Generalized Riffle Shuffles and Quasisymmetric functions

Richard P. Stanley[1]
Department of Mathematics 2-375
Massachusetts Institute of Technology
Cambridge, MA 02139

version of 31 May 2001

1 Introduction.

Let x_i be a probability distribution on a totally ordered set I, i.e., the probability of $i \in I$ is x_i. (Hence $x_i \geq 0$ and $\sum x_i = 1$.) Fix $n \in \mathbb{P} = \{1, 2, \ldots\}$, and define a random permutation $w \in \mathfrak{S}_n$ as follows. For each $1 \leq j \leq n$, choose independently an integer i_j (from the distribution x_i). Then *standardize* the sequence $\boldsymbol{i} = i_1 \cdots i_n$ in the sense of [34, p. 322], i.e., let $\alpha_1 < \cdots < \alpha_k$ be the elements of I actually appearing in $\boldsymbol{i}$, and let a_i be the number of α_i's in $\boldsymbol{i}$. Replace the α_1's in $\boldsymbol{i}$ by $1, 2, \ldots, a_1$ from left-to-right, then the α_2's in $\boldsymbol{i}$ by $a_1 + 1, a_1 + 2, \ldots, a_1 + a_2$ from left-to-right, etc. For instance, if $I = \mathbb{P}$ and $\boldsymbol{i} = 311431$, then $w = 412653$. This defines a probability distribution on the symmetric group $\mathfrak{S}_n$, which we call the *QS-distribution* (because of the close connection with quasisymmetric functions explained below). If we need to be explicit about the parameters $x = (x_i)_{i \in I}$, then we will refer to the *QS*(x)*-distribution*.

The QS-distribution is not new, at least when I is finite. It appears for instance in [10, pp. 153–154][14][20][24]. Although these authors anticipate some of our results, they don't make systematic use of quasisymmetric and symmetric functions. Some additional work using the viewpoint of our paper is due to Fulman [15].

1.1 Example. As an example of a general result to be proved later (Theorem 2.1), let us compute Prob(213), the probability that a random

[1]Partially supported by NSF grant DMS-9988459

permutation $w \in \mathfrak{S}_3$ (chosen from the QS-distribution) is equal to the permutation 213. A sequence $\boldsymbol{i} = i_1 i_2 i_3$ will have the standardization 213 if and only if $i_2 < i_1 \leq i_3$. Hence

$$\text{Prob}(213) = \sum_{a<b\leq c} x_a x_b x_c.$$

There is an alternative way to describe the QS-distribution. Suppose we have a deck of n cards. Cut the deck into random packets of respective sizes $a_i > 0,\ i \in I$, such that $\sum a_i = n$ and the probability of $(a_i)_{i\in I}$ is

$$\text{Prob}\left((a_i)_{i\in I}\right) = n! \prod_{i\in I} \frac{x_i^{a_i}}{a_i!}.$$

Then riffle shuffle these packets P_i together into a pile P in the manner described by Bayer and Diaconis [2] (see also [11][14]); namely, after placing k cards on P, the probability that the next card comes from the top of the current packet P_j is proportional to the current number of cards in P_j. This card is then placed at the bottom of P. The ordinary dovetail or riffle shuffle [2, p. 294] corresponds to the case $I = \{1, 2\}$ and $x_1 = x_2 = 1/2$. In this case the original deck is cut according to the binomial distribution. More generally, if for fixed $q \in \mathbb{P}$ we cut the deck into q packets (some possibly empty) according to the q-multinomial distribution, then we obtain the q-shuffles of Bayer and Diaconis [2, p. 299] (where they use a for our q). The q-shuffle is identical to the QS-distribution for $I = \{1, 2 \ldots, q\}$ and $x_1 = x_2 = \cdots = x_q = 1/q$. We will denote this distribution by U_q. If we q-shuffle a deck and then r-shuffle it, the distribution is the same as a single qr-shuffle [2, Lemma 1]. In other words, if $*$ denotes convolution of probability distributions then $U_q * U_r = U_{qr}$. We extend this result to the QS-distribution in Theorem 2.4.

In the next section we will establish the connection between the QS-distribution and the theory of quasisymmetric functions. In Section 3 we use known results from the theory of quasisymmetric and symmetric functions to obtain results about the QS-distribution, many of them direct generalizations of the work of Bayer, Diaconis, and Fulman mentioned above. For instance, we show that the probability that a random permutation $w \in \mathfrak{S}_n$ (always assumed chosen from the QS-distribution) has k inversions is equal to the probability that w has major index k. This is an analogue of MacMahon's corresponding result for the *uniform* distribution on $\mathfrak{S}_n$ (see e.g. [33, p. 23]). A further result is that if T' is a standard Young tableau of shape λ, then the probability that T' is the recording tableau of w under the Robinson-Schensted-Knuth algorithm is $s_\lambda(x)$, where $x = (x_i)_{i\in I}$ and s_λ denotes a Schur function.

2 The QS-distribution and quasisymmetric functions

In this section we show the connection between the QS-distribution and quasisymmetric functions. In general our reference for quasisymmetric functions and symmetric functions will be [34, Ch. 7]. Quasisymmetric functions may be regarded as bearing the same relation to compositions (ordered partitions of a nonnegative integer) as symmetric functions bear to partitions. Namely, the homogeneous symmetric functions of degree n form a vector space of dimension $p(n)$ (the number of partitions of n) and have many natural bases indexed by the partitions of n. Similarly, the homogeneous quasisymmetric functions of degree n form a vector space of dimension 2^{n-1} (the number of compositions of n) and have many natural bases indexed by the compositions of n.

