Maximilian Janisch

Sister Celine's method for binomial sums and generalizations

Student seminar on automatic proofs of binomial identities

Instructors: Prof. Valentin Féray and Raul Penaguiao

Contents

Date: April 23, 2020

1 Introduction

Figure 1: A picture of Sister Celine by Herbert Wilf

Mary Celine Fasenmyer was an American mathematician known mostly for her work in linear algebra and on hypergeometric sums.

In her Ph.D. thesis at the University of Michigan in 1945, she developed a method for systematically finding recurrence relations for (sums of) hypergeometric polynomials. Before her work, finding such a recurrence relation was a question of trial and error involving the use of various "algebraic tricks". Her methods were capable of finding recurrence relations for a specific subset of generalized hypergeometric polynomials including, among others, the Legendre- and Jacobi- polynomials. In her second publication [2], she illustrated her general method by going back to results from her Ph.D. Only in 1978 did Doron Zeilberger fall back on Sister Celine's "method". In my text I will also be following a part of his paper [3]. Her method (or extensions thereof, such as Zeilberger's algorithm which will also be the topic of three seminar talks) is now used by computer algebra systems such as Mathematica and Maple.

2 Problem statement

In general we are looking for a simplification of the sum

$$G(n) = \sum_{k=-\infty}^{\infty} F(n,k), \tag{1}$$

where $F : \mathbb{Z}^2 \to \mathbb{C}$ is multi-hypergeometric, according to the following definition. Sister Celine's Method provides tools described in 3 and formalised in 4 that find a linear recurrence relation satisfied by $G(n)$. The tools are guaranteed to work if the terms $F(n,k)$ are of a special form (namely if they are proper hypergeometric, see section 4). I will very briefly mention how to obtain a closed form from these recurrence relations in section 3.3.

Definition 1. A function $F : \mathbb{Z}^2 \to \mathbb{C}$ is called *multi-hypergeometric* if and only if there exist non-zero polynomials $P_1, Q_1, P_2, Q_2 : \mathbb{Z}^2 \to \mathbb{C}$ such that for all $(n,k) \in \mathbb{Z}^2$, we have

$$\begin{aligned} &P_1(n,k)F(n,k) - Q_1(n,k)F(n-1,k) = 0 \quad \text{and} \\ &P_2(n,k)F(n,k) - Q_2(n,k)F(n,k-1) = 0. \end{aligned} \tag{2}$$

In the following, χ_S will denote the characteristic function of a set S, given by $\chi_S(x) = 1$ if $x \in S$ and $\chi_S(x) = 0$ if $x \notin S$.

Example 1. The function $F(n,k) = 2^k \chi_{[0,n]}(k)$ is multi-hypergeometric. Indeed, we can choose $Q_1(n,k) = P_1(n,k)$ and $Q_2(n,k) = 2P_1(n,k)$ to be any non-zero polynomials which are 0 if $k = n$ and (2) will be satisfied. Another example is $F(n,k) = \chi_{[0,n]}(k) \cdot (c + kq)$ for any constants $c, q \in \mathbb{R}$ (this corresponds to the arithmetic series).

Example 2. If we define the binomial coefficient by $\binom{n}{k} = \chi_{[0,n]}(k) \cdot \frac{n!}{k!(n-k)!}$ (with the convention that $0 \cdot \text{undefined} = 0$), then any product of binomial coefficients is multi-hypergeometric. For example, for any $r > 0$, the function $F(n,k) = \binom{n}{k}^r$ satisfies (2) with $P_1(n,k) = n^r, Q_1(n,k) = (n-k)^r, P_2(n,k) = k^r, Q_2(n,k) = (n-k+1)^r$.

The idea of the algorithm is to first find a linear recurrence relation for the term F, such that the coefficients of this recurrence *do not rely on k*. Then, we sum this recurrence over all integers k in order to get a recurrence for G. In a final step (which is not part of Sister Celine's method), we solve this recurrence in order to obtain a simpler closed form for G.

