FPSAC 2013 Paris, France DMTCS proc. **AS**, 2013, 731–742

Cycles and sorting index for matchings and restricted permutations

Svetlana Poznanović

Department of Mathematical Sciences, Clemson University, Clemson, SC 29634, USA

Abstract. We prove that the Mahonian-Stirling pairs of permutation statistics $(\mathrm{sor}, \mathrm{cyc})$ and $(\mathrm{inv}, \mathrm{rlmin})$ are equidistributed on the set of permutations that correspond to arrangements of n non-atacking rooks on a fixed Ferrers board with n rows and n columns. The proofs are combinatorial and use bijections between matchings and Dyck paths and a new statistic, sorting index for matchings, that we define. We also prove a refinement of this equidistribution result which describes the minimal elements in the permutation cycles and the right-to-left minimum letters.

Résumé. Nous prouvons que les paires de statistiques de Mahonian-Stirling $(\mathrm{sor}, \mathrm{cyc})$ et $(\mathrm{inv}, \mathrm{rlmin})$ suivent la même distribution pour des permutations correspondant à des placements de n tours sur un tableau de Ferrer fixé avec n lignes et n colonnes. Les preuves sont combinatoires et utilisent des bijections entre les couplages et les chemins de Dyck. Nous définissons une nouvelle statistique, l'indice de tri pour les couplages. Nous prouvons également un résultat plus fin qui décrit les éléments minimaux dans les cycles des permutations et les lettres minimum droite á gauche.

Keywords: sorting index, cycle, matching, Ferrers board

1 Introduction

An inversion in a permutation σ is a pair $\sigma(i) > \sigma(j)$ such that $i < j$. The number of inversions in σ is denoted by $\mathrm{inv}(\sigma)$. The distribution of inv over the symmetric group S_n was first found by Rodriguez [9] in 1837 and is well known to be

$$\sum_{\sigma \in S_n} q^{\mathrm{inv}(\sigma)} = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}).$$

Much later, MacMahon [6] defined the major index maj and proved that it has the same distribution as inv. In his honor, all permutation statistics that are equally distributed with inv are called Mahonian. MacMahon's remarkable result initiated a systematic research of permutation statistics and in particular many more Mahonian statistics have been described in the literature since then.

Another classical permutation statistic is the number of cycles, cyc. Its distribution is given by

$$\sum_{\sigma \in S_n} t^{\mathrm{cyc}(\sigma)} = t(t+1)(t+2)\cdots(t+n-1)$$

1365–8050 © 2013 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

and the coefficients of this polynomial are known as the unsigned Stirling numbers of the first kind.

Given these two distributions, it is natural then to ask which "Mahonian-Stirling" pairs of statistics $(\mathrm{stat}_1, \mathrm{stat}_2)$ have the distribution

$$\sum_{\sigma\in S_n} q^{\mathrm{stat}_1(\sigma)} t^{\mathrm{stat}_2(\sigma)} = t(t+q)(t+q+q^2)\cdots(t+q+\cdots+q^{n-1}). \tag{1}$$

As proved by Björner and Wachs [1], $(\mathrm{inv}, \mathrm{rlmin})$ and $(\mathrm{maj}, \mathrm{rlmin})$ are two such pairs, where rlmin is the number of right-to-left minimum letters. A right-to-left minimum letter of a permutation σ is a letter $\sigma(i)$ such that $\sigma(i) < \sigma(j)$ for all $j > i$. The set of all right-to-left minimum letters in σ will be denoted by $\mathrm{Rlminl}(\sigma)$. In fact, Björner and Wachs proved the following stronger result

$$\sum_{\sigma\in S_n} q^{\mathrm{inv}(\sigma)} \prod_{i\in\mathrm{Rlminl}(\sigma)} t_i = \sum_{\sigma\in S_n} q^{\mathrm{maj}(\sigma)} \prod_{i\in\mathrm{Rlminl}(\sigma)} t_i = t_1(t_2+q)(t_3+q+q^2)\cdots(t_n+q+\cdots+q^{n-1}). \tag{2}$$

A natural Mahonian partner for cyc was found by Petersen [7]. For a given permutation $\sigma \in S_n$ there is a unique expression

$$\sigma = (i_1 j_1)(i_2 j_2)\cdots(i_k j_k)$$

as a product of transpositions such that $i_s < j_s$ for $1 \le s \le k$ and $j_1 < \cdots < j_k$. The sorting index of σ is defined to be

$$\mathrm{sor}(\sigma) = \sum_{s=1}^{k} (j_s - i_s).$$

The sorting index can also be described as the total distance the elements in σ travel when σ is sorted using the Straight Selection Sort algorithm [5] in which, using a transposition, we move the largest number to its proper place, then the second largest to its proper place, etc. For example, the steps for sorting $\sigma = 6571342$ are

$$65\mathbf{7}1342 \xrightarrow{(37)} \mathbf{6}521347 \xrightarrow{(16)} 4\mathbf{5}21367 \xrightarrow{(25)} \mathbf{4}321567 \xrightarrow{(14)} 1\mathbf{3}24567 \xrightarrow{(23)} 1234567$$

and therefore $\sigma = (2\ 3)(1\ 4)(2\ 5)(1\ 6)(3\ 7)$ and $\mathrm{sor}(\sigma) = (3-2)+(4-1)+(5-2)+(6-1)+(7-3) = 16$. The relationship to other Mahonian statistics and the Eulerian partner for sor were studied by Wilson [10] who called the sorting index DIS.

