Conjectured Statistics for the Higher q, t-Catalan Sequences

Nicholas A. Loehr*
Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104
nloehr@math.upenn.edu

Submitted: Oct 22, 2002; Accepted: Jan 24, 2005; Published: Feb 14, 2005
Mathematics Subject Classifications: 05A10, 05E05, 05E10, 20C30, 11B65

Abstract

This article describes conjectured combinatorial interpretations for the higher q, t-Catalan sequences introduced by Garsia and Haiman, which arise in the theory of symmetric functions and Macdonald polynomials. We define new combinatorial statistics generalizing those proposed by Haglund and Haiman for the original q, t-Catalan sequence. We prove explicit summation formulas, bijections, and recursions involving the new statistics. We show that specializations of the combinatorial sequences obtained by setting $t = 1$ or $q = 1$ or $t = 1/q$ agree with the corresponding specializations of the Garsia-Haiman sequences. A third statistic occurs naturally in the combinatorial setting, leading to the introduction of q, t, r-Catalan sequences. Similar combinatorial results are proved for these trivariate sequences.

1 Introduction

In [7], Garsia and Haiman introduced a q, t-analogue of the Catalan numbers, which they called the q, t-Catalan sequence. In the same paper, they introduced a whole family of "higher" q, t-Catalan sequences, one for each positive integer m. We begin by describing several equivalent characterizations of the original q, t-Catalan sequence. We then discuss analogous characterizations of the higher q, t-Catalan sequences.

In the rest of the paper, we present some conjectured combinatorial interpretations for the higher q, t-Catalan sequences. We prove some combinatorial formulas, recursions, and

*Work supported by an NSF Graduate Research Fellowship and an NSF Postdoctoral Research Fellowship

bijections and introduce a three-variable version of the Catalan sequences. We also show that certain specializations of our combinatorial sequences agree with the corresponding specializations of the higher q, t-Catalan sequences.

1.1 The Original q, t-Catalan Sequence

To give Garsia and Haiman's original definition of the q, t-Catalan sequence, we first need to review some standard terminology associated with integer partitions. A *partition* is a sequence $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k)$ of weakly decreasing positive integers, called the *parts* of λ. The integer $N = \lambda_1 + \lambda_2 + \cdots + \lambda_k$ is called the *area of* λ and denoted $|\lambda|$. In this case, λ is said to be a *partition of* N, and we write $\lambda \vdash N$. The number of parts k is called the *length of* λ and denoted $\ell(\lambda)$. We often depict a partition λ by its *Ferrers diagram.* This diagram consists of k left-justified rows of boxes (called *cells*), where the i'th row from the top has exactly λ_i boxes. Figure 1 shows the Ferrers diagram of $\lambda = (8, 7, 5, 4, 4, 2, 1, 1)$, which is a partition of 32 having eight parts.

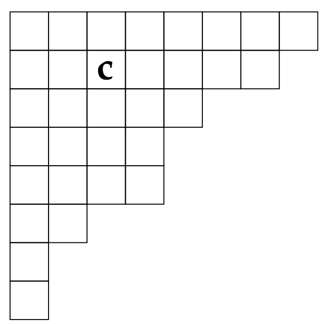


Figure 1: Diagram of a partition.

Let λ be a partition of N. Let c be one of the N cells in the diagram of λ. We make the following definitions.

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

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

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

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

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

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

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

Fix a positive integer n and a partition μ of n. Let μ' denote the transpose of μ, obtained by interchanging the rows and columns of μ. Define the following abbreviations:

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

In all but the last formula above, the sums and products range over all cells in the diagram of μ. In the product defining $\Pi_\mu(q,t)$, the northwest corner cell of μ 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.

Finally, we define the *original* q,t*-Catalan sequence* to be the following sequence of rational functions in the variables q and t:

$$OC_n(q,t) = \sum_{\mu\vdash n} \frac{t^{2n(\mu)}q^{2n(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)} \quad (n = 1,2,3,\ldots). \tag{1}$$

It turns out that, for all n, $OC_n(q,t)$ is a polynomial in q and t with nonnegative integer coefficients. But this fact is very difficult to prove. See Theorem 1 below.

1.2 Symmetric Function Version of the q,t-Catalan Sequence

This section assumes familiarity with basic symmetric function theory, including Macdonald polynomials. We begin by briefly recalling the definition of the modified Macdonald polynomials and the nabla operator.

Let Λ denote the ring of symmetric functions in the variables $x_1, \ldots, x_n, \ldots$ with coefficients in the field $K = \mathbb{Q}(q,t)$. Let α denote the unique automorphism of the ring Λ that interchanges q and t. Let ϕ denote the unique K-algebra endomorphism of Λ that sends the power-sum symmetric function p_k to $(1-q^k)p_k$. Let $\geq$ denote the usual dominance partial ordering on partitions. Then the *modified Macdonald basis* is the unique basis $\tilde{H}_\mu$ of Λ (indexed by partitions μ) such that:

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

The *nabla operator* is the unique linear operator on Λ defined on the basis $\tilde{H}_\mu$ by the formula

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

(The nabla operator was introduced by F. Bergeron and A. Garsia in [2]. See also [3] or [4] for more information about nabla).

Now, we define the *symmetric function version of the q, t-Catalan sequence* by the formula

$$SC_n(q,t) = \nabla(e_n)|_{s_{1^n}} \quad (n = 1, 2, 3, \ldots), \tag{2}$$

where e_n is an elementary symmetric function, s_{1^n} is a Schur function, and the vertical bar indicates extraction of a coefficient. In more detail, to calculate $SC_n(q,t)$, start with the elementary symmetric function e_n (regarded as an element of the K-vector space Λ), and perform the following steps:

1. Find the unique expansion of the vector e_n as a linear combination of the modified Macdonald basis elements $\tilde{H}_\mu$. The scalars appearing in this expansion are elements of $K = \mathbb{Q}(q,t)$.

2. Apply the nabla operator to this expansion by multiplying the coefficient of $\tilde{H}_\mu$ by $q^{n(\mu')} t^{n(\mu)}$, for every μ.

3. Express the resulting vector as a linear combination of the Schur function basis s_μ.

4. Extract the coefficient of s_{1^n} in this new expansion. This coefficient (an element of $\mathbb{Q}(q,t)$) is $SC_n(q,t)$.

1.3 The Representation-Theoretical q, t-Catalan Sequence

This section assumes familiarity with representation theory of the symmetric groups. Let $R_n = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be a polynomial ring over $\mathbb{C}$ in two independent sets of n variables. Let the symmetric group S_n act on the variables by

$$\sigma(x_i) = x_{\sigma(i)} \text{ and } \sigma(y_i) = y_{\sigma(i)} \text{ for } \sigma \in S_n.$$

Extending this action by linearity and multiplicativity, we obtain an action of S_n on R_n which is called the *diagonal action*. This action turns the vector space R_n into an S_n-module. We define a submodule DH_n of R_n, called the space of *diagonal harmonics*, as follows. A polynomial $f \in R_n$ belongs to DH_n iff f simultaneously solves the partial differential equations

$$\sum_{i=1}^{n} \frac{\partial^h}{\partial x_i^h} \frac{\partial^k}{\partial y_i^k} f = 0,$$

for all integers h, k with $1 \leq h+k \leq n$.

Let $R_{h,k}$ consist of polynomials in DH_n that are homogeneous of degree h in the x_i's, and homogeneous of degree k in the y_i's, together with the zero polynomial. Then each $R_{h,k}$ is a finite-dimensional submodule of DH_n, and we have

$$DH_n = \bigoplus_{h \geq 0} \bigoplus_{k \geq 0} R_{h,k}.$$

Thus, DH_n is a bigraded S_n-module.

Suppose we decompose each $R_{h,k}$ into a direct sum of irreducible modules (which correspond to the irreducible characters of S_n). Let $a_{h,k}(n)$ be the number of occurrences of the module corresponding to the sign character χ_{1^n} in $R_{h,k}$. Then we define the *representation-theoretical q,t-Catalan sequence* by

$$RC_n(q,t) = \sum_{h \geq 0} \sum_{k \geq 0} a_{h,k}(n) q^h t^k \quad (n = 1, 2, 3, \ldots).$$

Thus, $RC_n(q,t)$ is the generating function for occurrences of the sign character in DH_n. By the symmetry of x_i and y_i in the definition, we see that $RC_n(q,t) = RC_n(t,q)$.

1.4 The Two Combinatorial q,t-Catalan Sequences

We next present a combinatorial construction due to Haglund, and a related construction found later by Haiman, which interpret the q,t-Catalan sequence as a weighted sum of Dyck paths.

A *Dyck path of height n* is a path in the xy-plane from $(0,0)$ to (n,n) consisting of n north steps and n east steps (each of length one), such that the path never goes strictly below the diagonal line $y = x$. See Figure 2 for an example. Let $\mathcal{D}_n$ denote the collection of Dyck paths of height n. For $D \in \mathcal{D}_n$, let $\mathrm{area}(D)$ be the number of complete lattice squares (or *cells*) between the path D and the main diagonal.

For $0 \leq i < n$, define $\gamma_i(D)$ to be the number of cells between the path and the main diagonal in the i'th row of the picture, where we let the bottom row be row zero. Thus, $\mathrm{area}(D) = \sum_{i=0}^{n-1} \gamma_i(D)$. Following Haiman, we set

$$\mathrm{dinv}(D) = \sum_{i<j} \left[\chi(\gamma_i(D) = \gamma_j(D)) + \chi(\gamma_i(D) = \gamma_j(D) + 1) \right]. \tag{3}$$

Here and below, we set $\chi(A) = 1$ if A is a true statement, $\chi(A) = 0$ if A is a false statement.

Define *Haiman's combinatorial q,t-Catalan sequence* to be

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

Next, following Haglund (see [9]), we define a "bounce" statistic for each Dyck path D. Given D, we define a *bounce path* derived from D as follows. The bounce path begins

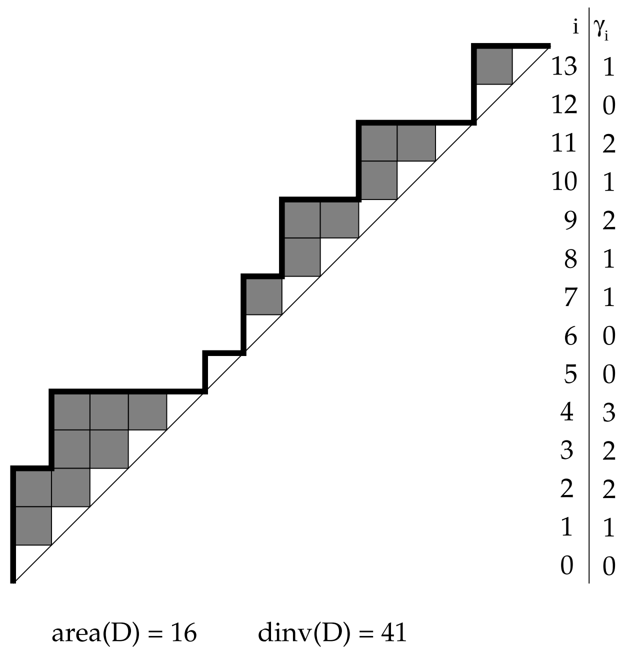


Figure 2: A Dyck path.

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 hits the Dyck path, then due south until it hits the main diagonal. The bounce path terminates when it reaches $(0, 0)$. See Figure 3 for an example.

Suppose the bounce path derived from D hits the main diagonal at the points

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

Then Haglund's bounce statistic is defined by

$$\text{bounce}(D) = \sum_{k=1}^{s} i_k.$$

We define *Haglund's combinatorial q, t-Catalan sequence* by

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

1.5 Equivalence of the q, t-Catalan Sequences

The five q, t-Catalan sequences discussed in the preceding sections have quite different definitions. In spite of this, we have the following theorem.

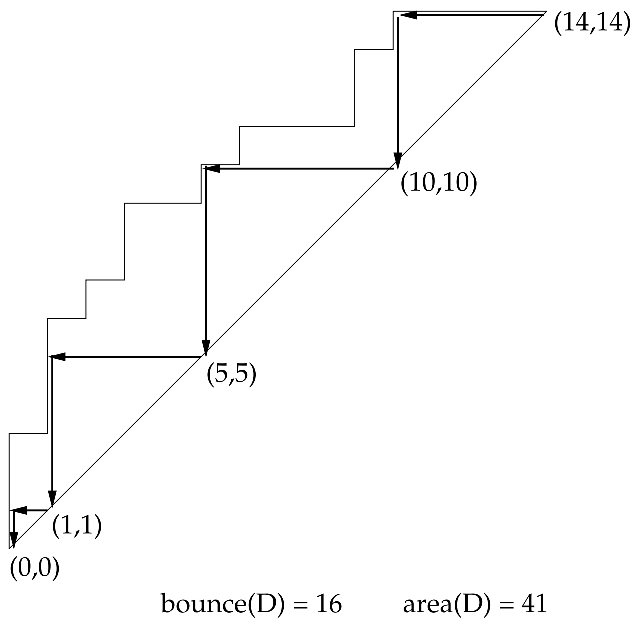


Figure 3: A Dyck path with its derived bounce path.

Theorem 1. For every positive integer n,

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

In particular, $OC_n(q,t)$ is a polynomial in q and t with nonnegative integer coefficients for all n.

This theorem was proved in various papers of Garsia, Haiman, and Haglund. In [7], Garsia and Haiman proved that $SC_n(q,t) = OC_n(q,t)$ using symmetric function identities. Haglund discovered the combinatorial sequence $C_n(q,t)$ (see [9]), and Haiman proposed his version $HC_n(q,t)$ shortly thereafter. Haiman and Haglund easily proved that $HC_n(q,t) = C_n(q,t)$ by showing that both satisfy the same recursion. We discuss this recursion later (§3). Similarly, Garsia and Haglund proved in [5, 6] that $C_n(q,t) = SC_n(q,t)$ by showing that both sequences satisfied the same recursion. This proof is much more difficult and requires substantial machinery from symmetric function theory. Finally, Haiman proved that $RC_n(q,t) = SC_n(q,t)$ using sophisticated algebraic geometric methods (see [16]).

A consequence of Theorem 1 is that $C_n(q,t) = C_n(t,q)$ for all n, since this symmetry property holds for RC_n. (It is also easily deduced from the formula for OC_n, by replacing the summation index μ by the conjugate of μ and simplifying.) An open question is to give a combinatorial proof that $C_n(q,t) = C_n(t,q)$. Later, we give bijections proving the weaker result that $C_n(q,1) = C_n(1,q) = HC_n(q,1) = HC_n(1,q)$. This says that the new statistics of Haiman and Haglund have the same univariate distribution as the area statistic on Dyck paths.

1.6 The Higher q,t-Catalan Sequences

We now discuss various descriptions of the higher q,t-Catalan sequences, also introduced by Garsia and Haiman in [7]. Fix a positive integer m. The *original higher q,t-Catalan sequence of order m* is defined by

$$OC_n^{(m)}(q,t) = \sum_{\mu\vdash n} \frac{t^{(m+1)n(\mu)}q^{(m+1)n(\mu')}(1-t)(1-q)\Pi_\mu(q,t)B_\mu(q,t)}{h_\mu(q,t)h'_\mu(q,t)} \quad (n=1,2,3,\ldots). \tag{4}$$

This formula is the same as (1), except that the factors $t^{2n(\mu)}q^{2n(\mu')}$ in $OC_n(q,t)$ have been replaced by $t^{(m+1)n(\mu)}q^{(m+1)n(\mu')}$. Clearly, $OC_n^{(1)}(q,t) = OC_n(q,t)$.

Next, the *symmetric function version* of the higher q,t-Catalan sequence of order m is defined by

$$SC_n^{(m)}(q,t) = \nabla^m(e_n)|_{s_{1^n}} \quad (n=1,2,3,\ldots), \tag{5}$$

where ∇^m means apply the nabla operator m times in succession. To calculate $SC_n^{(m)}(q,t)$ for a particular m and n, one should express e_n as a linear combination of the modified Macdonald basis elements $\tilde{H}_\mu$, multiply the coefficient of each $\tilde{H}_\mu$ by $t^{mn(\mu)}q^{mn(\mu')}$, express the result in terms of the Schur basis $\{s_\mu\}$, and extract the coefficient of s_{1^n}. Garsia and Haiman proved in [7] that $OC_n^{(m)}(q,t) = SC_n^{(m)}(q,t)$ using symmetric function identities.

A possible representation-theoretical version of the higher q,t-Catalan sequences is given in [7]; we will not discuss it here.

A problem mentioned but not solved in [7] is to give a combinatorial interpretation for the sequences $OC_n^{(m)}(q,t)$. That paper does give a simple interpretation for $OC_n^{(m)}(q,1)$, which we now describe. Given positive integers m and n, let us define an *m-Dyck path of height n* to be a path in the xy-plane from $(0,0)$ to (mn,n) consisting of n north steps and mn east steps (each of length one), such that the path never goes strictly below the slanted line $x=my$. See Figure 4 for an example with $m=3$ and $n=8$. Let $\mathcal{D}_n^{(m)}$ denote the collection of m-Dyck paths of height n. For $D\in\mathcal{D}_n^{(m)}$, let area(D) be the number of complete lattice squares strictly between the path D and the line $x=my$. For instance, area$(D)=23$ for the path D shown in Figure 4.

We then have (see [7])

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

2 Conjectured Combinatorial Interpretations for the Higher q,t-Catalan Sequences

Fix a positive integer m. We next describe two statistics defined on m-Dyck paths that each have the same distribution as the area statistic. The first statistic generalizes Haiman's statistic for Dyck paths; the second statistic generalizes Haglund's bounce statistic. We conjecture that either statistic, when paired with area and summed over m-Dyck paths of height n, will give a generating function that equals $OC_n^{(m)}(q,t)$.

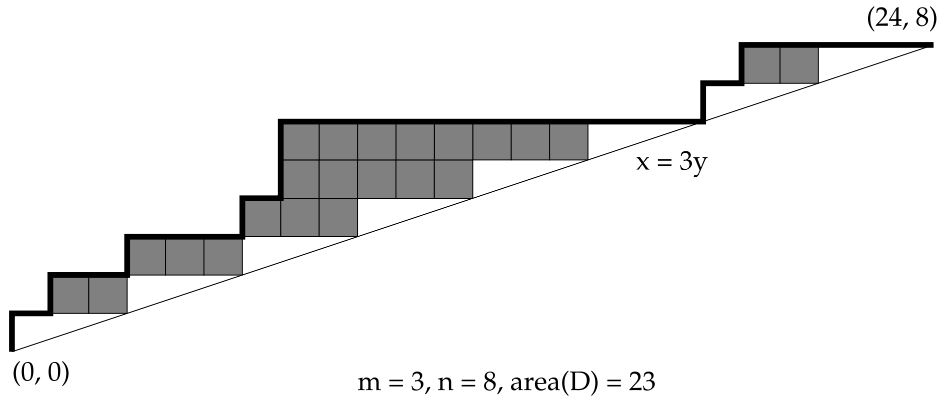


Figure 4: A 3-Dyck path of height 8.

2.1 A Version of Haiman's Statistic for m-Dyck Paths

The statistic discussed here was derived from a statistic communicated to the author by M. Haiman [15]. Let $D \in \mathcal{D}_n^{(m)}$ be an m-Dyck path of height n. As in §1.4, we define $\gamma_i(D)$ to be the number of cells in the i'th row that are completely contained in the region between the path D and the diagonal $x = my$, for $0 \leq i < n$. Here, the lowest row is row zero. Note that $\mathrm{area}(D) = \sum_{i=0}^{n-1} \gamma_i(D)$. Next, define a statistic $h(D)$ by

$$h(D) = \sum_{0 \leq i < j < n} \sum_{k=0}^{m-1} \chi\left(\gamma_i(D) - \gamma_j(D) + k \in \{0, 1, \ldots, m\}\right). \tag{6}$$

See Figure 5 for an example.

