# Conjectured Combinatorial Models for the Hilbert Series of Generalized Diagonal Harmonics Modules

Nicholas A. Loehr*

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104
nloehr@math.upenn.edu

Jeffrey B. Remmel

Department of Mathematics
University of California at San Diego
La Jolla, CA 92093
jremmel@math.ucsd.edu

Submitted: Jul 31, 2003; Accepted: Sep 5, 2004; Published: Sep 24, 2004
Mathematics Subject Classifications: 05A10, 05E05, 05E10, 20C30, 11B65

**Abstract**

Haglund and Loehr previously conjectured two equivalent combinatorial formulas for the Hilbert series of the Garsia-Haiman diagonal harmonics modules. These formulas involve weighted sums of labelled Dyck paths (or parking functions) relative to suitable statistics. This article introduces a third combinatorial formula that is shown to be equivalent to the first two. We show that the four statistics on labelled Dyck paths appearing in these formulas all have the same univariate distribution, which settles an earlier question of Haglund and Loehr. We then introduce analogous statistics on other collections of labelled lattice paths contained in trapezoids. We obtain a fermionic formula for the generating function for these statistics. We give bijective proofs of the equivalence of several forms of this generating function. These bijections imply that all the new statistics have the same univariate distribution. Using these new statistics, we conjecture combinatorial formulas for the Hilbert series of certain generalizations of the diagonal harmonics modules.

## 1 Introduction

A *Dyck path* of order $n$ is a path in the $xy$-plane from $(0,0)$ to $(n,n)$ consisting of $n$ vertical steps and $n$ horizontal steps, each of length one, such that no step goes strictly

*Supported by a National Science Foundation Graduate Research Fellowship

below the diagonal line $y = x$. A *labelled Dyck path* is a Dyck path whose vertical steps are labelled $1, 2, \ldots, n$ in such a way that the labels for vertical steps in a given column increase reading upwards. These labelled paths can be used to encode *parking functions* [17, 5, 6, 23], which are functions $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ such that $|f^{-1}(\{1, 2, \ldots, i\})| \geq i$ for $1 \leq i \leq n$.

In [11], J. Haglund and the first author introduced two pairs of statistics on labelled Dyck paths that give a conjectured combinatorial interpretation of the Hilbert series of the diagonal harmonics module studied by Garsia and Haiman [9]. This article introduces a third pair of statistics on labelled Dyck paths that has the same generating function as those considered in [11]. As a corollary, we obtain a simple bijective proof that all the statistics being discussed have the same univariate distribution. This result settles one of the open questions from [11].

We shall also define analogous pairs of statistics on other collections of labelled lattice paths corresponding to generalized parking functions [24, 25]. We study the combinatorial properties of these statistics, obtaining an explicit summation formula for their generating function and giving bijective proofs of the equivalence of different pairs of statistics. As before, these bijections imply that all the new statistics have the same univariate distribution.

To motivate our combinatorial study of labelled lattice paths, this introductory section will review the previous work of F. Bergeron, A. Garsia, J. Haglund, M. Haiman, G. Tesler, et al. regarding the diagonal harmonics module and its connections to representation theory, symmetric functions, Macdonald polynomials, and parking functions. This section also discusses the generalizations of the diagonal harmonics module, which were studied by the same authors. We conjecture that the new statistics introduced here for labelled lattice paths inside triangles give the Hilbert series for these generalized modules. Readers interested only in the combinatorics may safely skip much of this section, reading only §1.4, §1.5, and §1.7.

## 1.1 Notation

We assume the reader is acquainted with basic facts about partitions, symmetric functions, and representation theory, which can be found in standard references such as [22] or [21]. This section sets up the notation we will use when discussing these topics.

**Definition 1.** Let $\lambda = (\lambda_1 \geq \cdots \geq \lambda_k)$ be an integer partition. If $\lambda_1 + \cdots + \lambda_k = N$, we write $|\lambda| = N$ or $\lambda \vdash N$. We identify $\lambda$ with its *Ferrers diagram.* Figure 1 shows the Ferrers diagram of $\lambda = (8, 7, 5, 4, 4, 2, 1)$, which is a partition of 31 having seven parts. The *transpose* $\lambda'$ of $\lambda$ is the partition obtained by interchanging the rows and columns of the Ferrers diagram of $\lambda$. For example, the transpose of the partition in Figure 1 is

$$\lambda' = (7, 6, 5, 5, 3, 2, 2, 1).$$

**Definition 2.** Let $\lambda$ be a partition of $N$. Let $c$ be one of the $N$ cells in the diagram of $\lambda$.

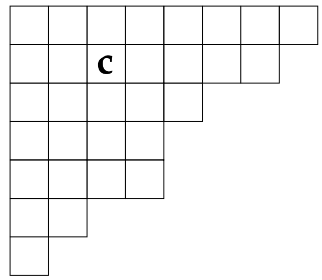


Figure 1: Diagram of a partition.

(1) The *arm of* $c$, denoted $a(c)$, is the number of cells strictly right of $c$ in the diagram of $\lambda$.

(2) The *coarm of* $c$, denoted $a'(c)$, is the number of cells strictly left of $c$ in the diagram of $\lambda$.

(3) The *leg of* $c$, denoted $l(c)$, is the number of cells strictly below $c$ in the diagram of $\lambda$.

(4) The *coleg of* $c$, denoted $l'(c)$, is the number of cells strictly above $c$ in the diagram of $\lambda$.

For example, the cell labelled $c$ in Figure 1 has $a(c) = 4$, $a'(c) = 2$, $l(c) = 3$, and $l'(c) = 1$.

**Definition 3.** We define the *dominance partial ordering* on partitions of $N$ as follows. If $\lambda$ and $\mu$ are partitions of $N$, we write $\lambda \geq \mu$ to mean that

$$\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i \text{ for all } i \geq 1.$$

**Definition 4.** Fix a positive integer $N$ and a partition $\mu$ of $N$. We introduce the following abbreviations to shorten upcoming formulas:

$$\begin{aligned}
h_\mu(q,t) &= \prod_{c\in\mu}(q^{a(c)} - t^{l(c)+1}) \\
h'_\mu(q,t) &= \prod_{c\in\mu}(t^{l(c)} - q^{a(c)+1}) \\
n(\mu) &= \sum_{c\in\mu} l(c) \\
n(\mu') &= \sum_{c\in\mu'} l(c) = \sum_{c\in\mu} a(c)
\end{aligned}$$

$$\begin{aligned}
B_\mu(q,t) &= \sum_{c\in\mu} q^{a'(c)}t^{l'(c)} \\
\Pi_\mu(q,t) &= \prod_{c\in\mu, c\neq(0,0)} (1 - q^{a'(c)}t^{l'(c)})
\end{aligned}$$

In all but the last formula above, the sums and products range over all cells in the diagram of $\mu$. In the product defining $\Pi_\mu(q,t)$, the northwest corner cell of $\mu$ is omitted from the product. This is the cell $c$ with $a'(c) = l'(c) = 0$; if we did not omit this cell, then $\Pi_\mu(q,t)$ would be zero.

**Definition 5.** Let $K = \mathbb{Q}(q,t)$ denote the field of rational functions in the variables $q$ and $t$ with rational coefficients. Let $\Lambda = \Lambda(K)$ denote the ring of symmetric functions in countably many indeterminates $x_n$ with coefficients in $K$. Let $\Lambda^N$ denote the ring of homogeneous symmetric functions of degree $N$ (together with zero). We let $m_\lambda$, $e_\lambda$, $h_\lambda$, $p_\lambda$, and $s_\lambda$ respectively denote the monomial symmetric function, elementary symmetric function, complete homogeneous symmetric function, power sum symmetric function, and Schur function indexed by the partition $\lambda$. Detailed definitions of these concepts appear in [21].

It is well known that the collections

$$\{m_\lambda : \lambda \vdash N\}, \{e_\lambda : \lambda \vdash N\}, \{h_\lambda : \lambda \vdash N\}, \{p_\lambda : \lambda \vdash N\}, \{s_\lambda : \lambda \vdash N\}$$

each constitute a $K$-basis for $\Lambda^N$. Moreover, $\{e_n : n \geq 1\}$ is an algebraically independent set, as is $\{h_n : n \geq 1\}$ and $\{p_n : n \geq 1\}$.

In particular, given any $K$-algebra $A$ and any function $\phi_0 : \{p_1, p_2, \ldots\} \to A$, there exists a unique $K$-algebra homomorphism $\phi : \Lambda(K) \to A$ extending $\phi_0$. When $\phi_0$ is the function sending each $p_k$ to $(1-q^k)p_k$, some authors denote $\phi(f)$ (for $f \in \Lambda$) by using the plethystic notation $f[X(1-q)]$.

**Definition 6.** For each $N$, introduce a scalar product on $\Lambda^N$ by requiring that

$$\langle s_\lambda, s_\mu \rangle = \chi(\lambda = \mu).$$

Here and below, for a logical statement $A$ we write $\chi(A) = 1$ if $A$ is true, $\chi(A) = 0$ if $A$ is false. If $f \in \Lambda^N$, the *coefficient of $s_\lambda$ in $f$* is

$$f|_{s_\lambda} = \langle s_\lambda, f \rangle.$$

**Definition 7.** Let $S_n$ denote the symmetric group on $n$ letters. Let $\mathbb{C}[S_n]$ denote the group algebra of $S_n$. Given a complex vector space $V$, a *representation* of $S_n$ on $V$ is a group homomorphism $A : S_n \to GL(V)$ from $S_n$ to the group of invertible linear transformations of $V$. The *character* of this representation is the function $\chi_A : S_n \to \mathbb{C}$ such that $\chi_A(\sigma) = trace(A(\sigma))$. Given a representation $A$, we can regard $V$ as an $S_n$-*module.* An $S_n$*-submodule of* $V$ is an $A$-invariant vector subspace $W$ of $V$. In symbols, $A(\sigma)(w) \in W$ for all $\sigma \in S_n$ and all $w \in W$. A nonzero space $V$ is called an *irreducible* $S_n$-module iff its only submodules are 0 and $V$ itself.

We recall the following well-known results from representation theory (see [22] for more details):

(1) Every $S_n$-module $V$ can be decomposed into a direct sum of irreducible $S_n$-modules.

(2) The isomorphism classes of irreducible $S_n$-modules correspond in a natural way to the partitions $\lambda \vdash n$. Thus, we may label these irreducible modules $M_\lambda$.

(3) An $S_n$-module $V$ is determined (up to isomorphism) by its character $\chi_V$.

(4) For any $S_n$-module $V$, the character $\chi_V$ belongs to the *center* of the group algebra $\mathbb{C}[S_n]$.

(5) The characters $\chi_\lambda \stackrel{def}{=} \chi_{M_\lambda}$ are a vector-space basis for the center of the group algebra.

(6) The center of the group algebra of $S_n$ is isomorphic to the ring $\Lambda(\mathbb{C})^n$ of homogeneous symmetric functions of degree $n$ under an isomorphism sending $\chi_\lambda$ to $s_\lambda$. This isomorphism is called the **Frobenius map**.

**Definition 8.** Let $V$ be an $S_n$-module. We can decompose $V$ into a direct sum of irreducible submodules, say

$$V = \bigoplus_{\lambda \vdash n} c_\lambda M_\lambda \text{ (where } c_\lambda \in \mathbb{N}) .$$

The **Frobenius characteristic of** $V$ is defined by

$$\mathcal{F}_V = \sum_{\lambda \vdash n} c_\lambda s_\lambda \in \Lambda^n.$$

Thus, $\mathcal{F}_V$ is a homogeneous symmetric function of degree $n$, and the coefficient of $s_\lambda$ in this function is just the multiplicity of the irreducible module $M_\lambda$ in $V$.

A similar procedure is possible for graded $S_n$-modules and doubly graded $S_n$-modules, which we now define.

**Definition 9.** Fix $n \geq 1$.

(1) An $S_n$-module $V$ is called a **graded** $S_n$-module if there is a direct sum decomposition

$$V = \bigoplus_{h \geq 0} V_h,$$

where each $V_h$ is an $S_n$-submodule of $V$.

(2) Let $V = \oplus_h V_h$ be a graded $S_n$-module. Decompose each $V_h$ into irreducible submodules, say $V_h = \oplus_{\lambda \vdash n} c_h(\lambda) M_\lambda$. The **Frobenius series of** $V$ is

$$\mathcal{F}_V(q) = \sum_{h \geq 0} \left( \sum_{\lambda \vdash n} c_h(\lambda) s_\lambda \right) q^h = \sum_{h \geq 0} \mathcal{F}_{V_h} q^h.$$

(3) Let $V = \oplus_h V_h$ be a graded $S_n$-module. The **Hilbert series of** $V$ is

$$H_V(q) = \sum_{h \geq 0} dim_{\mathbb{C}}(V_h) q^h.$$

(4) An $S_n$-module $V$ is called a **doubly graded** $S_n$-module if there is a direct sum decomposition

$$V = \bigoplus_{h \geq 0} \bigoplus_{k \geq 0} V_{h,k},$$

where each $V_{h,k}$ is an $S_n$-submodule of $V$.

(5) Let $V = \oplus_{h,k} V_{h,k}$ be a doubly graded $S_n$-module. Decompose each $V_{h,k}$ into irreducible submodules, say $V_{h,k} = \oplus_{\lambda \vdash n} c_{h,k}(\lambda) M_\lambda$. The **Frobenius series of** $V$ is

$$\mathcal{F}_V(q,t) = \sum_{h \geq 0} \sum_{k \geq 0} \left( \sum_{\lambda \vdash n} c_{h,k}(\lambda) s_\lambda \right) q^h t^k = \sum_{h \geq 0} \sum_{k \geq 0} \mathcal{F}_{V_{h,k}} q^h t^k.$$

(6) Let $V = \oplus_{h,k} V_{h,k}$ be a doubly graded $S_n$-module. The **Hilbert series of** $V$ is

$$H_V(q,t) = \sum_{h \geq 0} \sum_{k \geq 0} dim_{\mathbb{C}}(V_{h,k}) q^h t^k.$$

Given a doubly graded $S_n$-module $V$, there is a simple way to recover the Hilbert series of $V$ from the Frobenius series of $V$. Specifically, let $f_\lambda$ be the dimension of the irreducible $S_n$-module $M_\lambda$. A well-known theorem [22] states that $f_\lambda$ is the number of standard tableaux of shape $\lambda$, which is $n!$ divided by the product of the hook lengths of $\lambda$. It is immediate from the definitions that

$$H_V(q,t) = [\mathcal{F}_V(q,t)]|_{s_\lambda = f_\lambda},$$

where this notation indicates that we should replace every $s_\lambda$ by the integer $f_\lambda$.

Similarly, we can use the Frobenius series to obtain the generating function for the occurrences of any particular irreducible $S_n$-module inside $V$. For instance, $M_{1^n}$ is the irreducible module that affords the sign character of $S_n$. Thus, to find the generating function for the doubly graded submodule of $V$ that carries the sign representation, we would look at $\mathcal{F}_V(q,t)|_{s_{1^n}}$, the coefficient of $s_{1^n}$ in the Frobenius series.

## 1.2 Modified Macdonald Polynomials and the Nabla Operator

In this section, we define the modified Macdonald polynomials, which form another useful basis for the ring of symmetric functions. We also define the nabla operator, a linear operator on $\Lambda$ that has many important properties. The modified Macdonald polynomials were introduced by Garsia and Haiman [13] by modifying the definition in Macdonald's book [21]. The nabla operator was first introduced by F. Bergeron and Garsia [1]; see also [2, 3].

**Theorem 10.** *Let $\alpha : \Lambda(K) \to \Lambda(K)$ be the $K$-algebra automorphism that interchanges the variables $q$ and $t$. Abusing notation and writing $f \in \Lambda(K)$ as $f(x; q, t)$, we have $\alpha(f(x; q, t)) = f(x; t, q)$. Let $\phi : \Lambda \to \Lambda$ be the unique $K$-algebra homomorphism such that $\phi(p_k) = (1 - q^k)p_k$. There exists a unique basis $\tilde{H}_\mu$ of $\Lambda(K)$, called the* ***modified Macdonald polynomial basis****, with the following properties:*

*(1) $\phi(\tilde{H}_\mu) = \sum_{\lambda \geq \mu} c_{\lambda,\mu} s_\lambda$ for certain scalars $c_{\lambda,\mu} \in K$.*

*(2) $\alpha(\tilde{H}_\mu) = \tilde{H}_{\mu'}$.*

*(3) $\tilde{H}_\mu|_{s_{(n)}} = 1$.*

*(Here, $\mu$ ranges over all partitions, and $\geq$ is the dominance partial order on partitions.) Moreover, $\{\tilde{H}_\mu : \mu \vdash m\}$ is a basis of $\Lambda^m(K)$.*
*Some authors write the three properties in the definition using different notation, as follows:*

*(1) $\tilde{H}_\mu[(1 - q)X; q, t] = \sum_{\lambda \geq \mu} c_{\lambda,\mu}(q, t) s_\lambda(X)$ for certain scalars $c_{\lambda,\mu} \in K$.*

*(2) $\tilde{H}_\mu(X; q, t) = \tilde{H}_{\mu'}(X; t, q)$.*

*(3) $\langle \tilde{H}_\mu(X; q, t), s_{(n)}(X) \rangle = 1$.*

*Proof.* The proof for the original Macdonald polynomials can be found in [21]. For a discussion of the modified version, see e.g. [13]. □

For any $\mu \vdash n$, we can write

$$\tilde{H}_\mu = \sum_{\lambda \vdash n} \tilde{K}_{\lambda,\mu} s_\lambda$$

for unique coefficients $\tilde{K}_{\lambda,\mu} \in \mathbb{Q}(q, t)$. These coefficients are called the *modified Kostka-Macdonald coefficients*. The following theorem of Haiman resolves a long-standing conjecture of Macdonald regarding these coefficients.

**Theorem 11.** *[M. Haiman]*
*For every $\lambda \vdash n$ and $\mu \vdash n$, $\tilde{K}_{\lambda,\mu}$ is a polynomial in $q$ and $t$ with nonnegative integer coefficients.*

*Proof.* See [13]. □

In advance, one only knows that $\tilde{K}_{\lambda,\mu}$ is a *rational function* with *rational* coefficients. Haiman's proof uses sophisticated machinery from algebraic geometry. The proof provides an explicit interpretation for the coefficients of the polynomials $\tilde{K}_{\lambda,\mu}$. These coefficients count the multiplicities of irreducible modules in a certain doubly graded $S_n$-module. In particular, the coefficients must be nonnegative integers.

We now define the nabla operator of F. Bergeron and Garsia. Some of the special properties of this operator are developed in [1, 2, 3].

**Definition 12.** The **nabla operator** $\nabla$ is the unique linear operator on $\Lambda(K)$ that acts on the modified Macdonald basis as follows:

$$\nabla(\tilde{H}_\mu) = q^{n(\mu')}t^{n(\mu)}\tilde{H}_\mu.$$

Equivalently, $\nabla$ is the linear operator on $\Lambda$ with eigenvalues $q^{n(\mu')}t^{n(\mu)}$ and corresponding eigenfunctions $\tilde{H}_\mu$.

The next theorem, due to Garsia and Haiman, gives an explicit formula for $\nabla(e_n) = \nabla(s_{1^n})$ as an expansion in terms of the basis $(\tilde{H}_\mu)$.

**Theorem 13.**

$$\nabla(e_n) = \nabla(s_{1^n}) = \sum_{\mu \vdash n} \frac{\tilde{H}_\mu t^{n(\mu)} q^{n(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)}.$$

*Proof.* See [9], [7], [13]. □

## 1.3 The Diagonal Harmonics Module

The formula in the last theorem has a representation-theoretical interpretation, conjectured by Garsia and Haiman [9] and later proved by Haiman [13, 16]. This interpretation involves the diagonal harmonics modules, which we now define.

Fix a positive integer $n$. Consider the polynomial ring

$$R_n = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$$

in two sets of $n$ independent variables. $R_n$ is clearly an infinite-dimensional vector space over $\mathbb{C}$ with a basis given by the set of all monomials. We make $R_n$ into an $S_n$-module as follows. Given $\sigma \in S_n$, define an action of $\sigma$ on $R_n$ by setting

$$\sigma \cdot f(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}, y_{\sigma(1)}, \ldots, y_{\sigma(n)}).$$

This is called the *diagonal action* of $S_n$ on $R_n$, since $\sigma$ permutes the indices of the $x$-variables and the $y$-variables in the same way.

**Definition 14.** Define the *diagonal harmonics* in $R_n$ by

$$DH_n = \left\{ f \in R_n : \sum_{i=1}^{n} \frac{\partial^h}{\partial x_i^h}\frac{\partial^k}{\partial y_i^k} f = 0 \text{ for } 1 \le h+k \le n \right\}.$$

It is easy to see that $DH_n$ is an $S_n$-submodule of $R_n$. Furthermore, $DH_n$ is a doubly graded module: we can write

$$DH_n = \bigoplus_{h \ge 0} \bigoplus_{k \ge 0} V_{h,k}(n),$$

where $V_{h,k}(n)$ is the submodule of $DH_n$ consisting of zero and those polynomials $f$ that are homogeneous of degree $h$ in the $x$-variables and homogeneous of degree $k$ in the $y$-variables.

We can now form the Frobenius series $\mathcal{F}_{DH_n}(q,t)$, the Hilbert series $H_{DH_n}(q,t)$, and the generating function for the sign character $\mathcal{F}_{DH_n}(q,t)|_{s_{1^n}}$, as discussed earlier. For notational convenience, we will henceforth denote these three generating functions by $F_n(q,t)$, $H_n(q,t)$, and $RC_n(q,t)$, respectively.

To understand the representation theory of diagonal harmonics, we would like to have more explicit formulas for $F_n(q,t)$, $H_n(q,t)$, and $RC_n(q,t)$. As pointed out earlier, it is sufficient to find a formula for the Frobenius series. Garsia and Haiman conjectured such a formula involving the modified Macdonald polynomials [9]. The formula was proved much later by Haiman using advanced machinery from algebraic geometry. Our next theorem gives this formula.

**Theorem 15.**

$$F_n(q,t) = \sum_{\mu \vdash n} \frac{\tilde{H}_\mu t^{n(\mu)} q^{n(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)}.$$

*Proof.* See [13] and [16]. □

Combining this result with Theorem 13, we have

$$F_n(q,t) = \nabla(s_{1^n}).$$

**Definition 16.** Let $\mathcal{D}_n$ denote the collection of Dyck paths of order $n$. For $E \in \mathcal{D}_n$, define $area(E)$ to be the number of complete lattice cells between the path and the diagonal $y = x$. Define $maj(E) = \sum_{(x,y)}(x+y)$, where we sum over all points $(x,y)$ such that the line segments from $(x-1,y)$ to $(x,y)$ and from $(x,y)$ to $(x,y+1)$ both belong to $E$.

The following theorem of Garsia and Haiman can be used to compute the specializations $F_n(q,1)$ and $F_n(q,1/q)$ of the Frobenius series.

