Journal of Integer Sequences, Vol. 16 (2013),
Article 13.6.1

Permutations with Given Peak Set

Sara Billey[1] and Krzysztof Burdzy[2]
Department of Mathematics
University of Washington
Seattle, WA 98195-4350
USA
billey@math.washington.edu
burdzy@math.washington.edu

Bruce E. Sagan
Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027
USA
sagan@math.msu.edu

Abstract

Let $\mathfrak{S}_n$ denote the symmetric group of all permutations $\pi = a_1 \cdots a_n$ of $\{1, \ldots, n\}$. An index i is a *peak* of π if $a_{i-1} < a_i > a_{i+1}$ and we let $P(\pi)$ be the set of peaks of π. Given any set S of positive integers we define $\mathcal{P}(S; n) = \{\pi \in \mathfrak{S}_n \ : \ P(\pi) = S\}$. Our main result is that for all fixed subsets of positive integers S and all sufficiently large n we have $\#\mathcal{P}(S; n) = p(n)2^{n-\#S-1}$ for some polynomial $p(n)$ depending on S. We explicitly compute $p(n)$ for various S of probabilistic interest, including certain cases where S depends on n. We also discuss two conjectures, one about positivity of the coefficients of the expansion of $p(n)$ in a binomial coefficient basis, and the other about sets S maximizing $\#\mathcal{P}(S; n)$ when $\#S$ is fixed.

[1]Partially supported by grant DMS-1101017 from the NSF.
[2]Partially supported by grant DMS-1206276 from the NSF and by grant N N201 397137 from the MNiSW, Poland.

1 Introduction

Let $\mathbb{N}$ be the nonnegative integers and, for $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$. Consider the *symmetric group* $\mathfrak{S}_n$ of all permutations $\pi = a_1 \cdots a_n$ of $[n]$. Call an index i a *peak* of π if $a_{i-1} < a_i > a_{i+1}$ and define the *peak set* of π to be

$$P(\pi) = \{i \ : \ i \text{ is a peak of } \pi\}.$$

By way of illustration, if $\pi = a_1 \cdots a_7 = 1453276$ then $P(\pi) = \{3, 6\}$ because of $a_3 = 5$ and $a_6 = 7$. We will also apply the definitions and notation just given to any π which is a sequence of positive integers. Note that some authors call a_i a peak rather than i, but our convention is more consistent with what is used for other permutation statistics. Also note that if $\pi \in \mathfrak{S}_n$ then $P(\pi) \subseteq \{2, \ldots, n-1\}$. There has been a great deal of research into peaks of permutations, see [1, 2, 4, 6, 9, 10, 11, 12, 13, 15, 16, 17].

The purpose of the present work is to investigate permutations with a given peak set. To this end, define

$$\mathcal{P}(S; n) = \{\pi \in \mathfrak{S}_n \ : \ P(\pi) = S\}.$$

We will often omit the curly brackets around S in this notation. So, for example,

$$\mathcal{P}(2; 4) = \{1324, 1423, 1432, 2314, 2413, 2431, 3412, 3421\}.$$

Our main result will be about the cardinality $\#\mathcal{P}(S; n)$ as n varies, where S is a set of constants not depending on n. To state it, define a set $S = \{i_1 < \cdots < i_s\}$ to be *n-admissible* if $\#\mathcal{P}(S; n) \neq 0$. Note that we insist the elements be listed in increasing order. It is easy to see that S is n-admissible if and only if $i_1 > 1$, no two i_j are consecutive integers, and $n > i_s$. If we make a statement about an *admissible* set S, we mean that S is n-admissible for some n and the statement holds for every n such that S is n-admissible. We can now state our principal theorem.

Theorem 1. *If $S = \{i_1 < \cdots < i_s\}$ is admissible then*

$$\#\mathcal{P}(S; n) = p(n) 2^{n - \#S - 1}$$

where $p(n) = p(S; n)$ is a polynomial depending on S such that $p(n)$ is an integer for all integral n. In addition, $\deg p(n) = i_s - 1$ (when $S = \emptyset$ we have $\deg p(n) = 0$).

If S is not admissible, then $\#\mathcal{P}(S; n) = 0$ for all positive integers n, so we define the corresponding polynomial $p(S; n) = 0$. Thus, for all sets S of constants not depending on n, $p(S; n)$ is a well defined polynomial which we will call the *peak polynomial* for S and $\#\mathcal{P}(S; n) = p(S; n) 2^{n - \#S - 1}$ for all $n > \max S$ if $S \neq \emptyset$ or for all $n \geq 1$ if $S = \emptyset$.

Some of the motivation for our work comes from probability theory. A relationship between permutations and random data has been noticed for quite some time. We refer the reader to the 1937 paper of Kermack and McKendrick [9] and the references therein. Many probabilistic models are concerned with i.i.d. (independent identically distributed) sequences of data, or their generalization, exchangeable sequences. By definition, any permutation of an exchangeable sequence of data is as likely to be observed as the original sequence. One way to test whether a given sequence of n data points is in fact exchangeable is to analyze the order in which the data are arranged, starting from the highest value to the lowest value. Under the assumption of exchangeability, the order should be a randomly (uniformly) chosen permutation of $[n]$. Hence, probabilists are interested in probabilities of various events related to uniformly chosen permutations. This is equivalent to evaluating cardinalities of various subsets of $\mathfrak{S}_n$. This article is inspired by, and provides estimates for, a probabilistic project concerned with mass redistribution, to be presented in a forthcoming article [3]. The reader can consult that paper for details, including the specific i.i.d. sequence which will be used in our model.

