in \mathcal{A}$ is itself a vector space and we choose arbitrarily an orientation for it. Given a chamber C and a codimension 1 face A of C , we use the chosen orientations on V and on $H = \text{supp } A$ to define a sign $[A : C] = \pm 1$, as follows: Choose a positively-oriented ordered basis e_1, \dots, e_{n-1} for H , choose $v \in C$, and set

$$[A : C] = \varepsilon(e_1, \dots, e_{n-1}, v).$$

This is easily seen to be independent of the choice of e_1, \dots, e_{n-1} for fixed v , and it is independent of v because C is connected. In fact, $[A : C] = \varepsilon(e_1, \dots, e_{n-1}, v)$ for any vector v on the same side of H as C .

Note that if $A \in \mathcal{F}_1$ and C, C' are the two chambers having A as a face, then $[A : C'] = -[A : C]$ because C and C' are on opposite sides of $H = \text{supp } A$. The desired surjection

$$\partial_1 : \mathbb{R}\mathcal{F}_1 \rightarrow \ker \partial_0,$$

is now defined by $\partial_1(A) = [A : C]C + [A : C']C'$ for $A \in \mathcal{F}_1$, where C, C' are the chambers having A as a face.

Next, we define an action of \mathcal{F} on $\mathbb{R}\mathcal{F}_1$. Given $F \in \mathcal{F}$ and $A \in \mathcal{F}_1$, we wish to define $F * A \in \mathbb{R}\mathcal{F}_1$. Consider the product FA in the semigroup \mathcal{F} . If $F \subseteq H = \text{supp } A$, then FA is again in \mathcal{F}_1 (and has the same support H), and we set $F * A = FA$. If $F \not\subseteq H$, then FA is a chamber and we set $F * A = 0$. This product makes $\mathbb{R}\mathcal{F}_1$ an $\mathbb{R}\mathcal{F}$ -module.

⁵Alternatively, we could get π from the Perron–Frobenius theorem instead of probability theory.

⁶For our present purposes, we could set $\partial_1(A) = C - C'$, where C is on the (arbitrarily chosen)

“right side” of H (see §4). But we will give a more direct argument that will be much simpler.

Lemma 1. *The map $\partial_1 : \mathbb{R}\mathcal{F}_1 \rightarrow \ker \partial_0$ is a homomorphism of $\mathbb{R}\mathcal{F}$ -modules. In particular, it commutes with the action of $T = \Sigma w(F)F$.*

Proof. Given $F \in \mathcal{F}$ and $A \in \mathcal{F}_1$, we must check that $\partial_1(F * A) = F\partial_1(A)$. Let C and C' be the chambers having A as a face. If $F \not\subseteq H$, then $FA = FC = FC'$, and we have $F\partial_1(A) = \pm F(C - C') = 0$. Since $F * A = 0$, the desired equation holds. If $F \subseteq H$, then FC and FC' are the two chambers having FA as a face, so $\partial_1(F * A) = \pm(FC - FC')$ and $F\partial_1(A) = \pm(FC - FC')$. The two ambiguous signs agree because FC and C are on the same side of H . \square

Finally, we lump together the elements of \mathcal{F}_1 with the same support to obtain

$$(5.2) \quad \mathbb{R}\mathcal{F}_1 = \bigoplus_{H \in \mathcal{A}} \mathbb{R}\mathcal{C}_H,$$

where \mathcal{C}_H is the set of faces having support H . The notation “ \mathcal{C}_H ” serves as a reminder that \mathcal{C}_H is the set of chambers of the arrangement \mathcal{A}_H in H obtained by restriction (see §2A). The decomposition (5.2) is a decomposition of $\mathbb{R}\mathcal{F}$ -modules, and the action of T on the summand $\mathbb{R}\mathcal{C}_H$ is the same type of operator, relative to the arrangement \mathcal{A}_H , as the original action of T on $\mathbb{R}\mathcal{C}$. [The relevant measure here on the faces of \mathcal{A}_H , which are simply the $F \in \mathcal{F}$ such that $F \subseteq H$, is gotten by restricting w . This restriction will not generally be a probability measure, but a trivial argument shows that the results we are trying to prove are true for all positive measures if and only if they are true for probability measures.]

Arguing by induction on the dimension of the ambient space V , it follows that T is diagonalizable on $\mathbb{R}\mathcal{F}_1$, hence also on the homomorphic image $\ker \partial_0$, hence also on $\mathbb{R}\mathcal{C}$. The argument also shows that the eigenvalues are contained in $\{\lambda_W\}_{W \in \mathcal{S}}$.

Remark. The proof gives explicit formulas for a set of eigenvectors spanning $\mathbb{R}\mathcal{C}$, provided such formulas are already known inductively for T acting on each $\mathbb{R}\mathcal{C}_H$. The details have been worked out by Bidigare, Denham, and Hanlon [private communication] and will appear elsewhere.

