IDENTITIES

and their computer proofs

A series of lectures delivered by

Herbert S. Wilf
University of Pennsylvania

as "SPICE" lectures, to the U. S. National Security Agency*

August 31- September 2, 1993

These notes are available at no charge via anonymous ftp of the file `pub/wilf/lecnotes.ps` at the site `ftp.cis.upenn.edu`. Comments will be appreciatively received at `wilf@central.cis.upenn.edu`

Revised July 15, 1994

* Originally subtitled as *"... to a group of people who have unaccountably wandered into a shed near BWI to listen."*

Lecture notes: Identities, and their computer proofs.

Contents:

(I) The world of identities

There are beautiful identities in many branches of mathematics. Number theory, for instance, is one of their prime habitats:

$$\sum_{k\backslash n}\mu(k)=\begin{cases}1, & \text{if } n=1;\\ 0 & \text{else,}\end{cases}$$

$$\prod_{p}(1-p^{-s})^{-1}=\sum_{n\geq 1}\frac{1}{n^s}\qquad (\Re(s)>1),$$

$$\det\left((\gcd(i,j))_{i,j=1}^{n}\right)=\phi(1)\phi(2)\cdots\phi(n),$$

$$1+\sum_{m=1}^{\infty}\frac{x^{m^2}}{(1-x)(1-x^2)\cdots(1-x^m)}=\prod_{m=0}^{\infty}\frac{1}{(1-x^{5m+1})(1-x^{5m+4})}.$$

Combinatorics is one of the major producers of marvellous identities:

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

$$\sum_{i+j+k=n}\binom{i+j}{i}\binom{j+k}{j}\binom{k+i}{k}=\sum_{0\leq j\leq n}\binom{2j}{j},$$

$$\exp\left\{\sum_{m\geq 1}m^{m-1}\frac{t^m}{m!}\right\}=1+\sum_{n\geq 1}(n+1)^{n-1}\frac{t^n}{n!},$$

$$\sum_{s=0}^{2m}(-1)^s\binom{2m}{s}^3=(-1)^m\frac{(3m)!}{(m!)^3}.$$

Beautiful identities have often stimulated mathematicians to find beautiful proofs; proofs that have perhaps illuminated the combinatorial or other significance of the equality of the two members, or possibly just dazzled us with their unexpected compactness and elegance.

Here we will not, of course, be able to discuss all kinds of identities. Far from it. We are going to concentrate on one family of identities, called hypergeometric identities, that have been of great interest and importance, and include many of the famous binomial coefficient identities of combinatorics.

The proof theory of these identities has gone through roughly three phases of evolution. At first each identity was treated on its own merits. Combinatorial insights proved some, generating functions proved others, special tricks proved many, but unified methods of wide scope were lacking.

At the same time, the theory of hypergeometric functions, initiated by Gauss, was developing into a deep and rich area of mathematical analysis, but the realization that hypergeometric function theory could make serious contributions to understanding combinatorial identities was slow in coming. That realization, generally credited to George Andrews, finally arrived in 1974.

As a result of his observation, many binomial coefficient and other combinatorial identities were recognized to be special cases of a rather small set of hypergeometric identities, which we will call here *the* *hypergeometric database.* Hypergeometric identities that were not direct special cases of the database entries could often be found by additional *transformation rules*, that extended the database by moving the origins of the series involved. The unification and codification afforded by this recognition greatly stimulated the development of the field.

The third major phase of the proof theory was inaugurated by the recognition that yet another kind of work in hypergeometric function theory could be applied to the world of identities. That recognition is due to Doron Zeilberger, in 1982. He took note of the PhD dissertation of Sister Mary Celine Fasenmyer, in 1945, and he saw that the scope of the methods that she found extended far beyond finding recurrence relations for polynomial sequences, which is what they were originally used for. Zeilberger saw that these methods of Sister Celine, together, of course, with the development of fast computers, opened the door for a fully algorithmic paradigm for proving hypergeometric identities. Although the use of the hypergeometric database is moderately systematic, yet there is no universal agreement on the totality of identities that belong in the database, nor is there agreement on exactly what you should do if your identity seems not to be in the database. But now the way was clear for theorems and algorithms that would universally prove, certify, and discover identities in this area.

These talks are a survey* of some of what has been happening in this field recently.

A *geometric* series $\sum_{k\geq 0} t_k$ is one in which the ratio of two consecutive terms is constant, i.e., is a constant function of the summation index k. The k^{th} term, t_k, of a geometric series is of the form cx^k where c and x are constants, i.e., are independent of the summation index k. Therefore a general *geometric* series looks like

$$\sum_k cx^k.$$

A *hypergeometric* series $\sum_{k\geq 0} t_k$ is one in which the ratio of two consecutive terms is a *rational function* of the summation index k, i.e., in which

$$\frac{t_{k+1}}{t_k} = \frac{P(k)}{Q(k)},$$

where P and Q are polynomials in k. Examples of such hypergeometric terms t_k are $t_k = x^k$, or $k!$, or $(2k+7)!/(k-3)!$, or $(k^2-1)(3k+1)!/((k+3)!(2k+7))$.

Hypergeometric series are very important in mathematics. Many of the familiar functions of analysis are hypergeometric. These include the exponential, logarithmic, trigonometric, binomial, and Bessel functions, along with the classical orthogonal polynomial sequences of Legendre, Tschebycheff, Laguerre, Hermite, etc.

It is important to recognize that a given series is hypergeometric because the general theory of hypergeometric functions is very powerful, and we may gain a lot of insight into a function that concerns us by first recognizing that it is hypergeometric, and then by using known results about such functions.

We will see, beginning in the next chapter, how it can help us to do sums if we can recognize that the sums are hypergeometric. In this chapter we will concentrate on just the recognition problem.

In the ratio of consecutive terms, P/Q, let us imagine that the polynomials P and Q have been completely factored, in the form

$$\frac{t_{k+1}}{t_k} = \frac{P(k)}{Q(k)} = \frac{(k+a_1)(k+a_2)\cdots(k+a_p)}{(k+b_1)(k+b_2)\cdots(k+b_q)(k+1)}x, \tag{1.1}$$

where x is a constant. If we normalize by taking $t_0 = 1$, then we call the hypergeometric series whose terms are the t_k's the series

$${}_pF_q\left[\begin{matrix} a_1 & a_2 & \cdots & a_p \\ b_1 & b_2 & \cdots & b_q \end{matrix} \middle| x\right].$$

The a's and the b's are called, respectively, the *upper*, and the *lower* parameters of the series.

* I am indebted to Peter Paule and Marko Petkovšek for many helpful comments on these notes.

To put it another way, the hypergeometric series

$$ {}_pF_q\left[\begin{matrix} a_1 & a_2 & \cdots & a_p \\ b_1 & b_2 & \cdots & b_q \end{matrix} \middle| x\right] $$

is the series whose initial term is 1, and in which the ratio of the $(k+1)^{\mathrm{st}}$ term to the k^{th} is given by (1.1) above, for all $k \geq 0$.

The appearance of the factor $(k+1)$ in the denominator of (1.1) needs a few words of explanation. Why, one might ask, does it have to be segregated that way, when it might just as well be absorbed as one of the factors $(k+b_i)$? The answer is that there's no reason except that it has always been done that way, and far be it from me to try to reverse the course of the Nile. If there is not a factor of $(k+1)$ in the denominator of your term ratio, i.e., if there is no factor of $k!$ in the denominator of your term t_k, just put it in, and compensate for having done so by putting an extra factor in the numerator.

How to identify a given series as hypergeometric

Many of the famous functions of classical analysis have hypergeometric series expansions. In the exponential series, $e^x = \sum_{k\geq 0} x^k/k!$, the initial term is 1, and the ratio of the $(k+1)^{\mathrm{st}}$ term to the k^{th} is $x/(k+1)$, which is certainly of the form (1.1). So $e^x = {}_0F_0[\ |x]$.

What we want to do now is to show how one may identify a given hypergeometric series as a particular ${}_pF_q[\cdots]$. This process is at the heart of using the hypergeometric database as a database. If we have some hypergeometric series that interests us for some reason, we might wonder what is known about it. Is it possible to sum the series in simple form? Is it possible to transform the series into another form that is easier to work with? Is some result that we have just discovered about this series really new or is it well known? These questions can be answered by consulting the extensive literature on hypergeometric series. But the first step is to rewrite the series that interests us in the standard ${}_pF_q[\cdots]$ form, because the literature is cast in those terms.

The hypergeometric series lookup algorithm

1. Given a series $\sum_k t_k$. Shift the summation index k so that the sum starts at $k=0$ with a nonzero term. Extract the term corresponding to $k=0$ as a common factor so that the first term of the sum will be 1.
2. Simplify the ratio t_{k+1}/t_k to bring it into the form $P(k)/Q(k)$, where P, Q are polynomials. If this cannot be done, the series is not hypergeometric.
3. Completely factor the polynomials P and Q into linear factors, and write the term ratio in the form

$$ \frac{P(k)}{Q(k)} = \frac{(k+a_1)(k+a_2)\cdots(k+a_p)}{(k+b_1)(k+b_2)\cdots(k+b_q)(k+1)} x $$

   If the factor $k+1$ in the denominator wasn't there, put it in, and compensate by inserting an extra factor of $k+1$ in the numerator. Notice that all of the coefficients of k, in numerator and denominator, are $+1$. Whatever numerical factors are needed to achieve this are absorbed into the factor "x".
4. You have now identified the input series. It is (the common factor that you extracted in step 1 above, multiplied by) the hypergeometric series

$$ {}_pF_q\left[\begin{matrix} a_1 & a_2 & \cdots & a_p \\ b_1 & b_2 & \cdots & b_q \end{matrix} \middle| x\right]. \quad \blacksquare $$

Example. If the k^{th} term of the series is $t_k = 2^k/k!$, then we have $t_{k+1}/t_k = 2/(k+1)$ which is the form (1.1) with $p = q = 0$ and $x = 2$. Consequently, since $t_0 = 1$, the given series is

$$\sum_{k\geq 0} \frac{2^k}{k!} = {}_0F_0\left[\left.\,\right| 2\right].$$

Example. Consider the series $\sum_k t_k$ where $t_k = 1/((2k+1)(2k+3)!)$.

To identify this series, note that the smallest value of k for which the term t_k is nonzero is the term with $k = -1$. Hence we begin by shifting the origin of the sum as follows:

$$\sum_{k\geq -1} \frac{1}{(2k+1)(2k+3)!} = \sum_{k\geq 0} \frac{1}{(2k-1)(2k+1)!}$$

The ratio of two consecutive terms is

$$\frac{t_{k+1}}{t_k} = \frac{(k-\frac{1}{2})}{(k+\frac{1}{2})(k+\frac{3}{2})(k+1)}\frac{1}{4}. \tag{1.2.2}$$

Hence our given series is identified as

$$\sum_k \frac{1}{(2k+1)(2k+3)!} = -\,{}_1F_2\left[\left.\begin{matrix} -1/2 & \\ 1/2 & 3/2 \end{matrix}\right| 1/4\right].$$

Example. Suppose we define the symbol

$$[x,d]_n = \begin{cases} \prod_{j=0}^{n-1}(x-jd), & \text{if } n>0; \\ 1, & \text{if } n=0; \\ \prod_{j=1}^{-n}\frac{1}{x+jd}, & \text{if } n<0. \end{cases}$$

Now consider the series

$$\sum_{k=0}^{n}\binom{n}{k}[x,d]_k[y,d]_{n-k}.$$

Is this a hypergeometric series, and if so which one is it?

The term ratio is

$$\begin{aligned}
\frac{t_{k+1}}{t_k} &= \frac{\binom{n}{k+1}[x,d]_{k+1}[y,d]_{n-k-1}}{\binom{n}{k}[x,d]_k[y,d]_{n-k}} \\
&= \frac{(n-k)\prod_{j=0}^{k}(x-jd)\prod_{j=0}^{n-k-2}(y-jd)}{(k+1)\prod_{j=0}^{k-1}(x-jd)\prod_{j=0}^{n-k-1}(y-jd)} \\
&= \frac{(n-k)(x-kd)}{(k+1)(y-(n-k-1)d)} \\
&= \frac{(k-n)(k-\frac{x}{d})}{(k+(\frac{y}{d}-n+1))(k+1)}.
\end{aligned}$$

This is exactly in the standard form (1.1), so we have identitified our series as the hypergeometric series

$${}_2F_1\left[\left.\begin{matrix} -n & -\frac{x}{d} \\ \frac{y}{d}-n+1 & \end{matrix}\right| 1\right].$$

Example. Suppose we are wondering if the sum

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{k!}$$

can be evaluated in some simple form. A first step would be to identify it as a hypergeometric series. The next step would be to look up that hypergeometric series in the database to see if anything is known about it. Let's do the first step here.

The term ratio is

$$\begin{aligned} \frac{t_{k+1}}{t_k} &= \frac{\binom{n}{k+1}\frac{(-1)^{k+1}}{(k+1)!}}{\binom{n}{k}\frac{(-1)^k}{k!}} \\ &= \frac{k-n}{(k+1)^2}, \end{aligned}$$

and $t_0 = 1$. Hence by (1.1) our unknown sum is revealed to be a

$${}_1F_1\left[\begin{matrix}-n\\1\end{matrix}\middle|1\right].$$

Example. Is the Bessel function

$$J_p(x) = \sum_{k=0}^{\infty} \frac{(-1)^k(\frac{x}{2})^{2k+p}}{k!(k+p)!}$$

a hypergeometric function? The ratio of consecutive terms is

$$\begin{aligned} \frac{t_{k+1}}{t_k} &= \frac{(-1)^{k+1}(\frac{x}{2})^{2k+2+p}k!(k+p)!}{(k+1)!(k+p+1)!(-1)^k(\frac{x}{2})^{2k+p}} \\ &= \frac{-(\frac{x^2}{4})}{(k+1)(k+p+1)}. \end{aligned}$$

Here we must take note of the fact that $t_0 \neq 1$, whereas the standard hypergeometric series begins with a term equal to 1. Our conclusion is that the Bessel function is indeed hypergeometric, and it is in fact

$$J_p(x) = \frac{(\frac{x}{2})^p}{(p+1)!}\,{}_0F_1\left[\begin{matrix}\cdots\\p+1\end{matrix}\middle| -\frac{x^2}{4}\right].$$

The recognition of a given series as a hypergeometric series is handled nicely by Mathematica's `Algebra‘SymbolicSum‘` package, or Maple's `convert/hypergeom`, which also find the upper and lower parameters and argument of the function in question.

(II) The Hypergeometric Database

In these notes, which are primarily about sums, the hypergeometric sums occupy center stage. This class includes multitudes of sums that contain binomial coefficients and factorials, including virtually all of the familiar ones that have been summed in closed form.†

The fact is that quite a number of hypergeometric sums can be expressed in simple closed form. These are called identities. Typically when we look at such an identity we will see an equation which on the left hand side has a sum in which the summand contains a number of factorials, binomial coefficients, etc., and on the right has a considerably simpler function that is equal to the sum on the left.

If you have a sum that looks like it might be a known hypergeometric sum, then one way to find out if it is known is to look it up in the hypergeometric database. An entry in this database is a statement to the effect that a certain hypergeometric series is equal to a certain much simpler expression, for all values of the various free parameters that appear, or at least all values in a suitably restricted range.

In the sequel we will use the notation

$$(a)_n \stackrel{\text{def}}{=} \begin{cases} a(a+1)(a+2)\cdots(a+n-1) & \text{if } n \geq 1; \\ 1, & \text{if } n = 0. \end{cases}$$

for the *rising factorial* function. In terms of the rising factorial function, the general hypergeometric series looks like this-

$${}_pF_q \left[\begin{matrix} a_1 & a_2 & \ldots & a_p \\ b_1 & b_2 & \ldots & b_q \end{matrix} \middle| z \right] = \sum_{k \geq 0} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{z^k}{k!}. \tag{2.1}$$

The series is well defined as long as the *lower parameters* $b_1, b_2, \ldots, b_q$ are not negative integers. If the argument z happens to be equal to 1, we will often omit it altogether from the symbol on the left of (2.1). The series *terminates* automatically if any of the *upper parameters* $a_1, a_2, \ldots, a_p$ is a negative integer, otherwise it is *nonterminating*, i.e., it is an infinite series. If the series is well defined and nonterminating then questions of convergence or divergence become relevant. In these notes we will be concerned almost exclusively with terminating series.

The following are some of the most useful database entries. We will not prove any of them just now because *all* of their proofs will follow instantly from the computer certification methods that we will develop in chapter 13.

We use $n!$ and $\Gamma(n+1)$ interchangeably.

(I) Gauss's ${}_2F_1$ identity

If b is a nonpositive integer or $c - a - b$ has positive real part, then

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

† We will use the term "closed form" in an imprecise, colloquial sense only, to mean some form that is quite a lot simpler in its appearance than the original sum.

(II) Kummer's ${}_2F_1$ identity

If $a-b+c=1$ and $\Re(a)<1$ then

$${}_2F_1\left[\begin{matrix} a & b \\ c & \end{matrix}\middle| -1\right] = \frac{\Gamma(\frac{b}{2}+1)\Gamma(b-a+1)}{\Gamma(b+1)\Gamma(\frac{b}{2}-a+1)}.$$

If b is a negative integer then this identity should be used in the form

$${}_2F_1\left[\begin{matrix} a & b \\ c & \end{matrix}\middle| -1\right] = 2\cos\left(\frac{\pi b}{2}\right)\frac{\Gamma(|b|)\Gamma(b-a+1)}{\Gamma(\frac{|b|}{2})\Gamma(\frac{b}{2}-a+1)},$$

which follows from the first form by using the reflection formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

for the Γ-function and taking the limit as b approaches a negative integer.

(III) Saalschütz's ${}_3F_2$ identity

If $d+e=a+b+c+1$ and c is a negative integer then

$${}_3F_2\left[\begin{matrix} a & b & c \\ d & e & \end{matrix}\middle| 1\right] = \frac{(d-a)_{|c|}(d-b)_{|c|}}{d_{|c|}(d-a-b)_{|c|}}.$$

(IV) Dixon's identity

In prettier and easier-to-remember form this identity reads as

$$\sum_k (-1)^k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+a}{c+k} = \frac{(a+b+c)!}{a!b!c!}.$$

Translated into formal hypergeometric language it becomes the statement that if $1+\frac{a}{2}-b-c$ has positive real part, and if $d=a-b+1$ and $e=a-c+1$ then

$${}_3F_2\left[\begin{matrix} a & b & c \\ d & e & \end{matrix}\middle| 1\right] = \frac{(\frac{a}{2})!}{a!}\frac{(a-b)!}{(\frac{a}{2}-b)!}\frac{(a-c)!}{(\frac{a}{2}-c)!}\frac{(\frac{a}{2}-b-c)!}{(a-b-c)!}.$$

(V) Clausen's ${}_4F_3$ identity

If d is a nonpositive integer and $a+b+c-d=\frac{1}{2}$, and $e=a+b+\frac{1}{2}$, and $a+f=d+1=b+g$ then

$${}_4F_3\left[\begin{matrix} a & b & c & d \\ e & f & g & \end{matrix}\middle| 1\right] = \frac{(2a)_{|d|}(a+b)_{|d|}(2b)_{|d|}}{(2a+2b)_{|d|}a_{|d|}b_{|d|}}.$$

(VI) Dougall's ${}_7F_6$ identity

If $n+2a_1+1=a_2+a_3+a_4+a_5$ and

$$a_6 = 1+\frac{a_1}{2}; a_7=-n; \text{ and } b_i = 1+a_1-a_{i+1} \quad (i=1,\ldots,6),$$

then

$$
{}_7F_6\left[\begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & \end{matrix}\,\middle|\, 1\right] = \\
\frac{(a_1+1)_n(a_1-a_2-a_3+1)_n(a_1-a_2-a_4+1)_n(a_1-a_3-a_4+1)_n}{(a_1-a_2+1)_n(a_1-a_3+1)_n(a_1-a_4+1)_n(a_1-a_2-a_3-a_4+1)_n}.
$$

Let's review where we are. In the previous chapter we saw how to take a sum and to identify it, when possible, as a standard hypergeometric sum. In this chapter we have seen a list of most of the important hypergeometric sums that can be expressed in simple, closed form. We will now give a few examples of the whole process whereby one uses the hypergeometric database in order to try to "do" a given sum. The strengths and the limitations of the procedure should then be clearer.

