Proof of a subtle combinatorial identity

Idris Mercer

The identity

$$\binom{2n}{2m} = \sum_{k=0}^{m} \binom{2m+1}{2k+1}\binom{n+k}{2m} \tag{1}$$

appears in Problem 14 of Chapter 1 of Riordan's classic work [1], as well as in a 2016 Stack Exchange post [2]. At first glance, one might guess that this identity would be easy to prove, but it turns out to be surprisingly subtle.

Answers in the Stack Exchange post contain proofs using inclusion-exclusion, generating functions, and complex analysis. The instructions for Riordan's Problem 14 involve using results from earlier problems.

In this note, we prove the identity by a method that may take several steps, but is largely self-contained and is combinatorial in spirit.

Lemma

Suppose a, b, and c are positive integers. Suppose Z is a set of bitstrings of length $a+b$. Suppose that for every $\mathbf{x} \in Z$, at most $2c$ of the last b bits of $\mathbf{x}$ are 1. Then Z is the disjoint union $S_1 \cup S_2 \cup S_3$, where

$$\begin{aligned}
S_1 &= \{\mathbf{x} \in Z : \text{exactly } 2c \text{ of the last } b \text{ bits are } 1\},\\
S_2 &= \{\mathbf{x} \in Z : \text{the } (a{+}1)\text{th bit is } 0, \text{ and exactly } 2c{-}1 \text{ of the last } b{-}1 \text{ bits are } 1\},\\
S_3 &= \{\mathbf{x} \in Z : \text{at most } 2c{-}2 \text{ of the last } b{-}1 \text{ bits are } 1\}.
\end{aligned}$$

Proof. First note that Z can be written as the disjoint union $T_1 \cup T_2 \cup T_3$, where

$$\begin{aligned}
T_1 &= \{\mathbf{x} \in Z : \text{exactly } 2c \text{ of the last } b \text{ bits are } 1\},\\
T_2 &= \{\mathbf{x} \in Z : \text{exactly } 2c{-}1 \text{ of the last } b \text{ bits are } 1\},\\
T_3 &= \{\mathbf{x} \in Z : \text{at most } 2c{-}2 \text{ of the last } b \text{ bits are } 1\}.
\end{aligned}$$

Then note that T_2 can be written as the disjoint union $T_2' \cup T_2''$, where

$$\begin{aligned}
T_2' &= \{\mathbf{x} \in Z : \text{the } (a{+}1)\text{th bit is } 0, \text{ and exactly } 2c{-}1 \text{ of the last } b{-}1 \text{ bits are } 1\},\\
T_2'' &= \{\mathbf{x} \in Z : \text{the } (a{+}1)\text{th bit is } 1, \text{ and exactly } 2c{-}2 \text{ of the last } b{-}1 \text{ bits are } 1\}.
\end{aligned}$$

Finally, observe that we have $S_1 = T_1$, $S_2 = T_2'$, and $S_3 = T_2'' \cup T_3$. $\square$

We will apply this lemma repeatedly to the following situation. Let n and m be positive integers with $n \geq m$, and let Y be the set of all bitstrings of length $2n$ of which exactly $2m$ bits are 1, so $|Y| = \binom{2n}{2m}$. We first apply the lemma with

$$a = n,\ b = n,\ c = m,\ Z = Y.$$

(Note that for all $\mathbf{x} \in Y$, it is true that at most $2c = 2m$ of the last b bits of $\mathbf{x}$ are 1.) We conclude that Y is the disjoint union $S_1 \cup S_2 \cup S_3$, where

$$\begin{aligned}
S_1 &= \{\mathbf{x} \in Y : \text{exactly } 2m \text{ of the last } n \text{ bits are } 1\},\\
S_2 &= \{\mathbf{x} \in Y : \text{the } (n{+}1)\text{th bit is } 0, \text{ and exactly } 2m{-}1 \text{ of the last } n{-}1 \text{ bits are } 1\},\\
S_3 &= \{\mathbf{x} \in Y : \text{at most } 2m{-}2 \text{ of the last } n{-}1 \text{ bits are } 1\}.
\end{aligned}$$

We then apply the lemma with

$$a = n+1,\ b = n-1,\ c = m-1,\ Z = S_3$$

to conclude that S_3 is the disjoint union $S_1^* \cup S_2^* \cup S_3^*$, where

$$\begin{aligned}
S_1^* &= \{\mathbf{x} \in S_3 : \text{exactly } 2m-2 \text{ of the last } n{-}1 \text{ bits are } 1\},\\
S_2^* &= \{\mathbf{x} \in S_3 : \text{the } (n{+}2)\text{th bit is } 0\text{, and exactly } 2m{-}3 \text{ of the last } n{-}2 \text{ bits are } 1\},\\
S_3^* &= \{\mathbf{x} \in S_3 : \text{at most } 2m{-}4 \text{ of the last } n{-}2 \text{ bits are } 1\}.
\end{aligned}$$