The rest of this paper is organized as follows. Section 2 is devoted to a proof of Theorem 1 and its enumerative consequences. Sections 3 and 4 are devoted to computing the polynomial $p(n)$ for various sets of interest for the probabilistic applications. Section 5 investigates a conjecture about the expansion of $p(S;n)$ in a binomial coefficient basis for the space of polynomials. Section 6 states a conjecture about which S maximizes $\#\mathcal{P}(S;n)$ among all S with given cardinality. Section 7 shows how our methods can be applied to the enumeration of permutations with a fixed set of peaks and valleys. Finally, in Section 8 we use our results to prove a known formula for the number of permutations with a given number of peaks.

2 The main enumeration theorem

We need the following result as a base case for induction.

Proposition 2. *For $n \geq 1$ we have*

$$\#\mathcal{P}(\emptyset;n) = 2^{n-1}.$$

Proof. If $\pi \in \mathcal{P}(\emptyset;n)$ then write $\pi = \pi_1 1 \pi_2$ where π_1, π_2 are the portions of π to the left and right of 1, respectively. Now $P(\pi) = \emptyset$ if and only if π_1 is decreasing and π_2 is increasing. So $\#\mathcal{P}(\emptyset;n)$ is the number of choices of a subset of elements from $[2,n]$ to be in π_1 since after that choice is made the rest of π is determined. The result follows. $\square$

We now prove our principal theorem, restating it here for ease of reference.

Theorem 3. *If $S = \{i_1 < \cdots < i_s\}$ is admissible then*

$$\#\mathcal{P}(S;n) = p(n)2^{n-\#S-1}$$

where $p(n) = p(S;n)$ is a polynomial depending on S such that $p(n)$ is an integer for all integral n. In addition, $\deg p(n) = i_s - 1$ (when $S = \emptyset$ we have $\deg p(n) = 0$).

Proof. We induct on $i = i_1 + \cdots + i_s$. By Proposition 2 the result is true when $i = 0$. Now suppose $i \neq 0$. For ease of notation, let $k = i_s - 1$ and $S_1 = S - \{i_s\}$. For any fixed $n > i_s$, consider the set Π of permutations $\pi = a_1 \cdots a_n \in \mathfrak{S}_n$ such that $P(a_1 \cdots a_k) = S_1$ and $P(a_{k+1} \cdots a_n) = \emptyset$. Since S is n-admissible, we know S_1 is also n-admissible by the characterization of n-admissibility after its definition. Thus, $\#\Pi \neq 0$.

We can construct the elements $\pi \in \Pi$ by first picking the set of elements to be used for $a_1, \ldots, a_k$ and then arranging this set and its complement to have the prescribed peak sets. Note that when we concatenate the two sequences, then the resulting permutation π either has a peak at k, or a peak at $k+1$, or neither. Thus, by induction, the total number of choices is

$$\#\Pi = \binom{n}{k} \#\mathcal{P}(S_1; k) \#\mathcal{P}(\emptyset; n-k) = \binom{n}{k} p_1(k) 2^{k-s} \cdot 2^{n-k-1} = p_1(k) \binom{n}{k} 2^{n-s-1}$$

for some polynomial $p_1(n)$ with $\deg p_1(n) = i_{s-1} - 1 < k$. Also, $p_1(k)$ is an integer which must also be nonzero since $\#\Pi \neq 0$.

On the other hand, we can also count Π as follows. Let $S_2 = S_1 \cup \{i_s - 1\}$. Note that by the restrictions on $P(a_1 \cdots a_k)$ and $P(a_{k+1} \cdots a_n)$ we must have $P(\pi) = S_1$, $P(\pi) = S_2$, or $P(\pi) = S$ for all $\pi \in \Pi$. And by construction all such π appear, which shows that we must have the decomposition $\Pi = \mathcal{P}(S_1; n) \cup \mathcal{P}(S_2; n) \cup \mathcal{P}(S; n)$. Taking cardinalities and applying induction as well as the previous count for $\#\Pi$ yields

$$\#\mathcal{P}(S; n) = p_1(k) \binom{n}{k} 2^{n-s-1} - p_1(n) 2^{n-s} - p_2(n) 2^{n-s-1} = \left[p_1(k) \binom{n}{k} - 2p_1(n) - p_2(n) \right] 2^{n-s-1} \tag{1}$$

where $p_2(n)$ is a polynomial in n of degree $i_s - 2 < k$ which is integral at integers if S_2 is admissible, and 0 otherwise. To complete the proof, note that $p_1(k)\binom{n}{k}$ is a polynomial in n of degree k while $2p_1(n)$ and $p_2(n)$ have degree smaller than k. Thus the coefficient of 2^{n-s-1} above is a polynomial of degree $k = i_s - 1$. We also have that the functions $\binom{n}{k}$, $2p_1(n)$, and $p_2(n)$ all have integral values at integers, and we have previously established that $p_1(k)$ is an integer. Thus the same is true of the difference in brackets above. □

Examining the proof of Theorem 3, one sees that it goes through if $n = i_s$ in the sense that we will obtain $p(S; i_s) = 0$ using the fact that there are no permutations in $\mathcal{P}(S; i_s)$. (On the other hand, we may have $p(S; l) \neq 0$ for $l < i_s$ and so some bound is still needed.)

Equation (1) immediately yields the following recursive formula for this polynomial.

Corollary 4. *If $S \neq \emptyset$ is admissible and $m = \max S$ then*

$$p(S; n) = p_1(m-1) \binom{n}{m-1} - 2p_1(n) - p_2(n) \tag{2}$$

where $S_1 = S - \{m\}$, $S_2 = S_1 \cup \{m-1\}$, and $p_i(n) = p(S_i; n)$ for $i = 1, 2$.

