A CONJECTURED COMBINATORIAL FORMULA FOR THE HILBERT SERIES FOR DIAGONAL HARMONICS

J. HAGLUND AND N. LOEHR

Abstract. We introduce a conjectured way of expressing the Hilbert Series of diagonal harmonics as a weighted sum over parking functions. Our conjecture is based on a pair of statistics for the q,t-Catalan sequence discovered by M. Haiman and proven by the first author and A. Garsia. We show how our q,t-parking function formula for the Hilbert Series can be expressed more compactly as a sum over permutations. We also derive two equivalent forms of our conjecture, one of which is based on the original pair of statistics for the q,t-Catalan introduced by the first author and the other of which is expressed in terms of rooted, labelled trees.

Résumé. Nous présentons une façon conjecturelle d'exprimer les séries de Hilbert des harmoniques diagonales en tant que sommes pondérées de fonctions de stationnement. Notre conjecture repose sur une paire de statistiques (découverte par M. Haiman et ensuite prouvée par le premier auteur et A. Garsia) associé e à la séquence de q,t-Catalan. Nous montrons comment notre formule de q,t-stationnement pour la série de Hilbert peut s'exprimer de façon compacte en utilisant une somme sur les permutations. Nous dérivons aussi deux formes équivalentes de notre conjecture. La première, introduite par le premier auteur, est basée sur la paire originale de statistiques associée à la séquence de q,t-Catalan, alors que la seconde s'exprime en termes d'arbres enracinés et étiquetés.

1. Background

In the early 1990's Garsia and Haiman introduced the following conjecture [7].

Conjecture 1.1. *For each positive integer n define a rational function $C_n(q,t)$ by*

$$C_n(q,t) = \sum_{\mu\vdash n} \frac{t^{2\Sigma l} q^{2\Sigma a}(1-t)(1-q)(\sum q^{a'}t^{l'})\prod^{0,0}(1-q^{a'}t^{l'})}{\prod(q^a-t^{l+1})(t^l-q^{a+1})}, \tag{1}$$

where the outer sum is over all partitions μ of n, the products and sums in the inner summand are over the squares of the Ferrers diagram of μ, and the arm a, coarm a', leg l, and coleg l' of a square are as in Fig. 1. The $0,0$ above the product indicates we leave out the upper-left corner square of the diagram of μ. Then $C_n(q,t)$ is a polynomial in q and t with nonnegative integer coefficients.

Conjecture 1.1 is a special case of a more general conjecture in [7], that the Frobenius Series $\mathcal{F}_n(q,t)$ of the space H_n of "Diagonal Harmonics"(an S_n-module first introduced by Haiman in [9]) could be written as

$$\mathcal{F}_n(q,t) = \sum_{\mu\vdash n} \frac{\tilde{H}_\mu[X;q,t]t^{\Sigma l} q^{\Sigma a}(1-t)(1-q)(\sum q^{a'}t^{l'})\prod^{0,0}(1-q^{a'}t^{l'})}{\prod(q^a-t^{l+1})(t^l-q^{a+1})}. \tag{2}$$

Here $\tilde{H}_\mu[X;q,t] = \sum_\lambda \tilde{K}_{\lambda,\mu}(q,t)s_\lambda$ is the "modified" Macdonald polynomial, with $s_\lambda[X]$ the Schur function, $\tilde{K}_{\lambda,\mu}(q,t) = t^{\Sigma l'}K_{\lambda,\mu}(q,1/t)$, and $K_{\lambda,\mu}(q,t)$ is Macdonald's q,t-Kostka

1991 *Mathematics Subject Classification.* Primary 05A15, 05A30 Secondary 05E05.

Key words and phrases. Catalan Number, Parking Functions, Labelled Trees, Hilbert Series.

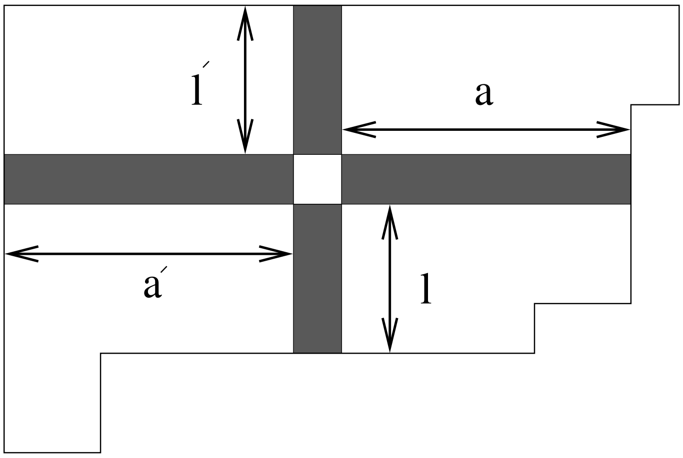


FIGURE 1. The arm, coarm, leg and coleg of a square

number, defined in [14]. They also showed that the right-hand-side of (2) equals $\nabla e_n[X]$, where ∇ is a linear operator defined on the modified Macdonald basis $\tilde{H}_\mu[X;q,t]$ by

$$\nabla \tilde{H}_\mu[X;q,t] = t^{\Sigma l} q^{\Sigma a} \tilde{H}_\mu[X;q,t],$$

and $e_n[X]$ is the nth elementary symmetric function. A special case of this conjecture is that the rational function expression for $C_n(q,t)$, which can be obtained by taking the coefficient of $s_{1^n}[X]$ in (2), equals the component of $\mathcal{F}_n(q,t)$ corresponding to the sign character χ^{1^n}, and hence has nonnegative coefficients.

A Dyck path is a lattice path from $(0,0)$ to (n,n) that never goes below the main diagonal $(i,i), 0 \le i \le n$. It is well-known that the number of such paths is the nth Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. We call the number of squares below the path but above the main diagonal the *area* of the Dyck path. Garsia and Haiman proved that

$$q^{\binom{n}{2}} C_n(q, 1/q) = \frac{1}{[n+1]} \begin{bmatrix} 2n \\ n \end{bmatrix}$$

where $[k] = (1-q^k)/(1-q)$ and $\begin{bmatrix} n \\ k \end{bmatrix}$ is the q-binomial coefficient $(q;q)_n/((q;q)_k(q;q)_{n-k})$. They also showed that

$$C_n(q,1) = \sum_{D \in \mathcal{D}_n} q^{\mathrm{area}(D)},$$

where the sum is over all Dyck paths from $(0,0)$ to (n,n). Based in part on these results, they called $C_n(q,t)$ the *q,t-Catalan sequence* [7].

