The Number of Degree Sequences of Graphs
by
Jason Matthew Burns
B.S., University of South Carolina, 1999
M.A., University of South Carolina, 2000
Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2007
© 2007 Jason Matthew Burns. All rights reserved.
The author hereby grants to MIT permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document
in whole or in part, in any medium now known or hereafter created.

Author.................................................................................................
Department of Mathematics
May 4, 2007

Certified by.................................................................
Richard P. Stanley
Levinson Professor of Applied Mathematics
Thesis Supervisor

Accepted by.................................................................
Alar Toomre
Chairman, Applied Mathematics Committee

Accepted by.................................................................
Pavel I. Etingof
Chairman, Department Committee on Graduate Students
The Number of Degree Sequences of Graphs

by

Jason Matthew Burns

Submitted to the Department of Mathematics
on May 4, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We give nontrivial upper and lower bounds for the total number of distinct degree sequences among all simple, unlabeled graphs on $n$ vertices (graphical partitions on $n$ vertices). Our upper bound is $4^n/(\log n)^C \sqrt{n}$ for some constant $C$, an improvement of $(\log n)^C$ over the trivial upper bound which is asymptotic to $4^n/2\sqrt{n}$. Our lower bound is $4^n/Cn$, an improvement of $\sqrt{n}$ over the trivial lower bound which is asymptotic to $4^n/Cn^{3/2}$.

Thesis Supervisor: Richard P. Stanley
Title: Levinson Professor of Applied Mathematics
Acknowledgments

It’s an incomplete list, but thanks go to:

- My thesis committee: Richard Stanley, Alex Postnikov, and Danny Kleitman.

- John Herron and Kevin Hope at the University of Montevallo, Alabama, who caught a mistake in an earlier draft of this thesis.

- MIT’s undergraduates, whose tuition paid my own; especially 18.03 (spring 2005) recitations 20 and 26, 18.02 recitation 10 (fall 2005), the fall 2006 class of 18.06, and Emily from 18.013A.

- The American taxpayer.

- The Department of Mathematics at MIT, who put up with me for longer than they had to; special thanks to Linda, Stevie, Shirley, Michelle, Anna, Debbie, and Joanne, among others.

- Mom and Dad, of course.
Contents

1 Introduction 7
   1.1 Background ..................................................... 7
   1.2 Definitions ..................................................... 11
       1.2.1 The Erdős-Gallai condition ................................. 11
       1.2.2 Enumerating graphical partitions ......................... 12
   1.3 Trivial bounds ................................................... 15
       1.3.1 A trivial upper bound ..................................... 15
       1.3.2 A trivial lower bound ..................................... 16

2 An Improved Lower Bound 20
   2.1 Strategy .......................................................... 20
   2.2 Permissible signs for dominance pairs .......................... 22
   2.3 Number of signs-and-orderings:
       the special case ................................................. 23
   2.4 Number of signs-and-orderings:
       the general case ................................................ 25
   2.5 Calculating the order of growth ................................ 27
   2.6 Even dominance-pairs ........................................... 29
List of Figures

1-1 The partition $\lambda = 666443221$ ................................. 13
1-2 The off-diagonal shape for $\lambda = 666443221$ ................................. 13
1-3 Counting inclusion-pairs in a $k \times l$ box as noncrossing paths ......... 16

2-1 Decomposing partition-pairs as inclusion pairs .............................. 21
2-2 Adding or removing a square to make our dominance-pair even .......... 30

3-1 The (upside-down) shape for partition $\lambda = (7, 6, 4, 4, 1, 0)$ ........... 35
3-2 $X_j$, $Y_j$, and $Z_j$ .................................................. 36

4-1 Graph of proportion of graphical partitions (log-log plot) ................. 47
4-2 Numbers of graphical partitions, out of 10,000 random partitions .... 48
4-3 Log number of shapes vs. dominance-pairs for each $r$, $k = l = 100$ .... 49
Chapter 1

Introduction

1.1 Background

Consider the set of graphs on \( n \) vertices (labeled 1, \ldots, \( n \)). For each graph \( G \), we can list the degrees \( d_i \) of each vertex in order \( d_1, \ldots, d_n \); this list of degrees is \( G \)'s degree sequence. Like the many other parameters graph theorists study, degree sequences give us a way to grasp the overall “shape” of large graphs too complicated to draw. For instance, graphs taken from a “real-world” source will have properties dramatically different (as statistical distributions) from truly randomly-generated graphs, but if we first choose a degree sequence (according to the observed “real-world” distribution) and then choose randomly from the set of graphs with that degree sequence, then our randomly-selected graphs will resemble the “real” graphs much more closely[5].

Degree sequences have been well studied, and there are a multitude of equivalent conditions for determining when a given sequence of integers is a degree sequence. (Several can be found in [14], for instance.) For example, one well-known condition is
the Havel-Hakimi condition. This says that a nonzero sequence is a degree sequence if and only if, when we remove the largest degree $d$ and reduce the $d$ largest remaining parts by 1 each, the resulting sequence is still a degree sequence. (In other words, we can inductively construct at least one graph corresponding to any degree sequence, by filling in edges for the highest-degree vertices first.)

Note that this condition depends only on the unlabeled graph; since the labels are irrelevant, we may well choose to ignore the ordering of our degrees, or place them into some standard order, say nondecreasing order. Such a sorted degree list we will call a graphical partition (a partition is a sequence of positive integers in nondecreasing order). This graphical partition describes the unlabeled graph, in much the same way as the degree sequence describes the labeled graph. It is natural to disregard trailing zeroes (corresponding to isolated vertices), so we will do so in this context.

It is thus quite efficient to test an individual sequence (or partition) to see if it is graphical; however, this is not something we want to test individually for each of $n^n$ sequences! So the asymptotic questions of how many such sequences are graphical for large numbers of vertices $n$ is a separate one, and quite a bit harder. There is also the related question for large numbers of edges $m$, and both questions may be asked for degree sequences on labeled graphs as well as for graphical partitions on unlabeled graphs.

**Labeled graphs on $n$ vertices**

The case of labeled graphs was tackled by Stanley [16] in 1991. Consider the possible degree sequences on $n$ vertices as integer points in $R^n$, and take the convex hull of these points; this $n$-dimensional polytope can be described explicitly, and the integer
points inside counted. These integer points correspond (more or less) to the possible degree sequences.

By writing this polytope as a union of its component parallelotopes, we can write the number \( f(n) \) of degree sequences as a sum:

\[
f(n) = \sum_X \max\{1, 2^{\#OC(X)} - 1\}
\]

where \( X \) is the set of “quasiforests” on the \( n \) labeled vertices, and \( \#OC(X) \) is the number of odd cycles in that quasiforest. (A quasiforest is like a forest (a disjoint union of trees), except that we allow some of the components to have exactly one cycle, of odd length.) No direct, combinatorial proof of this is known.

**Unlabeled graphs on \( m \) edges**

From a number-theoretical standpoint, partitions are interesting in their own right, so the question of how graphical partitions relate to partitions in general was a natural question. It is clear that the total number of partitions whose parts sum to \( 2m \) total, \( p(2m) \), forms an upper bound for the number of graphical partitions on \( m \) edges, i.e., \( g(2m) < p(2m) \). In 1982, Wilf conjectured that the true rate of growth of \( g(2m) \) was slower; that is,

\[
g(2m)/p(2m) \to 0 \quad \text{as} \quad m \to \infty
\]

(See [15].)

Like many other results about partitions, this was difficult to make progress on. In 1993, Erdős and Richmond [13] found a lower bound of \( C/\sqrt{m} \) for \( g(2m)/p(2m) \) where \( C \) may be taken asymptotic to \( \pi/\sqrt{12} \) as \( m \) gets large. Finally in 1997,
Pittel [15] showed Wilf’s conjecture to be true, though he could not give an explicit upper bound for $g(2m)/p(2m)$. (Empirical evidence suggests that this ratio decays only very slowly. Barnes and Savage [2] calculated $g(2m)/p(2m)$ for $m \leq 110$, and found the ratio decreased from 0.5 only to 0.3503 in this interval. More recently, Kohnert [8],[9] has calculated the ratio to $m = 585$; the ratio continues to decline monotonically, but only to 0.32188.)

**Labeled graphs on $m$ edges**

We might ask how many degree sequences there are for $m$-edge graphs, for large $m$. If we allow isolated vertices, there are infinitely many degree sequences for any $m$, all of which are distinct as labeled graphs. (For example, 11, 011, 0011, … are all possibilities for $m = 1$.) And if we fix the number of vertices $n$, then the answer is zero (for $m > \binom{n}{2}$). So in this case, we need to consider degree sequences of any length, but no zero entries. This problem has, to my knowledge, not been looked at yet.

**Unlabeled graphs on $n$ vertices: this thesis**

We will look at this case, and develop some asymptotic results, in this thesis. To my knowledge, the only previous work in this direction has been on listing and enumerating graphical partitions for specific values of $n$ (and $m$), with no real asymptotic results known. (See Stein [20], and Metropolis and Stein [10].)

We will begin, in this chapter, by giving some basic definitions and establishing trivial upper and lower bounds on the number of graphical partitions. (The upper bound is $\binom{2n-1}{n-1}$; the lower bound is about $4/3n$ times this.) In the second chapter, we will use an averaging technique to improve the lower bound by a factor of $\sqrt{n}$.
times a constant). In the third chapter, we will consider the graphical partition as a pair of random walks, and make some estimates of the probabilities that certain necessary conditions hold, improving the upper bound by a factor of \((\log n)^C\) for some constant \(C\). In the fourth and last chapter, we will discuss the accuracy of these bounds, and possibilities for further improvement.

## 1.2 Definitions

### 1.2.1 The Erdős-Gallai condition

To reiterate, a **partition** is a sequence of nonnegative integers (the *parts*) in nonincreasing order, such as \(\alpha = 4, 2, 2, 1, 0\) (but we will disregard trailing zeroes). A **graphical partition** is a partition whose parts can be interpreted as the degrees of the vertices of some graph. There are many equivalent tests for determining whether a given partition is graphical; one of them is the **Erdős-Gallai condition** [12] which is in many standard graph-theory textbooks (for example, [22], p. 42). This condition states that the partition \(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n\) is graphical if and only if (a) the sum of all parts is even and (b) for each \(k\),

\[
\lambda_1 + \cdots + \lambda_k \leq k(k - 1) + \min\{k, \lambda_{k+1}\} + \cdots + \min\{k, \lambda_n\}. \tag{1.1}
\]

We say that \(\alpha\) is **included in** \(\beta\) (\(\alpha \subseteq \beta\)) if every part \(\alpha_i\) is less than or equal to the corresponding part \(\beta_i\). We say that \(\alpha\) is **dominated by** \(\beta\) (\(\alpha \preceq \beta\)) if every partial sum \(\alpha_1 + \cdots + \alpha_i\) is less than or equal to the corresponding partial sum \(\beta_1 + \cdots + \beta_i\). (We do not require \(\alpha\) and \(\beta\) to sum to the same value.) These are the two standard order relations on partitions. (See [17] and [18], especially chapters 1 and 7, for more
background.)