It is well known that any sequence given by a polynomial of degree k can be completely determined by any consecutive $k+1$ values by the method of finite differences. See [7, Sections 2.6 and 5.3] or [14, Proposition 1.9.2]. If $f(n)$ is a polynomial function of n, then define another polynomial function of n by $\Delta f(n) = f(n+1) - f(n)$. Similarly, $\Delta^k f$ is a polynomial function of n obtained by applying Δ successively k times. Given the $d+1$ consecutive values $f(m), f(m+1), \ldots, f(m+d)$, then one can express $f(n)$ in the basis $\binom{n-m}{k}$ by the formula

$$f(n) = \sum_{k=0}^{d} \Delta^k f(m) \cdot \binom{n-m}{k}$$

where d is the degree of $f(n)$. Thus, Theorem 3 gives a way to find an explicit formula for $\#\mathcal{P}(S;n)$ and $p(S;n)$ for any admissible set S.

For example, if $S = \{2,5\}$ then $p(S,n)$ has degree 4. Thus, we can find $p(S,n)$ from the sequence $\#\mathcal{P}(S,n)/2^{n-3}$ for $n = 6,7,8,9,10$. Either by hand or computer we find $\#\mathcal{P}(S,6)/2^3 = 10$, $\#\mathcal{P}(S,7)/2^4 = 35$, $\#\mathcal{P}(S,8)/2^5 = 84$, $\#\mathcal{P}(S,9)/2^6 = 168$, $\#\mathcal{P}(S,10)/2^7 = 300$, etc. Taking successive consecutive differences 5 times gives the difference table

$$\begin{array}{ccccccccc}
10 & 35 & 84 & 168 & 300 & 495 & 770 & 1144 & 1638 \\
25 & 49 & 84 & 132 & 195 & 275 & 374 & 494 & \\
24 & 35 & 48 & 63 & 80 & 99 & 120 & & \\
11 & 13 & 15 & 17 & 19 & 21 & & & \\
2 & 2 & 2 & 2 & 2 & & & & \\
0 & 0 & 0 & 0 & & & & &
\end{array}.$$

Therefore,

$$\begin{aligned} p(2,5,n) &= 10\binom{n-6}{0} + 25\binom{n-6}{1} + 24\binom{n-6}{2} + 11\binom{n-6}{3} + 2\binom{n-6}{4} \\ &= \frac{1}{12} n\,(n-5)\,(n-2)\,(n-1)\,. \end{aligned} \tag{3}$$

We used the shifted basis $\binom{n-6}{k}$ here since the smallest value of n for which S is n-admissible is $n = 6$ and we want the sequences aligned properly.

Corollary 5. *If S is a nonempty admissible set and $m = \max S$, then $\#\mathcal{P}(S;n)$ has a rational generating function of the form*

$$\sum_{n\geq 1} \#\mathcal{P}(S;n)x^n = \frac{r(x)}{(1-2x)^m}$$

where $r(x) = r(S;x)$ is the polynomial

$$r(x) = (1-2x)^m \sum_{n\geq 1} \#\mathcal{P}(S;n)x^n = \sum_{k=m+1}^{2m-1} x^k \sum_{j=0}^{k-m-1} (-2)^j \binom{m}{j} \#\mathcal{P}(S;k-j).$$

Furthermore, for $n \geq 2m$ the following linear recurrence relation holds

$$\sum_{j=0}^{m} (-2)^j \binom{m}{j} \#\mathcal{P}(S; n-j) = 0.$$

Proof. These claims follow directly by applying the theory of rational generating functions in [14, Theorem 4.1.1 and Corollary 4.2.1] to Theorem 3 and the discussion of the $n = m$ case directly following it. □

Continuing with the example $S = \{2, 5\}$, we get the generating function

$$\sum_{n\geq 1} \#\mathcal{P}(S; n)x^n = \frac{80x^6 - 240x^7 + 288x^8 - 128x^9}{(1-2x)^5} = \frac{16x^6(1-x)(5-10x+8x^2)}{(1-2x)^5}$$

and recurrence relation

$$\#\mathcal{P}(S; n) = 10\#\mathcal{P}(S; n-1) - 40\#\mathcal{P}(S; n-2) + 80\#\mathcal{P}(S; n-3) - 80\#\mathcal{P}(S; n-4) + 32\#\mathcal{P}(S; n-5)$$

which holds for $n \geq 10$.

3 Specific peak sets with constant elements

We now derive formulas for $\#\mathcal{P}(S; n)$ for various sets S of elements which do not vary with n. These will be useful in proving the results needed for probabilistic applications. Peaks represent sites with no mass in the mass redistribution model analyzed in [3]. In that paper, we will use the results of this section to study the distance between "empty" sites.

Before stating the equations, we would like to indicate how they were originally derived as the proofs below are ones given in hindsight. Note that Corollary 4 expresses $p(S; n)$ in terms of polynomials for peak sets having smaller maxima than S. So by iterating this recursion, one can find an expression for $p(S; n)$ whose main contribution is from an alternating sum $\sum_k (-1)^k a_k \binom{n}{k}$ for certain coefficients $a_k \geq 0$. By next using the binomial recursion iteratively, one achieves substantial cancellation. The simplified formula can then be proved directly using Corollary 4 and these are the results and proofs given below.

