JOURNAL OF COMBINATORIAL THEORY, Series A 62, 280–298 (1993)

Lexicographic Ramsey Theory

P. C. FISHBURN AND R. L. GRAHAM

AT&T Bell Laboratories,
Murray Hill, New Jersey 07974

Communicated by the Managing Editors

Received September 14, 1990

Given positive integers d and n, there is an integer N such that for *every* injective map f from $\{1, ..., N\}^d$ into $\mathbb{R}$ there is a subset $A = A_1 \times A_2 \times \cdots \times A_d$ of $\{1, ..., N\}^d$ such that (1) each A_j has n elements, (2) the restriction of f to A is monotone in each coordinate, (3) there is an ordering of the coordinates such that f on A is lexicographic with respect to that ordering. Because injection f is otherwise arbitrary, the direction of monotonicity for each coordinate (increasing or decreasing) and the coordinate ordering for item (3) cannot be prespecified. Results on necessary sizes of N are included.

1. INTRODUCTION

Let d and n be fixed positive integers. We show that there is a suitably large integer $N(d, n)$ such that, when $N \geqslant N(d, n)$ and f is an injection $[x \neq y \Rightarrow f(x) \neq f(y)]$ from $\{1, ..., N\}^d$ into $\mathbb{R}$, there is an $n \times n \times \cdots \times n$ d-dimensional subcube A inside $\{1, ..., N\}^d$ on which f is monotone and lexicographic. Because f is arbitrary except for injectiveness, the sense of monotonicity on each coordinate of A and the ordering of coordinates under which f on A is lexicographic cannot be specified in advance. Suppose, for example, that $(d, n) = (3, 4)$ and that the desired conclusion holds on

$$A = \{a_1 < a_2 < a_3 < a_4\} \times \{b_1 < b_2 < b_3 < b_4\} \times \{c_1 < c_2 < c_3 < c_4\}$$

for a particular f on $\{1, ..., N\}^3$. It might then be the case that f decreases on the first coordinate of A $[i < j \Rightarrow f(a_i, b_k, c_h) > f(a_j, b_k, c_h)$ for all k and $h]$, increases on the second and third coordinates, and is lexicographic with respect to coordinate ordering 2 1 3. Then the second coordinate is lexicographically dominant and f increases there, so

$$i < j \Rightarrow f(a, b_i, c) < f(a', b_j, c') \qquad \text{for all } a, a', c, c';$$

0097-3165/93 $5.00

and the first coordinate is next dominant and f decreases on that coordinate, so

$$i<j \Rightarrow f(a_i, b, c) > f(a_j, b, c') \qquad \text{for all } b, c, c'.$$

Precise definitions and statements of main results appear in the next section. We note that the main theorem can be approached naturally by dealing with monotonicity and then with lexicography given monotonicity. Although this might not lead to the best $N(d, n)$ values, it is effective. Section 2 concludes with a short proof of the natural extension of the main theorem to m injections $f_1, ..., f_m$ considered simultaneously.

Section 3 sketches a proof of the monotonicity result for f. It is based on repeated applications of the theorem of Erdős and Szekeres [1], which says that every sequence of k^2+1 distinct numbers includes a monotone subsequence of length $k+1$. Other extensions and generalizations of the Erdős–Szekeres theorem appear in Kruskal [6].

Section 4 contains our proof of a lexicographically ordered subcube within a monotone cube. It is motivated by the fact that if f increases in each argument, if $a_{j1} < a_{j2} < \cdots < a_{jd}$ for $j = 1, ..., d$, and if

$$\begin{aligned}
x_1 &= (a_{1d}, a_{2,d-1}, a_{3,d-2}, ..., a_{d1})\\
x_2 &= (a_{11}, a_{2d}, a_{3,d-1}, ..., a_{d2})\\
x_3 &= (a_{12}, a_{21}, a_{3d}, ..., a_{d3})\\
&\vdots\\
x_d &= (a_{1,d-1}, a_{2,d-2}, a_{3,d-3}, ..., a_{dd})
\end{aligned}$$

then $f(x_i) > f(x_{i+1})$ for some $i \leqslant d-1$, or else $f(x_d) > f(x_1)$, since otherwise $f(x_1) < f(x_2) < \cdots < f(x_d) < f(x_1)$. Thus we get $f(x) > f(y)$ for an x, y pair that has $y_j > x_j$ for all but one j. Extensive use of the pigeon hole principle and associated results in Graham, Rothschild, and Spencer [3] build on this observation to produce a subcube on which f is lexicographically ordered according to some coordinate ordering.

Section 5 discusses how large N must be for $\{1, ..., N\}^d$ to (1) guarantee a monotone n^d subcube, or (2) guarantee a lexicographic n^d subcube given monotonicity, or (3) guarantee a monotone and lexicographic n^d subcube given only injectiveness. We show for (3) that

$$N > n^{(1-1/d)n^{d-1}}, \qquad d \geqslant 2, n \geqslant 3,$$

if every injection f on $\{1, ..., N\}^d$ is to have a monotone and lexicographic n^d subcube. In comparison to the Erdős–Szekeres result that $N = (n-1)^2+1$ is the smallest N that forces monotonicity on n points when $d=1$, we note that monotonicity in only one coordinate within some $n \times n$ subarray for $d=2$ is guaranteed only when

$$N > (n/e)^{1+n/2}.$$

A similar result holds for larger d. And the smallest N that serves for (2) when $d=2$ is between $(n-1)^2 + \lceil n/2 \rceil$ and $(2n-3)(n-1)+1$ inclusive.

The results of Section 5 are fragmentary and invite further research. A related problem of edge coloring on a planar grid is considered by Heinrich [4].

Our results are also related to the work of Nešetřil, Prömel, Rödl, and Voigt [7], which is based on a partition theorem for cubes due to Graham and Rothschild [2]. The main difference between their approach and ours in that we fix d and let N expand to obtain a desired substructure whereas they fix N and let d increase to generate a desired substructure. We comment further on their work in the next section.

The present research originated from a problem posed in 1984 by W. T. Trotter, who asked whether every finite three-dimensional partially ordered set is a circle order. Trotter's question is tantamount to:

> For every positive integer n is there a map C_n from $\{1, ..., n\}^3$ into planar disks such that, for all $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in $\{1, ..., n\}^3$, $(*)$
>
> $$x_j \leqslant y_j \quad \text{for} \quad j = 1, 2, 3 \Leftrightarrow C_n(x) \subseteq C_n(y)?$$

It is known (Scheinerman and Wierman [8]; see also Hurlbert [5]) that no such C exists when $\{1, ..., n\}^3$ is replaced by $\{1, ..., n\} \times \{1, ..., n\} \times \{1, 2, 3, ...\}$ for large n, but $(*)$ as stated is still unresolved.

