TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 288, Number 1, March 1985

A BIJECTIVE PROOF OF STANLEY'S SHUFFLING THEOREM

BY

I. P. GOULDEN

ABSTRACT. For two permutations σ and ω on disjoint sets of integers, consider forming a permutation on the combined sets by "shuffling" σ and ω (i.e., σ and ω appear as subsequences). Stanley [10], by considering P-partitions and a q-analogue of Saalschutz's ${}_3F_2$ summation, obtained the generating function for shuffles of σ and ω 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 σ and ω. A combinatorial proof of this result is obtained by finding bijections for lattice path representations of shuffles which reduce σ and ω 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 $\mathscr{D}(a)$, by $\mathscr{D}(a) = \{i | a_i > a_{i+1}\}$, the number of *descents* in a by $d(a) = |\mathscr{D}(a)|$, and the *greater index* of a by $I(a) = \sum_{i \in \mathscr{D}(a)} i$. 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 $\mathscr{N}_{m+n} = \{1, \ldots, m+n\}$, where $\alpha_1 < \cdots < \alpha_m$ and $\beta_1 < \cdots < \beta_n$. For any permutation σ of the elements of α, and any permutation ω of the elements of β, we say that σ and ω are (m, n)*-compatible*.

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

PROPOSITION 1.1. *If σ and ω are (m, n)-compatible, then*

$$|\mathscr{S}(\sigma, \omega)| = \binom{m+n}{m}.$$

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

The elements of ω appear, in order, in the set of positions of ρ complementary to α. $\square$

Received by the editors January 31, 1984.

1980 *Mathematics Subject Classification*. Primary 05A10, 05A15; Secondary 06A99, 68E05.

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

$$S_k(\sigma, \omega) = \sum_{\substack{\rho \in \mathscr{S}(\sigma, \omega) \\ d(\rho) = k}} q^{I(\rho)}.$$

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

$$\begin{bmatrix} i \\ j \end{bmatrix} = \frac{(1 - q^i) \cdots (1 - q^{i-j+1})}{(1 - q^j) \cdots (1 - q)},$$

and $\left[{i \atop j}\right] = 0$ otherwise.

THEOREM 1.2 (SHUFFLING THEOREM). *Let σ and ω be (m, n)-compatible, with $d(\sigma) = r$, $d(\omega) = s$. Then*

$$S_k(\sigma, \omega) = q^{I(\sigma) + I(\omega) + (k-s)(k-r)} \begin{bmatrix} m - r + s \\ k - r \end{bmatrix} \begin{bmatrix} n - s + r \\ k - s \end{bmatrix}. \qquad \square$$

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.

$$\sum_{i \geqslant 0} q^{(r-i)(s-i)} \begin{bmatrix} m + r - s \\ r - i \end{bmatrix} \begin{bmatrix} n + s - r \\ s - i \end{bmatrix} \begin{bmatrix} m + n + i \\ i \end{bmatrix} = \begin{bmatrix} m + r \\ m \end{bmatrix} \begin{bmatrix} n + s \\ n \end{bmatrix}. \qquad \square$$

This identity was proved by Gould [**4**], and is equivalent to Jackson's [**7**] q-analogue of Saalschutz's theorem (see [**9**, p. 243]). Combinatorial proofs of Saalschutz'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 $\left[{i \atop j}\right]$ 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.

$$\sum_{0 \leqslant \alpha_1 \leqslant \cdots \leqslant \alpha_j \leqslant i-j} q^{\alpha_1 + \cdots + \alpha_j} = \begin{bmatrix} i \\ j \end{bmatrix}.$$

2.

$$\sum_{1 \leqslant \beta_1 < \cdots < \beta_j \leqslant i} q^{\beta_1 + \cdots + \beta_j} = q^{\binom{j+1}{2}} \begin{bmatrix} i \\ j \end{bmatrix}.$$

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 $u = (u_1, u_2)$ ad $v = (v_1, v_2)$ 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 ith difference d_i is called the ith *step* and a_i is the ith *vertex* in the path a. We say that d_i follows a_{i-1} and precedes a_i. 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 $u_1 \leqslant x_0 \leqslant x_1 < \cdots < x_k \leqslant v_1$ and $u_2 \leqslant y_0 < y_1 < \cdots < y_{k-1} \leqslant y_k \leqslant v_2$ 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 $u_1 \leqslant x_0 < x_1 < \cdots < x_{k-1} \leqslant x_k \leqslant v_1$ and $u_2 \leqslant y_0 \leqslant y_1 < \cdots < y_k \leqslant v_2$ 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

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

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 $\leqslant$ *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 $a = UAU^2AUA$, then U, A^2U and $(UA^2U)_{(1,2)}$ are $\leqslant a$-paths, while $(U^2A)_{(0,1)}$ and $(UA^2)_{(0,3)}$ are $> a$-paths. The path U^2A is neither a $\leqslant a$-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 $(A^{v_1-u_1}U^{v_2-u_2})_u$ is called the *canonical cover* of $B(u, v)$.

