2007 Conference on Analysis of Algorithms, AofA 07 DMTCS proc. **AH**, 2007, 325–340

Expected values of statistics on permutation tableaux

Sylvie Corteel[1†] and Paweł Hitczenko[2‡]

[1] *LRI, CNRS Université Paris-Sud, 91405 Orsay, France, email:* `Sylvie.Corteel@lri.fr`
[2] *LRI, CNRS Université Paris-Sud, 91405 Orsay, France and Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA, email:* `phitczenko@math.drexel.edu`

received 15th Feb 2007, revised 23rd January 2008, accepted tomorrow.

Permutation tableaux are new objects that were introduced by Postnikov in the context of enumeration of the totally positive Grassmannian cells. They are known to be in bijection with permutations and recently, they have been connected to PASEP model used in statistical physics. Properties of permutation tableaux became a focus of a considerable research activity.

In this paper we study properties of basic statistics defined on permutation tableaux. We present a simple and unified approach based on probabilistic techniques and use it to compute the expected values of basic statistics defined on permutation tableaux. We also provide a non–bijective and very simple proof that there are $n!$ permutation tableaux of length n.

Keywords: permutation tableau, expected value, binomial distribution

1 Introduction

Permutation tableaux are new objects that are in bijection with permutations (Bur07; CN07; SW07). They come from the enumeration of the totally positive Grassmannian cells (Pos06; Wil05). Surprisingly they are also connected to a statistical physics model called the Partially ASymmetric Exclusion Process (PASEP) (Cor07; CW07; CW06). A *permutation tableau* (SW07) is a partition λ with non negative parts together with a filling of the boxes of the Ferrers diagram of λ with 0's and 1's such that the following properties hold:

1. Each column of the diagram contains at least one 1.
2. There is no 0 which has a 1 above it in the same column *and* a 1 to its left in the same row.

An example is given on Figure 1.

† supported in part by the NSF grant INT-0230800
‡ supported in part by the NSF grant INT-0230800 and by the National Security Agency under Grant H98230-05-1-0016.

1365–8050 © 2007 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

0	0	1	0	0	1	1
0	0	1	0	1		
0	1	1	1	1		
0	0	0				
1						

Fig. 1: Example of a permutation tableau

Different statistics on permutation tableaux were defined in (CW06; SW07). We list a few here. The *length* of a tableau is the sum of the number of parts and the largest part of the partition λ. A zero in a permutation tableau is *restricted* if there is a one above it in the same column. A row is *unrestricted* if it does not contain a restricted zero. A restricted zero is a *rightmost* restricted zero if it is restricted and it has no restricted zero to its right in the same row. A one is *superfluous* if it has a one above itself in the same column. For example the tableau of Figure 1 has length 12, 3 superfluous ones and 4 unrestricted rows.

The purpose of this work is to compute the expected value of several of these statistics. We list here our main results.

Theorem 1 *Let n be a fixed integer*

1. *The expected number of rows of a tableau of length n is $(n+1)/2$.*
2. *The expected number of unrestricted rows of a tableau of length n is H_n, the n^{th} harmonic number.*
3. *The expected number of ones in the first row of a tableau of length n is $H_n - 1$.*
4. *The expected number of superfluous ones of a tableau of length n is $(n-1)(n-2)/12$.*

The first three of these results are known as

1. The number of permutation tableaux of length n with k rows is equal to the number of permutation tableaux of length n with $n-k+1$ rows (CW06) or the number of permutation tableaux of length n with k rows is equal to the number of permutations of $[n] = \{1, 2, \ldots, n\}$ with k weak excedances (SW07).
2. The number of permutation tableaux of length n with k unrestricted rows is equal to the number of permutations of $[n]$ with k RL-minima (CN07; SW07). Moreover the number of permutations of n with k RL-minima is equal to the number of permutations of $[n]$ with k cycles and the expected number of cycles of a permutation of $[n]$ is H_n (Rio58).
3. The number of permutation tableaux of length n with k unrestricted rows is equal to the number of of permutation tableaux with $k-1$ ones in the first row (CW06).

These results come from bijections between permutation tableaux and permutations (CN07; SW07) and an involution on permutation tableaux (CW06). Nevertheless the combinatorics behind those results are not straightforward and all three required different methods. By contrast, in this paper we present an elementary approach that will enable us to derive all the results of Theorem 1 in a unified manner. The idea is to see how to construct a tableau of length n from a tableau of length $n-1$. This construction was already used in (CW06; Wil05) and is explained in Section 2. Here, we will reinterpret this construction in probabilistic language. This will allow us to use tools from probability theory to analyze properties of permutation tableaux. In particular, we provide a simple argument showing that the number of permutation tableaux of length n is $n!$. (The idea of extending the size of a tableau by one can also be translated into generating functions and gives another version of our proof of this enumeration result; we indicate it in Section 7.) All earlier proofs that we are aware of are based on establishing bijections between permutations and permutation tableaux. While there are at least three different bijections by now, none of them is particularly easy (of course, bijective proofs have their advantages as they give more information). Moreover thanks to this method, we can compute the expected number of superfluous ones of the tableaux of length n (item (4) of Theorem 1). This is a new and very interesting result. Indeed, using bijections between permutation tableaux and permutations (CN07; SW07), we know that the number of tableaux of length n with k superfluous ones is equal to the number of permutations of $[n]$ with k crossings and also to the number of permutations of $[n]$ with k occurrences of the generalized pattern 31-2. A crossing (Cor07) in a permutation $\sigma = (\sigma_1, \dots, \sigma_n)$ is a pair (i, j) such that $i < j \le \sigma_i < \sigma_j$ or $i > j > \sigma_i > \sigma_j$. This definition implies that the permutations with no crossings are in bijection with non-crossing partitions (Sim00) counted by the Catalan numbers. An occurrence of the generalized pattern 31-2 (BS00) in a permutation σ is a pair (i, j) with $1 < i < j$ and such that $\sigma_{i-1} > \sigma_j > \sigma_i$. It is also known that permutations with no occurrences of 31-2 are counted by the Catalan numbers. See (BS00; Cla01; CM05; CM02). Moreover it is shown in (Cor07) that permutations of $[n]$ with k crossings are in bijection with permutations of $[n]$ with k occurrences of 31-2. Therefore we also get a result on permutation enumeration:

