Contributions to Discrete Mathematics

Volume 17, Number 1, Pages 56–69
ISSN 1715-0868

PUSHES IN WORDS—A PRIMITIVE SORTING ALGORITHM

M. ARCHIBALD, A. BLECHER, C. BRENNAN, A. KNOPFMACHER, AND T. MANSOUR

Abstract. We define the statistic of a push for words on an alphabet $[k]$ and use this to obtain a generating function measuring the degree to which an arbitrary word deviates from sorted order. Several subsidiary concepts are investigated: the number of cells that are not pushed, the number of already sorted columns, the number of cells that coincide before and after pushing, the fixed cells in words and finally, the frictionless pushes.

1. Introduction

In this paper, we investigate a statistic which measures the extent to which a given word deviates from its representation in nondecreasing sorted order.

An earlier such statistic is inversions in words as first studied by MacMahon, (see Andrews [1], Theorem 3.6). Other statistics on words have recently been studied as seen in [2, 3, 4, 6, 7, 8, 9, 10]. This paper provides another of the deviations from sorted order by exploring the parameter "number of pushes" for an arbitrary word over an alphabet $[k]$.

For an arbitrary word W over an alphabet $[k]$, let us define a *push* in the following way: suppose the leftmost element in the word which is not a weak left-to-right maximum occurs in position i and has height $v(i)$. A weak left-to-right maximum is a part which is greater than or equal to all parts to its left. We define a cell to be a single square in any part as illustrated in Figure 1. Shift all cells in W which are to the left of i and above the level of $v(i)$, one horizontal position to the right. We call this shifting process a *push* and illustrate the concept in Figure 2. Also, we give an example in

Received by the editors March 14, 2019, and in revised form July 23, 2021.

2000 *Mathematics Subject Classification.* Primary: 05A15, Secondary: 05A05.

Key words and phrases. words, generating function, sorting, asymptotics.

This material is based upon work supported by the National Research Foundation under grant numbers 89147, 86329 and 81021 for M. Archibald, C. Brennan and A. Knopfmacher respectively.

Figure 1 of an original word and the sorted order after the push sequence is completed.

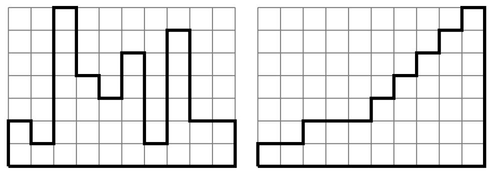

FIGURE 1. The sorted order of the word 2,1,7,4,3,5,1,6,2,2 resulting from the the pushing process.

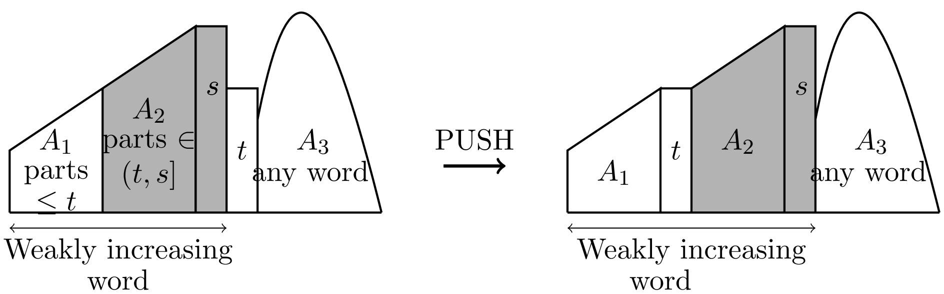

FIGURE 2. The pushing process illustrated.

We apply a sequence of pushes to a word and the successively pushed words that result. The sequence terminates once a weakly increasing word has been obtained. We obtain a generating function which tracks the number of pushes required for words of length n over an alphabet $[k]$.

Remark: We see that the total number of pushes required to generate a weakly increasing word is equal to the number of parts minus the number of weak left-to-right maxima.

For example, consider the word: **1,2**,1,**3,3**,2,**4**,2,1. It has 9 parts, 5 weak left-to-right maxima indicated in bold thus 4 pushes.

We also study associated concepts such as the generating function for the total number of cells which do not move in the pushing process (see section 3); the number of already sorted columns in arbitrary words (see section 4); the total number of cells that coincide before and after the full push sequence (see section 5); and finally a generating function for the so-called *frictionless pushes* (defined in section 6).