If a covers B and b is a path on B, then we define $\mathscr{C}_{B,a}(b)$ 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 $\mathscr{C}_{B,c}(b)$ 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 $e_1 = (e_{11}, e_{12})$ be $w(e_1) = e_{11} + e_{12}$, 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 ψ for paths is very important to our proof of Theorem 1.2. Suppose that a covers B and g is a $\leqslant a$-path on B from z_1 to z_2 with lower corners below a given by $(f_{11}, f_{21}),\ldots,(f_{1k}, f_{2k})$. Then we define $\psi_{B,a}(g)$ to be the path from z_1 to z_2 whose upper corners are $(f_{11} - 1, f_{21} + 1),\ldots,(f_{1k} - 1, f_{2k} + 1)$. (This path is unique, by Proposition 2.1.) If b is a path on B, then we can write b uniquely as $b = h_1g_1h_2g_2 \cdots h_lg_l$, for some $l \geqslant 1$, where $g_1,\ldots,g_{l-1}$ are $\leqslant a$-paths, g_l is either a $\leqslant a$-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_l is empty, then $\psi_{B,a}(g_l) = g_l$.)

For example, if $B = B((0,0),\ (7,6))$, $a = AU^2AUA^3UAU^2A$ and $b = (A^3U^3A^3UAU)_{(0,1)}$, 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 $\leqslant a$-paths. Now $\psi_{B,a}(g_1) = (AUAU)_{(1,1)}$ and $\psi_{B,a}(g_2) = (AU^2A)_{(5,4)}$, so

$$\psi_{B,a}(b) = (A^2UAU^2A^3U^2A)_{(0,1)}.$$

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

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

1. *$\psi_{B,a}\colon \mathscr{P}_B(x, y) \to \mathscr{P}_B(x, y) : b \mapsto \hat{b}$ is a bijection.*

Moreover, if c is the canonical cover of B, then
2. $|\mathscr{C}_{B,a}(b)| = |\mathscr{C}_{B,c}(\hat{b})|$,
3. $w(\mathscr{C}_{B,a}(b)) = w(\mathscr{C}_{B,c}(\hat{b}))$.

PROOF. 1. First note that if g is a $\leqslant a$-path from z_1 to z_2, then $\hat{g} = \psi_{B,a}(g)$ is also a $\leqslant a$-path from z_1 to z_2. This is because if the lower corners of g below a occur at $(f_{11}, f_{21}),\dots,(f_{1k}, f_{2k})$, then the upper corners of $\hat{g}$ occur at $(f_{11} - 1, f_{21} + 1),\dots,(f_{1k} - 1, f_{2k} + 1)$, 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 $\leqslant a$-path. Moreover g is recoverable from $\hat{g}$ as follows. Suppose that $\hat{g}$ is a $\leqslant a$-path from z_1 to z_2, whose upper corners are $(c_{11}, c_{21}),\dots,(c_{1k}, c_{2k})$. Let $\tilde{g}$ be the unique path from z_1 to z_2 whose lower corners are $(c_{11} + 1, c_{21} - 1),\dots,(c_{1k} + 1, c_{2k} - 1)$, given by Proposition 2.1. Now $\tilde{g}$ is not necessarily a $\leqslant a$-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 $m \geqslant 1$, where $d_1,\dots,d_m$ are $\leqslant a$-paths (d_1 and d_m can also be empty) and $e_1,\dots,e_{m-1}$ are $> a$-paths. Also, all lower corners of $\tilde{g}$ 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 $i = 1,\dots,k$. Thus the lower corners of $\tilde{g}$ must all be internal vertices in one of the paths $d_1,\dots,d_m$. The path e_i for $i = 1,\dots, m - 1$ is a path from a vertex on a, say a_{f_i}, to a vertex on a, say a_{l_i}, with a single corner (upper). For $i = 1,\dots,m - 1$, let r_i be the segment of a from a_{f_i} to a_{l_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 $\leqslant a$-paths.

Now, if $b = h_1 g_1 \cdots h_l g_l \in \mathscr{P}_B(x, y)$ in the notation of the definition of $\psi_{B,a}$ above, then $\hat{b} = h_1 \hat{g}_1 \cdots h_l \hat{g}_l \in \mathscr{P}_B(x, y)$, where $\hat{g}_i = \psi_{B,a}(g_i)$ for $i = 1,\dots,l$, the h_i's are $> a$-paths (by definition), and the $\hat{g}_i$'s are $\leqslant a$-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 $\leqslant a$-path, then $|\mathscr{C}_{B,a}(g)| = |\mathscr{C}_{B,c}(\hat{g})|$ ($= k$ above) and $w(\mathscr{C}_{B,a}(g)) = w(\mathscr{C}_{B,c}(\hat{g}))$ ($= f_{11} + f_{21} + \cdots + f_{1k} + f_{2k}$ 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

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

and

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

since the intersecting vertex of a $\leqslant a$-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 $\mathscr{C}_{B,a}(h_i) = \mathscr{C}_{B,c}(h_i)$ for $i = 1,\dots,l$ since h_i is a $> a$-path, and $|\mathscr{C}_{B,a}(g_i)| = |\mathscr{C}_{B,c}(\hat{g}_i)|$ from above, so $|\mathscr{C}_{B,a}(g)| = |\mathscr{C}_{B,c}(\hat{g})|$. The proof that $w(\mathscr{C}_{B,a}(g)) = w(\mathscr{C}_{B,c}(\hat{g}))$ is similar. □

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 $\mathscr{P}(m, n)$. (Note that $\mathscr{P}(m, n) = \mathscr{P}_G((0,0), (m, n))$ in the notation of Lemma 2.2.)

