When can the sum of $(1/p)$th of the binomial coefficients have closed form?

Marko Petkovšek
University of Ljubljana
Ljubljana, Slovenia

Herbert S. Wilf*
University of Pennsylvania
Philadelphia, PA 19104-6395

Submitted: May 23, 1996; Accepted: November 25, 1996

Abstract

We find all nonnegative integer r, s, p for which the sum $\sum_{k=rn}^{sn} \binom{pn}{k}$ has closed form.

Let

$$f_{p,r}(n) = \sum_{k=0}^{rn} \binom{pn}{k}.$$

where $0 \le r \le p$ are fixed integers. This is a *definite* sum in the sense that the summation limits and the summand are not independent. As we all know,

$$\begin{aligned} f_{r,r}(n) &= 2^{rn}, \\ f_{2r,r}(n) &= \frac{1}{2}\left(4^{rn} + \binom{2rn}{rn}\right). \end{aligned}$$

Thus $f_{r,r}(n)$ is a hypergeometric term, and $f_{2r,r}(n)$ is a linear combination of two hypergeometric terms.

*Supported in part by the Office of Naval Research

Following [PWZ], let us say that a function $f(n)$ has *closed form* if there is a fixed integer m and hypergeometric terms $\{t_i(n)\}_{i=1}^m$ such that $f(n) = \sum_{i=1}^m t_i(n)$ for all sufficiently large n. Our main results are as follows.

Theorem 1 *Let $0 < r < p$ and $p \neq 2r$. Then $f_{p,r}(n)$ does not have closed form.*

Theorem 2 *Let $0 \leq r < s \leq p$ be fixed integers. Then*

$$S_{p,r,s}(n) = \sum_{k=rn}^{sn} \binom{pn}{k}$$

does not have closed form, unless $r = 0, p = 2s$, or $p = s = 2r$, or $r = 0, p = s$.

1 Reduction to an indefinite sum

We begin by briefly discussing the method. One might anticipate that we would first find a recurrence formula that, say, $f_{p,r}(n)$ satisfies, using Zeilberger's algorithm, and then prove, using Petkovšek's theorem, that the recurrence has no closed form solution. As described in [PWZ], this method is quite effective in many cases.

In the present situation, however, the recurrence that $f_{p,r}(n)$ satisfies will grow in complexity with p, r. So for each *fixed* p, r the argument would work, but without further human input it could not produce a *general* proof, i.e., a proof for all p, r. This is somewhat analogous to the sums of the pth powers of all of the binomial coefficients of order n. There too, the methods described in [PWZ] can show that no closed form exists for specific fixed values of p, but the general question remains open for $p \geq 9$ or thereabouts.

Hence we resort to a somewhat different tactic. We will first reduce the definite sum $f_{p,r}(n)$ to an *indefinite* sum, and then we invoke Gosper's algorithm to show that the resulting indefinite sum is not Gosper summable.

Indeed, since $\binom{n}{k} = \sum_j \binom{p}{j}\binom{n-p}{k-j}$ by the Chu-Vandermonde convolution formula, we have

$$\begin{aligned}
f_{p,r}(n+1) &= \sum_{k=0}^{rn+r} \binom{pn+p}{k} = \sum_{k=0}^{rn+r} \sum_j \binom{p}{j}\binom{pn}{k-j} = \sum_j \binom{p}{j} \sum_{l=0}^{rn+r-j} \binom{pn}{l} \\
&= \left(\sum_{j=0}^{r} + \sum_{j=r+1}^{p} \right) \binom{p}{j} \sum_{l=0}^{rn+r-j} \binom{pn}{l} \\
&= \Sigma_I + \Sigma_{II},
\end{aligned}$$