Example 1 For a nonnegative integer n, consider the sum

$$
f(n) = \sum_k (-1)^k \binom{2n}{k}^2.
$$

Can this sum be evaluated in some simple form?

Here and in the sequel we make the convention that if the range of the summation variable is not explicitly given, then it is understood to be the set of all integers. In such cases the summand will usually vanish outside of some subset, often a finite subset, of the integers, and that subset will be the effective range of the summation. In this example the effective range of k is the interval $0 \le k \le 2n$.

The first step is to identify the sum $f(n)$ as a particular hypergeometric series. For that purpose we look at the term ratio

$$
\frac{t_{k+1}}{t_k} = \frac{(-1)^{k+1}\binom{2n}{k+1}^2}{(-1)^k\binom{2n}{k}^2} = -\frac{(k-2n)^2}{(k+1)^2}.
$$

Our series is thereby unmasked: it is

$$
f(n) = {}_2F_1\left[\begin{matrix} -2n & -2n \\ 1 & \end{matrix}\,\middle|\, -1\right].
$$

Next we check the database to see if we have any information about ${}_2F_1$'s whose argument is -1, and indeed *there is* such an entry in the database, namely Kummer's identity. If we use it in the form that is given there for negative integer b, then it tells us that our unknown sum $f(n)$ is

$$
f(n) = {}_2F_1\left[\begin{matrix} -2n & -2n \\ 1 & \end{matrix}\,\middle|\, -1\right] = \frac{(2n-1)!2(-1)^n}{(n-1)!n!} = (-1)^n\binom{2n}{n}
$$

which is a happy ending indeed.

Example 2 For one more example of a lookup in the hypergeometric database, consider the sum

$$
f(n) = \sum_k (-1)^k \binom{2n}{k}\binom{2k}{k}\binom{4n-2k}{2n-k}.
$$

The first step is, as always, to find the term ratio and resolve it into linear factors. We find that

$$
\frac{t_{k+1}}{t_k} = \frac{(-1)^{k+1}\binom{2n}{k+1}\binom{2k+2}{k+1}\binom{4n-2k-2}{2n-k-1}}{(-1)^k\binom{2n}{k}\binom{2k}{k}\binom{4n-2k}{2n-k}} = \frac{(k+\frac{1}{2})(k-2n)^2}{(k+1)^2(k-2n+\frac{1}{2})}.
$$

The first term of our sum is not 1, but is instead $\binom{4n}{2n}$, hence we now know that our desired sum is

$$
f(n) = \binom{4n}{2n}{}_3F_2\left[\begin{matrix} -2n & -2n & \frac{1}{2} \\ 1 & -2n+\frac{1}{2} & \end{matrix}\,\middle|\, 1\right].
$$

Since this is a ${}_3F_2$, we check the possibilities of Saalschütz's identity and Dixon's identity. It does not match the condition $d+e=a+b+c+1$ of Saalschütz, so we try Dixon's next. If we put $a=-2n$, $b=-2n$, $c=\frac{1}{2}$, $d=1$ and $e=-2n+\frac{1}{2}$ then we find that the conditions of Dixon's identity, namely that $d=a-b+1$ and $e=a-c+1$ are met. If we now use the right hand side of Dixon's identity, we find that if n is a nonnegative integer

$$f(n)=\binom{4n}{2n}\frac{(-n)!(-2n-\frac{1}{2})!(n-\frac{1}{2})!}{(-2n)!n!(-n-\frac{1}{2})!(-\frac{1}{2})!}.$$

This is a rather distressing development. We were expecting an answer in simple form, but the answer that we are looking at contains some factorials of negative numbers, and some of these negative numbers are negative *integers*, which are precisely the places where the factorial function is undefined.

Fortunately, there is a ratio of two factorials at negative integers, and if we take an appropriate limit the singularities will cancel and a pleasant limiting ratio will ensue. We will now do this calculation, urging the reader to take note of the fact that this kind of situation happens fairly frequently when one uses the database. The answers are formally correct, but we need some further analysis to transform them into readily useable form.

What about the ratio $(-n)!/(-2n)!$? Imagine, for a moment, that n is near a positive integer, but is not equal to a positive integer. Then we use the reflection formula for the Γ-function

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin \pi z}$$

once more, in the equivalent form

$$(-z)!=\frac{\pi}{(z-1)!\sin \pi z}.$$

When n is near, but not equal to, a positive integer we find that

$$\frac{(-n)!}{(-2n)!}=\frac{\pi}{(\sin n\pi)(n-1)!}\frac{(\sin 2n\pi)(2n-1)!}{\pi}=\frac{2(2n-1)!\cos n\pi}{(n-1)!}.$$

Thus as n approaches a positive integer, we have found that

$$\frac{(-n)!}{(-2n)!}\longrightarrow(-1)^n\frac{(2n)!}{n!}.$$

Our answer now has become

$$f(n)=\binom{4n}{2n}\frac{(-1)^n(2n)!(-2n-\frac{1}{2})!(n-\frac{1}{2})!}{n!^2(-n-\frac{1}{2})!(-\frac{1}{2})!}.$$

A similar argument shows that

$$\frac{(-2n-\frac{1}{2})!}{(-n-\frac{1}{2})!}=\frac{(-1)^n(n-\frac{1}{2})!}{(2n-\frac{1}{2})!},$$

which means that now

$$f(n)=\binom{4n}{2n}\binom{2n}{n}\frac{(n-\frac{1}{2})!^2}{(2n-\frac{1}{2})!(-\frac{1}{2})!}.$$

But for every positive integer m,

$$\begin{aligned}(m-\frac{1}{2})!&=(m-\frac{1}{2})(m-\frac{3}{2})\cdots(\frac{1}{2})(-\frac{1}{2})!\\&=\frac{(2m-1)(2m-3)\cdots 1}{2^m}(-\frac{1}{2})!\\&=\frac{(2m)!}{2^m m!}(-\frac{1}{2})!.\end{aligned}$$

So we can simplify our answer all the way down to $f(n) = \binom{2n}{n}^2$. The fruit of our labor is that we have found the identity

$$\sum_k (-1)^k \binom{2n}{k}\binom{2k}{k}\binom{4n-2k}{2n-k} = \binom{2n}{n}^2,$$

and we realize that it is a special case of Dixon's identity, and we further realize that the "lookup" in the database was not quite a routine matter!

The moral of the story: There *is* a hypergeometric database, and most binomial coefficient sums that are evaluable in simple form can be done by expressing the sum as a standard hypergeomtric series and looking up the series. This process is, like any other, an *algorithm* for doing hypergeometric sums, and it should be assessed the same way as other algorithms. How effective is it? Precisely when can we expect a pretty answer? How fast is it? What is the *complete* algorithm, including the simplifications at the end, and how costly are they? What are the alternatives?

In the sequel we will present another algorithm for doing hypergeometric sums in simple form. That algorithm can be rather easily programmed for a computer, and it works under conditions that are wider than those of the database lookup because for the latter one can never be sure that the database is complete, and under conditions that can, furthermore, be stated fairly clearly.

Zeilberger's Paradigm

To prove an identity

$$\sum_k \text{summand}(n,k) = \text{answer}(n)\ :$$

A. Find a recurrence relation that is satisfied by the sum on the left side.

B. Show that answer(n) satisfies the same recurrence, by substitution, and that enough corresponding initial values of both sides are equal to each other.

(III) Look for the recurrence!

Doron Zeilberger suggested another way to discover and to prove identities. If we wish to evaluate the sum (from chapter 2)

$$f(n) = \sum_k (-1)^k \binom{2n}{k}\binom{2k}{k}\binom{4n-2k}{2n-k},$$

suppose we could somehow (?) discover and prove that the sum $f(n)$ satisfies the recurrence relation

$$(n+1)^2 f(n+1) - 4(2n+1)^2 f(n) = 0 \qquad (n \geq 0; f(0) = 1).$$

That would be a boon indeed, for then we would have

$$f(n+1) = \frac{4(2n+1)^2}{(n+1)^2} f(n) \qquad (n \geq 0; f(0) = 1),$$

and we would quickly see that $f(n) = \binom{2n}{n}^2$, and prove it by induction on n.

Let's deal first with the question of how to find the recurrence relation, and when it exists.

The idea that one can *algorithmically*, i.e., by following a set recipe and doing no creative thinking, discover a recurrence relation that is satisfied by a hypergeometric sum, is due to Sister Mary Celine Fasenmyer, in her Michigan Ph.D. thesis of 1944. Sister Celine used her method to discover pure recurrence relations that were satisfied by sequences of hypergeometric polynomials. Doron Zeilberger recognized that her method could be made the basis for automated discovery and proof of summation identities according to the paradigm mentioned above.

We begin by illustrating her method on a simple sum. Let

$$f(n) = \sum_k k\binom{n}{k} \qquad (n = 0, 1, 2, \ldots),$$

and let's look for the recurrence that $f(n)$ satisfies. To do this we first look for the recurrence that the *summand*

$$F(n,k) = k\binom{n}{k}$$

satisfies. It is a function of two variables (n, k), so we try to find a recurrence of the form

$$a(n)F(n,k) + b(n)F(n+1,k) + c(n)F(n,k+1) + d(n)F(n+1,k+1) = 0, \tag{3.1}$$

in which the coefficients a, b, c, d *depend on n only, and not on k* (for reasons that will become clear presently).

To find the coefficients, if they exist, we divide (3.1) through by $F(n, k)$ getting

$$a + b\frac{F(n+1,k)}{F(n,k)} + c\frac{F(n,k+1)}{F(n,k)} + d\frac{F(n+1,k+1)}{F(n,k)} = 0.$$

Now each of the ratios of F's is a certain rational function. If we carry out the indicated divisions, our assumed recurrence becomes

$$a + b\frac{n+1}{n+1-k} + c\frac{n-k}{k} + d\frac{n+1}{k} = 0,$$

in which the factorials have all disappeared, and we see only rational functions of n and k.

The next step, following Sister Celine, is to put the whole thing over a common denominator. That denominator is $k(n+1-k)$ and the numerator is, after collecting by powers of k,

$$(d+(c+2d)n+(c+d)n^2)+(a+b-c-d+(a+b-2c-d)n)k+(c-a)k^2.$$

Our assumed recurrence will be true if this numerator identically vanishes, for all n and k, with the coefficients a,b,c,d being permitted to depend only on n. Thus the coefficient of each power of k must vanish. This gives us a system of three equations in four unknowns, namely

$$\begin{bmatrix} 0 & 0 & n(n+1) & (n+1)^2 \\ n+1 & n+1 & -2n-1 & -(n+1) \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

to solve for a,b,c,d.

Success in finding a nontrivial solution is now guaranteed simply because there are more unknowns than there are equations. If we actually solve these equations we find that

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = d\,[\,-1-1/n, 0, -1-1/n, 1\,].$$

We now substitute these values into the assumed form of the recurrence relation (3.1), and we have the desired "k-free" recurrence for the summand $F(n,k)$, namely

$$-(1+\frac{1}{n})F(n,k)-(1+\frac{1}{n})F(n,k+1)+F(n+1,k+1)=0. \tag{3.2}$$

From the recurrence for the *summand* $F(n,k)$ to the recurrence for the *sum* $f(n)$ is a very short step: just sum (3.2) over *all integers* k, noticing appreciatively that the coefficients in the recurrence are free of k's, so the summation over k can operate directly on the F in each term. We get instantly the recurrence

$$-(1+\frac{1}{n})f(n)-(1+\frac{1}{n})f(n)+f(n+1)=0,$$

i.e., the recurrence

$$f(n+1)=2\frac{n+1}{n}f(n) \qquad (n=1,2,\ldots;f(1)=1).$$

We can now easily find $f(n)$, the desired sum, since

$$f(n+1)=2\frac{n+1}{n}f(n)=2^2\frac{n+1}{n}\frac{n}{n-1}f(n-1)=\cdots=2^n(n+1)f(1),$$

and so

$$f(n)=n2^{n-1}$$

for all $n \geq 0$. ∎

Yes, this was a very simple example, but it illustrated many of the points of interest. The method works because one can prove (and we will!) that the number of unknowns can always be made larger than the number of equations, so a nontrivial solution must exist. The fact that the coefficients in the assumed recurrence are free of k is vital to the step in which we sum the F recurrence over all integers k.

The method does require some computation, and in some cases it requires a very large amount of computation. But that is what computers are for, and they do their job well on these recurrence relation problems.

(IV) Sister Mary Celine Fasenmyer

photo here

Sister Celine was born in Crown, in central Pennsylvania, October 4, 1906. Her parents were George and Cecilia Fasenmyer, though her mother, Cecilia, died when Mary was one year old. Her father worked his own oil lease in the area. He remarried three years later, a woman, Josephine, who was twenty five years his junior.

Her early education was at the St. Joseph's Academy in Titusville, PA, from which she was graduated in 1923, having been always "good in math." She then taught for ten years, and in 1933 received her AB degree from Mercyhurst College. She was sent to Pittsburgh by her order, to teach in the St. Justin School and to go to the University of Pittsburgh for her MA degree, which she received in 1937. Her major was mathematics and her minor was in physics.

The community told her to go to the University of Michigan for her doctorate, which she did from the fall of 1942 until June of 1946, when she received her degree. Her thesis was written under the direction of Earl Rainville, whom she remembers as having been quite accessible and helpful, as well as in a subject area which she liked. In it she showed how one can deduce recurrence relations that are satisfied by sums of hypergeometric terms, in a purely mechanical ("algorithmic") way. She used the method in her thesis to find pure recurrence relations that are satisfied by various hypergeometric polynomial sequences. In two later papers she developed the method further, and explained its workings to a broad audience in her paper [3]. For an exposition of her thesis results see [4].

Her method is the intellectual progenitor of the computerized methods that we use today to prove hypergeometric identities, thanks to the recognition by Doron Zeilberger that her method can be adapted to prove such identities.

Publications of Sister Mary Celine Fasenmyer

1. Ph.D. Dissertation, University of Michigan, 1945.
2. Some generalized hypergeometric polynomials, Bull. Amer. Math. Soc., **53** (1947), 806-812.
3. A note on pure recurrence relations, Amer. Math. Monthly, **56** (1949), 14-17.

See also Chapters 14 and 18 in

4. Earl D. Rainville, Special Functions, Macmillan, 1960.

(V) Algorithms and Programs

There is perhaps no better way to underscore the algorithmic nature of the foregoing than to give a complete computer program that will carry it out. We will do that here, giving two programs, one in Maple and one in Mathematica, for the same purpose. We note that these are not sophisticated programs, but they will suffice for the reader to carry out some enjoyable experiments. More powerful programs will be described later in this work.

Input to these programs are first, the summand $F(n,k)$, given as a function, and second, the orders of the recurrence equation that is being sought. The form of the recurrence that is found will be

$$\sum_{i=0}^{ii}\sum_{j=0}^{jj} a_{ij}(n)F(n-j,k-i)=0. \tag{5.1}$$

In these terms, input to the programs consists of F, ii, and jj.

The programs follow the method of chapter (III). Just to review, this method consists of the following steps:

A. Assume the recurrence formula in the form (5.1), with the coefficients $a_{ij}(n)$ to be determined, if possible.

B. Divide each term of (5.1) by $F(n,k)$, and reduce each ratio $F(n-j,k-i)/F(n,k)$ by simplifying the ratios of the factorials that it contains, so that only rational functions of n and k remain.

C. Place the entire expression over a single common denominator. Then collect the numerator as a polynomial in k.

D. Solve the system of linear equations that results from equating to zero the coefficients of each power of k in the numerator polynomial, for the unknown coefficients a_{ij}. If the system has no solution, try the whole thing again with larger values of ii and/or jj. That is, look for a bigger recurrence. ∎

The following Maple program will do the job. The job, however, can be very time-consuming if the $F(n,k)$ is at all complicated.

```
recur:=proc(f,ii,jj)
local m,j,l,i,tt,yy,zz,rr,xx,ss;
# Given f(n,k), the program finds the k-free recurrence for f
# if one of orders (ii,jj) exists.
# Place input at the end, then give `read` instruction in Maple
# In input, do not raise products of factorials to powers;
# instead raise each factorial to the power.
m:='m':j:='j':l:='l':i:='i':
iii:=ii*(jj+1)+jj:
xx:=seq(b[i],i=0..iii):i:='i':
ss:=numer(simplify(expand(sum(sum(b[i*(jj+1)+j]*f(n-j,k-i)/f(n,k),
i=0..ii),j=0..jj)))):
tt:=coeffs(collect(ss,k),k):
yy:=solve({tt},{xx}):
zz:=simplify(sum(sum(b[l*(jj+1)+m]*F(n-m,k-l),l=0..ii),m=0..jj)):
i:='i':j:='j':lprint(` ` ):lprint(` `):
lprint(`The full recurrence is`):
lprint(` `):
zz:=numer(simplify(subs(yy,zz))):
for i from 0 to ii do for j from 0 to jj do
zz:=collect(zz,F(n-j,k-i)) od od:
lprint(zz,`==0`);
end:
# Place input function definition f on the next line,
# with values of ii, jj, the trial orders of the recurrence.
f:=(n,k)->k*n!/(k!*(n-k)!):ii:=1:jj:=1:
recur(f,ii,jj);
```

To use the Maple program above, first store the program in a file whose name is, say, `recurrence`. As stored in the file, the definition of your summand function f and the sizes of the recurrence that you are looking for (ii and jj) should appear in the next-to-last line, as shown above for the example $F(n,k) = k\binom{n}{k}$, with $ii = 1$ and $jj = 1$, which corresponds to the problem that we worked out in full in chapter (III). After storing the file, then get into Maple and give the instruction `` read `recurrence`; ``.

In the case of the example shown above, the output that is given is the following:

```
The full recurrence is
 b[3] n F(n-1,k) - (n-1) b[3] F(n,k) + b[3] F(n-1,k-1) n == 0
```

In this output `b[3]` is an arbitrary constant, which can be ignored. The recurrence is identical with the one we had previously found by hand, in (3.2), as can be seen by replacing n and k by $n-1$ and $k-1$ in (3.2) and comparing with the output above.

