ELSEVIER

Discrete Mathematics 136 (1994) 309–346

DISCRETE MATHEMATICS

Binomial identities — combinatorial and algorithmic aspects

Volker Strehl*

Computer Science Department (Informatik), University of Erlangen-Nürnberg, Germany

Received 8 February 1993; revised 23 July 1993

Abstract

The problem of proving a particular binomial identity is taken as an opportunity to discuss various aspects of this field and to discuss various proof techniques in an exemplary way. In particular, the unifying role of the hypergeometric nature of binomial identities is underlined. This aspect is also basic for combinatorial models and techniques, developed during the last decade, and for the recent algorithmic proof procedures.

"Much of mathematics comes from looking at very simple examples from a more general perspective. Hypergeometric functions are a good example of this."

R. Askey

1. Introduction

In this article I want to highlight some aspects of 'binomial identities' or 'combinatorial sums' in an exemplary way. Writing such an article was motivated by a question that I was asked in spring 1992, and by my subsequent investigations on it:

Can you show that the binomial identity

$$\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2=\sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\sum_{j=0}^{k}\binom{k}{j}^3 \tag{1}$$

holds for all nonnegative integers n?

I had not seen this identity before, and was attracted by the amazing way in which it relates the famous *Apéry numbers* $a_n=\sum_k\binom{n}{k}^2\binom{n+k}{k}^2$ with the sums of cubes of the binomial coefficients $f_n=\sum_k\binom{n}{k}^3$. Recall that Apéry's original proof of the irrationality of $\zeta(3)$ made crucial use of the fact that the numbers $(a_n)_{n\geqslant 0}$ satisfy

*E-mail: strehl@informatik.uni-erlangen.de

SSDI 0012-365X(94)00118-3

a linear second-order recurrence with polynomial coefficients:

$$(n+1)^3 a_{n+1} - ((n+1)^3 + n^3 + 4(2n+1)^3) a_n + n^3 a_{n-1} = 0 \quad (n \geqslant 0). \tag{2}$$

I refer the reader to van der Poorten's [55] entertaining and highly instructive presentation of Apéry's proofs and the history(ies) around it. In particular, note that even for a clever mathematician it may be a hard task to *verify* (2), as can be seen from the following quotation [55, p. 200]:

> "...To convince ourselves of the validity of Apéry's proof we need only complete the following exercise:
> Let $a_n = \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2$, then $a_0 = 1$, $a_1 = 5$ and the sequence $\{a_n\}$ satisfies the recurrence (2). (3)
> ...Neither Cohen nor I had been able to prove this in the intervening two months..."

So that the task of *finding*[1] such a recurrence may be considered even harder (but see Section 2.3 below).

As to the numbers $f_n = \sum_k \binom{n}{k}^3$, recall that Franel has found a second-order recurrence long ago [26, 40, 10]:

$$(n+1)^2 f_{n+1} - (7n^2 + 7n + 2) f_n - 8n^2 f_{n-1} = 0 \quad (n \geqslant 0). \tag{4}$$

For this reason, I will call these numbers *Franel numbers* in this article.

Only after while I learned that problem (1) originated from a question that the number-theorist Schmidt from Copenhagen had asked:

"Define rational numbers $(c_k)_{k \geqslant 0}$, independent of n, by

$$\sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} c_k \quad (n \geqslant 0). \tag{5}$$

Is it then true that all these numbers c_k are integers?"

Note that the sequence $(c_k)_{k \geqslant 0}$ is uniquely determined by (5), but all one can see directly from their definition is the fact that the numbers $\binom{2k}{k} c_k (k \geqslant 0)$ are integers. Question (5) had been investigated by Deuber, Thumser, and Voigt from Bielefeld, who extracted (1) as a conjecture from numerical calculations. Thus (1) is much more than an arbitrary binomial identity, it is a statement of number-theoretical interest, claiming not only that the inverse of the sequence of Apéry numbers under *Legendre transform* (in the sense of Schmidt, see [45–47] is an integer sequence, but that this is precisely the sequence of Franel numbers. (The naming is due to the fact that the

[1] For *cleverly guessing* a recurrence on the basis of the first few values of the sequence and similar tasks, there is a package called gfun of procedures available in the Maple share library, accessible via anonymous ftp at daisy.uwaterloo.ca or neptune.inf.ethz.ch. It is due to Salvy and Zimmerman [44] and contains a nice set of tools for the combinatorialist working with sequences of numbers and their recurrences, the corresponding generating functions and the differential equations they satisfy.

numbers $\binom{n}{k}\binom{n+k}{k}$ are the coefficients of the Legendre polynomials, see Section 3.1 below.) The fact that both sequences, satisfying linear second-order recurrences of the same type, are related via this transform is an interesting result by itself, and this aspect is pursued in more detail in Sections 3.6, 4.3, and [52].

The present article gives a report on some of my investigations concerning identity (1) and the original question (5). Indeed, in Section 3 I offer six different proofs of (1), thus illustrating the many facets that such a simple (?) binomial identity can have. Even if I think that the beauty of (1) merits an investigation per se, I will take it here as an exemplary case. Each of the proofs and proof techniques of Section 3 bears its own potential for variation and generalization — some of the possibilities are outlined in Section 4.

The reader will notice that three aspects are of main interest: the *hypergeometric nature* of binomial identities, the *combinatorial models* involved, and the recent availability of *algorithmic tools* for verifying automatically such identities. In the next section I will comment on each of these, first in a completely informal way, then by discussing specific examples. In particular, the very classic Pfaff–Saalsschütz identity, which is used in one of the proofs of (1), will be considered under these various aspects. It is clear that in this article I cannot give a detailed treatment of any of these aspects in general, as they deserve. But I hope that the reader will get an idea of what is going on, and there are enough references given for further study.

2. Binomial identities

2.1. Generalities

Binomial summations, or 'combinatorial sums', their evaluations and identities involving them, 'binomial identities', for short, occur in many parts of mathematics, e.g. combinatorics, probability, number theory, analysis of algorithms, etc. Conceptually they are of a very simple nature, yet, if they occur 'in practice' they can pose considerable technical problems — quotation (3) is but one witness of that. The ubiquity of binomial coefficients has resulted in a vast body of literature devoted to those identities. Apart from the many ad-hoc approaches and straightforward, but often tedious, direct manipulations there are established techniques for dealing with them: generating functions, inverse relations, integral representations, etc., see [42, 36, 11, 33] as 'classics' in the field, and [31] for a comprehensive collection.

One particular feature of binomial identities is the fact that they are very often rediscovered in equivalent form. The reason for this is nicely described by Greene and Knuth [34, p. 6]:

> "One particularly confusing aspect of binomial identities is the ease with which a familiar formula can be rendered unrecognizable by a few transformations.

Because of their chameleon character there is no substitute for practice of manipulations with binomial coefficients ... "

and Riordan pessimistically states in [42, p. vii] that

"... identities are both inexhaustible and unpredictale; the age-old dream of putting order in this chaos is doomed to failure."

It is also instructive to re-read the introduction of [31], in particular in the light of the recent development of algorithmic methods, as mentioned below.

To give an informal definition, a *binomial sum* is an expression $\sum_k F(n,k)$, where $F(n,k)$ is a product of (positive and negative) powers of binomial coefficients, with numerator and denominator terms of these binomials involving n and k, and possibly other parameters.[2] Occasionally $F(n,k)$ will also contain a (simple) rational function of n and k and an argument z^k as factors. The summation usually runs over a 'natural' domain, i.e. a finite interval of the integers, depending on the parameter n, such that $F(n,k)$ vanishes outside this interval. This means that most often the range of summation need not be specified explicitly. Usually one considers $\sum_k F(n,k)$ as a function of n, where n is supposed to vary over the nonnegative integers.

Given such an $F(n,k)$, one may ask

• whether $f(n)=\sum_k F(n,k)$ has an evaluation in 'closed form', which roughly means that it can be written as a product of powers of binomials involving only n (and possibly the other parameters), together with a rational (in n) function as factor; or

• (if a closed form evaluation does not exist), what else can be said about $f(n)$ as a function of n — exactly or asymptotically;

• whether there are similar $G(n,k)$, with specified properties perhaps, such that $\sum_k F(n,k)=\sum_k G(n,k)$ for all $n\geqslant 0$.

Or, similar to the problem (1), we may have $F(n,k)$ and $G(n,k)$ given, and ask

• whether $\sum_k F(n,k)=\sum_k G(n,k)$ $(n\geqslant 0)$, and how this can be proved if equality has been checked for some initial values, even if we know that both sides are not evaluable in closed form.

Similar questions can be asked, for multiple sums, of course.

2.1.1. Hypergeometric aspects

It is a fundamental observation that binomial sums can, as a rule, be written as terminating hypergeometric series. Hence, 'closed form' evaluations of binomial sums will most likely correspond to the classical evaluations known for hypergeometric series, and identities $\sum_k F(n,k)=\sum_k G(n,k)$ involving binomials will most likely correspond to terminating cases of the classical transformation formulas known for these series — see e.g. [3, 17, 41, 33].

---

[2] It should be remarked that in the cases of interest for us these terms are always of the form $a\cdot n+b\cdot k+c$, where a and b are fixed *integers*, and c is a (complex) parameter — this restriction leads to the concept of a *proper hypergeometric term* mentioned in Section 2.2.3 below.

Although one can occasionally find references to the hypergeometric literature in older papers, the combinatorial literature in general does barely reflect this systematical point of view, which was perhaps first clearly spelled out by Andrews in [1], and vividly propagated by Askey in his talks. For example, hypergeometrics is almost nonexistent in [42, 31, etc.]. A nice, pedagogical article on the hypergeometric aspect of binomial identities was published by Roy in [43], and it is interesting to observe the transition from Knuth's [36] (where hypergeometrics is nonexistent) over Greene–Knuth's [34] (where it plays a marginal role) to the recent book [33] by Graham–Knuth–Patashnik (where it now occupies a prominent position).

Apart from the fact that many results from the literature on special functions contain binomial identities as special cases, the main virtue of the hypergeometric approach lies in its *normalizing character*—rewriting binomial sums as hypergeometric series establishes indeed a kind of *normal form* — up to standard transformations. Many results from the literature, proved by various ad-hoc methods, turn out to be equivalent when looked at from the hypergeometric point of view, even if they look very dissimilar at the surface — this is especially true for the many binomial incarnations of the Chu–Vandermonde identity, which all turn out to be consequences of Gauss evaluation of a ${}_2F_1$-series with unit argument in the terminating case.[3] But this applies also to the cryptomorphic versions of the Pfaff–Saalschütz identity, which for the hypergeometer is just the evaluation of a terminating 1-*balanced* ${}_3F_2$-series with unit argument — a case which will be used for illustrative purposes in Section 2.2.

There are many more instances of equivalent evaluations and binomial identities reflecting standard hypergeometric transforms, and a hypergeometer might be tempted to recompile and rewrite Gould's otherwise admirable and useful collection under this more systematic point of view. As to the classical, Riordan-style inversion techniques and their relation to hypergeometrics, I would like to draw attention to the forthcoming work of Chu [8].

2.1.2. *Combinatorial models*

Economic and efficient as the hypergeometric approach is, there is surely more to binomial identities than just scholarly application of these 'formal' techniques and tools. In many cases binomial sums arise in a 'natural' situation of counting, and binomial sums should reflect, on a numerical level, two ways of counting the same set of objects, possibly with the help of clever encodings of objects and bijections. So, even if a binomial identity is 'known' by being a special case of some hypergeometric identity or transformation, it is legitimate to ask for combinatorial proofs and interpretations. The result might not only illuminate the situation, but also lead to variations and generalizations (e.g. q-analogs) which are not easily 'seen' from hypergeometric work.

[3] For example, in [36] there are five versions, Eqs. (21)–(26) in Section 1.2.6, of Chu–Vandermonde without comment, in [33] (almost) the same table is reproduced with explicit reference to Chu and Vandermonde's convolution.

The usual combinatorial interpretation of binomial coefficients by counting subsets, or equivalently lattice path in a rectangular grid, or words over an alphabet with specified frequencies of letters, is at the basis of many of the interpretations from binomial sums and identities, especially the easier ones. To illustrate the benefit of combinatorics at this point, the more sophisticated model of the *free partially commutative monoid*, introduced by Cartier and Foata in [7] in order to provide a combinatorial approach to MacMahon's 'master theorem', hence to many of the known and not so known binomial identities, has not only led to new combinatorial results, but also helped to create a concept which was taken up much later by computer scientists in order to study models of concurrency (under the name of *traces* and *trace languages*). In combinatorics, the same concept is used in a very suggestive, pictorial way by Viennot [58] and his school under the naming *empilements* or *heaps of pieces* — for counting purposes which are, at times, very far from binomial identities — see [58] for the basics.

From recent studies of combinatorial interpretations and properties of classical (i.e. hypergeometrical) orthogonal polynomials, it has become clear that an interpretation a bit more general than the 'number-of-subsets' model is more flexible: counting of *injective functions*, where additional parameters can often be interpreted by cycle counting, see Section 2.2.2 below. Besides flexibility, there is another advantage: the combinatorial models thus obtained can be integrated into a general class of models, namely (multisorted) functions with a specified *local* structure. Confronting the local and the global description of such structures, and counting them accordingly in two different ways, leads to insight into and extension of many of the identities and generating functions existing as 'formal' results in the literature on special functions — beyond the class of identities considered here. As a (partial) explanation, a combinatorial view of multivariate Lagrange inversion (in the style of Gessel [27]) is the basic, unifying concept, but enriched with the feature of cycle counting. See Section 2.2.2 for a brief illustration, [51] for a compact survey, and [50] for a comprehensive treatment.

