The Major Index Specialization of the q,t-Catalan

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

September 9, 2004

Abstract

Various authors have defined statistics on Dyck paths that lead to generalizations of the Catalan numbers. Three such statistics are area, maj, and bounce. Haglund, who introduced the bounce statistic, gave an algebraic proof that $n(n-1)/2+\text{area}-\text{bounce}$ and maj have the same distribution on Dyck paths of order n. We give an explicit bijective proof of the same result.

1 Introduction

In [6], Garsia and Haiman introduced a two-variable analogue of the Catalan number called the q,t-Catalan. Haglund conjectured a combinatorial interpretation of the q,t-Catalan that involved enumerating Dyck paths relative to two statistics called area and bounce [10]. This conjecture was later proved by Garsia and Haglund [5]. In [10], Haglund proved that the statistic $n(n-1)/2+\text{area}-\text{bounce}$ had the same distribution as the major index statistic (maj) on Dyck paths of order n. Haglund's proof of this fact is a long algebraic manipulation, which shows that both distributions are given by the same explicit formula. Mark Haiman [11] posed the problem of finding a purely combinatorial proof of this result. We solve this problem by constructing a bijection on Dyck paths of order n that sends $n(n-1)/2+\text{area}-\text{bounce}$ to maj. We obtain our map by

*supported by a National Science Foundation Postdoctoral Research Fellowship.

assembling many standard bijections on lattice paths using a simple version of the Garsia-Milne Involution Principle that we call "combinatorial subtraction." Some authors refer to this technique as "sieve-equivalence"; see [1, 7, 8, 9, 12, 13, 14, 15] for discussion and applications.

The rest of the paper is organized as follows. Section 2 defines the statistics mentioned above for Dyck paths and similar statistics for lattice paths contained in rectangles. Section 3 sets up notation for discussing bijective proofs. Section 4 discusses the combinatorial subtraction principle. Section 5 reviews five well-known bijections on lattice paths from the literature. Section 6 shows how to assemble these bijections using the subtraction principle to obtain the desired map.

2 Statistics on Lattice Paths

A *lattice path* in a $c \times d$ rectangle is a path from $(0,0)$ to (c,d) consisting of c east steps and d north steps of length 1. Such a path can be encoded as a word with d zeroes (north steps) and c ones (east steps). If P is a lattice path with corresponding word $w = w_1 w_2 \cdots w_{c+d}$, we define the following statistics.

- $\text{inv}(P)$ is the number of pairs $i < j$ such that $w_i = 1$ and $w_j = 0$. Equivalently, $\text{inv}(P)$ is the area above the path in the rectangle with corners $(0,0)$, $(c,0)$, $(0,d)$, and (c,d).
- $\text{maj}(P)$ is the sum of all indices $i < c+d$ such that $w_i = 1$ and $w_{i+1} = 0$. Suppose we label the lattice points visited by P with the integers $0, 1, 2, \ldots$, starting at the origin. Then $\text{maj}(P)$ can also be defined as the sum of the labels of the "left-turns" where P goes east and then north.
- $a(P)$ is the number of pairs $i < j$ such that $w_i = 0$ and $w_j = 1$. This is the area below the path P in the rectangle with corners $(0,0)$, $(c,0)$, $(0,d)$, and (c,d).

For example, if P is the path encoded by $w = 0101110011$, then $c = 6$, $d = 4$, $\text{inv}(P) = 9$, $\text{maj}(P) = 2 + 6 = 8$, and $a(P) = 15$.

A *Dyck path of order n* is a lattice path in an $n \times n$ rectangle that never visits any point (x,y) with $y < x$. Such paths are contained in a triangle T_n with vertices $(0,0)$, $(0,n)$, and (n,n). Define the following statistics for a Dyck path D of order n.

- Set $\text{area}(D) = a(D) - n(n+1)/2$. This is the number of complete lattice cells between D and the diagonal line $y = x$.

- Define $\text{maj}(D)$ by the same formula used above for general paths.
- Define Haglund's statistic $\text{bounce}(D)$ by the following construction. A ball starts at the origin and moves north until blocked by an east step of D. The ball then moves east to the line $y = x$. The ball moves north again until blocked by the east step of D that starts on the vertical line now occupied by the ball. The ball then bounces east to the diagonal $y = x$. This process continues until the ball reaches (n, n). Call the east steps that block the ball *blocking east steps.* Define $\text{bounce}(D)$ to be the number of cells in T_n in the columns above the blocking east steps.

For example, if D is the Dyck path encoded by $w = 0010010011101101$, then $n = 8$, $\text{area}(D) = 10$, $\text{maj}(D) = 3+6+11+14 = 34$, and $\text{bounce}(D) = 6+2+1 = 9$. See Figure 4 below, where the cells contributing to $\text{bounce}(D)$ are marked by X's.

