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 accompanying Maple package `FindHypergeometric` explains the various miracles that account for the classical evaluations, and the more specialized Maple package `twoFone`, also accompanying this article, finds many "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.*

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 Simplified Zeilberger algorithm[MZ] inputs a *proper hypergeometric* term

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

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.* It outputs a non-negative integer L, polynomials (of n) $e_0(n), e_1(n), \ldots, e_L(n)$ 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

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

[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.

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 is a very popular Mathematica implementation by Paule and Schorn [PS]) can immediately discover the right hand side, if the left hand side is given. For example, if one inputs

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

one would get back a first-order recurrence for the sum, that immediately entails 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[P] celebrated algorithm (see also [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', of any bounded 'complexity'.

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'. Unfortunately, if that complexity gets higher, it runs out of time and memory, but it can outperform by orders of magnitude an exhaustive search.

WZ theory already has a mechanism for generating new identities by the process of *specializing and dualizing*, ([WZ], see also [PWZ]). This technique was elevated to an art form by Ira Gessel[Ge]. 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* ($POL(n,k)$ above) 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.

Robert Maier's Neo-Classical Approach

Hypergeometric started out as *solutions* of certain ordinary differential equations. Using this fact, Euler, Gauss, Kummer, Goursat and other giants found *transformation formulas*, that lead to the classical evaluations. On the other hand, WZ theory considers the erstwhile *parameters* as **active** *discrete* variables.

Quite recently, Robert Maier[M] found a brand-new transformation formula for (general!) ${}_{r+1}F_r$, that obeys *algebraic constraints*. I don't see anyway to prove them, let alone, discover them with WZ theory. So the moral is

Find new approaches but keep the old ones ....

Under the Hood of Zeilberger

The Zeilberger algorithm has been considerably simplified in [MZ] (for proper-hypergeometric summands, Abramov and Le ([A],[AL]) showed that Zeilberger's algorithm may even work for non-proper summands). 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 `DoronMiracles` in the Maple package `FindHyperGeometric` accompanying this article. `DoronMiracles` 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*, let's call it M, for $X(k)$, such that if one substitutes a generic polynomial of degree M in k, for $X(k)$, expands $(Gosper)$, and then sets all the coefficients of the resulting polynomial in k to 0, one gets a system of *homogeneous* linear equations with *more unknowns than equations*, and hence with a *guarantee* for a non-zero solution.

Miracles

However, sometimes pleasant surprises occur, and the guaranteed L can be made lower, thanks to *miracles.*

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* can only happen in the wake of miracle of the first kind. It is an extremely rare event. It happens when the potential degree of $X(k)$ can be made yet higher. To test for it, write the leading and second-to-leading coefficients of $X(k)$ in generic form, with *generic* degree, plug into (*Gosper*), and look at the leading coefficient and set it equal to 0. You will 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 proposed degree promised by the first miracle, 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 is 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 replaced it in (*Gosper*) by a generic polynomial of that degree, with *undetermined* coefficients, we expand everything, and set all the coefficients of the polynomial, in k, of the left side of (*Gosper*),to zero. We get a system of homogeneous linear equations for a certain set of unknowns. These unknowns consist of the coefficients of $X(k)$, as well as the coefficients, $e_i(n)$, of the desired recurrence. 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*.

Indeed, if in luck, a system of linear equations with more equations than unknowns it may 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.

The Miracles That Gauss, Kummer, Saaslchütz, Dixon, and Dougall Should Be Grateful For

By running `DoronMiracles` in the Maple package `FindHyperGeometric` (type `ezra(DoronMiracles):` there, for help), we are told that

Gauss's ${}_2F_1(a,b;c;1)$ happens because of the **first miracle**, that suffices.

Gauss's ${}_2F_1(2a,2b;a+b+1/2;1/2)$ happens because of the **third miracle**. The first miracle didn't happen, but the **third** one saved the day. This seems to be the case in all the **strange** evaluations, at least for ${}_2F_1$'s (see [E]).