We continue this process. We next apply the lemma with

$$a = n+2,\ b = n-2,\ c = m-2,\ Z = S_3^*$$

to conclude that S_3^* is the disjoint union $S_1^{**} \cup S_2^{**} \cup S_3^{**}$, where

$$\begin{aligned}
S_1^{**} &= \{\mathbf{x} \in S_3^* : \text{exactly } 2m-4 \text{ of the last } n{-}2 \text{ bits are } 1\},\\
S_2^{**} &= \{\mathbf{x} \in S_3^* : \text{the } (n{+}3)\text{th bit is } 0\text{, and exactly } 2m{-}5 \text{ of the last } n{-}3 \text{ bits are } 1\},\\
S_3^{**} &= \{\mathbf{x} \in S_3^* : \text{at most } 2m{-}6 \text{ of the last } n{-}3 \text{ bits are } 1\},
\end{aligned}$$

and so on. This means that Y is the disjoint union

$$S_1 \cup S_2 \cup S_1^* \cup S_2^* \cup S_1^{**} \cup S_2^{**} \cup \cdots$$

where the process terminates with a set in which at most $2m - 2m = 0$ of the last $n - m$ bits are 0. We may also note that the statements appearing in the definitions of the sets

$$S_1,\ S_2,\ S_1^*,\ S_2^*,\ S_1^{**},\ S_2^{**},\ \ldots$$

are pairwise incompatible for any element of Y. Hence the conditions $\mathbf{x} \in S_3$, $\mathbf{x} \in S_3^*$, … in the definitions of those sets can be replaced with the condition $\mathbf{x} \in Y$. Thus, if we define

$$V_{2m} = S_1,\ V_{2m-1} = S_2,\ V_{2m-2} = S_1^*,\ V_{2m-3} = S_2^*,\ V_{2m-4} = S_1^{**},\ V_{2m-5} = S_2^{**},\ \ldots,$$

then the above argument proves the following.

Corollary

Let n and m are positive integers with $n \geq m$, and let Y be the set of all bitstrings of length $2n$ of which exactly $2m$ bits are 1. Then Y is the disjoint union $V_{2m} \cup V_{2m-1} \cup \cdots \cup V_0$, where

$$\begin{aligned}
V_{2m} &= \{\mathbf{x} \in Y : \text{exactly } 2m \text{ of the last } n \text{ bits are } 1\},\\
V_{2m-1} &= \{\mathbf{x} \in Y : \text{the } (n{+}1)\text{th bit is } 0\text{, and exactly } 2m{-}1 \text{ of the last } n{-}1 \text{ bits are } 1\},\\
V_{2m-2} &= \{\mathbf{x} \in Y : \text{exactly } 2m{-}2 \text{ of the last } n{-}1 \text{ bits are } 1\},\\
V_{2m-3} &= \{\mathbf{x} \in Y : \text{the } (n{+}2)\text{th bit is } 0\text{, and exactly } 2m{-}3 \text{ of the last } n{-}2 \text{ bits are } 1\},\\
V_{2m-4} &= \{\mathbf{x} \in Y : \text{exactly } 2m{-}4 \text{ of the last } n{-}2 \text{ bits are } 1\},\\
&\ \vdots\\
V_{2m-2j+1} &= \{\mathbf{x} \in Y : \text{the } (n{+}j)\text{th bit is } 0\text{, and exactly } 2m{-}2j{+}1 \text{ of the last } n{-}j \text{ bits are } 1\},\\
V_{2m-2j} &= \{\mathbf{x} \in Y : \text{exactly } 2m{-}2j \text{ of the last } n{-}j \text{ bits are } 1\},\\
&\ \vdots\\
V_3 &= \{\mathbf{x} \in Y : \text{the } (n{+}m{-}1)\text{th bit is } 0\text{, and exactly } 3 \text{ of the last } n-m+1 \text{ bits are } 1\},\\
V_2 &= \{\mathbf{x} \in Y : \text{exactly } 2 \text{ of the last } n-m+1 \text{ bits are } 1\},\\
V_1 &= \{\mathbf{x} \in Y : \text{the } (n{+}m)\text{th bit is } 0\text{, and exactly } 1 \text{ of the last } n-m \text{ bits are } 1\},\\
V_0 &= \{\mathbf{x} \in Y : \text{exactly } 0 \text{ of the last } n-m \text{ bits are } 1\}.
\end{aligned}$$

Example

Consider the case where $n = 11$ and $m = 3$. Then the set Y is the disjoint union $V_6 \cup V_5 \cup \cdots \cup V_0$, where

$$\begin{aligned}
V_6 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{-----------} \text{ where exactly 6 of the } \texttt{-} \text{ are 1}\},\\
V_5 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{0----------} \text{ where exactly 5 of the } \texttt{-} \text{ are 1}\},\\
V_4 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{*----------} \text{ where exactly 4 of the } \texttt{-} \text{ are 1}\},\\
V_3 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{*0---------} \text{ where exactly 3 of the } \texttt{-} \text{ are 1}\},\\
V_2 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{**---------} \text{ where exactly 2 of the } \texttt{-} \text{ are 1}\},\\
V_1 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{**0--------} \text{ where exactly 1 of the } \texttt{-} \text{ is 1}\},\\
V_0 &= \{\mathbf{x} \in Y : \text{the last 11 bits are } \texttt{***--------} \text{ where exactly 0 of the } \texttt{-} \text{ are 1}\}.
\end{aligned}$$

