FOUR QUESTIONS ON BIRKHOFF POLYTOPE

IGOR PAK

Department of Mathematics,
Yale University,
New Haven, CT 06520
paki@math.yale.edu

January 8, 1999

ABSTRACT. We ask several questions on the structure of the polytope of doubly stochastic $n \times n$ matrices $\mathbf{P}_n$, known as a Birkhoff polytope. We discuss the volume of $\mathbf{P}_n$, the work of the simplex method, and the mixing of random walks $\mathbf{P}_n$.

Introduction

In this article we consider a Birkhoff polytope which is, arguably, one of the most important polytopes in many dimensions. Also known as transportation polytope and doubly stochastic matrices polytope, it miraculously appears in various branches of Mathematics from Geometry to Enumerative Combinatorics to Optimization Theory to Statistics. While Birkhoff polytope was thoroughly studied, there are still some interesting open questions related to its rich geometric and combinatorial structure. Here we will touch upon some of these questions.

A *Birkhoff polytope* $\mathbf{P}_n$ is a polytope defined by the following equations and inequalities:

$$\begin{cases} a_{i,j} \geq 0, \\ \sum_{i=1}^{n} a_{i,j} = 1, \\ \sum_{j=1}^{n} a_{i,j} = 1 \end{cases} \qquad \text{for all} \quad 1 \leq i,j \leq n$$

We think of $(a_{i,j})$ as of $n \times n$ doubly stochastic matrices. It is easy to see that $\mathbf{P}_n$ has a dimension $(n-1)^2$ since values of $a_{i,j}$, $1 \leq i,j \leq n-1$ determine the rest. Another easy observation shows that the vertices of $\mathbf{P}_n$ are the permutation matrices (exactly one 1 in each row and column; 0 elsewhere). From now on, we will denote vertices of $\mathbf{P}_n$ by permutations in a symmetric group S_n.

Key words and phrases. Birkhoff polytope, simplex method, random walk, symmetric group, mixing time.

Question 1. *What is the volume of* $\mathbf{P}_n$ *?*

This turns out to be a very hard question, of interest in various fields such as Enumerative Combinatorics, Statistics and Computational Geometry. It is still wide open as far as we know it. Not only that, even for relatively small values of n computing $\mathrm{vol}(P_n)$ represents a significant challenge. The exact values are currently known for $n \le 8$ (see [CR, St]). For more on the volume of the Birkhoff polytope, all transportation polytopes, and their applications see [DG,Mo,S1].

Rather than admit our failure to answer this question, let us prove that this volume is equal to the volume of a *different* polytope. Define a polytope $\mathbf{R}_n$ by the following equations and inequalities:

$$\begin{cases} b_{i,j} \ge 0, \ b_{i,j} \ge b_{i,j+1}, \ b_{i,j} \ge b_{i+1,j}, \\ \sum_{l=1}^{n-i} b_{i+l,l} = n-i, \\ \sum_{l=1}^{n-j} b_{l,j+l} = n-j \end{cases} \qquad \text{for all} \ \ 1 \le i,j \le n$$

We think of $(b_{i,j})$ as of $n \times n$ nonnegative real matrices which weakly decrease in rows and columns and have sums in diagonals equal to their lengths.

Theorem 1. $\mathrm{vol}(\mathbf{R}_n) = \mathrm{vol}(\mathbf{P}_n)$.

We prove this and other theorems in the next sections. We remark that a generalization of Theorem 1 to all transportation polytopes is also known (see [G1,PP]).

Question 2. *Does the simplex method work fast on* $\mathbf{P}_n$ *?*

Let us formalize this question using the following "toy" version of the simplex method.* Let ϕ be a linear function on $\mathbb{R}^{n^2}$ which takes distinct values on all vertices of $\mathbf{P}_n \subset \mathbb{R}^{n^2}$. Consider the following *algorithm*. Start at some vertex and move to any adjacent vertex along any edge on which ϕ decreases. When there are no such edges, we are at minimum.

To analyze this algorithm, recall that the vertices of $\mathbf{P}_n$ are in one-to-one correspondence with elements of the symmetric group S_n. What about the edges? One can show that the edges correspond to pairs of permutations (σ, ω) such that $\sigma^{-1}\omega$ is a single cycle (see e.g. [EKK]). Conclude from here that the diameter of a graph G_n of vertices and edges of $\mathbf{P}_n$ is 2 since every permutation can be presented as a product of two cycles.