say. Now

$$\begin{aligned}
\Sigma_I &= \sum_{j=0}^{r}\binom{p}{j}\left(\sum_{l=0}^{rn}+\sum_{l=rn+1}^{rn+r-j}\right)\binom{pn}{l} \\
&= f_{p,r}(n)\sum_{j=0}^{r}\binom{p}{j}+\sum_{j=0}^{r-1}\binom{p}{j}\sum_{i=1}^{r-j}\binom{pn}{rn+i}, \\
\Sigma_{II} &= \sum_{j=r+1}^{p}\binom{p}{j}\left(\sum_{l=0}^{rn}-\sum_{l=rn+r-j+1}^{rn}\right)\binom{pn}{l} \\
&= f_{p,r}(n)\sum_{j=r+1}^{p}\binom{p}{j}-\sum_{j=r+1}^{p}\binom{p}{j}\sum_{i=0}^{j-r-1}\binom{pn}{rn-i}.
\end{aligned}$$

Therefore,

$$f_{p,r}(n+1)=2^p f_{p,r}(n)+\sum_{j=0}^{r-1}\binom{p}{j}\sum_{i=1}^{r-j}\binom{pn}{rn+i}-\sum_{j=r+1}^{p}\binom{p}{j}\sum_{i=0}^{j-r-1}\binom{pn}{rn-i}.$$

For each fixed p and r this is a first-order inhomogeneous recurrence with a hypergeometric (and closed form) right hand side. Solving it, we find that $f_{p,r}(n)/2^{pn}$ can be written as an *indefinite* sum,

$$f_{p,r}(n)=2^{pn}\sum_{k=0}^{n}t_k,$$

where

$$t_k=2^{-pk}\left(\sum_{j=0}^{r-1}\binom{p}{j}\sum_{i=1}^{r-j}\binom{pk-p}{rk-r+i}-\sum_{j=r+1}^{p}\binom{p}{j}\sum_{i=0}^{j-r-1}\binom{pk-p}{rk-r-i}\right)$$

is a hypergeometric term for each fixed p and r. Note that this means $f_{p,r}(n)$ satisfies a homogeneous second-order recurrence with polynomial coefficients in n, which could be written down explicitly.

Example. Take $p=3$ and $r=1$. Then we have shown that

$$\begin{aligned}
f_{3,1}(n) &= \sum_{k=0}^{n}\binom{3n}{k}=8^n\sum_{k=0}^{n}8^{-k}\left(\binom{3k-3}{k}-4\binom{3k-3}{k-1}-\binom{3k-3}{k-2}\right) \\
&= 8^n\left(\frac{1}{2}-\sum_{k=2}^{n}\frac{5k^2+k-2}{2^{3k+1}(k-1)(2k-1)}\binom{3k-3}{k}\right) \qquad (n\geq 1)
\end{aligned}$$

2 Application of Gosper's algorithm

In view of the result of the previous section, we now have that $f_{p,r}(n)$ has a closed form if and only if t_k is Gosper-summable. To see if this is the case we "run" Gosper's algorithm on t_k.

In Step 1 of Gosper's algorithm[1] we rewrite t_k as

$$t_k = \frac{\binom{pk}{rk}}{2^{pk}\binom{pk}{p}} P_k, \quad k > 0,$$

where P_k is a polynomial in k,

$$P_k = \sum_{j=0}^{r-1} \binom{p}{j} \sum_{i=1}^{r-j} \frac{\binom{rk}{r-i}\binom{pk-rk}{p-r+i}}{\binom{p}{r-i}} - \sum_{j=r+1}^{p} \binom{p}{j} \sum_{i=0}^{j-r-1} \frac{\binom{rk}{r+i}\binom{pk-rk}{p-r-i}}{\binom{p}{r+i}},$$

and $t_0 = 1$. Then

$$\frac{t_{k+1}}{t_k} = \frac{\binom{p}{r}\binom{pk}{p}}{2^p\binom{r(k+1)}{r}\binom{(p-r)(k+1)}{p-r}} \frac{P_{k+1}}{P_k}, \quad k > 0,$$

is a rational function of k.

