Stanley’s Shuffle Theorem and Insertion Lemma

Kathy Q. Ji\(^1\) and Dax T.X. Zhang\(^2\)

\(^1\)\(^2\) Center for Applied Mathematics, Tianjin University, Tianjin 300072, P.R. China

\(^1\)\(^\text{kathyji@tju.edu.cn} \) and \(^2\)\(^\text{zhangtianxing6@tju.edu.cn}\)

Abstract. This note is devoted to a combinatorial proof of Stanley’s Shuffle Theorem by using the insertion lemma of Haglund, Loehr and Remmel.

Keywords: descent, major index, permutation, shuffle

AMS Classification: 05A05, 05A19, 11P81

1 Introduction

Stanley’s Shuffle Theorem states an explicit expression for the generating function of the number of shufflings of two disjoint permutation \(\sigma\) and \(\pi\) with \(k\) descents and the major index being \(t\). Let us recall some common notation and terminology on permutations as used in [8, Chapter 1]. We say that \(\pi = \pi_1 \pi_2 \cdots \pi_n\) is a permutation of length \(n\) if it is a sequence of \(n\) distinct letters—not necessarily from 1 to \(n\). For example, \(\pi = 9 3 8 10 12 3 7\) is a permutation of length 7. Let \(S_n\) denote the set of all permutations of length \(n\).

We say that \(i \in [n-1]\) is a descent of \(\pi \in S_n\) if \(\pi_i > \pi_{i+1}\). The set of descents of \(\pi\) is called the descent set of \(\pi\), denoted \(\text{Des}(\pi)\) and the number of its descents is called the descent number, denoted \(\text{des}(\pi)\). The major index of \(\pi\), denoted \(\text{maj}(\pi)\), is defined to be the sum of its descents. To wit,

\[ \text{maj}(\pi) := \sum_{k \in \text{Des}(\pi)} k. \]

Let \(\sigma \in S_m\) and \(\pi \in S_n\) be disjoint permutations, that is, permutations with no letters in common. We say that \(\alpha \in S_{n+m}\) is a shuffle of \(\sigma\) and \(\pi\) if both \(\sigma\) and \(\pi\) are subsequences of \(\alpha\). The set of shuffles of \(\sigma\) and \(\pi\) is denoted \(S(\sigma, \pi)\). For example,

\[ S(63,14) = 6314, 6134, 6143, 1463, 1634, 1643. \]

It is easy to see that the number of permutations in \(S(\sigma, \pi)\) is \(\binom{m+n}{n}\).

Define

\[ S_{k,q}(\sigma, \pi) = \sum_{\alpha \in S(\sigma, \pi)} q^{\text{maj}(\alpha)}. \]

In light of the \(q\)-Pfaff-Saalschütz identity in his setting of \(P\)-partitions, Stanley [7] obtained a compact expression for \(S_{k,q}(\sigma, \pi)\) in terms of the Gaussian polynomial (also called the \(q\)-binomial coefficients), as given by

\[ \binom{n}{m} = \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-m+1})}{(1-q^m)(1-q^{m-1})\cdots(1-q)}, \]

see Andrews [1, Chapter 1]. More precisely,

Theorem 1.1 (Stanley’s Shuffle Theorem). Let \(\sigma \in S_m\) and \(\pi \in S_n\) be disjoint permutations, where \(\text{des}(\sigma) = r\) and \(\text{des}(\pi) = s\). Then

\[ \sum_{\alpha \in S(\sigma, \pi) \atop \text{des}(\alpha) \geq k} q^{\text{maj}(\alpha)} = \binom{m-r+s}{k-r} \binom{n-s+r}{k-s} q^{\text{maj}(\sigma)+\text{maj}(\pi)+(k-s)(k-r)}. \quad (1.1) \]
Stanley asked for a proof of Theorem 1.1 which avoids the use of the \( q \)-Pfaff-Saalschütz identity (see [1, Eq.3.3.11]). The bijective proofs of Stanley’s shuffle theorem have been given by Goulden [3] and Stadler [6] respectively. Goulden’s proof is obtained by finding bijections for lattice path representations of shuffles which reduce \( \sigma \) and \( \pi \) to canonical permutations, for which the generating function is easily given. Stadler’s bijection is more elementary, but the inverse of Stadler’s map is not very explicit. In this paper, we give an explicit bijective proof of Theorem 1.1 by using the insertion lemma of Haglund, Leohr and Remmel [4]. It turns out that the insertion lemma of Haglund, Leohr and Remmel is equivalent to Stanley’s Shuffle Theorem in the case of \( n = 1 \). It should be mentioned that Novick [5] used the insertion lemma of Haglund, Leohr and Remmel to give a bijective proof of the following theorem due to Garsia and Gessel [2].

