The minimum number of monotone subsequences

Joseph Samuel Myers*

Department of Pure Mathematics and Mathematical Statistics
Centre for Mathematical Sciences
Wilberforce Road
Cambridge
CB3 0WB
United Kingdom
J.S.Myers@dpmms.cam.ac.uk

Submitted: Jul 4, 2002; Accepted: Nov 25, 2002; Published: Dec 13, 2002
MR Subject Classifications: 05A05, 05C35, 05D10

Abstract

Erdős and Szekeres showed that any permutation of length $n \geq k^2 + 1$ contains a monotone subsequence of length $k+1$. A simple example shows that there need be no more than $(n \bmod k)\binom{\lceil n/k \rceil}{k+1} + (k - (n \bmod k))\binom{\lfloor n/k \rfloor}{k+1}$ such subsequences; we conjecture that this is the minimum number of such subsequences. We prove this for $k = 2$, with a complete characterisation of the extremal permutations. For $k > 2$ and $n \geq k(2k-1)$, we characterise the permutations containing the minimum number of monotone subsequences of length $k+1$ subject to the additional constraint that all such subsequences go in the same direction (all ascending or all descending); we show that there are $2\binom{k}{n \bmod k}C_k^{2k-2}$ such extremal permutations, where $C_k = \frac{1}{k+1}\binom{2k}{k}$ is the k^{th} Catalan number. We conjecture, with some supporting computational evidence, that permutations with a minimum number of monotone $(k+1)$-subsequences must have all such subsequences in the same direction if $n \geq k(2k-1)$, except for the case of $k = 3$ and $n = 16$.

1 Introduction

A well-known result of Erdős and Szekeres [2] may be expressed as follows:

Theorem 1 (Erdős and Szekeres [2]) *Let n and k be positive integers. If $n \geq k^2 + 1$, then in any permutation of the integers $\{0, 1, \ldots, n-1\}$ there is a monotone subsequence of length $k+1$.*

*Research supported by EPSRC studentship 99801140.

This problem leads to many variations, a survey of which has been made by Steele [5]. Here we consider an extremal problem that arises as a variation; this problem was posed by Mike Atkinson, Michael Albert and Derek Holton. If $n \geq k^2+1$, then we know there is at least one monotone subsequence of length $k+1$; how many such sequences must there be? We write $m_k(S)$ for the number of monotone subsequences of length $k+1$ in the permutation S. This problem is related to a question of Erdős [1] in Ramsey theory asking for the minimum number of monochromatic K_t subgraphs in a 2-coloured K_n; Erdős's conjecture about the answer to that question (that the minimum was given by random colourings) was disproved by Thomason [6].

Some upper and lower bounds are obvious. For an upper bound, note that in a random permutation, any given subsequence of length $k+1$ is monotone with probability $2/(k+1)!$. Thus some permutation has at most

$$\frac{2}{(k+1)!}\binom{n}{k+1}$$

monotone subsequences of length $k+1$. For a lower bound, note that any subsequence of length k^2+1 must have a monotone subsequence of length $k+1$, and any sequence of length $k+1$ is in $\binom{n-k-1}{k^2-k}$ sequences of length k^2+1. Thus there are at least

$$\frac{\binom{n}{k^2+1}}{\binom{n-k-1}{k^2-k}} = \frac{1}{\binom{k^2+1}{k+1}}\binom{n}{k+1}$$

monotone subsequences of length $k+1$.

A simple example will, in fact, give a better upper bound than a random permutation; this bound is, for large k, half way (geometrically) between the upper and lower bounds just given. Consider the permutation

$$\begin{aligned}
&\lfloor n/k \rfloor - 1, \lfloor n/k \rfloor - 2, \ldots, 0, \\
&\lfloor 2n/k \rfloor - 1, \lfloor 2n/k \rfloor - 2, \ldots, \lfloor n/k \rfloor, \\
&\ldots, \\
&n-1, n-2, \ldots, \lfloor (k-1)n/k \rfloor.
\end{aligned}$$

(This permutation is illustrated in Figure 1 for $n = 17$ and $k = 3$.) This permutation is made up of k monotone descending subsequences, each of length $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$; clearly it has no monotone ascending subsequences of length $k+1$, and any monotone descending subsequences it has of length $k+1$ must lie entirely within just one of the k monotone descending subsequences into which it is divided. Thus the number of monotone subsequences of length $k+1$ is

$$(n \bmod k)\binom{\lceil n/k \rceil}{k+1} + (k - (n \bmod k))\binom{\lfloor n/k \rfloor}{k+1} \approx \frac{1}{k^k}\binom{n}{k+1}.$$

Let this number be known as $M_k(n)$. I conjecture that this is in fact the minimum number of monotone subsequences of length $k+1$.

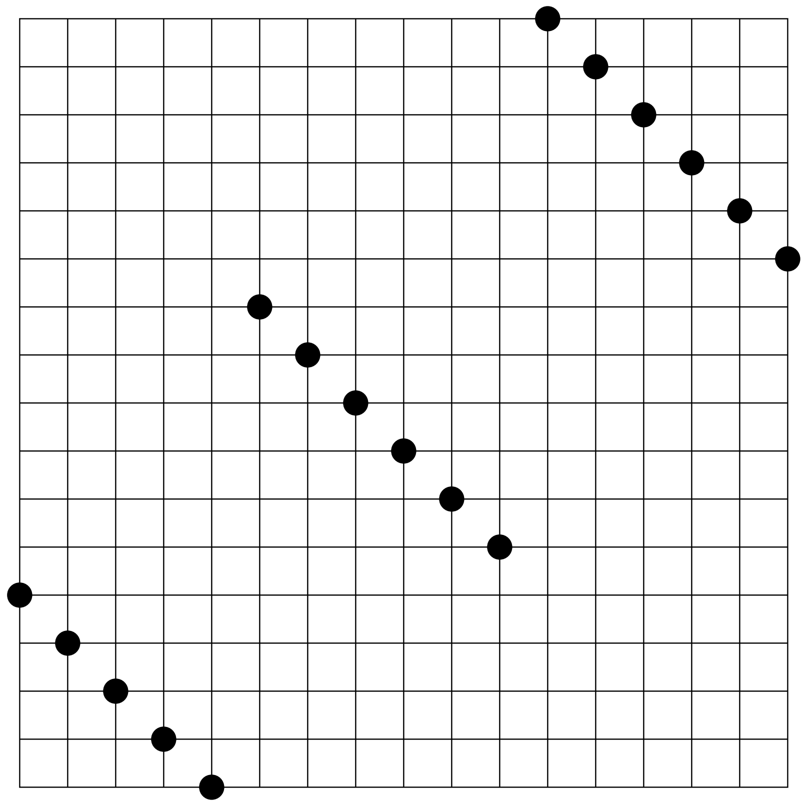

Figure 1: Canonical extremal permutation for $n = 17$ and $k = 3$

Conjecture 2 *Let n and k be positive integers. In any permutation of the integers $\{0, 1, \ldots, n-1\}$ there are at least $M_k(n)$ monotone subsequences of length $k+1$.*

A natural weaker conjecture is that this is asymptotically correct.