It is easy to see that $h(D)$ reduces to the statistic $\mathrm{dinv}(D)$ from §1.4 when $m = 1$. Here is another formula for $h(D)$ which will be useful later. Define a function $\mathrm{sc}_m : \mathbb{Z} \to \mathbb{Z}$ by

$$\mathrm{sc}_m(p) = \begin{cases} m + 1 - p & \text{if } 1 \leq p \leq m; \\ m + p & \text{if } -m \leq p \leq 0; \\ 0 & \text{for all other } p. \end{cases}$$

Note that, given the value of a particular difference $\gamma_i(D) - \gamma_j(D)$ for a fixed i and j, we can evaluate the inner sum $\sum_{k=0}^{m-1} \chi(\gamma_i(D) - \gamma_j(D) + k \in \{0, 1, \ldots, m\})$ in (6). By checking the various cases, one sees that the value of this sum is exactly $\mathrm{sc}_m(\gamma_i(D) - \gamma_j(D))$. For instance, if $\gamma_i(D) - \gamma_j(D)$ is 0 or 1, then we get a contribution for each of the m values of k, in agreement with the fact that $\mathrm{sc}_m(0) = \mathrm{sc}_m(1) = m$. Similarly, if $\gamma_i(D) - \gamma_j(D)$ is $-(m-1)$, then only the summand with $k = m-1$ will cause a contribution, in agreement with the fact that $\mathrm{sc}_m(-(m-1)) = 1$. The remaining cases are checked similarly. We conclude that

$$h(D) = \sum_{0 \leq i < j < n} \mathrm{sc}_m(\gamma_i(D) - \gamma_j(D)). \tag{7}$$

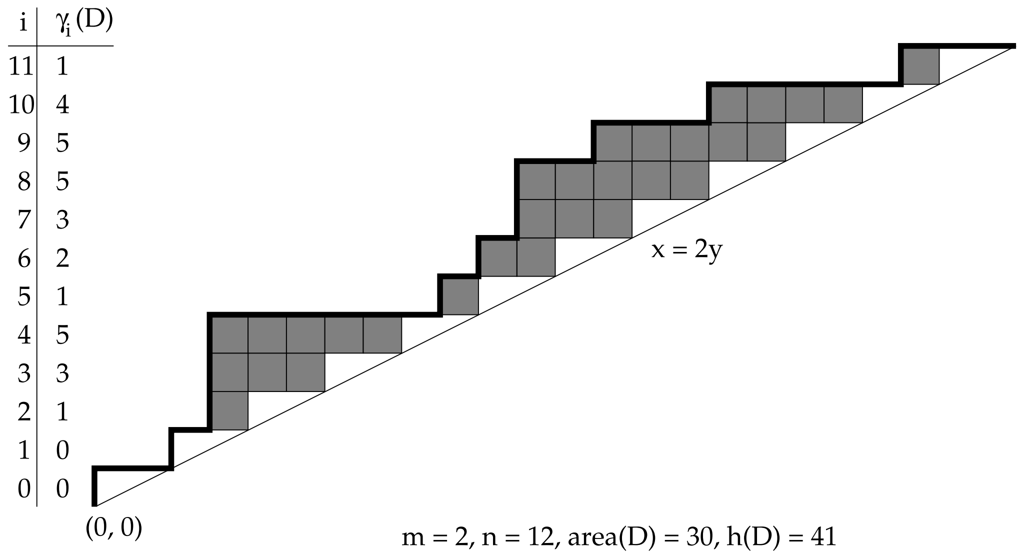


Figure 5: Defining the generalized Haiman statistic for a 2-path.

We now define the *first conjectured combinatorial version* of the higher q, t-Catalan sequence of order m by

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

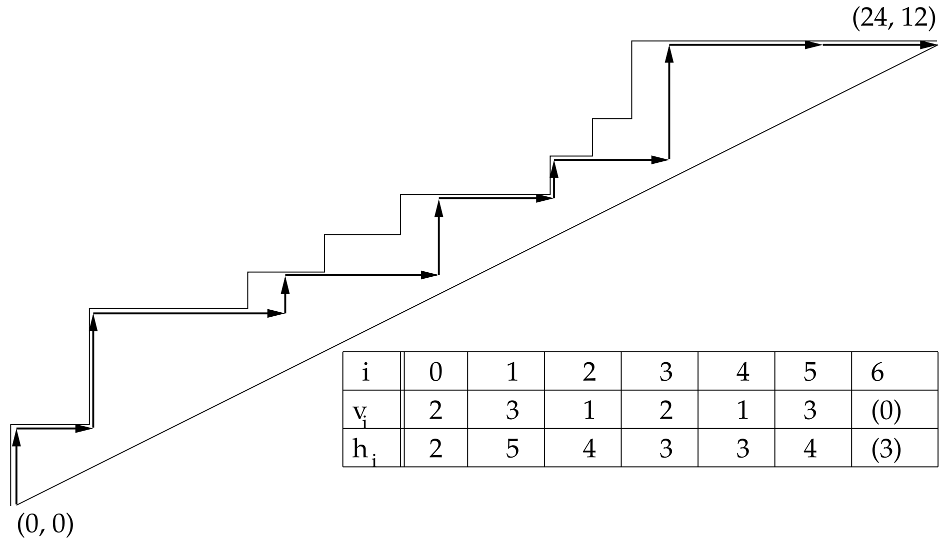
In §2.5, we will prove that $HC_n^{(m)}(q,1) = HC_n^{(m)}(1,q)$. This says that the statistic h has the same univariate distribution as the area statistic.

2.2 A Bounce Statistic for m-Dyck paths

We now discuss how to define a bounce statistic for m-Dyck paths that generalizes Haglund's statistic on ordinary Dyck paths. To define this statistic, we must first define the *bounce path* derived from a given m-Dyck path D.

In §1.4, we obtained the bounce path by starting at (n, n) and moving southwest towards $(0, 0)$ according to certain rules (see Figure 3). It is clear that, for ordinary Dyck paths, we could have obtained a similar statistic with the same distribution by starting at $(0, 0)$ and moving northeast. In the case of m-Dyck paths, it is more convenient to start the bouncing at $(0, 0)$.

Fix an integer $m \geq 2$. As before, the bounce path will consist of a sequence of alternating *vertical moves* and *horizontal moves*. We begin at $(0, 0)$ with a vertical move, and eventually end at (mn, n) after a horizontal move. Let $v_0, v_1, \ldots$ denote the lengths of the successive vertical moves in the bounce path, and let $h_0, h_1, \ldots$ denote the lengths of the successive horizontal moves. These lengths are calculated as follows. (Refer to Figures 6 and 7 for examples.)



i	0	1	2	3	4	5	6
v_i	2	3	1	2	1	3	(0)
h_i	2	5	4	3	3	4	(3)

Figure 6: Defining the bounce statistic for a 2-path.

To find v_0, move due north from $(0,0)$ until you reach an east step of the m-Dyck path; the distance traveled is v_0. Next, move due east v_0 units (so $h_0 = v_0$). Next, move north from the current position until you reach an east step of the m-Dyck path; let v_1 be the distance traveled. Next, move due east $v_0 + v_1$ units (so $h_1 = v_0 + v_1$). In general, for $i < m$, we move north v_i units from our current position until we are blocked by the m-Dyck path, and then move east $h_i = v_0 + v_1 + \cdots + v_i$ units.

For $i \geq m$, the rules change. At stage i, we still move north v_i units until we are blocked by the path. But we then move east $h_i = v_i + v_{i-1} + v_{i-2} + v_{i-(m-1)}$ units. In other words, the length of the next horizontal move is the sum of the m preceding vertical moves.

If we define $v_i = 0$ and $h_i = 0$ for all negative indices i, we can state a single rule that works for all the bounces. Start at $(0,0)$. Assuming inductively that v_j and h_j have been determined for all $j < i$ (where $i \geq 0$), move north from the current position until you are blocked by the m-Dyck path; define the distance traveled to be v_i. Then move east $h_i = v_i + v_{i-1} + \cdots + v_{i-(m-1)}$ units. We continue bouncing until we reach (mn, n). (In fact, it suffices to stop once we reach the top rim of the figure, which is the horizontal line $y = n$.) Finally, we define the *bounce statistic* $b(D)$ to be

$$b(D) = \sum_{k \geq 0} k \cdot v_k(D), \tag{8}$$

a weighted sum of the lengths of the vertical segments in the bounce path derived from D. For example, in Figure 6, we have

$$b(D') = 0 \cdot 2 + 1 \cdot 3 + 2 \cdot 1 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 3 = 30.$$

When $m = 1$, the new rule just says that $h_i = v_i$ for all i. In other words, we move north until we hit the Dyck path, and then move east the same distance, bringing us back to the main diagonal $y = x$. Thus, we obtain Haglund's bounce path (modified to start at $(0, 0)$, of course). To see that $b(D)$ agrees with the earlier statistic bounce(D), we first give an alternate formula for $b(D)$. Let s be the number of vertical moves needed to reach the top rim. Then $v_0 + v_1 + \cdots + v_{s-1} = n$, where n is the height of D. We claim that

$$b(D) = \sum_{k=0}^{s-1}(n - v_0 - v_1 - \cdots - v_k). \tag{9}$$

To see this, just replace n by $v_0 + \cdots + v_{s-1}$ in (9) and simplify the resulting sum. We get

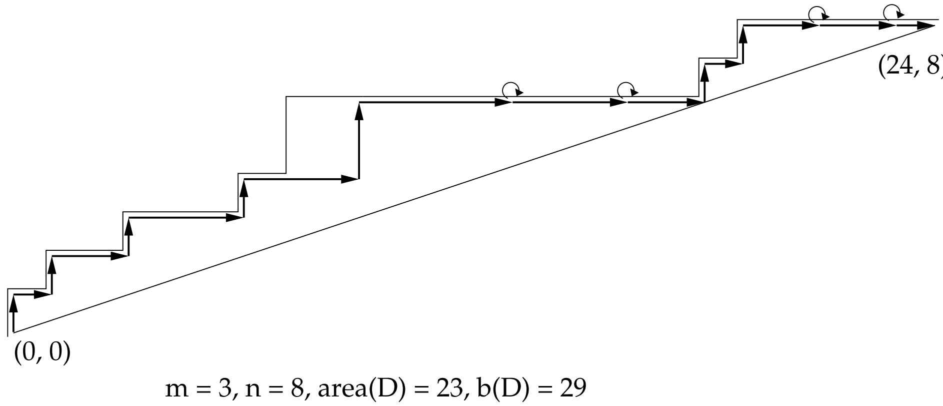
$$b(D) = (v_1 + v_2 + v_3 + \cdots + v_{s-1}) + (v_2 + v_3 + \cdots + v_{s-1}) + (v_3 + \cdots + v_{s-1}) + \cdots = \sum_{k\geq 0} k \cdot v_k,$$

which is formula (8). When $m = 1$, the numbers i_k in the definition of bounce(D) (see §1.4) are exactly the quantities $n - v_0 - v_1 - \cdots - v_k$. (Here we must reflect the shape in Figure 3 so that the bounce path starts at $(0, 0)$.) This shows that the new statistic does generalize the original one.

Note that, for $m > 1$, the bounce path does not necessarily return to the diagonal $x = my$ after each horizontal move. Consequently, it may occur that we cannot move north at all after making a particular horizontal move. This situation occurs for the bounce path shown in Figure 7, which is derived from the 3-path shown in Figure 4. In this case, we define the next v_i to be zero, and compute the next $h_i = v_i + v_{i-1} + \cdots + v_{i-(m-1)}$ just as before. In other words, vertical moves of length zero can occur, and are treated the same as nonzero vertical moves when computing the h_i's and the b statistic.

The possibility now arises that the bounce path could get "stuck" in the middle of the figure. To see why, suppose that m consecutive vertical moves $v_i, \ldots, v_{i+m-1}$ in the bounce path had length zero. Then the next horizontal move h_{i+m-1} would be zero also. As a result, our position in the figure at stage $i+m$ is exactly the same as the position at the beginning of stage $i+m-1$, since $v_{i+m-1} = h_{i+m-1} = 0$. From the bouncing rules, it follows that $v_{i+m} = 0$ also. But then $v_j = h_j = 0$ for all $j \geq i + m$, so that the bouncing path is stuck at the current position forever.

We now argue that the situation described in the last paragraph will never occur. Since the m-Dyck path must start with a north step, we have $v_0 > 0$, and so we do not get stuck at $(0, 0)$. The evolving bounce path will continue to make progress eastward with each horizontal step, unless $h_i = 0$ for some $i \geq 0$. Note that $h_i = 0$ iff $v_i + v_{i-1} + \cdots + v_{i-(m-1)} = 0$. Fix such an i, and consider the situation just after making the vertical move of length v_{i-1} and the horizontal move of length h_{i-1}. Let (x_0, y_0) denote the position of the bounce path at this instant. Then $y_0 = v_0 + v_1 + \cdots + v_{i-1}$ is the total vertical distance moved so far. Since $v_{i-1} = \cdots = v_{i-(m-1)} = 0$, we have $y_0 = v_0 + \cdots + v_{i-m}$. On the other hand, the total horizontal distance moved so far is $x_0 = h_0 + h_1 + \cdots + h_{i-1}$. From the definition of the h_j's and the fact that $v_{i-1} = \cdots = v_{i-(m-1)} = 0$, it follows that $x_0 = mv_0 + mv_1 + \cdots + mv_{i-m}$. In more detail, note that the last nonzero v_j, namely



i	0	1	2	3	4	5	6	7	8	9	10
v_i	1	1	1	1	2	0	0	1	1	(0)	(0)
h_i	1	2	3	3	4	3	2	1	2	(2)	(1)

Figure 7: A bounce path with vertical moves of length zero.

v_{i-m}, contributes to the m horizontal moves $h_{i-m}, \ldots, h_{i-1}$. Similarly, for $j < i-m$, v_j has contributed to m horizontal moves that have already occurred at the end of stage $i-1$. Since $v_j = 0$ for $i - m < j \leq i-1$, the stated formula for x_0 accounts for all the horizontal motion so far. Comparing the formulas for x_0 and y_0 gives $x_0 = my_0$, so that the bounce path has returned to the bounding diagonal $x = my$. If $y_0 = n$, the bounce path has reached its destination. If $y_0 < n$, the m-Dyck path continues above height y_0. But now $v_i > 0$ is forced; otherwise, the m-Dyck path must have gone east from (my_0, y_0), violating the requirement of always staying weakly above the line $x = my$. This argument is illustrated by the path in Figure 7.

Thus, the bounce path does not get stuck. The argument at the end of the last paragraph can be modified to show that the bounce path (like the m-Dyck path itself) never goes below the line $x = my$. For, after moving $v_0 + \cdots + v_{i-1}$ steps vertically at some time, we will have gone at most $mv_0 + \cdots + mv_{i-1}$ steps horizontally. Therefore, our position is on or above the line $x = my$.

Now that we know the bounce path is always well-defined, we can define the *second conjectured combinatorial version* of the higher q,t-Catalan sequence of order m by

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

In §2.5, we will give a bijective proof that $HC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$. Setting $t = 1$ or $q = 1$ here shows that both new statistics (h and b) have the same distribution on m-Dyck paths of height n as the area.

Conjecture: For all m and n, we have

$$OC_n^{(m)}(q,t) = HC_n^{(m)}(q,t) = C_n^{(m)}(q,t).$$

A possible approach to proving this conjecture will be indicated in §3.

2.3 A Formula for $C_n^{(m)}(q,t)$

In this section, we give an explicit algebraic formula (12) for $C_n^{(m)}(q,t)$ by analyzing bounce paths. This formula, while messy, is obviously a polynomial in q and t with nonnegative integer coefficients, unlike the formula defining $OC_n^{(m)}(q,t)$. A disadvantage of the new formula is that the (conjectured) symmetry $C_n^{(m)}(q,t) = C_n^{(m)}(t,q)$ is not evident from inspection of the formula.

Before stating the formula, we briefly review q-binomial coefficients. Let q be an indeterminate. Set $[n]_q = 1+q+q^2+\cdots+q^{n-1}$ for each positive integer n. Set $[0]_q! = 1$ and $[n]_q! = \prod_{i=1}^n [i]_q$ for $n>0$. Finally, set $\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q![n-k]_q!}$ for $0 \le k \le n$, and set $\begin{bmatrix} n \\ k \end{bmatrix}_q = 0$ for other values of k. When we replace q by 1, the expressions $[n]_q$, $[n]_q!$, and $\begin{bmatrix} n \\ k \end{bmatrix}_q$ evaluate to the numbers n, $n!$, and $\binom{n}{k}$, respectively. Note also that $\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q$. We will often write $\begin{bmatrix} a+b \\ a,b \end{bmatrix}_q$ to denote $\begin{bmatrix} a+b \\ a \end{bmatrix}_q = \begin{bmatrix} a+b \\ b \end{bmatrix}_q$ (multinomial coefficient notation).

We shall use the following well-known combinatorial interpretations of the q-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$. Let $R_{a,b}$ denote a rectangle of height a and width b. We write $\lambda \subset R_{a,b}$ for a partition λ if the Ferrers diagram of λ fits inside this rectangle. Then

$$\begin{bmatrix} a+b \\ a,b \end{bmatrix}_q = \sum_{\lambda \subset R_{a,b}} q^{|\lambda|} = \sum_{\lambda \subset R_{a,b}} q^{ab-|\lambda|}. \tag{10}$$

(The second equality follows from the first by rotating the rectangle 180° and considering the area cells inside the rectangle but outside λ.) We prefer the notation $\begin{bmatrix} a+b \\ a,b \end{bmatrix}_q$ because the bottom row displays both dimensions of the containing rectangle.

Here are two useful ways to rephrase (10). Let $\mathcal{P}_{a,b}$ denote the collection of all paths that proceed from the lower-left corner of $R_{a,b}$ to the upper-right corner by taking a north steps and b east steps of length one. (There is no other restriction on the paths.) If P is such a path, let $\operatorname{area}(P)$ be the number of cells in the rectangle lying below the path P. Then

$$\begin{bmatrix} a+b \\ a,b \end{bmatrix}_q = \sum_{P \in \mathcal{P}_{a,b}} q^{\operatorname{area}(P)} = \sum_{P \in \mathcal{P}_{a,b}} q^{ab-\operatorname{area}(P)}.$$

Similarly, let $R(0^a1^b)$ denote the collection of all rearrangements of a zeroes and b ones. If $w = (w_1w_2\ldots w_{a+b}) \in R(0^a1^b)$, define the *inversions of* w by $\operatorname{inv}(w) = \sum_{i<j} \chi(w_i > w_j)$ and the *coinversions of* w by $\operatorname{coinv}(w) = \sum_{i<j} \chi(w_i < w_j)$. Then

$$\begin{bmatrix} a+b \\ a,b \end{bmatrix}_q = \sum_{w \in R(0^a1^b)} q^{\operatorname{inv}(w)} = \sum_{w \in R(0^a1^b)} q^{\operatorname{coinv}(w)}. \tag{11}$$

This follows by representing w as a path $P \in \mathcal{P}_{a,b}$, which is obtained by replacing each zero in w by a north step and each one in w by an east step. Then the area above (resp., below) the path in $R_{a,b}$ is easily seen to be $\operatorname{inv}(w)$ (resp., $\operatorname{coinv}(w)$).

We are now ready to state the summation formula for $C_n^{(m)}(q,t)$. Let $\mathcal{V}_n^{(m)}$ denote the set of all sequences $v=(v_0,v_1,v_2,\ldots,v_s)$ such that: each v_i is a nonnegative integer; $v_0>0$; $v_s>0$; $v_0+v_1+v_2+\cdots+v_s=n$; and there is never a string of m or more consecutive zeroes in v. As usual, let $v_i=0$ for all negative i.

Theorem. With $\mathcal{V}_n^{(m)}$ defined as above, we have:

$$C_n^{(m)}(q,t)=\sum_{v\in\mathcal{V}_n^{(m)}} t^{\sum_{i\geq 0} iv_i} q^{m\sum_{i\geq 0}\binom{v_i}{2}} \prod_{i\geq 1} q^{v_i\sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i+v_{i-1}+\cdots+v_{i-m}-1 \\ v_i,\, v_{i-1}+\cdots+v_{i-m}-1\end{bmatrix}_q. \tag{12}$$

Equivalently, we may sum over all compositions v of n with zero parts allowed, if we identify compositions that differ only in trailing zeroes. The same formula holds for $HC_n^{(m)}(q,t)$, hence $C_n^{(m)}(q,t)=HC_n^{(m)}(q,t)$.

Remark: When $m=1$, this formula reduces to a formula for $C_n(q,t)$ given by Haglund in [9].