5C. Multiplicities. We continue to assume, for simplicity of notation, that \mathcal{A} is central. The argument above did not give us multiplicities for T acting on $\mathbb{R}\mathcal{C}$ because we ignored $\ker \partial_1$. In order to remedy this, we extend the exact sequence

$$\mathbb{R}\mathcal{F}_1 \xrightarrow{\partial_1} \mathbb{R}\mathcal{C} \xrightarrow{\partial_0} \mathbb{R} \rightarrow 0$$

to a longer exact sequence

$$(5.3) \quad \cdots \rightarrow \mathbb{R}\mathcal{F}_p \xrightarrow{\partial_p} \cdots \xrightarrow{\partial_2} \mathbb{R}\mathcal{F}_1 \xrightarrow{\partial_1} \mathbb{R}\mathcal{C} \xrightarrow{\partial_0} \mathbb{R} \rightarrow 0$$

(which is eventually 0 at the left), where \mathcal{F}_p is the set of faces of codimension p in V .

In order to define $\partial_p : \mathbb{R}\mathcal{F}_p \rightarrow \mathbb{R}\mathcal{F}_{p-1}$, we need numbers $[A : B] = \pm 1$ whenever A is a codimension 1 face of B . To this end we choose an orientation for each W in the intersection lattice \mathcal{S} . Then if we restrict \mathcal{A} to the support of B , the face B becomes a chamber, A becomes a face of codimension 1, and our chosen orientations give us a number $[A : B] = \pm 1$ by §5B applied to the restricted arrangement. We now define a linear map $\partial_p : \mathbb{R}\mathcal{F}_p \rightarrow \mathbb{R}\mathcal{F}_{p-1}$, by

$$\partial_p(A) = \sum_{B \triangleright A} [A : B]B \quad (A \in \mathcal{F}_p).$$

Here $B \triangleright A$ means that B covers A in the poset \mathcal{F} , i.e., A is a codimension 1 face of B . The following lemma implies that $\partial_p \circ \partial_{p-1} = 0$.

Lemma 2. *Let $A, C \in \mathcal{F}$, with $A < C$ and A of codimension 2 in C . Then there are exactly two faces B_1, B_2 in the open interval (A, C) , and we have*

$$[A : B_1][B_1 : C] = -[A : B_2][B_2 : C].$$

Proof. Replacing \mathcal{A} by its restriction to the support of C , we reduce to the case where C is a chamber. Let H_1, \dots, H_k be the walls of C , i.e., the supports of the codimension 1 faces. Let $H_i = \ker f_i$ ($f_i \in V^*$), where f_i is chosen so that $f_i > 0$ on C . Then C is defined by $f_i > 0$ for $1 \leq i \leq k$, and this is a minimal set of inequalities defining C ; moreover, the faces of C are the nonempty sets obtained by replacing some of these inequalities by equalities, cf. [11], §I.4B and last paragraph of §I.4A. The given face A of codimension 2 is contained in exactly two of the walls, say H_1, H_2 , because any three of the f_i are linearly independent and hence define a face of codimension 3. [A linear relation among three of the f_i would give one f_i as a positive linear combination of two others. But then the inequality $f_i > 0$ would be redundant in the description of C , and H_i would not be a wall.] It follows that the only faces between A and C are the faces B_1, B_2 of C supported by H_1, H_2 .

To prove $[A : B_1][B_1 : C] = -[A : B_2][B_2 : C]$, we may assume that H_i has been oriented so that $[A : B_i] = +1$. We may therefore get a positively-oriented ordered basis for H_i ($i = 1, 2$) by choosing a positively-oriented ordered basis e_1, \dots, e_{n-2} for $H_1 \cap H_2$ and adjoining a vector $v_i \in H_i$. Then $v_1 + v_2 \in C$, and $[B_i : C] = \varepsilon(e_1, \dots, e_{n-2}, v_i, v_1 + v_2)$. The lemma now follows from the fact that $e_1, \dots, e_{n-2}, v_1, v_1 + v_2$ and $e_1, \dots, e_{n-2}, v_2, v_1 + v_2$ are related by a matrix of determinant -1 . \square

Next we define an action of \mathcal{F} on $\mathbb{R}\mathcal{F}_p$ as in §5B: Given $F \in \mathcal{F}$ and $A \in \mathcal{F}_p$, set

$$F * A = \begin{cases} FA & \text{if } F \subseteq \text{supp } A \\ 0 & \text{otherwise.} \end{cases}$$

This makes $\mathbb{R}\mathcal{F}_p$ an $\mathbb{R}\mathcal{F}$ -module, which we may decompose according to supports:

$$\mathbb{R}\mathcal{F}_p = \bigoplus_{W \in \mathcal{S}_p} \mathbb{R}\mathcal{C}_W,$$

where $\mathcal{S}_p = \{W \in \mathcal{S} : \text{codim}(W, V) = p\}$ and \mathcal{C}_W is the set of faces with support W . The complex (5.3) now becomes

$$(5.4) \quad \cdots \rightarrow \bigoplus_{W \in \mathcal{S}_2} \mathbb{R}\mathcal{C}_W \rightarrow \bigoplus_{H \in \mathcal{A}} \mathbb{R}\mathcal{C}_H \rightarrow \mathbb{R}\mathcal{C} \rightarrow \mathbb{R} \rightarrow 0.$$

Lemma 3. *The complex (5.4) is a chain complex of $\mathbb{R}\mathcal{F}$ -modules, i.e., each boundary map commutes with the action of \mathcal{F} .*

