COMBINATORICS FOR THE EAST MODEL

by

Fan Chung
Persi Diaconis
Ronald Graham

September 2000

Department of Statistics
STANFORD UNIVERSITY
Stanford, California
COMBINATORICS FOR THE EAST MODEL

by

FAN CHUNG
Department of Mathematics
University of California, San Diego

PERSI DIACONIS
Department of Mathematics and Statistics
Stanford University

RONALD GRAHAM
Department in Computer Science
University of California, San Diego

SEPTEMBER 2000

This research was supported in part by
National Science Foundation grant DMS-9504379

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
Combinatorics for the East model

Fan Chung *    Persi Diaconis ††    Ronald Graham *

Abstract

We study the number of configurations in the East model of statistical physics. This may be pictured as sites in a line. The site at zero is always occupied. The site at \( i > 0 \) can only be changed if site \( i - 1 \) is occupied. If at most \( n \) occupied sites are permitted, we establish upper and lower bounds of the form \( 2^{(5)} n! c^n \) where \( c < 1 \) for the number of possible configurations.

1 Introduction

This paper is motivated by a variety of Markov chains used by chemists and physicists to study properties of glasses and super-cooled liquids. The chains are called ‘facilitated kinetic Ising spin models’. They are based on a graph or lattice with various sites occupied or empty. At each time, a site is chosen at random and changed or not according to the familiar Metropolis dynamics for a given stationary distribution. The difference is that the change is allowed only if the neighbors of the chosen site are in a prescribed configuration; otherwise, no change is made. These neighborhood restrictions do not change the long-term stationary distribution but can lead to dramatic changes in approach to equilibrium.

The earliest such chains were introduced by Andersen and Fredrickson [2, 3] who allowed a change when \( k \) neighbors on a \( d \)-dimensional lattice were occupied. Reiter, Jäckle and coworkers [10] studied asymmetric rules, e.g., on a two-dimensional lattice, change is allowed if sites North and East are occupied. The simplest such model is the East model; this takes place on a one-dimensional lattice or ring with a transition permitted only if the neighbor to the immediate left is occupied (This should probably be called the West model but historically East is East).

Reiter and Jäckle [10] studied how the kinematic “East” restriction changes relaxation and cor-

*University of California, San Diego
††Stanford University
*Research supported in part by NSF Grant No. DMS 95-04397
relation times. One of their conjectures was proved by Aldous and Diaconis [1]. Pitts, Young and Andersen [8] (following Pitts [9]) studied the autocorrelation function of a single site in the East model, started in stationarity. They derive various approximations paralleling mode-coupling approximations used in the study of real glasses and super-cooled liquids. They found that spin systems give illuminating toy models for studying the validity of mode-coupling – just as in more complex systems, mode-coupling works well in some regions but not in others.

The present paper studies the combinatorics of the East model if at most \( n \) occupied sites are allowed. We give bounds for the entropy (number of possible states). It is convenient to study the subset of occupied positions. Thus we consider a graph \( G(n) \) formed as follows. The vertex set \( V(n) \) of \( G \) is the set of all subsets \( X \subseteq \{1, 2, 3, \ldots \} \) of cardinality at most \( n \). A pair \( \{X, X'\} \) forms an edge of \( G \), written \( X \sim X' \), provided \( X' \) can be obtained from \( X \) by adjoining to (or removing from) \( X \) the element \( x + 1 \) for some \( x \in X \), or by adjoining (or removing) the element 1.

We will be interested in investigating various properties of \( G \). In particular, we will establish upper and lower bounds on \( |V(n)| \) of the form

\[
2^{(\frac{1}{2})} n! c^n
\]

for various constants \( c < 1 \) (see Theorems 2, 4 and 5).

In Figure 1, we show the graph \( G(3) \). With the help of Susan Holmes and Glenn Tesler, we have computed the first few values of \( |V(n)| \).

\[
\begin{array}{c|cccccc}
 n & 1 & 2 & 3 & 4 & 5 \\
|V(n)| & 2 & 5 & 26 & 373 & 15193 \\
\end{array}
\]

We did not find this sequence in standard lists of integer sequences. Our bounds show that \( |V(6)| \) is about \( 2.4 \times 10^6 \) which is too large for the brute force algorithm we employed. The exact value \( |V(5)| \) gives an estimate of \( c = .6583 \) if \( |V(n)| \sim 2^{(\frac{1}{2})} n! c^n \).