2. The number of pushes

Let $G_1(u,x)$ be the generating function for the number of pushes where x marks the length of the word and u the parts in the word that are not weak left-to-right maxima. The question of weak left-to-right maxima was previously studied in [11, 12]. For the convenience of the reader, we derive the result in full. By Remark 1.1 above, u also marks the number of pushes. We have the symbolic decomposition, see [13, 14],

$$\mathbf{1}^* \left(\mathbf{2}(<2)^*\right)^* \left(\mathbf{3}(<3)^*\right)^* \cdots \left(\boldsymbol{k}(<k)^*\right)^* \tag{2.1}$$

for all words in which all possible weak left-to-right maxima are shown in bold. Here $\boldsymbol{i}(<i)^*$ stands for a single element i followed by a sequence which may be empty that consist of parts in the set $\{1,2,\ldots,i-1\}$ with generating function

$$\frac{x}{1-u(i-1)x},$$

and thus $(\boldsymbol{i}(<i)^*)^*$ has the generating function

$$\frac{1}{1-\frac{x}{1-u(i-1)x}}.$$

Using the symbolic decomposition (2.1) we have

$$G_1(u,x) = \prod_{i=1}^{k} \frac{1}{1-\frac{x}{1-u(i-1)x}} = \prod_{i=1}^{k} \frac{1-(i-1)ux}{1-(i-1)ux-x}.$$

In order to find the total number of pushes for words of length n we compute

$$\begin{aligned}
\frac{\partial}{\partial u} G_1(u,x)\bigg|_{u=1} &= \frac{1}{1-kx}\times \\
&\quad \sum_{j=1}^{k} \frac{1-(j-1)ux-x}{1-(j-1)ux} \frac{\partial}{\partial u}\left(\frac{1-(j-1)ux}{1-(j-1)ux-x}\right)\bigg|_{u=1} \\
&= \frac{1}{1-kx} \sum_{j=1}^{k} \frac{1-jx}{1-(j-1)x} \frac{(j-1)x^2}{(1-jx)^2} \\
&= \frac{1}{1-kx} \sum_{j=1}^{k} \frac{(j-1)x^2}{(1-(j-1)x)(1-jx)}.
\end{aligned} \tag{2.2}$$

Partial fractions for the case $j=k$ in equation (2.2) gives

$$\frac{-1+k}{k(1-kx)^2} + \frac{(1-k)(1+k)}{k(1-kx)} - \frac{1-k}{1+y-kx}.$$

Extracting the coefficient $[x^n]$ gives

$$(-1+k)^{1+n} - (-1+k)\,k^{-1+n}\,(k-n).$$

Partial fractions for $1 \le j < k$ in equation (2.2) gives

$$-\frac{1-j}{(1-j+k)(1+(1-j)x)} + \frac{(1-j)}{(k-j)(1-jx)} + \frac{1-j}{(1-j+k)(j-k)(1-kx)}.$$

Extracting the coefficient $[x^n]$ gives

$$\frac{(1-j)(-(-1+j)^n j + j^{1+n} + (-1+j)^n k + k^n - j^n(1+k))}{(1-j+k)(j-k)}.$$

Replacing j by $k-i$ leads to

$$-(-1-i+k)^n + (-i+k)^n - \frac{k(k^n - (-1-i+k)^n)}{1+i} + \frac{(-1+k)(k^n - (-i+k)^n}{i}.$$

Thus the total number of pushes for all words of length n is

$$\begin{aligned}&(-1+k)^{1+n} + (1-k)k^{-1+n}(k-n)\\&+\sum_{i=1}^{k-1}\Bigg(-(-1-i+k)^n + (-i+k)^n\\&\qquad -\frac{k(k^n - (-1-i+k)^n)}{1+i} + \frac{(-1+k)(k^n - (-i+k)^n)}{i}\Bigg).\end{aligned}$$

This simplifies yielding the following result:

Theorem 2.1. *The total number of pushes in all words of length n over the alphabet $[k]$ is equal to*

$$k^n\left(\frac{(k-1)n}{k} - H_{k-1}\right) + \sum_{i=1}^{k-1}\frac{(k-i)^n}{i},$$

where $H_k = \sum_{j=1}^{k} 1/j$ is the Harmonic number.

