A BIJECTIVE PROOF
OF STANLEY'S SHUFFLING THEOREM

BY
I. P. GOULDEN

Abstract. For two permutations \( \sigma \) and \( \omega \) on disjoint sets of integers, consider forming a permutation on the combined sets by "shuffling" \( \sigma \) and \( \omega \) (i.e., \( \sigma \) and \( \omega \) appear as subsequences). Stanley [10], by considering \( P \)-partitions and a \( q \)-analogue of Saalschütz's \( \, _3F_2 \) summation, obtained the generating function for shuffles of \( \sigma \) and \( \omega \) with a given number of falls (an element larger than its successor) with respect to greater index (sum of positions of falls). It is a product of two \( q \)-binomial coefficients and depends only on remarkably simple parameters, namely the lengths, numbers of falls and greater indexes of \( \sigma \) and \( \omega \). A combinatorial proof of this result is obtained by finding bijections for lattice path representations of shuffles which reduce \( \sigma \) and \( \omega \) to canonical permutations, for which a direct evaluation of the generating function is given.

1. Introduction. For a sequence \( a = a_1 \cdots a_n \) of integers \( a_1, \ldots, a_n \), we define the descent set of \( a \), denoted by \( \mathcal{D}(a) \), by \( \mathcal{D}(a) = \{ i | a_i > a_{i+1} \} \), the number of descents in \( a \) by \( d(a) = |\mathcal{D}(a)| \), and the greater index of \( a \) by \( I(a) = \sum_{i \in \mathcal{D}(a)} i \). We say that \( a \) has length \( |a| = n \).

Let \( \alpha = (\alpha_1, \ldots, \alpha_m) \) and \( \beta = (\beta_1, \ldots, \beta_n) \) be disjoint subsets of \( \mathcal{N}_{m+n} = \{1, \ldots, m+n\} \), where \( \alpha_1 < \cdots < \alpha_m \) and \( \beta_1 < \cdots < \beta_n \). For any permutation \( \sigma \) of the elements of \( \alpha \), and any permutation \( \omega \) of the elements of \( \beta \), we say that \( \sigma \) and \( \omega \) are \((m, n)\)-compatible.

The shuffle product of \( \sigma \) and \( \omega \), denoted by \( \mathcal{S}(\sigma, \omega) \), is the set of all permutations of \( \mathcal{N}_{m+n} \) in which \( \sigma \) and \( \omega \) both appear as subsequences. The following result is worth recording, since it leads to the lattice path representation of shuffle products in \( \S 2 \).

Proposition 1.1. If \( \sigma \) and \( \omega \) are \((m, n)\)-compatible, then

\[ |\mathcal{S}(\sigma, \omega)| = \binom{m+n}{m}. \]

Proof. There is a bijection between elements \( \rho \) of \( \mathcal{S}(\sigma, \omega) \) and subsets \( \alpha = (\alpha_1, \ldots, \alpha_m) \) of \( \mathcal{N}_{m+n} \) defined as follows. If \( \alpha_1 < \cdots < \alpha_m \) and \( \sigma = \sigma_1 \cdots \sigma_m \), then \( \rho \) contains \( \sigma_i \) in position \( \alpha_i \) for \( i = 1, \ldots, m \).

The elements of \( \omega \) appear, in order, in the set of positions of \( \rho \) complementary to \( \alpha \).

Received by the editors January 31, 1984.
1980 Mathematics Subject Classification. Primary 05A10, 05A15; Secondary 06A99, 68E05.

©1985 American Mathematical Society
0002-9947/85 $1.00 + $ .25 per page

147
Consider the generating functions $S_k(\sigma, \omega)$ for the shuffle product $S(\sigma, \omega)$, defined by

$$S_k(\sigma, \omega) = \sum_{\rho \in S_k(\sigma, \omega)} q^{l(\rho)}.$$

Stanley [10] has obtained a compact expression for $S_k(\sigma, \omega)$ in terms of the Gaussian (or $q$-binomial) coefficient $[\cdot]$, defined for nonnegative integers $j$ by

$$[i] = \frac{(1 - q^i) \cdots (1 - q^{i+j-1})}{(1 - q^i) \cdots (1 - q)},$$

and $[\cdot] = 0$ otherwise.

**Theorem 1.2 (Shuffling Theorem).** Let $\sigma$ and $\omega$ be $(m, n)$-compatible, with $d(\sigma) = r$, $d(\omega) = s$. Then

$$S_k(\sigma, \omega) = q^{l(\sigma) + l(\omega) + (k-r)(k-s)} \left[ \begin{array}{c} m - r + s \\ k - r \\ k - s \\ k - s \\ n - s + r \\ m - r + s \end{array} \right].$$

The case $r = s = 0$ had been previously given by MacMahon [8, Vol. II, p. 210]. Stanley obtained the Shuffling Theorem by means of his theory of $P$-partitions, and by applying the following identity.

**Theorem 1.3.**

$$\sum_{i \geq 0} q^{r-i}(s-i) \left[ \begin{array}{c} m + r - s \\ r - i \\ s - i \\ m + s - r \\ n + s - r \\ m + n + i \\ i \\ i \\ i \end{array} \right] = \left[ \begin{array}{c} m + r \\ n + s \\ m \\ n \end{array} \right].$$

This identity was proved by Gould [4], and is equivalent to Jackson's [7] $q$-analogue of Saalschütz's theorem (see [9, p. 243]). Combinatorial proofs of Saalschütz's theorem (a $3F_2$ summation equivalent to the case $q = 1$ of Theorem 1.3) have been given by Andrews [1] and Cartier and Foata [3].

Stanley [10] has asked for a proof of Theorem 1.2 which avoids the use of Theorem 1.3. In this paper we present such a proof. Basic to our treatment is the combinatorial interpretation of the Gaussian coefficient $[\cdot]$ as the generating function for integer partitions with at most $j$ parts, and largest part at most $i - j$, where $i$ and $j$ are nonnegative integers.

**Lemma 1.4.**

1. \[ \sum_{0 \leq \alpha_1 \leq \cdots \leq \alpha_j \leq i-j} q^{\alpha_1 + \cdots + \alpha_j} = [i]. \]

2. \[ \sum_{1 \leq \beta_1 < \cdots < \beta_i \leq i} q^{\beta_1 + \cdots + \beta_i} = q^{\binom{i+1}{2}} [i]. \]

**Proof.** 1. See Andrews [2, p. 33] for a proof; historical references are given on p. 51.
2. Obtained from (1) by letting $\beta_m = \alpha_m + m$, for $m = 1, \ldots, j$, since $1 + 2 + \cdots + j = \binom{j+1}{2}$. □

In §2 we discuss lattice paths and their relationship to shuffle products. A bijective proof of the Shuffling Theorem is given in §3.

2. Lattice paths. Suppose that $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are pairs of integers with $u_1 < v_1$ and $u_2 < v_2$. Then $B(u, v) = \{u_1, u_1 + 1, \ldots, v_1\} \times \{u_2, \ldots, v_2\}$ is called a grid, and we shall denote $B(u, v)$ by $B$ in this section when the context allows. We consider lattice paths on a grid, with horizontal and vertical steps. In particular, let $a_i = (a_{i1}, a_{i2}) \in B$ for $i = 0, \ldots, k$, and let $d_i = a_i - a_{i-1}$ for $i = 1, \ldots, k$. Then if $d_i \in \{(1,0), (0,1)\}$ for $i = 1, \ldots, k$, $a = a_0 \cdots a_k$ is called a path on $B$, from $a_0$ to $a_k$, of length $|a| = k$. A path of length 0 (a single vertex) is called an empty path. The $i$th difference $d_i$ is called the $i$th step and $a_i$ is the $i$th vertex in the path $a$. We say that $d_i$ follows $a_{i-1}$ and precedes $a_{i+1}$. The difference $(1,0)$ is called a step across, and $(0,1)$ is a step up. The vertex $a_i$ is, for $i = 1, \ldots, k - 1$,

(i) an upper corner if $d_{i-1} = (0,1)$ and $d_i = (1,0)$,
(ii) a lower corner if $d_{i-1} = (1,0)$ and $d_i = (0,1)$,
(iii) a horizontal crossing of $x = a_{i1}$ if $d_{i-1} = d_i = (1,0)$,
(iv) a vertical crossing of $y = a_{i2}$ if $d_{i-1} = d_i = (0,1)$.