Suppose we were to use the program with the input $f(n,k) = \binom{n}{k}$. Then what recurrence would it find? If you guessed the Pascal triangle recurrence $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$, then you're right. In that same spirit, let's look for a recurrence that $f(n,k) = \binom{n}{k}^2$ satisfies.

So we take the Maple program above and we replace the "input" line by

```
f:=(n,k)->(n!/(k!*(n-k)!))^2:ii=2:jj=2:
```

It runs for a while, and then announces that

```
The full recurrence is
(n - 1) b[8] F(n - 2, k - 2) + b[8] (2 - 2 n) F(n - 2, k - 1)
+ (- 2 n + 1) b[8] F(n - 1, k - 1) + (n - 1) b[8] F(n - 2, k)
+ (- 2 n + 1) b[8] F(n - 1, k) + b[8] n F(n, k) == 0
```

Translated into conventional mathematical notation, this recurrence reads as

$$\begin{aligned} n\binom{n}{k}^2 - (2n-1)\left\{\binom{n-1}{k}^2 + \binom{n-1}{k-1}^2\right\} \\ + (n-1)\left\{\binom{n-2}{k}^2 - 2\binom{n-2}{k-1}^2 + \binom{n-2}{k-2}^2\right\} = 0, \end{aligned} \tag{5.2}$$

which is the "Pascal triangle" identity for the squares of the binomial coefficients.

Let's use the recurrence for the squares to find a recurrence for the sum of the squares. So let

$$f(n) = \sum_k \binom{n}{k}^2.$$

What is the recurrence for $f(n)$? To find it, just sum (5.2) over all integers k, obtaining

$$nf(n) - (2n-1)\{f(n-1) + f(n-1)\} + (n-1)\{f(n-2) - 2f(n-2) + f(n-2)\} = 0$$

which boils down to just

$$f(n) = \frac{2(2n-1)}{n} f(n-1) = \frac{2^2(2n-1)(2n-3)}{n(n-1)} f(n-2) = \cdots = \frac{(2n)!}{n!^2}.$$

We therefore have another illustration of how the computer can discover the evaluation in closed form of a hypergeometric sum.

The cubes and higher powers of the binomial coefficients also satisfy recurrences of this kind, but finding them with this program would require either an immense computer or immense patience. The *sums* of the cubes, for instance, of the binomial coefficients also satisfy a recurrence, which the method would discover. That recurrence is of second order, however, and its solution *provably* can not be expressed as a linear combination of a constant number of hypergeometric terms. Hence in the case of the sum of the cubes, what we get is a computer generated proof of the impossibility of finding a pleasant evaluation, with a reasonable definition of "pleasant." The theory of this remarkable chain of recent developments will be fully explained in the sequel.

Next let's do the same job in Mathematica. The first thing to do is to read in the package `DiscreteMath‘RSolve‘` in order to enable the `FactorialSimplify` instruction, which simplifies ratios of factorials. Next we define a Module that finds a recurrence relation satisfied by a given function `f`, the recurrence being of orders `ii, jj`.

```
<<DiscreteMath‘RSolve‘
findrecur[f_ ,ii_ ,jj_ ]:=
   Module[{yy,zz,ll,tt,uu,r,s,i,j},
     yy=Sum[Sum[a[i,j] FactorialSimplify[f[n-j,k-i]/f[n,k]],
        {i,0,ii}],{j,0,jj}];
     zz=Collect[Numerator[Together[yy]],k];
     ll=CoefficientList[zz,k];
     tt=Flatten[Table[a[i,j],{i,0,ii},{j,0,jj}]];
     uu=Flatten[Simplify[Solve[ll==0,tt]]];
     For[r=0,r<=ii,r++,
       For[s=0,s<=jj,s++,
         a[r,s]=Replace[a[r,s],uu]]];
     Sum[Sum[a[i,j] F[n-j,k-i],{i,0,ii}],{j,0,jj}]==0]
```

After that there's nothing to do but define the function

```
f[n_ ,k_ ]:=k n!/(k!(n-k)!)
```
and call the Module,
```
findrecur[f,1,1]
```
The resulting output is
```
a[1, 1] F[-1 + n, -1 + k] + a[1, 1] F[-1 + n, k] + (-1+1/n) a[1,1] F[n,k] ==0,
```
as before.

The reader is cautioned that this program is slower than its Maple counterpart, in its execution, and it should not be tried on recurrence relations of larger span.

Next let's try an example in the spirit of Sister Celine's original use of the recurrence method. We'll look for a recurrence that is satisfied by the classical Laguerre polynomials

$$L_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{x^k}{k!} \qquad (n = 0, 1, 2, \ldots).$$

The first step is to find a 2-variable recurrence that is satisfied by the summand itself, in this case by

$$F(n,k) = (-1)^k \binom{n}{k} \frac{x^k}{k!} \qquad (n = 0, 1, 2, \ldots). \tag{5.2}$$

The fundamental theorem guarantees that this F satisfies a recurrence. Before we go to the computer to find the recurrence, let's try to estimate its order in advance. To do this, identify the specific F in (5.2) with the general form (6.1) by taking $uu := 1$, $vv := 3$, $x := -x$, $(a_1, b_1, c_1) = (1, 0, 0)$ and for the three (u, v, w) vectors,

$$(0, 1, 0), (1, -1, 0), (0, 1, 0).$$

Then for the quantitiative estimates provided by the theorem, namely the (I^*, J^*) of eq. (6.3), we find $J^* = 3$, $I^* = 4$. Hence there is surely a recurrence for $F(n,k)$ of the form

$$\sum_{i=0}^{4} \sum_{j=0}^{3} a_{i,j}(n) F(n-j, k-i) = 0,$$

in which the $a_{i,j}$'s are polynomials in n, and are not all zero.

To find such a recurrence we use the Maple program of chapter 5. To search for a recurrence of orders 2, the line of Maple input will be

```
f:=(n,k)->(-1)^k*n!*x^k/(k!^2*(n-k)!):ii:=2:jj:=2:
```

The program runs briefly, and returns the following output:
```
The full recurrence is
(-b[3]*x+b[3]*x*n)*F(n-2,k-1)+(n**2*b[4]-n*b[4]+b[3]+2*b[3]*n**2
 -3*b[3]*n)*F(n-2,k)+(-2*n**2*b[4]+4*b[3]*n-b[3]+n*b[4]
 -4*b[3]*n**2)*F(n-1,k)+(n**2*b[4]+2*b[3]*n**2-b[3]*n)*F(n,k)
 +b[4]*F(n-1,k-1)*x*n+b[3]*x**2*F(n-1,k-2)+b[3]*F(n,k-1)*x*n==0
```

in which $b[3]$, $b[4]$ are arbitrary constants.

But the recurrence for $F(n,k)$ wasn't the object of the exercise. What we wanted was a recurrence for the Laguerre polynomial

$$L_n(x) = \sum_k F(n,k),$$

in which the sum on k is over *all* integers, i.e., extends from $-\infty$ to $+\infty$. This point is extremely important. The summand $F(n,k)$ of (7.1) will vanish automatically if $k<0$ or if $k>n$, i.e., it has *compact support*. Hence even if we sum over all integers, the sums will contain only finitely many nonvanishing terms.

Therefore, we can sum the output recurrence over *all integers* k. Further, since the constants $b[3], b[4]$ are arbitrary, let's take $b[3]=0$ and $b[4]=1$. This yields

$$(n^2-n)L_{n-2}(x)+(-2n^2+n)L_{n-1}(x)+n^2L_n(x)+nxL_{n-1}(x)=0$$

or finally

$$2nL_n(x)+(2x+1-2n)L_{n-1}(x)+(n-1)L_{n-2}(x)=0,$$

which is the classical three term recurrence relation for the Laguerre polynomials. ∎

It should be noted that while the Laguerre polynomials make a pleasant example, the main theorem assures us that every sequence of polynomials $\sum_k F(n,k)$, where F is a proper hypergeometric term, satisfies a recurrence relation, and it even gives us a bound on the order of the recurrence.

(VI) The fundamental theorem

The fundamental theorem states that every proper hypergeometric term $F(n,k)$ satisfies a recurrence relation of the kind we have found in the previous chapters, and it validates the procedure that we have used to find these recurrences in the sense that it guarantees that Sister Celine's method will work if the span of the assumed recurrence is large enough. The theorem also finds explicit precomputable upper bounds on the span.

Definition. *A function $F(n,k)$ is said to be a proper hypergeometric term if it can be written in the form*

$$F(n,k) = P(n,k)\frac{\prod_{i=1}^{uu}(a_i n + b_i k + c_i)!}{\prod_{i=1}^{vv}(u_i n + v_i k + w_i)!}x^k, \tag{6.1}$$

in which x is an indeterminate over, say, the complex numbers, and

1. *P is a polynomial, and*
2. *the a's, b's, u's, v's are specific integers, that is to say, they do not contain any additional parameters, and*
3. *the quantities uu and vv are finite, nonnegative, specific integers, and*
4. *the c's and the w's are integers, but they may contain additional parameters, that is, the validity of any results obtained will be restricted to parameter values for which the c's and w's are integers.*

An F of the form (6.1) is *well defined* at a point (n,k) if none of the numbers $\{a_i n + b_i k + c_i\}_{i=1}^{uu}$ is a negative integer. We will say that $F(n,k) = 0$ if F is well-defined at (n,k) and at least one of the numbers $\{u_i n + v_i k + w_i\}_{i=1}^{vv}$ is a negative integer, or $P(n,k) = 0$.

Some examples of proper hypergeometric terms are as follows.

$\binom{n}{k}2^k$ is proper hypergeometric because it can be written

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

which is exactly of the required form. Also $\binom{2n}{k}2^k$ is proper hypergeometric, but $\binom{an}{n}$ is not, if a is an unspecified parameter.

Consider, however, $F(n,k) = 1/(n+3k+1)$. This is not in proper hypergeometric form. It doesn't contain any of the factorials, and it isn't a polynomial. *However*, the definition says "... if it can be written in the form ..." This example $F(n,k)$ can be written in proper hypergeometric form, even though it was not given to us in that form! All we have to do is to write

$$\frac{1}{n+3k+1} = \frac{(n+3k)!}{(n+3k+1)!}.$$

On the other hand, if we take $F(n,k) = 1/(n^2+k^2+1)$ then no amount of rewriting will produce the form (6.1), so this f is not proper hypergeometric.

Now we can state the main theorem.

Theorem. *Let $F(n,k)$ be a proper hypergeometric term. Then F satisfies a k-free recurrence relation. That is to say, there exist positive integers I, J, and polynomials $a_{i,j}(n)$ for $i = 0, \ldots, I; j = 0, \ldots, J$, not all zero, such that the recurrence*

$$\sum_{i=0}^{I}\sum_{j=0}^{J} a_{i,j}(n)F(n-j,k-i) = 0 \tag{6.2}$$

holds at every point (n,k) at which $F(n,k) \neq 0$ and all of the values of F that occur in (6.2) are well-defined. Furthermore there is such a recurrence with $(I,J) = (I^*, J^*)$ where

$$J^* = \sum_s |b_s| + \sum_s |v_s|; \quad I^* = 1 + \deg(P) + J^*(\{\sum_s |a_s| + \sum_s |u_s|\} - 1). \tag{6.3}$$

The recurrence (6.2) is called *k-free* because the coefficients $a_{i,j}(n)$ may depend only on n, not on k.

Next we're going to prove the theorem. In order to do that it will be important to do a few simple exercises that relate to the behavior of *translates of a proper hypergeometric term.*

Suppose $f(n) = (2n+3)!$. What is $f(n-2)/f(n)$? it is

$$\frac{f(n-2)}{f(n)} = \frac{1}{4n(1+n)(1+2n)(3+2n)},$$

i.e., it is the reciprocal of a certain polynomial in n.

On the other hand if $f(n) = (3-2n)!$ then

$$\frac{f(n-2)}{f(n)} = 4(n-3)(n-2)(2n-7)(2n-5)$$

is a polynomial in n.

The conclusion here is that if $f(n) = (an+b)!$, then the ratio $f(n-j)/f(n)$ is, for $j \geq 0$, a polynomial in n, if $a \leq 0$, and the reciprocal of a polynomial in n, if $a > 0$.

For a two-variable case, consider $f(n,k) = (2n-3k+1)!$. Then

$$\frac{f(n-9,k-5)}{f(n,k)} = \frac{1}{(-1-3k+2n)(-3k+2n)(1-3k+2n)},$$

is the reciprocal of a polynomial in n and k, whereas

$$\frac{f(n-6,k-5)}{f(n,k)} = (2-3k+2n)(3-3k+2n)(4-3k+2n)$$

is a polynomial in n and k.

The general rule in the two-variable case is that if $F(n,k) = (an+bk+c)!$ then for $i,j \geq 0$ we have

$$\frac{F(n-j,k-i)}{F(n,k)} = \\ \begin{cases} \{(an+bk+c)\cdots(an+bk+c-aj-bi+1)\}^{-1}, & \text{if } aj+bi \geq 0; \\ (an+bk+c+|aj+bi|)\cdots(an+bk+c+1), & \text{if } aj+bi < 0. \end{cases} \tag{6.4}$$

Once again the result is either a polynomial in n and k, or is the reciprocal of such a polynomial, depending on the sign of $aj+bi$.

Let's introduce the following notation for the rising factorial (rf) and falling factorial (ff) polynomials, for nonnegative integer values of x (the empty product is =1):

$$\mathrm{rf}(x,y) = \prod_{j=1}^{x}(y+j),$$

$$\mathrm{ff}(x,y) = \prod_{j=0}^{x-1}(y-j).$$

In terms of these polynomials, we can rewrite (6.4) as

$$\frac{F(n-j,k-i)}{F(n,k)} = \begin{cases} 1/\mathrm{ff}(aj+bi, an+bk+c), & \text{if } aj+bi \geq 0; \\ \mathrm{rf}(|aj+bi|, an+bk+c) & \text{if } aj+bi < 0. \end{cases} \tag{6.5}$$

Now consider a function $F(n,k)$ which is not just a single factor $(an+bk+c)!$, but is a product of several such factors divided by another product of several such factors, as in (6.1) above,

$$F(n,k) = P(n,k)\frac{\prod_{s=1}^{uu}(a_s n + b_s k + c_s)!}{\prod_{s=1}^{vv}(u_s n + v_s k + w_s)!}x^k.$$

What can we say about the form of the ratio $\rho = F(n-j,k-i)/F(n,k)$ now? Well, each of the factorial factors in the numerator of F contributes a polynomial in n and k either to the numerator of ρ or to the denominator of ρ, as in (6.5). Hence, the ratio ρ will be a rational function of n and k, say $\nu(n,k)/\delta(n,k)$.

More precisely, the numerator $\nu(n,k)$ of the ratio ρ will be, according to (6.5),

$$P(n-j,k-i) \prod_{\substack{s=1 \\ a_s j + b_s i < 0}}^{uu} \mathrm{rf}(|a_s j + b_s i|, a_s n + b_s k + c_s) \prod_{\substack{s=1 \\ u_s j + b_s i \geq 0}}^{vv} \mathrm{ff}(u_s j + v_s i, u_s n + v_s k + w_s), \tag{6.6}$$

and its denominator $\delta(n,k)$ will be

$$P(n,k)x^i \prod_{\substack{s=1 \\ a_s j + b_s i \geq 0}}^{uu} \mathrm{ff}(a_s j + b_s i, a_s n + b_s k + c_s) \prod_{\substack{s=1 \\ u_s j + v_s i < 0}}^{vv} \mathrm{rf}(|u_s j + v_s i|, u_s n + v_s k + w_s). \tag{6.7}$$

OK, now to prove the theorem, let's assume the recurrence in the form (6.2) and try to solve for the coefficients $a_{i,j}(n)$. If we do that, then after dividing by $F(n,k)$, the left side of the assumed recurrence will be

$$\sum_{0 \leq i \leq I; 0 \leq j \leq J} a_{i,j}(n)\frac{\nu_{i,j}}{\delta_{i,j}}, \tag{6.8}$$

where each $\nu_{i,j}$ looks like (6.6) and each $\delta_{i,j}$ looks like (6.7).

The next step is to collect all of the terms in the sum (6.8) over a single least common denominator. But since we have such explicit formulas for the numerator and denominator polynomials of each term, we can write out explicitly what that common denominator will be.

The first thing to notice is that, by (6.7), each and every denominator in (6.8) contains the same factor $P(n,k)$, so $P(n,k)$ will be in the least common denominator that we are constructing.

(Notice that if we had permitted a denominator polynomial $Q(n,k)$ to appear in the definition (6.1) of a proper hypergeometric term, then the outlook would have been much bleaker. Indeed, in that case, the (i,j) term in (6.8) would have contained a factor $Q(n\text{-}j,k\text{-}i)$ also, and we would need to deal with the least common multiple of all such factors, when constructing the common denominator of all terms. That multiple would have been of an unacceptably high degree in k, and would have blocked the argument that follows from reaching a successful conclusion.)

We introduce the symbol $x^+ = \max(x,0)$, where x is a real number. Then for all real numbers a,b we have

$$\max\{|aj+bi| : aj+bi < 0; 0 \leq i \leq I; 0 \leq j \leq J\} = (-a)^+J + (-b)^+I,$$

and

$$\max\{aj+bi : aj+bi \geq 0; 0 \leq i \leq I; 0 \leq j \leq J\} = a^+J + b^+I,$$

so the 'x^+' notation is a device that saves the enumeration of many different cases.

Now we can address the question of finding the least common multiple of all of the $\delta_{i,j}$'s in (6.8). For each s, a common multiple of all of the *falling* factorials that appear there will be the one whose first argument is largest, i.e.,

$$\mathrm{ff}((a_s)^+J + (b_s)^+I, a_sn + b_sk + c_s),$$

and a common multiple of all of the *rising* factorials that appear there will similarly be

$$\mathrm{rf}((-u_s)^+J + (-v_s)^+I, u_sn + v_sk + w_s).$$

Consequently the least common denominator of the expression (6.8), when that expression is thought of as a rational function of k, with n as a parameter, surely divides

$$P(n,k)\prod_{s=1}^{uu}\mathrm{ff}((a_s)^+J + (b_s)^+I, a_sn + b_sk + c_s)\prod_{s=1}^{vv}\mathrm{rf}((-u_s)^+J + (-v_s)^+I, u_sn + v_sk + w_s). \tag{6.9}$$

Therefore we can clear (6.8) of fractions if we multiply it through by (6.9). The result of multiplying (6.8) through by (6.9) will be the *polynomial in k*

$$\sum_{0\le i\le I;0\le j\le J} a_{i,j}(n)\nu_{i,j}(n,k)\frac{\Delta}{\delta_{i,j}(n,k)}, \tag{6.10}$$

in which Δ is the common denominator (6.9).

In order to prove the theorem we must show that if I and J are large enough, then the system of linear equations in the unknown $a_{i,j}$'s that one obtains by equating to zero the coefficient of every power of k that appears in (6.10) actually has a nontrivial solution. This will surely happen if the number of unknowns exceeds the number of equations, and we claim that if I, J are large enough then this is exactly what happens.

