The Number of Linear Extensions of Ranked Posets

Graham R. Brightwell *
Department of Mathematics
London School of Economics
Houghton St.
London WC2A 2AE
U.K.

28 October 2003

CDAM Research Report LSE-CDAM-2003-18

Abstract

We revisit the question of counting the number of linear extensions of the Boolean lattice, relating this to the polyhedral methods of Kahn and Kim, and of Stanley. We give simpler proofs of various known results, and give an upper bound on the number of linear extensions of an arbitrary ranked poset satisfying the LYM condition.

[*Note: This preprint is not intended for journal publication, as it is a record of some alternative proofs of known theorems. The one result, Theorem 2.1, that does not appear in the literature will appear in the forthcoming paper of Brightwell and Tetali [1]. Other material here is likely to find its way into the book of Brightwell and Trotter.*]

1 Introduction

One frequently-asked question in the theory of partially ordered sets (posets) is to estimate the number of linear extensions of the Boolean lattice $\mathbf{2}^n$, consisting of the subsets of $[n] = \{1, 2, \ldots, n\}$, ordered by inclusion. This problem was attributed by Sha and Kleitman [5] to Richard Stanley, but the author has been asked it by several others independently.

A trivial lower bound is $\prod_{i=0}^{n} \binom{n}{i}!$, since this counts exactly those linear extensions of $\mathbf{2}^n$ in which all of the ith level comes before the $(i+1)$st, for each i. A trivial upper bound is

*email: g.r.brightwell@lse.ac.uk

$\binom{n}{\lfloor n/2 \rfloor}^{2^n}$; to see this, consider writing down the linear extension from the bottom up – at each stage the set of possible next elements is an antichain in $\mathbf{2}^n$, and so has size at most $\binom{n}{\lfloor n/2 \rfloor}^{2^n}$. For most purposes, these bounds are sufficiently close: the lower bound is of the form

$$\left(\binom{n}{n/2}\frac{1}{e^{3/2}}(1+o(1))\right)^{2^n}.$$

Sha and Kleitman [5] proved an appealing upper bound of $\prod_{i=0}^{n}\binom{n}{i}^{\binom{n}{i}}$ on the number of linear extensions of $\mathbf{2}^n$. This upper bound is of the form

$$\left(\binom{n}{n/2}\frac{1}{e^{1/2}}(1+o(1))\right)^{2^n}.$$

They suggested that the simple lower bound gives the right value of the constant. Recently, Brightwell and Tetali [1] have proved this conjecture, using the entropy method of Kahn.

The proof of Sha and Kleitman proceeds by considering, for each k, the size of the largest antichain in $\mathbf{2}^n$ with k elements below it; the point is that this is an upper bound on the number of choices available at the $(k+1)$st step when building a bottom-up linear extension, so the product of these numbers is an upper bound on the total number of linear extensions.

The analysis of Sha and Kleitman was extended by Shastri [6] to cover any ranked poset satisfying the LYM condition the sizes of whose ranks form a symmetric unimodal sequence. To be precise, let P be a ranked poset with rank numbers $N_0, N_1, \ldots, N_n$, and suppose that $N_0 \le N_1 \le \cdots N_{\lfloor n/2 \rfloor}$ and $N_i = N_{n-i}$ for each i, and also that P satisfies the LYM condition (we will discuss the LYM condition in due course). Then the number $e(P)$ of linear extensions of P satisfies

$$\prod_{i=0}^{n} N_i! \le e(P) \le \prod_{i=0}^{n} N_i^{N_i}.$$

Here, of course, the trivial lower bound requires none of the extra conditions on the ranked poset P. The proof gives a rather weaker conclusion if one removes the conditions on the ranks.

Our purpose in this note is to give an alternative proof (in two versions) of Shastri's result, without the conditions on the unimodality and symmetry of the sequence of rank numbers. The first version of the proof is a simple probabilistic argument: the second is a direct and extremely straightforward application of a result of Kahn and Kim [3] based on a theorem of Stanley [7] – for completeness we shall supply proofs from first principles. Our proof(s) *might* point the way to narrowing the gap between the lower and upper bounds, perhaps to providing an alternative proof of the result of Brightwell and Tetali, but so far we have not been able to make progress along these lines.

It is well-known that the LYM condition has several different equivalent formulations; we look at some of these in the context of polyhedra associated with P, which may also be of some interest.