Let’s visualize this condition (1.1). Start by visualizing our partition \( \lambda \) as a shape (Fig. 1-1), and move the squares on and to the right of the main diagonal (bolded) to make an “off-diagonal partition” (Fig. 1-2). (See [4]; [3], section 7.3, is also relevant.) Then the Erdős-Gallai condition (1.1) says that the total number of boxes in the first \( k \) rows is less than or equal to the total number of boxes in the first \( k \) columns. This still holds true if we ignore the first \( d - 1 \) boxes in each row and column (corresponding to ignoring the upper-left \( d \times d \) “Durfee square” in the off-diagonal partition). If we pick \( d \) to be as large as possible then the first rows and the first columns no longer overlap, and \( \lambda \) breaks up into two parts: One part (call it \( \alpha \)) is what’s left of the first \( d \) rows \((\lambda_1 - (d - 1), \lambda_2 - (d - 1), \ldots, \lambda_d - (d - 1))\); the other part (call it \( \beta' \) so its conjugate will be \( \beta \)) is the remaining \( n - d \) rows \((\lambda_{d+1}, \ldots, \lambda_n)\). Then the Erdős-Gallai condition (1.1) simply states that the partial sums of the rows of \( \alpha \) are less than or equal to the partial sums of the rows of \( \beta \) (the columns of \( \beta' \)) — that is, \( \alpha \preceq \beta \). (Note that the total number of boxes in \( \alpha \) and \( \beta \) together must be even.)

1.2.2 Enumerating graphical partitions

Using this, we can now rewrite all our graphical partitions in an easier-to-count form.

Call a pair of partitions \( \alpha \preceq \beta \), both contained in a single box, a dominance-pair. Denote the number of dominance pairs within a \( k \times l \) box by \( \#DP(k, l) \).

Call such a pair an even-dominance-pair if \( \alpha \) and \( \beta \) have the same parity, by which we will mean that both shapes contain an even, or both contain an odd, number of squares. Denote the number of even dominance pairs in \( k \times l \) by \( \#EDP(k, l) \).

If we can compute these, then the number of graphical partitions on \( n \) vertices
Figure 1-1: The partition $\lambda = 666443221$ can also be viewed as a shape, or as the path formed by the outside edge (in shadow). Its conjugate partition (swap rows and columns) is $\lambda' = 986533$. (The partitions included, $\mu \subseteq \lambda$, are precisely the shapes fitting inside the shape $\lambda$. The partitions dominated, $\nu \subseteq \lambda$, are harder to visualize, so we won't.)

Figure 1-2: The off-diagonal shape for $\lambda = 666443221$. Note that $d = 5$ (as shown) gives $\alpha = 222, \beta = 421$, but $d = 4$ is also possible.
follows. Let \( \lambda \) be a graphical partition on \( n \) vertices, draw it as an off-diagonal partition and remove a square of size \( d \times d \) from the upper left, leaving two partitions \( \alpha \) and \( \beta' \). Then \( \alpha \) and \( \beta \) form an even-dominance-pair within an \( (n - d) \times d \) box. So the number of graphical partitions on \( n \) vertices with a \( d \times d \) Durfee square (described above) is \( \#G_d(n) = \#EDP(n - d, d) \).

Unfortunately, if both \( \alpha \) and \( \beta \) have squares in fewer than \( d \) rows, then we could have used a \( (d - 1) \times (d - 1) \) Durfee square instead. So to avoid overcounting in our total number of graphical partitions, we will need to subtract out \( \#G'_d(n) = \#EDP(n - d, d - 1) \).

Hence, the total number of graphical partitions \( \#G(n) \) is

\[
\#G(n) = \sum_{0 \leq d \leq n} (\#G_d(n) - \#G'_d(n)) = \sum_{0 \leq d \leq n} (\#EDP(n - d, d) - \#EDP(n - d, d - 1)).
\]  

(1.2)

If we just want a lower bound, though, we can take a simpler approach: if we sum all the \( \#G_d(n) \)-terms we get an overestimate, but since no graphical partition is counted more than twice we may simply divide by 2 to get an underestimate for the true value. That is,

\[
\#G(n) \geq \frac{1}{2} \sum_{0 \leq d \leq n} \#G_d(n) = \frac{1}{2} \sum_{0 \leq d \leq n} \#EDP(n - d, d)
\]  

(1.3)

where \( \#EDP(n - d, d) \) is the number of even dominance-pairs within a \( (n - d) \times d \) box.
1.3 Trivial bounds

The trivial bounds for the number of dominance pairs (and hence the number of graphical partitions) are actually quite good; these will serve as a standard of comparison for our new bounds.

1.3.1 A trivial upper bound

A graphical partition is a partition of at most \( n \) parts, each of length at most \( n - 1 \); there are \( \binom{2n-1}{n-1} \) partitions meeting these requirements. Since \( \binom{2n}{n} = 2\binom{2n-1}{n-1} \) is asymptotic to \( 4^n/\sqrt{\pi n} \), this upper bound is asymptotic to half that, or \( 4^n/2\sqrt{\pi n} \):

\[
\#G(n) \leq \binom{2n-1}{n-1} \sim \frac{4^n}{2\sqrt{\pi n}} \tag{1.4}
\]

A trivial upper bound for the number of dominance-pairs in a \( k \times l \) box is the total number of pairs of partitions, whether related by dominance or not. Since there are \( \binom{k+l}{k} \) partitions that fit within such a box, there are \( \binom{k+l}{k}^2 \) pairs in all. This is also, of course, an upper bound on the number of even dominance-pairs:

\[
\#EDP(k,l) \leq \#DP(k,l) \leq \binom{k+l}{k}^2 \tag{1.5}
\]

Remark

We can also try using the trivial upper bound on dominance-pairs to obtain an upper bound on graphical partitions:

\[
\#G(n) = \sum_{0 \leq d \leq n} \#EDP(n-d,d) - \#EDP(n-d,d-1)
\]
Figure 1-3: Counting inclusion-pairs in a $k \times l$ box as noncrossing paths

\[
\leq \sum_{0 \leq d \leq n} \#EDP(n - d, d)
\]

\[
\leq \sum_{0 \leq d \leq n} \binom{n}{d}^2
\]

\[
= \binom{2n}{n}
\]

(See [17], Example 1.1.17.) This is actually worse than our original upper bound of \(\binom{2n-1}{n-1}\).

1.3.2 A trivial lower bound

Dominance-pairs

A trivial lower bound for the number of dominance-pairs $\alpha \preceq \beta$ within our $k \times l$ reference box is the number of inclusion-pairs $\alpha \subseteq \beta$, since inclusion implies dominance. (We will denote this number by $\#IP(k, l)$.) Actually counting these inclusion pairs is a simple application of the well-known Gessel-Viennot technique [17] for counting noncrossing paths.

We calculate this answer as follows. Draw $\alpha \subseteq \beta$ as paths on a $k \times l$ grid, then
displace $\beta$ by one unit to the south and east so that $\alpha$ runs from $(0,l)$ to $(k,0)$ and $\beta'$ runs from $(1,l+1)$ to $(k+1,1)$; these paths are noncrossing if and only if the original paths were an inclusion-pair. Now Gessel-Viennot tells us the total number of noncrossing path-pairs is

$$\#IP(k,l) = \det \begin{vmatrix} \binom{k+l}{k} & \binom{k+l}{k-1} \\ \binom{k+l}{k+1} & \binom{k+l}{k} \end{vmatrix} = \binom{k+l}{k}^2 \frac{k+l+1}{(k+1)(l+1)}$$

where the $(i,j)$ entry of the determinant is the number of paths from the $i$th starting-point to the $j$th ending point.

**Even dominance-pairs**

One might guess that about half of these have the correct parity to qualify as even dominance-pairs, but we will only argue $\#EDP(k,l) \geq \#IP(k,l)/3$. We associate every inclusion-pair with an even dominance-pair in such a way that no even dominance-pair in such a way that no even dominance-pair has more than three inclusion-pairs associated with it; this will then follow.

Specifically, take inclusion-pair $\alpha \subseteq \beta$, and:

1. If $\alpha$ and $\beta$ have the same parity, do nothing.

2. If not, add a square to the shortest row of $\alpha$.

This always yields a dominance-pair, though not necessarily an inclusion-pair, and is clearly even in parity. (It fails to be an inclusion-pair in case (2), if we augment an $\alpha_i$ which is already the same length as $\beta_i$. But since $\alpha$ and $\beta$ have different parities, we know $\alpha$ has fewer squares than $\beta$, and since $\alpha_i$ is the shortest row, and $\beta_i$ has the same length, we know the extra square(s) can’t be in row $i$—or any later row.
\( j > i \), since \( \beta_j \) can never be less than \( \alpha_j = \alpha_i = \beta_i \). Hence there’s an extra square in an earlier row of \( \beta \), which cancels out the added square in \( \alpha_j \) for all partial sums affected.)

Given an even dominance-pair \( \alpha' \preceq \beta \) we can reverse the process in at most three ways:

1. Do nothing.

2. Remove a square from the shortest row in \( \alpha' \)—which must be the last such row, in case of ties, since \( \alpha \) was a partition.

3. Remove a square from the row of second-shortest length in \( \alpha' \)—again, this must be the last such row.

These correspond to the cases, respectively, where \( \alpha \) and \( \beta \) already had matching parity; where \( \alpha \) had the wrong parity and there was only one row of shortest length in \( \alpha \); and where \( \alpha \) had the wrong parity and there was more than one row of shortest length in \( \alpha \).