In conjunction with Conjecture 1.1, they posed the question of finding a pair of statistics (qstat, tstat) so that $C_n(q,t)$ could be written in the form

$$\sum_{D \in \mathcal{D}_n} q^{\mathrm{qstat}(D)} t^{\mathrm{tstat}(D)}.$$

This problem was solved by the first author [8], who introduced a new statistic we here call *dmaj*. To define it, start by placing a ball at the upper right-hand corner (n,n) of a Dyck path D, then push the ball straight left. Once the ball intersects a vertical step of the path, it "ricochets" straight down until it intersects the diagonal, after which the process is iterated; the ball goes left until it hits another vertical step of the path, then down to the

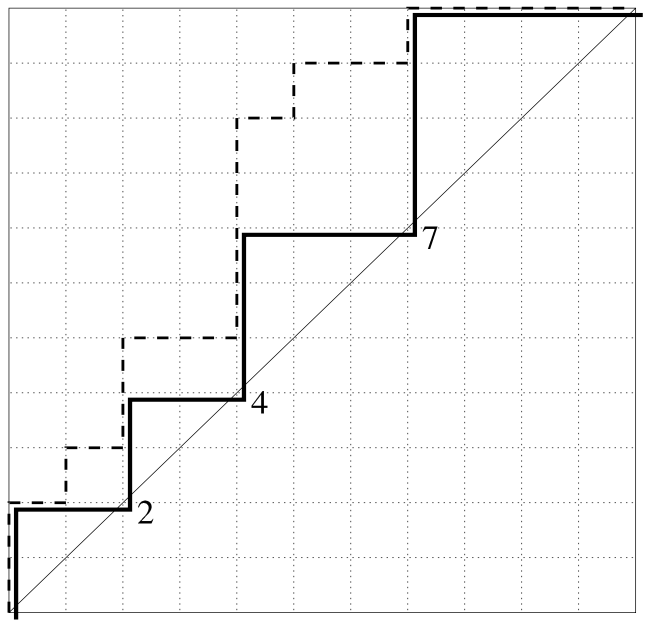


FIGURE 2. The statistic dmaj for a path. The Dyck path is the dashed line, the solid line is the bouncing ball. Here dmaj $= 2+4+7 = 13$ and area $= 22$.

diagonal, etc. On the way from (n, n) to $(0, 0)$ the ball will strike the diagonal at various points (i_j, i_j). We define $\mathrm{dmaj}(D)$ to be the sum of these i_j, and set

$$F_n(q,t) = \sum_{D\in\mathcal{D}_n} q^{\mathrm{area}(D)} t^{\mathrm{dmaj}(D)}.$$

See Fig. 1.

In [8], the first author conjectured that $F_n(q,t) = C_n(q,t)$, and also introduced a stratified function $F_{n,s}(q,t)$, defined as the sum, over all Dyck paths which end in exactly s horizontal steps, of $q^{\mathrm{area}(D)}t^{\mathrm{dmaj}(D)}$. He showed this function satisfies the recurrence relation

$$F_{n,s}(q,t) = \sum_{r=0}^{n-s} \begin{bmatrix} r+s-1 \\ r \end{bmatrix} q^{\binom{s}{2}} t^{n-s} F_{n-s,r}(q,t),$$

and by iterating this recurrence obtained the explicit formula

$$(3) \qquad F_n(q,t) = \sum_{k=1}^{n} \sum_{\substack{\alpha_1+\ldots+\alpha_k \\ \alpha_i>0}} t^{(k-1)\alpha_1+(k-2)\alpha_2+\ldots+\alpha_{k-1}} q^{\binom{\alpha_1}{2}+\ldots+\binom{\alpha_k}{2}} \begin{bmatrix} \alpha_1+\alpha_2-1 \\ \alpha_1 \end{bmatrix} \cdots \begin{bmatrix} \alpha_{k-1}+\alpha_k-1 \\ \alpha_{k-1} \end{bmatrix}.$$

Garsia and the first author found a conjectured expression for $F_{n,s}(q,t)$ in terms of the nabla operator, namely

$$(4) \qquad F_{n,s}(q,t) = t^{n-s} q^{\binom{s}{2}} \nabla e_{n-s}\left[X\frac{1-q^s}{1-q}\right].$$

They were then able to prove (4) by using extended versions of summation formulas for generalized Pieri coefficients and other plethystic identities that Garsia and a number of coauthors have developed over the last ten years [5, 6]. As a corollary they proved Haglund's conjecture that $F_n(q,t) = C_n(q,t)$.

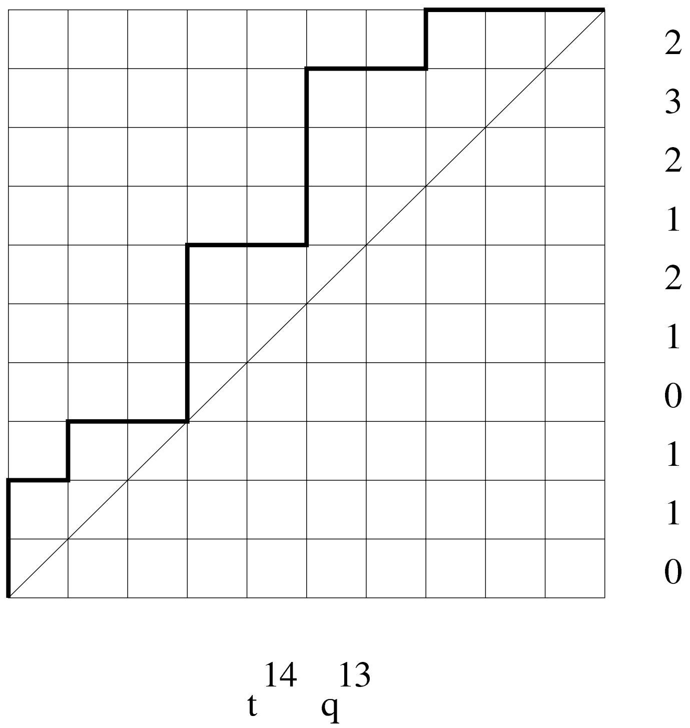


FIGURE 3. A Dyck path, with row lengths on the right

There is another pair of statistics for $C_n(q,t)$, discovered by M. Haiman while Garsia and the first author were still trying to prove (4). Given a Dyck path D, let $a_i(D)$ be the number of squares in the ith row, from the top, that are below D and strictly above the main diagonal. Note that the sum of the $a_i(D)$ equals area(D), that D ends in $a_1(D)+1$ horizontal steps, and that $a_n(D)=0$ for all D. Call the sequence $a_1(D), a_2(D), \ldots a_n(D)$ the *area sequence* of D. We then define a statistic dinv(D) to be the sum of the cardinalities of the two sets

$$\{(i,j) : i < j \text{ and } a_i(D) = a_j(D)\}$$

and

$$\{(i,j) : i < j \text{ and } a_i(D) + 1 = a_j(D)\}.$$

For example, for the path of Fig. 1, dinv $= 14$.