A *quasisymmetric function* $F(z)$ is a power series of bounded degree, say with rational coefficients, in the variables $z = (z_i)_{i\in I}$, with the following property: let $i_1 < i_2 < \cdots < i_n$ and $j_1 < j_2 < \cdots < j_n$, where $i_k, j_k \in I$. If $a_1, \ldots, a_k \in \mathbb{P}$, then the coefficient of $z_{i_1}^{a_1} \cdots z_{i_n}^{a_n}$ is equal to the coefficient of $z_{j_1}^{a_1} \cdots z_{j_n}^{a_n}$ in $F(z)$. The set of all quasisymmetric functions forms a graded $\mathbb{Q}$-algebra denoted $\mathcal{Q}$. The quasisymmetric functions that are homogeneous of degree n form a $\mathbb{Q}$-vector space, denoted $\mathcal{Q}_n$. If $|I| \geq n$ (including $|I| = \infty$), then $\dim \mathcal{Q}_n = 2^{n-1}$.

Let $\alpha = (\alpha_1, \ldots, \alpha_k)$ be a *composition* of n, i.e., $\alpha_i \in \mathbb{P}$ and $\sum \alpha_i = n$. Let $\mathrm{Comp}(n)$ denote the set of all compositions of n, so $\#\mathrm{Comp}(n) = 2^{n-1}$ [33, pp. 14–15]. Define

$$S_\alpha = \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{k-1}\} \subseteq \{1, \ldots, n-1\}.$$

The *fundamental quasisymmetric function* $L_\alpha = L_\alpha(z)$ is defined by

$$L_\alpha(z) = \sum_{\substack{i_1 \le \cdots \le i_n \\ i_j < i_{j+1} \text{ if } j \in S_\alpha}} z_{i_1} \cdots z_{i_n}.$$

It is not hard to show that when $|I| \ge n$, the set $\{L_\alpha : \alpha \in \mathrm{Comp}(n)\}$ is a $\mathbb{Q}$-basis for $\mathcal{Q}_n$ [34, Prop. 7.19.1]. More generally, if $|I| = k$ then the set $\{L_\alpha : \alpha \in \mathrm{Comp}(n),\ \mathrm{length}(\alpha) \le k\}$ is a $\mathbb{Q}$-basis for cq_n, and $L_\alpha = 0$ if $\mathrm{length}(\alpha) > k$.

If $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ then let $D(w)$ denote the *descent set* of w, i.e.,

$$D(w) = \{i : w_i > w_{i+1}\}.$$

Write $\mathrm{co}(w)$ for the unique composition of n satisfying $S_{\mathrm{co}(w)} = D(w)$. In other words, if $\mathrm{co}(w) = (\alpha_1, \ldots, \alpha_k)$, then w has descents exactly at α_1, $\alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{k-1}$. For any composition α of n it is easy to see that the (possibly infinite) series $L_\alpha(x)$ (where as always $x_i \ge 0$ and $\sum x_i = 1$) is absolutely convergent and therefore represents a nonnegative real number.

The main result on which all our other results depend is the following.

2.1 Theorem. *Let $w \in \mathfrak{S}_n$. The probability $\mathrm{Prob}(w)$ that a permutation in $\mathfrak{S}_n$ chosen from the QS-distribution is equal to w is given by*

$$\mathrm{Prob}(w) = L_{\mathrm{co}(w^{-1})}(x).$$

Proof. This result is equivalent to [18, Lemma 3.2] (see also [30, Lemma 9.39]). Because of its importance here we present the (easy) proof. The integers i for which $i+1$ appears somewhere to the left of i in $w = w_1 w_2 \cdots w_n$ are just the elements of $D(w^{-1})$. Let $i_1, \ldots, i_n$ be chosen independently from the distribution x_i. Let $a_j = i_{w^{-1}(j)}$, i.e., if we write down $i_1 \cdots i_n$ underneath $w_1 \cdots w_n$, then a_j appears below j. In order for the sequence $i_1 \cdots i_n$ to standardize to w it is necessary and sufficient that $a_1 \le a_2 \le \cdots \le a_n$ and that $a_j < a_{j+1}$ whenever $j \in D(w^{-1})$. Hence

$$\begin{aligned} \mathrm{Prob}(w) &= \sum_{\substack{a_1 \le a_2 \le \cdots \le a_n \\ a_i < a_{i+1} \text{ if } i \in D(w^{-1})}} x_{a_1} x_{a_2} \cdots x_{a_n} \\ &= L_{\mathrm{co}(w^{-1})}(x). \quad \square \end{aligned}$$

For some algebraic results closely related to Theorem 2.1, see [12].

Two special cases of Theorem 2.1 are of particular interest. The first is the q-shuffle distribution U_q. Let $\operatorname{des}(u)$ denote the number of *descents* of the permutation $u \in \mathfrak{S}_n$, i.e.,

$$\operatorname{des}(u) = \#\{i : w(i) > w(i+1)\}.$$

It follows easily from Theorem 2.1 or by a direct argument (see [34, p. 364]) that the probability $\operatorname{Prob}_{U_q}(w)$ that a random permutation in $\mathfrak{S}_n$ chosen from the distribution U_q is equal to w is given by

$$\operatorname{Prob}_{U_q}(w) = \binom{q - \operatorname{des}(w^{-1}) + n - 1}{n} q^{-n}.$$

This is just the probability of obtaining w with a q-shuffle, as defined in [2, §3], confirming that U_q is just the q-shuffle distribution.