### 2 Elementary facts

**Fact 1** (i) \( A(n) := \max\{x \in X \subseteq V(n) : |X| = 1\} = 2^{n-1} \);

(ii) \( B(n) := \max\{x \in X \subseteq V(n)\} = 2^n - 1 \).
Proof: (By induction on n.) The assertion certainly holds for $n = 1$ since $A(1) = 1 = B(1)$. Assume for some $n \geq 1$ that $A(k) = 2^{k-1}$ and $B(k) = 2^k - 1$ for all $k \leq n$. Observe that in general if $X \in V(n)$ with $|X| = r$ and $Y \in V(n-r)$ then $X \cup (x + Y) \in V(n)$ for any $x \in X$ (where $x + Y$ denotes $\{x + y : y \in Y\}$). In this case we can think of building a copy of $Y$ on the "base" $x \in X$. Thus, taking $X = \{2^{n-1}\} \in V(n) \subset V(n+1)$ and $Y = \{2^{n-1}\} \in V(n)$, we get $X' = \{2^{n-1}, 2^n\} \in V(n+1)$. Now we can reverse the process of generating the element $2^{n-1}$ in $V(n)$ to remove $2^{n-1}$ from $X'$, forming $X'' = \{2^n\} \in V(n+1)$, which shows that $A(n+1) \geq 2^n$.

Now, with $X = \{2^n\} \in V(n+1)$ (as we just showed) and $Y \in V(n)$ with $\max Y = 2^n - 1$ (by the induction hypothesis), we can construct $X' = X \cup (2^n + Y) \in V(n+1)$ with $\max X' = 2^n + 2^n - 1 = 2^{n+1} - 1$, which shows $B(n+1) \geq 2^{n+1} - 1$.

In the other direction, if $\{x_0\} \in V(n+1)$ with $x_0 \geq 2^n + 1$, then in order to remove it (i.e., reach $\emptyset$ through a sequence of edges), we would have to create a set $Y \in V(n)$ with $x_0 - 1 \in Y$. But since $x_0 - 1 \geq 2^n$ then by (ii), this is impossible. Thus, $A(n+1) = 2^n$. Finally, suppose $X \in V(n+1)$
where, without loss of generality, we can assume \( |X| = n + 1 \). Since by hypothesis there is a path in \( G(n + 1) \) from \( X \) to \( \emptyset \) then \( X \) must contain a pair of consecutive integers, say \( x_0 \) and \( x_0 + 1 \) (since otherwise we couldn't move at all from \( X \)). Removing \( x_0 + 1 \) to form \( X_1 \), we see (by induction) that \( X_1 \) much have a pair of elements \( x_1, x_1 + g_1 \), with \( g_1 \leq 2 \). (Again, since otherwise \( X_1 \) wouldn't be connected to \( \emptyset \)). Remove \( x_1 + g_1 \) to form \( X_2 \). The general step in this process forms the (sub)set \( X_k \subset X \) of size \( n + 1 - k \), which must then possess a pair of elements \( x_k, x_k + g_k \) with \( g_k \leq 2^k \). We remove \( x_k + g_k \) from \( X_k \) to form \( X_{k+1} \), etc. Eventually, we reach \( X_n \subset X \) of size 1, which must consist of a single element \( x_n \leq 2^n = A(n + 1) \). Combining all the preceding inequalities shows that

\[
\max X \leq 2^n + 2^{n-1} + \ldots + 2 + 1 = 2^{n+1} - 1
\]

Thus, \( B(n + 1) \leq 2^{n+1} - 1 \) and Fact 1 is proved. \( \Box \)

The same argument can be used to prove the more general fact:

**Fact 2** For \( 1 \leq k \leq n \),

\[
\max\{x \in X \in V(n) : |X| = k\} = 2^n - 2^{n-k}
\]

### 3 Upper bounds on \( |V(n)| \)

For a set \( X = \{X(1) < X(2) < \ldots < X(r)\} \in V(n) \), define the sequence of gaps of \( X \) to be the sequence \( g = g(X) = (g_1, g_2, \ldots, g_r) \) where \( g_i := X(i) - X(i - 1) \), and by convention, we always take \( X(0) = 0 \). The preceding considerations show that the following (polynomial-time) algorithm can always be used to decide whether a particular set \( X \subseteq \mathbb{P} \) is in \( V(n) \).

(1) If \( g(X) \) has no gap of size \( \leq 2^{n-|X|} \) then HALT. We can conclude that \( X \not\in V(n) \). Otherwise, if \( g_i = X(i) - X(i - 1) \in g(X) \) has \( g_i \leq 2^{n-|X|} \) then remove \( X(i) \) from \( X \) to form \( X' \).

