

# MULTIVARIATE FUSS-CATALAN NUMBERS

J.-C. AVAL

ABSTRACT. Catalan numbers  $C(n) = \frac{1}{n+1} \binom{2n}{n}$  enumerate binary trees and Dyck paths. The distribution of paths with respect to their number  $k$  of factors is given by ballot numbers  $B(n, k) = \frac{n-k}{n+k} \binom{n+k}{n}$ . These integers are known to satisfy simple recurrence, which may be visualised in a “Catalan triangle”, a lower-triangular two-dimensional array. It is surprising that the extension of this construction to 3 dimensions generates integers  $B_3(n, k, l)$  that give a 2-parameter distribution of  $C_3(n) = \frac{1}{2n+1} \binom{3n}{n}$ , which may be called order-3 Fuss-Catalan numbers, and enumerate ternary trees. The aim of this paper is a study of these integers  $B_3(n, k, l)$ . We obtain an explicit formula and a description in terms of trees and paths. Finally, we extend our construction to  $p$ -dimensional arrays, and in this case we obtain a  $(p-1)$ -parameter distribution of  $C_p(n) = \frac{1}{(p-1)n+1} \binom{pn}{n}$ , the number of  $p$ -ary trees.

## 1. CATALAN TRIANGLE, BINARY TREES, AND DYCK PATHS

We recall in this section well-known results about Catalan numbers and ballot numbers.

The *Catalan numbers*

$$C(n) = \frac{1}{n+1} \binom{2n}{n}$$

are integers that appear in many combinatorial problems. These numbers first arose in the work of Catalan as the number of triangulations of a polygon by mean of non-intersecting diagonals. Stanley [13, 14] maintains a dynamic list of exercises related to Catalan numbers, including (at this date) 127 combinatorial interpretations.

Closely related to Catalan numbers are *ballot numbers*. Their name is due to the fact that they are the solution of the so-called *ballot problem*: we consider an election between two candidates A and B, which respectively receive  $a > b$  votes. The question is: what is the probability that during the counting of votes, A stays ahead of B? The answer will be given below, and we refer to [5] for Bertrand’s first solution, and to [1] for André’s beautiful solution using the “reflection principle”.

Since our goal here is different, we shall neither define ballot numbers by the previous statement, nor by their explicit formula, but we introduce integers  $B(n, k)$  for  $(n, k) \in \mathbb{N} \times \mathbb{N}$  defined by the following recurrence:

- $B(1, 0) = 1$
- $\forall n > 1$  and  $0 \leq k < n$ ,  $B(n, k) = \sum_{i=0}^k B(n-1, i)$
- $\forall k \geq n$ ,  $B(n, k) = 0$ .

---

*Date:* April 26, 2005.

J.C. Aval is supported by EC’s IHRP Programme, within the Research Training Network “Algebraic Combinatorics in Europe,” grant HPRN-CT-2001-00272.

Observe that the recursive formula in the second condition is equivalent to:

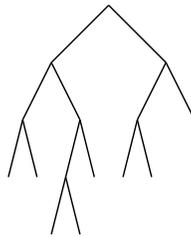
$$(1.1) \quad B(n, k) = B(n - 1, k) + B(n, k - 1).$$

We shall present the  $B(n, k)$ 's by the following triangular representation (zero entries are omitted) where moving down increases  $n$  and moving right increases  $k$ .

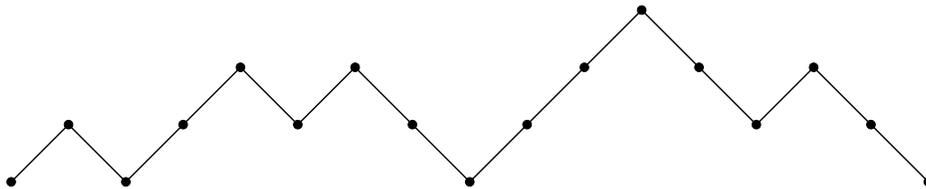
$$\begin{array}{cccccc} 1 & & & & & & \\ 1 & 1 & & & & & \\ 1 & 2 & 2 & & & & \\ 1 & 3 & 5 & 5 & & & \\ 1 & 4 & 9 & 14 & 14 & & \\ 1 & 5 & 14 & 28 & 42 & 42 & \end{array}$$

The crucial observation is that computing the horizontal sums of these integers give : 1, 2, 5, 14, 42, 132. We recognize the first terms of the Catalan series, and this fact will be proven in Proposition 1.1, after introducing combinatorial objects.

