A COMBINATORIAL PROOF OF SHAPIRO'S CATALAN CONVOLUTION

Gábor V. Nagy

Bolyai Institute, University of Szeged, Szeged, H-6720, Hungary

Abstract. Shapiro proved an elegant convolution formula involving Catalan numbers of even index. This paper gives a combinatorial proof of his formula. In addition, we show that it is equivalent to an alternating convolution formula of central binomial coefficients.

1. Introduction

In this paper C_n denotes the nth Catalan number and B_n denotes the nth central binomial coefficient, i.e. $C_n = \frac{1}{n+1}\binom{2n}{n}$ and $B_n = \binom{2n}{n}$. Unless otherwise stated, all indices (i, j, k, and so on) are nonnegative integers in our formulas.

In 2002, L. Shapiro found the following elegant identity [3; p. 123]:

Theorem 1.

$$\sum_{i+j=n} C_{2i}C_{2j} = 4^n C_n. \tag{1}$$

This can be easily proved using generating functions, but according to Stanley [5; p. 46], no simple direct combinatorial proof has been known (see [1] for another combinatorial proof). In Section 3 we will give a simple combinatorial proof of the the following equivalent version of Theorem 1:

Theorem 2.

$$\sum_{i+j=n} C_{2i}B_{2j} = 4^n B_n. \tag{2}$$

(1) and (2) are equivalent, because

$$\begin{aligned}\sum_{i+j=n} C_{2i}B_{2j} &= \frac{1}{2}\left(\sum_{i+j=n} C_{2i}(2j+1)C_{2j} + \sum_{i+j=n}(2i+1)C_{2i}C_{2j}\right)\\ &= \frac{1}{2}\sum_{i+j=n}(2n+2)C_{2i}C_{2j} = (n+1)\sum_{i+j=n} C_{2i}C_{2j}.\end{aligned}$$

1991 *Mathematics Subject Classification.* 05A10; 05A19.

Key words and phrases. Catalan numbers; Shapiro convolution.

Supported by the TÁMOP-4.2.1/B-09/1/KONV-2010-0005 project.

The publication is supported by the European Union and co-funded by the European Social Fund. Project title: "Broadening the knowledge base and supporting the long term professional sustainability of the Research University Centre of Excellence at the University of Szeged by ensuring the rising generation of excellent scientists." Project number: TÁMOP-4.2.2/B-10/1-2010-0012

The key observation of this paper is a non-standard interpretation of C_{2n}, that is discussed in the next section. Using that, we can give a new combinatorial meaning of the left-hand sides of (1) and (2). In Section 4, we show bijectively that (2) is equivalent with the alternating convolution formula of central binomial coefficients, which has a nice combinatorial interpretation, due to Spivey [4].

2. Even-zeroed balanced paths and C_{2n}

A *path of length* l is an l-element sequence of up-steps ($\nearrow$) and down-steps ($\searrow$). A *balanced* n*-path* is such a path of length $2n$ that has n up-steps and n down-steps. The number of balanced n-paths is clearly B_n. We denote by $\mathcal{B}_n$ the set of balanced n-paths. We visualize paths in the usual way: They start from the origin, $\nearrow$ is a step $(1,1)$ and $\searrow$ is a step $(1,-1)$; see the figures below. An n*-Dyck-path* (of length $2n$) is a balanced n-path such that it never falls below the x-axis. It is well-known that the number of n-Dyck-paths is C_n. We denote by $\mathcal{C}_n$ the set of n-Dyck-paths. A *signed* n*-Dyck-path* is an element of the set $\mathcal{S}_n := \{+,-\} \times \mathcal{C}_n$. The set of signed Dyck-paths is denoted by $\mathcal{S} := \bigcup_{i=0}^{\infty} \mathcal{S}_i$. The number of up-steps (or down-steps) in a signed or unsigned balanced path P is called the *parameter* of P and it is denoted by $\mathrm{par}(P)$. A special type of path plays a crucial role in our proofs, so we introduce a new terminology for them: We call a (balanced or non-balanced) path *even-zeroed*, if its x-intercepts are all divisible by 4.