For fixed n, the uniform distribution U on $\mathfrak{S}_n$ is given by

$$U = \lim_{q \to \infty} U_q. \tag{1}$$

Because of this formula many of our results below may be considered as generalizations of the corresponding results for the uniform distribution. These results for the uniform distribution are well-known and classical theorems in the theory of permutation enumeration.

The second interesting special case of Theorem 2.1 is defined by

$$x_i = (1-t)t^{i-1},$$

where $I = \mathbb{P}$ and $0 \leq t \leq 1$. Given $u = u_1 u_2 \cdots u_n \in \mathfrak{S}_n$, define $\bar{u} = v_1 v_2 \cdots v_n$, where $v_i = n + 1 - u_{n+1-i}$. (Equivalently, $\bar{u} = w_0 u w_0$, where $w_0(i) = n + 1 - i$.) Set $e(u) = \operatorname{maj}(\bar{u})$, where maj denotes *major index* (or *greater index*) of u [33, p. 216], i.e.,

$$\operatorname{maj}(u) = \sum_{i \in D(u)} i.$$

It follows from [34, Lemma 7.19.10] that

$$\operatorname{Prob}(w) = \frac{t^{e(w^{-1})}(1-t)^n}{(1-t)(1-t^2)\cdots(1-t^n)}. \tag{2}$$

The QS-distribution defines a Markov chain (or random walk) $\mathcal{M}_n$ on $\mathfrak{S}_n$ by defining the probability $\mathrm{Prob}(u, uw)$ of the transition from u to uw to be $\mathrm{Prob}(w) = L_{\mathrm{co}(w^{-1})}(x)$. The next result determines the eigenvalues of this Markov chain. Equivalently, define

$$\Gamma_n(x) = \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(x) w \in \mathbb{R}\mathfrak{S}_n,$$

where $\mathbb{R}\mathfrak{S}_n$ denotes the group algebra of $\mathfrak{S}_n$ over $\mathbb{R}$. Then the eigenvalues of $\mathcal{M}_n$ are just the eigenvalues of $\Gamma_n(x)$ acting on $\mathbb{R}\mathfrak{S}_n$ by right multiplication.

2.2 Theorem. *The eigenvalues of $\mathcal{M}_n$ are the power sum symmetric functions $p_\lambda(x)$ for $\lambda \vdash n$. The eigenvalue $p_\lambda(x)$ occurs with multiplicity $n!/z_\lambda$, the number of elements in $\mathfrak{S}_n$ of cycle type λ.*

The above theorem appears implicitly in [16, Thm. 4.4] and more explicitly in [23, Note 5.20]. See also [13, §7] and [34, Exer. 7.96]. It is also a special case of [3, Thm. 1.2], as we now explain.

Bidigare, Hanlon, and Rockmore [3] define a random walk on the set $\mathcal{R}$ of regions of a (finite) hyperplane arrangement $\mathcal{A}$ in $\mathbb{R}^n$ as follows. (Precise definitions of the terms used below related to arrangements may be found in [3]. For further information see [9].) Let $\mathcal{F}$ be the set of *faces* of $\mathcal{A}$, i.e., the nonempty faces of the closure of the regions of $\mathcal{A}$. Let wt be a probability distribution on $\mathcal{F}$. Given $R \in \mathcal{R}$ and $F \in \mathcal{F}$, define $F \cdot R$ to be that region R' of $\mathcal{A}$ closest to R that has F as a face. (There is always a unique such region R'.) Now given a region R, choose $F \in \mathcal{F}$ with probability $\mathrm{wt}(F)$ and move to the region $F \cdot R$.

Consider the special case when $\mathcal{A}$ is the braid arrangement $\mathcal{B}_n$, i.e., the set of hyperplanes $u_i = u_j$ for $1 \le i < j \le n$, where $u_1, \ldots, u_n$ are the coordinates in $\mathbb{R}^n$. The set of regions of $\mathcal{B}_n$ can be identified in a natural way with $\mathfrak{S}_n$, viz., identify the region $u_{a_1} < u_{a_2} < \cdots < u_{a_n}$ with the permutation w given by $w(a_i) = i$. The faces of $\mathcal{B}_n$ correspond to *ordered set partitions* π of $1, 2, \ldots, n$, i.e., $\pi = (B_1, \ldots, B_k)$ where $B_i \neq \emptyset$, $B_i \cap B_j = \emptyset$ for $i \neq j$, and $\cup B_i = \{1, 2, \ldots, n\}$; viz., π corresponds to the face defined by $u_r = u_s$ if $r, s \in B_i$ for some i, and $u_r < u_s$ if $r \in B_i$ and $s \in B_j$ with $i < j$. Define the *type* of π to be the composition

$$\mathrm{type}(\pi) = (\#B_1, \ldots, \#B_k) \in \mathrm{Comp}(n).$$

Given $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathrm{Comp}(n)$, define the *monomial quasisymmetric function* $M_\alpha(z)$ by

$$M_\alpha(z) = \sum_{i_1 < \cdots < i_k} z_{i_1}^{\alpha_1} \cdots z_{i_k}^{\alpha_k}.$$

(See [34, (7.87)].) Now if F is a face of $\mathcal{B}_n$ corresponding to the ordered partition π, then set

$$\mathrm{wt}(F) = M_{\mathrm{type}(\pi)}(x).$$

It is easy to see that wt is a probability distribution on the set of faces of $\mathcal{B}_n$; in fact, for indeterminates $(z_i)_{i \in I}$ we have

$$\sum_\pi M_{\mathrm{type}(\pi)}(z) = \left(\sum z_i\right)^n.$$

Thus wt defines a Bidigare-Hanlon-Rockmore random walk on the regions of $\mathcal{B}_n$. Let $P(R, R')$ denote the probability of moving from a region R to a region R'.