Haiman conjectured that

$$C_n(q,t) = \sum_{D \in \mathcal{D}_n} q^{\mathrm{area}(D)} t^{\mathrm{dinv}(D)}. \tag{5}$$

The discovery of (5) was evidently motivated by elements of his celebrated proof of the "$n!$" conjecture [11], which says that certain S_n-submodules $\mathbf{M}_\mu$ of H_n have dimension $n!$. Since he had previously shown [10] that this conjecture implied that $\tilde{H}_\mu[X;q,t]$ is the Frobenius Series for M_μ, this resolves Macdonald's famous conjecture that the $K_{\lambda,\mu}(q,t)$ are polynomials in q,t with nonnegative integer coefficients. More recently Haiman has also proven (2), the explicit formula for the Frobenius Series of H_n [12].

At first it seemed that (5) was quite different than Haglund's conjecture, but before long Garsia, Haiman, and the first author realized they are closely related. To see why, note that a sequence B of n nonnegative integers $b_1 \cdots b_n$ is the area sequence of a Dyck path if and only if $b_n = 0$ and B contains no "two descents", i.e. values of i, $1 \le i \le n-1$, with

$b_i - b_{i+1} \geq 2$. Given a multiset $A = \{0^{\alpha_k} 1^{\alpha_{k-1}} \cdots (k-1)^{\alpha_1}\}$ of α_k copies of 0, etc., consider the sum of $q^{\text{area}(D)} t^{\text{dinv}(D)}$ over all Dyck paths D whose area sequence is some multiset permutation of A. Note that for any such D, $\text{area}(D) = \alpha_1(k-1) + \ldots + \alpha_{k-1}$. Note also that the contribution to dinv coming from $|\{(i,j) : i < j \text{ and } a_i(D) = a_j(D)\}|$ will equal $\binom{\alpha_1}{2} + \ldots + \binom{\alpha_k}{2}$ for all these D.

We can construct these area sequences by first permuting the α_k 0's and α_{k-1} 1's in any fashion, with a 0 at the end, which can be done in $\binom{\alpha_{k-1}+\alpha_k-1}{\alpha_{k-1}}$ ways. When we take into account the contribution to dinv from these various permutations, we generate the term

$$\begin{bmatrix} \alpha_{k-1} + \alpha_k - 1 \\ \alpha_{k-1} \end{bmatrix}_t .$$

Next we insert the α_{k-2} 2's into the sequence, which cannot be placed in front of any of the 0's since we must avoid 2-descents. This generates the factor

$$\begin{bmatrix} \alpha_{k-2} + \alpha_{k-1} - 1 \\ \alpha_{k-2} \end{bmatrix}_t .$$

It is now clear from (3) that

$$F_n(t,q) = \sum_{D \in \mathcal{D}_n} q^{\text{area}(D)} t^{\text{dinv}(D)}.$$

It follows easily from (1) that $C_n(q,t) = C_n(t,q)$, since the terms for conjugate partitions interchange q and t. Thus Garsia and the first author's result that $F_n(q,t) = C_n(q,t)$ also implies (5).

2. Statistics for the Hilbert Series

Haiman's proof of (2) [12] implies an earlier conjecture of his, that the space H_n has dimension $(n+1)^{n-1}$. It also implies that an explicit expression, as a sum of rational functions, for the Hilbert Series of H_n can be obtained by substituting, for each partition λ, f^λ, the number of standard tableaux of shape λ, in for $s_\lambda[X]$ in the right-hand-side of (2).

The number $(n+1)^{n-1}$ also counts the number of *parking functions* on n cars. A parking function P is obtained by starting with a Dyck path $D(P)$ and placing n "cars", denoted by the integers 1 through n, in the squares immediately to the right of the vertical segments of D, with the restriction that if car i is placed immediately on top of car j, then $i < j$. An example of a parking function is given in Fig. 2.

In this section we introduce a conjectured combinatorial formula for the Hilbert Series, which involves pairing the area statistic with a natural extension of the statistic dinv to parking functions. Given a parking function P, we define $r_i(P)$ to be the number (car) in the ith row (from the top) of P. We then let $\text{dinv}(P)$ be the sum of the cardinalities of the two sets

$$\{(i,j) : i < j, r_i(P) > r_j(P) \text{ and } a_i(D) = a_j(D)\}$$

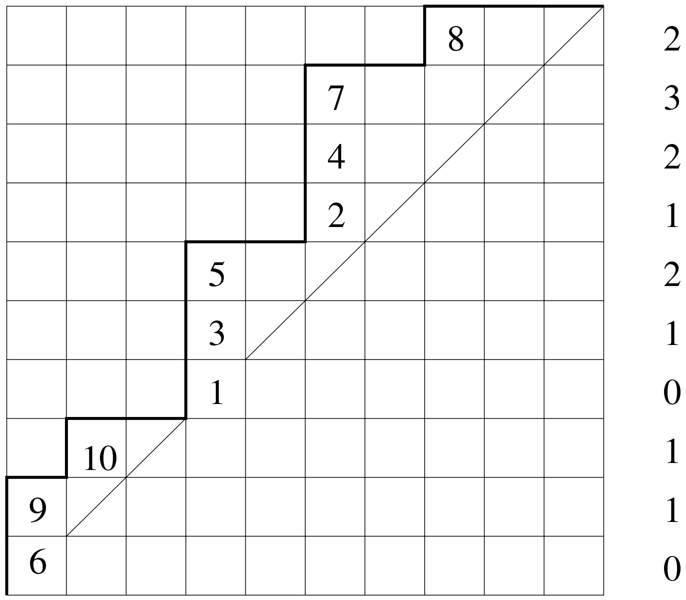
and

$$\{(i,j) : i < j, r_i(P) < r_j(P) \text{ and } a_i(D) + 1 = a_j(D)\}.$$

For the parking function of Fig. 2, dinv $= 6$, since "inversions" occur for pairs (i,j) of rows $(1,3), (1,5), (4,5), (7,8), (7,9),$ and $(8,9)$.

Conjecture 2.1. *Define*

$$R_n(q,t) = \sum_P q^{\text{area}(D(P))} t^{\text{dinv}(P)}, \tag{6}$$



$t^{6}\ q^{13}$

FIGURE 4. A parking function, with row lengths on the right

where the sum is over all parking functions on n cars. Then $R_n(q,t) = \mathcal{H}_n(q,t)$, where $\mathcal{H}_n(q,t)$ denotes the Hilbert Series of H_n.

