A COMBINATORIAL INTERPRETATION OF THE NUMBERS $6(2n)!/n!\,(n+2)!$

IRA M. GESSEL AND GUOCE XIN

Abstract. E. Catalan stated in 1874 that the numbers $(2m)!\,(2n)!/m!\,n!\,(m+n)!$ are integers. When $m = 0$ these numbers are the middle binomial coefficients $\binom{2n}{n}$. When $m = 1$ they are twice the Catalan numbers. In this paper, we give a combinatorial interpretation for these numbers when $m = 2$.

1. Introduction

The Catalan numbers

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!\,(n+1)!}$$

are well-known integers that arise in many combinatorial problems. In [7, pp. 219–229], 66 combinatorial interpretations of these numbers are given. Catalan numbers are generalized in several ways. The numbers

$$\frac{(2m)!\,(2n)!}{m!\,n!\,(m+n)!} \tag{1.1}$$

were first stated to be integers by E. Catalan in 1874 [1], and their number-theoretic properties were studied by several authors (see Dickson [2, pp. 265–266]). For $m = 0$, (1.1) is the middle binomial coefficient $\binom{2n}{n}$, and for $m = 1$ it is $2C_n$.

Except for $m = n = 0$, these integers are even, and it is convenient for our purposes to divide them by 2, so we consider the numbers

$$T(m,n) = \frac{1}{2}\,\frac{(2m)!\,(2n)!}{m!\,n!\,(m+n)!}. \tag{1.2}$$

The following identity (1.3) [3, Eq. 32], together with the symmetry property $T(m,n) = T(n,m)$, shows recursively that $T(m,n)$ is a positive integer for all m and n.

$$\sum_n 2^{p-2n}\binom{p}{2n} T(m,n) = T(m,m+p), \quad p \geq 0. \tag{1.3}$$

In principle, (1.3) gives a recursive combinatorial interpretation to $T(m,n)$. It is not too difficult to find combinatorial interpretations for (1.3) in the case $m = 1$ (Catalan numbers). See, e.g., Shapiro [6]. But it seems hard even for the case $m = 2$.

In this paper, we give a combinatorial interpretation for $T(2,n) = 6(2n)!/n!\,(n+2)!$, which starts with

$$3, 2, 3, 6, 14, 36, 99, 286, 858, 2652, \cdots.$$

Date: January 22, 2004.

Interpretations of these numbers by trees have been found by Gilles Schaeffer [5], and Nicholas Pippenger and Kristin Schleich [4, pp. 34].

The starting point of our interpretation is the formula

$$T(2,n) = 4C_n - C_{n+1}, \tag{1.4}$$

which can be easily checked. The idea is to construct an injection from one set of cardinality C_{n+1} to another set of cardinality $4C_n$, and then look at the difference.

2. The Main theorem

All paths in this paper have steps $(1,1)$ and $(1,-1)$. A step from a point u to a point v is denoted by $u \to v$. The *level* of a point in a path is defined to be its y-coordinate. A *Catalan path* of *length* n (or a *Dyck path* of length $2n$) is a path that starts at $(0,0)$, ends at $(2n,0)$, and never goes below level 0. Note that the length of a Catalan path is half the number of its steps. It is well-known that the number of Catalan paths of length n equals the Catalan number C_n. The *height* of a path P, denoted by $h(P)$, is the highest level it reaches.

Let $c(x)$ be the generating function for the Catalan numbers, so that

$$c(x) = \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n} x^n = \frac{1-\sqrt{1-4x}}{2x}.$$

Then $c(x)$ satisfies the functional equation $c(x) = 1 + xc(x)^2$. From equation (1.4), the generating function for $T(2,n)$ can be represented as

$$\sum_{n=0}^{\infty} T(2,n)x^n = \sum_{n=0}^{\infty} (4C_n - C_{n+1})x^n = 4c(x) - \frac{c(x)-1}{x} = 4c(x) - c(x)^2.$$

Now if we let $\mathbf{B}_n$ be the set of pairs of Catalan paths (P,Q) of total length n, then $|\mathbf{B}_n|$ has generating function $c(x)^2$, and we have $T(2,n) = 4C_n - |\mathbf{B}_n|$.

Theorem 2.1. *For $n \geq 1$, the number $T(2,n)$ counts pairs of Catalan paths (P,Q) of total length n with $|h(P) - h(Q)| \leq 1$.*

The proof of this theorem relies on the following Lemma 2.2. We will give two proofs of this lemma, one combinatorial and the other algebraic. The algebraic proof will be given in the next section.