Theorem 6. *If $S = \{m\}$ is admissible then*

$$p(S; n) = \binom{n-1}{m-1} - 1.$$

Proof. We induct on m and use the notation of Corollary 4. If $m = 2$ then $S_1 = \emptyset$ and $S_2 = \{1\}$. By Proposition 2 we have $p_1(n) = 1$. Also, $p_2(n) = 0$ since S_2 is not admissible for any n. Now applying Corollary 4 gives

$$p(2; n) = 1 \cdot \binom{n}{1} - 2 \cdot 1 = n - 2 = \binom{n-1}{1} - 1.$$

The induction step is similar, except that now $S_2 = \{m-1\}$ which is admissible, and has a corresponding polynomial which is known by induction. In particular

$$p(m;n) = \binom{n}{m-1} - 2 - \left[\binom{n-1}{m-2} - 1\right] = \binom{n-1}{m-1} - 1$$

as desired. $\square$

Theorem 7. *If $S = \{2, m\}$ is admissible then*

$$p(S;n) = (m-3)\binom{n-2}{m-1} + (m-2)\binom{n-2}{m-2} - \binom{n-2}{1}.$$

Proof. The proof is much like that of the previous proposition where $S_1 = \{2\}$ and $S_2 = \{2, m-1\}$. The details are left to the reader. $\square$

From this theorem, we immediately get that

$$p(2,5;n) = 2\binom{n-2}{4} + 3\binom{n-2}{3} - \binom{n-2}{1} = \frac{1}{12}\, n\,(n-5)\,(n-2)\,(n-1)\,.$$

Note that this agrees with our computations in (3).

Using the same technique, one can prove the following result whose demonstration is omitted.

Theorem 8. *If $S = \{2, m, m+2\}$ is admissible then*

$$p(S;n) = m(m-3)\binom{n}{m+1} - 2(m-3)\binom{n-2}{m-1} - 2(m-2)\binom{n-2}{m-2} + 2\binom{n-2}{1}.$$

4 Peak sets depending on a parameter

Consider sets of the form

$$S = \{i_1 < i_2 < \cdots < i_s < n - j_t < \cdots < n - j_2 < n - j_1\}$$

where $i_1, \ldots, i_s, j_1 \ldots, j_t$ are constants. Using exactly the same definition of admissibility as with sets of constants, one can show that Theorem 3 continues to hold for such S. The specific statement is as follows.

Theorem 9. *Let $S = \{i_1 < i_2 < \cdots < i_s < n - j_t < \cdots < n - j_2 < n - j_1\}$ be admissible. Then*

$$\#\mathcal{P}(S;n) = p(n)2^{n-\#S-1}$$

where $p(n) = p(S;n)$ is a polynomial depending on S such that $p(n)$ is an integer for all integral n. In addition,

$$\deg p(n) = \begin{cases} 0, & \text{if } s = t = 0; \\ i_s - 1, & \text{if } s > 0 \text{ and } t = 0; \\ j_t, & \text{if } s = 0 \text{ and } t > 0; \\ i_s + j_t - 1, & \text{otherwise.} \end{cases}$$

The demonstration is an induction on $i_1 + \cdots + i_s + j_1 + \cdots + j_t$ which is similar to the one given previously and so is omitted.

Next, we will use the results of the previous section to obtain formulas for peak sets S with $2, n-1 \in S$ which are useful probabilistically. Of all peak sets that a probabilist might consider, two peaks at the (almost) extreme points of a sequence, namely at 2 and $n-1$, are of the greatest interest because they represent the distribution of the distance between adjacent empty sites in the mass redistribution model in [3]. So a peak set which contains 2 and $n-1$ (with possibly other numbers as well) represents the joint distribution of several consecutive maximal sequences of consecutive empty sites.

For sets containing 2 and $n-1$, there is an alternative way to compute $p(S;n)$ which is simpler because it avoids alternating sums. If $\pi = a_1 \cdots a_n \in \mathcal{P}(S;n)$ where $2, n-1 \in S$ then $n = a_{i_j}$ for some $i_j \in S$. (One can also use similar reasoning for more general sets S which do not satisfy the given hypothesis, but one needs to worry about the possibility that $a_1 = n$ or $a_n = n$.) Note also that if we consider the reversal $\pi^r = a_n \cdots a_1$ then $P(\pi^r) = \{n+1-i_s, \ldots, n+1-i_1\}$ where $S = \{i_1, \ldots, i_s\}$. So if we write $\pi = \pi_L n \pi_R$ we have $\pi_L \in \mathcal{P}(S_L; i_j - 1)$ and $\pi_R^r \in \mathcal{P}(S_R^r; n - i_j)$ where $S_L = \{i_1, \ldots, i_{j-1}\}$ and $S_R^r = \{n+1-i_s, \ldots, n+1-i_{j+1}\}$. These observations yield the recursion

$$\#\mathcal{P}(S;n) = \sum_{j=1}^{s} \#\mathcal{P}(S_L; i_j - 1)\#\mathcal{P}(S_R^r; n - i_j)\binom{n-1}{i_j - 1}.$$

Using Theorem 3 and canceling powers of 2 gives

$$2p(S;n) = \sum_{j=1}^{s} p(S_L; i_j - 1)p(S_R^r; n - i_j)\binom{n-1}{i_j - 1}. \tag{4}$$

Because of the complexity of the formulas, we will often keep the 2 above on the left-hand side of the equation.

Theorem 10. *If $S = \{2, n-1\}$ is admissible then*

$$p(S;n) = (n-1)(n-4).$$

Remark 11. Andrew Crites pointed out that we get the same result if we substitute $m = n-1$ into the formula in Theorem 7.

Proof. In this case, equation (4) has two terms. In the first $S_L = \emptyset$ and $S_R^r = \{2\}$, while in the second $S_L = \{2\}$ and $S_R^r = \emptyset$. Applying Proposition 2 and Theorem 6, we obtain

$$2p(S;n) = 1 \cdot (n-4) \cdot \binom{n-1}{1} + (n-4) \cdot 1 \cdot \binom{n-1}{1}$$

from which the desired equation follows. $\square$

Theorem 12. *If $S = \{2, m, n-1\}$ is admissible then*

$$\begin{aligned} 2p(S;n) &= (m-3)(n-m-2)\binom{n-1}{m-1} + (n-1)\left[(m-3)\binom{n-4}{m-1} + (m-2)\binom{n-4}{m-2}\right. \\ &\qquad \left. +(n-m-1)\binom{n-4}{m-3} + (n-m-2)\binom{n-4}{m-4} - 2\binom{n-4}{1}\right]. \end{aligned}$$

For fixed n, the sequence $p(S;n)$ is symmetric and unimodal as m varies and only attains its maximum at $m = \left\lfloor \frac{n+1}{2} \right\rfloor$ and at $m = \left\lceil \frac{n+1}{2} \right\rceil$.