2.3 Theorem. *Let R correspond to $u \in \mathfrak{S}_n$ and R' to uw. Then*

$$P(R, R') = L_{\mathrm{co}(w^{-1})}(x).$$

Hence the Bidigare-Hanlon-Rockmore random walk just defined on the regions of $\mathcal{B}_n$ is isomorphic to the random walk on $\mathfrak{S}_n$ defined by the QS-distribution.

Proof. By symmetry we may suppose that $u = 12 \cdots n$, the identity permutation. The faces F of R' for which $F \cdot R = R'$ correspond to those ordered partitions π obtained from $w^{-1} = w_1 \cdots w_n$ by drawing bars between certain pairs w_i and w_{i+1}, where there must be a bar when $i \in D(w^{-1})$. The elements contained between the consecutive bars (including a bar before w_1 and after w_n) form the blocks of π, read from left-to-right. For instance, if $w^{-1} = 582679134$, then one way of drawing bars is $|58|267|9|1|34|$, yielding the ordered partition $\pi = (\{5,8\}, \{2,6,7\}, \{9\}, \{1\}, \{3,4\})$. It follows that

$$P(R, R') = \sum_\beta M_\beta(x),$$

where β runs over all compositions refining $\mathrm{co}(w^{-1})$. From the definition of $L_\alpha(x)$ we have

$$\sum_\beta M_\beta(x) = L_{\mathrm{co}(w^{-1})}(x),$$

and the proof follows. $\square$

The next result determines the convolution of the $\mathrm{QS}(x)$-distribution with the $\mathrm{QS}(y)$-distribution. We write xy for the variables x_iy_j in the order $x_iy_j < x_ry_s$ if either $i < r$ or $i = r$ and $j < s$.

2.4 Theorem. *Suppose that a permutation $u \in \mathfrak{S}_n$ is chosen from the QS(x)-distribution, and a permutation $v \in \mathfrak{S}_n$ is chosen from the QS(y)-distribution. Let $w \in \mathfrak{S}_n$. Then the probability that $uv = w$ is equal to $L_{\mathrm{co}(w^{-1})}(xy)$. In other words, if $*$ denotes convolution then*

$$\mathrm{QS}(x) * \mathrm{QS}(y) = \mathrm{QS}(xy).$$

Equivalently, in the ring $\mathbb{R}\mathfrak{S}_n$ we have

$$\Gamma_n(x)\Gamma_n(y) = \Gamma_n(xy).$$

Theorem 2.4 is equivalent to [23, (62)][30, Thm. 9.37]. For the U_q-distribution it was also proved in [2, Lemma 1], and this proof can be easily extended to prove Theorem 2.4 in its full generality.

3 Enumerative properties of the QS-distribution

Suppose that X is any subset of $\mathfrak{S}_n$. Let $w \in \mathfrak{S}_n$ be chosen from the QS-distribution. It follows from Theorem 2.1 that the probability that $w \in X$ is given by

$$\mathrm{Prob}(w \in X) = \sum_{u \in X} L_{\mathrm{co}(u^{-1})}(x). \tag{3}$$

There are many known instances of the set $X \subseteq \mathfrak{S}_n$ for which the right-hand side of (3) can be explicitly evaluated. Most of this section will be devoted to some examples of this type.

Our first result involves the symmetric function $l_n(z)$ defined by

$$l_n(z) = \frac{1}{n}\sum_{d|n} \mu(d) p_d^{n/d}(z),$$

where μ denotes the usual number-theoretic Möbius function. More information on $l_n(z)$ may be found in [30, Ch. 8] or [34, Exer. 7.88–7.89].

3.1 Theorem. *Let w be a random permutation in $\mathfrak{S}_n$, chosen from the QS-distribution. The probability* $\mathrm{Prob}(\rho(w) = \lambda)$ *that w has cycle type $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle \vdash n$ (i.e., m_i cycles of length i) is given by*

$$\mathrm{Prob}(\rho(w) = \lambda) = \prod_{i \geq 1} h_{m_i}[l_i](x),$$

where brackets denote plethysm [26, §I.8][34, p. 447].

The proof of Theorem 3.1 follows from [30, Thm. 8.23 and Thm. 9.41(a)]. In the special case of the U_q-distribution, the result appears in [11, Thm. A].

As above let $\mathrm{maj}(w)$ denote the major index of $w \in \mathfrak{S}_n$, and let $\mathrm{inv}(w)$ denote the number of inversions of w [33, pp. 20–21], i.e.,

$$\mathrm{inv}(w) = \#\{(i,j) \,:\, 1 \leq i < j \leq n,\ w(i) > w(j)\}.$$

Let $I_n(j)$ (respectively, $M_n(j)$) denote the probability that $w \in \mathfrak{S}_n$ satisfies $\mathrm{inv}(w) = j$ (respectively, $\mathrm{maj}(w) = j$) under the QS-distribution.