Since planar disks are specified by three functions based on centers and radii, the representation of $(*)$ at n amounts to a finite system of quadratic inequalities in functions f_1, f_2, and f_3 from $\{1, ..., n\}^3$ into $\mathbb{R}$, and it can be presumed without loss of generality that the f_i are injective. Because $(*)$ applies to all n, the extension for $m=3$ of our main theorem in the next section shows that no generality is lost by assuming also that each f_i is fully monotone and lexicographic on $\{1, ..., n\}^3$. The patterns of monotonicity and lexicographic coordinate ordering can of course differ for $i = 1, 2, 3$, but the number of pattern combinations for (f_1, f_2, f_3) are limited. It has been determined that most of these pattern combinations cannot satisfy the representation for suitably large n, but a few patterns remain to be resolved. If they also fail to satisfy the representation, then Trotter's question will have been answered in the negative.

2. Main Results

Several definitions are needed for our main theorem. Let d be a positive integer, for each $j \in \{1, ..., d\}$ let A_j be a nonempty finite subset of positive integers, and take $A = A_1 \times A_2 \times \cdots \times A_d$. Recall that $f: A \to \mathbb{R}$ is *injective* if for all $x, y \in A$, $x \neq y \Rightarrow f(x) \neq f(y)$. Every such injection induces a linear order $<_f$ on A defined by $x <_f y \Leftrightarrow f(x) < f(y)$.

Let f be an injection from A into $\mathbb{R}$. For $a_j \in A_j$ and $x \in A$ let $a_j \mid x = (x_1, ..., x_{j-1}, a_j, x_{j+1}, ..., x_d)$. We say that f is *monotone* if for each $j \in \{1, ..., d\}$ either

(i) $\forall a_j, b_j \in A_j, \forall x \in X, a_j < b_j \Rightarrow f(a_j | x) < f(b_j | x)$ or

(ii) $\forall a_j, b_j \in A_j, \forall x \in X, a_j < b_j \Rightarrow f(a_j | x) > f(b_j | x)$.

The *monotonicity pattern* of monotone f is $s = (s_1, ..., s_d)$ with $s_j = 1$ if (i) obtains for j, and $s_j = -1$ if (ii) obtains for j. If $d = 4$ and monotone f has $s = (1, -1, 1, -1)$, then f increases in its first and third arguments and decreases in its second and fourth arguments.

Let f be a monotone injection from A into $\mathbb{R}$ with monotonicity pattern s, and let σ be a permutation on $\{1, ..., d\}$. Then f is said to be *σ-lexicographic* if for all distinct x and y in A,

$$\{x_{\sigma(j)} < y_{\sigma(j)}, x_{\sigma(i)} = y_{\sigma(i)} \text{ for all } i < j\} \Rightarrow s_{\sigma(j)} f(x) < s_{\sigma(j)} f(y).$$

We say that f is *lexicographic* if it is σ-lexicographic for some permutation σ on $\{1, ..., d\}$. If $d = 3$ and f is a σ-lexicographic monotone injection with $s = (1, 1, -1)$ and $\{\sigma(1) = 3, \sigma(2) = 2, \sigma(3) = 1\}$, then

$$x <_f y \Leftrightarrow x_3 > y_3 \quad \text{or} \quad (x_3 = y_3, x_2 < y_2) \quad \text{or} \quad (x_3 = y_3, x_2 = y_2, x_1 < y_1).$$

In our main result and later we let $\mathbf{N} = \{1, ..., N\}$.

THEOREM 1. *Suppose d and n are positive integers. Then there is a positive integer $N(d, n)$ such that for every integer $N \geq N(d, n)$ and every injection $f: \mathbf{N}^d \to \mathbb{R}$ there exist $A_j \subseteq \mathbf{N}$ with $|A_j| = n$ for $j = 1, ..., d$ such that the restriction of f on $A = A_1 \times \cdots \times A_d$ is monotone and lexicographic.*

In other words, for any fixed dimension d and edge cardinality n, there is an N such that every linear order on $\mathbf{N}^d$ has a restriction on some d-dimensional $n \times n \times \cdots \times n$ subcube that is monotone and lexicographic.

Nešetřil *et al.* [7] also investigate lexicographic substructures in large cubes, but their restrictions for subcubes concern dimensionality rather than edge cardinality. They are motivated by the observation that the ordinary lexicographic order on $\mathbf{N}^d$ is inherited by every k-dimensional subcube in which $d-k$ of the original arguments for $\mathbf{N}^d$ are fixed at points in $\mathbf{N}$ and the other k arguments range over $\mathbf{N}$. We give only the flavor of their work.

Let N and k be fixed positive integers, and let $\{\lambda_1, ..., \lambda_k\}$ be a set of k designators disjoint from $\mathbf{N}$. For $d \geqslant k$ let S_d be the set of all $\beta = (\beta_1, ..., \beta_d)$ for which each $\beta_j \in \mathbf{N} \cup \{\lambda_1, ..., \lambda_k\}$, each $\{j: \beta_j = \lambda_i\}$ is nonempty, and (by convention) $\min\{j: \beta_j = \lambda_1\} < \min\{j: \beta_j = \lambda_2\} < \cdots < \min\{j: \beta_j = \lambda_k\}$. For each $\beta \in S_d$ let β^* map $\mathbf{N}^k$ into $\mathbf{N}^d$ according to: for all $x \in \mathbf{N}^k$ and all $j \in \{1, ..., d\}$,

$$\begin{aligned} \beta^*(x)_j &= \beta_j \qquad \text{if} \quad \beta_j \in \mathbf{N} \\ &= x_i \qquad \text{if} \quad \beta_j = \lambda_i. \end{aligned}$$

Each $\beta \in S_d$ or its corresponding $\beta^*(\mathbf{N}^k)$ is referred to as a *k-dimensional subcube* of $\mathbf{N}^d$. Every injection $f: \mathbf{N}^d \to \mathbb{R}$, or its corresponding linear order $<_f$, coupled with a $\beta \in S_d$ induces a linear order $<_{f\beta}$ on $\mathbf{N}^k$ defined by

$$x <_{f\beta} y \qquad \text{if} \quad f(\beta^*(x)) < f(\beta^*(y)).$$

With $\mathbf{N}$ and k fixed, Nešetřil *et al.* show that if d is sufficiently large then for every injection $f: \mathbf{N}^d \to \mathbb{R}$ there exists a k-dimensional subcube $\beta \in S_d$ and an "ordering schema" $\mathscr{F}$ for $\mathbf{N}$ such that for all $x, y \in \mathbf{N}^k$,

$$x <_{f\beta} y \Leftrightarrow x <_{\mathscr{F}} y.$$

Each ordering schema $\mathscr{F}$ is constructed from a hierarchical tree structure whose nodes are intervals of a linear ordering $<_0$ of $\mathbf{N}$. The corresponding linear order $<_{\mathscr{F}}$ on $\mathbf{N}^k$ is induced from $\mathscr{F}$ in a lexicographic manner that is faithful to the tree's hierarchical structure. An important aspect of $<_{\mathscr{F}}$ is that it depends only on $\mathbf{N}$ and k and is therefore independent of d. But only if d is sufficiently large does one have enough variety in $\mathbf{N}^d$ to guarantee that every linear order $<_f$ on $\mathbf{N}^d$ has a pair $(\beta, \mathscr{F})$ for which $<_{f\beta} = <_{\mathscr{F}}$ on $\mathbf{N}^k$.