In Step 2 we note that the roots r_i of $\binom{pk}{p}$ are $0, 1/p, \ldots, (p-1)/p$ while the roots s_j of $\binom{r(k+1)}{r}\binom{(p-r)(k+1)}{p-r}$ are $-1, -(r-1)/r, \ldots, -1/r$; $-1, -(p-r-1)/(p-r), \ldots, -1/(p-r)$. But $s_j - r_i$ is never a nonnegative integer. Hence

$$\frac{t_{k+1}}{t_k} = \frac{a_k c_{k+1}}{b_k c_k}$$

is a possible Gosper's normal form for t_{k+1}/t_k, where

$$\begin{aligned} a_k &= \binom{p}{r}\binom{pk}{p}, \\ b_k &= 2^p\binom{r(k+1)}{r}\binom{(p-r)(k+1)}{p-r}, \\ c_k &= P_k. \end{aligned}$$

[1] Our description of the steps of Gosper's algorithm follows the exposition of Chapter 5 of [PWZ].

In Step 3 we have to determine the degrees and leading coefficients of a_k, b_k and c_k. Obviously,

$$\deg a_k = \deg b_k = p,$$
$$\mathrm{lc}\ a_k = \binom{p}{r}\frac{p^p}{p!},$$
$$\mathrm{lc}\ b_k = 2^p\frac{r^r}{r!}\frac{(p-r)^{p-r}}{(p-r)!}.$$

When is lc a_k = lc b_k, or equivalently,

$$p^p = 2^p r^r (p-r)^{p-r}? \tag{1}$$

Claim: All integer solutions $0 < r < p$ of equation (1) are of the form

$$p = 2r.$$

To prove the claim, let $p = 2^k q$, $r = 2^m s$, where q, s are odd. Then (1) turns into

$$2^{kp} q^p = 2^{p+mr} s^r (2^k q - 2^m s)^{p-r}. \tag{2}$$

For an integer n and a prime u, let $\varepsilon_u(n)$ denote the largest exponent e such that u^e divides n. Let L and R denote the left and right sides of (2), respectively. So $\varepsilon_2(L) = kp$.

If $k < m$, $\varepsilon_2(R) = kp + p - r(k-m)$, so $p = r(k-m) < 0$, which is false.

If $k = m$, $\varepsilon_2(R) > mp + p$, so $k > m+1$, a contradiction.

If $k > m$, $\varepsilon_2(R) = mp + p$, so $k = m+1$ and (2) turns into

$$q^p = s^r (2q - s)^{p-r}.$$

Let u be an odd prime, $\varepsilon_u(q) = a$, $\varepsilon_u(s) = b$.

If $a < b$, $\varepsilon_u(q^p) = ap$ and $\varepsilon_u(s^r(2q-s)^{p-r}) = br + a(p-r)$, so $a = b$, contradiction.

If $a > b$, $\varepsilon_u(q^p) = ap$ and $\varepsilon_u(s^r(2q-s)^{p-r}) = br + b(p-r) = bp$, so $a = b$, contradiction.

It follows that $a = b$. So q and s have identical prime factorization and are therefore equal. Thus $p = 2^k q = 2^{m+1} s = 2r$, proving the claim.

Since we are assuming that $p \neq 2r$, the leading coefficients of a_k and b_k are different, and we are in Case 1 of Gosper's algorithm.

Obviously $\deg c_k = \deg P_k \leq p$, so any polynomial x_k satisfying Gosper's equation

$$a_k x_{k+1} - b_{k-1} x_k = c_k, \tag{3}$$

must be constant. After a little computation we find that the coefficient of k^p in P_k is

$$\frac{p-r}{(p-2r)p!}(p^p - 2^p r^r (p-r)^{p-r}),$$

which is non-zero. Comparing leading coefficients in Gosper's equation we find that

$$x_k = \frac{p-r}{(p-2r)\binom{p}{r}}.$$

But then one can verify that the coefficient of the first power of k in the polynomial on the left of (3) is $(-1)^{p-1}(p-r)/(p-2r)$, while the corresponding coefficient on the right is $(-1)^{p-1}(p-r)/p$. This discrepancy proves that Gosper's equation has no polynomial solution, and thus $f_{p,r}(n)$ no closed form, when $p \neq 2r$, completing the proof of Theorem 1.