**Theorem 17.** *(1) For a Dyck path $D$ of order $n$, define $v_i(D)$ to be the number of vertical steps taken by the path along the line $x = i$. Then*

$$\nabla(e_n)|_{t=1} = \sum_{D \in \mathcal{D}_n} q^{area(D)} \prod_{i=0}^{n-1} e_{v_i(D)},$$

*where $e_j$ denotes an elementary symmetric function, as usual.*

*(2)*

$$q^{n(n-1)/2}\nabla(e_n)\big|_{t=1/q} = \sum_{\mu \vdash n} s_\mu \frac{s_{\mu'}(1,q,q^2,\ldots,q^n)}{[n+1]_q},$$

*where we set $[j]_q = 1 + q + q^2 + \cdots + q^{j-1}$.*

*Proof.* See Theorem 1.2 and Corollary 2.5 in [9]. $\square$

Recall that the Hilbert series of $DH_n$ is given by $H_n(q,t) = F_n(q,t)|_{s_\lambda = f_\lambda}$. Haiman's work also implies the following specializations of the Hilbert series.

**Theorem 18.**

$$H_n(1,1) = (n+1)^{n-1}$$
$$q^{n(n-1)/2} H_n(q, 1/q) = [n+1]_q^{n-1}.$$

*Proof.* See [13] and [16]. $\square$

Note that the first statement just says that $dim(DH_n) = (n+1)^{n-1}$. Even this seemingly simple fact is very difficult to prove.

Next, consider $RC_n(q,t) = F_n(q,t)|_{s_{1^n}}$, the generating function for occurrences of the sign character in $DH_n$. Before Theorem 15 was proved, Garsia and Haiman [9] were able to compute the coefficient of $s_{1^n}$ in the conjectured character formula

$$\sum_{\mu \vdash n} \frac{\tilde{H}_\mu t^{n(\mu)} q^{n(\mu')} (1-t)(1-q) \Pi_\mu(q,t) B_\mu(q,t)}{h_\mu(q,t) h'_\mu(q,t)}.$$

In light of Theorem 13, this coefficient is just $\nabla(s_{1^n})|_{s_{1^n}}$, the entry in the lower-right corner of the matrix representing nabla relative to the Schur basis. This coefficient is the original version of the $q,t$-Catalan number, as defined by Garsia and Haiman in [9].

**Definition 19.** For $n \geq 1$, define the *original $q,t$-Catalan sequence* by

$$OC_n(q,t) = \sum_{\mu \vdash n} \frac{t^{2n(\mu)} q^{2n(\mu')} (1-t)(1-q) \Pi_\mu(q,t) B_\mu(q,t)}{h_\mu(q,t) h'_\mu(q,t)}.$$

**Theorem 20.** *For all $n$,*

$$OC_n(q,t) = \nabla(s_{1^n})|_{s_{1^n}}.$$

*Proof.* See [9]. $\square$

Of course, it is immediate from Haiman's Theorem 15 that $OC_n(q,t) = RC_n(q,t)$. However, since this equality is very difficult to prove, it is useful to maintain separate notation for the two expressions.

Garsia and Haiman also proved the following specializations of $OC_n(q,t)$, which explain why they called it the $q,t$-Catalan sequence.

**Theorem 21.** *For all $n$,*

$$OC_n(1,1) = \frac{1}{n+1}\binom{2n}{n} = C_n$$
$$q^{n(n-1)/2} OC_n(q,1/q) = \frac{1}{[n+1]_q} \begin{bmatrix} 2n \\ n \end{bmatrix}_q = \sum_{D \in \mathcal{D}_n} q^{maj(D)}$$
$$OC_n(1,q) = OC_n(q,1) = \sum_{D \in \mathcal{D}_n} q^{area(D)}.$$

*Proof.* See [9]. □

In light of this last result, it is natural to ask if there is a purely combinatorial interpretation for the bivariate sequence $OC_n(q,t)$. In other words, we would like to have a second statistic on Dyck paths, say *tstat*, such that

$$OC_n(q,t) = \sum_{D \in \mathcal{D}_n} q^{area(D)} t^{tstat(D)}.$$

Two different $t$-statistics were conjectured by Haglund and Haiman [10],[12]. Later, Garsia and Haglund proved that these conjectures really do give $OC_n(q,t)$ [7]. We discuss these statistics in the next subsection.

Similarly, we would like to have combinatorial interpretations for the Hilbert series $H_n(q,t)$ and the Frobenius series $F_n(q,t)$ by introducing suitable pairs of statistics on some collection of objects. Haglund, Haiman, and the present author conjectured such statistics for the Hilbert series (see [11] and §1.5 below). At this time, it is an open problem to prove that these conjectured statistics are correct.

## 1.4 Combinatorial Bivariate Catalan Numbers

In this section, we describe two different combinatorial versions of the bivariate Catalan sequence. These sequences are based on two statistics proposed by Haglund [10] and Haiman [12], respectively.

**Definition 22.** Let $E$ be a Dyck path of order $n$.

(1) Define a *bounce path* derived from $E$ as follows. The bounce path begins at $(n,n)$ and moves to $(0,0)$ via an alternating sequence of horizontal and vertical moves. Starting at $(n,n)$, the bounce path proceeds due west until it reaches the north step of the Dyck path going from height $n-1$ to height $n$. From there, the bounce path goes due south until it reaches the main diagonal line $y=x$. This process continues recursively. When the bounce path has reached the point $(i,i)$ on the main diagonal $(i>0)$, the bounce path goes due west until it is blocked by the north step of the Dyck path going from height $i-1$ to height $i$. From there, the bounce path goes due south until it hits the main diagonal. The bounce path terminates when it reaches $(0,0)$. See Figure 2 for an example.

Suppose the bounce path derived from $E$ hits the main diagonal at the points

$$(n,n),\quad (i_1,i_1),\quad (i_2,i_2),\quad \ldots,\quad (i_s,i_s),\quad (0,0).$$

The *bounce score* for $E$ is defined by

$$b(E) = \sum_{k=1}^{s} i_k.$$

For example, in Figure 2, the bounce path for $E$ hits the main diagonal at $(14,14)$, $(10,10)$, $(5,5)$, $(1,1)$, and $(0,0)$. Thus, $b(E) = 10+5+1 = 16$ for this path.

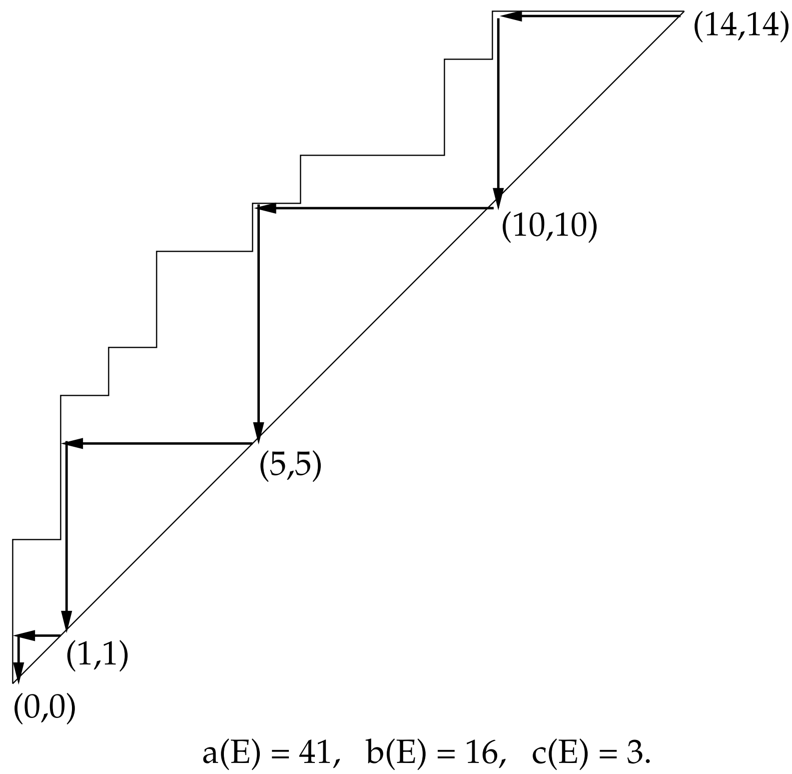



Figure 2: A Dyck path with its derived bounce path.

(2) Define *Haglund's combinatorial Catalan number* to be the bivariate generating function

$$C_n(q,t) = \sum_{P \in \mathcal{D}_n} q^{area(P)} t^{b(P)}.$$

(3) For $0 \le i < n$, define $g_i(E)$ to be the number of complete cells strictly between the path and the main diagonal in the $i$'th row of the picture, where the bottom row is row zero. Define the *area vector* $\vec{g}(E)$ to be the vector $(g_0(E), \ldots, g_{n-1}(E))$. For example, for the path $E$ shown in Figure 3, we have

$$\vec{g}(E) = (0,1,2,2,3,0,0,1,1,2,1,2,0,1).$$

Note that $area(E) = \sum_{i=0}^{n-1} g_i(E)$.

(4) We define Haiman's statistic $h$ by the formula

$$h(E) = \sum_{i<j} \left[\chi(g_i(E) = g_j(E)) + \chi(g_i(E) = g_j(E)+1)\right]. \tag{1}$$

For example, we have $h(E) = 41$ for the path in Figure 3.

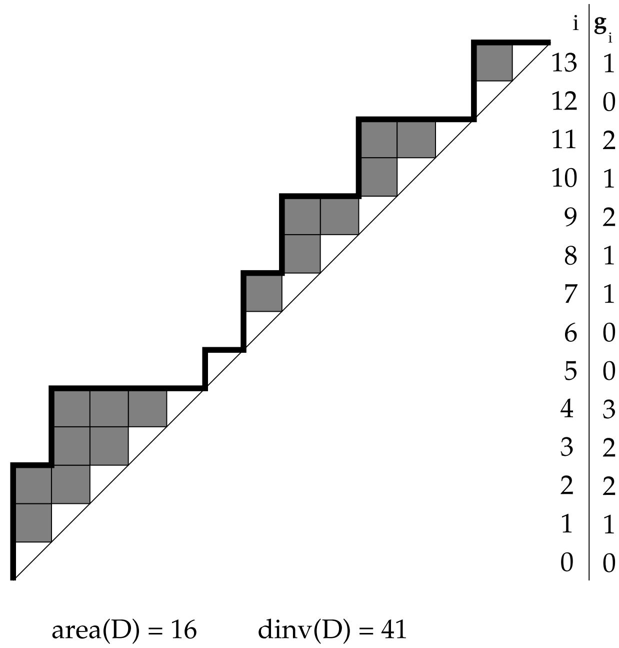



Figure 3: A Dyck path and the associated vector $\vec{g}$.

(5) We define *Haiman's combinatorial* $q, t$*-Catalan sequence* to be

$$HC_n(q,t) = \sum_{D \in \mathcal{D}_n} q^{h(D)} t^{area(D)} \quad (n = 1, 2, 3, \ldots).$$

Note that we use $t$, not $q$, to keep track of area in this sequence.

**Theorem 23.** *For all* $n \geq 1$*,*

$$C_n(q,t) = HC_n(q,t) = OC_n(q,t).$$

*Proof.* See [7, 11]. □

**Remark 24.** A variant of the bounce statistic is obtained by starting the bounce path at $(0,0)$ and bouncing north and east to $(n,n)$. This variant will be generalized in §1.7.

## 1.5 Combinatorial Hilbert Series

In this section, we describe two pairs of statistics on labelled Dyck paths (parking functions) of order $n$ that are conjectured to give the Hilbert series $H_n(q,t)$ of diagonal harmonics. These statistics were proposed by Haglund, Haiman, and the first author [11].

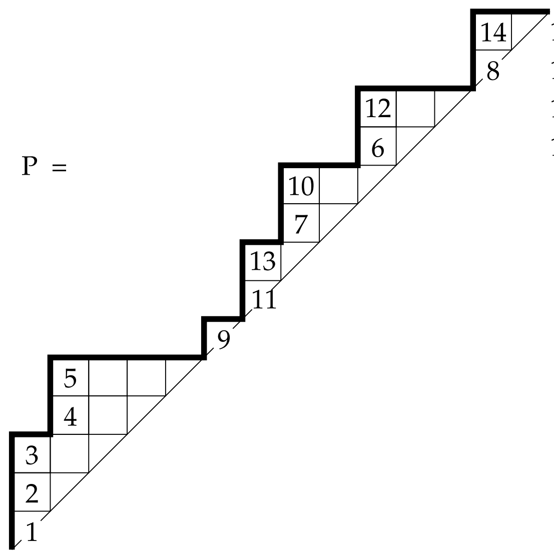
**Definition 25.** (1) Let $\mathcal{P}_n$ denote the set of labelled Dyck paths of order $n$. A typical object $P \in \mathcal{P}_n$ consists of a path $D \in \mathcal{D}_n$ and a labelling of the vertical steps of $D$

such that the labels in each column increase from bottom to top. It is convenient to regard $P$ as a pair of vectors

$$P = (\vec{g} = (g_0, \ldots, g_{n-1}), \vec{p} = (p_0, \ldots, p_{n-1})),$$

where $\vec{g}$ is the area vector for $P$, and $\vec{p}$ is obtained by reading the labels from bottom to top. The condition that labels increase in columns is equivalent to requiring that, for all $i < n-1$, $g_i(D) < g_{i+1}(D)$ implies $p_i < p_{i+1}$. See Figure 4 for an example.

P =




| i | $\gamma_i$ | $p_i$ |
|---|---|---|
| 13 | 1 | 14 |
| 12 | 0 | 8 |
| 11 | 2 | 12 |
| 10 | 1 | 6 |
| 9 | 2 | 10 |
| 8 | 1 | 7 |
| 7 | 1 | 13 |
| 6 | 0 | 11 |
| 5 | 0 | 9 |
| 4 | 3 | 5 |
| 3 | 2 | 4 |
| 2 | 2 | 3 |
| 1 | 1 | 2 |
| 0 | 0 | 1 |

area(P) = 16 dinv(P) = 18 dinv(D(P)) = 41

Figure 4: A labelled Dyck path (version 1).

(2) Given $P = (\vec{g}, \vec{p}) \in \mathcal{P}_n$, define the *area of* $P$ to be $area(P) = \sum_{i=0}^{n-1} g_i$. Also define

$$\begin{aligned} h(P) = \quad \textstyle\sum_{i<j} \quad & [\chi(g_i(P) = g_j(P) \text{ and } p_i < p_j) \\ & +\chi(g_i(P) = g_j(P) + 1) \text{ and } p_i > p_j)] \,. \end{aligned}$$

(3) Define the *first combinatorial Hilbert series* by

$$CH_n(q,t) = \sum_{P \in \mathcal{P}_n} q^{area(P)} t^{dinv(P)}. \tag{2}$$

(4) We now define another collection $\mathcal{Q}_n$ of labelled Dyck paths of order $n$. 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 $q_0 q_1 \cdots q_{n-1}$ be a permutation of the labels $\{1, 2, \ldots, n\}$. Place each label $q_i$ in the $i$'th row of the diagram for $D$, in the *main diagonal cell*. There

is one restriction: for each inner corner in the Dyck path consisting of an east step followed by a north step, 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 Figure 5 for an example. In the figure, capital letters mark the inner corners in the Dyck path. Since $4 < 5$, $6 < 12$, $7 < 10$, $2 < 3$, $8 < 14$, $11 < 13$, and $1 < 2$, the labelled path shown does belong to $\mathcal{Q}_{14}$.

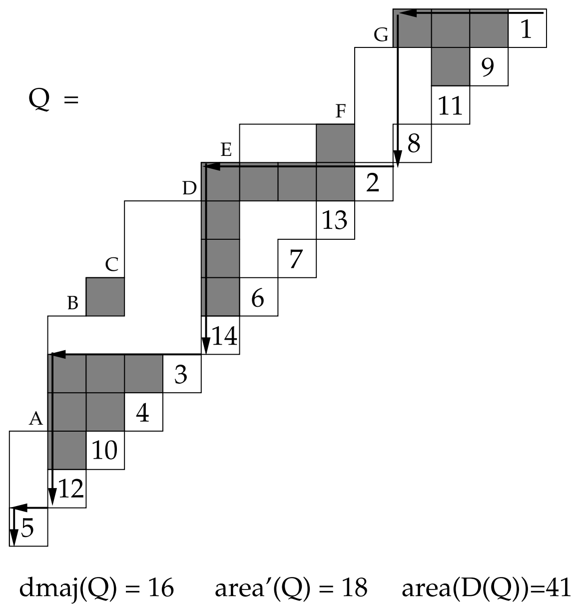



Figure 5: A labelled Dyck path (version 2).

(5) Given a labelled path $Q$ constructed from the ordinary Dyck path $D = D(Q)$, define $dmaj(Q)$ to be $b(D(Q))$, the bounce statistic for $D$ defined earlier. Also define $area'(Q)$ 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 Figure 5, only the shaded cells satisfy both conditions and hence contribute to $area'(Q)$.

(6) Define the *second combinatorial Hilbert series* by

$$CH'_n(q,t) = \sum_{Q \in \mathcal{Q}_n} q^{dmaj(Q)} t^{area'(Q)}. \tag{3}$$

**Theorem 26.** *For all $n \geq 1$,*

$$CH_n(q,t) = CH'_n(q,t).$$

*Proof.* This is proved via an explicit bijection in [11]. □

In §2, we will define a statistic $pmaj$ on $\mathcal{P}_n$ such that the generating function

$$CH''_n(q,t) \stackrel{def}{=} \sum_{P\in\mathcal{P}_n} q^{pmaj(P)}t^{area(P)}$$

is also equal to $CH_n(q,t)$. Using this result and the one just quoted, one obtains bijections that map any pair of statistics

$$(area, dinv),\ (dmaj, area'),\ (pmaj, area)$$

to any other. As a corollary, we obtain bijective proofs that all statistics in question have the same univariate distribution. This resolves one of the open questions from [11].

**Conjecture 27 (Haglund,Haiman,Loehr).** *For all* $n \geq 1$,

$$CH_n(q,t) = H_n(q,t) = \nabla(e_n)|_{s_\lambda = f_\lambda}.$$

This conjecture says that the generating function for statistics on labelled Dyck paths gives the Hilbert series of the diagonal harmonics module.

We now describe an explicit formula for $CH_n(q,t)$ as a summation over permutations $\sigma \in S_n$. First, we need some notation. Given $\sigma = \sigma_1\sigma_2\cdots\sigma_n$, a *descent* of $\sigma$ is an index $i < n$ such that $\sigma_i > \sigma_{i+1}$. Suppose $\sigma$ has descents $i_1, i_2, \ldots, i_s$, where $i_1 < i_2 < \cdots < i_s$. Then we call the lists of elements

$$\sigma_1\sigma_2\cdots\sigma_{i_1};\quad \sigma_{i_1+1}\cdots\sigma_{i_2};\quad \cdots;\quad \sigma_{i_s+1}\cdots\sigma_{i_n}$$

the *ascending runs* of $\sigma$. For example, if $\sigma = 4,7,1,5,8,3,2,6$, then the ascending runs of $\sigma$ are $4,7$ and $1,5,8$ and $3$ and $2,6$. We can display the runs more concisely by writing

$$\sigma = 4,7 > 1,5,8 > 3 > 2,6.$$

For $1 \leq i \leq n$, define a number $w_i(\sigma)$ as follows. Let $R_j$ be the ascending run of $\sigma$ containing $\sigma_i$. Let $R_{j+1}$ be the next ascending run of $\sigma$, if there is one. The number $w_i(\sigma)$ is the number of items in $R_j$ that are larger than $\sigma_i$, plus the number of items in $R_{j+1}$ that are smaller than $\sigma_i$ if $R_{j+1}$ exists, plus one if $R_{j+1}$ does not exist (i.e., if $R_j$ is the last ascending run of $\sigma$). For example, given $\sigma = 4,7 > 1,5,8 > 3 > 2,6$, we obtain

$$(w_1(\sigma), \ldots, w_8(\sigma)) = (2,2,2,2,1,1,2,1).$$

Also, if $v_1\cdots v_n$ is any sequence of integers, we define the usual *major index* statistic by

$$maj(v_1\cdots v_n) = \sum_{i=1}^{n-1} i\chi(v_i > v_{i+1}).$$

**Theorem 28.**

$$CH_n(q,t) = \sum_{\sigma \in S_n} q^{maj(\sigma)} \prod_{i=1}^{n} [w_i(\sigma)]_t. \tag{4}$$

*Proof.* This formula is proved in [11]. It also follows as a special case of formula (18), proved below. □

We end this subsection with a brief discussion of the connection between parking functions and labelled Dyck paths.

**Definition 29.** A *parking function* or *preference function* of order $n$ is a function $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ such that

$$|\{x : f(x) \leq i\}| \geq i \text{ for } 1 \leq i \leq n.$$

Let $\mathcal{P}'_n$ denote the collection of parking functions of order $n$.

As in [17], we think of the elements $x$ in the domain of $f$ as cars that wish to park on a one-way street with parking spots labelled $1, 2, \ldots, n$ (in that order). The number $f(x)$ represents the spot where car $x$ prefers to park. In the *standard parking policy*, cars 1 through $n$ arrive at the beginning of the street in increasing numerical order. Each car drives forward to the spot $f(x)$ it prefers. If this spot is available, the car parks there. If not, the car continues driving forward and parks in the next available spot. It can be shown that a function $f$ is a parking function iff all $n$ cars are able to park following this policy.

We can identify a parking function $f$ with a labelled Dyck path $P$ as follows. Let $S_i = \{x : f(x) = i\}$ be the set of cars preferring spot $i$. Starting in the bottom row of an $n$ by $n$ grid of lattice cells, place the elements of $S_1$ in increasing order in the first column of the diagram, one per row. Starting in the next empty row, place the elements of $S_2$ in increasing order in the second column of the diagram, one per row. Continue similarly: after listing all elements $x$ with $f(x) < i$, start in the next empty row and place the elements of $S_i$ in increasing order in column $i$. Finally, draw a lattice path from $(0, 0)$ to $(n, n)$ by drawing vertical steps immediately left of each label, and then drawing the necessary horizontal steps to get a connected path. It can be shown that the resulting labelled lattice path is a labelled Dyck path iff $f$ is a parking function. Furthermore, given a labelled Dyck path $P$, we can recover the parking function $f$ by setting $f(i) = j$ iff label $i$ occurs in column $j$. Thus, from now on, we will identify the set of parking functions $\mathcal{P}'_n$ with the set of labelled Dyck paths $\mathcal{P}_n$.

**Example 30.** Let $n = 8$, and define a function $f$ by

$$f(1) = 2,\ f(2) = 3,\ f(3) = 5,\ f(4) = 4,$$

$$f(5) = 1,\ f(6) = 4,\ f(7) = 2,\ f(8) = 6.$$

It is easy to check that $f$ is a parking function. The labelled path $P \in \mathcal{P}_8$ corresponding to $f$ is shown in Figure 6. Note that $area(P) = 9$.