Petersen showed that

$$\sum_{\sigma\in S_n} q^{\mathrm{sor}(\sigma)} t^{\mathrm{cyc}(\sigma)} = t(t+q)(t+q+q^2)\cdots(t+q+\cdots+q^{n-1}),$$

which implies equidistribution of the pairs $(\mathrm{inv}, \mathrm{rlmin})$ and $(\mathrm{sor}, \mathrm{cyc})$.

In this article we show that the pairs $(\mathrm{inv}, \mathrm{rlmin})$ and $(\mathrm{sor}, \mathrm{cyc})$ have the same distribution on the set of restricted permutations

$$S_{\mathbf{r}} = \{\sigma \in S_n : \sigma(k) \le r_k, 1 \le k \le n\}$$

for a nondecreasing sequence of integers $\mathbf{r} = (r_1, \ldots, r_n)$ with $1 \le r_1 \le r_2 \le \cdots \le r_n \le n$. These can be described as permutations that correspond to arrangements of n non-atacking rooks on a Ferrers board with rows of length $r_1, \ldots, r_n$. To obtain the results, in Section 2 we define a sorting index and cycles for perfect matchings and study the distributions of these statistics over matchings of fixed type. We use bijections between matchings and weighted Dyck paths which enable us to keep track of set-valued statistics and obtain more refined results similar to (2) for restricted permutations.

Analogously to sor, Petersen defined the sorting index for signed permutations of type B_n and D_n. Using algebraic methods he proved that

$$\sum_{\sigma \in B_n} q^{\mathrm{sor}_B(\sigma)} t^{\ell'_B(\sigma)} = \sum_{\sigma \in B_n} q^{\mathrm{inv}_B(\sigma)} t^{\mathrm{nmin}_B(\sigma)} = \prod_{i=1}^{n} (1 + t[2i]_q - t), \tag{3}$$

where for an element $\sigma \in B_n$, $\ell'_B(\sigma)$ denotes its reflection length, $\mathrm{inv}_B(\sigma)$ denotes the type B_n inversion number, and nmin is a signed permutation statistic similar to rlmin. Petersen also defined sor_D, a sorting index for type D_n permutations and showed that it is equidistributed with the number of type D_n inversions:

$$\sum_{\sigma \in D_n} q^{\mathrm{sor}_D(\sigma)} = \sum_{\sigma \in D_n} q^{\mathrm{inv}_D(\sigma)} = [n]_q \cdot \prod_{i=1}^{n-1} [2i]_q. \tag{4}$$

While space constraints prevent us from providing details in this extended abstract, we mention that in [8] we define a sorting index and cycle number for bicolored matchings in a fashion analogous to what we will show for ordinary matchings. In particular, this gives a combinatorial proof that the pairs $(\mathrm{sor}_B, \ell'_B)$ and $(\mathrm{inv}_B, \mathrm{nmin}_B)$ are equidistributed on the set of restricted signed permutations

$$B_{\mathbf{r}} = \{\sigma \in B_n : |\sigma(k)| \leq r_k, 1 \leq k \leq n\}$$

for a nondecreasing sequence of integers $1 \leq r_1 \leq r_2 \leq \cdots \leq r_n \leq n$. Using bijections between bicolored matchings and weighted Dyck paths with bicolored rises, we in fact prove equidistribution of set-valued statistics and their generating functions. Moreover, we find natural Stirling partners for sor_D and inv_D and prove equidistribution of the two Mahonian-Stirling pairs on sets of restricted permutations of type D_n:

$$D_{\mathbf{r}} = \{\sigma \in D_n : |\sigma(k)| \leq r_k, 1 \leq k \leq n\}.$$

2 Statistics on perfect matchings

A matching is a partition of a set in blocks of size at most two and if it has no single-element blocks the matching is said to be perfect. The set of all perfect matchings with n blocks is denoted by $\mathcal{M}_n$. All matchings in this work will be perfect and henceforth we will omit this adjective.

2.1 Statistics based on crossings and nestings

A matching in $\mathcal{M}_n$ can be represented by a graph with $2n$ labeled vertices and n edges in which each vertex has a degree 1. The vertices $1, 2, \ldots, 2n$ are drawn on a horizontal line in natural order and two vertices that are in a same block are connected by a semicircular arc in the upper half-plane. We will use $i \cdot j$ to denote an arc with vertices $i < j$. The vertex i is said to be the opener while j is said to be the closer of the arc. For a vertex i, we will denote by $M(i)$ the other vertex which is in the same block in the matching M as i. Two arcs $i \cdot j$ and $k \cdot l$ with $i < k$ can be in three different relative positions. We say that they form a crossing if $i < k < j < l$, they form a nesting if $i < k < l < j$, and they form an alignment if $i < j < k < l$. The arc with the smaller opener will be called the left arc of the crossing, nesting, or the alignment, respectively, while the arc with the larger opener will be called the right arc. The numbers of crossings, nestings, and alignements in a matching M are denoted by $\mathrm{cr}(M)$, $\mathrm{ne}(M)$, and $\mathrm{al}(M)$, respectively.

If $o_1 < \cdots < o_n$ and $c_1 < \cdots < c_n$ are the openers and the closers in M, respectively, let

$$\mathrm{Long}(M) = \{k : o_k \cdot M(o_k) \text{ is not a right arc in a nesting}\}$$

and

$$\mathrm{Short}(M) = \{k : M(c_k) \cdot c_k \text{ is not a left arc in a nesting}\}.$$

Similarly, let

$$\mathrm{Left}(M) = \{k : o_k \cdot M(o_k) \text{ is not a right arc in a crossing}\}.$$

We will use lower-case letters to denote the cardinalities of the sets. For example, $\mathrm{long}(M) = |\mathrm{Long}(M)|$.