Indeed, the number of unknown $a_{i,j}$'s is obviously $(I+1)(J+1)$. The number of equations that they must satisfy is the number of different powers of k that appear in (6.10). We claim that the number of different powers of k that appear there grows only linearly with I and J, that is, as $c_1I + c_2J + c_3$, where the c's are independent of I, J. This claim would be sufficient to prove the theorem because then number of unknowns would grow like IJ, for large I and J, whereas the number of equations would grow only as $c_1I + c_2J + c_3$. Hence for large enough I, J the latter would be less than the former.

Since the degree in k of each rising factorial and each falling factorial that appears in (6.10) grows linearly with I, J, and there are only a fixed number of each of them, the degrees in k of all of the ν's, δ's and Δ grow linearly with I, J. Hence the claim is clearly true, and the proof of the main theorem is complete. A more detailed argument, which we omit here, shows that the values I^* and J^* that are in the statement of the theorem are already sufficiently large. ∎

Suppose $\sum_k F(n,k) = 1$ is an identity of the type that we have been considering. Then by the fundamental theorem, there exists an integer n_0 with the following property: suppose that we have *numerically* verified that the claimed identity is correct for $n = 0, 1, \ldots, n_0$. Then the identity is thereby *proved* to be true in general.

What is the integer n_0 that validates such a proof by computation? Once we have found the recurrence that the left side, $f(n)$ satisfies, suppose that recurrence turns out to be of order J. A function that satisfies a recurrence of order J, and which is equal to 1 for J consecutive values of n, has *not* been thereby proved to be 1 for all n.

The reason is that in the recurrence relation for f, say

$$a_0(n)f(n) + a_1(n)f(n-1) + \cdots + a_J(n)f(n-J) = 0$$

the coefficient $a_0(n)$ might vanish for some large values of n, and then we would not be able to solve for $f(n)$ from its predecessors and sustain the induction.

For example, if a certain sum $f(n)$ satisfies

$$(n-100)f(n) - nf(n-1) + 100f(n-2) = 0$$

for all $n \geq 2$, and if we start checking that $f(n) = 1$, numerically from its definition as a sum, beginning with $n = 0$, then we will have to check up to $n = 102$ before we can safely conclude that $f(n) = 1$ for all $n \geq 0$.

So a safe estimate for n_0 is the order of the recurrence plus the size of the largest nonnegative zero of the leading coefficient of the recurrence.

It is possible to estimate, *a priori*, the roots of the leading coefficient of the recurrence relation. This has been done in a PhD dissertation at the University of Pennsylvania by Lily Yen, in 1993. Together with the a priori bounds on the order of the recurrence relation that we have already discussed here, this means that we can estimate n_0 a priori also. Theoretically, then, we can prove identities just by checking enough numerical values! As things now stand, however, the a priori estimates of n_0 are extremely large, so the algorithm is impractical. It is an interesting research question to ask how small we can make this a priori estimate of n_0.

(VII) The limits of the sum

We'll begin here with one more illustration of the recurrence method at work, this time on a sequence of orthogonal polynomials. It will be noted that a very important role is played by the limits of summation in the sum that is under study.

Many hypergeometric summands have compact support. More precisely, for each fixed n, the summand $F(n,k)$ vanishes outside of a finite interval of k. For instance, the summand $F(n,k) = \binom{n}{k}$ vanishes outside of the interval $0 \le k \le n$. When the summand has this kind of natural support we may write the sum without any limits of summation, and it should then be understood that the summation index k runs over the full support.

Thus $\sum_k \binom{3n}{k}$ means

$$\sum_{k=0}^{3n} \binom{3n}{k}.$$

In the sum from Dixon's identity

$$\sum_k (-1)^k \binom{a+b}{a+k} \binom{b+c}{b+k} \binom{c+a}{c+k},$$

the index k actually runs over the interval

$$-\min\{a,b,c\} \le k \le \min\{a,b,c\}.$$

When this situation holds, it is reflected in the notation that we use to designate the sum, in that the limits of the summation may be omitted. When the summation index runs over the full support of the summand, we will say that the sum has its *natural limits.*

The benefits that we get from using the natural limits are considerable. If we are trying to learn about the sum $f(n) = \sum_k F(n,k)$, we find the recurrence formula that F satisfies, as guaranteed by the main theorem, and we use strongly the fact that it is k-free, i.e., that the coefficients depend only on n and not on k.

To spell this out in more detail, suppose that F has compact support and that the recurrence that we find for F is

$$\sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n) F(n-j, k-i) = 0. \tag{7.1}$$

Sum this equation over *all values of* k, from $-\infty$ to ∞. This gives

$$\begin{aligned} \sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n) \sum_k F(n-j, k-i) &= \sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n) f(n-j) \\ &= \sum_{j=0}^{J} b_j(n) f(n-j) = 0, \end{aligned} \tag{7.2}$$

where

$$b_j(n) = \sum_{i=0}^{I} a_{i,j}(n). \tag{7.3}$$

Thus *we have found a recurrence relation of order* J that is satisfied by the sum $f(n)$ that we are looking for!

Theorem. *The sum over the full support of* k*, of a proper hypergeometric summand* $F(n,k)$*, regarded as a function of the running index* n*, always satisfies a homogeneous linear recurrence relation of order* J *whose coefficients are polynomials in* n*.*

Example 1 Let $F(n,k) = \binom{n}{k}$, and put $f(n) = \sum_k F(n,k)$. From the recurrence for the summand

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k},$$

we find by summing over k that

$$f(n) = f(n-1) + f(n-1) = 2f(n-1),$$

is the recurrence for the sum.

Could this recurrence for the sums turn out to be trivial? That is, might it happen that even though not all of the $a_{i,j}(n)$'s can be 0, which the fundamental theorem has guaranteed, their sums $b_j(n)$ might all turn out to vanish identically?

Not if we take mild precautions. Suppose we multiply the nontrivial 2-variable recurrence (7.1) by x^k and sum over (all integer values of) k. We find that

$$\begin{aligned} \sum_k \sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n) F(n-j, k-i) x^k &= 0 \\ &= \sum_{j=0}^{J} A_j(x) \phi_{n-j}(x), \end{aligned} \tag{7.4}$$

in which the F-polynomials $\phi_n(x)$ are

$$\phi_n(x) = \sum_r F(n,r) x^r, \tag{7.5}$$

and the coefficients A_j are

$$A_j(x) = \sum_{i=0}^{I} a_{i,j}(n) x^i. \tag{7.6}$$

The recurrence (7.4) shows that the infinite sequence of polynomials $\{\phi_n(x)\}$ associated with the summand F is guaranteed to satisfy a recurrence of order J whose coefficients are $J+1$ polynomials in n and x.

Now suppose that the doomsday scenario really happens, and that all of the coefficients $b_j(n)$ in (7.3) all vanish identically. Then the polynomials $A_j(x)$ that are displayed in (7.6) all vanish simultaneously at $x = 1$, though these polynomials themselves cannot all vanish identically. Since they all vanish at $x = 1$, they are all divisible by $(x-1)$. In fact, choose the largest integer $r \geq 1$ with the property that all of the $A_{i,j}(x)$ are simultaneously divisible by $(x-1)^r$. Then take the recurrence (7.4) whose coefficients are the $A_{i,j}(x)$ and divide it through by $(x-1)^r$. We then obtain a new recurrence that is satisfied by the ϕ_n's in which it is now no longer true that all of the coefficient polynomials vanish at $x = 1$. If we use this latter recurrence for the ϕ's, then after setting $x = 1$ we will find a recurrence for the sums $f(n)$ that is nontrivial and of order J. This proves

Theorem. *Let $F(n,k)$ be a proper hypergeometric term with compact support and let $f(n)$ be its natural sum. Then $f(n)$ satisfies a nontrivial homogeneous linear recurrence relation of order J, whose coefficients are polynomials in n. The order J is subject to the a priori upper bounds stated in the fundamental theorem.*

What can we do with this method when the limits of the sums are not the natural ones? Quite a lot, actually.

Thus we might like to know more about

$$f(n) = \sum_{k=0}^{n} \binom{3n}{k}, \tag{7.7}$$

for instance.

When the limits of the sum are not the natural ones, the procedure begins the same way: find the recurrence that is satisfied by the summand. Then sum that recurrence over the range of k that occurs in your sum. If the limits of your sum are *linear* functions of the running variable n then you'll get an *inhomogeneous* recurrence relation that your sum satisfies.

What happens next depends on your needs. If you want to *prove* a claimed identity, then just show that the claimed right hand side satisfies the same inhomogeneous recurrence relation, and that it agrees with the left side on enough initial values, and you'll be finished. If on the other hand, you don't have any candidate for the right hand side then your success will depend upon your being able to solve the recurrence relation, or to prove that it cannot be solved in simple form. For this purpose the theorems of Petkovšek, to be discussed in chapter 13, may be useful.

Example 1 Take the sum $f(n)$ of (7.7) as an example. The Maple program easily finds the recurrence

$$F(n,k) - F(n-1,k) - 3F(n-1,k-1) - 3F(n-1,k-2) - F(n-1,k-3) == 0$$

for the summand $F(n,k) = \binom{3n}{k}$. If we sum only over $k = 0,1,\ldots,n$, instead of over the full support, we obtain

$$\begin{aligned} 0 &= f(n) - (f(n-1) + F(n-1,n)) - 3f(n-1) - 3(f(n-1) \\ &\qquad - F(n-1,n-1) - (f(n-1) - F(n-1,n-1) - F(n-1,n-2)) \\ &= f(n) - 8f(n-1) - F(n-1,n) + 4F(n-1,n-1) + F(n-1,n-2) \\ &= f(n) - 8f(n-1) + \frac{(5n^2+n-2)}{(2n-1)(2n-2)}\binom{3n-3}{n}. \end{aligned}$$

If we had a candidate for an evaluation of $f(n)$ in simple form it would be a simple matter to prove or disprove it using this recurrence. The contrary, however, is true. Using such a recurrence we can prove, using Petkovšek's method, that it is impossible to evaluate this sum within a certain clearly defined class of "simple closed forms." Another use for recurrences of this kind is that one can do asymptotics with them (see Example 5 of chapter XII). Finally, after taking a new unknown function $f(n) = 8^n g(n)$, we can express our original sum $f(n)$ as another sum, this one in the form $\sum_{k=0}^{n} \phi(k)$, in which the running variable does not appear under the summation sign. Such a sum is called *indefinite.* ∎

Example 2 For a happier ending, consider the sum

$$f(n) = \sum_{k=0}^{n} \binom{n+k}{k} 2^{-k},$$

in which the limits are certainly not the natural ones since the summand is supported on the nonnegative integers.

This time Maple finds that the summand $F(n,k) = \binom{n+k}{k} 2^{-k}$ satisfies the recurrence

$$2F(n,k) = F(n,k-1) + 2F(n-1,k).$$

If we now sum from $k=0$ to $k=n$ we get

$$2f(n) = (f(n) - F(n,n)) + 2(f(n-1) + F(n-1,n)),$$

which simplifies (rather dramatically!) to $f(n) = 2f(n-1)$. This, together with the initial condition $f(0) = 1$ shows that $f(n) = 2^n$ for all $n \geq 0$. ∎

(VIII) Certification and telescoping

The fundamental theorem assures us that if $F(n,k)$ is a proper hypergeometric term then F satisfies a nontrivial recurrence of the form

$$\sum_{i=0}^{I}\sum_{j=0}^{J} a_{i,j}(n)F(n-j,k-i)=0. \tag{8.1}$$

If we are interested in proving that $\sum_k F(n,k)=r(n)$, where $r(n)$ is some simple evaluation then (1) will serve to *certify* the evaluation, as we now discuss.

The idea of certification is fundamental in mathematics and computer science. It arises whenever we would like to make it easy for some audience to believe that a certain assertion is true.

Suppose I would like to convince you of the following assertion:

"the number 1065481086209 is not a prime."

The statement by itself is not obviously true, so it requires proof. An obvious proof *algorithm*, though not the only imaginable one, would be to look for nontrivial integer divisors of 1065481086209. To *certify* the truth of this statement is to supply the audience with an extra piece of information such that with that information they can use some standard proof technique and verify the statement quickly.

Certification is all about saving the audience time. We do that by giving them some information beyond the bare-bones statement that is being made, so as to help them see that it is true.

In this example, our proof certificate reads as follows:

$$1065481086209 = 333041\cdot 3199249$$

Now our task is much easier, though there still is some work to do. We must

(a) verify that the certificate is true and

(b) verify that the truth of the certificate implies the truth of the assertion to be proved.

In this case, we do (a) by multiplying out the product shown on the certificate, which requires some work but not as much as searching for factors in the first place. Then we do (b) quite easily because the truth of the certificate is an immediate contradiction of the definition of primality of the integer 1065481086209.

The idea of certification is extremely useful in the study of identities. What we must do is to find a *general* proof technique, that is to say, one that does not rely on individual human ingenuity each time, that will prove every identity in a certain class, provided only that a small amount of extra information (the *certificate*) is supplied. We will study several techniques for certifying identities in this work. Our requirements will always be the same: it should be easy to verify that the information on the certificate is correct and it should be easy and standardized to show that the information on the certificate proves the truth of the identity.

Conspicuously absent from the above list is any requirement that it should be easy to find the certificate in the first place. Indeed, many kinds of certificates have a certain "yes-but-where-did-that-come-from" quality about them. The user of the certificate need not be concerned with that question, since that person's only tasks are to verify that the certificate carries correct information (which is usually a lot easier than finding the information in the first place), and then checking that the certificate implies the truth of the identity.

Let's take the last example of the previous chapter, which was the evaluation

$$f(n)=\sum_{k=0}^{n}\binom{n+k}{k}2^{-k}=2^n \qquad (n\geq 0).$$

To certify that identity we present, as a certificate, the recurrence formula

$$2F(n,k)=F(n,k-1)+2F(n-1,k), \tag{8.2}$$

that is satisfied by the *summand* $F(n,k)=\binom{n+k}{k}2^{-k}$. With this certificate, any reader can be quickly convinced of the truth of the evaluation $f(n)=2^n$ as follows. First, to check that the information on the certificate is correct, we would substitute $F(n,k)=\binom{n+k}{k}2^{-k}$ into (2) and verify that it is true. How hard is it to do that? Well, after substituting into (2) we would be looking at

$$2\binom{n+k}{k}2^{-k}=\binom{n+k-1}{k-1}2^{-k+1}+2\binom{n+k-1}{k}2^{-k}.$$

To check this, we could write out each binomial coefficient as a quotient of factorials, as

$$2\frac{(n+k)!}{k!n!2^k}=\frac{(n+k-1)!}{(k-1)!n!2^{k-1}}+2\frac{(n+k-1)!}{k!(n-1)!2^k}.$$

After cancelling all of the factorials and powers of 2 we will be left with a trivially verifiable *polynomial* identity.

Here's another remark that may shed light on the idea of certification. The famous four-color conjecture was finally proved, in 1978, by K. Appel and W. Haken, after over 100 years. The proof that was found can be split roughly into two parts. First there was the reduction of the original question to merely a finite number of cases, over 1000 of them, but a finite number nonetheless. This portion of the proof, while difficult and detailed, can be checked by those who are interested enough to invest the time and energy, in the customary way that mathematical proofs can be checked.

Beyond that lies the checking of the 'finite number' of cases. Each one of these cases required a huge amount of computer time. At the end of each computation, the output was simply 'yes'. That is, the computer did not find any certificate that would smooth the path of later generations who might like to check the proof in full. It is precisely in this lack of certification that one feels the wish for a more satisfactory proof. No one would mind if the computers had to compute for a very long time if, after all of that computation they would have left us a fairly easily verifiable certificate.

In contrast to that situation, the computer proofs of identities are in a quite satisfactory condition. *Although vast amounts of computation might need to be done in order to find these computer-proofs, the computation leaves behind a certificate with which any human can, with just a modest effort, verify the truth of the computation.*

The general procedure for certifying the truth of an identity

$$f(n)=\sum_k F(n,k)=r(n),$$

where F is a proper hypergeometric term, is as follows.

(i) On the certificate there is written the k-free recurrence that F satisfies.

(ii) To check that the recurrence is true, the user will substitute the summand $F(n,k)$ into the claimed recurrence, cancel out factorials and powers of constants, thereby reducing the recurrence to an easily checkable polynomial identity.

(iii) After checking the truth of the certificate, the user will next check that the claimed identity is true, as follows.

(a) Sum the recurrence that is written on the certificate over all integer values of k. The result is a recurrence formula that is satisfied by the sum $f(n)$ itself.
(b) Verify that the claimed right hand side $r(n)$ of the identity also satisfies this recurrence, by direct substitution.
(c) Check that the identity is true for enough consecutive values of n, numerically, so as to insure that the fact that both sides satisfy the same recurrence implies that the two sides are identical for all n. ■

The whole procedure is quite algorithmic, and its success is guaranteed if the identity is true and the summand F is a proper hypergeometric term. It relies on the fact that it is quite simple to check a recurrence that the *summand* satisfies, though it may be quite hard to check a recurrence that the *sum* satisfies.

In a book about identities, then, one might envisage a long list of identities, and the proof of each one would simply consist in giving the certificate that contains the recurrence relation that is satisfied by the summand. For example,

Theorem. *For* $n \geq 0$ *we have* $\sum_{k=0}^{n} \binom{n+k}{k} 2^{-k} = 2^n$.

Proof: $2F(n,k) = F(n,k-1) + 2F(n-1,k)$. ■
And that's it.

OK, all of the above amounts to *one* way of certifying identities. There are other ways of certifying them, and we're going to discuss a few of them too. The next certification method that we'll present is called *telescoping*, for reasons that will soon be clear. It is similar to the method above in that the proof certificate contains just a single recurrence relation. It is different in that the recurrence relation is a different sort, one that is faster to find, in the case of big complicated identities, and more compact to state.

Here is what a telescoping recurrence looks like-

$$a_0(n)F(n,k) + a_1(n)F(n-1,k) + \cdots + a_J(n)F(n-J,k) = G(n,k) - G(n,k-1). \tag{8.3}$$

In this recurrence, $F(n,k)$ is again the summand in the sum that we are trying to prove, and the coefficients $\{a_j(n)\}_{j=0}^{J}$ are polynomials in n. The function $G(n,k)$ will always be a rational multiple of $F(n,k)$, i.e., $G(n,k) = R(n,k)F(n,k)$, where R is a rational function of n and k, and the rational function R is also "printed on" the certificate.

So, (3) is somewhat shorter than the previous certificate (1) in that (1) has $(I+1)(J+1)$ terms on its left side, while (3) has just $J+1$ terms on its left side and two on its right side. It will eventually turn out that (3) is easier to find than (1) also, though this will not be apparent from our first derivation of it because we'll begin with (1). The point is, though, that it is possible to find the recurrence (3) directly, without finding (1) first.

The hallmark of the telescoping recurrence is that on the left side only n is varying, and the right side telescopes on k.

Let's see if (3) has the two useful qualities that a certificate should have (it certainly has the useless quality of "where-did-that-come-from?").

First, is it easy to check that an assertion of the form (3) is true? Yes. Substitute in the $F(n,k)$ in question. Multiply F by the rational function R to obtain G, and substitute it into (3) also. Now cancel out all of the factorials and powers of constants in all of the terms, as before, and you will be looking at a polynomial identity that will be trivial to check.