2.1.3. Automatic verification

Recent theoretical and algorithmical progress in hypergeometric summation leads to a totally new aspect of the field of binomial identities. Recall that there is an algorithm, a decision procedure, due to Gosper Jr. [30, 33, Section 5.7], which — in a sense — settles the problem of *indefinite* hypergeometric summation. Its *direct* application to *definite* hypergeometric sums, including binomial sums, has limited success, however. Most definite sums which have evaluations in 'closed form' do not correspond to indefinitely summable hypergeometric terms. So it comes as a surprise that Gosper's method can be used systematically and efficiently for problems of definite hypergeometric summation in a less direct way.

It follows from the theory of holonomic systems, as put into action for our purposes by Zeilberger in the fundamental article [63], that (under certain restrictions, which are satisfied for the binomial sums we consider) a hypergeometric summation

$f(n) = \sum_k F(n,k)$ is *P-recursive* or *holonomic* as a function of n, i.e. it satisfies a linear recurrence with polynomial coefficients — (2) and (4) are typical examples with recurrences of second order. If the recurrence is first-order, then an evaluation in 'closed form' is possible.

Fortunately, in the case of interest to us, i.e. when the terms $F(n,k)$ are *proper-hypergeometric* in the sense of [60, 59], as mentioned above, then a *constructive* proof of the existence of such a recurrence for $f(n) = \sum_k F(n,k)$ i.e. a difference operator with polynomial coefficients in n, annhilating the sum as a function of n, can be given. It provides bounds on the order of the recurrence and the degrees of the polynomial coefficients of $F(n,k)$, so that the recurrence can be obtained by the method of undetermined coefficients. Such a proof is possible even for multiple binomial sums, as elaborated by Wilf and Zeilberger, see [60] for a short presentation of the concepts and the result, [59] for a comprehensive treatment, and we shall make use of the corresponding algorithm in Section 3.6 below. In the case of a single summation such a proof has (apparently) first been given by Verbaeten in this work on 'Sister Celine's technique', which unfortunately seems to have been overlooked until recently ([56], [57], see [35] for a revival).

Note the way in which holonomicity is crucial for an *algorithmic* approach: the sequence $(f(n))_{n \geqslant 0}$ of values of a binomial sum is *specified* by the (minimal) polynomial recursion it satisfies together with an appropriate number of initial values. Thus closed form evaluations $\sum_k F(n,k) = g(n)$ $(n \geqslant 0)$ and binomial identities of the type $\sum_k F(n,k) = \sum_k G(n,k)$ $(n \geqslant 0)$ can be *proved* by determining the (minimal) polynomial recurrences (or annihilating difference operators) which both sides satisfy — which must be of first order in the former case. An essential aspect of this approach is the existence of *certificates*, i.e. a rational function in n, k, associated to $F(n,k)$, which permits an *easy* (possibly tedious, if done by hand) *verification* by rational arithmetic of the fact that $\sum_k F(n,k)$ indeed satisfies the recurrence found — remember the quotation (3), this is not trivial! What this means and how it works will be indicated in Section 2.2.3.

To come back to the beginning of this section, it was a brilliant observation by Zeilberger that in the case of single hypergeometric sums Gosper's method can be employed in a way which he calls *creative telescoping*, a term which goes back to van der Poorten's article [55, pp. 200, 201], where he mentions that Zagier confirmed (2) *with irritating speed* by virtue of *the method of creative telescoping*. Holonomicity and Gosper's algorithm together give a quite efficient proof and verification method for single hypergeometric/binomial sums, see e.g. [62, 65]. Various implementations,[4] also for the q-analogue, are available.

---

[4] Zeilberger provides *Maple* implementations for single sums, see [64], as well as for multiple sums (and integrals) with the general method described in [59], call zeilberg@euclid.math.temple.edu; Koornwinder has a *Maple* implementation of the ordinary and q-case for single summations, carefully described in [37], call thk@fwi.uva.nl; Hornegger has a very flexible implementation for single sums for the *Axiom*-system, together with related algorithms for hypergeometric sums, all documented in [35], call hornegge@informatik.uni-erlangen.de; Paule provides a *Mathematica* implementation, call ppaule@risc.uni-linz.ac.at, and there may be less well documented programs around.

It may seem, especially when looking at the numerous stunningly short articles published by Zeilberger's electronic servant ([12–16] and many more), that the availability of these powerful algorithmic methods in a sense *trivializes* the subject. This is expressed in the following quotation of Cartier, concluding his very instructive survey [6] of Zeilberger's work:

> "... Toutes les relations mentionnés ci-dessus, y compris l'extra-ordinaire récurrence d'Apéry, sont retrouvées de manière systématique et automatique, et l'on dispose d'un outil qui permet de découvrir et de démontrer des identités d'un certain type. Le jour est sans doute proche où les formulaires classiques sur les fonctions spéciales seront remplacées par un logiciel d'interrogation performant, une extension de Maple par exemple."

But even if, from this point of view, a collection like [31] has become obsolete, there is room and need for other proofs, especially of combinatorial nature, because the computer-made verifications, for all but very small cases, seem to be highly noninstructive, they apparently contain information indigestible for humans — as can be seen from the example Section A.2, and they 'only' confirm the truth of a certain statement.

2.2. Examples

2.2.1. Hypergeometric aspects

By definition, a series $\sum_k c_k$ is *hypergeometric* if the ratio c_{k+1}/c_k is a rational function of k (over a suitable field). This abstract definition can be made concrete if we imagine a complete factorization of the numerator and denominator polynomials, so that up to a suitable normalization one can write

$$\frac{c_{k+1}}{c_k}=\frac{(a_1+k)\dots(a_p+k)\cdot z}{(b_1+k)\dots(b_q+k)\cdot(k+1)}$$

for specific field elements $a_1,\dots,a_p$, the numerator parameters, $b_1,\dots,b_q$, the denominator parameters, and the factor z. Apart from a constant factor, $\sum_{k\geqslant 0}c_k$ can thus be written as

$$ {}_pF_q\left[\begin{matrix}a_1,\dots,a_p\\ b_1,\dots,b_q\end{matrix};z\right]:=\sum_{k\geqslant 0}\frac{(a_1)_k\dots(a_p)_k}{(b_1)_k\dots(b_q)_k}\cdot\frac{z^k}{k!}, \tag{6}$$

where the $(\alpha)_k$ are the *shifted factorials*, defined by

$$(\alpha)_0=1,\qquad (\alpha)_{k+1}=(\alpha+k)\cdot(\alpha)_k\quad (k\geqslant 0).$$

Another way of writing this is

$$(\alpha)_k=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$$

or, to point out the connection with binomial coefficients right away,

$$\binom{\alpha}{k}=\frac{(-\alpha)_k}{k!}(-1)^k.$$

Eq. (6) is the classical generalized hypergeometric series from the theory of special functions. The appearance of the Γ-function indicates that some care has to be taken if we want to consider (6) as an analytical object. These problems are well understood and are treated in the standard literature, e.g. in [41]. Here we are only interested in the *formal* properties of these series, and usually we will be concerned with the case where (6) *terminates* after a finite number of terms because one of the numerator parameters is a nonpositive integer, $a_1=-n$, say.

As a first illustration, note that the Franel numbers f_n and Apéry numbers a_n from the introduction can be written as

$$\begin{aligned} f_n=\sum_{k=0}^{n}\binom{n}{k}^3&=\sum_k\frac{(-n)_k(-n)_k(-n)_k}{k!k!k!}(-1)^k\\ &={}_3F_2\left[\begin{matrix}-n,-n,-n\\1,1\end{matrix};-1\right],\\ a_n=\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2&=\sum_k\frac{(-n)_k(-n)_k(n+1)_k(n+1)_k}{k!k!k!k!}\\ &={}_4F_3\left[\begin{matrix}-n,-n,n+1,n+1\\1,1,1\end{matrix};1\right].\end{aligned}$$

From the classical repertoire of hypergeometric evaluations and transformations, see e.g. [3, 17, 41, 33], I will mention only a few of the simplest ones. As already indicated, many binomial identities of varying appearance, such as[5]

$$\sum_k\binom{r}{k}\binom{s}{n-k}=\binom{r+s}{n} \tag{7}$$

and

$$\sum_k\binom{r}{k}\binom{s}{n+k}=\binom{r+s}{r+n}, \tag{8}$$

are in fact instances of the Chu–Vandermonde convolution

$$\sum_{k=0}^{n}\frac{(a)_k(b)_{n-k}}{k!(n-k)!}=\frac{(a+b)_n}{n!}, \tag{9}$$

which, written as a hypergeometric sum, is

$${}_2F_1\left[\begin{matrix}-n,a\\1-b-n\end{matrix};1\right]=\frac{(a+b)_n}{(b)_n}\quad\text{or}\quad {}_2F_1\left[\begin{matrix}-n,a\\c\end{matrix};1\right]=\frac{(c-a)_n}{(c)_n}. \tag{10}$$

[5] Assume n a nonnegative integer in (7) and n,r nonnegative integers in (8).

Note that (9) simply follows from the binomial theorem, applied to both sides of $(1-z)^{-(a+b)}=(1-z)^{-a}(1-z)^{-b}$, and that (10) is an instance of Gauss' celebrated

$$ {}_2F_1\left[\begin{matrix}a,b\\c\end{matrix};1\right]=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}. $$

In Section 3.4 below we will use the specializations

$$ \sum_k \binom{r}{k}\binom{r}{n-k}=\binom{2r}{n} \tag{11} $$

and

$$ \sum_k \binom{r}{k}\binom{s}{k}=\binom{r+s}{r} \tag{12} $$

of (7) and (8), respectively.

Typical examples of hypergeometric transformations are Pfaff's

$$ {}_2F_1\left[\begin{matrix}-n,b\\c\end{matrix};z\right]=\frac{(c-b)_n}{(c)_n}\,{}_2F_1\left[\begin{matrix}-n,b\\1+b-n-c\end{matrix};1-z\right], \tag{13} $$

which includes (10) for $z=1$, and Pfaff's

$$ {}_2F_1\left[\begin{matrix}a,b\\c\end{matrix};z\right]=(1-z)^{-a}\,{}_2F_1\left[\begin{matrix}a,c-b\\c\end{matrix};\frac{z}{z-1}\right], \tag{14} $$

which, when applied twice, gives Euler's

$$ {}_2F_1\left[\begin{matrix}a,b\\c\end{matrix};z\right]=(1-z)^{c-a-b}\,{}_2F_1\left[\begin{matrix}c-a,c-b\\c\end{matrix};z\right]. \tag{15} $$

Writing this as

$$ (1-z)^{a+b-c}\,{}_2F_1\left[\begin{matrix}a,b\\c\end{matrix};z\right]={}_2F_1\left[\begin{matrix}c-a,c-b\\c\end{matrix};z\right] \tag{16} $$

and comparing coefficients after expanding gives

$$ {}_3F_2\left[\begin{matrix}-n,a,b\\c,1+a+b-c-n\end{matrix};1\right]=\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}, \tag{17} $$

which is the Pfaff–Saalschütz identity. Again, this result appears in the literature on binomial identities in many disguises.

As an illustration, let us take a look at a short note published by Székely [53] not too long ago. He gives a combinatorial proof, by establishing a bijection between two sets of words with specific properties, of the following binomial identity:

$$ \binom{a+c+d+e}{a+c}\binom{b+c+d+e}{c+e}=\sum_k \binom{a+b+c+d+e-k}{a+b+c+d}\binom{a+d}{k+d}\binom{b+c}{k+c}, \tag{18} $$

which, by specialization of the five parameters, contains various identities to be found in the literature. They are attributed to Le Jen Shoo [31, Eq. (6.32)], Surány [31, Eq. (6.19)], Bizley [31, Eq. (6.42)], Nanjudiah [31, Eq. (6.17)], Stanley [31, Eq. (6.52)], Gould [31, Eq. (6.51)], and Takács [54].

But, not only are all these specializations instances of the Pfaff–Saalschütz identity, the same is true for (18) itself (see [66] for similar remarks and a q-analogue). To check this, replace k by $e-k$ in the summation on the r.h.s. of (18), then

$$\begin{aligned}
&\sum_k \binom{a+b+c+d+k}{k}\binom{a+d}{e-k+d}\binom{b+c}{e-k+c}\\
&=\frac{(a+d)!(b+c)!}{(a-e)!(b-e)!(e+d)!(e+c)!}\sum_k \frac{(a+b+c+d+1)_k(-e-d)_k(-e-c)_k}{k!(a-e+1)_k(b-e+1)_k}\\
&=\frac{(a+d)!(b+c)!}{(a-e)!(b-e)!(e+d)!(e+c)!}\,{}_3F_2\left[\begin{matrix}a+b+c+d+1,\,-e-d,\,-e-c\\ a-e+1,\,b-e+1\end{matrix};1\right]\\
&=\frac{(a+d)!(b+c)!}{(a-e)!(b-e)!(e+d)!(e+c)!}\,\frac{(b+c+1)_{e+d}(a+c+1)_{e+d}}{(a-e+1)_{e+d}(b-e+1)_{e+d}}\\
&=\binom{a+c+d+e}{a+c}\binom{b+c+d+e}{c+e}.
\end{aligned}$$

One of the variants, obtained by simultaneously substituting

$$a \leftarrow a-m, \quad b \leftarrow b-n, \quad c \leftarrow n, \quad d \leftarrow m, \quad e \leftarrow 0,$$

is "Stanley's identity"

$$\sum_k \binom{a}{m-k}\binom{b}{n-k}\binom{a+b+k}{k}=\binom{a+n}{m}\binom{b+m}{n} \tag{19}$$

for nonnegative integers m, n. This equivalent of the Pfaff–Saalschütz identity can be obtained from comparing coefficients in

$$\frac{(1+x)^a(1+y)^b}{(1-xy)^{1+a+b}}=\sum_{m,n}\binom{a+n}{m}\binom{b+m}{n}x^m y^n. \tag{20}$$

This generating function was obtained by Gessel and Stanton [28] in an elegant way as a constant-term identity for two-variable Laurent series — an approach inspired by multivariate Lagrange inversion technique. In the next section I will give a short combinatorial proof of (20), hence of (19) and (17).