Conjecture 3 *Let k be a positive integer and let $n \to \infty$. In any permutation of the integers $\{0, 1, \ldots, n-1\}$ there are at least $(1+o(1))\binom{n}{k+1}/k^k$ monotone subsequences of length $k+1$.*

It would also be interesting to know the extremal configurations. It appears from computation that the behaviour for $k = 2$ is entirely different from that for $k > 2$ (although I do not have a proof that $M_k(n)$ is the correct extremum, or that the conjectured sets of extremal configurations are complete, except for $k = 2$). For $k = 2$, n even, there are $2^{n/2}$ extremal configurations; for $k = 2$, n odd, there are 2^{n-1} extremal configurations. These configurations are described in Section 2. Some of these configurations have both ascending and descending monotone subsequences of length $k+1$. For $k > 2$, the extremal configurations, provided n is sufficiently large in terms of k, appear to be more restricted; it seems that no extremal configuration has both ascending and descending monotone subsequences of length $k+1$. These configurations are described in Section 3; it is shown that, if indeed no extremal configuration has both ascending and descending monotone subsequences of length $k+1$, the characterisation is complete and correct for $n \geq k(2k-1)$. (Computation suggests that—apart from the exceptional case of $k = 3$, $n = 16$, where there are also some extremal configurations not as described—all extremal configurations do indeed satisfy the given constraint.) The number of extremal configurations (under this assumption) may be described in terms of the Catalan numbers.

The problem may be seen to be equivalent to a problem on directed graphs as follows. Consider a permutation $p_0, p_1, \ldots, p_{n-1}$. Let A be a transitive tournament on n vertices, $v_0, v_1, \ldots, v_{n-1}$, with an edge $v_i \to v_j$ for all $i < j$. Let B be a transitive tournament on the same vertices, with an edge $v_i \to v_j$ if and only if $p_i < p_j$. Now a monotone ascending subsequence of length $k+1$ corresponds to a K_{k+1} subgraph on some subset of the same vertices, all of whose edges go in the same direction in both A and B; and a monotone descending subsequence of length $k+1$ corresponds to a K_{k+1} subgraph on some subset of the same vertices, all of whose edges go in opposite directions in A and B. Thus the problem is equivalent to: given two transitive tournaments on the same set of n vertices, what is the minimum number of K_{k+1} subgraphs on which the edge directions of the two tournaments entirely agree or entirely disagree? Furthermore, this formulation of the problem is symmetrical in A and B. In general, the problem has the following symmetries, which appear naturally in the formulation in terms of tournaments:

- The order of the permutation may be reversed (equivalent to reversing the order on A); the new permutation is $p_{n-1}, p_{n-2}, \ldots, p_0$.
- The value of p_i may be replaced by $n - 1 - p_i$ (equivalent to reversing the order on B).
- The permutation may be replaced by the permutation $q_0, q_1, \ldots, q_{n-1}$, where $q_{p_i} = i$ (equivalent to swapping A and B). This permutation is the inverse permutation to $p_1, p_2, \ldots, p_n$.

Combinations of these operations may also be applied; the symmetry group is that of the square, the dihedral group on 8 elements.

I would like to thank Andrew Thomason and an anonymous referee for their comments on earlier versions of this paper.

2 The case $k = 2$

We will see that, for $k = 2$, all permutations with a minimum number of monotone 3-sequences have the following form:

Theorem 4 *If $n = 1$, the extremal permutation is 0. If $n = 2$, the extremal permutations are $0, 1$ and $1, 0$. If $n > 2$, all extremal sequences have the form $L, 0, n-1, R$ or $L, n-1, 0, R$, where L and R have lengths $\lfloor n/2 \rfloor - 1$ or $\lceil n/2 \rceil - 1$ and L, R is an extremal permutation of $\{1, 2, \ldots, n-2\}$ (that is, the result of adding 1 to each element of an extremal permutation of $\{0, 1, \ldots, n-3\}$). All such permutations are extremal.*

Table 1: Extremal permutations for $n \leq 6$

<table>
<tr><td>$n = 1$</td><td colspan="4">0</td></tr>
<tr><td>$n = 2$</td><td colspan="2">0 1</td><td colspan="2">1 0</td></tr>
<tr><td>$n = 3$</td><td>0 2 1</td><td>1 0 2</td><td>1 2 0</td><td>2 0 1</td></tr>
<tr><td>$n = 4$</td><td>1 0 3 2</td><td>1 3 0 2</td><td>2 0 3 1</td><td>2 3 0 1</td></tr>
<tr><td>$n = 5$</td><td>1 0 4 3 2
1 3 0 4 2
1 3 4 0 2
1 4 0 3 2</td><td>2 0 4 1 3
2 0 4 3 1
2 1 0 4 3
2 1 4 0 3</td><td>2 3 0 4 1
2 3 4 0 1
2 4 0 1 3
2 4 0 3 1</td><td>3 0 4 1 2
3 1 0 4 2
3 1 4 0 2
3 4 0 1 2</td></tr>
<tr><td>$n = 6$</td><td>2 1 0 5 4 3
2 1 5 0 4 3</td><td>2 4 0 5 1 3
2 4 5 0 1 3</td><td>3 1 0 5 4 2
3 1 5 0 4 2</td><td>3 4 0 5 1 2
3 4 5 0 1 2</td></tr>
</table>

It is clear that this yields $2^{n/2}$ extremal permutations for n even and 2^{n-1} extremal permutations for n odd. For n even, there is a simple noninductive description: if the permutation is $p_0, p_1, \ldots, p_{n-1}$, then, for $0 \leq t < n/2$, we have that p_t and p_{n-1-t} take the values $(n/2)-1-t$ and $(n/2)+t$, in some order. Table 1 shows the extremal permutations for $n \leq 6$.