Now we relate lattice paths to shuffle products of permutations. For (m, n)-compatible permutations σ and ω, we represent the permutation $\rho \in \mathscr{S}(\sigma, \omega)$ by the path $\phi_{\sigma,\omega}(\rho) \in \mathscr{P}(m, n)$ as follows. If the ith element of ρ is in σ, then the ith step in $\phi_{\sigma,\omega}(\rho)$ is across, and if the ith element of ρ is in ω, then the ith step in $\phi_{\sigma,\omega}(\rho)$ is up for $i = 1,\dots,m+n$. We say that the ith step of $\phi_{\sigma,\omega}(\rho)$ *represents* the ith element of ρ. For example $\rho_0 = 647325819 \in \mathscr{S}(6358, 47219)$ is represented by $\phi(\rho_0) = AU^2AUA^2U^2 \in \mathscr{P}(4,5)$.

PROPOSITION 2.3. *If σ and ω are (m, n)-compatible permutations, then $\phi_{\sigma,\omega}$: $\mathscr{S}(\sigma, \omega) \to \mathscr{P}(m, n)$ is a bijection.*

PROOF. Immediate from Proposition 2.1, since a subset of $\mathscr{N}_{m+n}$ of cardinality m uniquely specifies the path in $\mathscr{P}(m, n)$ 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 $0 = t_0 < \cdots < t_{r+1} = m$, $0 = l_0 < \cdots < l_{s+1} = n$, $\mathbf{t} = \{t_1,\dots,t_r\}$ and $\mathbf{l} = \{l_1,\dots,l_s\}$. Let B_{ij} be the grid $B((t_i, l_j), (t_{i+1}, l_{j+1}))$ for $i = 0,\dots,r$, $j = 0,\dots,s$, so $\bigcup_{i=0}^{r}\bigcup_{j=0}^{s} B_{ij} = G$. The grids B_{ij} and B_{ij+1} 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+1j} 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 $b \in \mathscr{P}(m, n)$, define $\varepsilon_{ij}(b)$ to be the maximal subpath of b on B_{ij}. Let $\mathscr{V}_{\mathbf{l}}(b)$ be the set of vertical crossings of y-coordinates in $\mathbf{l}$, and $\mathscr{H}_{\mathbf{t}}(b)$ be the set of horizontal crossings of x-coordinates in $\mathbf{t}$. Suppose that $\mathbf{a}$ is an array with (i, j)-entry a_{ij} for $i = 0,\dots,r$, $j = 0,\dots,s$, where a_{ij} covers B_{ij}. Then define

$$\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b) = \mathscr{H}_{\mathbf{t}}(b) \,\dot{\cup}\, \mathbf{V}_{\mathbf{l}}(b) \,\dot{\cup}\, \dot{\bigcup_{i=0}^{r}} \dot{\bigcup_{j=0}^{s}} \mathscr{C}_{B_{ij},a_{ij}}(\varepsilon_{ij}(b)),$$

and

$$P_k(\mathbf{t}, \mathbf{l}, \mathbf{a}) = \sum_{\substack{b\in\mathscr{P}(m,n)\\ |\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b)|=k}} q^{w(\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b))}.$$

Finally we say that $\mathbf{a}$ 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 $m = 9$, $n = 6$, $r = 2$, $s = 1$, $\mathbf{t} = \{2, 6\}$, $\mathbf{l} = \{3\}$, $a_{00} = UA^2U^2$, $a_{01} = (U^3A^2)_{(0,3)}$, $a_{10} = (U^2A^4U)_{(2,0)}$, $a_{11} = (A^4U^3)_{(2,3)}$, $a_{20} = (UAUA^2U)_{(6,0)}$, $a_{21} = (A^3U^3)_{(6,3)}$ and $b = AU^2A^3U^2A^2U^2A^3$. Then $\mathscr{V}_{\mathbf{l}}(b) = \{(4,3)\}$, $\mathscr{H}_{\mathbf{t}}(b) = \{(2,2)\}$, $\varepsilon_{00}(b) = AU^2A$, $\varepsilon_{10}(b) = (A^2U)_{(2,2)}$, $\varepsilon_{11}(b) = (UA^2U^2)_{(4,3)}$, $\varepsilon_{21}(b) = (U^2A^3)_{(6,4)}$ and $\varepsilon_{01}(b) = \varepsilon_{20}(b) = \varnothing$. Thus $\mathscr{C}_{B_{00},a_{00}}(\varepsilon_{00}(b)) = \{(1,0),\ (1,2)\}$, $\mathscr{C}_{B_{11},a_{11}}(\varepsilon_{11}(b)) = \{(4,4)\}$, $\mathscr{C}_{B_{21},a_{21}}(\varepsilon_{21}(b)) = \{(6,6)\}$ and $\mathscr{C}_{B_{ij},a_{ij}}(\varepsilon_{ij}(b)) = \varnothing$ for other i, j. Note that $\mathbf{a}$ 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 $\varepsilon_{ij}(b)$, though distinct ε_{ij} can have a nonempty path (again a portion of mutual border) as their intersection. For example $\varepsilon_{11}(b)$ and $\varepsilon_{21}(b)$ 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* $\mathscr{D}(\sigma) = \mathbf{t}$ *and* $\mathscr{D}(\omega) = \mathbf{l}$. *If* a_{ij} *is the cover of* B_{ij} *representing the shuffle of* $\sigma_{t_i+1} \cdots \sigma_{t_{i+1}}$ *and* $\omega_{l_j+1} \cdots \omega_{l_{j+1}}$ *into increasing order, then*