3 Notation for Bijective Proofs

In later sections, we will be assembling many simple bijections to create more complicated maps. Here we introduce some notation to help organize this assembly process.

3.1 Sets of Weighted Objects

A *set of weighted objects* consists of a set A and a weight function $\text{wt}_A : A \to \mathbb{Z}$. It is helpful, for technical reasons, to assume that the image of wt_A is bounded below and that $A_k = \{a \in A : \text{wt}_A(a) = k\}$ is finite for each integer k. In all our work, A itself will be finite, so these assumptions automatically hold. Define a *generating function operator* γ on sets of weighted objects by setting

$$\gamma(A) = \sum_{a \in A} q^{\text{wt}_A(a)}.$$

(The assumptions above guarantee that this expression makes sense at least formally.)

We will use capital letters like $A, B, C, \ldots$ to denote sets of weighted objects, suppressing mention of the associated weight functions. If the definition of such a set depends on some numerical parameters, the parameters may be listed in parentheses or omitted when clear from context. For instance, we let $R(c, d)$ be the set of lattice paths in a $c \times d$ rectangle weighted by the area statistic a. If c and d are fixed, we write R instead of $R(c, d)$.

It will be convenient to let 0 denote the empty set and 1 denote any one-point set $\{x\}$ with $\text{wt}(x) = 0$.

3.2 Operations on Weighted Sets

Let A and B be two sets of weighted objects. Define $A+B$ to be the set that is the disjoint union of A and B, with weights given by

$$\mathrm{wt}_{A+B}(a) = \mathrm{wt}_A(a) \text{ for } a \in A;$$

$$\mathrm{wt}_{A+B}(b) = \mathrm{wt}_B(b) \text{ for } b \in B.$$

Define AB to be the set $A \times B = \{(a,b) : a \in A, b \in B\}$, with weights given by

$$\mathrm{wt}_{AB}((a,b)) = \mathrm{wt}_A(a) + \mathrm{wt}_B(b).$$

Finite sums and products of weighted sets are defined similarly. If s is a fixed integer, we define $q^s A$ to be the set with the same objects as A and weight function

$$\mathrm{wt}_{q^s A}(a) = \mathrm{wt}_A(a) + s \text{ for } a \in A.$$

This notation is related to the ordinary algebra of generating functions by the operator γ. One easily checks that

$$\gamma\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \gamma(A_i),$$

$$\gamma\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \gamma(A_i),$$

$$\text{and } \gamma(q^s A) = q^s \gamma(A).$$

3.3 The Equality Symbol

Suppose C and D are expressions that represent two sets of weighted objects. We write $C = D$ if we have explicitly constructed a particular weight-preserving bijection $f : C \to D$ and its inverse $f^{-1} : D \to C$. Note that the notation $C = D$ does not just mean that there exist weight-preserving bijections from C to D; it means that we have chosen and described a specific such bijection f. We may write $C \stackrel{f}{=} D$ to emphasize this choice. Note that $C = D$ implies that $\gamma(C) = \gamma(D)$, but the converse is not true unless we can construct a suitable map f.

3.4 Bijective Algebra

We can now write down algebraic expressions that encode bijective manipulations of various weighted sets. Many rules of algebra extend to the present setting. For instance, $A \stackrel{f}{=} B$ and $B \stackrel{g}{=} C$ clearly imply $A \stackrel{g\circ f}{=} C$. We have $AB + AC \stackrel{h}{=} A(B + C)$ via a canonical bijection h that is essentially the identity map on pairs (a, b) and (a, c). Similarly, $A + B = B + A$ and $(A+B)+C = A+(B+C)$ via identity maps. The following lemma is easy but useful.

Lemma 1 (Combinatorial Addition, Multiplication, and Scaling). *Let A, B, C, D be sets of weighted objects such that $A \stackrel{f}{=} C$ and $B \stackrel{g}{=} D$. Then there are canonical bijections*

$$A + B \stackrel{f+g}{=} C + D, \tag{1}$$

$$AB \stackrel{(f,g)}{=} CD. \tag{2}$$

$$q^s A \stackrel{f}{=} q^s C. \tag{3}$$

Proof. For (1), define $f+g : A+B \to C+D$ to be f on A and g on B. For (2), let $(f, g) : AB \to CD$ map (a, b) to $(f(a), g(b))$. For (3), just note that f is still weight-preserving if we add s to all weights in both A and C. □

Note that we can convert bijective proofs given in this notation to algebraic proofs of the same results by applying the γ operator.