The following definition is useful:

Definition 2. A function $G : \mathbb{Z} \to \mathbb{C}$ is called *P-recursive* iff there exists a $C \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ and polynomials $P_0, \dots, P_C$ such that P_0 and P_C are not identically 0 and

$$\sum_{i=0}^{C} P_i(n) G(n-i) = 0 \tag{3}$$

for all $n \in \mathbb{Z}$. The number C is called the *order* of the recursion. For example, hypergeometric functions are P-recursive functions of order 1.

Indeed, under certain circumstances described in section 4, the sum $G(n) = \sum_{k=-\infty}^{\infty} F(n,k)$, if well-defined, is P-recursive. *We are interested in finding the polynomials P_i such that* (3) *holds.*

Remark 1. Notice that if F is multi-hypergeometric and $F(n, \cdot)$ has compact support for all $n \in \mathbb{Z}$, i.e. if

$$\operatorname{supp}(F(n,\cdot)) \stackrel{\text{Def.}}{=} \overline{\{k \in \mathbb{Z} : F(n,k) \neq 0\}}$$

is compact (which in this case is equivalent to finite), then $G(n)$ is well-defined. Indeed, this means that we can study all finite sums of the form

$$\sum_{k=0}^{n} F(n,k) = \sum_{k=-\infty}^{\infty} \chi_{[0,n]}(k) \cdot F(n,k). \tag{4}$$

3 General algorithm

I will now try to outline the general algorithm for the simplification of $G(n) = \sum_{k=-\infty}^{\infty} F(n,k)$.

3.1 Finding a recurrence relation for the summands

First, we will be looking for non-negative integers M, N and coefficients $a_{r,s} : \mathbb{Z} \to \mathbb{C}$ for $r, s \in \mathbb{Z}$ such that

$$\sum_{r=0}^{M} \sum_{s=0}^{N} a_{r,s}(n) F(n-r, k-s) = 0 \tag{5}$$

for all $n, k \in \mathbb{Z}$. *It is important to note that the $a_{r,s}$ are independent of k, for reasons seen later.* In section 4, I will give some quantitative estimates for the choice of M, N.

Obtaining such a recurrence for given M, N is fairly straight-forward: By condition (2), we know that each $\frac{F(n-r,k-s)}{F(n,k)}$ is a quotient of two polynomials of n and k (provided that $F(n,k) \neq 0$). Indeed fix any n, k such that $F(n,k) \neq 0$ (we can assume that such n, k exist since the case $F(n,k) \equiv 0$ is not interesting). We can write (5) in the form

$$\sum_{r=0}^{M}\sum_{s=0}^{N} a_{r,s}(n) \underbrace{\frac{F(n-r,k-s)}{F(n,k)}}_{=\text{quotient of two polynomials}} = 0. \tag{6}$$

The left-hand side of this equation (consisting of the sum of the product of the $a_{r,s}$ and rational functions) can be put over a single common denominator. The numerator is now a polynomial in n and k so we can collect powers of k. Setting all coefficients of k to 0, we get a system of linear equations for each n where the unknowns are the $a_{r,s}(n)$.

If the only solution to this linear equation system for some n is $a_{r,s}(n) = 0$ for all r, s, then we have to try again with bigger constants M, N. Otherwise, we have succesfully found a (non-degenerate) recurrence relation for the $F(n,k)$ in the form of (5).

Example 3. Consider again the term $F(n,k) = \binom{n}{k}$. Then for $0 \leq k \leq n$ (for other k, $F(n,k) = 0$), we have $\frac{F(n-1,k)}{F(n,k)} = \frac{n-k}{n}$ and $\frac{F(n,k-1)}{F(n,k)} = \frac{k}{1-k+n}$. By performing the above steps for $M = N = 1$, we obtain the recurrence (note that this recurrence is well-known as *Pascal's rule*)

$$F(n,k) - F(n-1,k) - F(n-1,k-1) = 0. \tag{7}$$

3.2 Finding a recurrence relation for the sum

After we found a recurrence relation in the form of (5), we can quickly find a recurrence relation for the sum $G(n)$ with the following useful Lemma:

Lemma 1. *If F is multi-hypergeometric such that $G(n) = \sum_{k=-\infty}^{\infty} F(n,k)$ is well-defined for all n and if F satisfies a recurrence of the form (5), then $G(n)$ satisfies the linear recurrence relation*

$$\sum_{r=0}^{M} G(n-r) \cdot \left(\sum_{s=0}^{N} a_{r,s}(n)\right) = 0 \quad \textit{for all } n \in \mathbb{Z}. \tag{8}$$