A *binary tree* is a tree in which every internal node has exactly 2 sons. The number of binary trees with  $n$  internal nodes is given by the  $n$ -th Catalan number.



A *Dyck path* is a path consisting of steps  $(1, 1)$  and  $(1, -1)$ , starting from  $(0, 0)$ , ending at  $(2n, 0)$ , and remaining above the line  $y = 0$ . The number of Dyck paths of length  $2n$  is also given by the  $n$ -th catalan number. More precisely, the depth-first search of the tree gives a bijection between binary trees and Dyck paths: we associate to each external node (except the left-most one) a  $(1, 1)$  step and to each internal node a  $(1, -1)$  step by searching recursively the left son, then the right son, then the root. As an example, we show below the Dyck path corresponding to the binary tree given above.



An important parameter in our study will be the length of the right-most sequence of  $(1, -1)$  of the path. This parameter equals 2 in our example. Observe that under the correspondence between paths and trees, this parameter corresponds to the length of the right-most string of right sons in the tree. We shall use the expressions *last down sequence* and *last right string*, for these parts of the path and of the tree.

Now we come to the announced result. It is well-known and simple, but is the starting point of our work.

**Proposition 1.1.** *We have the following equality:*

$$\sum_{k=0}^{n-1} B(n, k) = C(n) = \frac{1}{n+1} \binom{2n}{n}.$$

*Proof.* Let us denote by  $\mathcal{C}_{n,k}$  the set of Dyck paths of length  $2n$  with a last down sequence of length equal to  $n - k$ .

We shall prove that  $B(n, k)$  is the cardinality of  $\mathcal{C}_{n,k}$ .

The proof is done recursively on  $n$ . If  $n = 0$ , this is trivial. If  $n > 0$ , let us suppose that  $B(n - 1, k)$  is the cardinality of  $\mathcal{C}_{n-1,k}$  for  $0 \leq k < n - 1$ . Let us consider an element of  $\mathcal{C}_{n,k}$ . If we erase the last step  $(1, 1)$  and the following step  $(1, -1)$ , we obtain a Dyck path of length  $2(n - 1)$  and with a last decreasing sequence of length  $n - l \geq n - k$ . If we keep track of the integer  $k$ , we obtain a bijection between  $\mathcal{C}_{n,k}$  and  $\cup_{l \leq k} \mathcal{C}_{n-1,l}$ . We mention that this process is very similar to the ECO method [3].

This is a combinatorial proof of Proposition 1.1.  $\square$

**Remark 1.2.** In fact ballot numbers are given by the explicit formula:

$$(1.2) \quad B(a, b) = \frac{a - b}{a + b} \binom{a + b}{a}.$$

This expression can be obtained shortly by checking the recurrence (1.1). We can alternatively use the reflection principle (see [9] for a clear presentation), or the cyclic lemma (*cf.* [6]), which will be used in the next sections to obtain a formula in the general case.

The expression (1.2) constitutes a solution of the ballot problem. Indeed, the integer  $\binom{a+b}{a}$  can be seen as the number of different countings of votes, and we use the interpretation in terms of paths: we represent a vote for A by an up step, and a vote for B by a down step, and erase the last down sequence. Then  $B(a, b)$  is the number of those countings for which A stays ahead of B, thus  $\frac{a-b}{a+b}$  is the required probability.

Of course, we could have used expression (1.2) to prove Proposition 1.1 by a simple computation, but our proof explains more about the combinatorial objects and can be adapted to ternary trees in the next section.

**Remark 1.3.** Our Catalan array presents similarities with Riordan arrays, but it is not a Riordan array. It may be useful to explicit this point. We recall (*cf.* [11]) that a Riordan array  $M = (m_{i,j}) \in \mathbb{C}^{\mathbb{N} \times \mathbb{N}}$  is relative to 2 generating functions

$$g(x) = \sum g_n x^n \quad \text{and} \quad f(x) = \sum f_n x^n$$

and is such that

$$M_j(x) = \sum_{n \geq 0} m_{n,j} x^n = g(x) f(x)^j.$$

We easily observe that a Riordan array with the first two columns  $M_0(x)$  and  $M_1(x)$  of our Catalan array is relative to  $g(x) = \frac{1}{1-x}$  and  $f(x) = \frac{x}{1-x}$ , which gives the Pascal triangle.