Now, since we can get to the desired minimum in at most two steps, what is the maximum number of steps the algorithm can take? Turns out, it can be very long. Consider the following linear function:

$$\phi = x_{1,1} + \alpha\, x_{1,2} + \ldots \alpha^{n-1} x_{1,n} + \alpha^n x_{2,1} + \cdots + \alpha^{n^2-1} x_{n,n}$$

* The real simplex method is more complex in many ways (see e.g. [Sc])

When $\alpha > 0$ is small enough, the linear order on $\mathbf{P}_n$ stabilizes and becomes reverse lexicographic order on permutations. Observe that $\alpha_0 = 1/(n+1)$ will already give such an order. Think of a graph G as of a partially ordered set with $\sigma \succcurlyeq \omega$ if there is a sequence of edges between σ and ω on which ϕ decreases. Then the problem of the length of the algorithm becomes the problem of finding the longest chain in this poset.

Theorem 2. *Let ϕ be as above, $0 < \alpha \le \alpha_0$. Then there exist a decreasing sequence of vertices of $\mathbf{P}_n$ of length $> Cn!$, for a universal constant $C > 0$.*

We will prove the theorem later by exhibiting such a sequence. Now suppose at each step we move along the edge chosen uniformly among all the increasing edges. Given the maximum running time of the algorithm is exponential, should the expected running time be also exponential? Not so.

Theorem 3. *Let ϕ be as above, $0 < \alpha \le \alpha_0$. Then for the expected running time E of the algorithm we have $E = O(n \log n)$.*

Question 3. *Does the nearest neighbor random walk mixes fast on* $\mathbf{P}_n$ *?*

The answer is yes, but we need to give some definitions first. Consider a Markov chain on a graph G, which moves along a uniform edge at each step. By the description of edges, this corresponds to a random walk on S_n generated by all cycles (cf. [D]). Say, we start at identity permutation. Denote the corresponding random walk by $\mathcal{W}$.

Denote by $Q^k(\sigma)$ the probability that after k steps the walk is in vertex $v \in G$. Define the *total variation distance* $d(k) = \|Q^k - U\|$ as follows:

$$d(k) = \frac{1}{2} \max_{A \subset S_n} |Q^k(A) - U(A)| = \sum_{\sigma \in S_n} \left| Q^k(\sigma) - \frac{1}{|S_n|} \right|,$$

where $Q^k(A) = \sum_{\sigma \in A} Q^k(\sigma)$ and $U(A) = |A|/|S_n|$. Now, here is the formal statement of the question: *Given n, $\varepsilon > 0$, what is the smallest $k = k(\varepsilon, n)$ such that $d(k) < \varepsilon$?*

Theorem 4. *For a random walk $\mathcal{W}$ we have $\mathbf{d}(2) \to 0$ as $n \to \infty$.*

In other words, when n is large, the random walk $\mathcal{W}$ mixed after just two steps, which is the diameter of the graph G. There are several other cases when the similar behavior of random walks was observed on S_n (see [L, FL]).

Often it is useful to define a *mixing time* $\mathbf{mix} = k(1/e) = \min\{k \mid d(k) < 1/e\}$ (see e.g. [AF]). The number $1/e$ here can be substituted by any constant and is chosen for convenience. Before we finish, let us recall an old conjecture which states that if the size of the generating set $R \subset S_n$ is polynomial in n, then a mixing time $\mathbf{mix}$ for a random walk on S_n is also polynomial (cf. [B, D, P1]).

Question 4. *Is it true that the nearest neighbor random walk mixes fast on all* $0-1$ *polytopes?*

The answer is no, and this is the easiest questions of all. Remarkably, the solution also involves a Birkhoff polytope. Here is how it can be formalized.

Let $\mathbf{T}$ be a polytope in $\mathbb{R}^d$. We say that $\mathbf{T}$ is a $0-1$ *polytope* if the coordinates of all vertices of $\mathbf{T}$ are either 0 or 1. Such polytopes are usually nicer than general polytopes, for example the diameter of graph of such polytope is at most $d+1$ which is achieved only on a cube (see e.g. [Z]). Now let $\mathcal{W}$ be a nearest neighbor random walk with holding probability $1/2$. Will the mixing time in such a case be polynomial in d?