Returning to our present focus on fixed d and n, we use two lemmas to carry the proof of Theorem 1.

LEMMA 1. *Given positive integers d and n, there is a positive integer N_1 such that for every injection $f: \mathbf{N}_1^d \to \mathbb{R}$ there exists $A = A_1 \times \cdots \times A_d \subseteq \mathbf{N}_1^d$ with $|A_j| = n$ for all j such that the restriction of f on A is monotone.*

LEMMA 2. *Given positive integers d and n, there is a positive integer N_2 such that for every monotone injection $f: \mathbf{N}_2^d \to \mathbb{R}$ with monotonicity pattern $s = (1, 1, ..., 1)$, there exists $A = A_1 \times \cdots \times A_d \subseteq \mathbf{N}_2^d$ with $|A_j| = n$ for all j such that the restriction of f on A is lexicographic.*

These are proved in the next two sections. We show here how they combine for Theorem 1.

Proof of Theorem 1. Suppose d and n are positive integers. Let N_2 be as stated for Lemma 2, and let N_1 be as in Lemma 1 when n there is replaced throughout by N_2. Let f be an injection from $\mathbf{N}_1^d$ into $\mathbb{R}$. By Lemma 1 for (d, N_2), let $A' = A'_1 \times \cdots \times A'_d \subseteq \mathbf{N}_1^d$ with $|A'_j| = N_2$ for all j be such that the restriction of f on A' is monotone with monotonicity pattern $s = (s_1, ..., s_d)$. For each j let g_j map A'_j onto $\mathbf{N}_2$ so that g_j is increasing if $s_j = 1$ and decreasing if $s_j = -1$, and define $g: A' \to \mathbf{N}_2^d$ by $g(a_1, ..., a_d) = (g_1(a_1), ..., g_d(a_d))$. Also define f^* on $\mathbf{N}_2^d$ by

$$f^*(i_1, ..., i_d) = f(g_1^{-1}(i_1), ..., g_d^{-1}(i_d)) = f(g^{-1}(i_1, ..., i_d)).$$

Then f^* is a monotone injection on $\mathbf{N}_2^d$ with monotonicity pattern $s^* = (1, ..., 1)$. Apply Lemma 2 to f^* to obtain $B = B_1 \times \cdots \times B_d \subseteq \mathbf{N}_2^d$ with $|B_j| = n$ for all j such that the restriction of f^* on B is σ-lexicographic for some permutation σ on $\{1, ..., d\}$. Finally, for each j let $A_j = g_j^{-1}(B_j)$ so that $|A_j| = n$ and $A_j \subseteq A'_j$. With $A = A_1 \times \cdots \times A_d$, it follows that the restriction of f on A is monotone with pattern s and is σ-lexicographic.

Theorem 1 clearly holds when $N(d, n)$ stated there is set equal to N_1 of the preceding paragraph. ▌

We conclude this section with the m injections extension of Theorem 1 alluded to earlier.

COROLLARY 1. *Suppose d, n, and m are positive integers. Then there is a positive integer $N_m(d, n)$ such that for every integer $N \geqslant N_m(d, n)$ and every list $f_1, f_2, ..., f_m$ of injections from $\mathbf{N}^d$ into $\mathbb{R}$ there exists $A = A_1 \times \cdots \times A_d \subseteq \mathbf{N}^d$ with $|A_j| = n$ for all j such that the restriction of every one of f_1 through f_m on A is monotone and lexicographic.*

Proof. We use the obvious fact that if injection f on $A = A_1 \times \cdots \times A_d$ is monotone and lexicographic, and if $B = B_1 \times \cdots \times B_d$ has $\varnothing \subset B_j \subseteq A_j$ for each j, then the restriction of f on B inherits monotonicity and lexicography from f with the same s and σ.

For fixed d and general n' let $N_1(n')$ denote a satisfactory value of $N(d, n')$ for d and n' in Theorem 1. Define

$$N_{k+1}(n) = N_1(N_k(n)), \qquad k = 1, 2, ...$$

Then $N_m(n)$ suffices for the value of $N_m(d, n)$ needed for Corollary 1. For example, if $N \geqslant N_m(n)$ and if $f_1, ..., f_m$ are injections from $\mathbf{N}^d$ into $\mathbb{R}$, Theorem 1 implies an $A = A_1 \times \cdots \times A_d \subseteq \mathbf{N}^d$ with $|A_j| = N_{m-1}(n)$ for all j such that the restriction of f_1 on A is monotone and lexicographic.

A second application of Theorem 1 yields $B = B_1 \times \cdots \times B_d$ with $B_j \subseteq A_j$ and $|B_j| = N_{m-2}(n)$ for all j such that the restriction of f_2 on B is monotone and lexicographic. The obvious continuation of this procedure yields an n^d subcube of $\mathbf{N}^d$ on which each f_k is monotone and lexicographic. ∎

3. Proof of Lemma 1

Lemma 3 (Erdős and Szekeres). *For each $k \geqslant 1$, every sequence of k^2+1 distinct numbers includes a monotone subsequence of length $k+1$, and there are sequences of k^2 distinct numbers that include no monotone subsequence of length $k+1$.*

Lemma 1 follows from this if $d=1$, so assume henceforth that $d \geqslant 2$. Assume also that $n \geqslant 2$ since otherwise there is nothing to prove. We consider $d=2$ first and then note that induction on d yields Lemma 1 for all d and n.

Suppose $d=2$. Let

$$T = n^{2^{2n}}, \qquad J = (2n)^{2^{2T}}, \qquad K = 2T,$$

and let f be an injection from $\mathbf{J} \times \mathbf{K}$ into $\mathbb{R}$. We show that f is monotone on an $n \times n$ subgrid within $\mathbf{J} \times \mathbf{K}$. Let

$$J_1 = \lceil \sqrt{J} \rceil \qquad \text{and} \qquad J_{t+1} = \lceil \sqrt{J_t} \rceil \qquad \text{for} \quad t = 1, 2, \ldots, 2T.$$

Observe that $J_t \geqslant J^{-2^t}$. Repeated applications of Lemma 3 give $\mathbf{J} \supset D_1 \supset D_2 \supset \cdots \supset D_{2T}$ with $|D_t| = J_t$ and f monotone on $D_t \times \{t\}$ for each t. Since

$$J_K = J_{2T} \geqslant J^{-2^{2T}} = 2n,$$

we have a subset D_K of at least $2n$ elements of $\mathbf{J}$ such that f is monotone on $D_K \times \{k\}$ for $k = 1, \ldots, K$. Since $T = K/2$, there is a $(2n) \times T$ subgrid G in $\mathbf{J} \times \mathbf{K}$ such that either f increases up each column of G or f decreases up each column of G. With no loss of generality, relabel the rows and columns of G by $1, \ldots, 2n$ and $1, \ldots, T$, respectively, in their original order. Let

$$T_1 = \lceil \sqrt{T} \rceil \qquad \text{and} \qquad T_{t+1} = \lceil \sqrt{T_t} \rceil \qquad \text{for} \quad t = 1, 2, \ldots, 2n.$$