If $b = b_0 b_1 \cdots b_j$ is a path on $B$, then the product $ab$ is defined when $a_k = b_0$ by $ab = a_0 a_1 \cdots a_k b_1 \cdots b_j$, and is not defined otherwise.

Note that a path is uniquely specified by its end-points and either its upper corners or lower corners.

**Proposition 2.1.** If $u_1 \leq x_0 \leq x_1 < \cdots < x_k \leq v_1$ and $u_2 \leq y_0 \leq y_1 < \cdots < y_{k-1} \leq y_k \leq v_2$ are integers, then there is a unique path on $B(u, v)$ from $(x_0, y_0)$ to $(x_k, y_k)$ with upper corners at $(x_1, y_1), \ldots, (x_{k-1}, y_{k-1})$, and no other upper corners.

If $u_1 \leq x_0 < x_1 < \cdots < x_{k-1} \leq x_k \leq v_1$ and $u_2 \leq y_0 < y_1 < \cdots < y_k \leq v_2$ are integers, then there is a unique path on $B(u, v)$ from $(x_0, y_0)$ to $(x_k, y_k)$ with lower corners at $(x_1, y_1), \ldots, (x_{k-1}, y_{k-1})$, and no other lower corners.

**Proof.** For upper corners, the path is $\rho_1 \cdots \rho_k$, where

$$\rho_i = (x_{i-1}, y_{i-1})(x_{i-1} + 1, y_{i-1}) \cdots (x_i, y_{i-1})(x_i, y_{i-1} + 1) \cdots (x_i, y_i).$$

For lower corners, the path is $\delta_1 \cdots \delta_k$, where

$$\delta_i = (x_{i-1}, y_{i-1})(x_{i-1}, y_{i-1} + 1) \cdots (x_{i-1}, y_i)(x_{i-1} + 1, y_i) \cdots (x_i, y_i).$$

For compactness, we also denote a path by its sequence of steps, using "A" for steps across, and "U" for steps up, subscripted by its initial vertex. If the initial vertex is $(0,0)$, then we suppress the subscript.

The path $a$ on $B(u, v)$ is said to cover $B$ (or to be a cover of $B$) if $a_0 = u$ and $a_k = v$. If $a$ covers $B$ then it partitions $B$ into 3 sets, consisting of the points in $B$ that are

(i) on $a$ (points $(t_1, t_2)$ such that $t_1 = a_{i1}, t_2 = a_{i2}$ for some $i = 0, \ldots, k$),
(ii) above $a$ (points $(t_1, t_2)$ such that $t_1 < a_{i1}, t_2 > a_{i2}$ for some $i = 0, \ldots, k$),
(iii) below $a$ (points $(t_1, t_2)$ such that $t_1 > a_{i1}, t_2 < a_{i2}$ for some $i = 0, \ldots, k$).
A path $b$ on $B$ is called a $\leq$ $a$-path if $b$ is nonempty, and all vertices of $b$ are on or below $a$. A path $b$ on $B$ is called a $>a$-path if $b$ is nonempty, and all vertices of $b$ are above $a$, except the first and last vertices of $b$, which may also be on $a$, but not both on $a$ if $|a| = 1$. For example, if $a = UAUA^2UA$, then $U$, $A^2U$ and $(UA^2U)_{(1,2)}$ are $\leq a$-paths, while $(UA^2)_{(0,1)}$ and $(UA^2)_{(0,3)}$ are $>a$-paths. The path $U^2A$ is neither a $\leq a$-path nor a $>a$-path.

Note that the use of "above" and "below" corresponds to the obvious meanings of these words in a geometric representation of a path. The constructions which are given later in this paper involve many parameters, and require a fair amount of notation and terminology to state accurately. It is intended that the terminology used throughout this paper be natural in the geometric representations of these constructions, though no pictures will be supplied by the author.

The cover $(A^v_{(v'-v)}U^{v_2-v_2})_a$ is called the canonical cover of $B(u, v)$.

If $a$ covers $B$ and $b$ is a path on $B$, then we define $\mathcal{C}_B(a)(b)$ to be the set of all upper corners of $b$ that are above $a$ and all lower corners of $b$ that are below $a$. If $c$ is the canonical cover of $B$, then $\mathcal{C}_B(c)(b)$ is more simply described as the set of all upper corners in $b$.

In order to define generating functions for sets of paths, we must define a weight function for lattice paths. Let the weight of a vertex $e = (e_1, e_2)$ be $w(e) = e_1 + e_2$, and the weight of a set $e = \{e_1, \ldots, e_k\}$ of vertices be $w(e) = w(e_1) + \cdots + w(e_k)$.

The following weight-preserving mapping $\psi$ for paths is very important to our proof of Theorem 1.2. Suppose that $a$ covers $B$ and $g$ is a $\leq a$-path on $B$ from $z_1$ to $z_2$ with lower corners below $a$ given by $(f_{11}, f_{12}), \ldots, (f_{k1}, f_{k2})$. Then we define $\psi_B(a)(g)$ to be the path from $z_1$ to $z_2$ whose upper corners are $(f_{11} - 1, f_{12} + 1), \ldots, (f_{k1} - 1, f_{k2} + 1)$. (This path is unique, by Proposition 2.1.) If $b$ is a path on $B$, then we can write $b$ uniquely as $b = h_xg_xh_2g_2\cdots h_lg_l$, for some $l \geq 1$, where $g_1, \ldots, g_{l-1}$ are $\leq a$-paths, $g_l$ is either a $\leq a$-path or empty, $h_1$ is either a $>a$-path or empty, and $h_2, \ldots, h_l$ are $>a$-paths. Then we define $\psi_B(a)(b) = h_1\psi_B(a)(g_1)\cdots h_l\psi_B(a)(g_l)$, where $\psi_B(a)(g_i), i = 1, \ldots, l$, are given above. (If $g_i$ is empty, then $\psi_B(a)(g_i) = g_i$.)

For example, if $B = B((0,0), (7,6))$, $a = AU^2AUA^3UAU^2A$ and $b = (A^3U^2A^3UAU)_{(0,1)}$, then $b = h_1g_1h_2g_2$, where $h_1 = (A)_{(0,1)}$ and $h_2 = (UA^2)_{(3,3)}$ are $>a$-paths, while $g_1 = (A^2U^2)_{(1,1)}$ and $g_2 = (AUAU)_{(5,4)}$ are $\leq a$-paths. Now $\psi_B(a)(g_1) = (AUAU)_{(1,1)}$ and $\psi_B(a)(g_2) = (AUA^2)_{(5,4)}$, so

$\psi_B(a)(b) = (A^2UAUA^3U^2A)_{(0,1)}$.

Note that $\mathcal{C}_B(a)(b) = \{(3,1), (3,4), (7,5)\}$ and $\mathcal{C}_B(c)(\hat{b}) = \{(2,2), (3,4), (6,6)\}$, where $\hat{b} = \psi_B(a)(b)$ and $c = A^T U^3$ is the canonical cover of $B$. Thus $|\mathcal{C}_B(a)(b)| = 3 = |\mathcal{C}_B(c)(\hat{b})|$ and $w(\mathcal{C}_B(a)(b)) = 4 + 7 + 12 = w(\mathcal{C}_B(c)(\hat{b}))$, equalities that are proved to hold in general in the following result.