(2) Repeat (1) with \( X \) replaced by \( X' \).

(3) If we succeed in reaching \( \emptyset \) this way then \( X \in V(n) \), and, in fact, by reversing the preceding steps (and using Fact 1), this shows how to construct it. Otherwise, we conclude \( X \not\in V(n) \). Notice that there may be many choices for the elements to be removed at each step. This reduction algorithm allows for any choice to be made at each step.

Let us assume for now that \( X \in V(n) \) with \( |X| = n \). We are going to specify a particular choice to be made at each of the removal steps. Namely, let \( R \) denote the preceding reduction algorithm
in which we always remove the largest possible integer satisfying the required gap size condition. This process results in the elements of \( X \) being removed in some particular order, generating a permutation \( \pi = \pi_X \) on \( \{1, 2, \ldots, n\} \). In particular, for \( X = \{X(1) < X(2) < \ldots < X(n)\} \in V(n) \), where \( X(i) \) is removed at step \( \pi(i) \).

It will be convenient to denote a set \( X \) by its corresponding gaps sequence \( g(X) = (g_1, g_2, \ldots, g_n) \) where \( g_i = X(i) - X(i - 1) \). What we will do is to derive upper bounds on the number \( N(\pi) \) of \( X \in V(n) \) which generate the permutation \( \pi = \pi_X \) for each permutation \( \pi \) of \( \{1, 2, \ldots, n\} \). We first illustrate this idea with several examples.

**Example 1:** \( n = 4, \pi = \left( \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{array} \right) \). \( g(X) = (g_1, g_2, g_3, g_4) \)

At the first step of the reduction, since \( \pi(4) = 1 \), then \( X(4) \) is removed, leaving \( g(X^{(1)}) := (g_1, g_2, g_3) \). This implies in particular that \( g_4 \leq 1 \).

At the second step, since \( \pi(3) = 2 \), then \( X(3) \) is removed (so \( g_3 \leq 2 \)), leaving \( g(X^{(2)}) := (g_1, g_2) \). We continue this process for two more steps, finally reaching \( \emptyset \). For the permutation \( \pi \) to be valid, we need the inequalities

\[
\begin{align*}
g_1 & \leq 1 \\
g_2 & \leq 2 \\
g_3 & \leq 4 \\
g_4 & \leq 8
\end{align*}
\]

Hence, the total number \( N(\pi) \) of possible \( X \in V(4) \) is at most \( g_1 g_2 g_3 g_4 \leq 1 \cdot 2 \cdot 4 \cdot 8 = 64 \). The same argument shows that for general \( n \), the reverse permutation \( \pi \) with \( \pi(k) = n + 1 - k, 1 \leq k \leq n \), has \( N(\pi) \leq \prod_{k=1}^{n} 2^{k-1} = 2^{n(n-1)/2} \). In general, since each \( X \) is determined by its gap sequence \( g(X) \), then in fact \( N(\sigma) \leq 2^{n(n-1)/2} \) for any permutation \( \sigma = \sigma_X \), which gives the (trivial) estimate

\[
|V(n)| \leq \sum_{\pi} N(\pi) \leq n! 2^{n(n-1)/2} \tag{1}
\]

Theorem 1 will improve upon this estimate by an exponential factor.

**Example 2:** \( n = 4, \pi = \left( \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{array} \right) \). \( g(X) = (g_1, g_2, g_3, g_4) \)

Proceeding as before we find \( X(3) \) is the first number removed, so that \( g_3 \leq 1 \). However, since \( X(4) \) was not removed (and is to the right of \( X(3) \)) then we must have \( g_4 > 1 \). Removing \( X(3) \) leaves us with the set \( X' \) with gap sequence \( g(X') = (g_1, g_2, g_3 + g_4) \). In general, whenever an internal
number $X(i)$ is removed, the new gap formed is the sum of the two gaps $X(i)$ is currently adjacent to. Now at the second step, $X(4)$ is removed, so we must have its (new) gap $g_3 + g_4 \leq 2$. However, this is not possible since $g_3 = 1$ and $g_4 > 1$. Hence, no $X$ can have this permutation, i.e., $N(\pi) = 0$.