2.2.2. Combinatorial models

The following observation is the basis of one way of combinatorially interpreting hypergeometric series:

Let a, b be nonnegative integers, γ a variable. Then the shifted factorial $(\gamma+b)_a$ is the generating polynomial with respect to cycle counting for the family of injective mappings $f\colon A \to A \cup B$, where A, B are disjoint sets with cardinalities $\# A = a$, $\# B = b$, i.e.

$$(\gamma + b)_a = \sum \{\gamma^{cyc(f)};\, f\colon A \to A \cup B,\, f \text{ injective}\},$$

where $cyc(f)$ denotes the number of cycles (within A) of f.

This concept, which gives a combinatorial meaning to the building blocks of hypergeometric series, generalizes simultaneously the usual counting of subsets by binomial coefficients ($\gamma = 1$) and the cycle enumeration for permutations ($b = 0$, which leads to the Stirling numbers of the first kind). It was introduced as a tool by Foata and myself in [25] in a combinatorial study of Laguerre polynomials, where it was used to obtain a combinatorial proof of the 'bilinear' generating function for these polynomials, the Hille–Hardy identity, and it opened the way to multivariate extensions from a combinatorial perspective. Just for illustration, note that the classical Laguerre polynomial $L_n^{(\alpha)}(x)$ can be defined as

$$\begin{aligned} n! \cdot L_n^{(\alpha)}(x) &= (1+\alpha)_n \, {}_1F_1\left[\begin{matrix} -n \\ 1+\alpha \end{matrix}; x\right] = (1+\alpha)_n \sum_{k \geq 0} \frac{(-n)_k}{(1+\alpha)_k} \frac{x^k}{k!} \\ &= \sum_{k=0}^{n} \binom{n}{k} (1+\alpha+k)_{n-k} (-x)^k = \sum \{(1+\alpha)^{cyc(f)} (-x)^{\# B};\, (A, B, f)\}, \end{aligned}$$

where the latter sum runs over all triples (A, B, f), called Laguerre configurations, where (A, B) is an ordered bipartition of $\{1, 2, \ldots, n\}$ and $f\colon A \to A \cup B$ is an injective mapping, as above. Note that the classical generating function

$$\sum_{n \geq 0} L_n^{(\alpha)}(x) t^n = (1-t)^{-1-\alpha} e^{-xt/(1+x)}$$

is an immediate consequence of the combinatorial picture, seen with the concepts of [18] or any other equivalent, by counting cyclic and linear components of these injective functions separately.

By the same approach, the Jacobi polynomials

$$\begin{aligned} n! \cdot P_n^{(\alpha, \beta)}(x) &= (1+\beta)_n \left(\frac{x+1}{2}\right)^n {}_2F_1\left[\begin{matrix} -n, -\alpha-n \\ 1+\beta \end{matrix}; \frac{x-1}{x+1}\right] \\ &= \sum_{k=0}^{n} \binom{n}{k} (1+\alpha+n-k)_k (1+\beta+k)_{n-k} \left(\frac{x-1}{2}\right)^k \left(\frac{x+1}{2}\right)^{n-k} \end{aligned} \tag{21}$$

can be understood as generating polynomials for quadruples $(A, B; f, g)$, where A, B are disjoint finite sets and $f\colon A \to A \cup B$, $g\colon B \to A \cup B$ are injective functions, with appropriate weights put on the cycles of f and g, and the points of A and B, respectively. These structures, called Jacobi configurations, can be seen as endofunctions of two-sorted (two-colored) sets, with the (local) restriction that the preimage of each point contains at most one point of each color. Hence Jacobi configurations are, in a sense, permutations of (ordered) binary trees, and if this idea is made precise, it leads to a short and most elegant proof for Jacobi's classical generating function for his polynomials. This approach — a gem in combinatorial enumeration — is due to Foata and Leroux [24], and it provides more insight than the traditional 'formal' or 'analytical' proofs.[6]

Observe that the success of the Foata–Leroux approach comes from the fact that one simultaneously has a *local* and a *global* view of the same class of structures, and does counting accordingly. This technique and the underlying ideas have been used and expanded in many studies on combinatorial properties of classical orthogonal polynomials and related functions of hypergeometric type, see [22] for an early survey and [50] for a systematic treatment.

Even though the model of Jacobi configurations is an elegant and most efficient one, it is not the best one in all situations where Jacobi polynomials are involved. There is a second classical presentation of the Jacobi polynomials, namely

$$\begin{aligned} n!\cdot P_n^{(\alpha,\beta)}(x) &= (1+\alpha)_n\,{}_2F_1\left[\begin{matrix}-n,1+\alpha+\beta+n\\ 1+\alpha\end{matrix};\frac{1-x}{2}\right] \\ &= \sum_{k=0}^{n}\binom{n}{k}(1+\alpha+\beta+n)_k(1+\alpha+k)_{n-k}\left(\frac{x-1}{2}\right)^k. \end{aligned} \tag{22}$$

As a side remark, both presentations (21) and (22) are related via a reversal of summation and an application of Pfaff's transformation (13). The reader may imagine a combinatorial interpretation of (22), similar to the one described for (21), and this is in fact the starting point for a combinatorial approach to the initial problem (1) and a generalization, see Sections 3.1, 3.2, and 4.1. The combinatorial meaning of this relation between the two different views of the Jacobi polynomials is fully described in [50].

As to the Pfaff–Saalschütz identity and its variations (including q-analogues) and specializations, there are several combinatorial proofs in the literature, see e.g. [7, 2, 32]. In particular, a proof by Foata [21], which was later q-generalized by Zeilberger in [61], uses precisely this cycle-counting approach for injective functions, as described above. I will not reproduce it here, instead I will outline a short proof of (20), hence of (19), also by means of enumerating injective functions. This proof is not an ad hoc construction. Even if this may not be evident at first sight, it represents, in a very simple situation, the general *local vs. global* counting technique for endofunctions as presented systematically in [50].

We start with the following situation: let A, B, M, N be mutually disjoint finite sets with respective cardinalities $\#A=a$, $\#B=b$, $\#M=m$, $\#N=n$. We consider pairs (f,g) of injective functions

$$f: M\to A\cup N, \qquad g: N\to B\cup M.$$

[6] It is not the shortest proof, however — that proof comes from an application of Zeilberger's algorithm, see [39], but is void of any specific 'insight'.

The number of injections f of that kind is precisely

$$(a+n)(a+n-1)\cdots(a+n-m+1)=(a+n-m+1)_m=\binom{a+n}{m}\cdot m!$$

and similarly for the g's, so that

$$\sum_{m,n\geq 0}\binom{a+n}{m}\binom{b+m}{n}x^m y^n$$

is the exponential generating functions for pairs (f,g) of injections as defined. A and B are fixed, M and N vary, and we associate the variable x (resp. y) with the points of M (resp. N).

Now look at the situation in terms of connected components. There are three types:

• f–g-cycles, i.e. (oriented) cycles of even length, with M-points and N-points alternating;

• f–g-chains, with M-points and N-points alternating, ending with a point of M (resp. N) which is mapped by f (resp. g) into A (resp. B) — let us call them ξ-structures (resp. η-structures) and denote by $\xi(x,y)$ (resp. $\eta(x,y)$) the corresponding exponential generating functions.

The exponential generating function for the cyclic components is clearly given by

$$\frac{1}{1-xy}$$

because M- and N-points come in pairs, and it is also clear from the combinatorial picture that $\xi(x,y)$ and $\eta(x,y)$ are related by the system

$$\xi(x,y)=x\cdot(1+\eta(x,y)),\qquad \eta(x,y)=y\cdot(1+\xi(x,y))$$

which shows that[7]

$$1+\xi(x,y)=\frac{1+x}{1-xy},\qquad 1+\eta(x,y)=\frac{1+y}{1-xy}.$$

Putting these informations together, we find the exponential generating function for pairs of injections to be

$$\frac{1}{1-xy}(1+\xi(x,y))^{\#A}(1+\eta(x,y))^{\#B}=\frac{(1+x)^a(1+y)^b}{(1-xy)^{1+a+b}},$$

as desired.

A similar combinatorial proof (working with ordinary instead of exponential generating functions) and a multivariate extension of the result (i.e. identity (20)) was

[7] This result can be obtained directly from the combinatorics, without using the implicit system, but in more complicated situations one does not have an explicit solution at hand and thus has to work with the system.

given by Gessel and Sturtevant [29], and independently by Constantineau [9]. In [50] the proof of this multivariate analogue comes out as a specialization of a very general approach, which includes all the series expansions given 'formally' by Carlitz in [5] — of which (20) is the simplest case.

2.2.3. Automatic verification

In this section I will give a short look at Zeilberger's method of proving and verifying hypergeometric identities. In order to outline the basic ideas, let us consider once again the Pfaff–Saalschütz identity in its original hypergeometric version (17). Put

$$F(n,k):=\frac{(-n)_k(a)_k(b)_k}{(c)_k(1+a+b-c-n)_k k!}, \qquad G(n,k):=-(a+k)(b+k)F(n,k). \tag{23}$$

Note that the following identity, after dividing both sides by $F(n,k)$,

$$\begin{aligned}&(c+n)(c-a-b+n)F(n+1,k)-(c-a+n)(c-b+n)F(n,k)\\&\quad=G(n,k)-G(n,k-1),\end{aligned} \tag{24}$$

is in fact an identity between *rational* functions, since $F(n+1,k)/F(n,k)$ and $F(n,k-1)/F(n,k)$ are rational functions of n and k. Hence (24) can be easily verified by any computer algebra system.

Using operator notation, with N denoting the shift in n, i.e. $Nf(n)=f(n+1)$, writing $S(n,N)=(c+n)(c-a-b+n)N-(c-a+n)(c-b+n)$ for the linear difference operator with polynomial coefficients (in n) on the l.h.s of (24), we get by summation over k, *telescoping* and taking care of the boundary conditions:

$$S(n,N)\sum_k F(n,k)=\sum_k S(n,N)F(n,k)=\sum_k G(n,k)-G(n,k-1))=0,$$

i.e. $S(n,N)$ is an annihilating difference operator for the l.h.s. of the Pfaff–Saalschütz identity. A proof of the Pfaff–Saalschütz identity now follows from the simple observation that

$$S(n,N)\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}=0, \tag{25}$$

i.e. both sides of the Pfaff–Saalschütz identity satisfy the same first-order recurrence, with the same initial values — hence (17) holds.

To summarize, once the difference operator $S(n,N)$ and the *certifying function* $z(n,k)=-(a+k)(b+k)$ are presented, the verification of (24) and (25), hence the proof of (17), can be routinely performed by any computer algebra system (or by tedious hand calculation). The essential point is, thus, *finding* or *constructing* the data $S(n,N)$ and $z(n,k)$, which encode all the relevant information in finite form.

As indicated in Section 2.1.3, the following holds: for 'proper' hypergeometric input, i.e. if $F(n,k)$ can be written in factorial notation as

$$F(n,k)=\frac{\prod_i(a_in+b_ik+c_i)!}{\prod_j(u_jn+v_jk+w_j)!}\,p(n,k)x^k,$$

where the a_i, b_i, u_j, v_j are integers, the c_i, w_j are parameters, and $p(n,k)$ is a polynomial in n and k, it can be shown that there exists a linear difference operator

$$S(n,N)=s_r(n)N^r+s_{r-1}(n)N^{r-1}+\cdots+s_1(n)N+s_0(n)$$

with polynomial coefficients $s_j(n)$, and a rational function $z(n,k)$, such that

$$S(n,N)F(n,k)=z(n,k)F(n,k)-z(n,k-1)F(n,k-1)$$

and hence $S(n,N)\left(\sum_k F(n,k)\right)=0$, as above. A similar assertion holds in the multisum case. Bounds for the order r of $S(n,N)$ and the degrees of the polynomial coefficients $s_j(n)$, in terms of the data entering into $F(n,k)$, can be effectively obtained. This allows for a determination of $S(n,N)$ and the certificate $z(n,k)$ by the method of undetermined coefficients.

In the case of single sums, one can apply Gosper's algorithms to

$$[s_rN^r+s_{r-1}N^{r-1}+\cdots+s_1N+s_0]F(n,k)$$

with *indeterminates* s_j, but fixed order r, and a hypergeometric term $F(n,k)$, the whole taken as a hypergeometric term in k, and decide whether there exists an indefinite hypergeometric sum w.r.t. k, a function $G(n,k)$, say. If such a solution exists, it is a rational multiple of $F(n,k)$, i.e. $G(n,k)=z(n,k)\cdot F(n,k)$. The Gosper procedure provides us with the rational function $z(n,k)$, the *certificate*, and will even solve for the s_j's, i.e. find polynomials $s_j(n)$ if the proposed order r is sufficiently high.