2 The proof: probabilistic version

From now on, P will always be a ranked poset (i.e., all maximal chains have the same length n) on $[N]$ with ranks $A_0, A_1, \ldots, A_n$ of sizes (*rank numbers*) $N_0, N_1, \ldots, N_n$ respectively. The *weight* $w(x)$ of an element $x \in A_i$ is defined to be $1/N_i$. The *LYM condition* is that, for any antichain A of P, the sum $w(A)$ of the weights of elements of A is at most 1. (Note that equality is achieved for each of the ranks A_i.)

Theorem 2.1 *Let P be a ranked poset satisfying the LYM condition, with rank numbers $N_0, \ldots, N_n$. Then $e(P) \leq \prod_{i=0}^{n} N_i^{N_i}$.*

Proof. Consider the following random procedure for building a linear extension of P from the bottom up. At each stage, from the set of available elements, choose the next element with probability proportional to its weight.

Now consider any single linear extension $\prec$ of P, say $x_1 \prec x_2 \prec \cdots \prec x_N$. The probability that our random procedure results in $\prec$ is exactly $\prod_{j=1}^{N} w(x_j)/w(B_j)$, where B_j is the set of elements minimal in $P|\{x_j, \ldots, x_N\}$. Since the B_j are antichains, each $w(B_j)$ is at most 1, and the probability of $\prec$ is at least $\prod_{j=1}^{N} w(x_j)$. This is independent of the particular linear extension $\prec$, so we conclude that the total number of possible linear extensions is at most $1/\prod w(x_j)$, which is exactly $\prod_{i=0}^{n} N_i^{N_i}$, since each of the N_i elements of A_i contributes a factor N_i to the product. $\square$

3 The proof: polyhedral version

The *chain polytope* $\mathcal{C}(P)$ of a poset P on $[N]$ is defined as

$$\{\mathbf{x} \in [0,1]^N : \sum_{j \in C} x_j \leq 1 \text{ for all chains } C \text{ of } P\}.$$

It is well-known that $\mathcal{C}(P)$ is the convex hull of the indicator vectors of antichains of P – since the comparability graph of P is perfect. A theorem of Stanley [7] states that the volume of $\mathcal{C}(P)$ is exactly $e(P)/N!$. (We shall discuss these issues in full later.)

Define the vector $\mathbf{y} \in [0,1]^N$ by $y_j = 1/N_i$ whenever $j \in N_i$. If $\mathbf{z}$ is the indicator function of any antichain in P, then by the LYM condition we have $\mathbf{z} \cdot \mathbf{y} \leq 1$. Therefore this linear inequality is satisfied by all $\mathbf{x} \in \mathcal{C}(P)$, and $\mathcal{C}(P) \subseteq \{\mathbf{x} \in [0,1]^N : \mathbf{x} \cdot \mathbf{y} \leq 1\}$, a simplex of volume $(1/N!) \prod_j (1/y_j) = (1/N!) \prod_{i=0}^{n} N_i^{N_i}$. Applying Stanley's Theorem now yields Theorem 2.1.

As we shall explain in the next section, this proof reveals Theorem 2.1 to be a special case of a more general observation due to Kahn and Kim [3].

4 The proof(s) in context

The *antichain polytope* $\mathcal{A}(P)$ of a poset P is defined as

$$\{\mathbf{x} \in [0,1]^N : \sum_{j \in A} x_j \leq 1 \text{ for all antichains } A \text{ of } P\}.$$

The LYM condition states exactly that the vector $\mathbf{y}$ in the previous section lies in $\mathcal{A}(P)$. It is clear that either of our two proofs of Theorem 2.1 could work just as well with any other point of $\mathcal{A}(P)$: if $\mathbf{z}$ is any vector in $\mathcal{A}(P)$, then $e(P) \leq \prod_{j=1}^N 1/z_j$. It is natural to wonder whether any of these other bounds could be better, but the answer is no.

The following material can all be found in the papers of Csiszar, Körner, Lovász, Marton and Simonyi [2] and/or Kahn and Kim [3]. A *convex corner* in $\mathbb{R}^N$ is a compact convex downset of full dimension contained within the positive quadrant. The *antiblocker* of a convex corner K is

$$K^* = \{\mathbf{x} \in \mathbb{R}_+^N : \mathbf{x} \cdot \mathbf{y} \leq 1 \text{ for all } \mathbf{y} \in K\}.$$

The antiblocker K^* is always a convex corner, and $K^{**} = K$. The *clique polytope* $\mathrm{Cl}(G)$ of a graph G is the convex hull of the indicator vectors of cliques in G; it is always a convex corner, and it is easy to see that $\mathrm{Cl}(\overline{G}) \subseteq (\mathrm{Cl}(G))^*$. In fact, we have equality if and only if the graph G is perfect. As the comparability graph of a poset P is perfect, the chain and antichain polytopes form an antiblocking pair.