We now consider the general case. We begin with a permutation $\pi$ on $\{1, 2, \ldots, n\}$ where $X(i)$ is removed at step $\pi(i)$ by the (greedy) algorithm R. Let $g_i(k)$ denote the gap associated with $X(i)$ at the beginning of step $k$ (i.e., when only $k - 1$ elements have been removed), assuming that $X(i)$ has not yet been removed. Thus, $g_i(k) = \sum_{j=0}^{r} g_{i-j}$ where $r$ is the largest index such that $\pi(i-r) < k$. In particular $g_i(1) = g_i$. Define $h_i = g_i(\pi(i))$. Then $h_i$ is the gap associated with $X(i)$ just prior to its being removed at step $\pi(i)$. By the definition of algorithm R, we always have

$$h_i \leq 2^{\pi(i)-1}, \quad 1 \leq i \leq n$$  \hspace{1cm} (2)

Now, suppose that for some $i$, we find there is a $j < i$ such that $\pi(j) = \pi(i) - 1$

\[ j \quad i \]

\[ \bullet \quad \bullet \]

\[ \pi(i)-1 \quad \pi(i) \]

Thus, at step $\pi(i)-1$, $X(i)$ was passed as a candidate for removal, and $X(j)$ was selected instead. This implies that

$$2^{\pi(i)-2} < g_i(\pi(i)-1) \leq g_i(\pi(i)) = h_i.$$  

Combining this with (2), we have

$$2^{\pi(i)-2} + 1 \leq h_i \leq 2^{\pi(i)-1}$$  \hspace{1cm} (3)

(i.e., we lose a factor of 1/2 over the trivial estimate of $2^{\pi(i)-1}$ for the number of choices for $h_i$). Hence, if there are $k$ such $i$'s for $\pi$, then total number of choices for all the $h_i$ is at most

$$2^{-k} \cdot 2^{0+1+\ldots+(n-1)} = 2^s \cdot 2^{-k}.$$  

It is easy to see by considering the inverse permutation $\pi^{-1}$ that the number of permutations $\pi$ having exactly $k$ values $i$ with $\pi(j) = \pi(i) - 1$ for some $j < i$ is just the Eulerian number $\left( \begin{array}{c} n \\ k \end{array} \right)$, which also counts the number of permutations $\pi$ of $\{1, 2, \ldots, n\}$ with $k$ rises, i.e., $k$ occurrences of an $s$ such that $\pi(s) < \pi(s+1)$ (see [4] for an in-depth discussion of Eulerian numbers). Hence, we have the estimate:
Theorem 1

\[ |V(n)| \leq 2^\left(\binom{4}{n}\right) \sum_k \binom{n}{k} 2^{-k} \]  \quad (4)

The sum \( S_n := \sum_k \binom{n}{k} 2^{-k} \) has occurred in various forms in the literature. In particular, one finds in page 627 of [7], the sum

\[ P_n := \sum_k \binom{n}{k} 2^{k-1} \]  \quad (5)

and references where it is shown that

\[ \sum_{n \geq 0} P_n \frac{z^n}{n!} = \frac{1}{1 - e^z} \]  \quad (6)

which implies

\[ \frac{P_n}{n!} = \frac{1}{2} (\ln 2)^{-n-1} + \sum_{k \geq 1} \operatorname{Re}((\ln 2 + 2\pi i k)^{-n-1}) \]  \quad (7)

One also finds the interesting equality of Gross [5]

\[ P_n = \sum_{k \geq 1} \frac{k^n}{2k+1}, \quad n \geq 1 \]  \quad (8)

Note that by the symmetry property of \( \binom{n}{k} = \binom{n}{n-k-1} \), we have

\[ \frac{1}{2n-2} P_n = \sum_k \binom{n}{k} 2^{-n+k+1} \]
\[ = \sum_k \binom{n}{n-k-1} 2^{-k} \]
\[ = \sum_k \binom{n}{k} 2^{-k} \]
\[ = S_n \]  \quad (9)

which implies

\[ S_n = \frac{1}{2n-1} \sum_{k \geq 1} \frac{k^n}{2^k} \]  \quad (10)

Using dominated convergence in (7) along with (8) shows

\[ S_n \sim \frac{n!}{(\ln 4)^n} \]  \quad (11)

Hence, we have
Theorem 2

\[ |V(n)| \leq 2^{\binom{n}{2}} S_n < 2^{\binom{n}{2}} n! - \frac{1}{(\ln 4)^n} \] (12)

for \( n \) sufficiently large.

A more refined version of this argument can be used to obtain the following stronger upper bound. For a permutation \( \pi \) of \( \{1, 2, \ldots, n\} \), define for \( 1 \leq i \leq n \) the quantity \( d_{\pi}(i) \) to be the least integer \( d \) (if it exists) such that \( \pi(i) < \pi(i + d) \). If \( d \) does not exist then set \( d_{\pi}(i) = \infty \). Finally, define

\[ d(\pi) := \prod_{i=1}^{n} \left( 1 - \frac{1}{2^{d_{\pi}(i)}} \right). \]

Then it can be shown that the following generalization of Theorem 1 holds.