For the average we divide by k^n to find

Corollary 2.2. *The average number of pushes in all words of length n over the alphabet $[k]$ is equal to*

$$\frac{n(k-1)}{k} - H_{k-1} + O\left(\left(\frac{k-1}{k}\right)^n\right), \text{ as } n \to \infty.$$

3. Number of cells that do not move

In this section, we calculate the total number of cells that do not move in the pushing process over all words of length n on alphabet $[k]$.

As an example, consider the word 61552142654 drawn in Figure 3. In this case, there are 22 cells that do not move in the pushing process. They are indicated with dots in the diagram.

It can be seen that the cells that do not move during the pushing process will necessarily form a weakly increasing word, and this is the basis of the proof. We divide this word into blocks ending in a strict right-to-left minimum. In the example in Figure 3, there are three strict right-to-left minima

(4, 2 and 1), indicated in bold. So there are three blocks (of length 3, 2 and 6 respectively), giving a total of $4 \times 3 + 2 \times 2 + 1 \times 6 = 22$.

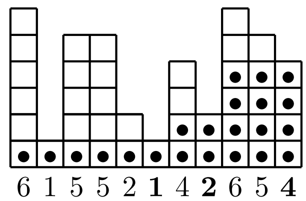

FIGURE 3. Word 61552142654 indicating cells not pushed.

Theorem 3.1. *Given the set of words of length n over alphabet $[k]$, the total number of cells that do not move during the pushing process is*

$$(n+1)k^n + (k-1)^{n+1} - k^{n+1}$$
$$- \sum_{m=2}^{k} \frac{m(k-m+1)^{n+1} - (m-1)(k-m)^{n+1} - k^{n+1}}{m-1}.$$

Asymptotically, the average number of cells that do not move during the pushing process is

$$n + 1 + k(H_{k-1} - 1) + O\left(\left(\frac{k-1}{k}\right)^n\right), \text{ as } n \to \infty.$$

Proof. To prove Theorem 3.1, decompose each word into blocks according to the size of the strict right-to-left minima. In each block, the number of cells that do not move are these left-to-right minima values multiplied by the number of parts in that block, as shown in Figure 3. A schema of the decomposition is given in Figure 4.

If there is a strict right-to-left minimum of size m, then its block $\mathcal{B}$ (see Figure 4), consists of a (possibly empty) sequence of letters of size between m and k, where the rightmost letter has size exactly m.

We let the variable z track the number of cells in the word and we let y track the number of columns (parts/letters) in the word. In addition, we introduce a variable u which will count the number of cells in the word that do not move. The generating function for the letter of size m (a strict right-to-left minimum) is $u^m y z^m$, and for any other letter is

$$\frac{u^m y(z^m - z^{k+1})}{1-z}.$$

The generating function for a block $\mathcal{B}$ is then

$$b_m(z, y, u) := \frac{u^m y z^m}{1 - \frac{u^m y(z^m - z^{k+1})}{1-z}}.$$

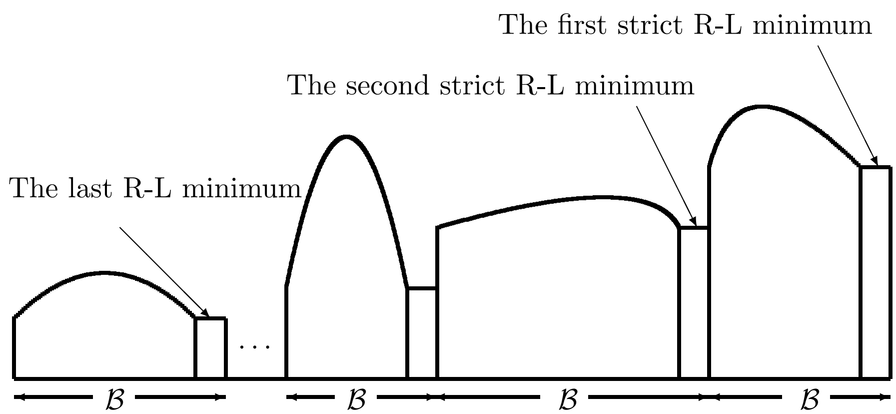

FIGURE 4. The decomposition schema for words indicating strict right-to-left minima.