**Theorem 1.2** (Garsia and Gessel). Let \( \sigma \) and \( \pi \) be disjoint permutations with length \( n \) and \( m \) respectively. Then

\[
\sum_{\alpha \in S(\sigma, \pi)} q^{\text{maj}(\alpha)} = \binom{n + m}{m} q^{\text{maj}(\sigma) + \text{maj}(\pi)}. \tag{1.2}
\]

In fact, Theorem 1.2 can be derived from Theorem 1.1 by employing \( q \)-analogue of the Chu-Vandermonde summation (see [1, Eq.3.3.10]),

\[
\sum_{k=0}^{h} \binom{n}{k} \binom{m}{h-k} q^{(n-k)(h-k)} = \binom{m + n}{h}. \tag{2.1}
\]

### 2 The Insertion Lemma of Haglund, Leohr and Remmel

Assume that \( \sigma = \sigma_1 \sigma_2 \cdots \sigma_n \in S_n \) and \( r \notin \sigma \) (that is, there does not exist \( 1 \leq j \leq n \) such that \( \sigma_j = r \)). For \( 0 \leq i \leq n \), let \( \sigma^{(i)}(r) \) denote the permutation of length \( n+1 \) obtained by inserting \( r \) into \( \sigma \) before \( \sigma_{i+1} \). Here we assume that \( \sigma^{(n)}(r) \) denotes the permutation of length \( n+1 \) obtained by inserting \( r \) into \( \sigma \) after \( \sigma_n \). The insertion lemma of Haglund, Leohr and Remmel [4] showed that no matter what is the relative value of \( r \) with respect to the elements in \( \sigma \),

\[
\sum_{i=0}^{n} q^{\text{maj}(\sigma^{(i)}(r))} = (1 + q + \cdots + q^n) q^{\text{maj}(\sigma)}. \tag{2.1}
\]

This relation can be used to establish the following celebrated formula due to MacMahon [9].

\[
\sum_{\sigma \in S_n} q^{\text{maj}(\sigma)} = [n]_q!, \tag{2.2}
\]

where \([n]_q! = [1]_q [2]_q \cdots [n]_q\) with \([k]_q = 1 + q + \cdots + q^{k-1}\).

Haglund, Leohr and Remmel [4] also classified the possible spaces where they can insert \( r \) into \( \sigma \) into two sets called the right-to-left spaces which denoted as RL-spaces and the left-to-right spaces which denoted as LR-spaces. That is, a space \( i \) is called a RL-space of \( \sigma \) relative to \( r \) if

1. \( i = n \) and \( \sigma_n < r \),
2. \( i = 0 \) and \( r \ < \sigma_1 \),
3. \( 0 < i < n \) and \( \sigma_i > \sigma_{i+1} \ > \sigma_r \),
4. \( 0 < i < n \) and \( r > \sigma_i > \sigma_{i+1} \), or
5. \( 0 < i < n \) and \( \sigma_i < r < \sigma_{i+1} \).

Then a space \( i \) is a LR-space of \( \sigma \) relative to \( r \) if it is not a RL-space of \( \sigma \) relative to \( r \). Assume that there are \( l \) RL-spaces for \( \sigma \) relative to \( r \). Then label the RL-spaces from right to left with \( 0, \ldots, l - 1 \) and label
the LR-spaces from left to right with \( l, \ldots, n \) and call this labeling the canonical labeling for \( \sigma \) relative to \( r \). For example suppose that \( r = 5 \) and \( \sigma = 10 \ 9 \ 8 \ 2 \ 4 \ 3 \ 6 \) is a permutation in \( S_9 \). By definition, we see the RL-spaces of \( \sigma \) relative to 5 are 0, 2, 3, 5, 7 and 8 and the LR-spaces of \( \sigma \) relative to 5 are 1, 4, 6 and 9. The canonical labeling of \( \sigma \) relative to 5 is
\[
\sigma = 10 \ 6 \ 1 \ 4 \ 9 \ 3 \ 8 \ 2 \ 7 \ 5 \ 3 \ \sigma = 1 \ 3 \ 6 \ 9.
\]
where the bold number in the subscript represents the labeling of the RL-spaces of \( \sigma \) relative to 5.

Haglund, Leohr and Remmel [4] established the following insertion lemma.

**Lemma 2.1 (The Insertion Lemma).** Assume that \( \sigma = \sigma_1 \sigma_2 \cdots \sigma_n \in S_n \) and \( r \not\in \sigma \), and let \( \sigma^{(i)}(r) \) denote the permutation obtained by inserting \( r \) into \( \sigma \) before \( \sigma_{i+1} \). If the label at the \( i \)-th space in the canonical labeling of \( \sigma \) relative to \( r \) is equal to \( a \), then
\[
\text{maj}(\sigma^{(i)}(r)) = a + \text{maj}(\sigma).
\]

A careful examination of definitions of the RL-spaces and the LR-spaces leads to the following lemma, which is useful in the proof of Stanley’s Shuffle Theorem.

**Lemma 2.2.** Let \( \sigma, r \) and \( \sigma^{(i)}(r) \) be given in Lemma 2.1. If \( i \) is a RL-space of \( \sigma \) relative to \( r \), then \( \text{des}(\sigma^{(i)}(r)) = \text{des}(\sigma) \). If \( i \) is a LR-space of \( \sigma \) relative to \( r \), then \( \text{des}(\sigma^{(i)}(r)) = \text{des}(\sigma) + 1 \). Moreover, the number of RL-spaces of \( \sigma \) relative to \( r \) is one more than the number of descents in \( \sigma \).