As a specimen, I include the output from the Zeilberger algorithm verifying the Pfaff–Saalschütz identity, taken directly from a program (written by Zeilberger himself) running under the Maple system (with input written in factorial form):

Theorem:

the sum of the following with respect to k

$$\frac{\mathrm{n!(-c-k)!(a+b-c-n)!}}{\mathrm{(n-k)!k!(-a-k)!(-b-k)!(a+b-c-n+k)!}}$$

satisfies the recurrence

$$\mathrm{-(n+c-b(-a+n+c)+(c+n)(-a-b+c+n)N}$$

Proof:

$$\mathrm{-(b+k)(a+k)}$$

$$\frac{\mathrm{rp(c-b,n)rp(c-a,n)}}{\mathrm{rp(c,n)rp(-b+c-a,n)}},\mathrm{-(b+k)(a+k)}$$

The 'theorem' part shows $F(n,k)$ from (23) (up to a constant factor) and the recurrence in the form of an operator $S(n,N)$. The 'proof' gives the certificate $z(n,k)$ and the closed form right-hand side of (17) (since the recurrence is of order 1), where rp(a, n) stands for the shifted factorial $(a)_n$.

A few more examples are given in Section A.1. See also the publications by Zeilberger, Wilf-Zeilberger, and Ekhad for many more examples.

3. Six proofs of one identity

3.1. Using Bailey's bilinear generating function

In this section I will present a proof of identity (1) based on a classical, nontrivial result of hypergeometric function theory — Bailey's bilinear generating function for the Jacobi polynomials. Recall from Section 2.2.2 that this family $(P_n^{(\alpha,\beta)}(x))_{n\geqslant 0}$ of orthogonal polynomials of hypergeometric type can be defined by

$$\frac{n!\,P_n^{(\alpha,\beta)}(x)}{(1+\alpha)_n}=\sum_{k=0}^{n}\binom{n}{k}\frac{(1+\alpha+\beta+n)_k}{(1+\alpha)_k}\left(\frac{x-1}{2}\right)^k={}_2F_1\left[\begin{matrix}-n,1+\alpha+\beta+n\\1+\alpha\end{matrix};\frac{1-x}{2}\right].$$

Now Bailey's result from [4] reads as follows:

$$\sum_{n\geqslant 0}\frac{n!(1+\alpha+\beta)_n}{(1+\alpha)_n(1+\beta)_n}P_n^{(\beta,\alpha)}(1+2x)P_n^{(\alpha,\beta)}(1+2y)t^n$$
$$=(1-t)^{-1-\alpha-\beta}F_4\left[\frac{1+\alpha+\beta}{2},\frac{2+\alpha+\beta}{2},1+\alpha,1+\beta;\,u,v\right],\tag{26}$$

where the F_4 is one of Appell's generalized hypergeometric functions [41, Section 139, 48, 17], viz.

$$F_4[a,b,c,d;\,x,y]:=\sum_{k,m\geqslant 0}\frac{(a)_{k+m}(b)_{k+m}}{(c)_k(d)_m}\cdot\frac{x^k y^m}{k!\,m!},$$

and where the parameters u and v are related to the variables x, y and t via

$$u=\frac{4y(1+x)}{(1-t)^2}t\quad\text{and}\quad v=\frac{4x(1+y)}{(1-t)^2}t.$$

For a more recent, elegant proof of (26), exploiting orthogonality, see [49]. A proof by computer is mentioned as Theorem 6.3.2 in [59]. Identity (26), related results and extensions can also be looked up in [48].

For the proof of (1) we will need Bailey's identity only in the special case $\alpha=\beta=0$. This is the case where the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ reduce to the Legendre polynomials $P_n(x)=P_n^{(0,0)}(x)$. For the sake of clarity, let us rewrite (26) in this case:

$$\sum_{n\geqslant 0}P_n(1+2x)P_n(1+2y)t^n=(1-t)^{-1}F_4[1/2,1,1,1;\,u,v].\tag{27}$$

Now, since

$$P_n(1+2x)=\sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}x^k,$$

we see that the Apéry number a_n may be written as the *constant term* (=coefficient of x^0) in the product $P_n(1+2x)\cdot P_n(1+(2/x))$.

In view of this, what should be done now is the following: expanding the right-hand side of (27), taking the coefficient of t^n in this expansion, replacing y by $1/x$ in this coefficient, and finally determining the constant term (with respect to the remaining variable x) in this expression.

$$\begin{aligned}F_4[1/2,1,1,1;\,u,v]&=\sum_{k,m}\frac{(\frac{1}{2})_{m+k}(m+k)!}{k!\,m!\,k!\,m!}\left[\frac{4y(1+x)t}{(1-t)^2}\right]^k\left[\frac{4x(1+y)t}{(1-t)^2}\right]^m\\&=\sum_{k,m}\frac{(2m+2k)!}{k!\,m!\,k!\,m!}[y(1+x)]^k[x(1+y)]^m\frac{t^{k+m}}{(1-t)^{2k+2m}}\\&=\sum_n\frac{t^n}{(1-t)^{2n}}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}^2[y(1+x)]^k[x(1+y)]^{n-k},\end{aligned}$$

hence

$$\begin{aligned}&(1-t)^{-1}F_4[1/2,1,1,1;\,u,v]\\&=\sum_n\frac{t^n}{(1-t)^{2n+1}}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}^2[y(1+x)]^k[x(1+y)]^{n-k}\\&=\sum_n t^n\cdot\sum_j t^j\frac{(2n+1)_j}{j!}\cdot\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}^2[y(1+x)]^k[x(1+y)]^{n-k}\\&=\sum_n t^n\cdot\sum_{k=0}^{n}\frac{(2k+1)_{n-k}}{(n-k)!}\cdot\binom{2k}{k}\sum_{i=0}^{k}\binom{k}{i}^2[y(1+x)]^i[x(1+y)]^{k-i}\\&=\sum_n t^n\cdot\sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\sum_{i=0}^{k}\binom{k}{i}^2[y(1+x)]^i[x(1+y)]^{k-i}.\end{aligned}$$

Now perform the substitution, replacing y by $1/x$, and observe that the constant term in

$$\left[\frac{1}{x}(1+x)\right]^i\left[x\left(1+\frac{1}{x}\right)\right]^{k-i}=\left(\frac{1}{x}+1\right)^i(x+1)^{k-i}=\frac{(1+x)^k}{x^i}$$

is $\binom{k}{i}$. The proof of identity (1) is now complete.

In Section 4.1 the same procedure will be applied to the general version of Bailey's identity, an the corresponding result will be given, together with some consequences.

3.2. A combinatorial approach to the Bailey identity

In the previous section we have seen a proof of (1) distilled out of Bailey's identity. The question about a possible combinatorial proof of (1) may thus be extended into a question about a combinatorial proof of (26). Such a proof has been given by me in [50].[8] To put this into the right perspective, recall from Section 2.2.2 that the combinatorial models described there were originally created in order to understand and extend the Hille–Hardy identity for the Laguerre polynomials from a combinatorial point of view. This approach itself was stimulated by the surprisingly simple and transparent combinatorial proof of the corresponding result for the Hermite polynomials, the well-known 'Mehler formula', by Foata, which led to multivariate extensions, see [19, 23, 20]. For Hermite polynomials, the underlying combinatorial structures are matchings of complete graphs, and the superposition of matchings — which is what has to be done combinatorially in a 'bilinear' situation — is very easy to understand. For Laguerre polynomials and Hille–Hardy, the superposition of injective functions is more difficult, but still manageable, see [25]. In [50] I revised the superposition technique so that many of the bilinear generating functions from the literature on special functions (see e.g. [48]) could be handled that way — but this approach is not able to deal with the more complicated case of Bailey's identity directly — which remained a challenge.

Without going too much into the details, let me mention that the combinatorial proof of (26), as given in [50], is based on the following 'binomial' statement, which can be proved by clever manipulation of *pairs* of injective functions and cycle counting, related to the view (22) of the Jacobi polynomials.

For nonnegative integers n, i, j and a parameter γ, we have

$$\binom{n}{i}(\gamma+n)_i\binom{n}{j}(\gamma+n)_j=\sum\binom{n}{a,b,c,d}(\gamma+n)_{a+b+d}(\gamma+a+b+d)_d,$$

where the summation is over all quadruples (a,b,c,d) of nonnegative integers such that $a+d=i$, $b+d=j$, and $a+b+c+d=n$.

From this one can deduce, by summing over the diagonal ($0\leqslant i=j\leqslant n$) and simplifying the r.h.s. using an appropriate order of summation into

$$\sum_k\binom{n}{k}^2\binom{\gamma+n+k}{k}^2=\sum_k\binom{n}{k}\binom{\gamma+n+k}{k}\sum_j\binom{k}{j}^2\binom{\gamma+2j}{2j-k}.$$

Thus we get, for $\gamma=0$, a proof of (1), provided we know that

$$\sum_{j=0}^{k}\binom{k}{j}^3=\sum_{j=0}^{k}\binom{k}{j}^2\binom{2j}{k}. \tag{28}$$

[8] Or, Zeilberger states in [59], where he raises the same question: "... although Strehl came close ...", which seems to indicate that the proof is not 100% combinatorial. Indeed, there are parts of the proof which are not really 'bijective', for legibility, but they could be turned into with some additional effort.

A simple proof of this identity is contained in the next section, and if one wishes it could be wired into the derivation given above.

In Section 4.1 I mention the binomial identity generalizing (1) which corresponds to the full Bailey identity with parametric α and β. This result can also be proved combinatorially by a slight extension of the approach outlined above.

3.3. Using a Legendre inverse pair — and the computer

In this section we start from the original problem (1), write it as

$$a_n=\sum_k \binom{2k}{k}^2\binom{n+k}{n-k}^2=\sum_k \binom{2k}{k}\binom{n+k}{n-k}c_k,$$

which means that the two sequences $(a_k)_{k\geqslant 0}$ and $(\binom{2k}{k}c_k)_{k\geqslant 0}$ form an inverse Legendre pair (cf. [42]). Thus by inversion we get

$$\binom{2n}{n}c_n=\sum_k (-1)^{n+k}d_{n,k}a_k,$$

where the coefficients $d_{n,k}$ are given by

$$d_{n,k}=\binom{2n}{n-k}-\binom{2n}{n-k-1}=\frac{2k+1}{n-k}\binom{2n}{n-k-1}=\frac{2k+1}{n+k+1}\binom{2n}{n-k}.$$

It should be noted in passing that the numbers $d_{n,k}$ have an interesting combinatorial significance: $d_{n,k}$ is the number initial segments of Dyck paths of length $2n$ ending on level $2k$ $(0\leqslant k\leqslant n)$.

An alternative way of writing the inverted relation is

$$\binom{2n}{n}c_n=(-1)^n\sum_{j=0}^{n}\binom{2j}{j}^2\sum_{k=j}^{n}(-1)^k d_{n,k}\binom{k+j}{k-j}^2.$$

Let us put

$$t_{n,j}:=\sum_{k=j}^{n}(-1)^k d_{n,k}\binom{k+j}{k-j}^2.$$

Fortunately, these numbers $t_{n,j}$ satisfy a first-order recurrence in n, as e.g. an application of Zeilberger's algorithm reveals:

$$[(n-j+1)^2N-4(n-2j)(n+\tfrac{1}{2})]\,t_{n,j}=0.$$

This allows us to obtain a closed form for these numbers:

$$t_{n,j}=(-1)^n\cdot\frac{\binom{n}{j}^2\binom{2j}{n}\binom{2n}{n}}{\binom{2j}{j}^2}=(-1)^n\binom{2n}{n-j,n-j,2j-n}\Big/\binom{2j}{j}$$

and from this we get

$$c_n = \sum_{j=0}^{n} \binom{n}{j}^2 \binom{2j}{n}.$$

Thus, what remains to be shown is the following identity, which we met already as (28) in the previous section.

$$\sum_k \binom{n}{k}^2 \binom{2k}{n} = \sum_k \binom{n}{k}^3. \tag{29}$$

Another application of Zeilberger's method — see Section A.1 — shows that the annihilating difference operator for the left-hand side is

$$(n+2)^2 N^2 - (7n^2+21n+16)N - 8(n+1)^2,$$

i.e. it is the same as the operator of the Franel recurrence (4). Since the initial values match, (29) is valid and hence (1) is proved again.

3.4. *Using the Pfaff–Saalschütz identity*

Writing (1) as a hypergeometric identity does not immediately suggest a way of proving it by using hypergeometric transformations or manipulations of binomial coefficients. But the result (29) obtained at the end of the last section suggests a way of proving (1) in two steps: first establishing (29), then obtaining (1) through a proof of

$$\sum_k \binom{n}{k}^2 \binom{n+k}{k}^2 = \sum_k \binom{n}{k}\binom{n+k}{k} \sum_j \binom{k}{j}^2 \binom{2j}{k}. \tag{30}$$

As to (29), I first had a proof using standard transformations, which I will only indicate, because after a while I found a simple way of getting this result by application of two instances of the Chu–Vandermonde formula.