3.2 Theorem. *We have*

$$\begin{aligned} M_n(j) &= I_n(j) \\ \sum_{n \geq 0} \sum_{j \geq 0} \frac{M_n(j) t^j z^n}{(1-t)(1-t^2)\cdots(1-t^n)} &= \prod_{i \geq 1} \prod_{j \in I} \left(1 - t^{i-1} x_j z\right)^{-1}. \end{aligned} \tag{4}$$

Proof. The standardization w of a sequence $\boldsymbol{i} = i_1 i_2 \cdots i_n$ satisfies $\mathrm{maj}(w) = \mathrm{maj}(\boldsymbol{i})$ and $\mathrm{inv}(w) = \mathrm{inv}(\boldsymbol{i})$. Hence by definition of the QS-distribution we have

$$\begin{aligned} \sum_j M_n(j) t^j &= \sum_{\boldsymbol{i} = i_1 \cdots i_n} t^{\mathrm{maj}(\boldsymbol{i})} x_{i_1} \cdots x_{i_n} \\ \sum_j I_n(j) t^j &= \sum_{\boldsymbol{i} = i_1 \cdots i_n} t^{\mathrm{inv}(\boldsymbol{i})} x_{i_1} \cdots x_{i_n}. \end{aligned}$$

Thus if

$$\begin{aligned} F_\lambda(t) &= \sum_v t^{\mathrm{maj}(v)} \\ G_\lambda(t) &= \sum_v t^{\mathrm{inv}(v)}, \end{aligned}$$

where v ranges over all permutations of the multiset $\{1^{\lambda_1}, 2^{\lambda_2}, \ldots\}$, then

$$\begin{aligned} \sum_j M_n(j)t^j &= \sum_{\lambda \vdash n} F_\lambda(t) m_\lambda(x) \\ \sum_j I_n(j)t^j &= \sum_{\lambda \vdash n} G_\lambda(t) m_\lambda(x), \end{aligned}$$

where $m_\lambda(x)$ denotes a monomial symmetric function. But it is known [32, (45), p. 97][33, Prop. 1.3.17] that

$$F_\lambda(t) = G_\lambda(t) = \left[\begin{array}{c} n \\ \lambda_1, \lambda_2, \ldots \end{array}\right]_t,$$

a q-multinomial coefficient in the variable t. Hence $M_n(j) = I_n(j)$.

Now let $h_\lambda(x)$ denote a complete homogeneous symmetric function [34, §7.5]. It is easy to see [34, Prop. 7.8.3] that

$$h_\lambda(1, t, t^2, \ldots) = \prod_i \frac{1}{(1-t)(1-t^2)\cdots(1-t^{\lambda_i})}.$$

Hence

$$\sum_j M_n(j)t^j = (1-t)\cdots(1-t^n) \sum_{\lambda \vdash n} h_\lambda(1, t, t^2, \ldots) m_\lambda(x).$$

Equation (4) then follows immediately from the expansion [34, (7.10)] of the Cauchy product $\prod_{i,j}(1 - x_i y_j)^{-1}$. Equation (4) is also equivalent to [17, (78)]. $\square$

Other expansions of the Cauchy product lead to further formulas for $\sum_j M_n(j)t^j$. In particular, from [34, (7.20), Thm. 7.12.1, and Cor. 7.21.3] there follows (using notation from [34])

$$
\begin{aligned}
\sum_j M_n(j)t^j &= \sum_{\lambda \vdash n} t^{b(\lambda)} \frac{(1-t)\cdots(1-t^n)}{\prod_{u\in\lambda}(1-t^{h(u)})} s_\lambda(x) \\
&= \sum_{\lambda \vdash n} \frac{(1-t)\cdots(1-t^n)}{(1-t^{\lambda_1})\cdots(1-t^{\lambda_l})} z_\lambda^{-1} p_\lambda(x). \qquad (5)
\end{aligned}
$$

A famous result of MacMahon asserts that

$$\#\{w \in \mathfrak{S}_n : \mathrm{maj}(w) = j\} = \#\{w \in \mathfrak{S}_n : \mathrm{inv}(w) = j\}.$$

Since the uniform distribution U on $\mathfrak{S}_n$ satisfies $U = \lim_{q\to\infty} U_q$, it follows that Theorem 3.2 is a generalization of MacMahon's result. In fact, the proof we have given of Theorem 3.2 shows that it is equivalent to the equidistribution of maj and inv on the set of permutations of any finite *multiset* (where the underlying set is linear ordered). This result was also known, at least implicitly, to MacMahon. For further references and some generalizations, see [5].

In principle one can use Theorem 3.2 to derive the moments of $\mathrm{maj}(w)$ (or equivalently $\mathrm{inv}(w)$). Let E denote expectation with respect to the QS-distribution and E_U expectation with respect to the uniform distribution. The next result obtains the first two moments of $\mathrm{maj}(w)$, stated for simplicity as the expectations of $\mathrm{maj}(w)$ and $\mathrm{maj}(w)(\mathrm{maj}(w)-1)$ (instead of $\mathrm{maj}(w)^2$).