Example 2.1. *For the matching M in Figure 1 we have* $\mathrm{ne}(M) = \mathrm{cr}(M) = \mathrm{al}(M) = 5$*,* $\mathrm{Long}(M) = \{1, 2\}$*,* $\mathrm{Short}(M) = \{1, 2, 3, 5\}$*, and* $\mathrm{Left}(M) = \{1, 5\}$.

The pair of sets $(\{o_1, \dots, o_n\}, \{c_1, \dots, c_n\})$ of openers and closers of a matching M is called the type of M. There is a natural one-to-one correspondence between types of matchings in $\mathcal{M}_n$ and Dyck paths of semilength n, i.e., lattice paths that start at $(0, 0)$, end at $(2n, 0)$, use steps $(1, 1)$ (rises) and $(1, -1)$ (falls), and never go below the x-axis. The set of all such Dyck paths will be denoted by $\mathcal{D}_n$. Namely, the openers in the type correspond to the rises in the Dyck path while the closers correspond to the falls. Therefore, for convenience, we will say that a matching in $\mathcal{M}_n$ is of type D, for some Dyck path $D \in \mathcal{D}_n$, and we will denote the set of all matchings of type D by $\mathcal{M}_n(D)$.

The height of a rise of a Dyck path is the y-coordinate of the right endpoint of the corresponding $(1, 1)$ segment. The sequence $(h_1, \dots, h_n)$ of the heights of the rises of a $D \in \mathcal{D}_n$ when read from left to right will be called shortly the height sequence of D. For example, the height sequence of the Dyck path in Figure 1 is $(1, 2, 3, 3, 3, 4)$. A weighted Dyck path is a pair $(D, (w_1, \dots, w_n))$ where $D \in \mathcal{D}_n$ with height sequence $(h_1, \dots, h_n)$ and $w_i \in \mathbb{Z}$ with $1 \leq w_i \leq h_i$. There is a well-known bijection φ from the set $\mathcal{WD}_n$ of weighted Dyck paths of semilength n to $\mathcal{M}_n$ [2]. Namely, the openers $o_1 < o_2 < \cdots < o_n$ of the matching that corresponds to a given $(D, (w_1, \dots, w_n)) \in \mathcal{WD}_n$ are determined according to the type D. To construct the corresponding matching M, we connect the openers from right to left, starting from o_n. After $o_n, o_{n-1}, \dots, o_{k+1}$ are connected to a closer, there are exactly h_k unconnected closers that are larger than o_k. We connect o_k to the w_k-th of the available closers, when they are listed in decreasing order (see Figure 1).

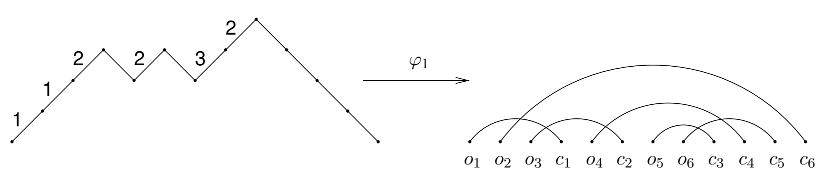


Fig. 1: The bijection φ between weighted Dyck paths and matchings.

Via the bijection φ we immediately get the following generating function.

Theorem 2.2. *If $D \in \mathcal{D}_n$ has a height sequence $(h_1, \ldots, h_n)$, then*

$$\sum_{M \in \mathcal{M}_n(D)} p^{\mathrm{cr}(M)} q^{\mathrm{ne}(M)} \prod_{i \in \mathrm{Left}(M)} s_i \prod_{i \in \mathrm{Long}(M)} t_i = \prod_{k=1}^{n} (t_k p^{h_k - 1} + p^{h_k - 2} q + \cdots + p q^{h_k - 2} + s_k q^{h_k - 1}). \tag{5}$$

Proof: The edge $o_k \cdot M(o_k)$ will be a right arc in exactly $w_k - 1$ nestings and exactly $h_k - w_k$ crossings in $M = \varphi(D, (w_1, \ldots, w_n))$. So, $k \in \mathrm{Long}(M)$ if and only if $w_k = 1$ while the closer that is connected to o_k is in $\mathrm{Left}(M)$ if and only if $w_k = h_k$. □

The map φ also has the following property. The definition of Rlminl was given for permutations but it extends to words in a straightforward way.

Proposition 2.3. *Let $(D, (w_1, \ldots, w_n)) \in \mathcal{WD}_n$ and $M = \varphi(D, (w_1, \ldots, w_n))$. Then*

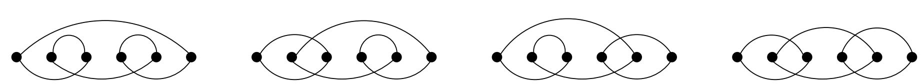
$$\mathrm{Short}(M) = \mathrm{Rlminl}(2 - w_1, 3 - w_2, \ldots, n + 1 - w_n). \tag{6}$$

2.2 Cycles and sorting index for matchings

Let M_0 be a matching in $\mathcal{M}_n(D)$. For $M \in \mathcal{M}_n(D)$ define $\mathrm{cyc}(M, M_0)$ as the number of cycles in the graph $G = (M, M_0)$ on $2n$ vertices in which the arcs from M are drawn in the upper half-plane as usual and the arcs of M_0 are drawn in the lower half-plane, reflected about the number axis. If the openers of M are $o_1 < \cdots < o_n$, we define

$$\mathrm{Cyc}(M, M_0) = \{k : o_k \text{ is a minimal vertex in a cycle in the graph } (M, M_0)\}.$$

Figure 2 shows the calculation of cyc and Cyc for all matchings of type with respect to the nonnesting matching of that type.