Theorem 3

\[ |V(G)| \leq 2^{\binom{n}{2}} \sum_{\pi} d(\pi) \] (13)

The bound in Theorem 1 comes from (13) by just taking account of those \( i \) in \( \pi \) for which \( d_{\pi}(i) = 1 \) (counted by Eulerian numbers). An intermediate result arises by just considering those \( i \) in \( \pi \) for which \( d_{\pi}(i) \leq 2 \) (and taking other factors in the product \( d(\pi) \) equal to 1). It is straightforward to show that this results in the following bound.

For a permutation \( \pi \) of \( \{1, 2, \ldots, n\} \), if \( \pi(i) < \pi(i + 1) \) we say that \( \pi \) has a rise at \( i \). Similarly, if \( \pi(i + 1) < \pi(i) < \pi(i + 2) \), we say that \( \pi \) has a “213” at \( i \).

Let \( \left\langle \begin{array}{cc} n \\ k \end{array} \right\rangle \) denote the number of permutations \( \pi \) of \( \{1, 2, \ldots, n\} \) which have \( k \) rises and \( l \) 213’s for \( 0 \leq l \leq k < n \). Thus, \( \sum_l \left\langle \begin{array}{cc} n \\ k, l \end{array} \right\rangle = \left\langle \begin{array}{c} n \\ k \end{array} \right\rangle \)

Theorem 4

\[ |V(n)| \leq 2^{\binom{n}{2}} \sum_{k,l} \left\langle \begin{array}{cc} n \\ k, l \end{array} \right\rangle 2^{-k}(4/3)^{-l} \] (14)
It is easy to see that these “generalized Eulerian” numbers \( \binom{n}{k,l} \) satisfy the recurrence

\[
\binom{n}{k,l} = (l+1)\binom{n-1}{k,l} + (l+1)\binom{n-1}{k-1,l+1} + (n-k-l)\binom{n-1}{k-1,l} + (k-l+1)\binom{n-1}{k,l-1}
\]

(15)

\[
\binom{0}{0} = 1, \quad \binom{a}{b,c} = 0 \text{ if } a, b \text{ or } c < 0
\]

We show some small values of \( \binom{n}{k,l} \) in Table 1.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( k )</th>
<th>( \binom{n}{k,l} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1:

We have not analyzed the asymptotic behavior of the sum in (14). However, preliminary computations indicate that

\[
\sum_{k,l} \binom{n}{k,l} 2^{-k}(4/3)^{-l} = O(n! c_2^n)
\]

(16)

where \( c_2 < 0.95/\ln 4 = 0.68528\ldots \), which represents a modest (but real) improvement over the bound (12).

4 Lower bounds on \( |V(n)| \)

To show that \( |V(n)| \) is relatively large, we will describe a method for constructing large subsets of \( V(n) \). We begin with a simple version of the construction. Suppose \( d = (d_1 > d_2 > \ldots > d_n) \) is a sequence of integers satisfying \( d_i \in (2^{n-i-1}, 2^{n-i}] \), \( 1 \leq i \leq n \). Form a set \( X = \{X(1), X(2), \ldots, X(n)\} \) from \( d \) as follows (where, as usual, we define \( X(0) = 0 \)):
For the first two steps, choose \( X(1) = d_1 \), and \( X(2) = X(1) + d_2 \). Now, in general, at the \( k \)-th step, select \( X(k) \) to be one of \( X(i) + d_k \), \( 0 \leq i < k \), where \( X(i) \) is required to be different from the \( X(i') \) used in forming \( X(k - 1) = X(i') + d_{k-1} \). Note that the intermediate set \( X_k = \{X(1), \ldots, X(k)\} \) has the property that the unique smallest gap between consecutive elements is just \( d_k \). This follows by induction since when \( X(k) \) is added then either it is the largest element of \( X_k \), or it falls between two consecutive elements of \( X_{k-1} \), say, \( X(i) < X(k) < X(i') \). Thus, the two new gaps created in this case are \( X(k) - X(i) = d_k \) and

\[
X(i') - X(k) = d_{i'} - d_k \\
\geq d_{k-2} - d_k \\
> 2^{n-k+1} - 2^{n-k} \\
= 2^{n-k} \geq d_k
\]

Hence, in either case, \( d_k \) is the unique minimum gap size of \( X_k \).