Notation	Set of Objects	Weight
$R(c,d)$	paths in $c \times d$ rectangle	a (area below)
$R'(c,d)$	paths in $c \times d$ rectangle	inv (area above)
$M(c,d)$	paths in $c \times d$ rectangle	maj
$F(n,k)$	Dyck paths of order n starting with exactly k north steps	area − bounce $+(n^2+n)/2 - nk$
$G(n,k)$	Dyck paths of order n ending with exactly k east steps	maj $+ 2k - 2n$
$H(m,n)$	Dyck paths of order n ending with at least $n-m$ east steps	maj
$S(n)$	all Dyck paths of order n	$\binom{n}{2}$ + area − bounce
$T(n)$	all Dyck paths of order n	maj

Table 1: Notation used for sets of weighted lattice paths.

3.5 Some Sets of Weighted Lattice Paths

Table 1 defines the notation we will use for certain sets of weighted lattice paths.

There are a number of simple relations between some entries of the table. For instance, $R(c,d) \stackrel{\text{flip}(c,d)}{=} R'(c,d)$ where flip(c,d) rotates a path 180° about the center of the $c \times d$ rectangle. Similarly, $R(c,d) \stackrel{\text{ref}}{=} R'(d,c)$ by reflecting a path inside a $c \times d$ rectangle through the line $y = x$ to produce a path in a $d \times c$ rectangle.

Our main goal is to construct a bijection $S(n) \stackrel{h(n)}{=} T(n)$ for each n. This will be a byproduct of the stronger result $F(n,k) \stackrel{h(n,k)}{=} G(n,k)$. To see why, note that we have $S(n) = F(n+1,1)$ via a map that adds a north step and an east step to the beginning of a Dyck path of order n (which increases bounce by n). We also have $T(n) = G(n+1,1)$ via a map that adds a north step and an east step to the end of a Dyck path of order n (which increases maj by $2n$). Both maps preserve the weights listed in Table 1. Thus, we can construct $h(n)$ by composing these maps with $h(n+1,1)$.

4 The Combinatorial Subtraction Principle

If x, y, z are numbers such that $x + y = x + z$, then $y = z$. The combinatorial subtraction principle stated next is a bijective version of this manipulation. It is one of the simplest instances of the general method of "sieve-equivalence." See [1, 7, 8, 9, 12, 13, 14, 15] for more examples and applications of this technique.

Theorem 2 (Combinatorial Subtraction). *Suppose that* $A+B \stackrel{f}{=} A+C$ *and* A *is a finite set. Then* $B \stackrel{g}{=} C$ *via a map* $g : B \to C$ *that can be canonically constructed from* f.

Proof. We define a weight-preserving bijection $g : B \to C$ as follows. Given an element $b \in B$, compute a sequence of elements

$$x_1 = f(b),\ x_2 = f(f(b)),\ x_3 = f(f(f(b))),\ x_4 = f(f(f(f(b)))),\ \cdots$$

until an element x_i belonging to C is reached for the first time; define $g(b) = x_i$. To see that this makes sense, note that each x_j (for $j > 0$) is in either A or C. In the former case, $x_{j+1} = f(x_j)$ can be computed. In the latter case, the sequence stops. It is easy to check by induction that the elements x_j are all distinct, since f is an injective map. Since A is finite, the sequence must terminate with an element of C after finitely many steps. See Figure 1. Finally, the inverse map $g^{-1} : C \to B$ acts on elements of C by repeatedly applying f^{-1} until an element of B is reached. □

This bijection is motivated by the same idea underlying the Garsia-Milne Involution Principle [7]: if the given bijection f doesn't map into the set you want, just keep iterating f until you get there.

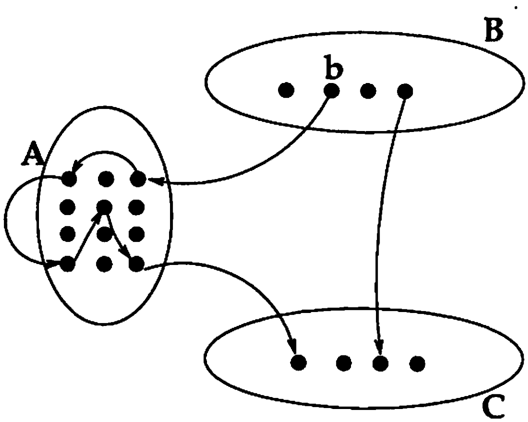

Figure 1: Bijective subtraction.

5 Five Easy Bijections

This section reviews five well-known bijections on lattice paths from the literature.