Second, it is easy to check that the assertion (3) proves the identity that you are trying to prove. To do that, sum the certificate (3) over all integer k. Notice that the right hand side then telescopes, and sums to zero because of the compact support of F. On the left side we obtain a recurrence for the sum $f(n)$, namely

$$\sum_{j=0}^{J} a_j(n) f(n-j) = 0.$$

We would then check that the claimed right hand side of the identity, $r(n)$, satisfies the same recurrence, as before, and check the initial values, and we would be finished. As an example, we have

Theorem. *For $n \geq 0$ we have*
$$\sum_k \frac{n! 2^{n-2k}}{(n-2k)! k!^2} = \binom{2n}{n}.$$

Proof. We have
$$(n+1)F(n+1,k) - 2(2n+1)F(n,k) = G(n,k) - G(n,k-1),$$
where $G = RF$ and $R(n,k) = 4k - 2n$. ∎

The reader should, as an exercise, go through the proof algorithm that uses this certificate, and thereby prove the identity.

As a second exercise, the reader is invited to look up this identity in the hypergeometric database and verify it that way.

It's time to prove that a telescoping recurrence always exists.

Theorem. *Let F be a proper hypergeometric term with compact support. Then there is a positive integer J, a rational function $R(n,k)$, and polynomials in n, $\{a_j(n)\}_{j=0}^{J}$, not all zero, such that*

$$\sum_{j=0}^{J} a_j(n)F(n-j,k) = G(n,k) - G(n,k-1)$$

holds at all points (n,k) where $F(n,k) \neq 0$ and all values of F and G that occur are well defined, where $G(n,k) = R(n,k)F(n,k)$.

The proof begins with the 2-variable recurrence

$$\sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n)F(n-j,k-i) = 0, \tag{8.4}$$

which we know exists nontrivially.

We claim that it can be assumed, without loss of generality, that not all of the sums $\sum_{i=0}^{I} a_{i,j}(n)$ $(j = 0, 1, \ldots, J)$ can vanish.

Indeed, the recurrence (4), after multiplication by t^k and summation over k, asserts that

$$\sum_{j=0}^{J} \psi_{n-j}(t)\phi_j(t) = 0 \tag{8.5}$$

for all n, where $\psi_n(t) = \sum_k F(n,k)t^k$ and $\phi_j(t) = \sum_{i=0}^{I} a_{i,j}(n)x^i$. Suppose that *all* of the ϕ_j's vanish at $x = 1$. then they are all divisible by at least one factor $(x-1)$. Divide through all of the ϕ's by the highest power of $x-1$ that will divide all of them, which will leave the relation (5) still true. Hence we can suppose that at least one of the ϕ's does not vanish at $x = 1$. If we take this reduced set of ϕ's, then if we put $x = 1$ in (5) we will recover a recurrence of the form (4) in which the claim made above holds true, which completes the proof of the claim.

Next, let us introduce the shift operators K, N defined by $Nh(n) = h(n-1)$ and $Ku(k) = u(k-1)$. Then (4) can be written in operator form as $P(N,K,n)F(n,k) = 0$. Suppose we take the polynomial $P(u,v,n)$ and expand it in a power series in v, about the point $v = 1$,

$$P(u,v,n) = P(u,1,n) + (v-1)Q(u,v,n)$$

where Q is a polynomial. Then (4) implies that

$$0 = P(N,K,n)F(n,k) = (P(N,1,n) + (K-1)Q(N,K,n))F(n,k), \tag{8.6}$$

i.e., that $P(N,1,n)F(n,k) = (1-K)Q(N,K,n)F(n,k)$. On the left side of this latter recurrence, only n varies. On the right side, if we put $G(n,k) = Q(N,K,n)F(n,k)$, then the right side is simply $G(n,k) - G(n,k-1)$, and G is itself a rational function multiple of F since any fixed number of shift operators, applied to a hypergeometric term, only multiply it by a rational function. Finally, the recurrence is nontrivial since the operator $P(N,1)$ is not the zero operator, because of the claim that we established at the beginning of the proof. ∎

A Maple program that will find and print this telescoped recurrence has been written by Zeilberger. It can be obtained by anonymous ftp from `euclid.math.temple.edu`. The program is quite fast, and it has many applications even beyond the problem of proving identities. The method has been implemented in Mathematica by Peter Paule and Markus Schorn. To obtain this version write to `Peter.Paule@risc.uni-linz.ac.at`.

The method (*fast algorithm*) of finding the telescoped form isn't at all like the proof of the existence theorem above. One could proceed, of course, as in the existence theorem, but that would lead through a system with $(I+1)(J+1)$ unknowns in order to find a recurrence which has only about $J+2$ parameters to be determined.

Instead, Zeilberger's idea was to go directly to the telescoped form. Since we have already proved that the form exists, the success of the algorithm is not in doubt. In the execution, the key to is the use of Gosper's algorithm, in modified form. Gosper's algorithm is the subject of the next chapter, so here we will only very briefly sketch the main idea.

We are given our hypergeometric term $F(n,k)$ and we are looking for numbers $a_j(n)$ and a hypergeometric term $G(n,k)$ such that

$$\sum_{j=0}^{J} a_j(n)F(n-j,k) = G(n,k) - G(n,k-1)$$

and we know that the $a_j(n)$ exist because we can guarantee that by choosing I, J large enough.

What Gosper's algorithm does is the following. Given a hypergeometric term $u(k)$, it will find, or prove the nonexistence of, a hypergeometric term $v(k)$ s.t. $u(k) = v(k+1) - v(k)$. Our present needs are just a bit different. Here our function $u(k)$, namely the entire left side of the equation above, is not "given". It contains some constants $a_{i,j}$ that are unknown. If we knew the a's then the use of Gosper would be routine and it would find the term v for us. But we don't know the a's.

Zeilberger's tactic was to carry along the unknown a's as parameters, and then do Gosper's algorithm, with a twist. Normally Gosper's method finds $v(k)$ by solving a set of linear simultaneous equations for coefficients that occur in it. The twist is that one solves a larger set of simultaneous equations, not only for the coefficients that must be determined in $v(k)$, but also for the unkown $a_{i,j}$'s, which also occur linearly in the equations. This point may be clearer after we discuss Gosper's method, which we will do next.

(IX) Gosper's algorithm

In 1978, R. W. Gosper, Jr. found a remarkable algorithm that solves the following problem:

> **Input.** *A hypergeometric term $f(k)$, that is, a function of a single variable k for which $f(k+1)/f(k)$ is a rational function of k.*
>
> **Question.** *Does there exist a hypergeometric term $g(k)$ such that $f(k) = g(k+1) - g(k)$? If so, output $g(k)$.*

The importance of this question is that an answer to it would open the way to having computers do *indefinite sums*. By an indefinite sum we mean one of the form

$$F(n) = \sum_{k=0}^{n} f(k),$$

in which the summand does not depend on the running index n, and in which there is no natural compact support. Thus, in a *definite* sum, such as we have been considering up till now, the limits of the sum are $-\infty$ to ∞, and the true limits of the summation process are dictated by the compact support of the summand. In an *indefinite* sum, we want the sum as a function of the upper limit of the summation.

If we want to do an indefinite sum $\sum_{k=0}^{n} f(k)$ just look how easy it is if Gosper's algorithm can provide us with a $g(k)$: since $f(k) = g(k+1) - g(k)$, we have

$$F(n) = \sum_{k=0}^{n} f(k) = \sum_{k=0}^{n} (g(k+1) - g(k)) = g(n+1) - g(0).$$

That is, of course, fully analogous to commenting on how easy it is to do integrals $\int_{t=0}^{x} f(t)dt$ if someone will kindly hand us an antiderivative of f, i.e., a function $g(x)$ such that $f(x) = g'(x)$, and will come as a shock to no one.

Things are just a tiny bit slippery, though. If we are given an $f(k)$, hypergeometric or not, there is *always* a function $g(k)$ for which $f(k) = g(k+1) - g(k)$, namely the function $g(k) = \sum_{j=0}^{k-1} f(j)$. The issue is *not* whether such a function exists. The issue is, when f is a hypergeometric term, can we find a g that is a hypergeometric term too?

Lemma 1. *Every rational function $R(k)$ can be expressed in the form*

$$R(k) = \frac{p(k+1)}{p(k)} \frac{q(k)}{r(k+1)}, \tag{9.1}$$

in which p, q, r are polynomials for which it is true that if $q(k_1) = 0$ and $r(k_2) = 0$, then $k_2 - k_1$ is not a positive integer.

Proof. Suppose $R(k) = N(k)/D(k)$, where N and D are polynomials. The following algorithm halts with a p, q, r with the required properties.

(I) Set $p(k) := 1$; $q(k) := N(k)$; $r(k) := D(k-1)$.
(II) **While** q, r have factors $k+\alpha$, $k+\beta$ for which $\alpha - \beta$ is a positive integer,
Do
$q(k) := q(k)/(k+\alpha)$; $r(k) := r(k)/(k+\beta)$;
$p(k) := p(k)(k+\alpha-1)(k+\alpha-2)\cdots(k+\beta+1)$. ■

Lemma 2. *There is a hypergeometric term $g(k)$ such that $f(k) = g(k+1) - g(k)$ if and only if there is one for which $g(k)/f(k)$ is a rational function.*

Proof. We have

$$\frac{g(k)}{f(k)} = \frac{g(k)}{g(k+1) - g(k)} = \frac{1}{\frac{g(k+1)}{g(k)} - 1},$$

which is a rational function since g is a hypergeometric term. ■

Lemma 3. *There is a hypergeometric term $g(k)$ such that $f(k) = g(k+1) - g(k)$ if and only if there is one of the form $g(k) = r(k)s(k)f(k)/p(k)$, where p, q, r are the polynomials of equation (1) above, for the rational function $R(k) = f(k+1)/f(k)$, and $s(k)$ is a polynomial which satisfies the equation*

$$q(k)s(k+1) - r(k)s(k) = p(k). \tag{9.2}$$

Proof. Since, by Lemma 2, g/f is a rational function, we can take that rational function in the form rs/p, where s is a rational function. If we substitute $g(k) = f(k)r(k)s(k)/p(k)$ into (1), we find, after cancelling the factor of $f(k)$, that $s(k)$ satisfies (2). It remains to show that in fact s is a polynomial.

Suppose, on the contrary, that $s(k) = \nu(k)/\delta(k)$, with degree$(\delta) \geq 1$ and ν, δ relatively prime. Then (2) becomes

$$q(k)\delta(k)\nu(k+1) - r(k)\nu(k)\delta(k+1) = p(k)\delta(k)\delta(k+1). \tag{9.3}$$

Let n be the largest integer such that δ vanishes at z and at $z-(n-1)$, for some complex z. Clearly $n \geq 1$. In (3), let $k = z$ and then let $k = z - n$. This yields

$$r(z)\nu(z)\delta(z+1) = q(z-n)\delta(z-n)\nu(z-n+1) = 0. \tag{9.4}$$

Since δ vanishes at z and at $z-n+1$, and since ν and δ are relatively prime, it follows that ν cannot vanish at either of these points. So (4) reduces to

$$r(z)\delta(z+1) = q(z-n)\delta(z-n) = 0. \tag{9.4}$$

Now neither $\delta(z+1)$ nor $\delta(z-n)$ can vanish, for suppose the former. Then δ would vanish at $z+1$ and at $z-n+1$, and n would not be maximal. A simlar remark holds for the latter possibility. Thus (4) reduces to $r(z) = q(z-n) = 0$. But this contradicts the properties of p, q, r that are guaranteed by Lemma 1. We have now contradicted our supposition that s is a non-polynomial rational function. ■

Where are we now? Well, there will be a hypergeometric term $g(k)$ for which $f(k) = g(k+1) - g(k)$ if and only if equation (2) has a polynomial solution $s(k)$. How can we discover whether or not it has such a solution? It's a matter of degree. Suppose we knew that if there were such a polynomial s then it must have degree 26. Then we could simply assume that s is the most general polynomial of degree 26, substitute in (2), and try to solve the system of linear equations for the coefficients of the polynomial s. If there is a nontrivial solution then we have a g, and if there is not then such a g doesn't exist.

So the whole question comes down to the problem of finding the degree of any polynomial $s(k)$ that satisfies (2).

Lemma 4. *For given polynomials p, q, r, equation (2) has a polynomial solution $s(k)$ if and only if it has a polynomial solution of degree Δ, where the possible values of Δ are determined as follows: Let d_+ and d_-, respectively, denote the degrees of $q+r$ and of $q-r$, and let h_+, h_-, respectively, be their highest coefficients.*

(1) If $d_- \geq d_+$ then the only possible degree of $s(k)$ is $\Delta = \deg(p(k)) - d_-$.

(2) Otherwise, if $d_- < d_+$ then there are either 0, 1 or 2 possible values of the degree of $s(k)$, namely:

(a) If $2h_- + h_+(\deg(p) - d_+ + 1) \neq 0$ then $\Delta = \deg(p) - d_+ + 1$ is possibly the degree of $s(k)$.

(b) If $-2h_-/h_+$ is a positive integer that is larger than $\deg(p) - d_+ + 1$ then $\Delta = -2h_-/h_+$ is possibly the degree of $s(k)$

Proof. First we rewrite (2) in the form

$$2p(k) = (s(k+1) + s(k))(q(k) - r(k)) + (s(k+1) - s(k))(q(k) + r(k)). \tag{9.5}$$

The degrees of the two polynomials on the right side are $\Delta + d_-$ and $(\Delta - 1) + d_+$. If $d_- \geq d_+$ then the larger of these two degrees is $\Delta + d_-$ which must therefore be equal to the degree of p, proving assertion (1) of the lemma.

Otherwise, $d_- \leq d_+ - 1$. Then the larger of the two degrees on the right is certainly $\Delta - 1 + d_+$. This, therefore, must be equal to $\deg(p(k))$ unless the two degrees on the right are equal and the terms of highest degree cancel. They do not cancel under the conditions stated in (2)(a) of the lemma, and they do cancel under the condition stated in (2)(b) of the lemma. ∎

The complete algorithm of Gosper can now be stated.

Given a hypergeometric term $f(k)$. To determine whether there exists a hypergeometric term $g(k)$ such that $f(k) = g(k+1) - g(k)$, and if so to find g, do the following:

(i) Apply the algorithm that appears in the proof of Lemma 1 to the rational function $f(k+1)/f(k)$ to obtain the polynomials p, q, r.

(ii) Compute the possible values of the number Δ, of Lemma 4.

(iii) For each of the possible values of Δ, look for a polynomial $s(k) = \sum_{j=0}^{\Delta} a_j k^j$ that satisfies (2), by substituting this assumed $s(k)$ into (2) and attempting to solve the system of linear equations in the $\Delta + 1$ unknowns.

(iv) If these linear equations have no solution, then exit with "No such $g(k)$ exists." Otherwise exit with $g(k) = r(k)s(k)f(k)/p(k)$. ∎

Example Take $f(k) = k^2$. That means we are looking for a $g(k)$ for which $g(k+1) - g(k) = k^2$. If we find such a g we will be able to write out a simple formula for $\sum_{j=0}^{n} j^2$.

The algorithm of Lemma 1 begins with $p = 1$, $q = (k+1)^2$ and $r = (k-1)^2$. The conditions for termination are violated, so we divide q by $(k+1)$, r by $(k-1)$ and multiply p by k. We now have $(p, q, r) = (k, k+1, k-1)$. The termination condition is again violated, so we repeat the process, this time getting $(p, q, r) = (k^2, 1, 1)$, and the algorithm halts.

Next we compute (note: the degree of the 0 polynomial is -1)

$$\Delta = \text{degree}(k^2) - \max\{\text{degree}(0), \text{degree}(2) - 1\} = 2 - (-1) = 3.$$

Hence we look for a polynomial $s(k) = a + bk + ck^2 + dk^3$ which satisfies the equation $s(k+1) - s(k) = k^2$. If we equate coefficients of like powers of k on both sides, and solve the resulting equations for a, b, c, d, we find that a is arbitrary, and $(b, c, d) = (1/6, -1/2, 1/3)$. So $s(k) = k/6 - k^2/2 + k^3/3$. By step (4) of the algorithm, we exit with

$$g(k) = r(k)s(k)f(k)/p(k) = 1 \cdot (k/6 - k^2/2 + k^3/3) \cdot k^2/k^2 = k/6 - k^2/2 + k^3/3.$$

Now we know that

$$\sum_{j=0}^{n} j^2 = g(n+1) - g(0) = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}. \ ∎$$

There is a somewhat subtle point that should be recognized. In many cases of interest, the input hypergeometric term $f(k)$ will contain more than just k and absolute constants. It will contain additional parameters as well. Gosper's algorithm relies heavily on knowing the degree Δ, in advance, for its success. Without Δ there is no assurance that the algorithm would terminate.

What would happen if the degree Δ turned out to depend on the parameters? That would be unacceptable, and the algorithm would fail. Hence the success of the algorithm is limited to hypergeometric terms for which the degree Δ is an absolute constant. Input hypergeometric terms for which that will be true, will look like

$$f(k) = R(k)\frac{\prod_{i=1}^{uu}(a_i k + b_i)!}{\prod_{i=1}^{vv}(c_i k + d_i)!}z^k, \tag{6}$$

in which R is a rational function, and all of the following must be absolute constants, containing no parameters: the limits uu, vv, the a_i's and the c_i's, and the degrees of the numerator and denominator of the rational function R.

In the previous chapter we remarked that Zeilberger's program for finding the telescoping recurrence can be obtained by anonymous ftp. Here we remark that that same program can be used to do just Gosper's algorithm, which is, of course, considerably less than what it is capable of doing. However, recall that one is given a term $F(n,k)$ and his program finds a recurrence of the form

$$\sum_{j=0}^{J} a_j(n)F(n-j,k) = G(n,k) - G(n,k-1).$$

for a given J, if such a recurrence exists. Take $J = 0$ and observe that in that case the algorithm will do the job of Gosper's. Hence, if you want a program in Maple that will implement Gosper's algorithm, get Zeilberger's program and use it with $J = 0$, and if you want one in Mathematica, get the program of Paule and Schorn.

(X) The WZ method

We come now to an amazingly short method for certifying the truth of combinatorial identities. With this method, the proof certificate for an identity contains just a single rational function. That's the whole proof. Thus we are able, for instance, to give proofs of every identity in the hypergeometric database of chapter II above, and each proof will consist of just a certain rational function $R(n,k)$.

Suppose we want to prove an identity $\sum_k F(n,k) = r(n)$. Well first, let's divide through by the right hand side (this proof technique will be valid only in a region of n where $r(n)$ does not vanish), and write the identity that is to be proved as

$$\sum_k \{\frac{F(n,k)}{r(n)}\} = 1.$$

That being done, we might as well just think of $f(n,k)/r(n)$ as having been the original summand. In other words, we can assume, without loss of generality, that the identity that we're trying to prove is

$$\sum_k F(n,k) = 1. \tag{10.1}$$

Let's call the left hand side $f(n)$. So $f(n) \stackrel{\text{def}}{=} \sum_k F(n,k)$, and we're trying to prove that $f(n) = 1$ for all n. One way to show that a function f is always equal to 1 is to show that $f(0) = 1$, say, and that $f(n+1) - f(n) = 0$ for all n. That would certainly do it.