3.3 Corollary. *Choose $w \in \mathfrak{S}_n$ from the QS-distribution. Then*

$$
\begin{aligned}
E(\mathrm{maj}(w)) &= E_U(\mathrm{maj}(w)) - \frac{1}{2}\binom{n}{2} p_2(x) \qquad (6) \\
E(\mathrm{maj}(w)(\mathrm{maj}(w)-1)) &= E_U(\mathrm{maj}(w)(\mathrm{maj}(w)-1)) \qquad (7) \\
&\quad -3\binom{n+1}{4} p_2(x) + \frac{4}{3}\binom{n}{3} p_3(x) + \frac{3}{2}\binom{n}{4} p_2(x)^2. \qquad (8)
\end{aligned}
$$

NOTE. It is easy to see (e.g. [22, p. 16]) that

$$
\begin{aligned}
E_U(\mathrm{maj}(w)) &= \frac{1}{2}\binom{n}{2} \\
E_U(\mathrm{maj}(w)(\mathrm{maj}(w)-1)) &= \frac{n(n-1)(n-2)(9n+13)}{144}.
\end{aligned}
$$

Proof of Corollary 3.3. Let

$$\Lambda_n(t) = \sum_j M_n(j) t^j.$$

Then

$$E(\mathrm{maj}(w)) = \Lambda_n'(1).$$

When we differentiate (5) term-by-term and set $t = 1$, the only surviving terms will be from $\lambda = \langle 1^n \rangle$ and $\lambda = \langle 21^{n-2} \rangle$. Hence

$$E(\mathrm{maj}(w)) = \frac{d}{dt} \left[\frac{(1-t)\cdots(1-t^n)}{(1-t)^n} \frac{1}{n!} p_1(x)^n \right.$$
$$\left. + \frac{(1-t)\cdots(1-t^n)}{(1-t^2)(1-t)^{n-2}} \frac{1}{2\cdot(n-2)!} p_1(x)^{n-2} p_2(x) \right]_{t=1}. \qquad (9)$$

Now $p_1(x) = \sum x_i = 1$ and

$$\frac{d}{dt} \left[\frac{(1-t)\cdots(1-t^n)}{(1-t)^n} \frac{1}{n!} \right]_{t=1} = E_U(\mathrm{maj}(w)).$$

It is routine to compute the second term on the right-hand side of (9), obtaining equation (6).

A similar computation using $E(\mathrm{maj}(w)(\mathrm{maj}(w)-1)) = \Lambda_n''(1)$ yields (8); we omit the details. $\square$

It is clear from the above proof that for general $k \geq 1$, $E(\mathrm{maj}(w)^k)$ is a linear combination, whose coefficients are polynomials in n, of power sum symmetric functions $p_\lambda(x)$ where $\lambda \vdash n$ and $\ell(\lambda) \geq n-k$. Here $\ell(\lambda)$ denotes the length (number of parts) of λ.

We next consider the relationship between the QS-distribution and the Robinson-Schensted-Knuth (RSK) algorithm [34, §7.11]. Recall that this algorithm associates a pair (T, T') of standard Young tableaux (SYT) of the same shape $\lambda \vdash n$ with a permutation $w \in \mathfrak{S}_n$. We call T' the *recording tableau* of w, denoted $\mathrm{rec}(w)$. The shape λ of T' is also called the *shape* of w, denoted $\mathrm{sh}(w) = \lambda$.

3.4 Theorem. *Choose $w \in \mathfrak{S}_n$ from the QS-distribution. Let T be an SYT of shape $\lambda \vdash n$. Then the probability that $\mathrm{rec}(w) = T$ is given by*

$$\mathrm{Prob}(\mathrm{rec}(w) = T) = s_\lambda(x),$$

where s_λ denotes a Schur function.

Proof. Let

$$X_T = \{w \in \mathfrak{S}_n \,:\, \mathrm{rec}(w) = T\}.$$

It follows from [34, Thm. 7.19.2 and Lemma 7.23.1] that

$$\sum_{w \in X_T} L_{\mathrm{co}(w^{-1})}(x) = s_\lambda(x).$$

The proof now follows from equation (3). $\square$

3.5 Corollary. *Choose $w \in \mathfrak{S}_n$ from the QS-distribution, and let $\lambda \vdash n$. Then*

$$\mathrm{Prob}(\mathrm{sh}(w) = \lambda) = f^\lambda s_\lambda(x),$$

where f^λ denotes the number of SYT of shape λ (given explicitly by the Frame-Robinson-Thrall hook-length formula [34, Cor. 7.21.6]).

Proof. There are f^λ SYT's T of shape λ, and by Theorem 3.4 they all have probability $s_\lambda(x)$ of being the recording tableau of w. $\square$

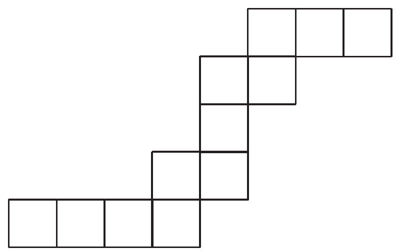
Note. If the RSK algorithm associates (T, T') with $w \in \mathfrak{S}_n$, then call T the *insertion tableau* of w, denoted $\mathrm{ins}(w)$. Since $\mathrm{ins}(w) = \mathrm{rec}(w^{-1})$ [34, Thm. 7.13.1], it follows that

$$\mathrm{Prob}(\mathrm{ins}(w) = T) = f^\lambda L_{\mathrm{co}(w)}(x).$$

Note. The probability distribution $\mathrm{Prob}(\lambda) = f^\lambda s_\lambda(x)$ on the set of all partitions λ of all nonnegative integers is a specialization of the z-measure of Borodin and Olshanski, as surveyed in [7] (see also [29]).