If $n \geq 1$, then every balanced n-path can be decomposed uniquely into a sequence of signed Dyck-paths in a very natural way (see Figure 1): The x-axis cuts the balanced path into nonempty subpaths so that every subpath is either a Dyck-path that never touches the x-axis (apart from its starting point and end point), or the reflection of such a Dyck-path across the x-axis. Every subpath is uniquely characterized by a signed Dyck-path that we get after removing the first and last steps ($+$: standard Dyck-path, $-$: reflected Dyck-path), and we can list these signed Dyck-paths (from left to right) in a sequence. It is very easy to see that we defined a bijection χ between $\mathcal{B}_n$ and $\mathrm{SEQ}(n)$, where

$$\mathrm{SEQ}(n) = \left\{ (P_1, \ldots, P_k) : k \in \mathbb{Z}^+, P_i \in \mathcal{S} \text{ for all } i, \text{ and } \sum_{i=1}^{k} (\mathrm{par}(P_i) + 1) = n \right\}.$$

Now we are ready to prove the key lemma of this paper:

Lemma 3. *C_{2n} counts the number of even-zeroed balanced $2n$-paths.*

Proof. The statement is true for $n = 0$. Now let us assume that $n \geq 1$.

Clearly, a balanced $2n$-path P is even-zeroed, if and only if all the signed Dyck-paths in its sequence $\chi(P)$ have odd parameter. So if we denote by $\tilde{\mathcal{B}}_{2n}$ the set of even-zeroed $2n$-paths, the restriction of χ to $\tilde{\mathcal{B}}_{2n}$ gives a bijection ϕ between $\tilde{\mathcal{B}}_{2n}$ and $\widetilde{\mathrm{SEQ}}(2n)$, where

$$\widetilde{\mathrm{SEQ}}(2n) = \left\{ (P_1, \ldots, P_k) : P_i \in \mathcal{S},\ \mathrm{par}(P_i) \text{ is odd } \forall i;\ \sum_{i=1}^{k} (\mathrm{par}(P_i) + 1) = 2n \right\}.$$

Now we define a bijection ψ between $\mathcal{C}_{2n}$ and $\widetilde{\mathrm{SEQ}}(2n)$ and so we give a bijective proof of the lemma ($\phi^{-1} \circ \psi$ is a bijection between $\mathcal{C}_{2n}$ and $\tilde{\mathcal{B}}_{2n}$). Consider an arbitrary $2n$-Dyck-path D. It is well-known from a standard proof of the Catalan recursion that D can be uniquely written as $\nearrow L \searrow R$, i.e. D can be decomposed into an ordered pair (L, R), where L and R are Dyck-paths, whose parameters sum to $2n-1$. Either $\mathrm{par}(L)$ or $\mathrm{par}(R)$ is odd. If $\mathrm{par}(L)$ is odd, then we define the first element of $\psi(D)$ to be $-L$, and we recursively repeat the process for $D' := R$ ($\mathrm{par}(R)$ is even) to get the other elements of $\psi(D)$. If $\mathrm{par}(R)$ is odd, then we define the first element of the $\psi(D)$ to be $+R$, and recursively repeat the process for $D' := L$. ($-$ means "left", $+$ means "right" here.) The process terminates when D' is the empty 0-Dyck-path. It is easy to check that the obtained $\psi(D)$ is in $\widetilde{\mathrm{SEQ}}(2n)$. See Figure 2 for a visualization.

One can easily compute $\psi^{-1}(S)$ for an arbitrary $S \in \widetilde{\mathrm{SEQ}}(2n)$, so we indeed defined a bijection. $\square$