Now consider a product $\mathbf{T} = \mathbf{P}_n \times I$ of a Birkhoff polytope and an interval $[0,1]$. Clearly, this is a $0-1$ polytope. Think of $\mathbf{T}$ as of a prism with $\mathbf{P}_n \times \{1\}$ as a *top* and $\mathbf{P}_n \times \{0\}$ as a *bottom*. Observe that the probability of moving from the bottom to the top is exactly $1/D$, where D is the degree of each vertex. Thus we need $O(D)$ steps to have a $1/2e$ chance to get to the top if we started on the bottom. This implies $\mathbf{mix} = \Omega(D)$. But $D-1$ is equal to the number of cycles in S_n. Therefore $D = O((n-1)!)$ and $\mathbf{mix}$ is mildly exponential in d, where $d = \dim(\mathbf{T}) = (n-1)^2$.

Note that our construction works for all $0-1$ polytopes with superpolynomial degrees. It is an important open conjecture to show that if vertex degrees are polynomial in the dimension d, then $\mathbf{mix}$ is also polynomial in d. The positive solution of this conjecture may have far reaching consequences in Theoretical Computer Science (see [Mi]).

Proof of Theorem 2

Let us construct the decreasing chain from the largest permutation $(1,2,\ldots,n)$ to the smallest $(n,n-1,\ldots,1)$. For $n=2$ the chain is trivial. Suppose we know the chain for $n \le m-1$. Let the chain for $n=m$ be as follows. It starts with $(1,2,\ldots,m-1,m) \to \cdots \to (1,m,m-1,\ldots,2)$ where the chain in the middle is the chain for $n=m-1$. Now let the chain go to some element $(2,*,\ldots,*)$ then some chain in the middle, then to $(2,n,n-1,\ldots,3,1)$, then to $(3,*,\ldots,*)$, etc. till we finally reach $(n,n-1,\ldots,1)$. Here by $*$ we denote some elements in $[1,n]$ value of which is irrelevant at the moment. All we need now is to describe where we move from $(i,n,n-1,\ldots,i+1,i-1,\ldots,1)$ and what is the chain afterwards before we hit $(i+1,n,n-1,\ldots,i+2,i,\ldots,1)$.

Let $k = \lfloor n/2 \rfloor$. Observe, that it is always possible to move from $\sigma_1 = (i,n,n-1,\ldots,1)$ to $\sigma_2 = (i+1,1,2,\ldots,k,*,*,\ldots,*)$ since $\sigma_2 \preccurlyeq \sigma_1$ and there is a rearrangement of $*$ such that $\sigma_1^{-1}\sigma_2$ is a long cycle. Now from σ_2 use any chain to get to $\sigma_3 = (i+1,1,2,\ldots,k,m,m-1,\ldots,k+1)$ (say, decompose the permutation $\sigma_2^{-1}\sigma_3$ into a product of cycles and use them one by one). Now use induction again for $n = m-k$. Eventually we reach $(i+1,n,n-1,\ldots,1)$ which completes the construction.

Now compute the length of our chain. Denote by L_n the length of the chain on S_n. The induction gives us

$$L_{n+1} = (n+1)\left(L_n - (\lfloor n/2 \rfloor + 1)!\right)$$

Dividing both sides by $(n+1)!$ we have

$$\frac{L_{n+1}}{(n+1)!} > \frac{L_n}{(n)!} - \frac{1}{\lfloor n/2 \rfloor)!}$$

Since $\sum_k 2/k! \to 2/e < 1$ we immediately have $L_n/n! > 1 - 2/e$ for all n. This implies the result. $\square$

Proof of Theorem 3

Let D_n be the total number of cycles in S_n, i.e. the degree of a graph G. We have

$$D_n = \sum_{l=2}^{n} \binom{n}{l} (l-1)! = n! \sum_{l=2}^{n} \frac{1}{l\,(n-l)!} < c\,(n-1)!$$

for some universal constant $c > 1$. Therefore with probability $> 1/c$ the randomly chosen cycle is a long cycle.

