All the Hypergeometric Series Evaluations that are Fit to Print (By Looking Under the Hood of Zeilberger's Algorithm)

Doron ZEILBERGER [1]

Abstract: By looking under the hood of Zeilberger's algorithm, as simplified by Mohammed and Zeilberger, it is shown that all the classical hypergeometric closed-form evaluations can be *discovered ab initio*, as well as many "strange" ones of Gosper, Gessel and Stanton. The Maple package twoFone, accompanying this article, also finds many apparently new strange ${}_2F_1$ evaluations, and these discoveries are in some sense exhaustive. Hence WZ theory is transgressing the boundaries of the *context of verification* into the *context of discovery*. An appendix by Shalosh B. Ekhad lists many such new strange evaluations, complete with proofs (of course!).

Prerequisites: We assume familiarity with [MZ].

"That's very Nice that you Computers can Prove Identities, But you Still Need Us Humans to Conjecture Them!", Well, No Longer!

Recall that the Zeilberger algorithm[Z] inputs a *proper hypergeometric* term $F(n,k)$ (in the two discrete variables n and k), and outputs a non-negative integer L, polynomials (of n) $e_0(n), e_1(n), \dots, e_L(n)$ and a rational function (of n and k) $R(n,k)$, such that, if $G(n,k) := R(n,k)F(n,k)$, then

$$\sum_{i=0}^{L} e_i(n)F(n+i,k) = G(n,k+1) - G(n,k) \quad . \qquad (Zpair)$$

Assuming that $F(n,\pm\infty) = 0$ (as is often the case), we have, by summing w.r.t. k, and noting that the sum on the right telescopes to 0, that

$$a(n) := \sum_{k=-\infty}^{\infty} F(n,k)$$

satisfies an *homogeneous linear recurrence equation with polynomial coefficients*:

$$\sum_{i=0}^{L} e_i(n)a(n+i) = 0 \quad . \qquad (Recurrence)$$

If the *order*, L happens to be 1, then the recurrence can be solved in *closed-form*, and we have a closed-form evaluations.

In particular, the Zeilberger algorithm (as implemented in my own Maple package `EKHAD`, and starting with Maple 6, in the built-in package `SumTools[Hypergeometric]`, and there are very

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg at math dot rutgers dot edu` , `http://www.math.rutgers.edu/~zeilberg/`.
First version: Oct. 3, 2004 Accompanied by the Maple packages `FindHyperGeometric` and `twoFone`, available from Zeilberger's website. Supported in part by the NSF.

popular Mathematica implementations by Paule and Schorn) can immediately discover the right hand side, if the left side is given. For example, if one inputs

$$\sum_{k=-n}^{n} (-1)^k \binom{2n}{n+k}^3 \quad ,$$

one would get back the closed-form evaluation $(3n)!/n!^3$.

Since the set of conceivable hypergeometric summands (that humans or computers can write down) is countable, one can arrange them in lexicographic order, and eventually, just like in Hilbert's dream, get to any specific hypergeometric sum, and get the recurrence it satisfies (and with Petkovsek's celebrated algorithm (see [PWZ]), one can guarantee that it is minimal). If we are only interested in finding *closed-form* evaluations, i.e. the cases when $L = 1$, then we can just discard all the times when we get $L > 1$, and then publish a book of 'closed-form evaluations'.

Alas, in this way it may take a thousand years to get to Saalschütz or Dixon, and a million years to get to Dougall. In this article I will outline an *efficient* algorithm for outputting *all* hypergeometric closed-form evaluations of any bounded 'complexity'.

Of course, the notion of *duality* of WZ-pairs ([WZ], see also [PWZ]) *explains* all "strange" evaluations, naturally, as duals of *specializations* of classical identities, and gives a way of generating hypergeometric closed-form evaluations *ad nauseum*, by *specializing and dualizing*. Alas, if we do this randomly, we would get lots of 'new' such evaluations, but the summands, while technically proper-hypergeometric, are usually extremely messy, in the sense, that they are far from being *purely-hypergeometric*, i.e. their *polynomial part* (what is called $POL(n, k)$ in [MZ]) is of high degree. Also, this approach needs the classical identities (Saalschütz, Dixon, Dougall etc.) as *starters*.

In the present approach we can rediscover *from scratch*, in a natural way, all the classical, and the so-called *strange* ([GS]) hypergeometric closed-form evaluations. The algorithm also *discovers* many **new** such strange identities.

Under the Hood of Zeilberger

The Zeilberger algorithm has been considerably simplified in [MZ]. There it is shown that there is a *sharp* upper bound for L, which is really what it should be, if $F(n, k)$ is replaced by $F(n, k)x^k$. But in special cases, L may be smaller. I strongly recommend that the reader experiment with procedure `DoronV` in the Maple packages `FindHyperGemetric` and `twoFone` accompanying this article. `DoronV` is a verbose rendition of the simplified Zeilbeger algorithm of Mohammed and Zeilberger[MZ], and lists any 'miracles' that happen that help reduce L from its generic promised value.

Recall from [MZ] that everything boils down to solving the *linear equation*

$$f(k)X(k+1) - g(k-1)X(k) - h(k) = 0 \quad , \qquad (Gosper)$$

where $f(k)$, $g(k)$, $h(k)$ are certain polynomials derivable from the input, $h(k)$ depends linearly on the unknowns $e_i(n)$'s, and the coefficients of $X(k)$ are also unknowns. The argument in [MZ] displays an L, and a *degree* for $X(k)$, such that if one substitutes a generic polynomial of the appropriate degree in k, for $X(k)$, expands, and compares coefficients of x, one gets a system of *homogeneous* linear equations with *more unknowns then equations*, and hence with a *guarantee* for a non-zero solution.