$\mathrm{cyc}(M_1, M_4) = 1$ $\quad$ $\mathrm{cyc}(M_2, M_4) = 2$ $\quad$ $\mathrm{cyc}(M_3, M_4) = 2$ $\quad$ $\mathrm{cyc}(M_4, M_4) = 3$

$\mathrm{Cyc}(M_1, M_4) = \{1\}$ $\quad$ $\mathrm{Cyc}(M_2, M_4) = \{1, 2\}$ $\quad$ $\mathrm{Cyc}(M_3, M_4) = \{1, 3\}$ $\quad$ $\mathrm{Cyc}(M_4, M_4) = \{1, 2, 3\}$

Fig. 2: Counting cycles in matchings.

For $M, M_0 \in \mathcal{M}_n(D)$, we define the sorting index of M with respect to M_0, denoted by $\mathrm{sor}(M, M_0)$, in the following way. Let $o_1 < o_2 < \cdots < o_n$ be the openers in M and M_0. We construct a sequence of matchings $M_n, M_{n-1}, \ldots, M_2, M_1$ as follows. First, set $M_n = M$. Then, if $M_k(o_k) = M_0(o_k)$, set $M_{k-1} = M_k$. Otherwise, set M_{k-1} to be the matching obtained by replacing the edges $o_k \cdot M_k(o_k)$ and $M_k(M_0(o_k)) \cdot M_0(o_k)$ in the matching M_k by the edges $o_k \cdot M_0(o_k)$ and $M_k(M_0(o_k)) \cdot M_k(o_k)$. It follows from the definition that $M_1 = M_0$. In other words, we gradually sort the matching M by reconnecting the openers to the closers as "prescribed" by M_0. Note that when swapping of edges takes place, it is always true that $M_k(M_0(o_k)) < o_k$ and therefore all the intermediary matchings we get in the process are of type

D. Define

$$\mathrm{sor}_k(M, M_0) = \begin{cases} |\{c : c > o_k, c \in [M_k(o_k), M_0(o_k)] \text{ and } M_0(c) < o_k\}|, & \text{if } M_k(o_k) \leq M_0(o_k) \\ |\{c : c > o_k, c \notin (M_0(o_k), M_k(o_k)) \text{ and } M_0(c) < o_k\}|, & \text{if } M_0(o_k) < M_k(o_k) \end{cases}$$

and

$$\mathrm{sor}(M, M_0) = \sum_{k=1}^{n} \mathrm{sor}_k(M, M_0).$$

Example 2.4. *Figure 3 shows the intermediate matchings that are obtained when $M = M_6$ is sorted to $M_0 = M_1$. So,*

$\mathrm{sor}_6(M, M_0) = |\{c_3, c_5, c_6\}| = 3$, $\mathrm{sor}_5(M, M_0) = |\{c_3, c_5\}| = 2$, $\mathrm{sor}_4(M, M_0) = |\{c_2, c_5\}| = 2$,
$\mathrm{sor}_3(M, M_0) = |\emptyset| = 0$, $\mathrm{sor}_2(M, M_0) = |\{c_5\}| = 1$, $\mathrm{sor}_1(M, M_0) = |\emptyset| = 0$,

and $\mathrm{sor}(M, M_0) = 0 + 1 + 0 + 2 + 2 + 3 = 8$.

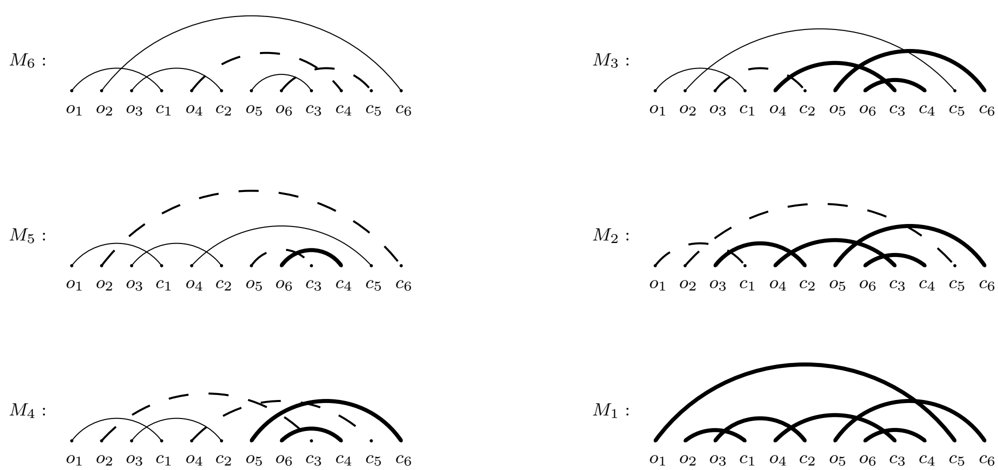


Fig. 3: Sorting of the matching $M = M_6$ to the matching $M_0 = M_1$. The dashed lines indicate arcs that are about to be swapped while the bold lines represent arcs that have been placed in correct position.

Theorem 2.5. *Let D be a Dyck path with height sequence $(h_1, \ldots, h_n)$. For each $M_0 \in \mathcal{M}_n(D)$, there is a bijection*

$$\phi : \{(w_1, w_2, \ldots, w_n) : 1 \leq w_i \leq h_i\} \to \mathcal{M}_n(D)$$

which depends on M_0 such that

(a) $\mathrm{sor}(\phi(w_1, \ldots, w_n), M_0) = \sum_{i=1}^{n}(w_i - 1)$,

(b) $\mathrm{Cyc}(\phi(w_1, \ldots, w_n), M_0) = \{k : w_k = 1\}$.