**Proof.** Assume that there are \( k \) descents in \( \sigma \). By the definitions of RL-spaces and LR-spaces, we find that if \( i \) is a RL-space of \( \sigma \) relative to \( r \), then \( \text{des}(\sigma^{(i)}(r)) = \text{des}(\sigma) = k \). Moreover, the major increment \( \text{maj}(\sigma^{(i)}(r)) - \text{maj}(\sigma) \) equals the number of descents of \( \sigma^{(i)}(r) \) after \( r \), and hence \( \text{maj}(\sigma^{(i)}(r)) - \text{maj}(\sigma) \leq k \). If \( i \) is a LR-space of \( \sigma \) relative to \( r \), then \( \text{des}(\sigma^{(i)}(r)) = \text{des}(\sigma) + 1 \). Moreover, the major increment \( \text{maj}(\sigma^{(i)}(r)) - \text{maj}(\sigma) \) is not smaller than the sum of \( i \) and the number of descents of \( \sigma \) after \( \sigma_{i+1} \), which is apparently greater than \( k \). Assume that the number of RL-spaces of \( \sigma \) relative to \( r \) is \( n_r \) and the number of LR-spaces of \( \sigma \) relative to \( r \) is \( n_l \). Clearly, \( n_r + n_l = n + 1 \). By Lemma 2.1, we see that the major increment at each space of \( \sigma \) is different, so we conclude that \( n_r \leq k + 1 \) and \( n_l \leq n - k \). Since \( n_r + n_l = n + 1 \), we derive that \( n_r = k + 1 \) and \( n_l = n - k \). This completes the proof.

We conclude this section with the proof of Stanley’s Shuffle Theorem in the case of \( n = 1 \) in view of Lemma 2.1 and Lemma 2.2.

Setting \( n = 1 \) in Theorem 1.1, we have \( s = 0 \) and \( \text{maj}(\pi) = 0 \), and thus Theorem 1.1 reads as follows:
\[
\sum_{\alpha \in S(\pi, \pi) \atop \text{des}(\alpha) = r} q^{\text{maj}(\alpha)} = \frac{m - r}{k - r} \left[ \frac{r + 1}{k} \right] q^{\text{maj}(\sigma) + k(k - r)}. \tag{2.3}
\]

From the right-hand side of (2.3), we find that \( k = r \) or \( k = r + 1 \). Hence identity (2.3) can be written as
\[
\sum_{\alpha \in S(\pi, \pi) \atop \text{des}(\alpha) = r} q^{\text{maj}(\alpha)} = \left[ \frac{r + 1}{r} \right] q^{\text{maj}(\pi)} = (1 + q + \cdots + q^r) q^{\text{maj}(\sigma)}. \tag{2.4}
\]

and
\[
\sum_{\alpha \in S(\pi, \pi) \atop \text{des}(\alpha) = r + 1} q^{\text{maj}(\alpha)} = \left[ \frac{m - r}{1} \right] q^{\text{maj}(\pi) + r + 1} = q^{r + 1} (1 + q + \cdots + q^{m - r - 1}) q^{\text{maj}(\sigma)}. \tag{2.5}
\]

Obviously, the identity (2.4) and the identity (2.5) are immediate consequences of Lemma 2.1 and Lemma 2.2.
3 The bijection

This section is devoted to a proof of Theorem 1.1 in the general case by means of Lemma 2.1 and Lemma 2.2. To state the proof, we need to recall some notation and terminology on partitions as in [1, Chapter 1]. A partition $\lambda$ of a positive integer $n$ is a finite nonincreasing sequence of positive integers $(\lambda_1, \lambda_2, \ldots, \lambda_s)$ such that $\sum_{i=1}^s \lambda_i = n$. Then $\lambda_i$ are called the parts of $\lambda$, where $\lambda_1$ is its largest part and $\lambda_s$ is its smallest part. The number of parts of $\lambda$ is called the length of $\lambda$, denoted by $l(\lambda)$. The weight of $\lambda$ is the sum of parts of $\lambda$, denoted $|\lambda|$. Let $P_{\leq m}(n)$ denote the set of partitions $\lambda$ such that $l(\lambda) \leq m$ and $\lambda_1 \leq n$. It is well-known that the Gaussian polynomial has the following partition interpretation [1, Theorem 3.1]:

$$\binom{n+m}{m} = \sum_{\lambda \in P_{\leq m}(n)} q^{\lambda},$$

Let $P_m(t, n)$ denote the set of partitions $\lambda$ such that $l(\lambda) = m$, $\lambda s \geq t$ and $\lambda_1 \leq n$. By (3.1), it is clear to see that

$$q^m \binom{n+m-t}{m} = \sum_{\lambda \in P_m(t, n)} q^{\lambda}.$$  (3.2)

Using (3.1) and (3.2), we see that Theorem 1.1 is equivalent to the following combinatorial statement.

**Theorem 3.1.** Let $\sigma$ and $\pi$ be given in Theorem 1.1 and let $S_k(\sigma, \pi)$ denote the set of all shuffles of $\sigma$ and $\pi$ with $k$ descents. Then there is a bijection $\Phi$ between $S_k(\sigma, \pi)$ and $P_{k-r}(k-s, s) \times P_{\leq n-k+r}(k-s)$, namely, for $\alpha \in S_k(\sigma, \pi)$, we have $(\lambda, \mu) = \Phi(\alpha) \in P_{k-r}(k-s, s) \times P_{\leq n-k+r}(k-s)$ such that

$$\text{maj}(\alpha) = |\lambda| + |\mu| + \text{maj}(\sigma) + \text{maj}(\pi).$$

(3.3)

To prove Theorem 3.1, we first give a description of the map $\Phi$ and then we show that the map $\Phi$ is a bijection as desired in Theorem 3.1. We denote by $d_k(\sigma)$ the number of descents in $\sigma$ greater than or equal to $k$. Obviously, $d_1(\sigma) = \text{des}(\sigma)$ and $\text{maj}(\sigma) = \sum_{k=1}^n d_k(n)$.