Observe that $T_t \geqslant T^{-2^t}$. Lemma 3 gives $\mathbf{T} \supset E_1 \supset E_2 \supset \cdots \supset E_{2n}$ with $|E_t| = T_t$ and f monotone on $\{t\} \times T_t$ for each t. Since

$$T_{2n} \geqslant T^{-2^{2n}} = n,$$

there is a subset E_{2n} of at least n elements of $\mathbf{T}$ such that f is monotone on $\{j\} \times E_{2n}$, $j = 1, \ldots, 2n$. Since f can either increase or decrease over E_{2n} for each j, we choose a majority pattern to conclude that G includes an $n \times n$ subgrid on which f is monotone. Since $K<J$, $N_1 = J$ suffices for Lemma 1.

Suppose $d \geqslant 3$. Assume that Lemma 1 holds at $d-1$, let $N(n') = N_1(d-1, n')$ satisfy the conclusion of the lemma for given $d-1$ and n', and let $N^{(1)}(n') = N(n')$ and $N^{(t+1)}(n') = N(N^{(t)}(n'))$ for $t = 1, 2, \ldots$. Define V, K, and J by

$$V = N^{(2n)}(n), \qquad K = 2^{d-1}V, \qquad J = N^{(K)}(V).$$

Let f be an injection from $\mathbf{J}^{d-1} \times \mathbf{K}$ into $\mathbb{R}$. Apply the $d-1$ version of Lemma 1 a total of K times, successively for $k = 1, 2, \ldots, K$, to generate a decreasing nested sequence of $(d-1)$-dimensional cubes within $\mathbf{J}^{d-1}$, say $D_1 \supset D_2 \supset \cdots \supset D_K$, such that D_K has V points on each edge and f on $D_k \times \{k\}$ is monotone for each k. Since there are 2^{d-1} monotonicity patterns at $d-1$, the same pattern, say $(s_1, \ldots, s_{d-1})$, occurs for at least $K/2^{d-1} = V$ of the $k \in K$. It follows that there is a d-dimensional subcube G of $\mathbf{J}^{d-1} \times \mathbf{K}$ with V points in each coordinate such that f is monotone on the first $d-1$ coordinates with pattern $(s_1, \ldots, s_{d-1})$ for each fixed point in the final coordinate.

We now apply the $d-1$ version of Lemma 1 to the final $d-1$ coordinates of G a total of $2n$ times for $2n$ successive values of the first coordinate, say $v_1, \ldots, v_{2n}$, to obtain a decreasing nested sequence $E_1 \supset E_2 \supset \cdots \supset E_{2n}$ of $(d-1)$-dimensional cubes such that E_{2n} has n points on each edge and f on $\{v_j\} \times E_j$ is monotone for $j = 1, \ldots, 2n$. At least n of the $\{v_j\} \times E_{2n}$ have the same monotonicity pattern for the final $d-1$ coordinates, namely $(s_2, \ldots, s_{d-1}, 1)$ or $(s_2, \ldots, s_{d-1}, -1)$. It follows that f is monotone on some $n \times n \times \cdots \times n$ d-dimensional subcube within G.

Hence Lemma 1 holds at d given that it holds at $d-1$, so it holds for all d by induction. ∎

4. Proof of Lemma 2

Throughout this section, $d \geqslant 2$, $d^* = d-1$, and f is an injection from a d-dimensional integer domain into $\mathbb{R}$ that increases in each coordinate. For convenience we often let 0 rather than 1 be the smallest point in a coordinate set.

We note two preliminary lemmas before embarking on the general proof of Lemma 2. Both take f on D^d with $D = \{0, 1, \ldots, Kd^*\}$, $K \geqslant 1$. Let

$$D_k = \{kd^* + i : i = 0, 1, \ldots, d^*\}, \qquad k = 0, 1, \ldots, K-1,$$

so that $D_0, \ldots, D_{K-1}$ divide D into K d-point intervals with $\max D_k = \min D_{k+1}$. We refer to the cubic subdomain

$$D_{k_1} \times D_{k_2} \times \cdots \times D_{k_d}, \qquad \text{each } k_j \in \{0, 1, \ldots, K-1\},$$

as *box* $(k_1, ..., k_d)$. D^d has K^d boxes. We distinguish d points in box $(k_1, ..., k_d)$ as follows:

$$b_1 = (k_1 d^* + d^*, k_2 d^* + d^* - 1, k_3 d^* + d^* - 2, ..., k_d d^*)$$
$$b_2 = (k_1 d^*, k_2 d^* + d^*, k_3 d^* + d^* - 1, ..., k_d d^* + 1)$$
$$b_3 = (k_1 d^* + 1, k_2 d^*, k_3 d^* + d^*, ..., k_d d^* + 2)$$
$$b_4 = (k_1 d^* + 2, k_2 d^* + 1, k_3 d^*, ..., k_d d^* + 3)$$
$$\vdots$$
$$b_d = (k_1 d^* + d^* - 1, k_2 d^* + d^* - 2, k_3 d^* + d^* - 3, ..., k_d d^* + d^*).$$

LEMMA 4. *For each box $(k_1, ..., k_d)$, either $f(b_i) > f(b_{i+1})$ for some $1 \leqslant i \leqslant d^*$, or $f(b_d) > f(b_1)$.*

Proof. Otherwise $f(b_1) < f(b_2) < \cdots < f(b_d) < f(b_1)$. ∎

Note that if $f(b_i) > f(b_{i+1})$ or $f(b_d) > f(b_1)$, then the b with the smaller f value is larger than the other b in every component except one. If the same exceptional coordinate obtains for a large number of boxes, it is a candidate for the lexicographically dominant coordinate, i.e., $\sigma(1)$.

We use the pigeon hole principle to obtain enough boxes of a specific type. Color box $(k_1, ..., k_d)$ with color c_i if $f(b_i) > f(b_{i+1})$, and with color c_d if $f(b_d) > f(b_1)$. A box can have several colors, but what matters is that it have at least one as guaranteed by Lemma 4.

For each j in $\{1, ..., d\}$ and fixed $k_{(j)} = (k_1, ..., k_{j-1}, k_{j+1}, ..., k_d)$, let $R(k_{(j)})$ denote the sequence of K boxes $(k_j = 0, k_{(j)})$, $(k_j = 1, k_{(j)})$, ..., $(k_j = K-1, k_{(j)})$. We refer to this sequence as *row* $R(k_{(j)})$. Each j has K^{d-1} rows since there are K^{d-1} choices for $k_{(j)}$. Suppose none of these rows for j has more than p boxes with color c_j. Then at most pK^{d-1} boxes in D^d have color c_j.