Now observe that we can reduce \( X \) to \( \emptyset \) by removing its elements sequentially, always choosing the point having the smallest current gap to be removed. Doing this will remove the \( X(k) \) exactly in the reverse order \( X(n), X(n - 1), \ldots, X(1) \) by the minimum gap size property of the \( d_k \) just mentioned. In fact, given the final set \( X \), this reduction will recover both the sequence \( d \), and the points \( X(i) \) on which each \( X(k) \) was “based”, (i.e., \( X(k) = X(i) + d_k \)). Hence, the total number of different \( X \)'s which can be constructed this way is

\[(n - 1)! \, 2^{1+2+\ldots+(n-2)} = (n - 1)! \, 2^{\binom{n-1}{2}} \]

This implies the estimate

\[|V(n)| \geq 2^{\left(\frac{n}{2}\right)} n! \left\{\frac{1}{2}\right\}^{n}/n\]  \hspace{1cm} (17)

For the next approximation, we will allow more choices for each \( d_k \) than before, but fewer choices for the number of ways that \( X(k) \) can be chosen, still however, so that when \( X(k) \) is selected, say \( X(k) = X(i) + d_k \), then \( d_k = X(k) - X(i) \) is always the unique smallest gap in \( X_k = \{X(1), X(2), \ldots, X(k)\} \).

Now for \( d = \{d_1 > d_2 > \ldots > d_n\} \), we will only require that \( d_i \in (2^{n-i-2}, 2^{n-i}] \), \( 1 \leq i \leq n \). However, we will now require in choosing \( X(k) = X(i) + d_k \) that \( X(k) \) is different from any \( X(i') \) used in defining \( X(k - 1) \) and \( X(k - 2) \). Thus, the number of ways of choosing the “base point” \( X(i) \) in forming \( X \) is now only \( (n - 2)! \) (instead of \( (n - 1)! \)) as in the preceding construction. However, we will more than make up for this with the increased number of choices of the \( d_i \). Our next job is
to estimate this number of choices, which we will denote by \( f_0(n) \). Further, define \( f_i(n) \) to be the number of choices of \( d = \{ d_1 > d_2 > \ldots > d_n \} \), with \( d_1 \in (2^{n-2}, 2^{n-1}] \) and \( d_i \in (2^{n-i-1}, 2^{n-i+1}] \), 2 \( \leq i \leq n \), where, for convenience, we will henceforth assume \( n \geq 10 \). Thus by considering where \( d_1 \) is chosen, we have the recurrences

\[
\begin{align*}
  f_0(n) &= 2^{n-2}f_0(n-1) + f_1(n-1), \\
  f_1(n) &= \binom{2^{n-2}}{2} f_0(n-1) + 2^{n-2}f_1(n-1), \quad n \geq 10
\end{align*}
\]

(18)

Let \( F_0(m) = \frac{f_0(m)}{(\begin{bmatrix} \frac{m}{2} \end{bmatrix})} \), \( F_1(m) = \frac{f_1(m)}{(\begin{bmatrix} \frac{m}{2} \end{bmatrix})} \), \( 1 \leq m \leq n \). Then (18) implies

\[
\begin{align*}
  F_0(n) &= F_0(n-1) + F_1(n-1), \\
  F_1(n) &= \left( \frac{1}{4} - \frac{1}{2^n} \right) F_0(n-1) + \frac{1}{2} F_1(n-1), \quad n \geq 10
\end{align*}
\]

(19)

Finally, for \( i = 0 \) and \( 1 \), define

\[
F_i'(n) = F_i(n) \prod_{j=5}^{n} \left( 1 - \frac{j}{2} \right) 2^{-j+2}^{-1}
\]

(20)

Substituting into (19), we obtain

\[
\begin{align*}
  F_0'(n)(1 - \left( \frac{n}{2} \right) 2^{-n+2}) &= F_0'(n-1) + F_1'(n-1), \\
  F_1'(n)(1 - \left( \frac{n}{2} \right) 2^{-n+2}) &= \left( \frac{1}{4} - \frac{1}{2^n} \right) F_0'(n-1) + \frac{1}{2} F_1'(n-1),
\end{align*}
\]

(21)

which implies

\[
\begin{align*}
  F_0'(n) &\geq F_0'(n-1) + F_1'(n-1), \\
  F_1'(n) &\geq \frac{1}{4} F_0'(n-1) + \frac{1}{2} F_1'(n-1),
\end{align*}
\]