**Definition 3.2 (The map $\Phi$).** Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_m$ be a permutation with $r$ descents and let $\pi = \pi_1 \pi_2 \cdots \pi_n$ be a permutation with $s$ descents. Assume that $\alpha = \alpha_1 \alpha_2 \cdots \alpha_{n+m}$ is the shuffle of $\sigma$ and $\pi$ with $k$ descents. The pair of partitions $(\lambda, \mu) = \Phi(\alpha)$ can be constructed as follows: Let $\alpha^{(i)}$ denote the permutation obtained by removing $\pi_1, \pi_2, \ldots, \pi_i$ from $\alpha$. Here we assume that $\alpha^{(0)} = \alpha$. Obviously, $\alpha^{(n)} = \sigma$. For $1 \leq i \leq n$, define

$$t(i) = \text{maj}(\alpha^{i-1}) - \text{maj}(\alpha^i) - d_i(\pi).$$

(3.4)

Observe that there are $k$ descents in $\alpha$ and there are $r$ descents in $\sigma$, so there exists $k - r$ permutations in $\alpha^{(i)}, \ldots, \alpha^{(n)}$, denoted by $\alpha^{(i_1)}, \ldots, \alpha^{(i_{k-r})}$ where $1 \leq i_1 < i_2 < \cdots < i_{k-r} \leq n$, such that $\text{des}(\alpha^{(i_l-1)}) = \text{des}(\alpha^{(i_l)}) + 1$ for $1 \leq l \leq k - r$. Let $\{j_1, \ldots, j_{n-k+r}\} \in \{1, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_{k-r}\}$, where $1 \leq j_1 < j_2 < \cdots < j_{n-k+r} \leq n$, obviously, $\text{des}(\alpha^{(j_l-1)}) = \text{des}(\alpha^{(j_l)})$ for $1 \leq l \leq n - k + r$. Then the pair of partitions $(\lambda, \mu) = \Phi(\alpha)$ is defined by

$$\lambda = (t(i_{k-r}), t(i_{k-r-1}), \ldots, t(i_1)),$$

and

$$\mu = (t(j_1), t(j_2), \ldots, t(j_{n-k+r})).$$

(3.5)

(3.6)

For example, let $\sigma = 9 3 8 1 0 1 2 4 7$, $\pi = 1 2 6 5 1 3 1 1$ and $\alpha = 1 9 2 6 3 5 1 3 8 1 0 1 2 1 1 4 7$, where $m = 7$, $r = 2$, $n = 6$, $s = 2$ and $k = 5$. The elements of $\pi$ in $\alpha^{(i)}$ are in boldface to distinguish them from the elements of $\sigma$. The pairs of partitions $(\lambda, \mu) = \Phi(\alpha)$ can be constructed as follows:
\[
\begin{array}{c|cccccc}
\alpha^{(i)} & d_i(\pi) & \text{maj}(\alpha^{(i-1)}) & \text{maj}(\alpha^{(i)}) & t(i) & \text{des}(\alpha^{(i-1)}) & \text{des}(\alpha^{(i)}) \\
\hline
6 & 938101247 & 0 & 6 & 6 & 1 \\
5 & 93810121147 & 1 & 5 & 4 & 1 \\
4 & 9313810121147 & 1 & 3 & 2 & 0 \\
3 & 93513810121147 & 2 & 5 & 3 & 1 \\
2 & 963513810121147 & 2 & 4 & 2 & 0 \\
1 & 9263513810121147 & 2 & 5 & 3 & 0 \\
0 & 19263513810121147 & & & & \\
\end{array}
\]

Hence \( \lambda = (6, 4, 3) \) and \( \mu = (3, 2, 2) \).

In order to prove that the map \( \Phi \) is a bijection, we shall reformulate the insertion lemma of Haglund, Leohr and Remmel. To this end, we first recall the notation of the major increment sequence. Let \( \sigma = \sigma_1 \sigma_2 \cdots \sigma_n \in S_n \) and \( r \notin \sigma \). Recall that \( \sigma^{(i)}(r) \) denotes the permutation obtained by inserting \( r \) before \( \sigma_{i+1} \). For \( 0 \leq i \leq n \), define the major increment
\[
\text{im}(\sigma, i, r) = \text{maj}(\sigma^{(i)}(r)) - \text{maj}(\sigma)
\]
and the major increment sequence
\[
\text{MIS}(\sigma, r) = (\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, n, r)).
\]

Combining Lemma 2.1 and Lemma 2.2, we have the following corollary.

**Corollary 3.3.** Let \( \sigma \in S_n \) with \( k \) descents and \( r \notin \sigma \). Then \( \text{MIS}(\sigma, r) \) is a shuffling of \( k + 1, k + 2, \ldots, n \) and \( k, k + 1, \ldots, 0 \). In particular, \( \text{im}(\sigma, i, r) \) is either \( \min\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i - 1, r)\} - 1 \) or \( \max\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i - 1, r)\} + 1 \). More precisely, if \( \text{des}(\sigma^{(i)}(r)) = \text{des}(\sigma) + 1 \), then \( \text{im}(\sigma, i, r) = \max\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i - 1, r)\} + 1 \), otherwise, \( \text{im}(\sigma, i, r) = \min\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i - 1, r)\} - 1 \).

