Pushes in permutations

Aubrey Blecher, Charlotte Brennan*
Arnold Knopfmacher*
The John Knopfmacher Centre for Applicable Analysis
and Number Theory, School of Mathematics
University of the Witwatersrand
Private Bag 3, Wits 2050, Johannesburg, South Africa
Email: *Aubrey.Blecher@wits.ac.za*, *Charlotte.Brennan@wits.ac.za*,
Arnold.Knopfmacher@wits.ac.za

Toufik Mansour
Department of Mathematics
University of Haifa, 3498838 Haifa, Israel
Email: *tmansour@univ.haifa.ac.il*

Mark Shattuck
Department of Mathematics
University of Tennessee, Knoxville, TN 37996, USA
Email: *shattuck@math.utk.edu*

January 23, 2020

Abstract

We define the push statistic on permutations and multipermutations and use this to obtain various results measuring the degree to which an arbitrary permutation deviates from sorted order. We study the distribution on permutations for the statistic recording the length of the longest push and derive an explicit expression for its first moment and generating function.

*This material is based upon work supported by the National Research Foundation under grant numbers 86329, 81021 respectively

Several auxiliary concepts are also investigated. These include the number of cells that are not pushed; the number of cells that coincide before and after pushing (i.e., the fixed cells of a permutation); and finally the number of groups of adjacent columns of the same height that must be reordered at some point during the pushing process.

1 Introduction

The combinatorial objects considered in this paper are permutations and multipermutations (collectively referred to as permutations). A permutation of $[n] = \{1, 2, \ldots, n\}$ is an ordering of the elements of $[n]$, and for any positive integer m, a multipermutation of the multiset $\{1^m, 2^m, \ldots, n^m\}$ is an ordering of $\{1^m, 2^m, \ldots, n^m\}$ where the m copies of each member of $[n]$ are considered indistinguishable (see, e.g., [12, Section 1.3]). It is well known that there are $n!$ permutations of $[n]$ and $\binom{nm}{m,\ldots,m}$ multipermutations of $\{1^m, 2^m, \ldots, n^m\}$, where it is understood that the m occurs n times in the multinomial coefficient. Let $\mathcal{S}_n$ denote the set of permutations of $[n]$ and $\mathcal{S}_{n,m}$ the set of multipermutations of $\{1^m, 2^m, \ldots, n^m\}$ for $m \geq 1$. Statistics related to left-to-right maxima on $\mathcal{S}_n$ and $\mathcal{S}_{n,m}$ have been previously studied in [10, 13]. See also [6] for related asymptotic properties on iid sequences of discrete random variables.

We define here a new statistic on permutations as follows. Consider an arbitrary member of $\mathcal{S}_{n,m}$ with its corresponding graphical representation where a column of height r represents a part of size r. We say that a column of size r consists of r unit squares called *cells*. Suppose that the leftmost element in the permutation which is not a weak left-to-right maximum (a weak left-to-right maximum is a part which is greater than or equal to all parts to its left) occurs in position i and has height $v(i)$. Shift all cells which are to the left of i and of height greater than $v(i)$ one position to the right. We call this shifting process a *push*. We apply a sequence of pushes, successively, to a permutation, which terminates once a weakly increasing permutation has been achieved. Illustrated in Figure 1 below is the multipermutation 322131 of $\{1^2, 2^2, 3^2\}$ and the sequence of four pushes leading to its counterpart 112233, which is the sorted order of any member of $\mathcal{S}_{3,2}$.

Remark 1. *We see that the total number of pushes required to generate a weakly increasing permutation is equal to the length minus the number of weak left-to-right maxima. In the example above, the length is 6 and there are 2 weak left-to-right maxima resulting in 4 pushes.*

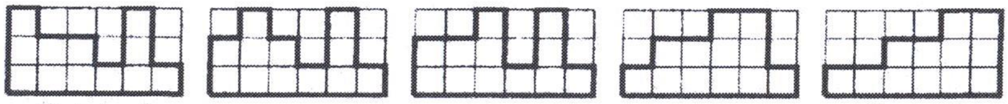


Figure 1: The multipermutation 322131 with its 4 pushes

Here, we study pushes and some associated concepts such as the total number of cells which do not move in the pushing process (see Section 2) and the total number of cells that coincide before and after the process (Section 3). In the fourth section, we consider the notion of a *frictionless push* where cells are shifted to the right not necessarily one position at a time but perhaps several positions corresponding to a group of adjacent columns of the same height strictly less than that of the cells being shifted. A simple formula for the average number of frictionless pushes in multipermutations is derived, and a combinatorial proof of this result is provided. In the fifth section, we consider a statistic on permutations recording the length of the longest push and derive an explicit formula for the generating function of the distribution and its first moment. Some concluding remarks are made in the final section.

2 Number of cells that do not move

The average number of left-to-right maxima in permutations of length n is well known to be the harmonic number $H_n = \sum_{j=1}^{n} \frac{1}{j}$ (see, e.g., [3, 4]). Thus, by Remark 1, we have the following result.

Theorem 1. *The average number of pushes over all permutations of* $[n]$ *is given by*

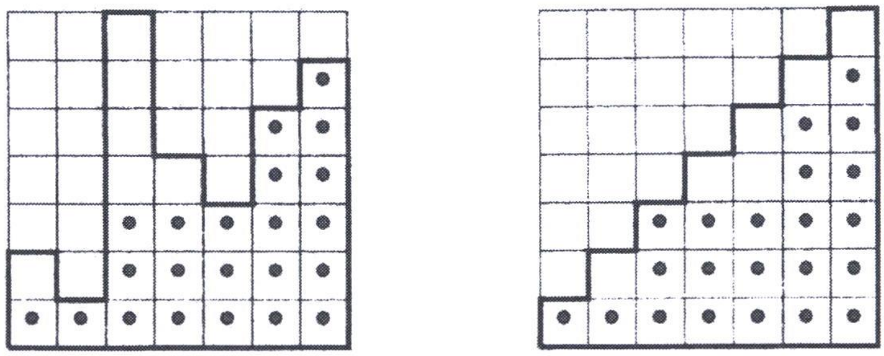
$$n - H_n.$$

The number of pushes in multipermutations is equal to the length minus the number of weak left-to-right maxima, which have been studied previously by Myers and Wilf in [10]. Using their result on multipermutations, we obtain the following.