1. **Pascal's begin/end maps.** For $c \geq 1$ and $d \geq 1$, we have bijections

$$R(c-1,d) + q^c R(c,d-1) \overset{\text{begin}(c,d)}{=} R(c,d);$$

$$q^d R(c-1,d) + R(c,d-1) \overset{\text{end}(c,d)}{=} R(c,d).$$

Also, $R(0,d) = 1 = R(c,0)$ for all c and d. The map begin(c,d) adds a horizontal step to the beginning of a path in $R(c-1,d)$ and adds a vertical step to the beginning of a path in $q^c R(c,d-1)$. The inverse map simply deletes the first step of a path in $R(c,d)$. Similarly, the map end(c,d) adds a horizontal step to the end of a path in $q^d R(c-1,d)$ and adds a vertical step to the end of a path in $R(c,d-1)$. The inverse map simply deletes the last step of a path in $R(c,d)$. See Figure 2. It is clear from the figure that the maps preserve weights.

2. **Chu-Vandermonde's splitting map.** For $c, d, e \geq 0$, we have bijections

$$R(c,d+1+e) \overset{\text{split}(c,d,e)}{=} \sum_{r=0}^{c} R(r,d)R(c-r,e)q^{r(e+1)}.$$

Given a path $P \in R(c,d+1+e)$, the map split(c,d,e) produces a triple (r, P_1, P_2) where $0 \leq r \leq c$, $P_1 \in R(r,d)$, and $P_2 \in R(c-r,e)q^{r(e+1)}$.

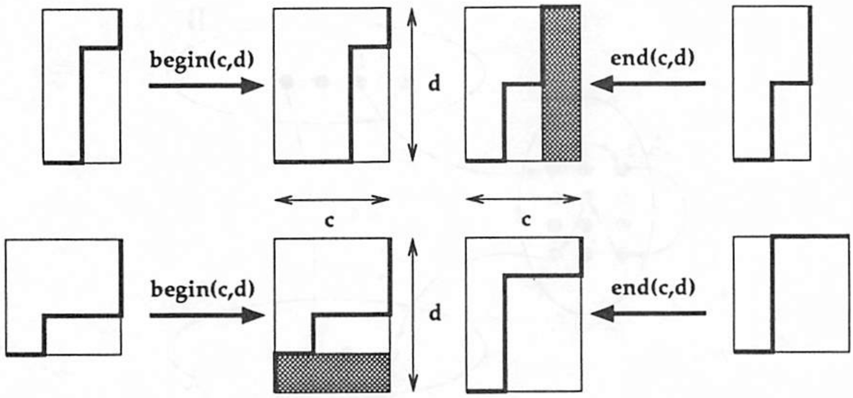

Figure 2: The begin and end maps.

Here, P_2 consists of the portion of P before the $(e+1)$'th north step of P, P_1 consists of the portion of P after the $(e+1)$'th north step, and r is the total number of east steps after this north step. The inverse of split(c,d,e) forms P by concatentating P_2, a north step, and P_1. See Figure 3. It is clear from the figure that $a(P) = a(P_1)+a(P_2)+r(e+1)$, so that the splitting map preserves weights.

3. **Haglund's bounce dissection map [10].** We have $F(n,n) = 1$ for $n \geq 0$ and $F(n,0) = 0$ for $n > 0$. Moreover, for $0 < k < n$ we have maps

$$F(n,k) \overset{\text{dis}(n,k)}{=} \sum_{r=0}^{n-k} R(k-1,r)q^{(n-k)(r-1)}F(n-k,r).$$

The map dis(n,k) "removes the first bounce" from a given path in $F(n,k)$. More explicitly, if $P \in F(n,k)$, we can write the word of P in the form $w = 0^k 1x1y$, where the two displayed ones encode the first two blocking east steps in P. Note that x has $k-1$ ones and r zeroes for some r between 0 and $n-k$. Here, r is the length of the second vertical move made by the bouncing ball. We map P to the triple (r, P_1, P_2), where $P_1 \in R(k-1,r)q^{(n-k)(r-1)}$ is the path encoded by the word x, and $P_2 \in F(n-k,r)$ is the Dyck path encoded by the word $0^r 1y$. This process is best understood pictorially; see Figure 4.

It is easy to see from the figure that

$$\text{area}(P) = k(k-1)/2 + a(P_1) + \text{area}(P_2),$$

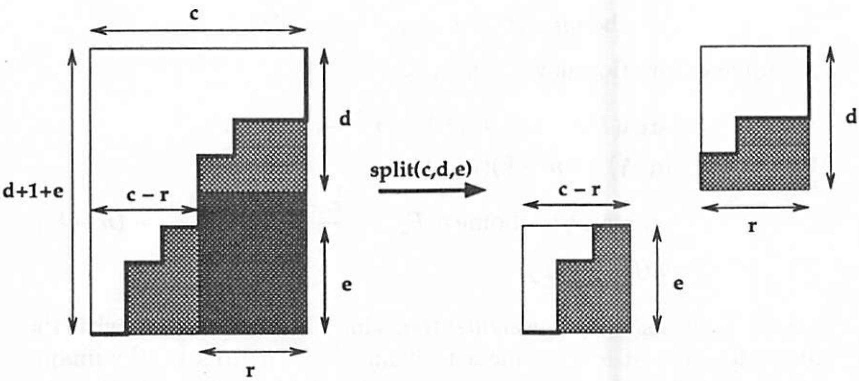

Figure 3: The splitting map.