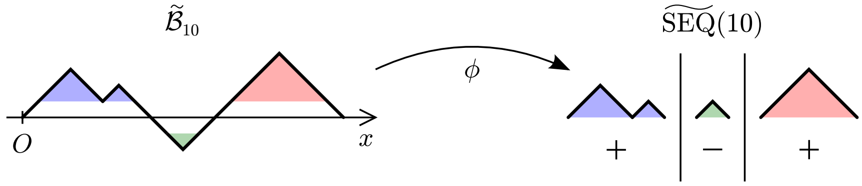

Figure 1: Illustration of χ and ϕ

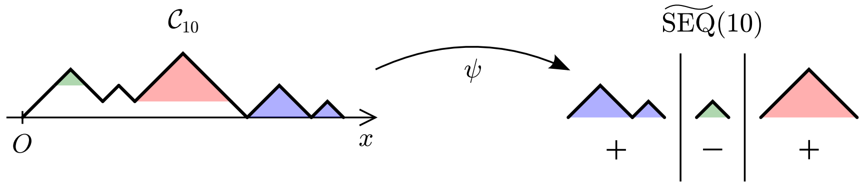

Figure 2: Illustration of ψ

Remarks. Roughly speaking, our bijection $\mathcal{C}_{2n} \to \tilde{\mathcal{B}}_{2n}$ converts the "left-right symmetry" of $\mathcal{C}_{2n}$ into the "up-down symmetry" of $\tilde{\mathcal{B}}_{2n}$.

When defining ψ, it might be slightly more natural to work with full binary trees that is an other representation of Catalan numbers (what we do here is to decompose even-parameter full binary trees into odd-parameter subtrees – the details are left to reader). But then $\phi^{-1} \circ \psi$ would become slightly less intuitive, since we would need an extra conversion between full binary trees and Dyck-paths.

If we already know or conjecture that $|\tilde{\mathcal{B}}_{2n}| = C_{2n}$, we can find a quicker (but recursive) argument for this. Namely, using the notations $X_n := |\tilde{\mathcal{B}}_{2n}|$ and $Y_n := C_{2n}$, one can quickly figure out that both $(X_n)_{n=0}^{\infty}$ and $(Y_n)_{n=0}^{\infty}$ satisfy the following recursion: $Z_0 = 1, \quad Z_n = 2\sum_{k=1}^{n} C_{2k-1} Z_{n-k}$ (if $n \geq 1$).

As an application, we prove a lemma, from which a recursive proof of Theorem 2 can be obtained.

Lemma 4.
$$2 \cdot \sum_{i+j+k=n} C_{2i} C_{2j} B_{2k} = B_{2n+1}.$$

Proof. Both sides count the number of balanced $(2n+1)$-paths. This is obvious for the right-hand side. In the left-hand side, we group the balanced $(2n+1)$-paths by the position (i) and length (j) of the leftmost such signed Dyck-path segment (cut by the x-axis) whose parameter is odd. Such a segment must exist, since the sum of the parameters is $2n+1$. If the starting point of that segment is $4i$ and its parameter is $2j+1$, then by Lemma 3, there are C_{2i} even-zeroed balanced paths from the origin to $4i$, there are $2C_{2j}$ possible choices for the segment in question, and there are B_{2k} possible endings for the rest of the path ($k = n-i-j$). $\square$

Remark. Using the well-known [2] convolution identity $\sum_{i+j=n} B_i B_j = 4^n$ and Lemma 4, it is easy to see that that both sides of (2) satisfy the following recursion: $X_0 = 1, \quad \sum_{s+t=n} X_s X_t = 16^n$. This is because

$$\sum_{s+t=n} 4^s B_s \cdot 4^t B_t = 4^n \cdot \sum_{s+t=n} B_s B_t = 4^n \cdot 4^n = 16^n,$$

and

$$\begin{aligned}
\sum_{s+t=n}\left(\sum_{i+k=s} C_{2i}B_{2k}\right)\left(\sum_{j+l=t} C_{2j}B_{2l}\right) &= \sum_{i+j+k+l=n} C_{2i}C_{2j}B_{2k}B_{2l} \\
&= \sum_{m+l=n}\left(\sum_{i+j+k=m} C_{2i}C_{2j}B_{2k}\right)B_{2l} = \frac{1}{2}\cdot\sum_{m+l=n} B_{2m+1}B_{2l} \\
&= \frac{1}{2}\cdot\frac{1}{2}\cdot 4^{2n+1} = 16^n.
\end{aligned}$$

3. The proof of Theorem 2

The following lemma is well-known, and it has several combinatorial proofs [2].