Proof. Consider a typical component $\partial_{U,W} : \mathbb{R}\mathcal{C}_U \rightarrow \mathbb{R}\mathcal{C}_W$ of ∂_p , where $U \in \mathcal{S}_p$ and $W \in \mathcal{S}_{p-1}$. We must show that the action of each $F \in \mathcal{F}$ commutes with $\partial_{U,W}$. If $U \not\subseteq W$, then $\partial_{U,V} = 0$ and there is nothing to prove. If $F \not\subseteq W$, then F acts as 0 on both $\mathbb{R}\mathcal{C}_U$ and $\mathbb{R}\mathcal{C}_W$, so again there is nothing to prove. Finally, if $F \subseteq W$, we may replace \mathcal{A} by its restriction to W and apply Lemma 1 of §5B. \square

Lemma 4. *The sequence (5.4) is exact.*

Proof. This follows from the homology theory of regular cell-complexes [13]: Recall that the zonotope Z associated to \mathcal{A} (§2E) is a contractible regular cell-complex whose face poset is the poset \mathcal{F}^{op} opposite to \mathcal{F} . The facts we have proved about the numbers $[A : B]$ say precisely that they form a system of “incidence numbers” for Z in the sense of [13] and hence that (5.3) is the augmented cellular chain complex of Z . The exactness of (5.3) and (5.4) therefore follows from the fact that Z has trivial homology. \square

We are now in a position to calculate eigenvalues and multiplicities by the same sort of induction used in §5B. We need the following well-known special case of the “Euler characteristic principle”:

Lemma 5. *Let*

$$0 \longrightarrow V_m \xrightarrow{\partial_m} \cdots \longrightarrow V_1 \xrightarrow{\partial_1} V_0 \rightarrow 0$$

be an exact sequence of finite-dimensional vector spaces. Let T_i be a linear operator on V_i such that $\partial_i T_i = T_{i-1} \partial_i$ for $1 \leq i \leq m$, and let φ_i be the characteristic polynomial of T_i . Then $\varphi_0 \varphi_1^{-1} \varphi_2 \varphi_3^{-1} \cdots = 1$.

Sketch of proof. For $m = 2$ this is proved by looking at matrices; the general case follows by induction on m . \square

Applying this to the action of $T = \Sigma w(F)F$ on (5.4), we may assume inductively that we have a decomposition of the characteristic polynomial φ_U of T acting on $\mathbb{R}\mathcal{C}_U$ for each $U \neq V$ in \mathcal{S} , say

$$\varphi_U(\lambda) = \prod_{W \in \mathcal{S}: W \subseteq U} (\lambda - \lambda_W)^{m(W, U)}$$

for some integers $m(W, U) \geq 0$. Lemmas 4 and 5 then give the characteristic polynomial $\varphi = \varphi_V$ for T acting on $\mathbb{R}\mathcal{C}$:

$$\begin{aligned} \varphi(\lambda) &= (\lambda - \lambda_V) \prod_{\substack{W, U: \\ W \subseteq U \subsetneq V}} (\lambda - \lambda_W)^{-(-1)^{\text{codim}(U, V)} m(W, U)} \\ &= \prod_{W \in \mathcal{S}} (\lambda - \lambda_W)^{m(W, V)}, \end{aligned}$$

where $m(V, V) = 1$ and, for $W \subsetneq V$,

$$(5.5) \quad m(W, V) = - \sum_{U: W \subseteq U \subsetneq V} (-1)^{\text{codim}(U, V)} m(W, U).$$

This recurrence formula is reminiscent of the recurrence (2.1) for the Möbius function $\mu = \mu_{\mathcal{S}}$. In fact, if we multiply (5.5) by $(-1)^{\text{codim}(W, V)}$ and compare the result to (2.1), we conclude, inductively, that

$$m(W, V) = (-1)^{\text{codim}(W, V)} \mu(W, V) = |\mu(W, V)|,$$

where the second equality is the “alternating sign” property of $\mu_{\mathcal{S}}$ ([31], 3.10.1, or [27], 2.4.7). Theorem 1 is now completely proved in the central case.

Remark. As a byproduct, we recover Zaslavski’s famous formula [35] for the number of chambers:

$$(5.6) \quad |\mathcal{C}| = \sum_{W \in \mathcal{S}} m(W, V) = \sum_{W \in \mathcal{S}} |\mu(W, V)|.$$

5D. The non-central case. There are two ways to treat the case where \mathcal{A} is not necessarily central.

Method 1. Go through §§5B and 5C, making changes to accommodate the general case. This involves minor changes in wording (e.g., one has to consider orientations of affine subspaces), and, more importantly, one has to replace the zonotope Z by the cell-complex Δ of §2F.