**Lemma 2.2.** Let $x, y \in B$, and define $\mathcal{P}_B(x, y)$ to be the set of paths on $B$ from $x$ to $y$. Then for any cover $a$ of $B$,

1. $\psi_B(a) : \mathcal{P}_B(x, y) \to \mathcal{P}_B(x, y) : b \mapsto \hat{b}$ is a bijection.
Moreover, if \( c \) is the canonical cover of \( B \), then

2. \( |\mathcal{C}_{B,a}(b)| = |\mathcal{C}_{B,c}(\hat{b})| \).
3. \( w(\mathcal{C}_{B,a}(b)) = w(\mathcal{C}_{B,c}(\hat{b})). \)

**Proof.** 1. First note that if \( g \) is a \( \leq a \)-path from \( z_1 \) to \( z_2 \), then \( \hat{g} = \psi_{B,a}(g) \) is also a \( \leq a \)-path from \( z_1 \) to \( z_2 \). This is because if the lower corners of \( g \) below \( a \) occur at \((f_{11}, f_{21}), \ldots, (f_{1k}, f_{2k})\), then the upper corners of \( \hat{g} \) occur at \((f_{11} - 1, f_{21} + 1), \ldots, (f_{1k} - 1, f_{2k} + 1)\), each of which must lie on or below \( a \). Thus \( \hat{g} \) is a path from a point \((z_1)\) on or below \( a \), to a point \((z_2)\) on or below \( a \), in which all upper corners are on or below \( a \). Thus Proposition 2.1 tells us that \( \hat{g} \) is unique, and is also a \( \leq a \)-path. Moreover \( g \) is recoverable from \( \hat{g} \) as follows. Suppose that \( \hat{g} \) is a \( \leq a \)-path from \( z_1 \) to \( z_2 \), whose upper corners are \((c_{11}, c_{21}), \ldots, (c_{1k}, c_{2k})\). Let \( \hat{g} \) be the unique path from \( z_1 \) to \( z_2 \) whose lower corners are \((c_{11} + 1, c_{21} - 1), \ldots, (c_{1k} + 1, c_{2k} - 1)\), given by Proposition 2.1. Now \( \hat{g} \) is not necessarily a \( \leq a \)-path, but we can write \( \tilde{g} \) uniquely as \( \tilde{g} = d_1 e_1 d_2 \cdots e_{m-1} d_m \) for some \( m \geq 1 \), where \( d_1, \ldots, d_m \) are \( \leq a \)-paths (\( d_1 \) and \( d_m \) can also be empty) and \( e_1, \ldots, e_{m-1} \) are \( > a \)-paths. Also, all lower corners of \( \hat{g} \) are below \( a \) since \((c_{1i}, c_{2i})\) is on or below \( a \), so \((c_{1i} + 1, c_{2i} - 1)\) must be below \( a \) for \( i = 1, \ldots, k \). Thus the lower corners of \( \hat{g} \) must all be internal vertices in one of the paths \( d_1, \ldots, d_m \). The path \( e_i \), for \( i = 1, \ldots, m - 1 \) is a path from a vertex on \( a \), say \( a_{i-1} \), to a vertex on \( a \), say \( a_i \), with a single corner (upper). For \( i = 1, \ldots, m - 1 \), let \( r_i \) be the segment of \( a \) from \( a_{i-1} \) to \( a_i \). Then \( r_i \), of course, has no lower corners below \( a \), since \( r_i \) has no vertices below \( a \), and we have \( g = d_1 r_1 d_2 \cdots r_{m-1} d_m \), so \( \psi_{B, a}^{-1} \) exists for \( \leq a \)-paths.

Now, if \( b = h_1 g_1 \cdots h_l g_l \in \mathcal{P}_{B}(x, y) \) in the notation of the definition of \( \psi_{B,a} \) above, then \( \hat{b} = h_1 \hat{g}_1 \cdots h_l \hat{g}_l \in \mathcal{P}_{B}(x, y) \), where \( \hat{g}_i = \psi_{B,a}(g_i) \) for \( i = 1, \ldots, l \), the \( h_i \)'s are \( > a \)-paths (by definition), and the \( \hat{g}_i \)'s are \( \leq a \)-paths (from above). Thus \( \psi_{B,a}^{-1} \) is well defined, so \( \psi_{B,a} \) is a bijection.

2 and 3. From the description of \( \psi_{B,a} \) above, we know that if \( \hat{g} = \psi_{B,a}(g) \), where \( g \) is a \( \leq a \)-path, then \( |\mathcal{C}_{B,a}(g)| = |\mathcal{C}_{B,c}(\hat{g})| \) \( (= k \) above) and \( w(\mathcal{C}_{B,a}(g)) = w(\mathcal{C}_{B,c}(\hat{g})) \) \( (= f_{11} + f_{21} + \cdots + f_{1k} + f_{2k} \) above). Again let \( b = h_1 g_1 \cdots h_l g_l \) and \( \hat{b} = h_1 \hat{g}_1 \cdots h_l \hat{g}_l \). Then

\[
|\mathcal{C}_{B,a}(g)| = \sum_{i=1}^{l} |\mathcal{C}_{B,a}(h_i)| + |\mathcal{C}_{B,a}(g_i)|
\]

and

\[
|\mathcal{C}_{B,c}(\hat{g})| = \sum_{i=1}^{l} |\mathcal{C}_{B,c}(h_i)| + |\mathcal{C}_{B,c}(\hat{g}_i)|
\]

since the intersecting vertex of a \( \leq a \)-path (like \( g_i \) or \( \hat{g}_i \)) and a \( > a \)-path (like \( h_i \)) must be on \( a \), and cannot be an upper corner. But \( \mathcal{C}_{B,a}(h_i) = \mathcal{C}_{B,c}(h_i) \) for \( i = 1, \ldots, l \) since \( h_i \) is a \( > a \)-path, and \( |\mathcal{C}_{B,a}(g_i)| = |\mathcal{C}_{B,c}(\hat{g}_i)| \) from above, so \( |\mathcal{C}_{B,a}(g)| = |\mathcal{C}_{B,c}(\hat{g})| \). The proof that \( w(\mathcal{C}_{B,a}(g)) = w(\mathcal{C}_{B,c}(\hat{g})) \) is similar. \( \square \)

Finally, denote the grid \( B(0,0), (m, n) \) by \( G \), and let the set of paths from \((0,0)\) to \((m, n)\), which are the covers of \( G \), be denoted by \( \mathcal{P}(m, n) \). (Note that \( \mathcal{P}(m, n) = \mathcal{P}_G((0,0), (m, n)) \) in the notation of Lemma 2.2.)
Now we relate lattice paths to shuffle products of permutations. For \((m, n)\)-compatible permutations \(\sigma, \omega\), we represent the permutation \(\rho \in \mathcal{P}(\sigma, \omega)\) by the path \(\phi_{\sigma, \omega}(\rho) \in \mathcal{P}(m, n)\) as follows. If the \(i\)th element of \(\rho\) is in \(\sigma\), then the \(i\)th step in \(\phi_{\sigma, \omega}(\rho)\) is across, and if the \(i\)th element of \(\rho\) is in \(\omega\), then the \(i\)th step in \(\phi_{\sigma, \omega}(\rho)\) is up for \(i = 1, \ldots, m + n\). We say that the \(i\)th step of \(\phi_{\sigma, \omega}(\rho)\) represents the \(i\)th element of \(\rho\). For example, \(\rho_0 = 647325819 \in \mathcal{P}(6358, 47219)\) is represented by \(\phi(\rho_0) = AU^2AU^2U^2 \in \mathcal{P}(4, 5)\).

**Proposition 2.3.** If \(\sigma\) and \(\omega\) are \((m, n)\)-compatible permutations, then \(\phi_{\sigma, \omega}: \mathcal{P}(\sigma, \omega) \to \mathcal{P}(m, n)\) is a bijection.

**Proof.** Immediate from Proposition 2.1, since a subset of \(\mathcal{N}_{m+n}\) of cardinality \(m\) uniquely specifies the path in \(\mathcal{P}(m, n)\) whose \(m\) steps across occur in positions belonging to that subset. \(\square\)

3. The Shuffling Theorem. In this section, we establish the Shuffling Theorem by a sequence of bijections for lattice paths and permutations. First we need some additional notation.