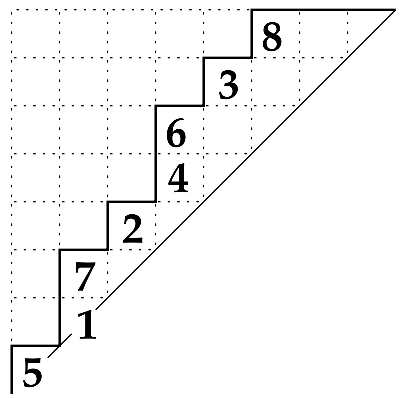



Figure 6: Diagram for a parking function.

If $P$ is the diagram for a parking function $f$, we can compute $area(P)$ as follows. Note that the triangle bounded by the lines $x = 0$, $y = n$, and $x = y$ contains $n(n-1)/2$ complete lattice cells. Since label $i$ occurs somewhere in column $f(i)$, there are $f(i) - 1$ lattice cells inside the triangle and left of label $i$. These lattice cells lie outside the Dyck path associated to $f$. Subtracting, we find that

$$area(P) = n(n-1)/2 - \sum_{i=1}^{n} [f(i) - 1] = n(n+1)/2 - \sum_{i=1}^{n} f(i). \tag{5}$$

For instance, in the example above we have

$$area(P) = 36 - (2 + 3 + 5 + 4 + 1 + 4 + 2 + 6) = 9.$$

## 1.6 Generalizations of the Diagonal Harmonics Module

In §3, we will discuss a generalization of Conjecture 27, based on pairs of statistics for generalized parking functions. The generalized conjecture involves modules introduced by Garsia and Haiman [9] that are natural extensions of the diagonal harmonics modules. We describe these modules now.

**Definition 31.** Fix integers $m, n \geq 1$. We define the *generalized diagonal harmonics module* $DH_n^{(m)}$ *of order* $m$ *in* $n$ *variables* as follows. As in §1.3, let $S_n$ act on the polynomial ring $R_n = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ via the diagonal action. Let $A_n$ denote the ideal in $R_n$ generated by all polynomials $P \in R_n$ for which

$$\sigma \cdot P = sgn(\sigma)P \text{ for all } \sigma \in S_n.$$

Let $A_n^m$ denote the ideal in $R_n$ generated by all products $P_1 P_2 \cdots P_m$, where each $P_i \in A_n$. Let $J_n$ denote the ideal in $R_n$ generated by all polarized power sums

$$\sum_{i=1}^{n} x_i^h y_i^k \quad (h + k \geq 1).$$

Finally, define

$$R_n^{(m)}[X;Y] = A_n^{m-1}/JA_n^{m-1}.$$

If $\sigma \in S_n$ and $f \in R_n^{(m)}[X;Y]$, the diagonal action induces an action of $S_n$ on this module, which we denote by $\sigma \cdot f$. Define a new action of $S_n$ by setting

$$\sigma \star f = (sgn(\sigma))^{m-1}\sigma \cdot f.$$

$DH_n^{(m)}$ is defined to be the doubly-graded module $R_n^{(m)}[X;Y]$ with this new action.

As with the original diagonal harmonics module, we would like to understand the Frobenius series $F_n^{(m)}(q,t)$, the Hilbert series $H_n^{(m)}(q,t)$, and the generating function for the sign character $RC_n^{(m)}(q,t)$ of $DH_n^{(m)}$. We have the following results, analogous to those in §1.3.

First, Haiman's results imply that the Frobenius series of $DH_n^{(m)}$ is given by

$$F_n^{(m)}(q,t) = \nabla^m(s_{1^n}).$$

By Theorem 13 and the definition of nabla, we have

$$\nabla^m(s_{1^n}) = \sum_{\mu\vdash n} \frac{\tilde{H}_\mu t^{mn(\mu)} q^{mn(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)}. \tag{6}$$

As in the case $m = 1$, there are nice formulas for the specializations at $t = 1$ and $t = 1/q$.

**Definition 32.** Let $\mathcal{D}_n^{(m)}$ denote the collection of lattice paths that go from $(0,0)$ to $(mn,n)$ by taking $n$ vertical steps and $mn$ horizontal steps and that never go below the line $x = my$. Such paths are called *m-Dyck paths of order n*. For $E \in \mathcal{D}_n^{(m)}$, define $area(E)$ to be the number of complete lattice cells between the path and the line $x = my$.

**Theorem 33.** *(1) For an m-Dyck path $D$ of order $n$, define $a_i(D)$ to be the number of vertical steps taken by the path along the line $x = i$. Then*

$$\nabla^m(s_{1^n})\big|_{t=1} = \sum_{D\in\mathcal{D}_n^{(m)}} q^{area(D)} \prod_{i=0}^{mn-1} e_{a_i(D)},$$

*where $e_j$ denotes an elementary symmetric function, as usual.*

*(2)*

$$q^{mn(n-1)/2}\nabla^m(s_{1^n})\big|_{t=1/q} = \sum_{\mu\vdash n} s_\mu \frac{s_{\mu'}(1,q,q^2,\ldots,q^{mn})}{[mn+1]_q}.$$

*Proof.* See Theorem 4.3 and Corollary 4.1 in [9]. □

Formula (6) gives the Frobenius series of $DH_n^{(m)}$ in terms of the symmetric functions $\tilde{H}_\mu$. To get the Hilbert series of $DH_n^{(m)}$, we can expand $\tilde{H}_\mu$ in terms of Schur functions and replace each $s_\lambda$ by $f_\lambda$. To get the generating function of the sign character, we extract the coefficient of $s_{1^n}$ in (6). What results is the following formula, which is called the *$n$'th bivariate Catalan number of order $m$*:

$$OC_n^{(m)}(q,t) = \sum_{\mu \vdash n} \frac{t^{(m+1)n(\mu)} q^{(m+1)n(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)}.$$

Haiman and the first author [12, 18, 19] defined combinatorial statistics on $m$-Dyck paths whose generating functions are conjectured to give $OC_n^{(m)}(q,t)$. These statistics will be generalized to labelled $m$-Dyck paths in §3 to give conjectured combinatorial interpretations for the higher-order Hilbert series $H_n^{(m)}(q,t)$.

## 1.7 Statistics for Trapezoidal Lattice Paths

This subsection discusses combinatorial statistics introduced by the first author [20, 19] on lattice paths contained in trapezoidal regions. These include the previously mentioned statistics on unlabelled Dyck paths and $m$-Dyck paths as special cases.

**Definition 34.** (1) Fix integers $n, k, m \geq 0$. Define a *trapezoidal lattice path of type* $(n, k, m)$ to be a lattice path that goes from $(0,0)$ to $(k+mn, n)$ by taking $n$ north steps and $k+mn$ east steps of length one, such that the path never goes strictly right of the line $x = k + my$. Let $\mathcal{T}_{n,k,m}$ be the set of all such paths.

(2) Given a path $P \in \mathcal{T}_{n,k,m}$, let $g_i(P)$ be the number of complete lattice squares between the path $P$ and the line $x = k + my$ in the $i$'th row from the bottom, for $0 \leq i < n$. Define the *area of $P$* by

$$area(P) = \sum_{i=0}^{n-1} g_i(P).$$

(3) For an integer $r$, set $r^+ = \max(r, 0)$. Define the *inversion statistic* for $P \in \mathcal{T}_{n,k,m}$ by

$$h(P) = \sum_{i<j}(m - |g_i - g_j|)^+ + \sum_{i<j}\chi(g_i - g_j \in \{1, 2, \ldots, m\}) + \sum_{i=0}^{n-1}(k - g_i)^+.$$

Alternatively, it can be shown [18, 19] that the following formula is equivalent to the previous one:

$$h(P) = \sum_{i<j}\sum_{d=0}^{m-1}\chi(g_i - g_j + d \in \{0, 1, \ldots, m\}) + \sum_{i=0}^{n-1}(k - g_i)^+.$$

(4) For $P \in \mathcal{T}_{n,k,m}$, define the *bounce path* $B(P)$ *associated to* $P$ as follows. A ball starts at $(0,0)$ and makes alternating vertical and horizontal moves until it reaches $(k+mn,n)$. Call the lengths of successive vertical and horizontal moves $v_i$ and $h_i$, for $i \geq 0$. These moves are determined as follows. At each step, the ball moves up $v_i \geq 0$ units from its current position until it is blocked by a horizontal step of the path $P$. The ball then moves right by $h_i$ units, where

$$h_i = v_i + v_{i-1} + \cdots + v_{i-(m-1)} + \chi(i < k). \tag{7}$$

In this formula, we let $v_i = 0$ for $i < 0$.

Finally, the *bounce score* for $P$ is the statistic

$$b(P) = \sum_{i \geq 0} i v_i.$$

(5) Define two generating functions

$$HC_{n,k,m}(q,t) = \sum_{P \in \mathcal{T}_{n,k,m}} q^{h(P)} t^{area(P)},$$

$$C_{n,k,m}(q,t) = \sum_{P \in \mathcal{T}_{n,k,m}} q^{area(P)} t^{b(P)}.$$

For a detailed combinatorial study of these statistics, see [20, 19]. In particular, it is shown there that the bounce path of $P$ always stays inside the trapezoid with vertices $(0,0)$, $(0,n)$, $(k,0)$ and $(k+mn,n)$. Also, the bounce path always reaches the upper-right corner $(k+mn,n)$, so that the algorithm for generating the bounce path always terminates.

We have the following identity, which has an explicit bijective proof:

$$HC_{n,k,m}(q,t) = C_{n,k,m}(q,t).$$

Furthermore, it is conjectured that

$$C_{n,0,m}(q,t) = OC_n^{(m)}(q,t).$$

**Example 35.** (1) Let $n = 6$, $k = 2$, and $m = 3$. Consider the unique path $P \in \mathcal{T}_{n,k,m}$ whose area vector is $g(P) = (1,4,4,0,3,1)$. This path is shown in Figure 7. We have $h(P) = 26$ and $area(P) = 13$.

(2) Figure 8 shows a trapezoidal path $P \in \mathcal{T}_{12,3,2}$ and its associated bounce path. We have $area(P) = 60$ and $b(P) = 31$.

**Remark 36.** When $k = 0$ and $m = 1$, the set $\mathcal{T}_{n,k,m}$ is exactly the set of Dyck paths of order $n$. Note that the bounce path described in this subsection starts at $(0,0)$ and ends at $(n,n)$. On the other hand, in Haglund's original bounce path construction for Dyck

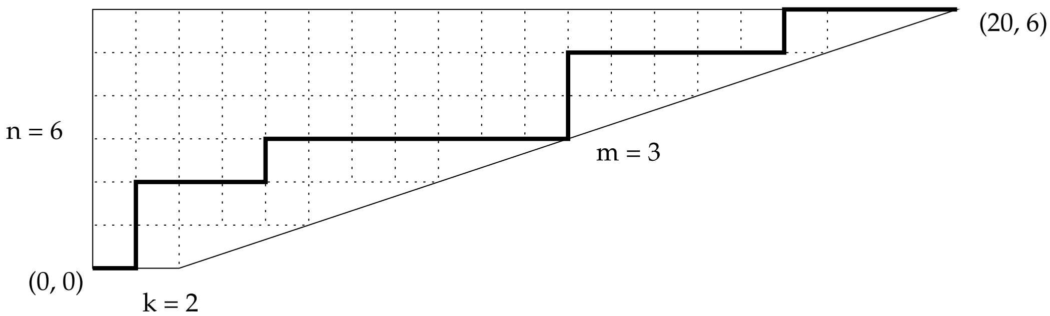
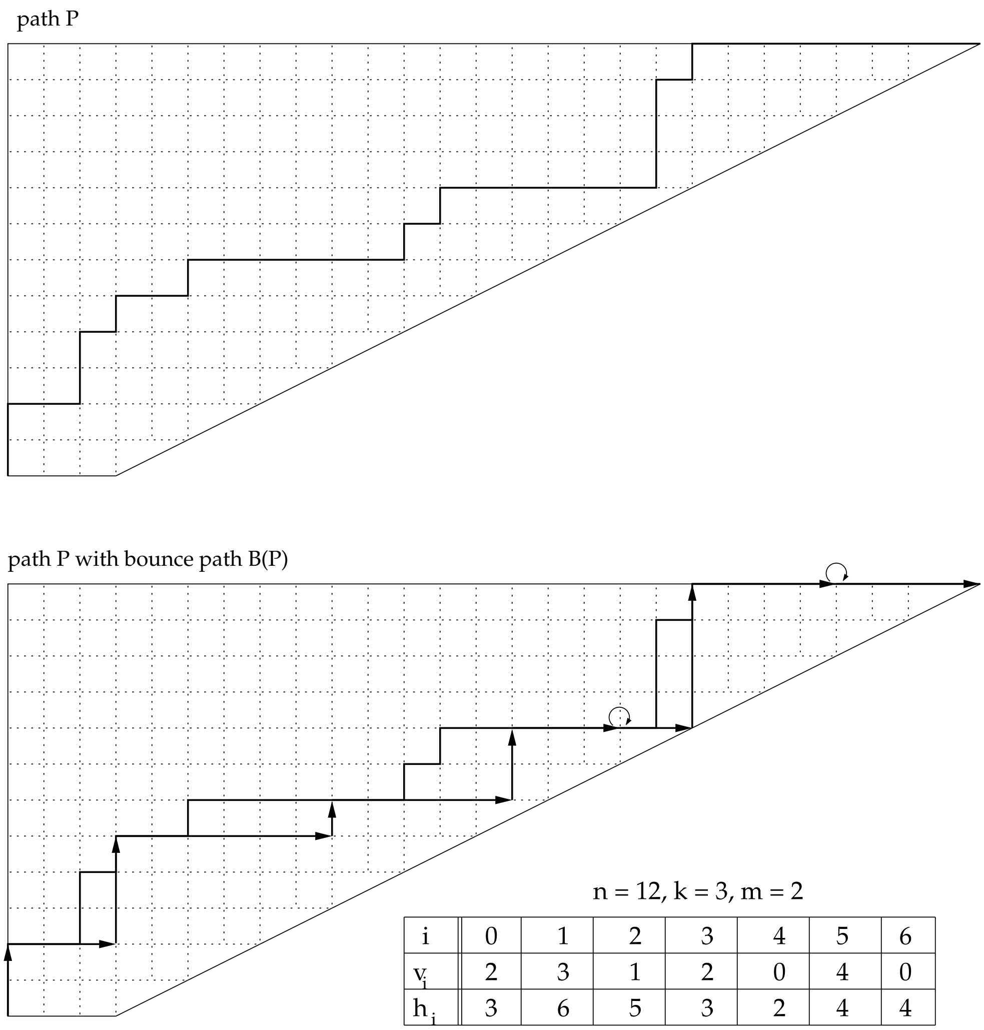



Figure 7: A trapezoidal lattice path.




| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $v_i$ | 2 | 3 | 1 | 2 | 0 | 4 | 0 |
| $h_i$ | 3 | 6 | 5 | 3 | 2 | 4 | 4 |

Figure 8: A trapezoidal path and its associated bounce path.

paths (see §1.4), the bounce path starts at $(n,n)$ and ends at $(0,0)$. It is easy to see that reflecting a Dyck path about the line $y=n-x$ transforms one bounce path to the other bounce path while preserving area. Hence, we have

$$C_n(q,t)=C_{n,0,1}(q,t).$$

In the rest of this paper, we will always compute bounce statistics using bounce paths starting at the origin, as described in this subsection.

# 2 Statistics based on Parking Policies

Recall from §1.5 that there are two pairs of statistics $(area, dinv)$ and $(dmaj, area')$ on parking functions that give conjectured combinatorial interpretations for the Hilbert series $H_n(q,t)$ of $DH_n$. This section introduces a third pair of statistics $(pmaj, area)$ on parking functions that has the same generating function as the previous two. In symbols, we have

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

Letting $q=1$ here shows that $area$, $dinv$ and $area'$ have the same univariate distribution, while letting $t=1$ shows that $pmaj$, $area$, and $dmaj$ have the same univariate distribution. Hence, all five individual statistics have the same univariate distribution. This result settles one of the open questions from [11].

Our starting point is the formula

$$CH_n(q,t)=\sum_{P\in\mathcal{P}_n} q^{area(P)}t^{dinv(P)}=\sum_{\sigma\in S_n} q^{maj(\sigma)}\prod_{i=1}^{n}\sum_{p=0}^{w_i(\sigma)-1} t^p. \tag{8}$$

It is convenient to represent this formula combinatorially. To do this, consider objects $I=(\sigma;u_1,\ldots,u_n)$, where $\sigma\in S_n$ and $u_i$ are integers satisfying $0\le u_i<w_i(\sigma)$. Let $\mathcal{I}_n$ denote the collection of such objects. Define $qstat(I)=maj(\sigma)$ and $tstat(I)=\sum_{i=1}^n u_i$. It is obvious from these definitions and formula (8) that

$$CH_n(q,t)=\sum_{I\in\mathcal{I}_n} q^{qstat(I)}t^{tstat(I)}. \tag{9}$$

In particular, letting $q=t=1$ here, we obtain

$$|\mathcal{I}_n|=|\mathcal{P}_n|=(n+1)^{n-1}. \tag{10}$$

We will define a statistic $pmaj$ on $\mathcal{P}_n$ and give a bijection $G:\mathcal{I}_n\to\mathcal{P}_n$ such that

$$qstat(I)=pmaj(G(I)) \text{ and } tstat(I)=area(G(I)).$$

It will then follow that

$$CH_n(q,t)=\sum_{P\in\mathcal{P}_n} q^{pmaj(P)}t^{area(P)}.$$

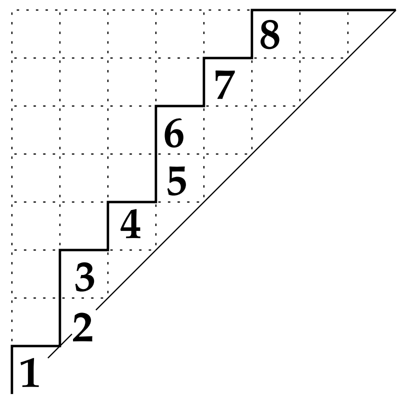



Figure 9: A labelled path with labels in increasing order.

The simplest way to define *pmaj* involves parking functions. Let $P \in \mathcal{P}_n$, and let $f$ be the associated parking function. Recall that $f(x) = j$ is interpreted to mean that car $x$ prefers spot $j$. Let $S_j = f^{-1}(j)$ be the set of cars that want to park in spot $j$. Let $T_j = \bigcup_{k=1}^{j} S_k$ be the set of cars that want to park at or before spot $j$. The definition of a parking function states that $|T_j| \geq j$ for $1 \leq j \leq n$.

We introduce the following new *parking policy*. Consider parking spots $1, \ldots, n$ in this order. These spots will be filled with cars $\tau_1, \ldots, \tau_n$ according to certain rules. The car $\tau_1$ that gets spot 1 is the largest car $x$ in the set $S_1 = T_1$. The car $\tau_2$ that gets spot 2 is the largest car $x$ in $T_2 - \{\tau_1\}$ such that $x < \tau_1$; if there is no such car, then $x$ is the largest car in $T_2 - \{\tau_1\}$. In general, the car $\tau_i$ that gets spot $i$ is the largest car $x$ in $T_i - \{\tau_1, \ldots, \tau_{i-1}\}$ such that $x < \tau_{i-1}$; if there is no such car, then $x$ is the largest car in $T_i - \{\tau_1, \ldots, \tau_{i-1}\}$. Since $|T_i| \geq i$, the set $T_i - \{\tau_1, \ldots, \tau_{i-1}\}$ is never empty. So this selection process makes sense. At the end of this process, we obtain a *parking order* $\tau = \tau_1, \ldots, \tau_n$, which is a permutation of $1, \ldots, n$. We let $\sigma = \sigma(P)$ be the reversal of $\tau$, so that $\sigma_j = \tau_{n+1-j}$ and $\tau_j = \sigma_{n+1-j}$ for $1 \leq j \leq n$. Finally, we define $pmaj(f) = pmaj(P) = maj(\sigma(P))$. Recall that $maj(\sigma_1 \cdots \sigma_n) = \sum_{i=1}^{n-1} i\chi(\sigma_i > \sigma_{i+1})$.

**Example 37.** For the parking function $f$ corresponding to the labelled path $P$ in Figure 6, the new parking policy gives

$$\tau = 5, 1, 7, 6, 4, 3, 2, 8.$$

Hence, $\sigma = 8 > 2, 3, 4, 6, 7 > 1, 5$, and so $pmaj(P) = maj(\sigma) = 1 + 6 = 7$.

**Example 38.** Consider the labelled path $P$ in Figure 9, in which the labels 1 to $n$ appear in order from bottom to top.

The new parking policy gives

$$\tau = 1, 3, 2, 6, 5, 4, 8, 7.$$

Hence, $\sigma = 7, 8 > 4, 5, 6 > 2, 3 > 1$, and so $pmaj(P) = maj(\sigma) = 14$. On the other hand, drawing the bounce path for the corresponding unlabelled path (starting at $(0, 0)$, as in

Remark 36) gives bounces of lengths $1, 2, 3, 2$. Thus, the bounce statistic for this path is also 14.

**Remark 39.** As in the previous example, it is easy to see that the *pmaj* statistic always reduces to the bounce statistic in the case where the labels 1 to $n$ increase from bottom to top. The proof, which is by induction on the number of bounces, is left to the reader.

We now define a map $G : \mathcal{I}_n \to \mathcal{P}_n$. Let $I = (\sigma; u_1, \ldots, u_n) \in \mathcal{I}_n$. We define $G(I)$ to be the function $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ such that

$$f(\sigma_i) = (n + 1 - i) - u_i \text{ for } 1 \leq i \leq n. \tag{11}$$

**Lemma 40.** *The function $G$ does map into the set $\mathcal{P}_n$.*

*Proof.* By definition, $w_i(\sigma)$ is no greater than the length of the list $\sigma_i, \sigma_{i+1}, \ldots, \sigma_n$. Hence,

$$0 \leq u_i < w_i(\sigma) \leq n + 1 - i,$$

which shows that

$$1 \leq f(\sigma_i) \leq n + 1 - i \leq n.$$

In particular, the image of $f$ is contained in the codomain $\{1, 2, \ldots, n\}$. This inequality also shows that the set $f^{-1}(\{1, 2, \ldots, i\})$ contains at least the $i$ elements $\sigma_n, \ldots, \sigma_{n+1-i}$, so that $f$ is a parking function. This shows that the image of $G$ is contained in the set $\mathcal{P}_n$. $\square$

We will see shortly that $G$ is a weight-preserving bijection.