For a convex corner K, set $V(K) = \max_{\mathbf{x} \in K} \prod x_j$. This is related to the *entropy* $H(K)$ of K; specifically $V(K) = 2^{-nH(K)}$. For any antiblocking pair K, K^* of convex corners in $\mathbb{R}^N$, a pair $\mathbf{x} \in K$, $\mathbf{y} \in K^*$ maximises $\prod x_j$ and $\prod y_j$ if and only if $x_j y_j = 1/N$ for each j. In other words, $\mathbf{x} \in K$ maximises $\prod x_j$ if and only if the vector $\mathbf{y}$ defined by $y_j = 1/(Nx_j)$ lies in K^*. This implies that $V(K)V(K^*) = 1/N^N$. In the case of a ranked poset P with rank numbers $N_0, \ldots, N_n$, the vector $\mathbf{y}$ given by $y_j = N_i/N$ whenever $j \in A_i$ is indeed clearly in the chain polytope, and therefore the LYM condition does indeed give rise to the 'best' upper bound of this type.

We have seen that $e(P) \leq \prod_{j=1}^N 1/z_j$ for any $\mathbf{z} \in \mathcal{A}(P)$ (although we only stated this for LYM posets, both our arguments go through for any poset P); evidently the best bound we can obtain is $1/V(\mathcal{A}(P))$, and we record this as a theorem.

Theorem 4.1 $e(P) \leq 1/V(\mathcal{A}(P)) = V(\mathcal{C}(P))N^N$.

Of course, Theorem 2.1 is a special case.

Theorem 4.1 is due to Kahn and Kim [3], who proved it exactly as in Section 3. This was the first paper to make use of Stanley's Theorem as a tool for estimating $e(P)$; in the paper, this is basically a side observation, while the main purpose was to use the ideas of entropy to prove powerful results about comparison sorting algorithms.

It is worth emphasising that our probabilistic argument from Section 2 goes through unchanged in the general setting, and thus yields an alternative proof of Theorem 4.1, perhaps simpler in that it does not appeal to Stanley's Theorem.

5 A proof of Stanley's Theorem

For completeness, we include a proof of Stanley's Theorem. This is somewhat different from Stanley's original [7], although ultimately the ideas are related. As we discuss below, the main idea of our proof can again be found in the influential paper of Kahn and Kim [3].

The *order polytope* $\mathcal{O}(P)$ of a partial order P is defined as

$$\{\mathbf{x} \in [0,1]^N : x_i \leq x_j \text{ whenever } i < j \text{ in } P\}.$$

Theorem 5.1

$$\mathrm{vol}(\mathcal{C}(P)) = \mathrm{vol}(\mathcal{O}(P)) = \frac{e(P)}{N!}.$$

Proof. To see the second identity, note that the unit cube $[0,1]^N$ breaks up into $N!$ pieces, according to the order of the co-ordinates. These pieces are disjoint up to a set of measure zero, and all have the same volume by symmetry; thus the volume of any piece is $1/N!$, and $\mathcal{O}(P)$ is the union of those pieces corresponding to linear extensions of P.

The main content of the theorem is the first identity; to establish this, we give a bijection from $\mathcal{O}(P)$ to $\mathcal{C}(P)$ that we shall demonstrate to be measure-preserving.

For $\mathbf{x} \in \mathcal{O}(P)$, define $y_j = x_j$ if j is minimal, and $y_j = x_j - \max_{i<j} x_i$ for all other $j \in [N]$. To see that $\mathbf{y} \in \mathcal{C}(P)$, note that all the y_i are non-negative since $\mathbf{x} \in \mathcal{O}(P)$, and that, for any chain $i_1 < i_2 < \cdots < i_k$, we have

$$y_{i_1} \leq x_{i_1}, y_{i_2} \leq x_{i_2} - x_{i_1}, \ldots, y_{i_k} \leq x_{i_k} - x_{i_{k-1}},$$

so $y_{i_1} + y_{i_2} + \cdots + y_{i_k} \leq x_{i_k} \leq 1$.