Proof, Part 1: Let $D\in\mathcal{D}_n^{(m)}$ be a typical object counted by $C_n^{(m)}(q,t)$. We can classify D based on the sequence $v(D)=(v_0,v_1,\ldots,v_s)$ of vertical moves in the bounce path derived from D. Call this sequence the *bounce composition* of D. By the discussion in the preceding section, the vector $v=v(D)$ belongs to $\mathcal{V}_n^{(m)}$. To prove the formula for $C_n^{(m)}(q,t)$, it suffices to show that

$$\sum_{D:\ v(D)=v} q^{\mathrm{area}(D)}t^{b(D)}=$$
$$t^{\sum_{i\geq 0} iv_i} q^{m\sum_{i\geq 0}\binom{v_i}{2}} \prod_{i=1}^{s} q^{v_i\sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i+v_{i-1}+\cdots+v_{i-m}-1 \\ v_i,\, v_{i-1}+\cdots+v_{i-m}-1\end{bmatrix}_q$$

for each $v=(v_0,\ldots,v_s)\in\mathcal{V}_n^{(m)}$. By our conventions for q-binomial coefficients, the right side of this expression is zero if any m consecutive v_i's are zero (in particular, this occurs if $v_0=0$). Thus, it does no harm in (12) to sum over all compositions v of n with zero parts allowed, not just the compositions v belonging to $\mathcal{V}_n^{(m)}$.

Now, fix $v\in\mathcal{V}_n^{(m)}$ and consider only the m-Dyck paths of height n having bounce composition v. By definition of the bounce statistic, every such path D will have the same t-weight, namely

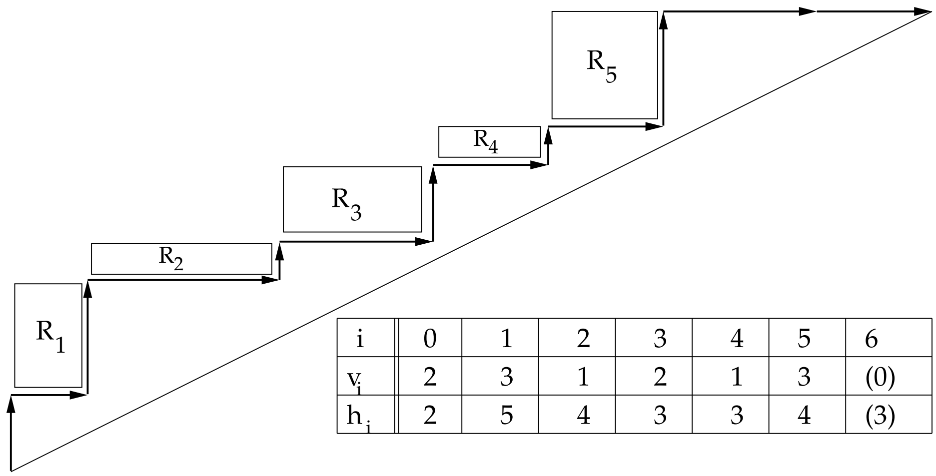
$$t^{b(D)}=t^{\sum_{i\geq 0} iv_i}.$$

To analyze the q-weights, note that we can construct all m-Dyck paths of height n having bounce composition v as follows.

1. Starting with an empty diagram, draw the bounce path with vertical segments $v_0,\ldots,v_s$. There is exactly one way to do this, since the horizontal moves h_i are completely determined by the vertical moves.

2. Having drawn the bounce path, there are now s empty rectangular areas just north-west of the "left-turns" in the bounce path. See Figure 8 for an example. Label

these rectangles $R_1, \ldots, R_s$, as shown. By definition of the bounce path, rectangle R_i has height v_i and width $h_{i-1} = v_{i-1} + \cdots + v_{i-m}$ for each i. To complete the m-Dyck path, draw a path in each rectangle R_i from the southwest corner to the northeast corner, *where each path begins with at least one east step.* The first east step in R_i must be present, by definition of v_{i-1}.



i	0	1	2	3	4	5	6
v_i	2	3	1	2	1	3	(0)
h_i	2	5	4	3	3	4	(3)

m = 2, n = 12.

Figure 8: Rectangles above the bounce path.

We can rephrase the second step as follows. Let R_i' denote the rectangle of height v_i and width $h_i = v_{i-1} + \cdots + v_{i-m} - 1$ obtained by ignoring the leftmost column of R_i. Then we can uniquely construct the path D by filling each shortened rectangle R_i' with an *arbitrary* path going from the southwest corner to the northeast corner.

The generating function for the number of ways to perform this second step, where the exponent of q records the total area above the bounce path, is

$$\prod_{i=1}^{s} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q$$

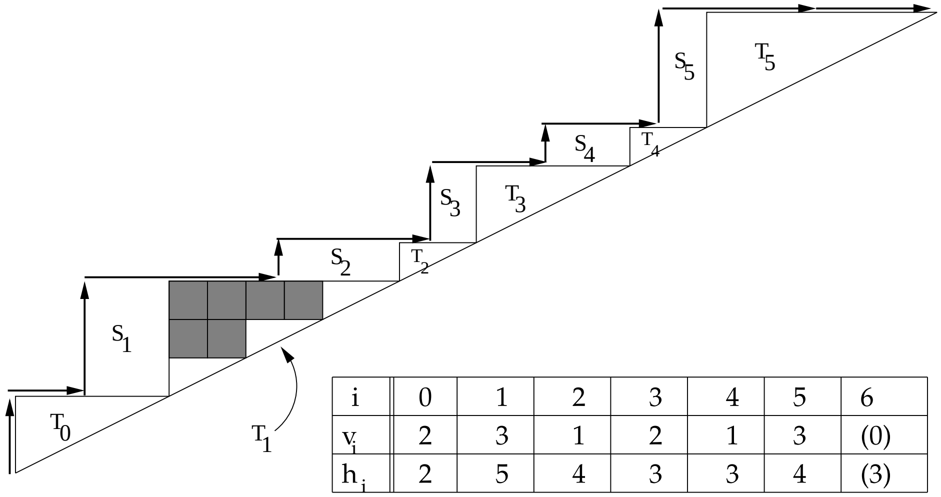
by the preceding discussion of q-binomial coefficients.

We still need to multiply by a power of q that records the area under the bounce path, which is independent of the choices in the second step. We claim that this area is

$$m \sum_{i=0}^{s} \frac{1}{2} v_i(v_i - 1) + \sum_{i=1}^{s} \left(v_i \sum_{j=1}^{m} (m-j) v_{i-j} \right),$$

which will complete the proof.

To establish the claim, dissect the area below the bounce path as shown in Figure 9. There are $s+1$ triangular pieces T_i, where the i'th triangle contains $0 + m + 2m + \cdots +$



i	0	1	2	3	4	5	6
v_i	2	3	1	2	1	3	(0)
h_i	2	5	4	3	3	4	(3)

m = 2, n = 12.

Figure 9: Dissecting the area below the bounce path.

$(v_i - 1)m = m\frac{v_i(v_i-1)}{2}$ complete cells. In Figure 9, for instance, where $v_1 = 3$, we have shaded the $0 + 2 + 4 = 6$ cells in T_1 that contribute to the area statistic. The total area coming from the triangles is

$$m \sum_{i=0}^{s} \frac{1}{2} v_i(v_i - 1).$$

There are also s rectangular slabs S_i (for $1 \leq i \leq s$). The height of slab S_i is v_i. What is the width of S_i? To answer this question, fix i, let (a, c) be the coordinates of the southeast corner of S_i, and let (b, c) be the coordinates of the southwest corner of S_i. First note that $c = v_0 + v_1 + \cdots + v_{i-1}$, the sum of the vertical steps prior to step i. Therefore,

$$a = mc = m(v_0 + \cdots + v_{i-1}) = mv_{i-1} + mv_{i-2} + \cdots + mv_{i-m} + mv_{i-m-1} + \cdots$$

since the southeast corner of S_i lies on the line $x = my$. Next, $b = h_0 + h_1 + \cdots + h_{i-1}$, the sum of the horizontal steps prior to step i. Recall that each h_j is the sum of the m preceding v_i's (starting with $i = j$). Substituting into the expression for b gives $b = 1v_{i-1} + 2v_{i-2} + \cdots + mv_{i-m} + mv_{i-m-1} + mv_{i-m-2} + \cdots$. We conclude that the width of S_i is

$$a - b = (m - 1)v_{i-1} + (m - 2)v_{i-2} + \cdots + (m - m)v_{i-m} + 0 + 0 + \cdots .$$

Finally, the area of S_i is the height times the width, which is

$$v_i(a - b) = v_i \sum_{j=1}^{m} (m - j)v_{i-j}.$$

Adding over all i gives the term

$$\sum_{i=1}^{s}\left(v_i\sum_{j=1}^{m}(m-j)v_{i-j}\right),$$

completing the proof of the claim and the first part of the theorem.

2.4 Proving the Formula for $HC_n^{(m)}(q,t)$

To finish the proof of the theorem, we now give a counting argument to show that $HC_n^{(m)}(q,t)$ is also given by the formula (12). This will show that $HC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$. In the next section, we combine the two different proofs of this formula to obtain a bijective proof of the identity $HC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$.

Recall that an m-Dyck path D can be represented by a vector

$$\gamma(D) = (\gamma_0(D), \ldots, \gamma_{n-1}(D)),$$

where $\gamma_i(D)$ is the number of area cells between the path and the diagonal in the i'th row from the bottom. Clearly, the path D is uniquely recoverable from the vector γ. Also, a vector $\gamma = (\gamma_0, \ldots, \gamma_{n-1})$ represents an element $D \in \mathcal{D}_n^{(m)}$ iff the following three conditions hold:

1. $\gamma_0 = 0$.
2. $\gamma_i \geq 0$ for all i.
3. $\gamma_{i+1} \leq \gamma_i + m$ for all $i < n-1$.

The first condition reflects the fact that the lowest row cannot have any area cells. The second condition ensures that the path D never goes below the diagonal $x = my$. The third condition follows since the path is not allowed to take any west steps.

Let $\mathcal{G}_n^{(m)}$ denote the set of all n-long vectors γ satisfying these three conditions. Then the preceding remarks show that

$$HC_n^{(m)}(q,t) = \sum_{\gamma\in\mathcal{G}_n^{(m)}} q^{h(\gamma)} t^{\sum_{i\geq 0}\gamma_i},$$

where $\sum_{i\geq 0}\gamma_i$ is the area of the path D corresponding to γ, and where we set

$$h(\gamma) = \sum_{0\leq i<j<n}\sum_{k=0}^{m-1}\chi(\gamma_i - \gamma_j + k \in \{0,1,\ldots,m\}),$$

so that $h(\gamma)$ is the h-statistic of the path D.

Given a vector $\gamma \in \mathcal{G}_n^{(m)}$, let $v_i(\gamma)$ be the number of times i occurs in the sequence $(\gamma_0, \ldots, \gamma_{n-1})$ for each $i \geq 0$. Let $v(\gamma) = (v_0(\gamma), v_1(\gamma), \ldots, v_s(\gamma))$ where s is the largest entry appearing in γ. We call $v(\gamma)$ the *composition of* γ. From the definitions of $\mathcal{G}_n^{(m)}$ and $v(\gamma)$, we see that $v_0 > 0$, $v_s > 0$, $v_0 + \cdots + v_s = n$, and there is never a string of m consecutive zeroes in v (lest $\gamma_{i+1} > \gamma_i + m$ for some i). In other words, v belongs to $\mathcal{V}_n^{(m)}$.

We now classify the objects γ in $\mathcal{G}_n^{(m)}$ based on their composition. To prove the summation formula for $HC_n^{(m)}(q,t)$, it suffices to show that

$$\sum_{\gamma:\ v(\gamma)=v} q^{h(\gamma)} t^{\sum_{i\geq 0}\gamma_i} = \\ t^{\sum_{i\geq 0} i v_i} q^{m\sum_{i\geq 0}\frac{1}{2}v_i(v_i-1)} \prod_{i=1}^{s} q^{v_i \sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q \tag{13}$$

for each $v = (v_0, \ldots, v_s) \in \mathcal{V}_n^{(m)}$. It is clear that the powers of t on each side of this equation agree, since v_i is the number of occurrences of the value i in the summation $\sum_{i\geq 0}\gamma_i$.

Before considering the powers of q, note that we can uniquely construct all vectors $\gamma \in \mathcal{G}_n^{(m)}$ having composition v as follows.

1. Initially, let γ be a string of v_0 zeroes.

2. Next, insert v_1 ones in the gaps to the right of these zeroes. There can be any number of ones in each gap, but no 1 may appear to the left of the leftmost zero.

3. Continue by inserting v_2 twos into valid locations, then v_3 threes, etc. The general step is to insert v_i copies of the symbol i into valid locations in the current string. Here, a "valid" location is one such that inserting i in that location will not cause a violation of the three conditions in the definition of $\mathcal{G}_n^{(m)}$.

How many ways are there to perform the i'th step of this insertion process, for $i > 0$? To answer this, note that a new symbol $i > 0$ can only be placed in a gap immediately to the right of the existing symbols $i-1, i-2, \ldots, i-m$ in the current string. There are $v_{i-1} + v_{i-2} + \cdots + v_{i-m}$ such symbols, and hence the same number of gaps. Since multiple copies of i can be placed in each gap, the number of ways to insert the v_i new copies of the symbol i is $\binom{v_i + v_{i-1} + \cdots + v_{i-m} - 1}{v_i, v_{i-1} + \cdots + v_{i-m} - 1}$. (To see this, represent a particular way of inserting the new i's by a string of v_i "stars" representing the i's and $v_{i-1} + \cdots + v_{i-m} - 1$ "bars" that separate the $v_{i-1} + \cdots + v_{i-m}$ available gaps.) Multiplying these expressions as i ranges from 1 to s, we see that formula (13) is correct when $q = 1$.

It remains to see that the power of q is correct as well. We prove this by induction on the largest symbol s appearing in γ. If $s = 0$, then $v = (n)$, and γ must consist of a string of n zeroes. From the definition, we see that $h(\gamma) = mn(n-1)/2$. This is the same as the power of q on the right side of (13), since $v_0 = n$ and $v_i = 0$ for $i > 0$.

Now assume that $s > 0$. Fix $v = (v_0, \ldots, v_s) \in \mathcal{V}_n^{(m)}$. Let $v' = (v_0, \ldots, v_{s-1})$, which is an element of $\mathcal{V}_{n-v_s}^{(m)}$ (ignore trailing zeroes in v' if necessary). Our induction hypothesis says that

$$\sum_{\delta:v(\delta)=v'} q^{h(\delta)} = q^{m\sum_{i=0}^{s-1} v_i(v_i-1)/2} \prod_{i=1}^{s-1} q^{v_i \sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q ;$$

note that any trailing zeroes in v' just contribute extra factors of 1 to the right side, which are harmless. We want to establish the analogous formula for

$$\sum_{\gamma:v(\gamma)=v} q^{h(\gamma)}.$$

For this purpose, recast the construction given in the $q = 1$ case as follows. We can uniquely produce every γ with $v(\gamma) = v$ by: first, choosing a δ with $v(\delta) = v'$; and second, choosing a way to insert v_s copies of s into δ in valid locations. The generating function for the number of ways to choose δ, where the power of q records $h(\delta)$, is by assumption

$$q^{m\sum_{i=0}^{s-1} v_i(v_i-1)/2} \prod_{i=1}^{s-1} q^{v_i \sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q .$$

To complete the proof, we need to show that the increase in the h-statistic caused by the second choice (namely, $h(\gamma) - h(\delta)$) has generating function

$$q^{mv_s(v_s-1)/2} q^{v_s \sum_{k=1}^{m}(m-k)v_{s-k}} \begin{bmatrix} v_s + v_{s-1} + \cdots + v_{s-m} - 1 \\ v_s, v_{s-1} + \cdots + v_{s-m} - 1 \end{bmatrix}_q ; \tag{14}$$

then the desired result will follow from the product rule for generating functions ([1], Ch. 10).

We encode the choice of how to insert the v_s copies of s into δ as a word

$$w \in R(0^{v_s} 1^{v_{s-1}+\cdots+v_{s-m}-1}).$$

To find w, read the symbols in the completed vector γ from left to right. Write down a zero in w every time an s occurs in γ; write down a one in w every time one of the symbols $s-1, \ldots, s-m$ occurs in γ; ignore all other symbols in γ. By the conditions on γ, the first symbol in w must be a one (since some symbol in $\{s-1, \ldots, s-m\}$ must appear just before the leftmost s in γ). Erase this initial 1 to obtain the word w.

We will prove that

$$h(\gamma) - h(\delta) = mv_s(v_s-1)/2 + v_s \sum_{k=1}^{m}(m-k)v_{s-k} + \mathrm{coinv}(w); \tag{15}$$

if this equation holds, then (14) immediately follows from it because of (11).

The proof of (15) proceeds by induction on the value of $\mathrm{coinv}(w)$. Suppose $\mathrm{coinv}(w) = 0$ first. This happens iff all v_s copies of s were inserted into δ immediately following the last occurrence of any symbol in the set $\{s-1,\ldots,s-m\}$. How do these v_s newly inserted symbols affect the h-statistic? To answer this, we must compute the sum (see (7))

$$\sum_{i<j} \mathrm{sc}_m(\gamma_i - \gamma_j)$$

over all pairs (i,j) such that $\gamma_i = s$ or $\gamma_j = s$.

First, consider the pairs (i,j) for which $i<j$ and $\gamma_i = s$ and $\gamma_j = s$. There are $\binom{v_s}{2}$ such pairs, and each contributes $\mathrm{sc}_m(s-s) = \mathrm{sc}_m(0) = m$ to the h-statistic. This gives the term $mv_s(v_s-1)/2$ in (15).

Second, consider the pairs (i,j) for which $i<j$ and $\gamma_i = s$ and $\gamma_j \neq s$. Since all the copies of s in γ occur in a contiguous group following all instances of the symbols $s-1,\ldots,s-m$, and since s is the largest symbol appearing in γ, $j>i$ implies that $\gamma_j < s-m$. Then $\mathrm{sc}_m(\gamma_i - \gamma_j) = 0$, since $\gamma_i - \gamma_j > m$. So these pairs contribute nothing to the h-statistic.

Third, consider the pairs (i,j) for which $i<j$ and $\gamma_i \neq s$ and $\gamma_j = s$. Since s is the largest symbol, we have $\gamma_i < s$. Write $\gamma_i = s-k$ for some $k>0$, and consider various subcases. Suppose $k \in \{1,2,\ldots,m\}$. Then $\mathrm{sc}_m(\gamma_i - \gamma_j) = \mathrm{sc}_m(-k) = m-k$. For how many pairs (i,j) does it happen that $i<j$, $\gamma_i = s-k$, and $\gamma_j = s$? There are v_s choices for the index j and v_{s-k} choices for the index i; the condition $i<j$ holds automatically, since all occurrences of s occur to the right of all occurrences of $s-k$. Thus, we get a total contribution to the h-statistic of $(m-k)v_s(v_{s-k})$ for this k. Adding over all k, we obtain the term

$$v_s \sum_{k=1}^{m} (m-k) v_{s-k}$$

appearing in (15). On the other hand, if $k>m$, then $\mathrm{sc}_m(\gamma_i - \gamma_j) = \mathrm{sc}_m(-k) = 0$, so there is no contribution to the h-statistic.

The three cases just considered are exhaustive, so we conclude that (15) is true when $\mathrm{coinv}(w)$ is zero.

For the inductive step, consider what happens when we replace two consecutive symbols 10 in w by 01, thus increasing $\mathrm{coinv}(w)$ by one. Let w' be the new word after the replacement, and let γ' be the associated vector obtained by inserting s's into δ according to the encoding w'. We may assume, by induction, that (15) is correct for γ and w. Passing from w to w' increases the right side of (15) by one. Hence, (15) will be correct for γ' and w', provided that $h(\gamma') = h(\gamma)+1$. To obtain γ' from γ, look at the symbols in γ corresponding to the replaced string 10 in w. The symbol corresponding to the 0 is an s. This s is immediately preceded in γ by a symbol in $\{s-1,\ldots,s-m\}$ which corresponds to the 1, by the conditions on γ and the fact that $s>0$. Say $s-k$ immediately precedes this s. The effect of replacing 10 by 01 in w is to move the s leftwards, past its predecessor $s-k$, and re-insert it in the next valid position in γ. This valid position occurs immediately to the right of the next occurrence of a symbol in $\{s,s-1,s-2,\ldots,s-m\}$ left of the symbol $s-k$. Pictorially, we have:

$$\text{original } \gamma = \ldots\ (s-j)\ z_1\ z_2\ \ldots\ z_\ell\ (s-k)\ s\ \ldots$$

where $0 \leq j \leq m$, $1 \leq k \leq m$, $\ell \geq 0$, and every $z_i < s - m$. After moving s left, we have

$$\text{new } \gamma' = \ldots \ (s-j)\ s\ z_1\ z_2\ \ldots\ z_\ell\ (s-k)\ \ldots.$$

Note that the symbol $s-j$ must exist, lest $\gamma'_0 = s > 0$.