The sequences in Theorem 4 all have 0 and $n-1$ adjacent. It is easy to see that Theorem 4 is a correct characterisation of extremal sequences with that property.

Lemma 5 *Suppose $n > 2$ and that some extremal permutation has 0 and $n-1$ adjacent. Then all extremal permutations with 0 and $n-1$ adjacent are as described in Theorem 4, and all such permutations are extremal.*

Proof Without loss of generality, suppose a permutation with 0 and $n-1$ adjacent is $L, 0, n-1, R$; call this permutation S. Suppose that L has length ℓ and R has length r. All monotone subsequences of length 3 in L, R are also such subsequences of S. There are no monotone subsequences of S containing both 0 and $n-1$. There are no monotone subsequences of S of the form $a, 0, b$ or $a, n-1, b$, with $a \in L$ and $b \in R$. If, however, a precedes b in L, exactly one of $a, b, 0$ and $a, b, n-1$ is monotone; likewise, if a precedes b in R, exactly one of $0, a, b$ and $n-1, a, b$ is monotone. Thus $m_2(S) = m_2(L, R) + \binom{\ell}{2} + \binom{r}{2}$. This is minimal when $|\ell - r| \leq 1$. $\square$

Consider again the relation to tournaments described in Section 1. Suppose we colour an edge red if the two tournaments agree on the direction of that edge, or blue if the two tournaments disagree on the direction of that edge. The problem is then to minimise the number of monochromatic triangles. (However, we cannot use any 2-colouring of K_n, only one arising from two tournaments in this manner.) Goodman [3] and Lorden [4] found that the number of monochromatic triangles depends only on the sequence of red (or blue) degrees:

Theorem 6 (Goodman [3] and Lorden [4]) *Let K_n be coloured in red and blue. Let $d_r(v)$ be the number of red edges from the vertex v. Then there are exactly*

$$\binom{n}{3} - \frac{1}{2}\sum_v d_r(v)(n-1-d_r(v))$$

monochromatic triangles.

This theorem allows us to prove correct our characterisation of extremal configurations.

Proof of Theorem 4 for n even The canonical extremum from Section 1 is of this form, and has $M_2(n) = 2\binom{n/2}{3}$ monotone subsequences of length 3. In the coloured graph corresponding to this permutation, each vertex has red degree equal to either $\lceil (n-1)/2 \rceil$ or $\lfloor (n-1)/2 \rfloor$, so the graph minimises the number of monochromatic triangles. Thus all the permutations for n even described in Theorem 4 are indeed extremal. Also, in the coloured graph corresponding to an extremal permutation $p_0, p_1, \ldots, p_{n-1}$, all vertices must have red degree either $\lceil (n-1)/2 \rceil$ or $\lfloor (n-1)/2 \rfloor$; in particular, the vertices corresponding to the values 0 and $n-1$ must have such red degrees. This means that 0 and $n-1$ must each be the value of one of $p_{(n/2)-1}$ and $p_{n/2}$, so they are adjacent, and the result follows by Lemma 5. □

This method does not apply quite so simply for n odd, where the graphs corresponding to extremal permutations do not minimise the number of monochromatic triangles over all colourings (that is, the colourings minimising the number of monochromatic triangles do not correspond to pairs of transitive tournaments). However, the colourings are sufficiently close to extremal that with a little more effort the method can be adapted.

Proof of Theorem 4 for n odd The canonical extremum from Section 1 is of this form, so $M_2(n)$ monotone subsequences of length 3 can be attained. We will show that this is indeed extremal, and that in all extremal permutations 0 and $n-1$ are adjacent, so that the result will then follow by Lemma 5.

Suppose we have some extremal permutation $p_1, p_2, \ldots, p_n$, and let $\ell(v)$ be the location of the value v; that is, $p_{\ell(v)} = v$. Let the vertex corresponding to the position $\ell(v)$ with value v also be known as v. Let $d_r(v)$ and $d_b(v)$ be the numbers of red and blue edges, respectively, from the vertex v; put $d_d(v) = \frac{1}{2}|d_r(v) - d_b(v)|$. Observe that $d_r(v)(n-1-d_r(v)) = d_r(v)d_b(v) = (\frac{n-1}{2})^2 - d_d(v)^2$, so, by Theorem 6, the number of monochromatic triangles then is

$$\binom{n}{3} - \frac{n(n-1)^2}{8} + \sum_v d_d(v)^2.$$

Thus, we wish to minimise $\sum_v d_d(v)^2$. In the canonical extremum this takes the value $\frac{n-1}{2}$.

Suppose $0 \le v \le (n-1)/2$. Let $L = \{\, u : \ell(u) < \ell(v) \,\}$ be the set of values to the left of v, and $R = \{\, u : \ell(u) > \ell(v) \,\}$ be the set of values to the right of v. Put further $L_r = \{ u \in L : u < v \}$, $L_b = \{ u \in L : u > v \}$, $R_r = \{ u \in R : u > v \}$ and $R_b = \{ u \in R : u < v \}$. Then we have $d_r(v) = |L_r| + |R_r|$ and $d_b(v) = |L_b| + |R_b|$, so

$$d_r(v) - d_b(v) = |R_r| - |R_b| - |L_b| + |L_r| = (|R| - |L|) + 2(|L_r| - |R_b|).$$

Now

$$|R| - |L| = (n - 1 - \ell(v)) - \ell(v) = 2\left(\tfrac{n-1}{2} - \ell(v)\right),$$

and

$$\Big||L_r| - |R_b|\Big| \leq |L_r \cup R_b| = v,$$

so $d_d(v) \geq \max\{0, |\frac{n-1}{2} - \ell(v)| - v\}$. Likewise, for $(n-1)/2 \leq v \leq n-1$, we have $d_d(v) \geq \max\{0, |\frac{n-1}{2} - \ell(v)| - (n-1-v)\}$. Define $r(j)$ by $r(j) = j$ for $0 \leq j \leq (n-1)/2$ and $r(j) = n-1-j$ for $(n-1)/2 \leq j \leq n-1$, so we have

$$d_d(v) \geq \max\left\{0, \left|\frac{n-1}{2} - \ell(v)\right| - r(v)\right\}.$$