For the generating function of the whole word, we have a sequence of these blocks (each with their corresponding right-to-left minimum), any of which may be empty. This gives an overall generating function of

$$G_2(z,y,u) := \prod_{m=1}^{k}\left(1+\frac{u^m y z^m}{1-\frac{u^m y(z^m-z^{k+1})}{1-z}}\right).$$

To calculate the total number of cells that do not move, we differentiate with respect to u and let $u=1$. This yields

$$\frac{\partial}{\partial u}G_2(z,y,u)\bigg|_{u=1}$$
$$=\sum_{m=1}^{k}\frac{my(1-z)^3z^m}{(yz^{k+1}-(y+1)z+1)(yz^m-yz^{k+1}+z-1)(yz^{m+1}-yz^{k+1}+z-1)}.$$

To count the cells that do not move, we no longer require the variable z, so we let $z\to 1$, and get

$$\frac{1}{1-ky}\sum_{m=1}^{k}\frac{my}{(1+(m-k-1)y)(1+my-ky)}.$$

The coefficient of y^n in the above two formulae gives the number of cells which do not move in a word of length n over alphabet $[k]$. For the term

$$\frac{y}{(1+(1-k)y)(1-ky)^2}$$

where $m=1$, the coefficient of y^n is

$$(n+1)k^n+(k-1)^{n+1}-k^{n+1}.$$

For $m \geq 2$, the coefficient of the term

$$\frac{1}{1-ky}\frac{my}{(1+(m-k-1)y)(1+my-ky)}$$

is

$$-\frac{m(k-m+1)^{n+1}-(m-1)(k-m)^{n+1}-k^{n+1}}{m-1}.$$

The final coefficient is

$$\begin{aligned}[y^n]&\frac{\partial G_2(1,y,u)}{\partial u}\bigg|_{u=1}\\ &= (n+1)k^n+(k-1)^{n+1}-k^{n+1}\\ &\quad -\sum_{m=2}^{k}\frac{m(k-m+1)^{n+1}-(m-1)(k-m)^{n+1}-k^{n+1}}{m-1}.\end{aligned}$$

To find the average we must divide this total by k^n. This concludes the proof of Theorem 3.1. □

4. The number of already sorted columns

We consider the number of nondecreasing columns to the left of the word prior to any pushing. For example, the word 1,1,4,2,5,4 has three such columns at the start of the word. These are the columns on the left that are already in sorted order. We have the following symbolic decomposition:

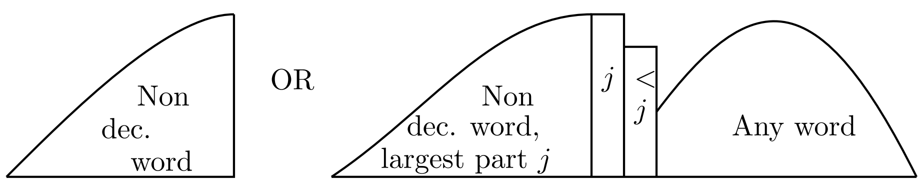

Figure 5. Decomposition of a word with maximum part of size k.

The generating function where v marks the number of initially sorted columns is

$$\begin{aligned}G_4(v,x) &= \frac{1}{(1-vx)^k}+\sum_{j=2}^{k}\frac{1}{(1-vx)^j}vx(j-1)x\frac{1}{1-kx}\\ &= \frac{1}{(1-vx)^k}+\frac{vx^2}{1-kx}\sum_{j=2}^{k}\frac{j-1}{(1-vx)^j}.\end{aligned}$$

For this we used the generating function for nondecreasing words, where the number of sequences $1 \leq a_1 \leq a_2 \leq \cdots \leq a_n \leq k$ is well known to be

$$\binom{n+k-1}{n} = [x^n]\frac{1}{(1-x)^k}.$$

Now

$$\frac{\partial}{\partial v}G_4(v,x)\Big|_{v=1} = \frac{(1-x)^{-k}}{1-kx} - \frac{1}{1-kx}.$$

Extracting $[x^n]$ gives

$$\sum_{j=0}^{n} k^j(-1)^{n-j}\binom{-k}{n-j} - k^n = \sum_{j=0}^{n} k^j \binom{n+k-j-1}{n-j} - k^n$$
$$\sim \left(\left(\frac{k}{k-1}\right)^k - 1\right)k^n, \text{ as } n \to \infty$$

by singularity analysis, see [5].