Now, let us examine the effect of this motion on the h-statistic. When we move the s left past its predecessor $s-k$ in γ, we get a net change in the h-statistic of

$$\mathrm{sc}_m(s - [s-k]) - \mathrm{sc}_m([s-k] - s) = \mathrm{sc}_m(k) - \mathrm{sc}_m(-k) = +1,$$

since $1 \leq k \leq m$ (see (7)). As before, since $|s - z_i| > m$, moving the s past each z_i will not affect the h-statistic at all. Thus, the total change in the h-statistic is $+1$, as desired.

We can obtain an arbitrary encoding word w from the word $11\ldots100\ldots0$ with no coinversions by doing a finite sequence of interchanges of the type just described. Thus, the validity of (15) for all words w follows by induction on the number of such interchanges required (this number is exactly $\mathrm{coinv}(w)$, of course). This completes the proof of the theorem.

2.5 A Bijection Proving that $HC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$

The two proofs just given to show that formula (12) holds for $C_n^{(m)}(q,t)$ and $HC_n^{(m)}(q,t)$ were completely combinatorial. Hence, we can combine these proofs to get a bijective proof that $HC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$. Fix m and n. We describe a bijection $\phi : \mathcal{D}_n^{(m)} \to \mathcal{D}_n^{(m)}$ such that

$$h(D) = \mathrm{area}(\phi(D)) \text{ and } \mathrm{area}(D) = b(\phi(D)) \text{ for } D \in \mathcal{D}_n^{(m)}$$

and a bijection $\psi = \phi^{-1} : \mathcal{D}_n^{(m)} \to \mathcal{D}_n^{(m)}$ such that

$$b(D) = \mathrm{area}(\psi(D)) \text{ and } \mathrm{area}(D) = h(\psi(D)) \text{ for } D \in \mathcal{D}_n^{(m)}.$$

These bijections will show that the three statistics area, h, and b all have the same univariate distribution on $\mathcal{D}_n^{(m)}$.

Description of ϕ. Let D be an m-Dyck path of height n. To find the path $\phi(D)$:

- Represent D by the vector of row lengths $\gamma(D) = (\gamma_0(D), \ldots, \gamma_{n-1}(D))$, where $\gamma_i(D)$ is the number of area cells in the i'th row from the bottom.

- Define $v = (v_0, \ldots, v_s)$ by letting v_j be the number of occurrences of the value j in the vector $\gamma(D)$.

- Starting with an empty triangle, draw a bounce path from $(0,0)$ with successive vertical segments $v_0, \ldots, v_s$ and horizontal segments $h_0, h_1, \ldots$, where $h_i = v_i + v_{i-1} + \cdots + v_{i-(m-1)}$ for each i.

- For $1 \leq i \leq s$, form a word w_i from $\gamma(D)$ as follows. Initially, w_i is empty. Read γ from left to right. Write down a zero every time the symbol i is seen in γ. Write down a one every time a symbol in $\{i-1, \ldots, i-m\}$ is seen in γ. Ignore all other symbols in γ. At the end, erase the first symbol in w_i (which is necessarily a 1).

- Let $R_1, \ldots, R_s$ be the empty rectangles above the bounce path. Let $R'_1, \ldots, R'_s$ be these rectangles with the leftmost columns deleted (as in §2.3). For $1 \leq i \leq s$, use the word w_i to fill in the part of the path lying in R'_i, from the southwest corner to the northeast corner, by taking a north step for each zero in w_i, and an east step for each one in w_i. Call the completed path $\phi(D)$.

The two preceding proofs have already shown that ϕ has the desired effect on the various statistics.

Example. Let D be the 2-Dyck path of height 12 depicted in Figure 5. We have

$$\gamma(D) = (0,0,1,3,5,1,2,3,5,5,4,1); \quad \mathrm{area}(D) = 30; \quad h(D) = 41.$$

Doing frequency counts on the entries of γ, we compute

$$v = (v_0, v_1, v_2, v_3, v_4, v_5) = (2,3,1,2,1,3).$$

Given v, we can draw the bounce path shown in Figure 8 with 5 empty rectangles above it. Now, we compute the words w_i:

$$w_1 = 1000; \quad w_2 = 11101; \quad w_3 = 01101; \quad w_4 = 110; \quad w_5 = 01001.$$

Using these words to fill in the partial paths, we obtain the path D' in Figure 6, which has $b(D) = 30$ and $\mathrm{area}(D) = 41$.

Here is a mild simplification of the bijection. Leave the first 1 at the beginning of each w_i instead of erasing it. Then the w_i tell us how to construct the partial paths in the full rectangles R_i (rather than the shortened rectangles R'_i). Every such partial path begins with an east step, as required by the bouncing rules.

Description of ψ. Let D be an m-Dyck path of height n. To find the path $\psi(D)$:

- Draw the bounce path derived from D according to the bouncing rules (see §2.2). Let $v = (v_0, \ldots, v_s)$ be the lengths of the vertical moves in this bounce path.

- Let $R_1, \ldots, R_s$ be the rectangular regions above the bounce path. These regions contain partial paths going from the southwest corner to the northeast corner. For $1 \leq i \leq s$, find the word w_i by traversing the partial path in R_i and writing a one for each east step and a zero for each north step. Note that every w_i has first symbol one.

- Build up γ as follows. Start with a string of v_0 zeroes. For $i = 1, 2, \ldots, s$, insert v_i copies of i into the current string γ according to w_i. More explicitly, read w_i left to right. When a 1 is encountered, scan γ from left to right for the next occurrence of a symbol in $\{i-1, \ldots, i-m\}$. When a 0 is encountered, place an i in the gap immediately to the right of the current symbol in γ. Continue until all symbols i have been inserted.

- Use γ to draw the picture of a new m-Dyck path D' of height n, by placing γ_i area cells in the i'th row of the figure. Since $\gamma \in \mathcal{G}_n^{(m)}$, the resulting picture will be a valid path.

Example. Let D be the 3-Dyck path of height 8 shown in Figure 7. From the bounce path drawn in that figure, we find that

$$v = (v_0, \ldots, v_9) = (1, 1, 1, 1, 2, 0, 0, 1, 1).$$

Examining the rectangles above the bounce path (several of which happen to be empty or have height zero), we get the words w_i:

$$w_1 = 10;\ w_2 = 110;\ w_3 = 1110;\ w_4 = 10011;\ w_5 = 1111;\ w_6 = 111;\ w_7 = 110;\ w_8 = 10.$$

Now, build up the vector γ as follows:

- Initially, $\gamma = 0$ (since $v_0 = 1$).
- Use $w_1 = 10$ to insert one 1 into γ to get $\gamma = 01$.
- Use $w_2 = 110$ to insert one 2 into γ to get $\gamma = 012$.
- Use $w_3 = 1110$ to insert one 3 into γ to get $\gamma = 0123$.
- Use $w_4 = 10011$ to insert two 4's into γ to get $\gamma = 014423$.
- Use $w_5 = 1111$ to insert zero 5's into γ to get $\gamma = 014423$.
- Use $w_6 = 111$ to insert zero 6's into γ to get $\gamma = 014423$.
- Use $w_7 = 110$ to insert one 7 into γ to get $\gamma = 0144723$.
- Use $w_8 = 10$ to insert one 8 into γ to get $\gamma = 01447823$.

Thus, the image path D' is the unique 3-Dyck path of height 8 such that $\gamma(D') = (0, 1, 4, 4, 7, 8, 2, 3)$. D' is pictured in Figure 10.

As this example indicates, the presence of vertical moves of length zero does not alter the validity of the preceding proofs and bijections.

Remark. The main difficulty involved in the combinatorial investigation of the original q,t-Catalan sequence $OC_n(q,t)$ was discovering the two statistics *dinv* and *bounce* defined in §1.4. The area statistic, on the other hand, is quite natural to consider once one notices that $OC_n(1,1)$ counts the number of Dyck paths of height n. Similar comments apply to the higher q,t-Catalan sequences.

Having introduced the bijections ϕ and $\psi = \phi^{-1}$, we can consider the problem of finding these statistics in a new light. It is natural to count Dyck paths (or m-Dyck paths) by constructing the associated γ-sequences through successive insertion of zeroes, ones, twos, etc., as done in §2.4. The map ϕ arises by representing the insertion choices

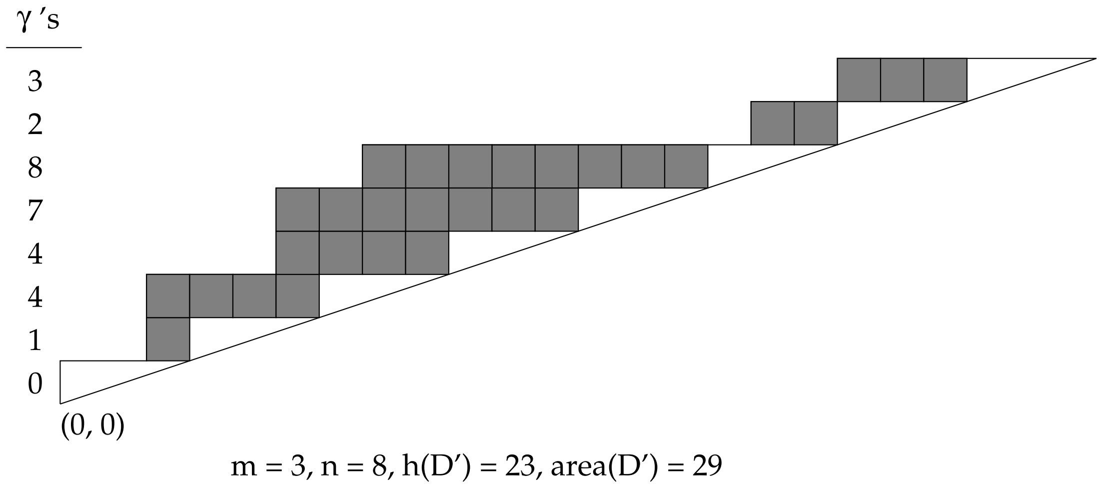


Figure 10: The image $\psi(D)$ for the path D from Figure 7.

geometrically as paths inside rectangles and positioning these rectangles in a nice way (as in Figure 8). The remarkable coincidence is that the resulting picture is another m-Dyck path.

We may thus regard the area statistic and the map ϕ as the "most fundamental" concepts. Then the two new statistics h and b can be "guessed" by simply looking at what happens to the area statistic when we apply ϕ (or ϕ^{-1})! We find that ϕ sends area to the bounce statistic b, and ϕ^{-1} sends area to the generalized Haiman statistic h.

This suggests a possible approach to other problems in which there are two variables with the same univariate distribution, but a combinatorial interpretation is only known for one of the variables. Finding a combinatorial interpretation for the Kostka-Macdonald coefficients (see [21]) provides an example of such a problem. There, the q-statistic is known (the so-called "cocharge statistic" on tableaux), but the t-statistic has not been discovered. For other examples of this technique of "guessing" new statistics, consult [17].

3 Recursions for $C_n^{(m)}(q,t)$

In this section, we prove several recursions for $C_n^{(m)}(q,t)$ and related sequences (see (23) and (36)). Of course, the same recursions hold for $HC_n^{(m)}(q,t)$. These recursions are more convenient for some purposes than the summation formula given in §2.3. As an example, we use the recursion to prove a formula for $C_n^{(m)}(q,1/q)$ which shows that $C_n^{(m)}(q,1/q) = OC_n^{(m)}(q,1/q)$ (see (24) and (34)).

We begin by describing Haglund's recursion for $C_n(q,t)$ (see [9]). This recursion is a key ingredient in the long proof that $SC_n(q,t) = C_n(q,t)$. We will just give the idea of the proof here; full details may be found in [5, 6].

3.1 Haglund's Recursion for $C_n(q,t)$

Fix n. Let $\mathcal{F}_{n,s}$ denote the set of Dyck paths of height n that terminate in exactly s east steps. For such a path, the length of the first bounce step will be s (see Figure 12 below). Define

$$F_{n,s}(q,t) = \sum_{D \in \mathcal{F}_{n,s}} q^{\mathrm{area}(D)} t^{\mathrm{bounce}(D)}.$$

These generating functions are related to $C_n(q,t)$ by the identities

$$C_n(q,t) = \sum_{s=1}^{n} F_{n,s}(q,t)$$

$$t^n C_n(q,t) = F_{n+1,1}(q,t).$$

The first identity follows by classifying Dyck paths of height n by the number s of east steps in the topmost row. To prove the second identity, augment the diagram of a Dyck path of height n by adding a new top row with no area cells. The result is a Dyck path of height $n+1$ terminating in one east step preceded by one north step. All elements of $\mathcal{F}_{n+1,1}$ arise uniquely in this way. The bounce path derived from this augmented Dyck path starts with a bounce of size 1 contributing n to the bounce statistic, and afterwards bounces in the same way that the original bounce path did. See Figure 11, and compare to Figure 3.

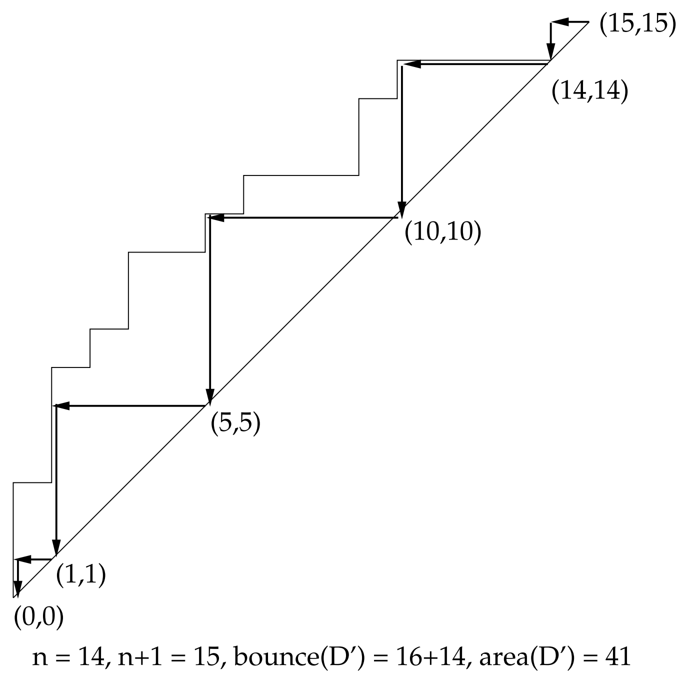


Figure 11: Adding an empty top row to a Dyck path.

Theorem (Haglund, [9]). The generating functions $F_{n,s}$ satisfy the recursion

$$F_{n,s}(q,t) = t^{n-s} q^{s(s-1)/2} \sum_{r=1}^{n-s} \begin{bmatrix} r+s-1 \\ r, s-1 \end{bmatrix}_q F_{n-s,r}(q,t) \text{ for } 1 \le s < n \tag{16}$$

with initial condition $F_{n,n}(q,t) = q^{n(n-1)/2}$.

Remark. Note that the initial condition and recursion uniquely determine the polynomials $F_{n,s}(q,t)$ and allow these polynomials to be computed rapidly.

Proof. Consider the initial condition first. If $D \in \mathcal{F}_{n,n}$, then D is a Dyck path of height n terminating in exactly n east steps in the top row. This can only happen if D is the path consisting of n north steps followed by n east steps. Then $\mathrm{area}(D) = n(n-1)/2$ (since $\gamma(D) = (0, 1, \ldots, n-1)$) and $\mathrm{bounce}(D) = 0$ (since the only bounce hits the diagonal at $(0,0)$). So, $F_{n,n}(q,t) = q^{n(n-1)/2} t^0$ as claimed.

The recursion for $F_{n,s}$ follows by "removing the first bounce" from a Dyck path to obtain a smaller Dyck path of height $n-s$. More precisely, let $D \in \mathcal{F}_{n,s}$. Then D ends in s east steps, so the derived bounce path starts with a bounce of size s ending at $(n-s, n-s)$. See Figure 12. If we ignore the top s rows of the figure, we see a smaller Dyck path D' of height $n-s$. Observe that the derived bounce path of D' is just the bounce path of D with the first bounce removed.

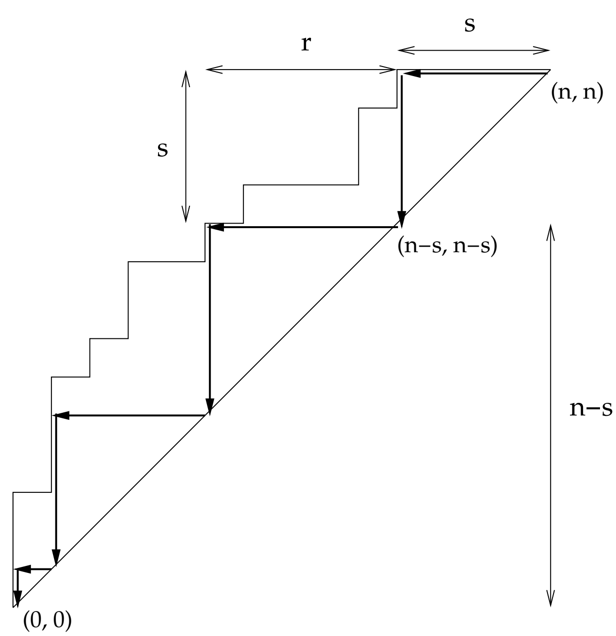


Figure 12: Proving the recursion by removing the first bounce.

We can uniquely construct a path $D \in \mathcal{F}_{n,s}$ as follows. Choose a number $r \in \{1, 2, \ldots, n-s\}$. Given r, build D by making a sequence of choices. First, choose a path $D' \in \mathcal{F}_{n-s,r}$. The generating function for this choice is $F_{n-s,r}(q,t)$. Second, draw a vertical and horizontal segment to create a triangle with vertices $(n-s, n-s)$ and $(n-s, n)$ and (n, n). This triangle adds $s(s-1)/2$ area cells to the path being constructed, giving a factor $q^{s(s-1)/2}$. Also, the path D will have a new bounce going from (n, n) to $(n-s, n-s)$, so we get a contribution of t^{n-s} as well. Third, draw a subpath ending with a north step in the rectangular region above the top row of D' and left of the triangle just drawn. This subpath does not change the bounce statistic (since it ends in a north step), but the area increases by the number of cells beneath the subpath in its rectangle. The generating function for this choice is thus $\begin{bmatrix} r+s-1 \\ r, s-1 \end{bmatrix}_q$. The recursion follows immediately from the sum and product rules for generating functionons ([1], Ch. 10).

Proving that $C_n(q,t) = SC_n(q,t)$.
In [5, 6], Garsia and Haglund used the recursion (16) to prove that $C_n(q,t) = SC_n(q,t)$ for all n. More specifically, they defined

$$Q_{n,s}(q,t) = t^{n-s} q^{s(s-1)/2} \nabla(e_{n-s}[X(1+q+\cdots+q^{s-1})])\big|_{s_{1^{n-s}}}.$$

Here, X is a formal infinite alphabet $X = x_1 + x_2 + \cdots$, and the square brackets denote plethystic substitution; in particular $e_n[X] = e_n$. Garsia and Haglund showed that

$$Q_{n,s}(q,t) = t^{n-s} q^{s(s-1)/2} \sum_{r=1}^{n-s} \begin{bmatrix} r+s-1 \\ r, s-1 \end{bmatrix}_q Q_{n-s,r}(q,t) \text{ and } Q_{n,n}(q,t) = q^{n(n-1)/2}.$$

In other words, $Q_{n,s}$ satisfies the same recursion and initial condition that $F_{n,s}$ does. By uniqueness, $Q_{n,s} = F_{n,s}$ for all n and s. In particular,

$$C_n(q,t) = F_{n+1,1}(q,t)/t^n = Q_{n+1,1}(q,t)/t^n = \nabla(e_n[X])\big|_{s_{1^n}} = SC_n(q,t).$$

3.2 A recursion based on removing the first bounce

Our goal here is to modify the idea in the proof of Haglund's recursion to get a recursion for $C_n^{(m)}(q,t)$. The main difficulty is that the bounce path depends on the prior bouncing history when $m > 1$, so that we cannot simply remove the first bounce and restart "from scratch." Consequently, we must add more subscripts that keep track of the lengths of the first m vertical moves in the bounce path.