For $0 \leq j \leq (n-1)$, put $S(j) = \{ i : |\frac{n-1}{2} - i| \leq r(j) \}$. That is, $S(j)$ is the set of possible value of $\ell(j)$ for which our lower bound on $d_d(j)$ would be 0. We then have

$$d_d(v) \geq |\{ (n-1)/2 \geq j \geq r(v) : \ell(v) \notin S(j) \}| = \sum_{\substack{(n-1)/2 \geq j \geq r(v) \\ \ell(v) \notin S(j)}} 1.$$

Adding over all v and reversing the order of summation then gives

$$\sum_v d_d(v) \geq \sum_{0 \leq j \leq (n-1)/2} |\{ v : r(v) \leq j,\, \ell(v) \notin S(j) \}|.$$

For $0 \leq j < (n-1)/2$, observe that $|S(j)| = 2j+1$, whereas $|\{ v : r(v) \leq j \}| = 2j+2$. Thus $\sum_v d_d(v) \geq \frac{n-1}{2}$, and equality requires that each $|\{ v : r(v) \leq j, \ell(v) \notin S(j) \}|$ equals 1, for $0 \leq j < \frac{n-1}{2}$. Now $\sum_v d_d(v)^2 \geq \sum_v d_d(v)$, with equality only if all terms are 0 or 1. So any extremum must have $\ell(0)$ and $\ell(n-1)$ both equal to $\frac{n-1}{2}$ or $\frac{n-1}{2} \pm 1$, with one of them equal to $\frac{n-1}{2}$. So 0 and $n-1$ are adjacent. $\square$

3 The case $k > 2$

For $k > 2$, it seems that, for n sufficiently large, the permutations with a minimum number of monotone $(k+1)$-subsequences have only descending, or only ascending, monotone subsequences of that length; making this assumption, we can give a characterisation of the extremal permutations for $n \geq k(2k-1)$ (which appears to be sufficiently large, except for $k = 3$, $n = 16$, where there are also some other extremal permutations). It is easy to see that this condition is equivalent to the permutation being divisible into (at most) k disjoint monotone descending subsequences, or k disjoint monotone ascending subsequences. If it can be divided into k disjoint monotone descending subsequences, there cannot be a monotone ascending $(k+1)$-subsequence, since such a sequence would have to contain two elements from one of the k descending subsequences. Conversely, if it contains only descending subsequences of length $k+1$, it can be divided into k descending subsequences explicitly; similarly to one proof of Theorem 1, form these subsequences by adding each element in turn to the first of the subsequences already present it can be added to without making that subsequence nondescending, or start a new subsequence if the element is greater than the last element of all existing subsequences. Any element added is at the end of an ascending subsequence, containing one element from each sequence up to the one to which the element was added, so having $k+1$ subsequences would imply the presence of a monotone ascending subsequence of length $k+1$, a contradiction.

The form of the extremal permutations (subject to the supposition described) is somewhat more complicated than that for $k=2$. We describe the form where all the monotone $(k+1)$-subsequences are descending; the sequences for which they are all ascending are just the reverse of those we describe. If the k subsequences are of lengths ℓ_1, ℓ_2, $\ldots$, ℓ_k (where some of the ℓ_i may be 0 if there are less than k subsequences), there are at least

$$\sum_{i=1}^{k}\binom{\ell_i}{k+1}$$

monotone subsequences of length $k+1$. For this to be minimal, convexity implies that $\lfloor n/k\rfloor \le \ell_i \le \lceil n/k\rceil$ for all i; in particular, there are k subsequences, and no ℓ_i is 0, for $n \ge k$. To make the ordering of the ℓ_i definite, order the k subsequences by the position of their middle element (the leftmost of two middle elements, if the sequence is of even length). There are $\binom{k}{n \bmod k}$ choices of the ℓ_i satisfying these inequalities. If they are satisfied, there are at least $M_k(n)$ monotone $(k+1)$-subsequences, and exactly that number if and only if there is no monotone descending $(k+1)$-subsequence that takes values from more than one of the k subsequences. Put $s_i = \sum_{1\le j\le i}\ell_i$. For each choice of the ℓ_i, we have a canonical extremum similar to that given in Section 1:

$$\begin{aligned}
&s_1-1, s_1-2, \ldots, 0,\\
&s_2-1, s_2-2, \ldots, s_1,\\
&\ldots,\\
&s_k-1, s_k-2, \ldots, s_{k-1}.
\end{aligned}$$

(where $0 = s_0$ and $s_k = n$).

We will describe the extrema with the given ℓ_i, supposing $n \ge k(2k-1)$. To do so we will need some more notation. Write $C_k = \frac{1}{k+1}\binom{2k}{k}$ for the k^{th} *Catalan number*. It will then turn out that there are exactly C_k^{2k-2} extrema with the given ℓ_i. Thus, the total number of extremal sequences, subject to the constraint that all monotone $(k+1)$-subsequences go in the same direction, and subject to $n \ge k(2k-1)$, will be

$$2\binom{k}{n \bmod k}C_k^{2k-2}.$$

The extrema are closely related to the canonical extremum with the given ℓ_i. In each extremum with those ℓ_i, the $\ell_i - (2k-2)$ middle values of each of the k monotone subsequences take the same values, in the same positions, as they do in the canonical extremum; the $k-1$ values at either end of each subsequence can vary, as can their positions.

The variation is described in terms of sets $C(k,p)$ of monotone descending sequences of $k-1$ integers; $|C(k,p)| = C_k$. This set is defined as follows: $C(k,p)$ is the set of monotone descending sequences $c_1, c_2, \ldots, c_{k-1}$ of integers, $p-2k+3 \le c_i \le p$ for all i, such that if $d_1, d_2, \ldots, d_{k-1}$ is the monotone descending sequence of all integers in $[p-2k+3, p]$ that are not one of the c_i, then $c_1, c_2, \ldots, c_{k-1}, d_1, d_2, \ldots, d_{k-1}$ has no monotone descending subsequence of length $k+1$.