Note that the hypergeometric version of (29) reads

$$\binom{2n}{n} {}_3F_2\left[\begin{matrix} -n, -n/2, -(n+1)/2 \\ 1, -n+1/2 \end{matrix}; 1\right] = {}_3F_2\left[\begin{matrix} -n, -n, -n \\ 1, 1 \end{matrix}; -1\right],$$

where the r.h.s. by use of Whipples identity (cf. e.g. [41, Section 52])

$${}_3F_2\left[\begin{matrix} -n, b, c \\ 1-b-n, 1-c-n \end{matrix}; x\right] = (1-x)^n {}_3F_2\left[\begin{matrix} -n/2, (1-n)/2, 1-b-c-n \\ 1-b-n, 1-c-n \end{matrix}; \frac{-4x}{(1-x)^2}\right]$$

transforms into

$$2^n {}_3F_2\left[\begin{matrix} -n/2, (1-n)/2, n+1 \\ 1, 1 \end{matrix}; 1\right].$$

The remaining task can be completed by suitable applications of Pfaff's identity (14) and related tools.

The interest of such a proof, when elaborated, lies in the question whether it q-generalizes routinely or not. As has been pointed out to me be Krattenthaler, Andrews and Paule, a q-analogization of such a hypergeometric proof of (29), as indicated, is not so obvious.

But here is a very simple proof of (29):

$$\begin{aligned}
\sum_k \binom{n}{k}^3 &= \sum_{i+j=n} \binom{n}{i}\binom{n}{j}\binom{i+j}{i} \\
&= \sum_{i+j=n} \binom{n}{i}\binom{n}{j} \sum_k \binom{i}{k}\binom{j}{k} \\
&= \sum_k \binom{n}{k}^2 \sum_{i+j=n} \binom{n-k}{n-i}\binom{n-k}{n-j} \\
&= \sum_k \binom{n}{k}^2 \sum_{i+j=n} \binom{k}{i}\binom{k}{j} \\
&= \sum_k \binom{n}{k}^2 \binom{2k}{n}.
\end{aligned}$$

Note the applications of the Chu–Vandermonde identity (11) and (12) at the beginning and at the end of this transformation.

To put this into the right context, observe that the same kind of derivation shows

$$\sum_k \binom{n}{k}^3 \alpha^k \beta^{n-k} = \sum_k \binom{n}{k}^2 \langle t^n \rangle (1+\alpha t)^k (1+\beta t)^k$$

(where $\langle t^n \rangle \dots$ means 'coefficient of t^n in $\dots$'), which in the case $\alpha = -1$, $\beta = 1$ reduces to the well-known 'Dixon's formula' ([31, Eq. (6.6)] see [41, Section 53] for the hypergeometric result):

$$\sum_k (-1)^k \binom{n}{k}^3 = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ (-1)^m \binom{2m}{m}\binom{3m}{m} & \text{if } n = 2m. \end{cases}$$

For the general case, note that consideration of $\langle t^n \rangle (1 + (\alpha+\beta)t + \alpha\beta t^2)^k$ and another application of Chu–Vandermonde leads to MacMahon's identity [38, Vol. 1, p. 122]

$$\sum_k \binom{n}{k}^3 \alpha^k \beta^{n-k} = \sum_k \binom{n}{2k}\binom{2k}{k}\binom{n+k}{k} (\alpha\beta)^k (\alpha+\beta)^{n-2k}.$$

Let us now show that (1) follows from (29) by an application of the Pfaff–Saalschütz identity. We take (17) with $n = j$, $c = 1$, $a = -n+j$, $b = n+j+1$, i.e.

$${}_3F_2\left[\begin{matrix} -j, -n+j, 1+n+j \\ 1, 1+j \end{matrix}; 1\right] = \frac{(1+n-j)_j(-n-j)_j}{j!(-2j)_j} = \binom{n+j}{n-j}.$$

Thus

$$\begin{aligned}
\sum_j \binom{n}{j}^2 \binom{n+j}{j} &= \sum_j \binom{n}{j}\binom{n+j}{j}\cdot\binom{2j}{j}\binom{n+j}{n-j} \\
&= \sum_j \binom{n}{j}\binom{n+j}{j}\cdot\binom{2j}{j}\,{}_3F_2\left[\begin{matrix}-j,\,-n+j,\,1+n+j\\ 1,\,1+j\end{matrix};1\right] \\
&= \sum_j \binom{n}{j}\binom{n+j}{j}\cdot\binom{2j}{j}\sum_{k\geqslant 0}\frac{(-j)_k(-n+j)_k(1+n+j)_k}{(1)_k(1+j)_k k!} \\
&= \sum_j \frac{(1+n-j)_j(1+n)_j(1+j)_j}{j!^3}\sum_{k\geqslant 0}\frac{(1+j-k)_k(1+n-k-j)_k(1+n+j)_k}{(1)_k(1+j)_k k!} \\
&= \sum_j \frac{(n+j)!}{(n-j)!}\,\frac{(1+j)_j}{j!^3}\sum_{k\geqslant j}\frac{(1+2j-k)_{k-j}(1+n-k)_{k-j}(1+n+j)_{k-j}}{(1)_{k-j}(1+j)_{k-j}(1)_{k-j}}.
\end{aligned}$$

Writing now

$$\frac{(n+j)!}{(n-j)!} \quad \text{as} \quad \frac{(n+k)!}{(n-k)!}\cdot\frac{1}{(1+n+j)_{k-j}(1+n-k)_{k-j}}$$

and splitting $(1+2j-k)_k$ into $(1+2j-k)_{k-j}(1+j)_j$ the last expression for our summation turns into

$$\sum_k \frac{(n+k)!}{(n-k)!}\sum_{j\leqslant k}\frac{(1+2j-k)_{k-j}(1+j)_j}{j!^2(k-j)!^2 k!} = \sum_k \binom{n}{k}\binom{n+k}{k}\sum_{j\leqslant k}\binom{k}{j}^2\binom{2j}{k},$$

as desired.

3.5. Applying Zeiberger's algorithm

The general algorithm of Wilf–Zeilberger [60, 59] for finding recurrence operators for hypergeometric multisums can be directly applied to the right-hand side of (1). Since an implementation of this algorithm was not yet available to me when I first encountered the problem, I asked Zeilberger (in June 1992) to put the double sum on his machine — and here is the result:[9] the algorithm gives back the difference operator

$$S(n,N) = (n+1)^3 - (2n+3)(17n^2+51n+39)N + (n+2)^3 N^2,$$

which is precisely what we want, namely the same difference operator as the one belonging to the Apéry numbers, cf. (2), and it further claims that if you put

$$F(j,k,n) := \frac{(n+k)!\,k!}{(n-k)!(k-j)!^3 j!^3}$$

[9] It must be admitted that things look even a bit nastier than necessary, due to a simplification bug for powers of factorials in Maple V, which had to be circumvented.

then

$$S(n,N)\sum_{j,k} F(j,k,n)=0,$$

and as a justification for this claim one is asked to 'routinely verify'

$$\begin{aligned}
&S(n,N)F(j,k,n)\\
&\quad=(J-1)\left(\frac{j(2n+3)p(j,k,n)}{(j+1)(n-k+2)(n-k+1)(k-j+3)(k-j+2)^2(k-j+1)^3}\right)F(j,k,n)\\
&\qquad+(K-1)\left(\frac{(2n+3)q(j,k,n)}{(n-k+2)(n-k+1)(j+1)^2(j+2)(k-j+1)^2(k-j+2)}\right)F(j,k,n),
\end{aligned}$$

where $p(j,k,n)$ and $q(j,k,n)$ are monstrously looking polynomials reproduced in Section A.2. Note that J and K are shift operators for the variables j and k, respectively.

3.6. Recurrences and simplification

In contrast to the proofs presented up to here, the last proof of (1) I present will work directly with the known recurrences (4) (resp. (2)) for the Franel (resp. Apery) numbers. Although it deals with difference operators, it is convenient to present it in matrix form.

To begin with, let us define a doubly infinite tridiagonal matrix $\boldsymbol{F}=(f_{i,j})_{i,j\geqslant 0}$ representing the difference operator of the Franel recurrence (4):

$$f_{i,j}=f_{i-j}(i) \quad \text{where} \quad f_k(z):=\begin{cases}(z+1)^2 & \text{if } k=-1,\\ -(7z^2+7z+2) & \text{if } k=0,\\ -8z^2 & \text{if } k=1,\\ 0 & \text{if } |k|>1,\end{cases}$$

so that

$$\boldsymbol{F}=\begin{pmatrix} f_0(0) & f_{-1}(0) & f_{-2}(0) & \cdots\\ f_1(1) & f_0(1) & f_{-1}(1) & \cdots\\ f_2(2) & f_1(2) & f_0(2) & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix}=\begin{pmatrix} -2 & 1 & 0 & 0 & \cdots\\ -8 & -16 & 4 & 0 & \cdots\\ 0 & -32 & -44 & 9 & \cdots\\ 0 & 0 & -72 & -86 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}.$$

If we denote by $\boldsymbol{f}=(f_n)_{n\geqslant 0}$ the infinite column vector with the Franel numbers $f_n=\sum_k \binom{n}{k}^3$ as entries, then Franel's result (4) is equivalent to

$$\boldsymbol{F}\cdot\boldsymbol{f}=\boldsymbol{0},$$

where $\boldsymbol{0}$ denotes the zero column vector.

Similarly, the difference operator occurring in the Apéry recurrence (2) has matrix form

$$A=\begin{pmatrix} a_0(0) & a_{-1}(0) & a_{-2}(0) & \cdots \\ a_1(1) & a_0(1) & a_{-1}(1) & \cdots \\ a_2(2) & a_1(2) & a_0(2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}=\begin{pmatrix} -5 & 1 & 0 & 0 & \cdots \\ 1 & -117 & 8 & 0 & \cdots \\ 0 & 8 & -535 & 27 & \cdots \\ 0 & 0 & 27 & -1463 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where

$$a_{i,j}=a_{i-j}(i) \quad \text{where} \quad a_k(z):=\begin{cases} (z+1)^3 & \text{if } k=-1, \\ -(z+1)^3-z^3-4(2z+1)^3 & \text{if } k=0, \\ z^3 & \text{if } k=1, \\ 0 & \text{if } |k|>1. \end{cases}$$

Then Apéry's recurrence (2) can be written as

$$\boldsymbol{A}\cdot\boldsymbol{a}=\boldsymbol{0},$$

where $\boldsymbol{a}=(a_n)_{n\geqslant 0}$ is the vector of the Apéry numbers, $a_n=\sum_k \binom{n}{k}^2\binom{n+k}{k}^2$.

Finally we introduce the doubly inifinite, lower triangluar matrix $\boldsymbol{P}=(p_{i,j})_{i,j\geqslant 0}$ of the *Legendre transform*:

$$\boldsymbol{P}=\left(\binom{i}{j}\binom{i+j}{j}\right)_{i,j\geqslant 0}=\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & 2 & 0 & 0 & \cdots \\ 1 & 6 & 6 & 0 & \cdots \\ 1 & 12 & 30 & 20 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

As mentioned in the introduction, Schmidt's question (5) asks for the inverse image of the sequence of Apéry's numbers under the Legendre transform, and the identity (2) claims that Apéry's sequence $\boldsymbol{a}$ is the image of Franel's sequence $\boldsymbol{f}$ under the Legendre transform, i.e. $\boldsymbol{a}=\boldsymbol{P}\cdot\boldsymbol{f}$. Put into matrix terms, what we would like to show is that

$$\boldsymbol{A}\cdot\boldsymbol{P}\cdot\boldsymbol{f}=\boldsymbol{0}.$$

Obviously, we would be done if we could exhibit a matrix $\boldsymbol{X}$ such that

$$\boldsymbol{A}\cdot\boldsymbol{P}=\boldsymbol{X}\cdot\boldsymbol{F}$$

Computing initial segments of this unknown (infinite) matrix $\boldsymbol{X}$ led me to the following (surprising?) *guess*:

$$\boldsymbol{X}=\boldsymbol{D}\cdot\boldsymbol{P} \text{ and hence } \boldsymbol{A}=\boldsymbol{D}\cdot\boldsymbol{P}\cdot\boldsymbol{F}\cdot\boldsymbol{P}^{-1}, \tag{31}$$

where $\boldsymbol{D}=(d_{i,j})_{i,j\geqslant 0}$ is a diagonal matrix given by $d_{i,i}=4i+2$ $(i\geqslant 0)$.

It remains to *prove* this claim. Note that it not only says that the Franel and the Apéry sequences are related via the Legendre transform, but also that the associated difference operators are related in the sense of *conjugation* via Legendre transform (up to multiplication with a diagonal matrix). Interestingly, even though we have to deal with infinite matrices containing binomial coefficients, i.e. nonrational terms, the proof can be established by rational arithmetic alone. For this to see, let us write the assertion (31) to be proved in the form:

$$\boldsymbol{A}\cdot\boldsymbol{P}=\boldsymbol{D}\cdot\boldsymbol{P}\cdot\boldsymbol{F}.$$

Consider the (i,j)-entry on both sides of this equation. Since both $\boldsymbol{A}$ and $\boldsymbol{F}$ are *tridiagonal* matrices, every such term involves three summands only. Write this as

$$lhs=a_1(i)p_{i-1,j}+a_0(i)p_{i,j}+a_{-1}(i)p_{i+1,j},$$

$$rhs=(4i+2)[p_{i,j-1}f_{-1}(j-1)+p_{i,j}f_0(j)+p_{i,j+1}f_1(j+1)],$$

where now i and j are treated as *variables*. Now ask for simplification of $lhs-rhs$. The Maple command expand(lhs − rhs) gives back an expression of considerable size. Of course, simplification could be done by hand, but this tedious task is better accomplished by your computer algebra system. On the command simplify (expand(lhs − rhs)) Maple responds with 0, thus *proving* that the Franel recurrence operator $\boldsymbol{F}$ and the Apéry recurrence operator $\boldsymbol{A}$ are Legendre conjugates in the following sense:

$$\boldsymbol{A}=\boldsymbol{D}\cdot\boldsymbol{P}\cdot\boldsymbol{F}\cdot\boldsymbol{P}^{-1},$$

with $\boldsymbol{A},\boldsymbol{F},\boldsymbol{D},\boldsymbol{P}$ as above.

