Filliman duality applied to the Birkhoff polytope

Greg Kuperberg*

UC Davis†

ABSTRACT

1. RESULTS

The Birkhoff polytope $B_n \subset \mathbb{R}^{n^2}$ is an $(n-1)^2$ dimensional polytope with two definitions: It is the convex hull of the $n \times n$ permutation matrices, and it is the set of $n \times n$ non-negative matrices whose row sums and columns sums are 1 (doubly stochastic matrices). It is also called the assignment polytope. An interesting problem is to exactly compute the volume of B_n for small values of n. Chan and Robbins [1] did it for $n \leq 7$ by triangulating B_n and for $n \leq 8$ by computing the Ehrhart polynomial of B_n. The triangulation method is based on a result of Stanley [8] that all simplices of iterated fan triangulations of B_n have the same volume. Instead of explicitly constructing a specific triangulation, the method recursively counts the number of simplices that would be incident to each face of B_n; these counts are invariant under the symmetries of B_n. The Ehrhart polynomial $e(B_n,t)$ counts $n \times n$ non-negative integer matrices whose row sums and column sums are t; such matrices can be counted for some values of n and t using dynamic programming.

In this article we propose Filliman duality as a third method compute $\operatorname{Vol} B_n$. Filliman duality expresses the volume of a polytope P as the sum of the (signed) volumes of the simplices that are dual to those in a triangulation of the polar body P^* [2]. We present a pair of theorems that realize our proposal as an explicit algorithm.

Theorem 1. *Let $E_{i,j}$ be the vertices of the dual Birkhoff polytope B_n^*, with $1 \leq i,j \leq n$. For each tree t with vertices numbered $1,\dots,n$ and edges numbered $1,\dots,n-1$, let Δ_t be the convex hull of all $E_{i,j}$ except those in which edge i contains vertex j. Then $T_n = \{\Delta_t\}$ is a triangulation of B_n^*.*

Theorem 2. *If $\Delta_t \in T$ is a simplex in the triangulation of B_n^* and X is a matrix, then*

$$\operatorname{Vol}\Delta_t = \frac{n}{(n-1)^2!}$$

and

$$\operatorname{Vol}(\Delta_t + X)^* = \frac{n}{(n-1)^2!}\prod ???.$$

Corollary 3.

$$\operatorname{Vol} B_n^* = (n-1)!n^{n-2}$$

and

$$\operatorname{Vol} B_n(X) = \sum_t ???. \tag{1}$$

The statements of Theorems 1 and 2 depend on some conventions for B_n^* and volumes that we describe now. Let V_n be the vector space of $n \times n$ matrices with vanishing row and column sums. Strictly speaking, V_n does not contain B_n, but we may remedy this by identifying B_n with $B_n - b$, where b is the centroid of B_n. $(n-1)^2$-dimensional plane containing B_n in $\mathbb{R}^{n^2}$. Strictly speaking V_n is an affine space rather than a vector space, but we can make it a vector space by declaring the centroid of B_n (the doubly stochastic marix with entries $\frac{1}{n}$) to be the origin. The dual space V_n^* is $\mathbb{R}^{n^2}$ modulo matrices with constant rows or constant columns. We define the inner product between V_n and V_n^* by

$$\langle Y,X\rangle = -n\operatorname{Tr}(Y^T X).$$

The dual polytope B_n^* is then the convex hull of the elementary matrices $E_{i,j}$, defined by $(E_{i,j})_{i,j} = 1$ and $(E_{i,j})_{k,\ell} = 0$ otherwise.

The vertices of B_n lie in a coset of the lattice L of integer matrices in V_n. We scale volume V_n so that $\det L = 1$; it follows that the volume of any $(n-1)^2$-simplex spanned by vertices of B_n is $\frac{k}{(n-1)^2!}$ for some integer k. (This convention differs slightly from Chan and Robbins [1], who omit the denominator.) The dual lattice L^* consists of integer matrices projected to V_n^* and we scale volume in V_n^* dually so that $\det L^* = 1$ as well. The vertices of B_n^* lie in L^* and affinely generate a sublattice of index n; thus each simplex Δ_t has minimum volume.

Finally Filliman duality does not directly apply because the origin lies on the boundary of each simplex Δ_t. As a workaround we translate B_n^* and its triangulation by a generic matrix X, and we let

$$B_n(X) = (B_n^* + X)^*.$$

By a basic property of polar bodies, $B_n(X)$ is the image of B_n under a projective transformation if X is in the interior of B_n.