There are various equivalent characterisations of $C(k,p)$:

Lemma 7 *Define $C_1(k,p)$ to be the set of monotone descending sequences $c_1, c_2, \ldots, c_{k-1}$ of integers, such that $p-k-i+2 \le c_i \le p-2i+2$ for all $1 \le i \le k-1$. Define $C_2(k,p)$ inductively as follows. Let $C_2(2,p) = \{p-1, p\}$. For $k > 2$, let $C_2(k,p) = \{\,(c_1, c_2, \ldots, c_{k-1}) : p-k+1 \le c_1 \le p,\ c_2 < c_1,\ (c_2, c_3, \ldots, c_{k-1}) \in C_2(k-1, p-2)\,\}$. Then $C_1(k,p) = C_2(k,p) = C(k,p)$. Furthermore, $|C(k,p)| = C_k$.*

Proof Of these definitions, C is the one that will be relevant later in proving the characterisation of extremal permutations correct. C_1 will be seen to be a direct description of C, and C_2 will be seen to be an inductive description of C_1. C_2 allows the number of such sequences to be calculated through recurrence relations, which will yield the last part of the lemma. Observe that all these definitions clearly have the property that $C(k,p_1)$ is related to $C(k,p_2)$ simply by adding $p_1 - p_2$ to all elements of all sequences in $C(k,p_2)$.

We first show that $C_1(k,p) = C(k,p)$. First consider a sequence $c_1, c_2, \ldots, c_{k-1}$ in $C_1(k,p)$, letting $d_1, d_2, \ldots, d_{k-1}$ be the monotone descending sequence of all integers in $[p-2k+3, p]$ that are not one of the c_i. If the sequence $c_1, c_2, \ldots, c_{k-1}, d_1, d_2, \ldots, d_{k-1}$ has a monotone descending subsequence of length $k+1$, suppose that subsequence has t values among the c_i. The last of these is at most $p-2t+2$. The interval $[p-2k+3, p]$ contains $2k-2t-1$ values smaller than $p-2t+2$; of these, at least $k-1-t$ must be among the c_i (namely, $c_{t+1}, c_{t+2}, \ldots, c_{k-1}$), so at most $k-t$ are among the d_i, so the monotone subsequence has length at most k, a contradiction. Thus $C_1(k,p) \subset C(k,p)$. Conversely, consider a sequence $c_1, c_2, \ldots, c_{k-1}$ in $C(k,p)$, and let d_i be as above. Clearly $c_i \ge p-k-i+2$ for all i; otherwise we would have $c_{k-1} < p-2k+3$. If we had $c_i > p-2i+2$, then there would be at least $2k-2i$ lesser values in the interval $[p-2k+3, p]$, of which $k-1-i$ are among the c_j, so at least $k-i+1$ are among the d_j; together with $c_1, c_2, \ldots, c_i$, this yields a monotone subsequence of length at least $k+1$, a contradiction. Thus $C(k,p) \subset C_1(k,p)$.

We now show that $C_1(k,p) = C_2(k,p)$. We do this by induction on k; it clearly holds for $k=2$ and all p. Suppose that $C_1(k-1,q) = C_2(k-1,q)$ for all q. If $c_1, c_2, \ldots, c_{k-1}$ is in $C_2(k,p)$, then $p-k+1 \le c_1 \le p$, and, since $c_1 > c_2$ and $c_2, c_3, \ldots, c_{k-1}$ is in $C_2(k-1, p-2) = C_1(k-1, p-2)$, the sequence of the c_i is descending and

$(p-2)-(k-1)-(i-1)+2 = p-k-i+2 \le c_i \le (p-2)-2(i-1)+2 = p-2i+2$ for all $2 \le i \le k-1$, so the sequence is in $C_1(k,p)$. Conversely, if $c_1, c_2, \ldots, c_{k-1}$ is in $C_1(k,p)$, then for $2 \le i \le k-1$ we have $p-k-i+2 = (p-2)-(k-1)-(i-1)+2 \le c_i \le p-2i+2 = (p-2)-2(i-1)+2$, so that $c_2, c_3, \ldots, c_{k-1}$ is in $C_1(k-1,p-2) = C_2(k-1,p-2)$, so the sequence is in $C_2(k,p)$.

Finally we show that $|C_2(k,p)| = C_k$. For $1 \le j \le k$, put

$$c_{k,j} = |\{ (c_1, c_2, \ldots, c_{k-1}) \in C_2(k,p) : c_1 = p-k+j \}|$$

(which as observed above does not depend on p). We then have

$$|C_2(k,p)| = \sum_{j=1}^{k} c_{k,j}$$

and the recurrence

$$c_{k,j} = \sum_{i=1}^{\min\{j,k-1\}} c_{k-1,i},$$

where $c_{2,1} = c_{2,2} = 1$. Observe that the recurrence implies that $c_{k,k-1} = c_{k,k} = |C_2(k-1,p)|$.

Put

$$d_{k,j} = \binom{k+j-3}{j-1} - \sum_{i=0}^{j-3} \binom{k+i-1}{i},$$

with $d_{k,1} = 1$. We claim that $c_{k,j} = d_{k,j}$ for all $k \ge j$; we prove this by induction on j. Clearly $c_{k,1} = 1$ and $c_{k,2} = k-1$. Suppose that $j > 2$ and $c_{k,j-1} = d_{k,j-1}$ for all k. For $k \ge j$ we then have $c_{k+1,j} - c_{k,j} = c_{k+1,j-1} = d_{k+1,j-1}$ and

$$d_{k+1,j} - d_{k,j} = \binom{k+j-3}{j-2} - \sum_{i=1}^{j-3} \binom{k+i-1}{i-1} = d_{k+1,j-1}.$$

Also, $d_{j,j} - c_{j,j} = d_{j,j} - c_{j,j-1} = d_{j,j} - d_{j,j-1} = \binom{2j-3}{j-1} - \binom{2j-4}{j-2} - \binom{2j-4}{j-3} = \binom{2j-3}{j-1} - \binom{2j-4}{j-2} - \binom{2j-4}{j-1} = 0$. Thus, by induction on k, $c_{k,j} = d_{k,j}$ for the given j and all k, and by induction on j this holds for all j.