1. $S_k(\sigma, \omega) = P_k(\mathbf{t}, \mathbf{l}, \mathbf{a})$.
2. $\mathbf{a}$ *is legitimate.*

PROOF. 1. Let $\rho \in \mathscr{S}(\sigma, \omega)$ and let $b = \phi_{\sigma,\omega}(\rho) \in \mathscr{P}(m, n)$. Suppose that ρ_i is the ith element of ρ, and $b_i = (b_{1i}, b_{2i})$ is the vertex of b that follows that ith 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 $b_{1i} \in \mathscr{D}(\sigma) = \mathbf{t}$, or equivalently, $b_i \in \mathscr{H}_{\mathbf{t}}(b)$. If b_i is a vertical crossing in b, then $\rho_i = \omega_{b_{2i}}$ and $\rho_{i+1} = \omega_{b_{2i}+1}$, so $\rho_i > \rho_{i+1}$ if and only if $b_{2i} \in \mathscr{D}(\omega) = \mathbf{l}$, or, equivalently, $b_i \in \mathscr{V}_{\mathbf{l}}(b)$.

If b_i is a corner in b, then b_i appears as a corner in $\varepsilon_{dl}(b)$, say, and in no other $\varepsilon_{ij}(b)$. If b_i is an upper corner, then $\rho_i = \omega_{b_{2i}}$ and $\rho_{i+1} = \sigma_{b_{1i}+1}$. Moreover, if b_i is above a_{dl}, then ρ_{i+1} occurs before ρ_i in a_{dl}, so $\rho_i > \rho_{i+1}$ by the construction of a_{dl}. However if b_i is on or below a_{dl}, then ρ_i occurs before ρ_{i+1} in a_{dl}, so $\rho_i < \rho_{i+1}$. Similarly, if b_i is a lower corner, then b_i is below a_{dl} if and only if $\rho_i > \rho_{i+1}$.

Thus we have a bijection between descents $i \in \mathscr{D}(\rho)$ and vertices $b_i \in \mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b)$, where $w(b_i) = i$. This immediately gives $d(\rho) = |\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b)|$ and $I(\rho) = w(\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b))$ so, from Proposition 2.3, we have

$$S_k(\sigma, \omega) = \sum_{\substack{\rho \in \mathscr{S}(\sigma, \omega) \\ d(\rho) = k}} q^{I(\rho)} = \sum_{\substack{b \in \mathscr{P}(m, n) \\ |\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b)| = k}} q^{w(\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{a}}(b))} = P_k(\mathbf{t}, \mathbf{l}, \mathbf{a}),$$

as required.

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

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

For $b \in \mathscr{P}(m,n)$, we define $b^{(i)} \in \mathscr{P}(m,n)$ 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 (α,β) be the ith element of $\mathscr{Q}$.

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

(iii) If $\varepsilon_{\alpha\beta}(b^{(i-1)}) = \delta$ is a path in $B_{\alpha\beta}$, then $b^{(i-1)} = \delta_1\delta\delta_2$ for unique paths δ_1, δ_2, each with a single vertex in $B_{\alpha\beta}$. Let $b^{(i)} = \delta_1\psi_{B_{\alpha\beta}a_{\alpha\beta}}(\delta)\delta_2$.

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

THEOREM 3.2. *If* $\mathbf{a}$ *is legitimate, then, for all* $k \geqslant 0$, $P_k(\mathbf{t},\mathbf{l},\mathbf{a}) = P_k(\mathbf{t},\mathbf{l},\mathbf{c})$.

PROOF. If, in the construction of $b^{(i)}$ from $b^{(i-1)}$ above, we have $\varepsilon_{\alpha\beta}(b^{(i-1)}) = \varnothing$, 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 $|\mathscr{F}^{(i)}| = |\mathscr{F}^{(i-1)}|$ and $w(\mathscr{F}^{(i)}) = w(\mathscr{F}^{(i-1)})$.