Theorem 2 *The expected value of the number of crossings or occurrences of the generalized pattern 31-2 in a random permutation of* $[n]$ *is equal to*

$$\frac{(n-1)(n-2)}{12}.$$

Most of the up to date enumeration research related to generalized patterns focused on enumerating permutations that avoid a given generalized pattern (see e. g. (BS00; Cla01) and references therein). To our knowledge the above theorem is the first result of this type for generalized permutation patterns of length larger than two.

We start in Section 2 by some definitions and general observations and then use Sections 3-6 to prove Theorem 1.

2 From tableaux of length $n-1$ to tableaux of length n

Given a fixed permutation tableau T one can increase its length by one by inserting a new row or a new column. We use the example of Figure 1.

If one inserts a new row (a south step to the shape) then the number of unrestricted rows increases by one. For example, the tableau below has now 5 unrestricted rows:

If one inserts a new column (a west step to the shape), this column needs to be filled with zeros and ones. Every restricted row of T needs a zero in the new column, as there is a zero in this row that has a one above itself. The unrestricted rows can be filled with zeros or ones in the new column. If the tableau T had $U = U(T)$ unrestricted rows, then there are $2^U - 1$ fillings of the new column as this new column must contain a one. Some unrestricted rows become restricted if they are filled with a zero that has a one

0	0	1	0	0	1	1
0	0	1	0	1		
0	1	1	1	1		
0	0	0				
1						

above itself. Therefore the new tableau can have up to U unrestricted rows. On the example below the tableau of length 13 has now 3 unrestricted rows.

0	0	0	1	0	0	1	1
1	0	0	1	0	1		
0	0	1	1	1	1		
0	0	0	0				
1	1						

All together, there are 2^U possible extensions of T and since we will consider the uniform probability measure each of those extensions should be assigned the same probability, namely $1/2^U$. Hence, if $\mathsf{P}(S)$ is the probability of adding a row (South step) and $\mathsf{P}(W)$ the probability of adding a column (West step) then

$$\mathsf{P}(S) = \frac{1}{2^U}; \quad \mathsf{P}(W) = \frac{2^U - 1}{2^U}. \tag{1}$$

Let $\mathsf{P}(U \to k)$ be the probability of going from U to k, $1 \le k \le U+1$, unrestricted rows. If $k = U+1$, the only way to increase the number of unrestricted rows by one is by adding a row. Therefore

$$\mathsf{P}(U \to U+1) = \frac{1}{2^U}.$$

Now let $k \le U$. That means that the step was West and so

$$\mathsf{P}(U \to k) = \mathsf{P}(U \to k, W) = \mathsf{P}_W(U \to k)\mathsf{P}(W), \tag{2}$$

where we have written $\mathsf{P}_W(\,\cdot\,)$ for the conditional probability $\mathsf{P}(\,\cdot\,|W)$.

If one adds a column, this decreases (potentially) the number of unrestricted rows according to the following rule. First, all restricted rows stay restricted (and the corresponding entries in the new column are filled with 0's). Of the remaining U entries those that are assigned 0's which are below the top 1 will become restricted. Let G be the number (counting only unrestricted rows) of the topmost row that is assigned 1 and let V be the number of 1's assigned to the remaining $U - G$ unrestricted rows. Then, conditionally on G, V is distributed like $\mathrm{Bin}(U - G)$ random variable. Here and in what follows $\mathrm{Bin}(m)$ denotes a binomial random variable with parameters m and $p = 1/2$. Consequently,

$$\begin{aligned}
\mathsf{P}_W(U \to k) &= \mathsf{P}_W(G + V = k) = \sum_{j=1}^{k} \mathsf{P}_W(G = j,\ V = k - j) \\
&= \sum_{j=1}^{k} \mathsf{P}_W(V = k - j | G = j)\mathsf{P}_W(G = j) = \sum_{j=1}^{k} \frac{1}{2^{U-j}} \binom{U-j}{k-j} \mathsf{P}_W(G = j).
\end{aligned}$$

Since G is the waiting time for the first 1 in a $\mathrm{Bin}(U)$ random variable conditioned that it is at least 1, for $1 \le k \le U$ we have

$$\begin{aligned}\mathsf{P}_W(G=j) &= \mathsf{P}(j-1\ 0'\text{s followed by } 1|\mathrm{Bin}(U)\ge 1)\\ &= \frac{\mathsf{P}(j-1\ 0'\text{s followed by } 1,\ \mathrm{Bin}(U)\ge 1)}{\mathsf{P}(\mathrm{Bin}(U)\ge 1)}\\ &= \frac{1/2^j}{1-1/2^U} = \frac{2^{U-j}}{2^U-1}.\end{aligned}$$

Combining with (2) and (1),

$$\mathsf{P}(U\to k) = \frac{\mathsf{P}(W)}{2^U-1}\sum_{j=1}^{k}\binom{U-j}{k-j} = \frac{1}{2^U}\sum_{j=1}^{k}\binom{U-j}{k-j} = \frac{1}{2^U}\binom{U}{k-1}.$$

Therefore

$$\mathsf{P}(U\to k) = \frac{1}{2^U}\binom{U}{k-1}, \quad k=1,\dots,U+1. \tag{3}$$

We will now single out two simple observations that are key to our arguments. Let $\mathcal{T}_n$ be the set of all permutation tableaux of length n. We denote the uniform probability measure on $\mathcal{T}_n$ by P_n and E_n will denote the integration with respect to P_n over $\mathcal{T}_n$. Let U_n be a random variable which is equal to the number of unrestricted rows in a random (i.e. chosen according to P_n) tableau of length n.