Theorem 2. *The average number of pushes over all multipermutations of* $\{1^m, 2^m, \ldots, n^m\}$ *is given by*

$$nm - \sum_{i=1}^{n} \frac{m}{(n-i)m+1}.$$

In the rest of this section, we calculate the total number of cells that do not move in the pushing process over all permutations of length n. In Figure 2 below, the cells that do not move are indicated by a dot.



Original permutation | Pushed permutation

Figure 2: The cells that do not move in the permutation 2174356

Let $P(n)$ denote the total number of cells that do not move, taken across all members of $\mathcal{S}_n$. We will construct (as illustrated below) the permutations of length $n+1$ from those of length n using the following process: For an arbitrary member of $\mathcal{S}_n$, place $n+1$ in the k-th position (from the left) for some $1 \leq k \leq n+1$. Let $\Delta_{n,k}$ be the additional number of unmoved cells across all members of $\mathcal{S}_{n+1}$ belonging to the column of height $n+1$.

We note that the insertion of $n+1$ does not change which cells did not move in the prior n-case. In other words, a cell to the right or left of the k-th position will not move under the pushing process of the new permutation if and only if it did not move in the pushing process for the member of $\mathcal{S}_n$ from which it arose. Furthermore, the number of additional cells that do not move and belonging to the new k-th part of size $n+1$ is seen to equal the size of the smallest part to the right of $n+1$.

As indicated in the sketch below, let i denote the smallest part in the permutation of $[n]$ which occurs to the right of the k-th position. Hence, the parts $1, 2, \ldots, i-1$ occur in any order to the left of the k-th position.

The number of permutations with this structure is then given by BCD as in (2.1) below, where B is the number of possible positions for i, C is the number of ways to arrange and order the parts less than i, and finally D is the number of ways to order the parts greater than i (excluding $n+1$) in their positions.

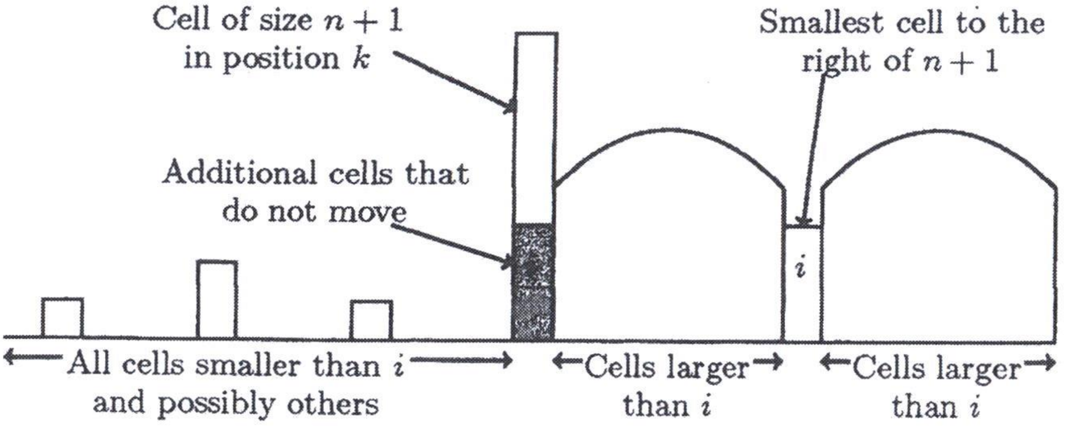


Figure 3: Decomposition of a permutation with regard to unmoved cells

$$\underbrace{(n-k+1)}_{B}\underbrace{\binom{k-1}{i-1}(i-1)!}_{C}\underbrace{(n-i)!}_{D} \tag{2.1}$$

Since each permutation with this structure contributes i cells, the total number is given by

$$i\,(n-k+1)\binom{k-1}{i-1}(i-1)!\,(n-i)!.$$

Considering all possible i implies

$$\Delta_{n,k}=\sum_{i=1}^{k}(n-k+1)\binom{k-1}{i-1}(i-1)!\,(n-i)!\,i.$$

Note that $i \leq k$, otherwise, it is not the minimum to the right of position k. Thus

$$\begin{aligned}\Delta_{n,k} &= (n-k+1)\sum_{i=1}^{k}\frac{i\,(k-1)!\,(i-1)!(n-i)!}{(i-1)!(k-i)!}\\ &= (n-k+1)(k-1)!\sum_{i=1}^{k}\frac{i(n-i)!}{(k-i)!}\\ &= (n-k+1)(k-1)!\frac{(n+1)!}{(n-k+1)(n-k+2)(k-1)!}=\frac{(n+1)!}{n+2-k}.\end{aligned}$$

Summing over all k, we obtain

$$\begin{aligned} P(n+1)-(n+1)P(n) &= \sum_{k=2}^{n} \frac{(n+1)!}{n+2-k} + n!(n+2) \\ &= (n+1)!(H_n - 1) + n!(n+2) \\ &= n!((n+1)H_n + 1), \end{aligned}$$

where the $n!(n+2)$ term accounts for the cases where $k=1$ and $k=n+1$. This last equation may be rewritten as

$$\frac{P(n+1)}{(n+1)!} = \frac{P(n)}{n!} + H_{n+1}, \qquad n \geq 1,$$

with initial condition $P(1) = 1$. Letting $av(n) := \frac{P(n)}{n!}$, we obtain the recursion $av(n+1) = av(n) + H_{n+1}$, which has solution

$$av(n) = (n+1)(H_{n+1} - 1).$$

Thus, we have the following result.