For example, let \( \sigma = 51624 \in S_5 \) and \( r = 3 \). Then \( \text{des}(\sigma) = 2, \text{maj}(\sigma) = 1 + 3 = 4 \) and
\[
\begin{array}{c|cccc}
\alpha^{(i)}(3) & \text{maj}(\alpha^{(i-1)}(3)) & \text{im}(\alpha, i, 3) & \text{des}(\alpha^{(i-1)}(3)) - \text{des}(\alpha) \\
\hline
0 & 351624 & 6 & 2 & 0 \\
1 & 531624 & 7 & 3 & 1 \\
2 & 513624 & 5 & 1 & 0 \\
3 & 516324 & 8 & 4 & 1 \\
4 & 516234 & 4 & 0 & 0 \\
5 & 516243 & 9 & 5 & 1 \\
\end{array}
\]
so \( \text{MIS}(\sigma, r) = (2, 3, 1, 4, 0, 5) \) which is a shuffle of 3, 4, 5 and 2, 1, 0.

Let \( \text{MIS}_i(\sigma, r) = (\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i - 1, r)) \) be the first \( i \) elements of \( \text{MIS}(\sigma, r) \). Employing the insertion lemma of Haglund, Leohr and Remmel, Novick [5] found the following interesting proposition about \( \text{MIS}_i(\sigma, r) \). It turns out that this proposition plays an important role in the proof that the map \( \Phi \) is a bijection. For completeness, we provide a simple proof of this proposition with the aid of Corollary 3.3. Here we use the common notation \( \chi(T) = 1 \) if the statement \( T \) is true and \( \chi(T) = 0 \) otherwise.

**Proposition 3.4 (Novick).** Let \( \sigma \) be a permutation of length \( m \) with \( r \) descents, \( p, q \notin \sigma \). Let \( \sigma^{(i-1)}(p) \) be the permutation by inserting \( p \) before \( \sigma_i \). Then \( \text{MIS}_i(\sigma^{(i-1)}(p), q) \) is a permutation of the set \( \{\text{im}(\sigma, j, p) + \chi(q > p) | 0 \leq j < i\} \).

**Proof.** Let \( \sigma[i] = \sigma_1 \sigma_2 \cdots \sigma_i \) be the permutation of first \( i \) elements of \( \sigma \). By Corollary 3.3, we find that \( \text{MIS}(\sigma[i], p) \) is a permutation of the set \( \{0, 1, \ldots, i\} \). Note that \( \text{im}(\sigma[i], i, p) = i \chi(\sigma_i > p) \), so \( \text{MIS}_i(\sigma[i], p) \) is a permutation of the set \( \{j - 1 + \chi(p > \sigma_i) | 0 < j < i\} \). By the definition of descents,
we find that \( \text{im}(\sigma[i], j, p) + d_i(\sigma) = \text{im}(\sigma, j, p) \) for \( 0 \leq j < i \). Hence \( MIS_i(\sigma, p) \) is a permutation of the set \( \{j - 1 + \chi(p > \sigma_i) + d_i(\sigma) \mid 0 < j \leq i \} \). Using the same argument, we derive that \( MIS_i(\sigma^{(i-1)}(p), q) \) is a permutation of the set \( \{j - 1 + \chi(q > \sigma_i) + d_i(\sigma^{(i-1)}(p)) \mid 0 < j \leq i \} \), where \( \sigma^{(i-1)}(p) = p \). Note that \( d_i(\sigma^{(i-1)}(p)) = \chi(p > \sigma_i) + d_i(\sigma) \), so we conclude that \( MIS_i(\sigma^{(i-1)}(p), q) \) is a permutation of the set \( \{j - 1 + \chi(q > \sigma_i) + \chi(p > \sigma_i) + d_i(\sigma) \mid 0 < j \leq i \} \). Then the proposition follows immediately by comparing \( MIS_i(\sigma^{(i-1)}(p), q) \) with \( MIS_i(\sigma, p) \). This completes the proof.

For example, let \( \sigma = 581462 \in S_6, p = 7, q = 9 \) and \( i = 5 \). Note that \( \sigma^{(4)}(7) = 5814762 \in S_7 \), and it can be evaluated that

\[
MIS(\sigma, 7) = (3, 2, 4, 5, 6, 1, 0)
\]

and

\[
MIS(\sigma^{(4)}(7), 9) = (4, 5, 3, 6, 7, 2, 1, 0).
\]

We find that \( MIS_i(\sigma, 7) \) is a permutation of \( \{2, 3, 4, 5, 6\} \) while \( MIS_i(\sigma^{(4)}(7), 9) \) is a permutation of \( \{3, 4, 5, 6, 7\} \). On the other hand,

\[
MIS(\sigma, 9) = (3, 4, 2, 5, 6, 1, 0),
\]

and \( \sigma^{(4)}(9) = 5814962 \in S_7 \)

\[
MIS(\sigma^{(4)}(9), 7) = (4, 3, 5, 6, 2, 7, 1, 0).
\]

It is clear that both \( MIS_i(\sigma, 9) \) and \( MIS_i(\sigma^{(4)}(9), 7) \) are permutations of \( \{2, 3, 4, 5, 6\} \).

With Corollary 3.3 and Proposition 3.4 in hand, we can now show that the map \( \Phi \) in Definition 3.2 is a map from \( S_k(\sigma, \pi) \) to \( P_{k-r}(k-s, m) \times P_{\leq n-k+r}(k-s) \).

**Lemma 3.5.** Assume that \( \sigma \) and \( \pi \) are given in Theorem 1.1 and let \( \alpha \in S_k(\sigma, \pi) \) and \( (\lambda, \mu) = \Phi(\alpha) \). Then \( \lambda \in P_{k-r}(k-s, m) \) and \( \mu \in P_{\leq n-k+r}(k-s) \). Furthermore, \( \text{maj}(\alpha) = |\lambda| + |\mu| + \text{maj}(\sigma) + \text{maj}(\pi) \).