So: the number of dominance-pairs that fit in a \( k \times l \) reference box has the trivial lower bound

\[
\#DP(k, l) \geq \#IP(k, l) = \binom{k + l}{k}^2 \frac{k + l + 1}{(k + 1)(l + 1)}
\]  

and the number of even dominance-pairs, the trivial lower bound

\[
\#EDP(k, l) \geq \frac{\#IP(k, l)}{3} = \frac{1}{3} \binom{k + l}{k}^2 \frac{k + l + 1}{(k + 1)(l + 1)}. 
\]
Graphical partitions

We can then (under)count graphical partitions using (1.3) and our estimate (1.7) above to get

\[
\#G(n) \geq \frac{1}{2} \sum_{0 \leq d \leq n} \frac{1}{3} \binom{n}{d}^2 \frac{n + 1}{(n - d + 1)(d + 1)}
\]

\[
= \frac{1}{6(n + 1)} \sum_{0 \leq d \leq n} \binom{n + 1}{d} \binom{n + 1}{d + 1}
\]

\[
= \frac{1}{6(n + 1)} \binom{2n + 2}{n}.
\]

([17], 1.1.17, again.) Compare this to our upper bound of \(\binom{2n-1}{n}\):

\[
\lim_{n \to \infty} \frac{1}{6(n + 1)} \frac{\binom{2n + 2}{n}}{\binom{2n - 1}{n}} = \frac{4}{3n}
\]

so our trivial lower bound is a factor of \(4/3n\) worse than our trivial upper bound; they are asymptotic to \(4^n/Cn^{3/2}\) and \(4^n/Cn^{1/2}\) respectively.
Chapter 2

An Improved Lower Bound

2.1 Strategy

Any pair of partitions \( \alpha, \beta \) in our \( k \times l \) reference box can be expressed equivalently as (i) an inclusion pair \( \lambda \subseteq \mu \), where \( \lambda \) is the intersection of, and \( \mu \) the union of, our two partitions \( \alpha, \beta \); together with (ii) a + or − sign on each component of \( \mu/\lambda \), indicating which of \( \alpha \) or \( \beta \) is the outside edge of that component.

Put another way, we can assemble our partition-pair in the following manner:

1. Choose the number of components \( r \).

2. Choose \( r \) single-component skew shapes \( \gamma^{(1)}, \ldots, \gamma^{(r)} \). (We may suppose for future reference they have respective dimensions \( (k_1+1, l_1+1), \ldots, (k_r+1, l_r+1) \), where the \( k_i \) and \( l_i \) are nonnegative.) There are \( \binom{k_i+l_i}{k_i}^2 \frac{k_i+l_i+1}{(k_i+1)(l_i+1)} \) ways to do this, using another Gessel-Viennot argument.

3. Choose a sign \( \epsilon_i \) for each component (+ or −). There are \( 2^r \) ways to do this in total, but if we only wanted dominance pairs we would need a more restrictive...
Figure 2-1: Decomposing partition-pairs as inclusion pairs. Here $\beta = 5441$ and $\alpha = 43322$, so $\lambda = 4331$, $\mu = 54422$. Our $r = 3$ components are $\gamma^{(1)} = 1$ (sign $+$), $\gamma^{(2)} = 11$ (sign $+$), and $\gamma^{(3)} = 22/1$ (sign $-$). (This is a dominance pair $\alpha \preceq \beta$. Notice we can pair every ($-$) with a ($+$) in an earlier row.)

condition. We will do so later.

4. Choose a single partition shape for the remaining $k_0$ columns and $l_0$ rows that we will get upon removing these $r$ components. There are $\binom{k_0 + l_0}{k_0}$ ways to do this. (If our overall box is $k \times l$, then $k_0 = k - (k_1 + 1) - \cdots - (k_r + 1)$, and likewise for $l_0$.) This uniquely determines the portion of the shapes $\mu$, $\lambda$ where the two shapes coincide.

5. Choose $r$ positions along this shape’s outer edge in which to insert these $r$ components. There are $\left(\binom{k_0 + l_0}{r}ight)$, or $\binom{k_0 + l_0 + r}{r}$, ways to do this. This, together with the number, shape, and sign of our components, uniquely determines the portions of the shapes $\lambda$, $\mu$ where the two shapes are different (= form a skew shape component).

So, for example, we could write the trivial upper bound $\binom{k+l}{k}^2$ (from (1.5)) as the
impressively cumbersome sum

\[ \sum_{r \geq 0} 2^r \sum_{k_0 \leq k-r \atop l_0 \leq l-r} \binom{k_0 + l_0}{k_0} \left( \binom{k_0 + l_0 + 1}{r} \right) \sum_{k_1 + \ldots + k_r = k-r-k_0 \atop l_1 + \ldots + l_r = l-r-l_0} \prod_i \binom{k_i + l_i}{k_i}^2 \frac{k_i + l_i + 1}{(k_i + 1)(l_i + 1)} \]

\[(2.1)\]

The trivial lower bound (for dominance pairs, (1.6)) has the same expansion, except that since \(\beta\) is always the outside partition, there is only one permissible sign choice (all +) and we replace the \(2^r\) with 1:

\[ \sum_{r \geq 0} 1 \sum_{k_0 \leq k-r \atop l_0 \leq l-r} \binom{k_0 + l_0}{k_0} \left( \binom{k_0 + l_0 + 1}{r} \right) \sum_{k_1 + \ldots + k_r = k-r-k_0 \atop l_1 + \ldots + l_r = l-r-l_0} \prod_i \binom{k_i + l_i}{k_i}^2 \frac{k_i + l_i + 1}{(k_i + 1)(l_i + 1)} \]

\[(2.2)\]

In short: if we want arbitrary partition-pairs, we choose our signs \(\epsilon_i\) arbitrarily, and if we want inclusion-pairs, we need to choose all \(\epsilon_i = 1\). This suggests we look at the corresponding criterion for dominance pairs, and try to find the (average) number of signs possible:

\[ \sum_{r \geq 0} f(r) \sum_{k_0 \leq k-r \atop l_0 \leq l-r} \binom{k_0 + l_0}{k_0} \left( \binom{k_0 + l_0 + 1}{r} \right) \sum_{k_1 + \ldots + k_r = k-r-k_0 \atop l_1 + \ldots + l_r = l-r-l_0} \prod_i \binom{k_i + l_i}{k_i}^2 \frac{k_i + l_i + 1}{(k_i + 1)(l_i + 1)} \]

\[(2.3)\]

We will show that \(f(r)\) is, on average, at least \((2r - 1)!!/r!\).

### 2.2 Permissible signs for dominance pairs

Let us suppose we have chosen \(r\) components \(\gamma^{(1)}, \ldots, \gamma^{(r)}\), in that order from top to bottom, with respective sizes \(c_1, \ldots, c_r\) and signs \(\epsilon_1, \ldots, \epsilon_r\). (That is, \(c_j\) is the number of boxes in \(\gamma^{(j)}\).)

By definition, the criterion for dominance \(\alpha \preceq \beta\) is that the partial sums of \(\alpha\) are
no greater than the partial sums of $\beta$: $\alpha_1 + \cdots + \alpha_i \leq \beta_1 + \cdots + \beta_i$, or rephrased slightly, $(\beta_1 - \alpha_1) + \cdots + (\beta_i - \alpha_i) \geq 0$.

We only need to check this for the last row of each component: if the $i$th row’s contribution $\beta_i - \alpha_i$ is nonnegative then the $i$th-row inequality follows from the previous row, and if the $i + 1$st row’s contribution will be nonpositive then it will follow from the next row. So the only partial sums we need check are where the sign of this contribution changes from negative to nonnegative: that is, rows corresponding to the last rows of components $\gamma^{(j)}$ with negative signs $\epsilon_j$. (We may further require the next component if it exists (i.e., $j < r$) to have positive sign $\epsilon_{j+1}$.)

So the condition (given these components) for having a dominance-pair is that the partial sums $\epsilon_1 c_1; \epsilon_1 c_1 + \epsilon_2 c_2; \ldots; \epsilon_1 c_1 + \cdots + \epsilon_r c_r$ are all nonnegative.

This suggests that we can improve our lower bound by focusing on the number of permitted signs, and in fact we can. Now, for any choice of $r$ components $\gamma^{(1)}, \ldots, \gamma^{(r)}$, the number of permitted signs may be as high as $2^r$ or as low as $1$ (consider the cases where $c_1 \gg c_2 \gg \cdots \gg c_r$ and where $c_1 \ll c_2 \ll \cdots \ll c_r$, respectively). But for any choice of $r$ components, we will see that the total number of permitted signs over all $r!$ permutations of these components is at least $(2r - 1)!!$, hence we may substitute the average value, $\frac{(2r-1)!!}{r!}$, for the $1$ in the lower bound.

### 2.3 Number of signs-and-orderings:

#### the special case

We need only consider the component sizes $c_i$, and since we will be totaling over all permutations, we may without loss of generality assume $c_1 \geq c_2 \geq \cdots \geq c_r$. Note also that we do not need to require the $c_i$ to be integers, as the inequality conditions
make sense for all real numbers. (But we will assume all $c_i$ are nonnegative, for convenience.)

It seems hard to list all possible ways of signing-and-ordering the $c_i$ in general, so we will consider the special case where $c_1 \gg c_2 \gg \cdots \gg c_r$ — more explicitly, we require that $c_i > c_{i+1} + \cdots + c_r$ for each $i$. We will prove by induction that, in this case, there are exactly $N(r) = (2r - 1)!! = (2r - 1)(2r - 3)\cdots(3)(1)$ permitted combinations of signs-and-orderings.

The partial sums we need to verify are of the form $\epsilon_1 c_{\pi(1)} + \cdots + \epsilon_j c_{\pi(j)}$ for $j = 1, 2, \ldots, r$; under the conditions above, this is positive exactly when the sign on the largest $c_i$ is positive. So, for example, if a given permutation $\pi$ places $c_1$ as the first term $c_{\pi(1)}$, $\epsilon_1$ is required to be positive, and this suffices for all partial sums. More generally, wherever $c_1$ appears, its sign must be positive, and this suffices for all partial sums that include the term $c_1$.