Let \(0 = t_0 < \cdots < t_{r+1} = m, 0 = l_0 < \cdots < l_{s+1} = n\), \(t = \{t_1, \ldots, t_r\}\) and \(l = \{l_1, \ldots, l_s\}\). Let \(B_{ij}\) be the grid \(B((t_i, l_j), (t_{i+1}, l_{j+1}))\) for \(i = 0, \ldots, r, j = 0, \ldots, s\), so \(\bigcup_{i=0}^{r} \bigcup_{j=0}^{s} B_{ij} = G\). The grids \(B_{ij}\) and \(B_{i+1,j}\) intersect in the segment of \(y = l_{j+1}\) from \((t_i, l_{j+1})\) to \((t_{i+1}, l_{j+1})\), and the grids \(B_{ij}\) and \(B_{i,j+1}\) intersect in the segment of \(x = t_{i+1}\) from \((t_{i+1}, l_j)\) to \((t_{i+1}, l_{j+1})\). These segments are called borders for the grids to which they belong. If \(b \in \mathcal{P}(m, n)\), define \(e_{ij}(b)\) to be the maximal subpath of \(b\) on \(B_{ij}\). Let \(\mathcal{Y}(b)\) be the set of vertical crossings of \(y\)-coordinates in \(l\), and \(\mathcal{X}(b)\) be the set of horizontal crossings of \(x\)-coordinates in \(t\). Suppose that \(a\) is an array with \((i, j)\)-entry \(a_{ij}\) for \(i = 0, \ldots, r, j = 0, \ldots, s\), where \(a_{ij}\) covers \(B_{ij}\). Then define

\[
\mathcal{F}_{t, a}(b) = \mathcal{X}(b) \cup \mathcal{Y}(b) \cup \bigcup_{i=0}^{r} \bigcup_{j=0}^{s} \mathcal{C}_{B_{ij}}(e_{ij}(b)),
\]

and

\[
P_k(t, l, a) = \sum_{b \in \mathcal{P}(m, n)} q^{|\mathcal{F}_{t, a}(b)|}.
\]

Finally we say that \(a\) is legitimate if no pair of distinct \(a_{ij}\)'s have a nonempty path as their intersection. Note that if a pair of \(a_{ij}\)'s have a nonempty path as their intersection, then the intersection path must lie on the mutual border of the corresponding \(B_{ij}\)'s.

For example, let \(m = 9, n = 6, \ r = 2, \ s = 1, \ t = \{2, 6\}, \ l = \{3\}, \ a_{00} = UA^2U^2, \ a_{01} = (U^2A^2)(2, 3), \ a_{10} = (U^2A^4U)(2, 0), \ a_{11} = (A^2U^3)(2, 3), \ a_{20} = (UAUA^2U)(6, 0), \ a_{21} = (A^3U^3)(6, 3)\) and \(b = AU^2A^3U^2A^2A^3\). Then \(\mathcal{Y}(b) = \{(4, 3)\}\), \(\mathcal{X}(b) = \{(2, 2)\}\), \(e_{00}(b) = AU^2A\), \(e_{10}(b) = (A^2U^2)(2, 3)\), \(e_{11}(b) = (AU^2U^2)(4, 3)\), \(e_{20}(b) = (U^2A^3)(6, 4)\) and \(e_{01}(b) = e_{20}(b) = \varnothing\). Thus \(\mathcal{C}_{B_{00, a_{00}}}(e_{00}(b)) = \{(1, 0), \ (1, 2)\}\), \(\mathcal{C}_{B_{11, a_{11}}}(e_{11}(b)) = \{(4, 4)\}\), \(\mathcal{C}_{B_{21, a_{21}}}(e_{21}(b)) = \{(6, 6)\}\) and \(\mathcal{C}_{B_{ij, a_{ij}}}(e_{ij}(b)) = \varnothing\) for other \(i, j\). Note that \(a\) is not legitimate since \(a_{00}\) and \(a_{10}\) have the path \((U)(2, 1)\) in common, where \((U)(2, 1)\) is on the border shared by \(B_{00}\) and \(B_{10}\). Also note that each corner in \(b\) occurs, as a
corner, in a unique \( e_{ij}(b) \), though distinct \( e_{ij} \) can have a nonempty path (again a portion of mutual border) as their intersection. For example \( e_{11}(b) \) and \( e_{21}(b) \) have the path \( (U^2)_{(6,4)} \) in common in the above example.

We now give the first of the bijections that will allow us to deduce the Shuffling Theorem. Examples of all of the results which lead to the Shuffling Theorem are contained in Example 3.6, at the end of this section.

**Lemma 3.1.** Let \( \sigma = \sigma_1 \cdots \sigma_m \) and \( \omega = \omega_1 \cdots \omega_n \) be \( (m, n) \)-compatible, with \( D(\sigma) = t \) and \( D(\omega) = 1 \). If \( a_{ij} \) is the cover of \( B_{ij} \) representing the shuffle of \( \sigma_{i+1} \cdots \sigma_{i+j} \) and \( \omega_{l+1} \cdots \omega_{l+j} \) into increasing order, then

1. \( S_k(\sigma, \omega) = P_k(t, l, a) \).

2. \( a \) is legitimate.

**Proof.** 1. Let \( p \in D(\sigma, \omega) \) and let \( b = \phi_{\sigma, \omega}(\rho) \in \mathcal{D}(m, n) \). Suppose that \( \rho_i \) is the \( i \)th element of \( \rho \), and \( b_i = (b_{1i}, b_{2i}) \) is the vertex of \( b \) that follows that \( i \)th step, so \( w(b_i) = i \).

If \( b_i \) is the horizontal crossing in \( b \), then \( \rho_i = \sigma_{b_{1i}} \) and \( \rho_{i+1} = \sigma_{b_{1i+1}} \), so \( \rho_i > \rho_{i+1} \) if and only if \( b_{1i} \in D(\sigma) = t \), or equivalently, \( b_i \in \mathcal{K}(b) \). If \( b_i \) is a vertical crossing in \( b \), then \( \rho_i = \omega_{b_{2i}}, \rho_{i+1} = \omega_{b_{2i+1}} \), so \( \rho_i > \rho_{i+1} \) if and only if \( b_{2i} \in D(\omega) = 1 \), or, equivalently, \( b_i \in \mathcal{K}(b) \).

If \( b_i \) is a corner in \( b \), then \( b_i \) appears as a corner in \( \varepsilon_{di}(b) \), say, and in no other \( \varepsilon_{ij}(b) \). If \( b_i \) is an upper corner, then \( \rho_i = \omega_{b_{2i}} \), and \( \rho_{i+1} = \omega_{b_{2i+1}} \), so \( \rho_i > \rho_{i+1} \) if and only if \( \rho_{i+1} < \rho_i \). Similarly, if \( b_i \) is a lower corner, then \( b_i \) is below \( a_{di} \), and \( \rho_{i+1} \) occurs before \( \rho_i \) in \( a_{di} \), so \( \rho_i > \rho_{i+1} \). Thus we have a bijection between descents \( i \in D(\rho) \) and vertices \( b_i \in \mathcal{F}_{1,1}(b) \), where \( w(b_i) = i \). This immediately gives \( d(\rho) = |\mathcal{F}_{1,1}(b)| \) and \( I(\rho) = w(\mathcal{F}_{1,1}(b)) \) so, from Proposition 2.3, we have

\[
S_k(\sigma, \omega) = \sum_{\rho \in D(\sigma, \omega)} q^d(\rho) = \sum_{b \in D(m, n)} q^w(\mathcal{F}_{1,1}(b)) = P_k(t, l, a),
\]
as required.

2. Suppose that \( a_{ij-1} \) and \( a_{ij} \) have a nonempty path in common. Then this path must be \( (f, l_1) \cdots (g, l_j) \) for some \( f, g \) with \( t_i < f < g < t_{i+1} \). But the next vertex in \( a_{ij} \) after \( (g, l_j) \) must be \( (g, l_j + 1) \), so by definition of \( a_{ij} \), \( \sigma_{f+1} < \cdots < \sigma_{g} < \omega_{l+1} \). Similarly \( \sigma_{f+1} > \omega_{l+1} \). By considering \( a_{ij-1} \), and we deduce that \( \omega_{l+1} < \sigma_{f+1} < \omega_{l+1} \), so \( \omega_{l+1} < \omega_{l+1} \). But, by definition, \( l_j \in D(\omega) \) so \( \omega_{l+1} > \omega_{l+1} \) and we have arrived at a contradiction. Thus \( a_{ij-1} \) and \( a_{ij} \) have at most one vertex in common, for all \( i, j \).

