# An Identity for the Central Binomial Coefficient

David Callan
Dept. of Statistics, University of Wisconsin-Madison, Madison, WI 53706
**callan@stat.wisc.edu**

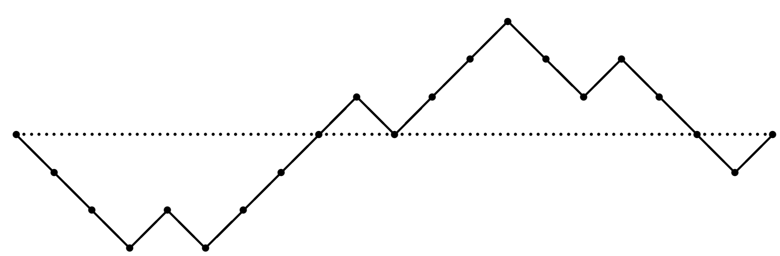
June 14, 2012

**Abstract**

We find the joint distribution of three simple statistics on lattice paths of n upsteps and n downsteps leading to a triple sum identity for the central binomial coefficient 2n-choose-n. We explain why one of the constituent double sums counts the irreducible pairs of compositions considered by Bender et al., and we evaluate some of the other sums.

## 1 Introduction

A *Grand-Dyck path* is a lattice path consisting of an equal number of upsteps $U = (1, 1)$ and downsteps $D = (1, -1)$. The horizontal line joining the endpoints is called *ground level* The number of upsteps is the *semilength* of a Grand-Dyck path, also known as its *size*. The number of Grand-Dyck paths of size $n$ is obviously the central binomial coefficient $\binom{2n}{n}$—choose locations for the upsteps among the $2n$ steps. A *Dyck path* is a Grand-Dyck path that never goes below ground level, and it is *primitive* if it is nonempty and its only vertices at ground level are its endpoints. The vertices at ground level of a nonempty Grand-Dyck path split it into *components*, each of which is a primitive Dyck path or an inverted primitive Dyck path. A *peak* in a Grand-Dyck path is an occurrence of $UD$ and a *low* peak is one that starts at ground level. A low peak is, in particular, a component of the path.



A Grand-Dyck path of semilength 10, with 1 low peak, 2 components above ground level, and 4 components altogether

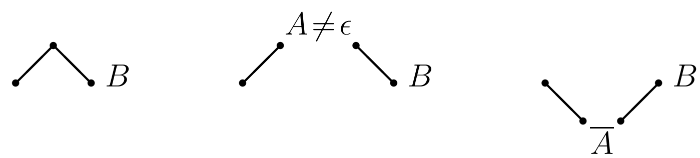
The generating function for Dyck paths counted by size is well known to be

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2x}.$$

In Section 2, we find the 4-variable generating function for Grand-Dyck paths counted by size, number of low peaks, number of components above ground level, and total number of components, and in Section 3, we find a closed formula for the joint distribution of these four statistics. In Section 4, we observe that the irreducible pairs of compositions considered by Bender et al. [1] are equinumerous with low-peak-free Grand-Dyck paths, and we give a bijective explanation.

## 2 A generating function

Let $F(x,y,z,w)$ denote the generating function for Grand-Dyck paths where $x$ marks size, $y$ marks number of low peaks, $z$ marks number of components above ground level, and $w$ marks total number of components. The first return to ground level partitions nonempty Grand-Dyck paths into the 3 classes illustrated below, where $A$ denotes a Dyck path, $\overline{A}$ an inverted Dyck path, and $B$ a Grand-Dyck path.