Before proving Theorems 1 and 2, we discuss the efficiency of equation (1) as an algorithm. Both sides of the equation are rational functions of the entries of the parameter matrix X. We want the value of the left side at $X = 0$, a point at which individual terms on the right are undefined. As a workaround we can choose a curve of matrices X that depends on a single parameter ε with the property that each $\Delta_t^*(X)$ is finite for $\varepsilon \neq 0$ and $X = 0$ when $\varepsilon = 0$. One suitable choice is

$$X_{i,j} = 2^{nj+i}\varepsilon.$$

---

*Supported by NSF grant DMS #0072342

†Electronic address: greg@math.ucdavis.edu

We can then compute the constant term of the Laurent series of $\operatorname{Vol}\Delta_t' + X$ in time $O(n^C)$. Alternatively we can let $\varepsilon = p^k$ be a prime power and evaluate equation (1) p-adically to determine $B_n(0) \bmod p^k$. Either way we can compute a single term of equation (1) in time $O(n^C)$, and the total computation time is

$$(n-1)!n^{n-2}O(n^C) = O(n^{2n}e^{-n}n^C).$$

(We allow the constant C to take more than one value in one equation.)

The Filliman duality method is asymptotically slower than the dynamic programming method to compute the Ehrhart polynomial of B_n. However, the dynamic programming method requires a lot of space. For estimate the space requirement we assume that n is even for simplicity. The last stage of the most difficult iteration of the dynamic programming method (with $t = \binom{n-1}{2}$) requires a table of the number of non-negative $n/2 \times n$ integer matrices whose row sums are t and whose column sums are a certain partition $\lambda \vdash tn/2$. The Young diagram of λ is constrained to lie in a $t \times n$ rectangle. t. There are $\binom{t+n}{n}$ partitions of all sizes in a $t \times n$ rectangle, and their generating function is the Gaussian binomial coefficient $\binom{t+n}{n}_q$. Since Gaussian binomial coefficients are unimodal [7], at least $1/tn$ of these are partitions of $tn/2$. Therefore the dynamic programming method uses

$$\binom{t+n}{n} n^C = O(n^n e^n 2^{-n} n^C)$$

space. Moreover it takes exponential time to compute each entry of the table; for realistic values of n the Ehrhart polynomial method might not be much faster than the Filliman duality method.

The Filliman duality method requires only polynomial space. Moreover, each term can be computed independently, so the algorithm is highly parallelizable.

One strange property of our version of the Filliman duality method is that it does not explicitly use the symmetry of the Birkhoff polytope. Two simplices Δ_{t_1} and Δ_{t_2} in our triangulation of B_n^+ are equivalent by symmetry if and only if the trees t_1 and t_2 are the same except for their labels. There are only $O(\alpha^n n^{-5/2})$ unlabelled trees, where $\alpha = 2.956\ldots$ [6]. Unfortunately, any value of the matrix X that desingularizes every term of the right side of equation (1) possesses none of the symmetry of the equation itself.

Question 4. *Is there a way to exploit the symmetry of B_n in the Filliman duality method to compute its volume?*

2. FILLIMAN DUALITY

In this section we state and prove our own version of Filliman duality.

An *orientation* of a d-simplex $\Delta \subset V$, where V is a d-dimensional vector space over $\mathbb{R}$, is an ordering of its vertices chosen up to an even permutation. An orientation of $\Delta \subset V$ is equivalent to an orientation of V itself in the more usual sense of an ordered (or sign-ordered) basis: If $v_0, \ldots, v_d$ are the vertices of Δ in order, then $v_1 - v_0, v_2 - v_0, \ldots, v_d - v_0$ is the corresponding ordered basis of V. Two simplices in V are said to have the same orientation if they induce the same orientation on V.

Whether oriented or not, a d-simplex Δ may be defined as the unique finite region defined by $d+1$ hyperplanes, provided that no d of them are parallel to a line and that not all of them meet at a point. If $v \in V$ is a point other than the origin, it is dual to a hyperplane $H_v \subset V^*$ defined by

$$H_v = \{w | \langle w, v\rangle = 1\}.$$

We call a simplex Δ *non-codegenerate* if none of its hyperplanes contain the origin. If Δ is non-codegenerate, we can define the dual simplex Δ^* bounded by the hyperplanes dual to the vertices of Δ. If Δ is oriented by an ordering $v_0, \ldots, v_d$ of its vertices, we give Δ^* the orientation $w_1, w_2, \ldots, w_0$, where w_i is the vertex opposite to the hyperplane H_{v_i}. (The change in ordering gives Δ and Δ^* the same orientation if they lie in $\mathbb{R}^d$ and contain the origin.) If Δ contains the origin, then Δ^* is the usual polar body of Δ. It is well-known that polar duality is an involution on non-codegenerate simplices.