It now remains only to show that $d_{k,k-1} = C_{k-1}$ for all k. For this, observe that $C_{k-1}/\binom{2k-4}{k-2} = \binom{2k-2}{k-1}/k\binom{2k-4}{k-2} = 2(2k-3)/k(k-1)$. We have

$$d_{k,k-1} = \binom{2k-4}{k-2} - \sum_{i=0}^{k-4} \binom{k+i-1}{i}$$

and

$$\sum_{i=0}^{k-4} \binom{k+i-1}{i} = \binom{2k-4}{k-4}$$

so that $d_{k,k-1}/\binom{2k-4}{k-2} = 1 - \binom{2k-4}{k-4}/\binom{2k-4}{k-2} = 1-(k-2)(k-3)/k(k-1) = 2(2k-3)/k(k-1) = C_{k-1}/\binom{2k-4}{k-2}$. Thus $d_{k,k-1} = C_{k-1}$. □

Table 2: Structure of an example extremal permutation

n	17
k	3
Extremum	5 4 2 12 1 0 9 8 7 16 6 3 15 14 13 11 10
ℓ_1, ℓ_2, ℓ_3	5, 6, 6
s_0, s_1, s_2, s_3	0, 5, 11, 17
Canonical extremum	4 3 2 1 0 10 9 8 7 6 5 16 15 14 13 12 11
Fixed and variable values	X X 2 X X X X 8 7 X X X X 14 13 X X
S_0, S_1, S_2, S_3	$\{0,1\}$, $\{3,4,5,6\}$, $\{9,10,11,12\}$, $\{15,16\}$
S	$\{0,1,3,4,5,6,9,10,11,12,15,16\}$
A_1, A_2	$\{5,4\}$, $\{12,9\}$
B_1, B_2	$\{5,4\}$, $\{11,10\}$
A'_1, A'_2	$\{6,3\}$, $\{11,10\}$
B'_1, B'_2	$\{6,3\}$, $\{12,9\}$
L_1, L_2, L_3	$\{1,0\}$, $\{6,3\}$, $\{11,10\}$
R_0, R_1, R_2	$\{5,4\}$, $\{12,9\}$, $\{16,15\}$
T_1, T_2, T_3	5 4 2 1 0, 12 9 8 7 6 3, 16 15 14 13 11 10

We now describe the conjectured extrema with given ℓ_i. We define sets S_j of integers: put $S_0 = \{\, i : 0 \le i \le k-2 \,\}$; put $S_k = \{\, i : n-k+1 \le i \le n-1 \,\}$; and for $1 \le j \le k-1$, put $S_j = \{\, i : s_j - k + 1 \le i \le s_j + k - 2 \,\}$. Put $S = \cup_{j=0}^{k} S_j$. Then S is the union of the sets of the $k-1$ values (or positions) at either end of each of the subsequences in the canonical extremum.

Write the canonical extremum as d_0, d_1, …, d_{n-1}. We describe an extremum c_0, c_1, …, c_{n-1}. For $i \notin S$, we have $c_i = d_i$; observe (as would be expected, given the symmetries of the problem) that $[0, n-1] \setminus S = \{\, d_i : i \notin S \,\}$.

For $i \le i \le k-1$, let A_i and B_i be arbitrary elements of $C(k, s_i + k - 2)$; let A'_i be $S_i \setminus A_i$ in descending order, and let B'_i be $S_i \setminus B_i$ in descending order. Given this choice of A_i and B_i (there being C_k^{2k-2} possible such choices), we can now describe the extremum associated with the A_i and B_i.

We will define sets L_i for $1 \le i \le k$ and R_i for $0 \le i \le k-1$. Put $L_1 = S_0$ and $R_{k-1} = S_k$. For $1 \le i \le k-1$, put $R_{i-1} = A_i$ and $L_{i+1} = A'_i$. Now, the values of c_i for $i \in S_0$ are the values of R_0 in descending order; the values of c_i for $i \in S_k$ are the values of L_k in descending order; the values of c_i for $i \in B_j$ are the values of L_j in descending order; and the values of c_i for $i \in B'_j$ are the values of R_j in descending order. Observe that this sequence can be divided into k disjoint monotone descending subsequences, of the required lengths; the i^{th} of them, for $1 \le i \le k$, contains R_{i-1}, the fixed values c_j for $s_{i-1} + k - 1 \le j \le s_i - k$, and L_i. Call this subsequence T_i.

An example extremum with $n = 17$ and $k = 3$ is shown in Table 2, along with the various parameters for its structure described above, and illustrated in Figure 2.

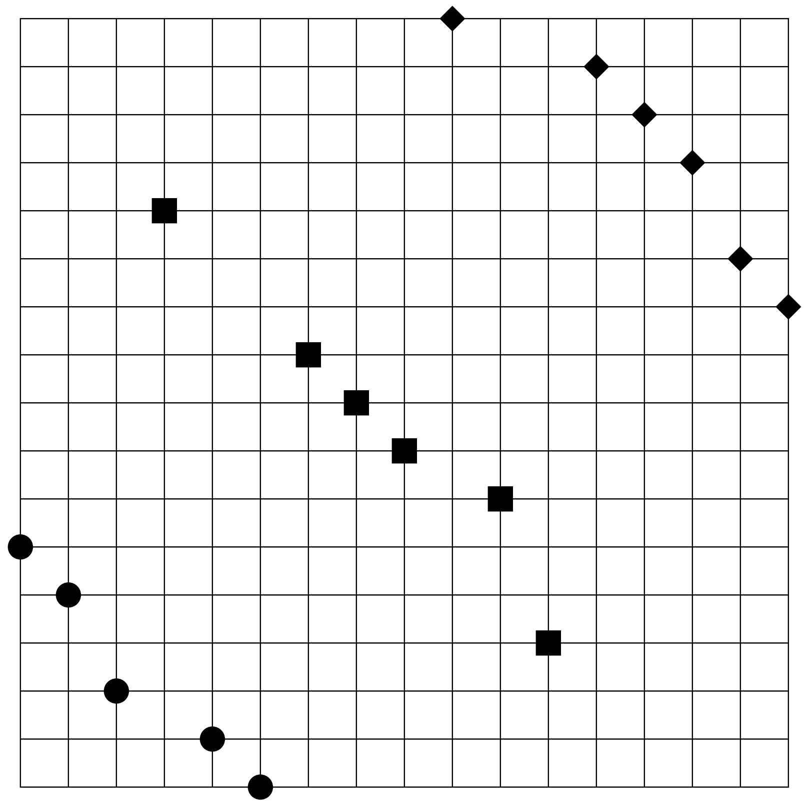

Figure 2: Illustration of an example extremal permutation

It remains to prove that this sequence has the expected number of monotone subsequences of length $k+1$, and that all extrema (subject to the sequence being divisible into k disjoint monotone descending subsequences) have that form. The description of the sequence makes sense for $n \geq k(2k-2)$, and Theorem 8 applies for all such n, but if $n < k(2k-1)$ there can be other extrema not of the form described.