Similarly, we can show that \( a_{i-1,j} \) and \( a_{ij} \) have at most one vertex in common, for all \( i, j \), and conclude that \( \mathcal{A} \) is legitimate.  

To proceed from here, it is convenient to define the following total order for the set \( \mathcal{A} = \{0, \ldots, r\} \times \{0, \ldots, s\} \). If \((r_1, s_1) \) and \((r_2, s_2) \) are in \( \mathcal{A} \), then we say that \((r_1, s_1) < (r_2, s_2) \) if \( s_1 < s_2 \) or if \( s_1 = s_2 \) and \( r_1 < r_2 \). Thus the arrangement of \( \mathcal{A} \) in increasing order is \( (r, 0), (r - 1, 0), (r, 1), \ldots, (0, s - 1), (1, s), (0, s) \). Now let \( \mathcal{C} \) be the array whose \((i, j)\)-entry is \( c_{ij} \), the canonical cover of \( B_{ij} \), for \( i = 0, \ldots, r, j = 0, \ldots, s \).
For $i = 0, \ldots, (r + 1)(s + 1)$, let $a^{(i)}$ be the array obtained from $a$ by replacing the first $i$ elements (in terms of the above total order) of $a$ by the first $i$ elements of $c$, so $a^{(0)} = a$, and $a^{(r + 1)(s + 1)} = c$.

For $b \in \mathcal{P}(m, n)$, we define $b^{(i)} \in \mathcal{P}(m, n)$ for $i = 0, \ldots, (r + 1)(s + 1)$, recursively as follows. Let $b^{(0)} = b$. For $i = 1, \ldots, (r + 1)(s + 1)$:

(i) Let $(\alpha, \beta)$ be the $i$th element of $b$.

(ii) If $\varepsilon_{\alpha\beta}(b^{(i - 1)}) = \emptyset$, then $b^{(i)} = b^{(i - 1)}$.

(iii) If $\varepsilon_{\alpha\beta}(b^{(i - 1)}) = \delta$ is a path in $B_{\alpha\beta}$, then $b^{(i - 1)} = \delta_1 \delta_2$ for unique paths $\delta_1, \delta_2$, each with a single vertex in $B_{\alpha\beta}$. Let $b^{(i)} = \delta_1 \psi_{\beta_\alpha\beta}(\delta) \delta_2$.

For example, if $m = 10$, $n = 8$, $r = s = 1$, $t = \{4\}$, $\ell = \{3\}$, $a_{00} = AUA^2U^2A$, $a_{01} = (A^2U^3A^2)_{(0,3)}$, $a_{10} = (U^3A^2)_{(4,0)}$, $a_{11} = (U^2A^3U^2A^3U)_{(4,3)}$ and $b = A^2U^3A^2U^2A^2U^3A^2U$, then $b^{(0)} = b$, $b^{(1)} = b^{(2)} = A^2U^3AUA^2U^3A^2U$ and $b^{(3)} = b^{(4)} = A^2U^4A^2U^3A^3U$. Now, for compactness, denote $\mathcal{F}(b^{(i)})$ by $\mathcal{F}(i)$. Then, in the above example, $\mathcal{F}(0) = ((4, 0), (6, 2), (6, 3), (8, 4))$, $\mathcal{F}(1) = \mathcal{F}(2) = ((4, 0), (5, 3), (6, 3), (8, 4))$, $\mathcal{F}(3) = \mathcal{F}(4) = ((4, 0), (5, 3), (5, 4), (7, 5))$, so $|\mathcal{F}(i)| = 4$, and $w(\mathcal{F}(i)) = 4 + 8 + 9 + 12 = 33$ for $i = 0, \ldots, 4$. This equality is proved to hold for general for any $a$ that is legitimate, as in this example, in the following result.

**Theorem 3.2.** If $a$ is legitimate, then, for all $k \geq 0$, $P_k(t, l, a) = P_k(t, l, c)$.

**Proof.** If, in the construction of $b^{(i)}$ from $b^{(i - 1)}$ above, we have $\varepsilon_{\alpha\beta}(b^{(i - 1)}) = \emptyset$, or $\varepsilon_{\alpha\beta}(b^{(i - 1)})$ is a single vertex (either top left or bottom right corner of $B_{\alpha\beta}$), then it is immediate that $|\mathcal{F}(i)| = |\mathcal{F}(i - 1)|$ and $w(\mathcal{F}(i)) = w(\mathcal{F}(i - 1))$.

Otherwise, we have $b^{(i - 1)} = \delta_1 \delta_2$, where $\delta = \varepsilon_{\alpha\beta}(b^{(i - 1)})$, and the final vertex, say $v_1$, of $\delta_1$ is in $B_{\alpha\beta}$, and the initial vertex, say $v_2$, of $\delta_2$ is in $B_{\alpha\beta}$. Suppose that $v_1$ is the $j$th vertex in $b^{(i - 1)}$ (and $b^{(i)}$) and that $v_2$ is the $k$th vertex in $b^{(i - 1)}$ (and $b^{(i)}$). Then, for $u = 0, \ldots, j - 1$ and $k + 1, \ldots, m + n$, the $u$th vertex of $b^{(i - 1)}$ is in $\mathcal{F}(i - 1)$ if and only if the $u$th vertex of $b^{(i)}$ is in $\mathcal{F}(i)$, from Lemma 2.2, with $x = v_1$, $y = v_2$, $B = B_{\alpha\beta}$, $a = a_{\alpha\beta}$. But the $u$th vertex in any path in $\mathcal{P}(m, n)$ has weight equal to $u$. Thus we prove $|\mathcal{F}(i)| = |\mathcal{F}(i - 1)|$ and $w(\mathcal{F}(i)) = w(\mathcal{F}(i - 1))$ by proving that $v_1 \in \mathcal{F}(i)$ if and only if $v_1 \in \mathcal{F}(i - 1)$, and $v_2 \in \mathcal{F}(i)$ if and only if $v_2 \in \mathcal{F}(i - 1)$.

Consider first $v_1$. If $a = \beta = 0$, then $v_1 = (0, 0)$, so $\delta_1$ is empty, and $v_1 \notin \mathcal{F}(i - 1)$, $v_1 \notin \mathcal{F}(i)$. Otherwise $v_1$ might lie on the lower border of $B_{\alpha\beta}$, with a step up immediately preceding it. This means that $v_1$ is either a vertical crossing (of $y = l_B$) or an upper corner in $b^{(i - 1)}$ and $b^{(i)}$. But if $v_1$ is an upper corner in either $b^{(i - 1)}$ or $b^{(i)}$, it is an upper corner of $B_{\alpha\beta - 1}$ which is above the canonical cover $c_{\alpha\beta - 1}$. Moreover, by our partial order, $(\alpha, \beta - 1) < (\alpha, \beta)$, so $c_{\alpha\beta}$ is contained in $a^{(i - 1)}$ and $a^{(i)}$. Thus, whether $v_1$ is a vertical crossing or upper corner in $b^{(i - 1)}$ and $b^{(i)}$, we have $v_1 \in \mathcal{F}(i - 1)$ and $v_1 \in \mathcal{F}(i)$.

The other choice for $v_1$ is that it lies on the left border of $B_{\alpha\beta}$, with a step across immediately preceding it. Then $v_1$ appears in $b^{(i - 1)}$ as

(i) a horizontal crossing of $x = t_\alpha$, so $v_1 \in \mathcal{F}(i - 1)$, or
(ii) a lower corner in $B_{a-1\beta}$ below $a_{a-1\beta}$, so $v_1 \in \mathcal{F}^{(i-1)}$, or
(iii) a lower corner in $B_{a-1\beta}$ on $a_{a-1\beta}$, so $v_1 \notin \mathcal{F}^{(i-1)}$.

In case (i) then either (a) $v_1$ is a horizontal crossing in $b^{(i)}$, so $v_1 \in \mathcal{F}^{(i)}$, or (b) $v_1$ is a lower corner in $b^{(i)}$. But (b) can only happen if $v_1$ and $v_1 + (0,1)$ are both on $a_{a\beta}$, which means that $v_1$ is below $a_{a-1\beta}$ (contained in $a^{(i)}$ since $(\alpha - 1, \beta) > (\alpha, \beta)$) since $a$ is legitimate, so $v_1 \in \mathcal{F}^{(i)}$.