Hence, if $c_1$ appears in the $j$th position (so $\pi(j) = 1$ and $\epsilon_j = +1$), the $r - j$ elements $c_i$ following can be placed in any order ($(r - j)!$ orderings) with any choice of signs ($2^{r-j}$ ways), and the $j - 1$ elements preceding can be signed-and-ordered in $N(j - 1) = (2(j - 1) - 1)!!$ ways by our induction hypothesis. We also need to select these $r - j$ elements, which we can do in $\binom{r-1}{r-j}$ ways, independent of these other choices. So we have a recurrence relation for the number $N(r)$ of signs-and-orderings, namely

$$N(r) = \sum_{1 \leq j \leq r} \binom{r-1}{j-1} N(j - 1) \cdot 2^{r-j}(r - j)!!$$

and we need only check that $N(r) = (2r - 1)!!$ satisfies it. (The initial conditions $N(0) = 1, N(1) = 1 = (1)!!$ are trivially satisfied.) This is routine:
\[ (2r + 1)N(r) \]
\[ = 2rN(r) + N(r) \]
\[ = 2r \left( \sum_{1 \leq j \leq r} \frac{r - 1}{r - j} N(j - 1) \cdot 2^{r-j}(r - j)! \right) + N(r) \]
\[ = \left( \sum_{1 \leq j \leq r} 2(r + 1 - j) \cdot \frac{r}{r + 1 - j} \cdot \left( \frac{r - 1}{r - j} \right) N(j - 1) \cdot 2^{r-j}(r - j)! \right) + N(r) \]
\[ = \left( \sum_{1 \leq j \leq r} \frac{r}{r + 1 - j} N(j - 1) \cdot 2^{r+1-j}(r + 1 - j)! \right) + \left( \frac{r}{r} \right) N(r) \cdot 2^0 \cdot 1! \]
\[ = \sum_{1 \leq j \leq r+1} \left( \frac{r}{r + 1 - j} \right) N(j - 1) \cdot 2^{r+1-j}(r + 1 - j)! \]
\[ = N(r + 1). \]

### 2.4 Number of signs-and-orderings:

the general case

Consider all the hyperplanes of the form \( c_{i_1} + \cdots + c_{i_a} = c_{j_1} + \cdots + c_{j_b} \) where all the indices \( i_1, \ldots, i_a, j_1, \ldots, j_b \) are distinct elements of \( \{1, \ldots, r\} \) (if to distinguish the two sides we assume \( i_1 < j_1 \), then each hyperplane is represented uniquely). These divide the \( r \)-dimensional space of possible (real) component sizes \( \vec{c} \) into chambers. And since all the partial sums we might need to verify are of this form (with \( \leq \) or \( \geq \) replacing the equality sign), within each chamber the set of signs-and-orderings possible doesn’t change, and in particular its size is constant.

We claim that this constant is equal for all chambers, hence equal to \( (2r - 1)!! \) as it is for the principal chamber \( c_1 \gg c_2 \gg \cdots \gg c_r \).
To see this, consider two chambers separated by only one hyperplane; they thus satisfy all the same relevant inequalities except for one. The only signs-and-orderings that are valid in chamber $A$ (where $c_i + \cdots + c_a > c_{j_1} + \cdots + c_{j_b}$) that are not valid in chamber $B$ (where $c_i + \cdots + c_a < c_{j_1} + \cdots + c_{j_b}$) are those having as a partial sum the quantity $P_{a+b}(\vec{c}) = (c_i + \cdots + c_a) - (c_{j_1} + \cdots + c_{j_b})$. In particular, the $c_i, \ldots$ appear with positive sign, and the $c_{j_1}, \ldots$ with negative sign, in some order as an initial sequence in such a sign-and-ordering. Conversely, the signs-and-orderings valid in chamber $B$ but not in chamber $A$ have the same $c_{j_1}, \ldots$ with negative sign, and the $c_i, \ldots$ with positive sign. So we can establish a bijection between these two subsets just by reversing the order and sign of the first $a+b$ indexes of $\vec{c}$.

Let $P_i(\vec{c})$ be the partial sums of a sign-and-ordering for chamber $A$ (so for any $\vec{c}_A \in A$, we have $P_i(\vec{c}_A) > 0$), and $P'_i(\vec{c})$ be the partial sums of the resulting sign-and-ordering (so we need to verify that for any $\vec{c}_B \in B$, we have $P_i(\vec{c}_B) > 0$). We may assume for convenience that $P_{a+b}(\vec{c})$ is at least twice as small as the other partial sums $P_i(\vec{c})$, i.e. that our points $\vec{c}_A \in A$ and $\vec{c}_B \in B$ are at least twice as close to the hyperplane $P_{a+b}(\vec{c}) = 0$ than to any other hyperplane. Then for $i < a + b$ we have $P'_i(\vec{c}_B) = P_{a+b-i}(\vec{c}_B) - P_{a+b}(\vec{c}_B) > 0$, for $i > a + b$ we have $P'_i(\vec{c}_B) = P_i(\vec{c}_B) - 2P_{a+b}(\vec{c}_B) > 0$, and of course for $i = a + b$ we have $P'_i(\vec{c}_B) = -P_i(\vec{c}_B) > 0$.

Remark

We have only shown $(2r-1)!!$ is the number of signs-and-orderings when the component sizes $\vec{c}$ lie within one of the chambers defined above. In general, this is only a lower bound.

If the point corresponding to our $c_i$ is not in any chamber, but lies on one or more hyperplanes, then the number of permissible signs-and-orderings is strictly
greater than \((2r - 1)!!\), since we can take the strict equalities to have either sign as an inequality (either \(\leq\) or \(\geq\)), and hence that point’s set of permissible signs-and-orderings is the union of all the adjacent chambers’ sets.

### 2.5 Calculating the order of growth

Comparing our upper bound \(\binom{k+l}{k}^2\) (as expressed in Equation (2.1)) to our lower bound (identical, but with \(2^r = (2r)!!/r!\) replaced by \((2r - 1)!!/r!\)), we see that all we’ve done is multiply the terms for a given \(r\) by a factor of

\[
A := \frac{(2r - 1)!!}{(2r)!!} = \frac{1}{2^r} \cdot \frac{3}{2^{r-1}} \cdots \frac{2r-1}{2r}.
\]

In other words, if we calculate the rate of growth of \(A\), we will know the rate of growth of our lower bound, as compared to the (known) rate of growth of our trivial upper bound.

Since we will want bounds on \(A\) even for small \(r\), we will save some work by avoiding Stirling’s approximation (and error estimates) in favor of the following more elementary approach.

Rather than working with \(A\) directly, we will estimate

\[
\log(A) = \log \left(1 - \frac{1}{2}\right) + \log \left(1 - \frac{1}{4}\right) + \cdots + \log \left(1 - \frac{1}{2r}\right).
\]

First estimate each term \(\log(1 - 1/2k)\). We know from basic calculus that

\[-\log(1 - x) = x + x^2/2 + x^3/3 + \cdots ,\]

and this is clearly greater than \(x + x^2/2\) for \(x > 0\). On the other hand, since \(x < 1/2\),
we have the upper bound

\[ x + x^2/2 + x^3/3 + \cdots + x^k/k + \cdots \leq x + x^2/2 + x^3/2 + \cdots + x^k/2 + \cdots, \]

which sums to \( x + x^2/2 \cdot (1 + x^2 + \cdots) \leq x + x^2/2 \cdot 2 \). So for each term we have the bounds

\[-\frac{1}{2k} - \left( \frac{1}{2k} \right)^2 \leq \log \left( 1 - \frac{1}{2k} \right) \leq -\frac{1}{2k} - \frac{1}{2} \left( \frac{1}{2k} \right)^2.\]

Now sum this from \( k = 1 \) to \( r \) to get \( \log(A) \). The sum of the quadratic terms

\[ \sum_{1 \leq k \leq r} \left( \frac{1}{2k} \right)^2 \]

is clearly bounded above by

\[ \sum_{1 \leq k} \left( \frac{1}{2k} \right)^2 = \frac{1}{4} \cdot \frac{\pi^2}{6} < 0.412, \]

and below by its first term \( 1/4 \). The sum of the linear terms

\[ \sum_{1 \leq k \leq r} \frac{1}{2k} \]

is between \((\log r)/2\) and \((1 + \log r)/2\). Taking the lower bounds on the left and the upper bounds on the right, we thus have

\[-\frac{1}{2} (1 + \log r) - 0.412 < \log(A) < -\frac{1}{2} \log(r) - \frac{1}{4}\]

or

\[ e^{-0.912r^{-1/2}} < A < e^{-1/4}r^{-1/2}. \]
If we’re counting dominance-pairs inside an $k \times l$ box, then $r$ is less than or equal to the smaller of $k$ and $l$, so our lower bound is within a factor of $e^{-0.912}/\sqrt{\min(k,l)}$ of the upper bound. This is the best we can do without considering different values of $r$ separately, up to a constant factor of at most $\exp(0.912 - 1/4) \approx 1.939 < 2$.

(For large $r$, the actual constant is asymptotic to $1/\sqrt{2\pi}$, since $A \sim 4^{-r}(2^r)$.)

### 2.6 Even dominance-pairs

One might guess, as before, that about half of these dominance-pairs $\alpha \trianglelefteq \beta$ have the correct parity to qualify as even dominance-pairs, but we will only argue that $\#EDP(k,l) \geq \#DP(k,l)/5.5$. This time, we associate every dominance-pair with an even dominance-pair in such a way that no even dominance-pair has more than $5.5$ dominance-pairs associated with it.

As before, let $r$ be the number of components in $\beta/\alpha$. If $r = 0$, that is if $\alpha = \beta$, then our dominance-pair is already even; do nothing. So we may assume $r > 0$. Now look at the sizes of each component $\beta/\alpha$. If there are an even number of odd components, then again our dominance-pair is already even; do nothing. So we may assume there are an odd number of odd components, and now we will have to do something.

Consider the last component in order from top to bottom, $\gamma^{(r)}$, of size $c_r > 0$. Since there are an odd number of odd components, the partial sum $\epsilon_1 c_1 + \cdots + \epsilon_r c_r$ is odd, so we may decrease (or increase) $c_r$ by one square without violating the dominance-pair property. We do so as follows. (In all cases $j$ is the last row of component $\gamma^{(r)}$, that is, the last $j$ for which $\alpha_j < \beta_j$.)

1. If $\beta_j > \alpha_j + 1$, add a square to the first row of $\alpha$ having length $\alpha_j$, hence
2. If $\beta_j = \alpha_j + 1$ and $\alpha_{j-1} = \alpha_j$, then remove a square from the $j$th row of $\beta$ (that is, $\beta_j' = \beta_j - 1$), hence shrinking $\gamma^{(r)}$ by one square.

3. If $\beta_j = \alpha_j + 1$ and $\alpha_{j-1} > \alpha_j + 1$, then $\beta_{j-1} > \alpha_j + 1$ also (note that $\gamma^{(r)}$ consists of a single square in this case) and we may increment $\beta_j' = \beta_j + 1$, adding a square to $\beta$ (hence to $\gamma^{(r)}$).