The binomial identity (1) we started with is just one of the consequences of this general fact — perhaps the most interesting one. Some further comments will be made in Section 4.3.

4. Variants and extensions

4.1. Using the Bailey identity in generality

Bailey's bilinear generating function (26) provides us with the family of identities (for $n\geqslant 0$)

$$\frac{n!}{(1+\beta)_n}P_n^{(\beta,\alpha)}(1+2x)\frac{n!}{(1+\alpha)_n}P_n^{(\alpha,\beta)}(1+2y)$$

$$=\sum_{\substack{k,m,r\geqslant 0\\ k+m+r=n}}\binom{n}{k,m,r}\frac{(1+\alpha+\beta+n)_{k+m}}{(1+\alpha)_k(1+\beta)_m}[y(1+x)]^k[x(1+y)]^m$$

or, if we use (22) as a way to write the Jacobi polynomials

$$\sum_j \binom{n}{j} \frac{(1+\alpha+\beta+n)_j}{(1+\beta)_j} x^j \cdot \sum_i \binom{n}{i} \frac{(1+\alpha+\beta+n)_i}{(1+\alpha)_i} y^i$$

$$= \sum_{\substack{k,m,r\geq 0\\ k+m+r=n}} \binom{n}{k,m,r} \frac{(1+\alpha+\beta+n)_{k+m}}{(1+\alpha)_k(1+\beta)_m} [y(1+x)]^k [x(1+y)]^m$$

$$= \sum_{k=0}^{n} \binom{n}{k} (1+\alpha+\beta+n)_k \sum_{j=0}^{k} \binom{k}{j} \frac{[y(1+x)]^j}{(1+\alpha)_j} \frac{[x(1+y)]^{k-j}}{(1+\beta)_{k-j}}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \frac{(1+\alpha+\beta+n)_k}{(1+\beta)_k} \sum_{j=0}^{k} \binom{k}{j} \frac{(1+\beta+k-j)_j}{(1+\alpha)_j} [y(1+x)]^j [x(1+y)]^{k-j}.$$

Replacing now y by $1/x$ and looking for the constant coefficient on both sides, as we did before, gives us

$$\sum_k \binom{n}{k}^2 \frac{(1+\alpha+\beta+n)_k^2}{(1+\alpha)_k(1+\beta)_k} = \sum_k \binom{n}{k} \frac{(1+\alpha+\beta+n)_k}{(1+\beta)_k} \sum_{j=0}^{k} \binom{k}{j}^2 \frac{(1+\beta+k-j)_j}{(1+\alpha)_j}. \tag{32}$$

Written as a binomial identity this reads

$$\sum_{k=0}^{n} \frac{\binom{n}{k}^2 \binom{\alpha+\beta+n+k}{k}^2}{\binom{\alpha+k}{k}\binom{\beta+k}{k}} = \sum_{k=0}^{n} \frac{\binom{n}{k}\binom{\alpha+\beta+n+k}{k}}{\binom{\beta+k}{k}} \sum_{j=0}^{k} \frac{\binom{k}{j}^2\binom{\beta+k}{j}}{\binom{\alpha+j}{j}}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \binom{\alpha+\beta+n+k}{k} \sum_{j=0}^{k} \frac{\binom{k}{j}^3}{\binom{\alpha+j}{j}\binom{\beta+k-j}{k-j}}.$$

Using standard hypergeometric notation, (32) could be equivalently written as

$${}_4F_3\left[\begin{matrix} -n, -n, 1+\alpha+\beta+n, 1+\alpha+\beta+n \\ 1, 1+\alpha, 1+\beta \end{matrix}; 1\right]$$

$$= \sum_k \binom{n}{k} \frac{(1+\alpha+\beta+n)_k}{(1+\beta)_k} \, {}_3F_2\left[\begin{matrix} -k, -k, -\beta-k \\ 1, 1+\alpha \end{matrix}; \; -1\right].$$

Indeed, what we have seen above is that Bailey's bilinear generating function can be stated as

$${}_2F_1\left[\begin{matrix} -n, 1+\alpha+\beta+n \\ 1+\beta \end{matrix}; -x\right] \cdot {}_2F_1\left[\begin{matrix} -n, 1+\alpha+\beta+n \\ 1+\alpha \end{matrix}; -y\right]$$

$$= \sum_{k=0}^{n} \binom{n}{k} \frac{(1+\alpha+\beta+n)_k}{(1+\beta)_k} [x(1+y)]^k \, {}_2F_1\left[\begin{matrix} -k, -k-\beta \\ 1+\alpha \end{matrix}; \frac{y(1+x)}{x(1+y)}\right].$$

To conclude this section, I will draw some consequences from the generalized identity (32). First note that obviously the setting $\alpha=\beta=0$ brings us back to (1).

Now let $\beta=\lambda\cdot\alpha$, substitute this into (32) and let $\alpha\to\infty$. Comparing the behavior of both sides leads to

$$\sum_{k=0}^{n}\binom{n}{k}^2\frac{(\lambda+1)^{2k}}{\lambda^k}=\sum_{k=0}^{n}\binom{n}{k}\left(\frac{\lambda+1}{\lambda}\right)^k\sum_{j=0}^{k}\binom{k}{j}^2\lambda^j. \tag{33}$$

If we take now $\lambda=1$, we will find

$$\sum_{k=0}^{n}\binom{n}{k}^2 4^k=\sum_{k=0}^{n}\binom{n}{k}\binom{2k}{k}2^k. \tag{34}$$

We may use Pfaff's transformation (13) to turn the left-hand side of (34) into

$$\sum_{k=0}^{n}\binom{n}{k}\binom{2n-k}{n}3^k \tag{35}$$

and the right-hand side of (34) into

$$\frac{1}{4^n}\sum_{k=0}^{n}\binom{2n-2k}{n-k}\binom{2k}{k}9^k \tag{36}$$

(see entries (3.65) and (3.84) in Gould's listing [31]). So all quantities in (34)–(36) are numerically equivalent. These consequences of (32) do not look too complicated, so they should have easier proofs than (32) itself. Indeed, it was pointed out to the author by Gessel that (34) is the special case $\alpha=1$ in the hypergeometric identity

$${}_2F_1\left[\begin{matrix}-n,\alpha-\frac{1}{2}\\2\alpha-1\end{matrix};-8\right]={}_2F_1\left[\begin{matrix}1-\alpha-n,n\\\alpha\end{matrix};4\right],$$

which in turn is nothing but a special case ($a=-n$, $b=1-\alpha-n$, and $w=2$) of Gauss' quadratic transform

$${}_2F_1\left[\begin{matrix}a,a-b+\frac{1}{2}\\2a-2b+1\end{matrix};\frac{-4w}{(1-w)^2}\right]={}_2F_1\left[\begin{matrix}a,b\\a-b+1\end{matrix};w^2\right]$$

(see [41, p. 65, Theorem 23]).

Let us now compare coefficients of λ^a (a any integer) in (33). This leads to the following family of identities:

$$\sum_{k=0}^{n}\binom{n}{k}^2\binom{2k}{k+a}=\sum_{k=0}^{n}\binom{n}{k}\sum_{j=0}^{k}\binom{k}{j}^2\binom{k}{j-a}, \tag{37}$$

which for $a=0$ turns into the beautiful

$$\sum_{k=0}^{n}\binom{n}{k}^2\binom{2k}{k}=\sum_{k=0}^{n}\binom{n}{k}\sum_{j=0}^{k}\binom{k}{j}^3. \tag{38}$$

Again, as Gessel pointed out to me, (37) and hence (38) could be proved by clever application of certain hypergeometric transformations, but he also states that "... I'm not sure that this proof would do any good, since it's probably no simpler than the proof you already have ...."

4.2. Higher exponents

The problem (1) raised in the beginning of this article can be generalized into the following:

For any positive integer $e \geqslant 1$, define a sequence of rational numbers $(c_k^{(e)})_{k\geqslant 0}$, independent of n, by

$$\sum_{k=0}^{n}\binom{n}{k}^e\binom{n+k}{k}^e=\sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}c_k^{(e)}.$$

Is it then true that all number $c_k^{(e)}$ are integers?

Trivially $c_k^{(1)}=1$ for all k, and Section 3 provided proofs of the less trivial fact that

$$c_k^{(2)}=\sum_{j=0}^{k}\binom{k}{j}^3 \quad (k\geqslant 0).$$

According to Schmidt there is number-theoretical interest in this question for general e, but it does not seem to be a simple task to find an equally concise presentation of $c_k^{(e)}$ for $e>2$ — if possible at all. [By the way, even though the numbers $c_n^{(0)}$ are not integers, $c_n^{(0)}=(-1)^{n+1}(4n-2)^{-1}$, the case $e=0$ is not uninteresting from a combinatorial point of view: combinatorics of lattice path and Catalan numbers are involved!]

Proceeding in exactly the same way as in Section 3.3 one has

$$\binom{2n}{n}c_n^{(e)}=(-1)^n\sum_{j=0}^{n}\binom{2j}{j}^e\sum_{k=j}^{n}(-1)^k d_{n,k}\binom{k+j}{k-j}^e,$$

which suggests to consider the numbers

$$t_{n,j}^{(e)}:=\sum_{k=j}^{n}(-1)^k d_{n,k}\binom{k+j}{k-j}^e.$$

In order to prove the integrality of the numbers $c_n^{(e)}$ it would be sufficient to show that

$$\binom{2n}{n}\,\Big|\,\binom{2j}{j}^e\cdot t_{n,j}^{(e)} \quad (0\leqslant j\leqslant n,\ e\geqslant 1).$$

But, according to numerical experiments, more seems to be true. I conjecture that even

$$\binom{2n}{n}\,\Big|\,\binom{2j}{j}\cdot t_{n,j}^{(e)} \quad (0\leqslant j\leqslant n,\ e\geqslant 1)$$

holds.

Zeilberger's algorithm shows that in the cases $e=1,2,3$ the annihilating difference operator for the $t_{n,j}^{(e)}$ (w.r.t. n) is of first order, but for $e>3$ its order is >1, as far as I checked. The same is true if we alternatively look for a difference operator w.r.t. the variable j. Hence we do not get a closed form expression for the $t_{n,j}^{(e)}$ in these cases, and thus we cannot draw the wanted conclusions about divisibility.

For $e=3$, however, we are in a similar situation as previously for $e=2$ in Section 3.3. The operator annihilating $t_{n,j}^{(3)}$ is

$$(n-j+1)^3 N - 4(n-3j)(n+1)(n+1/2).$$

This leads to the closed form

$$t_{n,j}^{(3)} = (-1)^n \cdot \frac{\binom{n}{j}^2 \binom{2j}{n-j} \binom{2n}{n}}{\binom{2j}{j}} = (-1)^n \binom{2n}{n-j, n-j, n-j, 3j-n}$$

and hence

$$c_n^{(3)} = \sum_{j=0}^{n} \binom{n}{j}^2 \binom{2j}{j}^2 \binom{2j}{n-j}.$$

If we now put Zeilberger's algorithm into action in order to obtain a difference operator in n which annihilates this sum, then we get a result which is quite a bit more voluminous than the result in the case $e=2$: it is an operator of degree 6 (!)

$$p_0(n) + p_1(n)N + \ldots + p_6(n)N^6,$$

the polynomial coefficients of which are given in Section A.3 — compare with the innocent-looking recurrence for the Franel numbers $f_n = c_n^{(2)}$!

4.3. *Recurrences and diophantine approximation*

In this section I would like to mention two aspects of identity (1) that have to do with the fact that it relates the recurrences (4) of Franel and (2) of Apéry.

4.3.1. *Transforming recurrence operators*

As in Section 3.6, let $\boldsymbol{P}$ denote the matrix of the Legendre transform. Further, let $\boldsymbol{G} = (g_{i,j})_{i,j \geqslant 0}$ be a matrix representing a second-order linear recurrence with polynomial coefficients, i.e. we have polynomials $g_{-1}(z)$, $g_0(z)$, $g_1(z)$ and

$$g_{i,j}(z) = \begin{cases} g_{i-j}(i) & \text{if } |i-j| \leqslant 1, \\ 0 & \text{else.} \end{cases}$$

Now let $\boldsymbol{H} = \boldsymbol{P} \cdot \boldsymbol{G} \cdot \boldsymbol{P}^{-1}$ be the matrix conjugate to $\boldsymbol{G}$ w.r.t. Legendre transform. In general, there is no reason for $\boldsymbol{H}$ to be a tridiagonal matrix representing a second-order linear recurrence, even though it is easy to see that the entries of $\boldsymbol{H}$ along diagonals are again values of polynomial or rational functions. A partial answer to the question about conditions under which a recurrence operator as $\boldsymbol{G}$ is transformed into a recurrence operator of the same kind, is given by Theorem 1 in [47], where it is shown that for polynomial $g_i(z)$, $i \in \{-1, 0, 1\}$, of degree $\leqslant 2$ the following holds.