In case (ii) $v_1$ appears in $b^{(i)}$ as either a lower corner (below $a_{a-1\beta}$) or perhaps a horizontal crossing, so $v_1 \in \mathcal{F}^{(i)}$.

In case (iii), $v_1$ is either on or above $a_{a\beta}$ in $b^{(i-1)}$, since $a$ is legitimate, so $v_1$ remains as a lower corner in $b^{(i)}$, on $a_{a-1\beta}$. But $a_{a-1\beta}$ is contained in $a^{(i)}$ (since $(\alpha - 1, \beta) > (\alpha, \beta)$) so $v_1 \notin \mathcal{F}^{(i)}$.

Thus, for all choices of $v_1$, we have $z_j \in 1^{(i)}$ if and only if $v_1 \in \mathcal{F}^{(i)}$. Similarly (by considering the above arguments reflected about the line $y = x$) we can show that $v_2 \in \mathcal{F}^{(i-1)}$ if and only if $v_2 \notin \mathcal{F}^{(i)}$.

Therefore, as noted above, we have $|\mathcal{F}^{(i)}| = |\mathcal{F}^{(i-1)}|$ and $w(\mathcal{F}^{(i)}) = w(\mathcal{F}^{(i-1)})$, and furthermore, Lemma 2.2 tells us that our construction of $b^{(i)}$ from $b^{(i-1)}$ is bijective. This gives $P_k(t, l, a^{(i)}) = P_k(t, l, a^{(i-1)})$ and the result follows by applying this result for $i = 1, \ldots, (r + 1)(s + 1)$, since $a^{(0)} = a$ and $a^{((r+1)(s+1)+1)} = c$. \hfill \Box

The above result allows us to consider only upper corners, and no lower corners, as well as horizontal crossings of arbitrary $x$-coordinates $t$ and vertical crossings of arbitrary $y$-coordinates $l$. The next result allows us to consider only horizontal crossings of $x$-coordinates in $m - r = \{m - r, \ldots, m - 1\}$ and vertical crossings of $y$-coordinates in $s = \{1, \ldots, s\}$. For compactness, we let $\Delta = (r^2 + 1) - (r^2 + 1) + mr$.

**Theorem 3.3.** For all $k \geq 0$,

$$P_k(t, l, c) = q^{\sum_{i=1}^{k} t_i + \sum_{i=1}^{k} h_i - \Delta} P_k(m - r, s, c).$$

**Proof.** First we prove that

$$P_k(t, l, c) = q^{\sum_{i=1}^{k} t_i - (r^2 + 1) + mr} P_k(m - r, l, c).$$

If $t = m - r$ this is obviously true. Otherwise, let $h$ be the largest value of $i$ such that $t_i < m - r - 1 + i$, so $t_{h+1} > t_h + 1$. Now take an arbitrary $b \in \mathcal{P}(m, n)$ and define $\xi(b) = \xi_{t_l}(b)$ as follows.

Let $\gamma$ be the maximal segment of $b$ with $x$-coordinates $t_h$ and $t_h + 1$, and $b = \gamma_1\gamma_2$, so $\gamma_1$ has its final vertex (and no others) with $x$-coordinate $t_h$, and $\gamma_2$ has its initial vertex (and no others) with $x$-coordinate $t_h + 1$. Moreover $\gamma = (t_h, y_1) \cdots (t_h, y_2)(t_{h+1}, y_2) \cdots (t_{h+1}, y_3)$, where $y_1 \leq y_2 \leq y_3$. We define $\xi(b)$ separately in three cases, depending on the values of $y_1$, $y_2$, $y_3$ and their interaction with $l$. Thus we have either

(i) $y_1 = y_2 = y_3$, or
(ii) $y_1 < y_2 = y_3$, or $y_1 < y_2 < y_3$ with $y_2 \in l$, or
(iii) $y_1 = y_2 < y_3$, or $y_1 < y_2 < y_3$ with $y_2 \notin l$.

In case (i), set $\xi(b) = b$.

In case (ii), let $\{y_1 + 1, \ldots, y_2 - 1\} \cap l = \{l_1, \ldots, l_{i_1}\}$, where $l_1 < \cdots < l_{i_1}$, and set $\xi(b) = \gamma_1(t_h, y_1) \cdots (t_h, l_{i_1})(t_{h+1}, l_{i_1}) \cdots (t_{h+1}, y_3)\gamma_2$. (If $\{y_1 + 1, \ldots, y_2 - 1\} \cap l = \emptyset$ then let $l_{i_1} = y_1$.)
In case (iii), let \( \{ y_2 + 1, \ldots, y_3 - 1 \} \cap \{ l_1, \ldots, l_j \} = \emptyset \). Let \( e \) be the maximum value of \( i \) such that \( l_i = y_2 + i \), and set \( \gamma (b) = y_1 (t_h, y_1) \cdots (t_h, y_2 + e + 1) (t_h + 1, y_2 + e + 1) \cdots (t_h + 1, y_3) y_2 \). (If there is no such \( i \), then let \( e = 0 \).)

Now it is routine to check that \( \xi \) is reversible, so \( \xi_{t_1} : \mathcal{P} (m, n) \to \mathcal{P} (m, n) : b \to b' \) is a bijection. Let \( t' = (t_1, \ldots, t_{h-1}, t_h + 1, t_{h+1}, \ldots, t_r) \), \( \mathcal{F}' = \mathcal{F}_{t, k, c} (b') \) and we examine the effect of \( \xi_{t, k, c} \) on \( \mathcal{F} = \mathcal{F}_{t, k, c} (b) \). Let \( \mathcal{F} \) consist of the elements of \( \mathcal{F} \) that are also in \( \mathcal{F}' \) and let \( \mathcal{F}' \) consist of the elements of \( \mathcal{F}' \) that are not also in \( \mathcal{F} \).

In cases (i) and (ii), \( \mathcal{F} = \{(t_h, y_2)\} \) and \( \mathcal{F}' = \{(t_h + 1, y_2)\} \), so \(|\mathcal{F}'| - |\mathcal{F}| = 1 - 1 = 0 \) and \( w(\mathcal{F}') - w(\mathcal{F}) = (t_h + 1 + y_2) - (t_h + y_2) = 1 \).

In case (iii), \( \mathcal{F} = \{(t_h, y_2), (t_h + 1, y_2 + 1), \ldots, (t_h + 1, y_2 + e)\} \) and \( \mathcal{F}' = \{(t_h, y_2 + 1), \ldots, (t_h, y_2 + e + 1)\} \), so \(|\mathcal{F}'| - |\mathcal{F}| = (e + 1) - (e + 1) = 0 \) and \( w(\mathcal{F}') - w(\mathcal{F}) = (t_h + y_2 + 1 + \cdots + t_h + y_2 + e + 1) - (t_h + y_2 + t_h + 1 + y_2 + 1 + \cdots + t_h + 1 + y_2 + e) = 1 \).

Thus in all cases \(|\mathcal{F}'| - |\mathcal{F}| = |\mathcal{F}'| - |\mathcal{F}| = 0 \) and \( w(\mathcal{F}') - w(\mathcal{F}) = w(\mathcal{F}') - w(\mathcal{F}) = 1 \), so for the bijection \( \xi_{t, k, c} : b \to b' \) we have \(|\mathcal{F}_{t, k, c} (b')| = |\mathcal{F}_{t, k, c} (b)| \) and \( w(\mathcal{F}_{t, k, c} (b')) = w(\mathcal{F}_{t, k, c} (b)) + 1 \). This immediately gives \( P_k (t', l, c) = P_k (t, l, c) \) and applying this \( m r - (t_2 + 1) - \sum_{i=1}^{t_2} t_i \) times yields

\[
P_k (m - r, l, c) = q^{m r - (t_2 + 1) - \sum_{i=1}^{t_2} l_i} P_k (t, l, c).
\]

But we can similarly show that

\[
P_k (t, s, c) = q^{s (t_2 + 1) - \sum_{i=1}^{t_2} l_i} P_k (t, l, c)
\]

(by applying \( \xi_{t, k, c} \) to the reflection of \( b \) about \( y = x \), then reflecting back, and applying this \( \sum_{i=1}^{t_2} l_i \) times). The result follows by combining these two results.