Method 2. Deduce the general case from the central case by coning (§2F). Namely, let $\overline{\mathcal{A}}$ be the central arrangement obtained from \mathcal{A} by coning, and let $\overline{\mathcal{F}}$, $\overline{\mathcal{C}}$, and $\overline{\mathcal{S}}$ be the associated face poset, set of chambers, and intersection lattice. Recall that \mathcal{F} may be identified with a subset $\overline{\mathcal{F}}^+ \subset \overline{\mathcal{F}}$. We then have $\mathbb{R}\mathcal{C} \subset \mathbb{R}\overline{\mathcal{C}}$, and if we view w as a measure on $\overline{\mathcal{F}}$ supported in $\overline{\mathcal{F}}^+$, we get an operator \overline{T} on $\mathbb{R}\overline{\mathcal{C}}$ extending T and mapping all of $\mathbb{R}\overline{\mathcal{C}}$ into $\mathbb{R}\mathcal{C}$.

Diagonalizability of T therefore follows from that of \overline{T} , and we get the same eigenvalues and multiplicities for T as for \overline{T} , except that the multiplicity of 0 has to be reduced by $|\overline{\mathcal{C}}| - |\mathcal{C}| = |\overline{\mathcal{C}}|/2$. To complete the proof we need to show that this reduction is accomplished by throwing out the eigenvalues of \overline{T} corresponding to the $W \in \overline{\mathcal{S}}$ such that $W \subseteq H_0 = V \times 0$. [These W are the elements of $\overline{\mathcal{S}}$ that do not correspond to anything in \mathcal{S} .] Note that if $\overline{W} \not\subseteq H_0$ in $\overline{\mathcal{S}}$ and W is the corresponding element of \mathcal{S} (gotten by intersecting with $V \times 1$), then $\mu_{\overline{\mathcal{S}}}(\overline{W}, \overline{V}) = \mu_{\mathcal{S}}(W, V)$ because the corresponding open intervals are isomorphic. What we need to complete the proof is:

$$\textbf{Lemma 6.} \sum_{\substack{W \in \overline{\mathcal{S}}: \\ W \subseteq H_0}} (-1)^{\text{codim}(W, \overline{V})} \mu_{\overline{\mathcal{S}}}(W, \overline{V}) = \sum_{\substack{U \in \overline{\mathcal{S}}: \\ U \not\subseteq H_0}} (-1)^{\text{codim}(U, \overline{V})} \mu_{\overline{\mathcal{S}}}(U, \overline{V}).$$

[Recall that the sum of the left and right sides is $|\overline{\mathcal{C}}|$ by (5.6), so this shows that the left side equals $|\overline{\mathcal{C}}|/2$, as desired.]

Proof. Fix $W \subseteq H_0$ in $\overline{\mathcal{S}}$ and note that

$$\mu(W, \overline{V}) = - \sum_{\substack{U \in \overline{\mathcal{S}}: \\ U \not\subseteq H_0, U \cap H_0 = W}} \mu(U, \overline{V})$$

by the Crapo complementation formula (see [14], Theorem 3, [8], 6.2, or [27], 2.40). Multiply by $(-1)^{\text{codim}(W, \overline{V})}$ and sum over W to obtain the lemma. \square

6. GENERALIZATION TO ORIENTED MATROIDS

We introduced oriented matroids in §3E, as generalizations of central hyperplane arrangements. There are also affine oriented matroids, which generalize arbitrary (affine) hyperplane arrangements. In this section we indicate how to extend Theorems 1 and 2 to the oriented matroid setting. We begin by introducing the relevant terminology in 6A. The extensions of Theorems 1 and 2 are then treated in 6B.

6A. Oriented matroids. Our basic reference here is [9], especially Chapter 4. See also [38]. Our treatment closely follows these sources, with one exception: We use the same geometric language (chambers, faces, intersection lattice, ...) in connection with oriented matroids that we used for hyperplane arrangements.

Recall from §3E that an oriented matroid is a set \mathcal{X} of sign sequences $(x_i)_{i \in I}$ satisfying 4 axioms. Here I is a finite index set. As an aid to the intuition, let \mathcal{F} be a set whose elements F (called faces) are in 1–1 correspondence with the sign sequences in \mathcal{X} . We denote the correspondence by $F \mapsto \sigma(F) = (\sigma_i(F))_{i \in I}$.

The oriented matroids arising from (central) hyperplane arrangements are said to be *realizable*. There are non-realizable oriented matroids, but it can be shown that all oriented matroids can be realized by “pseudo-hyperplane arrangements”; these are topological analogues of hyperplane arrangements, in which the “pseudo-hyperplanes” are not necessarily flat. See [9], Chapter 5.

The set \mathcal{F} of faces of an oriented matroid is a poset under the “face relation” defined in terms of sign sequences as in §2B. And axiom (3) for oriented matroids gives \mathcal{F} the semigroup structure that has played such an important role in the construction and study of the hyperplane chamber walk.

Each face F has a *support*, determined by the *zero set* $z(F) = \{i \in I : \sigma_i(F) = 0\}$. The set \mathcal{S} of all supports is a lattice in a natural way, which we call the *intersection lattice*. For any $W \in \mathcal{S}$, we write $z(W)$ for the zero set of any face F with support W . We denote by V the largest element of \mathcal{S} . This is the support of any maximal element $C \in \mathcal{F}$. These maximal elements are called *chambers*, and the set of all of them is denoted \mathcal{C} . For any $W \in \mathcal{S}$, the set of $F \in \mathcal{F}$ with support W is again the set of chambers of an oriented matroid \mathcal{X}_W , said to be obtained by *restriction* to W . Its face poset is $\mathcal{F}_W = \{F \in \mathcal{F} : \text{supp } F \leq W\}$.