Now, suppose the algorithm is at a permutation $\sigma = (n, n-1, \ldots, n-k, j, *, \ldots, *)$, $1 \le j < n-k-1$ Notice that elements $n, n-1, \ldots, n-k$ are "stuck" which means that no decreasing edge can ever change them. Let us compute the number of decreasing edges leaving σ which do not change j. Clearly, this number is not greater than D_{n-k-2}. On the other hand the total number of decreasing edges (σ, ω) is at least $(n-k-1-j)\cdot(n-k-2)!$, which is the number of all cycles on the last $n-k-1$ elements and such that $\omega(k+2) > j$. Therefore the probability that in a decreasing edge (σ,ω) we have $\omega(k+2) > j$ is at least $c' = 1/(1+c)$. We conclude that with probability $> c'$ the $(k+2)$-th element j in a permutation changes to a uniform element $\in [j+1, n-k-1]$. Therefore after $O(\log n)$ moves the $(k+2)$-th element in a permutation will become $n-k-1$ and "stuck". From here it takes $O(n \log n)$ for all elements to get "stucked", which means we reach $(n, n-1, \ldots, 1)$. This finishes the proof. $\square$

Proof of Theorem 1

Let us compute the number of integer points in $N \cdot \mathbf{P}_n$, which is a polytope $\mathbf{P}_n$ expanded in each direction by a factor of N. Clearly, this is the number of *Latin squares*, which is the number of nonnegative integer $n \times n$ matrices with row and column sums equal to N.

We will need few definitions. A partition of m is an integer sequence $\lambda = (\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_l > 0)$ such that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_l = m$. The number $l = l(\lambda)$ is called the *length* of λ. Let $\mathbb{Z}^2$ be an integer lattice with 1×1 squares (i,j), where increases down and j increases from left to the right. A set of squares (i,j) such that $1 \le i \le l$, $1 \le j \le \lambda_i$ is called a Young diagram which we also denote λ (see e.g. [Ma]). A Young tableau A of weight $(N)^n$ and of shape λ, $|\lambda = n\,N$, is a filling of a diagram λ with numbers $1, \ldots, n$, N copies of each, such that the numbers increase in columns and weakly increase in rows.

A famous *Robinson–Schensted–Knuth (RSK) correspondence* in a special case gives a bijection between Latin $n \times n$ squares with sums N and pairs of Young tableaux of weight $(N)^n$ and of the same shape (see [Kn,Sa]).

A *Gelfand–Tsetlin scheme* T of weight $(N)^n$ is a sequence of partitions $(\lambda^{(1)}, \ldots, \lambda^{(n)})$, such that $|\lambda^{(i)}| = N\,(n+1-i)$, $l(\lambda^{(i)}) = n+1-i$ and $\lambda^{(i)}_j \ge \lambda^{(i+1)}_j \ge \lambda^{(i)}_{j+1}$ for all i and j. There is an easy correspondence between Young tableaux and Gelfand–Tsetlin schemes. Simply think of the tableau as of a flag of diagrams $\lambda^{(i)}$ which contain numbers $1, \dots, n+1-i$, $i = 1, \dots, n$. Now arrange the scheme T as follows:

$$\begin{array}{lll} \lambda_1^{(1)} & & \\ \lambda_1^{(2)} & \lambda_2^{(1)} & \\ \lambda_1^{(3)} & \lambda_2^{(2)} & \lambda_3^{(1)} \\ \dots & & \end{array}$$

Observe that numbers now weakly decrease in rows and columns. and that the diagonal sums are $n\,N$, $(n-1)\,N$, $\dots$. Notice also that two Young tableaux with the same shame correspond to the schemes with the same first partition $\lambda^{(1)}$, i.e. with the same diagonal. Now convert two Young tableaux obtained by the RSK bijection into two Gelfand–Tsetlin schemes T_1, T_2. Arrange T_1 as a lower triangular shape (as above), and T_2 as an upper triangular shape. Now glue them together into a square shape (cf. [BK,G1,PP]). Observe that we obtained an integer point of $N \cdot \mathbf{R}_n$, and all integer points in $\mathbf{R}_n$ can be obtained by this correspondence.

We conclude that there is a bijection between integer points in $N \cdot \mathbf{P}_n$ and $N \cdot \mathbf{R}_n$. Therefore polytopes $\mathbf{P}_n$ and $\mathbf{R}_n$ have the same volume (cf. [S2]). $\square$

Proof of Theorem 4

We refer to [Ma, Sa] for the standard results in representation theory of the symmetric group. Let us recall the following notation. Let λ be a partition of n. By d_λ denote the dimension of the irreducible representation π_λ of the symmetric group S_n. By Burnside identity, $d_\lambda < \sqrt{n!}$. By $\chi_\lambda(\mu)$ we denote value of the character of π_λ on a conjugacy class with cycle structure (μ). Clearly, $d_\lambda = tr[\pi_\lambda(id)] = \chi_\lambda(n)$. Also $|\chi_\lambda(\mu)| \le d_\lambda$ for any μ, $|\mu| = |\lambda| = n$ (see e.g [D]).