Let a, b, c, d, e be parameters, and put

$$g_{-1}(z) = az^2 + bz + c,$$

$$g_0(z) = (a+d)z^2 - (a-b-d)z + e,$$

$$g_1(z) = dz^2.$$

Then $\boldsymbol{H}=(h_{i,j})_{i,j\geqslant 0}=\boldsymbol{P}\cdot\boldsymbol{G}\cdot\boldsymbol{P}^{-1}$ is again a tridiagonal matrix with

$$h_{i,j}(z)=\begin{cases}h_{i,j}(i) & \text{if } |i-j|\leqslant 1,\\ 0 & \text{else,}\end{cases}$$

where

$$h_{-1}(z)=\frac{(z+1)}{4z+2}(az^2+bz+c),$$

$$h_0(z)=\tfrac{1}{2}[(2d-a)z(z+1)+2e-c],$$

$$h_1(z)=\frac{z}{4z+2}[az^2+(2a-b)z+a-b+c].$$

Note that the Franel–Apéry conjugacy of Section 3.6 corresponds to the choice $a=c=1$, $b=2$, $d=-8$, $e=-2$ of the parameters.

The proof of this sufficient condition in [47] is given by creative telescoping, whereas in [52] this result is proved by exactly the same method — simplification of rationals expression — as applied in Section 3.6 to the Franel–Apéry case. The general problem of (Legendre) conjugacy is treated to some extent in [52], where it is also shown that the sufficient condition above, expressed in terms of the five parameters, is also necessary.

4.3.2. *Back to diophantine approximation*

To conclude, let me come back to a remark from the beginning and briefly mention the use which can be made of the Franel–Apéry conjugacy in the field of diophantine approximation. The material of this section is due to Schmidt [47].

Recall that the two sequences

$$(a_n)_{n\geqslant 0}=(1,5,73,1445,33001,\ldots),\qquad (b_n)_{n\geqslant 0}=\left(0,6,\frac{351}{4},\frac{62531}{36},\frac{11424695}{288},\ldots\right),$$

both satisfying Apéry's linear recurrence relation

$$n^3u_n-(34n^3-51n^2+27n-5)u_{n-1}+(n-1)^3u_{n-2}=0, \tag{39}$$

play a prominent role in Apéry proof of the irrationality of $\zeta(3)$, due to the fact that

$$\lim_{n\to\infty}\frac{b_n}{a_n}=\zeta(3)$$

and that the $a_n=\sum_k\binom{n}{k}^2\binom{n+k}{k}^2$ are integers, whereas the b_n are rational numbers with denominator dividing $2\,\mathrm{lcm}(1,2,\ldots,n)^3$, more precisely:

$$b_n=\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2 c_{n,k}\quad\text{where } c_{n,k}=\sum_{m=1}^{n}\frac{1}{m^3}+\sum_{m=1}^{k}\frac{(-1)^{m-1}}{2m^3\binom{n}{m}\binom{n+m}{m}}.$$

In terms of our matrix notation, writing $\boldsymbol{A}$ for the matrix representing the Apéry recurrence, $\boldsymbol{a}$ for the (column) vector of the a_n ($n \geqslant 0$) and similarly for $\boldsymbol{b}$, we have

$$\boldsymbol{A} \cdot \boldsymbol{a} = \mathbf{0} \quad \text{and} \quad \boldsymbol{A} \cdot \boldsymbol{b} = \begin{bmatrix} 6 \\ \mathbf{0} \end{bmatrix},$$

where on the right we have a vector with a 0 in all positions, except a 6 in the top position.

As to the Franel recurrence, it has been remarked by Cusick (see a footnote in [55]) that both sequences

$$(f_n)_{n \geqslant 0} = (1, 2, 10, 56, 346, \ldots), \qquad (g_n)_{n \geqslant 0} = \left(0, 3, 12, \frac{208}{3}, \frac{1280}{3}, \ldots\right)$$

satisfying the Franel recurrence, which we may write as

$$\boldsymbol{F} \cdot \boldsymbol{f} = \mathbf{0} \quad \text{and} \quad \boldsymbol{F} \cdot \boldsymbol{g} = \begin{bmatrix} 3 \\ \mathbf{0} \end{bmatrix},$$

lead to a diophantine approximation:

$$\lim \frac{g_n}{f_n} = \frac{\pi^2}{8}. \tag{40}$$

We have seen above that the equations

$$\boldsymbol{A} \cdot \boldsymbol{a} = \mathbf{0} \quad \text{and} \quad \boldsymbol{F} \cdot \boldsymbol{f} = \mathbf{0}$$

are related via Legendre conjugacy. Now conjugacy provides us with a solution of an inhomogeneous Franel recurrence associated to the second solution $\boldsymbol{b}$ of the (homogeneous) Apéry recurrence:

$$\boldsymbol{A} \cdot \boldsymbol{b} = \begin{bmatrix} 6 \\ \mathbf{0} \end{bmatrix} \quad \text{vs.} \quad \boldsymbol{F} \cdot \boldsymbol{h} = \boldsymbol{k},$$

where $\boldsymbol{h}$ and $\boldsymbol{k}$ are vectors belonging to the sequences

$$(h_n)_{n \geqslant 0} = \left(0, 3, \frac{93}{8}, \frac{1217}{18}, \frac{239429}{576}, \ldots\right),$$

$$(k_n)_{n \geqslant 0} = \left(3, -\frac{3}{2}, 1, -\frac{4}{3}, \ldots, (-1)^n \frac{3}{n+1}, \ldots\right).$$

Similarly, the second solution $\boldsymbol{g}$ of the (homogeneous) Franel recurrence provides us with a solution of an inhomogeneous Apéry recurrence:

$$\boldsymbol{F} \cdot \boldsymbol{g} = \begin{bmatrix} 3 \\ \mathbf{0} \end{bmatrix} \quad \text{vs.} \quad \boldsymbol{A} \cdot \boldsymbol{c} = \boldsymbol{d},$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are vectors belonging to the sequences

$$(c_n)_{n\geqslant 0}=\left(0,6,90,\frac{5348}{3},\frac{122130}{3}\cdots\right),\qquad (d_n)_{n\geqslant 0}=(6,18,30,\dots,3(4n+2),\dots).$$

As a result, we get the approximations

$$\lim_{n\to\infty}\frac{h_n}{f_n}=\zeta(3)\quad\text{and}\quad\lim_{n\to\infty}\frac{c_n}{a_n}=\frac{\pi^2}{8},$$

i.e. one approximation of $\zeta(3)$ in terms of Franel-recursive sequences, and one of $\pi^2/8$ in terms of Apéry-recursive sequences. From this, one may hope for a proof of independence of $\{1,\pi^2/8,\zeta(3)\}$ over the rationals.

Appendix A

A.1. Some operators and certificates

• The verification of the Apéry recurrence — annihilating difference operator for $\sum_k\binom{n}{k}^2\binom{n+k}{k}^2$ and certificate $z(n,k)$:

$$(n+1)^3-(17n^2+51n+39)N(2n+3)+(n+2)^3N^2,$$

$$-4\frac{(2n+3)(n+k+1)^2(9+12n+4n^2-k-2k^2)}{(n-k+1)^2}$$

• The verification of the Franel recurrence — annihilating difference operator for $\sum_k\binom{n}{k}^3$ and certificate $z(n,k)$:

$$-8(n+1)^2+(-7n^2-21n-16)N+(n+2)^2N^2,$$

$$-(n+1)^2$$

$$(14n^3+47n^2+53n+20-30k-57nk-27n^2k+18nk^2+18k^2-4k^3)$$

$$/\ (n-k+1)^3$$

• The verification of identity (28) — annihilating difference operator for $\sum_k\binom{n}{k}^2\binom{2k}{n}$ and certificate $z(n,k)$:

$$-8(n+1)^2+(-7n^2-21n-16)N+(n+2)^2N^2,$$

$$-2\frac{(n+1)(2k+1)(k+1)(-2k+3n+4)}{(n-k+1)^2}$$

Note that in this and the previous example we have same operator, but different certificates, of course.

A.2. Verification of (1) by the Wilf–Zeilberger method

The certifying polynomial $p(j,k,n)$ of Section 3.5 is given by

$$
\begin{aligned}
&108j^5k^3n-69k^2jn^4-12j^6kn-2944j^2k^2n^2+130j^3k^4n^2-960j^3kn^3\\
&+792j^4n^3+132j^4n^4-103670kj-13350k^2j+34043kj^2+5816j^3k^3\\
&-5266j^4k^2+10122j^3k^2+628j^5k-5040j^4k+8478j^3k-5691j^2k^3\\
&-11598j^2k^2+52jk^3+102j^3k^5-880j^4k^4+1772j^3k^4+1270j^5k^3\\
&-3876j^4k^3-568j^6k^2+2322j^5k^2-26j^6k-18j^2k^5-1532j^2k^4\\
&+30jk^5-48j^3k^6-30j^4k^5+150j^5k^4-166j^6k^3+60j^7k^2\\
&+72j^2k^6-48jk^6-3579kn-14604jn+13302j^2n+7749k^2n\\
&+2679j^3n-1641k^3n-588j^4k^3n+694j^4kn^2-12645n-414k^2jn^3\\
&+54j^3k^5n-196j^4k^3n^2+3006j^3k^2n-618j^4k^2n-12j^6k^2n-8274n^2\\
&-116052j-2309j^4+65507k-20093j^3-114j^6+1401j^5\\
&+29838k^2+87861j^2+4151k^3+1530k^5+12k^6-2706n^3\\
&+41338-451n^4+8370j^3kn+390j^3k^4n+54j^2k^5n+18j^3k^5n^2\\
&-834kj^2n+696jk^3n+354j^5kn+210j^5k^2n-2799k^2jn-360j^5n^3\\
&-1515j^2k^3n-216j^4k^4n-6780j^2k^2n+12j^2k^4n+756j^3k^3n-2724j^4kn\\
&-12837kjn+210j^3k^2n^2+280j^4k^2n^2-4j^6k^2n^2+70j^5k^2n^2-1554k^2jn^2\\
&+1350j^3kn^2+18j^2k^5n^2-71kJ^2n^2+232jk^3n^2+36j^5k^3n^2-206j^5kn^2\\
&-16j^2k^3n^4-649j^2k^3n^2-72j^4k^4n^2+4j^2k^4n^2+108j^3k^3n^2-4594kjn^2\\
&-4j^6kn^2+54j^4k^2n^4+324j^4k^2n^3-16j^3k^3n^4-96j^3k^3n^3-84j^6n\\
&+969j^5n-3573j^4n+324k^5n+12j^4k^6+6j^5k^5-6j^6k^4\\
&-4307kn^2-4562jn^2+5037j^2n^2+3042k^2n^2+317j^3n^2-1033k^3n^2\\
&+44j^6n^2-217j^5n^2-3j^4n^2+108k^5n^2-324k^3n^3-54k^3n^4\\
&+51k^2n^4+306k^2n^3-60j^5n^4+8j^6n^4+48j^6n^3+402j^2n^3\\
&+67j^2n^4-64j^3n^4-384j^3n^3+34jn^4+204jn^3-346kn^4\\
&-2076kn^3-88j^3k^2n^4-528j^3k^2n^3-160j^3kn^4+1068j^4kn^3+23j^2n^4\\
&+138kj^2n^3+178j^4kn^4-76j^2k^2n^4-456j^2k^2n^3-36j^5kn^4-216j^5kn^3\\
&-35kjn^4-210kjn^3-96j^2k^3n^3
\end{aligned}
$$

and similarly the certifying polynomial $q(j,k,n)$ is given by

$$
\begin{aligned}
&18j^5k^3n-66^{j6}kn-117k^4jn^2+38j^2k^2n^2-30228kj-16819k^2j\\
&+40207kj^2-342j^4k^2+1452j^3k^2+740j^5k-2767j^4k-5848j^3k\\
&+688j^2k^2-50j^4k^4+200j^5k^3-190j^4k^3-134j^6k^2+102j^5k^2\\
&-84j^6k+72j^5k^4-152j^6k^3+120j^7k^2+2268kn-5928jn
\end{aligned}
$$

$$+216j^2n+2430k^2n-3642j^3n-204j^4k^3n-363j^4kn^2-351k^4jn$$
$$-68j^4k^3n^2+156j^3k^2n-702j^4k^2n+30j^6k^2n-29180j+8348j^4$$
$$+10752k-15682j^3-300j^6+1092j^5+11520k^2+3466j^2$$
$$+768k^5-1518j^3kn+8217kj^2n+888j^5kn+330j^5k^2n-3537k^2jn$$
$$-18j^4k^4n+114j^2k^2n-1089j^4kn-6336kjn+52j^3k^2n^2-234j^4k^2n^2$$
$$+10j^6k^2n^2+110j^5k^2n^2-1179k^2jn^2-506j^3kn^2+2739kj^2n^2+6j^5k^3n^2$$
$$+296j^5kn^2-6j^4k^4n^2-2112kjn^2-22j^6kn^2-372j^6n-678j^5n$$
$$+2244j^4n+162k^5n+6j^4k^6-6j^5k^5-12j^6k^4+756kn^2$$
$$-1976jn^2+72j^2n^2+810k^2n^2-1214j^3n^2-124j^6n^2+226j^5n^2$$
$$+748j^4n^2+54k^5n^2-1643k^4j.$$

A.3. A difference operator for exponent e = 3

Here I reproduce the difference operator annihilating $\sum_k \binom{n}{k}^2 \binom{2k}{k}^2 \binom{2k}{n-k}$ of Section 4.2. I renounce the reproduction of the accompanying certificate, which is a rational function that has, apart from a few linear factors, a numerator polynomial of total degree 19 in n, k, with 176 terms and coefficients of similar size as the ones appearing in the operator — compare with the corresponding operator(s) for $e=2$, given in Section A.1!