The *rank* of an oriented matroid is the length of the interval $[0, V]$ in \mathcal{S} . (For a realizable oriented matroid associated to an essential central hyperplane arrangement, the rank is simply $\dim V$.) The length of the interval $[W, V]$ is the *codimension* of W ; it is equal to $\text{rank } \mathcal{X} - \text{rank } \mathcal{X}_W$. It is also the rank of an oriented matroid \mathcal{X}^W whose sign sequences are gotten from $\{\sigma(F) : F \in \mathcal{F}\}$ by considering only the components σ_i with $i \in z(W)$. The face poset of \mathcal{X}^W is isomorphic to $\mathcal{F}_{\geq F}$ for any F with support W . (In the realizable case, \mathcal{X}^W corresponds to the subarrangement $\mathcal{A}^W \subseteq \mathcal{A}$ given by the hyperplanes which contain W . This subarrangement corresponds to an essential arrangement in the quotient space V/W .)

The face poset \mathcal{F} of \mathcal{X} can be identified with the poset of cells (including the empty cell) of a regular cell-complex $\Sigma = \Sigma_{\mathcal{X}}$ homeomorphic to the sphere S^{n-1} ($n = \text{rank } \mathcal{X}$) by Theorem 4.3.3 of [9]. There is also a “dual” cell-complex Δ , which is topologically an n -ball, whose poset of nonempty cells is isomorphic to \mathcal{F}^{op} ([9], 4.3.4). This plays the role of the zonotope dual to a central hyperplane arrangement.

Finally, we have already remarked that there is a concept of *affine oriented matroid*. We omit the formal definition, which is discussed in [9], §4.5. Suffice it to say that the definition is cooked up so that the analogue of the coning construction (§2F) remains valid, reducing the affine case to the case of ordinary oriented matroids. Once again, \mathcal{F}^{op} is isomorphic in the affine case to the poset of cells of a regular cell-decomposition of a ball ([9], 4.5.8), generalizing the complex Δ dual to an affine hyperplane arrangement (§2F).

6B. Extension of Theorems 1 and 2. It is obvious how to formulate analogues of Theorems 1 and 2 for oriented matroids (including the affine case). The condition that the measure w in Theorem 2 be separating, for example, becomes: There is no $W < V$ in \mathcal{S} such that w is concentrated on \mathcal{F}_W . It is also a fairly routine matter to check that our proofs extend to the matroid setting, except for one technical

point: Our definition of the incidence numbers $[A : B]$ in §§5B and 5C made use of orientations of real vector spaces. In order to carry this over to oriented matroids, we could simply appeal to the topological representation theorem quoted above; topology provides the appropriate concept of orientation for a “pseudo-subspace” W . We prefer, however, the following approach, which is longer but does not rely on the topological representation theorem.

We wish to define the notion of orientation for an element $W \in \mathcal{S}$. It suffices to consider $W = V$, since this then applies to arbitrary W by the restriction operation described above. By an *orientation* for V we will mean a rule that associates to each maximal chain

$$0 = A_0 < \cdots < A_n$$

in \mathcal{F} a sign $\varepsilon = \pm 1$, in such a way that adjacent maximal chains get opposite signs. Here two distinct maximal chains are *adjacent* if they differ in exactly one position. We will also say in this situation that one maximal chain is obtained from the other by an *elementary move*.

Remark. The maximal chains in \mathcal{F} are in 1–1 correspondence with the maximal simplices in the simplicial complex $|\mathcal{F}_{>0}|$ associated to the poset $\mathcal{F}_{>0}$.⁷ Now $|\mathcal{F}_{>0}|$ is the barycentric subdivision of the cell-complex Σ discussed above; in particular, $|\mathcal{F}_{>0}|$ is topologically a sphere. And adjacency of maximal chains, as defined above, agrees with the usual notion of adjacency for triangulated manifolds: Two distinct maximal simplices are adjacent if and only if they have a common codimension 1 face. It follows that V admits an orientation, unique up to multiplication by -1 .

Underlying our definition of $[A : C]$ in §5B was the fact that if a vector space V and a hyperplane H are both oriented, then there is a canonical choice of “positive side” of H in V . The following technical lemma, though it does not mention orientations, will be crucial for generalizing this idea to oriented matroids.

Lemma 1. *Let \mathcal{X} be an oriented matroid of rank $n \geq 2$. Let $W \in \mathcal{S}$ have codimension 1 and choose $i \in z(W) - z(V)$. Let $0 = A_0 < \cdots < A_n$ be a maximal chain in \mathcal{F} with $\text{supp } A_{n-1} = W$. Let $0 = A'_0 < \cdots < A'_{n-1}$ be a maximal chain in \mathcal{F}_W adjacent to $A_0 < \cdots < A_{n-1}$. Let A'_n be a chamber such that $A'_n > A'_{n-1}$. Then the chain $A'_0 < \cdots < A'_n$ can be obtained from $A_0 < \cdots < A_n$ by an odd number of elementary moves if $\sigma_i(A'_n) = \sigma_i(A_n)$ and by an even number otherwise.*

[Note: In the realizable case, W is the hyperplane H_i , and the issue here is whether or not the chambers A_n and A'_n are on the same side of H_i .]