We will be using conditional distribution and expectations given the random variables U_k, $k \ge 1$. Since U_k is a discrete random variable, a book by Shiryaev is a very good reference as it contains a section on conditioning with respect to what Shiryaev calls "decompositions" ((Shi96, Chapter 1, Sec. 8)) that specifically covers that case. We refer to Shiryaev's book for more details, here we briefly explain the decomposition of $\mathcal{T}_n$ that we will be using. One natural possibility would be to put in one element of a decomposition all tableaux of length n that were generated by the same tableau from $\mathcal{T}_{n-1}$. However, since our process is fully described by the number of unrestricted rows, when we refer in the sequel to the conditional distributions or expectations given U_{n-1} we will mean a coarser decomposition generated by U_{n-1}, namely

$$\mathcal{T}_n = \bigcup_{j=1}^{n-1}\mathcal{D}_j,$$

where $\mathcal{D}_j$'s are pairwise disjoint subsets of $\mathcal{T}_n$ defined by

$$\mathcal{D}_j = \left\{T\in\mathcal{T}_n :\ U_{n-1}(T_{|_{n-1}}) = j\right\},$$

and where by $T_{|_{n-1}}$ we mean a tableau in $\mathcal{T}_{n-1}$ that is obtained from T by removing the nth piece of its length (and the leftmost column if the nth step was to the west).

With this notation our first observation is as follows: $U_1 \equiv 1$ and for $n \ge 2$ the conditional distribution of U_n given U_{n-1} is

$$\mathcal{L}(U_n|U_{n-1}) = 1 + \mathrm{Bin}(U_{n-1}). \tag{4}$$

That is a direct consequence of (3) for $k = 1, \ldots, U+1$:

$$\mathsf{P}_n(U_n = k | U_{n-1}) \quad = \quad \frac{1}{2^{U_{n-1}}}\binom{U_{n-1}}{k-1} = \mathsf{P}(\mathrm{Bin}(U_{n-1}) = k-1).$$

For the second observation, note that P_n induces a measure on $\mathcal{T}_{n-1}$, namely each tableau $T_{n-1} \in \mathcal{T}_{n-1}$ is assigned a measure that is proportional to the number of tableaux from $\mathcal{T}_n$ it generates when its length is increased from $n-1$ to n. This measure is different than P_{n-1} since different tableaux from $\mathcal{T}_{n-1}$ generate different number of tableaux in $\mathcal{T}_n$. Specifically, T_{n-1} generates $2^{U_{n-1}(T_{n-1})}$ tableaux of length n where $U_{n-1}(T_{n-1})$ is the number of unrestricted rows of T_{n-1}. Hence, the relationship between these two measures is given by

$$\mathsf{P}_n(T_{n-1}) = \frac{2^{U_{n-1}(T_{n-1})}}{|\mathcal{T}_n|} = 2^{U_{n-1}(T_{n-1})}\frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}\frac{1}{|\mathcal{T}_{n-1}|} = 2^{U_{n-1}(T_{n-1})}\frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}\mathsf{P}_{n-1}(T_{n-1}).$$

Of course, we know bijectively that $|\mathcal{T}_n| = n!$ but we will not use it yet since we want to provide an independent proof of that fact. Before doing this one comment about the notation is in order. As we will be switching between these two measures on $\mathcal{T}_{n-1}$, we need to carefully distinguish which of the two we mean. We adopt the following convention: for any $n \geq 2$ and any random variable X_{n-1} defined on $\mathcal{T}_{n-1}$, $\mathsf{E}_{n-1}X_{n-1}$ will mean the integration with respect to the uniform measure on $\mathcal{T}_{n-1}$, while E_nX_{n-1} will mean the integration with respect to the measure induced by P_n. Thus for example, E_nU_n is the expected number of unrestricted rows in a random tableau of length n, while E_nU_{n-1} denotes the expected number of unrestricted rows in a tableau of length $n-1$ computed with respect to the measure induced by P_n. In particular, the above relation between the measures means that

$$\mathsf{E}_nX_{n-1} = \mathsf{E}_{n-1}\left(\frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}2^{U_{n-1}}X_{n-1}\right) = \frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}\mathsf{E}_{n-1}\left(2^{U_{n-1}}X_{n-1}\right). \tag{5}$$

This together with (4) is all we need to prove Theorem 1. But first, we wish to illustrate our method by providing a direct proof of

Proposition 3 *For every $n \geq 0$ $|\mathcal{T}_{n+1}| = (n+1)!$.*

Proof: The elements of $\mathcal{T}_{n+1}$ may be counted as follows

$$|\mathcal{T}_{n+1}| = \sum_{T \in \mathcal{T}_n} 2^{U_n(T)}.$$

Therefore,

$$|\mathcal{T}_{n+1}| = |\mathcal{T}_n|\frac{1}{|\mathcal{T}_n|}\sum_{T \in \mathcal{T}_n} 2^{U_n(T)} = |\mathcal{T}_n| \cdot \mathsf{E}_n 2^{U_n} = |\mathcal{T}_n| \cdot \mathsf{E}_n\mathsf{E}(2^{U_n}|U_{n-1}), \tag{6}$$