Fix $m > 1$. Define $\mathcal{F}_{n;v_0,v_1,\ldots,v_{m-1}}$ to be the collection of m-Dyck paths of height n whose derived bounce paths start with vertical moves of lengths $v_0, v_1, \ldots, v_{m-1}$, in that order. Define $F_{n;v_0,\ldots,v_{m-1}}(q,t)$ to be the sum of $q^{\mathrm{area}(D)} t^{b(D)}$ over all paths $D \in \mathcal{F}_{n;v_0,\ldots,v_{m-1}}$. (An empty sum is defined to be zero.) We make the following observations about these definitions.

- If $\mathcal{F}_{n;v_0,v_1,\ldots,v_{m-1}}$ is a nonempty collection of paths, then we must have $v_0 > 0$, $v_i \geq 0$ for $i > 0$, and $v_0 + \cdots + v_{m-1} \leq n$.
- If $v_0 = n$ and $v_i = 0$ for $i > 0$, then $\mathcal{F}_{n;n,0,\ldots,0}$ consists of the single path D that goes north n steps and then east mn steps. Hence, $F_{n;n,0,\ldots,0}(q,t) = q^{mn(n-1)/2} t^0$.

- Consider the collection $\mathcal{F}_{n+1;1,0,\ldots,0}$. A path D in this collection starts by going north one unit and then east m units (since $v_1 = \cdots = v_{m-1} = 0$). At this point, D has returned to the diagonal $x = my$. If we look at the rest of the path beyond this point, we get an arbitrary m-Dyck path D' of height n. Also, the bounce path for D' is the same as the latter part of the bounce path for D (starting with v_m). Note that the prior history in D is immaterial, since $v_{m-1} = \cdots = v_1 = 0$. See Figure 13. We conclude that
$$t^{mn} C_n^{(m)}(q,t) = F_{n+1;1,0,\ldots,0}(q,t).$$
The extra factor of t^{mn} accounts for the contribution of the first m bounces to $b(D)$, which is not present in $b(D')$.

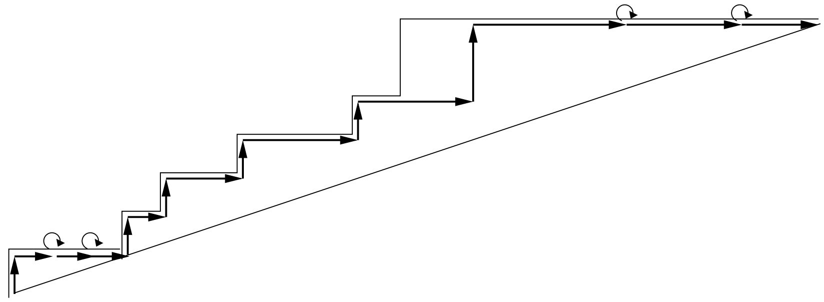


Figure 13: Removing a trivial bottom row of an m-Dyck path.

- There is a version of the formula (12) for $F_{n;v_0,\ldots,v_{m-1}}(q,t)$. Specifically,
$$F_{n;v_0,\ldots,v_{m-1}}(q,t) = \sum_{(v_m,v_{m+1},\ldots)} t^{\sum_{i\geq 0} iv_i} q^{m\sum_{i\geq 0}\binom{v_i}{2}} \prod_{i\geq 1} q^{v_i \sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q.$$
This equation follows immediately from the combinatorial interpretation of the summation index $v = (v_0, v_1 \ldots)$ appearing in (12) as the lengths of the vertical segments in the bounce path. Since $v_0, \ldots, v_{m-1}$ are fixed in advance, we need only sum over the remaining segments $v_m, v_{m+1}, \ldots$.

To state the new recursion, it is convenient to introduce a modified version of the generating functions $F_{n;v_0,\ldots,v_{m-1}}(q,t)$. Intuitively, we need to remove the influence of v_0 on the future bouncing history to obtain a recursion. Assume that $v_0 > 0$ first. Define $\mathcal{E}_{n;v_0,\ldots,v_{m-1}}$ to be the collection of all m-Dyck paths D of height n with the following properties. First, the bounce path derived from D starts with vertical moves of lengths $v_0, \ldots, v_{m-1}$. Second, the first $m-1$ rectangles $R_1, \ldots, R_{m-1}$ above the bounce path of D (see Figure 8) are all empty. This means that the subpath in each rectangle goes all the way east before turning north, so that there are no area cells in the rectangle. Then define
$$E_{n;v_0,\ldots,v_{m-1}}(q,t) = \sum_{D \in \mathcal{E}_{n;v_0,\ldots,v_{m-1}}} q^{\mathrm{area}(D)} t^{b(D)}.$$

By filling the empty rectangles $R_1, \ldots, R_{m-1}$ in an object $D \in \mathcal{E}_{n;v_0,\ldots,v_{m-1}}$ according to the bouncing rules, we deduce that

$$F_{n;v_0,\ldots,v_{m-1}}(q,t) = E_{n;v_0,\ldots,v_{m-1}}(q,t) \prod_{i=1}^{m-1} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q \text{ when } v_0 > 0. \quad (17)$$

This relation gives an exact formula for $E_{n;v_0,\ldots,v_{m-1}}(q,t)$ when $v_0 > 0$:

$$E_{n;v_0,\ldots,v_{m-1}}(q,t) = \sum_{(v_m,v_{m+1},\ldots)} t^{\sum_{i\geq 0} iv_i} q^{m\sum_{i\geq 0}\frac{1}{2}v_i(v_i-1)} \prod_{i\geq 1} q^{v_i \sum_{j=1}^{m\wedge i}(m-j)v_{i-j}} \prod_{i\geq m} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q. \quad (18)$$

Here, we have written $m \wedge i$ to denote the minimum of m and i. Note that the validity of equation (18) does not depend on the earlier convention that $v_i = 0$ for all negative i. Now, if $v_0 = 0$, we simply *define* $E_{n;v_0,\ldots,v_{m-1}}(q,t)$ by formula (18).

It follows from (17) that $E^{(m)}_{n+1;1,0,\ldots,0}(q,t) = F^{(m)}_{n+1;1,0,\ldots,0}(q,t)$. Therefore,

$$C_n^{(m)}(q,t) = t^{-mn} E^{(m)}_{n+1;1,0,\ldots,0}(q,t). \quad (19)$$

We can obtain a recursion for $E_{n;v_0,\ldots,v_{m-1}}$ by breaking up the summation in (18) based on the value of v_m. Consider a fixed choice of v_m in the range $\{0,1,\ldots,n-v_0-\cdots-v_{m-1}\}$. Write down (18) with n replaced by $n-v_0$ and v_k replaced by v_{k+1} for all $k \geq 0$:

$$E_{n-v_0;v_1,\ldots,v_m}(q,t) = \sum_{(v_{m+1},v_{m+2},\ldots)} t^{\sum_{i\geq 0} iv_{i+1}} q^{\mathrm{pow}_1} \prod_{i\geq m} \begin{bmatrix} v_{i+1} + v_i + \cdots + v_{i+1-m} - 1 \\ v_{i+1}, v_i + \cdots + v_{i+1-m} - 1 \end{bmatrix}_q, \quad (20)$$

$$\mathrm{pow}_1 = m \sum_{i\geq 0} \binom{v_{i+1}}{2} + \sum_{i\geq 1} v_{i+1} \sum_{j=1}^{m\wedge i} (m-j) v_{i+1-j}.$$

Replace i by $i-1$ in this formula to get

$$E_{n-v_0;v_1,\ldots,v_m}(q,t) = \sum_{(v_{m+1},v_{m+2},\ldots)} t^{\sum_{i\geq 1}(i-1)v_i} q^{\mathrm{pow}_2} \prod_{i> m} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q, \quad (21)$$

$$\mathrm{pow}_2 = m \sum_{i\geq 1} \binom{v_i}{2} + \sum_{i\geq 2} v_i \sum_{j=1}^{m\wedge(i-1)} (m-j) v_{i-j}.$$

In the original formula for $E_{n;v_0,\ldots,v_{m-1}}$, we can sum over v_m first and then sum over the remaining v_j's. The resulting formula is:

$$E_{n;v_0,\ldots,v_{m-1}}(q,t) = \sum_{v_m=0}^{n-v_0-\cdots-v_{m-1}} \sum_{(v_{m+1},v_{m+2},\ldots)} t^{\sum_{i\geq 0} iv_i} q^{\mathrm{pow}_3} \prod_{i\geq m} \begin{bmatrix} v_i + v_{i-1} + \cdots + v_{i-m} - 1 \\ v_i, v_{i-1} + \cdots + v_{i-m} - 1 \end{bmatrix}_q, \quad (22)$$

$$\mathrm{pow}_3 = m\sum_{i\geq 0}\binom{v_i}{2} + \sum_{i\geq 1} v_i \sum_{j=1}^{m\wedge i}(m-j)v_{i-j}.$$

To go from the formula in (21) to the corresponding summand in (22), we need to multiply the former by the expression

$$t^{0v_0+v_1+v_2+v_3+\cdots} q^{m\binom{v_0}{2}} q^{v_1 v_0(m-1)} \prod_{i=2}^{m} q^{v_i(m-i)v_0} \begin{bmatrix} v_m+\cdots+v_0-1 \\ v_m, v_{m-1}+\cdots+v_0-1 \end{bmatrix}_q .$$

Doing this multiplication and adding over all choices of v_m, we obtain the recursion

$$E_{n;v_0,\ldots,v_{m-1}}(q,t) =$$
$$t^{n-v_0} q^{m\binom{v_0}{2}} \prod_{i=1}^{m-1} q^{v_0 v_i(m-i)} \sum_{v_m=0}^{n-v_0-\cdots-v_{m-1}} \begin{bmatrix} v_m+\cdots+v_0-1 \\ v_m, v_{m-1}+\cdots+v_0-1 \end{bmatrix}_q E_{n-v_0;v_1,\ldots,v_{m-1},v_m}(q,t). \tag{23}$$

The initial conditions are

$$E_{n;n,0,\ldots,0}(q,t) = q^{mn(n-1)/2} t^0$$
$$E_{n;0,0,\ldots,0}(q,t) = 0.$$

Observe that we recover Haglund's original recursion when $m=1$.

It is hoped that (23) could be used to prove the conjecture $C_n^{(m)}(q,t) = SC_n^{(m)}(q,t)$. One difficulty is finding the analogues of $E_{n;v_0,\ldots,v_{m-1}}$ in the symmetric function setting. Computer experiments suggest that

$$E_{n;v_0,0,\ldots,0}(q,t) = q^{mv_0(v_0-1)/2} t^{(m-1)(n-v_0)}\ \nabla^m(e_{n-v_0}[X(1+q+q^2+\cdots+q^{v_0-1})])\big|_{s_{1^{n-v_0}}},$$

However, we have not found a conjectured formula for the general $E_{n;v_0,\ldots,v_{m-1}}$ in terms of the nabla operator.

It is clear that we could perform a similar manipulation of (12) to obtain a recursion based on removing the last nontrivial vertical bounce v_s. The inductive proof in §2.4 that (12) equals $HC_n^{(m)}(q,t)$ was based on this idea. There is a slight added complication because one must know s, not just v_s, to determine the effect of removing the last bounce on $b(D)$. On the other hand, v_s only affects the dimensions of one nontrivial rectangle in Figure 8.

3.3 Application: A Formula for the Specialization $C_n^{(m)}(q,1/q)$

We now use the recursion of the preceding subsection to derive an exact formula for the specialization

$$E^{(m)}_{n;v_0,\ldots,v_{m-1}}(q,1/q).$$

In particular, using this formula together with (19), we prove that

$$q^{mn(n-1)/2} C_n^{(m)}(q,1/q) = \frac{1}{[mn+1]_q} \begin{bmatrix} mn+n \\ mn, n \end{bmatrix}_q .$$

Garsia and Haiman proved the same formula for $OC_n^{(m)}(q,1/q)$ in [7]. It follows that

$$C_n^{(m)}(q,1/q) = OC_n^{(m)}(q,1/q).$$

The Formula for the E's. Fix m, N, and $v=(v_0,\dots,v_{m-1})$. Our formula for $E_{N;v}^{(m)}(q,1/q)$ will involve various intermediate quantities A, B, etc., depending on N, m, and v. If the dependence on the variables needs to be made explicit, we will write $A(N,m,v)$, $B(N,m,v)$, etc.

The basic formula is

$$E_{N;v}^{(m)}(q,1/q) = A_0 - B_1 - B_2 - \cdots - B_m, \tag{24}$$

where A_0 and each B_j is a certain q-binomial coefficient multiplied by a certain power of q. Specifically, define

$$\begin{aligned}
A = A(N,m,v_0,\dots,v_{m-1}) &= \begin{bmatrix} (m+1)N-1-\sum_{k=0}^{m-1}(m-k)v_k \\ N-\sum_{k=0}^{m-1} v_k \end{bmatrix}_q \\
B = B(N,m,v_0,\dots,v_{m-1}) &= \begin{bmatrix} (m+1)N-1-\sum_{k=0}^{m-1}(m-k)v_k \\ N-1-\sum_{k=0}^{m-1} v_k \end{bmatrix}_q \\
P_0 = P_0(N,m,v_0,\dots,v_{m-1}) &= -\frac{m}{2}(N^2+N) + \left[m\left(\sum_{k=0}^{m-1} v_k\right) - (m-1)\right] N \\
&\quad + \sum_{k=0}^{m-2}(m-1-k)v_k + \sum_{0\le j<k\le m-1}(j-k)v_j v_k \\
P_j = P_j(N,m,v_0,\dots,v_{m-1}) &= v_{m-1} + (j-1)N \\
&\quad - \sum_{\ell=0}^{m-2} \min(j-1, m-2-\ell)v_\ell \quad (1\le j\le m)
\end{aligned}$$

Finally, define

$$\begin{aligned}
A_0 = A_0(N,m,v_0,\dots,v_{m-1}) &= Aq^{P_0} \\
B_j = B_j(N,m,v_0,\dots,v_{m-1}) &= Bq^{P_0+P_j} \quad (1\le j\le m)
\end{aligned}$$

Examples. (1) Let $m=1$ and $v_0=w$. Then

$$E_{N;w}^{(1)}(q,1/q) = q^{-(N^2+N)/2+wN}\left\{\begin{bmatrix} 2N-w-1 \\ N-w, N-1 \end{bmatrix}_q - q^w \begin{bmatrix} 2N-w-1 \\ N-w-1, N \end{bmatrix}_q\right\}.$$

This is equivalent to a formula for $F_{n,s}(q,1/q)$ proved by Haglund in [9].

(2) Let $m=2$, $v_0=w$, $v_1=x$. Then

$$\begin{aligned} E^{(2)}_{N;w,x}(q,1/q) &= q^{\mathrm{pow}_2}\left\{\begin{bmatrix} 3N-2w-x-1 \\ N-w-x, 2N-w-1\end{bmatrix}_q \right.\\ &\qquad \left. -(q^x+q^{N+x})\begin{bmatrix} 3N-2w-x-1 \\ N-w-x-1, 2N-w\end{bmatrix}_q\right\}, \\ \mathrm{pow}_2 &= -(N^2+N)+(2w+2x-1)N-w(x-1). \end{aligned}$$

(3) Let $m=3$, $v_0=w$, $v_1=x$, $v_2=y$. Then

$$\begin{aligned} E^{(3)}_{N;w,x,y}(q,1/q) &= q^{\mathrm{pow}_3}\left\{\begin{bmatrix} 4N-3w-2x-y-1 \\ N-w-x-y, 3N-2w-x-1\end{bmatrix}_q \right.\\ &\qquad \left. -(q^y+q^{y+N-w}+q^{y+2N-w})\begin{bmatrix} 4N-3w-2x-y-1 \\ N-w-x-y-1, 3N-2w-x\end{bmatrix}_q\right\}, \\ \mathrm{pow}_3 &= -3(N^2+N)/2+(3w+3x+3y-2)N+(y-1)(-2w-x)-wx. \end{aligned}$$

(4) Let $m=5$ and $(v_0,v_1,v_2,v_3,v_4)=(v,w,x,y,z)$. Then

$$\begin{aligned} P_1 &= z \\ P_2 &= z+N-v-w-x \\ P_3 &= z+2N-2v-2w-x \\ P_4 &= z+3N-3v-2w-x \\ P_5 &= z+4N-3v-2w-x. \end{aligned}$$

Proof of the Formula. To prove (24), we need to check that the right side satisfies the same initial conditions and recursion that the specialization $E^{(m)}_{N;v}(q,1/q)$ satisfies. This check requires an inordinate amount of tedious manipulations of powers of q. Therefore, we only give an outline of the proof here, omitting routine algebraic manipulations. We refer the interested reader to the author's thesis [17] for more details.

Step 1. We begin by establishing the following identity, valid for nonnegative integers C, D, and E:

$$\sum_{i=0}^{D-E}\begin{bmatrix} C+i \\ C, i\end{bmatrix}_q \begin{bmatrix} D-i \\ E, D-i-E\end{bmatrix}_q q^{(E+1)i} = \begin{bmatrix} C+D+1 \\ D-E, C+1+E\end{bmatrix}_q. \tag{25}$$

Recall from §2.3 that

$$\begin{bmatrix} a+b \\ a, b\end{bmatrix}_q = \sum_{P\in\mathcal{P}_{a,b}} q^{\mathrm{area}(P)} = \sum_{P\in\mathcal{P}_{a,b}} q^{ab-\mathrm{area}(P)},$$

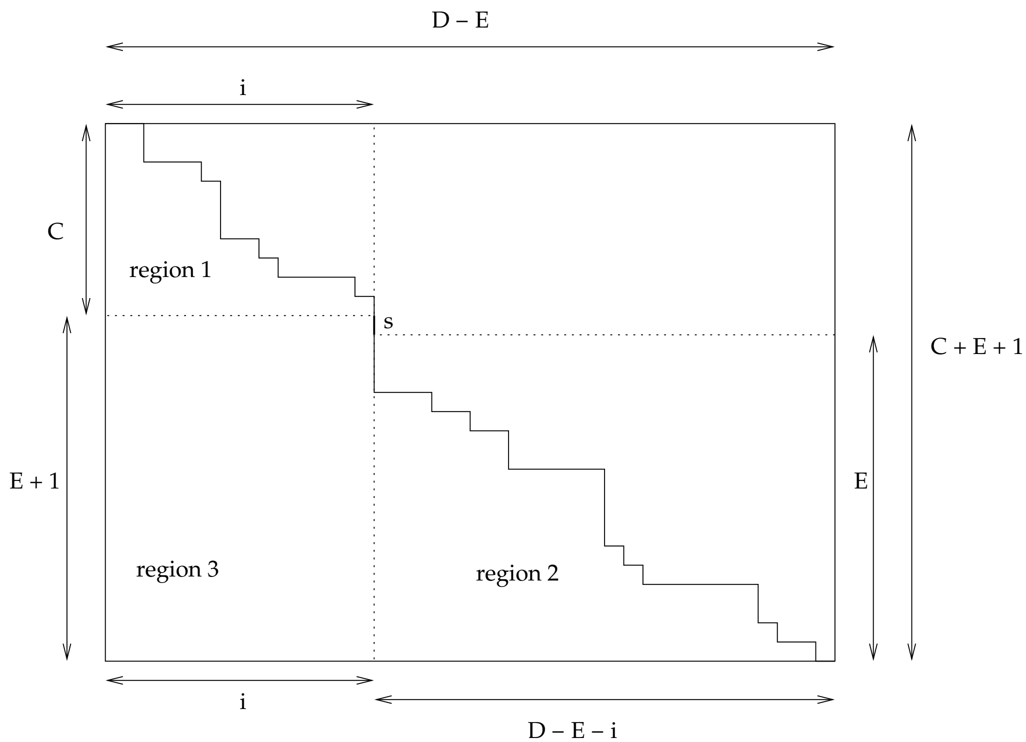


Figure 14: Picture used to prove (25).

where $\mathcal{P}_{a,b}$ is the set of lattice paths contained in an $a \times b$ rectangle. Using this fact, we can prove (25) by drawing a picture. See Figure 14.