Proof. If $A_0 < \cdots < A_{n-1}$ and $A'_0 < \cdots < A'_{n-1}$ differ at position $i < n-1$, the lemma is obvious. So assume we are in the situation

$$X \dashv Y$$

$$0 = A_0 \dashv A_1 \dashv \cdots \dashv A_{n-2}$$

$$X' \dashv Y'$$

with $\text{supp } X = \text{supp } X' = W$. Replacing \mathcal{F} by $\mathcal{F}_{\geq A_{n-2}}$, which is again the face poset of an oriented matroid as noted above, we reduce to the case $n = 2$. Now

⁷Recall that any finite poset P gives rise to a simplicial complex $|P|$, whose vertices are the elements of P and whose simplices are the chains in P .

every rank 2 oriented matroid is realizable ([9], p. 248). So we may identify \mathcal{F} with the poset of faces (including the empty face) of a $2m$ -gon Σ for some $m \geq 2$. Then X and X' correspond to opposite vertices and Y and Y' can be assumed to be opposite edges [this is the case $\sigma_i(Y) \neq \sigma_i(Y')$]. The maximal chains we have been considering correspond to the edges of the barycentric subdivision of Σ , and the lemma then states that in a $4m$ -gon, the gallery distance between opposite edges is even. (In fact, it is $2m$.) See Figure 11 below for the case $m = 3$. \square

Figure 11

We now proceed to define incidence numbers $[A : B]$ as in §§5B and 5C. Choose an orientation for each $W \in \mathcal{S}$. This means we can attach a sign $\varepsilon = \pm 1$ to any chain in \mathcal{F} of the form

$$0 = A_0 \lessdot A_1 \lessdot \cdots \lessdot A_r,$$

in such a way that the sign changes if an elementary move is performed. Here an elementary move changes exactly one A_i , keeping the support the same if $i = r$. Note that for each $i > 0$ there is exactly one candidate for an A'_i to which A_i can be changed ([9], 4.1.14(ii)).

Given $A \lessdot B$ in \mathcal{F} , define the *incidence number* $[A : B]$ by taking a chain $0 = A_0 \lessdot \cdots \lessdot A_r = A$ and setting $[A : B] = \pm 1$ depending on whether the augmented chain $0 = A_0 \lessdot \cdots \lessdot A_r \lessdot B$ has the same sign as the original one or the opposite sign. This is independent of the choice of the chain from 0 to A because any two can be related by a sequence of elementary moves,⁸ which affect the signs of the augmented chains in the same way they affect the signs of the original chains.

Note that if A has codimension 1 and C, C' are the chambers having A as a face, then $[A : C] = -[A : C']$. More generally, $[A : B] = -[A : B']$ if both are

⁸Because the barycentric subdivision of a $2m$ -gon is related to the one of a $4m$ -gon in a similar way as the one of a $2m$ -gon is related to the one of a $4m$ -gon.

defined, $B \neq B'$, and $\text{supp } B = \text{supp } B'$. Note also that the “diamond condition” for incidence numbers is satisfied as in Lemma 2 of §5C: In the situation

$$\begin{array}{ccc} & C & \\ B_1 & & B_2 \\ & A & \end{array}$$

we have $[A : B_1][B_1 : C] = -[A : B_2][B_2 : C]$. In fact, if we multiply both sides of this equation by the sign of any chain $0 = A_0 \lessdot \dots \lessdot A_r = A$, the resulting equation follows immediately from the definitions.

Lemma 2. *Let $W \in \mathcal{S}$ have codimension 1, and choose $i \in z(W) - z(V)$. Then either $[A : C] = \sigma_i(C)$ for all A with support W and all chambers $C > A$, or else $[A : C] = -\sigma_i(C)$ for all such A, C .*

Proof. We may assume $[A : C] = \sigma_i(C)$ for one pair A, C . Let A', C' be another. If $A = A'$ and $C \neq C'$, we know $[A : C'] = -[A : C] = -\sigma_i(C) = \sigma_i(C)$. [For the last equality, recall that $\mathcal{F}_{\geq A}$ is the face poset of a rank 1 oriented matroid, whose two nonzero faces are opposite one another.] So we may assume $A \neq A'$ and hence $\text{rank } \mathcal{X} \geq 2$. Choose maximal chains

$$\begin{aligned} 0 = A_0 < \dots < A_{n-1} = A \\ 0 = A'_0 < \dots < A'_{n-1} = A' \end{aligned}$$

in \mathcal{F}_W . If $d \geq 1$ elementary moves in \mathcal{F}_W change (A_i) to (A'_i) , then Lemma 1 implies that we can get from $A_0 < \dots < A_{n-1} < C$ to $A'_0 < \dots < A'_{n-1} < C'$ in d' moves, where $d \equiv d' \pmod{2}$ if and only if $\sigma_i(C) = \sigma_i(C')$. Suppose, for instance, that $\sigma_i(C) = \sigma_i(C')$. Then we have

$$\begin{aligned} \varepsilon(A_0 < \dots < A_{n-1})[A : C] &= (-1)^d \varepsilon(A'_0 < \dots < A'_{n-1})[A' : C'] \\ \varepsilon(A_0 < \dots < A_{n-1}) &= (-1)^d \varepsilon(A'_0 < \dots < A'_{n-1}), \end{aligned}$$

hence $[A' : C'] = [A : C] = \sigma_i(C) = \sigma_i(C')$, as required. \square

Armed with Lemma 2, the interested reader can now carry out for oriented matroids the arguments in §§5B and 5C. Finally, affine oriented matroids can be treated as in §5D.