Proof. The Lemma can be proven by switching the summation order after summing (5) over k (note that this is possible since I am "extracting" only finite sums which means that I am just using the linearity of the limit,´ i.e. $\lim_{N\to\infty} a_N + b_N = \lim_{N\to\infty} a_N + \lim_{N\to\infty} b_N$.)

$$\begin{aligned} 0 = \sum_{k=-\infty}^{\infty} \overbrace{\sum_{r=0}^{M}\sum_{s=0}^{N} a_{r,s}(n) F(n-r,k-s)}^{0} &= \sum_{r=0}^{M}\sum_{s=0}^{N} a_{r,s}(n) \overbrace{\sum_{k=-\infty}^{\infty} F(n-r,k-s)}^{G(n-r)} \\ &= \sum_{r=0}^{M}\sum_{s=0}^{N} a_{r,s}(n) G(n-r) \\ &= \sum_{r=0}^{M} G(n-r) \cdot \left(\sum_{s=0}^{N} a_{r,s}(n)\right). \end{aligned}$$

□

Example 4. Coming back to 3, let us consider the sum

$$G(n) = \sum_{k=-\infty}^{\infty} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k}. \tag{9}$$

We already know that the summand satisfies the recurrence

$$F(n,k) - F(n-1,k) - F(n-1,k-1) = 0.$$

From Lemma 1, we thus obtain that $G(n) = 2G(n-1)$ and hence from $G(0) = 1$ that $G(n) = 2^n$. (Note that this result can be achieved immediately with the smart observation that $G(n) = (1+1)^n$ by the Binomial Theorem.)

Example 5. Consider the sum with summands $F(n,k) = \binom{n}{k}^2$. Applying the procedure from 3.1 for $M = N = 2$, we get the recurrence

$$\begin{aligned} &a_{0,0}(n)F(n,k) + a_{1,0}(n)F(n-1,k) + a_{2,0}(n)F(n-2,k) \\ &\quad + a_{1,1}(n)F(n-1,k-1) + a_{2,1}(n)F(n-2,k-1) + a_{2,2}(n)F(n-2,k-2) = 0, \end{aligned} \tag{10}$$

where $a_{0,0}(n) = n, a_{1,0}(n) = a_{1,1}(n) = 1-2n, a_{2,0}(n) = a_{2,2}(n) = n-1$ and $a_{2,1}(n) = 2-2n$. Using Lemma 1, we obtain that $G(n) = \sum_{k=0}^{n} \binom{n}{k}^2$ satisfies

$$nG(n) + (2-4n)G(n-1) = 0, \quad \text{i.e. for } n \neq 0, \quad G(n) = \frac{4n-2}{n}G(n-1). \tag{11}$$

This, together with $G(0) = 0$, implies that $G(n) = \binom{2n}{n}$.

3.3 Obtaining a closed form for the sum

Suppose that we found, for example using the methods from the previous two sections, a recurrence of the form

$$\sum_{r=0}^{M} P(r,n)G(n-r) = 0. \tag{12}$$

We are now interested in finding a *closed form* for $G(n)$, i.e. an expression of the type $G(n) = \binom{2n}{n}$ (I will not give a precise definition of closed form here.)

There is an algorihmic approach called *algorithm hyper*, which will also be treated within this seminar.

4 The fundamental Theorem

In this section, I will try to give some quantitative estimates on when to expect the approach of the previous section to work, and what M, N need to be chosen. This section is following the fourth chapter of the great book *A=B* [4].

Definition 3. A function $F : \mathbb{Z}^2 \to \mathbb{C}$ is called *proper hypergeometric* iff there exist constants $U, V \in \mathbb{N}$, $a_i, b_i, c_i, u_j, v_j, w_j \in \mathbb{Z}$ for $i = 1, \ldots, U$ and $j = 1, \ldots, V$, $x \in \mathbb{C}$ and finally a polynomial $P : \mathbb{Z}^2 \to \mathbb{C}$ such that

$$F(n,k) = P(n,k) \cdot \frac{\prod_{i=1}^{U}(a_i n + b_i k + c_i)!}{\prod_{i=1}^{V}(u_i n + v_i k + w_i)!} x^k \tag{13}$$

for all $n, k \in \mathbb{Z}$.