Proof. The formula for $2p(S;n)$ follows from equation (4) and the results of the previous section similarly to the proof of Theorem 10. So we leave the details to the reader.

For fixed n, let us write $f(m) = 2p(2, m, n-1; n)$ where $4 \le m \le n-3$. The fact that this sequence is symmetric as a function of m follows directly from the form of S. To prove the rest of the theorem, it suffices to show that the first half of the sequence is strictly increasing. So consider the difference $f(m+1) - f(m)$ where $m \le (n-1)/2$. The first term of $f(m)$ contributes

$$(m-2)(n-m-3)\binom{n-1}{m} - (m-3)(n-m-2)\binom{n-1}{m-1} > 0$$

since, for the given range of m, we have $(m-2)(n-m-3) \ge (m-3)(n-m-2)$ by log concavity of the integers, and $\binom{n-1}{m} > \binom{n-1}{m-1}$ by unimodality of the binomial coefficients. Now consider the the terms with a factor of $n-1$. Combining terms corresponding to the same binomial coefficient and then using binomial coefficient unimodality gives a contribution to the difference of

$$\begin{aligned} &(m-2)\binom{n-4}{m} + 2\binom{n-4}{m-1} + (n-2m)\binom{n-4}{m-2} - 2\binom{n-4}{m-3} - (n-m-2)\binom{n-4}{m-4} \\ &> [(m-2) + (n-2m) - (n-m-2)]\binom{m-4}{m-4} + 2\left[\binom{n-4}{m-1} - \binom{n-4}{m-3}\right] \\ &= 2\left[\binom{n-4}{m-1} - \binom{n-4}{m-3}\right] \\ &> 0 \end{aligned}$$

which is what we wished to show. $\square$

The proof of the next theorem contains no new ideas and so is omitted.

Theorem 13. *If $S = \{2, m, m+2, n-1\}$ is admissible then*

$$\begin{aligned} 2p(S;n) &= (m-3)(n-m-4)\left[m\binom{n-1}{m+1} + (n-m-1)\binom{n-1}{m-1}\right] \\ &\quad +(n-1)\left[m(m-3)\binom{n-2}{m+1} + (n-m-1)(n-m-4)\binom{n-2}{m-2}\right. \\ &\quad \left. -2(n-6)\binom{n-4}{m-1} - 2(n-6)\binom{n-4}{m-2} + 4\binom{n-4}{1}\right]. \end{aligned}$$

5 A positivity conjecture

Given any integer m we have the following basis for $\mathbb{Q}[n]$, the ring of polynomials in a variable n with coefficients which are rational numbers,

$$\mathcal{B}_m = \left\{ \binom{n-m}{k} \ : \ k \geq 0 \right\}.$$

Consider a polynomial $p(n) \in \mathbb{Q}[n]$. It follows from Stanley's text [14, Corollary 1.9.3] that $p(n)$ is an integer for all integral n if and only if the coefficients in the expansion of $p(n)$ using $\mathcal{B}_0$ are all integral. In particular, this is true for $p(n) = p(S;n)$ by our main theorem.

One might wonder if the coefficients in the $\mathcal{B}_0$-expansion of $p(S;n)$ were also nonnegative. Unfortunately, it is easy to see from Theorem 6 that this is not always the case. However, we conjecture that $p(S;n)$ can be written as a nonnegative linear combination of the elements in another basis.

Throughout this section, let S be a nonempty admissible set of constants and $m = \max S$. Let c_k^S be the coefficient of $\binom{n-m}{k}$ in the expansion of $p(S;n)$, so

$$p(S;n) = \sum_{k=0}^{m-1} c_k^S \binom{n-m}{k},$$

where we know from Theorem 3 that c_{m-1}^S is a positive integer and $c_k^S = 0$ for $k \geq m$.

Conjecture 14. Each coefficient c_k^S is a positive integer for all $0 < k < m$ and all admissible sets S.

We have confirmed this conjecture for all admissible subsets S with $\max S \leq 20$. As further evidence for this conjecture, we investigate some special cases. We will first concern ourselves with c_0^S. In the following results, we use the usual convention that $\binom{a}{b} = 0$ unless $0 \leq b \leq a$.

Lemma 15. *For any nonempty set S with constant elements, we have $c_0^S = 0$.*

Proof. If S is not admissible then $c_k^S = 0$ for all k. Now suppose S is admissible and $m = \max S$. From our discussion following the proof of Theorem 3 about the case $n = i_S = m$, we see that $p(S; m) = 0$. Since our basis is $\binom{n-m}{k}$ for $k \geq 0$, we must therefore have $c_0^S = p(S; m) = 0$. □

Next we consider what happens for peak sets with one element.

Proposition 16. *If $S = \{m\}$ is admissible then*

$$c_k^S = \begin{cases} \binom{m-1}{k}, & \text{if } k \geq 1; \\ 0, & \text{if } k = 0. \end{cases}$$

Proof. Using Theorem 6 and Vandermonde's convolution give

$$\begin{aligned} p(S; n) &= -1 + \binom{n-1}{m-1} \\ &= -1 + \sum_{k \geq 0} \binom{m-1}{m-k-1} \binom{n-m}{k} \\ &= \sum_{k \geq 1} \binom{m-1}{k} \binom{n-m}{k} \end{aligned}$$

which is what we wished to prove. □

To deal with peak sets having two elements, we will need the characteristic function χ which evaluates to 1 on a true statement and 0 on a false one.