where $\mathsf{E}(\,\cdot\,|U_{n-1})$ is the conditional expectation given U_{n-1}. By (4) the conditional distribution of U_n given U_{n-1} is $1 + \mathrm{Bin}(U_{n-1})$ so that

$$\mathsf{E}(2^{U_n}|U_{n-1}) = \mathsf{E}(2^{1+\mathrm{Bin}(U_{n-1})}|U_{n-1}) = 2\mathsf{E}(2^{\mathrm{Bin}(U_{n-1})}|U_{n-1}) = 2\left(\frac{3}{2}\right)^{U_{n-1}},$$

where we have used an obvious fact that $\mathsf{E}a^{\mathrm{Bin}(m)} = \left(\frac{a+1}{2}\right)^m$. Hence by (5)

$$\mathsf{E}_n 2^{U_n} = 2\mathsf{E}_n \left(\frac{3}{2}\right)^{U_{n-1}} = 2\frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}\mathsf{E}_{n-1}2^{U_{n-1}}\left(\frac{3}{2}\right)^{U_{n-1}} = 2\frac{|\mathcal{T}_{n-1}|}{|\mathcal{T}_n|}\mathsf{E}_{n-1}3^{U_{n-1}}.$$

Putting this back into (6) we obtain

$$|\mathcal{T}_{n+1}| = |\mathcal{T}_n| \cdot \mathsf{E}_n 2^{U_n} = 2|\mathcal{T}_{n-1}| \cdot \mathsf{E}_{n-1}3^{U_{n-1}}.$$

This can be iterated and results in the following chain of equalities

$$\begin{aligned}
|\mathcal{T}_{n+1}| &= |\mathcal{T}_n| \cdot \mathsf{E}_n 2^{U_n} = 2 \cdot |\mathcal{T}_{n-1}| \cdot \mathsf{E}_{n-1}3^{U_{n-1}} = 2 \cdot 3 \cdot |\mathcal{T}_{n-2}| \cdot \mathsf{E}_{n-2}4^{U_{n-2}} \\
&= \ldots \\
&= 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot |\mathcal{T}_2| \cdot \mathsf{E}_2 n^{U_2} \\
&= 2 \cdot 3 \cdot \ldots \cdot n \cdot |\mathcal{T}_1| \cdot \mathsf{E}_1 (n+1)^{U_1} = (n+1)!
\end{aligned}$$

□

The above proof is typical for our arguments : we will condition, use (4), compute an expectation involving a function of a binomial random variable, and finally use (5) to reduce the length of the tableaux by one. In the next four sections we will further illustrate the method, by sketching the proofs of parts (1)–(4) of Theorem 1. Of course, now that Proposition 3 is proven, (5) simplifies to

$$\mathsf{E}_n X_{n-1} = \frac{1}{n}\mathsf{E}_{n-1}\left(2^{U_{n-1}} X_{n-1}\right) \tag{7}$$

and this is the form we will invoke in the sequel.

3 The expected number of unrestricted rows

In this section we will prove part (2) of Theorem 1. The relevant expectation involving a binomial random variable is isolated in the following lemma (observe that $b = 1$ and $c = 0$ gives the property used in the proof of Proposition 3).

Lemma 4 *For any numbers a, b, and c we have*

$$Ea^{\mathrm{Bin}(m)}(b + c\mathrm{Bin}(m)) = \left(\frac{a+1}{2}\right)^m (b + c\frac{a}{a+1}m). \tag{8}$$

Proof: Since $\mathrm{Bin}(m)$ is distributed like $\sum_{k=1}^m I_k$, where $I_1, \ldots, I_m$, are independent indicators with $\mathsf{P}(I_k = 1) = 1/2$, the left–hand side of (8) is

$$\mathsf{E}a^{\sum_{k=1}^m I_k}\left(b + c\sum_{k=1}^m I_k\right) = b\mathsf{E}a^{\sum_{k=1}^m I_k} + c\sum_{k=1}^m \mathsf{E}I_k a^{\sum_{j=1}^m I_j} = b\mathsf{E}\prod_{k=1}^m a^{I_k} + c\sum_{k=1}^m \mathsf{E}\left(I_k a^{I_k} \cdot a^{\sum_{j\neq k} I_j}\right).$$

By the independence of the indicators

$$\mathsf{E}\prod_{k=1}^m a^{I_k} = \prod_{k=1}^m \mathsf{E}a^{I_k} = \left(\frac{a+1}{2}\right)^m,$$

and, for $1 \leq k \leq m$,

$$\mathsf{E}\left(I_k a^{I_k} \cdot a^{\sum_{j\neq k} I_j}\right) = \left(\mathsf{E} I_k a^{I_k}\right) \mathsf{E} a^{\sum_{j\neq k} I_j} = \frac{a}{2}\left(\frac{a+1}{2}\right)^{m-1} = \frac{a}{a+1}\left(\frac{a+1}{2}\right)^{m}.$$

Putting the last two expressions in the earlier formula gives the right–hand side of (8). □