Conjecture 2.1 has been verified using Maple by the first author and A. Garsia for $n \leq 7$.

We have a more compact way of writing $R_n(q,t)$ as a sum over permutations, which is based on the following lemma.

Lemma 2.2. *Given sets $A = \{a_1, a_2, \ldots, a_s\}$, $B = \{b_1, \ldots, b_{n-s}\}$ with $a_1 < a_2 < \cdots < a_s$, $b_1 < b_2 < \cdots < b_{n-s}$, $A \cap B = \emptyset$ and $A \cup B = \{1, 2, \ldots, n\}$, define*

$$F(A,B) = \sum_P t^{\mathrm{dinv}(P)},$$

where the sum is over all parking functions P whose set of cars on the main diagonal (rows of length 0) consist of the elements of A, in any order, and whose cars on the diagonal just above the main diagonal (rows of length 1) consist of the elements of B, in any order. Then

$$F(A,B) = [s]!_t [b_{n-s} - (n-s)]_t [b_{n-s-1} - (n-s) + 2]_t \cdots [b_1 + n - s - 2]_t.$$

Proof. By reading the cars in a parking function starting with the car in the top row and moving down, a parking function considered in the sum above can be identified with a permutation (linear list) of elements of $\{a_1, \ldots, a_s, b_1, \ldots, b_{n-s}\}$ where b_j immediately precedes a_k implies $b_j > a_k$. We will construct such a sequence recursively by first placing the a's in any order, say $a_{\alpha_1} a_{\alpha_2} \cdots a_{\alpha_s}$. Now b_{n-s} can be placed into this sequence in any of $b_{n-s} - 1 - (n-s-1)$ places, since there are $b_{n-s} - 1$ numbers less than b_{n-s}, but $n-s-1$ are in $\{b_1, \ldots, b_{n-s-1}\}$ and thus can't be in $a_{\alpha_1} \cdots a_{\alpha_s}$. If we insert b_{n-s} in front of the leftmost a_{α_j} satisfying $a_{\alpha_j} < b_{n-s}$, then any a_i's to the left of b_{n-s} will be greater than b_{n-s}, and will not generate any inversions. If we insert b_{n-s} in front of the next-to-leftmost a_{α_j} satisfying $a_{\alpha_j} < b_{n-s}$, then there is one a_i to the left of b_{n-s} less than b_{n-s} and we get a contribution of t. Hence the various possible placements of b_{n-s} generate a factor of $[b_{n-s} - (n-s)]_t$, independent of the permutation $\alpha_1 \cdots \alpha_n$. Next we insert b_{n-s-1} in front of b_{n-s} or any a_j satisfying $a_j < b_{n-s-1}$. If we place b_{n-s-1} in front of the leftmost of these possibilities we do not generate any inversions. If we place b_{n-s-1} in front of the next-to-leftmost choice we will either have b_{n-s} or a_{α_j} (with $a_{\alpha_j} < b_{n-s-1}$) to the left of b_{n-s-1}, and in either case this pair contributes t. Thus we see the insertion of b_{n-s-1} will generate a factor of $[b_{n-s-1} - (n-s-1)+1]_t$ and by induction all the b's together will generate $[b_{n-s}-(n-s)]_t[b_{n-s-1}-(n-s)+2]_t \cdots [b_1+n-s-2]_t$. Since this was independent of $\alpha_1, \ldots, \alpha_s$, when we sum over all permutations of the a's, counting inversions amongst the a's only, we get the remaining $[s]!_t$ factor. □

Using Lemma 2.2we can derive a product formula for the sum of t^{dinv} over all parking functions whose cars on the ith diagonal are from the set A_i, $i = 0, \ldots, k$ for general k. Say for example $k = 2$ so we are considering parking functions with rows of length 0, 1, and 2 with cars from disjoint subsets A, B, and C respectively, of cardinalities u, v, w, with $A \cup B \cup C = \{1, \ldots, n\}$. We start with a permutation of the elements of A, then insert the largest of the elements of B. Since the number of inversions depends only on the relative order of the elements of A and B, we will get a factor of $[\tilde{b}_v - v]_t$, where $\tilde{b}_v$ is what b_v would become if we reduced the elements of A and B to the set $\{1, 2, \ldots u+v\}$, keeping the relative order of each element to the others intact. Then when inserting c_w into the sequence after inserting all the elements of B, we would get the factor $[\tilde{c}_w - w]_t$, where $\tilde{c}_w$ is what c_w would become if we reduced the elements of B and C to the set $\{1, 2, \ldots v+w\}$, keeping the relative order of each element to the others intact. Thus we end up with the final term of

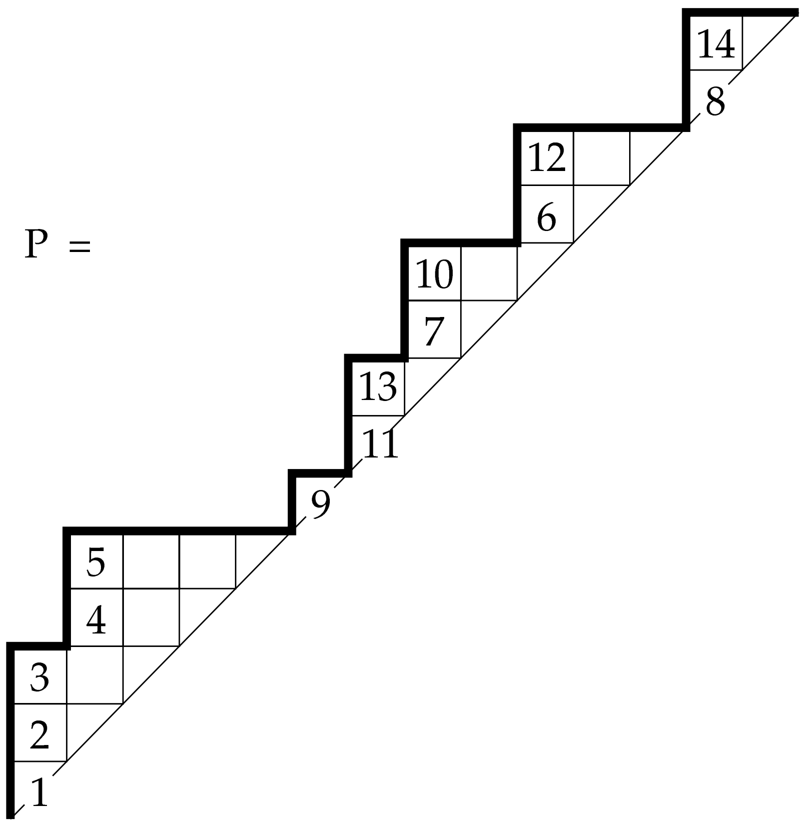
$$[u]!_t[\tilde{b}_v - v] \cdots [\tilde{b}_1 + v - 2]_t[\tilde{c}_w - w]_t \cdots [\tilde{c}_1 + w - 2]_t. \tag{7}$$