Proposition 17. *If $S = \{2, m\}$ is admissible then*

$$c_k^S = (m-3)\binom{m-2}{k-1} + (m-2)\binom{m-2}{k} - \binom{m-2}{k+m-3}. \tag{5}$$

Furthermore, $c_k^S \geq 0$ for all k.

Proof. To prove the formula for c_k^S, one first shows that

$$c_k^S = -2\binom{m-2}{k+m-3}\chi(m \text{ is even}) + \sum_{j=0}^{m-4} (-1)^j \left[\binom{m-j-2}{1} - 1\right] \binom{m}{k+j+1}. \tag{6}$$

Since this equality will be generalized in the next proposition, we will provide the details of the proof there.

To simplify this expression, consider the summation part which we will denote by Σ. Use the binomial recursion twice on $\binom{m}{k+j+1}$, each time reindexing the summation to combine terms, to get

$$\begin{aligned}\Sigma &= (m-3)\binom{m-1}{k} + \sum_{j=0}^{m-4}(-1)^j\binom{m-1}{k+j+1} \\ &= (m-3)\binom{m-2}{k-1} + (m-2)\binom{m-2}{k} + (-1)^{m-4}\binom{m-2}{k+m-3}.\end{aligned}$$

Adding in the term of equation (6) containing χ yields the desired formula.

We now prove positivity. If $k \geq 2$ then the last binomial coefficient in equation (5) is zero and the result is obvious. It is also easy to check the case $k = 1$, and $k = 0$ is Lemma 15. $\square$

We now consider the case of an arbitrary 2-element set. While we are able to obtain a general summation formula in this case, it does not seem to simplify readily and so we are only able to prove positivity for roughly half the coefficients.

Theorem 18. *If $S = \{l, m\}$ is admissible then $c_0^S = 0$, and for $k \geq 1$*

$$c_k^S = -2\binom{m-1}{k+m-l}\chi(m-l \text{ is even}) + \sum_{j=0}^{m-l-2}(-1)^j\left[\binom{m-j-2}{l-1} - 1\right]\binom{m}{k+j+1}.$$

Furthermore, $c_k^S \geq 0$ for $k \geq (m-2)/2$.

Proof. We first prove the formula for c_k^S. The case $k = 0$ is taken care of by Lemma 15. Let $p(n) = p(S; n)$. For $k \geq 1$ we will prove the formula for c_k^S by fixing l and inducting on m.

First consider the base case $m = l + 2$. Then $S_1 = \{l\}$ and $S_2 = \{l, l+1\}$ which is not admissible. Thus, using Theorem 6, Vandermonde's convolution, and the fact that $c_0^S = 0$, we see that equation (2) becomes

$$\begin{aligned}p(n) &= p_1(m-1)\binom{n}{m-1} - 2p_1(n) \\ &= \left[\binom{m-2}{l-1} - 1\right]\binom{n}{m-1} - 2\left[\binom{n-1}{l-1} - 1\right] \\ &= \sum_{k\geq 1}\left[\binom{m-2}{l-1} - 1\right]\binom{m}{m-1-k}\binom{n-m}{k} - 2\binom{m-1}{l-k-1}\binom{n-m}{k} \\ &= \sum_{k\geq 1}\left\{\left[\binom{m-2}{l-1} - 1\right]\binom{m}{k+1} - 2\binom{m-1}{k+m-l}\right\}\binom{n-m}{k}.\end{aligned}$$

The coefficient of $\binom{n-m}{k}$ in this expression agrees with the one given in the theorem when $m = l + 2$ and so we are done with the base case.

Now consider $m > l+2$. There are two subcases depending on whether $m-l$ is even or odd. Since they are similar, we will just do the latter. The computations in the base case remain valid except for the fact that $S_2 = \{l, m-1\}$ is now admissible and so we need to subtract off the $p_2(n)$ term in equation (2). For simplicity, let a_k denote the coefficient of $p_2(n)$ expanded in the basis $\mathcal{B}_{m-1}$. Since $m-l-1$ is even we have, by induction,

$$a_k = -2\binom{m-2}{k+m-l-1} + \sum_{j=0}^{m-l-3} (-1)^j \left[\binom{m-j-3}{l-1} - 1\right] \binom{m-1}{k+j+1}$$

when $k \geq 1$ and, as always, $a_0 = 0$. To convert to the basis $\mathcal{B}_m$, we compute

$$p_2(n) = \sum_{k\geq 0} a_k \binom{n-m+1}{k} = \sum_{k\geq 0} a_k \left[\binom{n-m}{k-1} + \binom{n-m}{k}\right] = \sum_{k\geq 0} (a_k + a_{k+1}) \binom{n-m}{k}.$$

It follows from the previous two displayed equations that the coefficient of $\binom{n-m}{k}$ in $-p_2(n)$ is

$$2\binom{m-1}{k-m-l} - \sum_{j=0}^{m-l-3} (-1)^j \left[\binom{m-j-3}{l-1} - 1\right] \binom{m}{k+j+2}.$$

Shifting indices in this last sum and adding in the contribution from the computation for $p(n)$ in the base case completes the induction step.

The proof of positivity breaks down into two cases depending on the parity of $m-l$. Since they are similar, we will only present the details when $m-l$ is odd. It suffices to show that the absolute values of the terms in the sum for c_k^S are weakly decreasing since then each negative term can be canceled into the preceding positive one. Clearly the term in square brackets is decreasing with j. And because $k+1 \geq m/2$ we have that $\binom{m}{k+j+1}$ is also decreasing by unimodality of the binomial coefficients. This completes the proof. $\square$

6 Equidistribution

Suppose one considers the distribution of $\#\mathcal{P}(S;n)$ over all possible peak sets $S = \{i_1, \ldots, i_s\}$ with s elements. We conjecture that a maximum will occur when the elements of S are as evenly spaced as possible.