Likewise, Kummer's ${}_2F_1(a,b;1+a-b;1)$ happens *only* because of the **third miracle**. So it deserves the name **strange**, even though it has two two-parameters.

The celebrated Pfaff-Saalchütz four-parameter ${}_3F_2(a,b,-n;c,1+a+b-c-n;1)$ evaluation happens because of the **first and second** miracles. The generosity of the second miracle produces less equations than unknowns (in the last, solving (*Gosper*), phase), hence there was no need for another miracle.

For equally celebrated Dixon three-parameter ${}_3F_2(a,b,c;1+a-b,1+a-c;1)$ evaluation the **first miracle** did happen. Alas, **no second miracle**. But cheer up, Dixon, even though (*Gosper*) demands that you solve a set with three equations and three unknowns, that has a priori probability 0 of success, nevertheless it canbe solved, thanks to a **miracle of the third kind**.

Last but not least, the Gigantic ${}_7F_6$ (which, in our notation is really a mere ${}_6F_5$ with a linear polynomial in front) owers its *raison d'être* to **all three miracles**.

To summarize:

Gauss: **1**;

GaussHalf, Kummer, all 'strange': ${}_2F_1$: **3**.

Pfaff-Saalscütz: **1,2**.

Dixon: **1,3**.

Dougall: **1,2,3**.

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* (polynomial) 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* $1 \leq a \leq K_1$, $-K \leq b, c \leq K$. For *any* inputted positive integers K_1, 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. We should mention again Maier's[Ma] recent work, that in particular produced two new 'strange' ${}_2F1$'s with argument $x = -1$, that were rediscovered by Ekhad in [E].

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

REFERENCES

[A] S.A. Abramov, *When does Zeilberger's algorithm succeed?*, Advances in Applied Mathematics **30** (2003), 424-441.

[AL] S.A. Abramov and H.Q. Le, *A criterion for the applicabilty of Zeilberger's algorithm to rational functions*, Discrete Math. **259** (2002), 1-17.

[AAR] G.E. Andrews, R. Askey, and R. Roy, *"Special Functions"*, Cambridge Univ. Press, 1999.

[E] Shalsoh B. Ekhad, *Forty "Strange" Computer-Discovered and Computer-Generated (of course!) Hypegeometric Series Evaluations, Personal Journal of Shalosh B. Ekhad and Doron Zeilbeger,* `http://www.math.rutgers.edu/~zeilberg/pj.html` .

[Ge] I. M. Gessel, Finding Identities with the WZ Method, J. Symbolic Computation **20***(1995), 537-566.*

[GS] I. Gessel and D. Stanton, Strange evaluations of hypergeometric series, SIAM J. Math. Anal. **13***(1982), 295-308.*

[Go] R. W. Gosper, Jr., Decision procedures for indefinite hypergeometric summation, Proc. Natl. Acad. Sci. USA **75** *(1975), 40-42.*

[M] R.S. Maier, A Generalization of Euler's hypergeometric transformation, preprint. (available from `arXiv.org`*).*

[MZ] M. Mohammed and D. Zeilberger, Sharp upper bounds for the orders of the recurrences out-

putted by the Zeilberger and q-Zeilberger algorithms, to appear in J. Symb. Comp., available from the authors' websites.

[PS] P. Paule and M. Schorn, A Mathematica version of Zeilberger's algorithm for proving binomial coefficients identities, J. Symbolic Comput. **20** *(1995), 673-698.*

[Pe] M. Petkovsek, Hypergeometric solutions of linear recurrences with polynomial coefficients, J. Symbolic Computation **14***(1992), 243-264.*

[PWZ] M. Petkovsek, H. S. Wilf and D. Zeilberger, A=B, AK Peters, Wellesley, (1996). [available on-line from the authors' websites.]

[WZ] H.S. Wilf and D. Zeilbegrer, Rational functions certify combinatorial identities, J. Amer. Math. Soc. **3** *(1990), 77-83.*

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