Let $\mathscr{A}(V)$ be the abelian group freely generated by oriented, non-codegenerate simplices in V. If Δ is a simplex we let $[\Delta]$ be the corresponding element of $\mathscr{A}(V)$. Evidently duality of non-codegenerate simplices extends to a map $\Phi : \mathscr{A}(V) \to \mathscr{A}(V^*)$. We define a quotient $\mathscr{B}$ of $\mathscr{A}$ by adjoining two kinds of relations:

1. If $\Delta_1, \Delta_2 \in \mathscr{A}$ differ only in orientation, then

$$[\Delta_1] + [\Delta_2] = 0.$$

2. Let $P \subset V$ be a region that can be tiled by finitely many non-codegenerate simplices $\Delta_1, \ldots, \Delta_n$. Suppose that P is decorated by an orientation of V and assume that each Δ_i has the same orientation. Then

$$[P] = \sum_i [\Delta_i]$$

depends only on P and not on the tiling.

We call the region P in the second relation a *non-codegenerate polyhedron.* It need not be convex, nor even a topological ball.

Theorem 5. *The involution Φ descends from $\mathscr{A}$ to $\mathscr{B}$.*

To prove Theorem 5 we will need a variant of the stellar subdivision theorem of M. H. A. Newman. We begin with a refinement also due to Newman [4, 5].

Theorem 6 (Newman). *Any two triangulations of a polyhedron $P \subset V$ are equivalent under stellar moves applied on edges.*

Corollary 7. *All non-codegenerate simplicial tilings of a polyehdron $P \subset V$ are equivalent under the elementary move of dissecting a simplex into two simplices.*

The warped join $P *_f Q$ is always a polyhedral ball and it is often a convex polytope. For example a 3-cube is a warped join of a hexagon and a line segment (exercise). Note that a triangulation of each facet of P joined with a triangulation of Q form a triangulation of $P *_f Q$, whether convex or not.

In our case let $\delta \subset B_n^*$ be the $n-1$-simplex with vertices $E_{n,1}, \ldots, E_{n,n}$. Express V_n and V_n^* as direct sums

$$V_n = Y_n \oplus W_n \qquad V_n^* = Y_n^* \oplus W_n^*,$$

where W_n^* is the $n-1$-plane containing δ and Y_n^* is its orthogonal complement. Recall that V_n is the space of matrices with vanishing row and column sums. It is easy to check that Y_n is the subspace of matrices with vanishing last row and W_n is the subspace of matrices supported on the last row.

The slice $B'_n = Y_n \cap B_n$ and the projection $(B'_n)^*$ of B_n^* onto Y_n^* are dual polytopes. The polytope B'_n is the transportation polytope of $(n-1) \times n$ non-negative matrices with row sums 1 and column sums $(n-1)/n$. (More precisely, B'_n is this polytope with its centroid moved to the origin, following our convention for B_n.) Of the n^2 vertices of B_n^*, n lie in δ and the other $n^2 - n$ must project bijectively to vertices of $(B'_n)^*$, since otherwise $(B'_n)^*$ would have too few vertices. Moreover δ itself projects to the centroid of $(B'_n)^*$. It follows that B_n^* is the convex hull of a warped join $J_n = (B'_n)^* *_f \delta$. Theorem 1 then follows from two final claims:

1. The polytope $(B'_n)^*$ is simplicial and has a facet Δ'_t for each simplex $\Delta_t \in T_n$.

2. The warped join J_n is convex, *i.e.*, $J_n = B_n^*$.

The first claim is equivalent to the characterization of the vertices of the transportation polytope B'_n [? ?]. Thus T_n is a triangulation of J_n. To prove the second claim, we claim that each external facet of each simplex in T_n lies on a facet of B_n^*. Let t be a tree with n numbered vertices and $n-1$ numbered edges so that $\Delta_t \in T_n$. In general if Γ is a bipartite graph with n numbered black vertices (corresponding to rows) and n numbered white vertices, we can associate to it a set of vertices of B_n^*, namely

$$S_\Gamma = \{E_{i,j} | (i,j) \notin E(\Gamma)\}.$$

If Γ is a perfect matching, then it is equivalent to a permutation and S_Γ is the vertex set of the corresponding facet of B_n^*. If $\Gamma = \Gamma_t$ is the barycentric subdivision of t plus an isolated black vertex labelled n, then S_Γ is the vertex set of Δ_t. The vertex set of a facet of Δ_t is S_Γ where Γ is obtained from Γ_t by adding an edge. The internal facets of Δ_t should be shared with another simplex $\Delta_{t'} \in T_n$. The external facets should lie on a facet of B_n^*. These two geometric facts follows from two combinatorial lemmas.