**Proof.** Recall that \( \sigma = \sigma_1 \sigma_2 \cdots \sigma_m \) be a permutation with \( r \) descents and let \( \pi = \pi_1 \pi_2 \cdots \pi_n \) be a permutation with \( s \) descents. Assume that \( \alpha = \alpha_1 \alpha_2 \cdots \alpha_{n+m} \) is the shuffle of \( \sigma \) and \( \pi \) with \( k \) descents. For \( 1 \leq i \leq n \), let \( \alpha^{(i)} \) denotes the permutation obtained by removing \( \pi_1, \pi_2, \ldots, \pi_i \) from \( \alpha \) and let \( k_i \) be the position at which \( \pi_i \) is inserted into \( \alpha^{(i)} \) to yield \( \alpha^{(i-1)} \). More precisely, \( \alpha^{(i-1)} \) is obtained by inserting \( \pi_i \) into \( \alpha^{(i)} \) before \( \alpha^{(i)}_{k_i} \). Evidently, \( k_1 \leq k_2 \leq \cdots \leq k_n \). Define

\[
T^{(i)} = (\text{im}(\alpha^{(i)}_{k_1}, 0, \pi_i) - d_i(\pi), \ldots, \text{im}(\alpha^{(i)}_{k_i - 1}, \pi_i) - d_i(\pi))
\]

for \( 1 \leq i \leq n \). From the definition (3.4) of \( t(i) \), it’s clear that \( t(i) \) is the last element in \( T^{(i)} \). By Corollary 3.3, we see that the elements are distinct in \( T^{(i)} \). So we may assume that \( T^{(i)} \) is a permutation of a set \( ST^{(i)} \). In terms of Proposition 3.4, Novick [5] showed that

\[
ST^{(1)} \subseteq ST^{(2)} \subseteq \cdots \subseteq ST^{(n)} \subseteq \{0, 1, \ldots, m-1\}.
\]

It implies that \( 0 \leq t(l) \leq m \) for any \( 1 \leq l \leq n \).

Recall that \( \alpha^{(i_1)}, \ldots, \alpha^{(i_{k-r})} \) are \( k-r \) permutations such that \( \text{des}(\alpha^{(i_1)}) = \text{des}(\alpha^{(i_1+1)}) + 1 \) for \( 1 \leq l \leq k-r \), where \( 1 \leq i_1 < i_2 < \cdots < i_{k-r} \leq n \) and \( \alpha^{(j_1)}, \ldots, \alpha^{(j_{k-r})} \) are permutations such that \( \text{des}(\alpha^{(j_{k-r})}) = \text{des}(\alpha^{(j_{k-r-1})}) + 1 \leq l \leq n-k+r \), where \( 1 \leq j_1 < j_2 < \cdots < j_{n-k+r} \leq n \). From Corollary 3.3, we see that if \( \text{des}(\alpha^{(i_1)}(r)) = \text{des}(\sigma) + 1 \), then \( \text{im}(\sigma, i, r) = \max\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i-1, r)\} + 1 \), otherwise, \( \text{im}(\sigma, i, r) = \min\{\text{im}(\sigma, 0, r), \ldots, \text{im}(\sigma, i-1, r)\} - 1 \). Since \( t(i) \) is the last elements in \( T_i \), it follows that \( t(i) \) is the largest element in \( T_i \) or the smallest element in \( T_i \). Hence, by (3.7), we get

\[
m \geq t(i_{k-r}) \geq \cdots \geq t(i_1) \geq t(j_1) \geq \cdots \geq t(j_{n-k+r}) \geq 0.
\]

To prove that \( \lambda \in P_{k-r}(k-s, m) \) and \( \mu \in P_{\leq n-k+r}(k-s) \), by (3.5) and (3.6), it suffices to show that \( t(i_1) \leq k-s \leq t(i_1) \). By the definition of \( i_1 \), we see that \( \text{des}(\alpha^{(i_1-1)}) = \text{des}(\sigma) = k \) and \( \text{des}(\alpha^{(i_1)}) = k-1 \), and so \( \text{maj}(\alpha^{(i_1-1)}) - \text{maj}(\alpha^{(i_1)}) \geq k \). But \( d_{i_1}(\pi) \leq \text{des}(\pi) = s \), so

\[
t(i_1) = \text{maj}(\alpha^{(i_1-1)}) - \text{maj}(\alpha^{(i_1)}) - d_{i_1}(\pi) \geq k-s.
\]
Combining (3.8) and (3.9), we conclude that \( \lambda = (t(i_{k-r}), t(i_{k-r-1}), \ldots, t(i_1)) \) is a partition in \( P_{k-r}(k-s, m) \). From the definition of \( j_1 \), we have \( \text{des}(\alpha(j_1-1)) = \text{des}(\alpha(j_1)) = k - j_1 + 1 \), and so \( \text{maj}(\alpha(j_1-1)) - \text{maj}(\alpha(j_1)) \leq k - j_1 + 1 \). Since \( d_{j_1}(\pi) \geq \text{des}(\pi) - j_1 + 1 = s - j_1 + 1 \), we have

\[
t(j_1) = \text{maj}(\alpha(j_1)) - \text{maj}(\alpha(j_1+1)) - d_{j_1}(\pi) \leq k - s. \tag{3.10}
\]

Combining (3.8) and (3.10), we arrive at \( \mu = (t(j_1), t(j_2), \ldots, t(j_{n-k+r})) \) is a partition in \( P_{\leq n-k+r}(k-s) \). Moreover, it is evident from (3.4), (3.5) and (3.6) that \( \text{maj}(\alpha) - \text{maj}(\sigma) = \text{maj}(\alpha(0)) - \text{maj}(\sigma(0)) = \sum_{i=1}^n t(i) + \text{maj}(\pi) = |\lambda| + |\mu| + \text{maj}(\pi) \). This completes the proof.