We classify paths P contained in a rectangle of height $C + E + 1$ and width $D - E$ based on what happens in row $(C+1)$ from the top. This row contains exactly one vertical step s of P; let i denote the distance of this vertical step from the left edge. Evidently, $0 \leq i \leq D - E$. Given i, we can uniquely construct such a path P as follows. First, choose a subpath P_1 in the rectangle R_1 northwest of s, which has height C and width i. Second, choose a subpath P_2 in the rectangle R_2 southeast of s, which has height E and width $D - E - i$. Then P is the concatenation of P_1 and the vertical step s and P_2.

Assume that the power of q records the area below the path P. This area is the sum of the area below P_1 inside R_1, the area below P_2 inside R_2, and the full area of the southwest rectangle of height $E + 1$ and width i. These three pieces of the area are accounted for by the factors $\left[{C+i \atop C,i}\right]_q$, $\left[{D-i \atop E,D-i-E}\right]_q$, and $q^{(E+1)i}$, respectively. Adding over all choices of i, we immediately obtain (25).

Step 2. We prove the identity

$$q^{-C}\left\{\begin{bmatrix} C+D \\ C,D \end{bmatrix}_q - \begin{bmatrix} C+D \\ C-1,D+1 \end{bmatrix}_q\right\} = \begin{bmatrix} C+D \\ C,D \end{bmatrix}_q - q^{D-C+1}\begin{bmatrix} C+D \\ C-1,D+1 \end{bmatrix}_q. \qquad (26)$$

This identity is equivalent to the relation

$$\begin{bmatrix} C+D \\ C,D \end{bmatrix}_q + q^{D+1} \begin{bmatrix} C+D \\ C-1,D+1 \end{bmatrix}_q = q^C \begin{bmatrix} C+D \\ C,D \end{bmatrix}_q + \begin{bmatrix} C+D \\ C-1,D+1 \end{bmatrix}_q ,$$

which can also be proved by drawing a picture. See Figure 15.

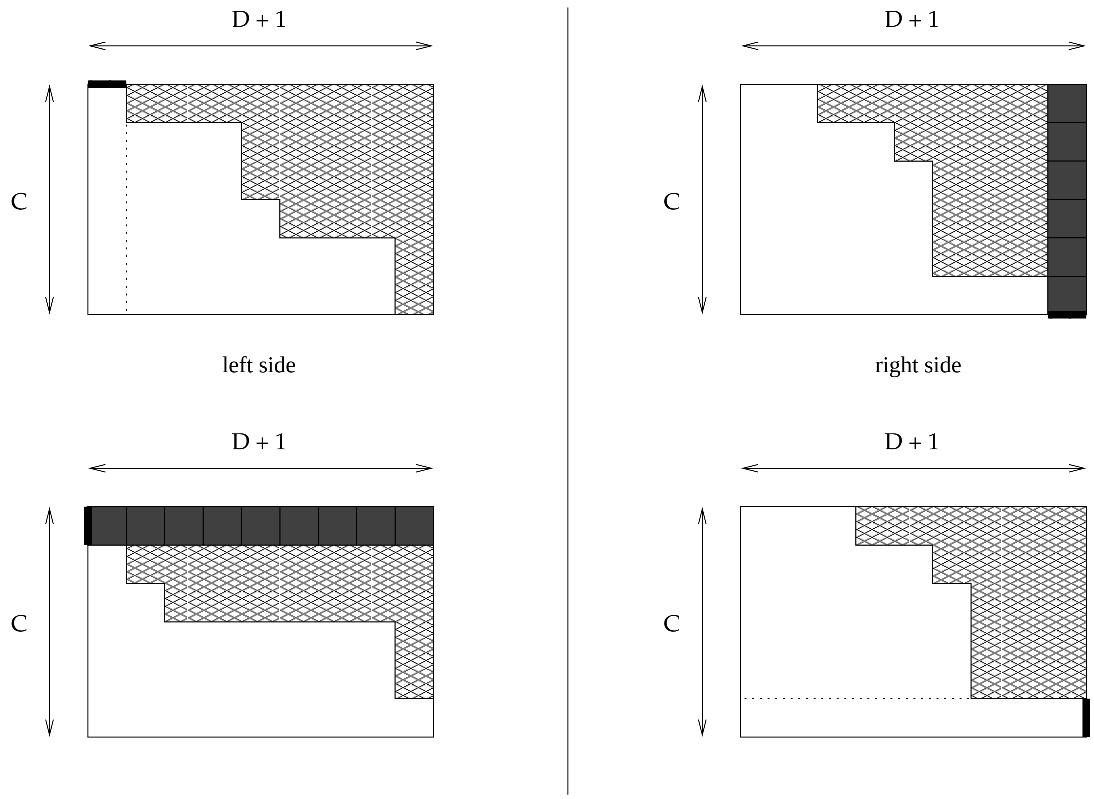


Figure 15: Picture used to prove (26).

Here, the power of q records the area above a path that goes from northwest to southeast in a rectangle of height C and width $D+1$. The left side classifies such paths by their *initial* step at the northwest corner. If this step is horizontal, the remainder of the path lies in a rectangle of height C and width D, giving the term $\begin{bmatrix} C+D \\ C,D \end{bmatrix}_q$. If this step is vertical, the remainder of the path lies in a rectangle of height $C-1$ and width $D+1$, giving the term $\begin{bmatrix} C+D \\ C-1,D+1 \end{bmatrix}_q$. However, we must also multiply by q^{D+1} to account for the $D+1$ area cells in the top row of the original rectangle.

The right side classifies the paths by their *final* step at the southeast corner. If this step is horizontal, the remainder of the path lies in a rectangle of height C and width D, giving the term $\begin{bmatrix} C+D \\ C,D \end{bmatrix}_q$. However, we must also multiply by q^C to account for the C

area cells in the rightmost column of the original rectangle. If the final step is vertical, the remainder of the path lies in a rectangle of height $C-1$ and width $D+1$, giving the term $\begin{bmatrix} C+D \\ C-1, D+1 \end{bmatrix}_q$.

Step 3. The next step is to check that the right side of (24) satisfies the specialized initial conditions

$$E^{(m)}_{N;N,0,\ldots,0}(q,1/q) = q^{mn(n-1)/2}$$
$$E^{(m)}_{N;0,0,\ldots,0}(q,1/q) = 0.$$

We leave this algebraic manipulation to the reader.

Step 4. The next step is to check that the right side of (24) satisfies the recursion (23) with t specialized to $1/q$. After setting $t = 1/q$ and simplifying, this recursion can be written

$$E^{(m)}_{N;v_0,\ldots,v_{m-1}}(q,1/q) = q^{\mathrm{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q E^{(m)}_{N-v_0;v_1,v_2\ldots,v_{m-1},i}(q,1/q), \tag{27}$$

where

$$\begin{aligned} \mathrm{pow} &= v_0 - N + m(v_0^2 - v_0)/2 + \sum_{k=1}^{m-1}(m-k)v_0 v_k \\ C &= v_0 + v_1 + \cdots + v_{m-1} - 1 \end{aligned}$$

The proof will be finished if we can show this same relation holds with the E's replaced by the appropriate formulas from the right side of (24). Specifically, write A' for $A(N, m, v_0, \ldots, v_{m-1})$, write P'_j for $P_j(N, m, v_0, \ldots, v_{m-1})$, and so forth. Write A'' for $A(N - v_0, m, v_1, \ldots, v_{m-1}, i)$, write P''_j for $P_j(N - v_0, m, v_1, \ldots, v_{m-1}, i)$, and so forth. Then we must show that the quantity

$$A'_0 - B'_1 - B'_2 - \cdots - B'_m \tag{28}$$

is equal to the quantity

$$q^{\mathrm{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q (A''_0 - B''_1 - B''_2 - \cdots - B''_m)\,. \tag{29}$$

To show this, we write the latter expression as the sum of $m+1$ smaller expressions, namely

$$q^{\mathrm{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q A''_0$$

and (for $1 \le j \le m$)

$$q^{\mathrm{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q (-B''_j).$$

Each of these $m+1$ expressions can be evaluated (see below) using the lemma from Step 1. The resulting sum is *almost* the desired quantity

$$A_0' - B_1' - B_2' - \cdots - B_m'.$$

More specifically, for $2 \leq j \leq m$, the expression involving $-B_j''$ will evaluate to $-B_{j-1}'$. On the other hand, the expression involving $-B_1''$ will evaluate to $-B_m'$ times an unwanted power of q. Similarly, the expression involving A_0'' will evaluate to A_0' times another unwanted power of q. Finally, the lemma from step 2 will show that these last two terms are in fact equal to $A_0' - B_m'$ without the unwanted powers of q! This will complete the proof of the formula (24).

Step 5. We indicate how to evaluate the expression

$$q^{\text{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q (A_0'') \tag{30}$$

from Step 4. The final answer will be $q^{-(N-v_0-\cdots-v_{m-1})}A_0'$.

One must first verify the algebraic identity

$$P_0'' + \text{pow} = P_0' - (N - v_0 - \cdots - v_{m-1}) + i\left[mN - \sum_{k=0}^{m-1}(m-k)v_k\right]. \tag{31}$$

Using this identity and expanding the definition of A_0'', the expression (30) can be written

$$q^{P_0'-(N-v_0-\cdots-v_{m-1})} \sum_{i=0}^{D-E} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q \begin{bmatrix} D-i \\ E, D-E-i \end{bmatrix}_q q^{i(E+1)},$$

where

$$\begin{aligned} D &= (m+1)N - 1 - \sum_{k=0}^{m-1}(m+1-k)v_k \\ E &= mN - 1 - \sum_{k=0}^{m-1}(m-k)v_k \\ D - E &= N - v_0 - v_1 - \cdots - v_{m-1} \end{aligned}$$

Using the identity from Step 1, this new expression becomes

$$q^{-(N-v_0-\cdots-v_{m-1})}q^{P_0'}\begin{bmatrix} C+D+1 \\ D-E, C+1+E \end{bmatrix}_q = q^{-(N-v_0-\cdots-v_{m-1})}A_0'.$$

Step 6. We indicate how to evaluate the expression

$$q^{\text{pow}} \sum_{i=0}^{N-v_0-\ldots-v_{m-1}} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q (-B_j'') \tag{32}$$

from Step 4. The answer will be $-B'_{j-1}$ for $j > 1$; it will be $-B' q^{P'_0-(N-v_0-\cdots-v_{m-1})}$ for $j = 1$.

The calculation is similar to the one in Step 5. Using the definition of B''_j and and the identity (31), we can rewrite (32) as

$$-q^{P'_0-(N-v_0-\cdots-v_{m-1})} \sum_{i=0}^{D-E+1} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q \begin{bmatrix} D-i \\ E, D-E-i \end{bmatrix}_q q^{iE} q^{P''_j}, \tag{33}$$

where we now set

$$\begin{aligned} D &= (m+1)N - 1 - \sum_{k=0}^{m-1} (m+1-k) v_k \\ E &= mN - \sum_{k=0}^{m-1} (m-k) v_k \\ D - E &= N - v_0 - v_1 - \cdots - v_{m-1} - 1. \end{aligned}$$

The summand where $i = D - E + 1$ is zero, so we may adjust the upper limit of the sum to be $i = D - E$ instead. To continue simplifying, one must first verify the identity

$$P'_{j-1} = P''_j - i - (N - v_0 - \cdots - v_{m-1}) \quad (j > 1).$$

Assume $j > 1$ first. Using the last identity to eliminate P''_j, the expression (33) becomes

$$-q^{P'_0+P'_{j-1}} \sum_{i=0}^{D-E} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q \begin{bmatrix} D-i \\ E, D-E-i \end{bmatrix}_q q^{i(E+1)}.$$

Using the identity from Step 1, the sum (without the outside power of q) evaluates to B'. Thus, when $j > 1$, the expression (32) evaluates to $-B'_{j-1}$ as claimed.

Now assume $j = 1$. Since $P''_1 = i$, the expression (33) becomes

$$-q^{P'_0-(N-v_0-\cdots-v_{m-1})} \sum_{i=0}^{D-E} \begin{bmatrix} C+i \\ C, i \end{bmatrix}_q \begin{bmatrix} D-i \\ E, D-E-i \end{bmatrix}_q q^{i(E+1)}.$$

Using the identity from Step 1, this becomes

$$-q^{P'_0-(N-v_0-\cdots-v_{m-1})} B'$$

as claimed.

Step 7. Let us recap the preceding calculations. We have evaluated the expression (29), hoping to obtain the answer

$$(A'_0 - B'_m) - B'_1 - B'_2 - \cdots - B'_{m-1}$$

from (28). Instead, we obtained the answer

$$q^{-(N-v_0-\cdots-v_{m-1})}(A'q^{P'_0}-B'q^{P'_0})-B'_1-B'_2-\cdots-B'_{m-1}.$$

Now, use the identity from Step 2, setting

$$\begin{aligned} C &= N-v_0-\cdots-v_{m-1} \\ D &= mN-1-\sum_{k=0}^{m-1}(m-1-k)v_k. \end{aligned}$$

The result is

$$q^{-(N-v_0-\cdots-v_{m-1})}(A'-B')=A'-q^{P'_m}B',$$

since one can check that $P'_m=D-C+1$ here. Multiplying by $q^{P'_0}$, we see that

$$q^{-(N-v_0-\cdots-v_{m-1})}(A'q^{P'_0}-B'q^{P'_0})=(A'_0-B'_m),$$

so that (29) does indeed evaluate to the desired answer (28). This completes the proof.

Proving the Formula for $C_n^{(m)}(q,1/q)$. We are now ready to verify that

$$q^{mn(n-1)/2}C_n^{(m)}(q,1/q)=\frac{1}{[mn+1]_q}\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q. \tag{34}$$

From (19) with $t=1/q$, we have

$$C_n^{(m)}(q,1/q)=q^{mn}E^{(m)}_{n+1;1,0,\dots,0}(q,1/q).$$

Now, we use the formula just proved for the E's with $N=n+1$, $v_0=1$, and $v_i=0$ for $i>0$. The reader may verify that, with these substitutions, we obtain

$$q^{mn(n-1)/2}C_n^{(m)}(q,1/q)=q^{n-nm}\left\{\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q-\begin{bmatrix} mn+n \\ n-1,mn+1 \end{bmatrix}_q\cdot\left(\sum_{j=0}^{m-1}q^{nj+\chi(j=m-1)}\right)\right\}.$$

The expression in the curly braces can be written

$$\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q\cdot\left(1-\frac{[n]_q\sum_{j=0}^{m-1}q^{nj+\chi(j=m-1)}}{[mn+1]_q}\right),$$

which in turn simplifies to

$$\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q\cdot\left(\frac{\sum_{k=0}^{mn}q^k-\sum_{k=0}^{mn}q^k\chi(k\neq mn-n)}{[mn+1]_q}\right)=\frac{1}{[mn+1]_q}\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q q^{mn-n}.$$

The leftover power of q is exactly what is needed to cancel the outside power q^{n-nm}. Thus, we obtain the desired result

$$q^{mn(n-1)/2}C_n^{(m)}(q,1/q)=\frac{1}{[mn+1]_q}\begin{bmatrix} mn+n \\ mn,n \end{bmatrix}_q.$$

3.4 Recursions for $C_n^{(m)}(q,t)$ based on removing the last row

We now present one more recursion (36) that is not based directly on formula (12). This recursion is simpler in form than (23) because it has only four terms. However, one must keep track of several new statistics in this recursion.

We need to introduce some temporary notation. Let D be an m-Dyck path of height n. Let the bounce path of D have successive vertical moves $(v_0, v_1, \dots, v_s)$ and horizontal moves $(h_0, h_1, \dots)$ as usual. Here, v_s is the last nonzero vertical move. Define

$$\begin{aligned}
Q(D) &= \text{area}(D);\\
T(D) &= b(D);\\
Y(D) &= s;\\
Z_i(D) &= v_{s-i} \text{ for } i \geq 0;\\
K(D) &= \text{the total number of area cells in the top row of } D;\\
W(D) &= \text{the number of area cells in the top row of } D\\
& \text{left of the last vertical move of the bounce path.}
\end{aligned}$$

Thus, $Y(D)$ is one less than the total number of bounces needed to reach the top rim; the statistics $Z_i(D)$ record the history of vertical moves near the end of the bounce path; and $W(D)$ counts the number of "extra" cells in the top row left of the bounce path. Define $\mathcal{D}_{n,k}^{(m)}$ to be the collection of paths $D \in \mathcal{D}_n^{(m)}$ with $K(D) = k$, for $0 \leq k \leq m(n-1)$.

For example, the 2-Dyck path E in Figure 6 has $Q(E) = 41$, $T(E) = 30$, $Y(E) = 5$, $Z_0(E) = 3$, $Z_1(E) = 1$, $W(E) = 1$, and $K(E) = 8$. The 3-Dyck path D in Figure 7 has $Q(D) = 23$, $T(D) = 29$, $Y(D) = 8$, $Z_0(D) = 1$, $Z_1(D) = 1$, $Z_2(D) = 0$, $W(D) = 0$, and $K(D) = 5$.

Now, define

$$C_{n,k}(q,t,y,z_0,\dots,z_{m-1},w) = \sum_{D \in \mathcal{D}_{n,k}^{(m)}} q^{Q(D)} t^{T(D)} y^{Y(D)} w^{W(D)} \prod_{i=0}^{m-1} z_i^{Z_i(D)}. \tag{35}$$

(We suppress the dependence on m from the notation.) If $k = m(n-1)$, there is only one path $D_0 \in \mathcal{D}_{n,m(n-1)}^{(m)}$, which goes north n steps and then east mn steps. Thus, we obtain the initial condition

$$C_{n,m(n-1)}(q,t,y,z_0,\dots,z_{m-1},w) = q^{mn(n-1)/2} z_0^n,$$

since, by inspection, D_0 has area $mn(n-1)/2$ and a single nontrivial bounce of height n.

Write $\vec{z}$ to denote $(z_0,\dots,z_{m-1})$. We will show combinatorially that, for $0 \leq k < m(n-1)$,

$$\begin{aligned}
C_{n,k}(q,t,y,\vec{z},w) = {} & z_0 q^k C_{n-1,k-m}(q,t,ty,\vec{z},w)\\
& + q^{-1} w^{-1} \left(C_{n,k+1}(q,t,y,\vec{z},w) - C_{n,k+1}(q,t,y,\vec{z},0) \right)\\
& + q^{-1} t y z_0 z_1^{-1} w^{-2} C_{n,k+1}(q,t,y,wz_1,wz_2,\dots,wz_{m-1},w,0).
\end{aligned} \tag{36}$$

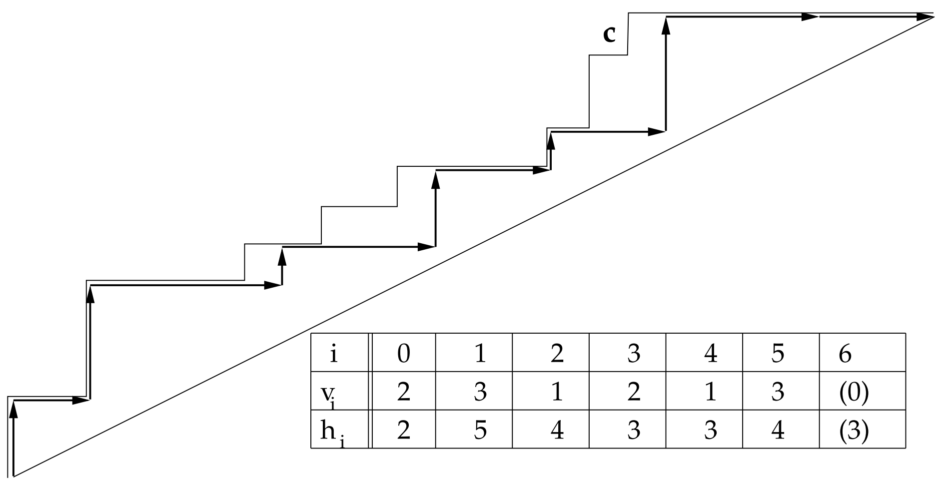
With the initial condition, this recursion uniquely determines the multivariable generating functions $C_{n,k}$ (by induction on n and backwards induction on k).

To prove this recursion, we classify a path $D \in \mathcal{D}_{n,k}^{(m)}$ based on what happens at the left edge of the top row of D. Exactly one of the following three cases must occur:

- **Case 1:** The path D reaches the top row by taking two consecutive north steps. See Figure 6 for an example.

- **Case 2:** The path D reaches the top row by taking a north step preceded by an east step, AND this east step did not block the progress of the next-to-last vertical bounce move. This means that adding one more area cell to the top row of D would not change the derived bounce path. See Figure 16 for an example.