By considering the first three results of this section, we obtain a theorem expressing the generating function for the shuffle product of an arbitrary pair of permutations in terms of the generating function for the shuffle product of the canonical pair \( \mu_r = r + 1 \cdots m r \cdots 1 \) and \( \nu_s = m + s + 1 \cdots m + 1 m + s + 2 \cdots m + n \).

**Theorem 3.4.** If \( \sigma \) and \( \omega \) are \((m, n)\)-compatible, with \( d (\sigma) = r \), \( d (\omega) = s \), then for all \( k \geq 0 \),

\[S_k (\sigma, \omega) = q^{l (\sigma) + l (\omega) - \Delta} S_k (\mu_r, \nu_s).
\]

**Proof.** Applying Theorems 3.2 and 3.3 to Lemma 3.1, we obtain

\[S_k (\sigma, \omega) = q^{l (\sigma) + l (\omega) - \Delta} P_k (m - r, s, c).
\]

But Lemma 3.1 also yields \( S_k (\mu_r, \nu_s) = P_k (m - r, s, c) \), since \( \mathcal{D} (\mu_r) = m - r \), \( \mathcal{D} (\nu_s) = s \), and all elements of \( \nu_s \) are larger than all elements of \( \mu_r \). The result follows immediately.

We now give a direct evaluation of the canonical generating function \( S_k (\mu_r, \nu_s) \).

**Theorem 3.5.** For all \( k \geq 0 \),

\[S_k (\mu_r, \nu_s) = q^{\Delta + (k - r)(k - s)} \left[ \frac{m - r + s}{k - r} \right] \left[ \frac{n - s + r}{k - s} \right].
\]
A BIJECTIVE PROOF OF STANLEY'S SHUFFLING THEOREM

Proof. In any \( \rho \in \mathcal{S}(\mu, \nu) \), each of \( m + s + 1, \ldots, m + 2 \) must be larger than the objects that immediately follow them, and each of \( r, \ldots, 1 \) must be smaller than the objects that immediately precede them. Thus \( S_k(\mu, \nu) = 0 \) unless \( k \geq s \) and \( k \geq r \), so the result holds when \( k < \max\{r, s\} \).

Now assume \( k \geq \max\{r, s\} \), and consider arbitrary \( \alpha = \{a_1, \ldots, a_{k-r}\} \subseteq \mathcal{N}_{m-r+s} \) and \( \beta = \{b_1, \ldots, b_{k-s}\} \subseteq \mathcal{N}_{n-s+r} \). We construct \( \rho \in \mathcal{S}(\mu, \nu) \) from \( \alpha, \beta \) as follows, considering two cases.

Case 1 (\( k > r + s \)). Place the first \( s + 1 \) elements of \( \nu \) in positions \( \alpha_1, \ldots, \alpha_{s+1} \) of \( \rho \), and put the first \( \alpha_{s+1} - s - 1 \) elements of \( \mu \) in the remaining positions from 1 to \( \alpha_{s+1} \). We have now filled the first \( \alpha_{s+1} \) positions of \( \rho \). We follow with the next \( \beta_{s+1} - 1 \) elements of \( \nu \), and then, for \( i = 2, \ldots, k - s - r \), we alternate blocks of the next \( \alpha_{s+i} - \alpha_{s+i-1} \) elements of \( \mu \), and the next \( \beta_{s+i} - \beta_{s+i-1} \) elements of \( \nu \). Then we place the next \( m - r + s + 1 - \alpha_{k-s} \) elements of \( \mu \), so that the first \( m - r + s + \beta_{k-s} \) positions of \( \rho \) are filled. Next we place the remaining \( r \) elements of \( \mu \), in positions \( m - r + s + \beta_{k-s-r+1}, \ldots, m - r + s + \beta_{k-s} \), and fill the remaining \( n - s - \beta_{k-s} \) positions of \( \rho \) with the final \( n - s - \beta_{k-s} \) elements of \( \nu \). For example if \( m = 5, n = 4, r = 2, s = 1, k = 5, \alpha = \{1, 2, 4\} \subseteq 4 \) and \( \beta = \{1, 2, 3, 5\} \subseteq 5 \), then \( \rho = 763485291 \in \mathcal{S}(34521, 7689) \).

Now the descents of \( \rho \) are the positions occupied by \( m + s + 1, \ldots, m + 2 \), the positions preceding those occupied by \( r, \ldots, 1 \), and the positions occupied by an element of \( \nu \), which is immediately followed by an element of \( \mu \) (these are not mutually exclusive). Thus for \( \rho \) constructed above, we have

\[
D(\rho) = \{\alpha_1, \ldots, \alpha_s, \alpha_{s+1} + \beta_1 - 1, \ldots, \alpha_{k-r} + \beta_{k-s} - 1, \\
m - r + s - 1 + \beta_{k-s-r+1}, \ldots, m - r + s - 1 + \beta_{k-s} \},
\]

so \( d(\rho) = k \) and \( I(\rho) = s - k + r(m - r + s) + \sum_{i=1}^{k-r} \alpha_i + \sum_{i=1}^{k-r} \beta_i \).

Case 2 (\( k \leq r + s \)). Let \( \{y_1, \ldots, y_{m-k+s}\} = \mathcal{N}_{m-r+s} - \alpha \) and \( \{\delta_1, \ldots, \delta_{n-k+r}\} = \mathcal{N}_{n-s+r} - \beta \), where \( y_1 < \cdots < y_{m-k+s} \) and \( \delta_1 < \cdots < \delta_{n-k+r} \). Place the first \( m - r \) elements of \( \mu \), in positions \( y_1, \ldots, y_{m-r} \), and put the first \( y_{m-r} \), \( m + r \) elements of \( \nu \) in the remaining positions from 1 to \( y_{m-r} \), so that the first \( y_{m-r} \) positions of \( \rho \) are filled. We follow with the next \( \delta_1 - 1 \) elements of \( \mu \), and then, for \( i = 1, s + r - k \), we alternate blocks of \( \gamma_{m-r+i} - \gamma_{m-r+i-1} \) elements of \( \nu \), and blocks of \( \delta_{s+r-k+i} - \delta_{s+r-k+i-1} \) elements of \( \mu \). Then we place the next \( m - r + s - \gamma_{m-k+s} \) elements of \( \nu \), so that the first \( m - r + s + \delta_{s+r-k+1} \) positions of \( \rho \) are filled. Then we place the remaining \( n - s \) elements of \( \nu \), in positions \( m - r + s + \delta_{s+r-k+1}, \ldots, m - r + s + \delta_{n-k+r} \), and fill the remaining \( r + 1 - \delta_{s+r-k+1} \) positions with the final \( r + 1 - \delta_{s+r-k+1} \) elements of \( \mu \). Thus we can identify positions that are not descents, and have

\[
D(\rho) = \mathcal{N}_{m+n-1} - \{y_1, \ldots, y_{m-r}, y_{m-r} + \delta_1 - 1, \ldots, \gamma_{m-k+s} + \delta_{s+r-k+1} - 1, \\
m - r + s - 1 + \delta_{s+r-k+2}, \ldots, m - r + s - 1 + \delta_{n-k+r} \},
\]
so \( d(\rho) = (m + n - 1) - (m + n - k - 1) = k \) and

\[
I(\rho) = \binom{m + n}{2} - \left( \sum_{i=1}^{m-k+s} \gamma_i + \sum_{i=1}^{n-k+r} \delta_i + (n - s - 1)(m - r + s) - (n - k + r) \right).
\]

But

\[
\sum_{i=1}^{m-k+s} \gamma_i = \binom{m - r + s + 1}{2} - \sum_{i=1}^{k-r} \alpha_i
\]

and

\[
\sum_{i=1}^{n-k+r} \delta_i = \binom{n - s + r + 1}{2} - \sum_{i=1}^{k-s} \beta_i,
\]

so simplifying gives

\[
I(\rho) = s - k + r(m - r + s) + \sum_{i=1}^{k-r} \alpha_i + \sum_{i=1}^{k-s} \beta_i.
\]