Lemma 5. *B_n counts the number of paths of length $2n$ that never return to the x-axis after the first step.*

With the help of Lemma 3, we can give an interesting combinatorial interpretation of the left-hand sides of (1) and (2).

Lemma 6. *a) $\sum_{i+j=n} C_{2i}B_{2j}$ is the number of even-zeroed paths of length $4n$.*
b) $\sum_{i+j=n} C_{2i}C_{2j}$ is the number of even-zeroed paths from the origin to $(4n+1,1)$.

Proof. a) By Lemmas 3 and 5, there are $C_{2i}B_{2(n-i)}$ such even-zeroed paths of length $4n$ whose rightmost x-intercept is $4i$.
b) There are $C_{2i}C_{2(n-i)}$ such even-zeroed paths from the origin to $(4n+1,1)$ whose rightmost x-intercept is $4i$ (followed by an up-step). The first factor comes from Lemma 3 and the second one comes from the standard interpretion of $C_{2(n-i)}$ by Dyck-paths. $\square$

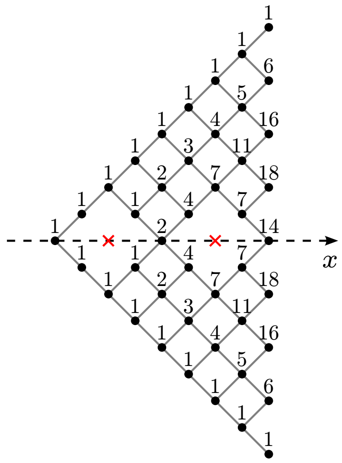

Figure 3: The number of even-zeroed paths

In Figure 3 the label of a node shows the number of even-zeroed paths from the origin to that node. These labels can be calculated recursively, since every label is the sum of its left neighbors. We already know that the label of $(4n, 0)$ is C_{2n}, the label of $(4n+1, \pm 1)$ is $L_n := \sum_{i+j=n} C_{2i}C_{2j}$ and the sums of the labels in the $4n$th column is $S_n := \sum_{i+j=n} C_{2i}B_{2j}$. In order to prove Theorem 2, we only have to show that $S_n = 4^n B_n$. The key observation is that S_{n+1} can be calculated from S_n and L_n easily, but we know from Section 1 that $L_n = \frac{1}{n+1}S_n$, so in fact S_{n+1} can be calculated from S_n easily. This calculation is done in the next lemma, which implies Theorem 2.

Lemma 7. *The number of even-zeroed paths of length* $4n$ *is* $4^n B_n$.

Proof. Let $\mathcal{P}_n$ denote the set of even-zeroed paths of length $4n$, and set $S_n := |\mathcal{P}_n|$. By induction on n, we prove that $S_n = 4^n B_n$. This is obviously true if $n = 0$.

Let us assume that $S_n = 4^n B_n$ holds. Clearly, every path of $\mathcal{P}_{n+1}$ is an extension of a path of $\mathcal{P}_n$ by 4 steps. For each path of $\mathcal{P}_n$ there are 16 possible extensions. But some of the $16S_n$ extensions are not in $\mathcal{P}_{n+1}$. These "wrong" extenstions are exactly the even-zeroed paths from the origin to $(4n+1, 1)$ followed by a down-step and two arbitrary steps, and the reflections of these paths across the x-axis. By Lemma 6.b, the number of these wrong extensions is $8\sum_{i+j=n} C_{2i}C_{2j}$, that equals to $\frac{8}{n+1}\sum_{i+j=n} C_{2i}B_{2j} = \frac{8}{n+1}S_n$, as seen in Section 1 and Lemma 6.a. By the induction hypothesis, $S_n = 4^n B_n$, thus $S_{n+1} = 16 \cdot 4^n B_n - \frac{8}{n+1}4^n B_n$. A quick calculation shows that $S_{n+1} = 4^{n+1}B_{n+1}$. $\square$

4. Alternating convolution of the central binomial coefficients

The following theorem has a nice combinatorial proof using random colored permutations, due to Spivey [4]:

Theorem 8.

$$\sum_{i+j=n} B_{2i}B_{2j} - \sum_{\substack{i+j=n\\ j\geq 1}} B_{2i+1}B_{2j-1} = 4^n B_n.$$

By proving the next theorem bijectively, we will see that Theorem 8 is equivalent with Theorem 2, so any combinatorial proof of Theorem 8 yields a combinatorial proof of Theroem 2. Conversely, our proof in the previous section can be interpreted as a new proof of Theorem 8.