It is easy to give an expression for the probability that a permutation $w \in \mathfrak{S}_n$ chosen from the QS-distribution has a fixed descent set $S \subseteq \{1, 2, \ldots, n-$

1}. In general, if λ and μ are partitions with $\mu \subseteq \lambda$ (i.e., $\mu_i \leq \lambda_i$ for all i), then λ/μ denotes the *skew shape* obtained by removing the diagram of μ from that of λ [34, p. 309]. One can then define (see [34, Def. 7.10.1]) the *skew Schur function* $s_{\lambda/\mu}(x)$. A *border strip* (or *rim hook* or *ribbon*) is a connected skew shape with no 2×2 square [34, p. 245]. Let $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathrm{Comp}(n)$. Then there is a unique border strip B_α (up to translation) with α_i squares in row i. For instance, the border strip $B_{(3,2,1,2,4)}$ looks like



.

Some special properties of border strip Schur functions s_{B_α} are discussed in [34, §7.23].

3.6 Theorem. *Let w be a random permutation in $\mathfrak{S}_n$, chosen from the QS-distribution. Let $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathrm{Comp}(n)$, and recall that $S_\alpha = \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{k-1}\} \subseteq \{1, 2 \ldots, n-1\}$. Then the probability that w has descent set S_α is given by*

$$\mathrm{Prob}(D(w) = S_\alpha) = s_{B_\alpha}(x).$$

Proof. Immediate from Theorem 2.1 and the fact [34, Cor. 7.23.4] that

$$s_{B_\alpha} = \sum_{\substack{w \in \mathfrak{S}_n \\ \alpha = \mathrm{co}(w)}} L_{\mathrm{co}(w^{-1})}. \qquad \square$$

4 Longest increasing subsequences.

It is a fundamental result of Schensted [31] (see also [34, Cor. 7.23.11]) that if $\mathrm{sh}(w) = \lambda = (\lambda_1, \lambda_2, \ldots)$, then λ_1 is the length $\mathrm{is}(w)$ of the longest increasing subsequence of w. (There is a generalization due to Greene [19][34, Thm. 7.23.13 or A1.1.1] that gives a similar interpretation of any λ_i.) Thus the expected value $E_U(n)$ of is(w) for $w \in \mathfrak{S}_n$ under the uniform distribution is given by

$$E_U(n) = E_U(\text{is}(w)) = \frac{1}{n!} \sum_{\lambda \vdash n} \lambda_1 \left(f^\lambda\right)^2.$$

(See [34, Exer. 7.109(b)].) It was shown by Vershik and Kerov [35] that $E_U(n) \sim 2\sqrt{n}$. Later Baik *et al.* [1] obtained a much stronger result, viz., the exact distribution (suitably normalized) of λ_1 (or equivalently is(w)) in the limit $n \to \infty$. (This result was extended to any λ_i in [8][28][21].) We can ask whether similar results hold for the QS-distribution. Many results along these lines appear in [4][6, Ex. 2][8][27][29]. Here we make some elementary comments pointing in a different direction from these papers.

It is clear from Corollary 3.5 that

$$E(\text{is}(w)) = \sum_{\lambda \vdash n} \lambda_1 f^\lambda s_\lambda(x), \tag{10}$$

where as usual E denotes expectation with respect to the QS-distribution. This formula can be made more explicit in the special case of the distribution U_q. Identify $\lambda \vdash n$ with its Young diagram, and let $c(u)$ denote the content of the square $u \in \lambda$ as defined in [34, p. 373].

4.1 Theorem. *We have*

$$E_{U_q}(\text{is}(w)) = \frac{1}{n!} \sum_{\lambda \vdash n} \lambda_1 \left(f^\lambda\right)^2 \prod_{u \in \lambda} \left(1 + q^{-1} c(u)\right). \tag{11}$$

Proof. Let $s_\lambda(1^q)$ denote $s_\lambda(x)$ where $x_1 = \cdots = x_q = 1$ and $x_i = 0$ for $i > q$. It is well-known (e.g. [34, Cor. 7.21.4 and 7.21.6]) that

$$s_\lambda(1^q) = \frac{f^\lambda}{n!} \prod_{u \in \lambda} (q + c(u)).$$

Since $s_\lambda(x)$ is homogeneous of degree n, the proof follows from equation (10). $\square$

Note that $E_{U_q}(\text{is}(w))$ is a polynomial in q^{-1}, which can be written

$$E_{U_q}(\text{is}(w)) = E_U(\text{is}(w)) + \frac{1}{n!} \sum_{\lambda \vdash n} \lambda_1 \left(f^\lambda\right)^2 \left(\sum_{u \in \lambda} c(u) \right) \frac{1}{q} + \cdots. \tag{12}$$

Let us mention that

$$\sum_{u\in\lambda} c(u) = \sum i(\lambda_i' - \lambda_i) = \sum \binom{\lambda_i}{2} - \sum \binom{\lambda_i'}{2},$$

where $\lambda' = (\lambda_1', \lambda_2', \ldots)$ denotes the conjugate partition to λ. If we regard n as fixed, then equation (12) shows the rate at which $E_{U_q}(\mathrm{is}(w))$ converges to $E_U(\mathrm{is}(w))$ as $q \to \infty$. It is therefore of interest to investigate the asymptotic behavior of the coefficient

$$F_1(n) = \frac{1}{n!} \sum_{\lambda \vdash n} \lambda_1 \left(f^\lambda\right)^2 \left(\sum_{u\in\lambda} c(u)\right)$$

as $n \to \infty$ (as well as the coefficient of q^{-i} for $i > 1$). Numerical evidence suggests that $F_1(n)/n$ is a slowly increasing function of n. The largest value of n for which we have done the computation is $n = 47$, giving $F_1(47)/47 \approx 0.6991$. Eric Rains suggests that there is some reason to suspect that $F(n)/n$ grows like $n^{1/6}$.