In fact, the Riordan array relative to  $g(x) = \mathbf{C}(x) = \sum C(n)x^n$  and  $f(x) = x \mathbf{C}(x)$  gives the ballot numbers, but requires to know the Catalan numbers.

## 2. FUSS-CATALAN TETRAHEDRON AND TERNARY TREES

2.1. **Definitions.** This section, which is the heart of this work, is the study of a 3-dimensional analogue of the Catalan triangle of the previous section. That is we consider exactly the same recurrence, and let the array grow, not in 2, but in 3 dimensions. More precisely, we introduce the sequence  $B_3(n, k, l)$  indexed by integers  $n, k$  and  $l$ , and defined recursively by:

- $B_3(1, 0, 0) = 1$
- $\forall n > 1, k + l < n, B_3(n, k, l) = \sum_{0 \leq i \leq k, 0 \leq j \leq l} B_3(n-1, i, j)$
- $\forall k + l \geq n, B_3(n, k, l) = 0.$

Observe that the recursive formula in the second condition is equivalent to:

$$(2.1) \quad B_3(n, k, l) = B_3(n-1, k, l) + B_3(n, k-1, l) + B_3(n, k, l-1) - B_3(n, k-1, l-1)$$

and this expression can be used to make some computations lighter, but the presentation above explains more about the generalization of the definition of the ballot numbers  $B(n, k)$ .

Because of the planar structure of the sheet of paper, we are forced to present the tetrahedron of  $B_3(n, k, l)$ 's by its sections with a given  $n$ .

$$\begin{array}{l}
 n = 1 \longrightarrow [ 1 ] \\
 n = 2 \longrightarrow \left[ \begin{array}{cc} 1 & 1 \\ 1 & \end{array} \right] \\
 n = 3 \longrightarrow \left[ \begin{array}{ccc} 1 & 2 & 2 \\ 2 & 3 & \\ 2 & & \end{array} \right] \\
 n = 4 \longrightarrow \left[ \begin{array}{cccc} 1 & 3 & 5 & 5 \\ 3 & 8 & 10 & \\ 5 & 10 & & \\ 5 & & & \end{array} \right] \\
 n = 5 \longrightarrow \left[ \begin{array}{ccccc} 1 & 4 & 9 & 14 & 14 \\ 4 & 15 & 30 & 35 & \\ 9 & 30 & 45 & & \\ 14 & 35 & & & \\ 14 & & & & \end{array} \right]
 \end{array}$$

It is clear that  $B_3(n, k, 0) = B_3(n, 0, k) = B(n, k)$ . The reader may easily check that when we compute  $\sum_{k,l} B_3(n, k, l)$ , we obtain: 1, 3, 12, 55, 273. These integers are the first terms of the following sequence (*cf.* [12]):

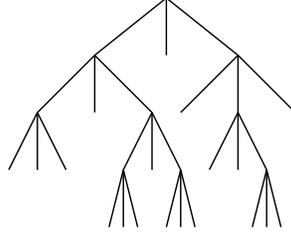
$$C_3(n) = \frac{1}{2n+1} \binom{3n}{n}.$$

This fact will be proven in Proposition 2.1.

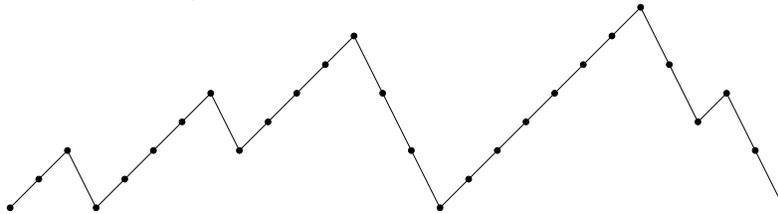
**2.2. Combinatorial interpretation.** Fuss<sup>1</sup>-Catalan numbers (*cf.* [7, 9]) are given by the formula

$$(2.2) \quad C_p(n) = \frac{1}{(p-1)n+1} \binom{pn}{n},$$

and  $C_3(n)$  appear as order-3 Fuss-Catalan numbers. The integers  $C_3(n)$  are known [12] to count *ternary trees*, *ie.* trees in which every internal node has exactly 3 sons.