Hence, the average number of already sorted columns is asymptotic to

$$\left(\frac{k}{k-1}\right)^k - 1, \text{ as } n \to \infty.$$

Thus

Theorem 4.1. *The number of already sorted columns in words of length n is equal to*

$$\sum_{j=0}^{n} k^j \binom{n+k-j-1}{n-j} - k^n.$$

Moreover the average number of already sorted columns is asymptotic to

$$\left(\frac{k}{k-1}\right)^k - 1, \text{ as } n \to \infty.$$

Note that as $k \to \infty$, we have

$$\left(\frac{k}{k-1}\right)^k - 1 \to e - 1 = 1.71828\ldots.$$

5. Fixed cells in a word

Consider a word over the alphabet $[k]$ with its corresponding graphical representation where the columns of height r represent the parts of size r. We say that the column of size r consists of r unit squares called cells. We now perform the push sequence from left-to-right and obtain a nondecreasing word. Any cell in the grid that is filled in both the original word and the final pushed word is called a *fixed cell.* We illustrate this in the following example: The word 2,1,7,4,3,5,1,6,2,2 is pushed, and the new pushed nondecreasing word is 1,1,2,2,2,3,4,5,6,7. The 21 fixed cells are shaded.

Let $C(n)$ be the total number of fixed cells in words of size n over the alphabet $[k]$. We shall consider words of size $n+1$ over the same fixed alphabet that start with a part of size j and shall denote by $C(n+1|j)$ the total number of fixed cells in such words. We compare a word starting with j with a word which differs only in that it starts with $j+1$. The diagram represents the situation *after* it has been pushed to sorted order.

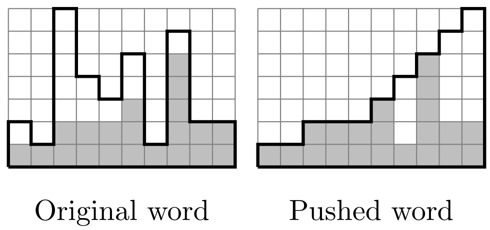

FIGURE 6. The fixed cells of the word 2,1,7,4,3,5,1,6,2,2.

In Figure 7 below, we show the two words after they have been pushed to sorted order. The subwords A_1 and A_2 are both weakly increasing subwords and the shaded block is in position s.

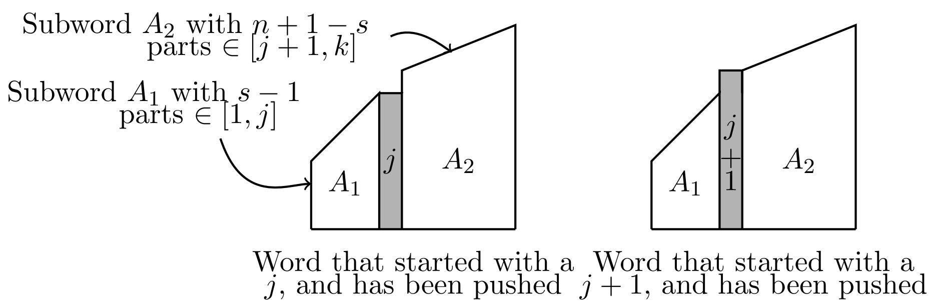

FIGURE 7. Comparison of two words that differed originally only in the first column, after they have been pushed.