Theorem 8 *The sequences just described have exactly $M_k(n)$ monotone subsequences of length $k+1$, all of them descending.*

Proof By the division into k disjoint monotone descending subsequences, of lengths ℓ_i, there are no monotone ascending subsequences of length $k+1$, and there are at least $M_k(n)$ monotone descending subsequences of length $k+1$ (that is, those subsequences entirely within any one of the k subsequences into which the sequence is divided). Thus it is only necessary to prove that there is no monotone descending subsequence of length $k+1$ containing values from more than one of the k subsequences.

If $j \geq i+2$, then the whole of T_j is to the right of the whole of T_i, and all the values in T_j are greater than all the values in T_i. Thus any additional monotone subsequence of length $k+1$ can contain values from only two of the T_j, say T_i and T_{i+1}. If it contains c_p from T_i and c_q from T_{i+1}, we still have $p < q$ except possibly for c_p from L_i and

c_q from R_i, and $c_p < c_q$ except possibly for c_p from R_{i-1} and c_q from L_{i+1}. Thus this sequence contains no values from the fixed central regions of T_i and T_{i+1}; if it contains a value from R_{i-1}, then it contains a value from L_{i+1}, and all values are from R_{i-1} and L_{i+1}; if it contains a value from L_i, then all values are from L_i and R_i. But a monotone descending subsequence of length $k+1$ in R_{i-1} followed by L_{i+1} would be such a subsequence in A_i followed by A_i', contradicting the definition of $C(k,p)$. Likewise, a monotone descending sequence (of values, as the position goes up) in L_i and R_i may be seen to be equivalent to a monotone descending sequence of positions, as the value goes up, in the positions (going down) of L_i followed by those of R_i; that is, in B_i followed by B_i', again a contradiction. Thus there are no such monotone subsequences. $\square$

Theorem 9 *For $n \geq k(2k-1)$, the sequences which contain no monotone ascending $(k+1)$-subsequences and a minimum number of monotone descending $(k+1)$-subsequences are exactly the $\binom{k}{n \bmod k} C_k^{2k-2}$ sequences described above. The sequences which contain no monotone descending $(k+1)$-subsequences and a minimum number of monotone ascending $(k+1)$-subsequences are those sequences, reversed.*

Proof The derivation of extremal sequences with only ascending $(k+1)$-subsequences from those with only descending $(k+1)$-subsequences is clear. As observed above, sequences with only descending $(k+1)$-subsequences are just those divisible into at most k disjoint monotone descending subsequences, and minimality requires that there be exactly k such subsequences, and that their lengths by $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. Thus the sequences described above are extremal (from Theorem 8), and it is only necessary to show that there are no more extremal sequences.

Suppose $c_0, c_1, \ldots, c_{n-1}$ is an extremal sequence. Suppose that one of the k monotone descending subsequences into which it is divided occupies positions $a_0 < a_1 < \cdots < a_{\ell_i-1}$ (so has values $c_{a_0} > c_{a_1} > \cdots > c_{a_{\ell_i-1}}$), and another occupies positions $b_0 < b_1 < \cdots < b_{\ell_j-1}$, where $a_0 < b_0$. Then $c_{a_0} < c_{b_0}$ (since otherwise $c_{a_0}, c_{b_0}, c_{b_1}, \ldots, c_{b_{k-1}}$ would be another monotone descending $(k+1)$-subsequence), so $c_{a_m} \leq c_{a_0} < c_{b_0}$ for all m. Thus $b_0 > a_{\ell_i-k}$, since otherwise $c_{b_0}, c_{a_{\ell_i-k}}, c_{a_{\ell_i-k+1}}, \ldots, c_{a_{\ell_i-1}}$ would be a monotone descending $(k+1)$-subsequence; and $a_{\ell_i-1} > b_{k-1}$, since otherwise either $c_{b_0}, c_{b_1}, \ldots, c_{b_{k-1}}, c_{a_{\ell_i-1}}$ or $c_{a_0}, c_{a_1}, \ldots, c_{a_{k-1}}, c_{b_{k-1}}$ would be a monotone descending $(k+1)$-subsequence (depending on the order of $c_{a_{\ell_i-1}}$ and $c_{b_{k-1}}$).

Thus, if we order our k subsequences by the position of the first element, we have seen that the only possible overlap in positions is between the last $k-1$ of one sequence and the first $k-1$ of a later sequence. Because $n \geq k(2k-1)$, each sequence has $\ell_i - 2(k-1) > 0$ central elements that are not in the first or last $k-1$; so the ordering by where the first elements are is the same as the ordering by where the central elements are (which was chosen previously as the ordering of the ℓ_i). In particular, we see that the only overlap in positions is between the last $k-1$ of one sequence and the first $k-1$ of the very next sequence in this order.

Likewise, we may consider the possible overlap in values. If as above we have $i < j$, a_p the positions of sequence i and b_q the positions of sequence j, then suppose for some p, q

we have $c_{a_p} > c_{b_q}$. If $p \geq k-1$, then $c_{a_0}, c_{a_1}, \ldots, c_{a_{k-1}}, c_{b_q}$ would be monotone descending; if $q \leq \ell_j - k$, then $c_{a_p}, c_{b_{\ell_j-k}}, c_{b_{\ell_j-k+1}}, \ldots, c_{b_{\ell_j-1}}$ would be monotone descending. Thus the only possible overlap in values is between the first $k-1$ of one sequence and the last $k-1$ of a later sequence, which again must be the very next sequence.

Given these restrictions on overlap of positions, the i^{th} sequence must include the positions from $s_{i-1}+k-1$ to $s_i - k$ (with $k-1$ positions to either side). The restrictions on overlap of values imply that in these central $\ell_i - 2(k-1)$ positions there must be the canonical values d_i. Thus all extrema have those fixed values that were fixed in our description of the extrema.