Ternary trees are in bijection with *2-Dyck paths*, which are defined as paths from  $(0, 0)$  to  $(3n, 0)$  with steps  $(1, 1)$  and  $(1, -2)$ , and remaining above the line  $y = 0$ . The bijection between these objects is the same as in the case of binary trees, *ie.* a depth-first search, with the difference that here an internal node is translated into a  $(1, -2)$  step. To illustrate this bijection, we give the path corresponding to the previous example of ternary tree:



We shall consider these paths with respect to the position of their down steps. Let  $\mathcal{D}_{n,k,l}$  denote the set of 2-Dyck paths of length  $3n$ , with  $k$  down steps at even height and  $l$  down steps at odd height. By convention, the last sequence of down steps is not considered (the number of these steps is by definition equal to  $n - k - l$ ).

**Proposition 2.1.** *We have*

$$\sum_{k,l} B_3(n, k, l) = C_3(n) = \frac{1}{2n+1} \binom{3n}{n}.$$

Moreover,  $B_3(n, k, l)$  is the cardinality of  $\mathcal{D}_{n,k,l}$ .

*Proof.* Let  $k$  and  $l$  be fixed. Let us consider an element of  $\mathcal{D}_{n,k,l}$ . If we cut this path after its  $(2n - 2)$ -th up step, and complete with down steps, we obtain a 2-Dyck path of length  $3(n - 1)$  (see figure below). It is clear that this path is an element of  $\mathcal{D}_{n,i,j}$  for some  $i \leq k$  and  $j \leq l$ . We can furthermore reconstruct the original path from the truncated one, if we know  $k$  and  $l$ . We only have to delete the last sequence of down steps (here the dashed line), to draw  $k - i$  down steps, one up step,  $l - j$  down steps,

---

<sup>1</sup>Nikolai Fuss (Basel, 1755 – St Petersburg, 1826) helped Euler prepare over 250 articles for publication over a period on about seven years in which he acted as Euler’s assistant, and was from 1800 to 1826 permanent secretary to the St Petersburg Academy.



- at odd height:  $n$  up steps and  $l$  down steps.

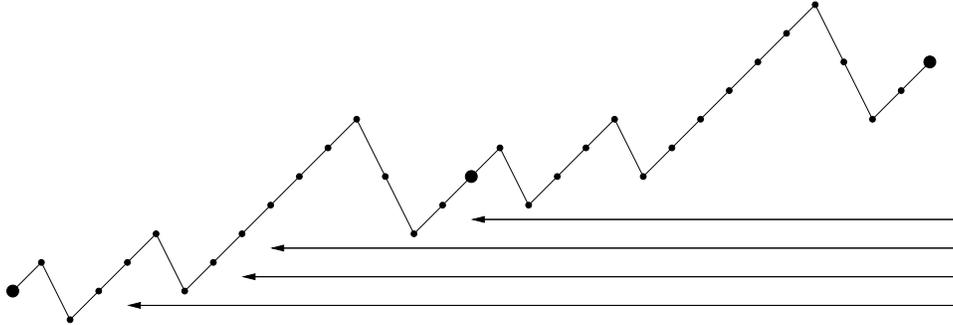
An important remark is to remember that an element of  $\mathcal{D}_{n,k,l}$  has an up step just before the last sequence down steps! If we suppose that all these steps are distinguished, we obtain:

- $\frac{(n+k)!}{n!k!}$  choices for even places;
- $\frac{(n-1+l)!}{(n-1)!l!}$  choices for odd places (we cannot put any odd down steps after the last odd step).

Now we group in this set  $\mathcal{P}$  the paths which are “even permutations” of a same path  $P$ . By even permutations, we mean cyclic permutations which preserves the parity of the steps.

We now want to prove that in any even orbit, the proportion of elements of  $\mathcal{D}_{n,k,l}$  is  $\frac{n-k-l}{n+k}$ .

We suppose first that the path  $P$  is acyclic (as a word). It is clear that such a path gives  $n+k$  different even permutations. Now we have to keep only those which give elements of  $\mathcal{D}_{n,k,l}$ . To do this we consider the concatenation of  $P$  and  $P'$ , which is a duplicate of  $P$ .



The cyclic permutations of  $P$  (not necessarily even) are the connex parts of  $P + P'$  of horizontal length  $2n + k + l$ . The number of such paths that remain above the horizontal axis, and ends with an up step, is the number of (up) steps of  $P$  in the light of an horizontal light source coming from the right. The only transformation is to put the enlightened up step at the end of the path. The number of enlightened up steps is  $2n - 2k - 2l$ , since every down steps put 2 up steps in the shadow. Among these  $2n - 2k - 2l$  permutations, only the half is an even one (observe that the heights of the enlightened steps are connex). Thus  $n - k - l$  paths among the  $n + k$  elements of the orbit are in  $\mathcal{D}_{n,k,l}$ .