Acknowledgment. I am grateful to Jinho Baik, François Bergeron, Ken Brown, Persi Diaconis, Alain Lascoux, Andrei Okounkov, Craig Tracy, and an anonymous referee for providing helpful suggestions and additional references.

References

[1] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.*, to appear.

[2] D. Bayer and P. Diaconis, Trailing the dovetail shuffle to its lair, *Ann. Applied Probability* **2** (1992), 294–313.

[3] P. Bidigare, P. Hanlon, and D. Rockmore, A combinatorial description of the spectrum of the Tsetlin library and its generalization to hyperplane arrangements, *Duke Math. J.* **99** (1999), 135–174.

[4] Ph. Biane, seminar talk at U.C. Berkeley, March 2000.

[5] A. Björner and M. Wachs, q-hook length formulas for forests, *J. Combinatorial Theory (A)* **52** (1989), 165–187.

[6] A. Borodin and A. Okounkov, A Fredholm determinant formula for Toeplitz determinants, *Integral Equations and Operator Theory* **37** (2000), 386–396, math.CA/9907165.

[7] A. Borodin and G. Olshanski, Z-measures on partitions, Robinson-Schensted-Knuth correspondence, and $\beta = 2$ random matrix ensembles, preprint, math.CO/9905189.

[8] A. Borodin, A. Okounkov, and G. Olshanski, On asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc.* **13** (2000), 481–515, math.CO/9905032.

[9] K. Brown and P. Diaconis, Random walks and hyperplane arrangements, *Ann. Probability* **26** (1998), 1813–1854.

[10] P. Diaconis, J. Fill, and J. Pitman, Analysis of top to random shuffles, *Combinatorics, Probability, and Computing* **1** (1992), 135–155.

[11] P. Diaconis, M. McGrath, and J. Pitman, Riffle shuffles, cycles, and descents, *Combinatorica* **15** (1995), 11–29.

[12] G. Duchamp, F. Hivert, and J.-Y. Thibon, Une généralisation des fonctions quasi-symétriques et des fonctions symétriques non commutatives, *C. R. Acad. Sci. Paris Sér. I Math.* **328** (1999), 1113–1116.

[13] G. Duchamp, A. Klyachko, D. Krob, and J.-Y. Thibon, Noncommutative symmetric functions III: Deformations of Cauchy and convolution algebras, *Discrete Math. Theor. Comput. Sci.* **1** (1997), 159–216.

[14] J. Fulman, The combinatorics of biased riffle shuffles, *Combinatorica* **18** (1998), 173–184.

[15] J. Fulman, Applications of symmetric functions to cycle and subsequence structure after shuffles, preprint, math.CO/0102176.

[16] A. Garsia and C. Reutenauer, A decomposition of Solomon's descent algebra, *Advances in Math.* **77** (1989), 189–262.

[17] I. M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, *Advances in Math.* **112** (1995), 218–348.

[18] I. Gessel and C. Reutenauer, Counting permutations with given cycle structure and descent set, *J. Combinatorial Theory (A)* **64** (1993), 189–215.

[19] C. Greene, An extension of Schensted's theorem, *Advances in Math.* **14** (1974), 254–265.

[20] A. R. Its, C. A. Tracy, and H. Widom, Random words, Toeplitz determinants and integrable systems. I, preprint, math.CO/9909169.

[21] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure, preprint, math.CO/9906120.

[22] D. E. Knuth, *The Art of Computer Programming*, vol. 3, *Sorting and Searching*, second ed., Addison-Wesley, Reading, Massachusetts, 1998.

[23] D. Krob, B. Leclerc, and J.-Y. Thibon, Noncommutative symmetric functions II: Tranformations of alphabets, *Internat. J. Algebra Comput.* **7** (1997), 181–264.

[24] G. Kuperberg, Random words, quantum statistics, central limits, random matrices, preprint, math.PR/9909104.

[25] B. F. Logan and L. A. Shepp, A variational problem for random Young tableaux, *Advances in Math.* **26** (1977), 206–222.

[26] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, second ed., Oxford University Press, Oxford, 1995.

[27] A. Okounkov, Infinite wedge and random partitions, preprint, math.RT/9907127.

[28] A. Okounkov, Random matrices and random permutations, *Internat. Math. Res. Notices* **2000**, 1043–1095, math.CO/9903176.

[29] A. Okounkov, SL(2) and z-measures, preprint, math.RT/0002135.

[30] C. Reutenauer, *Free Lie Algebras*, Oxford University Press, Oxford, 1993.

[31] C. E. Schensted, Longest increasing and decreasing subsequences, *Canad. Math. J.* **13** (1961), 179–191.

[32] R. Stanley, Ordered structures and partitions, *Mem. Amer. Math. Soc.*, no. 119 (1972).

[33] R. Stanley, *Enumerative Combinatorics*, vol. 1, second printing, Cambridge University Press, New York/Cambridge, 1996.

[34] R. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, New York/Cambridge, 1999.

[35] A. M. Vershik and S. V. Kerov, Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux, *Soviet Math. Dokl.* **18** (1977), 527–531; translated from *Dokl. Akad. Nauk SSSR* **233** (1977), 1024–1027.