Otherwise, we have $b^{(i-1)} = \delta_1\delta\delta_2$, where $\delta = \varepsilon_{\alpha\beta}(b^{(i-1)})$, and the final vertex, say v_1, of δ_1 is in $B_{\alpha\beta}$, and the initial vertex, say v_2, of δ_2 is in $B_{\alpha\beta}$. Suppose that v_1 is the jth vertex in $b^{(i-1)}$ (and $b^{(i)}$) and that v_2 is the kth vertex in $b^{(i-1)}$ (and $b^{(i)}$). Then, for $u = 0,\ldots,j-1$ and $k+1,\ldots,m+n$, the uth vertex of $b^{(i-1)}$ is in $\mathscr{F}^{(i-1)}$ if and only if the uth vertex of $b^{(i)}$ is in $\mathscr{F}^{(i)}$, because these are vertices internal to δ_1 and δ_2, unchanged in the construction. Also, for $u = j+1,\ldots,k-1$, the uth vertex of $b^{(i-1)}$ is in $\mathscr{F}^{(i-1)}$ if and only if the uth vertex of $b^{(i)}$ is in $\mathscr{F}^{(i)}$, from Lemma 2.2, with $x = v_1$, $y = v_2$, $B = B_{\alpha\beta}$, $a = a_{\alpha\beta}$. But the uth vertex in any path in $\mathscr{P}(m,n)$ has weight equal to u. Thus we prove $|\mathscr{F}^{(i)}| = |\mathscr{F}^{(i-1)}|$ and $w(\mathscr{F}^{(i)}) = w(\mathscr{F}^{(i-1)})$ by proving that $v_1 \in \mathscr{F}^{(i)}$ if and only if $v_1 \in \mathscr{F}^{(i-1)}$, and $v_2 \in \mathscr{F}^{(i)}$ if and only if $v_2 \in \mathscr{F}^{(i-1)}$.

Consider first v_1. If $\alpha = \beta = 0$, then $v_1 = (0,0)$, so δ_1 is empty, and $v_1 \notin \mathscr{F}^{(i-1)}$, $v_1 \notin \mathscr{F}^{(i)}$. 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 $y = l_\beta$) 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 $\mathbf{a}^{(i-1)}$ and $\mathbf{a}^{(i)}$. Thus, whether v_1 is a vertical crossing or upper corner in $b^{(i-1)}$ and $b^{(i)}$, we have $v_1 \in \mathscr{F}^{(i-1)}$ and $v_1 \in \mathscr{F}^{(i)}$.

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 $v_1 \in \mathscr{F}^{(i-1)}$, or

(ii) a lower corner in $B_{\alpha-1\beta}$, below $a_{\alpha-1\beta}$, so $v_1 \in \mathscr{F}^{(i-1)}$, or

(iii) a lower corner in $B_{\alpha-1\beta}$, on $a_{\alpha-1\beta}$, so $v_1 \notin \mathscr{F}^{(i-1)}$.

In case (i) then either (a) v_1 is a horizontal crossing in $b^{(i)}$, so $v_1 \in \mathscr{F}^{(i)}$, 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 $a_{\alpha\beta}$, which means that v_1 is below $a_{\alpha-1\beta}$ (contained in $\mathbf{a}^{(i)}$ since $(\alpha-1,\beta) > (\alpha,\beta)$) since $\mathbf{a}$ is legitimate, so $v_1 \in \mathscr{F}^{(i)}$.

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

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

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

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

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

THEOREM 3.3. *For all $k \geqslant 0$,*

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

PROOF. First we prove that

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

If $\mathbf{t} = m - \mathbf{r}$ 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 $b \in \mathscr{P}(m,n)$ and define $\xi(b) = \xi_{\mathbf{t},\mathbf{l}}(b)$ as follows.

Let γ be the maximal segment of b with x-coordinates t_h and $t_h + 1$, and $b = \gamma_1\gamma\gamma_2$, so γ_1 has its final vertex (and no others) with x-coordinate t_h, and γ_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 $y_1 \leqslant y_2 \leqslant y_3$. We define $\xi(b)$ separately in three cases, depending on the values of y_1, y_2, y_3 and their interaction with $\mathbf{l}$. 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 $y_2 \in \mathbf{l}$, or

(iii) $y_1 = y_2 < y_3$, or $y_1 < y_2 < y_3$ with $y_2 \notin \mathbf{l}$.

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

In case (ii), let $\{y_1 + 1,\dots,y_2 - 1\} \cap \mathbf{l} = \{l_{i_1},\dots,l_{i_g}\}$, where $l_{i_1} < \cdots < l_{i_g}$, and set $\xi(b) = \gamma_1(t_h, y_1) \cdots (t_h, l_{i_g})(t_h + 1, l_{i_g}) \cdots (t_h + 1, y_3)\gamma_2$. (If $\{y_1 + 1,\dots,y_2 - 1\} \cap \mathbf{l} = \varnothing$ then let $l_{i_g} = y_1$.)