For which set partitions $A, B, C, \ldots$ of $\{1, 2, \ldots, n\}$ is there at least one parking function with cars on the main diagonal from A, the next diagonal from B, and so on? A necessary condition is that the largest element of B be larger than the smallest element of A, the largest element of C be larger than the smallest element of B, and so on, since some b_i must be on top of some a_i, some c_i must be on top of some b_i, etc. This condition is also sufficient, since we could put the a_i in columns 1 through $|A|$, with the smallest of these (a_1) in column $|A|$, then put the b_i in columns $|A|$ through $|A \cup B| - 1$, with the largest b_i on top of a_1 in column $|A|$, and b_1 in column $|A \cup B| - 1$, etc. Call such a sequence of sets "valid". Assume for the moment $k = 2$, so we have sets A, B, and C of cardinalities u, v, w, respectively. In the permutation $\sigma = c_1c_2 \cdots c_wb_1 \cdots b_va_1 \cdots a_u$, there are descents at spots w (since $c_w > b_1$) and $w + v$. This argument shows that there is a bijection between valid sequences of $k+1$ sets and permutations of $\{1, \ldots, n\}$ with k descents. Note that the area of all the corresponding parking functions in our example is $2w+v$, which is also the major index of the permutation σ. It is easy to see this holds in general. Furthermore, M. Haiman has pointed out that the numbers $\tilde{c}_i$ somewhat awkwardly described above can be easily defined in terms of the elements of σ. To do so, define a sequence σ' by $\sigma'_i = \sigma_i$, $1 \le i \le n$, and $\sigma'_{n+1} = 0$. Then for each $i, 1 \le i \le n$, let $u_i(\sigma)$ be the length of the longest consecutive sequence $\sigma'_i\sigma'_{i+1} \cdots \sigma'_j$ that starts at σ'_i and has either no descents, or exactly one descent and $\sigma'_i > \sigma'_j$. For example, if $\sigma = 47358126 \in S_8$, then $(u_1, u_2, \ldots, u_8) = (3, 3, 5, 4, 4, 4, 3, 2)$. It is easy to check that $\tilde{c}_1 = u_1 - 1$, $\tilde{c}_2 = u_2 - 1$, etc. We finally arrive at the following.

Theorem 2.3.

$$R_n(q,t) = \sum_{\sigma \in S_n} q^{\mathrm{maj}(\sigma)} \prod_{i=1}^{n} [u_i(\sigma) - 1]_t.$$

P =



i	a_i	r_i
1	1	14
2	0	8
3	2	12
4	1	6
5	2	10
6	1	7
7	1	13
8	0	11
9	0	9
10	3	5
11	2	4
12	2	3
13	1	2
14	0	1

area(P) = 16 dinv(P)=19 dinv(D(P)) = 41

FIGURE 5. A labelled Dyck path (version 1).

3. Parking Functions and the dmaj Statistic

The relation

$$\sum_{D\in\mathcal{D}_n} q^{area(D)}t^{dinv(D)} = \sum_{D\in\mathcal{D}_n} q^{dmaj(D)}t^{area(D)}$$

suggests there should also be a way of extending the $(\mathrm{area}, \mathrm{dmaj})$ pair of statistics to get an alternate form of $R_n(q,t)$. We discuss one such extension in this section.

Let $\mathcal{P}_n$ denote the set of parking functions on n cars, as defined previously. We now view $\mathcal{P}_n$ as a collection of *labelled* Dyck paths. Fig. 5 shows a typical element of $\mathcal{P}_{14}$.

It is convenient to introduce another set of labelled Dyck paths, which we call $\mathcal{Q}_n$. To construct a typical object $Q \in \mathcal{Q}_n$, we attach labels to a path $D \in \mathcal{D}_n$ according to the following rules. Let $q_1q_2\cdots q_n$ be a permutation of the labels $\{1, 2, \ldots, n\}$. Place each label q_i in the ith row of the diagram for D, in the *main diagonal cell.* There is one restriction: For each "left-turn" in the Dyck path (i.e., an EAST step followed immediately by a NORTH step, reading from southwest to northeast), the label q_i appearing due east of the NORTH step must be less than the label q_j appearing due south of the EAST step. See Fig. 3 for an example. In the figure, capital letters mark the left-turns in the Dyck path. Since $4 < 5$,

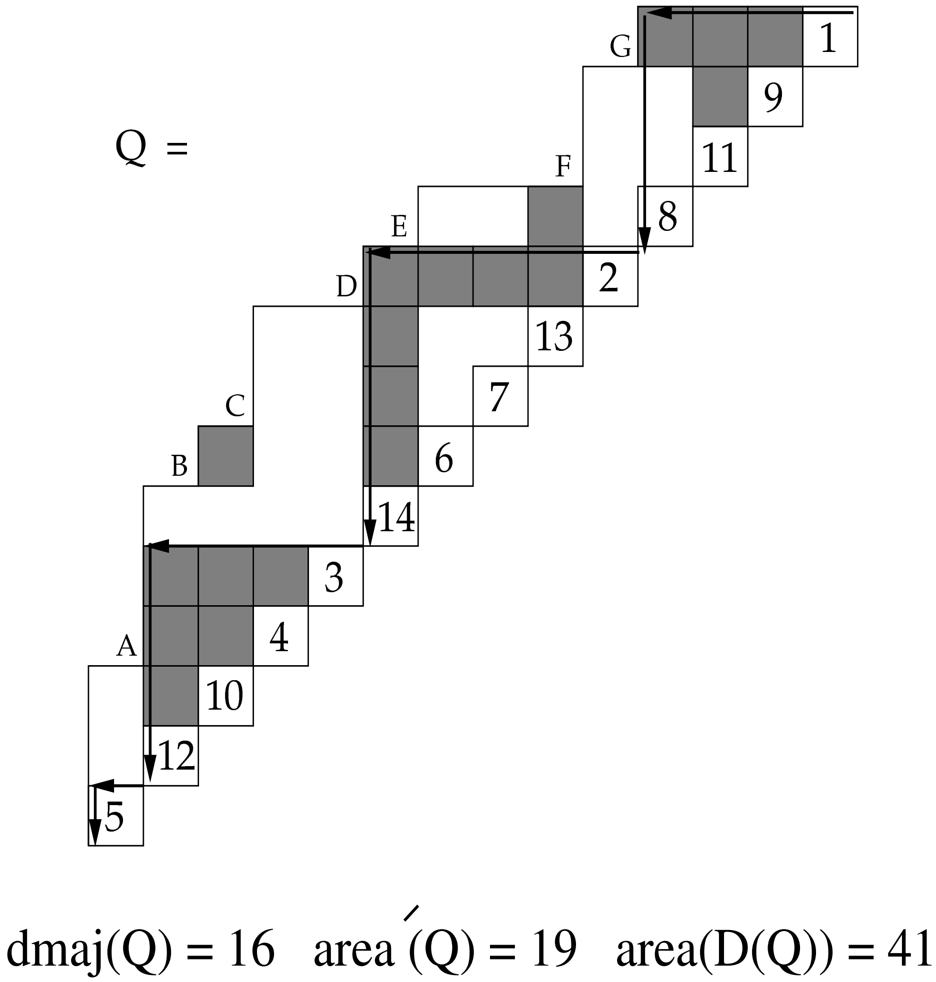


FIGURE 6. A labelled Dyck path (version 2).