Remark 2. It should be noted that (13) is well-defined only if $a_i n + b_i k + c_i \geq 0$ for all i. Also, the convention used here is that $F(n,k) = 0$ if $u_i n + v_i k + w_i < 0$ for any i.

Remark 3. The ideas described here arose initially in the original paper by Sister Celine [2] where she was studying quote "many of the classical polynomials, for example, the Laguerre, and [...] those of Jacobi, Legendre".

Example 6. The function

$$F(n,k) = \binom{n}{k} 2^k = \frac{n! 2^k}{k!(n-k)!} \tag{14}$$

is proper hypergeometric. So is the function

$$F(n,k) = \frac{1}{d_1 n + d_2 k + d_3 + 1} = \frac{(d_1 n + d_2 k + d_3)!}{(d_1 n + d_2 k + d_3 + 1)!} \tag{15}$$

for any constants $d_1, d_2, d_3 \in \mathbb{N}$.

Theorem 1. *Let F be a proper hypergeometric function written in the form* (13). *Then there exist $M, N \in \mathbb{N}_0$ and polynomials $a_{r,s} : \mathbb{Z} \to \mathbb{C}$ for $r = 0, \ldots, M$ and $s = 0, \ldots, N$ which are not all identical to* 0, *such that the recurrence* (5) *holds.*

Remark 4. Furthermore, it can be shown (I will omit the proof of this fact here) that such a recurrence holds for the particular choice

$$M = \sum_{s=1}^{U} |b_s| + \sum_{s=1}^{V} |v_s| \quad \text{and} \quad N = 1 + \deg(P) + M \cdot \left(-1 + \sum_{s=1}^{U} |a_s| + \sum_{s=1}^{V} |u_s| \right). \tag{16}$$

Proof. For this proof I will use the short-hand notation

$$\mathrm{rf}(x,y) = \prod_{j=1}^{x} (y+j), \text{ and } \mathrm{ff}(x,y) = \prod_{j=0}^{x-1} (y-j). \tag{17}$$

For illustrative purposes, let me first consider the case of a "simple" function $f(n,k) = (an + bk + c)!$. For this function,

$$\frac{f(n-j, k-i)}{f(n,k)} = \begin{cases} (\mathrm{ff}(aj + bi, an + bk + c))^{-1}, & \text{if } aj + bi \geq 0, \\ \mathrm{rf}(|aj + bi|, an + bk + c), & \text{if } aj + bi < 0. \end{cases} \tag{18}$$

Now I will get to the general case: For F written in the form (13), we have (assuming $F(n,k) \neq 0$) $\frac{F(n-j,k-i)}{F(n,k)} = \frac{\nu(n,k)}{\delta(n,k)}$, where

$$\nu_{r,s}(n,k) = P(n-r, k-s) \prod_{\substack{1 \leq i \leq U \\ a_i r + b_i s < 0}} \mathrm{rf}(|a_i r + b_i s|, a_i n + b_i k + c_i) \prod_{\substack{1 \leq i \leq V \\ u_i r + b_i s \geq 0}} \mathrm{ff}(u_i r + v_i s, u_i n + v_i k + w_i), \tag{19}$$

and

$$\delta_{r,s}(n,k) = P(n,k) x^s \prod_{\substack{1 \leq i \leq U \\ a_i r + b_i s \geq 0}} \mathrm{ff}(a_i r + b_i s, a_i n + b_i k + c_i) \prod_{\substack{1 \leq i \leq V \\ u_i r + v_i s < 0}} \mathrm{rf}(|u_i r + v_i s|, u_i n + v_i k + w_i). \tag{20}$$

If we try to apply the procedure in section 3.1, we get the equation

$$\sum_{r=0}^{M} \sum_{s=0}^{N} a_{r,s}(n) \frac{\nu_{s,r}(n,k)}{\delta_{s,r}(n,k)} = 0. \tag{21}$$

The next step is to put all terms in the sum (21) over a (least) common denominator. Looking at the explicit form of $\delta_{r,s}(n,k)$, we see that $P(n,k)$ must be in this common denominator. Let me introduce another shorthand notation: For $x \in \mathbb{R}$, let $x^+ = \max(x,0)$. Then we have for integers a, b,

$$\max\{|ar+bs| : ar+bs < 0, 0 \le s \le N, 0 \le r \le M\} = (-a)^+M + (-b)^+N \tag{22}$$

and

$$\max\{ar+bs : ar+bs \ge 0, 0 \le s \le N, 0 \le r \le M\} = a^+M + b^+N, \tag{23}$$

so that the notation gets much clearer.