For $1 \leq i \leq k$, let R_{i-1} be the set of the first $k-1$ values in the i^{th} sequence, and let L_i be the set of the last $k-1$ values. Then the i^{th} sequence contains the values R_{i-1}, the fixed values c_j for $s_{i-1}+k-1 \leq j \leq s_i - k$, and L_i, as in the above description of extrema. Further, the restriction on the overlap of values implies that $L_1 = S_0$ and $R_{k-1} = S_k$, and that, for $1 \leq i \leq k-1$, R_{i-1} and L_{i+1} are disjoint subsets of $[s_i - k + 1, s_i + k - 2]$. Put $A_i = R_{i-1}$ and $A'_i = L_{i+1}$. Similarly, the positions in our sequence of the values in L_i and R_i are disjoint subsets of $[s_i - k + 1, s_i + k - 2]$; let B_i be the set of positions of the values in L_i, and let B'_i be the set of positions of the values in R_i.

If A_i and B_i are indeed elements of $C(k, s_i + k - 2)$, then the sequence is of the given form, with those A_i and B_i. However, if A_i is not an element of $C(k, s_i + k - 2)$, then the sequence of the values of $A_i = R_{i-1}$ in descending order, followed by those of $A'_i = L_{i+1}$ in descending order, has a monotone descending subsequence of length $k+1$, which is such a subsequence in our original sequence, contradicting minimality. Likewise, if B_i is not an element of $C(k, s_i + k - 2)$, then the sequence of the values of B_i in descending order (the positions of L_i, in ascending order of value), followed by those of B'_i in descending order (the positions of R_i, in ascending order of value), has a monotone descending $(k+1)$-subsequence; that is, there is a monotone descending $(k+1)$-sequence of positions, the values in which are increasing, which gives a monotone descending sequence of values in the original sequence. $\square$

If $n < k(2k-1)$, the above proof no longer works, since some of the k subsequences have no fixed middle elements. However, for $k(2k-2) \leq n < k(2k-1)$, the construction still gives sequences with $M_k(n)$ monotone $(k+1)$-subsequences—but there can be other extrema (in which all monotone $(k+1)$-subsequences go in the same direction) as well.

Computation shows that, for some n and k, such other extrema do indeed exist. In particular, this applies for $k = 3$ and $12 \leq n < 15$: for each such n there are extrema, in which all monotone $(k+1)$-subsequences go in the same direction, that are not of the form described above. Further, if we remove the constraint that all monotone $(k+1)$-subsequences go in the same direction, the extremal function is as conjectured for $k = 3$ and $n \leq 18$, and for $k = 4$ and $n \leq 19$ (that is, there are no sequences with fewer than $M_k(n)$ monotone $(k+1)$-subsequences). For $k = 3$ and $15 \leq n \leq 18$, the extrema described above are found, but when $n = 16$ there are some additional extrema which contain both ascending and descending monotone $(k+1)$-subsequences. (The first such extremum lexicographically is '4 3 9 2 1 0 13 8 7 6 5 15 14 12 11 10'.) Table 3 shows the

Table 3: Number of extremal permutations for $3 \leq k \leq 4$

	$k=3$		$k=4$	
n	Total	Both	Total	Both
1	1	0	1	0
2	2	0	2	0
3	6	0	6	0
4	22	0	24	0
5	86	0	118	0
6	306	0	668	0
7	882	0	4124	0
8	1764	0	26328	0
9	1764	0	165636	0
10	8738	0	985032	0
11	6892	0	5323032	0
12	1682	0	25038288	0
13	14706	10092	97173648	0
14	4182	0	288576288	0
15	1250	0	577152576	0
16	6250	2500	577152576	0
17	3750	0	2855608848	0
18	1250	0	2330017568	0
19			710429200	0

number of extrema found in each case, in the columns headed 'Total', and the number of those which contain both ascending and descending monotone $(k+1)$-subsequences, in the columns headed 'Both'. The source code of the program that did the computations for Table 3 is in the C source file distributed with this paper.

For larger n exhaustive search could not be done, but heuristic computation, taking a random permutation and attempting to move from that to an apparent extremum, did not find any other cases of apparent extrema (i.e., permutations with $M_k(n)$ monotone subsequences of length $k+1$) not matching the form described above, nor any sequences with fewer than $M_k(n)$ monotone $(k+1)$-subsequences, for $n \geq k(2k-1)$.

The method for the heuristic computation started with a random permutation. Various operations were then applied to it: transposing a pair of values in the permutation; reversing the order of a block of values in the permutation; rotating a block of values (in consecutive positions) in the permutation left or right; and the dual operation of rotating a block of positions (of consecutive values). All possible operations that reduced the number of monotone $(k+1)$-subsequences were considered, if there were any; if there were none, operations that kept the number of monotone $(k+1)$-subsequences the same were considered; in that case, a completely random move was occasionally chosen instead (to try to avoid the problem of being stuck at a local minimum that was not a global minimum). This process was stopped when the permutation had no more than $M_k(n)$ monotone $(k+1)$-subsequences. In computations for various n and k with $n \geq k(2k-1)$, no cases were found with fewer than $M_k(n)$ monotone $(k+1)$-subsequences, and the only extrema found in which not all monotone $(k+1)$-subsequences went in the same direction were with $k=3$ and $n=16$. These computations were done for $k=3$ and $15 \leq n \leq 30$, and for $k=4$ and $28 \leq n \leq 40$.

References

[1] P. Erdős, *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), no. 3, 459–464.

[2] Pál Erdős and George Szekeres, *A combinatorial problem in geometry*, Compositio Math. **2** (1935), 463–470.

[3] A. W. Goodman, *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), no. 9, 778–783.

[4] Gary Lorden, *Blue-empty chromatic graphs*, Amer. Math. Monthly **69** (1962), no. 2, 114–120.

[5] J. Michael Steele, *Variations on the monotone subsequence theme of Erdős and Szekeres*, Discrete probability and algorithms (Minneapolis, MN, 1993) (David Aldous, Persi Diaconis, Joel Spencer, and J. Michael Steele, eds.), IMA Vol. Math. Appl. **72**, Springer-Verlag, 1995, pp. 111–131.

[6] Andrew Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, J. London Math. Soc. (2) **39** (1989), no. 2, 246–255.