4. If $\beta_j = \alpha_j + 1$ and $\alpha_{j-1} = \alpha_j + 1$ and $\beta_{j-1} = \alpha_{j-1}$ (note that $\gamma^{(r)}$ consists of a single square in this case) then we may decrease $\alpha_{j-1}' = \alpha_{j-1} - 1$, adding a square to $\gamma^{(r)}$.

5. Finally, if $\beta_j = \alpha_j + 1$ and $\alpha_{j-1} = \alpha_j + 1$ and $\beta_{j-1} > \alpha_{j-1}$ (note that $\gamma^{(r)}$ consists of a single square in this case) then component $\gamma^{(r-1)}$ ends at (row $j - 1$, column $\alpha_{j-1}$), and we will be able to locate the single-square $\gamma^{(r)}$ if we delete it. Do so in two different ways, each having weight $1/2$: either reduce $\beta_j$ by 1, or increase $\alpha_j$ by 1.

Now, given an even dominance-pair, in general there is:

- at most 1 way to leave it alone,
• at most 2 ways to reduce $\alpha$ (either we shrink the shortest row in the last component, or shrink the next-shortest row so it becomes equal in length to the shortest),

• at most 2 ways to increase $\beta$ (add a square to the last row in the last component, or to the row after it),

• and at most 1/2 way to add a component (of size 1) that got deleted,

for a total of 5 1/2.

If the last component consists of exactly two squares stacked vertically, then there is:

• at most 1 way to leave it alone,

• at most 1 way to reduce $\alpha$ (we can’t shrink the next-shortest row),

• at most 1 way to increase $\beta$ (we can’t add a square to the last row in the last component),

• at most 1 way to increase $\alpha$ (add the top square),

• and at most 1/2 way to add a component (of size 1) that got deleted.

for a total of 4 1/2, and similarly if it consists of exactly two horizontal squares.

Our conclusion then follows: and since

$$\#DP(k, l) \geq \frac{e^{-0.912}}{\sqrt{\min(k, l)}} \cdot \left(\frac{k + l}{k}\right)^2$$

we have

$$\#EDP(k, l) \geq \frac{1}{5.5} \frac{e^{-0.912}}{\sqrt{\min(k, l)}} \cdot \left(\frac{k + l}{k}\right)^2.$$
2.7 Graphical partitions

We can then (under)count graphical partitions using (1.3) and our estimate above to get

\[
\#G(n) = \sum_{0 \leq d \leq n} \#EDP(n - d, d) - \#EDP(n - d, d - 1)
\]

\[
\geq \frac{1}{2} \sum_{0 \leq d \leq n} \#EDP(n - d, d)
\]

\[
\geq \frac{1}{2} \sum_{0 \leq d \leq n} \frac{1}{5.5} \frac{e^{-0.912}}{\min(n - d, d)} \cdot \binom{n}{d}^2
\]

\[
\geq \frac{1}{11e \sqrt{n}} \sum_{0 \leq d \leq n} \binom{n}{d}^2
\]

\[
= \frac{4^n}{11e \sqrt{n}} \frac{1}{(2n - 1)!!}
\]

This last is just \(\frac{1}{11e \sqrt{n}}\) times our earlier upper bound for \(\#G(n)\), which had an order of \(\frac{4^n}{\sqrt{n}}\), so our new lower bound has order \(4^nn^{-1}\), which is an improvement over our earlier lower bound, which had order \(4^nn^{-3/2}\).
Chapter 3

Improving the Upper Bound

3.1 Strategy

Let us represent a graphical partition of $n$, as before, as a pair of partitions $\alpha$, $\beta$, each fitting in some box of size $(n-d) \times d$. As before, the conditions for this partition-pair to represent a graphical partition of $n$ are

1. $\sum \alpha + \sum \beta$ is even, and
2. $\alpha \preceq \beta$.

There are $2^n$ possible partitions fitting inside a box $(n-d) \times d$ for some $d$; choosing $\alpha$ uniformly at random from this set is the same as choosing an $n$-step random walk with $(n-d)$ “right” and $d$ “up” steps. (For more on random walks, see [6], especially chapter 3.)

Our strategy will be to represent $\alpha$ and $\beta$ as infinite random walks, extending the idea of dominance in a natural way. Then we will look at the state of the random walks after $m$ steps, for various values of $m$, and find conditions that occur...
with nonzero probability and guarantee us that dominance fails. By taking enough widely-spaced values of \( m \), we can then make the probability that dominance succeeds approach zero. Finally, to carry the infinite random walks back to the finite case, we need only truncate \( \alpha \) and \( \beta \) after \( n \) steps, taking care that

3. both \( \alpha \) and \( \beta \) end at the same point \((n - d, d)\)

also holds (with, again, some probability we can determine).

For reference, we now define some random variables for the (infinite) random walk \( \alpha \):

\[
\begin{align*}
X_m^{(\alpha)} &= \pm 1 \text{ with probability } 1/2 \\
Y_m^{(\alpha)} &= X_1 + X_2 + \ldots + X_m \\
Z_m^{(\alpha)} &= Y_1 + Y_2 + \ldots + Y_m \\
&= mX_1 + (m-1)X_2 + \ldots + X_m
\end{align*}
\]

where the \( X_j \) are the random steps (we will think of +1 as an “up” and −1 as a “right” step) and \( Y_m \) is the usual random walk (representing our position as we cross the \( m \)th diagonal \( k+l = m \), ranging by twos from \( +m \) for all up steps \((k = m, l = 0)\) to \( -m \) for all rightward steps \((k = 0, l = m)\)). \( Z_m \) then corresponds roughly to the area above this random walk; we will use it mainly to define a fourth random variable

\[
U_m^{(\alpha)} = 2Z_m^{(\alpha)} - (m+1)Y_m^{(\alpha)}
= (m-1)X_1^{(\alpha)} + (m-3)X_2^{(\alpha)} + \cdots + (3-m)X_{m-1}^{(\alpha)} + (1-m)X_m^{(\alpha)}
\]

which is (almost) independent of \( Y_m^{(\alpha)} \).

It is worth noting that \( U_m \) is less than \( m^2 \) in absolute value.
Since dominance is defined in terms of the largest parts of our partitions, it will be convenient to keep these largest parts in a consistent place: so in this chapter we will rotate the usual shape of $\alpha$ 180 degrees, to be the region below the path $\alpha$.

Even when $\alpha$ and $\beta$ are infinite paths and the rows $\alpha_i$, $\beta_i$ are infinitely long, the differences $\beta_i - \alpha_i$ are still well-defined, as the (signed) number of squares between the two paths in that row. So we’ll define $\alpha \leq \beta$, in the infinite case, to mean that $(\beta_1 - \alpha_1) + \ldots + (\beta_i - \alpha_i)$ is nonnegative for all $i$. Note that these differences $(\beta_i - \alpha_i)$ aren’t affected when we truncate our paths, provided both paths are still defined on that row. (So “dominance” for the infinite box will indeed carry over to the finite $(n - d) \times d$ case.)

### 3.2 The infinite box

Now we consider two infinite random walks $\alpha$, $\beta$, as described above. What happens when we truncate them at, say, the $m$th diagonal?

The “triangle” below $\alpha$ and below this $m$th diagonal has area $Z_m^{(\alpha)}/2 - Y_m^{(\alpha)}/4 + m^2/4$. (Argue as follows: If instead of following the path $\alpha$ we traveled due northeast
from (0, 0) along the line $k = l$ to the $m$th diagonal, the area of the triangle would be 
$(m/\sqrt{2})^2/2$, yielding the $m^2/4$ term. The net area enclosed by our partition relative to that line is given by the trapezoidal rule: width is $m/\sqrt{2}$, and height is $1/\sqrt{2}$ times 
$\left(\frac{1}{2}Y_0 + Y_1 + \ldots + Y_{m-1} + \frac{1}{2}Y_m\right)/m$, or $Z_m/m\sqrt{2} - Y_m/(2m\sqrt{2})$.

Suppose $\alpha$ ends at point $(K, L)$ (corresponding to $Y_m = L - K$ or $K = (m - Y_m)/2$, $L = (m + Y_m)/2$); then the area in the $K \times L$ box below $\alpha$ is

$$
\text{Area}(\alpha|_{K \times L}) = \left(\frac{Z_m^{(\alpha)}}{2} - \frac{Y_m^{(\alpha)}}{4} + \frac{m^2}{4}\right) - \frac{L^2}{2} \\
= \left(\frac{Z_m^{(\alpha)}}{2} - \frac{Y_m^{(\alpha)}}{4} + \frac{m^2}{4}\right) - \left(\frac{m^2}{8} + \frac{mY_m}{4} + \frac{(Y_m^{(\alpha)})^2}{8}\right)
$$
\[
\begin{align*}
&= \left( \frac{Z_m^{(\alpha)}}{2} - (m+1)\frac{Y_m^{(\alpha)}}{4} \right) + \frac{m^2}{8} - \frac{(Y_m^{(\alpha)})^2}{8} \\
&= \frac{1}{4} \left( 2Z_m^{(\alpha)} - (m+1)Y_m^{(\alpha)} \right) + \frac{1}{8}(m^2 - (Y_m^{(\alpha)})^2) \\
&= \frac{1}{4} U_m^{(\alpha)} + \frac{1}{8}(m^2 - (Y_m^{(\alpha)})^2).
\end{align*}
\]

Likewise for \( \beta \): supposing \( \beta \) ends at \((K', L')\) we have in the \( K' \times L' \) box

\[
\text{Area}(\beta_{|K' \times L'}) = \frac{1}{4} U_m^{(\beta)} + \frac{1}{8}(m^2 - (Y_m^{(\beta)})^2).
\]

This isn’t quite the \( K \times L \) box we really want, but it’s close. In particular, if \( Y^{(\alpha)} < Y^{(\beta)} \), then \( \text{Area}(\beta_{|K' \times L'}) > \text{Area}(\beta_{|K \times L}) \), because in going from \( K' \times L' \) to \( K \times L \) the columns we add have length zero (and the rows we remove can only decrease area).

If in addition \( U^{(\alpha)} > U^{(\beta)} \), then \( \text{Area}(\alpha_{|K \times L}) - \text{Area}(\beta_{|K' \times L'}) = \frac{1}{4} (U_m^{(\alpha)} - U_m^{(\beta)}) + \frac{1}{8}((Y_m^{(\alpha)})^2 - (Y_m^{(\beta)})^2) > 0 \), so \( \text{Area}(\alpha_{|K \times L}) > \text{Area}(\beta_{|K' \times L'}) > \text{Area}(\beta_{|K \times L}) \) and, in particular, \( \alpha_1 + \cdots + \alpha_L > \beta_1 + \cdots + \beta_L \), which demonstrates that \( \alpha \not\sim \beta \) and hence that the corresponding partition is not graphical. If we do this enough times, we’ll get our improved lower bound.