Now we observe that if  $P$  is  $p$ -cyclic, then its orbit has  $p$  times less elements, and we obtain  $p$  times less different paths, whence the proportion of elements of elements of  $\mathcal{D}_{n,k,l}$  in this orbit is:

$$\frac{(n - k - l)/p}{(n + k)/p} = \frac{n - k - l}{n + k}.$$

Finally, we obtain that the cardinality of  $\mathcal{D}_{n,k,l}$  is

$$\frac{(n + k)!}{n!k!} \frac{(n - 1 + l)!}{(n - 1)!l!} \frac{n - k - l}{n + k}$$



**Proposition 2.3.** *The generating function  $F$  of the  $B_3$ 's is given by:*

$$F(t, x, y) = \frac{1}{1 - tG(t, x, y)G(t, y, x)}$$

with  $G(t, x, y)$  solution of the algebraic equation

$$tx^2G^3 + (y - x)G^2 + (x - 2y)G + y = 0.$$

An alternate approach to the generating function is to use formula 2.3 to obtain what MacMahon called a ‘‘redundant generating function’’ (cf. [8]), since it contains terms other than those which are combinatorially significant.

To do this we extend the recursive definition of  $B_3(n, k, l)$  as follows: we define  $B'_3(n, k, l)$  for nonnegative integers  $n, k, l$  by:

$$(2.9) \quad B'_3(n, k, l) = \binom{n+k-1}{k} \binom{n+l-1}{l} \frac{n-k-l}{n}.$$

Of course when  $k+l > n$ , the integer  $B'_3(n, k, l)$  is negative.

As an example, here is the ‘‘section’’ of the array of  $B'_3(n, k, l)$  with  $n = 3$ .

$$\begin{bmatrix} 1 & 2 & 2 & 0 & -5 & \dots \\ 2 & 3 & 0 & -10 & -30 & \dots \\ 2 & 0 & -12 & -40 & -90 & \dots \\ 0 & -10 & -40 & -100 & -200 & \dots \\ -5 & -30 & -90 & -200 & -375 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The equation (2.9) is equivalent to the following recursive definition:  $\forall n \geq 0$  or  $k, l < 0$ ,  $B'_3(n, k, l) = 0$  and

$$B'_3(n, k, l) = B'_3(n-1, k, l) + B'_3(n, k-1, l) + B'_3(n, k, l-1) - B'_3(n, k-1, l-1) + c(n, k, l)$$

$$\text{with } c(n, k, l) = \begin{cases} +1 & \text{if } (n, k, l) = (1, 0, 0) \\ -1 & \text{if } (n, k, l) = (0, 1, 0) \\ -1 & \text{if } (n, k, l) = (0, 0, 1) \\ +2 & \text{if } (n, k, l) = (0, 1, 1) \\ 0 & \text{otherwise.} \end{cases}$$

The interest of this presentation is to give a simple (rational!) generating series. Indeed, it is quite simple to deduce from the recursive definition of  $B'_3(n, k, l)$ 's that:

$$(2.10) \quad \sum_{n,k,l} B'_3(n, k, l) t^n x^k y^l = \frac{t - x - y + 2xy}{1 - t - x - y + xy}.$$

This formula encloses the  $B'_3(n, k, l)$  since  $\forall k+l < n$ ,  $B_3(n, k, l) = B'_3(n, k, l)$ .

### 3. FUSS-CATALAN $p$ -SIMPLEX AND $p$ -ARY TREES

The aim of this final section is to present an extension of the results of Section 2 to  $p$ -dimensional recursive sequences. In the same spirit, we define the sequence  $B_p(n, k_1, k_2, \dots, k_{p-1})$  by the recurrence:

- $B_p(1, 0, 0, \dots, 0) = 1$
- $\forall n > 1$  and  $0 \leq k_1 + k_2 + \dots + k_{p-1} < n$ ,

$$B_p(n, k_1, k_2, \dots, k_{p-1}) = \sum_{0 \leq i_1 \leq k_1, \dots, 0 \leq i_{p-1} \leq k_{p-1}} B(n-1, i_1, i_2, \dots, i_{p-1})$$

- $\forall k_1 + k_2 + \dots + k_{p-1} \geq n$ ,  $B_p(n, k_1, k_2, \dots, k_{p-1}) = 0$ .

Every result of Section 2 extends to general  $p$ . We shall only give the main results, since the proofs are straight generalizations of the proofs in the previous section.