**Example 41.** Let $n = 8$ and let $I = (\sigma; u_1, \ldots, u_n)$, where

$$\sigma = 8 > 2, 3, 4, 6, 7 > 1, 5;$$

$$w_1 = 5,\ w_2 = 5,\ w_3 = 4,\ w_4 = 3,$$

$$w_5 = 3,\ w_6 = 2,\ w_7 = 2,\ w_8 = 1;$$

$$u_1 = 2,\ u_2 = 4,\ u_3 = 1,\ u_4 = 1,$$

$$u_5 = 0,\ u_6 = 1,\ u_7 = 0,\ u_8 = 0.$$

Using the formula above, we have $G(I) = f$, where

$$f(1) = f(\sigma_7) = 2,\ f(2) = f(\sigma_2) = 3,\ f(3) = f(\sigma_3) = 5,\ f(4) = f(\sigma_4) = 4,$$

$$f(5) = f(\sigma_8) = 1,\ f(6) = f(\sigma_5) = 4,\ f(7) = f(\sigma_6) = 2,\ f(8) = f(\sigma_1) = 6.$$

The labelled path $P$ corresponding to this $f$ appears in Figure 6. Note that

$$qstat(I) = 6 = pmaj(f) \text{ and } tstat(I) = 9 = area(f).$$

We now define a map $H : \mathcal{P}_n \to \mathcal{I}_n$ that will turn out to be the inverse of $G$. Let $P \in \mathcal{P}_n$, and let $f$ be the associated parking function. Construct a permutation $\sigma$, as in the definition of $pmaj$, by reversing the parking permutation $\tau$. Define

$$u_i = n + 1 - i - f(\sigma_i) \text{ for } 1 \le i \le n. \tag{12}$$

Finally, set $H(P) = H(f) = (\sigma; u_1, \ldots, u_n)$.

**Lemma 42.** *$H$ does map $\mathcal{P}_n$ into the set $\mathcal{I}_n$. Moreover,*

$$G \circ H = Id_{\mathcal{P}_n}, \tag{13}$$

$$pmaj(P) = qstat(H(P)) \text{ and } area(P) = tstat(H(P)).$$

*Proof.* Let $f \in \mathcal{P}_n$. As usual, we set $S_j = f^{-1}(j)$ and $T_j = f^{-1}(\{1, 2, \ldots, j\})$. To see that $H$ maps into $\mathcal{I}_n$, we need only show that $0 \le u_i < w_i(\sigma)$. Observe that $\sigma_i = \tau_{n+1-i}$ is an element of $T_{n+1-i}$, and so $1 \le f(\sigma_i) \le n + 1 - i$. Hence, $u_i = n + 1 - i - f(\sigma_i)$ always satisfies the inequalities

$$0 \le u_i \le n - i < n + 1 - i. \tag{14}$$

We now consider several cases.

(I) $\sigma_i$ occurs in the rightmost ascending run of $\sigma$. By definition of $w_i$, this implies $w_i(\sigma) = n + 1 - i$. In this case, inequality (14) immediately gives the desired conclusion $0 \le u_i < w_i(\sigma)$.

(II) $\sigma_i$ is not in the rightmost ascending run of $\sigma$, and $\sigma$ can be written

$$\sigma = \cdots \sigma_i \cdots \sigma_k > \sigma_{k+1} \cdots \sigma_j \cdots,$$

where: $\sigma_k$ is the last entry in the ascending run containing $\sigma_i$ (so $k \ge i$); $\sigma_j$ and $\sigma_{k+1}$ are in the same ascending run; $\sigma_j < \sigma_i$; and either: (a) $j = n$, or (b) $\sigma_j > \sigma_{j+1}$, or (c) $\sigma_j < \sigma_{j+1}$ and $\sigma_{j+1} > \sigma_i$. By definition, $w_i(\sigma) = j - i$. It suffices to check that $u_i < w_i(\sigma)$. Substituting $u_i = n+1-i-f(\sigma_i)$ and $w_i(\sigma) = j-i$, it suffices to check that $f(\sigma_i) > n+1-j$. If this inequality did not hold, we would have $f(\sigma_i) \le n+1-j$, hence $\sigma_i \in T_{n+1-j}$. This will contradict the definition of the parking policy used to create $\tau$, as follows. Consider $\sigma_j = \tau_{n+1-j}$. In subcase (a), $\sigma_j = \sigma_n = \tau_1 = \max T_1$. But our assumption gives $\sigma_i \in T_1$ and $\sigma_i > \sigma_j$, a contradiction. In subcase (b), $\sigma_j > \sigma_{j+1}$ means that $\tau_{n+1-j} > \tau_{n-j}$, which implies that all elements of the set

$$T_{n+1-j} - \{\tau_1, \ldots, \tau_{n-j}\} = T_{n+1-j} - \{\sigma_{j+1}, \ldots, \sigma_n\}$$

are larger than $\tau_{n-j} = \sigma_{j+1}$, and $\sigma_j$ is the largest element in this set. But $\sigma_i$ is also an element of this set, and it is larger than $\sigma_j$, a contradiction. In subcase (c), $\sigma_j < \sigma_{j+1}$ implies that $\sigma_j$ is the largest element in the set

$$T_{n+1-j} - \{\tau_1, \ldots, \tau_{n-j}\} = T_{n+1-j} - \{\sigma_{j+1}, \ldots, \sigma_n\}$$

that is smaller than $\sigma_{j+1}$. But our assumption gives that $\sigma_i$ is in this set and satisfies $\sigma_j < \sigma_i < \sigma_{j+1}$, a contradiction. Thus, the desired inequality must hold in all subcases.

(III) $\sigma_i$ is not in the rightmost ascending run of $\sigma$, and $\sigma$ can be written

$$\sigma = \cdots \sigma_i \cdots \sigma_j > \sigma_{j+1} \cdots ,$$

where: $\sigma_j$ is the last entry in the ascending run containing $\sigma_i$ (so $j \geq i$); and $\sigma_i < \sigma_{j+1}$. These inequalities force $\sigma_i < \sigma_j$. By definition, $w_i(\sigma) = j - i$. As in case (II), the desired inequality $u_i < w_i(\sigma)$ is equivalent to the inequality $f(\sigma_i) > n+1-j$. If the latter inequality fails, then $\sigma_i \in T_{n+1-j}$. As in case (II) subcase (b), $\sigma_j > \sigma_{j+1}$ means that $\tau_{n+1-j} > \tau_{n-j}$, which implies that all elements of the set

$$T_{n+1-j} - \{\tau_1, \ldots, \tau_{n-j}\} = T_{n+1-j} - \{\sigma_{j+1}, \ldots, \sigma_n\}$$

are larger than $\tau_{n-j} = \sigma_{j+1}$, and $\sigma_j$ is the largest element in this set. But $\sigma_i$ is an element of this set that is smaller than $\sigma_{j+1}$, which is a contradiction. So the desired inequality must hold.

This completes the proof that $H$ maps into $\mathcal{I}_n$.

Next, the definitions of $u_i$ and $G$ in (12) and (11) make it clear that

$$G \circ H = Id_{\mathcal{P}_n}.$$

It is also obvious from the definition of $H$ that

$$pmaj(P) = qstat(H(P)).$$

On the other hand, note that

$$\begin{aligned} tstat(H(P)) &= \textstyle\sum_{i=1}^n u_i = \sum_{i=1}^n (n+1-i) - \sum_{i=1}^n f(i) \\ &= n(n+1)/2 - \textstyle\sum_{i=1}^n f(i) = area(P), \end{aligned}$$

where the last equality is formula (5). □

**Example 43.** Let $n = 8$ and let $f \in \mathcal{P}_8$ be given by

$$f(1) = 2,\ f(2) = 3,\ f(3) = 5,\ f(4) = 4,$$

$$f(5) = 1,\ f(6) = 4,\ f(7) = 2,\ f(8) = 6.$$

As in Example 37, we compute $\sigma = 8 > 2, 3, 4, 6, 7 > 1, 5$. We then compute

$$u_1 = 2,\ u_2 = 4,\ u_3 = 1,\ u_4 = 1,$$

$$u_5 = 0,\ u_6 = 1,\ u_7 = 0,\ u_8 = 0.$$

Note that $H(f) = I$, where $I$ is the object in $\mathcal{I}_n$ from Example 41. We have $G(H(f)) = f$ and $H(G(I)) = I$.

**Theorem 44.** *The maps $G : \mathcal{I}_n \to \mathcal{P}_n$ and $H : \mathcal{P}_n \to \mathcal{I}_n$ are bijections with $H = G^{-1}$. $G$ and $H$ are weight-preserving in the sense that*

$$pmaj(P) = qstat(H(P)) \text{ and } area(P) = tstat(H(P)); \tag{15}$$

$$qstat(I) = pmaj(G(I)) \text{ and } tstat(I) = area(G(I)). \tag{16}$$

*Consequently,*

$$\sum_{P \in \mathcal{P}_n} q^{pmaj(P)} t^{area(P)} = CH_n(q,t) = \sum_{P \in \mathcal{P}_n} q^{area(P)} t^{dinv(P)} = \sum_{Q \in \mathcal{Q}_n} q^{dmaj(Q)} t^{area'(Q)}, \tag{17}$$

*and so all these statistics have the same univariate distribution.*

*Proof.* We have already shown that $G$ maps into $\mathcal{P}_n$, $H$ maps into $\mathcal{I}_n$, and $G \circ H = Id_{\mathcal{P}_n}$. The last equation implies that $H$ is an injection and $G$ is a surjection. But we have seen in (10) that

$$|\mathcal{I}_n| = |\mathcal{P}_n| = (n+1)^{n-1} < \infty.$$

Since the sets are finite, $H$ is automatically a surjection, $G$ is automatically an injection, and $H = G^{-1}$. The properties in (15) were proved in the previous lemma, and (16) follows by replacing $P$ by $G(I)$ and simplifying. Finally, the first equality in (17) follows from (9) and the existence of the weight-preserving map bijection $G$. The other formulas for $CH_n(q,t)$ have already been discussed. Letting $q = 1$ or $t = 1$ in (17) gives the final assertion of the theorem. ☐

**Remark 45.** It can be shown directly from the definitions of $H$ and $G$ that $H \circ G = Id_{\mathcal{I}_n}$, without using the identity $|\mathcal{P}_n| = |\mathcal{I}_n|$. Given a labelled path of the form $G(I)$, where $I = (\sigma; u_1, \ldots, u_n)$, one shows by backwards induction that the algorithm defining $H(G(I))$ correctly recovers $\sigma_n, \sigma_{n-1}, \ldots, \sigma_1$. The argument is similar to the case analysis in the proof of Lemma 42, and is left to the interested reader.

# 3 Statistics for Labelled Trapezoidal Lattice Paths

This section describes statistics for labelled trapezoidal paths, which lead to a conjectured combinatorial interpretation for the Hilbert series of the modules $DH_n^{(m)}$.

**Definition 46.** Fix integers $n, k, m \geq 0$.

(1) A *labelled lattice path of height* $n$ consists of a lattice path having $n$ vertical steps labelled $1, 2, \ldots, n$ and an unspecified number of unlabelled horizontal steps. When drawing a labelled path, our convention is to place the label for each vertical step in the lattice square directly right of that vertical step. We call a labelled lattice path *valid* iff the labels in each column increase from bottom to top.

(2) A *labelled trapezoidal path* of type $(n, k, m)$ is a valid labelled lattice path whose underlying unlabelled path $P$ lies in $\mathcal{T}_{n,k,m}$. Let $\mathcal{P}_{n,k,m}$ denote the collection of all such labelled paths.

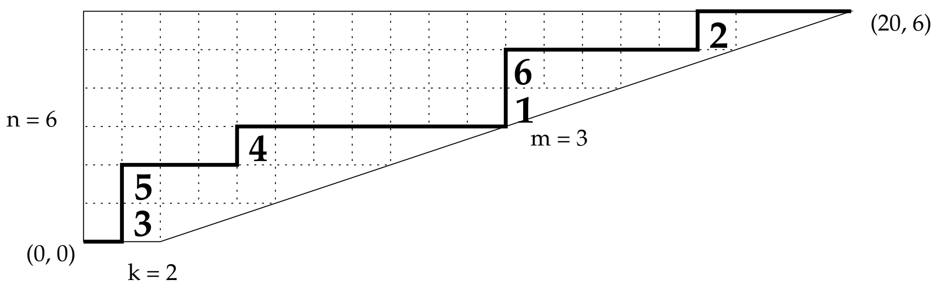



Figure 10: A labelled trapezoidal path.

As in the case of labelled Dyck paths, we can specify a labelled trapezoidal path $P$ by giving a pair of vectors

$$\vec{g}(P) = (g_0, g_1, \ldots, g_{n-1}), \quad \vec{p}(P) = (p_0, p_1, \ldots, p_{n-1}),$$

where $g_i(P)$ is the number of area cells in the $i$'th row from the bottom, and $p_i$ is the label of the vertical step in the $i$'th row from the bottom. It is easy to see that a vector of $n$ integers $(g_0, \ldots, g_{n-1})$ corresponds to a legal path in $\mathcal{T}_{n,k,m}$ iff the following conditions hold:

(A) $g_0 \in \{0, 1, \ldots, k\}$.

(B) $g_i \geq 0$ for all $i$.

(C) $g_{i+1} \leq g_i + m$ for all $i$.

Moreover, the associated vector of integers $\vec{p}(P)$ represents a valid labelling iff:

(D) $p_0, \ldots, p_{n-1}$ is a permutation of $1, 2, \ldots, n$.

(E) For all $i$, if $g_{i+1} = g_i + m$, then $p_i < p_{i+1}$.

Thus, when convenient, we may regard $\mathcal{P}_{n,k,m}$ as the set of all pairs of vectors $(\vec{g}, \vec{p})$ satisfying (A)—(E).

**Example 47.** Figure 10 shows a typical labelled path in $\mathcal{P}_{6,2,3}$. This object corresponds to the vector pair

$$((1, 4, 4, 0, 3, 1), (3, 5, 4, 1, 6, 2)).$$

We have the following analogues of the *area* and *dinv* statistics.

**Definition 48.** (1) The *area* of $P = (\vec{g}, \vec{p}) \in \mathcal{P}_{n,k,m}$ is defined by

$$area(P) = \sum_{i=0}^{n-1} g_i.$$

This is the number of area cells in the diagram of $P$, as usual.

(2) As above, set $r^+ = \max(r,0)$ for any integer $r$. The *inversion statistic* of $P$ is defined by letting

$$\begin{aligned}
h(P) &= h_1(P)+h_2(P)+h_3(P)-h_4(P), \text{ where:}\\
h_1(P) &= \sum_{i<j}(m-|g_i-g_j|)^+\\
h_2(P) &= \sum_{i=0}^{n-1}(k-g_i)^+\\
h_3(P) &= \sum_{i<j}\chi(g_i-g_j\in\{1,2,\ldots,m\} \text{ and } p_i>p_j)\\
h_4(P) &= \sum_{i<j}\chi(g_i-g_j\in\{0,-1,-2,\ldots,-(m-1)\} \text{ and } p_i>p_j)
\end{aligned}$$

Equivalently, we can define

$$h(P) = h_2(P) + \sum_{i<j}\sum_{d=0}^{m-1}\chi(A_{i,j,d}),$$

where $A_{i,j,d}$ is the logical statement

$$\begin{aligned}
&(g_i-g_j+d=0 \text{ and } p_i<p_j) \text{ or}\\
&(g_i-g_j+d\in\{1,2,\ldots,m-1\}) \text{ or}\\
&(g_i-g_j+d=m \text{ and } p_i>p_j).
\end{aligned}$$

The verification of this equivalence involves checking that the summands corresponding to a fixed choice of $i$ and $j$ in $h_1(P)+h_3(P)-h_4(P)$ always add up to the corresponding summand $\sum_{d=0}^{m-1}\chi(A_{i,j,d})$. This is done by considering cases based on the value of $g_i-g_j$ and whether $p_i>p_j$ or $p_i<p_j$ holds. These cases are checked in Table 1.

(3) Define

$$CH_{n,k,m}(q,t) = \sum_{P\in\mathcal{P}_{n,k,m}} q^{area(P)}t^{h(P)}.$$

**Example 49.** For the path

$$P = ((1,4,4,0,3,1),(3,5,4,1,6,2)).$$

shown in Figure 10, where $n=6$, $k=2$, $m=3$, the values of $g_i-g_j$ for $i<j$ are:

$$\begin{aligned}
i=1:&\quad -3,-3,1,-2,0;\\
i=2:&\quad 0,4,1,3;\\
i=3:&\quad 4,1,3;\\
i=4:&\quad -3,-1;\\
i=5:&\quad 2.
\end{aligned}$$

| Value of $z = g_i - g_j$ | Order of labels $p_i,\ p_j$ | Contribution to $h_1(P) + h_3(P) - h_4(P)$ | Value of $\sum_{d=0}^{m-1} A_{i,j,d}$ |
|---|---|---|---|
| $z > m$ | $p_i > p_j$ | 0 | 0 |
| $z > m$ | $p_i < p_j$ | 0 | 0 |
| $1 \leq z \leq m$ | $p_i > p_j$ | $m - z + 1$ | $m - z + 1$ |
| $1 \leq z \leq m$ | $p_i < p_j$ | $m - z$ | $m - z$ |
| $-m < z \leq 0$ | $p_i > p_j$ | $m - \lvert z \rvert - 1$ | $m - \lvert z \rvert - 1$ |
| $-m < z \leq 0$ | $p_i < p_j$ | $m - \lvert z \rvert$ | $m - \lvert z \rvert$ |
| $z \leq -m$ | $p_i > p_j$ | 0 | 0 |
| $z \leq -m$ | $p_i < p_j$ | 0 | 0 |

Table 1: Checking the equivalence of the two formulas for $h(P)$.

Hence, we compute:

$$area(P) = 13,\ h_1(P) = 16,\ h_2(P) = 4,\ h_3(P) = 4,\ h_4(P) = 2,\ h(P) = 22.$$

**Conjecture 50.** *For all $n, m \geq 1$, we have*

$$CH_{n,0,m}(q,t) = H_n^{(m)}(q,t).$$

*In other words, the statistics for labelled paths inside the triangle with vertices $(0,0)$, $(0,n)$, and $(mn,n)$ give a combinatorial interpretation for the Hilbert series of the generalized diagonal harmonics module $DH_n^{(m)}$.*

This conjecture has been confirmed for small values of $n$ and $m$ by computer, using the formula

$$H_n^{(m)}(q,t) = \nabla^m(s_{1^n})|_{s_\lambda = f_\lambda}$$

mentioned in the Introduction.

**Conjecture 51.** *For all $n, m \geq 1$, we have the specializations*

$$q^{mn(n-1)/2} CH_{n,0,m}(q,1/q) = [mn+1]_q^{n-1};$$

$$q^{n+mn(n-1)/2} CH_{n,1,m}(q,1/q) = (1+q^{n+1}) \cdot [mn+2]_q^{n-1}.$$

At present, there are no conjectures for the corresponding specializations when $k > 1$.

**Conjecture 52.** *For all $n, k, m$, we have the joint symmetry*

$$CH_{n,k,m}(q,t) = CH_{n,k,m}(t,q).$$

As evidence for this conjecture, we will prove the univariate symmetry

$$CH_{n,k,m}(q,1) = CH_{n,k,m}(1,q).$$

The proof will use an analogue of the *pmaj* statistic, which is defined later. First, we need to establish the analogue of the summation formula (4).

# 4 Summation Formula for $CH_{n,k,m}(q,t)$

In this section, we will derive a formula for the generating function $CH_{n,k,m}(q,t)$ as a summation over a collection of functions (equation (18) below). This formula is the extension of formula (4) to the cases $k > 0$ and $m > 1$.

Here are some remarks to motivate this new formula. We proved that the original formula (4) is the common generating function for the pairs of statistics $(pmaj, area)$ and $(area, dinv)$ on Dyck paths. In particular, this formula was the key ingredient in the proof that $pmaj$, $area$, and $dinv$ have the same univariate distribution. We will see that formula (18) plays a similar role in proving that statistics defined on $\mathcal{P}_{n,k,m}$ have the same distribution.

Examining the proof of (4), which appears in [11], suggests that we should look at subcollections of $\mathcal{P}_{n,k,m}$ where the labels appearing on each "diagonal" are fixed in advance. More precisely, suppose we are given an ordered partition $S_0, S_1, \ldots, S_{k+m(n-1)}$ of the set of labels $\{1, 2, \ldots, n\}$ into pairwise disjoint subsets, some of which may be empty. Then we can consider only those labelled paths $P = (\vec{g}, \vec{p})$ in $\mathcal{P}_{n,k,m}$ such that $p_i \in S_j$ implies $g_i = j$. In other words, the set of labels in $S_j$ must appear in rows of $P$ that contain exactly $j$ area cells.

In the original formula (4), where $k = 0$ and $m = 1$, it was convenient to represent the set partition $S_0, S_1, \ldots,$ as a permutation $\sigma$ as follows. First, write down the word

$$w = \mid S_n \mid S_{n-1} \mid \cdots S_3 \mid S_2 \mid S_1 \mid S_0$$

in which the elements of each $S_j$ (read from left to right) appear in increasing order, and a bar symbol is drawn between consecutive sets $S_j$. Now, it is easy to see that conditions (A)—(E) imply the following properties of $w$ when $k = 0$ and $m = 1$:

- $S_j = \emptyset$ implies $S_k = \emptyset$ for all $k > j$.
- The largest element of $S_j$ is greater than the smallest element of $S_{j-1}$ whenever both sets are nonempty.

Let $\sigma$ denote $w$ with all bar symbols erased; clearly, $\sigma$ is a permutation of $\{1, 2, \ldots, n\}$. The first property says that there are never two or more consecutive bar symbols, except possibly at the beginning of the word $w$. The second property says that the *descents* of $w$ occur precisely at the locations of the erased bars (occurring after the beginning of the word). Therefore, $w$ is recoverable from $\sigma$: given $\sigma$, we simply draw bars wherever descents occur, and then draw extra bars at the beginning of $w$ until there are $n$ bars total. Of course, the sets $S_0, S_1, \ldots$ are recoverable from $w$.

Unfortunately, the two properties above are no longer guaranteed in the case where $k > 0$ or $m > 1$. Hence, we are led to seek another representation for the set partition $S_0, S_1, \ldots$. It is convenient to introduce *functions* for this purpose. Let $f : \{1, 2, \ldots, n\} \to \{0, 1, \ldots, k + m(n-1)\}$ be a function. Then we obtain a set partition of $\{1, 2, \ldots, n\}$ by setting $S_j = f^{-1}(\{j\})$ for $0 \leq j \leq k + m(n-1)$. In this notation, we wish to consider the subcollection of paths $P = (\vec{g}, \vec{p})$ in $\mathcal{P}_{n,k,m}$ such that $f(p_i) = g_i$ for $1 \leq i \leq n$. It is convenient to introduce further notation to describe these functions.

**Definition 53.** Fix $n, k, m$. Let $f : \{1, 2, \ldots, n\} \to \{0, 1, \ldots, k + m(n-1)\}$ be any function.

(1) Define the subcollection of *labelled paths of type* $(n, k, m)$ *associated to* $f$ by

$$\mathcal{P}_{n,k,m}(f) = \{P = (\vec{g}, \vec{p}) \in \mathcal{P}_{n,k,m} : f(p_i) = g_i \text{ for } 1 \leq i \leq n\}.$$

Note that, for certain choices of $f$, this subcollection may be empty.