Our first step will be restating the problem in term of characters of S_n. Use an Upper Bound Lemma of Diaconis and Shahshahani (see [DS, D]). We obtain

$$\mathbf{d}^2(m) \le \frac{1}{4} \sum_{\lambda : |\lambda| = n, \lambda \ne (n)} d_\lambda^2 \left(p + \sum_{l=2}^{n} p_l \frac{\chi_\lambda(l, 1^{n-l})}{d_\lambda} \right)^{2\,m} \; = \; (\circ)$$

where p is the holding probability and

$$p_l = (1-p) \frac{\binom{n}{l}(l-1)!}{D_n} \le \frac{C\,n}{l\,(n-l)!}\,,$$

D_n is the total number of cycles, and C is a universal constant.

To estimate the right hand side of $(\circ)$ we break the inner summation inside into two parts : $l \le n/2$ and $l > n/2$. In the first case we simply bound each ratio by 1. In the second case, the value of the character turns out to be 0 except in one case. This will simplify the problem and eventually give us the desired bound.

Denote by $\widetilde{\lambda}$ the Young diagram obtained by throwing away the first column and the first row of a diagram λ. As before, let $l = l(\lambda) = \lambda_1'$ be the number of rows in λ. Now recall the Murnaghan-Nakayama rule, which is a combinatorial rule for computing the characters $\chi_\lambda(\mu)$ (see e.g. [Ma, Sa]). Without being explicit, observe that the Murnaghan-Nakayama rule immediately implies that if $l > n/2$, then we have $|\chi_\lambda(k, 1^{n-k})| = d_{\widetilde{\lambda}}$ if $k = \lambda_1 + l(\lambda) - 1$ and $|\chi_\lambda(k, 1^{n-k})| = 0$ otherwise. We have:

$$(\circ) \le \sum_{k=1}^{n} \sum_{\lambda \ne (n), l+\lambda_1-1=k} d_\lambda^2 \left(p_k \frac{d_{\widetilde{\lambda}}}{d_\lambda} + \sum_{j=2}^{\lceil n/2 \rceil} p_j \right)^{2\,m}$$

Now let $m = 2$. Using the formula for p_k, and $\sum_{j=2}^{\lceil n/2 \rceil} p_j \le \frac{Cn^2}{\lceil n/2 \rceil !}$, we have

$$(\circ) \le \sum_{k=1}^{n} \sum_{\lambda \ne (n), l+\lambda_1-1=k} d_\lambda^2 \left(\frac{Cn}{k\,(n-k)!} \cdot \frac{\sqrt{(n-k)!}}{d_\lambda} + \frac{Cn^2}{\lceil n/2 \rceil !} \right)^4$$

$$\le \sum_{k=1}^{n} \sum_{\lambda \ne (n), l+\lambda_1-1=k} \left(\frac{C\,n}{k\sqrt{(n-k)!}\sqrt{d_\lambda}} + \frac{Cn^2\sqrt{d_\lambda}}{\lceil n/2 \rceil !} \right)^4$$

where the last inequality follows from $d_{\widetilde{\lambda}} < \sqrt{(n-k)!}$. Recall that $(a+b)^4 \le 8(a^4 + b^4)$. This breaks each summand into two, which breaks our double sum into two double sums. The second double sum $\to 0$ as $n \to \infty$, which easily follows from $\sqrt{d_\lambda} < (n!)^{1/4}$, Stirling formula, and since the total number of partitions $\sim \frac{1}{4n\sqrt{3}} \cdot e^{\pi\sqrt{2/3n}}$ (see e.g. [O]). For the first double sum, it is known (see [FL,L]) that

$$\sum_{\lambda \ne (n), (1^n);\, |\lambda| = n} \frac{1}{d_\lambda} \to 0 \quad \text{as} \quad n \to \infty$$

It is not hard to see that this implies that the first double sum also $\to 0$ as $n \to \infty$ and proves the result. $\square$

Acknowledgments

We would like to thank A. Berenstein, P. Diaconis, R. Kannan, A.N. Kirillov, L. Lovasz, N. Lulov, M. Mihail, R. Stanley, B. Sturmfels, P. Winkler and A. Zelevinsky for helpful remarks. Special thanks to Richard Stanley for pointing out the paper [G2] which gives a shorter alternative proof of Theorem 1.

