A Very Short (Bijective!) Proof of Touchard’s Catalan Identity

By Amitai REGEV, Nathaniel SHAR, and Doron ZEILBERGER

Recall that one of the almost infinitely many definitions of the ubiquitous Catalan Numbers, \( C_n \), is as the cardinality the set of Dyck paths of length \( 2n \), let’s call it \( C_n \):

\[
C_n := \{w_1 \ldots w_{2n} \in \{-1, 1\}^{2n} | \sum_{j=1}^{i} w_j \geq 0 (1 \leq i < 2n), \sum_{j=1}^{2n} w_j = 0 \} .
\]

In the 1924 Toronto ICM, Touchard [T] announced (and proved) the elegant identity

\[
C_{n+1} = \sum_{k \leq 0} \binom{n}{2k} 2^{n-2k} C_k . \tag{Touchard}
\]

Here is a very short, purely bijective, proof, even nicer than Lou Shapiro’s [S].

Let \( f(1, 1) = 1, f(1, -1) = 0, f(-1, 1) = 0, f(-1, -1) = -1 \), where 0 is a twin-sister of 0, whose value is also 0. Define

\[
T(w_1 \ldots w_{2n+2}) := f(w_1, w_2)f(w_3, w_4)\ldots f(w_{2n+1}, w_{2n+2}) .
\]

This is a bijection from \( C_{n+1} \) onto the set

\[
S'_{n+1} := \{w_1 \ldots w_{n+1} \in \{-1,0,0,1\}^{n+1} | \sum_{j=1}^{i} w_j \geq 0 (1 \leq i \leq n), \sum_{j=1}^{n+1} w_j = 0 , \sum_{j=1}^{i} w_j = 0 \Rightarrow w_{i+1} \neq 0 \} .
\]

But this latter set is in bijection with the set

\[
S_n := \{w_1 \ldots w_n \in \{-1,0,0,1\}^n | \sum_{j=1}^{i} w_j \geq 0 (1 \leq i < n), \sum_{j=1}^{n} w_j = 0 \} ,
\]

as follows. For \( w = w_1 \ldots w_{n+1} \in S'_{n+1} \), if \( w_{n+1} = 0 \) just chop that last letter, mapping it to \( w_1 \ldots w_n \). Otherwise, of course \( w_{n+1} = -1 \) (it can’t be 1, and it can’t be 0), so write \( w \) as \( 1 := -1 \) \( w' \) \( 1 \) \( w'' \) (where \( w' \in S'_k \) and \( w'' \in S'_{n-k} \) for some \( 0 \leq k \leq n \), and map it to \( w'0w'' \)).

Now the cardinality of \( S_n \) is obviously given by the right side of Eq. (Touchard). Indeed, let the number of ones be \( k \) \( (0 \leq k \leq n/2) \), then there are also \( k \) minus-ones. There are \( \binom{n}{2k} \) ways to choose the locations of the 1’ and −1’s, \( C_k \) ways of forming them into a member of \( C_k \), and \( 2^{n-2k} \) ways of deciding which kind of zero (0 or 0) will occupy the remaining \( n - 2k \) slots. \( \square \)
Remarks

1. While it is nice to give nice bijective proofs, let us note that today, thanks to WZ proof theory, the epistemological stature of identities like Touchard’s is the same as that of the identity $2 + 2 = 1 + 3$. Indeed just go to Maple, and type:

```
SumTools[Hypergeometric][Zeilberger](binomial(n,2*k)*2**(n-2*k)*binomial(2*k,k)/(k+1)/binomial(2*n+2,n+1)*(n+2),n,k,N);
```

2. Another way of counting $S_n$ is to partition it according to the number of occurrences of 0, say $n - k$, then choose the $\binom{n}{n-k}$ locations of the 0 and ‘fill-in’ the remaining $k$ slots by a so-called Motzkin word of length $k$, i.e. a word in the alphabet $\{-1, 0, 1\}$, that add-up to 0, and whose partial sums are non-negative, yielding the equally elegant identity

$$C_{n+1} = \sum_{k=0}^{n} \binom{n}{k} M_k.$$

While this identity is ‘trivially equivalent’ to quite a few known identities, and is ‘well-known to the experts’ we were unable to find it in the literature.

3. We intentionally avoided drawing diagrams, but most human readers will probably better appreciate the beauty of the proof by drawing a random Dyck path in $C_{n+1}$, and then replace 11 (alias up-up) by and Up Step, replace II by a Down Step, replace 11 by a green horizontal step, and replace I1 by a red horizontal step. Then $S'_{n+1}$ are generalized Motzkin paths of length $n$, with two types of horizontal steps, green and red, where a red horizontal step may not lie on the x-axis, and $S_n$ is the set of such $n$-letter words without this restriction. The bijection between $S'_{n+1}$ and $S_n$ consists of removing the last step, if it is a green horizontal step (of course it can’t be a red horizontal step), and otherwise looking at the ‘Up-mate’ of the last step (that must necessarily be a Down step), and replace that Up-Mate by a red horizontal step, and at the same time delete the above-mentioned last Down step.

4. We thank Lou Shapiro for telling us that we rediscovered Touchard’s identity (in its almost-equivalent form in terms of Motzkin numbers cited above), and telling us about [S]. While we admire Shapiro’s combinatorial proof, it it is not purely bijective, and makes use of generating functions.

5. Our bijection is a renormalization-group transformation, where we ‘renormalized’ a word of length $2n + 2$ into a word half as long, but with more letters in the underlying alphabet. It may be interesting to see if one can get less trivial identities by considering generalized Dyck words where the fundamental steps are drawn from a larger set then just $\{(1, -1), (1, 1)\}$.
References

[S] L. W. Shapiro, *A short proof of an identity of Touchard’s concerning Catalan Numbers*


Amitai Regev, Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. amitai.regev at weizmann dot ac dot il ; http://www.wisdom.weizmann.ac.il/~regev/.

Nathaniel Shar and Doron Zeilberger, Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. [nshar, zeilberg] at math dot rutgers dot edu ,

March 10, 2015