In case (iii), let $\{y_2+1,\ldots,y_3-1\}\cap \mathbf{l}=\{l_{j_1},\ldots,l_{j_f}\}$, where $l_{j_1}<\cdots<l_{j_f}$. Let e be the maximum value of i such that $l_{j_i}=y_2+i$, and set $\xi(b)=\gamma_1(t_h,y_1)\cdots(t_h,y_2+e+1)(t_h+1,y_2+e+1)\cdots(t_h+1,y_3)\gamma_2$. (If there is no such i, then let $e=0$.)

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

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

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

Thus in all cases $|\mathscr{F}'|-|\mathscr{F}|=|\mathscr{G}'|-|\mathscr{G}|=0$ and $w(\mathscr{F}')-w(\mathscr{F})=w(\mathscr{G}')-w(\mathscr{G})=1$, so for the bijection $\xi_{\mathbf{t},\mathbf{l}}\colon b\to b'$ we have $|\mathscr{F}_{\mathbf{t}',\mathbf{l},\mathbf{c}}(b')|=|\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{c}}(b)|$ and $w(\mathscr{F}_{\mathbf{t}',\mathbf{l},\mathbf{c}}(b'))=w(\mathscr{F}_{\mathbf{t},\mathbf{l},\mathbf{c}}(b))+1$. This immediately gives $P_k(\mathbf{t}',\mathbf{l},\mathbf{c})=qP_k(\mathbf{t},\mathbf{l},\mathbf{c})$, and applying this $mr-\binom{r+1}{2}-\sum_{i=1}^r t_i$ times yields

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

But we can similarly show that

$$P_k(\mathbf{t},\mathbf{s},\mathbf{c})=q^{\binom{s+1}{2}-\sum_{i=1}^s l_i}P_k(\mathbf{t},\mathbf{l},\mathbf{c})$$

(by applying $\xi_{\mathbf{l},\mathbf{t}}^{-1}$ to the reflection of b about $y=x$, then reflecting back, and applying this $\sum_{i=1}^s l_i-\binom{s+1}{2}$ 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 $\mu_r=r+1\cdots mr\cdots 1$ and $\nu_s=m+s+1\cdots m+1\ m+s+2\cdots m+n$.

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

$$S_k(\sigma,\omega)=q^{I(\sigma)+I(\omega)-\Delta}S_k(\mu_r,\nu_s).$$

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

$$S_k(\sigma,\omega)=q^{I(\sigma)+I(\omega)-\Delta}P_k(m-\mathbf{r},\mathbf{s},\mathbf{c}).$$

But Lemma 3.1 also yields $S_k(\mu_r,\nu_s)=P_k(m-\mathbf{r},\mathbf{s},\mathbf{c})$, since $\mathscr{D}(\mu_r)=m-\mathbf{r}$, $\mathscr{D}(\nu_s)=\mathbf{s}$, and all elements of ν_s are larger than all elements of μ_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 $k\geq 0$,*

$$S_k(\mu_r,\nu_s)=q^{\Delta+(k-r)(k-s)}\begin{bmatrix} m-r+s\\ k-r\end{bmatrix}\begin{bmatrix} n-s+r\\ k-s\end{bmatrix}.$$

PROOF. In any $\rho \in \mathscr{S}(\mu_r, \nu_s)$, each of $m+s+1,\dots,m+2$ must be larger than the objects that immediately follow them, and each of $r,\dots,1$ must be smaller than the objects that immediately precede them. Thus $S_k(\mu_r, \nu_s) = 0$ unless $k \geqslant s$ and $k \geqslant r$, so the result holds when $k < \max\{r, s\}$.

Now assume $k \geqslant \max\{r, s\}$, and consider arbitrary $\alpha = \{\alpha_1,\dots,\alpha_{k-r}\} \subseteq \mathscr{N}_{m-r+s}$ and $\beta = \{\beta_1,\dots,\beta_{k-s}\} \subseteq \mathscr{N}_{n-s+r}$. We construct $\rho \in \mathscr{S}(\mu_r, \nu_s)$ from α, β as follows, considering two cases.

Case 1 ($k > r+s$). Place the first $s+1$ elements of ν_s in positions $\alpha_1,\dots,\alpha_{s+1}$ of ρ, and put the first $\alpha_{s+1} - s - 1$ elements of μ_r in the remaining positions from 1 to α_{s+1}. We have now filled the first α_{s+1} positions of ρ. We follow with the next $\beta_1 - 1$ elements of ν_s and then, for $i = 2,\dots,k-s-r$, we alternate blocks of the next $\alpha_{s+i} - \alpha_{s+i-1}$ elements of μ_r, and the next $\beta_i - \beta_{i-1}$ elements of ν_s. Then we place the next $m-r+s+1-\alpha_{k-r}$ elements of μ_r, so that the first $m-r+s+\beta_{k-s-r}$ positions of ρ are now filled. Next we place the remaining r elements of μ_r in positions $m-r+s+\beta_{k-s-r+1},\dots,m-r+s+\beta_{k-s}$, and fill the remaining $n-s-\beta_{k-s-r}$ positions of ρ with the final $n-s-\beta_{k-s-r}$ elements of ν_s. For example if $m=5$, $n=4$, $r=2$, $s=1$, $k=5$, $\alpha=\{1,2,4\} \subseteq \mathscr{N}_4$ and $\beta=\{1,2,3,5\} \subseteq \mathscr{N}_5$, then $\rho = 763485291 \in \mathscr{S}(34521, 7689)$.