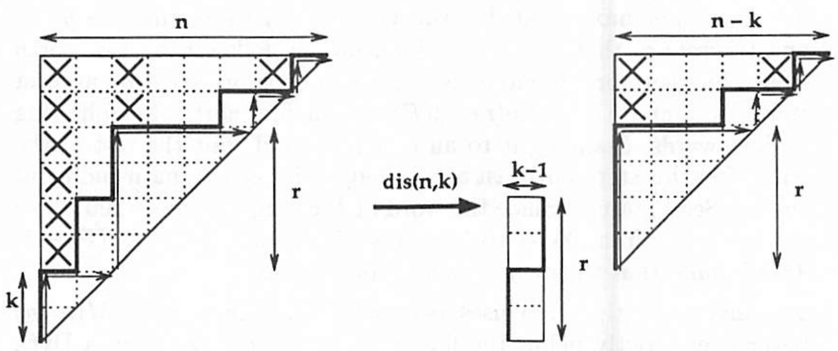

Figure 4: The bounce dissection map.

$$\text{bounce}(P) = n - k + \text{bounce}(P_2).$$

A simple calculation now confirms that

$$\begin{aligned}\text{wt}(P) &= \text{area}(P) - \text{bounce}(P) + (n^2+n)/2 - nk \\ &= a(P_1) + (n-k)(r-1) \\ &\quad +\text{area}(P_2) - \text{bounce}(P_2) + \frac{(n-k)^2+n-k}{2} - (n-k)r \\ &= \text{wt}((r, P_1, P_2)).\end{aligned}$$

So the map dis(n, k) is weight-preserving. For consistency with the above formula, it is convenient to define dis(n, n) to send the unique element of $F(n, n)$ to the unique element of $R(n-1, 0)F(0, 0)$.

4. **Fürlinger and Hofbauer's tipping map [4].** For all $m \le n$, we have maps

$$qM(m-1, n+1) + H(m, n) \overset{\text{tip}(m,n)}{=} M(m, n).$$

If $P \in H(m, n)$, then tip(m, n) deletes the last $n-m$ steps of P (which are horizontal steps) to produce an element of $M(m, n)$ with the same major index. If $P \in qM(m-1, n+1)$, scan the lattice points (x, y) on P (starting from the origin) and choose the last such point Q for which $x - y$ is maximized. For this Q, it is easy to see that $x - y \ge 0$ and (therefore) that Q must be immediately followed by two north steps. Furthermore, Q either is the origin or is preceded by an east step. To compute $P' = \text{tip}(m, n)(P)$, tip the first north step following Q downwards, changing it to an east step, and shift the rest of the path after this step southeast accordingly. This gives a path ending at (m, n). See Figure 5. Since the word of P changes from $\cdots 100 \cdots$ to $\cdots 110 \cdots$ (or from $00\cdots$ to $10\cdots$), we have maj(P') = maj$(P) + 1$. This implies that tip(m, n) is weight-preserving.

The inverse of tip(m, n) uses two cases. If a path $P \in M(m, n)$ never goes strictly below the line $y = x$, then enlarge P to a Dyck path in $H(m, n)$ by adding $n - m$ horizontal steps at the end. If P does go below $y = x$, find the first point $R = (x, y)$ on P (starting from the origin) such that $x - y$ is maximized. Change the east step that precedes R to a north step, shifting the rest of the path after R northwest accordingly. It is easy to check that this reverses the procedure above.