$6 < 12$, $7 < 10$, $2 < 3$, $8 < 14$, $11 < 13$, and $1 < 2$, the labelled path shown does belong to $\mathcal{Q}_{14}$.

Given a labelled path Q constructed from the ordinary Dyck path $D = D(Q)$, define $\mathrm{dmaj}(Q)$ to be $\mathrm{dmaj}(D(Q))$, which was defined earlier. Also define $\mathrm{area}'(Q)$ to be the number of cells c in the diagram for Q such that:

(1) Cell c is strictly between the Dyck path D and the main diagonal; AND
(2) The label on the main diagonal due east of c is less than the label on the main diagonal due south of c.

In Fig. 3, only the shaded cells satisfy both conditions and hence contribute to $\mathrm{area}'(Q)$. Evidently, $\mathrm{area}'(Q) \leq \mathrm{area}(D(Q))$ for all Q, and strict inequality can occur.

We conjecture that

$$S_n(q,t) = \sum_{Q \in \mathcal{Q}_n} q^{dmaj(Q)} t^{area'(Q)} \tag{8}$$

also gives the Hilbert series for H_n. We will show this conjecture is equivalent to the previous one by giving a bijective proof that $R_n(q,t) = S_n(q,t)$.

Bijections. We begin with the case of unlabelled Dyck paths. Fix a path $D \in \mathcal{D}_n$. We will construct a new path $E \in \mathcal{D}_n$ such that $\mathrm{area}(D) = \mathrm{dmaj}(E)$ and $\mathrm{dinv}(D) = \mathrm{area}(E)$, which proves that

$$\sum_{D \in \mathcal{D}_n} q^{area(D)} t^{dinv(D)} = \sum_{E \in \mathcal{D}_n} q^{dmaj(E)} t^{area(E)}.$$

The bijection is essentially a combinatorial version of the proof of this formula given in Section 1.

Consider the area sequence $a(D) = (a_1(D), \ldots, a_n(D))$. It is easy to see that such a list of numbers corresponds to a valid Dyck path iff $a_i \geq 0$ for all i, $a_n = 0$, and $a_i \leq a_{i+1} + 1$ for all $i < n$. Set $s = \max_{1 \leq i \leq n} a_i$. For $0 \leq j \leq s$, let b_j be the number of occurrences of j in $a(D)$. It follows from the above conditions on $a(D)$ that $b_j > 0$ for all j; moreover, $b_0 + \cdots + b_s = n$.

To construct E, we first draw a bounce path B whose successive horizontal moves (starting from (n, n)) have lengths $b_0, \ldots, b_s$. This bounce path, together with the main diagonal line $y = x$, creates a sequence of $s+1$ triangles which we shall call $T_0, \ldots, T_s$. For $1 \leq i \leq s$, there is an empty rectangular region R_i located north of triangle T_i and west of triangle T_{i-1}. Note that rectangle R_i has width b_i and height b_{i-1}.

We now describe how to construct the portion of the path E located in rectangle R_i. Fix i, and let w_i be the word obtained from $a(D)$ by deleting all symbols other than $i-1$ and i. Then w_i consists of b_{i-1} occurrences of $i-1$ and b_i occurrences of i; also, by the conditions on $a(D)$, the last symbol in w_i must be $i-1$. Read the symbols in w_i from left to right. Starting at the northwest tip of triangle T_i, draw an EAST step when the symbol i is read; draw a NORTH step when the symbol $i-1$ is read. Note that this partial path must terminate in a NORTH step. For later use, we remark that the "left-turns" of E in the region R_i correspond precisely to the *descents* in the word w_i. Because of the condition $a_i \leq a_{i+1} + 1$, the set of descents in all the words w_i corresponds bijectively with the set of descents in the full word $a(D)$.

After filling all the rectangular regions in this way, we obtain the Dyck path E. Observe that, because the paths within each R_i ended in NORTH steps, B is the bounce path derived from E. Therefore,

$$\begin{aligned}
\mathrm{dmaj}(E) &= (n - b_0) + (n - b_0 - b_1) + \cdots + (n - b_0 - b_1 - \cdots - b_s) \\
&= n(s+1) - (s+1)b_0 - sb_1 - \cdots - (s+1-j)b_j - \cdots - 1b_s \\
&= (s+1)(n - b_0 - \cdots - b_s) + \sum_{j=0}^{s} jb_j \\
&= \sum_{j=0}^{s} jb_j = \sum_{i=1}^{n} a_i \\
&= \mathrm{area}(D).
\end{aligned}$$

Furthermore, from the definitions of b_j and w_i, it is easy to see that the formula for dinv can be rewritten as

$$\mathrm{dinv}(D) = \sum_{j=0}^{s} \binom{b_j}{2} + \sum_{i=1}^{s} \mathrm{coinv}(w_i),$$

where $\mathrm{coinv}(w_i)$ is the number of coinversions in the word w_i. Now $\binom{b_j}{2}$ is the number of area cells in the triangle T_j, and $\mathrm{coinv}(w_i)$ is the number of cells beneath the path E in the rectangle R_i. Hence, $\mathrm{dinv}(D) = \mathrm{area}(E)$.

The process used to construct E from D is reversible. First, we obtain $b_0, \ldots, b_s$ by examining the bounce path of E. Next, we recover $a(D)$ from E by starting with b_0 zeroes and successively inserting b_1 ones, then b_2 twos, etc., according to the partial paths in R_1, R_2, etc. The condition that $a_i \leq a_{i+1}+1$ ensures that there will be a unique way to perform this insertion procedure. Hence, we obtain the desired bijection.

As an example, if we take D to be the path shown in Fig. 5 (ignoring labels), then E will be the path shown in Fig. 3 (ignoring labels).