(2) For any set $T$, define the usual *inverse image of* $T$ *under* $f$ by

$$f^{-1}(T) = \{x \in \{1, 2, \ldots, n\} : f(x) \in T\}.$$

Also, for any integer $i$, define

$$f^{-1}_{<i}(T) = \{x : x < i \text{ and } f(x) \in T\},$$

$$f^{-1}_{>i}(T) = \{x : x > i \text{ and } f(x) \in T\}.$$

For brevity, we may write $f^{-1}(j)$ instead of $f^{-1}(\{j\})$, etc.

(3) Define the *set partition associated to* $f$ to be the list $(S_0, \ldots, S_{k+m(n-1)})$, where $S_j = f^{-1}(j)$ for each $j$.

(4) Define the *word of* $f$ by

$$w(f) = S_0 \mid S_1 \mid S_2 \mid \cdots \mid S_{k+m(n-1)} \mid$$

where the elements of each $S_j$ appear in *decreasing* order from left to right, followed by a bar symbol. Consecutive bar symbols appear in the word iff some $S_j$ is empty. Note that this is the reversal of the word $w$ described in the special case $k = 0$, $m = 1$ above. Also note that $f$ is recoverable from $w(f)$, thanks to the bar symbols. We may safely omit bar symbols that occur together at the far right of the word of $f$.

(5) Let $\mathcal{F}_{n,k,m}$ denote the set of all functions $f : \{1, 2, \ldots, n\} \to \{0, 1, \ldots, k+m(n-1)\}$. For $f \in \mathcal{F}_{n,k,m}$, define

$$\begin{aligned}
maj(f) &= \sum_{j=1}^{n} f(j) \\
count(f, j) &= \chi(f(j) \leq k) + |f^{-1}_{<j}(f(j) - m)| + |f^{-1}_{>j}(f(j))| \\
&\quad + |f^{-1}(\{f(j) - 1, \ldots, f(j) - (m-1)\})|. \\
x_0(f) &= \sum_{j_1 < j_2} (m - |f(j_1) - f(j_2)|)^+ + \sum_{j=1}^{n} (k - f(j))^+ \\
x_j(f) &= -|f^{-1}_{>j}(\{f(j), f(j) - 1, \ldots, f(j) - (m-1)\})|
\end{aligned}$$

**Example 54.** Let $n = 7$, $k = 2$, $m = 2$. Let the function $f$ be given by

$$f(1) = 2,\ f(2) = 0,\ f(3) = 0,\ f(4) = 3,\ f(5) = 0,\ f(6) = 2,\ f(7) = 2.$$

The word of $f$ is

$$w(f) = 5\ 3\ 2\ |\ |\ 7\ 6\ 1\ |\ 4\ ||||||||||||,$$

where there are 12 trailing bar symbols. Also

$$maj(f) = 9,\ count(f,1) = 3,\ count(f,2) = 3,\ count(f,3) = 2,$$

$$count(f,4) = 3,\ count(f,5) = 1,\ count(f,6) = 5,\ count(f,7) = 4,$$

$$x_0(f) = 15 + 6 = 21,\ x_1(f) = -2,\ x_2(f) = -2,\ x_3(f) = -1,$$

$$x_4(f) = -2,\ x_5(f) = 0,\ x_6(f) = -1,\ x_7(f) = 0.$$

The goal of the rest of this section is to establish the following formula.

**Theorem 55.**

$$CH_{n,k,m}(q,t) = \sum_{f \in \mathcal{F}_{n,k,m}} q^{maj(f)} t^{x_0(f)} \prod_{j=1}^{n} t^{x_j(f)} [count(f,j)]_t. \tag{18}$$

In the coming proofs, it will be convenient to use the following notation. Given a labelled path $P = (\vec{g}, \vec{p})$, we can think of $P$ as a single list of $n$ "tiles"

$$P = \boxed{\begin{matrix} g_1 \\ p_1 \end{matrix}} \boxed{\begin{matrix} g_2 \\ p_2 \end{matrix}} \cdots \boxed{\begin{matrix} g_n \\ p_n \end{matrix}}. \tag{19}$$

In this notation, specifying a function $f \in \mathcal{F}_{n,k,m}$ is equivalent to specifying a collection of $n$ tiles, namely $\boxed{\begin{matrix} f(i) \\ i \end{matrix}}$ for $1 \le i \le n$. The subcollection $\mathcal{P}_{n,k,m}(f)$ consists precisely of all rearrangements of these $n$ tiles that satisfy the restrictions (A)—(E) above. Note that (B) and (D) are guaranteed to hold, by definition of $f$.

We will be interested in building the object $P$ by putting down tiles one at a time. Thus, it is of interest to consider "partial" objects $Q$ satisfying (A)—(C) and (E) but not necessarily (D).

**Lemma 56.** *Suppose $P \in \mathcal{P}_{n,k,m}(f)$, so $P$ satisfies conditions (A)—(E). Let $i_1, \ldots, i_n$ be the word of $f$ with all bar symbols erased. Let $S = \{i_s, \ldots, i_n\}$ be any suffix of this word, where $2 \le s \le n$. Let $Q$ be obtained from $P$ by removing all tiles of the form $\boxed{\begin{matrix} f(i) \\ i \end{matrix}}$ for $i \in S$. Then $Q$ satisfies conditions (A)—(C) and (E).*

*Proof.* We prove the contrapositive in each case. If $Q$ does not satisfy condition (A), then it begins with a tile of the form $\begin{array}{|c|}\hline x \\ y \\ \hline\end{array}$, where $x > k$. By definition of the word of $f$ and $S$, all the removed tiles must have had top entries $x' \geq x$. Thus, the first tile of $P$ must have had top entry at least $x$, and so $P$ does not satisfy condition (A).

If $Q$ does not satisfy condition (B), then $P$ does not satisfy (B) either, since every tile in $Q$ is a tile in $P$.

Suppose $Q$ does not satisfy condition (C), so that there are consecutive tiles $\begin{array}{|c|c|}\hline x & x' \\ y & y' \\ \hline\end{array}$ in $Q$ with $x' > x + m$. As before, the definitions of $w(f)$ and $S$ show that any tiles in $P$ that were between these two tiles of $Q$ before being removed must have had top entry $x'' \geq x'$. Hence, the tile immediately following $\begin{array}{|c|}\hline x \\ y \\ \hline\end{array}$ in $P$ still has top entry larger than $x + m$. So $P$ does not satisfy condition (C).

Finally, suppose $Q$ does not satisfy condition (E), so that there are consecutive tiles $\begin{array}{|c|c|}\hline x & x' \\ y & y' \\ \hline\end{array}$ in $Q$ with $x' = x + m$ and $y > y'$. Let $\begin{array}{|c|}\hline x'' \\ y'' \\ \hline\end{array}$ be the tile immediately following $\begin{array}{|c|}\hline x \\ y \\ \hline\end{array}$ in $P$. If $x'' = x'$ and $y'' = y'$, then $P$ fails condition (E) already. Otherwise, by definition of $w(f)$ and $S$, we must have $x'' \geq x' = x + m$. Since $P$ satisfies condition (C), we in fact have $x'' = x + m = x'$. Now, since the elements of $S_{x+m}$ appear in *decreasing* order in the word of $f$, the presence of the tile $\begin{array}{|c|}\hline x+m \\ y' \\ \hline\end{array}$ in $Q$ is only possible if the value $y''$ in the removed tile is less than $y'$. But then $y > y''$, so that $P$ fails condition (E) in this case too. $\square$

**Lemma 57.** *Given $n, k, m$ and $f \in \mathcal{F}_{n,k,m}$, we have*

$$|\mathcal{P}_{n,k,m}(f)| = \prod_{j=1}^{n} count(f, j).$$

*Proof.* We can uniquely construct every object $P \in \mathcal{P}_{n,k,m}(f)$ as follows. Start with a pool of $n$ available tiles $\begin{array}{|c|}\hline f(i) \\ i \\ \hline\end{array}$, for $1 \leq i \leq n$. Let $i_1, \ldots, i_n$ be the word of $f$ with all bar symbols erased. Starting with an empty list of tiles, form the object $P$ by successively inserting each tile

$$\begin{array}{|c|}\hline f(i_1) \\ i_1 \\ \hline\end{array}, \quad \begin{array}{|c|}\hline f(i_2) \\ i_2 \\ \hline\end{array}, \quad \ldots, \quad \begin{array}{|c|}\hline f(i_n) \\ i_n \\ \hline\end{array}$$

into the list of previously inserted tiles. At each step, the new tile may be inserted anywhere in the existing list, provided that conditions (A), (C), and (E) hold. The previous lemma guarantees that all objects in $\mathcal{P}_{n,k,m}(f)$ can be constructed under these restrictions on tile insertions. Since the tiles are distinct, it is clear that there is a unique insertion order that will produce any given object $P$.

Thus, we need only count how many legal positions are available when each tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is inserted. Fix $j$. First, observe that the insertion order ensures that $f(j) \geq f(i)$ for all previously inserted tiles $\boxed{\begin{array}{c} f(i) \\ i \end{array}}$. This means that conditions (C) and (E) automatically hold for the tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ and the tile immediately following it (if any). Thus, to check that conditions (A), (C), and (E) continue to hold after the insertion of tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$, we need only check that: (i) $f(j) \leq k$ if $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is inserted in the leftmost position; or (ii) conditions (C) and (E) hold for the pair of tiles $\boxed{\begin{array}{c|c} f(\ell) & f(j) \\ \ell & j \end{array}}$, where $\boxed{\begin{array}{c} f(\ell) \\ \ell \end{array}}$ is the tile immediately preceding $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$. By condition (C), we must have $f(\ell) \in \{f(j), f(j)-1, \ldots, f(j)-m\}$. By condition (E), if $f(\ell) = f(j) - m$, then we must also have $\ell < j$.

Now, consider the various places where the new tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ may be inserted.

- The tile may be inserted at the far left position, becoming the new first tile in the list. By condition (A), this is allowable iff $f(j) \leq k$. So, we get a contribution of $\chi(f(j) \leq k)$ to the position count.

- The tile may be inserted immediately after a tile of the form $\boxed{\begin{array}{c} f(j)-m \\ \ell \end{array}}$, where we need $\ell < j$ by condition (E). By definition of $w(f)$ and the tile insertion order, all such tiles have already been placed when tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is being inserted. Therefore, the number of such tiles is
$$|f_{<j}^{-1}(f(j)-m)|.$$

- The tile may be inserted immediately after a tile of the form $\boxed{\begin{array}{c} f(j)-u \\ \ell \end{array}}$, where $1 \leq u < m$ and $\ell$ is arbitrary. By definition of $w(f)$ and the tile insertion order, all such tiles have already been placed when tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is being inserted. Therefore, the number of such tiles is
$$|f^{-1}(\{f(j)-1, \ldots, f(j)-(m-1)\})|.$$

- The tile may be inserted immediately after a tile of the form $\boxed{\begin{array}{c} f(j) \\ \ell \end{array}}$, where $\ell$ is arbitrary. However, by definition of $w(f)$ and the tile insertion order, only those tiles with $\ell > j$ have been inserted prior to the insertion of tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$. Therefore, the number of such tiles is only

$$|f^{-1}_{>j}(f(j))|.$$

- The new tile can *only* be inserted in positions of the type described in the last four cases, thanks to condition (C).

In summary, for each $j$ between 1 and $n$, the number of ways to place tile $\boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is precisely

$$\chi(f(j) \leq k) + |f^{-1}_{<j}(f(j) - m)| + |f^{-1}(\{f(j) - 1, \ldots, f(j) - (m-1)\})| + |f^{-1}_{>j}(f(j))|,$$

which is just $count(f, j)$. The formula in the statement of the lemma now follows from the product rule. $\square$

**Corollary 58.**

$$CH_{n,k,m}(q, 1) = \sum_{f \in \mathcal{F}_{n,k,m}} q^{maj(f)} \prod_{j=1}^{n} count(f, j). \tag{20}$$

*Proof.* Note that $\mathcal{P}_{n,k,m}$ is the disjoint union of the sets $\mathcal{P}_{n,k,m}(f)$ over all $f \in \mathcal{F}_{n,k,m}$. Fix $f$, and consider any $P \in \mathcal{P}_{n,k,m}$. We have

$$area(P) = \sum_{i=0}^{n-1} g_i = \sum_{i=0}^{n-1} f(p_i) = \sum_{i=1}^{n} f(i) = maj(f),$$

since the labels $p_i$ are a permutation of $1, 2, \ldots, n$. Thus, all paths in $\mathcal{P}_{n,k,m}(f)$ contribute a summand $q^{maj(f)}$ to the generating function $CH_{n,k,m}(q, 1)$. The stated formula then follows immediately from the previous lemma. $\square$

**Example 59.** Let $n$,$k$,$m$, and $f$ be as in the previous example. To construct an object $P \in \mathcal{P}_{n,k,m}(f)$, we should insert tiles in the following order:

$$\boxed{\begin{array}{c} 0 \\ 5 \end{array}}\ \boxed{\begin{array}{c} 0 \\ 3 \end{array}}\ \boxed{\begin{array}{c} 0 \\ 2 \end{array}}\ \boxed{\begin{array}{c} 2 \\ 7 \end{array}}\ \boxed{\begin{array}{c} 2 \\ 6 \end{array}}\ \boxed{\begin{array}{c} 2 \\ 1 \end{array}}\ \boxed{\begin{array}{c} 3 \\ 4 \end{array}}$$

An example of an object created in this way is

$$P = \boxed{\begin{array}{c} 2 \\ 6 \end{array}}\ \boxed{\begin{array}{c} 0 \\ 3 \end{array}}\ \boxed{\begin{array}{c} 2 \\ 7 \end{array}}\ \boxed{\begin{array}{c} 3 \\ 4 \end{array}}\ \boxed{\begin{array}{c} 2 \\ 1 \end{array}}\ \boxed{\begin{array}{c} 0 \\ 5 \end{array}}\ \boxed{\begin{array}{c} 0 \\ 2 \end{array}}$$

Note that $area(P) = 9 = maj(f)$.

Our next goal is to extend formula (20) to keep track of the statistic $h(P) = h_1(P) + h_2(P) + h_3(P) - h_4(P)$. The final formula, equation (18), will be proved in two steps. The first (easier) step involves analyzing the contribution of $h_1(P) + h_2(P)$. The next result shows that this quantity is constant for all objects $P$ in a given subcollection $\mathcal{P}_{n,k,m}(f)$.

**Lemma 60.**

$$\sum_{P \in \mathcal{P}_{n,k,m}} q^{area(P)} t^{h_1(P)+h_2(P)} = \sum_{f \in \mathcal{F}_{n,k,m}} q^{maj(f)} t^{x_0(f)} \prod_{j=1}^{n} count(f,j). \tag{21}$$

*Proof.* In light of formula (20) and its proof, we need only show that

$$h_1(P) + h_2(P) = x_0(f) \text{ for all } P \in \mathcal{P}_{n,k,m}(f).$$

Recall that for $P \in \mathcal{P}_{n,k,m}(f)$, we have $g_i = f(p_i)$ for all $i$. Also, $p_0, \ldots, p_{n-1}$ is a rearrangement of $1, 2, \ldots, n$, so we have

$$h_2(P) = \sum_{i=0}^{n-1} (k - g_i)^+ = \sum_{i=0}^{n-1} (k - f(p_i))^+ = \sum_{j=1}^{n} (k - f(j))^+.$$

Next, recall that

$$h_1(P) = \sum_{i<j} (m - |g_i - g_j|)^+.$$

This sum extends over all ordered pairs $(i,j)$ with $0 \le i < j \le n-1$. However, since $|g_i - g_j| = |g_j - g_i|$, we could equally well sum over all *unordered* pairs $\{i,j\}$ with $0 \le i, j \le n-1$ and $i \ne j$. Hence,

$$\begin{aligned} h_1(P) &= \sum_{\{i_1,i_2\}: i_1 \ne i_2} (m - |g_{i_1} - g_{i_2}|)^+ \\ &= \sum_{\{i_1,i_2\}: i_1 \ne i_2} (m - |f(p_{i_1}) - f(p_{i_2})|)^+ \\ &= \sum_{\{j_1,j_2\}: j_1 \ne j_2} (m - |f(j_1) - f(j_2)|)^+ \\ &= \sum_{j_1 < j_2} (m - |f(j_1) - f(j_2)|)^+. \end{aligned}$$

Combining these calculations and comparing to the definition of $x_0(f)$, we get $h_1(P) + h_2(P) = x_0(f)$ as desired. □

The next step is to analyze the value of $h_3(P) - h_4(P)$ for $P \in \mathcal{P}_{n,k,m}(f)$. Consider the partial objects

$$P_0, P_1, \ldots, P_n = P$$

that are constructed in Lemma 57 by inserting tiles in the order given by the word of $f$. We think of each newly inserted tile as contributing a certain increment to the statistic $h_3(P) - h_4(P)$. More specifically, let $y_0 = 0$ and, for $1 \leq i \leq n$, let

$$y_i = [h_3(P_i) - h_4(P_i)] \; - \; [h_3(P_{i-1}) - h_4(P_{i-1})] \, .$$

Then $h_3(P) - h_4(P) = \sum_{i=1}^{n} y_i$; note that $y_i$ is the change in the statistic $h_3 - h_4$ (which may be positive or negative) resulting from the insertion of the $i$'th tile.

It will be convenient to alter the indexing scheme slightly, as follows. Suppose the $i$'th tile in the insertion order is $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$. Then define $z_j = y_i$. In words, $z_j$ is the change in the statistic $h_3 - h_4$ due to the insertion of tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$. Note that $h_3(P) - h_4(P) = \sum_{j=1}^{n} z_j$.

We have shown, in the proof of Lemma 57, that there are exactly $c_j = \mathit{count}(f, j)$ valid positions in which tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$ may be inserted. Temporarily number these valid positions $0, 1, \ldots, c_j - 1$ reading from right to left. We will show later that, if the tile is placed in the valid position numbered $p$, then

$$z_j = x_j(f) + p \quad (0 \leq p < c_j). \tag{22}$$

Thus, the contribution to $h_3 - h_4$ due to this particular tile insertion can be accounted for by the polynomial

$$t^{x_j(f)} \sum_{p=0}^{c_j - 1} t^p = t^{x_j(f)} \cdot [\mathit{count}(f, j)]_t,$$

which is a $t$-analogue of the number $\mathit{count}(f, j)$ in (20). By the product rule for generating functions, we conclude that

$$\sum_{P \in \mathcal{P}_{n,k,m}} t^{h_3(P) - h_4(P)} = \sum_{f \in \mathcal{F}_{n,k,m}} \prod_{j=1}^{n} x_j(f) \cdot [\mathit{count}(f, j)]_t.$$

Combining this with the previous analysis for $\mathit{area}$ and $h_1 + h_2$, the desired formula (18) will follow immediately.

To prove the claims in the last paragraph, we need to consider the effect of inserting tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$ in *arbitrary* positions in the current tile configuration, not just the *valid* positions. We will now label each position with the change in $h_3 - h_4$ caused by inserting tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$ in this position, regardless of the validity of the resulting partial object.

**Example 61.** Continuing Example 59, consider the partial object

$$Q = \boxed{\begin{matrix} 2 \\ 6 \end{matrix}} \; \boxed{\begin{matrix} 0 \\ 3 \end{matrix}} \; \boxed{\begin{matrix} 2 \\ 7 \end{matrix}} \; \boxed{\begin{matrix} 0 \\ 5 \end{matrix}} \; \boxed{\begin{matrix} 0 \\ 2 \end{matrix}},$$

which has $h_3(Q) - h_4(Q) = 3$. The next tile to be inserted is $\boxed{\begin{matrix}2\\1\end{matrix}}$. The following diagram shows the change in $h_3 - h_4$ when we insert this tile in all possible positions. We have also labelled which positions are valid.

$$\begin{array}{rccccccccccc}
 & — & \boxed{\begin{matrix}2\\6\end{matrix}} & — & \boxed{\begin{matrix}0\\3\end{matrix}} & — & \boxed{\begin{matrix}2\\7\end{matrix}} & — & \boxed{\begin{matrix}0\\5\end{matrix}} & — & \boxed{\begin{matrix}0\\2\end{matrix}} & — \\
\text{validity:} & \text{yes} & & \text{yes} & & \text{no} & & \text{yes} & & \text{no} & & \text{no} \\
\text{change in } h_3 - h_4\text{:} & 0 & & -1 & & -1 & & -2 & & -2 & & -2
\end{array}$$

Note that if we look at only the *valid* positions, from right to left, the changes we get are $-2$, $-1$, and $0$, which are exactly the numbers $x_1(f) + p$ for $0 \le p < 3 = \mathit{count}(f, 1)$.

Now, as in the previous example, assume that we choose to insert tile $\boxed{\begin{matrix}2\\1\end{matrix}}$ immediately after $\boxed{\begin{matrix}2\\7\end{matrix}}$, producing the partial object

$$Q' = \boxed{\begin{matrix}2\\6\end{matrix}}\ \boxed{\begin{matrix}0\\3\end{matrix}}\ \boxed{\begin{matrix}2\\7\end{matrix}}\ \boxed{\begin{matrix}2\\1\end{matrix}}\ \boxed{\begin{matrix}0\\5\end{matrix}}\ \boxed{\begin{matrix}0\\2\end{matrix}},$$

with $h_3(Q') - h_4(Q') = 1$. The next tile to be inserted is $\boxed{\begin{matrix}3\\4\end{matrix}}$. The following diagram shows the change in $h_3 - h_4$ when we insert this tile in all possible positions. We have also labelled which positions are valid.

$$\begin{array}{rccccccccccccc}
 & — & \boxed{\begin{matrix}2\\6\end{matrix}} & — & \boxed{\begin{matrix}0\\3\end{matrix}} & — & \boxed{\begin{matrix}2\\7\end{matrix}} & — & \boxed{\begin{matrix}2\\1\end{matrix}} & — & \boxed{\begin{matrix}0\\5\end{matrix}} & — & \boxed{\begin{matrix}0\\2\end{matrix}} & — \\
\text{validity:} & \text{no} & & \text{yes} & & \text{no} & & \text{yes} & & \text{yes} & & \text{no} & & \text{no} \\
\text{change in } h_3 - h_4\text{:} & 1 & & 0 & & 0 & & -1 & & -2 & & -2 & & -2
\end{array}$$

Note that if we look at only the *valid* positions, from right to left, the changes we get are $-2$, $-1$, and $0$, which are exactly the numbers $x_4(f) + p$ for $0 \le p < 3 = \mathit{count}(f, 4)$.

Let us make some observations about these examples. First, note that there are *blocks* of consecutive insertion positions for which the change in $h_3 - h_4$ is the same. Each such block (except possibly the leftmost block) consists of zero or more invalid positions terminated by one valid position, scanning from right to left. The leftmost block may or may not end with a valid position, depending on the value of $k$. Next, note that the change in $h_3 - h_4$ for positions in the rightmost block is $x_j(f)$. As we pass from one block to the next, scanning from right to left as always, the change in $h_3 - h_4$ increases by 1 each time. Finally, we have already proved (in Lemma 57) that the number of valid positions is exactly $\mathit{count}(f, j)$. Combining all these observations, we deduce that the claim (22) does hold in these two examples. The next lemma shows that these observations are true in general, and hence claim (22) always holds.