5. **Foata's maj-to-inv map [2, 3].** For all $c, d \ge 0$, we have maps

$$M(c, d) \overset{\text{Foata}(c,d)}{=} R'(c, d).$$

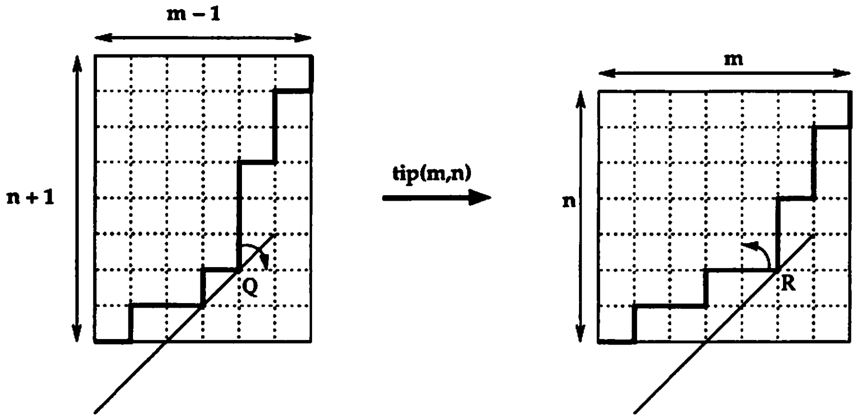

Figure 5: The tipping map.

These maps are defined by induction on $c + d$. They have the additional property of preserving the direction of the last step in the path. If $c + d \leq 1$, then Foata(c, d) is the identity map. Now consider the case $c + d > 1$. Let $P \in M(c, d)$ be encoded by the word w. We will specify the image of P under Foata(c, d) by giving its encoding word w'. There are three cases.

Case 1: $w = v1$. Let $v' = \text{Foata}(c-1, d)(v)$, and define $w' = v'1$. We have $\text{maj}(w) = \text{maj}(v) = \text{inv}(v') = \text{inv}(w')$.

Case 2: $w = v00$. Applying Foata$(c, d-1)$ to the word $v0$ gives some word v' that also ends in zero. Write v' uniquely as

$$v' = 1^{a_1}01^{a_2}0 \cdots 1^{a_k}0 \quad (a_i \geq 0),$$

and set

$$w' = 01^{a_1}01^{a_2} \cdots 01^{a_k}0.$$

In other words, we simultaneously move each zero in v' to the left of the block of ones immediately preceding it (if any), and then append a zero. We have

$$\begin{aligned} \text{maj}(w) &= \text{maj}(v0) = \text{inv}(v') \\ &= \text{inv}(w') - a_1 - \cdots - a_k + a_1 + \cdots + a_k \\ &= \text{inv}(w'). \end{aligned}$$

Case 3: $w = v10$. Applying Foata$(c, d-1)$ to the word $v1$ gives some word v' that also ends in one. Write v' uniquely as

$$v' = 0^{a_1}10^{a_2}1\cdots 0^{a_k}1 \quad (a_i \geq 0),$$

and set

$$w' = 10^{a_1}10^{a_2}\cdots 10^{a_k}0.$$

In other words, we simultaneously move each one in v' to the left of the block of zeroes immediately preceding it (if any), and then append a zero. We have

$$\begin{aligned} \mathrm{maj}(w) &= \mathrm{maj}(v1) + c + d - 1 = \mathrm{inv}(v') + (d-1) + c \\ &= \mathrm{inv}(v') + a_1 + \cdots + a_k + c = \mathrm{inv}(w'). \end{aligned}$$

Foata's original bijection is even more general; see [2].

Since we have $M(c,d) \overset{\mathrm{Foata}(c,d)}{\underline{=}} R'(c,d) \overset{\mathrm{flip}(c,d)}{\underline{=}} R(c,d)$ for all c and d, it is easy to construct bijections

$$qR(m-1, n+1) + H(m,n) \overset{\mathrm{tip}'(m,n)}{\underline{=}} R(m,n)$$

from the maps $\mathrm{tip}(m,n)$ and combinatorial addition.

6 Assembling the Bijections

We are now ready to construct bijections $F(n,k) \overset{h(n,k)}{\underline{=}} G(n,k)$ for all n and k. The algebraic version of this proof consists of showing that $\gamma(F(n,k))$ and $\gamma(G(n,k))$ are both given by the same formula, and hence are equal to each other. Accordingly, we divide our construction into three stages.

6.1 Stage 1: Maps Involving $F(n,k)$

The bijections in this subsection will show that:

$$\gamma(F(n,k)) = \begin{bmatrix} 2n-k-1 \\ n-k, n-1 \end{bmatrix}_q - q^k \begin{bmatrix} 2n-k-1 \\ n-k-1, n \end{bmatrix}_q;$$

$$\gamma(F(n,k)) = q^{k-n} \begin{bmatrix} 2n-k-1 \\ n-k, n-1 \end{bmatrix}_q - q^{k-n} \begin{bmatrix} 2n-k-1 \\ n-k-1, n \end{bmatrix}_q.$$

We will simultaneously construct, by induction on $n \geq 1$, two bijections

$$q^k R(n-k-1, n) + F(n,k) \overset{a(n,k)}{\underline{=}} R(n-k, n-1) \text{ and} \tag{4}$$

$$q^{k-n}R(n-k-1,n)+F(n,k) \overset{b(n,k)}{=} q^{k-n}R(n-k,n-1). \tag{5}$$

Base cases occur when $k=0$ or $k=n$ (which includes the case $n=1$). The case $k=n$ is easy to handle, since the sets on each side of (4) and (5) have exactly one element. When $k=0$, $F(n,k)$ is empty, and we can take $a(n,0)=b(n,0)=\mathrm{ref}\circ\mathrm{flip}(n-1,n)$.