One way to certify the fact that $f(n+1) - f(n) = 0$ for all n would be to display a function $G(n,k)$ such that

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k), \tag{10.2}$$

for then we would simply sum (10.2) over all integer k to find that indeed $f(n+1) - f(n)$ is always 0.

A pair of functions (F,G) that satisfy (10.2) is called *a WZ pair.*

Example 1 Let's convince ourselves that the function

$$f(n) = \sum_k \frac{\binom{n}{k}^2}{\binom{2n}{n}} \tag{10.3}$$

is always equal to 1 for $n \geq 0$. To do that we state that (10.2) is indeed true, where F is the summand (not the sum! the *summand*) in (10.3), and $G = RF$, where $R(n,k)$ is the rational function

$$R(n,k) = -\frac{k^2(3n-2k+3)}{2(2n+1)(n-k+1)^2}. \tag{10.4}$$

As usual, when confronted with such things, we suppress the natural whereonearthdidthatcomefrom reaction, and we confine ourselves to verifying the certificate. To verify it we check that if $R(n,k)$ is given by (10.3), we multiply R by $F(n,k)$, the summand in (10.1), to get a certain function $G(n,k)$. We then take that function and check that (10.2) is satisfied, and we're all finished. ∎

Now let's discuss where on earth it came from. Begin with your summand $F(n,k)$, from (10.1). Now form the difference $D = F(n+1,k) - F(n,k)$, and collect and simplify it as much as you can (after all, it's only a rational function times $F(n,k)$). This difference D depends on n and k and the various parameters that may have been in your summand. Let's think of n, for the time being, as one of the parameters, so we won't explicitly name it as one of the variables that D depends on. That means that D is a function of k alone, so call it $D(k)$.

Take $D(k)$ and give it to Gosper's algorithm. If possible, Gosper's algorithm will produce a function $g(k)$ such that $D(k) = g(k+1) - g(k)$. That function $g(k)$ will of course have the 'parameter' n in it, so let's rename $g(k)$ to $G(n,k)$. This $G(n,k)$ does exactly what we wanted it to do, namely it is the WZ-mate for F in equation (10.2). Furthermore, by Lemma 2 of the previous chapter, G/F is a rational function.

There's just one problem.

There is no assurance that Gosper's algorithm will find a g. After all, Gosper's algorithm might just return "No such g exists." There is no theorem which says that such a g must exist. In fact, there are cases where the identity $\sum_k F(n,k) = 1$ is true, but the g really doesn't exist.

Despite that, the simple fact remains that among hundreds of hypergeomtric identities for which it has been tried, the WZ certification does indeed work for all but a handful of them (just two, at last count).

Of course the earlier results *always* work. So for every proper hypergeometric term $F(n,k)$ we always have a ("telescoping") certification of (10.1) that looks like

$$\sum_{j=0}^{J} a_j(n)F(n-j,k) = G(n,k+1) - G(n,k). \tag{10.5}$$

The observed fact is that 99% of the time, this reduces to just two terms on the left, and becomes the WZ equation (10.2).

When this happens it is important that it be recognized. One reason is the following. If we use the telescoping certification algorithm to look for a certificate like (10.5), it will certainly find one. But the one it finds might not be the one of the WZ form (10.2), even though such a certificate may exist! The algorithm finds *some* telescoping certificate, but need not find the one of the form (10.2), even in cases where that form can be satisfied. Gosper's algorithm, on the other hand, is guaranteed to find the WZ certificate if there is one.

How to find the WZ certificate of an identity

To certify an identity that is in the form $\sum_k f(n,k) = r(n)$:

A. Let $F(n,k) := f(n,k)/r(n)$.

B. Let $f(k) := F(n+1,k) - F(n,k)$. Input $f(k)$ to Gosper's algorithm. If that algorithm fails then this one does too.

C. Otherwise, the output $G(n,k)$ of Gosper's algorithm is the WZ mate of F. The rational function $R(n,k) := G(n,k)/F(n,k)$ is the WZ certificate of the identity $\sum_k F(n,k) = 1$. ∎

How to prove an identity from its WZ certificate

To prove an identity $\sum_k f(n,k) = r(n)$ from its WZ certificate $R(n,k)$:

A. Construct $F(n,k) := f(n,k)/r(n)$ and $G(n,k) := R(n,k)F(n,k)$.

B. Verify that equation (10.2) is true. To do this, write it all out, divide out all of the factorials, and verify the resulting polynomial identity.

C. Verify that the given identity is true when $n = 0$. ∎

The theorem that underlies these procedures makes use of these hypotheses:

(F1) For each integer k, the limit

$$f_k = \lim_{n\to\infty} F(n,k) \tag{10.6}$$

exists and is finite.

(G1) For each integer $n \geq 0$, we have

$$\lim_{k \to \pm\infty} G(n,k) = 0.$$

(G2) We have $\lim_{L\to\infty} \sum_{n\geq 0} G(n,-L) = 0$.

The theorem itself is the following.

Theorem [WZ]. *Let (F,G) satisfy equation (10.2). If (G1) holds then*

$$\sum_k F(n,k) = const. \qquad (n = 0,1,2,\ldots). \tag{10.7}$$

If (F1) and (G2) both hold then we have the companion identity

$$\sum_{n\geq 0} G(n,k) = \sum_{j\leq k-1} (f_j - F(0,j)), \tag{10.8}$$

where f is defined by (10.6).

The theorem not only validates the certification procedure, it shows that we get a free new identity every time we use the method. The new identity is the companion identity (10.8). Roughly the reason that it is there is that the fuinctions F, G play very symmetrical roles in the WZ equation (10.2). If they are so symmetric, why should there be an identity associated with F and not with G? Well there *is* one associated with G, and it is (10.8).

This is not the only *free* identity that we get from the procedure. We will discuss at least two more before we're finished. None of them appear as results of any of the earlier certification procedures that we have discussed. That is, if we use the raw 2-variable recurrence for F, or the telescoping recurrence for F as our certification method, we get the advantage that they are guaranteed to work on every proper hypergeometric summand, but a disadvantage relative to the WZ method is that we get no extra identities that we didn't know before.

Proof. We use the symbol Δ_n for the forward difference operator on n: $\Delta_n h(n) = h(n+1) - h(n)$. Sum both sides of equation (10.2) from $k = -L$ to $k = K$, getting

$$\begin{aligned}\Delta_n\left\{\sum_{k=-L}^{K} F(n,k)\right\} &= \sum_{k=-L}^{K} \{\Delta_k G(n,k)\} \\ &= G(n,K+1) - G(n,-L).\end{aligned}$$

Now let $K, L \to \infty$ and use (G1), to find that $\Delta_n \sum_k F(n,k) = 0$, i.e., that $\sum_k F(n,k)$ is independent of $n \geq 0$, which establishes (10.7).

If we sum both sides of (10.2) from $n = 0$ to N, we get

$$F(N+1,k) - F(0,k) = \Delta_k\left\{\sum_{n=0}^{N} G(n,k)\right\}.$$

Now let $N \to \infty$ and use (F1) to get

$$f_k - F(0,k) = \Delta_k\left\{\sum_{n\geq 0} G(n,k)\right\}.$$

Replace k by k', sum from $k' = -L$ to $k' = k-1$, let $L \to \infty$, and use (G2), to obtain the companion identity (8), completing the proof. ∎

(XI) WZ proofs of the hypergeometric database

To illustrate the scope of the WZ method, we now present one-line proofs of every identity in the hypergeometric database, of section 2.

(I) Proof of Gauss's $_2F_1$ identity

To prove Gauss's identity in the form $\sum_k F(n,k) = 1$ where

$$F(n,k) = \frac{(n+k)!(b+k)!(c-n-1)!(c-b-1)!}{(n-1)!(c-n-b-1)!(k+1)!(c+k)!(b-1)!},$$

take

$$R(n,k) = \frac{(k+1)(k+c)}{n(n+1-c)}.$$

(II) Proof of Kummer's $_2F_1$ identity

To prove that

$$_2F_1\left[\begin{matrix} 1-c-2n & -2n \\ & c \end{matrix} \middle| -1\right] = (-1)^n \frac{(2n)!(c-1)!}{n!(c+n-1)!}$$

rewrite it as $\sum_k F(n,k) = 1$ where

$$F(n,k) = (-1)^{n+k} \frac{(2n+c-1)!n!(n+c-1)!}{(2n+c-1-k)!(2n-k)!(c+k-1)!k!}.$$

Then take

$$R(n,k) = \frac{k(k+c-1)(2+4c+c^2-3k-2ck+k^2+10n+7cn-6kn+10n^2)}{2(2n-k+c+1)(2n-k+c)(2n-k+2)(2n-k+1)}. \blacksquare$$

(III) Proof of Saalschütz's $_3F_2$ identity

To prove Saalschütz's $_3F_2$ identity in the form $\sum_k F(n,k) = 1$ where

$$F(n,k) = \frac{(a+k-1)!(b+k-1)!n!(n+c-a-b-k-1)!(n+c-1)!}{k!(n-k)!(k+c-1)!(n+c-a-1)!(n+c-b-1)!},$$

take

$$R(n,k) = \frac{k(-1+c+k)(a+b-c+k-n)}{(b-c-n)(-a+c+n)(1-k+n)}. \blacksquare$$

(IV) Proof of Dixon's identity

To prove that

$$\sum_k (-1)^k \binom{n+b}{n+k}\binom{n+c}{c+k}\binom{b+c}{b+k} = \frac{(n+b+c)!}{n!b!c!},$$

take

$$R(n,k) = \frac{(k+b)(k+c)}{2(k-n-1)(n+b+c+1)}. \blacksquare$$

(V) Proof of Clausen's ${}_4F_3$ identity

To prove Clausen's ${}_4F_3$ identity in the form $\sum_k F(n,k) = 1$, where

$$F(n,k) = \phi(k)\phi(n-k)\psi(n),$$

and

$$\phi(t) = \frac{(a+t-1)!(b+t-1)!}{t!(-\frac{1}{2}+a+b+t)!},$$

$$\psi(t) = \frac{(t+a+b-\frac{1}{2})!t!(t+2a+2b-1)!}{(t+2a-1)!(t+a+b-1)!(t+2b-1)!},$$

take $R(n,k)$ to be

$$\frac{(1-2a-2b-2k)k(a-k+n)(b-k+n)(2+2a+2b-2k+3n)}{(2a+n)(a+b+n)(2b+n)(1-k+n)(1+2a+2b-2k+2n)}. \blacksquare$$

(VI) Proof of Dougall's ${}_7F_6$ identity

To prove Dougall's ${}_7F_6$ identity

$$\begin{aligned} {}_7F_6\left[\begin{matrix} d, & 1+\frac{d}{2}, & d+b-a & d+c-a & 1+a-b-c & n+a & -n \\ & \frac{d}{2}, & 1+a-b, & 1+a-c & b+c+d-a & 1+d-a-n & 1+d+n \end{matrix}\right] \\ = \frac{(d+1)_n(b)_n(c)_n(1+2a-b-c-d)_n}{(a-d)_n(1+a-b)_n(1+a-c)_n(b+c+d-a)_n}, \end{aligned}$$

take $R(n,k)$ to be

$$\frac{(a-c+k)(a-b+k)(a-b-c+k)(-1-a+b+c+d+k)(a-d-k+n)(1+a+2n)}{(-a+b+c-k)(d+2k)(a+n)(b+n)(c+n)(-1+2a-b-c-d+n)(1-k+n)}.$$

(XII) Spinoffs from the WZ method

The main business of the WZ method is the certification of identities. However, in the course of getting that job done it gives as a dividend at least three additional kinds of new identities for each one that it certifies. These are (1) the companion identity (2) the dual identity and (3) the indefinite-sum-made-definite. Here we will discuss and illustrate these three kinds of spinoffs.

The Companion Identity

The companion identity is equation (8) of section (X). It states that

$$\sum_{n\geq 0} G(n,k) = \sum_{j\leq k-1} (f_j - F(0,j)), \tag{12.1}$$

in which

$$f_k \stackrel{\text{def}}{=} \lim_{n\to\infty} F(n,k).$$

Example 1 Take the famous identity

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

Here $F(n,k) = \binom{n}{k}^2/\binom{2n}{n}$, and all f_k's are 0. To do the WZ procedure we apply Gosper's algorithm to the input $F(n+1,k) - F(n,k)$, and it outputs

$$G(n,k) = \frac{(-3n+2k-3)}{2(2n+1)} \frac{n!^2}{(k-1)!^2(n-k+1)!^2\binom{2n}{n}}.$$

If we substitute in (1) and use the fact that $F(0,j) = \delta_{0,j}$ we obtain the companion identity

$$\sum_{n\geq 0} \frac{(-3n+2k-3)}{2(2n+1)} \frac{n!^2}{(k-1)!^2(n-k+1)!^2\binom{2n}{n}} = \begin{cases} 0 & \text{if } k=0; \\ -1 & \text{if } k\geq 1, \end{cases}$$

which simplifies to

$$\sum_{n\geq 0} \frac{(3n-2k+1)}{(2n+1)} \frac{\binom{n}{k}^2}{\binom{2n}{n}} = 2 \qquad (k=1,2,3,\ldots). \tag{12.2}$$

Equation (2) is a new identity, in the sense that it is not immediately reducible to any known identity in the hypergeometric database. This comment is made here not so much to impress the reader with what a spectacular identity (2) is, but rather to underline the incompleteness of a fixed database. It seems that whatever finite list of general identities one may incorporate into a database will be incomplete. One can add, of course, a number of hypergeometric transformation rules, which map given identities onto other ones. These greatly extend the scope of any database, but still there is no fixed list of identities and fixed algorithm that will prove every preassigned identity.

Practically every small binomial coefficient identity *is* a special case of some known hypergeometric identity. But the other side of that coin is that virtually all of them can be certified directly by WZ certificates which yield a companion identity that is new. Of course not all of these companion identities will be esthetically delightful. Many of them are quite messy. But they are mostly beyond the reach of any fixed database and searching and transforming algorithm that is known. The WZ approach provides a systematic way to find and to prove them by computer.

Another point is the following. Suppose identity x is a special case of identity X. *It does not follow that the companion of identity x is necessarily a special case of the companion of identity X.* In fact this is usually false. A similar remark will hold for the idea of a *dual* identity that we will treat later in this section.

Example 1. Consider the identity

$$\sum_{k=0}^{n}(-1)^k\binom{x}{k}\binom{2n-k}{n}=\binom{2n-x}{n}, \tag{12.3}$$

which is a triviality to prove, even without the hypergeometric database.

The WZ certificate of this identity is

$$R(n,k)=\frac{(x+1-k)(k-x)}{(2n-x+1)(2n-x+2)},$$

so the WZ mate is

$$G(n,k)=R(n,k)F(n,k)=(-1)^k\frac{(2n-k+1)!(n-x)!x!}{(k-1)!(n-k+1)!(2n-x+2)!(x-k-1)!}.$$

Using the reflection formula for the Gamma function

$$\Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin\pi z},$$

we find the quantities required for the companion identity to be

$$f_k\stackrel{\text{def}}{=}\lim_{n\to\infty}F(n,k)=(-1)^k\binom{x}{k}2^{x-k},$$

$$F(0,k)=(-1)^k\binom{x}{k}.$$

Now we can write down the companion identity of (3). It is

$$\sum_{n\geq 0}\frac{(2n-k+1)!(n-x)!}{(n-k+1)!(2n-x+2)!}=\frac{(-1)^k(k-1)!(x-k-1)!}{x!}\sum_{j=0}^{k-1}(-1)^j\binom{x}{j}(2^{x-j}-1). \tag{12.4}$$

This identity is valid for noninteger x and positive integer k, and it should be noted that the sum on the left is an infinite series.

The summand is a standard ${}_3F_2$ summand of Saalschütz type. That is, the sum of the lower parameters exceeds the sum of the upper parameters by 1. In fact the identity can be rewritten as

$$\begin{aligned}&{}_3F_2\left[\begin{matrix}k/2, & (k+1)/2, & k-x\\ k+1-x/2, & k+(1-x)/2 & \end{matrix}\right]\\&\qquad=\frac{(k-1)!(x-k-1)!}{x!}\sum_{j=0}^{k-1}(-1)^{j+k}\binom{x}{j}(2^{x-j}-1)\\&\qquad\qquad(\text{noninteger }x;\text{integer }k>0).\end{aligned} \tag{12.5}$$

A fair description of this result, then, is that it is a nonterminating extension of a special case of Saalschütz's identity to noninteger x. Saalschütz's identity itself would give the value of this sum if x were an integer. ∎

Example 2. If we begin with the full Saalschütz identity, as stated in section 2, then we find the companion identity

$$\begin{aligned}{}_3F_2&\left[\begin{matrix}k, & c-a-b, & c+k-1\\ & c-b+k, & c-a+k\end{matrix}\right]\\ &=\frac{(c-b)_k(c-a)_k}{(a)_k(b)_k}\left\{1-\frac{(c-a-b)_a}{(c-a)_a}\sum_{j=0}^{k-1}\frac{(a)_j(b)_j}{j!(c)_j}\right\}.\end{aligned}$$

This is valid for $c>a+b$, integer $k>0$, and the sum is again nonterminating. It is interesting in that on the right side we see a partial sum of Gauss's original ${}_2F_1[1]$ identity. ∎

Example 3. This time let's start with Vandermonde's identity

$$\sum_k\binom{a}{k}\binom{n}{k}=\binom{n+a}{a}.$$

For this we find the WZ certificate

$$R(n,k)=\frac{k^2}{(-1+k-n)(1+a+n)},$$

and the companion identity

$$\sum_n\frac{\binom{n}{k}}{\binom{n+a+1}{n}}=\frac{a+1}{(k+1)\binom{a}{k+1}},$$

which is valid for integer $a>k\geq 0$. ∎

The dual identity

The mapping from an identity to its companion is not an involution. But there is a true dual identity in the WZ theory. It is obtained as follows.

Suppose $F(n,k)$ is a summand, of the form

$$F(n,k)=x^ny^k\rho(n,k)\frac{\prod_i(a_in+b_ik+c_i)!}{\prod_i(u_in+v_ik+w_i)!},$$

where ρ is a rational function of n,k. Now formally replace each factor $(an+bk+c)!$ according to the shadowing rule

$$(an+bk+c)!\longrightarrow\begin{cases}(-1)^{an+bk}/(-an-bk-c-1)!, & \text{if } a+b\neq 0;\\ \text{unchanged}, & \text{if } a+b=0.\end{cases}\tag{12.6}$$

The result of this operation, applied to $F(n,k)$, is *the shadow* of F, written $\overline{F}$. The shadow of the binomial coefficient $\binom{n}{k}$ is $(-1)^{n+k}\binom{-k-1}{-n-1}$, for example.