Theorem 3. *The number of cells that do not move in all permutations of length n is*

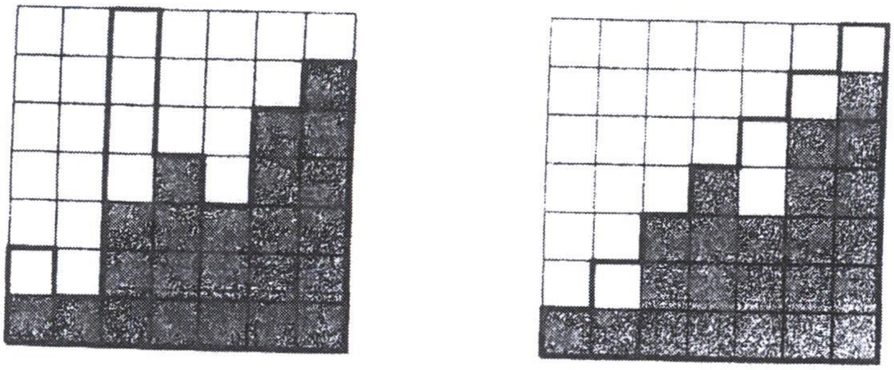
$$P(n) = av(n)n! = (n+1)!(H_{n+1} - 1).$$

From this, we see that the average number of cells that do not move is asymptotically equal to $n \log n$.

3 Fixed cells in permutations

Consider performing the push sequence from left-to-right until an increasing permutation is obtained. Any cell in the grid that is filled in both the original and final permutation will be termed a *fixed cell*. We illustrate this in the following example: The permutation 2174356 is pushed, resulting in the identity permutation of length 7. The 23 fixed cells are shaded in Figure 4 below.

By comparing the original arbitrary permutation with the pushed identity permutation, we notice that for a particular column of size r in position k of the original permutation, the number of fixed cells contributed by this column is the minimum of r and k. This is because after the pushing process, the part of size k ends up in position k and so the number of fixed cells contributed is obtained by comparing r with k. Hence, considering



Original permutation   Pushed permutation

Figure 4: The fixed cells of the permutation 2174356

all possible parts r in the k-th position, the contribution to the number of fixed cells from this position is given by

$$(n-1)!\left(\sum_{r=1}^{k} r + \sum_{r=k+1}^{n} k\right).$$

So the total contribution over all positions k is

$$(n-1)!\sum_{k=1}^{n}\left(\sum_{r=1}^{k} r + \sum_{r=k+1}^{n} k\right)$$

$$= (n-1)!\sum_{k=1}^{n}\left(\binom{k}{2} + (n-k+1)k\right) = (n-1)!\frac{n}{6}(n+1)(2n+1).$$

Let $Av(n)$ be the average number of fixed cells over all members of $\mathcal{S}_n$. From the previous equation, we have the following result.

Theorem 4. *The average number of fixed cells in a permutation of* $[n]$ *is*

$$Av(n) = \frac{1}{6}(n+1)(2n+1).$$

Asymptotically, the number of fixed cells on average is $\frac{n^2}{3}$ compared to $\frac{n^2}{2}$ cells in total. Thus, the proportion tends to $\frac{2}{3}$, i.e., two thirds of the cells are fixed.

The preceding may be generalized to members of $\mathcal{S}_{n,m}$. In this case, the reordered multipermutation with which we compare an arbitrary permutation is $1^m 2^m \cdots n^m$. For each part k in the final sorted multipermutation, the number of fixed cells is found by comparison of k with the original number i of cells in the position now occupied by the k in question. If $i < k$, then the position corresponding to this k contributes i fixed cells, whereas if $i \geq k$, then the contribution is k. For example, the multipermutation 322131 has 9 fixed cells which are shaded in the figure below.

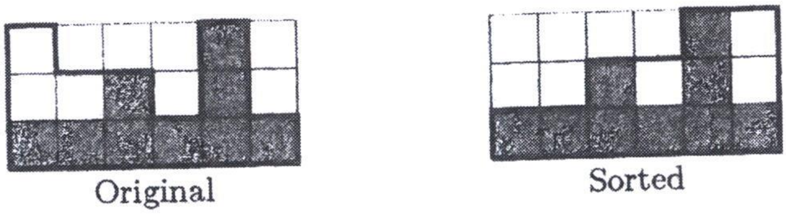


Figure 5: The permutation 322131 with its fixed cells

For each such i, the number of permutations of the remaining $nm-1$ positions is given by $\binom{nm-1}{m,\ldots,m,m-1}$, where the multinomial coefficient has $n-1$ copies of m. This evaluates to $A := \frac{(nm-1)!}{(m!)^{n-1}(m-1)!}$. Hence, the total number of fixed cells is

$$\sum_{k=1}^{n} mA(\sum_{r=1}^{k} r + \sum_{r=k+1}^{n} k) = m\frac{(nm-1)!}{(m!)^{n-1}(m-1)!}\frac{n}{6}(n+1)(2n+1).$$

The factor m arises because there are m copies of each letter k, and hence m positions to consider. Again, letting $Av(n)$ denote the average number of fixed cells over all members of $\mathcal{S}_{n,m}$, we obtain

$$Av(n) = \frac{m}{6}(n+1)(2n+1).$$

4 Frictionless pushes in multipermutations

A *frictionless push* is counted from the bottom of the multipermutation upward and within each row from left to right. It is defined recursively as follows: The leftmost group of adjacent 1's (single or multiple) in the bottom row count one frictionless push constituted by moving the maximal block consisting of parts ≥ 2 from the left of these 1's to their immediate right. This is then repeated until all 1's are at the beginning of the multipermutation. Then the same procedure is applied to the second row, i.e., for multiple or single 2's, and so on, recursively, to all other rows. The procedure ends once the multipermutation is in sorted order.