Because we are aiming at the number of such words before they are pushed which contribute an extra fixed cell, we exclude the case where A_2 is empty. In such a case the j or $j+1$ is the largest letter in the word and changing this letter does not add an extra fixed cell. So the number of letters in A_1 (counted by $s-1$) has a maximum value of $n-1$. This accounts for the fact that the binomial coefficient in the sum below is $\binom{n-1}{s-1}$. The case $s=1$ (i.e., when $A_1=\emptyset$) will be handled in the next paragraph. We do a direct count of the number of words with the attributes described in Figure 7. For $s=2,\dots,n$ we have

$$\begin{aligned}\sum_{s=2}^{n}\binom{n-1}{s-1}j^{s-1}(k-j)^{n-(s-1)} &= \sum_{r=1}^{n-1}\binom{n-1}{r}j^{r}(k-j)^{n-r}\\ &= (k-j)k^{n-1}-(k-j)^{n}\end{aligned}\tag{5.1}$$

where the term $\binom{n-1}{s-1}j^{s-1}$ is for the subword A_1 where the parts are $\leq j$ before pushing, and the remaining term $(k-j)^{n-r}$ is for the subword A_2 where the parts are $\geq j+1$.

For $s = 1$ there are

$$(k-j)^n \tag{5.2}$$

such words. This is also the number of words which had $l \geq j+1$ in the s-th position in the original words (before pushing). Each of these contributes one extra fixed cell caused by the change in the shaded column. Thus the extra number of fixed cells obtained when adding 1 to the first part j, i.e., the difference between $C(n+1|j+1)$ and $C(n+1|j)$ is the sum of equations (5.1) and (5.2) which equals $(k-j)k^{n-1}$. Thus we have the recursion

$$C(n+1|j+1) = C(n+1|j) + (k-j)k^{n-1}, \tag{5.3}$$

with the following initial condition

$$C(n+1|1) = C(n) + k^n. \tag{5.4}$$

We use equation (5.3) with $j = 1$ and substitute equation (5.4) to obtain

$$C(n+1|2) = C(n) + k^n + (k-1)k^{n-1}. \tag{5.5}$$

We repeat the process, and use the expression for $C(n+1|2)$ to obtain

$$C(n+1|3) = C(n) + k^n + (k-1)k^{n-1} + (k-2)k^{n-2}. \tag{5.6}$$

Continuing the process until $j = k-1$, we get an expression for $C(n+1|k)$

$$C(n+1|k) = C(n) + k^n + k^{n-1}\sum_{i=1}^{k-1} i = C(n) + k^n\frac{k+1}{2}. \tag{5.7}$$

Thus summing all k cases (5.4), (5.5), (5.6), ...,(5.7) gives

$$\sum_{j=1}^{k} C(n+1|j) = kC(n) + k^{n+1} + k^{n-1}\sum_{i=1}^{k-1} i^2 = kC(n) + \frac{k^n(2k+1)(k+1))}{6}.$$

Since $C(n+1) = \sum_{j=1}^{k} C(n+1|j)$, we have the recursion

$$C(n+1) = kC(n) + \frac{k^n(2k+1)(k+1))}{6} \tag{5.8}$$

with

$$C(1) = \binom{k+1}{2} = \frac{k(k+1)}{2}.$$

Solving the recursion (5.8) yields the result

Theorem 5.1. *The total number of fixed cells in words of size n is*

$$C(n) = \frac{k^{n-1}(k+1)(2nk+n+k-1)}{6},$$

and the average number of fixed cells is

$$\frac{(k+1)(2k+1)n}{6k} + \frac{k^2-1}{6k}.$$

6. Frictionless pushes

A frictionless push differs from an ordinary push in that if two or more pushes occur in adjacent positions on the same level, they are thought of as constituting one frictionless push.

With this definition, weak left-to-right maxima in the word collectively contribute no frictionless pushes, and each maximal set of adjacent equal letters which are not left-to-right maxima contributes one frictionless push.

In order to track these, we decompose a word of length n over an alphabet $[k]$ to the first occurrence of each left-to-right maximum, m_i. Let $1 \leq m_1 < m_2 < m_3 < \cdots < m_s \leq k$ with $s \leq k$.