Now suppose $(F(n,k),G(n,k))$ is a WZ pair. The dual pair (F',G') is defined to be

$$(F'(n,k),G'(n,k))\stackrel{\text{def}}{=}(\overline{G}(-k-1,-n),\overline{F}(-k,-n-1)).\tag{12.7}$$

Lemma. *If (F,G) is a WZ pair, then so is $(\overline{F},\overline{G})$.*

Indeed, to say that (F,G) is a WZ pair is to say that

$$F(n+1,k)-F(n,k)=G(n,k+1)-G(n,k).\tag{12.8}$$

But $F(n+1,k)/F(n,k)$ is a rational function, and so is $F(n,k+1)/F(n,k)$. Let's call these two rational functions $R_1(n,k;F)$ and $R_2(n,k;F)$. A third rational function, inherent in the WZ construction, is $R(n,k) = G(n,k)/F(n,k)$.

Let's divide (8) through by $F(n,k)$, to get

$$\begin{aligned} R_1(n,k;F) - 1 &= \frac{G(n,k+1)}{F(n,k)} - R(n,k) \\ &= \frac{G(n,k+1)}{F(n,k+1)} \frac{F(n,k+1)}{F(n,k)} - R(n,k) \\ &= R(n,k+1)R_2(n,k;F) - R(n,k). \end{aligned} \tag{12.9}$$

Now the shadow functions $(\overline{F},\overline{G})$ *have the same rational functions associated with them as do* (F,G). That is, $R_1(n,k;\overline{F}) = R_1(n,k;F)$, etc. To verify that, notice that if the property holds for each multiplicative factor in F it holds for F, so it is enough to verify the property for $F^*(n,k) := (an+bk+c)!$. That is, one must check that $\overline{F^*}(n+1,k)/\overline{F^*}(n,k) = F^*(n+1,k)/F^*(n,k)$, and three other relations like that, all trivial to check.

It follows that (9) can be rewritten as

$$R_1(n,k;\overline{F}) - 1 = R(n,k+1)R_2(n,k;\overline{F}) - R(n,k).$$

If we now multiply through by $\overline{F}(n,k)$ we find that $(\overline{F},\overline{G})$ is a WZ pair, as claimed. ∎

Remark. The argument shows that the rational function $R(n,k)$ is the same for the original and the shadow pairs. It also shows that one does not even have to do the shadow mapping exactly as prescribed, namely on all factors of F except those for which $a+b=0$. One can map some of the factors and not others at one's whim, with the same result. The choice that was made above seem to give fruitful results, but others are possible.

The final step in the dualization is to pass from the shadow pair to the dual pair (F',G'). That is done simply by changing the variables as shown in (7), which of course does not affect the fact that the functions form a WZ pair, though it does alter the rational function $R(n,k)$ that is their ratio, via the same change of variables.

We have now established that the mapping

$$(F(n,k),G(n,k)) \longrightarrow (F'(n,k),G'(n,k)) = (\overline{G}(-k-1,-n),\overline{F}(-k,-n-1)), \tag{12.10}$$

is an involution of WZ pairs.

Example 4. We return to the sum of the squares of the binomial coefficients identity of example 1 of this section, where the WZ pair was

$$F(n,k) = \frac{\binom{n}{k}^2}{\binom{2n}{n}}; \quad G(n,k) = \frac{-3+2k-3n}{2(2n+1)} \frac{\binom{n}{k-1}^2}{\binom{2n}{n}}.$$

We apply the shadowing rule (6) and discover that the shadow pair is

$$\overline{F}(n,k) = \frac{(-2n-1)!(-k-1)!^2}{(-n-1)!^4(n-k)!^2}; \quad \overline{G}(n,k) = \frac{(-3+2k-3n)}{2(2n+1)} \frac{\binom{-k}{-n-1}^2}{\binom{-n-1}{-2n-1}}.$$

Finally we change variables as in (10) to pass to the dual pair,

$$F'(n,k) = \frac{-3k+2n}{2}\binom{n}{k}^2\binom{2k}{k}; \quad G'(n,k) = \frac{k}{2}\binom{n}{k-1}^2\binom{2k}{k}.$$

We now have a formal WZ pair. We check that the hypotheses (G1), (G2) of the theorem are satisfied, and then we know that $\sum_k F'(n,k)$ is independent of $n \geq 0$. Since it is 0 when $n = 0$ we have a proof of the dual identity

$$\sum_k (3k-2n)\binom{n}{k}^2\binom{2k}{k} = 0 \qquad (n = 0, 1, \ldots).$$

Again, while this identity is hardly spectacular in itself, it is new in that there is no immediate, algorithmic way to deduce it from the standard hypergeometric database (though we certainly would not want to challenge human mathematicians to give independent proofs of it!) . ∎

We summarize a few other instances of duality. The dual of $\sum_k \binom{n}{k} = 2^n$ is

$$\sum_k (-1)^{n+k}\binom{n}{k}2^k = 1 \quad (n = 0, 1, \ldots)$$

The Saalschütz identity is self dual.

The dual of Dixon's identity is a special case of Saalschütz's. Note that since dualization is symmetric, it follows that we can prove Dixon's identity from Saalschütz's by dualization of a special case.

The dual of Vandermonde's identity

$$\sum_k \binom{a}{k}\binom{n}{k} = \binom{n+a}{a}$$

is

$$\sum_k (-1)^{n+k}\binom{n}{k}\binom{k+b}{k} = \binom{b}{n},$$

which is Vandermonde's identity again. So it is self-dual.

The process of dualization does not in general commute with specialization. Thus we can imagine beginning with some identity, passing to some special case, dualizing, passing to some special case, etc., thereby generating a whole chain of identities from a given one. If the original identity has a large number of free parameters, like Dougall's ${}_7F_6$ for instance, then the chain might be fairly long. Whether interesting new identities can be found this way is not known.

The definite sum made indefinite

The following observation of Zeilberger generates an interesting family of identities and is of help in finding asymptotic estimates of hypergeometric sums.

Imagine a sum $\sum_k F(n,k) = 1$ of the type that we have been considering, in which the support of F properly contains the interval $[0,n]$. Now consider the restricted sum

$$h(n) := \sum_{k=0}^{n} F(n,k).$$

A prototype of this situation is a sum like $\sum_{k=0}^{n}\binom{3n}{n}$.

Suppose now that F has a WZ mate G, so that

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k).$$

Sum this equation over $k \leq n$, to obtain

$$h(n+1) - F(n+1,n+1) - h(n) = G(n,n+1).$$

Next, replace n by j, and sum over $j = 0, \ldots, n-1$ to get

$$h(n) = h(0) + \sum_{j=1}^{n} (F(j,j) + G(j-1,j)),$$

which is to say that

$$\sum_{k=0}^{n} F(n,k) = F(0,0) + \sum_{j=1}^{n} (F(j,j) + G(j-1,j)). \tag{12.11}$$

On the left side we have a definite sum, and on the right we have one in which the running index n does not appear under the summation sign.

Example 5. Take $F(n,k) = \binom{3n}{k}/8^n$, so that $\sum_k F(n,k) = 1$. The WZ mate of F is

$$G(n,k) = \frac{(-32 + 22k - 4k^2 - 93n + 30kn - 63n^2)}{(3n-k+3)(3n-k+2)} \frac{\binom{3n}{k-1}}{8^{n+1}}.$$

We find that (be sure to use a computer algebra package to do things like this!)

$$F(j,j) + G(j-1,j) = \frac{2(2 - j - 5j^2)}{3(3j-1)(3j-2)} 8^{-j} \binom{3j}{j}.$$

Hence (11) reads as

$$8^{-n} \sum_{k=0}^{n} \binom{3n}{k} = 1 + \frac{2}{3} \sum_{j=1}^{n} \frac{(2 - j - 5j^2)}{(3j-1)(3j-2)} 8^{-j} \binom{3j}{j}. \tag{12.12}$$

We can find some asymptotic information as follows. It is easy to see that the left side of (12) approaches 0 as $n \to \infty$. So the infinite sum on the right must be 0. Thus we have

$$\frac{2}{3} \sum_{j=1}^{\infty} \frac{(5j^2 + j - 2)}{(3j-1)(3j-2)} 8^{-j} \binom{3j}{j} = 1,$$

an identity that is perhaps not instantly obvious by inspection, and therefore (12) can be rewritten as

$$8^{-n} \sum_{k=0}^{n} \binom{3n}{k} = \frac{2}{3} \sum_{j=n+1}^{\infty} \frac{(5j^2 + j - 2)}{(3j-1)(3j-2)} 8^{-j} \binom{3j}{j}.$$

On the right side we replace j by $n+k$, factor out $8^{-n}\binom{3n}{n}$, and divide each term of the sum by that factor. We then have, as $n \to \infty$,

$$\begin{aligned} 8^{-n} \sum_{k=0}^{n} \binom{3n}{k} &\sim 8^{-n} \binom{3n}{n} \frac{2}{3} \sum_{k=1}^{\infty} \frac{5}{9} 8^{-k} (\frac{27}{4})^k \\ &= 2 \cdot 8^{-n} \binom{3n}{n}. \end{aligned}$$

So the sum in question is asymptotic to twice its last term. ∎

(XIII) Which recurrences have simple solutions?

Let's slow down for a moment to see where matters stand. Here are some things that you might like to do about identities:

Prove a given identity. Here we are on strong ground. If you have an identity

$$\sum_k F(n,k) = r(n),$$

and if F is a proper hypergeometric term, and if the summation is over the full support of F, then the left side definitely satisfies a homogeneous recurrence relation with polynomial-in-n coefficients, and your computer can find that recurrence and certify it so you will believe what it says. You can then check, by substitution, that the right side, $r(n)$, satisfies the same recurrence. It then remains only to show, by calculation, that the left and right sides agree with each other on enough initial values to sustain the recurrence, and your proof is complete.

Certify a true identity. Under the same assumptions, we are on very strong ground here too. To convince an audience, with minimal effort on their part, that your identity is true, you might certify it by displaying the full two-variable recurrence that $F(n,k)$ satisfies, or the telescoped "fast-form" of that recurrence, both of which are guaranteed to exist. Not guaranteed, but more spectacular when it does exist, is the WZ certificate, so probably one ought to look for that first. Any of these provides a true certification: it is easy for the audience to check that the information on the certificate is correct, and it is easy to check that that information implies the truth of the given identity.

Find a simple evaluation of a given sum. Now it's getting harder. Given a sum $f(n) = \sum_k F(n,k)$, under the same assumptions as above. Can we find a "simple" explicit evaluation of the sum? Suppose for a moment that the summand $F(n,k)$ does not contain any parameters other than n and k. Then *the purpose of the present chapter is to show that we can discover whether or not the sum has a "simple" evaluation algorithmically, under a very reasonable definition of "simple."*

More precisely, the following is the case. The first step will be to find the recurrence that your unknown sum satisfies. We know that that can always be done, and we know how to do it. So suppose we are now looking at the recurrence that $f(n)$ satisfies. What we see is a linear, homogeneous recurrence relation with polynomial coefficients. Thanks to results of Marko Petkovšek, some of which we will now discuss, there is an algorithm that will discover whether or not there exist a fixed number of hypergeometric terms,

$$t_1(n), t_2(n), \ldots, t_q(n)$$

such that your recurrence relation is satisfied by

$$f(n) = \sum_{i=1}^{q} c_i t_i(n)$$

for some constants c_i.

So if your sum evaluates to $n^5\binom{7n}{3n}^3$, or to 2^n+5 or to n^4+n^2+1, you'll be sure to find out, and your computer will do the searching. If the computer returns "No" then that will constitute a proof that no such solution exists, rather than an abject admission of failure. If the computer returns "Yes," then it will also return to you the hypergeometric basis $\{t_i(n)\}_{i=1}^n$. This procedure is clearly self-certifying: to convince yourself that the computer solution is true, just take each member t_i of the claimed basis, substitute it in your recurrence formula, and see if it works.

To continue, suppose we have an unknown sum that satisfies our conditions but it *does* contain additional parameters aside from the summation index k and one running index n. Of course the sum still satisfies a recurrence relation and your computer can find it. So find it. Now, if that recurrence is of order 1, like $a(n)f(n+1)-b(n)f(n)=0$, then f is indeed a hypergeometric term† and you will have explicitly found it.

It may happen that even though your $f(n)$ is a hypergeometric term, i.e., satisfies a first order recurrence, the algorithms that we have developed might not find that recurrence of lowest order, but perhaps they will find one of higher order. At this moment the quest becomes nonalgorithmic. I would suggest that you interrogate the hypergeometric database and all of its appendages as thoroughly as possible, and wish you good luck.

Please notice that if your summand is proper hypergeometric and the sum is over the full support of the summand, then you haven't been reduced to relying on your good luck unless the summand does contain extra parameters, and the recurrence formulas that our algorithms find for the sum are of orders greater than 1.

To increase our chances in this case, it is worthy of note that the algorithm for finding a two-variable recurrence satisfied by the summand actually returns not just one recurrence, but a whole vector space of them. That is because we are solving a system of equations in which there are fewer equations than unknowns. It is possible that by choosing the arbitrary constants correctly one can lower the order of the recurrence that the sum satisfies.

In more detail, the fundamental theorem guarantees a recurrence of the form

$$\sum_{i=0}^{I}\sum_{j=0}^{J} a_{i,j}(n)F(n-j,k-i)=0.$$

The next step is to find the recurrence for the sum $f(n)$, by summing this two-variable recurrence over integer k. This yields

$$\sum_{j=0}^{J} f(n-j)\{\sum_{i=0}^{I} a_{i,j}(n)\}.$$

If $J=1$, we are finished, because the recurrence is of first order. Suppose $J=2$. It still might happen (and often does) that $\sum_{i=0}^{I} a_{i,2}(n)=0$, in which case the recurrence for the sum will be first order (good news) even though the recurrence for the summand was of 2nd order (apparent bad news). If the $a_{i,j}(n)$'s have some arbitrary constants in them, it may be possible to choose these to force, or at any rate, encourage, the good news to happen.

Petkovšek's algorithm Hyper

We are now going to develop the outlines of the theory of hypergeometric solutions of linear recurrences with polynomial coefficients. We will do this in detail for recurrences of order two, and will quote the results for higher order recurrences.

Let p, q, r be polynomials in n, and consider the recurrence

$$p(n)a(n+2)+q(n)a(n+1)+r(n)a(n)=0. \tag{13.1}$$

† In fact, a good *definition* of a hypergeometric term is that it is a function that satisfies a first order homogeneous recurrence with polynomial coefficients.

Our objective is to describe an algorithm for deciding whether this recurrence has hypergeometric term solutions, and if so, finding them. This will be done in two major stages. First we will describe how to decide if such a recurrence has a *polynomial* solution. Then we will consider hypergeometric solutions.

To decide whether polynomial equations exist we have again a question of the sort that we considered in developing Gosper's algorithm. We first decide what the possible degrees of such polynomial solutions must be, and then we try to find such polynomials by exhaustive computation within the specified degree limits, which is a finite task.

So suppose (1) has a polynomial solution $a(n)$, of degree N. What are possible, and impossible, values of N? If m is the maximum degree of p, q, r, then we can suppose that

$$\begin{aligned} p(n) &= u_0 n^m + u_1 n^{m-1} + \cdots, \\ q(n) &= v_0 n^m + v_1 n^{m-1} + \cdots, \\ r(n) &= w_0 n^m + w_1 n^{m-1} + \cdots, \\ a(n) &= \alpha_N n^N + \alpha_{N-1} n^{N-1} + \cdots. \end{aligned}$$

Here, at least one of $u_0 . v_0, w_0$ is nonzero, and $\alpha_N \neq 0$.

It follows, by the binomial theorem, that

$$\begin{aligned} a(n+j) = \alpha_N &\left(n^N + jNn^{N-1} + j^2 \binom{N}{2} n^{N-2} \right) + \alpha_{N-1} \left(n^{N-1} + j(N-1) n^{N-2} \right) \\ &+ \alpha_{N-2} n^{N-2} + O(n^{N-3}). \end{aligned}$$

We substitute this expansion, with $j = 1$ and $j = 2$ into (1) and equate the coefficients of like powers of n, beginning with the highest power.

The coefficient of n^{N+m} gives

$$u_0 + v_0 + w_0 = 0. \tag{13.2}$$

If (2) does not holds then (1) has no nontrivial polynomial solution. If it does hold, then the coefficient of n^{N+m-1} requires that

$$(2u_0 + v_0)N + u_1 + v_1 + w_1 = 0. \tag{13.3}$$

If $2u_0 + v_0 \neq 0$ then (3) determines a single possible value for the unknown degree N. If, on the other hand

$$2u_0 + v_0 = 0, \tag{13.4}$$

then (3) requires that

$$u_1 + v_1 + w_1 = 0 \tag{13.5}$$

also.

Now suppose that both (4) and (5) hold. Then matching the coefficient of n^{N+m-2} gives, after some rearrangement, the quadratic equation

$$u_0 N^2 + (2u_1 - u_0 + v_1)N + u_2 + v_2 + w_2 = 0. \tag{13.6}$$

We claim that $u_0 \neq 0$. Indeed if $u_0 = 0$ then (2), (4) imply that $u_0 = v_0 = w_0 = 0$, a contradiction. Hence $u_0 \neq 0$, and (6) determines at most two possible values for the degree N.

Theorem 1. *There is a finite algorithm that will decide whether the recurrence (1) has a nontrivial polynomial solution, and will find all of them, if they exist.* ∎

The algorithm that has just been described, for finding all polynomial solutions to a recurrence relation with polynomial coefficients, will be called *Algorithm Poly*. In fact, we have described only the case of *Poly* that deals with second order recurrences, but it can be extended to recurrences of every order.

That settles the question for polynomial solutions of recurrences of order 2. We will now consider *hypergeometric* solutions of recurrences of order 2. To do this we need a sharpening of the representation of rational functions that we proved in connection with Gosper's algorithm.

Theorem 2. *Let $r(x)$ be a nonzero rational function. Then there exist a constant $Z \neq 0$ and monic polynomials $A(x), B(x), C(x)$ such that*

$$r(x) = Z \frac{A(x)}{B(x)} \frac{C(x+1)}{C(x)}, \tag{13.7}$$

in which

1. *$A(x), B(x+k)$ are relatively prime, for every integer $k \geq 0$,*
2. *$A(x), C(x)$ are relatively prime,*
3. *$B(x), C(x+1)$ are relatively prime.*

Remark. Actually the factorization of $r(x)$ that is described in the theorem is unique, though we shall not prove this here.

Proof. Let $r = p/q$, where p, q are polynomials with highest coefficients a, b, respectively, and let $Z = a/b$, which is clearly nonzero. Let N be the largest nonnegative integer root of the polynomial

$$R(k) := \mathrm{Resultant}_x(p(x), q(x+k)),$$

or -1 if no such roots exist. Then for integer $k > N$, the polynomials $p(x)$ and $q(x+k)$ are relatively prime.

We construct a sequence of polynomials $\{p_i, q_i\}_{i=-1}^{N}$, as follows: First, $p_{-1}(x) = p(x)/a$ and $q_{-1}(x) = q(x)/b$. Then, for each $i = 0, 1, \ldots, N$, do

$$\begin{aligned} s_i(x) &:= \gcd(p_{i-1}(x), q_{i-1}(x+i)), \\ p_i(x) &:= \frac{p_{i-1}(x)}{s_i(x)}, \\ q_i(x) &:= \frac{q_{i-1}(x)}{s_i(x-i)}. \end{aligned}$$

Now put $A(x) := p_N(x)$, $B(x) := q_N(x)$ and

$$C(x) := \prod_{i=1}^{N} \prod_{k=1}^{i} s_i(x-k).$$

We claim that A, B, C fulfill the conditions stated in the theorem.