Note that the statistic recording the number of frictionless pushes differs from the push statistic as defined in the introduction when applied to members of $\mathcal{S}_{n,m}$ where $m > 1$. Though the two statistics coincide when $m = 1$, the order in which the various elements are shifted forward is different. If $m > 1$, then a member of $\mathcal{S}_{n,m}$ will have strictly more pushes than it does frictionless pushes, unless each element that is shifted occurs singly (i.e., as a run of length one), in which case they are equal in number. Thus, counting frictionless pushes involves keeping track of certain runs of elements in a multipermutation whose elements are not weak left-to-right maxima. See, e.g., [1, 2, 8, 9] for other kinds of statistics on permutations related to runs or various types of maxima.

We define a generating function $F_n^m(q)$ which enumerates members of $\mathcal{S}_{n,m}$ according to the number of frictionless pushes required to produce sorted order. Note that the distribution for frictionless pushes on multipermutations may also be obtained, equivalently, by first considering the groups of columns of height n and shifting them, if necessary, so that they all occur at the end, and then considering columns of height $n-1$ and shifting them, and so on for columns of lesser height, until sorted order is achieved. To realize this equivalence, consider applying to multipermutations the reverse complement operation.

We now find a recursion for $F_n^m(q)$ as follows where $m \geq 1$ is fixed. An arbitrary member of $\mathcal{S}_{n+1,m}$ is obtained from a precursor in $\mathcal{S}_{n,m}$ by inserting the m $(n+1)$'s in any of the positions of the precursor. To count these, we assume that j of the $(n+1)$'s where $0 \leq j \leq m$ are placed together in the rightmost position of the precursor. These are already in sorted order and therefore contribute no frictionless pushes. Let j_i denote the number of runs of $n+1$ of length i (in each run, precisely i $n+1$'s are inserted adjacent to one another somewhere within the precursor). Each run contributes one frictionless push, and the number of choices regarding the relative positions of the runs for each precursor is given by $\binom{nm}{j_1, j_2, \ldots, j_r, nm-\sum_{i=1}^r j_i}$. To see this, first consider the number j_1 of possible positions for the runs of $n+1$ of length one, then the number j_2 of positions for runs of $n+1$ of length two, and so on, up to m. We obtain the total number of choices regarding the positions of the runs as

$$\binom{nm}{j_1} \cdot \binom{nm-j_1}{j_2} \cdot \binom{nm-(j_1+j_2)}{j_3} \cdots \binom{nm-\sum_{i=1}^{r-1} j_i}{j_r}$$
$$= \binom{nm}{j_1, j_2, \ldots, j_r, nm-\sum_{i=1}^r j_i}.$$

Note, for instance, that once the positions for the j_1 runs of length one have been chosen, there are $(nm + j_1) - 2j_1 = nm - j_1$ possible positions for the j_2 runs of length two, which accounts for the $\binom{nm-j_1}{j_2}$ factor. This development is then repeated when considering the subsequent run lengths.

From the prior considerations, we thus obtain the following recursion:

$$F_{n+1}^m(q) = \left(1+\sum_{j=0}^{m-1}\sum_{\sum_{i\geq 1} ij_i=m-j}\binom{nm}{j_1,j_2,\ldots,j_r,nm-\sum_{i=1}^r j_i}q^{\sum_{i\geq 1} j_i}\right)F_n^m(q). \tag{4.1}$$

Iterating this equation yields

$$F_n^m(q) = \prod_{i=1}^{n-1}\left(1+\sum_{j=0}^{m-1}\sum_{\sum_{i\geq 1} ij_i=m-j}\binom{im}{j_1,j_2,\ldots,j_r,nm-\sum_{i=1}^r j_i}q^{\sum_{i\geq 1} j_i}\right), \tag{4.2}$$

where $F_1^m(q)=1$. When $q=1$, we have the identity

$$\binom{nm}{m,m,\ldots,m} = \prod_{i=1}^{n-1}\left(1+\sum_{j=0}^{m-1}\sum_{\sum_{i=1}^r ij_i=m-j}\binom{im}{j_1,j_2,\ldots,j_r,nm-\sum_{i=1}^r j_i}\right). \tag{4.3}$$

For later use, we define the factors in (4.2) by

$$f_\ell^m := 1+\sum_{j=0}^{m-1}\sum_{\sum_{i=1}^r ij_i=m-j}\binom{\ell m}{j_1,j_2,\ldots,j_r,\ell m-\sum_{i=1}^r j_i}q^{\sum_{i\geq 1} j_i}, \tag{4.4}$$

where $1\leq \ell\leq n-1$. For example, here is a calculation for the generating function F_n^4, i.e., for arbitrary n where $m=4$. In this case, the solutions to the equation

$$j_1+2j_2+\cdots+rj_r=4-j$$

are given in Table 1 below.

j					
3	$j_1 = 1$				
2	$j_1 = 2$	$j_2 = 1$			
1	$j_1 = 3$	$j_1 = 1, \quad j_2 = 1$	$j_3 = 1$		
0	$j_1 = 4$	$j_1 = 2, \quad j_2 = 1$	$j_1 = 1, \quad j_3 = 1$	$j_2 = 2$	$j_4 = 1$

Table 1: Possible values of j_i

Each j_i not specified in the solutions is taken to be zero. So (4.1) becomes:

$$F_{n+1}^4(q) = \left[1 + 4\binom{4n}{1}q + 2\binom{4n}{2}q^2 + \binom{4n}{3}q^3 + \binom{4n}{4}q^4 \right.$$
$$\left. +2\binom{4n}{1,1,4n-2}q^2 + \binom{4n}{2,1,4n-3}q^3\right] F_n^4(q).$$

Simplifying this and substituting into (4.2), we obtain

$$\begin{aligned} &F_n^4(q) \\ &= \prod_{i=1}^{n-1}\left(1 + 16iq - 12i(1-4i)q^2 + \frac{16}{3}i(1-2i)(1-4i)q^3\right. \\ &\qquad \left. -\frac{1}{3}i(1-2i)(3-4i)(1-4i)q^4\right). \end{aligned} \tag{4.5}$$

Checking equations (4.3) and (4.5), we insert $q = 1$ and $n = 3$ into the latter and obtain 34650. This is indeed $\binom{12}{4,4,4}$, as asserted by the former.