Miracles

However, sometimes miracles *do* occur. A *miracle of the first kind* happens when there is a potential cancellation in the left of (*Gosper*). That happens when the degrees (in k) of $f(k)$ and $g(k)$ are the same, and the leading coefficients of $f(k)$ and $g(k)$ are the also the same. Then the potential degree gets upped by 1.

A *miracle of the second kind*, that is extremely rare results when it is possible for the degree of $X(k)$ to be even higher. By writing the leading and second-to-leading in generic form, with *generic* degree, plugging into (*Gosper*) and looking at the leading coefficient, we get a certain equation for that guessed degree. Usually (and certainly generically) the solution would be *symbolic*, and hence impossible, but if it happens to be *numeric* and exceeds the generic degree above (or one more than that, in case a miracle of the first kind has already happened), then we do indeed have a miracle of the second kind. Interestingly, Apéry's celebrated sum

$$\sum_{k=0}^{n} \binom{n+k}{k}^2 \binom{n}{k}^2 \quad ,$$

whose generic order should be 4, actually has order 2, because it is a beneficiary of this extremely rare miracle of the second kind.

Finally, once we settled for the highest-possible-degree for $X(k)$, and replace it in (*Gosper*) by a generic polynomial of that degree with *undetermined* coefficients, expand everything, and compare coefficients of x, we get a system of linear equations for the coefficients of $X(k)$ and the $e_i(n)$'s. Sometimes having and first and/or second miracle is already enough to have more unknowns than equations, but in the contrary case, there is still hope. Sometimes a system of linear equations with more equations than unknowns it does have a non-zero solution. All we need is that a certain determinant (or determinants) vanish! In that case, we have a *miracle of the third kind.*

Note that the third miracle may still happen even if the first and second ones did not. Sometimes the first miracle suffices by itself, sometimes we need the first and the second, sometimes we need the first and the third, and sometimes (in fact most of the time) the third miracle by itself suffices.

How to Manufacture Miracles by Tweaking the Zeilberger algorithm

Recall that the input has the form

$$F(n,k) = POL(n,k) \cdot H(n,k) \quad , \qquad (ProperHypergeometric)$$

where $POL(n,k)$ is a polynomial in (n,k) and

$$H(n,k) = \frac{\prod_{j=1}^{A} (a''_j)_{a'_j n + a_j k} \prod_{j=1}^{B} (b''_j)_{b'_j n - b_j k}}{\prod_{j=1}^{C} (c''_j)_{c'_j n + c_j k} \prod_{j=1}^{D} (d''_j)_{d'_j n - d_j k}} z^k \quad , \qquad (PureHypergeometric)$$

where the $a_j, a'_j, b_j, b'_j, c_j, c'_j, d_j, d'_j$ are *non-negative integers*, and z, $a''_j, b''_j, c''_j, d''_j$ are *commuting indeterminates*.

Now fix the $a_j, a'_j, b_j, b'_j, c_j, c'_j, d_j, d'_j$ (sorry about that, this can't be helped, at least for the present approach), but keep the $a''_j, b''_j, c''_j, d''_j$ and z as *indeterminates* and ask for what *specialization* will one or more of the above miracles happen. It is very easy for the computer to find conditions for the first miracle, and also for the second (usually it does not happen). The hardest miracle to perform (computationally) is the third. We have to set a certain determinant (or determinants) to zero, and get, this time a set of *non-linear* (polynomials) equations for the $a''_j, b''_j, c''_j, d''_j$'s and z. A priori, there may be no solution (and indeed often no miracle is possible), but whenever there is a solution, the computer can find it, since it knows, thanks to Bruno Buchberger and his Gröbner bases, how to solve a system of polynomial equations. For now we are using Maple's built-in implementation, but it may be a good idea to use special-purpose programs like Macaulay, SINGULAR, or MAGMA.

Of course, we are unable to guarantee that we found all hypergeometric identities, even not all ${}_2F_1$'s, but what the package `twoFone` (to be hopefully followed by packages like `threeFtwo`) finds *all* tuples (a,b,c,b',c',z) such that

$${}_2F_1\left(\begin{matrix} -an\,,\, bn+b' \\ cn+c' \end{matrix}\; ;\, z\right)$$

admits a closed-form evaluation and a,b,c lie in the *range* $1 \le a \le K$, $-K \le b,c \le K$. For *any* inputted positive integer K. The computer also discards all specializations of the classical identities of Gauss and Kummer, as well as any consequences of previously discovered identities via the Euler and the two Pfaff Transformations (see [AAR], ch.2, equations (2.2.6), (2.2.7) and (2.3.14)). So the final listing contains *mutually independent* genuinely new "strange" identities. Of course, some of them were already discovered and/or proved by Gosper and Gessel & Stanton (see [GS]), but many of them seem brand-new.

Acknowledgement: Many thanks are due to Mohamud Mohammed for many stimulating discussions.

Appendix by Shalosh B. Ekhad

REFERENCES

[MZ] M. Mohammed and D. Zeilberger, *Sharp upper bounds for the orders of the recurrences outputted y the Zeilberger and q-Zeilberger algorithms*, submitted. Available from the authors' websites.