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

$$\mathscr{D}(\rho) = \{\alpha_1,\dots,\alpha_s, \alpha_{s+1}+\beta_1-1,\dots,\alpha_{k-r}+\beta_{k-s-r}-1,$$
$$m-r+s-1+\beta_{k-s-r+1},\dots,m-r+s-1+\beta_{k-s}\},$$

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

Case 2 ($k \leqslant r+s$). Let $\{\gamma_1,\dots,\gamma_{m-k+s}\} = \mathscr{N}_{m-r+s} - \alpha$ and $\{\delta_1,\dots,\delta_{n-k+r}\} = \mathscr{N}_{n-s+r} - \beta$, where $\gamma_1 < \cdots < \gamma_{m-k+s}$ and $\delta_1 < \cdots < \delta_{n-k+r}$. Place the first $m-r$ elements of μ_r in positions $\gamma_1,\dots,\gamma_{m-r}$, and put the first $\gamma_{m-r}-m+r$ elements of ν_s in the remaining positions from 1 to γ_{m-r}, so that the first γ_{m-r} positions of ρ are filled. We follow with the next $\delta_1 - 1$ elements of μ_r and then, for $i = 1,\dots,s+r-k$, we alternate blocks of $\gamma_{m-r+i} - \gamma_{m-r+i-1}$ elements of ν_s and blocks of $\delta_{i+1} - \delta_i$ elements of μ_r. Then we place the next $m-r+s-\gamma_{m-k+s}$ elements of ν_s, so that the first $m-r+s-1+\delta_{s+r-k+1}$ positions of ρ are filled. Then we place the remaining $n-s$ elements of ν_s in positions $m-r+s+\delta_{s+r-k+1},\dots,m-r+s+\delta_{n-k+r}$, and fill the remaining $r+1-\delta_{s+r-k+1}$ positions with the final $r+1-\delta_{s+r-k+1}$ elements of μ_r. Thus we can identify positions that are not descents, and have

$$\mathscr{D}(\rho) = \mathscr{N}_{m+n-1} - \{\gamma_1,\dots,\gamma_{m-r-1}, \gamma_{m-r}+\delta_1-1,\dots,\gamma_{m-k+s}+\delta_{s+r-k+1}-1,$$
$$m-r+s-1+\delta_{s+r-k+2},\dots,m-r+s-1+\delta_{n-k+r}\},$$

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 α and β associated with each $\rho \in \mathscr{S}(\mu_r, \nu_s)$ with $d(\rho) = k$, so the construction is bijective. Thus

$$\begin{aligned} S_k(\mu_r, \nu_s) &= \sum_{\substack{\rho \in \mathscr{S}(\mu_r, \nu_s) \\ d(\rho) = k}} q^{I(\rho)} \\ &= q^{s-k+r(m-r+s)} \sum_{1 \leqslant \alpha_1 < \cdots < \alpha_{k-r} \leqslant m-r+s} q^{\alpha_1 + \cdots + \alpha_{k-r}} \\ &\quad \times \sum_{1 \leqslant \beta_1 < \cdots < \beta_{k-s} \leqslant n-s+r} q^{\beta_1 + \cdots + \beta_{k-s}} \\ &= q^{s-k+r(m-r+s)} q^{\binom{k-r+1}{2}} \begin{bmatrix} m - r + s \\ k - r \end{bmatrix} q^{\binom{k-s+1}{2}} \begin{bmatrix} n - s + r \\ k - s \end{bmatrix} \end{aligned}$$

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

$$\rho = 5\ \underline{10}\ 8\ 4\ \underline{12}\ 2\ 7\ 6\ \underline{13}\ \underline{11}\ 3\ 1\ 9 \in \mathscr{S}(5\ \underline{10}\ \underline{12}\ 2\ 7\ \underline{13}\ \underline{11}\ 9, 8\ 4\ 6\ 3\ 1),$$

so $\sigma = 5\ \underline{10}\ \underline{12}\ 2\ 7\ \underline{13}\ \underline{11}\ 9$, $\omega = 84631$, $m = 8$, $n = 5$, $r = s = 3$, $I(\sigma) = 16$, $I(\omega) = 8$, $d(\rho) = k = 7$ and $I(\rho) = 47$.