(22)

for \( n \geq 10 \). Hence, if we define \( F_0'' \) and \( F_1'' \) recursively by

\[
\begin{align*}
  F_0''(n) &= F_0''(n-1) + F_1''(n-1), \\
  F_1''(n) &= \frac{1}{4} F_0''(n-1) + \frac{1}{2} F_1''(n-1),
\end{align*}
\]

(23)

then we find

\[
F_0''(n) > c \left( \frac{3 + \sqrt{5}}{4} \right)^n
\]

for a suitable constant \( c > 0 \) as \( n \to \infty \). This implies

\[
F_0(n) > c' \left( \frac{3 + \sqrt{5}}{4} \right)^n
\]

11
for some $c' > 0$, and so,

\[
    f_0(n) > c'(\frac{3 + \sqrt{5}}{4})^n 2^{(n-1)}
    = c'(\frac{3 + \sqrt{5}}{8})^n 2^{\frac{n}{2}}
\]

Thus, by the previous remark on the number of choices for base points, we have the lower bound:

\[
    |V(n)| \geq c'(\frac{3 + \sqrt{5}}{8})^n 2^{\frac{n}{2}}(n - 2)!
\]

for a suitable constant $c' > 0$.

Before proceeding to the general construction, we will sketch the next stage in this approach. Now, we will relax the constraints on choosing $d = \{d_1 > d_2 > \ldots > d_n\}$ even further, while at the same time, increasing the constraints on selecting the $X(i)$. Namely, we now only require that $d_i \in (2^{n-i-3}, 2^{n-i}]$. However, in choosing $X(k) = X(i) + d_k$, we require that $X(i)$ is different from any $X(i')$ used in defining $X(k - j)$ for $j = 1, 2, 3$. As usual, this will guarantee that $d_k$ is always the current smallest gap (and consequently, the $d_k$ and (something) where they are attached can be recovered uniquely from $X$). However, the number of choices for the $X(i)$ is now only $(n - 3)!$. To count the number of choices for $d$, define

\[
    g_0(n) = \# \text{ of choices for } d \text{ with } d_i \in (2^{n-i-3}, 2^{n-i}], 1 \leq i \leq n.
\]

\[
    g_1(n) = \# \text{ of choices for } d \text{ with } d_i \in (2^{n-i-3}, 2^{n-i}],
    \quad d_i \in (2^{n-i-2}, 2^{n-i+1}], 2 \leq i \leq n.
\]

\[
    g_2(n) = \# \text{ of choices for } d \text{ with } d_i \in (2^{n-i-2}, 2^{n-i}],
    \quad d_i \in (2^{n-i-3}, 2^{n-i+1}],
    \quad d_i \in (2^{n-i-1}, 2^{n-i+2}], 3 \leq i \leq n.
\]

Again, by considering where $d_1$ and $d_2$ are chosen, we have the recurrences:

\[
    g_0(n) = 2^{n-2} g_0(n-1) + g_1(n-1), \quad (25)
\]

\[
    g_1(n) = \binom{2^{n-2}}{2} g_0(n-1) + 2^{n-2} g_1(n-1) + g_2(n-1),
\]

\[
    g_2(n) = \binom{2^{n-2}}{3} g_0(n-1) + \binom{2^{n-2}}{2} g_1(n-1) + 2^{n-2} g_2(n-1), \quad n \geq 10
\]

As before, setting $G_i(n) = g_i(n) 2^{-\frac{(n-1)^i}{2}}$, $1 \leq i \leq 3$, and defining
\[ G'_i(n) = G_i(n) \prod_{j=5}^{n} \left( 1 - \left( \frac{j}{2} \right)^{2^{-j+2}} \right)^{-1}, \]

we obtain the system of inequalities:

\[ G'_0(n) \geq G'_0(n-1) + G'_1(n-1), \]

\[ G'_1(n) \geq \frac{1}{4} G'_0(n-1) + \frac{1}{2} G'_1(n-1) + G'_2(n-1), \]

\[ G'_2(n) \geq \frac{1}{48} G'_0(n-1) + \frac{1}{16} G'_1(n-1) + \frac{1}{4} G'_3(n-1), \]

This implies that

\[ G'_0(n) > c \rho^n \]

for a suitable \( c > 0 \) where \( \rho \approx 1.34259 \ldots \) is the largest root of \( x^3 - \frac{7}{4} x^2 + \frac{9}{16} x - \frac{1}{48} \), i.e., \( \rho \) is the largest eigenvalue of the matrix:

\[
\begin{pmatrix}
1 & 1 & 0 \\
\frac{1}{4} & \frac{1}{2} & 1 \\
\frac{1}{48} & \frac{1}{16} & \frac{1}{4}
\end{pmatrix}
\]