First, we have

$$\begin{aligned} Z \frac{A(x)}{B(x)} \frac{C(x+1)}{C(x)} &= \frac{a}{b} \frac{p_N(x)}{q_N(x)} \prod_{i=1}^{N} \prod_{k=1}^{i} \frac{s_i(x-k+1)}{s_i(x-k)} \\ &= \frac{a}{b} \frac{p_{-1}(x)}{\prod_{i=0}^{N} s_i(x)} \frac{\prod_{i=0}^{N} s_i(x-i)}{q_{-1}(x)} \prod_{i=1}^{N} \frac{s_i(x)}{s_i(x-i)} \\ &= \frac{ap_{-1}(x)}{bq_{-1}(x)} = \frac{p(x)}{q(x)} = r(x), \end{aligned}$$

as required. Next, A, B, C are monic, by construction. Now we check that properties 1-3 in the conclusion of the theorem are satisfied.

For each $k = 0, 1, \ldots, N$, we have

$$\gcd(p_k(x), q_k(x+k)) = \gcd\left(\frac{p_{k-1}(x)}{s_k(x)}, \frac{q_{k-1}(x+k)}{s_k(x)}\right) = 1.$$

Further, suppose that $0 \le k \le i, j \le N$. Then $p_i, p_j \backslash p_k$, and $q_i, q_j \backslash q_k$, and therefore

$$\gcd\bigl(p_i(x), q_j(x+k)\bigr) \backslash \gcd\bigl(p_k(x), q_k(x+k)\bigr).$$

Thus, for such i, j, k, we have

$$\gcd\bigl(p_i(x), q_j(x+k)\bigr) = 1. \tag{13.8}$$

Proof of 1: $A(x)$ and $B(x+k)$ are relatively prime for all $0 \le k \le N$, as we see by putting $i = j = N$ in (8). By definition of N, this implies that they are relatively prime for all $k \ge 0$.

Proof of 2: Suppose A, C have a common factor. Then for some $1 \le k \le i \le N$, so do $p_N(x)$ and $s_i(x-k)$. Then $p_N(x)$ and $q_{i-1}(x+i-k)$ likewise have a common factor, by definition of q_i. However, since $0 \le i-k \le i-1 < N$, this contradicts (8).

Proof of 3: Suppose B, C have a common factor. Then for some i, k such that $0 \le k \le i-1 \le N-1$, so do $q_N(x)$ and $s_i(x-k)$. Then $q_N(x)$ and $p_{i-1}(x-k)$ likewise have a common factor, by definition of p_i. Therefore $q_N(x+k)$ and $p_{i-1}(x)$ have a common factor, which contradicts (8), since $0 \le k \le i-1 < N$. ∎

Now we can discuss hypergeometric solutions of second order recurrences with polynomial coefficients. Once more, we consider the recurrence

$$p(n)a(n+2) + q(n)a(n+1) + r(n)a(n) = 0, \tag{13.9}$$

with p, q, r given polynomials. Suppose there is a hypergeometric solution of (9). Then there is a rational function $S(n)$ such that $a(n+1) = S(n)a(n)$. Substituting in (9) we find that

$$p(n)S(n+1)S(n) + q(n)S(n) + r(n) = 0. \tag{13.10}$$

Let

$$S(n) = Z\frac{A(n)}{B(n)}\frac{C(n+1)}{C(n)}$$

be the factorization of $S(n)$ that was described in Theorem 2. Then we have

$$Z^2p(n)A(n+1)A(n)C(n+2) + Zq(n)B(n+1)A(n)C(n+1) + r(n)B(n+1)B(n)C(n) = 0. \tag{13.11}$$

We will show that there are only finitely many possibilities (in fact, very few indeed) for each of the components A, B, C, Z of the factorization of $S(n)$, so there is a finite algorithm that will find them all, or prove that there are none.

Since $A(n)$ divides the first two terms, it also divides $r(n)B(n+1)B(n)C(n)$. By properties 1, 2 of theorem 2, $A(n)$ is relatively prime to $C(n)$, $B(n)$, and $B(n+1)$. Therefore $A(n)\backslash r(n)$.

Similarly, $B(n+1)\backslash Z^2p(n)A(n+1)A(n)C(n+2)$, and by properties 1 and 3 of theorem 2, $B(n+1)\backslash p(n)$.

Thus *there is only a finite set of possibilities for the factors* A, B, namely the set of monic divisors of $r(n)$ and of $p(n-1)$, respectively.

Furthermore there are at most two possible choices for Z, because (11) can be divided by $A(n)B(n+1)$ to get

$$Z^2\frac{p(n)}{B(n+1)}A(n+1)C(n+2) + Zq(n)C(n+1) + \frac{r(n)}{A(n)}B(n)C(n) = 0, \tag{13.11}$$

in which the coefficients are still polynomials. But now we can determine the value of Z by looking at the highest coefficient of the left hand side. We find that Z satisfies a quadratic equation, so there are at most two possibilities for it, for each fixed choice of A, B.

Finally, if A, B, Z are fixed, we can use algorithm Poly to decide whether (11) has any polynomial solutions $C(n)$. If so, then we will have found a hypergeometric solution of (9). If (11) has no polynomial solution, then (9) has no hypergeometric solution.

Algorithm Hyper, for recurrences of order 2

Input are polynomials p, q, r. Output is a hypergeometric solution $a(n)$ of (9), if one exists, or else 0.

(1) For all monic factors $A(n)$ of $r(n)$ and $B(n)$ of $p(n-1)$ do
  $P(n) := A(n+1)(p(n)/B(n+1))$;
  $Q(n) := q(n)$;
  $R(n) := (r(n)/A(n))B(n)$;
  $m := \max \deg\{P, Q, R\}$;
  Let α,β,γ be the coefficients of n^m in P, Q, R, resp.;
  for all nonzero Z such that $\alpha Z^2 + \beta Z + \gamma = 0$ do
    If the equation
    $Z^2 P(n)C(n+2) + ZQ(n)C(n+1) + R(n)C(n) = 0$
    has a nonzero polynomial solution $C(n)$ then
      set $S(n) := Z(A(n)/B(n))(C(n+1)/C(n))$;
      return a nonzero solution of $a(n+1) = S(n)a(n)$;
      exit.
(2) Return 0 ∎

Theorem 3. *There is a finite algorithm (Hyper) that will decide if a given homogenenous linear recurrence relation with polynomial coefficients has a hypergeometric solution, and will output one if they exist.* ∎

Example. The sequence of van der Poorten

$$a(n) = \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2$$

is easily found, by our earlier algorithms, to satisfy the recurrence

$$(n+2)^3 a(n+2) - (2n+3)(17n^2 + 51n + 39)a(n+1) + (n+1)^3 a(n) = 0.$$

We claim that the sequence $a(n)$ does not satisfy a recurrence of order 1, i.e., is not hypergeometric.

Indeed, the coefficients are all of the same degree, so the equation for Z will have a nonzero solution iff A, B are of equal degrees. But they are both monic factors of $(n+1)^3$, so they must be equal. Then $P = p, Q = q, R = r$, and Z satisfies $Z^2 - 34Z + 1 = 0$, i.e., $Z = 17 \pm 12\sqrt{2}$. In both cases Algorithm Poly shows that the auxiliary equation has no nonzero polynomial solutions, whence the given sequence is not hypergeometric. ∎

(XIV) Multi- and q-sparklers

The theory that we have discussed so far has concerned only hypergeometric identities with a single discrete summation variable. In fact the results and algorithms have been generalized far beyond that stage. The scope of the ideas extends to multivariate summation identities, i.e., to $\sum_{r,s} F(n,r,s) = f(n)$, and generally to summations over any fixed number of summation indices ("fixed" means that the number of indices must not be one of the parameters involved in the problem). The result likewise extend to q-hypergeometric identities, to multivariate q-identities, to integral-sum identities, and to constant term identities as well. In all of these cases, computer methods that are analogous to the ones that we have been discussing are able, unaided by human hands, to prove, to certify and to discover identities.

We will not be able here to develop these ideas fully, but instead refer the interested reader to the literature. We will confine orselves to stating a few of the results, and then giving some of the very pretty examples ("sparklers") that have resulted from applying the theory.

Multivariate recurrences.

In the case of sums that involve more than one summation variable, a proper-hypergeometric term looks like this:

$$F(n,\mathbf{k}) = P(n,\mathbf{k})\frac{\prod_{s=1}^{p}(a_s n + \mathbf{b}_s\cdot\mathbf{k} + c_s)!}{\prod_{s=1}^{q}(u_s n + \mathbf{v}_s\cdot\mathbf{k} + w_s)!}z^k.$$

The main theorem still holds, which is to say that every such function satisfies a k-free recurrence of the form

$$\sum_{j=0}^{J}\sum_{0\le i\le I}\alpha(\mathbf{i},j,n)F(n-j,\mathbf{k}-\mathbf{i}) = 0,$$

where the α's are polynomials in n. Also, it turns out that the existence of a recurrence of order J is guaranteed, provided that

$$J \ge \left\lfloor \frac{1}{r!}\left\{\sum_{s=1}^{p}\sum_{r'=1}^{r}|(\mathbf{b}_s)_{r'}| + \sum_{s=1}^{q}\sum_{r'=1}^{r}|(\mathbf{v}_s)_{r'}|\right\}^r\right\rfloor. \tag{14.1}$$

The above is the certification by the fundamental theorem. The two forms of compressed certification have r-variable analogues. First the telescoping certificate is guaranteed to exist, and it looks like

$$a_0(n)F(n,\mathbf{k}) + \cdots + a_J(n)F(n-J,\mathbf{k}) = \sum_{i=1}^{r}\Delta_i G_i(n,\mathbf{k}),$$

in which Δ_i is the forward difference operator w.r.t. k_i, and in which each

$$G_i(n,\mathbf{k})/F(n,\mathbf{k}) = R_i(n,\mathbf{k})$$

is a rational function.

Second, the r-variable WZ certificate, while not guaranteed to exist, usually does, and is of the form

$$F(n+1,\mathbf{k}) - F(n,\mathbf{k}) = \sum_{i=1}^{r}\Delta_i G_i(n,\mathbf{k}).$$

Example. Here is an example of of a 2-variable sum and its certification.

The summand $F(n,j,m)$ is

$$\frac{(-1)^{j+n}(2n-2m+j)!(2k+1+j)!i!(2k-2i+1)!n!(k-i-n)!}{2^{2n}j!(n-m)!(i+m)!(2k+1)!(k-i)!(m-j)!(n-m+j)!(2k-2i-2m+1+j)!}.$$

We claim that $\sum_{j,m} F(n,j,m) = 1$ for all $n \geq 0$. If $f(n)$ denotes the sum, then clearly $f(0)=1$, so it suffices to show that $f(n+1)-f(n)=0$ for all $n \geq 0$. For this purpose we construct two rational functions $R_1(n,j,m)$, $R_2(n,j,m)$.

R_1 has numerator

$$\begin{aligned}(&-4j^2k-3nj-4ji^2+4mj-2n-j+2m-j^2-4m^2-4n^2-4ni^2-8m^2k+4mi^2\\ &-4jk^2+4n^2j-4n^2m+8nm+3j^2n+2ni-2mi+2mk+8m^2i+4mk^2-4nk^2\\ &+4m^2n-2nk+2ji-2jk+4j^2i-8mjn+8nmk-8nmi+8kji-8kmi+8kni\\ &-4njk-12mji+4nji+12mjk)(-j/4),\end{aligned}$$

and denominator $(2n-2m+j)(-1-m+j)(n+1+i)(-k+i+n)$.

The rational function $R_2(n,j,m)$ has numerator

$$\begin{aligned}&-(i+m)(2n-2m+j+1)(4n^2j-4n^2m+j^2n-2njk+4m^2n-6mjn-4nm\\ &+6nj+4nmk-4nmi+2nji+2j+4mjk-4mji-2jk+2ji-4m^2k+4m^2i\\ &+4mk-4mi-2mj+j^2),\end{aligned}$$

and denominator

$$4(n+1-m)(n+1-m+j)(2k-2i-2m+2+j)(n+1+i)(-k+i+n).$$

It is then routinely verifiable that

$$\Delta_n F = \Delta_j(R_1 F) + \Delta_m(R_2 F),$$

where the Δ's are the usual forward difference operators, and the assertion follows upon double-summing w.r.t. j and m. ∎

Example. Here is another example of a double sum certification, this time of the following identity of Carlitz:

$$\sum_{i+j+k=n} \binom{i+j}{i}\binom{j+k}{j}\binom{k+i}{k} = \sum_{l=0}^{n}\binom{2l}{l}.$$

The computer found and proved a recurrence satisfied by the left side that is easily seen to be also satisfied by the right side, since the recurrence operator factorizes nicely into $((n+2)N-(4n+6))(N-1)$.

Theorem. *Let*

$$F(i,j,n) := \frac{(i+j)!(n-j)!(n-i)!}{i!^2 j!^2 (n-i-j)!^2},$$

and let $a(n)$ be its double sum w.r.t. j and i. Let N be the forward shift operator in n. The sequence $a(n)$ satisfies the recurrence

$$(4n+6+(-8-5n)N+(2+n)N^2)a(n)=0.$$

Proof. Let Δ_i and Δ_j be the forward difference operators in i and j respectively. Define two rational functions as follows. First, $R_1(i,j)$ has for its numerator

$$\begin{aligned}(&-6i^2-2i^2n+n^3+5n^2+4+8n-4ij+nij+4i^2j+2i^3-8j-14nj-5n^2j\\ &+4nj^2+6j^2-n^2i-ni+2i)(n-i+1)i^2\end{aligned}$$

and its denominator is $(i+j-n-1)^2(i+j-n-2)^2(i+j)$. Next define

$$R_2(i,j) = \frac{(j-n-1)j^2(-n+i-1)(4i^2+2ij-4i-2ni+nj-4n-n^2+2j-4)}{(i+j-n-1)^2(i+j-n-2)^2(i+j)}.$$

It is then routinely verifiable that

$$(4n+6+(-8-5n)N+(2+n)N^2)F(i,j,n) = \Delta_i(R_1F)+\Delta_j(R_2F)$$

and the result follows by summing w.r.t. i and j. ∎

q-identities

A q-proper-hypergeometric term is one that is of the form

$$F(n,k) = \frac{\prod_s Q(a_sn+b_sk,c_s)}{\prod_s Q(u_sn+v_sk,w_s)} q^{an^2+bnk+ck^2+dk+en}\xi^k,$$

where

$$Q(m,c) = (1-cq)(1-cq^2)\cdots(1-cq^m).$$

Once again, such terms satisfy k-free recurrences, and the certification machinery is all available. As an example, let's take the q analogue of the Chu-Vandermonde identity,

$$\sum_k q^{k^2}\binom{n}{k}_q^2 = \binom{2n}{n}_q,$$

in which $\binom{n}{k}_q$ is based on the q-factorial $[m!]_q = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{m-1})$.

Our rational functions $R(n,k)$ are now rational functions of q^n and q^k instead of n and k.

Define

$$R(n,k) = \frac{q^{n+1}(q^{2n+1}+q^n-2q^k)}{(1-q^{2n+1})(1+q^{n+1})}.$$

If we then define $G(n,k)=R(n,k)F(n,k)$, then it is routine to check that $F(n+1,k)-F(n,k)=G(n,k)-G(n,k-1)$, which is a q-WZ certificate that certifies the given identity.

If we follow the shadowing and dualizing procedure for WZ certifiable identities, we find the dual of the q-Chu-Vandermonde identity to be

$$\sum_k \binom{2k}{k}_q \binom{n}{k}_q^2 \frac{(q^{n-k}+q^n-2q^k)}{(1+q^k)} = 0.$$

Likewise, the companion identity is found to be

$$\sum_{n\geq 0} \frac{q^{n+1}(q^{2n+1}+q^n-2q^k)}{(1-q^{2n+1})(1+q^{n+1})} \frac{\binom{n}{k}_q^2}{\binom{2n}{n}_q}$$

$$= q^{-k^2}\left((q)_\infty \sum_{j=0}^{k} \frac{q^{j^2}}{(q)_j^2} - 1\right), \qquad (k\geq 0),$$

in which $(q)_\infty = \prod_i (1-q^i)$. The case $k=0$, which reads as

$$(q)_\infty = 1+\sum_{n\geq 0} \frac{q^{n+1}(q^{2n+1}+q^n-2)}{(1-q^{2n+1})(1+q^{n+1})} \frac{(q)_n^2}{(q)_{2n}}$$

is already of some interest. All three of these spinoffs of WZ appear to be new identities.

References

The study of hypergeometric functions was initiated by C. F. Gauss, in 1812. See his *Werke*, volume 3, 123-163. An exposition of the classical theory is in

W. N. Bailey, *Generalized hypergeometric series*, Cambridge Tracts in Mathematics, Vol. 32, Cambridge University press, London, 1935 (reprinted by Hafner, New York, 1964).

Sister Celine's creation of the algorithmic method for finding hypergeometric polynomial recurrence relations is in

Fasenmyer, Mary Celine, PhD dissertation, University of Michigan, 1945.

George Andrews's observation that identities from the theory of hypergeometric functions can unify the study of combinatorial identities, is contained in

George E. Andrews, *Applications of basic hypergeometric functions*, SIAM Rev. **16** (1974), 441-484.

Doron Zeilberger's recognition that Sister Celine's methods could be adapted to automated proofs of identities appears in

D. Zeilberger, Sister Celine's technique and its generalizations, J. Math. Anal. Appl. **85** (1982), 114-145.

Gosper's algorithm is in

R. William Gosper, Jr., *Decision procedure for indefinite hypergeometric summation*, Proc. Nat. Acad. Sci. U.S.A. **75** (1978), 40-42.

The first proof of the validity of Sister Celine's algorithm in a large class of hypergeometric summands is due to

P. Verbaeten, The automatic construction of pure recurrence relations, Proc. EUROSAM '74, ACM-SIGSAM Bulletin **8**, 96-98 (1974).

A proof of the above in the setting of holonomic function theory is in

D. Zeilberger, *A holonomic systems approach to special function identities*, J. Comput. Appl. Math. **32** (1990), 321-368.

The telescoping certification algorithm, and the recognition of the importance of certification itself, are in

D. Zeilberger, *The method of creative telescoping*, J. Symbolic Computation **11** (1991), 195-204.

D. Zeilberger, *A fast algorithm for proving terminating hypergeometric identities*, Discrete Math. **80** (1990), 207-211;

The WZ certification algorithms and the companion and dual identities are found in

Herbert S. Wilf and Doron Zeilberger, *Rational functions certify combinatorial identities*, J. Amer. math. Soc. **1** (1990), 147-158.

The extension of the methods to multivariate, "q", constant term, and multisum-integral identities is due to

Herbert S. Wilf and Doron Zeilberger, *An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities*, Invent. Math. **108** (1992), 575-633.

The decision procedure for finding hypergeometric solutions of recurrence relations is due to

Marko Petkovšek, *Finding closed-form solutions of difference equations by symbolic methods*, Doctoral dissertation, Carnegie Mellon University (1990), CMU-CS-91-103.