### 3.3 Limiting distributions for \( U \) and \( Y \)

We first estimate the probability distribution for \( U_m \) and \( Y_m \) for large \( m \). \( Y_m \) is easy: it’s a sum of independent, identically distributed random variables \( X_i \) with mean 0 and variance 1, so by the central limit theorem \( Y_m/\sqrt{m} \to N(0,1) \), the standard normal distribution. For \( U_m = (m-1)X_1 + (m-3)X_2 + \cdots + (3-m)X_{m-1} + (1-m)X_m \), we will use the following generalized central-limit theorem.
Theorem 1 (Lindeberg)  Suppose $S_n = W_1 + \cdots + W_n$, where $W_1, \ldots, W_n$ are independent random variables with mean 0 and respective variances $\omega_j^2$ (so $S_n$ has mean 0 and variance $\sigma_n^2 = \sum j \omega_j^2$). Then $\frac{S_n}{\sigma_n} \to N(0, 1)$, the standard normal distribution, provided
\[
\frac{1}{\sigma_n^2} \sum_{j=1}^{n} \int_{|x| > t\sigma_n} x^2 \, dW_j \to 0, \quad \text{for all } t > 0.
\]
(See Feller [7], VIII.4 Theorem 3.)

We can’t apply this directly to $U_m$, since the coefficients depend on $m$; but for $Z_m$, we can take the terms $W_j = jX_{m-j}$. (All the $X_j$ are identically distributed, so the distribution of $W_j$ doesn’t really depend on $m$.) The variance is $1^2 + 2^2 + \cdots + m^2 = m(m+1)(2m+1)/6 \sim m^3/3$. The condition holds automatically: for any fixed $t$, as $m$ becomes large $tm\sqrt{m/3} \gg m$, but $W_j$ is zero except at $x = \pm j$ which is less than $m$. Conclude that $Z_m/\sigma_m \to N(0, 1)$, for $\sigma_m = m\sqrt{m/3}$.

Now we can show that $U_m/\sigma_m \to N(0, 1)$ also. First, split $U_m$ into two identically-distributed, independent halves:
\[
U_m = (m-1)X_1 + (m-3)X_2 + \cdots + X_{m/2}
\]

Then write $U_m^+ = ((m-2)X_1 + (m-4)X_2 + \cdots + 2X_{m/2-1}) + (X_1 + \cdots + X_{m/2}) = 2\hat{Z}_{m/2-1}^+ + Y_{m/2}^+$ (where $\hat{Z}_{m/2-1}^+$ is just $Z_{m/2-1}$ with the variables $X_i$ shuffled; since the $X_i$ are identically distributed, so are $Z$ and $\hat{Z}$). In particular, $\hat{Z}_{m/2-1}^+$ has variance

\[\text{Shown for } m \text{ even. If } m \text{ is odd, the middle term is zero.}\]
\[\text{For } m \text{ odd, replace } m/2 - 1 \text{ with } (m-1)/2 \text{ and ignore the } Y_{m/2} \text{ term.}\]
Finally, take the limit:

\[
\frac{U_m}{\sigma_m} = \frac{U^+_m - U^-_m}{\sqrt{m^3/3}} = \frac{2\hat{Z}^+_m/2 - Y^+_m/2 - 2\hat{Z}^-_{m/2} - Y^-_{m/2}}{\sqrt{m^3/3}}
\]

\[
\to 2N(0,1/8) + 0 - 2N(0,1/8) = N(0,1).
\]

3.4 Independence of U and Y

Since we ultimately want to pick both \(Y^{(\alpha)} < Y^{(\beta)}\) and \(U^{(\alpha)} > U^{(\beta)}\), we need to make sure we can do both with positive probability. In fact, we can show that their limiting distributions are independent! We will do this by taking characteristic functions. Specifically, we claim that the characteristic function

\[
\phi_m(u, v) = \mathbb{E} \left[ \exp \left( iu \frac{U_m}{\sigma_m} \right) \cdot \exp \left( iv \frac{Y_m}{\sqrt{m}} \right) \right]
\]

of the joint distribution converges, as \(m \to \infty\), to the characteristic function

\[
\phi(u, v) = \exp \left( -\frac{u^2}{2} \right) \cdot \exp \left( -\frac{v^2}{2} \right)
\]

of the product of two independent \(N(0,1)\) normal distributions.

First, expand \(\phi_m\) in terms of the independent \(X_i\):

\[
\phi_m(u, v) = \mathbb{E} \left[ \prod_{j=1}^{m} \exp \left( iu \frac{(m + 1 - 2j)X_j}{\sigma_m} \right) \cdot \exp \left( iv \frac{X_j}{\sqrt{m}} \right) \right]
\]

\[
= \mathbb{E} \left[ \prod_{j=1}^{m} \exp \left( i\psi_{mj}X_j \right) \right] \quad \text{where } \psi_{mj}(u, v) = \left( \frac{u(m + 1 - 2j)}{\sigma_m} + \frac{v}{\sqrt{m}} \right)
\]

\[
= \prod_{j=1}^{m} \mathbb{E} \left[ \exp \left( i\psi_{mj}X_j \right) \right]
\]
\[
\prod_{j=1}^{m} \exp(i\psi_{mj}) + \exp(-i\psi_{mj}) = \prod_{j=1}^{m} \cos(\psi_{mj})
\]

Now take logarithms, and use the power series expansion to simplify:

\[
\ln \phi_m(u, v) = \ln \prod_{j=1}^{m} \cos(\psi_{mj})
\]

\[
= \sum_{j=1}^{m} \ln(\cos \psi_{mj})
\]

\[
= \sum_{j=1}^{m} \ln(1 - \psi_{mj}^2/2 + O(m^{-2}))
\]

\[
= \sum_{j=1}^{m} -\psi_{mj}^2/2 + O(m^{-2})
\]

Multiplying out \(\psi_{mj}^2\) and distributing the sum, we get

\[
\ln \phi_m(u, v) = -\frac{u^2}{2\sigma_m^2} \sum_{j=1}^{m} (m + 1 - 2j)^2 - \frac{v^2}{2m} \sum_{j=1}^{m} 1 - \frac{uv}{\sigma_m \sqrt{m}} \sum_{j=1}^{m} (m + 1 - 2j) - O(m^{-1})
\]

In the limit as \(m \to \infty\), the \(O\)-term goes to zero, and the first sum is asymptotic to \(\sigma_m^2 = m^3/3\); the second sum is \(m\), and the third sum simplifies to zero, giving us

\[
\ln \phi_m(u, v) = -\frac{u^2}{2} - \frac{v^2}{2} = \ln \phi(u, v)
\]

as claimed. In particular, \(U_m/\sigma_m\) and \(Y_m/\sqrt{m}\) are asymptotically independent.
3.5 Probabilistic Estimates

Now we need to argue that, when we pick $U_m$ and $Y_m$ to meet our needs on one diagonal $k + l = m$, we don’t impair our ability to pick appropriate values $U_n$ and $Y_n$ for a later diagonal $k + l = n$.

We may write $Y_n = (X_1 + \cdots + X_m) + (X_{m+1} + \cdots + X_n) = Y_m + \hat{Y}_{n-m}$, where $\hat{Y}_{n-m} = X_{m+1} + \cdots + X_n$ is identically distributed to $Y_{n-m}$, but independent of the first $m$ variables $X_i$. To guarantee $Y_n^{(\alpha)} < Y_n^{(\beta)}$, pick $n > Am^2$. Then

$$Y_n^{(\beta)} - Y_n^{(\alpha)} = (Y_m^{(\beta)} + \hat{Y}_{n-m}^{(\beta)}) - (Y_n^{(\alpha)} + \hat{Y}_{n-m}^{(\alpha)})$$
$$> (m \hat{Y}_{n-m}^{(\beta)}) - (m \hat{Y}_{n-m}^{(\alpha)})$$
$$= Y_{n-m}^{(\beta)} - \hat{Y}_{n-m}^{(\alpha)} - 2\sqrt{n/A}$$

Since $Y_{n-m}$ has standard deviation $\sqrt{n-m} > \sqrt{n/A}$ (use any $A > 1$ and large $n$), this is guaranteed positive if $Y_{n-m}^{(\beta)}$ is more than two standard deviations ($= 2\sqrt{n-m}$) larger than $Y_{n-m}^{(\alpha)}$; for example, if $Y_{n-m}^{(\beta)} > +\sqrt{n-m}$ and $Y_{n-m}^{(\alpha)} < -\sqrt{n-m}$. Since $Y_{n-m}$ is approximately normal, the probability for both of these is about $\Phi(-1) = 0.158655$, and both conditions can be satisfied with probability $(0.158655)^2 = 0.02517 > 1/40$.

Similarly,

$$U_n = 2Z_n - (n + 1)Y_n = 2(Z_m + (n - m)Y_m + \hat{Z}_{n-m}) - (n + 1)(Y_m + \hat{Y}_{n-m})$$
$$= \underbrace{2Z_m - (m + 1)Y_m}_{U_m} + \underbrace{(2\hat{Z}_{n-m} - (n - m + 1)\hat{Y}_{n-m})}_{\hat{U}_{n-m}}$$
$$+ (n - m)Y_m - m\hat{Y}_{n-m}$$
hence (for large \( n \))

\[
U_n^{(\alpha)} - U_n^{(\beta)} = (U_m^{(\alpha)} + \hat{U}_{n-m}^{(\alpha)} + (n-m)Y_m^{(\alpha)} - m\hat{Y}_{n-m}^{(\alpha)})
- (U_m^{(\beta)} + \hat{U}_{n-m}^{(\beta)} + (n-m)Y_m^{(\beta)} - m\hat{Y}_{n-m}^{(\beta)})
> (-m^2 + \hat{U}_{n-m}^{(\alpha)} - 2m(n-m)) - (m^2 + \hat{U}_{n-m}^{(\beta)} + 2m(n-m))
> \hat{U}_{n-m}^{(\alpha)} - \hat{U}_{n-m}^{(\beta)} - \sqrt{n^3/A}
\]

Since \( U_{n-m} \) has standard deviation asymptotic to \( \sigma_{n-m} = \sqrt{(n-m)^3/3} \) and \( \sigma_{n-m} > \sqrt{n^3/A} \) (for say \( A = 4 \), and large \( n \)), this is guaranteed positive if \( U_{n-m}^{(\alpha)} \) is more than two standard deviations (\( \approx 2\sigma_{n-m} \)) larger than \( U_{n-m}^{(\beta)} \). For example, if \( U_{n-m}^{(\alpha)} > +\sigma_{n-m} \) and \( U_{n-m}^{(\beta)} < -\sigma_{n-m} \). Since \( U_{n-m} \) is approximately normal, the probability for both of these is about \( \Phi(-1) = 0.158655 \), and both conditions can be satisfied with probability \((0.158655)^2 = 0.02517 > 1/40\).