We still want to find the least common multiple of all the $\delta_{r,s}$. It is convenient that the rising factorials rf and the falling factorials ff share many factors. Indeed, for each i in the product of (20), a common multiple of all the falling factorials is the falling factorial with the largest argument, that is

$$\mathrm{ff}((a_i)^+M + (b_i)^+N, a_in + b_ik + c_i). \tag{24}$$

Analogously, a common multiple of all the rising factorials is

$$\mathrm{rf}((-u_i)^+M + (-v_i)^+N, u_in + v_ik + w_i). \tag{25}$$

It follows that if we define the polynomial

$$\Delta(n,k) = P(n,k) \cdot \prod_{i=1}^{U} \mathrm{ff}((a_i)^+M+(b_i)^+N, a_in+b_ik+c_i) \prod_{i=1}^{V} \mathrm{rf}((-u_i)^+M+(-v_i)^+N, u_in+v_ik+w_i), \tag{26}$$

then by rewriting (21) as

$$\sum_{r=0}^{M}\sum_{s=0}^{N} a_{r,s}(n)\, \underbrace{\nu_{s,r}(n,k)\frac{\Delta}{\delta_{s,r}(n,k)}}_{\text{polynomial in } n \text{ and } k} = 0, \tag{27}$$

I get a polynomial equation in the $a_{r,s}(n)$. All that is left to do is to collect powers of k and set all coefficients of the powers of k to 0. This amounts in a linear equation system.

It is thus enough to show that if M and N are large enough, then there will be more unknowns $a_{r,s}(n)$ than equations. By construction, there are $(M+1)(N+1)$ unknowns. The number of equations is equal to the number of (distinct) powers of k that appear in the sum (27). So I will now study the number of powers of k appearing in (27).

Indeed, the degree in k of each rising and falling factorial in (27) grows linearly with M and N and hence the degrees in k of the ν's, δ's and Δ also grow linearly with M, N. So there is a linear function $L : \mathbb{N}^2 \to \mathbb{R}$ and a constant $C \ge 0$ such that

$$\text{number of powers of } k \text{ in (27)} \le L(M,N) + C. \tag{28}$$

Since for M, N large enough we have $(M+1)(N+1) > MN \gg L(M,N) + C$, the claim follows. □

5 Some examples

5.1 Example 1

As another example, I want to consider the sum $G(n) = \sum_{k=0}^{n} k^3 \binom{n}{k}$.

Remark 5. By repeatedly using the identity $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ and performing algebraic manipulations, one can also obtain the result proven below with Sister Celine's method. See [5].

Using the attached Mathematica script, more specifically the function `findRecur`, we find that the $F(n,k) = k^3\binom{n}{k}$ satisfy the recurrence (5) with $M = N = 2$ and the coefficients

$$(a_{r,s}(n))_{r,s\in\{0,1,2\}} = \begin{pmatrix} \frac{\frac{\left(-5n^2+11n-6\right)a_{1,1}(n)}{n} - \frac{\left(10n^3-33n^2+29n-2\right)a_{0,1}(n)}{(n-1)^2}}{4(n+2)} & a_{0,1}(n) & 0 \\ \frac{\frac{n\left(10n^3-37n^2+37n-4\right)a_{0,1}(n)}{(n-1)^3}+(5n-8)a_{1,1}(n)}{2(n+2)} & a_{1,1}(n) & -\frac{n^3a_{0,1}(n)}{(n-1)^3} \\ -\frac{n\left(10n^2-21n+3\right)a_{0,1}(n)+5(n-1)^3a_{1,1}(n)}{4(n-1)^2(n+2)} & -\frac{n\left(n^2-6n+1\right)a_{0,1}(n)+(n-1)^3a_{1,1}(n)}{(n-1)^2(n+2)} & \frac{\frac{n\left(6n^2+3n-1\right)a_{0,1}(n)}{(n-1)^2}+(n-1)a_{1,1}(n)}{4(n+2)} \end{pmatrix} \tag{29}$$