We are finally in a position to give a proof of Theorem 3.1.

**Proof of Theorem 3.1.** To prove \( \Phi \) defined in Definition 3.2 is the desired bijection in Theorem 3.1, we shall define a map \( \Psi \) from \( P_{k-r}(k-s, m) \times P_{\leq n-k+r}(k-s) \) to \( S_k(\sigma, \pi) \). Let \( \lambda = (\lambda_1, \ldots, \lambda_{k-r}) \) and \( \mu = (\mu_1, \ldots, \mu_{n-k+r}) \) where

\[
m \geq \lambda_1 \geq \cdots \geq \lambda_{k-r} \geq k - s \geq \mu_1 \geq \cdots \geq \mu_{n-k+r} \geq 0. \tag{3.11}
\]

Let

\[
M^{(n)} = \{\lambda_1, \ldots, \lambda_{k-r}, \mu_1, \ldots, \mu_{n-k+r}\} \tag{3.12}
\]

be a multi-set consisting of \( \lambda \) and \( \mu \). All elements in \( M^{(n)} \) are listed in non-increasing order. We wish to construct a shuffle \( \alpha = \Psi(\lambda, \mu) \) by inserting \( \pi_n, \pi_{n-1}, \ldots, \pi_1 \) in turn into \( \sigma \) based on \( M^{(n)} \). Recall that \( \sigma = \sigma_1 \sigma_2 \cdots \sigma_m \) is a permutation with \( r \) descents and \( \pi = \pi_1 \pi_2 \cdots \pi_n \) is a permutation with \( s \) descents.

Assume that \( \alpha^{(n)} = \sigma \) and \( k_{n+1} = m + 1 \). Set \( b = 0 \) and carry out the following procedure.

(A) We intend to insert \( \pi_{n-b} \) into \( \alpha^{(n-b-1)} \) to obtain \( \alpha^{(n-b)} \). The key step is to determine the position \( k_{n-b} \) at which \( \pi_n-b \) is inserted into \( \alpha^{(n-b)} \). Define

\[
T^{(n-b)}(\pi) = (\text{im}(\alpha^{(n-b)}, 0, \pi_{n-b}) - d_{n-b}(\pi), \ldots, \text{im}(\alpha^{(n-b)}, k_{n-b+1} - 1, \pi_{n-b}) - d_{n-b}(\pi)). \tag{3.13}
\]

Using Corollary 3.3, we find that when \( b = 0 \), \( T^{(n)} \) is a permutation of the set \{0, 1, \ldots, m\} since \( d_n(\pi) = 0 \). From (3.12), we see that all elements in \( M^{(n)} \) are in \( T^{(n)} \). When \( b \geq 1 \), in light of Proposition 3.4, we derive that the elements in \( T^{(n-b)} \) are the same as the first \( k_{n-b+1} \) elements in \( T^{(n-b+1)} \). Under the precondition that the elements in \( M^{(n-b)} \) are on the right side of \( T^{(n-b+1)}_{k_{n-b+1}} \), we deduce that elements in \( M^{(n-b)} \) are in \( T^{(n-b)} \). We use \( T^{(n-b)}_{k_{n-b}} \) to denote the \( i \)-th element in \( T^{(n-b)} \). Let \( k_{n-b} \) be the largest positive integer such that \( T^{(n-b)}_{k_{n-b}} \in M^{(n-b)} \). By definition, we see that

\[
k_{n-b} \leq k_{n-b+1}. \tag{3.14}
\]

Let \( \alpha^{(n-b-1)} \) be the permutation obtained by inserting \( \pi_{n-b} \) into \( \alpha^{(n-b)} \) before \( \alpha_{k_{n-b}}^{(n-b)} \). From (3.13), we find that

\[
\text{maj}(\alpha^{(n-b-1)}) - \text{maj}(\alpha^{(n-b)}) = T^{(n-b)}_{k_{n-b}} + d_{n-b}(\pi). \tag{3.15}
\]

Define

\[
M^{(n-b-1)} = M^{(n-b)} \setminus \{T^{(n-b)}_{k_{n-b}}\}, \tag{3.16}
\]

which is a multiset of length \( n - b - 1 \). By the choice of \( k_{n-b} \), we see that the elements in \( M^{(n-b-1)} \) are on the right side of \( T^{(n-b)}_{k_{n-b}} \) in \( T^{(n-b)} \).

(B) Replace \( b \) by \( b + 1 \). If \( b = n \), then we are done. Otherwise, go back to (A).