Finally, since \( U_n \) and \( Y_n \) are asymptotically independent, we can further choose \( n \) large enough that the deviations from independence don’t affect these estimates.

We can now prove the infinite case. Let \( n_0 \) be sufficiently large that the normal approximations for \( Y_n \) and \( U_n \) hold to within the bounds above, and that both random variables are approximately independent. Then let \( n_1 = 4n_0^2, \ldots, n_{j+1} = 4n_j^2, \ldots \). For each \( n = n_j \), the probability that \( Y_n^{(\alpha)} < Y_n^{(\beta)} \) and \( U_n^{(\alpha)} > U_n^{(\beta)} \) — hence, the probability that \( \text{area}(\alpha|_{K\times L}) > \text{area}(\beta|_{K\times L}) \) and \( \alpha \not\sqsubseteq \beta \) — is at least \( 1/1600 \), independent of previous \( n_j \). If we take \( j \) of these diagonals, the probability that all \( j \) of them are compatible with \( \alpha \sqsubseteq \beta \) declines to \((1 - 1/1600)^j \), which goes to zero as \( j \to \infty \). (More specifically, we can fit \( j + 1 \) diagonals, spaced accordingly, in a box of size less than \( N = (4n_0)^2 \); \( j \) is on the order of \( \log \log N \).)
3.6 The finite box

Recall that we ultimately want to count the number of graphical partitions of $n$; this is (almost) the number of “dominance-pairs” of partitions $\alpha \preceq \beta$ on an $(n - d) \times d$ box, for all values of $d$. If we ignore the dominance requirement and only count the total number of partition-pairs on the same-size box ($(n - d) \times d$, for all $d$) we get the trivial upper bound $\sum_d \binom{n}{d}^2 = \binom{2n}{n}$, asymptotic to $4^n/\sqrt{n}$.

If we truncate a single infinite partition after $n$ steps, we wind up with a partition of just this type: a partition on an $(n - d) \times d$ box, for some value of $d$. Each of these $2^n$ possibilities is equally likely, so our probabilities translate directly into counts of these partitions. We want to count pairs of these partitions that both meet our infinite-box conditions, and have the same value of $d$ for both partitions.

3.6.1 Values of $d$ not near $n/2$

For small values of $d$, say $d \leq n/3$ (and by symmetry, for $d \geq 2n/3$), even the total number of paths grows slowly, so we needn’t consider the infinite-box conditions. Specifically,

$$\binom{n}{d} \leq \binom{n}{n/3} \sim \frac{\sqrt{2\pi n}(n/e)^n}{\sqrt{2\pi n}(n/3e)^n/3} \sqrt{2\pi n/3e}^{2n/3}$$

$$= \frac{1}{\sqrt{2\pi} \sqrt{2n^2/9} 2^{2n/3}/3^n}$$

$$< \frac{1}{\sqrt{n}} \left( \frac{3}{2^{2/3}} \right)^n$$

$$\ll \left( \frac{3}{2^{2/3}} \right)^n \ll \frac{2^n}{n^j} \text{ for any } j$$
so in particular

\[
\sum_{d < n/3 \text{ or } d > 2n/3} \binom{n}{d}^2 \ll \frac{4^n}{n}
\]

is smaller than our lower bound for graphical partitions (at least for large \(n\)), and may safely be ignored.

### 3.6.2 Values of \(d\) near \(n/2\)

We now may assume \(n/3 < d < 2n/3\), and count the number of pairs \(\alpha, \beta\) that both satisfy the infinite-box conditions and end at the same point \((d, n - d)\). We’ve already done the first stage of this: finding the proportion of pairs that satisfy the infinite-box conditions for selected values of \(m\). Now we find the proportion of these for which \(Y_n^{(\alpha)} = Y_n^{(\beta)}\).

Observe first that the value of \(U\) (or \(Z\)) is irrelevant here, as are the values of \(Y_j\) at values \(j\) prior to \(m\); if we know \(Y_m^{(\alpha)}\) and \(Y_m^{(\beta)}\), we know everything we need, since \(Y\) is a random walk. It thus suffices to prove that \(Y_n^{(\alpha)} = Y_n^{(\beta)}\) is rare for any specific values of \(Y_m^{(\alpha)}\) and \(Y_m^{(\beta)}\); then it follows that for all values of \(Y_j\) and \(U_j\) satisfying the infinite-box conditions, \(Y_n^{(\alpha)} = Y_n^{(\beta)}\) is rare.

Suppose \(Y_m^{(\alpha)}\) corresponds to the point \((K^{(\alpha)}, L^{(\alpha)})\). Of the \(2^{n-m}\) paths from here to the \(n\)th diagonal, only \(\binom{n-m}{K-K^{(\alpha)}}\) run to the specific point \((K, L)\) at which \(\beta\) meets the \(n\)th diagonal. Now, assuming that \(m\) is small, say \(m < n/2\), we have the probability

\[
\left(\frac{n-m}{K-K^{(\alpha)}}\right) / 2^{n-m} \leq \frac{n'}{(n'/2)} / 2^{n'} \quad \text{(where \(n' = n - m > n/2\))}
\]

\[
\sim \frac{1}{2^{n'}} \cdot \frac{\sqrt{2\pi n' n'^{n'}/e^{n'}}}{(\sqrt{2\pi n'(n'/2)^{n'/2}/e^{n'/2}})^2}
\]

44
\[
\begin{align*}
&< \frac{1}{\sqrt{2\pi n'}} \\
&< \frac{1}{\sqrt{\pi n}}
\end{align*}
\]

### 3.7 The upper bound

For any given \(n\) we can accumulate on the order of \(\log \log n\) diagonals; the probability any two paths meet the requirements at any one of these is at most \(1599/1600\), so the overall probability is \((1599/1600)^{C \log \log n} = e^{C \log \log n \log(1599/1600)} \approx (\log n)^{-C/1600}\). Multiplying by the independent probability of about \(1/\sqrt{n}\) that both paths end at the same place, and by the \(4^n\) total path-pairs of length \(n\), we have a final upper bound for the number of graphical partitions:

\[
\#G(n) < \frac{4^n}{\sqrt{n}(\log n)^{C'}}
\]  \hspace{1cm} (3.1)

where \(C\), and hence \(C' = C/1600 \geq 0\), could be determined explicitly.

This is an asymptotic improvement over the trivial bound \(#G(n) < 4^n/\sqrt{n}\).
Chapter 4

Further Improvements

4.1 Improving the Upper Bound

The ratio of the improved upper bound (3.1) to the trivial bound (1.4) is \((\log n)^{-C}\) for some power \(C\). Can we improve on this logarithmic-factor improvement on the trivial bound?

It’s impractical to explicitly count the number of graphical partitions on \(n\) vertices for even moderately large \(n\), but since we’re only looking at asymptotic results anyway, we can just estimate them. This is easy to do: Since all our graphical partitions have at most \(n\) parts, each of size at most \(n - 1\), we may pick partitions uniformly at random from the \(\binom{2n-1}{n-1}\) partitions satisfying those requirements, and count the fraction that are graphical. Recall that \(\binom{2n-1}{n-1}\) was our trivial upper bound, so what we’re doing is actually estimating the number of graphical partitions as a fraction of that trivial upper bound.

If we do this (see Figure (4-2)) and plot the results (see (4-1)), what we find is that this fraction decays only very slowly, say logarithmically. So the actual number
Figure 4-1: Graph of proportion of graphical partitions (log-log plot). The solid lines are $1/\log n$ and $.44/\log n$ (bold); this last fits well, better than competitors like $1/(\log n)^2$ (dashed) or $.29/x^{1/4}$ (dotted).
Figure 4-2: Numbers of graphical partitions, out of 10,000 random partitions

4.2 Improving the Lower Bound

As we’ve seen, the lower bound is almost certainly not optimal. This raises the question: Can we improve our lower bound without an entirely new strategy?

There are two main respects in which our estimate is clearly not tight. One is the order-of-growth estimate in Section 2.5: $r$, the number of components in our shape $\mu/\lambda$, should rarely be close to its maximum value $\min(k,l)$. For example, think about the case when $k = l$: the only way we can have $k$ distinct components is if no component takes up more than one row or column, and the only way this can happen is when $\mu/\lambda$ consists of the squares on the main antidiagonal of our $k \times l$ box. (This corresponds to the graphical partition $(2k-1, 2k-2, \ldots, 1)$.)

In order to make this approach work, we need to have an idea of what typical values of $r$ will be. See Figure 4-3 for a plot of (the $\log_2$ of) the number of graphical
Figure 4-3: (Natural log of) number of shapes vs. dominance-pairs for each $r$, $k = l = 100$. The lower (dotted) line is the trivial lower bound, the number of shapes; the upper (solid) line is the trivial upper bound, $2^r$ times the number of shapes; the middle (dashed) line is our improved lower bound, $(2r - 1)!!/r!$ times the number of shapes.
partitions with \( r \) components for each \( r \), taking \( k = l = 100 \).

Note that, if we had to guess from this graph alone, we might think that our new lower bound was accurate even to within a constant factor! The factor by which \( r! \) and \( (2r - 1)!!/r! \) differ increases only very slowly, to a maximum of about 18 at \( r = 100 \). Worse yet for our hopes of improvement, this graph has a very “flat” top, so we will have difficulty neglecting the high-\( r \) components where the difference is greatest.

There is a regularity worth noting, and possibly exploitable: for small \( r \), the numbers of components for sizes \( r \) and \( r + 1 \) approach small-integer ratios. For example, \( r = 1 \) and \( r = 2 \) agree to dozens of places, \( r = 3 \) is about \( 3/4 \) the size of \( r = 2 \), and \( r = 4 \) is about \( 2/3 \) the size of \( r = 3 \). These regularities seem to persist for other values of \( k, l \). If these held for, say, \( r < \log(n) \), then we could use that to estimate the number of components in that range and show that they constitute most of the possible shapes; then we could neglect the \( r > \log(n) \) shapes and improve the estimate we made in Section (2.5) of our new lower bound.