Next, we consider the case of labelled Dyck paths. We will give a bijection from $\mathcal{P}_n$ to $\mathcal{Q}_n$ that sends area to dmaj and sends dinv to area$'$. This bijection proves that

$$R_n(q,t) = \sum_{P \in \mathcal{P}_n} q^{area(P)} t^{dinv(P)} = \sum_{Q \in \mathcal{Q}_n} q^{dmaj(Q)} t^{area'(Q)} = S_n(q,t).$$

Fix $P \in \mathcal{P}_n$. We shall construct $Q \in \mathcal{Q}_n$ with $\mathrm{dmaj}(Q) = \mathrm{area}(P)$ and $\mathrm{area}'(Q) = \mathrm{dinv}(P)$. As an example, the labelled path P in Fig. 5 will map to the labelled path Q in Fig. 3.

Let $D = D(P)$ denote the underlying unlabelled Dyck path of P. Let E be the unlabelled Dyck path produced by the above bijection, with $\mathrm{dmaj}(E) = \mathrm{area}(D)$ and $\mathrm{area}(E) = \mathrm{dinv}(D)$. E will be the underlying unlabelled path for Q (i.e., $D(Q) = E$).

We obtain Q by attaching labels to E, as follows. Scan each of the diagonals of P, from southwest to northeast, starting with the main diagonal and proceeding upward. Enter the labels of P, in the order in which they are encountered, on the main diagonal of Q going from northeast to southwest. For instance, in Fig. 5, P has the labels 1, 9, 11, 8 on the main diagonal, followed by the labels 2, 13, 7, 6, 14 on the first superdiagonal, etc. Hence (see Fig. 3), the labels on the main diagonal of Q are 1, 9, 11, 8, 2, 13, 7, 6, 14, ... starting from (n,n). Clearly, we can recover the labelling of P from the labelling of Q.

Here is an equivalent way of describing the relation between the labels in P and Q. Recall that $E = D(Q)$ can be dissected into triangles $T_0, \ldots, T_s$ and rectangles $R_1, \ldots, R_s$. For $0 \leq j \leq s$, the b_j labels on the main diagonal of Q inside triangle T_j (read from top to bottom) are the labels appearing in the leftmost cells of the b_j rows of $D = D(P)$ for which $a_i(D) = j$ (read from bottom to top).

Recall that the labels of P in a given column must increase from bottom to top. To check the validity of a given labelling, it clearly suffices to check that *adjacent* labels in the same column are always properly ordered. Suppose that the labels r_i and r_{i+1} in rows i and $i+1$ both occur in column j. This occurs iff $a_i = a_{i+1}+1$ iff there is a *descent* of $a(D)$ at position i (recall that $a_i \leq a_{i+1}+1$). We observed earlier that the descents of $a(D)$ correspond bijectively to the left-turns of $E = D(Q)$. From here, it is easy to verify that label r_{i+1} appears in Q due east of the left-turn corresponding to the descent $a_{i+1} > a_i$, and the label r_i appears in Q due south of this left-turn. Hence, the labelling restrictions on P imply the corresponding labelling restrictions on Q, and conversely.

Clearly, $\mathrm{dmaj}(Q) = \mathrm{dmaj}(E) = \mathrm{area}(D) = \mathrm{area}(P)$. We now show that $\mathrm{area}'(Q) = \mathrm{dinv}(P)$. Consider a typical area cell c of the path $E = D(Q)$. Suppose first that c is inside triangle T_k. Let $x_1, x_2, \ldots, x_{a_k}$ be the labels on the diagonal of Q inside T_k, from top to bottom. As noted above, the labels $x_1, x_2, \ldots, x_{a_k}$ are just the numbers r_i in all positions i for which $a_i = k$, *written in reverse order.* Thus, the cells in T_k that contribute to $\mathrm{area}'(Q)$ correspond precisely to inversions in the word $x_{a_k}, \ldots, x_2, x_1$. We obtain a bijection between the contributing cells in T_k and the set

$$\{(i,j) : i < j, r_i(P) > r_j(P) \text{ and } a_i(D) = a_j(D) = k\} \quad (0 \leq k \leq s).$$

A similar argument applies to a cell c in rectangle R_k. The horizontal position of the cell determines a unique r_j such that $a_j = k$, and the vertical position of the cell determines a unique r_i such that $a_i = k-1$. All pairs (i,j) for which $a_i + 1 = a_j$ are accounted

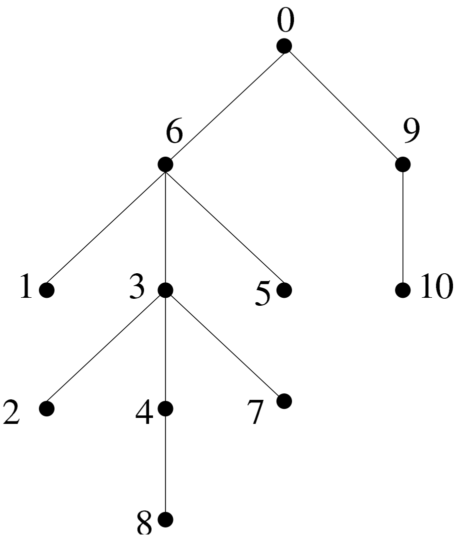


FIGURE 7. The tree for the parking function of Figure 2

for exactly once in this fashion (as k ranges from 1 to s). Now, we get a contribution to $\mathrm{dinv}(P)$ iff $i < j$ and $r_i < r_j$; this occurs precisely when the associated cell c satisfies the two conditions for contributing to $\mathrm{area}'(Q)$. We conclude that the number of contributing cells in all the rectangular regions R_k is exactly the cardinality of the set

$$\{(i, j) : i < j, r_i(P) < r_j(P) \text{ and } a_i(D) + 1 = a_j(D)\}.$$

Combining this result with the one in the preceding paragraph, we conclude that $\mathrm{area}'(Q) = \mathrm{dinv}(P)$. This completes the proof.

4. Labelled Trees

There are a number of known bijections between parking functions on n cars and forests of labelled trees on n vertices [2, 4, 13]; see also [16, p.140]. These typically use an alternate definition of a parking function, as a function $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ with the property that the number of values of j with $f(j) \leq i$ is at least i for all $1 \leq i \leq n$. Such a function can be obtained from the geometric representation of a parking function from section 2 by letting $f(i)$ be the column containing the car i. It isn't easy to translate Conjecture 2.1 into a statement about trees using these bijections since it is hard to keep track of what happens to the area and dinv statistics. In this section we describe a simple bijection between forests of rooted labelled trees, and parking functions in the geometric form of section 2 which makes it easy to describe versions of these statistics for trees.