**Lemma 62.** *Fix $n$, $k$, $m$, and $f \in \mathcal{F}_{n,k,m}$. Let $Q$ be a partial object, constructed as in the proof of Lemma 57 by inserting tiles in the order given by the word of $f$. Suppose $T_j = \boxed{\begin{array}{c} f(j) \\ j \end{array}}$ is the next tile to be inserted into $Q$. Label each insertion position with the change in $h_3 - h_4$ caused by inserting the new tile in that position. Then we have the following properties:*

*(1) The rightmost position is labelled $x_j(f)$.*

*(2) Suppose two consecutive insertion positions in $Q$ are separated by a tile $T_p = \boxed{\begin{array}{c} f(p) \\ p \end{array}}$.*

*(a) If the position just right of $T_p$ is an invalid position for $T_j$, then the position just left of $T_p$ has the same label as the position just right of $T_p$.*

*(b) If the position just right of $T_p$ is a valid position for $T_j$, then the label of the position just left of $T_p$ is one more than the label of the position just right of $T_p$.*

*(3) Hence, when scanning the count$(f,j)$* **valid** *insertion positions from right to left, their labels are precisely the numbers*

$$x_j(f) + p \quad (0 \le p < \mathit{count}(f,j)).$$

*Proof.* We begin by observing that, for any tile $T_p = \boxed{\begin{array}{c} f(p) \\ p \end{array}}$ in the partial object $Q$, we must have $f(j) \ge f(p)$; moreover, if $f(j) = f(p)$, then $p > j$. This follows directly from the definition of the word of $f$ and the tile insertion order.

To prove (1), recall the definitions of $h_3$ and $h_4$:

$$\begin{aligned} h_3(Q) &= \sum_{i_1<i_2} \chi(g_{i_1} - g_{i_2} \in \{1,2,\ldots,m\} \text{ and } p_{i_1} > p_{i_2}) \\ h_4(Q) &= \sum_{i_1<i_2} \chi(g_{i_1} - g_{i_2} \in \{0,-1,-2,\ldots,-(m-1)\} \text{ and } p_{i_1} > p_{i_2}) \end{aligned}$$

Suppose we insert tile $T_j$ in the far right position, after all the tiles $T_p = \boxed{\begin{array}{c} f(p) \\ p \end{array}}$ in $Q$. The change in $h_3$ caused by this insertion is

$$\sum_{T_p \in Q} \chi(f(p) - f(j) \in \{1,2,\ldots,m\} \text{ and } p > j) = 0,$$

since the observation above gives $f(p) - f(j) \le 0$. On the other hand, the change in $h_4$ caused by the insertion is

$$\sum_{T_p \in Q} \chi(f(p) - f(j) \in \{0,-1,\ldots,-(m-1)\} \text{ and } p > j)$$

$$= |f_{>j}^{-1}(\{f(j), f(j)-1, \ldots, f(j)-(m-1)\})|.$$

The last equality uses the fact that all tiles with lower entries in the set

$$f_{>j}^{-1}(\{f(j), \ldots, f(j)-(m-1)\})$$

have already been inserted in $Q$ before tile $T_p$ is inserted; this again follows from the definition of the insertion order. To summarize, the change in $h_3 - h_4$ caused by inserting $T_p$ at the far right is

$$0 - |f_{>j}^{-1}(\{f(j), \ldots, f(j)-(m-1)\})| = x_j(f),$$

which proves (1).

In the proof of (2), we will consider configurations where the new tile $T_j$ is inserted immediately right or left of an existing tile $T_p$. Note that passing from the configuration $\cdots T_p T_j \cdots$ to $\cdots T_j T_p \cdots$ simply amounts to interchanging the two adjacent tiles $T_p$ and $T_j$. This interchange will only affect a single term in the formulas for $h_3$ and $h_4$. Specifically, in $h_3$, the term

$$t_1 = \chi(f(p) - f(j) \in \{1, 2, \ldots, m\} \text{ and } p > j)$$

will be replaced by the term

$$t_2 = \chi(f(j) - f(p) \in \{1, 2, \ldots, m\} \text{ and } j > p).$$

In $h_4$, the term

$$t_3 = \chi(f(p) - f(j) \in \{0, -1, \ldots, -(m-1)\} \text{ and } p > j)$$

will be replaced by the term

$$t_4 = \chi(f(j) - f(p) \in \{0, -1, \ldots, -(m-1)\} \text{ and } j > p).$$

The net change in the statistic $h_3 - h_4$ due to the interchange is therefore $(t_2 - t_4) - (t_1 - t_3) = t_2 + t_3 - t_1 - t_4$.

To prove (2a), assume that $p$ and $j$ are such that the position just right of $T_p$ is an invalid position for $T_j$. This situation occurs in the following two cases.

(i) We have $f(p) - f(j) < -m$, so that the position right of $T_p$ is invalid because condition (C) fails. We have $t_1 = t_2 = t_3 = t_4 = 0$, so the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is zero.

(ii) We have $f(p) - f(j) = -m$ and $p > j$, so that the position right of $T_p$ is invalid because condition (E) fails. We have $t_1 = t_2 = t_3 = t_4 = 0$, so the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is zero.

To prove (2b), assume that $p$ and $j$ are such that the position just right of $T_p$ is a valid position for $T_j$. This situation occurs in the following four cases.

(iii) We have $f(p) - f(j) = -m$ and $p < j$. Then $t_1 = t_4 = 0$, while $t_2 = 1$ and $t_3 = 0$. Hence, the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is $+1$.

(iv) We have $-(m-1) \leq f(p) - f(j) \leq -1$ and $p > j$. Then $t_1 = t_4 = 0$, while $t_2 = 0$ and $t_3 = 1$. Hence, the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is $+1$.

(v) We have $-(m-1) \leq f(p) - f(j) \leq -1$ and $p < j$. Then $t_1 = t_4 = 0$, while $t_2 = 1$ and $t_3 = 0$. Hence, the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is $+1$.

(vi) We have $f(p) - f(j) = 0$, which forces $p > j$ by definition of $w(f)$ and the tile insertion order. Then $t_1 = t_2 = 0$, while $t_3 = 1$ and $t_4 = 0$. Hence, the change in $h_3 - h_4$ when we move $T_j$ to the left of $T_p$ is $+1$.

Note that the cases (i)—(vi) are exhaustive, since the tile insertion order rules out the possibility that $f(p) - f(j) > 0$. This completes the proof of (2).

To prove (3), note that (1) shows the rightmost position has label $x_j(f)$. Reading the positions from right to left, (2a) implies that there will be a block of positions with label $x_j(f)$, consisting of zero or more invalid positions followed by one valid position. By (2b), the next position to the left will have label $x_j(f) + 1$. Then (2a) implies that there is another block of positions labelled $x_j(f) + 1$, consisting of zero or more invalid positions followed by one valid position. This process continues until all valid positions have been encountered. We saw in the proof of Lemma 57 that the number of valid positions is exactly $count(f, j)$. Note that the leftmost block of positions may or may not end with a valid position, depending on $k$. This ambiguity does not affect the correctness of the present argument, since we stop as soon as the last (leftmost) valid position has been scanned. (This is illustrated by the two preceding examples, where the leftmost position is valid in one case and invalid in the other.) $\square$

This lemma, together with the discussion preceding it, completes the proof of formula (18). We leave to the reader the task of showing that this formula reduces to formula (4) from [11] in the case $m = 1$, $k = 0$. This is merely a matter of notation translation, keeping in mind that the permutation $\sigma$ corresponds to the reversal of the word of $f$.

# 5 Univariate Symmetry of $CH_{n,k,m}(q, t)$

This section generalizes the constructions of §2 to labelled trapezoidal paths of type $(n, k, m)$. We obtain another combinatorial interpretation of the right side of formula (18) in which $t$ keeps track of area and $q$ keeps track of a new statistic $pmaj$. As in §2, we can conclude that the ordered pairs of statistics $(area, h)$ and $(pmaj, area)$ have the same bivariate distribution on labelled paths. Therefore, all three statistics have the same univariate distribution. Unfortunately, the arguments given here are not strong enough to prove the conjectured joint symmetry of $CH_{n,k,m}(q, t)$.

## 5.1 Combinatorial Model of the Generating Function

We begin by introducing a simple combinatorial model for the formula (18).

**Definition 63.** (1) Given $n$, $k$, $m$, and $f \in \mathcal{F}_{n,k,m}$, define the *right limit* of $j$ relative to $f$ by

$$R_j(f) = |f_{>j}^{-1}(\{f(j), \ldots, f(j) - (m-1)\})| = |x_j(f)|,$$

and define the *left limit* of $j$ relative to $f$ by

$$L_j(f) = x_j(f) + count(f, j) - 1.$$

Formula (18) can then be rewritten

$$CH_{n,k,m}(q,t) = \sum_{f \in \mathcal{F}_{n,k,m}} q^{maj(f)} t^{x_0(f)} \prod_{j=1}^{n} \sum_{p=-R_j(f)}^{p=L_j(f)} t^p.$$

(2) Fix $n$, $k$, and $m$. Define an *intermediate object* of type $(n,k,m)$ to be a pair

$$I = (f;\ u_1, u_2, \ldots, u_n),$$

where $f \in \mathcal{F}_{n,k,m}$ and where $u_j$ are integers such that $-R_j(f) \leq u_j \leq L_j(f)$ for all $j$. Denote the collection of such intermediate objects by $\mathcal{I}_{n,k,m}$.

(3) Define the *intermediate q-statistic* for $I$ to be

$$qstat(I) = maj(f) = \sum_{j=1}^{n} f(j).$$

Define the *intermediate t-statistic* for $I$ to be

$$tstat(I) = x_0(f) + \sum_{j=1}^{n} u_j.$$

It is obvious from the definition of the intermediate objects and statistics that

$$\sum_{I \in \mathcal{I}_{n,k,m}} q^{qstat(I)} t^{tstat(I)} = \sum_{f \in \mathcal{F}_{n,k,m}} q^{maj(f)} t^{x_0(f)} \prod_{j=1}^{n} \sum_{p=-R_j(f)}^{p=L_j(f)} t^p = CH_{n,k,m}(q,t).$$

**Theorem 64.** *There exists a bijection $F : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$ such that*

$$area(P) = qstat(F(P)) \text{ and } h(P) = tstat(F(P)) \text{ for all } P \in \mathcal{P}_{n,k,m}.$$

*Proof.* The bijection $F$ is based on the tile insertion process from the last section (see Lemma 57). If $P = (\vec{g}, \vec{p})$ is a labelled path, we define $f \in \mathcal{F}_{n,k,m}$ by setting $f(p_i) = g_i$, and we define $u_j$ to be the change in the statistic $h_3 - h_4$ caused by the insertion of the tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$. We then set $F(P) = (f; u_1, \ldots, u_n)$. Lemma 62 shows that each $u_j$ satisfies the required inequalities

$$-R_j(f) \leq u_j \leq L_j(f).$$

The discussion in the last section shows that $area(P) = qstat(F(P))$ and $h(P) = tstat(F(P))$. The map $F^{-1}$ is defined similarly: given $I = (f; u_1, \ldots, u_n)$, the function $f$ tells us which tiles to use, and the numbers $u_j$ tell us where to insert each tile to reconstruct $P$. Lemma 62 shows that there exists a unique valid insertion position for tile $\boxed{\begin{matrix} f(j) \\ j \end{matrix}}$ that causes a change of $u_j$ in the statistic $h_3 - h_4$, so that $F^{-1}$ is well-defined. Thus $F$ is a bijection. $\square$

**Corollary 65.**

$$|\mathcal{I}_{n,k,m}| = |\mathcal{P}_{n,k,m}| \text{ for all } n, k, m. \tag{23}$$

*Proof.* This is immediate from the existence of the bijection $F : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$. $\square$

**Example 66.** Let us compute $F(P)$, where $P$ is the path given in tile notation by

$$P = \boxed{\begin{matrix} 2 \\ 6 \end{matrix}}\ \boxed{\begin{matrix} 0 \\ 3 \end{matrix}}\ \boxed{\begin{matrix} 2 \\ 7 \end{matrix}}\ \boxed{\begin{matrix} 3 \\ 4 \end{matrix}}\ \boxed{\begin{matrix} 2 \\ 1 \end{matrix}}\ \boxed{\begin{matrix} 0 \\ 5 \end{matrix}}\ \boxed{\begin{matrix} 0 \\ 2 \end{matrix}}$$

and $(n, k, m) = (7, 2, 2)$. First, examination of the tiles in $P$ gives

$$f(1) = 2,\ f(2) = 0,\ f(3) = 0,\ f(4) = 3,\ f(5) = 0,\ f(6) = 2,\ f(7) = 2.$$

Recall that $w(f)$, $count(f, j)$, etc., were computed before in Example 54. To find the numbers $u_j$, we build up $P$ by inserting tiles in the order given in Example 59. For instance, Example 61 discusses the last two steps of the tile insertion. From that example, we see that $u_1 = -2$ and $u_4 = -1$. Similar analysis of the earlier tile insertions shows that

$$u_5 = 0, \quad u_3 = 0, \quad u_2 = -2, \quad u_7 = 2, \quad u_6 = 3.$$

We conclude that

$$F(P) = (f; -2, -2, 0, -1, 0, 3, 2).$$

The reader should consult Example 54 to confirm that $-R_j(f) \leq u_j \leq L_j(f)$ for $1 \leq j \leq 7$.

Our goal in the rest of this section is to describe another bijection $G : \mathcal{I}_{n,k,m} \to \mathcal{P}_{n,k,m}$ such that *tstat* maps to *area*. The definition of the new statistic *pmaj* on labelled paths is engineered so that *qstat* maps to *pmaj* under $G$. Indeed, we will use the equation $pmaj(P) = qstat(G^{-1}(P))$ as the definition of *pmaj*. Except in the case $m = 1$ and $k = 0$ considered earlier, the *pmaj* statistic does not seem to have a particularly simple direct definition (not relying on the bijection $G$).

## 5.2 Generalized Parking Functions

It is convenient to introduce the notion of *generalized parking functions*, which give an alternate notation for describing labelled trapezoidal paths. Some combinatorial properties of generalized parking functions were studied by C. Yan in [24, 25].

**Definition 67.** Fix integers $n \geq 1$, $k \geq 0$, and $m \geq 1$.

(1) Let $TZ_{n,k,m}$ denote the region bounded by the lines $x = 0$, $y = 0$, $x = k + my$, and $y = n$. Number the rows of this region 1 to $n$, starting at the bottom. Number the columns in each row of this region $1, 2, 3, \dots$ from left to right. Define

$$B(i) = k + m(i - 1) + 1.$$

Note that a labelled lattice path with $n$ labels stays within the region $TZ_{n,k,m}$ iff the label in row $i$ appears in one of the columns $1, 2, \dots, B(i)$ for $1 \leq i \leq n$.

(2) Given any function $g$ with domain $\{1, 2, \dots, n\}$, set

$$S_j(g) = g^{-1}(j) \text{ and } T_i(g) = g^{-1}(\{1, 2, \dots, i\}) = \bigcup_{j=1}^{i} S_j(g).$$

(3) A *generalized parking function* or *generalized preference function* of *type* $(n, k, m)$ is a function $g : \{1, 2, \dots, n\} \to \{1, 2, \dots, B(n)\}$ such that

$$|T_{B(i)}(g)| \geq i \text{ for } 1 \leq i \leq n.$$

Let $\mathcal{P}'_{n,k,m}$ denote the collection of parking functions of type $(n, k, m)$.

**Lemma 68.** *There exists a bijection $D_0$ between functions $g : \{1, 2, \dots, n\} \to \{1, 2, \dots\}$ and valid labelled lattice paths of height n starting at the origin and ending with a vertical step. This bijection yields a bijection $D : \mathcal{P}'_{n,k,m} \to \mathcal{P}_{n,k,m}$ between generalized parking functions and labelled trapezoidal paths.*

*Proof.* Let $g$ be any function mapping $\{1, 2, \dots, n\}$ into the positive integers. Starting in the bottom row of the region $\{(x, y) : x \geq 0,\ 0 \leq y \leq n\}$, place the elements of $S_1(g)$ in increasing order in the first column of the diagram, one per row. Starting in the next empty row, place the elements of $S_2(g)$ in increasing order in the second column of the diagram, one per row. Continue similarly: after listing all elements $x$ with $g(x) < i$, start in the next empty row and place the elements of $S_i$ in increasing order in column $i$. Finally, draw a lattice path starting at $(0, 0)$ by drawing vertical steps immediately left of each label, and then drawing the necessary horizontal steps to get a connected path. $D_0(g)$ is defined to be the resulting labelled path.

The inverse of $D_0$ is defined as follows. Let $P$ be any valid labelled lattice path of height $n$ starting at the origin and ending with a vertical step. For $1 \leq j \leq n$, define $g(j)$

to be the number of the column in which label $j$ appears. This construction obviously gives an inverse to $D_0$, hence $D_0$ is a bijection.

Now, consider a function $g$ and its associated path $P = D_0(g)$. Note that each row in the diagram of $P$ contains exactly one label. We claim that $|T_x(g)| \geq i$ iff the label $\ell$ in row $i$ of $P$ appears in one of the columns $1, 2, \ldots, x$. We prove the contrapositive of each direction. First, assume that label $\ell$ appears in some column $z > x$. By definition of $D_0$, there can be at most $i-1$ numbers $c$ such that $g(c) < z$. In particular, the size of $T_x(g)$ is at most $i-1$. Conversely, assume that $|T_x(g)| < i$. In the construction of $P$, we will have exhausted all numbers $c$ with $g(c) \leq x$ before reaching row $i$. Thus, the label $\ell$ in row $i$ must satisfy $g(\ell) > x$, so it appears in a column $z > x$. This proves the claim.

Letting $x = B(i)$ in the claim for $1 \leq i \leq n$, we see that $g$ belongs to $\mathcal{P}'_{n,k,m}$ iff the associated path $D_0(g)$ belongs to $\mathcal{P}_{n,k,m}$. (Technically, we may need to add some horizontal steps to the path $D_0(g)$ at the top level $y = n$ to get a path in $\mathcal{P}_{n,k,m}$ ending at $(k + mn, n)$. These extra steps are obviously harmless.) Hence, restricting $D_0$ to the set of generalized parking functions gives the desired bijection $D : \mathcal{P}'_{n,k,m} \to \mathcal{P}_{n,k,m}$. □

From now on, we will identify the set of generalized parking functions $\mathcal{P}'_{n,k,m}$ with the set of labelled trapezoidal paths $\mathcal{P}_{n,k,m}$.

**Example 69.** For the labelled trapezoidal path $P \in \mathcal{P}_{6,2,3}$ shown in Figure 10, the associated parking function $g$ is

$$g(1) = 12,\ g(2) = 17,\ g(3) = 2,$$

$$g(4) = 5,\ g(5) = 2,\ g(6) = 12.$$

**Remark 70.** It is easy to get a recurrence for labelled trapezoidal paths by removing the steps in the first column and their associated labels. If there are $\ell \geq 0$ vertical steps in this column, the associated increasing sequence of labels can be chosen in $\binom{n}{\ell}$ ways. What remains in the upper-right part of the diagram is a labelled trapezoidal path of height $n - \ell$ with the same value of $m$ and a new base length of $k + m\ell - 1$. Setting $P(n, k, m) = |\mathcal{P}_{n,k,m}|$, we obtain the recurrence

$$P(n, k, m) = \sum_{\ell=0}^{n} \binom{n}{\ell} P(n - \ell, k + m\ell - 1, m)$$

with initial conditions

$$P(n, k, m) = 0 \text{ if } n < 0 \text{ or } k < 0;$$

$$P(0, k, m) = 1 \text{ for all } k \geq 0, m \geq 1.$$

From this recurrence, it is easily verified by induction that

$$P(n, k, m) = (k + 1) \cdot (mn + k + 1)^{n-1}.$$

These calculations (and other more general ones) appear in [25].

**Lemma 71.** *Let $P \in \mathcal{P}_{n,k,m}$ correspond to the generalized parking function $g$. Then*

$$area(P) = n(k+1) + mn(n-1)/2 - \sum_{i=1}^{n} g(i). \tag{24}$$

*Proof.* It is easy to see that the region $TZ_{n,k,m}$ contains $nk+mn(n-1)/2$ complete lattice cells. Since label $i$ occurs somewhere in column $g(i)$, there are $g(i)-1$ lattice cells inside the region $TZ_{n,k,m}$ and left of label $i$. These lattice cells lie outside the labelled path associated to $g$. Subtracting, we find that

$$area(P) = kn + mn(n-1)/2 - \sum_{i=1}^{n}[g(i)-1] = n(k+1) + mn(n-1)/2 - \sum_{i=1}^{n} g(i). \tag{25}$$

□

For instance, in the example above we have

$$area(P) = 63 - (12+17+2+5+2+12) = 13.$$

## 5.3 Formal Bounce Paths

Before defining the map $G$, we need to prove a few technical facts about bounce paths. The basic idea is that a bounce path can be constructed from any sequence of vertical moves $v_j$ by using the usual rule to determine the horizontal moves $h_j$.

§1.7 discussed the *bouncing algorithm* that assigns to each trapezoidal path $P \in \mathcal{T}_{n,k,m}$ its associated bounce path $B(P)$. Recall that $B(P)$ consists of a sequence of alternating vertical and horizontal moves, which we will denote here as $v_j(P)$ and $h_j(P)$. Each vertical move $v_j(P)$ was determined from the path $P$ (and the partial bounce path already constructed), while the horizontal move $h_j(P)$ was calculated from the formula

$$h_j(P) = \sum_{i=0}^{m-1} v_{j-i}(P) + \chi(j<k). \tag{26}$$

In the last paragraph, a given path $P \in \mathcal{T}_{n,k,m}$ was used to construct the lists of numbers $v_j(P)$ and $h_j(P)$. Suppose, instead, that we are given only a list of numbers $v_j$ that does not necessarily come from executing the bouncing algorithm on a path $P$. Then we can still create a "formal bounce path" from the list $v_j$ by using a formula like (26) to define numbers $h_i$ in terms of $n$, $k$, $m$, and the $v_j$'s. The precise construction is as follows.

**Definition 72.** Fix integers $n \geq 1$, $k \geq 0$, and $m \geq 1$. Suppose $\{v_j : j \in \mathbb{Z}\}$ is an indexed family of nonnegative integers satisfying the following conditions:

(a) For all $j < 0$, $v_j = 0$.

(b) There exists $j^* \geq 0$ such that $v_{j^*} > 0$ and $v_\ell = 0$ for all $\ell > j^*$.

| $j$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $v_j$ | 3 | 0 | 3 | 1 | 0 |
| $h_j$ | 4 | 4 | 3 | 4 | 1 |

Table 2: The vertical and horizontal moves of a formal bounce path.