LEMMA 5. *For some $j \in \{1, ..., d\}$ and some $k_{(j)} \in \{0, 1, ..., K-1\}^{d-1}$, at least K/d boxes in row $R(k_{(j)})$ have color c_j.*

Proof. Otherwise there are fewer than $d(K/d)\, K^{d-1} = K^d$ boxes. ∎

By increasing K we force the number of c_j boxes in some $R(k_{(j)})$ to be arbitrarily large. Let $p = \lceil K/d \rceil$ and with no loss of generality assume on the basis of Lemma 5 that row $R(k^*_{(1)})$ with $k^*_{(1)} = (0, 0, ..., 0)$ has p boxes of color c_1. Denote p values of k_1 for which $(k_1, k^*_{(1)})$ has color c_1 by $0 \leqslant a_1 < a_2 < \cdots < a_p < K$. Then for $i = 1, ..., p$,

$$f(a_i d^* + d^*, d^* - 1, d^* - 2, ..., 0) > f(a_i d^*, d^*, d^* - 1, ..., 1).$$

Since $a_i - 1 \geqslant a_{i-1}$ for each $i > 1$, monotonicity yields $f(a_i d^* + d^*, d^* - 1, ..., 0) > f(a_{i-1} d^* + d^*, d^*, ..., 1)$ for each $i > 1$. For notational convenience, relabel $a_1 d^*$, $a_1 d^* + d^*$, $a_2 d^* + d^*$, ..., $a_p d^* + d^*$ on coordinate 1 as 0, 1, 2, ..., p. Then

$$f(j, d-2, d-3, ..., 0) > f(j-1, d-1, d-2, ..., 1) \qquad \text{for} \quad j = 1, ..., p. \tag{1}$$

We use (1) as our point of departure for a general proof of Lemma 2. The proof has $d-1$ steps. The first establishes a lexicographically dominant coordinate which, by convention as in the preceding paragraph, will be coordinate 1. The second establishes a next dominant coordinate, and so forth. In each step we make a uniform expansion of the main cube of the preceding step (step 0 uses D^d as above) by inserting W points between each pair of adjacent points on each coordinate of that cube. Hence if the main cube at step j has E points on each coordinate, the main cube at step $j+1$ will have $E + W(E-1)$ points on each coordinate. The value of W will change from step to step. The ensuing subsections focus on a dominant coordinate, a next dominant coordinate, and a later coordinate in the lexicographic hierarchy.

A Dominant Coordinate

Given D^d with $D = \{0, 1, ..., Kd^*\}$ as above, insert W new points between each two adjacent points on each coordinate of D^d. We do this uniformly over the entire domain since it is not known at the start which j and $k_{(j)}$ will emerge from Lemma 5 as the basis of the refinements described just prior to (1). The preceding analysis is unchanged for the special points used there and their associated K^d boxes. But we now have $Kd^* + 1 + WKd^*$ points in each coordinate of the new cube instead of the original $Kd^* + 1$.

Having made the insertions for W, we assume with no loss of generality that the preceding analysis based on D^d yields (1). However, because of the insertions there will now be W additional points for each coordinate $j \geqslant 2$ between the two points for coordinate j shown in (1). For convenience, relabel these $W+2$ points as 0, 1, ..., $W+1$. Then (1) becomes

$$f(j, 0, ..., 0) > f(j-1, W+1, ..., W+1), \qquad j = 1, ..., p. \tag{2}$$

Here, and in later expansions, we assume of course that f is a monotone increasing injection on the largest cube defined .thus far. Successive relabelings of points that preserve order on each coordinate and give integer domains are tacitly assumed. Monotonicity with (2) gives

$$x_1 > y_1 \Rightarrow f(x) > f(y), \qquad \forall x, y \in \{0, ..., p\} \times \{0, ..., W+1\}^{d-1},$$

so that coordinate 1 is lexicographically dominant within the indicated subdomain.

Suppose $d=2$. The conclusion of Lemma 2 then follows from $p=n-1$ and $W=n-2$. Since $p=\lceil K/d \rceil$, we can take $K=2n-3$. A suitable N_2 for Lemma 2 is

$$N_2 = K+1+WK = (2n-3)(n-1)+1.$$

A Next Dominant Coordinate

Suppose henceforth that $d \geqslant 3$. We proceed from (2) and now denote K, p, and W by K_1, p_1, and W_1 as a reminder of their use in the first step. Later we insert W_2 points uniformly in the main cube of the preceding step, but for the time being we work within

$$\mathscr{S}_1 = \{0, 1, ..., p_1\} \times \{0, 1, ..., W_1+1\}^{d-1}.$$

No generality is lost in doing this since the same type of substructure would emerge if later insertions were suppressed.

To obtain a second most important coordinate within $\mathscr{S}_1$ we focus on the last $d-1$ coordinates, take

$$K_2(d-2) = W_1 + 1,$$

which is used later to define W_1 on the basis of K_2, define

$$D'_k = \{k(d-2)+i : i=0, 1, ..., d-2\}, \qquad k=0, 1, ..., K_2-1,$$

and refer to the cubic subdomain

$$D'_{k_2} \times D'_{k_3} \times \cdots \times D'_{k_d}$$

within the final $d-1$ coordinates of $\mathscr{S}_1$ as *box* $(k_2, ..., k_d)$. There are K_2^{d-1} boxes of this type, each with $d-1$ points along each edge. We proceed in the manner described earlier for Lemmas 4 and 5 but need to pay attention to the lexicographically dominant first coordinate of (2).

Fix $x_1 \in \{0, 1, ..., p_1\}$. We suppress the first coordinate and repeat the analysis around Lemmas 4 and 5 for rows $R'(k'_{(j)})$ with $j \geqslant 2$ and $k'_{(j)} = (k_2, ..., k_{j-1}, k_{j+1}, ..., k_d)$ to obtain a $j \geqslant 2$ and a $k'_{(j)}$ such that at least $K_2/(d-1)$ new boxes in row $R'(k'_{(j)})$ have the same color c_j. Let

$$p_2 = \lceil K_2/(d-1) \rceil.$$

This p_2-monochromatic result holds for each x_1 in $\{0, 1, ..., p_1\}$. Therefore at least $\lceil (p_1+1)/(d-1) \rceil$ rows of type R' for some $j \geqslant 2$, corresponding to different x_1, have p_2 boxes of color c_j. Assume for definiteness that this is true for $j=2$. That is, there are $\lceil (p_1+1)/(d-1)\rceil$ values of x_1 in $\{0, 1, ..., p_1\}$ such that some $R'(k'_{(2)})$, where $k'_{(2)}$ can depend on x_1, contains p_2 new boxes of color c_2. We refer to these x_1's as *special.*

There are K_2^{d-2} possibilities for $k'_{(2)}$. The average number of special x_1's per $k'_{(2)}$ is

$$\lceil (p_1+1)/(d-1)\rceil / K_2^{d-2} > \frac{K_1}{d(d-1)\,K_2^{d-2}}$$

since $p_1 = \lceil K_1/d \rceil$. Let

$$H = \lceil \lceil (p_1+1)/(d-1)\rceil / K_2^{d-2} \rceil$$

and suppose with no loss of generality that $k'_{(2)} = (0, ..., 0)$ has H special x_1's. By making K_1 large, we can get H as large as we please.