A first return decomposition for nonempty Grand-Dyck paths

From this decomposition, we see that

$$F = 1 \text{ [for the empty path]} + xyzwF + x(C(x)-1)zwF + xC(x)wF,$$

an equation with solution

$$F(x,y,z,w) = \frac{2}{2+2wxz(1-y)-w(1+z)(1-\sqrt{1-4x})}\,. \tag{1}$$

In particular, the generating function to count Grand-Dyck paths with no low peaks is

$$F(x,0,1,1) = \frac{1}{x+\sqrt{1-4x}}\,, \tag{2}$$

and the generating function to count Grand-Dyck paths by number of components above ground level is

$$F(x,1,z,1) = \frac{2}{(z+1)\sqrt{1-4x}-z+1} \tag{3}$$

## 3 An explicit count

To obtain an explicit expression for the number $u(n,i,j,k)$ of Grand-Dyck paths of size $n$ with $i$ low peaks, $j$ components above ground level, and $k$ components altogether, first observe that there are $\binom{k}{i}$ ways to place the low peaks among the components. This reduces the problem to finding an expression for $v(n, j, k)$, the number of Grand-Dyck paths of size $n$ with no low peaks, $j$ components above ground level, and $k - j$ components below ground level. There are $\binom{k}{j}$ ways to arrange the above- and below-ground level components, so we may assume all components above ground level precede those below ground level. Each component above ground level has the form $UUPDQD$ where $P$ and $Q$ are (possibly empty) Dyck paths; each component below ground level has the form $D\overline{R}U$ where $\overline{R}$ is a Dyck path $R$ flipped over. Make the reversible transformations $UUPDQD \to UPDUQD$ and $D\overline{R}U \to URD$. Thus we see that the Grand-Dyck paths in question are equinumerous with Dyck paths of size $n$ and $2j + (k - j) = j + k$ components. It is well known that the number of Dyck paths of size $n$ with $k$ components is the generalized Catalan number $\frac{k}{2n-k}\binom{2n-k}{n-k}$ (arising as a $k$-fold convolution of Catalan numbers). Hence,

$$v(n, j, k) = \binom{k}{j} \frac{j + k}{2n - j - k} \binom{2n - j - k}{n - j - k},$$

and, as noted above, $u(n, i, j, k) = \binom{k}{i} v(n - i, j - i, k - i)$, yielding

$$u(n, i, j, k) = \binom{k}{i} \binom{k - i}{j - i} \frac{k - 2i + j}{2n - j - k} \binom{2n - j - k}{n - i}.$$

When $i = n$, as for the "sawtooth" path $(UD)^n$, we have $j = k = n$ and the expression for $u(n, i, j, k)$ is indeterminate; we must interpret it as 1.

Summing over $i, j, k$, we have the identity

$$\binom{2n}{n} = 1 + \sum_{\substack{i,j,k \\ 0 \le i \le j \le k \le n \\ j+k<2n}} \binom{k}{i} \binom{k - i}{j - i} \frac{k - 2i + j}{2n - j - k} \binom{2n - j - k}{n - i}. \tag{4}$$

## 4 Irreducible pairs of compositions

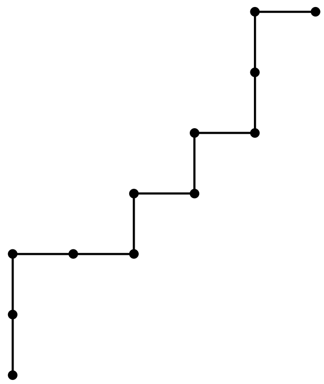
Bender et al. [1] made the following definition.

Let $n = b_1 + \cdots + b_k = b'_1 + \cdots + b'_k$ be a pair of compositions of $n$ into $k$ positive parts. We say this pair is *irreducible* if there is no positive $j < k$ for which $b_1 + \cdots + b_j = b'_1 + \cdots + b'_j$.