It is easy to check that there is a unique such pair of subsets \( \alpha \) and \( \beta \) associated with each \( \rho \in \mathcal{S}(\mu_r, \nu_s) \) with \( d(\rho) = k \), so the construction is bijective. Thus

\[
S_k(\mu_r, \nu_s) = \sum_{\rho \in \mathcal{S}(\mu_r, \nu_s), \ d(\rho) = k} q^{I(\rho)} = q^s - k + r(m - r + s) \sum_{1 \leq \alpha_1 < \cdots < \alpha_{k-r} \leq m-r+s} q^{\alpha_1 + \cdots + \alpha_{k-r}} \\
\times \sum_{1 \leq \beta_1 < \cdots < \beta_{k-s} \leq n-s+r} q^{\beta_1 + \cdots + \beta_{k-s}} = q^{s - k + r(m - r + s)} q^{\binom{m - r + s}{k-r}} q^{\binom{n - s + r}{k-s}}
\]

from Lemma 1.4, and the result follows since

\[
\binom{k-r+1}{2} + \binom{k-s+1}{2} + s - k + r(m - r + s) = \Delta + (k - s)(k - r). \quad \square
\]

MacMahon [8, Vol. I, p. 169] has given a direct evaluation of \( S_k(\mu_0, \nu_0) \) at \( q = 1 \); one of his proofs involved the lattice path representation given in Proposition 2.3. The special case \( s = 0 \) of Theorem 3.5 has been used in Goulden [5] as one of three ingredients in a combinatorial proof of an identity equivalent to Theorem 1.3.

We now have completed all ingredients for a proof of the Shuffling Theorem.

**Proof of Theorem 1.2.** The result follows immediately from Theorems 3.4 and 3.5. \( \square \)

We conclude with an example that illustrates all of the results of this section.

**Example 3.6.** Let

\[
\rho = 5 \quad 10 \quad 8 \quad 4 \quad 12 \quad 2 \quad 7 \quad 6 \quad 13 \quad 11 \quad 3 \quad 1 \quad 9 \quad \in \mathcal{S}(5, 10, 12, 2, 7, 6, 13, 11, 9, 8, 4, 6, 3, 1),
\]

so \( \sigma = 5 \quad 10 \quad 12 \quad 2 \quad 7 \quad 13 \quad 11 \quad 9 \), \( \omega = 84631 \), \( m = 8, \ n = 5, \ r = s = 3, \ I(\sigma) = 16, \ I(\omega) = 8, \ d(\rho) = k = 7 \) and \( I(\rho) = 47 \).
Then, in the notation of Lemma 3.1, we represent \( p \) by \( b = A^2U^2A^3U^2A^2U^2A. \) Moreover \( t = \{3, 6, 7\}, \ell = \{1, 3, 4\}, a_{00} = AUA^2, a_{10} = A^2UA(3, 0), a_{20} = UA(6, 0), a_{30} = UA(7, 0), a_{01} = UAUA(0, 1), a_{11} = AU^2A^2(3, 1), a_{21} = U^2A(6, 1), a_{31} = U^2A(7, 1), a_{02} = UA^3(0, 3), a_{12} = AU^2A^2(3, 3), a_{22} = UA(6, 3), a_{32} = UA(7, 3), a_{03} = UA^3(0, 4), a_{13} = UA^3(3, 4), a_{23} = UA(6, 4) \) and \( a_{33} = UA(7, 4) \). Also \( \mathcal{F}_{1,\ell}(b) = \{(2, 0), (2, 1), (3, 2), (5, 2), (6, 3), (7, 3), (7, 4)\} \), so indeed \( |\mathcal{F}_{1,\ell}(b)| = 7 = d(p) \) and \( w(\mathcal{F}_{1,\ell}(b)) = 47 = I(p) \).

In the notation of Theorem 3.2, we have \( b^{(6)} = \cdots = b^{(0)} = b, b^{(7)} = AUAUA^3UA^2U^2A, b^{(8)} = AUAUA^2UA^2U^2A, b^{(9)} = AUAUA^2UA^2UAUA, b^{(10)} = AUAUA^2UA^2UA^2U, b^{(11)} = AU^2A^3UA^2UA^2U \) and \( b^{(16)} = \cdots = b^{(12)} = AU^2A^3UAUA^3U \). Now \( \mathcal{F}_{1,\ell}(b^{(16)}) = \{(1, 1), (1, 2), (3, 2), (4, 3), (5, 4), (6, 4), (7, 4)\} \), so \( |\mathcal{F}_{1,\ell}(b^{(16)})| = 7 = d(p) \) and \( w(\mathcal{F}_{1,\ell}(b^{(16)})) = 47 = I(p) \), as required.

In Theorem 3.3, by applying \( \xi \) twice \( (mr - (r^2 + 1) - \Sigma_{i=1}^r t_i = 2) \), we get the path \( b'' = AU^2A^2UA^2U^2U \) and obtain \( \mathcal{F}_{m-r,\ell}(b'') = \{(1, 1), (1, 2), (3, 3), (4, 4), (5, 4), (6, 4), (7, 4)\} \), so \( |\mathcal{F}_{m-r,\ell}(b'')| = 7 = d(p) \) and \( w(\mathcal{F}_{m-r,\ell}(b'')) = 49 = I(p) + 2 \), as required. Finally, by similarly reducing \( \ell \) to \( s \), we obtain the path \( b_0 = AU^2AUAUA^3U \) with \( \mathcal{F}_{m-r,s}(b_0) = \{(1, 1), (1, 2), (2, 3), (3, 4), (5, 4), (6, 4), (7, 4)\} \), so \( |\mathcal{F}_{m-r,s}(b_0)| = 7 = d(p) \) and \( w(\mathcal{F}_{m-r,s}(b_0)) = 47 = I(p) + \Delta - \Sigma_{i=1}^r t_i - \Sigma_{i=r}^s \), as required.

In Theorem 3.4, we finally obtain that \( p \in \mathcal{P}_0(\sigma, \omega) \) corresponds to

\[
\rho' = 4 \quad 12 \quad 11 \quad 10 \quad 9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 13 \in \mathcal{P}(\mu, v),
\]

where \( \mu = 45678321 \) and \( v = 12 \quad 11 \quad 10 \quad 9 \quad 13 \), \( r(p') = 7 = d(p), I(p') = 47 = I(p) + \Delta - I(\sigma) - I(\omega). \)

Finally in Theorem 3.5 we have \( k = 7 > 3 + 3 = r + s \), so we have Case 1, with \( \alpha = \{2, 3, 5, 7\}, \beta = \{1, 2, 3, 4\} \) corresponding to \( \rho' \). Of course \( I(p') = 47 = s - k + r(m - r + s) + \Sigma_{i=1}^r a_i + \Sigma_{i=r}^s b_i. \)

We say that this proof of the Shuffling Theorem is bijective because we are able to explicitly give a bijection between elements of \( \mathcal{P}_0(\sigma, \omega) \) and pairs of subsets of \( \mathcal{N}_{m-s+r} \) and \( \mathcal{N}_{m-r+s} \), the existence of which is implicit in its statement. Thus, in Example 3.6 we have demonstrated that

\[
p = 5 \quad 10 \quad 8 \quad 4 \quad 12 \quad 2 \quad 7 \quad 6 \quad 13 \quad 11 \quad 3 \quad 9 \quad 14 \in \mathcal{P}(5 \quad 10 \quad 12 \quad 2 \quad 7 \quad 13 \quad 11 \quad 9, 8 \quad 4 \quad 6 \quad 3 \quad 1)
\]

corresponds to \( \alpha = \{2, 3, 5, 7\} \subseteq \mathcal{N}_{m-r+s} \) and \( \beta = \{1, 2, 3, 4\} \subseteq \mathcal{N}_{m-s+r}. \)

Acknowledgements. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada, and was carried out while the author was visiting the Department of Mathematics, M.I.T. The author would like to thank Ira Gessel for suggesting this problem.

References


DEPARTMENT OF COMBINATORICS AND OPTIMIZATION, UNIVERSITY OF WATERLOO, WATERLOO, ONTARIO, CANADA N2L 3G1