The proof of Theorem 1(2) can now be completed as follows:

$$\begin{aligned}\mathsf{E}_n U_n &= \mathsf{E}_n \mathsf{E}\left(U_n | U_{n-1}\right) = \mathsf{E}_n \mathsf{E}(1 + \mathrm{Bin}(U_{n-1}) | U_{n-1}) = \mathsf{E}_n(1 + \frac{U_{n-1}}{2}) \\ &= \frac{1}{n} E_{n-1} 2^{U_{n-1}}(1 + \frac{U_{n-1}}{2})\end{aligned}$$

where we have used (in that order) conditioning, (4), Lemma 4 (with $a = b = c = 1$), and (7). Applying the same procedure to this expression (this time with $a = 2$, $b = H_2$, $c = 1/2$) we get

$$\begin{aligned}\mathsf{E}_n U_n &= \frac{1}{n}\mathsf{E}_{n-1}\mathsf{E}\left(2^{1+\mathrm{Bin}(U_{n-2})}(1 + \frac{1+\mathrm{Bin}(U_{n-2})}{2})\big|U_{n-2}\right) \\ &= \frac{2}{n}\mathsf{E}_{n-1}\mathsf{E}\left(2^{\mathrm{Bin}(U_{n-2})}(H_2 + \frac{\mathrm{Bin}(U_{n-2})}{2})\big|U_{n-2}\right) \\ &= \frac{2}{n}\mathsf{E}_{n-1}\left(\frac{3}{2}\right)^{U_{n-2}}(H_2 + \frac{1}{2}\cdot\frac{2}{3}U_{n-2}) = \frac{2}{n(n-1)}\mathsf{E}_{n-2} 3^{U_{n-2}}(H_2 + \frac{1}{3}U_{n-2}).\end{aligned}$$

Further iterations lead to

$$\mathsf{E}_n U_n = \frac{(n-1)!}{n \cdot \ldots \cdot 2}\mathsf{E}_1 n^{U_1}(H_{n-1} + \frac{U_1}{n}) = H_{n-1} + \frac{1}{n} = H_n,$$

which completes the proof.

Remark: By modifying the above argument one can obtain a more general statement, namely, that the number of unrestricted rows U_n is distributed like

$$U_n \stackrel{d}{=} \sum_{k=1}^{n} J_k,$$

where $J_1, J_2, \ldots$, are independent indicators with $\mathsf{P}(J_k = 1) = 1/k$. In particular,

$$\mathsf{E}_n U_n = \sum_{k=1}^{n}\mathsf{P}(J_k) = H_n, \quad \mathsf{var}_n(U_n) = \sum_{k=1}^{n}\mathsf{var}(J_k) = \sum_{k=1}^{n}\frac{1}{k}\left(1 - \frac{1}{k}\right) = H_n - H_n^{(2)},$$

where $H_n^{(2)} = \sum_{k=1}^{n} k^{-2}$ is the generalized nth harmonic number. The details together with further results concerning more precise information about the other statistics on permutation tableaux will be included in the subsequent manuscript, now in preparation. Here, we just mention that this recovers the full strength of the second comment after Theorem 1 which in probabilistic language says that the number of unrestricted rows is distributed like the number of cycles in a random permutation. The latter distribution has been known to coincide with that of the sum $\sum_{k=1}^{n} J_k$ for a long time, see for example (Fel68, Chapter X, Section 6(b)) or (Rio58, Chapter 4, Section 3).

4 Expected number of rows

Let R_n be the number of rows in a random permutation tableau of length n. To prove part (1) of Theorem 1 we have to show that $\mathsf{E}_n R_n = \frac{n+1}{2}$.

Recall form the construction that rows of a permutation tableau correspond to south steps in the process of its construction. To be more precise, we let M_k indicate the direction of the move at the kth step (i.e. when the length is increased from $k-1$ to k). We refer to M_k as the kth move and will write $M_k = S$ or $M_k = W$ to indicate its direction. With this notation we have

$$R_n = \sum_{k=1}^{n} I_{M_k=S},$$

and the proof will be complete once we show

Proposition 5 *For* $1 \le k \le n$

$$\mathsf{E}_n I_{M_k=S} = \mathsf{P}_n(M_k = S) = 1 - \frac{k-1}{n}.$$

Proof: Let a be any number and consider more generally the expression $\mathsf{E}_m I_{M_k=S} a^{U_m}$ for $m \ge k$. Then

$$\mathsf{E}_m I_{M_k=S} a^{U_m} = \mathsf{E}_m \mathsf{E}(I_{M_k=S} a^{U_m} | U_{m-1}).$$

If $m > k$, then $k \le m-1$ and thus $I_{M_k=S}$ is measurable with respect to σ–algebra generated by U_{m-1}. Hence, using properties of conditional expectations and then (4), Lemma 4, and (7) we can further write

$$\begin{aligned}\mathsf{E}_m \mathsf{E}(I_{M_k=S} a^{U_m} | U_{m-1}) &= \mathsf{E}_m I_{M_k=S} \mathsf{E}(a^{U_m} | U_{m-1}) = a \mathsf{E}_m I_{M_k=S} \left(\frac{a+1}{2}\right)^{U_{m-1}} \\ &= \frac{a}{m} \mathsf{E}_{m-1} I_{M_k=S} (a+1)^{U_{m-1}}.\end{aligned}$$

If, on the other hand, $m = k$, then $M_k = S$ implies that $U_k = U_{k-1} + 1$ and thus

$$\begin{aligned}\mathsf{E}_k \mathsf{E}(I_{M_k=S} a^{U_k} | U_{k-1}) &= \mathsf{E}_k I_{M_k=S} \mathsf{E}(a^{1+U_{k-1}} | U_{k-1}) = a \mathsf{E}_k a^{U_{k-1}} \mathsf{E}(I_{M_k=S} | U_{k-1}) \\ &= a \mathsf{E}_k a^{U_{k-1}} \mathsf{P}(M_k = S | U_{k-1}) = a \mathsf{E}_k a^{U_{k-1}} \frac{1}{2^{U_{k-1}}} = \frac{a}{k} \mathsf{E}_{k-1} a^{U_{k-1}}.\end{aligned}$$