- **Case 3:** The path D reaches the top row by taking a north step preceded by an east step, AND this east step did block the progress of the next-to-last vertical bounce move. This means that adding one more area cell to the top row of D would enable the next-to-last bounce to reach the top rim, so that the total number of bounces would decrease by one. See Figure 7 for an example.



i	0	1	2	3	4	5	6
v_i	2	3	1	2	1	3	(0)
h_i	2	5	4	3	3	4	(3)

m = 2, n = 12, area(D) = 42, b(D) = 30

Figure 16: A path satisfying case 2 in the recursion analysis.

The three terms on the right side of (36) are the respective generating functions for the paths in the three cases above.

To see this, first consider paths satisfying Case 1. We can uniquely construct each such path D by first picking a path D' of height $n-1$ with $k-m$ area cells in row $n-1$, and then placing k new area cells in row n to obtain D. See Figures 17 and 8 (where $D = E$). The generating function for the choice of D' is $C_{n-1,k-m}(q,t,y,\vec{z},w)$. Adding the new row influences the statistics as follows. The power of q increases by k since we added k new area cells. Let $(v'_0, \ldots, v'_{s'})$ be the vertical moves in the bounce path for D'. It is clear from Figure 17 that the bounce path of D will have vertical moves $(v_0, \ldots, v_s)$, where $s = s'$, $v_i = v'_i$ for $i < s$, and $v_s = v'_{s'} + 1$. Since only the last vertical move changed, all horizontal moves before reaching the top rim are the same. Since $v_s = v'_{s'} + 1$, the power of z_0 should increase by one when we pass from D' to D. Since $v_i = v'_i$ for $i < s$, the powers of $z_1, z_2, \ldots$ should not change. Similarly, since $s = s'$, the power of y does not change in the passage from D' to D. The power of w does not change either, since there are the same number of extra cells left of the last vertical move after adding the new row. Finally, we have $b(D) = \sum_{i \geq 0} iv_i = \sum_{i \geq 0} iv'_i + s = b(D') + s$, since $v_s = v'_s + 1$. We can increase the power of t in the generating function by exactly s if we replace y by ty in $C_{n-1,k-m}(q, t, y, \vec{z}, w)$. To see this, recall that $Y(D') = s' = s$ and compare to definition (35) [with D there replaced by D']. Putting all this together, we see that the generating function for paths in Case 1 is precisely $z_0 q^k C_{n-1,k-m}(q, t, ty, \vec{z}, w)$.

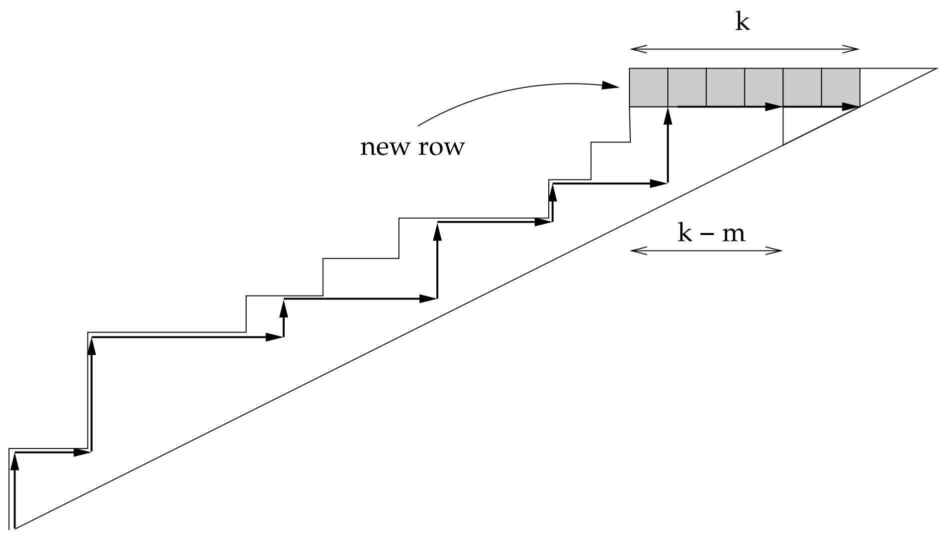


Figure 17: Constructing a path in Case 1 by adding a row.

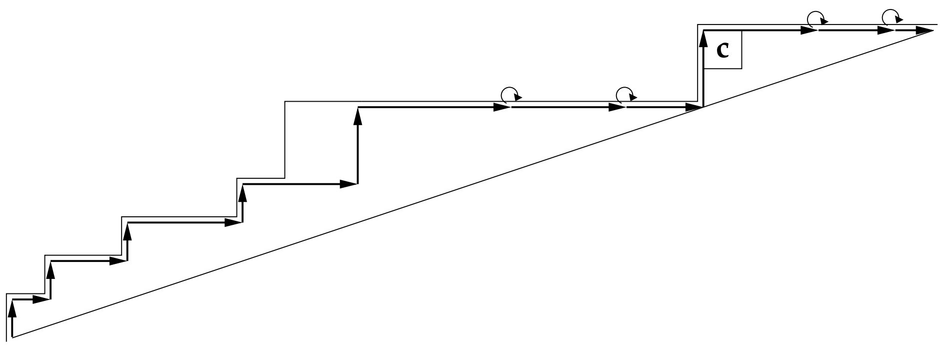
We will treat the next two cases together. Note that all paths D satisfying Case 2 or 3 can be uniquely constructed by choosing a path $D' \in \mathcal{D}^{(m)}_{n,k+1}$ and then *removing* the leftmost area cell in the top row of D'. The generating function for the paths D' is $C_{n,k+1}(q, t, y, \vec{z}, w)$. However, to determine the effect of the cell removal on the bounce statistic, we must know whether the removed cell was an "extra" cell or one that was part of the bounce path. This complication forces the introduction of two separate cases.

If $w(D') = 0$, then D' has no extra area cells in its top row. The path D constructed from D' therefore belongs to case 3. Consider the definition (35) with D replaced by D' and k replaced by $k+1$. If we substitute $w = 0$ in that definition (with the usual convention that $0^0 = 1$), we are left with the generating function for just those paths D' with $w(D') = 0$. By the sum rule, the generating function for just those paths D' with $w(D') > 0$ must be $C_{n,k+1}(q, t, y, \vec{z}, w) - C_{n,k+1}(q, t, y, \vec{z}, 0)$.

In case 2, we start with a path D' counted by the latter generating function. For example, D' could be the path D shown in Figure 16 with the cell c adjoined. To go from D' to D, we remove the cell in position c. This clearly decreases $Q(D')$ and $W(D')$ by 1, but does not affect the other statistics that are determined by the bounce path. It immediately follows that the generating function for the paths D in case 2 is

$$q^{-1}w^{-1}\left(C_{n,k+1}(q,t,y,\vec{z},w) - C_{n,k+1}(q,t,y,\vec{z},0)\right)$$

To get a path D belonging to case 3, on the other hand, we must have started with a path D' such that $w(D') = 0$. For example, the path D' in Figure 18 is used to construct the path D in Figure 7.



i	0	1	2	3	4	5	6	7	8	9
v_i	1	1	1	1	2	0	0	2	(0)	(0)
h_i	1	2	3	3	4	3	2	2	(2)	(2)

Figure 18: Constructing a path in Case 3 by deleting one cell.

The generating function for the choice of D' is $C_{n,k+1}(q,t,y,\vec{z},0)$. We obtain D from D' by removing the leftmost area cell c in the top row of D'. To see how this affects the statistics, compare Figure 18 to Figure 7. Clearly, the area $Q(D) = Q(D') - 1$ because of the removed cell. Let $(v'_0,\ldots,v'_{s'})$ be the lengths of the vertical moves in the bounce path for D'; let $(v_0,\ldots,v_s)$ be the lengths of the vertical moves in the bounce path for D. In this case, removing the cell forces the last vertical move in D' to be shortened by 1 unit, so that there must be a new vertical move of length 1 afterwards in D. Thus, $v_i = v'_i$ for $i < s'$, $v_{s'} = v'_{s'} - 1$, $s = s'+1$, and $v_s = 1$. We find that $b(D) - b(D') = (s'+1)\cdot 1 - s'\cdot 1 = 1$, so that the bounce statistic has increased by 1. We also have $Y(D) = Y(D')+1$, $Z_0(D) = 1$, $Z_1(D) = Z_0(D') - 1$, and $Z_i(D) = Z_{i-1}(D')$ for $i \geq 2$. Finally, we must compute the new value $W(D)$. After the bounce path for D takes the vertical step of length $v_{s'} = v'_{s'} - 1$ (this step is blocked by the east step introduced by the removed cell), the bounce path moves east

$$Z_1(D) + Z_2(D) + \cdots + Z_m(D) = Z_0(D') - 1 + Z_1(D') + \cdots + Z_{m-1}(D') \text{ units.}$$

All the area cells above this horizontal move were present in D'; in D, all these cells exist except the leftmost cell c. This implies that

$$W(D) = Z_0(D') + \cdots + Z_{m-1}(D') - 2.$$

Consider the last term on the right side of (36):

$$q^{-1}tyz_0z_1^{-1}w^{-2}C_{n,k+1}(q,t,y,wz_1,wz_2,\ldots,wz_{m-1},w,0).$$

By the definition in (35) and the comments above,

$$C_{n,k+1}(q,t,y,z_0,\ldots,z_{m-1},0) = \sum_{D' \text{ as in Case 3}} q^{Q(D')}t^{T(D')}y^{Y(D')}\prod_{i=0}^{m-1} z_i^{Z_i(D')}.$$

Therefore, making the indicated substitutions for the variables,

$$\begin{aligned}
&q^{-1}tyz_0z_1^{-1}w^{-2}C_{n,k+1}(q,t,y,wz_1,wz_2,\ldots,wz_{m-1},w,0)\\
&= \sum_{D'} q^{Q(D')-1}t^{T(D')+1}y^{Y(D')+1}z_0^1 z_1^{Z_0(D')-1}z_2^{Z_1(D')}\cdots z_{m-1}^{Z_{m-2}(D')}w^{\text{pow}}\\
&\quad(\text{pow} = Z_0(D') + \cdots + Z_{m-2}(D') + Z_{m-1}(D') - 2)\\
&= \sum_{D} q^{Q(D)}t^{T(D)}y^{Y(D)}z_0^{Z_0(D)}\cdots z_{m-1}^{Z_{m-1}(D)}w^{W(D)},
\end{aligned} \tag{37}$$

where the sums extend over the paths D' and D appearing in the description of case 3 above. Thus, the third term in (36) is the correct generating function for the paths belonging to case 3. This completes the proof of the recursion.

Remarks. The given recursion (36) keeps track of the last m vertical bounces $Z_0(D)$, ... , $Z_{m-1}(D)$. This is necessary to determine what happens to the other statistics in certain cases. Though it is not necessary here, we clearly could add even more variables $z_m,\ldots$ to keep track of the earlier bounce moves $Z_m(D),\ldots$ if we wished. Later (§4), we shall consider a more general recursion in which it becomes necessary to keep track of $Z_m(D)$.

We remark that a similar recursion can be proved for a suitable generalization of $HC_n^{(m)}(q,t)$. We do not give the details of the proof, which are quite messy, but merely list the appropriate reinterpretations of the statistics. In this setting, one should take

$$\begin{aligned}
Q(D) &= h(D);\\
T(D) &= \text{area}(D);\\
Y(D) &= \max_{0\le i<n} \gamma_i(D);\\
Z_i(D) &= |\{j : \gamma_j(D) = Y(D) - i\}| \text{ for } i \ge 0;\\
K(D) &= h(D) - h(D'), \text{where } D' \text{ is obtained from } D \text{ by}\\
&\quad\ \text{removing the rightmost value } Y(D);\\
W(D) &= \text{the number of symbols in } \{Y(D)-1,\ldots,Y(D)-m\} \text{ appearing}\\
&\quad\ \text{in } \gamma(D) \text{ after the last occurrence of } Y(D).
\end{aligned}$$

This gives an alternate way of proving that $C_n^{(m)}(q,t) = HC_n^{(m)}(q,t)$.

4 Trivariate Catalan Sequences

We now introduce three-variable sequences $C_n^{(m)}(q,t,r)$ that generalize the higher q,t-Catalan sequences. Our point of departure is the observation that

$$r^{ab}\begin{bmatrix} a+b \\ a,b \end{bmatrix}_{q/r} = \sum_{\lambda\subset R_{a,b}} q^{|\lambda|} r^{ab-|\lambda|} = \sum_{w\in R(0^a1^b)} q^{\mathrm{coinv}(w)} r^{\mathrm{inv}(w)}.$$

In other words, given a lattice path from the southwest corner to the northeast corner of the rectangle $R_{a,b}$, we can keep track of both the area in $R_{a,b}$ below the path and the area in $R_{a,b}$ above the path by making the indicated substitution in the q-binomial coefficient. For convenience, set

$$\begin{bmatrix} a+b \\ a,b \end{bmatrix}_{q,r} = r^{ab}\begin{bmatrix} a+b \\ a,b \end{bmatrix}_{q/r}.$$

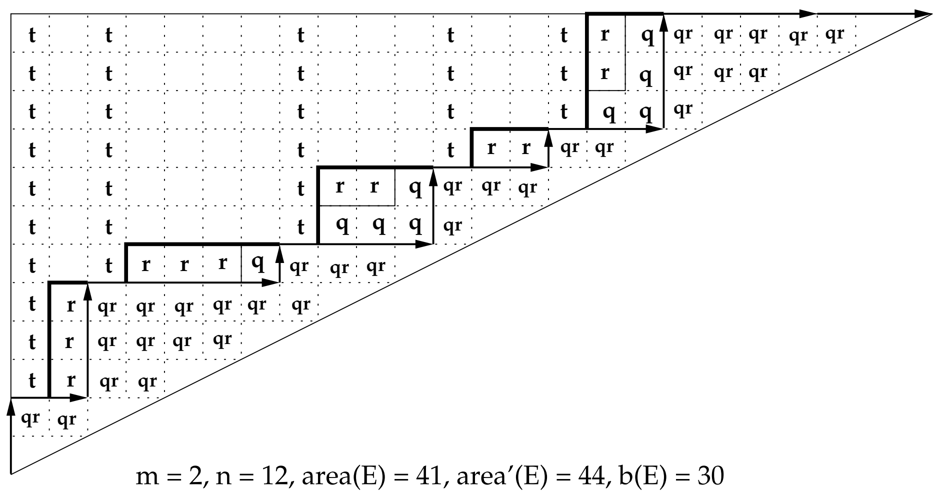


Figure 19: Visualizing the three statistics as counting cells.

We introduce a new statistic area′ on m-Dyck paths D of height n as follows. Given D, draw the bounce path of D and the associated rectangles R_i as in Figure 8. Let R'_i denote the rectangle R_i without its leftmost column. Define $\mathrm{area}'(D)$ to be the number of complete cells below the bounce path of D plus the number of cells inside the rectangles R'_i and *above* the path D. By contrast, $\mathrm{area}(D)$ is the number of complete cells below the bounce path of D plus the number of cells inside the rectangles R'_i and *below* the path D. For each m and n, define

$$C_n^{(m)}(q,t,r) = \sum_{D\in\mathcal{D}_n^{(m)}} q^{\mathrm{area}(D)} t^{b(D)} r^{\mathrm{area}'(D)}.$$

See Figure 19 for an example.

In this figure, cells below the bounce path contributing to both area and area$'$ are labelled by their weight qr. Cells above the bounce path but below the m-Dyck path contribute only to area and are labelled q. Cells inside the rectangles R_i' but above the m-Dyck path are labelled r. Finally, Figure 19 shows how we can interpret the bounce statistic $b(D)$ as counting certain cells in the picture as well. Specifically, we label each cell in the column above a vertical bounce move with t. Equation (9) shows that the number of such factors t is exactly $b(D)$.

From Figure 19, we immediately deduce the symmetry result

$$C_n^{(m)}(q,t,r) = C_n^{(m)}(r,t,q).$$

For, we can interchange the number of cells labelled q and the number of cells labelled r by merely rotating the contents of each shortened rectangle R_i' by 180°. Note that this rotation will not affect the bounce path, since it does not affect the leftmost columns of the full rectangles R_i. The image of the path in Figure 19 under this involution is shown in Figure 20.

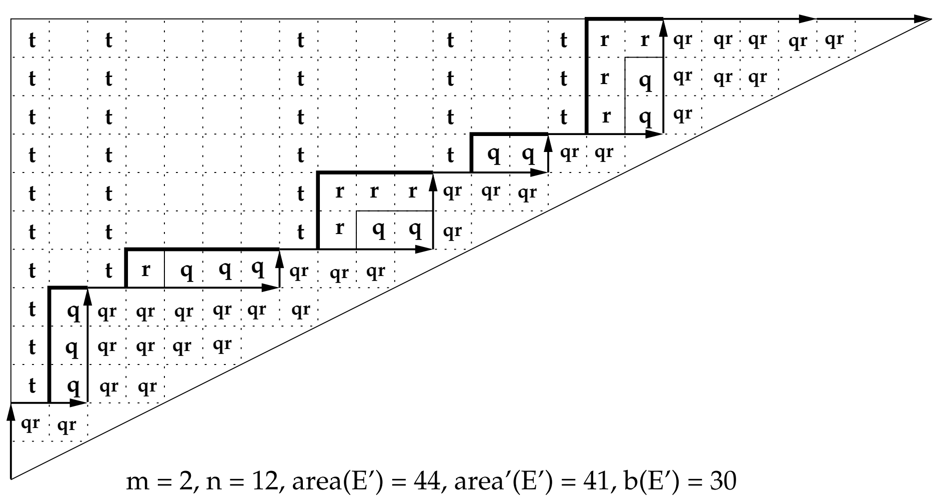


Figure 20: Interchanging area and area$'$ by flipping rectangles.

It is easy to incorporate area$'$ into formula (12). We have

$$C_n^{(m)}(q,t,r) = \sum_{v\in\mathcal{V}_n^{(m)}} t^{\sum_{i\geq 0} iv_i} (qr)^{m\sum_{i\geq 0}\frac{1}{2}v_i(v_i-1)} \prod_{i\geq 1} (qr)^{v_i\sum_{j=1}^{m}(m-j)v_{i-j}} \begin{bmatrix} v_i+v_{i-1}+\cdots+v_{i-m}-1 \\ v_i, v_{i-1}+\cdots+v_{i-m}-1 \end{bmatrix}_{q,r}. \tag{38}$$

The new formula follows by recalling that the factors $\begin{bmatrix} v_i+v_{i-1}+\cdots+v_{i-m}-1 \\ v_i, v_{i-1}+\cdots+v_{i-m}-1 \end{bmatrix}_q$ keep track of the area cells below the path in the rectangles R_i', whereas the remaining powers of q in (12) count the cells below the bounce path. Hence, to keep track of area$'$, it suffices to replace the latter occurrences of q by qr and to use q,r-binomial coefficients in place of q-binomial coefficients.

The recursion in §3.2 is also easily modified. Define

$$E_{n;v_0,\ldots,v_{m-1}}(q,t,r) = \sum_{(v_m,v_{m+1},\ldots)} t^{\sum_{i\geq 0} iv_i}(qr)^{\mathrm{pow}_1} r^{\mathrm{pow}_2} \prod_{i\geq m} \begin{bmatrix} v_i+v_{i-1}+\cdots+v_{i-m}-1 \\ v_i, v_{i-1}+\cdots+v_{i-m}-1 \end{bmatrix}_{q,r}, \tag{39}$$

where

$$\mathrm{pow}_1 = m\sum_{i\geq 0}\frac{1}{2}v_i(v_i-1)+\sum_{i\geq 1}v_i\sum_{j=1}^{m\wedge i}(m-j)v_{i-j},$$

$$\mathrm{pow}_2 = \sum_{i=1}^{m-1} v_i\left(\left(\sum_{j=0}^{i-1}v_j\right)-1\right).$$

The extra power r^{pow_2} accounts for the cells in the first $m-1$ rectangles R'_i, which all contribute to the r-statistic.