One can further compute $\frac{\left.\frac{df_i^4}{dq}\right|_{q=1}}{\left.f_i^4\right|_{q=1}} = \frac{4i}{1+i}$.

Lemma 1. *For each fixed* $m \geq 1$, $\frac{\left.\frac{df_i^m}{dq}\right|_{q=1}}{\left.f_i^m\right|_{q=1}} = \frac{mi}{1+i}$.

Proof. We begin by simplifying $\left.f_i^m\right|_{q=1}$. From (4.3) and (4.4),

$$\left.f_i^m\right|_{q=1} = \frac{\binom{(i+1)m}{m,\ldots,m}}{\binom{im}{m,\ldots,m}} = \frac{1}{m!}(mi+1)\cdots(mi+m). \tag{4.6}$$

Next, we provide a combinatorial interpretation of $\left.\frac{df_i^m}{dq}\right|_{q=1}$. We represent by f_i^m the number of ways of positioning the m $(i+1)$'s when constructing an arbitrary member of $\mathcal{S}_{i+1,m}$ from one in $\mathcal{S}_{i,m}$, where if a run of $i+1$ is placed in some position other than the last, this produces a single frictionless push marked by the variable q. Then $\frac{df_i^m}{dq}\big|_{q=1}$ is the total number of frictionless pushes and thus

$$\frac{df_i^m}{dq}\bigg|_{q=1} = \sum_{j=1}^{m} \binom{mi}{j} j \binom{m}{j} = m \sum_{j=1}^{m} \binom{mi}{j} \binom{m-1}{j-1}. \tag{4.7}$$

To see this, note that the first binomial coefficient in the first sum represents the number of ways of choosing the positions for j distinct runs of $i+1$ (not counting a possible terminal run of $i+1$), with the factor of j accounting for the number of frictionless pushes which result from this choice. The second binomial coefficient $\binom{m}{j}$ then gives the number of ways that the m $(i+1)$'s may be split into $j+1$ groups where each of the first j groups is nonempty. By (4.6) and (4.7), our lemma is true if and only if

$$m(1+i) \sum_{j=1}^{m} \binom{mi}{j} \binom{m-1}{j-1} = \frac{1}{m!}(mi)(mi+1)\cdots(mi+m),$$

or equivalently

$$\begin{aligned} \sum_{j=1}^{m} \binom{mi}{j} \binom{m-1}{j-1} &= \frac{1}{m!}(mi)(mi+1)\cdots(mi+m-1) \\ &= \binom{mi+m-1}{m}, \end{aligned}$$

which holds by Vandermonde's identity (see, e.g., [5, Formula 5.23]), completing the proof. □

Theorem 5. *The average number of frictionless pushes in multipermutations of* $\{1^m, 2^m, \ldots, n^m\}$ *is* $m(n - H_n)$.

Proof. Differentiating both sides of (4.2) with respect to q, setting $q = 1$ and using Lemma 1, we have

$$\begin{aligned} \frac{dF_n^m(q)}{dq} &= \binom{nm}{m, \ldots, m} m \sum_{i=0}^{n-1} \frac{i}{1+i} = \binom{nm}{m, \ldots, m} m \sum_{i=0}^{n-1} \left(1 - \frac{1}{1+i}\right) \\ &= \binom{nm}{m, \ldots, m} m(n - H_n). \end{aligned}$$

Dividing by $\binom{nm}{m,\ldots,m}$ yields the desired result. □

We conclude this section by providing a bijective proof of the prior result.

Combinatorial proof of Theorem 5.

We first count the frictionless pushes where it is assumed that the m copies of each letter $i \in [n]$ are distinct (which we denote by $i_1, i_2, \ldots, i_m$). Assume that $i_a < j_b$ if $i < j$ and that the i_a for a given i are equal when comparing the entries of a multipermutation. Then there are $(mn)(mn)!$ letters in all of multiset permutations π of $\{1^m, 2^m, \ldots, n^m\}$ where the letters are considered distinct from which we will subtract a certain set of letters to obtain the stated result. By a *block ender*, we will mean the final letter of a run (i.e., block) of the same element which is moved in some frictionless push when π is reordered (where here we start by moving blocks involving the letter n, if needed, and then $n-1$, and so on). That is, a block ender is the last letter of some block for which there is at least one element to the right of the block that is smaller than the elements in the block. Note that the total number of block enders within all the permutations is the same as the total number of frictionless pushes. Other letters will be referred to as *non block enders*. So we will count the total number of non block enders and subtract from the total number of letters to obtain the number of frictionless pushes. Within a given multipermutation π of $\{1^m, 2^m, \ldots, n^m\}$, let $S = S_i$ denote the subsequence (of length mi) comprising copies of all letters in $[i]$ where $1 \le i \le n$. Fix a copy i_a of the letter i, where $1 \le a \le m$. Note that i_a is a non block ender of π if and only if either (a) i_a is the rightmost letter of S, or (b) i_a is not the final letter of its run.