Both cases can be further iterated and if we begin with $m = n$ and $a = 1$ we obtain

$$\begin{aligned}\mathsf{E}_n I_{M_k=S} &= \frac{2 \cdot 3 \cdot \ldots \cdot (n-k)}{n(n-1) \cdot \ldots \cdot (k+1)} \mathsf{E}_k I_{M_k=S} (n-(k-1))^{U_k} \\ &= \frac{2 \cdot 3 \cdot \ldots \cdot (n-(k-1))}{n(n-1) \cdot \ldots \cdot (k+1)} \cdot \frac{n-(k-1)}{k} \mathsf{E}_{k-1} (n-(k-1))^{U_{k-1}} \\ &= \frac{(n-2)!(n-(k-1))}{n!} \mathsf{E}_1 (n-1)^{U_1} = \frac{(n-1)!(n-(k-1))}{n!} = 1 - \frac{k-1}{n},\end{aligned}$$

as claimed. □

5 The expected number of ones in the first row

Let F_n be the number of ones in the top row of a random permutation tableau of length n. Define G_k to be the position of the topmost (counting only unrestricted rows) 1 on the kth move. Note that G_k is undefined if this move is south. While it is inconsequential in this section as we are interested in the event $\{G_k = 1\}$, for the purpose of the next section it is convenient to set $G_k = U_{k-1} + 1$ if the kth move is south. Then we have

$$E_n F_n = \sum_{k=2}^{n} \mathsf{E}_n I_{G_k=1} = \sum_{k=2}^{n} \mathsf{P}_n(G_k = 1) = \sum_{k=2}^{n} \frac{1}{n-k+2} = H_n - 1,$$

where the third equality follows by iterating the following identity: If $n \geq m \geq k \geq 2$ then

$$\mathsf{E}_m I_{G_k=1} a^{U_m} = \begin{cases} \frac{a}{m} \mathsf{E}_{m-1} I_{G_k=1} (a+1)^{U_{m-1}}, & \text{if} \quad m > k, \\ \frac{1}{k} \frac{a}{a+1} \mathsf{E}_{k-1} a^{U_{k-1}} & \text{if} \quad m = k. \end{cases}$$

The relevant property of a binomial random variable is

$$\mathsf{E} I_{G=1} a^{G+\mathrm{Bin}(m-G)} = \frac{a}{a+1} \left(\frac{a+1}{2} \right)^m,$$

where G is the position of the first 1 in a $\mathrm{Bin}(m)$ random variable. We omit further details.

6 The expected number of superfluous ones

Let S_n be the number of superfluous ones in a random permutation tableau of length n. We write

$$S_n = \sum_{k=1}^{n} V_k,$$

where V_k is the increase in the number of superfluous ones on the kth move (thus, in particular, $V_k = 0$ if this move is south). Let G_k be as in the previous section. Then for $1 \leq j \leq U_{k-1}$ we have

$$\begin{aligned} \mathsf{P}_k(G_k = j | U_{k-1}) &= \mathsf{P}_k(G_k = j, M_k = W | U_{k-1}) = \mathsf{P}_k(G_k = j | M_k = W, U_{k-1}) \mathsf{P}_k(M_k = W | U_{k-1}) \\ &= \mathsf{P}(\mathrm{Bin}(U_{k-1}) = j | \mathrm{Bin}(U_{k-1}) \geq 1) \frac{2^{U_{k-1}} - 1}{2^{U_{k-1}}} = \frac{2^{-j}}{1 - 2^{-U_{k-1}}} \frac{2^{U_{k-1}} - 1}{2^{U_{k-1}}} = \frac{1}{2^j}, \end{aligned}$$

and, for $j = 1 + U_{k-1}$,

$$\mathsf{P}_k(G_k = U_{k-1} + 1 | U_{k-1}) = \frac{1}{2^{U_{k-1}}}.$$

Moreover, after some additional calculations, one finds that the joint conditional distribution of (U_k, V_k) given U_{k-1} is

$$\mathcal{L}((U_k, V_k) | U_{k-1}) = (G_k + \mathrm{Bin}(U_{k-1} - G_k), \mathrm{Bin}(U_{k-1} - G_k)).$$

We will need the following property of a binomial random variable.

Lemma 6 *G is the position of the first one in a binomial random variable. Then*

$$\begin{aligned} E\mathrm{Bin}(m-G) &= \frac{m}{2} - 1 + \frac{1}{2^m} \\ Ea^{G+\mathrm{Bin}(m-G)}\mathrm{Bin}(m-G) &= a^2\left(\left(\frac{m}{a+1}-1\right)\left(\frac{a+1}{2}\right)^m + \left(\frac{a}{2}\right)^m\right). \end{aligned}$$

Proof: For the first expression write

$$\begin{aligned} \mathsf{E}\mathrm{Bin}(m-G) &= \sum_{j=1}^{m-1}\frac{1}{2^j}\mathrm{Bin}(m-j) = \sum_{j=1}^{m-1}\frac{1}{2^j}\frac{m-j}{2} = \frac{1}{2}\sum_{k=1}^{m-1}\frac{k}{2^{m-k}} = 2^{-m}\sum_{k=1}^{m-1}k2^{k-1} \\ &= 2^{-m}\frac{d}{dx}\left(\sum_{k=0}^{m-1}x^k\right)\Big|_{x=2} = 2^{-m}(m2^{m-1}-2^m+1) = \frac{m}{2}-1+\frac{1}{2^m}, \end{aligned}$$

and for the second

$$\begin{aligned} &\mathsf{E}a^{G+\mathrm{Bin}(m-G)}\mathrm{Bin}(m-G) = \sum_{j=1}^{m-1}\frac{1}{2^j}a^j\mathsf{E}a^{\mathrm{Bin}(m-j)}\mathrm{Bin}(m-j) = \sum_{j=1}^{m-1}\left(\frac{a}{2}\right)^j\left(\frac{a+1}{2}\right)^{m-j}(m-j)\frac{a}{a+1} \\ &= \frac{a}{a+1}\sum_{k=1}^{m-1}k\left(\frac{a+1}{a}\right)^k\left(\frac{a}{2}\right)^{m-k} = \left(\frac{a}{2}\right)^m\sum_{k=1}^{m-1}k\left(\frac{a+1}{a}\right)^{k-1} = \left(\frac{a}{2}\right)^m\frac{d}{dx}\left(\sum_{k=0}^{m-1}x^k\right)\Big|_{x=\frac{a+1}{a}} \\ &= \left(\frac{a}{2}\right)^m a^2\left(m\left(\frac{a+1}{a}\right)^{m-1}\frac{1}{a} - 1\left(\frac{a+1}{a}\right)^m + 1\right) = a^2\left(\left(\frac{m}{a+1}-1\right)\left(\frac{a+1}{2}\right)^m + \left(\frac{a}{2}\right)^m\right). \end{aligned}$$