(c) $\sum_{j=0}^{j^*} v_j = n$.

We introduce the following notation.

(1) Let $J = \max(j^* + (m-1), k-1)$.

(2) For $0 \leq j \leq J$, let

$$h_j = \sum_{i=0}^{m-1} v_{j-i} + \chi(j < k). \tag{27}$$

(3) For $0 \leq j \leq J$, let

$$H_j = \sum_{i=0}^{j} h_i \text{ and } V_j = \sum_{i=0}^{j} v_i.$$

It will be convenient to set $H_{-1} = h_{-1} = V_{-1} = 0$.

(4) Let $Q = Q(\{v_j\})$ be a path constructed as follows. $Q$ starts at the origin and makes alternating vertical moves and horizontal moves. For $0 \leq j \leq J$, $Q$ moves up $v_j$ units from its current position and then right $h_j$ units. We refer to this move as the "$j$'th bounce." After the $j$'th bounce, $Q$ has reached coordinates $(H_j, V_j)$. $Q$ is called the *formal bounce path associated to the sequence* $\{v_j\}$.

**Example 73.** Let $(n, k, m) = (7, 2, 2)$. Suppose we are given $v_0 = 3$, $v_1 = 0$, $v_2 = 3$, $v_3 = 1$, and $v_j = 0$ for all other $j$. Here, $j^* = 3$ and $J = \max(3+1, 1) = 4$. Table 2 shows the vertical moves and horizontal moves for the formal bounce path $Q(\{v_j\})$.

The path $Q = Q(\{v_j\})$ is shown in Figure 11. Observe that $Q$ happens to lie in the trapezoid $TZ_{n,k,m}$, and $Q$ ends exactly at the upper-right corner $(k+mn, n)$ of this trapezoid. We have $B(Q) = Q$, i.e., the bounce path associated to $Q$ is $Q$ itself. Furthermore, the vertical moves $v_j(Q)$ of this bounce path are precisely the numbers $v_j$ that we were originally given.

**Example 74.** Let $(n, k, m) = (4, 0, 2)$. Suppose we are given $v_0 = 1$, $v_1 = 0$, $v_2 = 0$, $v_3 = 3$, and $v_j = 0$ for all other $j$. Here, $j^* = 3$ and $J = \max(3+1, 0) = 4$. Table 3 shows the vertical moves and horizontal moves for the formal bounce path $Q(\{v_j\})$.

The path $Q = Q(\{v_j\})$ is shown in Figure 12. As before, $Q$ lies in the trapezoid $TZ_{n,k,m}$, and ends exactly at the upper-right corner $(k + mn, n)$ of this trapezoid. The vertical moves $v_j(Q)$ of the bounce path $B(Q)$ are

$$v_0(Q) = 1, \quad v_1(Q) = 0, \quad v_2(Q) = 3, \quad v_3(Q) = 0.$$

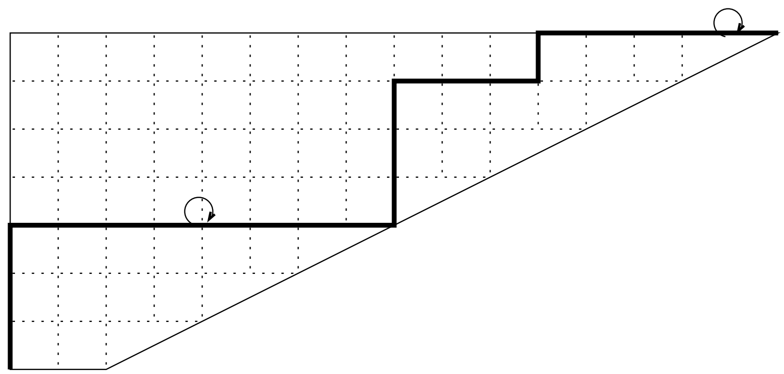


Figure 11: A formal bounce path.

| $j$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $v_j$ | 1 | 0 | 0 | 3 | 0 |
| $h_j$ | 1 | 1 | 0 | 3 | 3 |

Table 3: The vertical and horizontal moves of another formal bounce path.

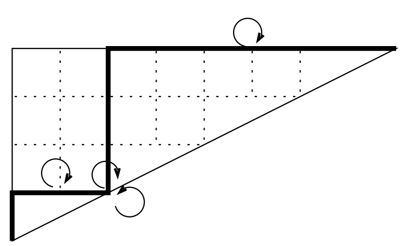


Figure 12: Another formal bounce path.

This is almost the same as the original sequence $v_j$, except that the element $v_2 = 0$ has disappeared. This occurred because the corresponding horizontal move $h_2$ was zero.

The phenomenon in the two examples above is typical. We will show that the path $Q = Q(\{v_j\})$ is always a valid trapezoidal path. Furthermore, under the additional condition that $h_j > 0$ for $0 \le j \le J$, the vertical moves $v_j(Q)$ for the bounce path $B(Q)$ are precisely the original numbers $v_j$. Hence, in this situation, the list $v_j$ can be recovered from the path $Q$ by performing the bouncing algorithm.

**Lemma 75.** *Let $n$, $k$, $m$, and $\{v_j\}$ be given as in Definition 72. Let $h_j$, $V_j$, $H_j$, and $Q = Q(\{v_j\})$ be given as in that definition. Then:*

*(1) For $0 \le j \le J$, we have*

$$H_j = \min(k, j+1) + \sum_{i=0}^{j} \min(m, j+1-i) v_i.$$

*(2) For $0 \le j \le J$, we have*

$$H_j \le k + mV_j,$$

*with equality iff $j \ge k-1$ and $v_{j-i} = 0$ for $0 \le i < m-1$.*

*(3) $Q$ is a path from $(0,0)$ to $(k+mn, n)$ that always stays inside the trapezoid $TZ_{n,k,m}$.*

*(4) The path $Q$ reaches the right edge of the trapezoid $TZ_{n,k,m}$ after the $j$'th bounce iff $j \ge k-1$ and $v_{j-i} = 0$ for $0 \le i < m-1$.*

*(5) Assume that $h_j > 0$ for $0 \le j \le J$. Then $B(Q) = Q$ and $v_j(Q) = v_j$ for $0 \le j \le J$, so that the original list $\{v_j\}$ can be recovered by performing the bouncing algorithm on $Q$.*

*Proof.* To prove (1), use (27) to compute

$$\begin{aligned} H_j &= \sum_{u=0}^{j} h_u = \sum_{u=0}^{j} \left[ \chi(u<k) + \sum_{i=u-(m-1)}^{u} v_i \right] \\ &= \min(k, j+1) + \sum_{i=0}^{j} v_i \sum_{u=0}^{j} \chi(u-(m-1) \le i \le u) \\ &= \min(k, j+1) + \sum_{i=0}^{j} \min(m, j+1-i) v_i. \end{aligned}$$

To justify the last equality, fix $i$ with $0 \le i \le j$. If $0 \le i \le j-(m-1)$, then there are exactly $m$ choices of the index $u$ for which $\chi(u-(m-1) \le i \le u) = 1$, namely $u = i$, $u = i+1$, $\cdots$, $u = i+(m-1)$. In this case, $j+1-i \ge m$, so that $\min(m, j+1-i) = m$. On the other hand, if $j-(m-1) < i \le j$, then there are exactly $j+1-i$ choices of the index $u$ for which $\chi(u-(m-1) \leq i \leq u) = 1$, namely $u = i$, $u = i+1$, $\cdots$, $u = j$. In this case, $j+1-i < m$, so that $\min(m, j+1-i) = j+1-i$.

Now, we use (1) to compute

$$H_j = \min(k, j+1) + \sum_{i=0}^{j} \min(m, j+1-i) v_i \leq k + m \sum_{i=0}^{j} v_i = k + mV_j.$$

Equality is attained here iff $\min(k, j+1) = k$ and $v_i = 0$ for all $i$ such that $\min(m, j+1-i) = j+1-i < m$. In other words, equality is attained iff $j \geq k-1$ and $v_{j-i} = 0$ for $0 \leq i < m-1$, giving (2).

Recall that the right boundary of the trapezoid $TZ_{n,k,m}$ is the line $x = k + my$. $Q$ lies inside this trapezoid iff all the points $(H_j, V_j)$ lie weakly left of this line, for $0 \leq j \leq J$. This is exactly what the inequality in (2) asserts. Also, equality holds in (2) for some $j$ iff $(H_j, V_j)$ lies exactly on the line $x = k + my$. The definition of $J$ guarantees that equality holds in (2) for $j = J$ and that $V_J = n$. Hence, $Q$ ends at the upper-right corner $(k + mn, n)$. We have now proved (3) and (4).

To prove (5), let $0 \leq j \leq J$. We can assume by induction on $j$ that $v_i(Q) = v_i$ and $h_i(Q) = h_i$ for all $i$ with $0 \leq i < j$. In particular, just before the $j$'th bounce, both $Q$ and $B(Q)$ are at coordinates $(H_{j-1}, V_{j-1})$. Where does $Q$ go from here? By definition of $Q$, $Q$ goes up $v_j \geq 0$ units and then over $h_j$ units. Now, by the definition of the bouncing algorithm in §1.7, $B(Q)$ also goes up $v_j$ units, since it is blocked there by a horizontal step of $Q$. This step must exist because of the assumption that $h_j > 0$. Therefore, $v_j(Q) = v_j$. Comparing formulas (26) and (27) and using the induction hypothesis, it is immediate that $h_j(Q) = h_j$ also. This completes the induction. □

## 5.4 The Map $G : \mathcal{I}_{n,k,m} \to \mathcal{P}_{n,k,m}$

We are now ready to define the map $G$ from intermediate objects to generalized parking functions. Let $I = (f; u_1, \ldots, u_n)$ be an intermediate object in $\mathcal{I}_{n,k,m}$. For all integers $j$, let $S_j = f^{-1}(j)$ and $v_j = |f^{-1}(j)| = |S_j|$. Note that, for $0 \leq j \leq k + m(n-1)$, $S_j$ consists of the $v_j$ labels appearing in the $j$'th descending block of the word of $f$. Let $j^*$ be the largest value of $j$ for which $v_j > 0$. It is clear that the sequence $\{v_j : j \in \mathbb{Z}\}$ satisfies conditions (a), (b), and (c) in Definition 72.

We will define $G(I)$ in three steps.

- First, draw the formal bounce path $Q = Q(\{v_j\})$ associated to the sequence $\{v_j\}$.
- Second, attach labels to the path $Q$. Place the $v_j$ labels in $S_j$ in the cells to the right of the $v_j$ vertical segments in the $j$'th vertical move of the path $Q$, in increasing order from bottom to top. Let $g_0$ be the function associated to this labelled lattice path via $D_0^{-1}$.
- Third, define a function $g$ by

$$g(i) = g_0(i) - u_i \text{ for } 1 \leq i \leq n, \tag{28}$$

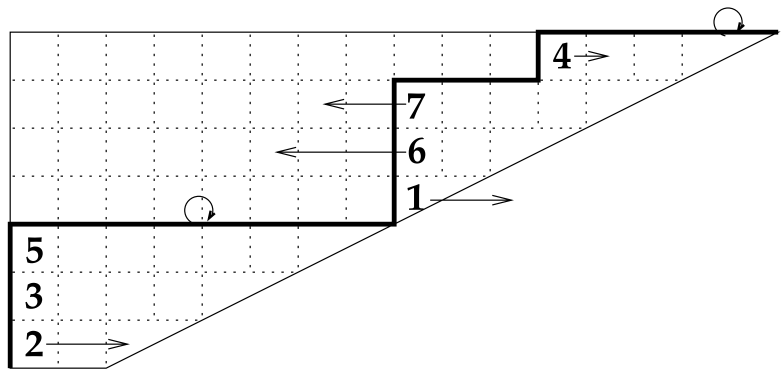



Figure 13: Labelled path diagram for $g_0$.

and set $G(I) = g$.

In terms of labelled paths, the diagram for $g$ is obtained from the diagram of $g_0$ as follows. For each label $i$ in the diagram of $g_0$, move the label $|u_i|$ cells to the right if $u_i < 0$, or $u_i$ cells to the left if $u_i \geq 0$. Then reorder the rows of the diagram to produce a valid configuration of labels (in which labels in a given column increase from bottom to top, and for $r < s$, the labels in column $r$ appear in lower rows than the labels in column $s$). This construction explains why we called $R_j(f)$ and $L_j(f)$ the right and left limits of $j$ relative to $f$.

**Example 76.** Let $(n, k, m) = (7, 2, 2)$, and let $I = (f; -2, -2, 0, -1, 0, 3, 2) \in \mathcal{I}_{n,k,m}$, where

$$f(1) = 2,\ f(2) = 0,\ f(3) = 0,\ f(4) = 3,\ f(5) = 0,\ f(6) = 2,\ f(7) = 2.$$

Let us compute $G(I)$. By looking at $f$, we find that

$$v_0 = 3,\ v_1 = 0,\ v_2 = 3,\ v_3 = 1,\ v_j = 0 \text{ for other } j.$$

The unlabelled path $Q = Q(\{v_j\})$ is shown in Figure 11. The corresponding labelled path is shown in Figure 13. The arrows in this figure show the motion of the labels caused by the numbers $u_j$.

Applying $D_0^{-1}$, we compute

$$g_0(1) = 9, g_0(2) = 1, g_0(3) = 1, g_0(4) = 12, g_0(5) = 1, g_0(6) = 9, g_0(7) = 9.$$

By (28), we get

$$g(1) = 11, g(2) = 3, g(3) = 1, g(4) = 13, g(5) = 1, g(6) = 6, g(7) = 7.$$

The labelled path diagram for $g$ is shown in Figure 14. This figure can be obtained from the previous one by moving labels as indicated by the arrows, and then rearranging the rows as explained above. Note that $g$ and $g_0$ are indeed generalized parking functions of type $(n, k, m)$.

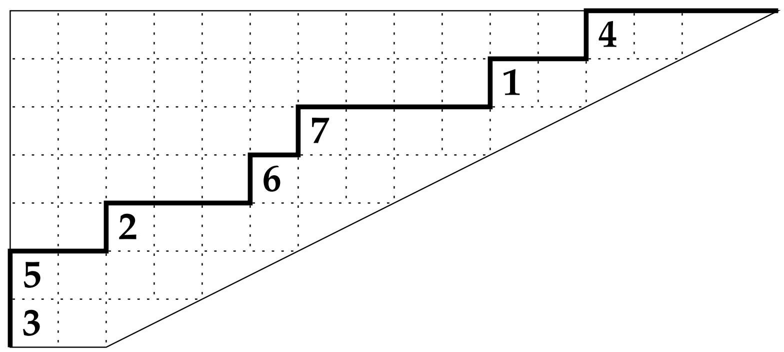



Figure 14: Labelled path diagram for $g$.

**Lemma 77.** *Let $I = (f; u_1, \ldots, u_n) \in \mathcal{I}_{n,k,m}$, and let $Q$ and $g_0$ be computed from $I$ as indicated above.*

*(1) $Q$ is a path from $(0,0)$ to $(k+mn, n)$ that always stays inside $TZ_{n,k,m}$. Hence, $g_0$ is a parking function of type $(n,k,m)$.*

*(2) For $0 \le j \le J$, the horizontal moves $h_j$ of $Q$ satisfy $h_j > 0$.*

*(3) The function $f$ can be uniquely recovered from $g_0$.*

*(4) For $1 \le x \le n$, we have*

$$g_0(x) = 1 + H_{f(x)-1}.$$

*Proof.* Statement (1) is immediate from Lemma 75(3) and Lemma 68.

To prove (2), we suppose that $h_j = 0$ for some $j$ with $0 \le j \le J$ and derive a contradiction. First note that the existence of the object $I = (f; u_1, \ldots, u_n) \in \mathcal{I}_{n,k,m}$ implies that $count(f,i) > 0$ for $1 \le i \le n$, by definition of $u_i$. By (27), the assumption $h_j = 0$ forces $j \ge k$ and

$$v_j = v_{j-1} = \cdots = v_{j-(m-1)} = 0.$$

Assume that $j^* \le j \le J$. Since $v_{j^*} > 0$, the last condition forces $j^* + m \le j \le J$. But we also have $j \ge k$, so that $J \ge \max(j^* + m, k)$. This contradicts the definition of $J$. Therefore, $0 \le j < j^*$. Since $v_{j^*} > 0$, there exists some $\ell > j$ with $v_\ell > 0$. Choose the minimal $\ell$ with this property; note that $\ell > k$ since $j \ge k$, and we have

$$v_{\ell-1} = \cdots = v_{\ell-m} = 0.$$

Next, choose $i$ to be the maximum element of the nonempty set $f^{-1}(\ell)$. Recall that

$$\begin{aligned} count(f,i) \;&=\; \chi(f(i) \le k) + |f^{-1}_{<i}(f(i)-m)| + |f^{-1}_{>i}(f(i))| \\ &\quad + |f^{-1}(\{f(i)-1, \ldots, f(i)-(m-1)\})|. \end{aligned}$$

By our choice of $i$ and $\ell$, we get $count(f,i) = 0$, which is a contradiction.

Now we can prove that $f$ is uniquely recoverable from $g_0$. Given $g_0$, draw the path $Q$ corresponding to $g_0$ and perform the bouncing algorithm to compute the vertical moves $v_j(Q)$. By Lemma 75(5) and part (2), we have $v_j(Q) = v_j$ for $0 \leq j \leq J$. In particular, $v_j(Q) = v_j$ for $0 \leq j \leq j^*$. So, we can recover the numbers $v_j = |f^{-1}(\{j\})|$ from $g_0$. The labels attached to the $j$'th vertical move of $Q$ are the elements of $f^{-1}(j)$, so we can now recover $f$ itself. Of course, it is possible that $v_j = 0$ for some $j$; in this case, $f^{-1}(j) = \emptyset$.

To prove (4), consider the labelled path diagram for $g_0$. In that diagram, all the labels in the set $S_j = f^{-1}(j)$ occur in the column numbered $1 + H_{j-1}$, since the $j$'th vertical move of $Q$ is drawn just to the left of this column, beginning at $(H_{j-1}, V_{j-1})$. By definition of $D_0^{-1}$, we must have $g_0(x) = 1 + H_{j-1}$ whenever $f(x) = j$, i.e.,

$$g_0(x) = 1 + H_{f(x)-1} \text{ for } 1 \leq x \leq n.$$

□

The next lemma shows that $G$ does map into the set $\mathcal{P}_{n,k,m}$.

**Lemma 78.** *For each $I \in \mathcal{I}_{n,k,m}$, $g = G(I)$ is a parking function of type $(n, k, m)$.*

*Proof.* We must check that $g(i) > 0$ for all $i$ and that

$$|T_{B(i)}(g)| \geq i \text{ for } 1 \leq i \leq n.$$

Recall the following definitions:

$$\begin{aligned}
R_j(f) &= |f_{>j}^{-1}(\{f(j), \ldots, f(j) - (m-1)\})| = -x_j(f); \\
L_j(f) &= x_j(f) + count(f, j) - 1; \\
count(f, j) &= \chi(f(j) \leq k) + |f_{<j}^{-1}(f(j) - m)| + |f_{>j}^{-1}(f(j))| \\
&\quad + |f^{-1}(\{f(j) - 1, \ldots, f(j) - (m-1)\})|.
\end{aligned}$$

Comparing these formulas, we see that

$$L_j(f) = |f_{<j}^{-1}(\{f(j) - 1, \ldots, f(j) - m\})| - \chi(f(j) > k). \tag{29}$$

Recall that $g_0$ was constructed from the formal bounce path $Q = Q(\{v_i\})$, where $v_i = |f^{-1}(i)|$ and

$$h_i = v_i + v_{i-1} + \cdots + v_{i-(m-1)} + \chi(i < k) \text{ for } 0 \leq i \leq J. \tag{30}$$

Let $1 \leq j \leq n$. If $f(j) \geq 1$, we may take $i = f(j) - 1$ in (30). Comparing to formula (29), we find that

$$L_j(f) \leq v_{f(j)-1} + \cdots + v_{f(j)-m} \leq h_{f(j)-1}. \tag{31}$$

If $f(j) = 0$, then (31) holds trivially since $h_{-1} = 0$. Thus, (31) holds for all $j$ with $1 \leq j \leq n$.

Now, since $R_j(f) \geq -u_j \geq -L_j(f)$ by definition of $u_j$, we have

$$g(j) = g_0(j) - u_j \geq g_0(j) - L_j(f) = 1 + H_{f(j)-1} - L_j(f) \geq 1.$$

The last inequality follows since $L_j(f) \le h_{f(j)-1} \le H_{f(j)-1}$.

Recall that $B(i) = k + m(i-1) + 1$, so $B(1) \le B(2) \le \cdots$. Let $w_1, w_2, \ldots, w_n$ be the permutation of $1, 2, \ldots, n$ obtained from the word of $f$ by erasing all bar symbols. To check that $|T_{B(i)}(g)| \ge i$ for all $i$, it suffices to show that $g(w_i) \le B(i)$ for all $i$, for this will imply that

$$\{w_1, \ldots, w_i\} \subset T_{B(i)}(g).$$

Fix $i_0 \in \{1, 2, \ldots, n\}$, and set $j_0 = f(w_{i_0})$. By definition of the word of $f$, we can write $i_0 = r + t$, where

$$r = |\{x : f(x) < j_0\}| = V_{j_0-1} \text{ and } t = |\{x : f(x) = j_0 \text{ and } x \ge w_{i_0}\}|.$$

Now consider two cases.

**Case 1:** $H_{j_0-1} \neq k + mV_{j_0-1}$. Let us construct a new formal bounce path $Q'$ from a sequence $\{v'_j : j \in \mathbb{Z}\}$, as follows. Set $v'_j = v_j$ for $0 \le j < j_0$, and set $v'_j = 0$ for all other $j$. Let $Q' = Q(\{v'_j\})$, and let $h'_j$ be the horizontal moves of $Q'$. (Here, $n' = \sum_{j<j_0} v_j$.) Using Lemma 75(2) and the assumption $H_{j_0-1} \neq k + mV_{j_0-1}$, it is easy to see that $J' \ge j_0$. From (27), we have $h'_j = h_j$ for $-1 \le j < j_0$, so $H'_j = H_j$ for $-1 \le j < j_0$. On the other hand,

$$h'_{j_0} = 0 + v_{j_0-1} + \cdots + v_{j_0-(m-1)} + \chi(j_0 < k).$$

Lemma 75(2), applied to $Q'$ with $j = j_0 \le J'$, states that $H'_{j_0} \le k + mV'_{j_0}$. In other words, $H_{j_0-1} + h'_{j_0} \le k + mr$, which yields

$$1 + H_{j_0-1} + h'_{j_0} \le 1 + k + m(r + 1 - 1) = B(r + 1).$$

**Case 2:** $H_{j_0-1} = k + mV_{j_0-1}$. In this case, lemma (75) says that $v_{j_0-1} = \cdots = v_{j_0-(m-1)} = 0$ and $j_0 - 1 \ge k - 1$, hence $j_0 \ge k$. So, if we define $h'_{j_0}$ by the same equation

$$h'_{j_0} = 0 + v_{j_0-1} + \cdots + v_{j_0-(m-1)} + \chi(j_0 < k)$$

used above, then $h'_{j_0} = 0$. So we trivially have $H_{j_0-1} + h'_{j_0} \le k + mr$, and

$$1 + H_{j_0-1} + h'_{j_0} \le B(r + 1)$$

in this case as well.