Here, $*$ represents a bit that can be either 1 or 0. This partitions Y based on the 'endings' of $\mathbf{x} \in Y$, where 'ending' means the last 11 bits. The total number of possible endings is $\sum_{j=0}^{6} \binom{11}{j} = 1486$, because the number of 1s in the last 11 bits can be anything from 0 to 6. We can count the number of different endings corresponding to each V_j as follows:

$$\begin{aligned}
&V_6 \text{ corresponds to } 2^0\tbinom{11}{6} = 462 \text{ endings}\\
&V_5 \text{ corresponds to } 2^0\tbinom{10}{5} = 252 \text{ endings}\\
&V_4 \text{ corresponds to } 2^1\tbinom{10}{4} = 420 \text{ endings}\\
&V_3 \text{ corresponds to } 2^1\tbinom{9}{3} = 168 \text{ endings}\\
&V_2 \text{ corresponds to } 2^2\tbinom{9}{2} = 144 \text{ endings}\\
&V_1 \text{ corresponds to } 2^2\tbinom{8}{1} = 32 \text{ endings}\\
&V_0 \text{ corresponds to } 2^3\tbinom{8}{0} = 8 \text{ endings}
\end{aligned}$$

The numbers of endings add to 1486, as they should.

We now consider not just the endings of the bitstrings, but the total number of bitstrings in Y and in the V_j for a general n and m. Notice that the condition defining $V_{2m-2j+1}$ is equivalent to

- exactly $2j-1$ of the first $n+j-1$ bits are 1,
- the $(n+j)$th bit is 0, and
- exactly $2m-2j+1$ of the last $n-j$ bits are 1,

and the condition defining V_{2m-2j} is equivalent to

- exactly $2j$ of the first $n+j$ bits are 1, and
- exactly $2m-2j$ of the last $n-j$ bits are 1.

It follows that we have

$$\begin{aligned}
|V_{2m-2j+1}| &= \binom{n+j-1}{2j-1}\binom{n-j}{2m-2j+1},\\
|V_{2m-2j}| &= \binom{n+j}{2j}\binom{n-j}{2m-2j}.
\end{aligned}$$

Since Y is the disjoint union

$$Y = \bigcup_{j=1}^{m} V_{2m-2j+1} \;\cup\; \bigcup_{j=0}^{m} V_{2m-2j},$$

it follows that we have the identity

$$\binom{2n}{2m} = \sum_{j=1}^{m} \binom{n+j-1}{2j-1}\binom{n-j}{2m-2j+1} + \sum_{j=0}^{m} \binom{n+j}{2j}\binom{n-j}{2m-2j}. \tag{2}$$

Now the identity (2) is not the identity (1) from the beginning of this note, but we will show how (1) can be proved using (2). The bulk of the work is already done, and the remainder of our argument uses some well-known facts such as

$$\begin{aligned} \binom{a+1}{b+1} &= \binom{a}{b} + \binom{a}{b+1}, \\ \binom{a}{b}\binom{b}{c} &= \binom{a}{c}\binom{a-c}{b-c}, \qquad \text{and} \\ \binom{a}{b} &= 0 \qquad \text{if } b < 0 \text{ or } b > a. \end{aligned}$$

We begin with the right-hand side of (1) and proceed as follows.

$$\begin{aligned} \sum_{k=0}^{m} \binom{2m+1}{2k+1}\binom{n+k}{2m} &= \sum_{k=0}^{m} \left[\binom{2m}{2k} + \binom{2m}{2k+1}\right]\binom{n+k}{2m} \\ &= \sum_{k=0}^{m} \binom{n+k}{2m}\binom{2m}{2k} + \sum_{k=0}^{m} \binom{n+k}{2m}\binom{2m}{2k+1} \\ &= \sum_{k=0}^{m} \binom{n+k}{2k}\binom{n-k}{2m-2k} + \sum_{k=0}^{m} \binom{n+k}{2k+1}\binom{n-k-1}{2m-2k-1}. \end{aligned}$$

If we let $j = k$ in the first sum and $j = k+1$ in the second sum, this becomes

$$\sum_{j=0}^{m} \binom{n+j}{2j}\binom{n-j}{2m-2j} + \sum_{j=1}^{m+1} \binom{n+j-1}{2j-1}\binom{n-j}{2m-2j+1}$$

and the term with $j = m+1$ in the second sum can be ignored, because it satisfies $2m-2j+1 = -1$. We thus have something identical to the right-hand side of (2), which completes the proof of (1).

References

[1] John Riordan, *Combinatorial identities.* Reprint of the 1968 original. Robert E. Krieger Publishing Co., Huntington, N.Y., 1979.

[2] waterman butterfly, *Combinatorial proofs of the following identities.* https://math.stackexchange.com/q/1900578