The p_2 values of k_2 for each of these H special x_1's are, say, $0 \leqslant a_1(x_1) < a_2(x_1) < \cdots < a_{p_2}(x_1) < K_2$ and can depend on x_1 as indicated. There are

$$\binom{K_2}{p_2} \text{ possibilities for } (a_1, a_2, ..., a_{p_2}).$$

To ensure that at least r of the $(a_1(x_1), ..., a_{p_2}(x_1))$ are identical within the set of H special x_1's, it suffices to have

$$\left[H \Big/ \binom{K_2}{p_2} \right] \geqslant r.$$

Define r by equality in the preceding expression and for convenience relabel the r of the H x_1's that have the same $(a_1, ..., a_{p_2})$ as $0, 1, ..., r-1$. Also relabel the relevant p_2+1 values of coordinate 2 as $0, 1, ..., p_2$, and the two adjacent values on each coordinate $j \geqslant 3$ that pertain to $k_{(2)} = (0, ..., 0)$ as 0 and 1. Then

$$f(i, 0, 0, ..., 0) > f(i-1, p_2, 1, ..., 1), \qquad i = 1, ..., r-1 \tag{3}$$

for lexicographically dominant coordinate 1, and

$$f(i, j, 0, ..., 0) > f(i, j-1, 1, ..., 1), \qquad 0 \leqslant i \leqslant r-1,\ j = 1, ..., p_2, \tag{4}$$

for next dominant coordinate 2.

We now insert W_2 new points between each two adjacent points on each coordinate of the main cube of the preceding step. Since that cube had $K_1(d-1)(W_1+1)+1$ points on each coordinate, our new main cube has $K_1(d-1)(W_1+1)(W_2+1)+1$ points on each coordinate. Relabel the

W_2+2 points on each coordinate $j \geqslant 3$ between 0 and 1 inclusive in (3) and (4) as $0, 1, ..., W_2+1$ to obtain

$$\begin{aligned} f(i, 0, 0, ..., 0) &> f(i-1, p_2, W_2+1, ..., W_2+1) \\ f(i, j, 0, ..., 0) &> f(i, j-1, W_2+1, ..., W_2+1) \end{aligned} \tag{5}$$

with the domains for i and j as shown in (3) and (4).

Suppose $d=3$. We work backward to determine a suitable N_2 for Lemma 2 that guarantees an $n \times n \times n$ subdomain on which f is lexicographic. For (5) we require

$$W_2 = n-2, \qquad p_2 = n-1, \qquad r=n.$$

Satisfactory values of other parameters are

$$\begin{aligned} H &= n\binom{K_2}{n-1} \\ p_1 &= 2nK_2\binom{K_2}{n-1} \\ K_2 &= p_2(d-1) = 2(n-1) \\ K_1 &= p_1 d = 12n(n-1)\binom{2(n-1)}{n-1} \\ W_1 &= K_2(d-2)-1 = 2n-3. \end{aligned}$$

Finally, since the main cube for (5) has $K_1(d-1)(W_1+1)(W_2+1)+1$ points on each coordinate, we use this for N_2 to get

$$N_2 = 48n(n-1)^3\binom{2(n-1)}{n-1}+1.$$

Hence Lemma 2 holds at $d=3$ for this N_2.

A Later Coordinate

Suppose henceforth that $d \geqslant 4$. N_2 becomes very large indeed for these cases, and we will not give an explicit value for N_2 when $d \geqslant 4$. An outline for step q, $3 \leqslant q \leqslant d-1$, follows.

We begin with an array like (5) that has $q-1$ rows from the successive dominance, by convention, of coordinates 1 through $q-1$. No generality is lost by assuming that the subdomain at this point is

$$\mathscr{S}_{q-1} = \{0, 1, ..., r-1\}^{q-1} \times \{0, 1, ..., W_{q-1}+1\}^{d-q+1},$$

so the first and last rows of the array are

$$f(i, 0, ..., 0) > f(i-1, r-1 \ (q-2 \text{ times}), W_{q-1}+1 \ (d-q+1 \text{ times}))$$

and

$$f(x_1, ..., x_{q-2}, i, 0, ..., 0) > f(x_1, ..., x_{q-2}, i-1, W_{q-1}+1, ..., W_{q-1}+1)$$

for all $i \in \{1, ..., r-1\}$ and all $(x_1, ..., x_{q-2}) \in \{0, ..., r-1\}^{q-2}$.

Boxes and rows for coordinates q through d are defined in the usual way. Each box is denoted by $(k_q, ..., k_d)$. There are K_q^{d-q+1} such boxes, where

$$K_q(d-q) = W_{q-1}+1.$$

A row of boxes for $j \geqslant q$ can be expressed as $(k_j = 0, k_{(j)})$, $(k_j = 1, k_{(j)})$, ..., $(k_j = K_q - 1, k_{(j)})$, where $k_{(j)}$ is a $(d-q)$-tuple in $\{0, ..., K_q - 1\}^{d-q}$ that omits the argument for coordinate j.

Let $p_q = \lceil K_q/(d-q+1) \rceil$. For each $(x_1, ..., x_{q-1})$ in the leading r^{q-1} cube of $\mathscr{S}_{q-1}$ there is a $j \geqslant q$ and a row for some $k_{(j)}$ such that p_q boxes in that row have color c_j. The values of k_j for these boxes are, say,

$$0 \leqslant a_1(x_1, ..., x_{q-1}) < a_2(x_1, ..., x_{q-1}) < \cdots < a_{p_q}(x_1, ..., x_{q-1}) < K_q.$$

Such an a_i sequence can occur in $\binom{K_q}{p_q}$ ways. Moreover, since there are $d-q+1$ choices for j and K_q^{d-q} possibilities for $k_{(j)}$ with fixed j, there are

$$M = \binom{K_q}{p_q} (d-q+1) \, K_q^{d-q}$$

possibilities for $(j, k_{(j)}, a_i$ sequence$)$ for the monochromatic c_j result for each $(x_1, ..., x_{q-1})$.