Recall that we estimated it term-by-term to be about \( r^{-1/2} \) worse than the trivial upper bound; if we can show that we can neglect \( r > \log(n) \), then this becomes only \( \log(n)^{-1/2} \), which is of the same form as our improved upper bound! So we would have proven that the number of graphical partitions is about \( \log(n)^{-C} \) times the total number of partitions being considered. Probably this is not true; notice that the maximum on this graph is about \( r = \sqrt{n} \), and we can’t very well neglect the largest term! But even this would substantially improve our lower bound.

The other respect in which our lower bound is clearly suboptimal is in our estimate of the number of signs-and-orderings that yield graphical partitions for a given unordered list \( \vec{c} \) of component sizes (see Section 2.4). Our estimate of \( (2r - 1)!! \) is only a lower bound, which is strict only if there are no “coincidences” among the
component sizes that place that particular \( \vec{c} \) on a chamber boundary. This is increasingly likely to happen as \( r \) increases, since not only does the number of components increase, but the average component size decreases as well (so all \( r \) components can fit inside the \( k \times l \) box), making coincidences more likely, even guaranteed.

This is less promising, since it has most room for improvement in the high-\( r \) range where few graphical partitions live. We will take a brief look at it anyway, to see if we can hope to better characterize the arrangement of hyperplanes dividing the space of component-size vectors into chambers, which would be of some interest in its own right.

The most interesting combinatorial synopsis of a hyperplane arrangement is the characteristic polynomial, which we may define as

\[
\chi(q) = \sum_{S \subseteq H} (-1)^{\#S} q^{\dim \cap S}
\]

provided the collection of hyperplanes \( H \) have a point 0 in common, as they do in our case. (This is Theorem 2.3.8 in [11], or 2.4 in the notes [19]; consult either of these references for more background on arrangements, and for the usual, more general, definition.) This provides us with information about the intersections of hyperplanes and the chambers created by them (for example, the number of chambers is \( |\chi(-1)| \), and the number of chambers of finite size is \( |\chi(1)| \)).

Most of the “good parts” of the theory of hyperplane arrangements expect the arrangement to be nice enough that its characteristic polynomial factors linearly. (This includes the “free arrangements” of Terao in [11].) In order to determine what these factors should be, we can calculate some characteristic polynomials numerically, using another interpretation of \( \chi(q) \) first used by Athanasiadis in [1]: when \( q \) is a sufficiently large prime (or power thereof), the number of elements not on any
hyperplane of $\mathcal{H}$ is $\chi(q)$.

So, we consider the hyperplane arrangement

$$\mathcal{H}_r := \left\{ \sum_{i \in I} x_i = \sum_{j \in J} x_j : \emptyset \neq I \cup J \subseteq \{1, \ldots, r\} \right\}$$

for some small $r$-values, let the computer count points for us for suitable primes $p$, and then interpolate to find the polynomial itself. To guarantee that $p$ is not so small that degeneracies are introduced, it is necessary and sufficient that the prime $p$ not divide any minor of the matrix representing $D_r$. Since this matrix’s rows are the coefficients of the hyperplanes constituting $D_r$, all of which are either 0, 1, or $-1$, the maximum possible value of any minor is given by the Hadamard bound $r^{r/2}$. (This is achieved, since all possible rows of this form appear, up to a factor $\pm 1$, as a row of $D_r$; pick $r$ rows forming a Hadamard matrix.) Hence, if we choose any $p$ greater than this bound, reduction mod $p$ will not reduce any minor’s value to zero, so $\chi_r(p)$ may be determined by counting points; since $\chi$ is a polynomial of degree $r$, interpolation at $r + 1$ such primes suffices to determine $\chi$ uniquely. (Actually, we only need $r$ points, since the number of bounded regions is $\chi(-1) = 0$, meaning $(q - 1)$ must be a factor.)

$$k = 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8$$

$$\left\lfloor k^{k/2} \right\rfloor = 5 \quad 16 \quad 55 \quad 216 \quad 907 \quad 4096$$

$$p \geq 7 \quad 17 \quad 59 \quad 223 \quad 911 \quad 4099$$

We then have:
so that \( \chi_3(q) = (q - 1)(q - 5)(q - 7) \), \( \chi_4(q) = (q - 1)(q - 11)(q - 13)(q - 15) \),
and \( \chi_5(q) = (q - 1)(q - 29)(q - 31)(q^2 - 60q + 971) \). Notice that \( \chi_5 \) doesn’t factor completely; in general we expect that “nice” structure in hyperplane arrangements will be reflected in the characteristic polynomial, so this suggests that the family of arrangements we’ve defined is unlikely to have further helpful properties.
Appendix A

Programs

Program 1:
sample_gp(n,q), a Maple program to count the graphical partitions on n vertices, out of q randomly selected partitions with at most n parts and at most (n−1) in height

# Generate a random graphical partition (using a random permutation)
rand_gp := proc(n::integer)
local height, index, randperm, randpart, i;

height:=n-1;
index:=1;
randperm:=combinat[randperm](2*n-1);
randpart := [seq(0,i=1..n)];
for i from 1 to 2*n-1 do
if (randperm[i] <= n-1) then
    height := height-1;
else
    randpart[index] := height;
    index := index + 1;
end if;
end do;
return(randpart);
end proc; # rand_gp(n)
# Count graphical partitions by random sampling

\[
sample\_gp:= \text{proc}(n::\text{integer}, \text{samplesize}::\text{integer}) \\
\text{local} \ count, i, \ this\_gp, \ this\_graph; \\
\text{count} := 0; \\
\text{for} \ i \ \text{from} \ 1 \ \text{to} \ \text{samplesize} \ \text{do} \\
\quad \this\_gp := \text{rand\_gp}(n); \\
\quad \this\_graph := \text{networks}[\text{graphical}](this\_gp); \\
\quad \# \ this\_graph \ is \ either \ an \ edgelist \ or \ the \ Boolean \ FAIL \\
\quad \text{if} \ (\text{not} \ \text{type}(\this\_graph, \text{boolean})) \ \text{then} \\
\quad \quad \text{count} := \text{count} + 1; \\
\quad \text{end \ if}; \\
\text{end \ do}; \\
\text{return} \ \text{count}; \\
\text{end \ proc};
\]

Program 2:

\[
\text{numshapes}(k,l,r), \ a \ PARI-GP[21] \ \text{script for computing the number of} \ r\text{-component skew shapes inside a} \ k\text{-by-}l \ \text{box.}
\]

```
\\ trap();
\n\n\PARI \indexes \start \at \ 1.
\C = \text{matrix}(100,100,i,j, \\
\quad \text{binomial}(i+j-2,i-1)^2*(i+j-1)/(i*j) );
\n\n\\ P\text{storage for }n1() \\
\N1\_TABLE = \text{matrix}(100*100,100);
```

```
/* \text{n1}(A,B,r) \ Choose \ r \ components \ in \ order, \ total \ size \ A\text{-by-}B. */
n1(A,B,r) = 
\{
\text{local}(i,j,s);
```

55
```
trap();
if( (r>A)||(r>B), return(0); );
if( (0==A)&&(0==B)&&(0==r), return(1); );
if( (0>=A)||(0>=B)||(0>=r), return(0); );
if( (0!=N1_TABLE[A*100+B,r]),
    return(N1_TABLE[A*100+B,r]);
else
    s=0;
    for(i=1,A,
        for(j=1,B,
            s += C[i,j]*n1(A-i,B-j,r-1);
        );
    );
    N1_TABLE[A*100+B,r]=s;
    return(s);
);

/* n(k,l,r) counts skew-shapes in a k-by-l box with r components */
n(k,l,r) =
{
    local(count,k0,10);

    trap();
    count=0;
    for(k0=0,k-r,
        for(10=0,1-l-r,
            count += binomial(k0+10,10)*binomial(k0+10+r,r)*
                n1(k-k0,k-10,r);
        ); \for 10
    ); \for k0
    return(count);
}

/* numshapes(K,L) lists n(k,l,r) for all possible r */
numshapes(K,L) =
{
    local(i,j);
```

56
Program 3:
chi(N,P), a PARI-GP script for computing the characteristic polynomial of the
planes of symmetry used in the lower bound, for small numbers of vertices N and
evaluated at small primes P.

/* PARI-GP script for computing chi(N,P) : */
/* note N <= 10, as if you’d try longer... */
trap();

chi(N,P) =
{
local(i,k,Count,incCount,incPoint,nextPoint);
local(point,setPlus,setMinus,setDiff);
local(TRUE,FALSE);

/* Parameters: */
TRUE=1; FALSE=0;

/* initialize point[] as a length-N column vector -- */
/* note point[1] isn’t used so set = 0 */
point=vectorv(N,i,2*i-2);

Count=0;
nexK=TRUE;
while(nextPoint ,
   /* Check this point for hyperplanes: */
   incCount=TRUE; /* Turn off incCount if we discover an equality.*/
   for( setPlus=1 , 2^(N-1)-1 , /* was 1 to 2^(N-1)-1 */
       for( setMinus=0 , 2^(N-2)-1 , /* was 0 to 2^(N-2)-1 */
       /* Only test sets if disjoint. */

57
if( ! bitand(setPlus,setMinus) ,
setDiff = (binary(2^(N-1)+setPlus)*point) -
(binary(2^(N-1)+setMinus)*point);
setDiff %= P;
if( (0==setDiff)||1==setDiff)||(P-1==setDiff) ,
incCount=FALSE;
brown break(2);
); \if setdiff...
); \if ! bitand(setPlus,setMinus)
); \for setMinus
if(incCount ,
Count++;
print(point,Count);
); \if incCount
/* Move on to next point: */
icnPoint=TRUE;
i=N;
while(incPoint ,
    if( (point[i] < P-2*(N-i+1) ),
        point[i]++;
        incPoint=FALSE;
    if( N=i,
        break;
    ,\else
        for( k=i+1 , N , point[k]=point[k-1]+2 );
    ); \if N==i
,\else
    i--;
    if( 1==i,
   nextPoint=FALSE;
        incPoint=FALSE;
    ); \if 1==i
); \if point[i]...
); \while incPoint
); \while nextPoint
return(Count*(P-1)*(N-1)!);
Bibliography