Additionally, if M_0 is the unique nonnesting matching of type D, then

(c) $\mathrm{Short}(\phi(w_1,\dots,w_n)) = \mathrm{Rlminl}(2-w_1, 3-w_2, \dots, n+1-w_n)$.

Proof: Fix $M_0 \in \mathcal{M}_n(D)$. We construct the bijection ϕ in the following way. Draw the matching M_0 with arcs in the lower half-plane. Suppose $o_1 < \cdots < o_n$ are the openers of M_0. To construct $M = \phi(w_1,\dots,w_n)$, we draw arcs in the upper half plane by connecting the openers from right to left to closers as follows.

Suppose that the openers $o_n, o_{n-1}, \dots, o_{k+1}$ are already connected to a closer and denote the partial matching in the upper half-plane by N_k. To connect o_k, we consider all the closers c with the property $c > o_k$ and $M_0(c) \le o_k$. There are exactly h_k such closers, call them candidates for o_k.

Let c_{k_0} be the closer which is w_k-th on the list when all those h_k candidates are listed starting from $M_0(o_k)$ and then going cyclically to left. If c_{k_0} is not connected to an opener by an arc in the upper half-plane, draw the arc $o_k \cdot c_{k_0}$. Otherwise, there is a maximal path in the graph of the type: $c_{k_0}, N_k(c_{k_0}), M_0(N_k(c_{k_0})), N_k(M_0(N_k(c_{k_0}))), \dots, c^*$ which starts with c_{k_0}, follows arcs in N_k and M_0 alternately and ends with a closer c^* which has not been connected to an opener yet (see Figure 4). Due to the order in which we have been drawing the arcs in the upper half-plane, all vertices in the aforementioned path are to the right of o_k. In particular, c^* is to the right of o_k and is not one of the candidates for o_k. Draw an arc in the upper half-plane connecting o_k to c^*. After all openers are connected in this manner, the resulting matching in the upper half-plane is $M = \phi(w_1,\dots,w_n)$.

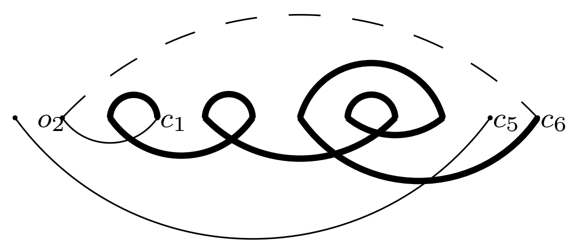


Fig. 4: The solid arcs in the top half-plane represent the partial matching N_2. The candidates for o_2 are c_1 and c_5. If $w_2 = 1$, o_2 will try to connect to c_1, but since it is already connected to an opener, we follow the bold path that starts with c_1 to reach $c^* = c_6$ and connect it to o_2.

Let $M_n = M, M_{n-1}, \dots, M_2, M_1 = M_0$ be the intermediary sequence of matchings constructed when sorting M to M_0. Then $M_k(o_k)$ is exactly the closer c_{k_0} defined above. This means that $\mathrm{sor}_k(M, M_0) = w_k - 1$ and therefore $\mathrm{sor}(M, M_0) = \sum_{k=1}^{n}(w_k - 1)$. This property also gives us a way of finding the sequence $(w_1,\dots,w_n)$ which corresponds to a given $M \in \mathcal{M}_n(D)$. Namely, $w_k = \mathrm{sor}_k(M, M_0) + 1$.

To prove the second property of ϕ, we analyze when connecting o_k by an arc will close a cycle. There are two cases.

1. The closer c_{k_0} which was w_k-th on the list of candidates for o_k was not incident to an arc in the partial matching N_k and we drew the arc $o_k \cdot c_{k_0}$. If $w_k = 1$, then $c_{k_0} = M_0(o_k)$ and the arcs connecting o_k and c_{k_0} in the upper and lower half-planes close a cycle. Otherwise, $M_0(c_{k_0}) < o_k$ and therefore $M_0(c_{k_0})$ is not incident to an arc in N_k and the arc $o_k \cdot c_{k_0}$ will not close a cycle.

2. The closer c_{k_0} which was w_k-th on the list of candidates for o_k was incident to an arc in the partial matching N_k and we drew the arc $o_k \cdot c^*$. If $w_k = 1$, the path traced from c_{k_0} to c^*, the arc $o_k \cdot c_{k_0}$ in M_0, and the newly added arc $o_k \cdot c^*$ form a cycle. Otherwise, connecting o_k to c^* does not close a cycle since the opener $M_0(c_{k_0})$ is in the same connected component of the graph (M, M_0) as o_k, but is not connected to a closer yet, since $M_0(c_{k_0}) < o_k$.

We conclude that a cycle is closed exactly when $w_k = 1$ and therefore

$$\mathrm{Cyc}(\phi(w_1, \ldots, w_n), M_0) = \{k : w_k = 1\}.$$

Finally, we prove the third property of ϕ. If M_0 is a nonnesting matching, its edges are $o_k \cdot c_k$ where the openers and closers are indexed in ascending order. Let $M = \phi(w_1, \ldots, w_n)$. The following observations are helpful. When connecting o_k in the construction of M, the first choice for o_k, i.e., the w_k-th candidate for o_k is exactly c_{k+1-w_k}. Also, $M(o_k) \geq c_{k+1-w_k}$. Furthermore, if c_k was not a candidate for $M(c_k)$, i.e. if the edge c_k was chosen as a partner for $M(c_k)$ by following a path in the graph as described above, then $k \notin \mathrm{Short}(M)$. Namely the edge $M(c_{k_0}) \cdot c_{k_0}$, where c_{k_0} was the first choice when the opener $M(c_k)$ was connected in the construction of M, is nested below it.