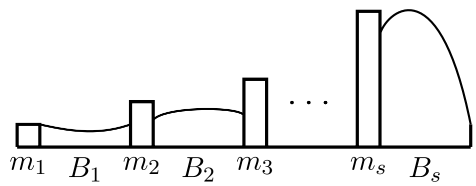

Figure 8. Decomposition for frictionless pushes

Then the subword B_i (possibly empty) has all parts less than or equal to m_i. The number of frictionless pushes (starting from the left and proceeding right) is the sum of the frictionless pushes in each subword $(m_i B_i)$. This is the product of the generating functions for each $m_i B_i$. Let $m_r = i$ be fixed. If B_r contains a nonempty sequence (indicated by $+$), $(< i)^+$, this has generating function

$$B_{(<i)^+}(u,x) := \frac{(i-1)ux}{1-x} \frac{1}{1 - \frac{(i-2)ux}{1-x}} := \frac{(i-1)ux}{1-x-(i-2)ux}, \tag{6.1}$$

where u marks the frictionless pushes from the left. From now on, to simplify the notation, we will write $B_{(<i)^+}$ instead of $B_{(<i)^+}(u,x)$. The first term is for the first repeated part and the second term is for a possibly empty sequence whose components are themselves maximal sequences of parts of a fixed size. If B_r contains a nonempty sequence

$$\left((< i)^+(i)^+\right)^+, \tag{6.2}$$

one such part $(< i)^+(i)^+$ has generating function

$$C_{(<i)^+} := B_{(<i)^+} \frac{x}{1-x}. \tag{6.3}$$

And so the sequence in (6.2) has generating function

$$D := \frac{C_{(<i)^+}}{1 - C_{(<i)^+}}.$$

This sequence can be preceded by a possible empty sequence $(i)^*$ and followed by a possibly empty sequence $\epsilon+(<i)^+$. Altogether these possibilities have generating function

$$\frac{1}{1-x}D(1+B_{(<i)^+}) \tag{6.4}$$

and this represents all possibilities that contain at least one term of the form $((<i)^+(i)^+)^+$.

Unaccounted for sequences that do not contain one term of the form $((<i)^+(i)^+)^+$ have generating function either

$$\frac{1}{1-x} \text{ or } B_{(<i)^+}$$

or

$$\frac{x}{1-x}B_{(<i)^+}.$$

On the other hand, when there is no left-to-right maximum i, this is represented by 1.

So all possibilities have generating function

$$\begin{aligned} &F(x,u)\\ &\quad= \prod_{i=1}^{k}\left[1+x\left(\frac{1}{1-x}D(1+B_{(<i)^+})+\frac{1}{1-x}+B_{(<i)^+}+\frac{x}{1-x}B_{(<i)^+}\right)\right]\\ &\quad= \prod_{i=1}^{k}\frac{1-x+(2-i)ux}{1-2x+(2-i)ux+(1-u)x^2}, \end{aligned} \tag{6.5}$$

which is our required generating function, stated in the theorem below.

Theorem 6.1. *The generating function for the number of frictionless pushes in words over alphabet* $[k]$ *where* x *marks the size of the word and* u *the number of frictionless pushes is*

$$F(x,u)=\prod_{i=1}^{k}\frac{1-x+(2-i)ux}{1-2x+(2-i)ux+(1-u)x^2}.$$

We now differentiate with respect to u and then set $u=1$ to obtain

$$\left.\frac{\partial F(x,u)}{\partial u}\right|_{u=1}=\underbrace{\prod_{i=1}^{k}\frac{1+(1-i)x}{1-ix}}_{1}\ \underbrace{\sum_{j=1}^{k}\frac{x^2(1-x)(j-1)}{(1-jx)(1+x-jx)}}_{2}.$$

The product labeled 1 simplifies to $1/(1-kx)$. The sum labeled 2 is expressed in terms of partial fractions as

$$\sum_{j=1}^{k}\frac{x^2(1-x)(j-1)}{(1-jx)(1+x-jx)}=(1-x)^2\sum_{j=1}^{k}\frac{1}{1-jx}-(1-x)\sum_{j=1}^{k}\frac{1}{1-(j-1)x}.$$

Thus

$$\left.\frac{\partial F(x,u)}{\partial u}\right|_{u=1} = \frac{1}{1-kx}\left((1-x)^2\sum_{j=1}^{k}\frac{1}{1-jx} - (1-x)\sum_{j=1}^{k}\frac{1}{1-(j-1)x}\right)$$
$$= \frac{1}{1-kx}\left(\left(\frac{1}{1-kx}-1\right)(1-x) + (x^2-x)\sum_{j=1}^{k}\frac{1}{1-jx}\right)$$
$$= \frac{kx(1-x)}{(1-kx)^2} - \frac{x(1-x)}{1-kx}\sum_{j=1}^{k}\frac{1}{1-jx}.$$