We use basic probability to compute the number of π for which either (a) or (b) holds. The probability that (a) holds for a randomly chosen permutation π is seen to be $\frac{1}{mi}$ and thus there are $\frac{1}{mi}(mn)!$ possibilities. In order for (b) to occur, it must be the case that i_a is not the last letter of S and, given this fact, the next letter in S is another copy of i. Upon conditioning on the first of these events, the probability that (b) holds for a randomly selected permutation is $\left(\frac{mi-1}{mi}\right)\left(\frac{m-1}{mi-1}\right) = \frac{m-1}{mi}$, and thus there are $\left(\frac{m-1}{mi}\right)(mn)!$ such permutations. Combining (a) and (b) yields $\left(\frac{1}{mi} + \frac{m-1}{mi}\right)(mn)! = \frac{1}{i}(mn)!$ possible permutations π, and thus the number of times i_a occurs as a non block ender. By symmetry, the same is true of any copy of i giving $\frac{m}{i}(mn)!$ non block enders involving a copy of i. Summing over $1 \le i \le n$ yields $\sum_{i=1}^{n} \frac{m}{i}(mn)! = mH_n(mn)!$ non block enders in total. The desired result now follows from subtracting from $(mn)(mn)!$, and dividing by $(m!)^n$ (since copies of letters are not to be distinct). □

It is also possible to give a combinatorial proof of Theorem 2 above, which provides such a proof as well for the comparable result from [10] on weak left-to-right maxima.

We first count weak left-to-right maxima (wlrm) in multipermutations where it is assumed once again that the m copies of each letter $i \in [n]$ are distinct (which we will denote by $i_1, i_2, \ldots, i_m$). Note that if i_j for some $j \in [m]$ corresponds to a wlrm, then it must occur to the left of all copies of elements in $[i+1, n]$. The probability that i_j occurs to the left of all elements in $[i+1, n]$ within a randomly chosen permutation is given by $\frac{1}{(n-i)m+1}$ (which can also be realized by cyclic rotation of i_j and the elements of $[i+1, n]$ within their positions). So there are $\frac{1}{(n-i)m+1}(nm)!$ wlrm involving i_j and thus $\frac{m}{(n-i)m+1}(nm)!$ wlrm involving any copy of i, by symmetry. Summing over i gives $\sum_{i=1}^{n} \frac{m}{(n-i)m+1}(nm)!$ wlrm in total, and dividing this by $(m!)^n$ gives all the wlrm in members of $\mathcal{S}_{n,m}$. Subtracting from $nm\binom{nm}{m,\ldots,m}$ yields the total number of pushes in members of $\mathcal{S}_{n,m}$, and dividing by $\binom{nm}{m,\ldots,m}$ implies the result. □

5 Push of greatest length

By the length of the longest push within $\pi \in \mathcal{S}_n$, we mean the greatest number of positions an individual element $i \in [n]$ must be moved when π is reordered. For example, if $\pi = 6514372 \in \mathcal{S}_7$, then we have the sequence of pushes

$$\pi = \underline{6}514372 \to \underline{56}14372 \to 1\underline{56}4372 \to 1\underline{456}372 \to 1\underline{3456}72 \to 1234567,$$

where the underlined entries in each step are those responsible for the push. Note that we have push lengths of one, two, two, three and five corresponding to the elements 5, 1, 4, 3 and 2, respectively. Thus, the longest push is of length five (achieved by the element 2). Clearly, it is possible for the maximal length to be achieved by more than one element. Let $a(n,i)$ denote the number of permutations of $[n]$ where the length of the longest push is i for $0 \le i \le n-1$. Note for example that $a(n,0) = 1$, since only the identity permutation is counted, whereas $a(n,n-1) = (n-1)!$ since all permutations of the form $\pi = \pi'1$ are counted.

From the definitions, it is seen that the statistics recording the length of the longest push and the longest frictionless push are the same. Thus, for convenience, to ascertain $a(n,i)$, we will consider the length of the longest frictionless push. As before, when discussing frictionless pushes, we will start by shifting n to the right, if necessary, and then work downwards recursively. Define the distribution polynomial $a(n;y) = \sum_{i=0}^{n-1} a(n,i)y^i$ for $n \geq 1$. The $a(n;y)$ satisfy the following recurrence formula.

Lemma 2. *If $n \geq 2$, then*

$$a(n;y) = \frac{a(n-1;y)}{1-y} + y\frac{d}{dy}a(n-1;y) - \frac{y^n}{1-y}(n-1)!, \qquad (5.1)$$

with $a(1;y) = 1$.

Proof. We start by writing a recurrence for the numbers $a(n,i)$. Suppose permutations of length n are obtained by inserting the element n within a precursor permutation of length $n-1$. Note that the number of positions required to move n (so that it occurs at the end of the permutation) does not affect the length j of the longest frictionless push of the precursor. Furthermore, observe that if $j < i$, then n must correspond to the longest frictionless push (and thus be inserted in the $(i+1)$-st position from the right), whereas if $j = i$, then n can be inserted in any of the $i+1$ rightmost positions. Combining the observations above, and noting that $j > i$ is not possible, yields the following recurrence for $n \geq 2$:

$$a(n,i) = \sum_{j=0}^{i-1} a(n-1,j) + (i+1)a(n-1,i), \qquad 1 \leq i \leq n-1, \qquad (5.2)$$

with $a(n,0) = 1$ and $a(n,n) = 0$ for all $n \geq 1$. Multiplying both sides of (5.2) by y^i, and summing over $1 \leq i \leq n-1$, yields

$$\begin{aligned} a(n;y) - 1 &= \sum_{i=1}^{n-1} y^i \sum_{j=0}^{i-1} a(n-1,j) + \sum_{i=1}^{n-1}(i+1)a(n-1,i)y^i \\ &= \sum_{j=0}^{n-2} a(n-1,j) \sum_{i=j+1}^{n-1} y^i + \sum_{i=0}^{n-2}(i+1)a(n-1,i)y^i - 1 \\ &= \sum_{j=0}^{n-2} a(n-1,j)\frac{y^{j+1}-y^n}{1-y} + \frac{d}{dy}(ya(n-1;y)) - 1. \end{aligned}$$