Let m be a positive integer. It then follows from Graham, Rothschild, and Spencer [3, pp. 95–97] that for suitably large r every $(x_1, ..., x_{q-1})$ in an m^{q-1} subcube of the leading cube in $\mathscr{S}_{q-1}$ has the same one of the M possibilities for $(j, k_{(j)}, a_i$ sequence$)$ as its monochromatic realization. Let $j = q$ be the next dominant coordinate thus identified, set m equal to $p_q + 1$, insert W_q points uniformly in the preceding main cube, and relabel in the usual way to arrive at

$$f(i, 0, ..., 0) > f(i-1, m-1 \ (q-1 \text{ times}), W_q + 1 \ (d-q \text{ times}))$$
$$f(x_1, i, 0, ..., 0) > f(x_1, i-1, m-1 \ (q-2 \text{ times}), W_q + 1 \ (d-q \text{ times}))$$
$$\vdots$$
$$f(x_1, ..., x_{q-1}, i, 0, ..., 0) > f(x_1, ..., x_{q-1}, i-1, W_q + 1 \ (d-q \text{ times}))$$

on subdomain $\mathscr{S}_q = \{0, 1, ..., m-1\}^q \times \{0, 1, ..., W_q+1\}^{d-q}$. If $d=q+1$, the process terminates with $W_q = n-2$, $m=n$, and preceding parameter values that support these terminal values. If $d>q+1$, we begin the next step at this point. ∎

5. Comments on N

Let $F(d, N)$ be the set of all injections from $\mathbf{N}^d$ into $\mathbb{R}$. Also let

$$N_1(d, n) = \min\{N: \text{for every } f \in F(d, N) \text{ there is an } n^d \text{ subcube of } \mathbf{N}^d \text{ on which } f \text{ is monotone}\},$$

$$N_2(d, n) = \min\{N: \text{for every monotone } f \in F(d, N) \text{ there is an } n^d \text{ subcube of } \mathbf{N}^d \text{ on which } f \text{ is lexicographic}\},$$

$$N_3(d, n) = \min\{N: \text{for every } f \in F(d, N) \text{ there is an } n^d \text{ subcube of } \mathbf{N}^d \text{ on which } f \text{ is monotone and lexicographic}\}.$$

We know from Lemma 3 and constructions in preceding sections that

$$N_1(1, n) = (n-1)^2 + 1$$

$$N_1(2, n) \leqslant (2n)^{2^{2T}} \quad \text{with} \quad T = n^{2^{2n}}$$

$$N_2(2, n) \leqslant (2n-3)(n-1) + 1$$

$$N_2(3, n) \leqslant 48n(n-1)^3 \binom{2(n-1)}{n-1} + 1 \sim 12n^{7/2}\, 4^n/\sqrt{\pi}.$$

An upper bound for $N_3(2, n)$ results when n in the upper bound on $N_1(2, n)$ is replaced by $(2n-3)(n-1)+1$.

Because the constructions of the two preceding sections proceed in a greedy manner, the upper bounds obtained there are likely to be much larger than the $N_i(d, n)$. We therefore derive lower bounds on some $N_i(d, n)$ to get an idea of how large N must be to guarantee various conclusions. We begin with N_3, followed by N_1 and N_2.

Theorem 2. *For all $d \geqslant 2$ and $n \geqslant 3$,*

$$N_3(d, n) > n^{(1-1/d)\, n^{d-1}}.$$

Proof. Given d and n, we consider for each $N > n$ the probability distribution that assigns probability $1/N^d!$ to each linear order on $\mathbf{N}^d$. Let $K = \binom{N}{n}^d$ denote the number of n^d subcubes of $\mathbf{N}^d$, enumerate these as $1, ..., K$, and let G_k be the event that a randomly chosen linear order on $\mathbf{N}^d$

is monotone and lexicographic on subcube k. The probability for each linear order on the subcube is $1/n^d!$, there are 2^d monotonicity patterns (s) and $d!$ coordinate orderings (σ), and each (s, σ) characterizes exactly one linear order on the subcube. Therefore

$$\Pr(G_k) = 2^d(d!)/n^d!.$$

Let $P_3(N)$ be the probability that a randomly chosen linear order on $\mathbf{N}^d$ is monotone and lexicographic on at least one n^d subcube. Then

$$P_3(N) = \Pr\left(\bigcup_{k=1}^{K} G_k\right) < \sum_{k=1}^{K} \Pr(G_k) = \binom{N}{n}^d 2^d(d!)/n^d!.$$

Clearly $P_3(N) = 1$ if and only if $N \geqslant N_3(d, n)$. Hence if

$$\binom{N}{n}^d 2^d(d!)/n^d! < 1 \tag{6}$$

then $P_3(N) < 1$ and $N < N_3(d, n)$ so that some linear order on $\mathbf{N}^d$ has no n^d subcube on which it is monotone and lexicographic. Since $N!/(N-n)! < N^n$, (6) holds if

$$N^{dn} \leqslant \frac{(n!)^d\,(n^d!)}{2^d(d!)}. \tag{7}$$

Consider

$$N \leqslant n^{(1-1/d)\,n^{d-1}} \leqslant \frac{(n!)^{1/n}\,(n^d!)^{1/dn}}{2^{1/n}(d!)^{1/dn}}.$$

We claim that the second inequality holds when $d \geqslant 2$ and $n \geqslant 3$. Hence if the first inequality also holds, then (7) holds, so (6) holds and $N < N_3(d, n)$. Therefore, if (6) is to fail, it must be true that

$$N > n^{(1-1/d)n^{d-1}},$$

and the conclusion of Theorem 2 follows.

It remains to verify the preceding claim. By the lower bound version of Stirling's inequality, i.e., $m! > \sqrt{2\pi m}\,(m/e)^m$,

$$\frac{(n!)^{1/n}\,(n^d!)^{1/dn}}{2^{1/n}(d!)^{1/dn}} > n^{(1-1/d)\,n^{d-1}}\left[\frac{(2\pi)^{(d+1)/2dn}\,n^{1+1/n+n^{d-1}/d}}{2^{1/n}(d!)^{1/dn}\,e^{1+n^{d-1}/d}}\right].$$

Hence the claim is true if the ratio in brackets is $\geqslant 1$. Taking logarithms, we have $[\cdots]\geqslant 1$ if and only if

$$\left(1+\frac{n^{d-1}}{d}\right)(\log n-1)+\frac{\log n}{n}$$
$$\geqslant \frac{1}{n}\left[\frac{1}{d}\{\log(d!)-(\log 2\pi)/2\}+\log 2-\frac{\log 2\pi}{2}\right].$$

Given $n\geqslant 3$, the left side is positive since $\log 3=1.098\cdots$. If $d=2$, the right side is negative, so the inequality holds at $d=2$. If $d\geqslant 3$, the right side is positive and decreases in n, and the left side increases in n, so the inequality holds for all $n\geqslant 3$ if it holds at $n=3$. It is routine to check that it holds when $n=3$ and $d\geqslant 3$, so the claim is true. ∎

Since the full monotone and lexicographic case is the most restrictive of those considered, the probabilistic method of the preceding proof will not give a lower bound on $N_1(d,n)$ essentially larger than that on $N_3(d,n)$ in Theorem 2. In fact, a substantial part of that bound obtains when f is restricted to be monotone in only one coordinate. Let

$N_1'(d,n)=\min\{N$: for every $f\in F(d,N)$ there is an n^d subcube of $\mathbf{N}^d$, say $\{(x_1,...,x_d)\}$, such that f is either increasing in x_1 for every $(x_2,...,x_d)$ or decreasing in x_1 for every $(x_2,...,x_d)\}$.