Given a parking function P as in section 2 and a given car i, travel northeast, staying in the same diagonal, until we either leave the $n \times n$ square or run into another car. If we run into a car at the bottom of a column, then we say all the cars in that column are "children" of car i. If we leave the square or run into a car which isn't at the bottom of a column, car i has no children. We define a rooted, labelled tree $T(P)$ with root labelled 0 by the condition that the node with label i (call this node i) is a child of the root node if and only if car i is in the first column of P, and for a non-root node node i, i has j as a child if and only if car j is a child of car i in P. Also, when we draw a tree T we put the children of node i below node i, increasing from left to right. For example, if P is the parking function from Fig. 2, $T(P)$ is represented in Fig. 4.

We recursively define the *preference order* of such a tree as the sequence of n numbers, beginning with the children of the root node, listed left to right (smallest to largest) followed by the preference order of the descendants of the largest child of the root node, followed by the preference order of the descendants of the next-to-largest child of the root node, etc. We also define a function $d_i(T)$ recursively by $d_0(T) = 0$, and if j is the kth-smallest child of i, then $d_j(T) = d_i(T) + k - 1$. For example, the tree of Fig. 4 has preference order $6, 9, 10, 1, 3, 5, 2, 4, 7, 8$, and the d values of these (non-root) nodes are $0, 1, 1, 0, 1, 2, 1, 2, 3, 2$, respectively. Given a tree T, we construct a Dyck path $D(T)$ by the condition that the length of the k row, from the bottom, of $D(T)$ is the d value of the kth element of the preference order of T. We then construct a parking function $P(T)$ by placing the kth number from the preference sequence of T in the kth row, from the bottom, immediately to the right of $D(T)$. Next define $\mathrm{area}(T)$ to be the sum of the d-values of the nodes of T, and $\mathrm{dinv}(T)$ to be the sum of the cardinalities of the two sets

$$\{(i, j) : i < j, d(i) = d(j) \text{ and } i \text{ occurs before } j \text{ in preference order}\}$$

and

$$\{(i, j) : i < j, d(i) + 1 = d(j), \text{ and } i \text{ occurs after } j \text{ in preference order}\}.$$

It is not hard to see that $P(T(P)) = P$, $T(P(T)) = T$ and furthermore that $\mathrm{area}(T(P)) = \mathrm{area}(P)$ and $\mathrm{dinv}(T(P)) = \mathrm{dinv}(P)$. Thus Conjecture 2 is equivalent to the following.

Conjecture 4.1.

$$\mathcal{H}_n(q, t) = \sum_T q^{\mathrm{area}(T)} t^{\mathrm{dinv}(T)}$$

where the sum is over all labelled, rooted trees T with $n + 1$ vertices and root node labelled 0.

5. Open Problems

The main obstacle to proving Conjecture 2.1 by the methods of [5, 6] is the lack of a recurrence relation for $R_n(q, t)$. We have also been unable to resolve a number of interesting bijective questions related to Conjecture 2.1, most notably to prove $R_n(q, t)$ is symmetric in q and t by finding an involution on parking functions which interchanges area and dinv. In fact, we can't even prove the much weaker statement that $R_n(q, 1) = R_n(1, q)$. We have also been unable to show $q^{\binom{n}{2}} R_n(q, 1/q) = (1 + q + \ldots + q^n)^{n-1}$, which is the value for $q^{\binom{n}{2}} \mathcal{H}_n(q, 1/q)$ conjectured by Stanley [9]. One could also hope to refine Conjecture 2.1 to find a pair of statistics on parking functions that generate $\mathcal{F}_n(q, t)$.

The statistics and bijections discussed in section 3 are specializations of more general constructs that apply to variations of the q, t-Catalan numbers and parking functions. In particular, the second author has discovered combinatorial results concerning the "extended family" of q, t-Catlalan sequences $C_n^m(q, t)$ introduced by Garsia and Haiman [7]. These generalizations are discussed in the second author's thesis [to appear].

References

[1] F. Bergeron, A. Garsia, M. Haiman, and G. Tesler, Identities and positivity conjectures for some remarkable operators in the theory of symmetric functions, *Asian J. Math.* **6** (1999), 363–420.
[2] J.S. Beissinger and U.N. Peled, A note on major sequences and external activity in trees, The Wilf Festschrift (Philadelphia, PA 1996) *Electron. J. Combin.* **4** (1997), Research Paper 4, approx. 10 pp.
[3] Ö. Eğecioğlu and J. Remmel, Bijections for Cayley trees, spanning trees, and their q-analogues, *J. Combin. Theory Ser. A* **42** (1986), 15–30.
[4] D. Foata and J. Riordan, Mappings of acyclic and parking functions, *Aequationes Math.* **10** (1974), 10–22.

[5] A. Garsia and J. Haglund, A positivity result in the theory of Macdonald polynomials, *Proc. Nat. Acad. Sci. U.S.A* **98** (2001), 4313–4316.
[6] A. Garsia and J. Haglund, A proof of the (q, t)-Catalan positivity conjectur , *Discrete Math.*, to appear (the LACIM 2000 issue).
[7] A. Garsia and M. Haiman, A remarkable q, t-Catalan sequence and q-Lagrange inversion, *J. Algebraic Combin.* **5** (1996), 191–244.
[8] J. Haglund, Conjectured statistics for the (q, t)-Catalan numbers, *Adv. in Math.* (To appear).
[9] M. Haiman, Conjectures on the quotient ring by diagonal invariants, *J. Algebraic Combin.* **3** (1994), 17–76.
[10] M. Haiman, Macdonald polynomials and geometry, in *New Perspectives in Algebraic Combinatorics* 37, Billers, et. al. (eds.), MSRI Publication, 1999, 207–254.
[11] M. Haiman, Hilbert Schemes, polygraphs, and the Macdonald positivity conjecture, *J. Amer. Math. Soc.* **14** (2001), 941–1006
[12] M. Haiman, Vanishing theorems and character formulas for the Hilbert scheme of points in the plane, (preprint, 2001), available from `http://www.math.berkeley.edu/ mhaiman/`.
[13] G. Krewerau, Une famille de polynômes ayant plusieurs proprietes enumeratives, **11** (1980), 309–320.
[14] I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, The Clarendon Press, Oxford University Press, New York, 1995.
[15] J. Riordan, Ballots and trees, *C. Combinatorial Theory* **6** (1969), 408–411.
[16] R.P. Stanley, *Enumerative Combinatorics Volume 2* Cambridge University Press, Cambridge, United Kingdom, 1999.

Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104-6395
E-mail address: `jhaglund@math.upenn.edu`

Department of Mathematics, University of California at San Diego, La Jolla, CA 92093-0112
E-mail address: `nloehr@math.ucsd.edu`