For the induction step, assume that bijections $a(n',k')$ and $b(n',k')$ have already been constructed for all $n'<n$ and $0\le k'\le n'$. We must now construct the maps $a(n,k)$ and $b(n,k)$, where $0<k<n$. We have a chain of bijections

$$\begin{aligned}
&F(n,k)+q^{k-n}R(n-k-1,n)\\
&\overset{\mathrm{dis}(n,k)+\mathrm{split}(n-k-1,k-1,n-k)}{=}\sum_{r=0}^{n-k}R(k-1,r)q^{(n-k)(r-1)}F(n-k,r)\\
&+\sum_{r=0}^{n-k}R(r,k-1)q^{(n-k)(r-1)+r}R(n-k-r-1,n-k)\\
&=\sum_{r=0}^{n-k}R(r,k-1)q^{(n-k)(r-1)}\left(F(n-k,r)+q^{r}R(n-k-r-1,n-k)\right)\\
&\overset{\Sigma_r(\mathrm{id},a(n-k,r))}{=}\sum_{r=0}^{n-k}R(r,k-1)q^{(n-k)(r-1)}R(n-k-r,n-k-1)\\
&\overset{\mathrm{split}(n-k,k-1,n-k-1)^{-1}}{=} q^{k-n}R(n-k,n-1).
\end{aligned}$$

(Here id denotes an identity map, and the unlabelled bijection in this chain uses $R(k-1,r)\overset{\mathrm{ref}\circ\mathrm{flip}}{=}R(r,k-1)$ and the distributive law.) Composing all the bijections gives us the map $b(n,k)$.

Introduce temporary notation:

$$\begin{aligned}
A&=q^{k-n}R(n-k,n); & B&=q^{k-n}R(n-k-1,n);\\
C&=q^{k-n}R(n-k,n-1); & D&=q^{k}R(n-k-1,n);\\
E&=R(n-k,n-1); & F&=F(n,k).
\end{aligned}$$

We have just shown that $F+B\overset{b(n,k)}{=}C$. Also, the Pascal begin-end maps show that $C+D\overset{\mathrm{end}(n-k,n)}{=}A\overset{\mathrm{begin}(n-k,n)^{-1}}{=}B+E$. By combinatorial addition, we get a bijection $(B+C)+D+F=(B+C)+E$. The subtraction principle now gives a map $D+F=E$, i.e., a map

$$q^{k}R(n-k-1,n)+F(n,k)\overset{a(n,k)}{=}R(n-k,n-1).$$

This completes the inductive definition of $a(n,k)$ and $b(n,k)$.

Remark: If we apply γ everywhere, we get an algebraic proof of the formulas at the beginning of this subsection for $\gamma(F(n,k))$. This proof is shorter and simpler than Haglund's original proof.

6.2 Stage 2: Maps Involving $G(n,k)$

The maps in this subsection will show that

$$\gamma(G(n,k)) = q^{k-n}\begin{bmatrix}2n-k-1\\ n-k,n-1\end{bmatrix}_q - q^{k-n}\begin{bmatrix}2n-k-1\\ n-k-1,n\end{bmatrix}_q.$$

We will construct bijections

$$q^{n-k}R(n-k-1,n) + q^{2n-2k}G(n,k) \overset{d(n,k)}{=} q^{n-k}R(n-k,n-1). \qquad (6)$$

The bijection is trivial if $k=0$ or $k=n$, so assume $k<n$. Introduce the following temporary notation:

$$\begin{array}{ll} A = H(n-k,n); & B = H(n-k-1,n); \\ C = R(n-k,n); & D = qR(n-k-1,n+1); \\ E = R(n-k-1,n); & F = qR(n-k-2,n+1); \\ G = q^{2n-2k}G(n,k); & U = q^{n-k}R(n-k,n-1); \\ V = q^{n-k}R(n-k-1,n). & \end{array}$$

Then we have bijections

$$\begin{array}{lcll} B+G & = & A & \text{(via the identity map)} \\ A+D & = & C & (\text{use } \mathrm{tip}'(n-k,n)) \\ E & = & B+F & (\text{use } \mathrm{tip}'(n-k-1,n)^{-1}) \\ C & = & E+U & (\text{use } \mathrm{begin}(n-k,n)^{-1}) \\ F+V & = & D & (\text{use } \mathrm{begin}(n-k-1,n+1)). \end{array}$$