To prove Theorem 2, we see that if $r = 0$ then $S_{p,r,s}(n) = f_{p,s}(n)$, and if $s = p$ then $S_{p,r,s}(n) = 2^{pn} - f_{p,r}(n) + \binom{pn}{rn}$, so in these two cases the assertion follows immediately from Theorem 1.

If $r \neq 0$ and $s \neq p$ then write

$$S_{p,r,s}(n) = f_{p,s}(n) - f_{p,r}(n) + \binom{pn}{rn}.$$

As in the proof of Theorem 1, $f_{p,s}(n) - f_{p,r}(n)$ can be written as the indefinite sum of two hypergeometric terms, one similar to $\binom{pn}{rn}$ and the other to $\binom{pn}{sn}$. Since $r < s$, these two terms are not similar to each other, hence $S_{p,r,s}(n)$ has a closed form if and only if both $f_{p,s}(n)$ and $f_{p,r}(n)$ have it[2]. According to Theorem 1, this is possible only if $p = 2s = 2r$, contradicting the assumption $r < s$. $\square$

3 Discussion

A number of interesting combinatorial sequences have already been proved not to be of closed form. In [PWZ] there are several examples, including the number of involutions

[2]See section 5.6 of [PWZ]

of n letters, the "central trinomial coefficient," and others. The arguments there were made sometimes with Gosper's algorithm, and sometimes with Petkovšek's algorithm, which decides whether a linear recurrence with polynomial coefficients does or does not have closed form solutions.

In the earlier literature there are one or two related results. One elegant and difficult theorem of de Bruijn [Bru] asserts that the sums $\sum_k (-1)^k \binom{2n}{k}^s$ do not have closed form if s is an integer ≥ 3. The idea of his proof was to compare the actual asymptotic behavior of the given sum, for fixed s and $n \to \infty$, with the asymptotic behavior of a hypothetical closed form, and to show that the two could never be the same.

In Cusick [Cus] there is a method that can, in principle, yield the recurrence that is satisfied by the sum $f_p(n) = \sum_k \binom{n}{k}^p$, for fixed p, and a few examples are worked out. Zeilberger's algorithm (see, e.g., [PWZ]) can do the same task very efficiently. Using these recurrences, it has been shown, by Petkovšek's algorithm, that these sums $f_p(n)$ do not have closed form if $p \leq 8$ (but, starting with 6th powers, we have proved this only over fields which are at most quadratic extensions of the rational number field). The general case for these pth power sums remains open, as far as we know. McIntosh [McI] has investigated the order of some related recurrences, as a function of p, and also showed that the Apéry numbers cannot be expressed in a certain form which is a restriction of our notion of closed form. Again, with Petkovšek's algorithm it is quite simple to show that the Apéry numbers are not of closed form, in the wider sense that we use here.

References

[Bru] N. G. de Bruijn, Asymptotic Methods in Analysis, Bibliotheca Mathematica, Vol. 4, North Holland Publishing Co., Amsterdam, 1958 (reprinted by Dover Publications, Inc., 1981).

[Cus] T. W. Cusick, Recurrences for sums of powers of binomial coefficients, *J. Comb. Theory Ser. A* **52** (1989), 77–83.

[McI] Richard J. McIntosh, Recurrences for alternating sums of powers of binomial coefficients, *J. Comb. Theory Ser. A* **63** (1993), 223-233.

[PWZ] Marko Petkovšek, Herbert S. Wilf and Doron Zeilberger, *A = B*, A K Peters, Ltd., Wellesley, MA, 1996.