Notice that, since we have more unknowns than linear equations, we have the functions $a_{0,1}(n)$ and $a_{1,1}(n)$ that are freely choosable. For example, I will choose (see also the attached Mathematica script) $a_{0,1}(n) = (n-1)^2$ and $a_{1,1}(n) = n$. This arbitrary choice was made in order to cancel some of the denominators. From Lemma 1, which was implemented as `findRecurForSum` in the Mathematica script, we get the linear recurrence relation

$$4(n-5)n^2G(n-2) + 2(13-4n)n^2G(n-1) + \left(3n^3 - 14n^2 + 15n - 2\right)G(n) = 0. \tag{30}$$

Using Mathematica's `RSolve` function, which seems to be an implementation of algorithm Hyper, we get that the general solution to this recurrence is

$$G(n) = C_1 2^{n-3}n^2(3+n) + C_2 \cdot \text{something}(n), \tag{31}$$

where C_1 and C_2 are arbitrary constants.The solution we are looking for is with $C_1 = 1$ and $C_2 = 0$ so that

$$\boxed{G(n) = \sum_{k=0}^{n} k^3\binom{n}{k} = 2^{n-3}n^2(3+n).} \tag{32}$$

Remark 6. Since the choice $C_1 = 1$ and $C_2 = 0$ is independent of Sister Celine's method, I will not further explain it here (it should be noted that the term something(n) by Mathematica is quite "ugly".)

5.2 Example 2

As a final example, I will consider the sum $G(n) = \sum_{k=0}^{n}\binom{n}{k}\binom{2k}{k}(-2)^{n-k}$. Using the `findRecur` function (see again the attached program), we find that the $F(n,k) = \binom{n}{k}\binom{2k}{k}(-2)^{n-k}$ satisfy the recurrence (5) with $M = 2, N = 1$ and the coefficients

$$(a_{r,s}(n))_{r,s} = \begin{pmatrix} a_{0,0}(n) & 0 \\ \frac{2(2n-1)a_{0,0}(n)}{n} & -\frac{2(2n-1)a_{0,0}(n)}{n} \\ \frac{4(n-1)a_{0,0}(n)}{n} & -\frac{8(n-1)a_{0,0}(n)}{n} \end{pmatrix}, \tag{33}$$

where $a_{0,0}(n)$ is again freely choosable. I will choose $a_{0,0}(n) = n$, since this cancels all the denominators in the other coefficients, and by Lemma 1, I obtain the recurrence

$$nG(n) = 4(n-1)G(n-2). \tag{34}$$

Indeed it follows that (the following formula is expressed a bit more clumsily by Mathematica)

$$\boxed{G(n) = \begin{cases} 0, & \text{if } n \text{ is odd} \\ \binom{n}{\frac{n}{2}}, & \text{if } n \text{ is even} \end{cases}.} \tag{35}$$

References

[1] Sister Marie Celine Fasenmyer, *Some Generalized Hypergeometric Polynomials.* Ph.D. thesis, University of Michigan, November 1945.

[2] Sister Marie Celine Fasenmyer, *A Note on Pure Recurrence Relations*, American Mathematical Monthly **56**, pages 14-17 (1949).-

[3] Doron Zeilberger, *Sister Celine's Technique and Its Generalizations*, Journal of Mathematical Analysis and Applications **85**, pages 114-145 (1982).

[4] Marko Petkovesk, Herbert Wilf, Doron Zeilberger, *A=B*, Publisher: A K Peters, Ltd 1996. This book has its own homepage: `https://www.math.upenn.edu/~wilf/AeqB.html`.

[5] Markus Scheuer, *If* $m = \sum_{k=1}^{2017} k^3\binom{2017}{k}$*, find* m*.* URL (version: 17. 02. 2020): `https://math.stackexchange.com/q/3550534`.