Recalling that $f(w_{i_0}) = j_0$ and using Lemma 77(4), we now compute

$$g(w_{i_0}) = g_0(w_{i_0}) - u_{w_{i_0}} \le g_0(w_{i_0}) + R_{w_{i_0}}(f) = 1 + H_{j_0-1} + f^{-1}_{>w_{i_0}}(\{j_0, j_0-1, \ldots, j_0-(m-1)\}).$$

Now, using the definition of $t$,

$$\begin{aligned} f^{-1}_{>w_{i_0}}(\{j_0, j_0-1, \ldots, j_0-(m-1)\}) &= t - 1 + f^{-1}_{>w_{i_0}}(\{j_0-1, \ldots, j_0-(m-1)\}) \\ &\le t - 1 + v_{j_0-1} + \cdots + v_{j_0-(m-1)} + \chi(j_0 < k) \\ &= t - 1 + h'_{j_0}, \end{aligned}$$

and so

$$g(w_{i_0}) \le t - 1 + (1 + H_{j_0-1} + h'_{j_0}) \le B(r+1) + (t-1) \le B(r+t) = B(i_0),$$

as desired. ☐

**Lemma 79.** *For each $I = (f; u_1, \ldots, u_n) \in \mathcal{I}_{n,k,m}$, we have $area(G(I)) = tstat(I)$.*

*Proof.* Let $g_0$ and $g = G(I)$ be defined as above. It is shown in §3.3 of [20] that the area of the bounce path corresponding to $g_0$ is precisely $x_0(f)$. Let $C$ denote the constant $n(k+1) + mn(n-1)/2$. Using Lemma 71 for $g$ and for $g_0$, we get

$$\begin{aligned} area(g) &= C - \sum_{i=1}^{n} g(i) = C - \sum_{i=1}^{n} g_0(i) + \sum_{i=1}^{n} u_i \\ &= area(g_0) + \sum_{i=1}^{n} u_i = x_0(f) + \sum_{i=1}^{n} u_i \\ &= tstat(I). \end{aligned}$$

□

## 5.5 The map $G^{-1} : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$

We now define a map $H : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$ that will turn out to be the inverse of $G$. The basic idea is to use a labelled path $P \in \mathcal{P}_{n,k,m}$ to recover the symbols in the word of $f$ (including bars) from left to right. As $w(f)$ is being reconstructed, we obtain partial information about $f$ and $g_0$ that is used to continue the reconstruction process. When the full word has been found, we know $f$ and $g_0$. We then define

$$u_i = g_0(i) - g(i), \tag{32}$$

where $g$ is the parking function corresponding to $P$, and set $H(P) = (f; u_1, \ldots, u_n)$. Comparing (32) to (28) makes it clear that $G \circ H = Id_{\mathcal{P}_{n,k,m}}$. It is less clear that $H \circ G$ is also an identity map, and that $H$ does map into $\mathcal{I}_{n,k,m}$. The former will follow from the latter by a pigeonhole-type argument, as in the $m = 1$ case worked out earlier.

A key observation is the following "prefix property" of $w(f)$: if $j$ is any label, then the quantities

$$x_j(f),\ count(f, j),\ L_j(f),\ R_j(f)$$

can all be computed using only the symbols preceding $j$ (including bars) in the word of $f$. This observation is immediate from the definitions of these quantities and the word of $f$. In particular, here we use the fact that the elements of each set $f^{-1}(i)$ appear in *decreasing* order in $w(f)$.

We now give the detailed definition of $H$. Let $P \in \mathcal{P}_{n,k,m}$ be a given labelled path, and let $g = g(P)$ be the corresponding generalized parking function. We compute $H(P)$ using the algorithm given below. The algorithm uses variables $pw(f)$, $pf$, and $pg_0$ to represent partially reconstructed versions of $w(f)$, $f$, and $g_0$, respectively. The observation in the previous paragraph says that if $j$ occurs in $pw(f)$ at some step, then we can compute $x_j(f) = x_j(pf)$, etc., and the answer obtained is independent of how $pw(f)$ is extended in later steps to $w(f)$. Note that every time we add a label $j$ to $pw(f)$, the definitions of $w(f)$ and $g_0$ allow us to deduce the values of $f(j)$ and $g_0(j)$. Thus, $pf$ and $pg_0$ can now be defined for input $j$. The following example illustrates this process.

**Example 80.** Let $(n, k, m) = (7, 2, 2)$. Suppose that we are given the following partial reconstruction of the word of $f$:

$$pw(f) = 5\ 3\ 2 \mid \mid 7.$$

From this prefix of $w(f)$, we can deduce that $v_0 = 3$, $v_1 = 0$, and $v_2 \geq 1$. We can further deduce that $h_0 = H_0 = 4$, $h_1 = 4$, and $H_1 = 8$. Therefore, regardless of the value of $v_2$ or $h_2$, Lemma 77(4) gives

$$pg_0(5) = 1, pg_0(3) = 1, pg_0(2) = 1, pg_0(7) = 9.$$

Moreover,

$$pf(5) = 0, pf(3) = 0, pf(2) = 0, pf(7) = 2.$$

The domain of definition for $pf$ and $pg_0$ is currently $\{2, 3, 5, 7\}$.

Now suppose that we are told the next symbol in $w(f)$ is 6. Then we can conclude that $v_2 \geq 2$, that $pg_0(6) = 9$, and that $pf(6) = 2$.

Figure 15 gives the algorithm defining $H$. It is obvious from this algorithm and the prefix property of $w(f)$ that the required inequalities

$$-R_j(f) \leq u_j \leq L_j(f)$$

all hold. To see that $H(P)$ does belong to $\mathcal{I}_{n,k,m}$, we must still prove that the algorithm produces a function $f \in \mathcal{F}_{n,k,m}$, i.e., that $0 \leq f(x) \leq k + m(n-1)$ for $1 \leq x \leq n$. We must also show that the algorithm always terminates and that it never encounters the error condition.

**Lemma 81.** *If $P \in \mathcal{P}_{n,k,m}$, then the algorithm defining $H(P)$ never declares an error.*

*Proof.* The proof is by induction on the value of the variable $v$. First, assume $v = 0$. Suppose that the processing of label $x$ causes an error. Note that $pg_0(x) = 1$, since the bounce path starts in the first column. On the other hand, $g(x) \geq 1$, so that $u_x = pg_0(x) - g(x) \leq 0$. But the error occurred because $u_x > L_x(pf)$, where $L_x(pf) = 0$ by (29). This is a contradiction.

Next, assume by induction that $v = j > 0$ and that the algorithm has executed the loop iterations for $v = 0, 1, \ldots, j-1$ with no error. Suppose that the algorithm declares an error in the loop iteration for $v = j$ while processing label $x$. Then, in particular, label $x$ was not added to $pw(f)$ in the previous iteration when $v = j-1$. So, in iteration $j-1$ we must have had $u_x < -R_x(pf)$. Using Lemma 77(4) and expanding the definitions, this says that

$$H_{j-2} + 1 - g(x) < -|pf^{-1}_{>x}(\{j-1, \ldots, j-m\})|. \tag{33}$$

Multiplying by $-1$ and noting that all quantities are integers, we can rewrite this as

$$-H_{j-2} - 1 + g(x) \geq |pf^{-1}_{>x}(\{j-1, \ldots, j-m\})| + 1. \tag{34}$$

**Algorithm for $H$:** Input: a labelled path $P \in \mathcal{P}_{n,k,m}$.

1. Initialize $pw(f)$ be the empty word. Let $pf$ and $pg_0$ be undefined for every input. Initialize a variable $v$ to be 0. Let $g = D^{-1}(P)$ be the parking function associated to the labelled path $P$.

2. While $pw(f)$ does not contain all the labels from 1 to $n$, perform the following steps.

   (a) Loop through all labels $x \in \{1, 2, \ldots, n\}$ that do not yet appear in $pw(f)$, from largest to smallest. For each such label $x$, perform the following step.

   Temporarily assume that the next symbol in $pw(f)$ is $x$. Use this assumption to compute $pg_0(x)$ using the formal bouncing rules. Set $u_x = pg_0(x) - g(x)$.

   - If $u_x > L_x(pf)$, declare an error condition and abort the algorithm.
   - If $u_x < -R_x(pf)$, assert that $f(x) > v$ and discard the assumption that the next symbol in $pw(f)$ is $x$. At this point, reset $pg_0(x)$ and $u_x$ to be undefined again.
   - Otherwise, we must have $-R_x(pf) \leq u_x \leq L_x(pf)$. In this case, assert that $f(x) = v$. Retain the assumption that the next symbol in $pw(f)$ is $x$, and retain the values of $pg_0(x)$ and $u_x$ already computed. Define $pf(x) = v$.

   (b) We have now (greedily) determined all values $x$ for which $f(x) = v$. Append a bar symbol at the end of $pw(f)$, and increment $v$ by 1.

3. Add the appropriate number of trailing bar symbols to the end of $pw(f)$, so that the total number of bars is $k + m(n-1) + 1$. We now know $w(f)$ and $f$ and $g_0$. The output of the algorithm is the object

$$H(P) = (f; u_1, \ldots, u_n),$$

where $u_i = g_0(i) - g(i)$ for $1 \leq i \leq n$.

Figure 15: Definition of $H$.

Next, the assumption that $x$ caused an error in iteration $j$ means that $u_x > L_x(pf)$ in iteration $j$. Translating the definitions gives

$$H_{j-1} + 1 - g(x) > |pf_{<x}^{-1}(\{j-1, \ldots, j-m\})| - 1 + \chi(j \leq k). \tag{35}$$

Adding the last two inequalities, we conclude that

$$h_{j-1} > |pf^{-1}(\{j-1, \ldots, j-m\})| + \chi(j < k+1). \tag{36}$$

(To justify the simplification of the right side, observe that we cannot have $pf(x) \in \{j-1, \ldots, j-m\}$; otherwise the algorithm would not be considering label $x$ during iteration $j$.) But, on the other hand, the definition of the bounce path gives

$$h_{j-1} = |pf^{-1}(\{j-1, \ldots, j-m\})| + \chi(j-1 < k), \tag{37}$$

which contradicts the preceding inequality and completes the induction proof. Note that the prefix property of the word of $f$ is needed to ensure that certain quantities appearing in the equations above do not change from one iteration to the next. ☐

**Lemma 82.** *Let $P \in \mathcal{P}_{n,k,m}$. Suppose that, at some stage of the algorithm defining $H(P)$, $pw(f)$ contains exactly $i$ numbers, where $0 \leq i \leq n-1$. Then $pw(f)$ contains at most $k + mi$ bar symbols.*

*Proof.* We use induction on the length $\ell$ of $pw(f)$. The result obviously holds when $pw(f)$ is the empty word. Suppose that the result holds when $pw(f)$ has length $\ell \geq 0$. Let $pw(f)$ have $i$ numbers and $b$ bar symbols. If $b < k + mi$ and the algorithm appends a number next, then the result still holds since $b < k + m(i+1)$. If $b < k + mi$ and the algorithm appends a bar symbol next, then the result still holds since $b + 1 \leq k + mi$. We are reduced to the case where $b = k + mi$. It suffices to show that, in this case, the next symbol appended by step 2 of the algorithm will be a number, not a bar.

To prove this, we establish a number of claims.

Claim 1: If $i > 0$ and $y$ is the rightmost label in $pw(f)$, then $pw(f)$ has at least $m$ bar symbols following $y$. Proof: If, instead, there were $s < m$ bar symbols after $y$, consider the prefix $p'$ with $y$ and these $s$ bar symbols erased. This shorter prefix has $i - 1 \geq 0$ numbers and $k + mi - s > k + m(i-1)$ bar symbols in it, which contradicts the induction hypothesis.

For the next few claims, assume $x$ is a label not already appearing in $pw(f)$. Let us tentatively append $x$ to $pw(f)$ to obtain a new partial word $pw'(f)$, as in step 2 of the algorithm defining $H$. Then $pf(x) = b$, since there are $b$ bars preceding $x$ in the word of $f$.

Claim 2: $V_{b-1} = i$. Recall that $V_{b-1} = |\{x : pf(x) \leq b-1\}|$. The claim is clear when $i = 0$, since $x$ is the first number in $pw'(f)$ and $pf(x) = b > b - 1$. If $i > 0$, claim 1 shows that $pw(f)$ ends in a bar symbol. So, the $i - 1$ numbers $y$ preceding $x$ in $pw'(f)$ must satisfy $pf(y) \leq b - 1$. The numbers following $x$ (and $x$ itself) have function values at least $b$, so claim 2 follows.

Claim 3: $b-1 \geq k-1$ and $v_{b-1-u} = 0$ for $0 \leq u < m-1$. We have $b-1 = k+mi-1 \geq k-1$. If $i = 0$, so that $x$ is the first number in $pw'(f)$, then we certainly have $v_{b-1-u} = 0$ for all $u \geq 0$. If $i > 0$, claim 1 shows that $pw(f)$ ends in $m$ (or more) bar symbols. It again follows that $v_{b-1-u} = 0$ for $0 \leq u < m-1$.

Claim 4: $pg_0(x) = b+1$. Recalling that $pf(x) = b$, Lemma 77(4) gives $pg_0(x) = 1 + H_{b-1}$. Next, Lemma 75(2) and claim 3 show that $H_{b-1} = k + mV_{b-1}$. Combining this with claim 2, we get $pg_0(x) = 1 + k + mi = b + 1$.

Now we can prove the earlier assertion that the next symbol appended to $pw(f)$ by the algorithm will be a number, not a bar. By claim 1, the last symbol (if any) generated by the algorithm was a bar symbol. So, without loss of generality, we can assume the algorithm is at the beginning of step 2(a).

To get a contradiction, suppose that all labels considered in this iteration of step 2(a) are rejected. This happens iff $u_x < -R_x(pf)$ for all unused labels $x$. Now, by definition of $\mathcal{P}_{n,k,m}$, $|T_{B(i+1)}(g)| \geq i+1$. So there exist at least $i+1$ labels $x \in \{1, \ldots, n\}$ such that

$$g(x) \leq B(i+1) = 1 + k + mi = b + 1.$$

Choose such an $x$ that does not already appear in $pw(f)$. Consider what happens when step 2(a) tentatively appends this $x$ to $pw(f)$ to give $pw'(f)$. We have $pg_0(x) = b+1$ by claim 4, and so

$$u_x = pg_0(x) - g(x) \geq b + 1 - (b+1) = 0.$$

But the assumption that $x$ was rejected means that

$$u_x < -R_x(pf) \leq 0.$$

We obtain the contradiction $u_x < 0$ and $u_x \geq 0$. ☐

**Corollary 83.** *Let $P \in \mathcal{P}_{n,k,m}$.*

*(1) When executing the algorithm defining $H(P)$, all $n$ labels in $\{1, 2, \ldots, n\}$ are eventually added to $pw(f)$. Consequently, the algorithm always terminates.*

*(2) If $f$ is the function produced by the algorithm defining $H(P)$, then*

$$f(x) \in \{0, 1, \ldots, k + m(n-1)\} \text{ for } 1 \leq x \leq n.$$

*Consequently, $H$ is a well-defined map from $\mathcal{P}_{n,k,m}$ to $\mathcal{I}_{n,k,m}$.*

*Proof.* To prove (1), suppose that the algorithm only adds $i < n$ labels to $pw(f)$. After the $i$'th label is appended, each subsequent iteration of step 2 of the algorithm will add one more bar symbol to $pw(f)$. Eventually, there will be more than $k + mi$ bar symbols, contradicting the previous lemma. Thus, all $n$ labels are eventually added to $pw(f)$, at which point the algorithm exits the loop in step 2 and terminates after step 3.

To prove (2), consider the value of $pw(f)$ just before the $n$'th label $x$ is appended to it. This prefix of $w(f)$ contains $i = n-1$ labels. By the lemma, the number of bars in $pw(f)$ is at most $k + m(n-1)$. Since $f(x)$ is always the number of bars preceding $x$ in $w(f)$, we have $f(x) \leq k+m(n-1)$. For the same reason, we have $f(y) \leq k+m(n-1)$ for all labels $y$ preceding $x$ in $w(f)$. So, the image of $f$ is contained in $\{0,1,\ldots,k+m(n-1)\}$. This also shows, incidentally, that step 3 of the algorithm defining $H$ makes sense. We observed earlier that the numbers $u_j$ produced by the algorithm satisfy the required inequalities. Hence, we finally conclude that $H$ is a well-defined map from $\mathcal{P}_{n,k,m}$ to $\mathcal{I}_{n,k,m}$. □

As remarked earlier, it is clear that $G \circ H = Id_{\mathcal{P}_{n,k,m}}$. Our final theorem says that $H$ is the two-sided inverse for $G$.

**Theorem 84.** *The maps $G : \mathcal{I}_{n,k,m} \to \mathcal{P}_{n,k,m}$ and $H : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$ are bijections with $H = G^{-1}$. For $P \in \mathcal{P}_{n,k,m}$, define $pmaj(P) = qstat(H(P))$. Then:*

$$pmaj(P) = qstat(H(P)) \text{ and } area(P) = tstat(H(P)); \tag{38}$$

$$qstat(I) = pmaj(G(I)) \text{ and } tstat(I) = area(G(I)). \tag{39}$$

*Consequently,*

$$\sum_{P\in\mathcal{P}_{n,k,m}} q^{pmaj(P)}t^{area(P)} = \sum_{I\in\mathcal{I}_{n,k,m}} q^{qstat(I)}t^{tstat(I)} = \sum_{P\in\mathcal{P}_{n,k,m}} q^{area(P)}t^{h(P)} = CH_{n,k,m}(q,t), \tag{40}$$

*and so all these statistics have the same univariate distribution.*

*Proof.* We have already shown that $G$ maps into $\mathcal{P}_{n,k,m}$, $H$ maps into $\mathcal{I}_{n,k,m}$, and $G \circ H = Id_{\mathcal{P}_{n,k,m}}$. The last equation implies that $H$ is an injection and $G$ is a surjection. But Corollary 65 showed that

$$|\mathcal{I}_{n,k,m}| = |\mathcal{P}_{n,k,m}| < \infty.$$

Since the sets are finite, $H$ is automatically a surjection, $G$ is automatically an injection, and $H = G^{-1}$. The properties in (39) follow from Lemma 79 and the very definition of $pmaj$, and (38) follows by replacing $I$ by $H(P)$ and simplifying. The equalities in (40) follow from the existence of the weight-preserving bijections $G$ and $F$. Letting $q = 1$ or $t = 1$ in (40) gives the final assertion of the theorem. □

**Acknowledgement:** The authors would like to thank James Haglund for many helpful discussions regarding the material presented here.

# References

[1] F. Bergeron and A. Garsia, "Science Fiction and Macdonald Polynomials," *CRM Proceedings and Lecture Notes AMS* **6** (1999), 363—429.

[2] F. Bergeron, N. Bergeron, A. Garsia, M. Haiman, and G. Tesler, "Lattice Diagram Polynomials and Extended Pieri Rules," *Adv. in Math.* **2** (1999), 244—334.

[3] F. Bergeron, A. Garsia, M. Haiman, and G. Tesler, "Identities and Positivity Conjectures for some remarkable Operators in the Theory of Symmetric Functions," *Methods and Applications of Analysis* **7** (1999), 363—420.

[4] L. Carlitz and J. Riordan, "Two element lattice permutation numbers and their $q$-generalization," *Duke Math. J.* **31** (1964), 371—388.

[5] D. Foata and J. Riordan, "Mappings of acyclic and parking functions," *Aequationes Math.* **10** (1974), 10—22.

[6] J. Françon, "Acyclic and parking functions," *J. Combinatorial Theory Series A* **18** (1975), 27—35.

[7] A. Garsia and J. Haglund, "A proof of the $q,t$-Catalan positivity conjecture," *Discrete Math.* **256** (2002), 677—717.

[8] A. Garsia and J. Haglund, "A positivity result in the theory of Macdonald polynomials," *Proc. Nat. Acad. Sci.* **98** (2001), 4313—4316.

[9] A. Garsia and M. Haiman, "A remarkable $q,t$-Catalan sequence and $q$-Lagrange Inversion," *J. Algebraic Combinatorics* **5** (1996), 191—244.

[10] J. Haglund, "Conjectured Statistics for the $q,t$-Catalan numbers," *Adv. in Math.* **175** (2003), 319—334.

[11] J. Haglund and N. Loehr, "A Conjectured Combinatorial Formula for the Hilbert Series for Diagonal Harmonics," to appear in *Discrete Mathematics.*

[12] Mark Haiman. Personal communication.

[13] M. Haiman, "Notes on Macdonald Polynomials and the Geometry of Hilbert Schemes," *Symmetric Functions 2001: Surveys of Developments and Perspectives,* Proceedings of the NATO Advaced Study Institute. Sergey Fomin, ed. Kluwer, Dordrecht (2002), 1—64.

[14] M. Haiman, "Combinatorics, symmetric functions, and Hilbert schemes," *CDM 2002: Current Developments in Mathematics*, Intl. Press Books (2003), 39—112.

[15] M. Haiman, "Hilbert schemes, polygraphs, and the Macdonald positivity conjecture." *J. Amer. Math. Soc.* **14** (2001), 941—1006.

[16] M. Haiman. "Vanishing theorems and character formulas for the Hilbert scheme of points in the plane," *Invent. Math.* **149** (2002), 371—407.

[17] A. Konheim and B. Weiss, "An occupancy discipline and applications," *SIAM J. Applied Math.* **14** (1966), 1266—1274.

[18] N. Loehr, "Conjectured Statistics for the Higher $q,t$-Catalan Sequences." Preprint, 2002.

[19] N. Loehr, *Multivariate Analogues of Catalan Numbers, Parking Functions, and their Extensions*. Ph.D. thesis, University of California at San Diego, June 2003.

[20] N. Loehr, "Trapezoidal lattice paths and multivariate analogues," *Adv. in Appl. Math.* **31** (2003), 597—629.

[21] I. G. Macdonald, *Symmetric Functions and Hall Polynomials.* 2nd ed. Oxford University Press, 1995.

[22] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions.* Wadsworth and Brooks/Cole, 1991.

[23] M. P. Schützenberger, "On an enumeration problem," *J. Combinatorial Theory* **4** (1968), 219—221.

[24] C. Yan, "Generalized Parking Functions, Tree Inversions and Multicolored Graphs." *Advances in Applied Mathematics* **27** (2001), 641—670.

[25] C. Yan, "On the Enumeration of Generalized Parking Functions." *Proceedings of the 31st Southeastern Intl. Conference on Combinatorics, Graph Theory, and Computing. Congressus Numerantium* **147** (2000), 201—209.