We claim that the inverse of this map $\mathbf{x} \mapsto \mathbf{y}$ is given by $x_j = \max_{C(j)} \sum_{i \in C(j)} y_i$, where $C(j)$ runs over all chains in P with top element j. Indeed, this is trivial for all minimal j: if it is true for all elements ℓ below j in P, then

$$\max_{C(j)} \sum_{i \in C(j)} y_i = y_j + \max_{\ell<j} \max_{C(\ell)} \sum_{i \in C(\ell)} y_i = y_j + \max_{\ell<j} x_\ell = x_j.$$

As this inverse map $\mathbf{y} \mapsto \mathbf{x}$ evidently takes any point of $\mathcal{C}(P)$ to a point of $\mathcal{O}(P)$, this establishes that the original map is a bijection.

To check that it is measure-preserving, consider the map $\mathbf{x} \mapsto \mathbf{y}$ acting on the piece of $\mathcal{O}(P)$ where $x_{i_1} < x_{i_2} < \cdots < x_{i_N}$. This is a linear map, represented by a matrix which, when written with co-ordinates in the order $i_1, i_2, \ldots, i_N$, has 1s on the leading diagonal, some -1s below it, and 0s everywhere else. Therefore this matrix has determinant 1, and the map is measure-preserving on each such piece.

Therefore, as claimed, our map is a measure-preserving bijection from $\mathcal{O}(P)$ to $\mathcal{C}(P)$. $\square$

There is another way to look at this proof, as we sketch below, which links it to the well-known correspondence between antichains and up-sets in a poset.

For each linear extension $i_1 < i_2 < \cdots < i_N$ of P, let A_j be the set of minimal elements in the restriction of P to $\{i_j, \ldots, i_N\}$. Now for any non-negative multipliers $\lambda_1, \ldots, \lambda_N$ summing to 1, the vector $\sum_j \lambda_j \mathbf{e}_{A_j}$ is in $\mathcal{C}(P)$, where $\mathbf{e}_A$ is the indicator vector of the set A. The volume of $\mathcal{C}(P)$ associated with a given linear extension is $1/N!$, and (up to a set of measure zero) every vector in $\mathcal{C}(P)$ has a unique representation in this form – this is Theorem 2.4 of Kahn and Kim [3], they call this the *laminar decomposition*. Now the map taking $\sum_j \lambda_j \mathbf{e}_{A_j}$ to $\sum_j \lambda_j \mathbf{e}_{\{x_j,\ldots,x_N\}}$ is a measure-preserving bijection from $\mathcal{C}(P)$ to $\mathcal{O}(P)$. Indeed it is the same map as we studied in our proof of Theorem 5.1.

In other words, each extreme point of $\mathcal{C}(P)$ is the indicator vector of an antichain A, we map this to the extreme point of $\mathcal{O}(P)$ given by the indicator vector of the up-set of elements above or in A, we associate each point of $\mathcal{C}(P)$ or $\mathcal{O}(P)$ to a canonical simplex spanned by extreme points, and interpolate linearly. Note that this approach does in fact represent $\mathcal{C}(P)$ as the union of $e(P)$ essentially disjoint simplices of volume $1/N!$; this means that Stanley's Theorem can be deduced quickly from Theorem 2.4 of Kahn and Kim.

6 The LYM condition in a polyhedral context

In 1974, Kleitman [4] proved that the LYM condition was equivalent to various others, one of which is the existence of a *regular covering* of P by chains, i.e., a non-empty collection of maximal chains such that, for each i, every element of rank i occurs in the same number of chains. A moment's thought reveals that this says that the vector $\mathbf{y}$ given by $y_j = 1/N_i$ for $j \in A_i$ can be written as a convex combination of indicator vectors of chains – the polyhedral theory assures us that this is equivalent to $\mathbf{y}$ being in $\mathcal{A}(P)$, which is exactly the LYM condition.

References

[1] Graham R. Brightwell and Prasad Tetali, The number of linear extensions of the Boolean lattice, to appear in *Order*.

[2] I.Csiszár, J.Körner, L.Lovász, K.Marton and G.Simonyi, Entropy splitting for antiblocking corners and perfect graphs, *Combinatorica* **10** (1990) 27–40.

[3] J. Kahn and J.H. Kim, Entropy and sorting, *J. Comput. System Sci.* **51** (1995) 390–399.

[4] Daniel J. Kleitman, On an extremal property of antichains in partial orders, in *Combinatorics* (M.Hall and J.H.van Lint eds) Math. Centre Tracts 55, 77-90.

[5] Jichang Sha and D.J.Kleitman, The number of linear extensions of subset ordering, *Disc. Math.* **63** (1987) 271–278.

[6] Aditya Shastri, Number of linear extensions of symmetric KLYM posets, *Util. Math.* **53** (1998) 23–35.

[7] R.Stanley, Two poset polytopes, *Disc. Comp. Geom.* **1** (1986) 9–23.