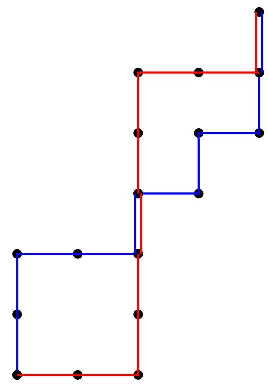
They showed that the number $f(n)$ of irreducible ordered pairs of compositions of $n$ into the same (unspecified) number of parts has the generating function

$$\sum_{n \ge 0} f(n + 1) x^n = \frac{1}{x + \sqrt{1 - 4x}}. \tag{5}$$

The generating functions in (2) and (5) are the same, which implies that irreducible pairs of compositions of $n+1$ are equinumerous with low-peak-free Grand-Dyck paths of size $n$. It is not too hard to show this bijectively using a lattice path representation of compositions as illustrated below. Each entry $a_i$ in a composition $(a_1, \ldots, a_k)$ contributes $a_i - 1$ North steps followed by 1 East step.



The lattice path of the composition $(3\,1\,2\,2\,3)$



The paths for the irreducible pair $\big((3\,1\,2\,2\,3),\, (1\,1\,6\,1\,2)\big)$ omitting the last steps, necessarily East

The diagram on the right above represents an irreducible ordered pair of compositions. By definition of irreducible, no East step in the first (blue) path coincides with an East step in the second (red) path. The vertices common to both paths split the diagram into path pairs that form parallelogram polyominoes, possibly the degenerate polyomino consisting of 2 coincident North steps. There are several bijections [2, Ex. 6.19 $\ell$] from parallelogram polyominoes of size (semiperimeter) $k$ to Dyck paths of size $k-1$. By elevating the resulting Dyck path (prepend an upstep, append a downstep), we get a size-preserving bijection from parallelogram polyominoes to primitive Dyck paths. So apply this bijection to each parallelogram polyomino to get a primitive Dyck path and use the color of the upper path to determine whether to flip it over. The degenerate polyomino corresponds to the Dyck path $UD$ and we always flip this over because we don't want any low peaks. Lastly, concatenate the Dyck paths to obtain a low-peak-free Grand-Dyck path. This is the desired bijection from irreducible pairs of compositions of $n+1$ to low-peak-free Grand-Dyck paths of size $n$.

Emanuele Munarini notes in the OEIS [3] entry for the sequence A081696 that the number of irreducible pairs of compositions of $n+1$ can be expressed as $\sum_{j=0}^{n} \frac{3j+1}{n+j+1}\binom{2n-j}{n-2j}$. In fact, our results can be used to get the following counts.

1. The number of low-peak-free Grand-Dyck paths with $j$ components above ground level is Munarini's summand $\frac{3j+1}{n+j+1}\binom{2n-j}{n-2j}$.

2. The number of low-peak-free Grand-Dyck paths with $j$ components above ground level and $k$ components altogether is $\frac{j+k}{2n-j-k}\binom{2n-j-k}{n-j-k}\binom{k}{j}$.

3. The number of unrestricted Grand-Dyck paths with $j$ components above ground level is $\frac{2j+1}{2n+1}\binom{2n+1}{n-j}$.

4. The number of unrestricted Grand-Dyck paths with $j$ components above ground level and $k$ components altogether is $\frac{k}{2n-k}\binom{2n-k}{n-k}\binom{k}{j}$.

5. The number of unrestricted Grand-Dyck paths with $j$ big components above ground level is $\frac{2j+1}{n+1}\binom{2n+2}{n-2j}$ (a big component is one of size $\geq 2$).

Are there bijective proofs?

# References


[1] Edward A. Bender, Gregory F. Lawler, Robin Pemantle, and Herbert S. Wilf, Irreducible compositions and the first return to the origin of a random walk, *Séminaire Lotharingien de Combinatoire* **50** (2004), Article B50h.

[2] Richard P. Stanley, *Enumerative Combinatorics* Vol. 2, Cambridge University Press, 1999. Exercise 6.19 and related material on Catalan numbers are available online at http://www-math.mit.edu/~rstan/ec/ .

[3] The On-Line Encyclopedia of Integer Sequences, published electronically at http://oeis.org/, 2012.