There are two natural probabilistic conjectures which one could make about the peak distribution, assuming a small number of peaks in a long sequence. First, one could guess that the places where peaks occur is an approximation to Poisson process arrivals, and hence locations of the peaks are distributed approximately uniformly over the whole sequence and are approximately independent. Available evidence points to an alternative conjecture that the peaks have a tendency to repel each other. This phenomenon is found in some random models, for example, under certain assumptions, eigenvalues of random matrices have a tendency to repel each other. We do not see a direct connection with that model at the technical level, but the repelling nature of peaks invites further exploration.

It will be useful to pass from the set S to the corresponding composition. A *composition of n into k parts* is a sequence of positive integers $\kappa = (\kappa_1, \ldots, \kappa_k)$ where $\sum_j \kappa_j = n$. We also write $\kappa = (a^{m_a}, b^{m_b}, \ldots)$ for the composition that starts with m_a copies of the part a, then m_b copies of the part b, and so forth. Given any set $S = \{i_1, \ldots, i_s\}$ of $[n]$ there is a corresponding composition $\kappa(S)$ of $n+1$ into $s+1$ parts given by $\kappa_j = i_j - i_{j-1}$ for $1 \le j \le s+1$ where we let $i_0 = 0$ and $i_{s+1} = n+1$. This construction is bijective. Given any composition $\kappa = (\kappa_1, \ldots, \kappa_{s+1})$ of $n+1$ we can recover $S = \{i_1, \ldots, i_s\} \subseteq [n]$ where $i_j = \kappa_1 + \cdots + \kappa_j$ for $1 \le j \le s$.

A composition is *Turán* if $|\kappa_a - \kappa_b| \le 1$ for all a, b. This terminology is in reference to Turán's famous theorem in graph theory (about maximizing the number of edges in a graph with no complete subgraph of a given order) where these compositions play an important role. There is another description of Turán compositions which will be useful. Suppose we wish to form a Turán composition of n with k parts. Apply the Division Algorithm to write $n = qk + r$ where $0 \le r < k$. Then the desired compositions are exactly those gotten by permuting $k-r$ copies of the part q and r copies of the part $q+1$. We will call q the *quotient* corresponding to the Turán composition.

Conjecture 19 (Equidistribution Conjecture)**.** If n and s are fixed positive integers, then we conjecture the following two statements.

(a) If $S \subseteq [n]$ maximizes $\#\mathcal{P}(S; n)$ among all subsets with $\#S = s$, then $\kappa(S)$ is Turán.

(b) The maximizing Turán compositions in (a) are precisely those of the form

$$((q+1)^{m_1}, q^{m_2}, (q+1)^{m_3})$$

where q is the quotient of $\kappa(S)$ and as many of the multiplicities m_1 and m_3 are positive as possible. (If there is only one copy of $q+1$, then one of these two multiplicities is zero and the other equals one, and if there are no copies then both multiplicities are zero.)

If n is fixed and $s = \#S$ is allowed to vary, then we conjecture that the peak sets maximizing $\#\mathcal{P}(S; n)$ over all $S \subseteq [n]$ are the Turán compositions satisfying (b) with the maximum number of 3's.

Note that for $s = 1$ this conjecture is true because of Theorem 6. It has also been verified by computer for $n \le 13$. The part of the conjecture about maximization over all $S \subseteq [n]$ is consistent with a result of Kermack and McKendrick [9] stating that the mean size of a part in all $\kappa(S)$ with S admissible is 3.

7 Peaks and valleys

For some applications, it will be useful to know the number of permutations with peaks at 2 and $n-1$ and a valley at a given position m. In the mass redistribution model analyzed in [3], valleys represent the oldest sites, where age is measured since the last mass redistribution. It is a natural question to investigate the relationship between the oldest sites (valleys) and the sites most recently affected by the mass redistribution process (peaks).

In this section we derive the desired formula. To set up notation, let

$$\begin{aligned} PV(\pi) &= \text{set of peaks and valleys of permutation } \pi, \\ \mathcal{PV}(i_1,\ldots,i_s;n) &= \{\pi \in \mathfrak{S}_n \ :\ PV(\pi) = \{i_1,\ldots,i_s\} \text{ and } i_1 \text{ is a peak}\}. \end{aligned}$$

As usual, we require $i_1 < \cdots < i_s$. Of course, peaks and valleys must alternate. So $\mathcal{PV}(i_1,\ldots,i_s;n)$ also counts permutations π with $PV(\pi) = \{i_1,\ldots,i_s\}$ and i_1 being a valley, a fact which will be useful in the sequel. The definition of admissible is as before.

It is easy to adapt the proof of Theorem 3 to this setting, so the proof of the next result is omitted.

Theorem 20. *If $S = \{i_1 < \cdots < i_s\}$ is admissible then*

$$\#\mathcal{PV}(S;n) = q(n)$$

where $q(n)$ is a polynomial depending on S such that $q(n)$ is an integer for all integral n. In addition, if S is a set of constants not depending on n then $\deg q(n) = i_s - 1$ (when $S = \emptyset$ we have $\deg q(n) = 0$).

We now derived a formula for $\#\mathcal{PV}(2,m,n-1;n)$ via a sequence of results. Since the techniques are much like those we have used before, the proofs will only be sketched.