We have the recursion

$$E_{n;v_0,\ldots,v_{m-1}}(q,t,r) = t^{n-v_0}(qr)^{\mathrm{pow}_3} \sum_{v_m=0}^{n-v_0-\cdots-v_{m-1}} r^{\mathrm{pow}_4} \begin{bmatrix} v_m+\cdots+v_0-1 \\ v_m, v_{m-1}+\cdots+v_0-1 \end{bmatrix}_{q,r} E_{n-v_0;v_1,\ldots,v_{m-1},v_m}(q,t,r), \tag{40}$$

where

$$\mathrm{pow}_3 = m\binom{v_0}{2}+\sum_{i=1}^{m-1}v_0v_i(m-i),$$

$$\mathrm{pow}_4 = v_0(v_1+\cdots+v_{m-1})-v_1-v_m(v_{m-1}+\cdots+v_0-1).$$

The initial condition is

$$E_{n;n,0,\ldots,0}(q,t,r) = (qr)^{mn(n-1)/2}t^0.$$

The recursion from §3.4 requires a bit more work. For an m-Dyck path D, define $R(D) = \mathrm{area}'(D)$, and set

$$C_{n,k}(q,r,t,y,z_0,\ldots,z_m,w) = \sum_{D\in\mathcal{D}^{(m)}_{n,k}} q^{Q(D)}r^{R(D)}t^{T(D)}y^{Y(D)}w^{W(D)}\prod_{i=0}^{m} z_i^{Z_i(D)}. \tag{41}$$

Observe that this generating function, unlike the original, keeps track of $Z_m(D)$ as well as $Z_i(D)$ for $i<m$. We need to make one technical adjustment in the definition of Z_m. If D_0 is the special path that goes north n steps and east mn steps, set $Z_m(D_0)=1$; for all other paths, define $Z_m(D)$ as in §3.4.

With this adjustment, the initial condition is

$$C_{n,k}(q,r,t,y,\vec{z},w) = (qr)^{mn(n-1)/2}z_0^n z_m^1 \text{ when } k=m(n-1).$$

The new recursion, valid for $0 \le k < m(n-1)$, is:

$$\begin{aligned}
&C_{n,k}(q,r,t,y,\vec{z},w) = \\
&z_0 q^k r^{k-1} C_{n-1,k-m}(q,r,t,ty,z_0,rz_1,\dots,rz_m,r^{-2}w) \\
&\ + q^{-1}w^{-1}r^{+1}\left(C_{n,k+1}(q,r,t,y,\vec{z},w) - C_{n,k+1}(q,r,t,y,\vec{z},0)\right) \\
&\ + q^{-1}tyz_0z_1^{-1}w^{-2}r^2 C_{n,k+1}(q,r,t,y,r^{-1}wz_1,r^{-2}wz_2,\dots,r^{-2}wz_m,r^{-1},0).
\end{aligned} \tag{42}$$

To verify this equation, we need only check the correctness of the powers of r and z_m. We look at three cases, as in §3.4. In case 1, we go from $D' \in \mathcal{D}^{(m)}_{n-1,k-m}$ to $D \in \mathcal{D}^{(m)}_{n,k}$ by adding a new top row with k area cells. By definition, $Z_m(D') = Z_m(D)$. [Note that the technical adjustment made to $Z_m(D_0)$ has no effect here, since $k < m(n-1)$ implies that $(k-m) < m((n-1)-1)$, hence $D' \neq D_0$ and $D \neq D_0$.] What happens to area$'$ when we pass from D' to D? In the new top row, $k - W(D)$ of the k new area cells are below the bounce path for D, hence contribute to area$'$. The last rectangle R_s has also gained a new top row, which contains $h_{s-1} = v_{s-1} + \dots + v_{s-m}$ cells. Of these cells, the one in the leftmost column does not count towards area$'$, nor do the $W(D)$ new cells below the path D. These observations explain why we replace $z_1,\dots,z_m$ by $rz_1,\dots,rz_m$ (leaving z_0 alone) and multiply by r^{k-1} in the term

$$z_0 q^k r^{k-1} C_{n-1,k-m}(q,r,t,ty,z_0,rz_1,\dots,rz_m,r^{-2}w).$$

For, the net gain in the power of r is

$$Z_1(D')+\dots+Z_m(D')+k-1-2W(D') = v_{s-1}(D)+\dots+v_{s-m}(D)+(k-W(D))-(W(D)+1),$$

as required.

The term from Case 2, namely

$$q^{-1}w^{-1}r^{+1}\left(C_{n,k+1}(q,r,t,y,\vec{z},w) - C_{n,k+1}(q,r,t,y,\vec{z},0)\right),$$

is the easiest to derive. Recall that we go from D' to D by removing the leftmost ordinary area cell in the top row of D', which is not below the bounce path of D' or D. But "removing" this cell from D' causes the cell to contribute to area$'$ instead, since it belonges to one of the rectangles R' and is now above D. Thus, we have an extra factor r^{+1} in the generating function. As for z_m, note that $D' \neq D_0$ since $W(D') > 0 = W(D_0)$, and $D \neq D_0$ since $k < k+1 \le m(n-1)$. Thus, $Z_m(D') = Z_m(D)$.

Finally, consider the term from Case 3, namely

$$q^{-1}tyz_0z_1^{-1}w^{-2}r^2 C_{n,k+1}(q,r,t,y,r^{-1}wz_1,r^{-2}wz_2,\dots,r^{-2}wz_m,r^{-1},0).$$

In this case, we go from D' to D by removing the leftmost ordinary area cell c in the top row of D', causing a change in the end of the bounce path. (See Figures 18 and 7.) Specifically, the bounce path of D has a new terminating vertical move v_s of length 1, and the previous vertical move v_{s-1} is one less than the corresponding move v'_{s-1} in D'. Note that the top row of the last rectangle R^*_{s-1} in D' does not belong to the rectangle R_{s-1}

in D. Every cell in the top row of R^*_{s-1}, except the leftmost one, contributed to $R(D')$, because $w(D')=0$. The number of contributing cells is one less than the horizontal dimension of R^*_{s-1}; this dimension is $Z_1(D')+\cdots+Z_m(D')$. The conclusion is that $R(D)$ drops by

$$(Z_1(D')+\cdots+Z_m(D')-1) \tag{43}$$

as a result of the lost row in R_{s-1}.

On the other hand, consider cells in the top row of D' that are to the right of the bounce path in D'. After removing cell c from D', the new bounce path for D stops at the southwest corner of c, then goes east for a distance

$$Z_0(D)+\cdots+Z_{m-1}(D)=(Z_0(D')-1)+Z_1(D')+\cdots+Z_{m-1}(D'),$$

then goes north one unit. The cells in the top row above this last east step used to count towards $\text{area}'(D')$, being below the bounce path of D', but will no longer count towards $\text{area}'(D)$. In more detail, cell c does not count towards $\text{area}'(D)$ because it is in the leftmost column of its rectangle. The other cells do not count towards $\text{area}'(D)$ because they count towards ordinary area instead. We conclude that $R(D)$ drops by an additional

$$Z_0(D')+\cdots+Z_{m-1}(D')-1 \tag{44}$$

as a result of the change in this part of the bounce path. The total change is

$$R(D)-R(D')=-(1Z_0(D')+2Z_1(D')+\cdots+2Z_{m-1}(D')+1Z_m(D'))+2.$$

This change is modelled algebraically by the additional occurrences of r in the expression

$$q^{-1}tyz_0z_1^{-1}w^{-2}r^2C_{n,k+1}(q,r,t,y,r^{-1}wz_1,r^{-2}wz_2,\ldots,r^{-2}wz_m,r^{-1},0).$$

The argument in the last two paragraphs is correct, unless R^*_{s-1} has width zero. In this situation, there is no leftmost column in R^*_{s-1}, so we should not have subtracted 1 in (43). But it is easy to see that this situation occurs iff $D'=D_0$. Then our technical convention that $Z_m(D_0)=+1$ causes (43) to be correct after all, and the validity of (44) is not affected either. Since the new value $Z_m(D)$ comes from the old value $Z_{m-1}(D')$ (not from $Z_m(D')$), the technical convention for $Z_m(D_0)$ does not affect the correctness of the values of $Z_m(D)$ calculated using the recursion. This completes the proof of the new recursion.

Finally, we describe an analogous way of adding a third statistic to the other combinatorial sequence $HC_n^{(m)}(q,t)$. We can "guess" what this statistic should be by seeing what happens to area$'$ when we apply the bijection ψ from §2.5. We are led to the following formula. For an m-Dyck path of height n, define

$$h'(D)=\sum_{0\le i<j<n}\sum_{k=0}^{m-1}\chi\left(\gamma_i(D)-\gamma_j(D)+k\in\{-1,0,1,\ldots,m-1\}\right)-\sum_{i=0}^{n-1}\chi(\gamma_i(D)>0).$$

The first sum is similar to the one appearing in $h(D)$. The second sum that is subtracted may look surprising, but it arises from the fact that the leftmost column of each rectangle

R_i does *not* count towards area$'$. Note that the total number of cells in these columns is $n - v_0(\phi(D)) = \sum_{i=0}^{n-1} \chi(\gamma_i(D) > 0)$.

There is a formula analogous to (7) for $h'(D)$. Define $\mathrm{sc}'_m : \mathbb{Z} \to \mathbb{Z}$ by

$$\mathrm{sc}'_m(p) = \begin{cases} m - p & \text{for } 0 \leq p \leq m; \\ m + 1 + p & \text{for } -m \leq p \leq -1; \\ 0 & \text{for other } p. \end{cases} \tag{45}$$

Define $adj' : \mathbb{Z} \to \mathbb{Z}$ by $adj'(p) = -1$ for $p > 0$ and $adj'(p) = 0$ for other p. Then

$$h'(D) = \sum_{0 \leq i < j < n} \mathrm{sc}'_m(\gamma_i(D) - \gamma_j(D)) + \sum_{i=0}^{n-1} adj'(\gamma_i).$$

The proof is the same as the corresponding proof of (7).

Now, define

$$HC_n^{(m)}(q,t,r) = \sum_{D \in \mathcal{D}_n^{(m)}} q^{h(D)} t^{\mathrm{area}(D)} r^{h'(D)}.$$

We show that $HC_n^{(m)}(q,t,r)$ equals the right side of (38) by modifying the earlier proof to include r. It will follow that the bijection ϕ introduced in §2.5 maps the ordered triple of statistics (h, area, h') to the ordered triple $(\mathrm{area}, b, \mathrm{area}')$ (similarly for $\psi = \phi^{-1}$).

As in §2.4, we proceed by induction on the largest symbol s appearing in $\gamma(D)$. When $s = 0$, γ must consist of n zeroes, and $h'(D) = mn(n-1)/2$. This is the same as the power of r on the right side of (38).

For the induction step, it suffices to prove the following formula, which is the analogue of (15) for h':

$$h'(\gamma) - h'(\delta) = mv_s(v_s - 1)/2 + v_s \sum_{k=1}^{m} (m-k) v_{s-k} + \mathrm{inv}(w). \tag{46}$$

Here, $\gamma = \gamma(D)$ has largest symbol $s > 0$; v_i is the number of occurrences of i in γ for $0 \leq i \leq s$; δ is obtained from γ by erasing all the symbols s; and the word w records how to insert the v_s copies of s into δ to recover γ.

We still proceed by induction on $\mathrm{coinv}(w)$. If $\mathrm{coinv}(w) = 0$, all v_s copies of s were inserted into δ just after the last occurrence of any symbol in the set $\{s-1, \ldots, s-m\}$. The change $h'(\gamma) - h'(\delta)$ caused by this insertion is

$$\sum_{i<j} \mathrm{sc}'_m(\gamma_i - \gamma_j) - v_s$$

where the sum extends over all pairs (i,j) such that $\gamma_i = s$ or $\gamma_j = s$. We subtract v_s since we introduced v_s new positive entries (all equal to s) in γ.

First, consider the pairs (i,j) for which $i < j$ and $\gamma_i = s = \gamma_j$. There are $\binom{v_s}{2}$ such pairs, and each contributes $\mathrm{sc}'_m(s - s) = \mathrm{sc}'_m(0) = m$ to the h'-statistic. This gives the term $mv_s(v_s - 1)/2$ in (46).

Second, consider the pairs (i, j) for which $i < j$ and $\gamma_i = s$ and $\gamma_j \neq s$. Since all the copies of s in γ occur in a contiguous group following all instances of the symbols $s-1, \ldots, s-m$, and since s is the largest symbol appearing in γ, $j > i$ implies that $\gamma_j < s-m$. Then $\mathrm{sc}'_m(\gamma_i - \gamma_j) = 0$, since $\gamma_i - \gamma_j > m$. So these pairs contribute nothing to the h'-statistic.

Third, consider the pairs (i, j) for which $i < j$ and $\gamma_i \neq s$ and $\gamma_j = s$. Since s is the largest symbol, we have $\gamma_i < s$. Write $\gamma_i = s-k$ for some $k > 0$, and consider various subcases. Suppose $k \in \{1, 2, \ldots, m\}$. Then $\mathrm{sc}'_m(\gamma_i - \gamma_j) = \mathrm{sc}'_m(-k) = m+1-k$. For how many pairs (i, j) does it happen that $i < j$, $\gamma_i = s-k$, and $\gamma_j = s$? There are v_s choices for the index j and v_{s-k} choices for the index i; the condition $i < j$ holds automatically, since all occurrences of s occur to the right of all occurrences of $s-k$. Thus, we get a total contribution to the h'-statistic of $(m+1-k)v_s(v_{s-k})$ for this k. Adding over all k, we obtain

$$v_s \sum_{k=1}^{m} (m+1-k) v_{s-k} = v_s \sum_{k=1}^{m} (m-k) v_{s-k} + \sum_{k=1}^{m} v_s v_{s-k}.$$

On the other hand, if $k > m$, then $\mathrm{sc}'_m(\gamma_i - \gamma_j) = \mathrm{sc}_m(-k) = 0$, so there is no contribution to the h'-statistic.

Finally, recall that w is a rearrangement of v_s zeroes and $v_{s-1} + \cdots + v_{s-m} - 1$ ones. Since $\mathrm{coinv}(w) = 0$, all zeroes in w occur at the end, and hence

$$\mathrm{inv}(w) = v_s(v_{s-1} + \cdots + v_{s-m} - 1) = \left(\sum_{k=1}^{m} v_s v_{s-k} \right) - v_s.$$

Thus, the change $h'(\gamma) - h'(\delta)$ is precisely the expression on the right side of (46). So we are done when $\mathrm{coinv}(w) = 0$.

To finish the induction step, it suffices to show that replacing 10 by 01 in w decreases h' by one (since this replacement also decreases $\mathrm{inv}(w)$ by one). Let w' be the new word after the replacement, with corresponding vector γ' As in §2.4, we have

$$\text{original } \gamma = \ldots\ (s-j)\ z_1\ z_2\ \ldots\ z_\ell\ (s-k)\ s\ \ldots$$

where $0 \leq j \leq m$, $1 \leq k \leq m$, $\ell \geq 0$, and every $z_i < s-m$. Replacing 10 by 01 in w causes the s to move left, resulting in:

$$\text{new } \gamma' = \ldots\ (s-j)\ s\ z_1\ z_2\ \ldots\ z_\ell\ (s-k)\ \ldots.$$

Note that the symbol $s-j$ must exist, lest $\gamma'_0 = s > 0$.

Let us examine the effect of this motion on the h'-statistic. When we move the s left past its predecessor $s-k$ in γ, we get a net change in the h'-statistic of

$$\mathrm{sc}'_m(s - [s-k]) - \mathrm{sc}'_m([s-k] - s) = \mathrm{sc}'_m(k) - \mathrm{sc}'_m(-k) = -1,$$

since $1 \leq k \leq m$ (see (45)). As before, since $|s - z_i| > m$, moving the s past each z_i will not affect the h'-statistic at all. Thus, the total change in the h'-statistic is -1, as desired.

5 Open Problems and Recent Developments

There are several open problems involving the combinatorial polynomials $C_n^{(m)}(q,t)$. The main open problem is to prove that the conjectured combinatorial interpretation of the higher q,t-Catalan sequences is correct, i.e., that $OC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$. It may be possible to prove the equivalent assertion that $SC_n^{(m)}(q,t) = C_n^{(m)}(q,t)$ by proving that both sides satisfy the same recursion. A recursion characterizing $C_n^{(m)}(q,t)$ appears in §3.2. The first difficulty with this approach is finding the symmetric function formulas that correspond to the generating functions $E_{n;v_0,v_1,\ldots,v_{m-1}}(q,t)$ when some v_i (besides v_0) is nonzero.

Another, purely combinatorial problem is to prove that the q and t statistics introduced here for m-Dyck paths are jointly symmetric. This conjecture is only known to be true when $m = 1$ by invoking the long proof in [5, 6]. Here, we have only proved the weaker statement that the univariate distributions of the q and t statistics are the same.

Recall that the higher q,t-Catalan sequences all have (conjectural) interpretations in representation theory, symmetric function theory, and algebraic geometry. It would be interesting to find analogous interpretations for the trivariate q,t,r-Catalan sequences.

Remark: Since the initial submission of this article, a number of related combinatorial developments have appeared in the literature. The present author has generalized many of the combinatorial constructs presented here to paths inside certain trapezoidal shapes [17, 18]. Haglund, Haiman, Loehr, and Remmel introduced statistics on labelled Dyck paths and m-Dyck paths, which are conjectured to give the Hilbert series of certain doubly graded S_n-modules [13, 20, 19, 17]. The same four authors and A. Ulyanov recently conjectured a combinatorial formula for the monomial expansion of the Frobenius series of these same modules [12]. Using ideas from these papers, Haglund conjectured a combinatorial formula for the monomial expansion of the modified Macdonald polynomials $\tilde{H}_\mu$ [8]. This conjecture has been proved by Haglund, Haiman, and the present author [10, 11]. Finding combinatorial statistics for the Kostka-Macdonald coefficients (which arise in the Schur expansion of $\tilde{H}_\mu$) remains open. It seems likely that the new combinatorial formula for Macdonald polynomials may soon shed additional light on the conjectures in this paper and in [12].

Acknowledgement: The author thanks Jeffrey Remmel and James Haglund for many helpful discussions regarding this problem.

References

[1] E. Bender and S. G. Williamson, *Foundations of Applied Combinatorics.* Electronic version available online at `http://math.ucsd.edu/~ebender`.

[2] F. Bergeron and A. Garsia, "Science Fiction and Macdonald Polynomials," *CRM Proceedings and Lecture Notes AMS* VI **3** (1999), 363—429.

[3] 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.

[4] 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* VII **3** (1999), 363—420.

[5] A. Garsia and J. Haglund, "A proof of the q,t-Catalan positivity conjecture," *LACIM 2000 Conference on Combinatorics, Computer Science, and Applications* (Montreal), *Discrete Math.* **256** (2002), 677—717.

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

[7] A. Garsia and M. Haiman, "A remarkable q,t-Catalan sequence and q-Lagrange Inversion," *J. Algebraic Combinatorics* 5 (1996), no. 3, 191—244.

[8] J. Haglund, "A combinatorial model for the Macdonald polynomials," *Proc. Nat. Acad. Sci. USA*, **101** #46 (2004), 16127–16131.

[9] J. Haglund, "Conjectured Statistics for the q,t-Catalan numbers," *Advances in Mathematics* **175** (2003), 319—334.

[10] J. Haglund, M. Haiman, and N. Loehr, "A combinatorial formula for Macdonald polynomials," `arXiv:math.CO/0409538`, 2004.

[11] J. Haglund, M. Haiman, and N. Loehr, "Combinatorial Theory of Macdonald Polynomials I: Proof of Haglund's Formula," *Proc. Natl. Acad. Sci. USA*, in press.

[12] J. Haglund, M. Haiman, N. Loehr, J. Remmel, and A. Ulyanov, "A combinatorial formula for the character of the diagonal coinvariants," *Duke Math. J.* **126** #2 (2005), 195—232.

[13] J. Haglund and N. Loehr, "A Conjectured Combinatorial Formula for the Hilbert Series for Diagonal Harmonics." *Proceedings of FPSAC 2002 Conference* (Melbourne, Australia), to appear in *Discrete Mathematics*.

[14] 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.

[15] Mark Haiman. Personal communication.

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

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

[18] N. Loehr, "Trapezoidal Lattice Paths and Multivariate Analogues," *Adv. in Appl. Math.* **31** (2003), 597—629.

[19] N. Loehr, "Combinatorics of q, t-Parking Functions," to appear in *Adv. in Appl. Math.*.

[20] N. Loehr and J. Remmel, "Conjectured Combinatorial Models for the Hilbert Series of Generalized Diagonal Harmonics Modules," *Electronic Journal of Combinatorics* **11** (2004), R68. 64 pages.

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