This implies

\[ |V(n)| > c_2 \left( \frac{\rho}{2} \right)^n 2^{\left( \frac{5}{2} \right)} (n-3)! \]

Now, for the general case of this construction, we choose a fixed integer \( r > 0 \), and we want to estimate the number of \( d = \{d_1 > d_2 > \ldots > d_n\} \), this time with \( d_i \in (2^{n-i-r}, 2^{n-i}] \), \( 1 \leq i \leq n \), where \( n \geq 10 \). Correspondingly, in choosing \( X(k) = X(i) + d_k \), we require that \( X(i) \) is different from any \( X(i') \) used in defining \( X(k-j) \) for \( 1 \leq j \leq r \). Thus, we will have a factor of \( (n-r)! \) when counting the number of choices for \( X \).

Next, for \( 0 \leq u \leq r-1 \), let \( h_u(n) \) denote the number of ways of choosing \( d = \{d_1 > d_2 > \ldots > d_n\} \) with

\[ d_i \in (2^{n-r+u-i}, 2^{n-i}] \text{ for } 1 \leq i \leq u, \]

\[ d_i \in (2^{n-r+u-i}, 2^{n+u-i}] \text{ for } u+1 \leq i \leq n. \]
By analyzing where the initial $u$ $d_i$'s are chosen, we obtain the following recurrence equations:

$$h_u(n) = \sum_{i=0}^{u+1} \left( \binom{2n-2}{u-i+1} h_i(n-1) \right), \quad 0 \leq u \leq r - 1$$  \hspace{1cm} (28)

Substituting

$$H_i(n) = h_i(n)2^{-\binom{n}{2}}$$

we obtain

$$H_u(n) = \sum_{i=0}^{u+1} \left( \prod_{j=0}^{u-i} \left( 1 - \frac{j}{2n-2} \right) \right) \frac{1}{(u-i+1)!} \frac{2^{(i)}}{2^{\binom{i+1}{2}}} H_i(n-1)$$  \hspace{1cm} (29)

As before, if we make the substitution

$$H_i'(n) = H_i(n) \prod_{j=5}^{n} \left( 1 - \binom{j}{2} 2^{-j+2} \right)$$

then we find

$$H_u'(n) \geq \sum_{i=0}^{u+1} \frac{1}{(u-i+1)!} \frac{2^{(i)}}{2^{\binom{i+1}{2}}} H_i'(n-1)$$  \hspace{1cm} (30)

This implies that for a suitable constant $c_r > 0$,

$$H_0(n) > c_r \frac{\rho_r^{n}}{2^{\binom{r}{2}}} (n-r)!$$

where $\rho_r$ is the largest eigenvalue of the $r \times r$ matrix

$$M_r = \frac{2^{(i)}}{(i+1-j)! 2^{\binom{i+1}{2}}} \left( 0 \leq i, j \leq r-1 \right)$$

Note that

$$M_r = U_r A_r U_r^{-1}$$

where $U_r$ is the $r \times r$ diagonal matrix with $i$-th entry $2^{-\binom{i}{2}}$ and

$$A_r = \frac{1}{(2(i+1-j)!)} \left( 0 \leq i, j \leq r-1 \right)$$

Thus, $\rho_r$ is just the largest eigenvalue of $A_r$. We note that $\rho_r$, $r \to \infty$, is an increasing sequence.

Computation produces the following bounds on the $\rho_r$:
\begin{tabular}{c|c}
\hline
$r$ & $\rho_r$ \\
\hline
1 & $1.309 \ldots = \frac{3 + \sqrt{5}}{4}$ \\
2 & 1.34259 \ldots \\
3 & 1.34399 \ldots \\
4 & 1.344014945 \ldots \\
5 & 1.344015076 \ldots \\
6 & 1.344015076 \ldots \\
\hline
\end{tabular}

This rapid convergence is to be expected because of the smallness of the entries of $A_r$ as their row indices increase.

Thus, we have the lower bound:

**Theorem 5**

$$|V(n)| > (0.672)^n 2^{\left(\frac{n}{2}\right)} n! \quad \text{for} \quad n > n_0$$

(31)

Recall the bound in (14) gives (via (16) which we do not prove)

$$|V(n)| < (0.6852)^n 2^{\left(\frac{n}{2}\right)} n!$$

**Acknowledgments**

We thank Hans Andersen, Susan Holmes, Don Knuth and Glenn Tessler for their help.

**References**