□

This, (4), and (7) is enough to establish the following lemma which contains all ingredients needed for the iteration

Lemma 7

$$\begin{aligned} E_kS_ka^{U_k} &= \frac{a}{k}E_{k-1}S_{k-1}(a+1)^{U_{k-1}} + \frac{a^2}{k}E_{k-1}\left(\left(\frac{U_{k-1}}{a+1}-1\right)(a+1)^{U_{k-1}} + a^{U_{k-1}}\right) \\ E_ka^{U_k}(b+cU_k) &= \frac{a}{k}E_{k-1}(a+1)^{U_{k-1}}\left(b+c+c\frac{a}{a+1}U_{k-1}\right). \end{aligned}$$

□

Starting with $k=n$ and $a=1$ and using successive conditioning and the above lemma, we obtain the following chain of equalities for E_nS_n.

$$\begin{aligned} \mathsf{E}_nS_n &= \frac{1}{n}\mathsf{E}_{n-1}S_{n-1}2^{U_{n-1}} + \frac{1}{n}\mathsf{E}_{n-1}\left(\left(\frac{U_{n-1}}{2}-1\right)2^{U_{n-1}}+1\right) \\ &= \frac{2}{n(n-1)}\mathsf{E}_{n-2}S_{n-2}3^{U_{n-2}} + \frac{2^2}{n(n-1)}\mathsf{E}_{n-2}\left(\left(\frac{U_{n-2}}{3}-1\right)3^{U_{n-1}}+2^{U_{n-2}}\right) \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{n}\mathsf{E}_{n-1}\left(\left(\frac{U_{n-1}}{2}-1\right)2^{U_{n-1}}+1\right)\\
=\;& \frac{2}{n(n-1)}\left\{\frac{3}{n-2}\mathsf{E}_{n-3}S_{n-3}4^{U_{n-3}}+\frac{3^2}{n-2}\mathsf{E}_{n-3}\left(\left(\frac{U_{n-3}}{4}-1\right)4^{U_{n-3}}+3^{U_{n-3}}\right)\right\}\\
&+\frac{2^2}{n(n-1)}\mathsf{E}_{n-2}\left(\left(\frac{U_{n-2}}{3}-1\right)3^{U_{n-1}}+2^{U_{n-2}}\right)+\frac{1}{n}\mathsf{E}_{n-1}\left(\left(\frac{U_{n-1}}{2}-1\right)2^{U_{n-1}}+1\right).
\end{aligned}
$$

Continuing in the same manner and using $\mathsf{E}_3 S_3 a^{U_3} = \frac{a^2}{6}$ in the last step we obtain

$$
\begin{aligned}
\mathsf{E}_n S_n \;=\;& \frac{(n-2)(n-2)!}{n!}\\
&+\sum_{j=3}^{n-1}\frac{(n-j)(n-j)!}{n(n-1)\dots(j+1)}\mathsf{E}_j\left(\left(\frac{U_j}{n+1-j}-1\right)(n+1-j)^{U_j}+(n-j)^{U_j}\right).
\end{aligned}
$$

Splitting the above expectation and successively applying the second part of Lemma 7 to each of the resulting parts (starting with $a=n+1-j$, $b=-1$, $c=1/(n+1-j)$, and $a=n-j$, and $b=1$, $c=0$, respectively) yields

$$
\begin{aligned}
\mathsf{E}_n S_n \;=\;& \frac{n-2}{n(n-1)}+\frac{(n-3)(n-2)}{2}(H_n-H_{n-3})\\
&+\frac{(n-4)(n-3)}{4}-\frac{(n-3)(n-2)}{2}+\frac{(n-3)(n-2)(2n-5)}{6n},
\end{aligned}
$$

which simplifies to $\frac{(n-1)(n-2)}{12}$. □

Remark. We could also have tried to attack this problem with techniques developed for labeled Motzkin paths (FPV86; Lou87) as it was shown in (Cor07) that permutations of $[n]$ with k crossings (or permutation tableaux of length n with k superfluous ones) are in bijection with certain labeled Motzkin paths of length n such that the sum of the labels is k. This method is easy to use for crossings for matchings but appears to be hard in this case (Lou06).