Then, in the notation of Lemma 3.1, we represent ρ by $\phi_{\sigma,\omega}(\rho) = b = A^2U^2A^3UA^2U^2A$. Moreover $\mathbf{t} = \{3,6,7\}$, $\mathbf{l} = \{1,3,4\}$, $a_{00} = AUA^2$, $a_{10} = A^2UA_{(3,0)}$, $a_{20} = UA_{(6,0)}$, $a_{30} = UA_{(7,0)}$, $a_{01} = UAUA^2_{(0,1)}$, $a_{11} = AU^2A^2_{(3,1)}$, $a_{21} = U^2A_{(6,1)}$, $a_{31} = U^2A_{(7,1)}$, $a_{02} = UA^3_{(0,3)}$, $a_{12} = AUA^2_{(3,3)}$, $a_{22} = UA_{(6,3)}$, $a_{32} = UA_{(7,3)}$, $a_{03} = UA^3_{(0,4)}$, $a_{13} = UA^3_{(3,4)}$, $a_{23} = UA_{(6,4)}$ and $a_{33} = UA_{(7,4)}$. Also $\mathscr{F}_{\mathbf{t,l,a}}(b) = \{(2,0), (2,1), (3,2), (5,2), (6,3), (7,3), (7,4)\}$, so indeed $|\mathscr{F}_{\mathbf{t,l,a}}(b)| = 7 = d(\rho)$ and $w(\mathscr{F}_{\mathbf{t,l,a}}(b)) = 47 = I(\rho)$.

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

In Theorem 3.3, by applying ξ twice ($mr - \binom{r+1}{2} - \sum_{i=1}^{r} t_i = 2$), we get the path $b'' = AU^2A^2UAUA^4U$, and obtain $\mathscr{F}_{m-\mathbf{r,l,c}}(b'') = \{(1,1), (1,2), (3,3), (4,4), (5,4), (6,4), (7,4)\}$, so $|\mathscr{F}_{m-\mathbf{r,l,c}}(b'')| = 7 = d(\rho)$ and $w(\mathscr{F}_{m-\mathbf{r,l,c}}(b'')) = 49 = I(\rho) + 2$, as required. Finally, by similarly reducing $\mathbf{l}$ to $\mathbf{s}$, we obtain the path $b_0 = AU^2AUAUA^5U$, with $\mathscr{F}_{m-\mathbf{r,s,c}}(b_0) = \{(1,1), (1,2), (2,3), (3,4), (5,4), (6,4), (7,4)\}$, so $|\mathscr{F}_{m-\mathbf{r,s,c}}(b_0)| = 7 = d(\rho)$ and $w(\mathscr{F}_{m-\mathbf{r,s,c}}(b_0)) = 47 = I(\rho) + \Delta - \sum_{i=1}^{r} t_i - \sum_{i=1}^{s} l_i$, as required.

In Theorem 3.4, we finally obtain that $\rho \in \mathscr{S}(\sigma, \omega)$ corresponds to

$$\rho' = 4\ \underline{12}\ \underline{11}\ 5\ \underline{10}\ 6\,9\,7\,8\,3\,2\,1\ \underline{13} \in \mathscr{S}(\mu_r, \nu_s),$$

where $\mu_r = 45678321$ and $\nu_s = \underline{12}\ \underline{11}\ \underline{10}\ 9\ \underline{13}$, and $d(\rho') = 7 = d(\rho)$, $I(\rho') = 47 = I(\rho) + \Delta - I(\sigma) - I(\omega)$.

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 ρ'. Of course $I(\rho') = 47 = s - k + r(m - r + s) + \sum_{i=1}^{k-r} \alpha_i + \sum_{i=1}^{k-s} \beta_i$. $\square$

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

$$\rho = 5\ \underline{10}\ 8\,4\ \underline{12}\ 2\,7\,6\ \underline{13}\ \underline{11}\ 3\,1\,9 \in \mathscr{S}(5\ \underline{10}\ \underline{12}\ 2\,7\ \underline{13}\ \underline{11}\ 9, 8\,4\,6\,3\,1)$$

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

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

1. G. E. Andrews, *Identities in combinatorics*, I: *On sorting two ordered sets*, Discrete Math. **11** (1975), 97–106.

2. ______, *The theory of partitions*, Encyclopedia of Mathematics and its Applications, Vol. 2, Addison-Wesley, Reading, Mass., 1976.

3. P. Cartier and D. Foata, *Problèmes combinatoires de commutation et rearrangements*, Lecture Notes in Math., vol. 85, Springer, Berlin, 1969.

4. H. W. Gould, *A new symmetrical combinatorial identity*, J. Combin. Theory **13** (1972), 278–286.
5. I. P. Goulden, *A bijective proof of the q-Saalschutz theorem* (preprint).
6. I. P. Goulden and D. M. Jackson, *Combinatorial enumeration*, Wiley, New York, 1983.
7. F. H. Jackson, *Transformations of q-series*, Messenger of Math. **39** (1910), 145–153.
8. P. A. MacMahon, *Combinatory analysis*, Vols. I and II, Chelsea, New York, 1960.
9. L. J. Slater, *Generalized hypergeometric functions*, Cambridge Univ. Press, Cambridge, 1966.
10. R. P. Stanley, *Ordered structures and partitions*, Mem. Amer. Math. Soc. No. 119 (1972), 104 pp.

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