Lemma 9. *If Γ is obtained from Γ_t by connecting the black vertex n to a white vertex by an edge, then it contains a unique perfect matching.*

Lemma 10. *If Γ is obtained from Γ_t by connecting two non-isolated vertices by an edge, then it contains Γ'_t for exactly one other labelled tree t'.*

In lieu of the proofs, which are elementary, we give an example of each lemma in Figure **??**. (Moreover we don't really need Lemma 10.)

Remark. An iterated fan triangulation of a d-polytope P is one whose simplices come from sequences

$$P = F_0 \ni v_0 \notin F_1 \ni v_1 \notin F_2 \ni \ldots \notin F_d \ni v_d.$$

Here each F_i is a facet of F_{i-1} that does not contain v_{i-1} and each v_i is a vertex of F_i. The corresponding simplex has vertices $v_0, \ldots, v_d$. A triangulation is obtained by generating the sequence from the left and at each stage taking one choice for each v_i and all choices for each F_i. The author first found the triangulation T_n of B_n^* as an iterated fan with $v_i = E_{n,i+1}$ for $i < n$. It turns out that every choice for F_n is then a simplex (a warped facet of $(B'_n)^*$ in our terminology here). Thus the remaining iterations of the fan construction are trivial. Interestingly, the fan construction at once establishes the triangulation T_n, the simplicial structure of $(B'_n)^*$, and the warped joined structure of B_n^*.

Question 11. *What other standard polytopes are warped joins or dual to warped joins?*

4. DETERMINANTS

ACKNOWLEDGMENTS

We would like to thank Yael Karshon, Colin Rourke, and Günter Ziegler for useful discussions.

REFERENCES

[1] Clara S. Chan and David P. Robbins, *On the volume of the polytope of doubly stochastic matrices*, Experiment. Math. **8** (1999), 291–300, arXiv:math.CO/9806076.
[2] Paul Filliman, *The volume of duals and sections of polytopes*, Mathematika **39** (1992), 67–80.
[3] Jim Lawrence, *Polytope volume computation*, Math. Comp. **57** (1991), 259–271.
[4] W. B. R. Lickorish, *Simplicial moves on complexes and manifolds*, Geom. Topol. Monogr. **2** (1999), 299–320, arXiv:math.GT/9911256.
[5] M. H. A. Newman, *A theorem in combinatorial topology*, J. London Math. Soc. **6** (1931), 186–192.
[6] Richard Otter, *The number of trees*, Ann. of Math. (2) **49** (1948), 583–599.
[7] Robert A. Proctor, *Solution of two difficult combinatorial problems with linear algebra*, Amer. Math. Monthly **89** (1982), 721–734.
[8] Richard P. Stanley, *Combinatorics and commutative algebra*, 2nd ed., Birkhäuser, Boston, MA, 1996.

Proof. Let T_1 and T_2 be two non-codegenerate simplicial tilings of P. If we intersect every element of T_1 with every element of T_2, the result is a tiling T_3 of P by convex polytopes. The tiling T_3 may not be vertex-to-vertex, but we can interpret it as a vertex-to-vertex tiling by adding degenerate vertices and other facets to some of its elements. The barycentric subdivision T_4 of T_3 is then a triangulation of P that refines both tilings T_1 and T_2. Figure **??** shows an example. If Δ is a simplex in either T_1 or T_2, then the simplices in T_4 that lie in Δ form triangulation T_Δ. It suffices to obtain each such triangulation T_Δ from Δ. Since T_Δ is a triangulation, Theorem 6 says that we can get T_Δ from the triangulation of Δ by itself by stellar moves on edges.

Finally, a stellar move on an edge e consists of dividing each simplex containing e into two simplices. It is therefore realized by a sequence of elementary dissections. If in each stellar move we divide the edge e at a generic point, and if we similarly put the barycentric subdivision in generic position, none of the new facets created are coplanar with the origin. This guarantees that all simplices that ever appear are non-codegenerate. □