Proposition 21. *If $\{m\}$ is admissible then*

$$\#\mathcal{PV}(m;n) = \binom{n-1}{m-1}.$$

Proof. We have $\pi \in \mathcal{PV}(m;n)$ if and only if $\pi = \pi_L n \pi_R$ where π_L is increasing, $\#\pi_L = m-1$, π_R is decreasing, and $\#\pi_R = n-m$. □

Proposition 22. *If $\{2,m\}$ is admissible then*

$$\#\mathcal{PV}(2,m;n) = \binom{n-2}{m-2} + (m-2)\binom{n-1}{m-1}.$$

Proof. If $a_1 \ldots a_n \in \mathcal{PV}(2,m;n)$ then either $a_1 = 1$ or $a_m = 1$. In the first case, the number of π is given by $\binom{n-2}{m-2}$ by the previous proposition. In the second case, there are $\binom{n-1}{m-1}$ ways to pick the elements to the left of 1 and then $m-2$ ways to pick a_1. □

Proposition 23. *If $\{2,m,n-1\}$ is admissible then*

$$\#\mathcal{PV}(2,m,n-1;n) = 2(n-1)\left[\binom{n-4}{m-2} + (m-2)\binom{n-3}{m-1}\right].$$

Proof. By symmetry, it suffices to double the number of $a_1 \ldots a_n \in \mathcal{PV}(2,m,n-1;n)$ where $a_{n-1} = n$. There are $n-1$ ways to choose a_n. Using the previous proposition, we see that the number of ways to pick the remaining elements is given by the expression in the square brackets. □

8 Fixing the number of peaks

We now use our theorems to prove a result already in the literature. In general, there does not seem to be a simple explicit formula for the number $f(s,n)$ of permutations in $\mathfrak{S}_n$ with s peaks. For more information about these numbers, see sequence A008303 in the Online Encyclopedia of Integer Sequences (OEIS). However, David and Barton [5, p. 163] give the recurrence

$$f(s,n) = (2s+2)f(s,n-1) + (n-2s)f(s-1,n-1)$$

with the initial conditions that $f(0,n) = 2^{n-1}$ and $f(s,n) = 0$ whenever $s \geq \frac{n}{2}$. In addition, for small s, one can write down an explicit expression for $f(s,n)$. In fact, the sequence $f(1,n)$ appears as sequence A000431 in the OEIS where the following result is attributed to Mitchell Harris.

Proposition 24. *For $n \geq 1$*

$$f(1,n) = 2^{2n-3} - n2^{n-2}.$$

Proof. Using Theorems 3 and 6 we have

$$f(1,n) = \sum_{m=2}^{n-1} \left[\binom{n-1}{m-1} - 1 \right] 2^{n-2} = 2^{n-2} \sum_{m=1}^{n} \left[\binom{n-1}{m-1} - 1 \right] = 2^{n-2}(2^{n-1} - n)$$

which multiplies out to the formula we want. □

9 Acknowledgments

We would like to thank Andrew Crites and Soumik Pal for stimulating discussions about peak sets. We would also like to thank Andrew Sills for information about hypergeometric series. We credit the OEIS for assisting us with our research on this project. Finally, we would like to thank the two anonymous referees for useful suggestions.

Note added in proof. Recently, Kasraoui [8] has proved parts (a) and (b) of the Equidistribution Conjecture.

References

[1] M. Aguiar, K. Nyman, and R. Orellana, New results on the peak algebra, *J. Algebraic Combin.* **23** (2006) 149–188.

[2] L. J. Billera, S. K. Hsiao, and S. van Willigenburg, Peak quasisymmetric functions and Eulerian enumeration, *Adv. Math.* **176** (2003) 248–276.

[3] S. Billey, K. Burdzy, S. Pal, and B. E. Sagan, On meteors and earthworms, forthcoming, 2013.

[4] P. Bouchard, H. Chang, J. Ma, J. Yeh, and Y.-N. Yeh, Value-peaks of permutations, *Electron. J. Combin.* **17** (2010) Research Paper 46, 20.

[5] F. N. David and D. E. Barton, *Combinatorial Chance*, Hafner Publishing Co., 1962.

[6] J. Françon and G. Viennot, Permutations selon leurs pics, creux, doubles montées et double descentes, nombres d'Euler et nombres de Genocchi, *Discrete Math.* **28** (1979) 21–35.

[7] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 2nd edition, 1994.

[8] A. Kasraoui, The most frequent peak set in a random permutation. Preprint available at http://arxiv.org/abs/1210.5869.

[9] W. O. Kermack and A. G. McKendrick. Tests for randomness in a series of numerical observations. *Proc. Roy. Soc. Edinburgh* **57** (1937) 228–240.

[10] S.-M. Ma, Derivative polynomials and enumeration of permutations by number of interior and left peaks, *Discrete Math.* **312** (2012) 405–412.

[11] K. L. Nyman, The peak algebra of the symmetric group, *J. Algebraic Combin.* **17** (2003) 309–322.

[12] T. K. Petersen, Enriched P-partitions and peak algebras, *Adv. Math.* **209** (2007) 561–610.

[13] M. Schocker, The peak algebra of the symmetric group revisited, *Adv. Math.* **192** (2005) 259–309.

[14] R. P. Stanley, *Enumerative Combinatorics, Volume 1*, Vol. 49 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, 2nd edition, 2012.

[15] J. R. Stembridge, Enriched P-partitions, *Trans. Amer. Math. Soc.* **349** (1997) 763–788.

[16] V. Strehl, Enumeration of alternating permutations according to peak sets, *J. Combin. Theory Ser. A* **24** (1978) 238–240.

[17] D. Warren and E. Seneta, Peaks and Eulerian numbers in a random sequence, *J. Appl. Probab.* **33** (1996) 101–114.

2010 *Mathematics Subject Classification*: Primary 05A05; Secondary 05A10, 05A15.
Keywords: binomial coefficient, peak, permutation.

(Concerned with sequences A000431 and A008303.)

Received December 3 2012; revised version received May 22 2013. Published in *Journal of Integer Sequences*, June 5 2013.

Return to Journal of Integer Sequences home page.