For a number $k \in [n]$ there are three possibilities:

1. $k \notin \{2 - w_1, 3 - w_2, \ldots, n + 1 - w_n\}$

   In this case, c_k was not a first choice for any of the openers and therefore must have been connected to an opener by following a path in the graph (M, M_0). It follows from the observation above that $k \notin \mathrm{Short}(M)$.

2. $k \in \{2 - w_1, 3 - w_2, \ldots, n + 1 - w_n\}$ and $k \in \mathrm{Rlminl}(2 - w_1, 3 - w_2, \ldots, n + 1 - w_n)$

   Then c_k was a first choice for at least one opener. Let o be the largest one. Then all openers to the right of o got connected to a closer which is greater than c_k, so no edge is nested below $o \cdot c_k \in M$. Consequently, $k \in \mathrm{Short}(M)$.

3. $k \in \{2 - w_1, 3 - w_2, \ldots, n + 1 - w_n\}$ but $k \notin \mathrm{Rlminl}(2 - w_1, 3 - w_2, \ldots, n + 1 - w_n)$

   In this case, let $m + 1 - w_m$ be the rightmost number in the sequence $(2 - w_1, \ldots, n + 1 - w_n)$ which is smaller than k. It is necessarily to the right of k in this sequence and belongs to $\mathrm{Rlminl}(2 - w_1, \ldots, n + 1 - w_n)$. This implies that the edge $o_m \cdot c_{m+1-w_m}$ is in M, while $M(o_l) > c_k$ for all $l > m$. So, $M(c_k) < o_m$ and therefore the edge $o_m \cdot c_{m+1-w_m}$ is nested below $M(c_k) \cdot c_k$, which means that $k \notin \mathrm{Rlminl}(2 - w_1, 3 - w_2, \ldots, n + 1 - w_n)$.

□

As a consequence, we get the following generating functions. Note that their explicit formulas imply that in fact the distributions are independent of the choice of M_0.

Corollary 2.6. *Let $M_0 \in \mathcal{M}_n(D)$ and let $(h_1, \ldots, h_n)$ be the height sequence of D. Then*

$$\sum_{M \in \mathcal{M}_n(D)} q^{\mathrm{sor}(M, M_0)} \prod_{i \in \mathrm{Cyc}(M, M_0)} t_i = \prod_{k=1}^{n} (t_k + q + \cdots + q^{h_k - 1}). \tag{7}$$

Combining Theorem 2.2 and Corollary 2.6 we get the following corollary.

Corollary 2.7. *Let $M_0 \in \mathcal{M}_n(D)$ and let $(h_1, \dots, h_n)$ be the height sequence of D. Then*

$$\sum_{M\in\mathcal{M}_n(D)} q^{\mathrm{sor}(M,M_0)} \prod_{i\in\mathrm{Cyc}(M,M_0)} t_i = \sum_{M\in\mathcal{M}_n(D)} q^{\mathrm{ne}(M)} \prod_{i\in\mathrm{Long}(M)} t_i.$$

Corollary 2.8. *If M_0 is the unique nonnesting matching of type D then the multisets*

$$\{(\mathrm{sor}(M, M_0), \mathrm{Cyc}(M, M_0), \mathrm{Short}(M)) : M \in \mathcal{M}_n(D)\}$$

and

$$\{(\mathrm{ne}(M), \mathrm{Long}(M), \mathrm{Short}(M)) : M \in \mathcal{M}_n(D)\}$$

are equal.

2.3 Connections with restricted permutations

For a fixed n, let $\mathbf{r}$ denote the non-decreasing sequence of integers $1 \le r_1 \le r_2 \le \cdots \le r_n \le n$. Let

$$S_{\mathbf{r}} = \{\sigma \in S_n : \sigma(k) \le r_k, 1 \le k \le n\}.$$

Note that $S_{\mathbf{r}} \ne \emptyset$ precisely when $r_k \ge k$, for all k, so we will consider only the sequences that satisfy this condition without explicitly mentioning it. Let $D(\mathbf{r})$ be the unique Dyck path whose k-th fall is preceded by exactly r_k rises. Consider the following bijection $f_{\mathbf{r}} : S_{\mathbf{r}} \to \mathcal{M}_n(D(\mathbf{r}))$. If $\sigma \in S_{\mathbf{r}}$, then $f_{\mathbf{r}}(\sigma)$ is the matching in $\mathcal{M}_n(D(\mathbf{r}))$ with edges $o_{\sigma(k)} \cdot c_k$, where $o_1 < \cdots < o_n$ are the openers and $c_1 < \cdots < c_n$ are the closers. It is not difficult to see that $f_{\mathbf{r}}$ is well defined and that it is a bijection.

Two arcs $o_{\sigma(j)} \cdot c_j$ and $o_{\sigma(k)} \cdot c_k$ in $f_{\mathbf{r}}(\sigma)$ with $j < k$ form a nesting if and only if $\sigma(j) > \sigma(k)$. So, $\mathrm{ne}(f_{\mathbf{r}}(\sigma)) = \mathrm{inv}(\sigma)$. Moreover, $\sigma(j) \in \mathrm{Rlminl}(\sigma)$ if and only if $\sigma(j)$ does not form an inversion with a $\sigma(k)$ for any $k > j$, which means if and only if $o_{\sigma(j)} \cdot c_j$ is not nested within anything in $f_{\mathbf{r}}(\sigma)$, i.e., $\sigma(j) \in \mathrm{Long}(f_{\mathbf{r}}(\sigma))$. From Theorem 2.2 we get the following corollary.