Now

$$[x^n]\left.\frac{\partial F(x,u)}{\partial u}\right|_{u=1}$$
$$= nk^n - (n-1)k^{n-1} - \sum_{j=1}^{k-1}\frac{j^n-k^n}{j-k} - nk^{n-1} + \sum_{j=1}^{k-1}\frac{j^{n-1}-k^{n-1}}{j-k}$$
$$+ k^{n-2}(n-1)$$
$$= nk^{n-1}(k-1) - k^{n-2}(n-1)(k-1) + \sum_{j=1}^{k-1}\frac{j^{n-1}-k^{n-1}}{j-k} - \sum_{j=1}^{k-1}\frac{j^n-k^n}{j-k}$$
$$= nk^{n-1}(k-1) - k^{n-2}(n-1)(k-1) - (k^n-k^{n-1})H_{k-1} + \sum_{j=1}^{k-1}\frac{j^n-j^{n-1}}{k-j}.$$

This leads to our final result:

Theorem 6.2. *The total number of frictionless pushes in words of size n over an alphabet $[k]$ is*

$$nk^{n-1}(k-1) - k^{n-2}(n-1)(k-1) - (k^n-k^{n-1})H_{k-1} + \sum_{j=1}^{k-1}\frac{j^n-j^{n-1}}{k-j}.$$

Hence, as $n\to\infty$, the asymptotic average number of frictionless pushes is

$$\frac{(k-1)^2 n}{k^2} - \frac{(1-k)(1-kH_{k-1})}{k^2} + O\left(\left(\frac{k-1}{k}\right)^n\right).$$

References

[1] G. Andrews, *The Theory of Partitions*, Cambridge University Press, Cambridge Mathematical Library, 1998.

[2] M. Archibald, A. Blecher, C. Brennan, A. Knopfmacher, and T. Mansour, *Shedding light on words*, Appl. Anal. Discrete Math. **11** (2017), 216–231.

[3] A. Blecher, C. Brennan, A. Knopfmacher, and T. Mansour, *The perimeter of words*, Discrete Math. **340** (2017), 2456–2465.

[4] A. Burstein, and T. Mansour, *Words restricted by patterns with at most 2 distinct letters*, Electron. J. Combin. **9:2**, (2002-3), ♯ R3.

[5] P. Flajolet, and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
[6] S. Heubach, and T. Mansour, *Combinatorics of Compositions and Words* Discrete Mathematics and its applications, CRC press, Taylor and Francis group, 2010.
[7] A. Knopfmacher, T. Mansour, and A. Munagi, *Smooth compositions and smooth words*, Pure Math. Appl. **22** (2011), 209–226.
[8] A. Knopfmacher, A. Munagi, and S. Wagner, *Successions in words and compositions*, Ann. Comb. **16** (2012), 277–287.
[9] S.-M. Ma, and T. Mansour, *Pattern restricted Stirling k-array words, plateau statistic, and kernel method*, Discrete Appl. Math. (2016), 100–108.
[10] T. Mansour, and M. Shattuck, *A statistic related to trees and words on a finite alphabet*, Discrete Mathematics, Algorithms and Applications **8:2** (2016), 1650031.
[11] A. Myers, and H. Wilf, *Left-to right maxima in words and multiset permutations*, Israel J. Math. **166** (2008), 167–183.
[12] H. Prodinger, *Combinatorics of geometrically distributed random variables: Left-to-right maxima*, Discrete Math. **153** (1996), 253–270.
[13] R. Stanley, *Enumerative combinatorics*, Wadsworth and Brooks/Cole, ISBN: 0-534-06546-5, Advanced books and software, Monterey, California. 1986.
[14] H. Wilf, *Generatingfunctionology*, ISB: 0-12-751956-4, 1994, Academia Press, INC. United Kingdom.

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
E-mail address: Margaret.Archibald@wits.ac.za

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
E-mail address: Aubrey.Blecher@wits.ac.za

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
E-mail address: Charlotte.Brennan@wits.ac.za

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
E-mail address: Arnold.Knopfmacher@wits.ac.za

Department of Mathematics, University of Haifa, 3498838 Haifa, Israel
E-mail address: tmansour@univ.haifa.ac.il