Set \( \alpha = \Psi(\lambda, \mu) = \alpha^{(0)} \). By (3.14), we find that \( \alpha \) is a shuffle of \( \sigma \) and \( \pi \). Moreover, it is clear from (3.15) and (3.16) that

\[
\text{maj}(\alpha) - \text{maj}(\sigma) = \text{maj}(\alpha^{(0)}) - \text{maj}(\alpha^{(n)}) = |\lambda| + |\mu| - \sum_{b=1}^n d_b(\pi) = |\lambda| + |\mu| + \text{maj}(\pi).
\]
It remains to show that there are $k$ descents in $\alpha^{(i)}$. Suppose to the contrary that $\des(\alpha^{(i)}) \neq k$. Assume that $\des(\alpha^{(i)}) = l < k$. Let $(\mathcal{L}, \mathcal{P}) = \Phi(\alpha^{(i)})$, by Lemma 3.5, we derive that $\mathcal{L} \in P_{l-r}(l-s, m)$ and $\mathcal{P} \in P_{n-l+r}(l-s, s)$, that is,

$$m \geq \mathcal{L}_1 \geq \cdots \geq \mathcal{L}_{l-r} \geq l-s \geq \mathcal{P}_1 \geq \cdots \geq \mathcal{P}_{n-l+r} \geq 0.$$ 

Let

$$\overline{M}^{(n)} = \{\mathcal{L}_1, \ldots, \mathcal{L}_{n-r}, \mathcal{P}_1, \ldots, \mathcal{P}_{n-l+r}\}$$

be a multiset consisting of $\mathcal{L}$ and $\mathcal{P}$. All elements in $\overline{M}^{(n)}$ are listed in non-increasing order. From the definitions of $\Phi$ and $\Psi$, it is easy to see that $\overline{M}^{(n)}$ equals $M^{(n)}$ defined in (3.12). Since $k > l$, we derive that $\mathcal{P}_{k-l} = \lambda_{k-l} \geq k-s$ which contradicts the fact that $\mathcal{P}_{k-l} \leq l-s < k-s$. Applying the same argument yields that $\des(\alpha^{(i)}) = l > k$ is also impossible. Hence we arrive at the conclusion that $\des(\alpha^{(i)}) = k$. Therefore, $\Psi$ is a map from $P_{k-r}(k-s, m) \times P_{l-r}(l-s, s)$ to $S_{k}(\sigma, \pi)$.

From the definitions of $\Phi$ and $\Psi$, it is easy to see that $\Psi(\Phi(\alpha)) = \alpha$ for each $\alpha \in S_{k}(\sigma, \pi)$ and $\Phi(\Psi(\lambda, \mu)) = (\lambda, \mu)$ for each $(\lambda, \mu) \in P_{k-r}(k-s, m) \times P_{l-r}(l-s, s)$. Hence $\Phi$ is a bijection between $S_{k}(\sigma, \pi)$ and $P_{k-r}(k-s, m) \times P_{l-r}(l-s, s)$. This completes the proof.

For example, let $\sigma = 938101247 \in S_7, \pi = 12651311 \in S_6$, where $\des(\sigma) = 2$ and $\des(\pi) = 2$. Given $k = 5, \lambda = (6, 4, 3)$ and $\mu = (3, 2, 2)$, we will recover the shuffle $\alpha$ of $\sigma$ and $\pi$ as follows. The elements of $\pi$ in $\alpha^{(i)}$ are in boldface to distinguish them from the elements of $\sigma$.

| $i$ | $\pi_i$ | $T^{(i)}$ | $M^{(i)}$ | $k_i$ | $\alpha^{(i)}$ |
|-----|---------|-----------|----------|------|----------------|
| 6   | 11      | (3, 2, 4, 5, 1, 6, 7, 0) | {6, 4, 3, 3, 2, 2} | 6    | 938101247     |
| 5   | 13      | (3, 2, 4, 5, 6, 1, . . .) | {4, 3, 3, 2, 2} | 3    | 938101247     |
| 4   | 5       | (3, 4, 2, . . .) | {3, 3, 2, 2} | 3    | 931381012     |
| 3   | 6       | (2, 3, 4, . . .) | {3, 3, 2} | 2    | 9351381012     |
| 2   | 2       | (3, 2, . . .) | {3, 2} | 2    | 96351381012    |
| 1   | 1       | (3, 2, . . .) | {3} | 1    | 926351381012   |
| 0   |         |          | $\emptyset$ |     | 1926351381012 |

Hence $\alpha = \alpha^{(i)} = 19263513810121147$.

Acknowledgment. This work was supported by the National Science Foundation of China.

References

[1] G.E. Andrews, The Theory of Partitions, Addison-Wesley Publishing Co., 1976.
[2] A. M. Garsia and I. M. Gessel, Permutation statistics and partitions, Adv. in Math. 31 (1979) 288–305.
[3] I. P. Goulden, A bijective proof of Stanley’s shuffling theorem, Trans. Amer. Math. Soc. 288 (1985), no. 1, 147—160.
[4] J. Haglund, N. Loehr and J.B. Remmel, Statistics on wreath products, perfect matchings, and signed words, European J. Combin. 26 (2005) 835–868.
[5] M. Novick, A bijective proof of a major index theorem of Garsia and Gessel, Electron. J. Combin. 17 (2010), no. 1, Research Paper 64, 12 pp.
[6] J. D. Stadler, Stanley’s shuffling theorem revisited, J. Combin. Theory Ser. A 88 (1999), no. 1, 176–187.
[7] R. P. Stanley, Ordered structures and partitions, Memoirs of the American Mathematical Society, No. 119. American Mathematical Society, Providence, R.I., 1972. iii+104 pp.
[8] R.P. Stanley, Enumerative Combinatorics, Vol. I, 2nd ed., Cambridge University Press, Cambridge, 2012.
[9] P.A. MacMahon, Collected Papers. Vol. I, MIT Press, Cambridge, MA, 1978.