Corollary 2.9. *Let $\mathbf{r}$ be a non-decreasing sequence of integers $1 \le r_1 \le r_2 \le \cdots \le r_n \le n$ with $r_k \ge k$, for all k. Then*

$$\sum_{\sigma\in S_{\mathbf{r}}} q^{\mathrm{inv}(\sigma)} \prod_{i\in\mathrm{Rlminl}(\sigma)} t_i = \prod_{k=1}^{n} (t_k + q + q^2 + \cdots + q^{h_k - 1})$$

where $(h_1, \dots, h_n)$ is the height sequence of $D(\mathbf{r})$. In particular,

$$\sum_{\sigma\in S_{\mathbf{r}}} q^{\mathrm{inv}(\sigma)} t^{\mathrm{rlminl}(\sigma)} = \prod_{k=1}^{n} (t + q + q^2 + \cdots + q^{r_k - k}).$$

Proof: The first result follows directly from the discussion above and Theorem 2.2. For the second equality, note that the height sequence $(h_1, \dots, h_n)$ of the Dyck path $D(\mathbf{r})$ is a permutation of the sequence of the heights of the falls in $D(\mathbf{r})$, where the height of a fall is the y-coordinate of the higher end of the corresponding $(1, -1)$ step. The height of the k-th fall is easily seen to be $r_k - k + 1$. □

In particular, when $r_1 = r_2 = \cdots = r_n = n$, we have $S_{\mathbf{r}} = S_n$. The height sequence of $D(\mathbf{r})$ is $(1, 2, \dots, n)$ and we recover the result of Björner and Wachs about the distribution of $(\mathrm{inv}, \mathrm{Rlmin})$ given in (2).

If $M_0 \in \mathcal{M}(D(\mathbf{r}))$ the sorting index $\mathrm{sor}(\,\cdot\,, M_0)$ induces a permutation statistic on $S_{\mathbf{r}}$. Namely, if $\sigma, \sigma_0 \in S_{\mathbf{r}}$, define

$$\mathrm{sor}_{\mathbf{r}}(\sigma, \sigma_0) = \mathrm{sor}(f_{\mathbf{r}}^{-1}(\sigma), f_{\mathbf{r}}^{-1}(\sigma_0)).$$

Equivalently, the statistic $\mathrm{sor}_{\mathbf{r}}(\sigma, \sigma_0)$ on $S_{\mathbf{r}}$ can be defined directly via a sorting algorithm similar to Straight Selection Sort. Namely, permute the elements in $\sigma \in S_{\mathbf{r}}$ by applying transpositions which place the largest element n in position $\sigma_0^{-1}(n)$, then the element $n-1$ in position $\sigma_0^{-1}(n-1)$, etc. Let $\sigma_n = \sigma, \sigma_{n-1}, \ldots, \sigma_1 = \sigma_0$, be the sequence of permutations obtained in this way. Specifically, $\sigma_k^{-1}(i) = \sigma_0^{-1}(i)$ for $i > k$, and σ_{k-1} is obtained by swapping k and $\sigma_k(\sigma_0^{-1}(k))$ in σ_k.

Let $l = \sigma_k^{-1}(k)$ and $m = \sigma_0^{-1}(k)$. Define

$$a_k = \begin{cases} |\{i : l \le i \le m, \sigma_0(i) < k\}|, & l < m \\ 0, & l = m \\ |\{i : r_i \ge k, i \notin (m, l), \sigma_0(i) < k\}|, & l > m. \end{cases} \tag{8}$$

Then

$$\mathrm{sor}_{\mathbf{r}}(\sigma, \sigma_0) = \sum_{k=1}^{n} a_k.$$

Note that, $\mathrm{sor}_{\mathbf{r}}(\sigma, \sigma_0)$ in general depends on $\mathbf{r}$. However, the case when σ_0 is the identity permutation is an exception.

Lemma 2.10. *Let* $\mathbf{r}$ *be a non-decreasing sequence of integers* $1 \le r_1 \le r_2 \le \cdots \le r_n \le n$ *with* $r_k \ge k$, *for all* k. *Let* $\sigma \in S_{\mathbf{r}}$. *Then*

$$\mathrm{sor}_{\mathbf{r}}(\sigma, \mathbf{id}) = \mathrm{sor}(\sigma).$$

Proof: First note that the case $l > m$ in (8) cannot occur. Namely, in the case when $\sigma_0 = \mathbf{id}$, we have $m = k$ and if $l > k$, $\sigma_k^{-1}(l) = \sigma_0^{-1}(l) = l$. This contradicts $l = \sigma_k^{-1}(k)$. Therefore, the definition of a_k simplifies to

$$a_k = |\{i : l \le i < k\}|.$$

This is precisely the "distance" that k travels when being placed in its correct position with the Straight Selection Sort algorithm. $\square$

Corollary 2.11. *Let* $\mathbf{r}$ *be a non-decreasing sequence of integers* $1 \le r_1 \le r_2 \le \cdots \le r_n \le n$ *with* $r_k \ge k$, *for all* k. *Let* $\sigma_0 \in S_{\mathbf{r}}$. *Then*

$$\sum_{\sigma \in S_{\mathbf{r}}} q^{\mathrm{sor}_{\mathbf{r}}(\sigma, \sigma_0)} \prod_{i \in \mathrm{Cyc}(\sigma\sigma_0^{-1})} t_i = \prod_{i=1}^{n} (t_i + q + \cdots + q^{h_i - 1}), \tag{9}$$

where $(h_1, \ldots, h_n)$ *is the height sequence of* $D(\mathbf{r})$ *and* $\mathrm{Cyc}(\sigma)$ *is the set of the minimal elements in the cycles of* σ. *In particular,*

$$\sum_{\sigma \in S_{\mathbf{r}}} q^{\mathrm{sor}(\sigma)} \prod_{i \in \mathrm{Cyc}(\sigma)} t_i = \prod_{i=1}^{n} (t_k + q + \cdots + q^{h_k - 1}) \tag{10}$$

and

$$\sum_{\sigma\in S_{\mathbf{r}}} q^{\mathrm{sor}(\sigma)}t^{\mathrm{cyc}(\sigma)} = \sum_{\sigma\in S_{\mathbf{r}}} q^{\mathrm{inv}(\sigma)}t^{\mathrm{rlminl}(\sigma)} \tag{11}$$