Lemma 2.2. *For $n \geq 1$, C_n equals the number of pairs of Catalan paths (P,Q) of total length n, with P nonempty and $h(P) \leq h(Q) + 1$.*

Proof. Let $\mathbf{D}_n$ be the set of Catalan paths of length n, and let $\mathbf{E}_n$ be the set of pairs of Catalan paths (P,Q) of total length n, with P nonempty and $h(P) \leq h(Q) + 1$.

For a given pair (P,Q) in $\mathbf{E}_n$, since P is nonempty, the last step of P must be a down step, say, $u \to v$. By replacing $u \to v$ in P with an up step $u \to v'$, we get a path F_1. Now raising Q by two levels, we get a path F_2. Thus $F := F_1F_2$ is a path that ends at level 2 and never goes below level 0. The point v' belongs to both F_1 and F_2, but we treat it as a point only in F_2, even if F_2 is the empty path. The condition that $h(P) \leq h(Q) + 1$ yields $h(F_1) < h(F_2)$, which implies that the highest point of F must belong to F_2. See Figure 1 below.

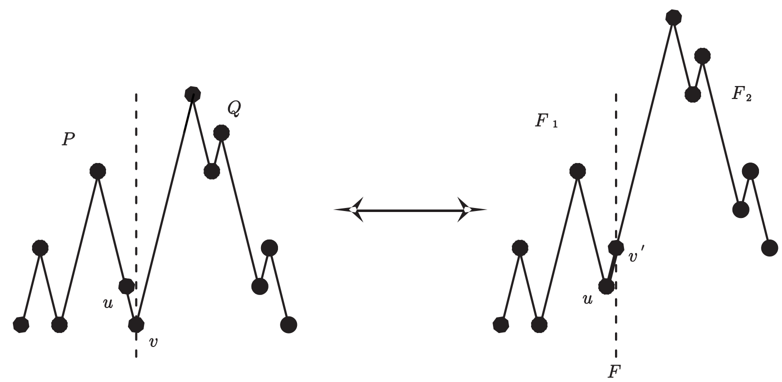

Figure 1. First step of the bijection

Now let y be the leftmost highest point of F (which is in F_2), and let $x \to y$ be the step in F. Then $x \to y$ is an up step. By replacing $x \to y$ with a down step $x \to y'$, and lowering the part of F_2 after y by two levels, we get a Catalan path $D \in \mathbf{D}_n$. See Figure 2 below.

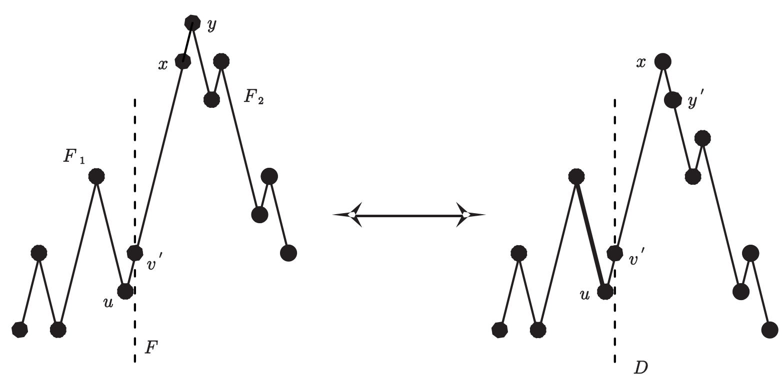

Figure 2. Second step of the bijection

With the following two key observations, it is easy to see that the above procedure gives a bijection from $\mathbf{E}_n$ to $\mathbf{D}_n$. First, x in the final Catalan path D is the rightmost highest point. Second, u in the intermediate path F is the rightmost point of level 1 in both F and F_1.

□

Proof of Theorem 2.1. Let F be the set of pairs of Catalan paths (P, Q) with $h(P) \le h(Q) + 1$, and let G be the set of pairs of Catalan paths (P, Q) with $h(Q) \le h(P) + 1$. By symmetry, we see that $|F| = |G|$. Now we claim that the cardinality of F is $2C_n$. This claim follows from Lemma 2.2 and the fact that if P is the empty path, then $h(P) \le h(Q) + 1$ for every $Q \in \mathbf{D}_n$.