APPENDIX

We construct here the convex polytope $\widehat{\Sigma}$ mentioned in §2D. Recall the setup: \mathcal{A} is an essential, central hyperplane arrangement in V , whose hyperplanes H_i ($i \in I$) are defined by homogeneous linear equations $f_i = 0$. Since \mathcal{A} is essential, the f_i span the dual space V^* . For any sequence $\tau = (\tau_i)_{i \in I}$ with $\tau_i = \pm 1$, set $g_\tau = \sum_{i \in I} \tau_i f_i$. We then define $\widehat{\Sigma} \subset V$ by the $2^{|I|}$ inequalities $g_\tau \leq 1$, one for each τ .

It will follow from what we do below that these inequalities are redundant in general, and that $\widehat{\Sigma}$ is actually defined by the inequalities $g_\sigma \leq 1$ in which σ is the sign sequence of a chamber. Notice that $\sigma_i > 0$ on C if $\tau_i = \tau_i(C)$, so we can

Figure 12. The facets of $\widehat{\Sigma}$ cut across the chambers.

visualize the facets of $\widehat{\Sigma}$ (given by $g_\sigma = 1$) as cutting across the chambers and matching up correctly along faces. See Figure 12.

Note first that $\widehat{\Sigma}$ is indeed a convex polytope; for the inequalities $g_\tau \leq 1$ imply that $|f_i| \leq 1$ on $\widehat{\Sigma}$ and hence that $\widehat{\Sigma}$ is bounded (because the f_i span V^*). Note also that 0 is an interior point of $\widehat{\Sigma}$. We wish to show that the proper, nonempty faces of $\widehat{\Sigma}$ are in 1–1 correspondence with the faces $F \neq \{0\}$ in the face poset \mathcal{F} of \mathcal{A} .

Given $F \neq \{0\}$ in \mathcal{F} , let $\sigma = \sigma(F)$ and let $g_\sigma = \sum_{i \in I} \sigma_i f_i$. (Note that some of these terms are 0 if F is not a chamber.) Then $g_\sigma > 0$ on F , and F is the cone over

$$F_1 \stackrel{\text{def}}{=} F \cap \{g_\sigma = 1\}.$$

We claim that F_1 is a (relatively open) face of $\widehat{\Sigma}$. In fact, F_1 is the face defined by

$$(A.1) \quad \begin{cases} g_\tau = 1 & \text{if } \tau \text{ is consistent with } \sigma \\ g_\tau < 1 & \text{otherwise.} \end{cases}$$

Here τ is *consistent* with σ if $\tau_i = \sigma_i$ for all i such that $\sigma_i \neq 0$. [Recall that $\tau_i = \pm 1$ for all i .]

To verify that (A.1) defines F_1 , suppose τ is consistent with σ , and write $g_\tau = g_\sigma + \sum_{i \in I_0} \tau_i f_i$, where $I_0 = \{i \in I : \sigma_i = 0\}$. It is then immediate that the equalities in (A.1) are equivalent to “ $f_i = 0$ for $i \in I_0$ and $g_\sigma = 1$ ”. And in the presence of these equalities, the inequalities in (A.1) are equivalent to “ $\sigma_i f_i > 0$ for $i \notin I_0$ ”. Thus we have transformed (A.1) to the set of equalities and inequalities defining F_1 , whence the claim.

We now have $V - \{0\}$ partitioned into the cones over some of the (relatively open) faces F_1 of the boundary Σ of $\widehat{\Sigma}$. It follows that the F_1 are in fact all of the faces of Σ , and we have established the desired 1–1 correspondence between the

(nonempty) faces of Σ and the elements of \mathcal{F} (other than $\{0\}$). It is easy to check that this correspondence is a poset isomorphism, i.e., it preserves the face relation.

Finally, our assertion that $\widehat{\Sigma}$ is defined by the inequalities $g_\sigma \leq 1$ (one for each chamber) follows from the fact that a convex polytope with nonempty interior can always be defined by one inequality for each facet (cf., [38], Theorem 2.15).