Proof: Let $f_{\mathbf{r}}(\sigma_0) = M_0$ and $f_{\mathbf{r}}(\sigma) = M$. The cycle $k \to \sigma_0\sigma^{-1}(k) \to \cdots \to (\sigma_0\sigma^{-1})^s(k) = k$ of the permutation $\sigma_0\sigma^{-1}$ corresponds to the cycle $o_k \frown M(o_k) \smile M_0(M(o_k)) \frown \cdots \smile o_k$ in the graph (M, M_0). So, $k \in \mathrm{Cyc}(\sigma_0\sigma^{-1})$ if and only if $k \in \mathrm{Cyc}(M, M_0)$. Now, (9) follows from (7) and the fact that the cycles of $\sigma\sigma_0^{-1}$ are equal to the cycles of $\sigma_0\sigma^{-1}$ reversed. Since **id** $\in S_{\mathbf{r}}$ for every sequence $\mathbf{r}$, we get (10) as a corollary of Lemma 2.10. $\square$

Let $\mathrm{Lrmaxp}(\sigma)$ denote the set of left-to-right maximum places in the permutation σ, i.e,

$$\mathrm{Lrmaxp}(\sigma) = \{k : \sigma(k) > \sigma(j) \text{ for all } j < k\}.$$

From Corollary 2.8 we get the following result for restricted permutations.

Corollary 2.12. *The triples* $(\mathrm{inv}, \mathrm{Rlminl}, \mathrm{Lrmaxp})$ *and* $(\mathrm{sor}, \mathrm{Cyc}, \mathrm{Lrmaxp})$ *are equidistributed on* $S_{\mathbf{r}}$. *That is, the multisets*

$$\{(\mathrm{inv}(\sigma), \mathrm{Rlminl}(\sigma), \mathrm{Lrmaxp}(\sigma)) : \sigma \in S_{\mathbf{r}}\}$$

and

$$\{(\mathrm{sor}(\sigma), \mathrm{Cyc}(\sigma), \mathrm{Lrmaxp}(\sigma)) : \sigma \in S_{\mathbf{r}}\}$$

are equal.

The equidistribution of the pairs $(\mathrm{Rlminl}, \mathrm{Lrmaxp})$ and $(\mathrm{Cyc}, \mathrm{Lrmaxp})$ on $S_{\mathbf{r}}$ for the special case when the corresponding Dyck path $D(\mathbf{r})$ is of the form $u^{k_1}d^{k_1}u^{k_2}d^{k_2}\cdots u^{k_s}d^{k_s}$ was shown by Foata and Han [3] .

Corollary 2.13. *Let* $\sigma_0 \in S_{\mathbf{r}}$. *Then*

$$\sum_{\sigma\in S_{\mathbf{r}}} t^{\mathrm{cyc}(\sigma\sigma_0^{-1})} = \prod_{k=1}^{n}(t + r_k - k). \tag{12}$$

In particular, the left-hand side of (12) *does not depend on* σ_0.

We remark that the sets $\{\sigma\sigma_0^{-1} : \sigma \in S_{\mathbf{r}}\}$ and $S_{\mathbf{r}}$ are in general not equal. For example, let $\sigma_0 = 143265 \in S_{[4,4,4,6,6,6]}$. Then $\sigma = 231546 \in S_{[4,4,4,6,6,6]}$ but $\sigma\sigma_0^{-1} = 251364 \notin S_{[4,4,4,6,6,6]}$.

The polynomial $\prod_{k=1}^{n}(t + r_k - k)$ is well-known in rook theory. It is equal [4] to the polynomial

$$\sum_{k=0}^{n} r_{n-k}(t-1)(t-2)\cdots(t-k)$$

where r_k is the number of placements of k non-atacking rooks on a Ferrers board with rows of length $r_1, r_2, \ldots, r_n$.

References

[1] A. Björner, M. Wachs, Permutation statistics and linear extensions of posets, J. Combin. Theory Ser. A 58 (1991) 85–114.

[2] M. de Sainte-Catherine, Couplages et Pfaffiens en Combinatoire, Physique et Informatique, Ph.D. thesis, University of Bordeaux I, Talence, France, 1993.

[3] D. Foata, G.-N. Han, New permutation coding and equidistribution of set valued statistics, Theoret. Comput. Sci. 410 (2009) 3743–3750.

[4] J.R. Goldman, J.T. Joichi, D.E.White Rook theory I: Rook equivalence of Ferrers boards, Proc. Amer. Math. Soc. 52 (1975) 485–492.

[5] D. Knuth, The art of computer programming, vol.3, Addison-Wesley, 1998.

[6] P.A. MacMahon, Two applications of general theorems in combinatory analysis, Proc. London Math. Soc. 15 (1916) 314–321.

[7] T.K. Petersen, The sorting index, Adv. in Appl. Math. 47 (2011) 615–630.

[8] S. Poznanović, Cycles and sorting index for matchings and restricted permutations, arXiv:1206.1301.

[9] O. Rodriguez, Note sur les inversions, ou dérangements produits dans les permutations, J. de Math. 4 (1839), 236–240.

[10] M.C. Wilson, An interesting new Mahonian permutation statistic, Electron. J. Combin. 17 (2010), #R147.