7 Conclusion

In this paper we used simple probabilistic techniques to get the expected values of the different parameters of interest of permutation tableaux of length n: the number of unrestricted rows, the number of ones in the first row, the number of rows and the number of superfluous ones. Our approach is very elementary and gives some exact as opposite to asymptotic expressions for the expected values of these parameters. Thanks to this approach we can also compute the variances and the distributions of those parameters and this will be reported in the long version of this work. To finish we want to highlight that as the method used here is just to build tableaux of length n from tableaux of length $n-1$, this can be translated into recurrences and generating functions. Indeed let $t_{n,k,j}$ be the number of tableaux of length n with k unrestricted rows and j ones in the first row. Then $t_{n,k,j}=0$ if $k>n$ or $j>n-1$ or $n<0$, $t_{1,1,0}=1$, and

$$
t_{n,k,j}=\sum_{\ell=k}^{n-1}\binom{\ell-1}{k-1}t_{n-1,\ell,j-1}+\sum_{\ell=k-1}^{n-1}\binom{\ell-1}{k-2}t_{n-1,\ell,j},
$$

otherwise. This implies that if $T_n(x, y) = \sum_{k,j} t_{n,k,j} x^k y^j$, then $T_1(x, y) = x$ and

$$T_n(x, y) = \frac{x(x+y)}{1+x} T_{n-1}(x+1, y).$$

Therefore for $n \geq 1$,

$$T_n(x, y) = x \prod_{i=0}^{n-2} (x + y + i).$$

This gives another straightforward proof that there are $n!$ permutation tableaux of length n and that permutation tableaux of length n and with k unrestricted rows (or $k-1$ ones in the first row) are in bijection with permutations of $\{1, 2, \dots n\}$ with k cycles. This generating function approach therefore looks very easy. Computations get way more complicated for superfluous ones as it was shown by Williams in (Wil05). This is why we chose to present in this extended abstract the details of the probabilistic approach that works well for all the parameters. Combinatorial studies showed the links between the permutation tableaux and the Partially ASymmetric Exclusion process (PASEP) (Cor07; CW07; CW06). As suggested by one of the referees, it would be very interesting to study the typical shape of tableaux not under the uniform distribution but under the weighted distribution that underlies the stationary distribution of the PASEP. This will be done in a later work.

Acknowledgements

The research of the second author was carried out while he was visiting the Algorithms and Complexity Group at the LRI, Université Paris-Sud in 2006-2007 academic year. He would like to express his thanks to Sylvie Corteel, Dominique Gouyou-Beauchamps and other members of the group for their support and hospitality.

References

[BCE+06] R. Brak, S. Corteel, J. Essam, R. Parviainen, and A. Rechnitzer. A combinatorial derivation of the PASEP stationary state. *Electron. J. Combin.*, 13(1):Research Paper 108, 23 pp. (electronic), 2006.

[BE04] R. Brak and J. W. Essam. Asymmetric exclusion model and weighted lattice paths. *J. Phys. A*, 37(14):4183–4217, 2004.

[BS00] E. Babson and E. Steingrímsson. Generalized permutation patterns and a classification of the Mahonian statistics. *Sém. Lothar. Combin.*, 44:Art. B44b, 18 pp. (electronic), 2000.

[Bur07] A. Burstein. On some properties of permutation tableaux, 2007. preprint.

[Cla01] A. Claesson. Generalized pattern avoidance. *European J. Combin.*, 22(7):961–971, 2001.

[CM02] A. Claesson and T. Mansour. Counting occurrences of a pattern of type $(1, 2)$ or $(2, 1)$ in permutations. *Adv. in Appl. Math.*, 29(2):293–310, 2002.

[CM05] A. Claesson and T. Mansour. Enumerating permutations avoiding a pair of Babson-Steingrímsson patterns. *Ars Combin.*, 77:17–31, 2005.

[CN07] S. Corteel and P. Nadeau. Permutation tableaux and permutation descents, 2007. preprint.

[Cor07] S. Corteel. Crossings and alignments of permutations. *Adv. in Appl. Math.*, 38(2):149–163, 2007.

[CW06] S. Corteel and L. K. Williams. A Markov chain on permutations which projects to the PASEP, 2006. math.CO/0609188.

[CW07] S. Corteel and L. K. Williams. Tableaux combinatorics for the asymmetric exclusion process. *Adv. in Appl. Math.*, 2007. to appear.

[DDM92] B. Derrida, E. Domany, and D. Mukamel. An exact solution of a one-dimensional asymmetric exclusion model with open boundaries. *J. Statist. Phys.*, 69(3-4):667–687, 1992.

[DEHP93] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A*, 26(7):1493–1517, 1993.

[DS05] E. Duchi and G. Schaeffer. A combinatorial approach to jumping particles. *J. Combin. Theory Ser. A*, 110(1):1–29, 2005.

[Fel68] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[FPV86] P. Flajolet, C. Puech, and J. Vuillemin. The analysis of simple list structures. *Inform. Sci.*, 38(2):121–146, 1986.

[Lou87] G. Louchard. Random walks, Gaussian processes and list structures. *Theoret. Comput. Sci.*, 53(1):99–124, 1987. Eleventh colloquium on trees in algebra and programming (Nice, 1986).

[Lou06] G. Louchard, 2006. Private communication.

[Pos06] A. Postnikov. Total positivity, Grassmannians, and networks, 2006. arXiv:math.CO/0609764.

[Rio58] J. Riordan. *An introduction to combinatorial analysis*. Wiley Publications in Mathematical Statistics. John Wiley & Sons Inc., New York, 1958.

[Sas99] T. Sasamoto. One-dimensional partially asymmetric simple exclusion process with open boundaries: orthogonal polynomials approach. *J. Phys. A*, 32(41):7109–7131, 1999.

[Shi96] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996.

[Sim00] R. Simion. Noncrossing partitions. *Discrete Math.*, 217(1-3):367–409, 2000. Formal power series and algebraic combinatorics (Vienna, 1997).

[SW07] E. Steingrímsson and L. K. Williams. Permutation tableaux and permutation patterns. *J. Combin. Theory, Ser. A*, 114(2):211–234, 2007.

[USW04] M. Uchiyama, T. Sasamoto, and M. Wadati. Asymmetric simple exclusion process with open boundaries and Askey-Wilson polynomials. *J. Phys. A*, 37(18):4985–5002, 2004.

[Wil05] L. K. Williams. Enumeration of totally positive Grassmann cells. *Adv. Math.*, 190(2):319–342, 2005.