By combinatorial addition (and commutativity), we obtain a bijection

$$(A+B+C+D+E+F)+G+V = (A+B+C+D+E+F)+U.$$

The subtraction principle applies to give the map $G+V=U$, i.e.,

$$q^{n-k}R(n-k-1,n) + q^{2n-2k}G(n,k) \overset{d(n,k)}{=} q^{n-k}R(n-k,n-1).$$

Multiplying by q^{2k-2n}, we have bijections

$$q^{k-n}R(n-k-1,n) + G(n,k) \overset{d(n,k)}{=} q^{k-n}R(n-k,n-1).$$

6.3 Stage 3: Maps Showing that $F(n,k) = G(n,k)$

We have now constructed maps

$$b(n,k) : q^{k-n}R(n-k-1,n) + F(n,k) \to q^{k-n}R(n-k,n-1),$$

$$d(n,k) : q^{k-n}R(n-k-1,n) + G(n,k) \to q^{k-n}R(n-k,n-1),$$

Thus we have

$$q^{k-n}R(n-k-1,n) + F(n,k) \overset{d(n,k)^{-1}\circ b(n,k)}{=} q^{k-n}R(n-k-1,n) + G(n,k).$$

The combinatorial subtraction principle applies at once to give bijections $F(n,k) \overset{h(n,k)}{=} G(n,k)$ for all n and k.

We remark that there is another derivation of the weaker result $S(n) = T(n)$ that avoids the complicated maps from stage 2. Just note that there are obvious maps $S(n) = F(n+1,1)$ and $T(n) = H(n,n)$, and we have constructed maps

$$qR(n-1,n+1) + F(n+1,1) \overset{a(n+1,1)}{=} R(n,n),$$

$$qR(n-1,n+1) + H(n,n) \overset{\mathrm{tip}'(n,n)}{=} R(n,n).$$

Composing these maps gives $qR(n-1,n+1)+S(n) = qR(n-1,n+1)+T(n)$. Applying the subtraction principle one last time, we get a map $S(n) \overset{j(n)}{=} T(n)$.

All the bijections we have produced are quite complicated to compute by hand, although they are built up in a natural way from very simple component maps. Generally, the more times the subtraction principle is used, the more computations are required to execute the final bijection on a given input. Of course, these bijections are straightforward to program on a computer.

Since the construction of $j(n)$ avoided the use of the subtraction principle in stage 2, it is a bit simpler to compute than the maps $h(n,k)$. However, $j(n)$ need not have the additional property of sending the number of initial north steps to the number of final east steps. For example, $j(8)$ sends the path D with word 0010010011101101 to the path with word 0011010100100111, while $h(8,2)$ sends D to the path with word 0001010011011011.

A notorious unsolved problem is to give a bijection on Dyck paths of order n interchanging area and bounce. It is likely that any such bijection will have complexity comparable to that of the maps constructed above.

References

[1] D. Cohen, "PIE-sums: A combinatorial tool for partition theory," *J. Comb. Theory A* **31** (1981), 223—236.

[2] D. Foata, "On the Netto inversion number of a sequence," *Proc. Amer. Math. Soc.* **19** (1968), 236—240.

[3] D. Foata and M. Schützenberger, "Major index and inversion number of permutations," *Math. Nachr.* **83** (1978), 143—159.

[4] J. Fürlinger and J. Hofbauer, "q-Catalan numbers," *J. Comb. Theory A* **40** (1985), 248—264.

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

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

[7] A. Garsia and S. Milne, "Method for constructing bijections for classical partition identities," *Proc. Nat. Acad. Sci. USA* **78** (1981), 2026—2028.

[8] A. Garsia and S. Milne, "A Rogers-Ramanujan bijection," *J. Comb. Theory A* **31** (1981), 289—339.

[9] B. Gordon, "Sieve-equivalence and explicit bijections," *J. Comb. Theory A* **34** (1983), 90—93.

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

[11] Mark Haiman, personal communication. August 11, 2002.

[12] J. Joichi and D. Stanton, "Bijective Proofs of Basic Hypergeometric Series Identities," *Pacific J. of Mathematics* **127** (1987), 103—120.

[13] J. Remmel, "Bijective proofs of some classical partition identities," *J. Comb. Theory A* **33** (1982), 273—286.

[14] H. Wilf, "Sieve-equivalence in generalized partition theory," *J. Comb. Theory A* **34** (1983), 80—89.

[15] D. Zeilberger, "Garsia and Milne's bijective proof of the inclusion-exclusion principle," *Discrete Math.* **51** (1984), 109—110.