Clearly we have that $F \cup G = \mathbf{B}_n$, and that $F \cap G$ is the set of pairs of Catalan paths (P, Q), with $|h(P) - h(Q)| \le 1$. The theorem then follows from the following easy computation:

$$|F \cap G| = |F| + |G| - |F \cup G| = 4C_n - |\mathbf{B}_n| = 4C_n - C_{n+1}.$$ □

The formula (1.3) in the case $m = 2$ is given as follows.

$$\sum_n 2^{p-2n}\binom{p}{2n} T(2,n) = T(2,2+p), \quad p \geq 0. \tag{2.1}$$

We have not found an explanation of this formula using our interpretation.

3. An Algebraic Proof and Further Results

In this section we give an algebraic proof of Lemma 2.2. Let $C = xc(x)^2 = c(x) - 1$ and let G_k be the generating function for Catalan paths of height at most k.

Lemma 3.1. *For all $k \geq -1$,*

$$G_k = (1+C)\frac{1-C^{k+1}}{1-C^{k+2}}. \tag{3.1}$$

Proof. Let P be a path of height at most $k \geq 1$. If P is nonempty then P can be factored as uP_1dP_2, where u is an up step, P_1 is a Catalan path of height at most $k-1$ (shifted up one unit), d is a down step, and P_2 is a Catalan path of height at most k. Thus $G_k = 1 + xG_{k-1}G_k$, so $G_k = 1/(1 - xG_{k-1})$. Equation (3.1) clearly holds for $k = -1$ and $k = 0$. Now suppose that for some $k \geq 1$,

$$G_{k-1} = (1+C)\frac{1-C^k}{1-C^{k+1}}.$$

Then the recurrence, together with the formula $x = C/(1+C)^2$, gives

$$\begin{aligned} G_k &= \left[1 - x(1+C)\frac{1-C^k}{1-C^{k+1}}\right]^{-1} = \left[1 - \frac{C}{1+C}\,\frac{1-C^k}{1-C^{k+1}}\right]^{-1} \\ &= \left[\frac{1-C^{k+2}}{(1+C)(1-C^{k+1})}\right]^{-1} = (1+C)\frac{1-C^{k+1}}{1-C^{k+2}}. \end{aligned}$$

□

We can prove Lemma 2.2 by showing that $\sum_{n=0}^{\infty} G_{n+1}(G_n - G_{n-1}) = 1+2C$; this is equivalent to the statement that the number of pairs (P, Q) of Catalan paths of length $n > 0$ with $h(P) \leq h(Q) + 1$ is $2C_n$.

Theorem 3.2.

$$\sum_{n=0}^{\infty}(G_n - G_{n-1})G_{n+1} = 1 + 2C.$$

Proof. Let

$$\Psi_k = \sum_{n=k}^{\infty} \frac{C^n}{1-C^n}.$$

Thus if $j < k$ then

$$\Psi_j = \sum_{n=j}^{k-1} \frac{C^n}{1-C^n} + \Psi_k. \tag{3.2}$$

We have

$$G_n G_{n+1} = (1+C)^2 \frac{1-C^{n+1}}{1-C^{n+3}}$$

and

$$G_{n-1}G_{n+1} = (1+C)^2\frac{(1-C^n)(1-C^{n+2})}{(1-C^{n+1})(1-C^{n+3})}.$$

Let

$$S_1 = \sum_{n=0}^{\infty}\left(\frac{1-C^{n+1}}{1-C^{n+3}} - 1\right)$$

and

$$S_2 = \sum_{n=0}^{\infty}\left(\frac{(1-C^n)(1-C^{n+2})}{(1-C^{n+1})(1-C^{n+3})} - 1\right).$$

Then $\sum_{n=0}^{\infty}(G_n - G_{n-1})G_{n+1} = (1+C)^2(S_1 - S_2)$. We have

$$\frac{1-C^{n+1}}{1-C^{n+3}} - 1 = -\frac{(1-C^2)C^{n+1}}{1-C^{n+3}},$$

so $S_1 = -(1-C^2)C^{-2}\Psi_3$, and

$$\frac{(1-C^n)(1-C^{n+2})}{(1-C^{n+1})(1-C^{n+3})} - 1 = -\frac{(1-C)C^n}{(1+C)(1-C^{n+1})} - \frac{(1-C^3)C^{n+1}}{(1+C)(1-C^{n+3})},$$

so

$$S_2 = -\frac{1-C}{1+C}C^{-1}\Psi_1 - \frac{1-C^3}{1+C}C^{-2}\Psi_3.$$

Therefore

$$\begin{aligned} S_1 - S_2 &= \frac{1-C}{1+C}C^{-1}\Psi_1 + \left(\frac{1-C^3}{1+C} - (1-C^2)\right)C^{-2}\Psi_3 \\ &= \frac{1-C}{1+C}C^{-1}(\Psi_1 - \Psi_3) \\ &= \frac{1-C}{1+C}C^{-1}\left(\frac{C}{1-C} + \frac{C^2}{1-C^2}\right) = \frac{1+2C}{(1+C)^2}. \end{aligned}$$