Proof of Theorem 5. In light of Corollary 7, we only need to check that Φ preserves an elementary dissection of a simplex. Suppose that x_1, x_2, and x_3 are 3 collinear points in V, and suppose that $v_0, \ldots, v_{d-2}$ are $d-1$ other points affinely independent from any two of x_0, x_1, and x_2. By a slight abuse of notation we write an oriented simplex as the product of its vertices listed in a compatible order. Then

$$[v_0 \cdots v_{d-2}x_1x_2] + [v_0 \cdots v_{d-2}x_2x_3] + [v_0 \cdots v_{d-2}x_3x_1] = 0 \quad (2)$$

expresses an elementary dissection. Applying Φ to both sides produces

$$[w_1 \cdots w_{d-2}y_2y_1w_0] + [w_1 \cdots w_{d-2}y_3y_2w_0] + [w_1 \cdots w_{d-2}y_1y_3w_0] = 0. \quad (3)$$

Here each point w_i lies in the hyperplanes H_{v_j} for $i \neq j$ and in the hyperplanes H_{x_j} for all j. Each point y_i lies in the hyperplane H_{v_j} for all j and in the hyperplane H_{x_i}. Evidently equation (3) is the same equation in V^* up to sign as equation (2) in V. Figure **??** shows an example. □

Proposition 8. *If $P \subset V$ is a convex polytope that strictly contains the origin, then $\Phi([P]) = [P^*]$, the polar body of P.*

Proof. We first assume that any d vertices of P are linearly independent, or equivalently affinely independent from the origin. Let T be a triangulation of P with no vertices in the interior of P. We claim, first, that any point w in the interior of P^* is covered by the dual of exactly one simplex. A unique simplex $\Delta_0 \in T$ contains the origin. The inclusions $0 \in \Delta_0 \subset P$ imply that $\Delta_0^* \supset P^*$, so Δ^* covers w. If $\Delta \in T$ is another simplex, then there exists a vertex v of Δ which is separated from the origin by the opposite face of Δ. It follows that Δ^* is separated from the origin by the hyperplane H_v. Since v is also a vertex of P, H_v is a supporting hyperplane of P^*. Therefore H_v separates Δ^* from P^* and Δ^* does not contain w. This establishes the first claim.

We claim, second, that if w is in the exterior of P^*, there exists a triangulation T of P such that w is not covered by Δ^* for any $\Delta \in T$. There exists a vertex v of P such that the hyperplane H_v separates w from P^*. Let T be a fan triangulation all of whose simplies contain v. If $\Delta \in T$, then H_v separates Δ from w, as desired. The two claims together with Theorem 5 establish the proposition under the independence assumption on P.

Finally we assume that P is arbitrary. The argument so far establishes the proposition for the polytope $P - v$ for a dense set of vectors v in the interior of P. Namely v can be any point that does not lie on a hyperplane affinely spanned by vertices of P. But if v is in the interior of P, then $P - v$ has a non-codegenerate triangulation whether or not v lies on such a hyperplane. It follows that $\Phi([P-v])$ varies continuously in a neighborhood of v. Thus the truth of the proposition for a dense set of v implies its truth for all v in the interior of P. In particular the proposition holds for $v = 0$. □

Theorem 5 and Proposition 8 together imply Filliman's theorem [2]: If T is a triangulation of P by non-codegenerate simplices, then $[P^*]$ is the sum of the duals of the simplices in T. In our version, T can be any collection of oriented non-codegenerate simplices that sum to $[P]$, not necessarily a triangulation:

$$[P] = \sum_{\Delta \in T} [\Delta] \quad \Longrightarrow \quad [P^*] = \sum \Delta \in T[\Delta^*].$$

Nonetheless, T is a triangulation in our application.

Remark. If P is a convex polytope with $0 \in \operatorname{Int} P$, then we can consider a triangulation T of P with a single extra vertex εv and take the limit $\varepsilon \to 0$. The limit of the Filliman dual of T expresses $[P^*]$ as a signed sum of affine orthants. This special case was found independently by Lawrence [3]. Lawrence's orthant decomposition leads to a convenient way to express the Fourier transform of uniform measure on P^*, even when P^* is a simplex. The author first considered $\operatorname{Vol} B_n$ as the value at 0 of the n-fold convolution of uniform measure on a simplex centered at 0. This led to the Fourier transform of this measure, which led to the Lawrence decomposition, and finally to Filliman duality.

3. A TRIANGULATION

We will establish Theorem 1 by realizing B_n^* as a warped join of two other polytopes. Let $P \subset V$ and $Q \subset W$ be two convex polytopes in possibly distinct vector spaces. Let $f : \partial P \to W$ be a continuous function from the boundary of P which is linear on facets. Let $\mathscr{F}$ be the set of facets of the graph of f in $V \oplus W$. Then the *f-warped join* of P and Q is defined as

$$P *_f Q = \bigcup_{F \in \mathscr{F}} \operatorname{Hull} F \cup Q.$$