Theorem 9.

$$\sum_{i+j=n} B_{2i}B_{2j} - \sum_{\substack{i+j=n\\ j\geq 1}} B_{2i+1}B_{2j-1} = \sum_{i+j=n} C_{2i}B_{2j}.$$

Proof. Using Lemma 3, we will prove the following equivalent form:

$$\sum_{i+j=n} (B_{2i} - C_{2i})B_{2j} = \sum_{\substack{i+j=n\\ j\geq 1}} B_{2i+1}B_{2j-1}. \tag{3}$$

The right-hand side counts the number of pairs (O_1, O_2), where O_1 and O_2 are balanced paths with odd parameters, and $\mathrm{par}(O_1)+\mathrm{par}(O_2) = 2n$. Let $\mathcal{O}$ be the set of these pairs. By Lemma 3, the left-hand side counts the number of pairs (E_1, E_2), where E_1 and E_2 are balanced paths with even parameters, E_1 has an x-intercept of the form $4t + 2$ (for some integer t), and $\mathrm{par}(E_1) + \mathrm{par}(E_2) = 2n$. Let $\mathcal{E}$ be the set of these pairs.

We will give a bijection between $\mathcal{E}$ and $\mathcal{O}$, which means that $|\mathcal{E}| = |\mathcal{O}|$, as stated. Pick an arbitrary element (E_1, E_2) of $\mathcal{E}$. Let L be the subpath of E_1 which is identical with E_1 from the origin to its leftmost x-intercept of the form $4t+2$, and let R be the rest of E_1. Then the image of (E_1, E_2) is defined as (LE_2, R), where LE_2 is the concatenation of L and E_2 in this order. It is easy to see that this mapping is bijective. □

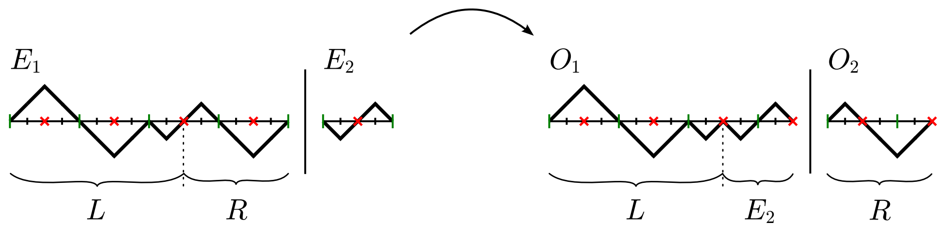

Figure 4: Illustration of the proof of Theorem 9

If we write $C_{2i} = \binom{4i}{2i} - \binom{4i}{2i-1}$ in (3), we get the following identity:

Corollary 10.

$$\sum_{i=1}^{n} \binom{4i}{2i-1}\binom{4n-4i}{2n-2i} = \sum_{i=0}^{n-1} \binom{4i+2}{2i+1}\binom{4n-4i-2}{2n-2i-1}.$$

References

[1] G. E. Andrews, On Shapiro's Catalan convolution, Adv. in Appl. Math. 46 (2011) 15–24.
[2] Ö. Eğecioğlu, A. King, Random walks and Catalan factorization, Congr. Numer. 138 (1999) 129–140.
[3] T. Koshy, Catalan Numbers with Applications, Oxford University Press, New York, 2009.
[4] M. Spivey, Combinatorial interpretation of the alternating convolution of the central binomial coefficients, http://mikespivey.wordpress.com/2012/03/16/altconvcentralbinom/, 2012.
[5] R. P. Stanley, Bijective proof problems, http://www-math.mit.edu/~rstan/bij.pdf, 2009.