Thus $(1+C)^2(S_1 - S_2) = 1 + 2C$. □

Although the fact that the series in Theorem 3.2 telescopes may seem surprising, it is in fact a special case of a theorem that applies to much more general sums of generating functions for Catalan paths with restricted height, given by Theorem 3.4

Lemma 3.3. *Let $R(z, C)$ be a rational function of z and C of the form*

$$\frac{zN(z,C)}{\prod_{i=1}^{m}(1 - zC^{a_i})},$$

where $N(z, C)$ is a polynomial in z of degree less than m, with coefficients that are rational functions of C, and the a_i are distinct positive integers. Let $L = -\lim_{z\to\infty} R(z, C)$. Then

$$\sum_{n=0}^{\infty} R(C^n, C) = Q(C) + L\Psi_1,$$

where $Q(C)$ is a rational function of C.

Proof. First we show that the lemma holds for $R(z, C) = z/(1 - zC^a)$. In this case, $L = -\lim_{z\to\infty} R(z, C) = C^{-a}$ and

$$\sum_{n=0}^{\infty} R(C^n, C) = \sum_{n=0}^{\infty} \frac{C^n}{1 - C^{n+a}} = C^{-a} \sum_{n=a}^{\infty} \frac{C^n}{1 - C^n} = -\sum_{n=0}^{a-1} \frac{C^{n-a}}{1 - C^n} + C^{-a}\Psi_1.$$

Now we consider the general case. Since $R(z, C)/z$ is a proper rational function of z, it has a partial fraction expansion

$$\frac{1}{z} R(z, C) = \sum_{i=1}^{m} \frac{U_i(C)}{1 - zC^{a_i}}$$

for some rational functions $U_i(C)$, so

$$R(z, C) = \sum_{i=1}^{m} U_i(C) \frac{z}{1 - zC^{a_i}}.$$

The general theorem then follows by applying the special case to each summand. □

Theorem 3.4. *If $i_1, i_1', \ldots, i_r$ are distinct integers, then*

$$\sum_{n=0}^{\infty} (G_{n+i_1} - G_{n+i_1'}) G_{n+i_2} G_{n+i_3} \cdots G_{n+i_r}$$

is a rational function of C.

Proof. Apply Lemma 3.3 to

$$\sum_{n=0}^{\infty} (G_{n+i_1} G_{n+i_2} G_{n+i_3} \cdots G_{n+i_r} - 1)$$

and

$$\sum_{n=0}^{\infty} \left(G_{n+i_1'} G_{n+i_2} G_{n+i_3} \cdots G_{n+i_r} - 1\right).$$

□

References

[1] E. Catalan, *Nouvelles Annales de Mathématiques, J. des Candidats aux École Polytechnic et Normale* (2), 13 (1874), 207.

[2] L. E. Dickson, *History of the Theory of Numbers*, New York: Chelsea. Vol. 1, 1966. Originally published in 1919 by the Carnegie Institute of Washington.

[3] I. Gessel, “Super ballot numbers”, *J. Symb. Comput.* 14 (1992), 179–194.

[4] N. Pippenger and K. Schleich, “Topological characteristics of random surfaces generated by cubic interactions” gr-qc/0306049.

[5] G. Schaeffer, “A combinatorial interpretation of super-Catalan numbers of order two”, http://www.loria.fr/~schaeffe/Pub/Conjugacy/superCat.ps, (2001).

[6] L. W. Shapiro, “A short proof of an identity of Touchard’s concerning Catalan numbers”, *J. Combin. Theory Ser. A* 20 (1976) 375–376.

[7] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, 1999.

Department of Mathematics, Brandeis University, Waltham MA 02454-9110

E-mail address: gessel@brandeis.edu

Department of Mathematics, Brandeis University, Waltham MA 02454-9110

E-mail address: maxima@brandeis.edu