REFERENCES

1. D. Aldous, *An introduction to covering problems for random walks on graphs*, J. Theoret. Probab. **2** (1989), 87–89.
2. D. Bayer and P. Diaconis, *Trailing the dovetail shuffle to its lair*, Ann. Appl. Probab. **2** (1992), 294–313.
3. J. S. Beissinger and U. N. Peled, *Enumeration of labelled threshold graphs and a theorem of Frobenius involving Eulerian polynomials*, Graphs Combin. **3** (1987), 213–219.
4. F. Bergeron, N. Bergeron, R. B. Howlett, and D. E. Taylor, *A decomposition of the descent algebra of a finite Coxeter group*, J. Algebraic Combin. **1** (1992), 23–44.
5. T. P. Bidigare, *Hyperplane arrangement face algebras and their associated Markov chains*, dissertation, U. of Michigan, 1997.
6. T. P. Bidigare, P. Hanlon and D. Rockmore, *A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements*, preprint (1997).
7. L. J. Billera and B. Sturmfels, *Fiber polytopes*, Ann. of Math. (2) **135** (1992), 527–549.
8. A. Björner and J. W. Walker, *A homotopy complementation formula for partially ordered sets*, European J. Combin. **4** (1983), 11–19.
9. A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler, *Oriented matroids*, Encyclopedia of Mathematics and its Applications, 46, Cambridge University Press, Cambridge, 1993.
10. N. Bourbaki, *Groupes et Algèbres de Lie*, Chapitres 4–6, Masson, Paris, 1981.
11. K. S. Brown, *Buildings*, Springer–Verlag, New York–Berlin, 1989.
12. J–F. Chamayou and G. Letac, *Explicit stationary distributions for compositions of random functions and products of random matrices*, J. Theoret. Probab. **4** (1991), 3–36.
13. G. E. Cooke and R. L. Finney, *Homology of cell complexes*, Based on lectures by Norman E. Steenrod, Princeton University Press, Princeton, 1967.
14. H. H. Crapo, *The Möbius function of a lattice*, J. Combinatorial Theory **1** (1966), 126–131.
15. P. Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11, Institute of Mathematical Statistics, Hayward, CA, 1988.
16. P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996), 1659–1664.
17. P. Diaconis, J. A. Fill, and J. Pitman, *Analysis of top to random shuffles*, Combin. Probab. Comput. **1** (1992), 135–155.
18. P. Diaconis, R. L. Graham, and J. A. Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random Structures Algorithms **1** (1990), 51–72.
19. J–E. Dies, *Chânes de Markov sur les permutations*, Lecture Notes in Mathematics, 1010, Springer-Verlag, Berlin-New York, 1983.
20. R. P. Dobrow and J. A. Fill, *The move-to-front rule for self-organizing lists with Markov dependent requests*, Discrete probability and algorithms (Minneapolis, MN, 1993), IMA Vol. Math. Appl., 72, Springer, New York, 1995, pp. 57–80.
21. P. H. Edelman and V. Reiner, *Free arrangements and rhombic tilings*, Discrete Comput. Geom. **15** (1996), 307–340, Erratum, **17** (1997), 359.
22. P. Ehrenfest and T. Ehrenfest, *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*, Physikalische Zeitschrift **8** (1907), 311–314.
23. J. A. Fill, *An exact formula for the move-to-front rule for self-organizing lists*, J. Theoret. Probab. **9** (1996), 113–160.
24. M. Kac, *Random walk and the theory of Brownian motion*, Amer. Math. Monthly **54** (1947), 369–391.
25. G. Letac, *A contraction principle for certain Markov chains and its applications*, Random matrices and their applications (Brunswick, Maine, 1984), Contemp. Math., 50, Amer. Math. Soc., Providence, RI, 1986, pp. 263–273.

26. N. V. R. Mahadev and U. N. Peled, *Threshold graphs and related topics*, Annals of Discrete Mathematics, 56, North-Holland Publishing Co., Amsterdam, 1995.
27. P. Orlik and H. Terao, *Arrangements of hyperplanes*, Grundlehren der Mathematischen Wissenschaften, 300, Springer-Verlag, Berlin, 1992.
28. R. M. Phatarfod, *On the matrix occurring in a linear search problem*, J. Appl. Probab. **28** (1991), 336–346.
29. R. M. Phatarfod and D. Dye, *The linear search problem with Markov-dependent requests*, preprint (1993).
30. L. Solomon, *A Mackey formula in the group ring of a Coxeter group*, J. Algebra **41** (1976), 255–264.
31. R. P. Stanley, *Enumerative combinatorics, Vol. 1*, Corrected reprint of the 1986 original. Cambridge Studies in Advanced Mathematics, 49., Cambridge University Press, Cambridge, 1997.
32. R. P. Stanley, *A zonotope associated with graphical degree sequences*, Applied geometry and discrete mathematics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 4, Amer. Math. Soc., Providence, RI, 1991, pp. 555–570.
33. R. P. Stanley, *Hyperplane arrangements, interval orders, and trees*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996).
34. D. B. Wilson and J. G. Propp, *How to get an exact sample from a generic Markov chain and sample a random spanning tree from a directed graph, both within the cover time*, Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996), ACM, New York, 1996, pp. 448–457.
35. T. Zaslavsky, *Facing up to arrangements: Face-count formulas for partitions of space by hyperplanes*, Mem. Amer. Math. Soc. **1** (1975), no. 154.
36. G. M. Ziegler, *The face lattice of hyperplane arrangements*, Discrete Math. **73** (1989), 223–238.
37. G. M. Ziegler, *Higher Bruhat orders and cyclic hyperplane arrangements*, Topology **32** (1993), 259–279.
38. G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, 152, Springer-Verlag, New York, 1995.