**Proposition 3.1.**

$$\sum_{k_1, \dots, k_{p-1}} B_p(n, k_1, \dots, k_{p-1}) = C_p(n) = \frac{1}{(p-1)n+1} \binom{pn}{n}$$

The integer  $C_p(n)$  are order- $p$  Fuss-Catalan numbers and enumerate  $p$ -ary trees, or alternatively  $p$ -Dyck paths (the down steps are  $(1, -p)$ ). In this general case, the recursive definition of  $B_p(n, k_1, k_2, \dots, k_{p-1})$  gives rise  $p-1$  to statistics on trees and paths analogous to the one defined in section 3.

**Remark 3.2.** By the same method as in the previous section, it is possible to obtain an explicit formula for these multivariate Fuss-Catalan numbers:

$$B_p(n, k_1, k_2, \dots, k_{p-1}) = \left( \prod_{i=1}^{p-1} \binom{n+k_i-1}{k_i} \right) \frac{n - \sum_{i=1}^{p-1} k_i}{n}.$$

**Comment.** The numbers  $B_3(n, k, l)$  first arose in a question of algebraic combinatorics (*cf.* [2]). Let  $\mathcal{I}_n$  be the ideal generated by  $B$ -quasisymmetric polynomials in the  $2n$  variables  $x_1, \dots, x_n$  and  $y_1, \dots, y_n$  (*cf.* [4]) without constant term. We denote by  $\mathbf{Q}_n$  the quotient  $\mathbb{Q}[x_1, \dots, x_n, y_1, \dots, y_n]/\mathcal{I}_n$  and by  $\mathbf{Q}_n^{k,l}$  the bihomogeneous component of  $\mathbf{Q}_n$  of degree  $k$  in  $x_1, \dots, x_n$  and degree  $l$  in  $y_1, \dots, y_n$ . It is proven in [2] that:

$$\dim \mathbf{Q}_n^{k,l} = B_3(n, k, l) = \binom{n+k-1}{k} \binom{n+l-1}{l} \frac{n-k-l}{n}.$$

REFERENCES

- [1] D. ANDRÉ, *Solution directe du problème résolu par M. Bertrand*, Comptes Rendus Acad. Sci. Paris, **105** (1887), 436–437.
- [2] J.-C. AVAL, *Ideals and quotients of B-quasisymmetric functions*, preprint.
- [3] E. BARCUCCI, A. DEL LUNGO, E. PERGOLA, R. PINZANI, *ECO: a methodology for the enumeration combinatorial objects*, J. Difference Equations Appl. 5 (1999) 435–490.
- [4] P. BAUMANN AND C. HOHLWEG, *A Solomon theory for the wreath product  $G \wr \mathfrak{S}_n$* , preprint.
- [5] J. BERTRAND, *Calcul des probabilités. Solution d'un problème*, Comptes Rendus Acad. Sci. Paris, **105** (1887), 369.
- [6] N. DERSHOWITZ AND S. ZAKS, *The Cycle Lemma and Some Applications*, Europ. J. Combinatorics, **11** (1990), 35–40.
- [7] S. FOMIN AND N. READING, *Generalized cluster complexes and Coxeter combinatorics*, In preparation.
- [8] I. M. GESSEL, *Super Ballot Numbers*, J. Symbolic Computation, **14** (1992), 179–194.

- [9] P. HILTON AND J. PEDERSEN, *Catalan Numbers, their Generalizations, and their Uses*, Math. Intelligencer, **13** (1001), 64–75.
- [10] S. G. MOHANTY, *Lattice paths counting and application*, New York and London: Academic Press (1979).
- [11] L. W. SHAPIRO, S. GETU, W.-J. WOAN AND L. C. WOODSON, *The Riordan Group*, Discrete Appl. Math., **34** (1991), 229–239.
- [12] N. J. A. SLOANE, *A Handbook of integer sequences*, New York and London: Academic Press (1973).
- [13] R. STANLEY, *Enumerative Combinatorics, volume 2*, no. 62 in Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999.
- [14] R. STANLEY, *Catalan Addendum*, (web) <http://www-math.mit.edu/~rstan/ec/catadd.pdf>.

(Jean-Christophe Aval) LABRI, UNIVERSITÉ BORDEAUX 1, 351 COURS DE LA LIBÉRATION,  
33405 TALENCE CEDEX, FRANCE

*E-mail address:* [aval@labri.fr](mailto:aval@labri.fr)

*URL:* <http://www.labri.fr/~aval/>