Formula (5.1) now follows from the last equality, upon noting

$$\sum_{j=0}^{n-2} a(n-1,j) = a(n-1;1) = (n-1)!.$$

□

Define the exponential generating function

$$F(x,y)=\sum_{n\geq 1} a(n;y)\frac{x^n}{n!}=\sum_{n\geq 1}\left(\sum_{i=0}^{n-1} a(n,i)y^i\right)\frac{x^n}{n!}.$$

Then multiplying both sides of (5.1) by $\frac{x^{n-1}}{(n-1)!}$, and summing over $n \geq 2$, implies $F(x,y)$ is a solution to the following linear first order pde:

$$\frac{\partial}{\partial x}F(x,y)-y\frac{\partial}{\partial y}F(x,y)=\frac{1-y-xy}{(1-y)(1-xy)}+\frac{F(x,y)}{1-y}, \qquad (5.3)$$

with $F(x,0)=e^x-1$. By direct calculation, one can obtain the following solution of (5.3).

Theorem 6. *We have*

$$F(x,y)=(y-1)\int_0^x \frac{e^{x-t}(1-y(t+1)e^{x-t})}{(ye^{x-t}-1)^2(yte^{x-t}-1)}dt.$$

Let $F_i(x)=\sum_{n\geq i+1} a(n,i)\frac{x^n}{n!}$ for $i \geq 0$ be the exponential generating function for $a(n,i)$. Then $F_i(x)$ is the coefficient of y^i in $F(x,y)$. Thus, by Theorem 6, we have

$$\begin{aligned}F(x,y)&=(y-1)\int_0^x \frac{1-y(t+1)e^{x-t}}{(y-e^{t-x})(1-y(t+1)e^{x-t}+y^2te^{2(x-t)})}dt\\ &=\int_0^x \frac{t-e^{t-x}}{(yt-e^{t-x})(1-t)^2}dt\\ &\quad-\int_0^x\left(\frac{te^{t-x}(1-e^{t-x})}{(y-e^{t-x})^2(1-t)}+\frac{1+e^{t-x}(t^2-t-1)}{(y-e^{t-x})(1-t)^2}\right)dt.\end{aligned}$$

Therefore, we can state the following result.

Theorem 7. *If $i \geq 0$, then*

$$F_i(x)=\int_0^x\left(\frac{(1-te^{x-t})t^i}{(1-t)^2}+\frac{(i+1)t(1-e^{x-t})}{1-t}+\frac{e^{x-t}+t^2-t-1}{(1-t)^2}\right)e^{i(x-t)}dt.$$

From the previous theorem, one can obtain the following generating function formula concerning the first moment of the distribution $a(n;y)$:

$$\sum_{i\geq 0} iF_i(x)=\int_0^x \frac{((t+1)e^{x-t}-1)e^{x-t}}{(1-e^{x-t})^2(1-te^{x-t})}dt. \qquad (5.4)$$

Let α denote the longest push statistic on $\mathcal{S}_n$ and β the statistic on $\mathcal{S}_n$ given by $\beta(\rho) = \max\{\rho_i - i : 1 \leq i \leq n\}$ for $\rho = \rho_1\rho_2\cdots\rho_n$. By induction, upon considering the possible positions of n, one can show $\alpha(\rho) = \beta(\rho')$ for all $\rho \in \mathcal{S}_n$, where ρ' denotes the reverse complement of ρ. Thus α, being equally distributed to β, may be viewed as a kind of maximum deviation statistic. Note further that the sequence whose generating function is given by (5.4) corresponds to the sum total of either the α or β statistics on $\mathcal{S}_n$ and occurs as entry A018927 in the OEIS [11]. Using the formula from [11], we obtain the following result.

Theorem 8. *The average length of the longest push in a permutation of* $[n]$ *is given by*

$$\sum_{k=0}^{n-1} \frac{\frac{k}{(n-k)!}\left((k+1)^{n-k} - k^{n-k}\right)}{\binom{n}{k}}.$$

Corollary 1. *The average length of the longest push in a permutation of* $[n]$ *is also given by the simpler formula*

$$n - \frac{1}{n!}\sum_{k=1}^{n} k!k^{n-k}. \tag{5.5}$$

Asymptotically as $n \to \infty$, *this is* $n - \sqrt{\frac{\pi n}{2}} + \frac{2}{3} + O\left(\frac{1}{\sqrt{n}}\right)$.

Proof. We can rewrite the formula of Theorem 8 as

$$\frac{1}{n!}\sum_{k=0}^{n-1} kk!((k+1)^{n-k} - k^{n-k}).$$

Ignoring for the moment the factor $1/n!$, this is

$$\sum_{k=0}^{n-1}(k+1)(k+1)!(k+1)^{n-(k+1)} - \sum_{k=0}^{n-1}(k+1)!(k+1)^{n-(k+1)} - \sum_{k=1}^{n-1} kk!k^{n-k}.$$

Rewrite this as

$$\sum_{k=1}^{n} kk!k^{n-k} - \sum_{k=0}^{n-1} kk!k^{n-k} - \sum_{k=1}^{n} k!k^{n-k},$$

which gives (5.5) after division by $n!$. Now

$$\frac{1}{n!}\sum_{k=1}^{n} k!k^{n-k} = P(n),$$

where $P(n)$ is the function studied by Knuth in Chapter 1 of [7] defined as

$$P(n) := \frac{1}{n!} \sum_{k=0}^{n} (n-k)^k (n-k)!.$$

The asymptotic result follows from the asymptotics of $P(n)$ derived in Chapter 1 of [7]. □

One can consider the analogous statistic on multipermutations as follows. If π is a multiset permutation of $\{1^m, 2^m, \ldots, n^m\}$, then the length of the longest push of π is defined as the maximum number of positions a copy of some letter ℓ must be moved before it is ordered. Thus, it is the number of copies of letters that are larger than ℓ that ℓ must move past when ordered. Let $a_m(n,i)$ denote the number of members of $\mathcal{S}_{n,m}$ whose longest push has length i. Note that $a_m(n,i)$ is nonzero only for $0 \le i \le m(n-1)$. The $a_m(n,i)$ are determined recursively as follows.