$$\begin{aligned}
p_0(n) = {} & -8281(6620250n^7+194374905n^6+2404907720n^5+16182725885n^4 \\
& +63538882220n^5+144000168368n^2+171051956992n \\
& +78847954240)(n+3)^2(n+2)^2(n+1)^2,
\end{aligned}$$

$$\begin{aligned}
p_1(n) = {} & -(163672440750n^9+5624687410065n^8+84404024171040n^7 \\
& +723615434540555n^6+3888554690243330n^5+13490882373318699n^4 \\
& +29878295174795508n^3+39883006527213557n^2 \\
& +27783753376637444n+6664869698310180)(n+3)^3(n+2)^2,
\end{aligned}$$

$$\begin{aligned}
p_2(n) = {} & -(162401352750n^{11}+6718774773555n^{10}+124711194036660n^9 \\
& +1368380628262070n^8+9836894865699870n^7+48472570983304643n^6 \\
& +166181428133860302 0n^5+393051049322866468n^4 \\
& +619419550265900144n^3+601996733156814672n^2 \\
& +303106129360889920n+46189584994221760)(n+3)^2,
\end{aligned}$$

$$\begin{aligned}
p_3(n) = {} & (-53385696000n^{13}-2689035368670n^{12}-61906699547820n^{11} \\
& -861476777355560n^{10}-8072186328339580n^9-53672105152884912n^8 \\
& -259786052073295348n^7-923038053700083216 4n^6
\end{aligned}$$

$$-2392735085813946388n^5-4425366900944064930n^4$$
$$-5577505932198043888n^3-4376732891104921364n^2$$
$$-1744617790556183136n-168635012492565360),$$

$$p_4(n)=(-139025250n^{13}-6365859255n^{12}-130710111720n^{11}$$
$$-1580390020385n^{10}-12395319285010n^9-65207080885953n^8$$
$$-227510670047436n^7-486141379391615n^6-447630919087976n^5$$
$$+442524695319912n^4+1299863359935856n^3-296727667685808n^2$$
$$-3271290691708800n-2661031246809600),$$

$$p_5(n)=-(297911250n^{11}+13235400225n^{10}+263446626120n^9$$
$$+3094268495855n^8+23755870990890n^7+124635949498195n^6$$
$$+453029629903388n^5+1128916484999241n^4$$
$$+1854976781234016n^3+1841060829487412n^2+893652306644064n$$
$$+89391519915264)(n+5)^2,$$

$$p_6(n)=(6620250n^7+148033155n^6+1377683540n^5+6842102110n^4$$
$$+19201266530n^3+29207398143n^2+19841994296n$$
$$+2222856216)(n+5)^2(n+6)^4.$$

Printing out the recurrence (of degree 6) and the certificate for $\sum_k \binom{n}{k}^3 \binom{n+k}{k}^3$ would require several pages, and the corresponding data for exponent $e=4$ would require about as much space as this whole article.

Acknowledgements

I am grateful to G. Andrews, I. Gessel, C. Krattenthaler, P. Paule, A. Schmidt, D. Stanton and D. Zeilberger for discussions on aspects of identity (1) and its proofs.

References

[1] G.E. Andrews, Applications of basic hypergeometric functions, SIAM Rev. 16 (1974) 441–484.
[2] G.E. Andrews, Identities in combinatorics, I: on sorting two ordered sets. Discrete Math. 11 (1975) 97–106.
[3] W.N. Bailey, Generalized Hypergeometric Series (Cambridge Univ. Press, London, 1935).
[4] W.N. Bailey, The generating function for Jacobi polynomials, J. London Math. Soc. 13 (1938) 243–246.
[5] L. Carlitz, Multiple binomial and power sums, Houston J. Math 6 (1980) 331–354.
[6] P. Cartier, Démonstration 'automatique' d'identités et fonctions hypergéometriques (d'après D. Zeilberger), in: Séminare Bourbaki, Vol. 746, 1991.
[7] P. Cartier and D. Foata, Problèmes combinatoires de commutation et de réarrangements, Lecture Notes in Mathematics, Vol. 85 (Springer, Berlin, 1969).

[8] W. Chu, Inversion techniques and combinatorial identities, preprint, Lehrstuhl II für Mathematik, Universität Bayreuth, 1992.

[9] I. Constantineau, A generalization of the Pfaff–Saalschütz theorem, Stud. Appl. Math. 85 (1991) 243–248.

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

[11] G.P. Egorychev, Integral Representation and Computation of Combinatorial Sums, Transl. Math. Monographs, Vol. 59 (Amer. Mathematical Soc., Providence, RI, 1984).

[12] S.B. Ekhad, Short proofs of two hypergeometric summation formulas of Karlsson, Proc. Amer. Math. Soc. 107 (1989) 1143–1144.

[13] S.B. Ekhad, A very short proof of Dixon's theorem, J. Combin. Theory Ser. A 54 (1990) 141–142.

[14] S.B. Ekhad, A one-line proof of the Habsieger–Zeilberger G_2 constant term identity, J. Comput. Appl. Math. 90 (1991) 133–134.

[15] S.B. Ekhad, Short proof of a strange combinatorial identity conjectured by Gosper, Discrete Math. 90 (1991) 319–320.

[16] S.B. Ekhad and D. Zeilberger, A 21st century proof of Dougall's hypergeometric sum identity, J. Math. Anal. Appl. 147 (1990) 610–611.

[17] A. Erdélyi et al., Higher Transcendental Functions (McGraw-Hill, New York, 1953) (The Bateman Manuscript Project, 3 vols.).

[18] D. Foata, La série génératrice exponentielle dans les problèmes d'énumération (Presses de l'Université de Montréal, 1974) pp. 186.

[19] D. Foata, A combinatorial proof of the Mehler formula, J. Combin. Theory Ser. A 24 (1978) 367–376.

[20] D. Foata, Some Hermite polynomial identities and their combinatorics, Adv. Appl. Math. 2 (1981) 250–259.

[21] D. Foata, Une démonstration combinatoire de l'identité de Pfaff–Saalschütz, C.R. Acad. Sci. Paris 297 (1983) 221–224.

[22] D. Foata, Combinatoire des identités sur les polynômes orthogonaux, in: Proc. Internat. Congr. of Math., Warsaw, 16–24 August 1983, Vol. 2 (PWN — Polish Scientific Publications and North-Holland, Amsterdam, 1985) 1541–1553.

[23] D. Foata and A. Garsia, A combinatorial approach of the Mehler formulas for Hermite polynomials, in: Relations between Combinatorics and other Parts of Mathematics, Columbus, 1978 (Amer. Mathematical Soc., Providence, RI, 1979) 163–179.

[24] D. Foata and P. Leroux, Polynômes de Jacobi, interprétation combinatoire et fonction génératrice, Proc. Amer. Math. Soc. 87 (1983) 47–53.

[25] D. Foata and V. Strehl, Combinatorics of Laguerre polynomials, in: D.M. Jackson and S.A. Vanstone, eds., Enumeration and Design, Waterloo, 1982 (Academic Press, Toronto, 1984) 123–140.

[26] J. Franel, notes published in: L'Intermédiaire des Mathématiciens, Vol. 1 (1894) 45–47, and Vol. 2 (1895) 33–35.

[27] I. Gessel, A combinatorial proof of the multivariable Lagrange inversion formula, J. Combin. Theory Ser. A 45 (1987) 178–195.

[28] I. Gessel and D. Santon, Short proofs of Saalschütz's and Dixon's theorems, J. Combin. Theory Ser A 38 (1985) 87–90.

[29] I. Gessel and D. Sturtevant, A combinatorial proof of Saalschütz theorem, preprint, 1989.

[30] R.W. Gosper, Decision procedure for indefinite hypergeometric summation, Proc. Nat. Acad. Sci. USA 75 (1978) 40–42.

[31] H.W. Gould, Combinatorial Identities: A Standard Set of Tables Listing 500 Binomial Coefficient Identities (rev. ed.) (Morgantown, VA, 1972).

[32] I.P. Goulden, A bijective proof of the q-Saalschütz identity, Discrete Math. 57 (1985) 39–44.

[33] R.L. Graham, D.E. Knuth and O. Patashnik, Concrete Mathematics (Addison-Wesley, Reading, MA, 1989).

[34] D.H. Greene and D.E. Knuth, Mathematics for the Analysis of Algorithms (Birkhäuser, Basel, 1981).

[35] J. Hornegger, Hypergeometrische Summation and polynomiale Rekursion, Diplomarbeit, IMMD-I, Erlangen, 1992.

[36] D.E. Knuth, The Art of Computer Programming, Vol. I: Fundamental Algorithms (Addison-Wesley, Reading, MA, 1969).

[37] T.H. Koornwinder, On Zeilberger's algorithm and its q-analogue: a rigorous description, Tech. Report AM-R9207, Centre for Mathematics and Computer Science, Amsterdam, 1992.

[38] P.A. MacMahon, Combinatory Analysis (Cambridge Univ. Press Cambridge, 1915) (Reprinted by Chelsea, New York, 1955).

[39] S. Parnes and S.B. Ekhad, A WZ-style proof of Jacobi polynomial's generating function, Discrete Math. 110 (1992) 263–264.

[40] M.A. Perlstadt, Some recurrences for sums of powers of binomial coefficients, J. Number Theory 27 (1987) 304–309.

[41] E. Rainville, Special Functions (Chelsea, Bronx, NY, 1971).

[42] J. Riordan, Combinatorial Identities (Wiley, New York, 1968).

[43] R. Roy, Binomial identities and hypergeometric functions, Amer. Math. Monthly 94(1) (1987) 36–46.

[44] B. Salvy and P. Zimmerman, GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable, preprint, October 1992.

[45] A.L. Schmidt, Generalized Legendre polynomials, J. Reine Angew. Math. 404 (1990) 192–202.

[46] A.L. Schmidt, Generalized Legendre and q-Legendre polynomials, in: VII Simposium sobre polinomios oriogonales y applicaciones, Granada, 1991: J. Comp. Appl. Math., to appear.

[47] A.L. Schmidt, Legendre transforms and Apéry's sequences, preprint series 14, Matematisk Institut, University of Copenhagen, August 1992.

[48] H.M. Srivastava and H.L. Manocha, A Treatise on Generating Functions (Ellis Horwood, Chichester, 1984).

[49] D. Stanton, A short proof of a generating function for the Jacobi polynomials, Proc. Amer. Math. Soc. 80 (1980) 398–400.

[50] V. Strehl, Zykel-Enumeration bei lokal-strukturierten Funktionen, Habilitationsschrift. Erlangen, 1990.

[51] V. Strehl, Combinatorics of special functions: facets of Brock's identity, in: P. Leroux and C. Reutenauer, eds., Actes 4e Colloque Séries Formelles et Combinatoire Algébrique, Montréal, 1992 (Université du Québec à Montréal, LACIM, 1992) 363–378.

[52] V. Strehl, Recurrences and Legendre transform, Actes du Séminaire Lotharingien de Combinatoire, 29^e session, Thurnau, 1992, A. Kerber (ed.). Publications de l'IRMA 1993/033, Strasbourg, 81–100.

[53] L.A. Székely, Common origin of cubic binomial identities; a generalization of surányi's proof of Le Jen Shoo's formula, J. Combin Theory Ser. A 40 (1985) 171–174.

[54] L. Takács, On an identity of Shih-Chieh Chu, Acta Sci. Math. (Szeged) 34 (1973) 383–391.

[55] A.J. van der Poorten, A proof that Euler missed ... Apérys proof of the irrationality of $\zeta(3)$, Math. Intelligencer 1 (1978) 195–203.

[56] P. Verbaeten, The automatic construction of pure recurrence relations, in: Proc. EUROSAM '74; ACM-SIGSAM Bull. 8 (1974) 96–98.

[57] P. Verbaeten, Rekursiebetrekkingen voor lineaire hypergeometrische funkties, Proefschrift voor het doctoraat in de toegepaste wetenschappen, Katholieke Universiteit Leuven, Belgium, 1976.

[58] G. Viennot, Heaps of pieces I: basic definitions and combinatorial lemmas, in: G. Labelle and P. Leroux, eds., Combinatoire Énumérative, Montréal, 1985, Lecture Notes in Mathematics, Vol. 1234 (Springer, Berlin, 1986) 210–245.

[59] H.S. Wilf and D. Zeilberger, An algorithmic proof theory for hypergeometric (ordinary and 'q') multisum/integral identities, Invent. Math. 108 (1992) 575–633.

[60] H.S. Wilf and D. Zeilberger, Rational function certification of multisum/integral/'q' identities, Bull. Amer. Math. Soc. (new series) 27 (1992) 148–153.

[61] D. Zeilberger, A q-Foata proof of the q-Saalschütz identity. European J. Combin. 8 (1987) 461–463.

[62] D. Zeilberger, A fast algorithm for proving terminating hypergeometric identities, Discrete Math. 80 (1990) 207–211.

[63] D. Zeilberger, A holonomic systems approach to special function identities, J. Comput. Appl. Math. 32 (1990) 321–368.

[64] D. Zeilberger, A Maple program for proving hypergeometric identities, SIGSAM Bull. 25(3) (1991) 4–13.

[65] D. Zeilberger, The method of creative telescoping, J. Symbolic Comput. 11 (1991) 195–205.

[66] J. Zeng, Pfaff–Saalschütz revisited, J. Combin. Theory Ser. A 51 (1989) 141–143.