The author was partially supported by the NSF Postdoctoral Research Fellowship.

References

[A] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, Lecture Notes in Mathematics **986** (1983).

[AF] D. Aldous, J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, 1996.

[B] L. Babai, *Automorphism groups, isomorphism, reconstruction*, in Handbook of Combinatorics (R. L. Graham, M. Groetschel, and L. Lovasz, eds.) (1996), Elsevier.

[BK] A. Berenstein, A. N. Kirillov, *Groups generated by involutions, Gelfand-Tsetlin patterns, and combinatorics of Young tableaux*, Algera i Analiz (in Russian) **7** (1995), 92–152.

[CR] C. S. Chan, D. P. Robbins, *On the volume of the polytope of doubly stochastic matrices* (1998), preprint.

[C] F. R. K. Chung, *Spectral Graph Theory* (Regional conference series in mathematics, no 92), American Mathematical Society, Providence, RI, 1997.

[D] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, Hayward, California, 1988.

[DG] P. Diaconis, A. Gangolli, *Rectangular arrays with fixed margins*, IMA series **72** (1995), Springer, 15–41.

[DS] P. Diaconis, M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahr. verw. Gebiete **57** (1981), 159–179.

[EKK] V. A. Emelichev, M. M. Kovalev, M. K. Kravtsov, *Polytopes, Graphs, Optimization*, Nauka, Moscow, 1981.

[FL] S. Fomin, N. Lulov, *On the number of rim hook tableaux*, Zap. Nauch. Sem. POMI **223** (1995), 219–226.

[G1] E. Gansner, *Matrix correspondences of plane partition*, Pacific J. Math. **92** (1981), 295–315.

[G2] E. Gansner, *The Hillman-Grassl correspondence and the enumeration of reverse plane partitions*, J. Combin. Theory, Ser. A **30** (1981), 71–89.

[K] D. Knuth, *Permutations, matrices and generalized Young tableaux*, Pacific J. Math. **34** (1970), 709–727.

[L] N. Lulov, *Random walks on groups*, Ph.D. Thesis, Harvard U., 1996.

[Ma] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford U. Press, 2nd ed., 1995.

[Mi] M. Mihail, *On the expansion of combinatorial polytopes*, Proc. Mathematical Foundations of Computer Science (Prague 1992), in Lecture Notes in Comp. Sci. , vol 629 (1992), Springer, Berlin, 37–49.

[Mo] J. Mount, *Application of Convex Sampling to Optimization and Contingency Table Generation/Counting*, Ph.D. Thesis, CMU, 1995.

[O] A. Odlyzko, *Asymptotic Enumeration Methods*, in Handbook of Combinatorics, vol. 2 (R. L. Graham, M. Groetschel, and L. Lovasz, eds.) (1996), Elsevier.

[P1] I. Pak, *Random walks on groups: strong uniform time approach*, Ph.D. Thesis, Harvard U., 1997.

[P1] I. Pak, *When and how n choose k,* in Randomization Methods in Algorithmic Design (P. Pardalos et al., eds.) (1998), A.M.S. DIMACS Series, vol. 43, Providence, RI.

[PP] I. Pak, A. Postnikov, *Oscillating tableaux and generalized RSK correspondence* (1996), Proc. of Eighth FPSAC Conference, UMN, Minneapolis.

[Sa] B. Sagan, *The symmetric group*, Wadsworth, Pacific Grove, CA, 1991.

[Sc] A. Schrijer, *Theory of Integer and Linear Programming*, John Wiley, New York, NY, 1988.

[S1] R. Stanley, *Combinatorics and Commutative Algebra*, Birkhouser, Boston, 1996.

[S2] R. Stanley, *Decompositions of rational convex polytopes*, Ann. Discrete Math. **6** (1980), 333–342.

[St] B. Sturmfels, *Equations defining toric varieties*, A.M.S. Proceeding of Symposia in Pure Mathematics **62** (1998), Providence, RI, 447–449.

[Z] G. Ziegler, *Lectures on Polytopes, Graduate Texts in Mathematics 152*, Springer, New York, 1995.