Proposition 1. *If $n \ge 2$ and $m \ge 1$, then*

$$a_m(n,i) = \binom{i+m-1}{m-1} \sum_{j=0}^{i-1} a_m(n-1,j) + \binom{i+m}{m} a_m(n-1,i), \quad (5.6)$$

for $0 \le i \le m(n-2)$, with

$$a_m(n,i) = \binom{i+m-1}{m-1} \binom{m(n-1)}{m,\ldots,m}, \quad (5.7)$$

for $m(n-2)+1 \le i \le m(n-1)$.

Proof. To determine $a_m(n,i)$, one may consider equivalently the length of the longest frictionless push within members of $\mathcal{S}_{n,m}$ and again start by moving copies of n as needed. If $m(n-2)+1 \le i \le m(n-1)$, then a copy of n must be responsible for the longest frictionless push, in which case the leftmost n must be preceded by exactly $m(n-1)-i$ copies of letters in $[n-1]$. Thus, there are $\binom{i+m-1}{m-1}$ choices for the positions of the other n's and $\binom{m(n-1)}{m,\ldots,m}$ ways to arrange the letters less than n in their positions, which implies (5.7). If $0 \le i \le m(n-2)$, then consider the length j of the longest frictionless push in the precursor member of $\mathcal{S}_{n-1,m}$ into which m copies of n are to be inserted. Note that $j > i$ is not possible since inserting copies of a larger element cannot decrease the longest push length, whence $j \le i$. If $j < i$, then at least one n must be inserted into the $(i+1)$-st rightmost position of the precursor, with the other copies of n restricted to this position and those to its right. Thus, there are $\binom{i+m-1}{m-1}$ ways in which to insert the n's into the precursor and considering all $0 \le j \le i-1$ gives the first part of (5.6). If $j = i$, then again the n's are confined to the rightmost $i+1$ positions of the precursor except now there is no requirement that an n be placed in the first of these positions as before. This gives $\binom{i+m}{m}a(n-1,i)$ additional multipermutations, which completes the proof of (5.6). □

We were unable to find an extension of Theorem 6 to $a_m(n,i)$ for general m and leave this as a challenge to the reader.

6 Conclusion

In this paper, we have explored various concepts related to pushes on permutations and multipermutations. Explicit formulas have been found for the total number of cells that are not moved in any step of the pushing process as well as for the number of cells that belong to both the original and final sorted permutations. The distribution for frictionless pushes on multipermutations was found and a simple formula obtained for the first moment. It is possible to extend some of the results above to an arbitrary multiset $\{1^{a_1}, 2^{a_2}, \ldots, n^{a_n}\}$, where $a_i \ge 0$ for all i. For example, extending the combinatorial proof given for Theorem 2 yields the following result.

Theorem 9. *The average number of pushes over all multipermutations of* $\{1^{a_1}, 2^{a_2}, \ldots, n^{a_n}\}$ *is given by*

$$a_1 + \cdots + a_n - \sum_{i=1}^{n} \frac{a_i}{a_{i+1} + \cdots + a_n + 1}.$$

See [10] for the comparable result concerning weak left-to-right maxima in an arbitrary multiset. It is also possible to generalize Theorem 4 above as follows.

Theorem 10. *The average number of fixed cells in a multipermutation of* $\{1^{a_1}, 2^{a_2}, \ldots, n^{a_n}\}$ *is given by*

$$\frac{\sum_{i=1}^{n} a_i \left(\sum_{j=1}^{i} j a_j + \sum_{j=i+1}^{n} i a_j\right)}{a_1 + \cdots + a_n}.$$

Note that the formula stated above for the average number of fixed cells on $\mathcal{S}_{n,m}$ corresponds to the special case $a_1 = \cdots = a_n = m$ of the prior theorem. The distribution that was found for frictionless pushes can also be extended to arbitrary multisets, though the general formula is more complicated. On the other hand, it appears to be more difficult to generalize Theorems 3, 6 and 7 above, which we leave as further problems to explore.

References

[1] J.-L. Baril and V. Vajnovszki, A permutation code preserving a double Eulerian bistatistic, *Discrete Appl. Math.* **224** (2016), 9–15.

[2] A. M. Baxter, Refining enumeration schemes to count according to permutation statistics, *Electron. J. Combin.* **21(2)** (2014), #P2.50.

[3] A. T. Benjamin and J. J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington DC, 2003.

[4] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, 1974.

[5] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Second Edition, Addison-Wesley, Boston, 1994.

[6] E. S. Key, On the number of records in an iid discrete sequence, *J. Theoret. Probab.* **18(1)** (2005), 99–108.

[7] D. E. Knuth, *The Art of Computer Programming: Volume 1*, Third Edition, Addison-Wesley, Boston, 1997.

[8] S.-M. Ma, An explicit formula for the number of permutations with a given number of alternating runs, *J. Combin. Theory Ser. A* **119(8)** (2012), 1660–1664.

[9] S.-M. Ma, Enumeration of permutations by number of alternating runs, *Discrete Math.* **313(18)** (2013), 1816–1822.

[10] A. Myers and H. Wilf, Left-to-right maxima in words and multiset permutations, *Israel J. Math.* **166** (2008), 167–183.

[11] N. J. Sloane, On-line Encyclopedia of Integer Sequences, published electronically at http://oeis.org, 2010.

[12] R. Stanley, *Enumerative Combinatorics, Vol. I*, Cambridge University Press, 1997.

[13] H. S. Wilf, On the outstanding elements of permutations, preprint, http://www.cis.upenn.edu/wilf, 1995.