THEOREM 3. *For all $d\geqslant 2$ and $n\geqslant 2$,*

$$N_1'(d,n)\geqslant (n/e)^{1+n^{d-1}/d}\, n^{n^{d-2}/(2d)}.$$

Proof. The proof begins as the preceding proof, but in place of G_k let H_k be the event that a randomly chosen linear order on $\mathbf{N}^d$ is monotone in the first coordinate on subcube k. Within subcube k, say $\{(x_1,...,x_d)\}$, we refer to the n points $(x_1,x_2^0,...,x_d^0)$ for all x_1 and fixed $(x_2^0,...,x_d^0)$ as a *row*. The probability that a given row increases in x_1 is $1/n!$. Since monotonicity events for different rows are independent and there are n^{d-1} rows, the probability that all rows increase in x_1 or all rows decrease in x_1 is

$$\Pr(H_k)=2/(n!)^{n^{d-1}}.$$

Let $P_1(N)$ be the probability that a randomly chosen linear order on $\mathbf{N}^d$ is monotone in the first coordinate of at least one n^d subcube. Then

$$P_1(N)=\Pr\left(\bigcup_{k=1}^{K} H_k\right)<\sum_{k=1}^{K}\Pr(H_k)=\binom{N}{n}^d 2/(n!)^{n^{d-1}}.$$

If

$$\binom{N}{n}^d < (n!)^{n^{d-1}}/2 \tag{8}$$

then $N < N_1'(d, n)$. Inequality (8) holds if

$$N \leqslant \frac{(n!)^{(d+n^{d-1})/dn}}{2^{1/dn}}.$$

It is easily checked that the right side of this inequality exceeds the lower bound in Theorem 3 when $d \geqslant 2$ and $n \geqslant 2$. Hence if N is no greater than that bound, then (8) holds. Therefore the conclusion of Theorem 3 must hold if (8) is to fail. ∎

In the lowest-dimensional case of $d = 2$, Theorem 3 gives

$$N_1'(2, n) > (n/e)^{1+n/2},$$

which suffices to show that $N_1(2, n)$ is not polynomial; that is, there are no constants c_0 and c_1 such that $N_1(2, n) \leqslant c_0 n^{c_1}$ for all n.

A comparison between Theorems 2 and 3 suggests also that the lion's share of $N_3(d, n)$ is borne by $N_1(d, n)$, i.e., by monotonicity, hence that the lexicographic restriction for N_3 plays a secondary role. Although this is conjectural, it is supported by our present inability to obtain a larger than polynomial lower bound on $N_2(d, n)$. We prove a lower bound for $d = 2$ and then comment on higher dimensions.

THEOREM 4. *For all* $n \geqslant 2$,

$$(n-1)^2 + \lceil n/2 \rceil \leqslant N_2(2, n) \leqslant (2n-3)(n-1) + 1.$$

Proof. The upper bound was noted after (2) in Section 4. To verify the lower bound let $N = (n-1)^2 + \lceil n/2 \rceil - 1$. We refer to $\{(i, j) \in \mathbf{N}^2 : i + j = p\}$ as *diagonal* p and say that we *go down* diagonal p as i increases in $(i, p - i)$. Let f be a monotone increasing injection on $\mathbf{N}^2$ that satisfies

$$f(i, j) = i + j + \varepsilon_{ij}, \qquad |\varepsilon_{ij}| \leqslant \tfrac{1}{2},$$

where $\varepsilon_{1j} = 0$ for all j, ε_{ij} decreases as we go down each of the first $n - 1$ diagonals, increases as we go down each of the next $n - 1$ diagonals, decreases as we go down each of the next $n - 1$ diagonals, and so forth.

We show that $\mathbf{N}^2$ includes no $n \times n$ subarray on which f is lexicographic with coordinate 1 dominant. A similar proof yields the same conclusion with coordinate 2 dominant, so $N_2(2, n) > N$.

Suppose in fact that $\mathbf{N}^2$ includes an $n \times n$ subarray $X_1 \times X_2$ on which f is lexicographic with coordinate 1 dominant. We can do no better than to have X_2 as a set of n consecutive integers, say b through $b+n-1$. Then, with $X_1 = \{i_1 < i_2 < \cdots < i_n\}$, we require

$$f(i_{k+1}, b) > f(i_k, b+n-1) \qquad \text{for} \quad k = 1, ..., n-1.$$

The definition of f implies $i_{k+1} - i_k \geqslant n-1$ for each $k < n$. If $i_{k+1} - i_k = n-1$ then ε increases as we go down diagonal $i_k + b + n - 1$ since $(i_k + n - 1, b)$ and $(i_k, b+n-1)$ are on the same diagonal. Moreover, if $i_{k+1} - i_k = n-1$, then $i_{k+2} - i_{k+1} \geqslant n$ since ε decreases down the $(n-1)$st diagonal to the right of diagonal $i_{k+1} + b$. Therefore, since there are $n-1$ differences $i_{k+1} - i_k$,

$$i_n - i_1 = \sum_{k=1}^{n-1} (i_{k+1} - i_k) \geqslant (n-1)^2 + \left\lfloor \frac{n-1}{2} \right\rfloor.$$

Since $i_1 \geqslant 1$, $i_n \geqslant (n-1)^2 + \lfloor (n-1)/2 \rfloor + 1 = (n-1)^2 + \lceil n/2 \rceil$. But then $i_n > N$, for a contradiction. ∎

When $n = 3$, the lower and upper bounds on $N_2(2, 3)$ in Theorem 4 are 6 and 7, respectively. We have checked that $N_2(2, 3) = 6$ but do not know whether $N_2(2, n)$ equals the lower bound for larger n.

Straightforward extension of the preceding proof that balances the coordinates in the sense that $f(x_1, ..., x_d)$ is approximately $x_1 + x_2 + \cdots + x_d$ shows that $N_2(d, n)$ exceeds $(n-1)^d$. It is not known at this time whether $N_2(d, n)$ is substantially larger than $(n-1)^d$ for $d \geqslant 3$.

References

1. P. Erdös and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 464–470.
2. R. L. Graham and B. L. Rothschild, Ramsey's theorem for n-parameter sets, *Trans. Amer. Math. Soc.* **159** (1971), 257–292.
3. R. L. Graham, B. L. Rothschild, and J. H. Spencer, "Ramsey Theory," Wiley, New York, 1980.
4. K. Heinrich, Coloring the edges of $K_m \times K_m$, *J. Graph Theory* **14** (1990), 575–583.
5. G. H. Hurlbert, A short proof that N^3 is not a circle containment order, *Order* **5** (1988), 235–237.
6. J. B. Kruskal, Monotonic subsequences, *Proc. Amer. Math. Soc.* **4** (1953), 264–274.
7. J. Nešetřil, H. J. Prömel, V. Rödl, and B. Voigt, Canonizing ordering theorems for Hales Jewett structures, *J. Combin. Theory Ser. A* **40** (1985), 394–408.
8. E. R. Scheinerman and J. C. Wierman, On circle containment orders, *Order* **4** (1988), 315–318.